

22 Στατιστικές Μέθοδοι Ανάλυσης Πειραματικών Δεδομένων Συνεργασίας

Χρήστος Κατσάνος και Νικόλαος Αβούρης
Πανεπιστήμιο Πατρών

Σκοπός

Το παρόν κεφάλαιο, συμπληρωματικό του κυρίως υλικού του βιβλίου, περιλαμβάνει μια επισκόπηση των βασικών στατιστικών εννοιών και μεθόδων που χρησιμοποιούνται στα πλαίσια επιστημονικών ερευνών και μελετών σχετικών με τη συνεργατική αλληλεπίδραση σε εργασιακά και μαθησιακά περιβάλλοντα. Είναι γνωστό ότι η επιστημονική έρευνα βασίζεται στη διατύπωση υποθέσεων, στη σχεδίαση πειραμάτων παρατήρησης και στην παρατήρηση φαινομένων, ώστε να γίνει εφικτή η συλλογή και ανάλυση πληροφοριών και μετρήσεων που προέρχονται από αυτές τις παρατηρήσεις, και στη συνέχεια η εξαγωγή χρήσιμων και γενικεύσιμων συμπερασμάτων προς επιβεβαίωση ή απόρριψη των αρχικών υποθέσεων. Στο κεφάλαιο αυτό γίνεται μια επισκόπηση των βασικών μεθόδων της στατιστικής ώστε να υποστηριχθεί ο μελετητής σχετικά με την επιλογή της κατάλληλης μεθόδου. Αν και η έμφαση δίνεται σε μελέτες που αφορούν στη χρήση συνεργατικών τεχνολογιών στην εργασία και την μάθηση, οι μέθοδοι που παρουσιάζονται μπορούν να χρησιμοποιηθούν σε οποιαδήποτε διαδικασία επιστημονικής έρευνας.

Έννοιες – Κλειδιά

- | | |
|-------------------------------------|-------------------------------------|
| • Περιγραφική Στατιστική | • Έλεγχος υποθέσεων |
| • Επαγωγική Στατιστική | • Μηδενική υπόθεση |
| • Πληθυσμός - Δείγμα | • Επίπεδο σημαντικότητας |
| • Ανεξάρτητη - Εξαρτημένη μεταβλητή | • Στατιστικός δείκτης ελέγχου |
| • Κατανομές συχνότητας | • Παραμετρικά – μη παραμετρικά τεστ |
| • Μέση τιμή | • Σύσχέτιση |
| • Διακύμανση - Τυπική απόκλιση | • Στατιστικά πακέτα λογισμικού |
| | • Section Break (Continuous) |

Εισαγωγικές Παρατηρήσεις

Είναι γνωστό ότι η επιστημονική έρευνα βασίζεται στην παρατήρηση φαινομένων για να συλλέξει και να αναλύσει πληροφορίες και μετρήσεις που προέρχονται από αυτές τις παρατηρήσεις, και στη συνέχεια να προχωρήσει στην εξαγωγή χρήσιμων και γενικεύσιμων συμπερασμάτων. Η αναφορά σε αυτό το σύνολο των πληροφοριών και μετρήσεων συνήθως γίνεται με τον επιστημονικό όρο *δεδομένα (data)*. Ένας ερευνητής διεξάγει επιστημονικές μελέτες και έρευνες (σαν αυτές που παρουσιάστηκαν στην ενότητα 14.6.4), με σκοπό να δώσει απαντήσεις σε συγκεκριμένα ερωτήματα και γενικούς προβληματισμούς που σχετίζονται με μία συγκεκριμένη ομάδα ή ομάδες ατόμων ή αντικειμένων (ή άλλων οντοτήτων).

Η επιστημονική έρευνα επιτρέπει στους ερευνητές να προβαίνουν σε παρατηρήσεις των φαινομένων και των προβλημάτων που εξετάζουν. Η *Στατιστική* παρέχει τα κατάλληλα εργαλεία για το σχεδιασμό της διαδικασίας συλλογής, περιγραφής, ανάλυσης και επεξεργασίας αυτών των παρατηρήσεων. Δηλαδή, ένα σύνολο από μεθόδους και κανόνες που επιτρέπουν την οργάνωση μιας συλλογής δεδομένων, τη συνοπτική και αποτελεσματική παρουσίαση και ανάλυση τους, και τέλος την εξαγωγή χρήσιμων συμπερασμάτων.

Εύκολα γίνεται αντιληπτό λοιπόν, ότι η επιστήμη της στατιστικής παρέχει εργαλεία για οποιαδήποτε επιστημονική έρευνα. Το παρόν κεφάλαιο αποσκοπεί στην εξοικείωση του αναγνώστη με τις βασικές στατιστικές έννοιες και μεθόδους που χρησιμοποιούνται στα πλαίσια των επιστημονικών ερευνών, και πιο συγκεκριμένα των ερευνών που μελετούν την επίδραση της χρήσης συνεργατικών τεχνολογιών στην εργασία και στη μάθηση. Είναι εμφανές ότι στα πλαίσια ενός κεφαλαίου είναι αδύνατο να γίνει μία αναλυτική παρουσίαση όλων των τεχνικών που περιλαμβάνει η στατιστική ανάλυση δεδομένων. Για τους αναγνώστες που ενδιαφέρονται να ασχοληθούν περισσότερο με τον επιστημονικό κλάδο της στατιστικής υπάρχει πληθώρα διαθέσιμων συγγραμμάτων (Gravetter & Wallnau 1996, Howell, 1989, Klugh, 1986, Keppel and Saufley, 1980).

22.1 Βασικές Στατιστικές Έννοιες και Όροι

Σε αυτό το σημείο, και πριν προχωρήσουμε σε μία σύντομη περιγραφή των μεθόδων που περιλαμβάνει η στατιστική για την οργάνωση και ανάλυση των δεδομένων, κρίνεται σκόπιμο να παρουσιαστούν οι βασικές έννοιες και οι όροι που χρησιμοποιούνται στη στατιστική.

22.1.1 Πληθυσμοί και Δείγματα

Όπως αναφέρεται και παραπάνω, η επιστημονική έρευνα σχετίζεται με ένα γενικό προβληματισμό ή μια συγκεκριμένη ερώτηση που αφορά ομάδες ατόμων, αντικειμένων ή άλλων οντοτήτων. Οι ομάδες αυτές αποτελούν το επίκεντρο του ενδιαφέροντος σε μία μελέτη, και είναι γνωστές με τον όρο *πληθυσμός (population)* στην επιστήμη της στατιστικής. Ο πληθυσμός μπορεί να αποτελείται από έμψυχα αντικείμενα (ομάδες ανθρώπων), από άψυχα

αντικείμενα (εξαρτήματα αυτοκινήτων), ή από οτιδήποτε άλλο ο ερευνητής θέλει να εξετάσει (αφηρημένες έννοιες). Είναι εμφανές ότι ένας πληθυσμός μπορεί να είναι από πολύ μικρός (π.χ. οι έφηβοι σε μία τάξη σε ένα σχολείο), έως πολύ μεγάλος (π.χ. το σύνολο όλων των εφήβων του πλανήτη). Επειδή η επιστημονική έρευνα οδηγεί σε ένα σύνολο από παρατηρήσεις για τα υποκείμενα (*subjects*) του πληθυσμού, τις περισσότερες φορές ο όρος πληθυσμός αναφέρεται σε ένα σύνολο παρατηρήσεων που περιγράφουν τα υπό εξέταση υποκείμενα.

Όμως συχνά η μελέτη ολόκληρου του πληθυσμού δεν είναι πρακτικά δυνατή, γιατί το μέγεθος του είναι πολύ μεγάλο. Για αυτό το λόγο, τις περισσότερες φορές ένα υποσύνολο του πληθυσμού, που ονομάζεται *δείγμα* (*sample*), αποτελεί το αντικείμενο της μελέτης. Το δείγμα λοιπόν, είναι ένα σύνολο από υποκείμενα που έχει επιλεγεί κατάλληλα ώστε να αντιπροσωπεύει έναν ολόκληρο πληθυσμό. Προφανώς, η διαδικασία επιλογής του δείγματος αποτελεί μία εξαιρετικά σημαντική διαδικασία, καθώς καθορίζει την εγκυρότητα των αποτελεσμάτων της μελέτης. Αν το δείγμα δεν είναι αντιπροσωπευτικό, υπάρχει μεγάλη πιθανότητα η μελέτη να οδηγήσει σε εσφαλμένα συμπεράσματα.

Στην πράξη, ένα δείγμα είναι αντιπροσωπευτικό του πληθυσμού όταν έχει χρησιμοποιηθεί η διαδικασία της *τυχαίας δειγματοληψίας* (*random sampling*) για την απόκτηση του. Η τυχαία δειγματοληψία απαιτεί κάθε υποκείμενο του πληθυσμού να έχει την ίδια πιθανότητα να επιλεγεί. Όπως και οι πληθυσμοί, ένα δείγμα μπορεί να είναι αντίστοιχα από πολύ μικρό έως και πολύ μεγάλο. Εύκολα γίνεται αντιληπτό ότι όσο μεγαλύτερο είναι το δείγμα, τόσο πιο αντιπροσωπευτικό θα είναι, καθώς αυξάνεται ο αριθμός των υποκειμένων που επιλέγονται από τον πληθυσμό.

Στην στατιστική όταν χρησιμοποιούμε δεδομένα κρίνεται αναγκαίο να προσδιορίζουμε εάν τα δεδομένα προέρχονται από ένα πληθυσμό ή από ένα δείγμα. Για να εξυπηρετηθεί αυτός ο διαχωρισμός, η στατιστική χρησιμοποιεί τον όρο *παράμετρος* (*parameter*) για να περιγράψει δεδομένα που αναφέρονται στον πληθυσμό, και τον όρο *στατιστικός δείκτης* (*statistic*) για τα δεδομένα που σχετίζονται με ένα δείγμα.

22.1.2 Κατηγορίες Στατιστικών Μεθόδων

Αν και η επιστήμη της στατιστικής προτείνει ένα μεγάλο σύνολο από τεχνικές για την οργάνωση και την ανάλυση των δεδομένων, αυτές οι τεχνικές μπορούν να κατηγοριοποιηθούν σε δύο γενικές κατηγορίες.

Η πρώτη κατηγορία, η *περιγραφική στατιστική* (*descriptive statistics*), περιλαμβάνει μεθόδους για την οργάνωση, απλοποίηση και συνοπτική παρουσίαση των δεδομένων. Αν και υπάρχουν πολλές τεχνικές που ανήκουν σε

αυτήν την κατηγορία, η πιο διαδεδομένη είναι ο υπολογισμός της μέσης τιμής (*mean*) και της τυπικής απόκλισης (*standard deviation*). Άλλοι περιγραφικοί στατιστικοί δείκτες θα παρουσιαστούν συνοπτικά στη συνέχεια του κεφαλαίου.

Όπως ήδη αναφέρθηκε, τις περισσότερες φορές είναι αδύνατο να συλλέξουμε μετρήσεις για κάθε υποκείμενο του πληθυσμού, και γι αυτό χρησιμοποιούμε ένα αντιπροσωπευτικό υποσύνολο του, τα δείγματα. Η δεύτερη κατηγορία, η επαγωγική στατιστική (*inferential statistics*), περιλαμβάνει τεχνικές που επιτρέπουν την ανάλυση τέτοιων δεδομένων ώστε να γίνει εφικτή η εξαγωγή χρήσιμων και γενικεύσιμων συμπερασμάτων για τους πληθυσμούς, με βάση τις πληροφορίες που συλλέγονται από τα δείγματα.

Ωστόσο, είναι προφανές ότι όσο αντιπροσωπευτικό και αν θεωρείται ένα δείγμα, δε μπορεί να παρέχει μια πιστή αναπαράσταση του πληθυσμού. Έτσι, πάντα θα υπάρχει μία ασυμφωνία, ή με άλλα λόγια ένα ποσοστό λάθους, ανάμεσα στο στατιστικό δείκτη που προκύπτει από το δείγμα, και την αντίστοιχη τιμή της παραμέτρου του πληθυσμού. Αυτό το ποσοστό λάθους ονομάζεται σφάλμα δειγματοληψίας (*sampling error*), και αποτελεί ένα από τα κύρια προβλήματα που καλείται να αντιμετωπίσει ο ερευνητής όταν προσπαθεί να εξάγει γενικά συμπεράσματα για πληθυσμούς έχοντας στη διάθεση του ένα αριθμό από δείγματα. Το σφάλμα δειγματοληψίας τονίζει το γεγονός ότι ένας στατιστικός δείκτης αποτελεί μόνο μία εκτίμηση της αντίστοιχης παραμέτρου του πληθυσμού.

22.1.3 Κατηγορίες Μεταβλητών και Δεδομένων

Τυπικά, αυτό που ενδιαφέρει έναν ερευνητή είναι να απαντήσει σε ερωτήματα που σχετίζονται με έναν ή περισσότερους πληθυσμούς, έχοντας στη διάθεση του κάποιο αριθμό δειγμάτων. Τέτοια ερωτήματα, για παράδειγμα, παρουσιάζονται στην ενότητα 22.3.5 υπό μορφή σεναρίων έρευνας.

Είναι φανερό ότι τις περισσότερες φορές ο στόχος της έρευνας είναι να δημιουργήσει μία σχέση αιτίου-αποτελέσματος ανάμεσα σε δύο μεταβλητές (*variables*).⁶⁸ Με άλλα λόγια, ο στόχος σε μία έρευνα είναι να αποδειχθεί ότι οι αλλαγές που εμφανίζονται στην τιμή μιας μεταβλητής οφείλονται στην αλλαγή της τιμής μιας άλλης μεταβλητής. Για να γίνει αυτό εφικτό, η επιστημονική έρευνα χρησιμοποιεί δείγματα και προσεκτικά σχεδιασμένα πειράματα, όπου ο ερευνητής μπορεί να ελέγχει την τιμή μιας μεταβλητής και ταυτόχρονα να παρατηρεί τις τιμές της άλλης μεταβλητής, χωρίς η διαδικασία του πειράματος

⁶⁸ Μεταβλητή ονομάζεται οποιοδήποτε χαρακτηριστικό ή κατάσταση παρουσιάζει αλλαγή, ή έχει διαφορετική τιμή για διαφορετικά υποκείμενα τα οποία ανήκουν σε ένα πληθυσμό ή ένα δείγμα. Δεν είναι ασυνήθιστο ο ερευνητής να ελέγχει και να παρατηρεί πολλές μεταβλητές ταυτόχρονα, αλλά για ακαδημαϊκούς λόγους χρησιμοποιούμε την απλούστερη περίπτωση όπου ελέγχει μία μεταβλητή και παρατηρεί μία άλλη.

να επηρεάζεται από εξωτερικούς παράγοντες ή μεταβλητές. Η μεταβλητή που ελέγχει ο ερευνητής ονομάζεται *ανεξάρτητη μεταβλητή* (independent variable), ενώ αυτή που αποτελεί το αντικείμενο παρατήρησης ονομάζεται *εξαρτημένη μεταβλητή* (dependent variable). Το ζητούμενο για τον ερευνητή είναι να εξετάσει την επίδραση των διαφορετικών τιμών της ανεξάρτητης μεταβλητής στις τιμές της εξαρτημένης.

Όπως έχει ήδη αναφερθεί, μία επιστημονική μελέτη απαιτεί τη συλλογή ενός συνόλου δεδομένων. Τα δεδομένα που συλλέγονται μπορεί να είναι ποιοτικά ή ποσοτικά. Ποιοτικά ονομάζονται τα δεδομένα που δεν επιδέχονται μέτρηση, όπως για παράδειγμα το επίπεδο της μόρφωσης, το φύλο, το χρώμα των ματιών, κλπ, ενώ ποσοτικά αυτά που μπορούν να μετρηθούν με τη βοήθεια κάποιων κλίμακας, όπως για παράδειγμα το βάρος, το ύψος, ο βαθμός επίδοσης σε κάποιες εξετάσεις, κλπ. Τα ποιοτικά δεδομένα μπορεί να είναι οργανωμένα απλά σαν ονομαστικές κατηγορίες (nominal data), όπως για παράδειγμα το φύλο, ή σαν ταξινομημένες κατηγορίες (ordinal data), όπως για παράδειγμα η κατάταξη των μεταλλίων στους Ολυμπιακούς Αγώνες. Τα ποσοτικά δεδομένα διακρίνονται σε συνεχή (continuous), όπως για παράδειγμα ο χρόνος επίλυσης ενός μαθηματικού προβλήματος, και διακριτά (discrete), όπως για παράδειγμα ο αριθμός των παιδιών που έχει μία οικογένεια.

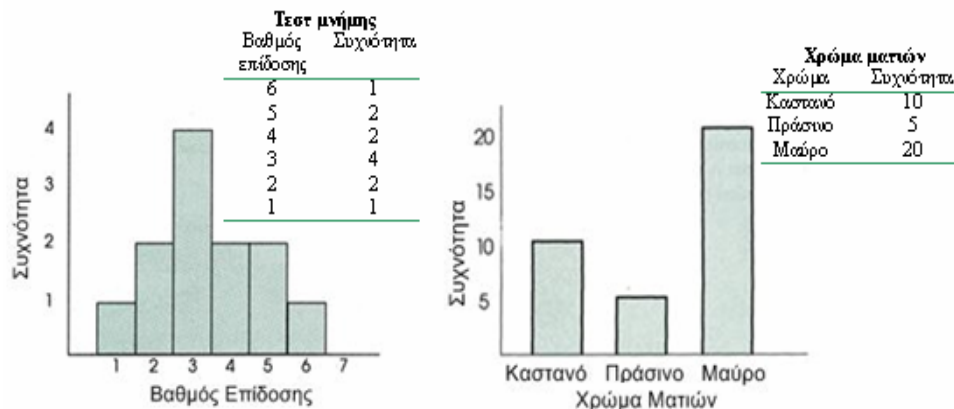
22.2 Περιγραφική Στατιστική

Τα αποτελέσματα ενός πειράματος συνήθως δημιουργούν ένα μεγάλο αριθμό δεδομένων. Κρίνεται λοιπόν αναγκαία η εύρεση διαδικασιών, με τις οποίες τα αποτελέσματα αυτά μπορούν να οργανωθούν και να παρουσιαστούν με απλό και εύληπτο τρόπο. Αυτός ακριβώς είναι και ο στόχος των περιγραφικών στατιστικών δεικτών, δηλαδή να παρέχουν μεθόδους που απλοποιούν και διευκολύνουν την οργάνωση και παρουσίαση των αποτελεσμάτων. Στην ενότητα αυτή παρουσιάζονται σύντομα ορισμένοι από τους πιο γνωστούς περιγραφικούς στατιστικούς δείκτες. Στην ενότητα 22.3.4 περιλαμβάνεται ένας συνοπτικός οδηγός για τη διαδικασία υπολογισμού των περιγραφικών στατιστικών δεικτών που περιγράφονται στην συνέχεια.

22.2.1 Κατανομές Συχνότητας

Οι *κατανομές συχνότητας* (frequency distributions) αποτελούν μία από τις πιο γνωστές μεθόδους για την παρουσίαση ενός συνόλου δεδομένων. Οι κατανομές συχνότητας μπορούν να δομηθούν είτε σαν πίνακες είτε σαν γραφικές παραστάσεις (Σχήμα 22.1). Και στις δύο περιπτώσεις, ο στόχος είναι να δοθεί μία πλήρης εικόνα για την κατανομή των δεδομένων στην κλίμακα μέτρησης. Ανάλογα με την κατηγορία των δεδομένων (ποσοτικά ή ποιοτικά),

χρησιμοποιούνται διαφορετικές μορφές αυτών των κατανομών (ιστογράμματα ή ραβδόγραμμα, αντίστοιχα).



Σχήμα 22.1 Κατανομές Συχνότητας υπό μορφή Πίνακα και Γραφικής Παράστασης για Ποσοτικά (ιστόγραμμα στα αριστερά) και Ποιοτικά (ραβδόγραμμα στα δεξιά) Δεδομένα

22.2.2 Μέτρα Κεντρικής Τάσης

Ο σκοπός των *μέτρων κεντρικής τάσης (measures of central tendency)* είναι να προσδιοριστεί ένα στατιστικό μέγεθος το οποίο μπορεί να χρησιμοποιηθεί για να αντιπροσωπεύσει ένα σύνολο δεδομένων. Για την επίτευξη αυτού του στόχου χρησιμοποιούνται συνήθως τρία μέτρα κεντρικής τάσης: η *μέση τιμή (mean)*, η *διάμεσος (median)* και η *επικρατούσα τιμή (mode)*. Τα μέτρα κεντρικής τάσης επιτρέπουν τη συνοπτική περιγραφή και τη σύγκριση των διαθέσιμων δεδομένων.

Στις περισσότερες περιπτώσεις χρησιμοποιείται η μέση τιμή σαν μέτρο κεντρικής τάσης, καθώς δίνει την πιο αντιπροσωπευτική εικόνα για ένα σύνολο δεδομένων. Η μέση τιμή ορίζεται ως το άθροισμα των παρατηρήσεων διά του πλήθους τους, δηλαδή ουσιαστικά πρόκειται για τον υπολογισμό του αριθμητικού μέσου όρου.

Ωστόσο, σε ορισμένες περιπτώσεις ο υπολογισμός της μέσης τιμής δεν είναι δυνατός (π.χ. υπάρχουν απροσδιόριστες τιμές στα δεδομένα, ή έχουμε διαθέσιμα ποιοτικά δεδομένα), ή δεν παρέχει την πιο αντιπροσωπευτική εικόνα (π.χ. η μέση τιμή μπορεί να επηρεαστεί σημαντικά από μία πολύ μεγάλη ή πολύ μικρή παρατήρηση). Σε αυτές τις περιπτώσεις χρησιμοποιείται σαν μέτρο κεντρικής τάσης η διάμεσος, που αντιστοιχεί σε εκείνη την τιμή για την οποία το 50% των παρατηρήσεων έχει τιμή ίση ή μικρότερη. Για τον υπολογισμό της διαμέσου διατάσσουμε τις παρατηρήσεις σε αύξουσα σειρά, και αν το πλήθος

τους είναι περιττό επιλέγουμε τη μεσαία παρατήρηση, ενώ αν είναι άρτιο η τιμή της διαμέσου δίνεται από το ημιάθροισμα των μεσαίων παρατηρήσεων.

Τέλος, σαν μέτρο κεντρικής τάσης μπορεί να χρησιμοποιηθεί και η επικρατούσα τιμή. Η επικρατούσα τιμή ουσιαστικά αντιστοιχεί στην παρατήρηση με τη μεγαλύτερη συχνότητα, και χρησιμοποιείται συνήθως συμπληρωματικά με τα δύο προηγούμενα μέτρα κεντρικής τάσης, ή όταν είναι διαθέσιμα ποιοτικά δεδομένα με τη μορφή ονομαστικών κατηγοριών, όπου δεν έχει νόημα ο υπολογισμός της μέσης τιμής ή της διαμέσου.

22.2.3 Μέτρα Διασποράς

Τα μέτρα διασποράς (*measures of variability*) δίνουν μία εικόνα σχετικά με το πόσο συγκεντρωμένες είναι οι παρατηρήσεις σε ένα σύνολο δεδομένων. Τα κυριότερα μέτρα διασποράς είναι το εύρος (*range*), το ήμι-ενδοτεταρτομοριακό εύρος (*semi-interquartile range*), η τυπική απόκλιση (*standard deviation*) και η διακύμανση (*variance*).

Το εύρος ορίζεται ως η διαφορά της μεγαλύτερης παρατήρησης από τη μικρότερη. Το εύρος θεωρείται ως ένα πρόχειρο μέτρο διασποράς, καθώς βασίζεται μόνο στις ακραίες παρατηρήσεις και επηρεάζεται σημαντικά από το μέγεθος του δείγματος, καθώς κάθε νέα παρατήρηση μπορεί να αλλάξει την τιμή της μικρότερης ή της μεγαλύτερης παρατήρησης στο δείγμα.

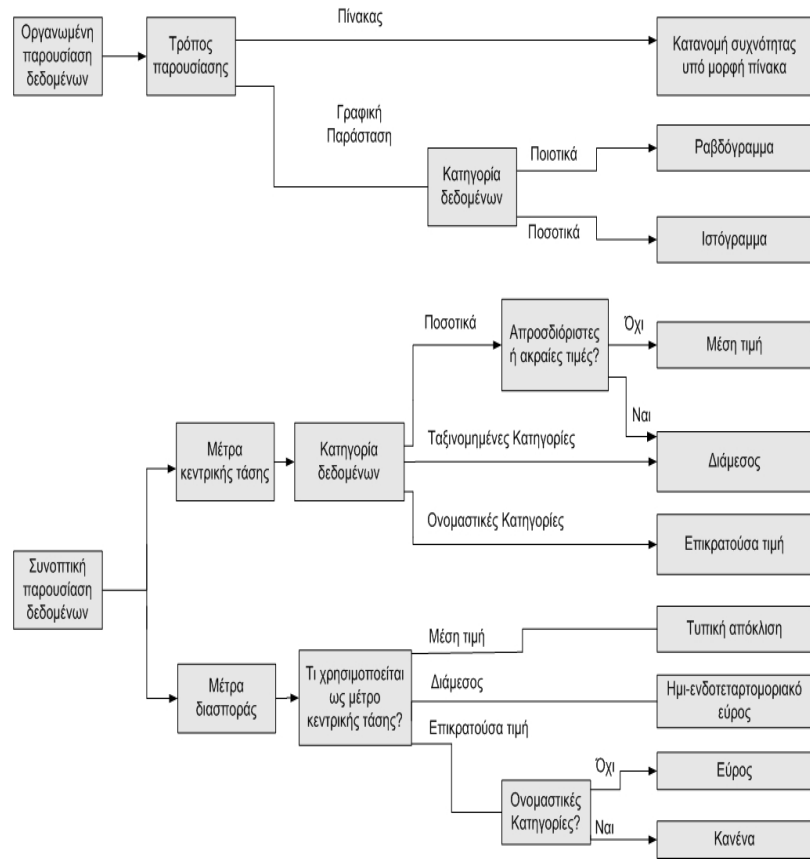
Το ήμι-ενδοτεταρτομοριακό εύρος ορίζεται σαν το μισό του εύρους που καλύπτει το κεντρικό 50% της κατανομής παρατηρήσεων. Επειδή το ήμι-ενδοτεταρτομοριακό εύρος επικεντρώνεται στο κεντρικό 50% της κατανομής, είναι λιγότερο πιθανό να επηρεαστεί από ακραίες παρατηρήσεις. Το ήμι-ενδοτεταρτομοριακό εύρος χρησιμοποιείται συνήθως σε συνδυασμό με τη διάμεσο ως μέτρο της κεντρικής τάσης των παρατηρήσεων.

Ωστόσο, και το ήμι-ενδοτεταρτομοριακό εύρος δε δίνει μία ακριβή εικόνα για τη διασπορά των παρατηρήσεων, καθώς δεν λαμβάνει υπόψη του τις πραγματικές αποστάσεις μεταξύ όλων των παρατηρήσεων. Η διακύμανση και η τυπική απόκλιση αποτελούν το πιο αξιόπιστο και το πιο συνηθισμένο μέτρο διασποράς. Αυτά τα μέτρα διασποράς χρησιμοποιούν ως σημείο αναφοράς τη μέση τιμή, και λαμβάνουν υπόψη την απόσταση όλων των παρατηρήσεων από αυτήν.

22.2.4 Διαδικασία και Τύποι Υπολογισμού Περιγραφικών Στατιστικών

Στη συνέχεια παρατίθεται ένα διάγραμμα αποφάσεων που μπορεί να βοηθήσει τον αναγνώστη στην επιλογή του πιο κατάλληλου περιγραφικού στατιστικού δείκτη για κάθε περίπτωση, και ένας πίνακας που συνοψίζει τη

διαδικασία υπολογισμού των περιγραφικών στατιστικών δεικτών που παρουσιάστηκαν στις προηγούμενες ενότητες.



Σχήμα 22.2 Διάγραμμα Αποφάσεων για την Επιλογή Περιγραφικών Στατιστικών Δεικτών

Περιγραφικός Στατιστικός Δείκτης	Τύποι – Διαδικασία Υπολογισμού		
Μέση Τιμή (mean)	Πληθυσμός	$\mu = \frac{\sum X}{N}$	X : τιμή παρατήρησης N : μέγεθος πληθυσμού
	Δείγμα	$\bar{X} = \frac{\sum X}{n}$	n : μέγεθος δείγματος
Διάμεσος (median)	Διάταξη παρατηρήσεων σε αύξουσα σειρά. Η διάμεσος ισούται με: <ul style="list-style-type: none"> • την μεσαία τιμή για πλήθος τιμών περιττό • το ημίαθροισμα των μεσαίων τιμών για πλήθος τιμών άρτιο 		
Επικρατούσα Τιμή (mode)	Ισούται με την τιμή με την μεγαλύτερη συχνότητα		
Εύρος (range)	Ισούται με την διαφορά της μεγαλύτερης από την μικρότερη τιμή		
Ημι-ενδοτεταρτομοριακό εύρος (semi-interquartile range)	Ισούται με το μισό του εύρους που καλύπτει το κεντρικό 50% της κατανομής παρατηρήσεων		
Διακύμανση (variance)	Πληθυσμός	$\sigma^2 = \frac{SS}{N}$	$SS = \sum X^2 - \frac{(\sum X)^2}{N}$
	Δείγμα	$s^2 = \frac{SS}{n-1}$	$SS = \sum X^2 - \frac{(\sum X)^2}{n}$
Τυπική απόκλιση (standard deviation – SD)	Πληθυσμός	$\sigma = \sqrt{\frac{SS}{N}}$	$SS = \sum X^2 - \frac{(\sum X)^2}{N}$
	Δείγμα	$s = \sqrt{\frac{SS}{n-1}}$	$SS = \sum X^2 - \frac{(\sum X)^2}{n}$

Πίνακας 22.1 Διαδικασία και Τύποι Υπολογισμού Τυπικών Περιγραφικών Στατιστικών Δεικτών για το Συνολικό Πληθυσμό και για Δείγμα

22.2.5 Παραδείγματα Επιστημονικών Μελετών και ο Ρόλος της Στατιστικής

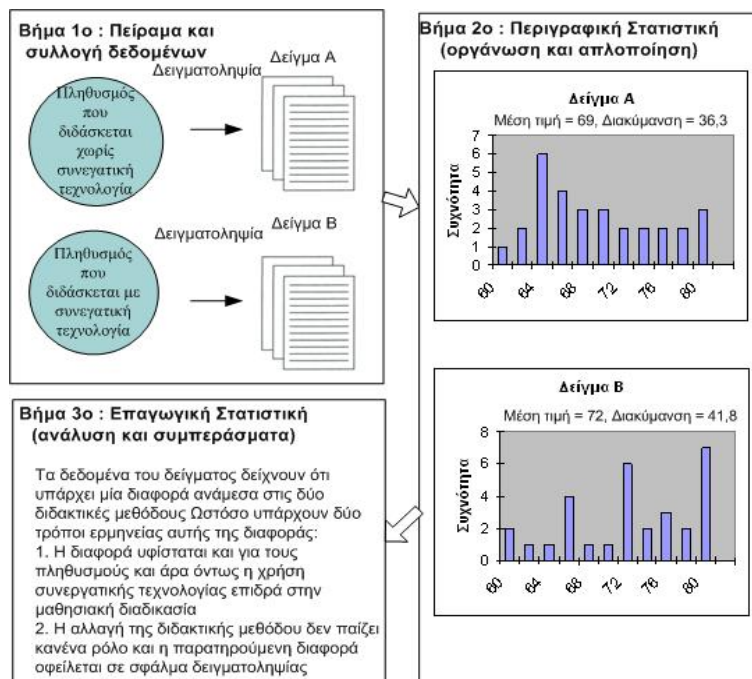
Σε αυτήν την ενότητα περιγράφονται ορισμένα τυπικά σενάρια μελετών και προβληματισμών που μπορεί να συναντήσει ένας ερευνητής ο οποίος μελετά την επίδραση της χρήσης συνεργατικών τεχνολογιών σε εργασιακά και

μαθησιακά περιβάλλοντα. Τα σενάρια αυτά θα μας βοηθήσουν να κατανοήσουμε το ρόλο της στατιστικής στην επιστημονική έρευνα.

- *Σενάριο 1^ο*: ένας καθηγητής Πανεπιστημίου διδάσκει Μαθηματικά για 20 χρόνια. Χρησιμοποιώντας παλαιότερες βαθμολογίες φοιτητών έχει υπολογίσει ότι η βαθμολογία τους ακολουθεί κανονική κατανομή με μέση τιμή 6 και διακύμανση 22. Ο καθηγητής θέλει τώρα να εξετάσει αν η χρήση συστημάτων ασύγχρονης τηλεκπαίδευσης θα επηρεάσει θετικά τη βαθμολογία των φετινών φοιτητών του. Ο καθηγητής χρησιμοποιεί τυχαίο δείγμα 15 μαθητών
- *Σενάριο 2^ο*: το σενάριο αυτό αποτελεί μία παραλλαγή του πρώτου. Τώρα ο καθηγητής θέλει να εξετάσει την επίδραση της χρήσης ασύγχρονης τηλεκπαίδευσης στο μάθημα των Μαθηματικών για όλα τα τμήματα του Πανεπιστημίου. Επειδή ο πληθυσμός τώρα είναι πολύ μεγαλύτερος επιλέγει ένα δείγμα 100 φοιτητών. Από τις βαθμολογίες όλων των καθηγητών που παρακολουθούν Μαθηματικά γνωρίζει μόνο ότι ο μέσος όρος της βαθμολογίας των φοιτητών είναι 6.
- *Σενάριο 3^ο*: ένας ερευνητής θέλει να εξετάσει εάν μια νέα διδακτική μέθοδος για μαθητές Γυμνασίου, η οποία χρησιμοποιεί συνεργατική τεχνολογία, επιδρά στη μαθησιακή διαδικασία. Ο ερευνητής επιλέγει τυχαία δύο δείγματα 30 μαθητών και εφαρμόζει στο πρώτο δείγμα την κλασική μέθοδο διδασκαλίας, ενώ στο δεύτερο δείγμα τη νέα μέθοδο διδασκαλίας. Για να διαπιστώσει την επίδραση της αλλαγής διδακτικής μεθόδου (ανεξάρτητη μεταβλητή) στη μαθησιακή διαδικασία, υποβάλλει τους μαθητές κάθε δείγματος σε ένα βαθμολογημένο τεστ με άριστα το 100, και καταγράφει τις βαθμολογίες τους (εξαρτημένη μεταβλητή). Επιπρόσθετα, από τα δεδομένα που έχει συλλέξει υπολογίζει ότι η διακύμανση του πρώτου δείγματος είναι 36,3, ενώ του δεύτερου είναι 41,8. Στο Σχήμα 22.3 παρουσιάζεται συνοπτικά αυτή η μελέτη και ο ρόλος που διαδραματίζει η στατιστική.
- *Σενάριο 4^ο*: το σενάριο αυτό αποτελεί μία παραλλαγή του προηγούμενου. Τώρα, επειδή ο ερευνητής πιστεύει ότι οι νοητικές ικανότητες κάθε μαθητή μπορεί να αλλοιώσουν τα αποτελέσματα, δεν επιλέγει με τελείως τυχαίο τρόπο τα δύο δείγματα των 30 μαθητών, αλλά χρησιμοποιεί ένα τεστ IQ, και για κάθε μαθητή με ένα συγκεκριμένο δείκτη IQ που τοποθετεί στο πρώτο δείγμα, τοποθετεί και ένα ακόμη με τον ίδιο δείκτη IQ στο δεύτερο δείγμα, δηλαδή "ζευγαρώνει" τα υποκείμενα που ανήκουν στα δείγματα, χρησιμοποιώντας ως κριτήριο τον δείκτη IQ τους.
- *Σενάριο 5^ο*: ένας μάνατζερ θέλει να εξετάσει αν η χρήση χώρων σύσκεψης με υποστήριξη από υπολογιστικά συστήματα επιδρά στο χρόνο λήψης της

τελικής απόφασης. Για το σκοπό αυτό επιλέγει ένα αντιπροσωπευτικό δείγμα 30 εργαζομένων, και τους τοποθετεί τη μία μέρα σε ένα "κλασικό" χώρο σύσκεψης, και την επόμενη μέρα σε ένα χώρο σύσκεψης που υποστηρίζεται από υπολογιστικά συστήματα, χρονομετρώντας κάθε φορά τη διαδικασία λήψης της τελικής απόφασης.

- *Σενάριο 6^ο*: ένας ερευνητής θέλει να μελετήσει την επίδραση τριών διαφορετικών μεθόδων διδασκαλίας Μαθηματικών στην ικανότητα επίλυσης προβλημάτων από φοιτητές. Η πρώτη μέθοδος διδασκαλίας είναι η "κλασική" μέθοδος διδασκαλίας, όπου ο καθηγητής διδάσκει ένα σύνολο φοιτητών. Στη δεύτερη μέθοδο οι φοιτητές αναλαμβάνουν με μία προκαθορισμένη σειρά να παρουσιάσουν την ύλη κάθε διάλεξης, και ο καθηγητής παρακολουθεί και επεμβαίνει όποτε κρίνει ότι είναι απαραίτητο. Στην τρίτη μέθοδο ο καθηγητής παρουσιάζει τη διάλεξη χρησιμοποιώντας υποστηρικτικά ασύρματες τεχνολογίες συνεργασίας. Ο ερευνητής επιλέγει τρία δείγματα 50 φοιτητών, και εφαρμόζει σε καθένα μία από τις τρεις διαφορετικές μεθόδους διδασκαλίας. Στο τέλος του εξαμήνου χρησιμοποιεί για κάθε δείγμα ένα τεστ μαθηματικών και συγκεντρώνει τις βαθμολογίες όλων των φοιτητών. Με τα δεδομένα αυτά υπολογίζει ότι οι διακυμάνσεις των τριών δειγμάτων είναι 30, 45 και 40 αντίστοιχα.
- *Σενάριο 7^ο*: ένας ερευνητής θέλει να κάνει μία έρευνα για να διαπιστώσει ποια από τις τέσσερις κατηγορίες συνεργατικής αλληλεπίδρασης (ασύγχρονη-από απόσταση, σύγχρονη-από απόσταση, σύγχρονη-τοπική, ασύγχρονη-τοπική) είναι η πιο δημοφιλής. Για το σκοπό αυτό επιλέγει ένα τυχαίο δείγμα 1000 ατόμων και χρησιμοποιεί ένα ερωτηματολόγιο στο οποίο τα υποκείμενα δηλώνουν την προτίμησή τους.



Σχήμα 22.3 Ο Ρόλος της Στατιστικής στην Επιστημονική Έρευνα

22.3 Επαγωγική Στατιστική, Έλεγχος Υποθέσεων

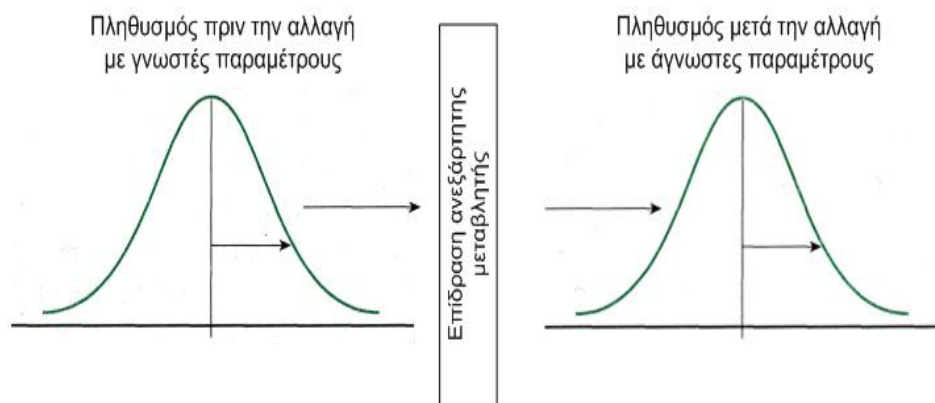
Συνήθως όταν ένας ερευνητής σχεδιάζει κάποιο πείραμα έχει κάποια θεωρία που θέλει να επαληθεύσει, και επομένως κάποια ένδειξη για τα τελικά αποτελέσματα. Αυτή η πρόβλεψη του ερευνητή για το τελικό αποτέλεσμα ονομάζεται στην *στατιστική υπόθεση* (hypothesis). Ουσιαστικά μία υπόθεση αποτελεί την πρόβλεψη του ερευνητή για την επίδραση της αλλαγής της ανεξάρτητης μεταβλητής στην εξαρτημένη. Ο έλεγχος υποθέσεων (hypothesis testing) αποτελεί ένα σημαντικό κομμάτι της επαγωγικής στατιστικής και ένα πολύ χρήσιμο εργαλείο για τον ερευνητή, καθώς του επιτρέπει να ελέγξει την εγκυρότητα της θεωρίας του.

22.3.1 Η Λογική του Ελέγχου Υποθέσεων

Αν και όπως θα δούμε στη συνέχεια υπάρχουν πολλές παραλλαγές της μεθόδου ελέγχου υποθέσεων για κάθε ερευνητική περίπτωση, η βασική λογική που ακολουθείται είναι πάντα η ίδια (Σχήμα 22.4). Αξίζει να σημειωθεί ότι συνήθως ένας ερευνητής μελετά πληθυσμούς οι οποίοι προσεγγίζουν μία

κανονική κατανομή (*normal distribution*), δηλαδή οι περισσότερες παρατηρήσεις συγκεντρώνονται γύρω από τη μέση τιμή, και ο αριθμός τους μειώνεται συμμετρικά και προς τις δύο μεριές (Σχήμα 22.4).

Η λογική του ελέγχου υποθέσεων μπορεί να συνοψιστεί στα εξής τέσσερα βήματα: διατύπωση υποθέσεων, διαμόρφωση των κριτηρίων για την λήψη μιας απόφασης, συλλογή δεδομένων από δείγματα και αξιολόγηση της μηδενικής υπόθεσης.



Σχήμα 22.4 Η Λογική του Ελέγχου Υποθέσεων

Βήμα 1^ο: Διατύπωση των Υποθέσεων

Το πρώτο βήμα σε έναν έλεγχο υποθέσεων είναι να διατυπωθεί η *μηδενική υπόθεση* H_0 (*null hypothesis*). Η μηδενική υπόθεση είναι μία πρόταση που προβλέπει ότι η αλλαγή της ανεξάρτητης μεταβλητής (μεταβλητή που ελέγχω) δεν έχει καμία επίδραση στην εξαρτημένη μεταβλητή (μεταβλητή που παρατηρώ). Με αυτόν τον τρόπο μπορούμε να θεωρήσουμε γνωστές τις παραμέτρους του νέου πληθυσμού που δημιουργείται από την αλλαγή της τιμής της ανεξάρτητης μεταβλητής, καθώς υποθέτουμε ότι αυτή η αλλαγή δεν έχει καμία επίδραση. Επομένως, οι τιμές των παραμέτρων του πληθυσμού μετά την αλλαγή της ανεξάρτητης μεταβλητής είναι ίδιες με αυτές πριν την αλλαγή.

Η δεύτερη υπόθεση που διατυπώνει ο ερευνητής ονομάζεται *εναλλακτική υπόθεση* H_1 (*alternative hypothesis*), και είναι η ακριβώς αντίθετη από τη μηδενική υπόθεση, δηλαδή δηλώνει ότι η αλλαγή της ανεξάρτητης μεταβλητής επιδρά στην εξαρτημένη μεταβλητή. Ο στόχος της μεθόδου ελέγχου υποθέσεων είναι να αποδείξει ότι η μηδενική υπόθεση μπορεί να απορριφθεί με κάποιο βαθμό βεβαιότητας.

Βήμα 2°: Διαμόρφωση των Κριτηρίων για τη Λήψη μιας Απόφασης

Ο ερευνητής τελικά θα συλλέξει κάποια δεδομένα από δείγματα για να αξιολογήσει την αξιοπιστία της μηδενικής υπόθεσης. Με άλλα λόγια, θα χρησιμοποιήσει στατιστικούς δείκτες για να βγάλει συμπεράσματα για την τιμή μιας παραμέτρου του πληθυσμού. Ωστόσο, όπως προαναφέρθηκε, όσο αντιπροσωπευτικό και αν είναι το δείγμα, πάντα θα υπάρχει μια ασυμφωνία (σφάλμα δειγματοληψίας) ανάμεσα στην πραγματική τιμή της παραμέτρου και του στατιστικού δείκτη.

Το ερώτημα λοιπόν που προκύπτει είναι αν η διαφορά ανάμεσα στην τιμή του στατιστικού δείκτη και στην τιμή που προβλέπει η μηδενική υπόθεση για την παράμετρο του πληθυσμού οφείλεται πραγματικά στην επίδραση της ανεξάρτητης μεταβλητής, ή οφείλεται απλά σε σφάλματα δειγματοληψίας. Για να απαντήσει σε αυτό το ερώτημα ο ερευνητής χρειάζεται να θέσει κάποια κριτήρια που καθορίζουν επακριβώς πόση διαφορά χρειάζεται να υφίσταται ανάμεσα σε αυτές τις δύο τιμές, ώστε να δικαιολογείται η απόφαση να απορριφθεί η μηδενική υπόθεση. Αν τελικά αποφασίσει να απορρίψει τη μηδενική υπόθεση, θα πρέπει να καθορίσει το βαθμό βεβαιότητας με τον οποίο προβαίνει σε μία τέτοια απόφαση, ή με άλλα λόγια την πιθανότητα τα αποτελέσματα της ανάλυσής του να είναι εσφαλμένα. Η πιθανότητα αυτή ονομάζεται *επίπεδο σημαντικότητας ή επίπεδο α (level of significance or alpha level)*, και καθορίζει τη μέγιστη πιθανότητα το αποτέλεσμα μιας στατιστικής ανάλυσης να οφείλεται σε σφάλματα ή τυχαίους παράγοντες.

Βήμα 3°: Συλλογή Δεδομένων από Δείγματα

Το επόμενο βήμα στη μέθοδο ελέγχου υποθέσεων είναι η συλλογή των δεδομένων από το δείγμα και ο υπολογισμός των κατάλληλων περιγραφικών στατιστικών δεικτών. Είναι εξαιρετικά σημαντικό να επιλεγεί ένα αντιπροσωπευτικό δείγμα, ώστε η μέθοδος να παρέχει αξιόπιστα αποτελέσματα. Όπως ήδη αναφέρθηκε, η μέθοδος της τυχαίας δειγματοληψίας εξασφαλίζει αυτή την απαίτηση. Σε αυτό το σημείο αξίζει να σημειωθεί ότι η διαδικασία της συλλογής των δεδομένων από το δείγμα γίνεται αφού ο ερευνητής διαμορφώσει τα κριτήρια για την λήψη μιας απόφασης, ώστε να είναι αμερόληπτος και ανεπηρέαστος από τα δεδομένα αυτά.

Βήμα 4°: Αξιολόγηση της Μηδενικής Υπόθεσης

Στο τελευταίο βήμα της μεθόδου ελέγχου υποθέσεων ο ερευνητής μελετά την εγκυρότητα της μηδενικής υπόθεσης με τη βοήθεια στατιστικών δεικτών, ώστε να λάβει μια απόφαση σύμφωνα με τα κριτήρια που έχει θέσει στο δεύτερο βήμα. Υπάρχουν λοιπόν δύο δυνατότητες, είτε αποφασίζει να *απορρίψει τη μηδενική υπόθεση (reject the null hypothesis)* όταν τα αποτελέσματα που προκύπτουν από το δείγμα είναι *σημαντικά (significantly)*

διαφορετικά από αυτά που προβλέπει η μηδενική υπόθεση, είτε οδηγείται στην απόφαση ότι *απέτυχε να απορρίψει τη μηδενική υπόθεση (fail to reject the null hypothesis)* όταν τα αποτελέσματα του πειράματος δεν παρέχουν επαρκείς αποδείξεις ότι η μηδενική υπόθεση δεν ισχύει.

Ο ερευνητής για να οδηγηθεί σε μία από τις δύο παραπάνω αποφάσεις χρησιμοποιεί κάποιο *στατιστικό δείκτη ελέγχου (test statistic)*. Αν και όπως θα δούμε στη συνέχεια υπάρχουν διάφοροι στατιστικοί δείκτες ελέγχου (βλέπε ενότητα 22.5) ανάλογα με τις διαθέσιμες πληροφορίες, ο στόχος τους σε όλες τις περιπτώσεις είναι να καθορίσουν αν και κατά πόσο τα αποτελέσματα του πειράματος οφείλονται αποκλειστικά και μόνο σε τυχαίους παράγοντες και σφάλματα (π.χ. σφάλμα δειγματοληψίας, ανεξέλεγκτες μεταβλητές). Με άλλα λόγια, κάθε στατιστικός δείκτης ελέγχου διαμορφώνει τον εξής λόγο:

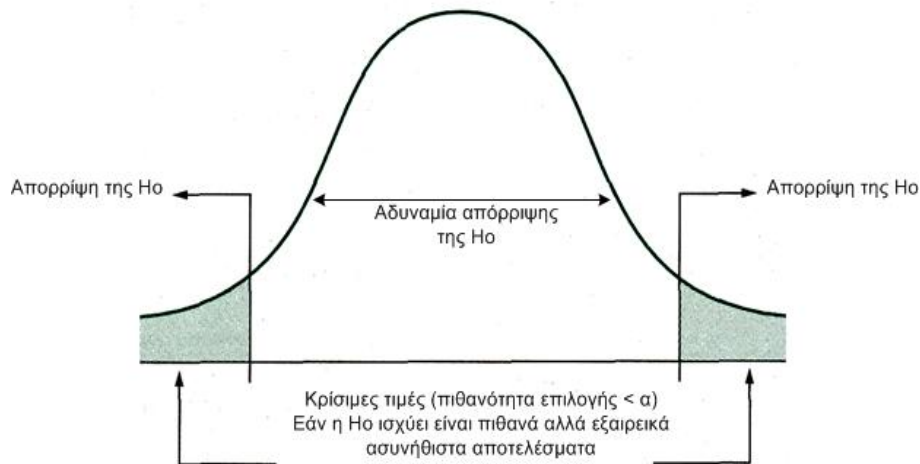
$$\text{στατιστικό ελέγχου} = \frac{\text{αποτελέσματα οφείλονται στην επίδραση της ανεξάρτητης μεταβλητής}}{\text{αποτελέσματα οφείλονται σε τυχαίους παράγοντες και σφάλματα}}$$

Έτσι, οποιαδήποτε τιμή για το στατιστικό δείκτη ελέγχου είναι μεγαλύτερη από μονάδα υποδηλώνει ότι η πιθανότητα λήψης των συγκεκριμένων αποτελεσμάτων είναι μεγαλύτερη από την πιθανότητα αυτά τα δεδομένα να οφείλονται σε τυχαίους παράγοντες. Ωστόσο, ένας ερευνητής θέλει να αποδείξει ότι η πιθανότητα λήψης αυτών των αποτελεσμάτων είναι σημαντικά μεγαλύτερη, και όχι απλά μεγαλύτερη, από το να είναι τυχαία. Αυτό το επίπεδο σημαντικότητας καθορίζεται από το επίπεδο α .

Για παράδειγμα, αν ο ερευνητής έχει καθορίσει επίπεδο σημαντικότητας $\alpha=0.05$, αυτό σημαίνει ότι θέλει να είναι σίγουρος κατά 95% ότι τα συμπεράσματα του δεν είναι εσφαλμένα. Αν και η ακριβής τιμή κάθε στατιστικού δείκτη ελέγχου διαφέρει, σε γενικές γραμμές για $\alpha=0.05$ θέλουμε τα αποτελέσματα μας να εμφανίζονται με διπλάσια πιθανότητα από την πιθανότητα να οφείλονται απλά στην τύχη ή σε σφάλματα (για $\alpha=0.001$ η πιθανότητα γίνεται περίπου τριπλάσια). Στις επιστημονικές μελέτες η μεγαλύτερη αποδεκτή τιμή για το α είναι 0.05, δηλαδή η μέγιστη αποδεκτή πιθανότητα τα αποτελέσματα του ερευνητή να είναι εσφαλμένα είναι 5%.

Στην *Σχήμα 22.5* φαίνεται η κατανομή όλων των δυνατών μέσων τιμών των παρατηρήσεων που μπορούν να ληφθούν από τα δείγματα (*sample means distribution*) εάν ισχύει η μηδενική υπόθεση. Η κατανομή αυτή είναι κανονική, καθώς, εφόσον ισχύει η μηδενική υπόθεση, τα περισσότερα δείγματα αναμένεται να έχουν την ίδια μέση τιμή με τον αρχικό πληθυσμό, αφού η αλλαγή της ανεξάρτητης μεταβλητής δεν έχει καμία επίδραση. Είναι φανερό ότι οι περιοχές στις οποίες μπορεί να απορριφθεί η μηδενική υπόθεση (ονομάζονται και *κρίσιμες περιοχές*) αντιστοιχούν σε μεγάλες τιμές του στατιστικού δείκτη ελέγχου, οι οποίες εξασφαλίζουν ότι τα αποτελέσματα που

έλαβε ο ερευνητής είναι αρκετά ασυνήθιστα (έχουν μεγάλη διαφορά από τη μέση τιμή του αρχικού πληθυσμού) για το επίπεδο σημαντικότητας που έχει θέσει. Για την εύρεση της τιμής που ορίζει τις κρίσιμες περιοχές της κατανομής, ο ερευνητής χρειάζεται να ανατρέξει στον κατάλληλο στατιστικό πίνακα σύμφωνα με το στατιστικό δείκτη ελέγχου που χρησιμοποιεί.



Σχήμα 22.5 Η Κατανομή των Μέσων Τιμών των Δειγμάτων (όλα τα δυνατά αποτελέσματα) όταν η Μηδενική Υπόθεση H_0 ισχύει

22.3.2 Κατευθυνόμενος και Μη-κατευθυνόμενος Έλεγχος Υποθέσεων

Ο στόχος σε μία διαδικασία ελέγχου υποθέσεων, όπως έχει παρουσιαστεί μέχρι στιγμής, είναι να εξετάσει αν η αλλαγή της ελεγχόμενης μεταβλητής έχει οποιαδήποτε επίδραση στην παρατηρούμενη μεταβλητή. Ωστόσο, πολλές φορές ο ερευνητής έχει κάποιο "προαίσθημα" για τα τελικά αποτελέσματα, και δε θέλει απλά να εξετάσει αν η ελεγχόμενη μεταβλητή επιδρά στην παρατηρούμενη, αλλά θέλει να εξετάσει και τον τρόπο με τον οποίο επιδρά. Σε αυτές τις περιπτώσεις διατυπώνει μία διαφορετική μηδενική πρόταση, που δηλώνει ότι η παρατηρούμενη μεταβλητή θα μειωθεί εάν έχει προαίσθημα ότι θα αυξηθεί, και αντίστοιχα ότι θα αυξηθεί εάν έχει προαίσθημα ότι θα μειωθεί.⁶⁹ Αυτό το είδος ελέγχου υποθέσεων ονομάζεται *κατευθυνόμενος*

⁶⁹ Είναι προφανές ότι τα πειράματα γίνονται για να αποδείξουμε ότι η αλλαγή μιας ελεγχόμενης μεταβλητής επιδρά σε μία παρατηρούμενη μεταβλητή, δηλαδή να αποδείξουμε ότι ισχύει η εναλλακτική υπόθεση. Ωστόσο, στον έλεγχο υποθέσεων προσπαθούμε να αποδείξουμε ότι δεν ισχύει το αντίθετο, δηλαδή να απορρίψουμε την μηδενική υπόθεση. Η επιλογή αυτή οφείλεται στους περιορισμούς της επαγωγικής λογικής, και στο γεγονός ότι γενικά είναι πιο εύκολο να αποδείξουμε ότι κάτι δεν ισχύει παρά να αποδείξουμε ότι ισχύει.

έλεγχος υποθέσεων (*directional or one-tailed hypothesis tests*), σε αντιδιαστολή με το μη-κατευθυνόμενο έλεγχο υποθέσεων (*non-directional or two-tailed hypothesis tests*) ο οποίος εξετάζει απλά την επίδραση της ελεγχόμενης μεταβλητής στην παρατηρούμενη.

Ένα χαρακτηριστικό παράδειγμα μη-κατευθυνόμενου ελέγχου υποθέσεων αποτελεί το πρώτο σενάριο που παρουσιάστηκε στην ενότητα 22.3.5, όπου ο καθηγητής επιθυμεί απλά να εξετάσει την επίδραση της χρήσης συστημάτων ασύγχρονης τηλεκπαίδευσης στη βαθμολογία των φοιτητών στα Μαθηματικά. Ωστόσο, με μία μικρή τροποποίηση το σενάριο αυτό μπορεί εύκολα να μετατραπεί σε σενάριο κατευθυνόμενου ελέγχου υποθέσεων. Εάν για παράδειγμα ο καθηγητής είχε διαβάσει από παρόμοιες μελέτες ότι η εφαρμογή τεχνολογιών ασύγχρονης τηλεκπαίδευσης επιδρά θετικά στη βαθμολογία των φοιτητών, θα μπορούσε να διατυπώσει σαν μηδενική υπόθεση μία πρόταση της μορφής: "Η εφαρμογή των τεχνολογιών ασύγχρονης τηλεκπαίδευσης επιδρά αρνητικά στη βαθμολογία των φοιτητών στα Μαθηματικά", και να προσπαθήσει να την απορρίψει στη συνέχεια χρησιμοποιώντας τη μέθοδο του κατευθυνόμενου ελέγχου υποθέσεων.

Ωστόσο, χρειάζεται να τονιστεί ότι ο κατευθυνόμενος έλεγχος υποθέσεων είναι γενικά λιγότερο αξιόπιστος από τον μη-κατευθυνόμενο (για το ίδιο επίπεδο α), και μπορεί να οδηγήσει σε εσφαλμένη απόρριψη της μηδενικής υπόθεσης. Ο κατευθυνόμενος έλεγχος υποθέσεων πρέπει να χρησιμοποιείται μόνο όταν υπάρχει μία θεωρία ή άλλα πειράματα που να υποστηρίζουν το προαίσθημα του ερευνητή. Αντίθετα, ο μη-κατευθυνόμενος έλεγχος υποθέσεων μπορεί να χρησιμοποιηθεί σε οποιαδήποτε περίπτωση εξίσου αποτελεσματικά.

22.4 Τυπικές Στατιστικές Μέθοδοι Ελέγχου Υποθέσεων

Είναι φανερό ότι ο έλεγχος υποθέσεων αποτελεί ένα απαραίτητο εργαλείο για τον ερευνητή, καθώς του επιτρέπει να βγάλει γενικά συμπεράσματα για πληθυσμούς χρησιμοποιώντας πληροφορίες που προέρχονται από δείγματα. Η επαγωγική στατιστική περιλαμβάνει μία πληθώρα στατιστικών μεθόδων που βασίζονται στον έλεγχο υποθέσεων και χρησιμοποιούνται σε διαφορετικές ερευνητικές περιστάσεις. Σε κάθε ερευνητική μελέτη οι πληροφορίες που χρειάζεται να λάβουμε υπόψη για την επιλογή της κατάλληλης μεθόδου ελέγχου υποθέσεων είναι ο στόχος του πειράματος, το πλήθος των ανεξάρτητων μεταβλητών και των διαφορετικών τιμών τους, το πλήθος των εξαρτημένων μεταβλητών, ο αριθμός των διαθέσιμων δειγμάτων, η κατηγορία των δεδομένων, και τέλος το είδος του στατιστικού ελέγχου σε συνδυασμό με τις προϋποθέσεις που πρέπει να ικανοποιούνται για τη χρήση του.

Ο στόχος αυτής της ενότητας είναι να παρουσιάσει τις πιο σημαντικές στατιστικές μεθόδους ώστε να γνωρίζει ο αναγνώστης ποια μέθοδο πρέπει να χρησιμοποιήσει κάθε φορά, πότε μπορεί να την χρησιμοποιήσει, τον τρόπο εφαρμογής της, αλλά και τους περιορισμούς της. Σε αυτό το σημείο αξίζει να σημειωθεί ότι στην ενότητα 22.5.8 περιλαμβάνεται ένας συνοπτικός οδηγός για τη διαδικασία και τους τύπους υπολογισμού των στατιστικών μεθόδων που παρουσιάζονται σε αυτήν την ενότητα.

22.4.1 Το z Στατιστικό Τεστ

Το z στατιστικό τεστ (z-test) χρησιμοποιεί δεδομένα που προέρχονται από ένα δείγμα για να ελέγξει υποθέσεις που σχετίζονται με την μέση τιμή ενός πληθυσμού όταν η διακύμανση του αρχικού πληθυσμού είναι γνωστή. Σε αυτή τη στατιστική μέθοδο η μηδενική υπόθεση δηλώνει μία συγκεκριμένη τιμή για τη μέση τιμή του άγνωστου πληθυσμού ο οποίος προκύπτει από την αλλαγή της ανεξάρτητης μεταβλητής. Για την εύρεση των κρίσιμων περιοχών ο ερευνητής επιλέγει ένα επίπεδο σημαντικότητας και συμβουλευεται το στατιστικό πίνακα τυπικής κανονικής κατανομής (unit normal table – βλέπε παράρτημα κεφαλαίου), ενώ για τον υπολογισμό της τιμής του στατιστικού δείκτη ελέγχου z, ο οποίος τον βοηθάει να αξιολογήσει την μηδενική υπόθεση, χρησιμοποιεί τις εξής μαθηματικές σχέσεις:

$$z = \frac{\bar{X} - \mu}{\sigma_{\bar{x}}}, \text{ με } \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

Το z στατιστικό τεστ μπορεί να χρησιμοποιηθεί μόνο όταν τηρούνται ορισμένες προϋποθέσεις:

- τα υποκείμενα του δείγματος έχουν επιλεγεί με τυχαίο τρόπο ώστε το δείγμα να είναι αντιπροσωπευτικό. Η προϋπόθεση αυτή ικανοποιείται όταν έχει χρησιμοποιηθεί τυχαία δειγματοληψία για την απόκτηση του δείγματος
- οι τιμές που απαρτίζουν το δείγμα οφείλονται σε ανεξάρτητες παρατηρήσεις. Δύο παρατηρήσεις είναι ανεξάρτητες όταν δεν υπάρχει μία προβλέψιμη σχέση μεταξύ τους. Και αυτή η προϋπόθεση ικανοποιείται όταν χρησιμοποιείται η τυχαία δειγματοληψία
- η διακύμανση του πληθυσμού δε μεταβάλλεται από την επίδραση της ανεξάρτητης μεταβλητής. Αν και αυτή η συνθήκη δεν ικανοποιείται πλήρως σε πραγματικά πειράματα, συνήθως οι ερευνητές υποθέτουν ότι ισχύει για να εφαρμόσουν το z στατιστικό τεστ.
- η κατανομή των τιμών του δείγματος είναι κανονική. Αυτή η συνθήκη ικανοποιείται όταν το δείγμα προέρχεται από ένα πληθυσμό με κανονική κατανομή (όπως τις περισσότερες φορές), ή όταν το μέγεθος του δείγματος

είναι αρκετά μεγάλο (τυπικά μεγαλύτερο του 30). Έτσι, ο ερευνητής αν έχει αμφιβολία ότι ο πληθυσμός ακολουθεί κανονική κατανομή, το μόνο που χρειάζεται να κάνει είναι να επιλέξει ένα μεγαλύτερο δείγμα.⁷⁰

Μία χαρακτηριστική περίπτωση ερευνητικής μελέτης στην οποία αρμόζει η χρήση του στατιστικού τεστ z είναι το πρώτο σενάριο που παρουσιάστηκε στην ενότητα 22.3.5. Στο σενάριο αυτό ο καθηγητής έχει στη διάθεση του τη διακύμανση του αρχικού πληθυσμού, και θέλει να εξετάσει την επίδραση της χρήσης ασύγχρονης τηλεκπαίδευσης στο μέσο όρο της βαθμολογίας των φοιτητών στα Μαθηματικά. Παρατηρούμε ότι ικανοποιούνται όλες οι απαραίτητες προϋποθέσεις, καθώς χρησιμοποιείται τυχαία δειγματοληψία για την απόκτηση του δείγματος, και ο καθηγητής γνωρίζει ότι ο αρχικός πληθυσμός των βαθμολογιών ακολουθεί κανονική κατανομή. Έτσι μπορεί να διατυπώσει σαν μηδενική υπόθεση την πρόταση: "Η χρήση ασύγχρονης τηλεκπαίδευσης δεν επιδρά στη βαθμολογία των φοιτητών στα Μαθηματικά", και να προχωρήσει σε ένα μη-κατευθυνόμενο έλεγχο υποθέσεων με τη χρήση του z στατιστικού τεστ ώστε να την αξιολογήσει.

22.4.2 Το t Στατιστικό Τεστ ενός Δείγματος

Το t στατιστικό τεστ ενός δείγματος (*single sample t-test*) χρησιμοποιεί δεδομένα που προέρχονται από ένα δείγμα για να ελέγξει υποθέσεις που σχετίζονται με τη μέση τιμή ενός πληθυσμού όταν η διακύμανση του αρχικού πληθυσμού είναι άγνωστη. Σε αυτή τη στατιστική μέθοδο γίνεται μία εκτίμηση της διακύμανσης του πληθυσμού με τη βοήθεια της διακύμανσης του δείγματος. Η μηδενική υπόθεση δηλώνει και πάλι μία συγκεκριμένη τιμή για την μέση τιμή του άγνωστου πληθυσμού που προκύπτει από την αλλαγή της ανεξάρτητης μεταβλητής. Για την εύρεση των κρίσιμων περιοχών ο ερευνητής επιλέγει ένα επίπεδο σημαντικότητας, υπολογίζει του βαθμούς ελευθερίας που δίνονται από τον τύπο $df = n - 1$ (όπου n το μέγεθος τους δείγματος), και συμβουλευεται το στατιστικό πίνακα κατανομής t (*t-distribution table* – βλέπε παράρτημα κεφαλαίου).

Ο υπολογισμός της τιμής του στατιστικού δείκτη ελέγχου t βασίζεται στις παρακάτω μαθηματικές σχέσεις:

$$t = \frac{\bar{X} - \mu}{s_{\bar{x}}}, \text{ με } s_{\bar{x}} = \frac{s}{\sqrt{n}}$$

Το t στατιστικό τεστ ενός δείγματος μπορεί να χρησιμοποιηθεί μόνο όταν τηρούνται ορισμένες προϋποθέσεις:

⁷⁰ Υπάρχουν διάφορα στατιστικά τεστ για τον έλεγχο της κανονικότητας μιας κατανομής αλλά δεν θα μας απασχολήσουν στα πλαίσια μιας σύντομης εισαγωγής στην Στατιστική.

- το δείγμα είναι αντιπροσωπευτικό και οι τιμές που το απαρτίζουν οφείλονται σε ανεξάρτητες παρατηρήσεις (ενότητα 22.5.1)
- η κατανομή των τιμών του δείγματος είναι κανονική (ενότητα 22.5.1)

Μία χαρακτηριστική περίπτωση ερευνητικής μελέτης στην οποία αρμόζει η χρήση του στατιστικού τεστ t ενός δείγματος είναι το δεύτερο σενάριο που παρουσιάστηκε στην ενότητα 22.3.5. Στο σενάριο αυτό ο καθηγητής δεν έχει στην διάθεση του τη διακύμανση του αρχικού πληθυσμού, και θέλει και πάλι να εξετάσει την επίδραση της χρήσης ασύγχρονης τηλεκπαίδευσης στο μέσο όρο της βαθμολογίας των φοιτητών στα Μαθηματικά. Παρατηρούμε ότι ικανοποιούνται όλες οι απαραίτητες προϋποθέσεις, καθώς χρησιμοποιείται τυχαία δειγματοληψία για την απόκτηση του δείγματος, και το μέγεθος του δείγματος είναι αρκετά μεγάλο (>30). Έτσι ο ερευνητής μπορεί να διατυπώσει σαν μηδενική υπόθεση την πρόταση: "Η χρήση ασύγχρονης τηλεκπαίδευσης δεν επιδρά στην βαθμολογία των φοιτητών στα Μαθηματικά", και να προχωρήσει σε ένα μη-κατευθυνόμενο έλεγχο υποθέσεων με τη χρήση του t στατιστικού τεστ ενός δείγματος ώστε να την αξιολογήσει.

22.4.3 Το t Στατιστικό Τεστ Ανεξάρτητων Δειγμάτων

Το t στατιστικό τεστ ανεξάρτητων δειγμάτων (independent measures t -test) χρησιμοποιεί δεδομένα που προέρχονται από δύο ξεχωριστά δείγματα για να ελέγξει υποθέσεις που σχετίζονται με την διαφορά των μέσων τιμών δύο πληθυσμών. Σε αυτή την περίπτωση η μηδενική υπόθεση δηλώνει ότι δεν υφίσταται διαφορά ανάμεσα στις μέσες τιμές των δύο πληθυσμών. Για τον υπολογισμό της τιμής του στατιστικού δείκτη ελέγχου t και του βαθμού ελευθερίας που καθορίζει τις κρίσιμες περιοχές στο στατιστικό πίνακα κατανομής t , χρησιμοποιούνται οι εξής μαθηματικές σχέσεις:

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{s_{\bar{x}-\bar{y}}}, \text{ με } s_{\bar{x}-\bar{y}} = \sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}} \text{ και } s_p^2 = \sqrt{\frac{SS_1 + SS_2}{df_1 + df_2}}$$

$$df = df_1 + df_2 = (n_1 - 1) + (n_2 - 1) = n_1 + n_2 - 2$$

Οι προϋποθέσεις που πρέπει να ισχύουν για να χρησιμοποιηθεί το t στατιστικό τεστ ανεξάρτητων δειγμάτων είναι:

- τα δείγματα είναι αντιπροσωπευτικά και οι τιμές που τα απαρτίζουν οφείλονται σε ανεξάρτητες παρατηρήσεις (ενότητα 22.5.1)
- η κατανομή των τιμών των δειγμάτων είναι κανονική (ενότητα 22.5.1)
- οι δύο πληθυσμοί από τους οποίους έχουν επιλεγεί τα δύο δείγματα έχουν την ίδια διακύμανση (δηλαδή $\sigma_1 \approx \sigma_2$). Τυπικά αυτή η προϋπόθεση

ικανοποιείται όταν η τιμή της μεγαλύτερης διακύμανσης ενός δείγματος είναι το πολύ διπλάσια από την τιμή της μικρότερης διακύμανσης ενός άλλου δείγματος. Αν και αυτός ο κανόνας είναι εμπειρικός, μπορεί να δώσει μια ικανοποιητική εικόνα για τον αν παραβιάζεται ή όχι η συγκεκριμένη προϋπόθεση. Ωστόσο, αν ο ερευνητής επιθυμεί μία καλύτερη εικόνα μπορεί να χρησιμοποιήσει το F στατιστικό τεστ το οποίο είναι σχεδιασμένο για αυτόν ακριβώς τον σκοπό (στα πλαίσια ενός εισαγωγικού κεφαλαίου δε θα μας απασχολήσει).

Σαν παράδειγμα ερευνητικής μελέτης στην οποία αρμόζει η χρήση του στατιστικού τεστ t ανεξάρτητων δειγμάτων θα χρησιμοποιήσουμε το τρίτο σενάριο που παρουσιάστηκε στην ενότητα 22.3.5. Στο σενάριο αυτό ο ερευνητής χρησιμοποιεί δύο ανεξάρτητα δείγματα για να εξετάσει την επίδραση της χρήσης συνεργατικής τεχνολογίας στη μαθησιακή διαδικασία μαθητών Γυμνασίου. Παρατηρούμε ότι ικανοποιούνται όλες οι απαραίτητες προϋποθέσεις, καθώς χρησιμοποιείται τυχαία δειγματοληψία για την απόκτηση του δείγματος, το μέγεθος του δείγματος είναι αρκετά μεγάλο ($n=30$), και η τιμή της μεγαλύτερης διακύμανσης, δηλαδή του ενός δείγματος ($s_1 = 41,8$), είναι μόλις 1.1 φορές μεγαλύτερη από την τιμή της μικρότερης διακύμανσης, δηλαδή του άλλου δείγματος ($s_2 = 36,3$). Έτσι ο ερευνητής μπορεί να διατυπώσει τη μηδενική υπόθεση: "Η χρήση συνεργατικής τεχνολογίας δεν επιδρά στη μαθησιακή διαδικασία των μαθητών Γυμνασίου", και να προχωρήσει σε ένα μη-κατευθυνόμενο έλεγχο υποθέσεων με τη χρήση του t στατιστικού τεστ ανεξάρτητων δειγμάτων ώστε να την αξιολογήσει.

22.4.4 Το t Στατιστικό Τεστ Εξαρτημένων Δειγμάτων

Εύκολα γίνεται αντιληπτό ότι η χρήση ανεξάρτητων δειγμάτων δεν είναι ο μοναδικός τρόπος για να σχεδιαστεί ένα πείραμα το οποίο εξετάζει την επίδραση μιας ελεγχόμενης μεταβλητής σε μία παρατηρούμενη. Μια διαφορετική προσέγγιση αποτελεί η χρήση ενός μόνο δείγματος υποκειμένων, το οποίο εξετάζεται σε όλες τις διαφορετικές τιμές της ελεγχόμενης μεταβλητής. Με άλλα λόγια, επαναλαμβάνοντας τις μετρήσεις στο ίδιο σύνολο υποκειμένων και για διαφορετικές συνθήκες, προσπαθούμε να εντοπίσουμε διαφορές στη συμπεριφορά τους. Οι ερευνητικές μελέτες που βασίζονται σε πειράματα σχεδιασμένα με αυτή την προσέγγιση ονομάζονται *μελέτες επαναλαμβανόμενων μετρήσεων (repeated-measures studies)*.

Το σημαντικό πλεονέκτημα αυτών των μελετών, συγκριτικά με τις μελέτες που βασίζονται σε πειράματα με ανεξάρτητα δείγματα, είναι το γεγονός ότι δεν επιτρέπουν στη διαφορετικότητα των υποκειμένων (σε επίπεδο μόρφωσης, κοινωνικότητας, κλπ) να επιδράσει στα αποτελέσματα της μελέτης, αφού χρησιμοποιούν το ίδιο σύνολο υποκειμένων σε όλα τα πειράματα. Επίσης, οι μελέτες επαναλαμβανόμενων μετρήσεων, λόγω του τρόπου σχεδιασμού τους,

χρειάζονται μικρότερο συνολικό αριθμό υποκειμένων, κάτι που μπορεί να είναι πολύ σημαντικό σε ορισμένες περιπτώσεις. Το μόνο μειονέκτημα αυτών των μελετών είναι ότι, επειδή ακριβώς χρησιμοποιούν τον ίδιο αριθμό υποκειμένων σε διαφορετικά πειράματα, υπάρχει ο κίνδυνος να επιδράσουν άλλοι ανεπιθύμητοι παράγοντες (*carry-over effects and progressive error*) στα αποτελέσματα, όπως για παράδειγμα η κούραση.

Το t στατιστικό τεστ με επαναλαμβανόμενες μετρήσεις (repeated-measures t-test) χρησιμοποιείται για τον έλεγχο υποθέσεων σε μελέτες επαναλαμβανόμενων μετρήσεων. Το τεστ αυτό χρησιμοποιεί δύο σύνολα μετρήσεων που προέρχονται από ένα δείγμα για να ελέγξει υποθέσεις που σχετίζονται με τη διαφορά των μέσων τιμών τους. Σε αυτήν την περίπτωση, η μηδενική υπόθεση δηλώνει ότι δεν υφίσταται διαφορά ανάμεσα στις μέσες τιμές των δύο συνόλων μετρήσεων. Για την απλοποίηση των μαθηματικών τύπων που επιτρέπουν τον υπολογισμό της τιμής του στατιστικού ελέγχου t, συνήθως εισάγεται ο όρος "διαφορά μετρήσεων D" (difference score D) που εκφράζει για κάθε υποκείμενο τη διαφορά ανάμεσα στις μετρήσεις κάθε πειράματος. Οι μαθηματικές σχέσεις που χρειάζονται σε ένα t στατιστικό τεστ με επαναλαμβανόμενες μετρήσεις είναι οι εξής:

$$t = \frac{\bar{D} - \mu_D}{s_{\bar{D}}}, \text{ με } D = X_2 - X_1 \text{ και } \bar{D} = \frac{\sum D}{n} \text{ και } s_{\bar{D}} = \sqrt{\frac{s^2}{n}}$$

$$df = n - 1$$

Οι προϋποθέσεις που πρέπει να ισχύουν για να χρησιμοποιηθεί το t στατιστικό τεστ με επαναλαμβανόμενες μετρήσεις είναι:

- το δείγμα είναι αντιπροσωπευτικό και οι τιμές οφείλονται σε ανεξάρτητες παρατηρήσεις (ενότητα 22.5.1)
- η κατανομή των τιμών του D είναι κανονική.

Σαν παράδειγμα ερευνητικής μελέτης στην οποία αρμόζει η χρήση του t στατιστικού τεστ με επαναλαμβανόμενες μετρήσεις θα χρησιμοποιήσουμε το πέμπτο σενάριο που παρουσιάστηκε στην ενότητα 22.3.5. Στο σενάριο αυτό, ο μάνατζερ χρησιμοποιεί ένα δείγμα για να εξετάσει την επίδραση της χρήσης χώρων σύσκεψης που υποστηρίζονται από υπολογιστικά συστήματα στο χρόνο λήψης μιας απόφασης. Παρατηρούμε ότι ικανοποιούνται όλες οι απαραίτητες προϋποθέσεις, καθώς χρησιμοποιείται τυχαία δειγματοληψία για την απόκτηση του δείγματος, και το μέγεθος του δείγματος είναι αρκετά μεγάλο (=30). Έτσι, ο ερευνητής μπορεί να διατυπώσει τη μηδενική υπόθεση: "Η χρήση χώρων σύσκεψης που υποστηρίζονται από υπολογιστικά συστήματα δεν επιδρά στον χρόνο λήψης μιας απόφασης", και να προχωρήσει σε ένα μη-κατευθυνόμενο

έλεγχου υποθέσεων με τη χρήση του t στατιστικού τεστ επαναλαμβανόμενων μετρήσεων ώστε να την αξιολογήσει.

Ορισμένες φορές, ο ερευνητής μπορεί να ανησυχεί ότι μία ανεξέλεγκτη μεταβλητή μπορεί να αλλοιώσει σημαντικά τα αποτελέσματά του. Για να το αποφύγει αυτό, μπορεί να σχεδιάσει με διαφορετικό τρόπο τα πειράματά του, χρησιμοποιώντας την τεχνική του ζευγαρώματος υποκειμένων (*matching subjects*). Σε μία μελέτη με ζευγαρωμένα υποκείμενα ο ερευνητής τοποθετεί τα υποκείμενα σε δύο δείγματα, ώστε σε κάθε δείγμα να υπάρχουν υποκείμενα με παρόμοια χαρακτηριστικά. Για το t στατιστικό τεστ με ζευγαρωμένα υποκείμενα χρησιμοποιούνται οι ίδιες μαθηματικές σχέσεις με το t στατιστικό τεστ με επαναλαμβανόμενες μετρήσεις, και πρέπει να ικανοποιούνται οι ίδιες προϋποθέσεις. Αυτό που διαφέρει είναι ο τρόπος με τον οποίο σχεδιάζεται το πείραμα, και κατ' επέκταση ο αριθμός δειγμάτων.

Χαρακτηριστική περίπτωση ερευνητικής μελέτης στην οποία αρμόζει η χρήση του t στατιστικού τεστ με ζευγαρωμένα υποκείμενα αποτελεί το τέταρτο σενάριο που παρουσιάστηκε στην ενότητα 22.3.5. Στο σενάριο αυτό ο ερευνητής χρησιμοποιεί δύο δείγματα ζευγαρωμένων (ως προς το IQ) υποκειμένων, για να εξετάσει την επίδραση της χρήσης συνεργατικής τεχνολογίας στη μαθησιακή διαδικασία μαθητών Γυμνασίου. Παρατηρούμε ότι ικανοποιούνται όλες οι απαραίτητες προϋποθέσεις, καθώς χρησιμοποιείται τυχαία δειγματοληψία για την απόκτηση του δείγματος, και το μέγεθός του είναι αρκετά μεγάλο ($=30$). Έτσι, ο ερευνητής μπορεί να διατυπώσει σαν μηδενική υπόθεση την πρόταση: "Η χρήση συνεργατικής τεχνολογίας δεν επιδρά στη μαθησιακή διαδικασία των μαθητών Γυμνασίου", και να προχωρήσει σε ένα μη-κατευθυνόμενο έλεγχο υποθέσεων με τη χρήση του t στατιστικού τεστ ζευγαρωμένων υποκειμένων ώστε να την αξιολογήσει.

Τέλος, αξίζει να σημειωθεί ότι επειδή και τα στατιστικά τεστ με επαναλαμβανόμενες μετρήσεις και τα στατιστικά τεστ με ζευγαρωμένα υποκείμενα δημιουργούν κάποια μορφή σχέσης ανάμεσα στις μετρήσεις, συνήθως αναφέρονται με τον όρο *στατιστικά τεστ εξαρτημένων δειγμάτων* (*related-samples tests*).

22.4.5 Ανάλυση Διακύμανσης

Όλα τα στατιστικά τεστ που έχουμε εξετάσει μέχρι τώρα χρησιμοποιούν ένα ή δύο δείγματα για να εξετάσουν υποθέσεις για δύο τιμές της ελεγχόμενης μεταβλητής. Αν και αυτές οι στατιστικές μέθοδοι είναι οι πιο συχνά χρησιμοποιούμενες, υπάρχουν και περιπτώσεις στις οποίες ο ερευνητής θέλει να εξετάσει την επίδραση δύο ή περισσότερων εξαρτημένων μεταβλητών με δύο ή περισσότερες δυνατές τιμές στην παρατηρούμενη μεταβλητή. Η *ανάλυση*

διακύμανσης (*ANalysis Of VAriance-ANOVA*) εξυπηρετεί αυτόν ακριβώς το σκοπό.

Σε αυτό το σημείο μπορεί να αναρωτηθεί κανείς γιατί χρειαζόμαστε μία καινούρια στατιστική μέθοδο για να ελέγξουμε την επίδραση δύο ή περισσοτέρων τιμών της ανεξάρτητης μεταβλητής, αφού θα μπορούσαμε κάλλιστα να εξετάσουμε την επίδραση όλων αυτών των δυνατών τιμών ανά δύο, χρησιμοποιώντας όσες φορές χρειάζεται μία από τις στατιστικές μεθόδους που παρουσιάστηκαν προηγουμένως. Ωστόσο, εύκολα γίνεται αντιληπτό ότι η χρήση πολλαπλών στατιστικών τεστ αυξάνει τη συνολική πιθανότητα να οδηγηθούμε σε λανθασμένα συμπεράσματα, και για αυτό χρησιμοποιείται η ανάλυση διακύμανσης σε αυτές τις περιπτώσεις.⁷¹

Όπως και τα *t* στατιστικά τεστ, η ανάλυση διακύμανσης μπορεί να χρησιμοποιηθεί είτε με ανεξάρτητα δείγματα είτε με εξαρτημένα δείγματα. Επίσης, η ανάλυση διακύμανσης μπορεί να χρησιμοποιηθεί και σε πειράματα που ο ερευνητής θέλει να ελέγχει δύο ή περισσότερες ανεξάρτητες μεταβλητές, οι οποίες στην ορολογία της ανάλυσης διακύμανσης ονομάζονται *παράγοντες (factors)*. Επειδή αυτές οι ερευνητικές περιπτώσεις αυξάνουν σημαντικά το βαθμό πολυπλοκότητας της στατιστικής ανάλυσης και ξεφεύγουν από τους στόχους αυτού του κεφαλαίου, θα περιγράψουμε στη συνέχεια μόνο την *ανάλυση διακύμανσης ενός παράγοντα (single-factor ANOVA)*. Ο αναγνώστης που ενδιαφέρεται και για τις υπόλοιπες μεθόδους μπορεί να ανατρέξει στη βιβλιογραφία που αναφέρθηκε στην εισαγωγή.

Η ανάλυση διακύμανσης ενός παράγοντα χρησιμοποιείται για τον έλεγχο υποθέσεων σε μελέτες ανεξάρτητων δειγμάτων. Ο ερευνητής επιλέγει ένα δείγμα για κάθε διαφορετική τιμή της ανεξάρτητης μεταβλητής, και ελέγχει υποθέσεις που συγκρίνουν τις μέσες τιμές των δειγμάτων αυτών. Η μηδενική υπόθεση δηλώνει ότι δεν υφίσταται καμία διαφορά ανάμεσα σε όλες τις μέσες τιμές των δειγμάτων. Είναι εμφανές λοιπόν, ότι στην περίπτωση που ο ερευνητής απορρίψει τελικά τη μηδενική υπόθεση, το μόνο που μπορεί να ισχυριστεί είναι ότι τα δείγματα διαφέρουν μεταξύ τους, αλλά δεν είναι σε θέση να γνωρίζει ποια συγκεκριμένα δείγματα διαφέρουν. Αν θέλει να εξακριβώσει τέτοιου είδους πληροφορίες μπορεί να χρησιμοποιήσει μετά την ανάλυση διακύμανσης ειδικά στατιστικά τεστ, τα οποία ονομάζονται *post hoc tests*, και δε θα παρουσιαστούν στα πλαίσια αυτού του εισαγωγικού κεφαλαίου.

⁷¹ Θυμηθείτε ότι σε κάθε τεστ υπάρχει μία πιθανότητα τα συμπεράσματα να είναι εσφαλμένα, η οποία εκφράζεται από το συντελεστή σημαντικότητας α . Αν χρησιμοποιήσουμε διαδοχικά τεστ, τότε ο συνολικός συντελεστής σημαντικότητας ισούται με το άθροισμα των επιμέρους συντελεστών του κάθε τεστ, και επομένως αυξάνεται η συνολική πιθανότητα να οδηγηθούμε σε εσφαλμένα συμπεράσματα

Η ανάλυση διακύμανσης ενός παράγοντα χρησιμοποιεί λίγο διαφορετική σημειολογία σε σχέση με τις στατιστικές μεθόδους που έχουν παρουσιαστεί μέχρι στιγμής, η οποία δικαιολογείται από την αύξηση της πολυπλοκότητας. Το γράμμα k χρησιμοποιείται για να προσδιορίσει τον αριθμό των δυνατών τιμών της ανεξάρτητης μεταβλητής ($k \geq 2$), το T για να προσδιορίσει το άθροισμα των μετρήσεων κάθε δείγματος ($T = \sum X$), το N για να προσδιορίσει το συνολικό αριθμό υποκειμένων ($N = \sum n$), και το G για να προσδιορίσει το συνολικό άθροισμα των μετρήσεων όλων των δειγμάτων ($G = \sum T$). Η ανάλυση διακύμανσης χρησιμοποιεί την *κατανομή F* (*F-distribution* – βλέπε παράρτημα κεφαλαίου) για την εύρεση των κρίσιμων τιμών, και τους παρακάτω μαθηματικούς τύπους για τον υπολογισμό του στατιστικού δείκτη ελέγχου F :

$$SS_{\text{between_treatments}} = \sum \frac{T^2}{n} - \frac{G^2}{N}, \quad df_{\text{between_treatments}} = k - 1,$$

$$SS_{\text{within_treatments}} = \sum_N SS, \quad df_{\text{within_treatments}} = N - k,$$

$$F = \frac{\frac{SS_{\text{between_treatments}}}{df_{\text{between_treatments}}}}{\frac{SS_{\text{within_treatments}}}{df_{\text{within_treatments}}}} = \frac{MS_{\text{between_treatments}}}{MS_{\text{within_treatments}}}$$

Οι προϋποθέσεις που πρέπει να ισχύουν για να χρησιμοποιηθεί η ανάλυση διακύμανσης ενός παράγοντα είναι:

- τα δείγματα είναι αντιπροσωπευτικά και οι τιμές που τα απαρτίζουν οφείλονται σε ανεξάρτητες παρατηρήσεις (ενότητα 22.5.1),
- η κατανομή των τιμών των δειγμάτων είναι κανονική (ενότητα 22.5.1), και
- οι πληθυσμοί από τους οποίους έχουν επιλεγεί τα δείγματα έχουν την ίδια διακύμανση (ενότητα 22.5.3).

Σαν παράδειγμα ερευνητικής μελέτης στην οποία αρμόζει η χρήση της ανάλυσης διακύμανσης ενός παράγοντα θα εξετάσουμε το έκτο σενάριο που παρουσιάστηκε στην ενότητα 22.3.5. Στο σενάριο αυτό ο ερευνητής χρησιμοποιεί τρία δείγματα για να εξετάσει την επίδραση τριών διαφορετικών τρόπων διδασκαλίας στη βαθμολογία των φοιτητών στα Μαθηματικά. Παρατηρούμε ότι ικανοποιούνται όλες οι απαραίτητες προϋποθέσεις, καθώς χρησιμοποιείται τυχαία δειγματοληψία για την απόκτηση των δειγμάτων, το μέγεθός τους είναι αρκετά μεγάλο (>30), και μπορεί να θεωρηθεί ότι οι διακυμάνσεις των πληθυσμών δε διαφέρουν σημαντικά σύμφωνα με τα εμπειρικά κριτήρια της ενότητας 22.5.3. Έτσι, ο ερευνητής μπορεί να

διατυπώσει τη μηδενική υπόθεση: "Δεν υπάρχει διαφορά στην επίδραση των τριών μεθόδων διδασκαλίας στη βαθμολογία των φοιτητών στα Μαθηματικά", και να προχωρήσει σε μία ανάλυση διακύμανσης ενός παράγοντα ώστε να την αξιολογήσει.

22.4.6 Το χ^2 Στατιστικό Τεστ

Τα στατιστικά τεστ που έχουν παρουσιάζονται παραπάνω ανήκουν στην κατηγορία των *παραμετρικών τεστ (parametric tests)*. Τα παραμετρικά τεστ χρησιμοποιούν ποσοτικά δεδομένα και μπορούν να εφαρμοστούν μόνο όταν ικανοποιούνται ορισμένες προϋποθέσεις. Τι γίνεται όμως αν ο ερευνητής έχει στη διάθεση του ποιοτικά δεδομένα, ή δεν ικανοποιούνται οι προϋποθέσεις των παραμετρικών τεστ; Για αυτές τις περιπτώσεις η επαγωγική στατιστική παρέχει ένα άλλο σύνολο από στατιστικά τεστ ελέγχου υποθέσεων, τα οποία ονομάζονται *μη-παραμετρικά (non-parametric tests)*. Όταν λοιπόν ένα παραμετρικό τεστ δε μπορεί να χρησιμοποιηθεί, συνήθως υπάρχει ένα μη-παραμετρικό τεστ που μπορεί να το αντικαταστήσει. Αν και τα μη-παραμετρικά τεστ μπορούν να χρησιμοποιηθούν και σε περιπτώσεις που είναι δυνατή η εφαρμογή ενός παραμετρικού τεστ, συνήθως αυτό αποφεύγεται γιατί τα παραμετρικά τεστ δίνουν πιο αξιόπιστα αποτελέσματα.

Το στατιστικό τεστ χ^2 (Chi-Square test) είναι ίσως το πιο δημοφιλές μη-παραμετρικό τεστ. Αν και υπάρχουν διάφορες παραλλαγές του, ανάλογα με το διαθέσιμο αριθμό δειγμάτων και το στόχο της ανάλυσης, εμείς θα εξετάσουμε την πιο απλή μορφή του. Σε αυτές τις περιπτώσεις, ο ερευνητής έχει στη διάθεση του ένα δείγμα ποιοτικών δεδομένων οργανωμένο σε ονομαστικές κατηγορίες, και ο στόχος είναι να χρησιμοποιηθούν τα δεδομένα αυτά ώστε να προσδιοριστεί η αναλογία (ή το ποσοστό) του πληθυσμού που ανήκει στην κάθε κατηγορία. Για την επίτευξη αυτού του στόχου διατυπώνεται μια μηδενική υπόθεση, που είτε δηλώνει ότι δεν υπάρχει κάποια συγκεκριμένη προτίμηση στις διαθέσιμες ονομαστικές κατηγορίες (*no-preference null-hypothesis*), είτε δηλώνει ότι τα ποσοστά που προτιμώνται από τα υποκείμενα δε διαφέρουν από τα ποσοστά άλλων πληθυσμών οι οποίοι αποτελούν το σημείο αναφοράς (*no-difference from a comparison population*).

Και στις δύο περιπτώσεις, αυτό που προσδιορίζει η μηδενική υπόθεση είναι ο αναμενόμενος αριθμός (expected frequency – f_e) των υποκειμένων που ανήκει σε κάθε ονομαστική κατηγορία. Ο έλεγχος υποθέσεων που ακολουθεί αξιολογεί αυτή τη μηδενική υπόθεση, συγκρίνοντας τον αριθμό των υποκειμένων που αναμένεται σε κάθε ονομαστική κατηγορία με τον αριθμό των υποκειμένων που παρατηρείται ότι ανήκει σε κάθε ονομαστική κατηγορία (observed frequency - f_o), με βάση τις μετρήσεις του δείγματος. Ο στατιστικός δείκτης ελέγχου που χρησιμοποιείται για αυτή την αξιολόγηση είναι το χ^2 . Οι κρίσιμες τιμές για το στατιστικό δείκτη ελέγχου χ^2 δίνονται από την κατανομή

χ^2 (βλέπε παράρτημα κεφαλαίου), ενώ οι μαθηματικές σχέσεις υπολογισμού του είναι οι εξής:

$$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e}, \text{ όπου } f_e \text{ και } f_o \text{ η αναμενόμενη και η παρατηρούμενη} \\ \text{συχνότητα}$$

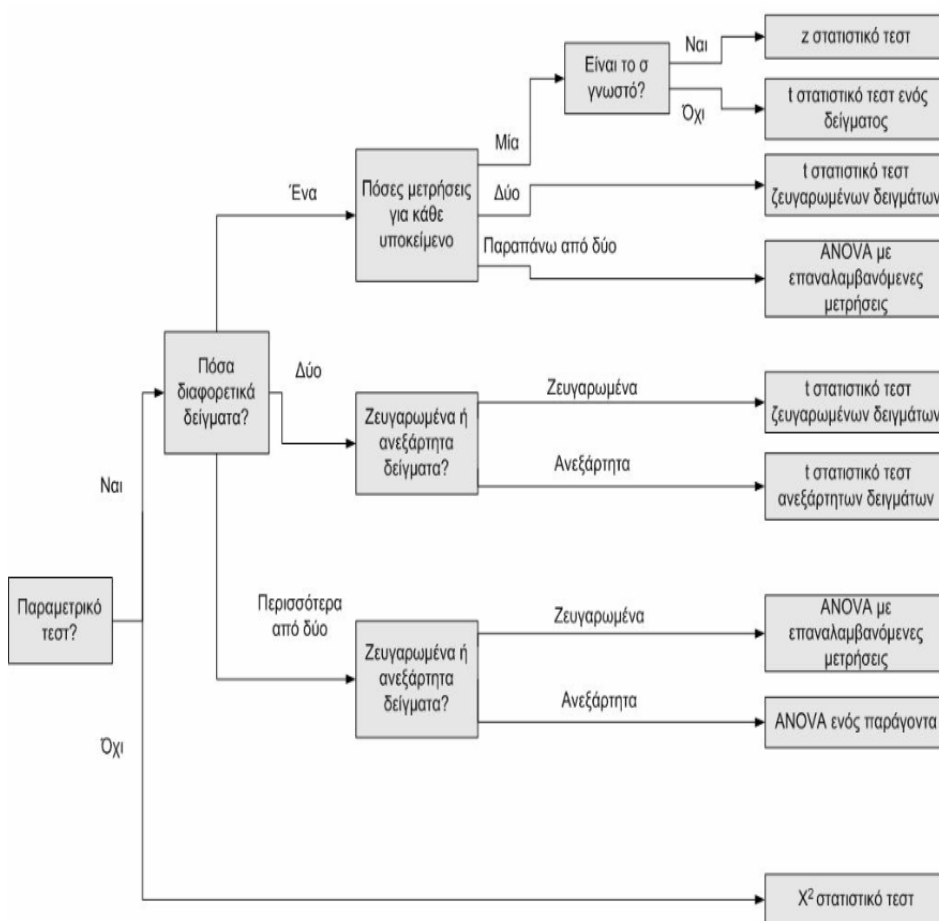
$$df = C - 1, \text{ όπου } C \text{ ο αριθμός των ονομαστικών κατηγοριών}$$

Αν και, όπως αναφέραμε, τα μη-παραμετρικά τεστ δε χρειάζεται να ικανοποιούν συγκεκριμένες προϋποθέσεις, τα συμπεράσματα που προκύπτουν από την εφαρμογή του στατιστικού τεστ χ^2 είναι πιο αξιόπιστα όταν χρησιμοποιείται αντιπροσωπευτικό δείγμα, και η αναμενόμενη συχνότητα όλων των κατηγοριών είναι μεγαλύτερη από πέντε.

Σαν παράδειγμα ερευνητικής μελέτης στην οποία αρμόζει η χρήση του στατιστικού τεστ χ^2 θα εξετάσουμε το έβδομο σενάριο που παρουσιάστηκε στην ενότητα 22.3.5. Στο σενάριο αυτό ο ερευνητής επιθυμεί να μάθει ποιος τρόπος συνεργατικής αλληλεπίδρασης είναι περισσότερο δημοφιλής (4 ονομαστικές κατηγορίες). Για το σκοπό αυτό χρησιμοποιεί ένα αρκετά μεγάλο δείγμα ανθρώπων, καθώς ο υπό εξέταση πληθυσμός είναι επίσης μεγάλος (όλοι οι άνθρωποι του πλανήτη). Έτσι, ο ερευνητής μπορεί να διατυπώσει σαν μηδενική υπόθεση την πρόταση: "Δεν υπάρχει κάποια ιδιαίτερη προτίμηση σε έναν από τους τρόπους συνεργατικής αλληλεπίδρασης", και να προχωρήσει σε ένα στατιστικό τεστ χ^2 ώστε να την αξιολογήσει.

22.4.7 Διαδικασία και Τύποι Υπολογισμού Μεθόδων Ελέγχου Υποθέσεων

Στη συνέχεια παρατίθεται ένα διάγραμμα αποφάσεων που μπορεί να βοηθήσει τον αναγνώστη να επιλέξει το πιο κατάλληλο στατιστικό τεστ ελέγχου υποθέσεων, και ένας πίνακας που συνοψίζει τη διαδικασία υπολογισμού των στατιστικών δεικτών ελέγχου που παρουσιάστηκαν στις προηγούμενες ενότητες.



Σχήμα 22.6 Διάγραμμα Αποφάσεων για την Επιλογή μιας Μεθόδου Ελέγχου Υποθέσεων

Τεστ ελέγχου υποθέσεων	Πότε χρησιμοποιείται	Στατιστικό ελέγχου	Κρίσιμες περιοχές
z στατιστικό τεστ	Υποθέσεις για την μέση τιμή ενός πληθυσμού με γνωστή διακύμανση όταν έχω ένα δείγμα	$z = \frac{\bar{X} - \mu}{\sigma_{\bar{x}}}, \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$	Τυπική κανονική κατανομή
t στατιστικό τεστ ενός δείγματος	Υποθέσεις για την μέση τιμή ενός πληθυσμού με άγνωστη διακύμανση όταν έχω ένα δείγμα	$t = \frac{\bar{X} - \mu}{s_{\bar{x}}}, s_{\bar{x}} = \frac{s}{\sqrt{n}}$	Κατανομή t $df = n - 1$

Τεστ ελέγχου υποθέσεων	Πότε χρησιμοποιείται	Στατιστικό ελέγχου	Κρίσιμες περιοχές
t στατιστικό τεστ ανεξάρτητων δειγμάτων	Υποθέσεις για την διαφορά των μέσων τιμών δύο πληθυσμών όταν έχω δύο δείγματα	$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{s_{\bar{x}-\bar{y}}}$ $s_{\bar{x}-\bar{y}} = \sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}$ $s_p^2 = \sqrt{\frac{SS_1 + SS_2}{df_1 + df_2}}$	Κατανομή t $df = n_1 + n_2 - 2$
t στατιστικά τεστ εξαρτημένων δειγμάτων	<p>(με επαναλαμβανόμενες μετρήσεις)</p> <p>Υποθέσεις για την διαφορά των μέσων τιμών δύο συνόλων μετρήσεων που προκύπτουν από ένα δείγμα με επαναλαμβανόμενες μετρήσεις</p> <p>(με ζευγαρωμένα υποκείμενα)</p> <p>Υποθέσεις για την διαφορά των μέσων τιμών δύο συνόλων μετρήσεων που προκύπτουν από δύο ζευγαρωμένα δείγματα</p>	$t = \frac{\bar{D} - \mu_D}{s_{\bar{D}}}$ $D = X_2 - X_1$ $\bar{D} = \frac{\sum D}{n}$ $s_{\bar{D}} = \sqrt{\frac{s^2}{n}}$	Κατανομή t $df = n - 1$
Ανάλυση διακύμανσης ενός παράγοντα	Υποθέσεις για την διαφορά των μέσων τιμών δύο ή περισσότερων πληθυσμών όταν έχω δύο ή περισσότερα ανεξάρτητα δείγματα	$SS_{between} = \sum \frac{T^2}{n} - \frac{G^2}{N}$ $SS_{within} = \sum SS$ $F = \frac{\frac{SS_{between}}{df_{between}}}{\frac{SS_{within}}{df_{within}}} = \frac{MS_{between}}{MS_{within}}$	Κατανομή F $df_{between} = k - 1$ $df_{within} = N - k$
χ^2 στατιστικό τεστ	Υποθέσεις για την αναλογία ή το ποσοστό ενός πληθυσμού που ανήκει στην κάθε ονομαστική κατηγορία	$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e}$	Κατανομή χ^2 $df = C - 1$

Πίνακας 22.2 Διαδικασία και Τύποι Υπολογισμού Μεθόδων Ελέγχου Υποθέσεων

22.5 Επαγωγική Στατιστική – Συσχέτιση

Οι επαγωγικές στατιστικές διαδικασίες που έχουμε εξετάσει βασίζονται στον προσεκτικό σχεδιασμό πειραμάτων, όπου ο ερευνητής ελέγχει μία μεταβλητή και παρατηρεί κάποια άλλη για να ανακαλύψει αν υπάρχει κάποια επίδραση της ελεγχόμενης μεταβλητής στην παρατηρούμενη. Ωστόσο, αυτές οι διαδικασίες συχνά δεν επαρκούν, καθώς ο ερευνητής μπορεί να μην είναι σε θέση να σχεδιάσει πειράματα που δεν επηρεάζονται από εξωτερικούς παράγοντες, ή μπορεί να επιθυμεί να διερευνήσει και να περιγράψει με ακρίβεια τη σχέση που συνδέει δύο μεταβλητές, και όχι απλά να εξετάσει αν τις συνδέει κάποια σχέση. Η *συσχέτιση (correlation)* είναι μία στατιστική τεχνική που εξυπηρετεί τον ερευνητή σε αυτές τις περιπτώσεις, καθώς του επιτρέπει να παρατηρεί δύο μεταβλητές στο "φυσικό τους περιβάλλον" (μη-ελεγχόμενο περιβάλλον), και να προσδιορίζει και να μετρήσει επακριβώς τη σχέση που τις συνδέει. Σε αυτήν την εισαγωγική ενότητα θα επικεντρώσουμε το ενδιαφέρον μας μόνο στην αναζήτηση γραμμικών σχέσεων ανάμεσα σε μεταβλητές, και επομένως στον υπολογισμό γραμμικών συντελεστών συσχέτισης.

Ο υπολογισμός του συντελεστή συσχέτισης απαιτεί τη μέτρηση δύο μεταβλητών για κάθε υποκείμενο (τα οποία συνήθως συμβολίζονται με X, Y), σε ένα μη-ελεγχόμενο περιβάλλον. Ο στόχος της τεχνικής είναι να προσδιορίσει τη σχέση που συνδέει τις δύο μεταβλητές, υπολογίζοντας την τιμή του συντελεστή συσχέτισης. Για παράδειγμα, ένας ερευνητής μπορεί να θέλει να προσδιορίσει και να ποσοτικοποιήσει τη σχέση που υπάρχει (αν υπάρχει) ανάμεσα στο ύψος και το βάρος των ανθρώπων.

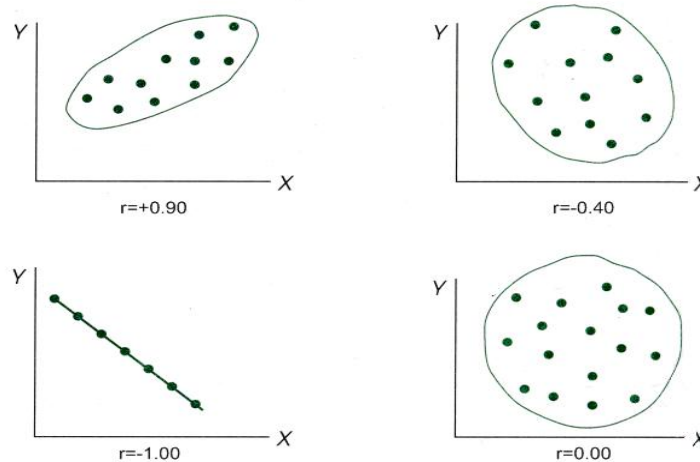
Οι τιμές που μπορεί να πάρει ο συντελεστής συσχέτισης κυμαίνονται από -1 έως $+1$. Αν η τιμή του είναι θετική, τότε οι δύο μεταβλητές τείνουν να μεταβάλλονται προς την ίδια κατεύθυνση, δηλαδή όταν αυξάνεται η μία αυξάνεται και η άλλη. Αν αντίθετα η τιμή του είναι αρνητική, τότε οι μεταβλητές τείνουν να κινούνται προς την αντίθετη κατεύθυνση, δηλαδή όταν η μία αυξάνεται η άλλη μειώνεται. Όσο μεγαλύτερη είναι η τιμή του συντελεστή συσχέτισης (ανεξαρτήτως προσήμου), τόσο μεγαλύτερη είναι η σχέση που συνδέει τις δύο μεταβλητές. Στην ακραία περίπτωση που οι δύο μεταβλητές έχουν συντελεστή συσχέτισης 0 , ο ερευνητής μπορεί να συμπεράνει ότι δεν υπάρχει κάποια σχέση ανάμεσα τους (Σχήμα 22.7).

Αν και υπάρχουν διάφορες τεχνικές υπολογισμού του γραμμικού συντελεστή συσχέτισης ανάλογα με τα διαθέσιμα δεδομένα (ποσοτικά και ποιοτικά), η πιο συχνά χρησιμοποιούμενη μέθοδος είναι ο υπολογισμός του συντελεστή συσχέτισης Pearson (Pearson correlation). Ο συντελεστής συσχέτισης Pearson συμβολίζεται με το γράμμα r , και δίνεται από τη μαθηματική σχέση:

$$r = \frac{SP}{\sqrt{SS_x SS_y}}, \text{ με } SP = \sum XY - \frac{(\sum X)(\sum Y)}{n}$$

Αξίζει να σημειωθεί ότι η τιμή του συντελεστή συσχέτισης υψωμένη στο τετράγωνο (r^2) εκφράζει το ποσοστό της διακύμανσης (αλλαγές στην τιμή) της Y μεταβλητής που μπορεί να προβλεφθεί από τη διακύμανση της X μεταβλητής. Για παράδειγμα, αν ο συντελεστής συσχέτισης έχει υπολογιστεί σε $r=0.80$, τότε το $r^2=0.64=64\%$ των αλλαγών της Y μεταβλητής μπορεί να προβλεφθεί από τις αλλαγές στην X μεταβλητή.

Τέλος, αξίζει να αναφέρουμε ότι όταν ο ερευνητής θέλει να βγάλει συμπεράσματα και να ποσοτικοποιήσει τις σχέσεις ανάμεσα σε παραμέτρους του πληθυσμού, μπορεί να χρησιμοποιήσει το συντελεστή συσχέτισης ως στατιστικό δείκτη ελέγχου σε μία μέθοδο ελέγχου υποθέσεων. Σε αυτήν την περίπτωση, η μηδενική υπόθεση δηλώνει ότι δεν υπάρχει καμία συσχέτιση ανάμεσα στις δύο υπό εξέταση παραμέτρους, και οι κρίσιμες περιοχές ορίζονται από τιμές που βρίσκονται σε στατιστικούς πίνακες με κρίσιμες τιμές για το συντελεστή συσχέτισης Pearson.



Σχήμα 22.7 Παραδείγματα Τιμών της Συσχέτισης και Κατανομή των Μετρήσεων

22.6 Στατιστικά Εργαλεία και Πακέτα Λογισμικού

Ολοκληρώνοντας αυτό το εισαγωγικό κεφάλαιο στη Στατιστική, αξίζει να γίνει μία σύντομη αναφορά σε στατιστικά εργαλεία και πακέτα λογισμικού που διευκολύνουν σε μεγάλο βαθμό τη στατιστική ανάλυση πειραματικών δεδομένων. Υπάρχει πληθώρα διαθέσιμων στατιστικών εργαλείων και πακέτων

λογισμικού, τα οποία καλύπτουν όλες τις ανάγκες μιας στατιστικής ανάλυσης, από τις πιο απλές (όπως οι περιγραφικές και επαγωγικές τεχνικές που αναφέραμε σε αυτό το κεφάλαιο), μέχρι και τις πιο σύνθετες (π.χ. μη γραμμικές συσχετίσεις, παλινδρόμηση, εξόρυξη γνώσης από δεδομένα, κλπ).

Τα πιο γνωστά στατιστικά εργαλεία είναι τα εμπορικά πακέτα λογισμικού (π.χ. SPSS, SAS, Statistica, κλπ) τα οποία προσφέρουν μία μεγάλη ποικιλία στατιστικών αναλύσεων. Αν και τα πακέτα αυτά μπορούν να χρησιμοποιηθούν είτε για γενικού είτε για ειδικού σκοπού στατιστικές αναλύσεις, είναι συνήθως αρκετά πολύπλοκα και ακριβά. Εκτός από αυτά τα ολοκληρωμένα πακέτα λογισμικού, αρκετά διαδεδομένη είναι και η χρήση πρόσθετων εργαλείων (add-ins) σε εμπορικά πακέτα γενικού σκοπού, τα οποία στηρίζονται στη μεταφορά του λογιστικού φύλλου. Τα πρόσθετα εργαλεία (π.χ. πρόσθετα του Ms Excel) συνήθως απαιτούν τη δαπάνη μικρότερων χρηματικών ποσών, αλλά προσφέρουν ένα υποσύνολο των στατιστικών αναλύσεων συγκριτικά με τα εμπορικά πακέτα λογισμικού. Τέλος, αξίζει να σημειωθεί ότι υπάρχουν πολλά στατιστικά εργαλεία στον παγκόσμιο ιστό, τα οποία μπορεί κανείς να προμηθευτεί δωρεάν για να καλύψει τις ανάγκες του. Στον Πίνακα 22.3 γίνεται μία επισκόπηση μερικών επιλεγμένων συστημάτων της κατηγορίας αυτής.

Στατιστικό εργαλείο	Περιγραφή	Πρόσθετες πληροφορίες
CoStat (www.cohort.com/costat.html)	Περιλαμβάνει εργαλεία για τον υπολογισμό περιγραφικών στατιστικών δεικτών, την δημιουργία γραφημάτων, τον έλεγχο υποθέσεων με παραμετρικά και μη παραμετρικά τεστ και τον υπολογισμό συντελεστών συσχέτισης.	Ο μόνος περιορισμός στην έκδοση που διατίθεται δωρεάν είναι ότι δεν επιτρέπει την εκτύπωση και την απευθείας αντιγραφή των δεδομένων. Απαιτεί εγκατάσταση (setup)
Gnumeric (www.gnome.org/projects/gnumeric/)	Περιλαμβάνει πληθώρα τεχνικών περιγραφικής και επαγωγικής στατιστικής ανάλυσης. Προσφέρει τις ίδιες και περισσότερες λειτουργίες με τα πρόσθετα του Ms Excel	Διατίθεται δωρεάν η πλήρως λειτουργική έκδοση χωρίς κανέναν περιορισμό. Απαιτεί εγκατάσταση (setup)
InStat - GraphPad (www.graphpad.com/instat.html)	Ένα πολύ απλό και χρήσιμο στατιστικό εργαλείο υπό μορφή οδηγού τεσσάρων βημάτων. Ενδείκνυται για άτομα που δεν έχουν εξειδικευμένες γνώσεις Στατιστικής. Παρέχει πολύ καλή βοήθεια και παραδείγματα.	Ο μόνος περιορισμός στην έκδοση που διατίθεται δωρεάν είναι ότι δεν επιτρέπει την εκτύπωση και την απευθείας αντιγραφή των δεδομένων. Απαιτεί εγκατάσταση (setup)

Past (http://folk.uio.no/ohamer/past/)	Περιλαμβάνει απλές και προχωρημένες τεχνικές στατιστικής ανάλυσης. Αν και έχει σχεδιαστεί για παλαιοντολόγους μπορεί να χρησιμοποιηθεί εξίσου αποδοτικά από όλους.	Διατίθεται δωρεάν η πλήρως λειτουργική έκδοση χωρίς κανέναν περιορισμό. Δεν απαιτεί εγκατάσταση (εκτελέσιμο αρχείο exe)
SSP (www.economics.pomona.edu/StatSite/SSP.html)	Περιλαμβάνει όλες τις βασικές στατιστικές τεχνικές ανάλυσης δεδομένων.	Διατίθεται δωρεάν η πλήρως λειτουργική έκδοση χωρίς κανέναν περιορισμό. Δεν απαιτεί εγκατάσταση (εκτελέσιμο αρχείο exe)
WebStat (www.staterunch.com/)	Περιλαμβάνει μεγάλη ποικιλία περιγραφικών και επαγωγικών στατιστικών ανάλυσης δεδομένων. Παρέχει πολύ καλή βοήθεια και παραδείγματα. Απαιτεί απλά την ύπαρξη ενός διαφυλλιστή (browser) για να εκτελεστεί.	Διατίθεται δωρεάν η πλήρως λειτουργική έκδοση χωρίς κανέναν περιορισμό. Δεν απαιτεί εγκατάσταση (Web-based εφαρμογή)

Πίνακας 22.3 Επιλεγμένα Στατιστικά Εργαλεία που Διατίθενται Δωρεάν

Σύνοψη

Στο παρόν κεφάλαιο μελετήθηκαν οι βασικές στατιστικές έννοιες, όροι και μέθοδοι που χρησιμοποιούνται στα πλαίσια επιστημονικών ερευνών. Η επιστήμη της στατιστικής αποτελεί ένα αναντικατάστατο εργαλείο για κάθε ερευνητή, καθώς επιτρέπει την οργάνωση μιας συλλογής δεδομένων, τη συνοπτική και αποτελεσματική παρουσίαση τους, την ανάλυση τους, και τέλος την εξαγωγή χρήσιμων και γενικεύσιμων συμπερασμάτων. Στα πλαίσια αυτού του κεφαλαίου γίνεται μια επισκόπηση των βασικών περιγραφικών και επαγωγικών μεθόδων της στατιστικής, ώστε να γνωρίζει ο αναγνώστης ποια μέθοδο πρέπει να χρησιμοποιήσει κάθε φορά, πότε μπορεί να την χρησιμοποιήσει, τον τρόπο εφαρμογής της, αλλά και τους περιορισμούς της. Επιπρόσθετα, δίνεται ιδιαίτερη έμφαση στην περιγραφή παραδειγμάτων ερευνητικών μελετών που σχετίζονται με τη χρήση συνεργατικών τεχνολογιών σε εργασιακά και μαθησιακά περιβάλλοντα, ώστε να γίνουν κατανοητά τα οφέλη της χρήσης στατιστικών μεθόδων σε αυτό το σύνθετο και πολυδιάστατο τομέα επιστημονικής έρευνας. Τέλος, γίνεται μια σύντομη αναφορά σε στατιστικά εργαλεία και πακέτα λογισμικού που διευκολύνουν σημαντικά τη στατιστική ανάλυση πειραματικών δεδομένων.

Ερωτήματα και Θέματα για συζήτηση

1. Δώστε έναν ορισμό της επιστήμης της Στατιστικής.
2. Ποιους κλάδους περιλαμβάνει η Στατιστική, και που αποσκοπεί ο καθένας;
3. Τι εννοούμε με τους όρους πληθυσμός, δείγμα και σφάλμα δειγματοληψίας;

4. Τι κατηγορίες μεταβλητών και δεδομένων χρησιμοποιούμε στη Στατιστική;
5. Περιγράψτε σύντομα τα τυπικά περιγραφικά στατιστικά. Σε ποιες περιπτώσεις χρησιμοποιείται το καθένα;
6. Ποια βήματα περιλαμβάνει ο έλεγχος υποθέσεων; Περιγράψτε τα σύντομα.
7. Τι εννοούμε με τους όρους μηδενική και εναλλακτική υπόθεση;
8. Τι είναι το επίπεδο σημαντικότητας και τι εκφράζει;
9. Τι κριτήρια χρησιμοποιούνται για να καταλήξει σε ασφαλή συμπεράσματα ένας ερευνητής που χρησιμοποιεί τον έλεγχο υποθέσεων;
10. Περιγράψτε σχηματικά τις περιπτώσεις στις οποίες αποτυγχάνουμε να απορρίψουμε τη μηδενική υπόθεση, και εκείνες στις οποίες απορρίπτουμε τη μηδενική υπόθεση.
11. Τι διαφορά έχει ο κατευθυνόμενος έλεγχος από το μη-κατευθυνόμενο;
12. Ποιες είναι οι τυπικές στατιστικές μέθοδοι ελέγχου υποθέσεων, και πότε χρησιμοποιείται η κάθε μία;
13. Πότε χρησιμοποιείται η στατιστική τεχνική της συσχέτισης;

Ασκήσεις

Ένας ερευνητής επιθυμεί να εξετάσει αν οι ενήλικοι χρήστες συστημάτων σύγχρονης συνεργατικής αλληλεπίδρασης συνεργάζονται πιο αποτελεσματικά από ότι τα παιδιά. Για το σκοπό αυτό επιλέγει τυχαία ένα δείγμα 30 παιδιών (δείγμα 1) και 30 ενηλίκων (δείγμα 2). Ο ερευνητής υποθέτει ότι ο χρόνος ολοκλήρωσης μιας απλής εργασίας αποτελεί δείκτη της αποτελεσματικότητας της συνεργασίας των ατόμων. Για να διαπιστώσει λοιπόν την επίδραση της ηλικίας στην αποτελεσματικότητα της συνεργασίας καταγράφει για κάθε δείγμα τον χρόνο ολοκλήρωσης της εργασίας συγκεντρώνοντας τα παρακάτω δεδομένα:

Δείγμα 1 – Χρόνος ολοκλήρωσης της εργασίας (min)

20	25	40	35	25	19	20	22	27	45	50	29	37	41	23
48	38	35	33	31	29	50	26	54	23	33	39	41	32	19

Δείγμα 2 – Χρόνος ολοκλήρωσης της εργασίας (min)

15	17	19	20	26	18	14	16	15	19	20	21	22	25	27
20	17	18	21	22	26	30	33	40	23	27	41	32	33	25

1. Οργανώστε τον τρόπο παρουσίασης των παραπάνω δεδομένων χρησιμοποιώντας τον κατάλληλο τύπο γραφικής παράστασης.
2. Ποιους περιγραφικούς στατιστικούς δείκτες (μέτρα κεντρικής τάσης και μέτρα διασποράς) θα επιλέγατε για τα παραπάνω δεδομένα; Δικαιολογείστε την απάντησή σας και υπολογίστε τους.
3. Αν υποθέσουμε ότι κάποιες από τις παραπάνω μετρήσεις δεν είναι διαθέσιμες (π.χ κάποιος δεν κατάφερε να ολοκληρώσει την εργασία στον μέγιστο διαθέσιμο χρόνο)

τι θα άλλαζε στην επιλογή των περιγραφικών στατιστικών δεικτών; Δικαιολογείστε την απάντησή σας.

4. Ποια είναι η ανεξάρτητη και ποια η εξαρτημένη μεταβλητή;. Διατυπώστε μία μηδενική υπόθεση κατάλληλη για το παραπάνω πείραμα;
5. Εφαρμόστε την κατάλληλη στατιστική μέθοδο ελέγχου υποθέσεων (θεωρήστε επίπεδο σημαντικότητας $\alpha=0.01$). Σε τι συμπέρασμα οδηγείστε;

Ένας ερευνητής ενδιαφέρεται να εξετάσει τους παράγοντες που επηρεάζουν την απόφαση ενός χρήστη κατά την επιλογή και χρήση εικονικών περιβαλλόντων συνεργασίας (βλέπε ενότητα 14.3.5). Για την επίτευξη αυτού του στόχου επιλέγεται και ερωτάται ένα σύνολο από 50 χρήστες τέτοιων συστημάτων. Οι χρήστες καλούνται να επιλέξουν μόνο μία από τις εξής διαθέσιμες επιλογές:

1. Ρεαλισμός των εικονικών χώρων
2. Αριθμός συνδεδεμένων χρηστών
3. Ποικιλία στην επιλογή «προσωποποίησης»
4. Χρόνος εκμάθησης των βασικών λειτουργιών
5. Τα αποτελέσματα της έρευνας συνοψίζονται στον παρακάτω πίνακα:

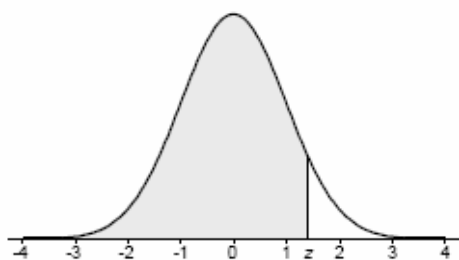
Ρεαλισμός των εικονικών χώρων	Αριθμός χρηστών	Ποικιλία στην επιλογή προσωποποίησης	Χρόνος εκμάθησης λειτουργιών
18	17	7	8

1. Με βάση τα παραπάνω δεδομένα εφαρμόστε την κατάλληλη στατιστική μέθοδο για να διαπιστώσετε ένα κάποιο από τους παραπάνω παράγοντες επιδρά σε μεγαλύτερο βαθμό στην απόφαση ενός χρήστη για την επιλογή και χρήση ενός εικονικού περιβάλλοντος συνεργασίας. (θεωρήστε επίπεδο σημαντικότητας $\alpha=0.05$). Σε τι συμπέρασμα οδηγείστε;
2. Επιλέξτε την κατάλληλη στατιστική μέθοδο ανάλυσης πειραματικών δεδομένων για τις εξής περιπτώσεις ερευνητικών μελετών:
3. Ένας ερευνητής θέλει να εξετάσει αν η ύπαρξη ετερογενούς μαθησιακού υλικού σε περιβάλλοντα συνεργατικής μάθησης επιδρά στην ποιότητα της συνεργασίας. Για το σκοπό αυτό επιλέγει ένα σύνολο 30 μαθητών και διεξάγει δύο πειράματα στα οποία οι μαθητές καλούνται να επιλύσουν το ίδιο πρόβλημα. Στο πρώτο πείραμα όλοι οι μαθητές έχουν πρόσβαση στο ίδιο μαθησιακό υλικό ενώ στο δεύτερο υπάρχει ετερογένεια στο διαθέσιμο μαθησιακό υλικό του κάθε μαθητή. Μετά την ολοκλήρωση του κάθε πειράματος ο ερευνητής χρησιμοποιεί ένα κατάλληλα βαθμολογημένο ερωτηματολόγιο για να καταγράψει έναν ποσοτικό δείκτη για την ποιότητα της συνεργασίας όπως την αντιλαμβάνεται ο κάθε μαθητής.

Ένας ερευνητής θέλει να μελετήσει την επίδραση τριών εναλλακτικών τρόπων σχεδίασης ενός συστήματος συνεργατικής αλληλεπίδρασης στον χρόνο επίλυσης ενός προβλήματος. Για το σκοπό αυτό ο ερευνητής διεξάγει τρία ανεξάρτητα πειράματα σε ένα σύνολο 300 φοιτητών (χωρίστηκαν σε 3 ομάδες με 100 φοιτητές σε κάθε πείραμα). Στα πειράματα αυτά οι φοιτητές καλούνται να επιλύσουν ανά ζεύγη το ίδιο πρόβλημα αλλά η κάθε ομάδα χρησιμοποιεί διαφορετική έκδοση του ίδιου συστήματος συνεργατικής αλληλεπίδρασης. Σε κάθε ένα από τα τρία πειράματα ο ερευνητής καταγράφει το χρονικό διάστημα που χρειάστηκε το κάθε ζευγάρι φοιτητών για την επίλυση του δοθέντος προβλήματος.

Μετά από σειρά ερευνών ένας ερευνητής έχει καταλήξει στο συμπέρασμα ότι ο δείκτης νοημοσύνης ενός ατόμου επιδρά στην ικανότητα του να επιλύει προβλήματα με την χρήση συστημάτων συνεργατικής αλληλεπίδρασης. Στα πλαίσια της τρέχουσας έρευνας του ο μελετητής επιθυμεί να διερευνήσει και να περιγράψει με ακρίβεια αυτή την σχέση. Για το σκοπό αυτό επιλέγει ένα δείγμα 100 ατόμων και τους υποβάλλει αρχικά σε ένα τεστ IQ για να μετρήσει τον δείκτη νοημοσύνης τους. Στην συνέχεια διεξάγει ένα πείραμα στο οποίο οι συμμετέχοντες προσπαθούν να επιλύσουν ένα πρόβλημα σε ομάδες των δύο ατόμων με την βοήθεια ενός συστήματος συνεργατικής αλληλεπίδρασης. Μετά την ολοκλήρωση του πειράματος ο ερευνητής αξιολογεί την ποιότητα της παραγόμενης λύσης και την βαθμολογεί με άριστα το 10.

22.7 Παράρτημα Κεφαλαίου



Significance level

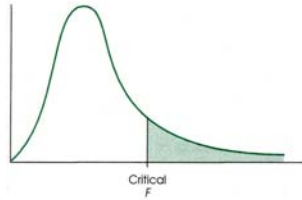
z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999							

Πίνακας 22.4 Κρίσιμες τιμές στην Τυπική Κανονική Κατανομή

df	PROPORTION IN ONE TAIL					
	0.25	0.10	0.05	0.025	0.01	0.005
df	PROPORTION IN TWO TAILS					
	0.50	0.20	0.10	0.05	0.02	0.01
1	1.000	3.078	6.314	12.706	31.821	63.657
2	0.816	1.886	2.920	4.303	6.965	9.925
3	0.765	1.638	2.353	3.182	4.541	5.841
4	0.741	1.533	2.132	2.776	3.747	4.604
5	0.727	1.476	2.015	2.571	3.365	4.032
6	0.718	1.440	1.943	2.447	3.143	3.707
7	0.711	1.415	1.895	2.365	2.998	3.499
8	0.706	1.397	1.860	2.306	2.896	3.355
9	0.703	1.383	1.833	2.262	2.821	3.250
10	0.700	1.372	1.812	2.228	2.764	3.169
11	0.697	1.363	1.796	2.201	2.718	3.106
12	0.695	1.356	1.782	2.179	2.681	3.055
13	0.694	1.350	1.771	2.160	2.650	3.012
14	0.692	1.345	1.761	2.145	2.624	2.977
15	0.691	1.341	1.753	2.131	2.602	2.947
16	0.690	1.337	1.746	2.120	2.583	2.921
17	0.689	1.333	1.740	2.110	2.567	2.898
18	0.688	1.330	1.734	2.101	2.552	2.878
19	0.688	1.328	1.729	2.093	2.539	2.861
20	0.687	1.325	1.725	2.086	2.528	2.845
21	0.686	1.323	1.721	2.080	2.518	2.831
22	0.686	1.321	1.717	2.074	2.508	2.819
23	0.685	1.319	1.714	2.069	2.500	2.807
24	0.685	1.318	1.711	2.064	2.492	2.797
25	0.684	1.316	1.708	2.060	2.485	2.787
26	0.684	1.315	1.706	2.056	2.479	2.779
27	0.684	1.314	1.703	2.052	2.473	2.771
28	0.683	1.313	1.701	2.048	2.467	2.763
29	0.683	1.311	1.699	2.045	2.462	2.756
30	0.683	1.310	1.697	2.042	2.457	2.750
40	0.681	1.303	1.684	2.021	2.423	2.704
60	0.679	1.296	1.671	2.000	2.390	2.660
120	0.677	1.289	1.658	1.980	2.358	2.617
∞	0.674	1.282	1.645	1.960	2.326	2.576

Πίνακας 22.5 Κρίσιμες Τιμές στην Κατανομή t (ή κατανομή Student)

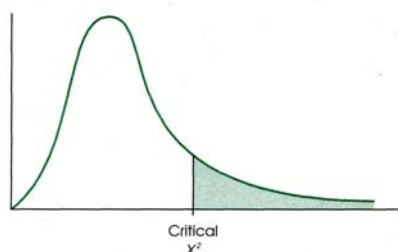
*Table entries in lightface type are critical values for the .05 level of significance. Bold-face type values are for the .01 level of significance.



DEGREES OF FREEDOM: DENOMINATOR	DEGREES OF FREEDOM: NUMERATOR															
	1	2	3	4	5	6	7	8	9	10	11	12	14	16	20	
1	161	200	216	225	230	234	237	239	241	242	243	244	245	246	248	
	4052	4999	5403	5625	5764	5859	5928	5981	6022	6056	6082	6106	6142	6169	6208	
2	18.51	19.00	19.16	19.25	19.30	19.33	19.36	19.37	19.38	19.39	19.40	19.41	19.42	19.43	19.44	
	98.49	99.00	99.17	99.25	99.30	99.33	99.34	99.36	99.38	99.40	99.41	99.42	99.43	99.44	99.45	
3	10.13	9.55	9.28	9.12	9.01	8.94	8.88	8.84	8.81	8.78	8.76	8.74	8.71	8.69	8.66	
	34.12	30.92	29.46	28.71	28.24	27.91	27.67	27.49	27.34	27.23	27.13	27.05	26.92	26.83	26.69	
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.93	5.91	5.87	5.84	5.80	
	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.54	14.45	14.37	14.24	14.15	14.02	
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.78	4.74	4.70	4.68	4.64	4.60	4.56	
	16.26	13.27	12.06	11.39	10.97	10.67	10.45	10.27	10.15	10.05	9.96	9.89	9.77	9.68	9.55	
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.03	4.00	3.96	3.92	3.87	
	13.74	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.79	7.72	7.60	7.52	7.39	
7	5.59	4.47	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.63	3.60	3.57	3.52	3.49	3.44	
	12.25	9.55	8.45	7.85	7.46	7.19	7.00	6.84	6.71	6.62	6.54	6.47	6.35	6.27	6.15	
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.34	3.31	3.28	3.23	3.20	3.15	
	11.26	8.65	7.59	7.01	6.63	6.37	6.19	6.03	5.91	5.82	5.74	5.67	5.56	5.48	5.36	
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.13	3.10	3.07	3.02	2.98	2.93	
	10.56	8.02	6.99	6.42	6.06	5.80	5.62	5.47	5.35	5.26	5.18	5.11	5.00	4.92	4.80	
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.97	2.94	2.91	2.86	2.82	2.77	
	10.04	7.56	6.55	5.99	5.64	5.39	5.21	5.06	4.95	4.85	4.78	4.71	4.60	4.52	4.41	
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.86	2.82	2.79	2.74	2.70	2.65	
	9.65	7.20	6.22	5.67	5.32	5.07	4.88	4.74	4.63	4.54	4.46	4.40	4.29	4.21	4.10	
12	4.75	3.88	3.49	3.26	3.11	3.00	2.92	2.85	2.80	2.76	2.72	2.69	2.64	2.60	2.54	
	9.33	6.93	5.95	5.41	5.06	4.82	4.65	4.50	4.39	4.30	4.22	4.16	4.05	3.98	3.86	
13	4.67	3.80	3.41	3.18	3.02	2.92	2.84	2.77	2.72	2.67	2.63	2.60	2.55	2.51	2.46	
	9.07	6.70	5.74	5.20	4.86	4.62	4.44	4.30	4.19	4.10	4.02	3.96	3.85	3.78	3.67	
14	4.60	3.74	3.34	3.11	2.96	2.85	2.77	2.70	2.65	2.60	2.56	2.53	2.48	2.44	2.39	
	8.86	6.51	5.56	5.03	4.69	4.46	4.28	4.14	4.03	3.94	3.86	3.80	3.70	3.62	3.51	
15	4.54	3.68	3.29	3.06	2.90	2.79	2.70	2.64	2.59	2.55	2.51	2.48	2.43	2.39	2.33	
	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.73	3.67	3.56	3.48	3.36	
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.45	2.42	2.37	2.33	2.28	
	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.61	3.55	3.45	3.37	3.25	

Πίνακας 22.6 Κρίσιμες Τιμές στην Κατανομή F

*The table entries are critical values of χ^2 .



df	PROPORTION IN CRITICAL REGION				
	0.10	0.05	0.025	0.01	0.005
1	2.71	3.84	5.02	6.63	7.88
2	4.61	5.99	7.38	9.21	10.60
3	6.25	7.81	9.35	11.34	12.84
4	7.78	9.49	11.14	13.28	14.86
5	9.24	11.07	12.83	15.09	16.75
6	10.64	12.59	14.45	16.81	18.55
7	12.02	14.07	16.01	18.48	20.28
8	13.36	15.51	17.53	20.09	21.96
9	14.68	16.92	19.02	21.67	23.59
10	15.99	18.31	20.48	23.21	25.19
11	17.28	19.68	21.92	24.72	26.76
12	18.55	21.03	23.34	26.22	28.30
13	19.81	22.36	24.74	27.69	29.82
14	21.06	23.68	26.12	29.14	31.32
15	22.31	25.00	27.49	30.58	32.80
16	23.54	26.30	28.85	32.00	34.27
17	24.77	27.59	30.19	33.41	35.72
18	25.99	28.87	31.53	34.81	37.16
19	27.20	30.14	32.85	36.19	38.58
20	28.41	31.41	34.17	37.57	40.00
21	29.62	32.67	35.48	38.93	41.40
22	30.81	33.92	36.78	40.29	42.80
23	32.01	35.17	38.08	41.64	44.18
24	33.20	36.42	39.36	42.98	45.56
25	34.38	37.65	40.65	44.31	46.93
26	35.56	38.89	41.92	45.64	48.29
27	36.74	40.11	43.19	46.96	49.64
28	37.92	41.34	44.46	48.28	50.99
29	39.09	42.56	45.72	49.59	52.34
30	40.26	43.77	46.98	50.89	53.67
40	51.81	55.76	59.34	63.69	66.77
50	63.17	67.50	71.42	76.15	79.49
60	74.40	79.08	83.30	88.38	91.95
70	85.53	90.53	95.02	100.42	104.22
80	96.58	101.88	106.63	112.33	116.32
90	107.56	113.14	118.14	124.12	128.30
100	118.50	124.34	129.56	135.81	140.17

Πίνακας 22.7 Κρίσιμες Τιμές στην Κατανομή χ^2