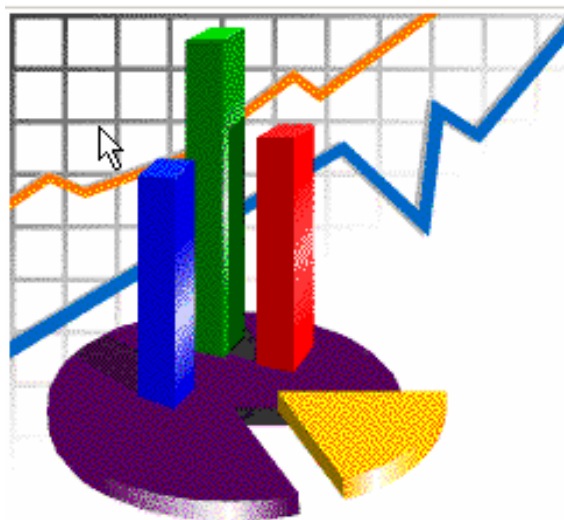


Α. ΕΜΒΑΛΩΤΗΣ Α. ΚΑΤΣΗΣ Γ. ΣΙΔΕΡΙΑΔΗΣ

**ΣΤΑΤΙΣΤΙΚΗ
ΜΕΘΟΔΟΛΟΓΙΑ ΕΚΠΑΙΔΕΥΤΙΚΗΣ ΕΡΕΥΝΑΣ**



ΠΙΛΟΤΙΚΗ ΕΚΔΟΣΗ 0.9

ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

Πρόλογος	3
1. ΕΙΣΑΓΩΓΗ	4
1.1. Η έννοια της μεταβλητής.....	4
1.2. Κλίμακες μέτρησης.....	6
1.3. Γραφική παρουσίαση των δεδομένων – ευρημάτων	7
2. ΑΝΑΛΥΣΗ ΣΥΣΧΕΤΙΣΗΣ	13
2.1. Συντελεστής συσχέτισης Pearson r	14
2.2. Προϋπόθεση της γραμμικότητας.....	19
2.3. Προϋπόθεση της κανονικής κατανομής.....	21
2.4. Υπολογισμός και ερμηνεία του συντελεστή συσχέτισης Pearson.....	23
2.5. Μη-παραμετρικοί συντελεστές συσχέτισης (Spearman).....	28
3. ΑΝΑΛΥΣΗ ΠΑΛΙΝΔΡΟΜΗΣΗΣ.....	30
3.1. Υπολογισμός της γραμμής παλινδρόμησης με τη χρήση του SPSS.....	34
3.2. Εφαρμογή της προβλεπτικής εξίσωσης	39
4. ΕΛΕΓΧΟΣ ΓΙΑ ΤΟ ΜΕΣΟ ΟΡΟ.....	41
4.1. Εισαγωγή	41
4.2. Στατιστική μεθοδολογία.....	43
4.2.1. Έλεγχος για το μέσο όρο ενός πληθυσμού.....	43
4.3. Σχόλια για τον έλεγχο του μέσου όρου ενός πληθυσμού.....	47
4.4. Εφαρμογές του ελέγχου για το μέσο όρο ενός πληθυσμού.....	49
4.5. Έλεγχος για τη σύγκριση μέσων όρων δύο πληθυσμών	57
4.5.1. Ανεξάρτητα δείγματα	57
4.5.2. Εξαρτημένα δείγματα: Κριτήριο ελέγχου και εφαρμογή.....	61
4.6. Σχόλια για τον έλεγχο μεταξύ των μέσων όρων 2 πληθυσμών.....	64
5. ΑΝΑΛΥΣΗ ΔΙΑΚΥΜΑΝΣΗΣ.....	67
5.1. Εισαγωγή	67
5.2. Μεθοδολογία ANOVA.....	68
5.3. Σχόλια.....	70
5.4. Εφαρμογή.....	71
6. ΤΟ ΚΡΙΤΗΡΙΟ χ^2.....	76
6.1. Εφαρμογή.....	77
6.2. Τυπολόγιο ανάγνωσης ενδείξεων για τους πίνακες του SPSS.....	82
ΠΑΡΑΡΤΗΜΑΤΑ.....	85
A: Έλεγχος κανονικότητας.....	86
B. Επιλογή κριτηρίου ελέγχου	89
ΒΙΒΛΙΟΓΡΑΦΙΑ-ΔΙΑΔΙΚΤΥΟ	93

Πρόλογος

Στη σύγχρονη έρευνα η Στατιστική είναι ένα σημαντικότατο εργαλείο για οποιονδήποτε ασχολείται με την (ποσοτική) ανάλυση δεδομένων. Οι σημειώσεις που κρατάτε στα χέρια σας απευθύνονται σε φοιτητές, αλλά και ερευνητές που κάνουν χρήση των εφαρμογών της Στατιστικής στο χώρο κυρίως των Κοινωνικών Επιστημών. Στόχος τους είναι η παρουσίαση και ανάλυση των κυριότερων εφαρμογών της Στατιστικής Μεθοδολογίας με έμφαση στα παρακάτω σημεία:

- Επεξήγηση της θεωρίας με παραδείγματα από το χώρο των Κοινωνικών Επιστημών
- Επιλογή και παρουσίαση του κατάλληλου κάθε φορά στατιστικού κριτηρίου και
- Ερμηνεία των αποτελεσμάτων.

Ένας τομέας στον οποίο δόθηκε ιδιαίτερη προσοχή ήταν η (περιορισμένη) χρήση μαθηματικών εννοιών και όρων. Εκτιμήθηκε ότι για τη μελέτη και αξιοποίηση του σχετικού υλικού δεν θα πρέπει να θεωρείται ως προϋπόθεση η σε βάθος προσέγγιση μαθηματικών εννοιών. Για το σκοπό αυτό, καταβλήθηκε προσπάθεια να περιοριστούν στο ελάχιστο οι αναφορές στη μαθηματική ορολογία και τυπολογία. Παραθέτουμε τα σημαντικότερα και απολύτως απαραίτητα μαθηματικά εργαλεία με τρόπο όσο γίνεται πιο κατανοητό. Ο αναγνώστης καλό είναι να μην επιμείνει στην πλήρη κατανόηση ενός μαθηματικού τύπου, αν αυτό του δημιουργεί πρόβλημα, αλλά στην ποιοτική διάσταση της μεθόδου και των αποτελεσμάτων της. Από την άλλη πλευρά όμως αυτή η προσπάθεια "απο-μαθηματικοποίησης" δε σημαίνει ότι το κείμενο δεν είναι έγκυρο και δεν απαιτείται σοβαρή και συστηματική προσπάθεια για την κατανόησή του.

Ελπίζουμε να έχουμε προσεγγίσει τους στόχους που θέσαμε εξ' αρχής για τη συγγραφή αυτών των σημειώσεων. Ενδεχομένως έχουν παρεισφρήσει λάθη και παραλείψεις, τα οποία όπως πάντα βαρύνουν τους συγγραφείς.

1. ΕΙΣΑΓΩΓΗ

1.1. Η έννοια της μεταβλητής

Η έρευνα στις κοινωνικές επιστήμες και ειδικότερα στις επιστήμες της αγωγής, βασίζεται στη μελέτη παραγόντων-χαρακτηριστικών που συμμετέχουν στα υπό διερεύνηση φαινόμενα. Ονομάζουμε αυτούς τους παράγοντες «**μεταβλητές**» (variables). Γενικώς μεταβλητή είναι κάθε τι το οποίο επιδέχεται περισσότερες από μία τιμές. Η σχολική επίδοση είναι μια κλασική μεταβλητή που χρησιμοποιείται στην εκπαιδευτική έρευνα. Επιδέχεται περισσότερες από μία τιμές, οι οποίες μάλιστα δεν χρειάζεται να είναι αριθμητικές. Στην περίπτωση που οι τιμές μιας μεταβλητής είναι αριθμοί (όπως στην περίπτωση του βάρους ή της ηλικίας) η μεταβλητή ονομάζεται αριθμητική. Οι τιμές της μεταβλητής μπορεί επίσης να είναι κάποια ιδιότητα, όπως στην περίπτωση της μεταβλητής του φύλου.

Στην κοινωνική έρευνα έχουν προταθεί και καθιερωθεί μια σειρά ταξινομικά συστήματα, τα οποία διακρίνουν τύπους μεταβλητών. Η συνηθέστερη ταξινόμηση ορίζει «ποιοτικές» (qualitative) και «ποσοτικές» (quantitative) μεταβλητές.

Ποιοτικές μεταβλητές (qualitative variables) ονομάζονται όσες μεταβάλλονται σε είδος, όπως το φύλο, το θρήσκευμα, η εθνικότητα, η οικογενειακή κατάσταση, κ.τ.ό. Επισημαίνεται ότι οι τιμές των ποιοτικών μεταβλητών θα πρέπει να συγκροτούν ευδιάκριτες και μη επικαλυπτόμενες κατηγορίες. Έτσι στη μεταβλητή με την ενδεικτική περιγραφή «πολιτική τοποθέτηση», (θα πρέπει να) ανήκουν σαφώς προσδιορισμένες τιμές. Οι γενικές διατυπώσεις «προοδευτικός», «κεντρώος» και «συντηρητικός» ως τιμές στην παραπάνω μεταβλητή, εάν δεν περιγραφούν/οριστούν με ακρίβεια, δημιουργούν συνθήκες ανασφάλειας ως προς την εγκυρότητα των σχετικών μετρήσεων.

Ποσοτικές μεταβλητές (quantitative variables) ονομάζονται όσες μεταβάλλονται ποσοτικά, όπως το βάρος, η ηλικία, το εισόδημα (όταν αποτυπώνεται σε αριθμητική τιμή), κ.τ.ό. Ιδιαίτερη σημασία έχει η επιλογή της καταλληλότερης μονάδας μέτρησης, προκειμένου να είναι συγκρίσιμες οι επιμέρους τιμές. Έτσι λοιπόν, η θερμοκρασία σε μια μέτρηση θα πρέπει αποτυπώνεται για το σύνολο των μετρήσεων σε μια και μόνο κλίμακα (κλίμακα Celsius, κλίμακα Kelvin ή κλίμακα Fahrenheit), η οποία μάλιστα καλό θα ήταν να επιλεγεί πριν την έναρξη των μετρήσεων.

Οι ποσοτικές μεταβλητές διακρίνονται σε **συνεχείς** (continuous) και **διακριτές** (discrete) μεταβλητές.

Η συνεχής (ποσοτική) μεταβλητή επιδέχεται όλες τις τιμές ενός εύρους τιμών με ελάχιστο κάτω όριο την ελάχιστη τιμή και ανώτατο πάνω όριο τη μέγιστη τιμή ενός συνόλου πραγματικών αριθμών (real numbers). Παράδειγμα συνεχούς ποσοτικής μεταβλητής είναι το βάρος ενός υποκειμένου, το οποίο δέχεται τιμές από το δυνατό εύρος τιμών μέτρησης βάρους. Έτσι οι τιμές μπορεί να είναι 67.2 (Kgr), 67.3 (Kgr), 67.4 (Kgr), κ.τ.ό.

Οι διακριτές (ποσοτικές) μεταβλητές επιδέχονται επιλεγμένες τιμές, οι οποίες πρέπει να είναι ακέραιοι αριθμοί. Παράδειγμα διακριτής μεταβλητής είναι η σύνθεση της οικογένειας, όταν η αποτύπωση αναφέρεται στον αριθμό των προσώπων που την αποτελούν. Έτσι, τιμές της παραπάνω μεταβλητής μπορεί να είναι 3 άτομα, 4 άτομα, 5 άτομα, κ.τ.ό, αλλά ποτέ 2.5 άτομα.

Ένας άλλος τύπος διάκρισης και ταξινόμησης των μεταβλητών ορίζει **ανεξάρτητες** (independent) και **εξαρτημένες** (dependent) μεταβλητές. Η συγκεκριμένη τυπολογία συναντάται συχνά στην εκπόνηση ερευνητικών σχεδίων και ιδιαίτερα στην ενότητα της διατύπωσης ερευνητικών υποθέσεων ή ερωτημάτων.

Ανεξάρτητη μεταβλητή ονομάζεται η μεταβλητή η οποία επηρεάζει αιτιωδώς μια ή περισσότερες εξαρτημένες μεταβλητές. Σε ορισμένες (στατιστικές) αναλύσεις η ανεξάρτητη μεταβλητή ορίζεται και ως προβλεπτικός παράγοντας (predictor μεταβλητή), με την έννοια ότι προβλέπει την τιμή της μεταβλητής που επηρεάζει (της εξαρτημένης μεταβλητής).

Εξαρτημένη μεταβλητή ονομάζεται η μεταβλητή οι τιμές της οποίας εξαρτώνται από τις μεταβολές της ανεξάρτητης μεταβλητής. Σε ορισμένες (στατιστικές) αναλύσεις η εξαρτημένη μεταβλητή ορίζεται και ως αποτέλεσμα (outcome/response μεταβλητή).

Συμπληρωματικά της παραπάνω ταξινόμησης στις αναλύσεις εντοπίζουμε μεταβλητές **ελέγχου** (control), **αλλότριες** (extraneous) μεταβλητές, **κατασταλτικές** (suppressor) μεταβλητές και **παρεμβαλλόμενες** (intervening) μεταβλητές.

Μεταβλητές ελέγχου ονομάζονται όσες ουδετεροποιούνται ή των οποίων η δραστική επιρροή περιορίζεται από τον ερευνητή, προκειμένου να μετρηθεί η επίδραση της ανεξάρτητης στην εξαρτημένη μεταβλητή.

Αλλότριες μεταβλητές ονομάζονται όσες ανεξάρτητες μεταβλητές δεν απασχολούν τον ερευνητή, παρότι ενδέχεται να επηρεάζουν την εξαρτημένη μεταβλητή.

Κατασταλτικές μεταβλητές ονομάζονται όσες ανεξάρτητες μεταβλητές επηρεάζουν τόσο δραστικά στη σχέση ανεξάρτητης – εξαρτημένης μεταβλητής ώστε να περιορίζεται η εμβέλεια της ανεξάρτητης μεταβλητής.

Παρεμβαλλόμενες μεταβλητές ονομάζονται όσες μεταβλητές παρεμβάλλονται κατά τη διάρκεια της έρευνας και ενδεχομένως η επίδρασή τους δεν είχε προβλεφθεί ή δεν είναι δυνατό να περιοριστεί ή να ακυρωθεί.

1.2. Κλίμακες μέτρησης

Η συζήτηση για τις κλίμακες μέτρησης ορίζεται με αναφορά στις εργασίες του Stevens (1951). Οι ταξινομήσεις αναφορικά με τις μετρήσεις που προκύπτουν στο ερευνητικό πεδίο οργανώνονται σε δύο τύπους, ανάλογα με τον τύπο της μεταβλητής που διερευνάται. Η βασική διάκριση αναγνωρίζει **ποσοτικές & ποιοτικές** μετρήσεις. Οι ποσοτικές μετρήσεις προκύπτουν από την επεξεργασία των ποσοτικών (quantitative) μεταβλητών, ενώ οι ποιοτικές μετρήσεις προκύπτουν από την επεξεργασία των ποιοτικών (qualitative) μεταβλητών. Στη μέτρηση ποιοτικών μεταβλητών εντοπίζουμε **κατηγορικές ή ονομαστικές** (nominal) και **διατακτικές ή ιεραρχικές** (ordinal) κλίμακες μέτρησης, ενώ στις ποσοτικές μεταβλητές εντοπίζουμε **κλίμακες ίσων διαστημάτων** (interval) και **κλίμακες αναλογίας** (ratio).

Η ονομαστική κλίμακα εφαρμόζεται όταν επιχειρείται απλή ταξινόμηση χαρακτηριστικών. Το σύνολο τιμών δεν έχει κάποια ιδιότητα και η σημασία αναγνωρίζεται μόνο στο πλήθος των (διακριτών) τιμών. Παραδείγματα ονομαστικής κλίμακας αναγνωρίζονται στη μεταβλητή του φύλου ή της οικογενειακής κατάστασης ενός υποκειμένου. Οφείλει να επισημάνει κανείς ότι στις μετρήσεις που προκύπτουν με χρήση ονομαστικής κλίμακας απαιτείται ιδιαίτερη προσοχή στην εκτίμηση των μέτρων θέσης και διασποράς.

Η διατακτική κλίμακα εφαρμόζεται όταν οι τιμές της μεταβλητής αναπτύσσονται κατά τρόπο ώστε να είναι δυνατή η ανίχνευση κάποιας σχέσης διάταξης. Για παράδειγμα η σειρά τερματισμού σε κάποιο αγώνισμα ή η αξιολόγηση των κινηματογραφικών ταινιών της εβδομάδας είναι περιπτώσεις εφαρμογής διατακτικής κλίμακας. Στον έλεγχο διατακτικής κλίμακας ίσες διαφορές μεταξύ των τιμών της μεταβλητής δεν συνεπάγεται

και ίσες διαφορές για το χαρακτηριστικό που συνοδεύει τη μεταβλητή. Για παράδειγμα, η διαφορά στο χρόνο μεταξύ του πρώτου και του δεύτερου αθλητή σε ένα μαραθώνιο δρόμο μπορεί να είναι 2 δευτερόλεπτα ενώ η διαφορά του δεύτερου από τον τρίτο 30 λεπτά! Με βάση αυτή την κλιμάκωση το μέγεθος της διαφοράς δεν έχει σημασία, ενώ εξετάζεται μόνο η σειρά (πρώτος, δεύτερος, τρίτος, κλπ.). Στις μετρήσεις που προκύπτουν με χρήση διατακτικής κλίμακας ορισμένα μέτρα θέσης μπορούν να χρησιμοποιηθούν με προσοχή ως δείκτες τάσης.

Στην κλίμακα ίσων διαστημάτων οι τιμές της μεταβλητής επιδέχονται όχι μόνο σχέση διάταξης, αλλά και ίσες διαφορές μεταξύ των τιμών συνεπάγονται και ίσες διαφορές του χαρακτηριστικού που μετράει η μεταβλητή. Παράδειγμα εφαρμογής κλίμακας διαστήματος έχουμε στην περίπτωση ελέγχου θερμοκρασιακών μετρήσεων. Η διαφορά ανάμεσα στους 10°C και 20°C (10°C) στη στήλη του θερμομέτρου, είναι ίση με τη διαφορά ανάμεσα στους 80°C και 90 °C). Ιδιαίτερο γνώρισμα των μεταβλητών διαστήματος είναι ότι η αρχική τιμή ορίζεται αυθαίρετα. Αυτό επιτρέπει σε κάθε γραμμικό μετασχηματισμό της μορφής $y=a+bx$ (όπου x η αρχική τιμή, y η μετασχηματισμένη b μια θετική και a κάποια τυχαία σταθερά) να διατηρεί τις ιδιότητες της κλίμακας (Τσάντας, κ.ά. 1999:1-28)¹.

Η κλίμακα αναλογίας εφαρμόζεται όταν οι τιμές της μεταβλητής αντιστοιχούν αναλογικά στην ποσότητα του χαρακτηριστικού της μεταβλητής που μετρούν. Στην περίπτωση αυτή η αρχική τιμή (το μηδέν επί παραδείγματι), ορίζεται με σαφήνεια και καθιστά την αφετηρία βαθμονόμησης (και βαθμολόγησης) των μεταβολών στις τιμές. Κλίμακες αναλογίες επιτρέπουν μετασχηματισμούς της μορφής $y=bx$ (Τσάντας, κ.ά. 1999: 1-28). Παραδείγματα τέτοιων μετασχηματισμών είναι οι μετατροπές δραχμών σε € (1 € = 340,750 δραχμές)

1.3. Γραφική παρουσίαση των δεδομένων – ευρημάτων

Στις αρχικές ενέργειες του ερευνητή που ενδιαφέρεται να παρουσιάσει τις κατανομές των δεδομένων του είναι η συνοπτική, περιεκτική και εποπτική απόδοσή τους με τη βοήθεια κατάλληλων γραφικών παραστάσεων. Μια γραφική παράσταση οφείλει να είναι:

¹ Μελετήστε το κεφάλαιο της ανάλυσης παλινδρόμησης σε αυτές τις σημειώσεις για την επεξήγηση αυτής της συνάρτησης.

- (α) παραστατική, διευκολύνοντας την αναγνώριση και κατανόηση των βασικών χαρακτηριστικών της εξεταζόμενης μεταβλητής,
- (β) σαφής, αποφεύγοντας τις συγχύσεις ανάμεσα στις τιμές κάθε μεταβλητής και
- (γ) ακριβής ώστε να αποφεύγονται οι πλάνες.

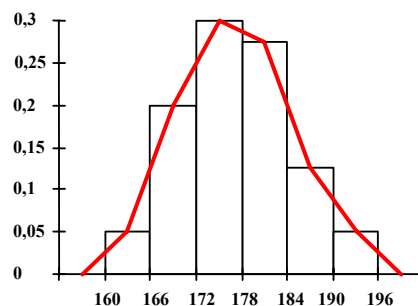
Οι συνηθέστερες γραφικές παραστάσεις είναι το ιστόγραμμα συχνοτήτων (histogram), το πολύγωνο συχνοτήτων (frequency polygon), η καμπύλη συχνοτήτων (frequency curve), το ραβδόγραμμα (bar chart) και το κυκλικό διάγραμμα (pie chart) και οι σπανιότερες το διάγραμμα μίσχου – φύλλων (stem and leaf diagram) & το σημειόγραμμα (dot diagram).

- **Το ιστόγραμμα (histogram)**

Πρόκειται για τη συνηθέστερη επιλογή γραφικής παράστασης ποσοτικών μεταβλητών. Οι τιμές της μεταβλητής διατάσσονται (συνήθως) στον οριζόντιο άξονα κατ' αύξουσα σειρά με αφετηρία την αρχή των αξόνων. Το ύψος κάθε στήλης αντιστοιχεί στη συχνότητα κάθε τιμής² και δίνεται πάντοτε σημασία στην καταλληλότητα της κλίμακας μέτρησης στους άξονες αναφοράς³ (βλ. ενδεικτικά και <http://www.shodor.org/interactivate/activities/histogram/>)

- **Το πολύγωνο συχνοτήτων (frequency polygon)**

Το πολύγωνο συχνοτήτων σχηματίζεται συνδέοντας την κεντρική τιμή κάθε στήλης με ευθύγραμμα τμήματα. Το βασικό χαρακτηριστικό του είναι ότι το εμβαδόν της περιοχής που περικλείεται από την πολυγωνική γραμμή και τον οριζόντιο άξονα x , ισούται με το σύνολο των συχνοτήτων.

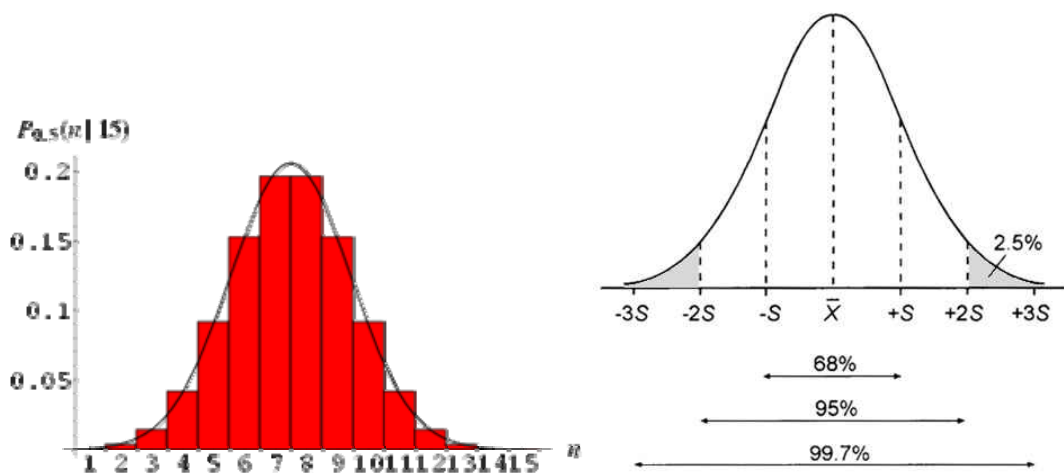


Σχήμα 1: Πολύγωνο (σχετικών) συχνοτήτων

² Εάν πρόκειται για ομαδοποιημένη κατανομή αναφέρεται το εύρος τιμών.

- **Η καμπύλη συχνοτήτων (frequency curve)**

Η καμπύλη συχνοτήτων παράγει γραφική παράσταση της θεωρητικής κατανομής των συχνοτήτων⁴ και μπορεί να εξαχθεί από το πολύγωνο συχνοτήτων. Οι συνηθέστερες (και γνωστότερες) κατανομές είναι η ομοιόμορφη κατανομή (rectangular distribution), η κανονική κατανομή (normal distribution), η κατανομή student και η F κατανομή. Ιδιαίτερη αναφορά χρειάζεται να γίνει στην κανονική κατανομή (ή κατανομή Gauss) η οποία θεωρείται από τις σημαντικότερες στην ανάλυση ερευνητικών δεδομένων λόγω της συμμετρίας και της ευρείας χρήσης της. Η κανονική κατανομή αξιοποιείται κυρίως προκειμένου να αναλυθούν μετρήσεις οι οποίες αναπτύσσονται σε ένα συμμετρικό ιστόγραμμα σε σχήμα καμπάνας.



Σχήματα 2 & 3: Ιστόγραμμα και καμπύλη κανονικής κατανομής

- **Το ραβδόγραμμα (bar chart)**

Το ραβδόγραμμα επιτρέπει τη γραφική απεικόνιση ποιοτικών μεταβλητών. Η συνηθέστερη πρακτική είναι η τοποθέτηση στον οριζόντιο άξονα (x) των κατηγοριών της μεταβλητής και η ανάπτυξη ράβδων με ύψος ανάλογο της συχνότητας για κάθε κατηγορία. Συχνά το ραβδόγραμμα αξιοποιείται από τους ερευνητές προκειμένου να συγκρίνουν ομάδες περιπτώσεων (Norusis 2002:570-573). Μια διασκευή του

³ Σε πολλές περιπτώσεις το βήμα (scale) επιλέγεται από τον ερευνητή μετά την επεξεργασία των δεδομένων και την παραγωγή του σχετικού πίνακα συχνοτήτων.

ραβδογράμματος είναι το διάγραμμα Pareto (Pareto's chart), το οποίο εμφανίζει τις κατηγορίες της ποιοτικής μεταβλητής σε φθίνουσα σειρά (βλ. ενδεικτικά και http://nces.ed.gov/nceskids/Graphing/bar_pie_data.asp?ChartType=bar)

- **Το κυκλικό διάγραμμα (pie chart)**

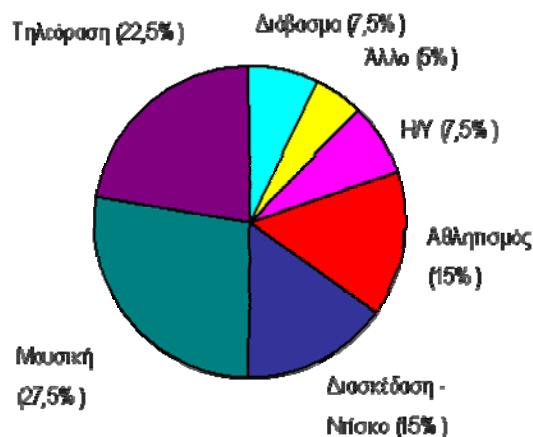
Πρόκειται για μια εναλλακτική (του ραβδογράμματος) γραφική απεικόνιση κατηγορικών δεδομένων και βασίζεται στην απλή παραδοχή ότι κάθε κύκλος μπορεί να διαιρεθεί σε 360 τμήματα (360 μοίρες). Για τη δημιουργία του η συχνότητα κάθε τιμής πολλαπλασιάζεται επί 3.6 (όση και η γωνία που αντιστοιχεί στο 1% του κυκλικού δίσκου), με αποτέλεσμα το γινόμενο να αντιστοιχεί στη γωνία του αντίστοιχου τμήματος στο κυκλικό διάγραμμα. Κατ' αυτό τον τρόπο η απεικόνιση της (σχετικής) συχνότητας 34 είναι η γωνία με άνοιγμα $34 \times 3.6 = 122.4$ μοίρες. Η χρήση των κυκλικών διαγραμμάτων προτείνεται όταν (α) οι τιμές της μεταβλητής αθροίζουν σε σταθερό άθροισμα (το 100% των σχετικών συχνοτήτων), (β) όταν οι κατηγορίες αναπτύσσονται κατά τρόπο αισθητά διαφορετικό⁵ και (γ) όταν ο αριθμός των επιμέρους κατηγοριών δεν είναι ιδιαίτερα μεγάλος⁶.

(βλ. ενδεικτικά http://nces.ed.gov/nceskids/Graphing/bar_pie_data.asp?ChartType=pie, <http://www.shodor.org/interactivate/activities/piechart/index.html>).

⁴ Οι θεωρητικές κατανομές συχνοτήτων διακρίνονται σε συνεχείς (κανονική, εκθετική, student, X^2 , κ.λπ.) και ασυνεχείς (διωνυμική, γεωμετρική, υπεργεωμετρική, ομοιόμορφη, κ.λπ).

⁵ Ένα κυκλικό διάγραμμα το οποίο εμφανίζει οκτώ τομείς (κατηγορίες) με την ίδια συχνότητα δεν είναι παραστατικό.

⁶ Ένα κυκλικό διάγραμμα με πολλούς τομείς δεν είναι (επίσης) παραστατικό.



Σχήμα 4: Κυκλικό διάγραμμα

- **Το θηκόγραμμα (boxplot)**

Η συνήθης εφαρμογή του θηκογράμματος είναι ο σχεδιασμός του προκειμένου να συγκριθούν κατανομές. Για κάθε κατανομή δημιουργείται με τη βοήθεια πέντε στατιστικών στοιχείων θέσεως (ελάχιστο, μέγιστο, 1ο-2^ο & 3ο τεταρτημόριο) ένα ορθογώνιο πλαίσιο, το μήκος του οποίου αντιστοιχεί στο ενδοτεταρτημοριακό εύρος της μεταβλητής. Στα άκρα κάθε (ορθογώνιου) πλαισίου (βλ. σχήμα) αναπτύσσονται γραμμές (απολήξεις⁷) οι οποίες συνδέουν τις μέγιστες και ελάχιστες τιμές της μεταβλητής⁸. Με τη βοήθεια του θηκογράμματος ο ερευνητής έχει άμεσα τις πρώτες ενδείξεις αναφορικά με τη συμμετρία της κατανομής.

Παρατηρήσεις αναφορικά με τις γραφικές παρουσιάσεις των δεδομένων (διαγράμματα):

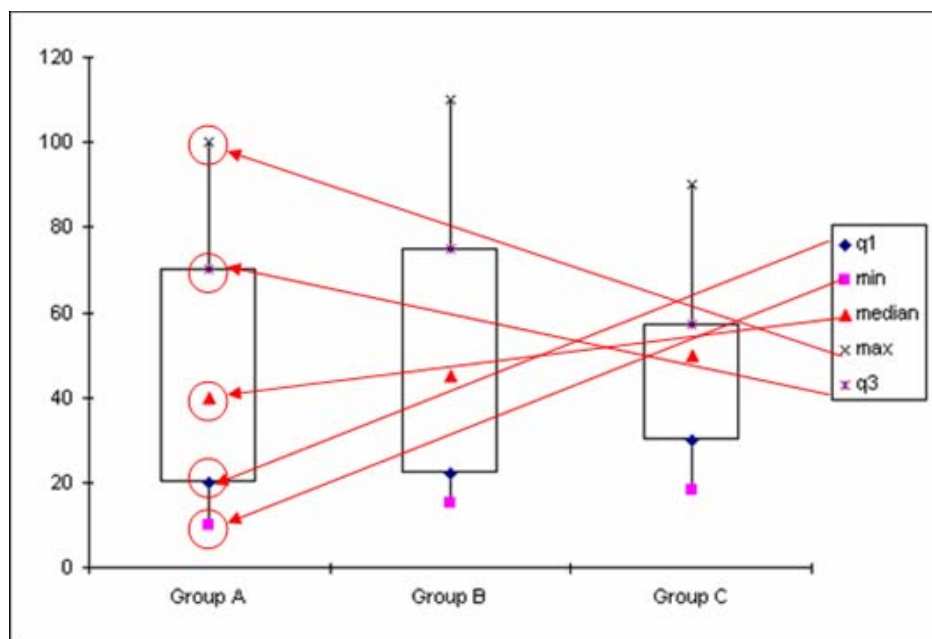
1. Τα διαγράμματα παρέχουν την οπτικοποίηση⁹ των δεδομένων ενός πίνακα. Πίνακας και διάγραμμα θα πρέπει να συμφωνούν (ως προς τις τιμές) απόλυτα.

⁷ Στη σχετική βιβλιογραφία το boxplot αναφέρεται και ως διάγραμμα πλαισίου – απολήξεων.

⁸ Εκτός από τις έκκεντρες (outlier) ή τις ακραίες (extreme) τιμές

⁹ Η οπτικοποίηση των δεδομένων έχει απασχολήσει τους ερευνητές ορισμένοι εκ των οποίων έχουν επισημάνει την ιδιαίτερη προσοχή που οφείλει ο ερευνητής να έχει προκειμένου να αποφευχθούν οπτικές πλάνες

2. Κάθε διάγραμμα θα πρέπει να είναι πλήρες. Θα πρέπει δηλαδή ο αναγνώστης μελετώντας το διάγραμμα να ενημερώνεται πλήρως για το περιεχόμενό του γραφήματος χωρίς να χρειάζεται να δοθούν σχετικές διευκρινίσεις.
3. Κάθε διάγραμμα θα πρέπει να συνοδεύεται από:
 - a. αρίθμηση
 - b. (κατανοητό και πλήρη) τίτλο
 - c. πηγή (όταν τα δεδομένα δεν είναι εμπειρικά)



Box-plot

Σχήμα 4: Θηκόγραμμα

2. ΑΝΑΛΥΣΗ ΣΥΣΧΕΤΙΣΗΣ

Υπάρχει σχέση¹⁰ (και αν ναι, ποια) μεταξύ διαβάσματος και επίδοσης, ευφυΐας και επαγγελματικής αποκατάστασης, καπνίσματος και πρόκλησης καρκίνου; Σε σχετικές ερευνητικές ερωτήσεις καλείται να απαντήσει η έρευνα (και) στις κοινωνικές επιστήμες. Προκειμένου να απαντηθούν τα ερωτήματα χρειαζόμαστε αξιόπιστα στατιστικά εργαλεία, με τα οποία θα αξιολογηθούν τα ερωτήματα και θα επιχειρηθεί να δοθούν πειστικές απαντήσεις.

Ας υποθέσουμε ότι η ερευνητική μας υπόθεση αφορά στις διαφορές μεταξύ αγοριών και κοριτσιών αναφορικά με την επίδοση τους στην ανάγνωση. Σε σχετική μέτρηση διαπιστώθηκε ότι τα αγόρια έχουν μέσο όρο 5 ενώ τα κορίτσια 7. Είμαστε βέβαιοι ότι τα κορίτσια έχουν καλύτερη επίδοση από τα αγόρια; Ενδεχομένως οι περισσότεροι ισχυριστούν «ναι», άλλοι όμως μπορεί να είναι περισσότερο επιφυλακτικοί. Αν το άριστα δεν είναι το 10 αλλά το 20; Θα υπάρχουν περισσότεροι που θα απαντήσουν «όχι»; Αν το άριστα είναι το 50; Προκειμένου να δοθεί απάντηση στα παραπάνω ερωτήματα η στατιστική παρέχει στους ερευνητές «αντικειμενικά κριτήρια» επιτρέποντας να απαντήσουν με τρόπο κατηγορηματικό στο αν οι διαφορές που παρατηρήθηκαν δεν είναι «τυχαίες» (random -δηλαδή δεν οφείλονται σε τυχαίους παράγοντες αλλά σε συστηματικούς, που αξίζει να μελετηθούν) ή οφείλονται σε τυχαίους παράγοντες και δεν επιδέχονται περαιτέρω ανάλυση. Για τον καθορισμό αυτής της τυχαιότητας απαιτείται η χρήση πιθανοτήτων. Χωρίς να μπούμε σε μαθηματικές αναλύσεις, αναφέρουμε ότι στη στατιστική κάτι θεωρείται τυχαίο όταν είναι αρκετά συνηθισμένο, συμβαίνει δηλαδή αρκετές φορές, τουλάχιστον στο 5% των περιπτώσεων. Αντιθέτως, όταν συμβαίνει λιγότερες από 5 στις 100 φορές, τότε θεωρούμε ότι το συγκεκριμένο αποτέλεσμα δε συμβαίνει κατά τύχη, είναι «σπάνιο» και αξίζει να συζητηθεί. Είναι φανερό ότι διχοτομικές αποφάσεις του τύπου «μη τυχαίο – τυχαίο» (ή αντιστοίχως «σπάνιο – συνηθισμένο») δημιουργούν άλλα προβλήματα. Για παράδειγμα κάτι που συμβαίνει 4 φορές στις 100 είναι μη τυχαίο-σπάνιο ενώ κάτι που συμβαίνει 5

¹⁰ Σε αυτό το κεφάλαιο χρησιμοποιούμε τους όρους «σχέση» και «συνάφεια» σαν να είναι ταυτόσημοι.

στις 100 φορές δεν είναι. Επί τοις ουσίαις κάθε ερευνητής αποφασίζει για το επίπεδο τυχειότητας που ταιριάζει στο «πρόβλημα» που μελετά. Τα τυπικά επίπεδα τυχειότητας που χρησιμοποιούνται (συνήθως) είναι το 10% το 2.5%, το 1% το 1% κλπ. Πιο «αυστηρά» (δηλαδή μικρότερα του 5%) επίπεδα σημαντικότητας ορίζονται (συχνά) στην ιατρική και στη φαρμακευτική.

Επομένως, στο παραπάνω παράδειγμα που εξετάζουμε αν υπάρχουν ουσιαστικές διαφορές επίδοσης μεταξύ αγοριών και κοριτσιών, ο στατιστικός δείκτης που θα χρησιμοποιηθεί πρέπει να σχετίζεται με μια πιθανότητα εμφάνισης του συγκεκριμένου αποτελέσματος η οποία συγκρίνεται με το 5% (εκτός και αν οι ερευνητές έχουν αποφασίσει να ορίσουν πιο αυστηρά κριτήρια). Μόνο αν η πιθανότητα είναι μικρότερη του 5% έχουμε στατιστικά σημαντικό εύρημα, το οποίο (υπο)δηλώνει μη-τυχαίες διαφορές και άρα υπάρχει σχέση μεταξύ φύλου και επίδοσης. Σε αντίθετη περίπτωση θεωρούμε ότι δεν υπάρχει στατιστικά σημαντική σχέση. Η υπόθεση της μη στατιστικά σημαντικής σχέσης ονομάζεται και «αρχική» ή «μηδενική» υπόθεση (H_0) ενώ η υπόθεση της στατιστικά σημαντικής σχέσης ονομάζεται και «εναλλακτική» ή «ερευνητική» υπόθεση ή «υπόθεση εργασίας» (H_1 , ή H_A , ή H_e). Η διατύπωση των υποθέσεων είναι απαραίτητη για την απεικόνιση των μεταβλητών και την επιλογή των κριτηρίων στατιστικού ελέγχου.

2.1. Συντελεστής συσχέτισης Pearson r

Ο συντελεστής συσχέτισης Pearson r είναι ο κατάλληλος παραμετρικός στατιστικός δείκτης προκειμένου να αξιολογηθεί αν υπάρχει συνάφεια μεταξύ δύο μεταβλητών (π.χ., ωρών διαβάσματος και επίδοσης). Με τον όρο παραμετρικός αναφερόμαστε σε μια σειρά στατιστικών δεικτών που ικανοποιούν συγκεκριμένες προϋποθέσεις. Αν αυτές οι προϋποθέσεις «απουσιάζουν», τότε γίνεται χρήση μη-παραμετρικών στατιστικών δεικτών, οι οποίοι δεν επηρεάζονται από «καταπάτηση» των προϋποθέσεων. Για τον συντελεστή συσχέτισης Pearson r αυτές είναι:

- (α) οι μεταβλητές πρέπει να είναι συνεχείς σε κλίμακα ίσων διαστημάτων και να κατανέμονται κανονικά &
 - (β) να υπάρχει γραμμική σχέση μεταξύ των δύο μεταβλητών.
-

Αυτό που απαιτείται για την ανίχνευση των σχέσεων είναι να υπάρχει αναλογική σχέση μεταξύ των δύο μεταβλητών για όλα τα επίπεδά τους. Δηλαδή όσο μεγαλώνει η μία να μεγαλώνει και η άλλη και το αντίθετο. Οι προϋποθέσεις αυτές γίνονται κατανοητές στη συνέχεια με τη χρήση του διαγράμματος σκεδασμού και με σχετικές αναλύσεις κανονικότητας.

Το διάγραμμα σκεδασμού είναι ένα πάρα πολύ χρήσιμο διάγραμμα, το οποίο απεικονίζει το είδος και (κατά προσέγγιση), το μέγεθος της σχέσης των δύο μεταβλητών που μελετά ο ερευνητής. Μπορεί να χαρακτηριστεί και ως μια «φωτογραφία» της υπό μελέτη σχέσης. Στο διάγραμμα αυτό κάθε μια από τις δύο μεταβλητές απεικονίζεται σε έναν από τους δύο άξονες (σε ποιον άξονα δεν έχει σημασία για την ανάλυση της συσχέτισης) και μέσα στο σχήμα κάθε «κουκίδα» αποτελεί μια παρατήρηση (ένα άτομο), η οποία είναι συνάρτηση των τιμών που έχει το άτομο για (στις) δύο μεταβλητές (πρόκειται για το σημείο τομής αν χαράξουμε παράλληλες γραμμές από την μια τιμή προς τον ένα άξονα και από την άλλη τιμή προς τον άλλο άξονα). Ας δούμε ένα παράδειγμα:

Υποθέσεις (μηδενική και εναλλακτική):

H_0 : ΔΕΝ υπάρχει σχέση μεταξύ ωρών διαβάσματος και επίδοσης

H_e : Υπάρχει σχέση μεταξύ ωρών διαβάσματος και επίδοσης

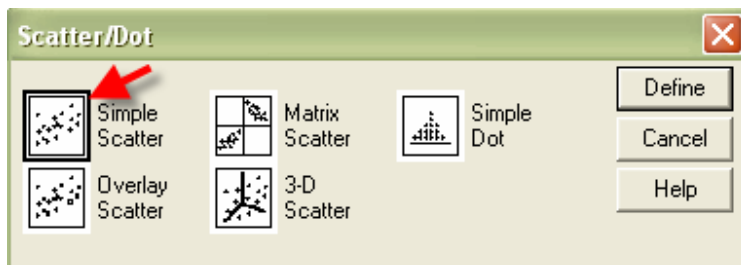
Αν υποθέσουμε ότι έχουμε 10 μαθητές, για τους οποίους έχουμε τα εξής δεδομένα:

Πίνακας 1: Δεδομένα από υποθετική μελέτη για τη σχέση ωρών διαβάσματος και επίδοσης

Συμμετέχοντες	1	2	3	4	5	6	7	8	9	10
Ώρες Διαβάσματος	2	4	6	8	1	0	3	5	7	9
Επίδοση	12	15	16	20	9	6	13	16	18	19

Η δημιουργία του διαγράμματος σκεδασμού (scatterplot) για την οπτική απεικόνιση του είδους και του μεγέθους της σχέσης στο SPSS έχει ως εξής: Από το

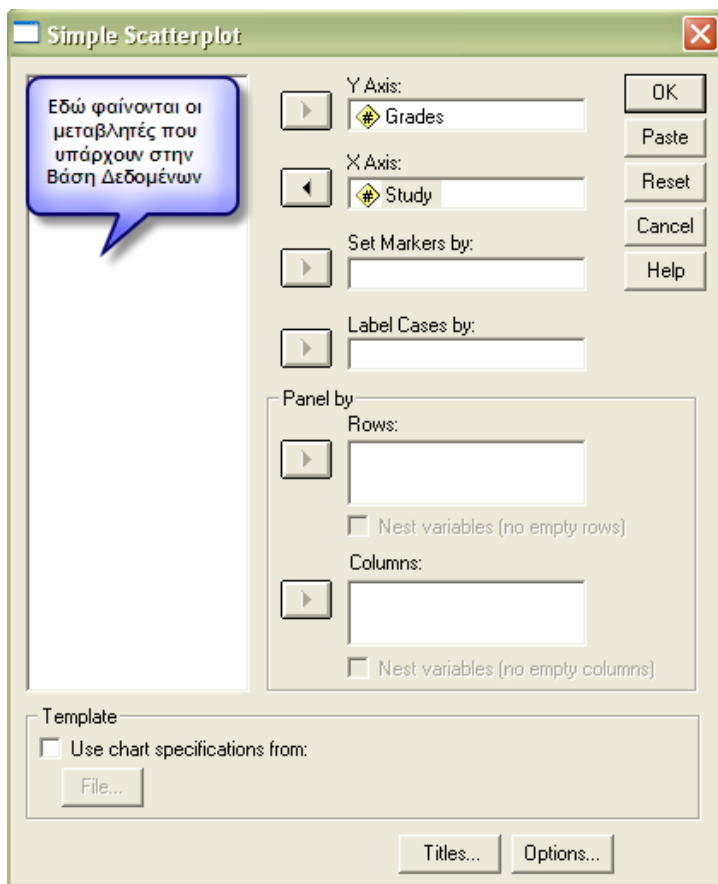
μενού Graphs του SPSS και κάντε κλικ στο scatter/dot. Στη συνέχεια κάντε κλικ στο Define, έχοντας επιλεγμένο το simple scatter



Σχήμα 4: Επιλογές για τη δημιουργία απλού διαγράμματος σκεδασμού

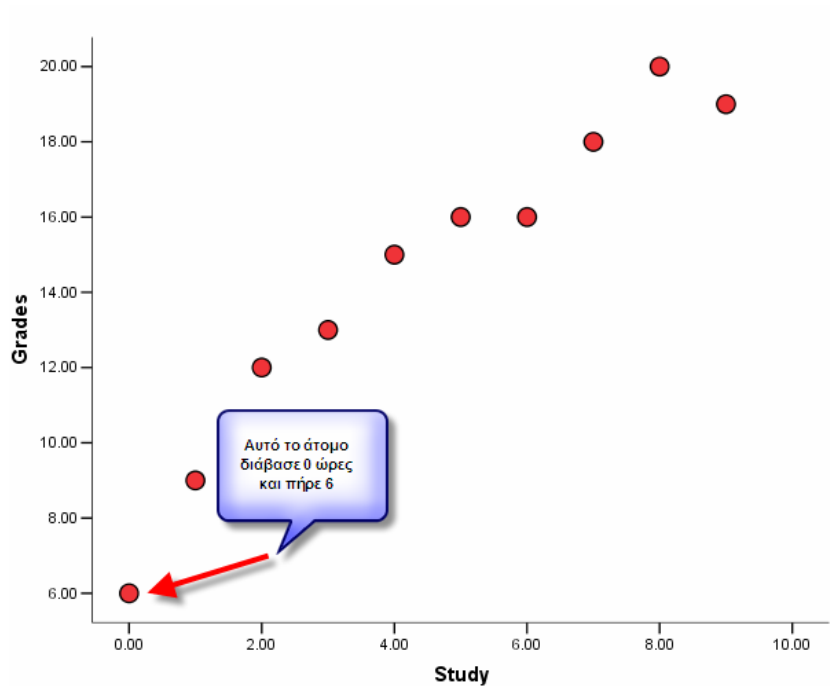
Στη συνέχεια επιλέξτε τις μεταβλητές που θα τοποθετηθούν στον οριζόντιο και τον κάθετο άξονα. Τοποθετείστε τις ώρες διαβάσματος (Study) στον οριζόντιο άξονα και την επίδοση (Grades) στον κάθετο άξονα¹¹. Στη συνέχεια επιλέξτε OK.

¹¹ Στην περίπτωση της απλής συσχέτισης δεν έχει ιδιαίτερη σημασία το ποια μεταβλητή θα τοποθετηθεί σε ποιόν άξονα αφού η σχέση $A - B$ είναι ίδια με την $B - A$.

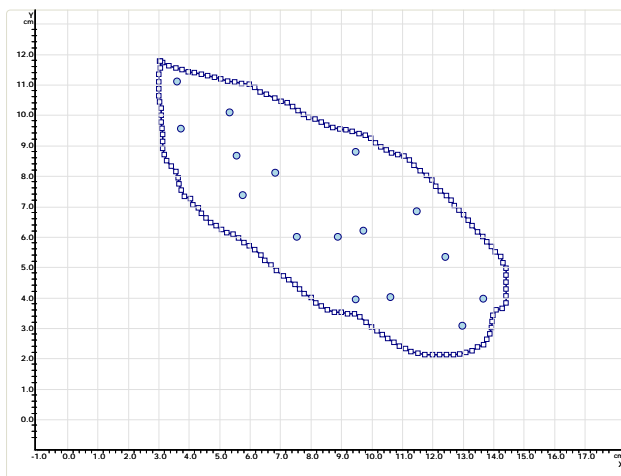


Σχήμα 5: Επιλογή των μεταβλητών σε άξονες για τη δημιουργία απλού διαγράμματος σκεδασμού.

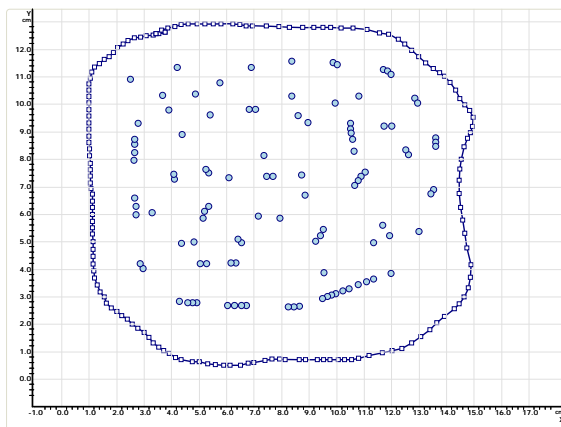
Το Σχήμα 6 δείχνει το διάγραμμα σκεδασμού όπως προκύπτει από το SPSS για τη σχέση μεταξύ ωρών διαβάσματος και επίδοσης. Όπως φαίνεται στο Σχήμα 6, όσο περισσότερο διάβασε κάποιος τόσο μεγαλύτερο βαθμό πήρε (με κάποιες διακυμάνσεις). Για την ακρίβεια, όποτε το σχήμα με τις παρατηρήσεις στο διάγραμμα σκεδασμού έχει την κατεύθυνση από αριστερά και κάτω προς τα δεξιά και πάνω, τότε υπάρχει θετική σχέση μεταξύ των δύο μεταβλητών. Όταν το σχήμα έχει κατεύθυνση από αριστερά και πάνω προς τα δεξιά και κάτω, τότε έχουμε αρνητική σχέση (βλέπε Σχήμα 7), ενώ όποτε μας είναι δύσκολο να εντοπίσουμε οπτικά κάποια κατεύθυνση (όπως η περίπτωση να είναι οι παρατηρήσεις διάσπαρτες στο χώρο) τότε έχουμε μηδενική σχέση (Σχήμα 8).



Σχήμα 6: Διάγραμμα σκεδασμού για τη σχέση ωρών διαβάσματος και επίδοσης.



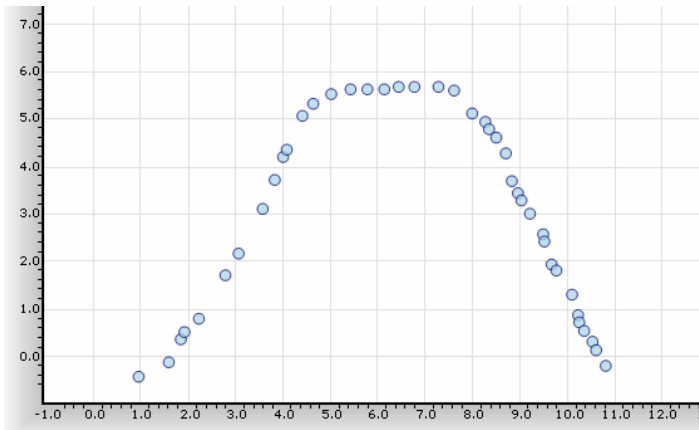
Σχήμα 7: Διάγραμμα σκεδασμού για την ύπαρξη αρνητικής σχέσης μεταξύ 2 μεταβλητών.



Σχήμα 8: Διάγραμμα σκεδασμού για την απουσία σχέσης μεταξύ 2 μεταβλητών.

2.2. Προϋπόθεση της γραμμικότητας

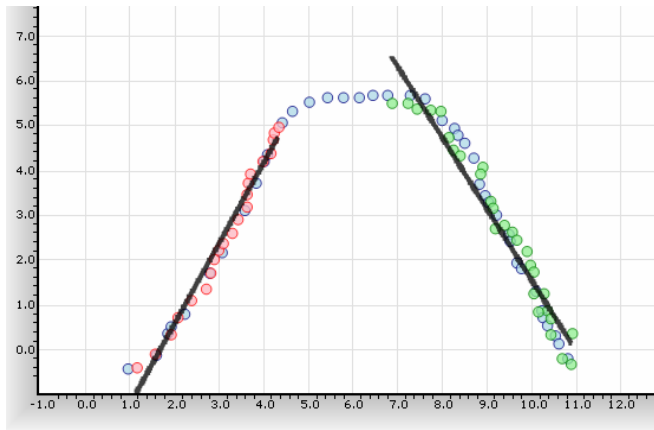
Όπως αναφέρθηκε προκειμένου να υπολογιστεί ο συντελεστής συσχέτισης θα πρέπει να ισχύει η προϋπόθεση της γραμμικότητας. Με βάση αυτή την προϋπόθεση, θα πρέπει το μοντέλο της ευθείας γραμμής να ταιριάζει καλύτερα στα δεδομένα. Με άλλα λόγια, αν τραβήξουμε μια ευθεία γραμμή στα δεδομένα του Σχήματος 6, θα δούμε ότι τα περισσότερα από αυτά είναι πάνω, ή πολύ κοντά στη γραμμή. Επομένως φαίνεται και οπτικά ότι υπάρχει αναλογική (γραμμική) σχέση μεταξύ των δύο μεταβλητών. Μια από τις περιπτώσεις μη ύπαρξης γραμμικής σχέσης αναφέρεται στη σχέση μεταξύ άγχους και επίδοσης. Με βάση σχετικές θεωρίες (π.χ., Yerkes & Dowdson, 1908) χαμηλά επίπεδα άγχους σχετίζονται με χαμηλές επιδόσεις, μεσαία επίπεδα άγχους μπορεί να σχετίζονται με εξαιρετικές επιδόσεις (επειδή μπορεί να παρέχουν ιδανικά επίπεδα ενεργοποίησης), ενώ πολύ ψηλά επίπεδα άγχους είναι καταστροφικά για την επίδοση (εξαιτίας της ανικανότητας του ατόμου να αυτό-ρυθμίσει τη συμπεριφορά του και τα συναισθήματά του). Αυτό το είδος σχέσης, φαίνεται να εκφράζεται καλύτερα με την ύπαρξη καμπυλόγραμμης σχέσης όπου η σχέση άγχους και επίδοσης είναι γραμμική ως ένα σημείο. Στη συνέχεια, αύξηση των επιπέδων άγχους συνδυάζεται με μείωση της επίδοσης (βλέπε Σχήμα 9).



Σχήμα 9: Διάγραμμα σκεδασμού που εκφράζει την υποθετική σχέση μεταξύ άγχους και επίδοσης.

Όπως φαίνεται στο Σχήμα 9, υπάρχουν κάποια ιδανικά επίπεδα άγχους (περίπου στο μέσο του οριζόντιου άξονα) στα οποία επιτυγχάνονται οι υψηλότερες επιδόσεις. Αντίθετα όταν ξεπεραστούν αυτά τα επίπεδα, τότε η επίδοση μειώνεται δραματικά. Η απουσία της προϋπόθεσης της γραμμικότητας γίνεται φανερή στο Σχήμα 10 όπου είναι σχεδόν αδύνατο να τραβήξουμε μια ευθεία γραμμή η οποία να ταιριάζει στα περισσότερα δεδομένα. Βλέπουμε ότι η πρώτη γραμμή ταιριάζει στο πρώτο μισό των παρατηρήσεων ενώ η δεύτερη γραμμή στο δεύτερο μισό. Αν λοιπόν εφαρμόζαμε την πρώτη ευθεία γραμμή θα βγάγαμε το συμπέρασμα της θετικής συνάφειας ενώ με τη δεύτερη το συμπέρασμα της αρνητικής. Επειδή λοιπόν η σχέση των δύο μεταβλητών αρχίζει να διαφοροποιείται (από θετική προς αρνητική) μετά από κάποιο επίπεδο της μιας μεταβλητής (μεσαία επίπεδα άγχους) δεν είναι κατάλληλη η εφαρμογή του συντελεστή συσχέτισης Pearson το οποίο απαιτεί την ευθύγραμμη σχέση μεταξύ των 2 μεταβλητών.

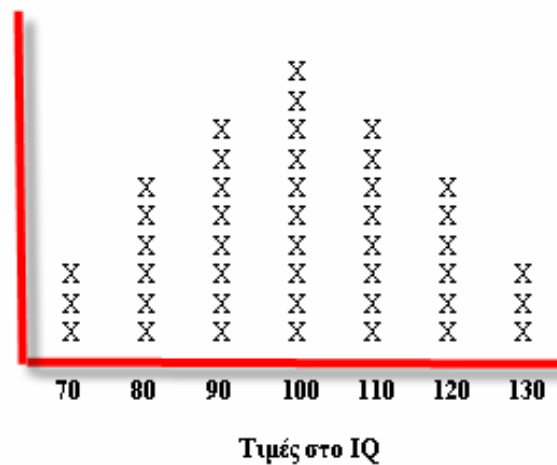
Η παραπάνω καμπυλόγραμμη σχέση δεν είναι η μοναδική. Η διερεύνηση μη γραμμικών σχέσεων δεν θα μας απασχολήσει. Για όσους ενδιαφέρονται περισσότερο, το SPSS μπορεί πολύ εύκολα να μοντελοποιήσει μια σειρά μη γραμμικών σχέσεων όπως στην δεύτερη και τρίτη δύναμη, εκθετικές, κλπ. Αυτή η σειρά αναλύσεων ανήκει στο μενού «Regression-curve estimation». Εναλλακτικά, οι σχέσεις αυτές μπορούν να μοντελοποιηθούν απευθείας στο διάγραμμα σκεδασμού (fit functions).



Σχήμα 10: Διάγραμμα σκεδασμού που εκφράζει τη σχέση μεταξύ άγχους και επίδοσης με την εφαρμογή προβλεπτικών γραμμών.

2.3. Προϋπόθεση της κανονικής κατανομής

Επίσης, προκειμένου να είναι έγκυρα τα αποτελέσματα του συντελεστή συσχέτισης θα πρέπει οι παρατηρήσεις (δηλ. τα σκορ των συμμετεχόντων) σε κάθε μια από τις 2 μεταβλητές να κατανέμονται κανονικά. Αν και η έννοια της κανονικής κατανομής είναι προαπαιτούμενη της κατανόησης της συσχέτισης μια μικρή συζήτηση ώστε να γίνει κατανοητή η προϋπόθεση ακολουθεί. Η κανονική κατανομή αποτελεί ένα διάγραμμα συχνοτήτων και θα πρέπει να ακολουθεί ένα σχήμα σαν αυτό της «καμπάνας», του οποίου το περίγραμμα εκφράζει τον αριθμό περιπτώσεων που εμπίπτουν σε κάθε επίπεδο της μεταβλητής που μελετάμε. Έτσι για παράδειγμα αν θέλουμε να δούμε πως κατανέμεται το φαινόμενο της ευφυΐας θα περίμενε κανείς να δει μια εικόνα, περίπου όπως την παρακάτω, γνωρίζοντας ότι ο μέσος όρος ευφυΐας με τη χρήση σταθμισμένων τεστ είναι 100 και η τυπική απόκλιση περίπου 15:



Σχήμα 11: Διάγραμμα συχνοτήτων όπου δείχνει μια υποθετική κανονική κατανομή.

Από το Σχήμα 11 φαίνεται ότι οι περισσότερες παρατηρήσεις (δηλ. συμμετέχοντες) βρίσκονται κοντά στο μέσο όρο (μεταξύ 90-100) και ότι όσο απομακρυνόμαστε από το μέσο όρο τόσο μειώνεται και ο αριθμός των παρατηρήσεων. Δηλαδή υπάρχουν λιγότερα άτομα που είτε είναι πάρα πολύ έξυπνα ή το αντίθετο, ενώ για τους περισσότερους από εμάς περιμένουμε (και ελπίζουμε) ότι θα είμαστε κοντά στο μέσο όρο. Η καταπάτηση αυτής της προϋπόθεσης έχει επίσης σημαντικές αρνητικές επιπτώσεις στον υπολογισμό του συντελεστή συσχέτισης, οι οποίες γίνονται φανερές αν μειωθεί πολύ το εύρος κάποιας από τις δύο μεταβλητές που μελετώνται ή και των δύο. Για παράδειγμα αν φανταστούμε στο Σχήμα 11 ότι όλες οι παρατηρήσεις είναι μεταξύ 120 και 130 (επειδή π.χ., κάναμε μετρήσεις IQ σε πολύ έξυπνα άτομα) τότε δεν θα υπάρχει αρκετό εύρος τιμών για να μοντελοποιηθεί και να αναδειχθεί πιθανή σχέση μεταξύ των δύο μεταβλητών. Το ίδιο θα συνέβαινε αν οι περισσότερες παρατηρήσεις είχαν ομαδοποιηθεί σε χαμηλά επίπεδα διαβάθμισης της ευφυΐας (π.χ., μεταξύ 70-80).

Αξίζει να αναφερθεί ότι το εύρος των μεταβλητών και η σημαντικότητά του στον υπολογισμό του συντελεστή συσχέτισης είναι κάτι το οποίο καταπατείται συχνά στην έρευνα στις κοινωνικές επιστήμες. Θεωρητικά οι μεταβλητές που θα συσχετιστούν πρέπει να μπορούν να πάρουν τιμές από το $-\infty$ (άπειρο) έως και τα $+\infty$ (άπειρο). Τις περισσότερες φορές όμως διαβαθμίσεις τύπου Likert (συμφωνώ-διαφωνώ) όπως συνηθίζονται για τη μέτρηση στάσεων, χαρακτηριστικών προσωπικότητας, κλπ., έχουν

εξαιρετικά περιορισμένο εύρος με αποτέλεσμα να επηρεάζεται το μέγεθος του συντελεστή συσχέτισης. Εντούτοις, ο συντελεστής συσχέτισης Pearson r είναι σχετικά ακριβής όταν το εύρος των μεταβλητών είναι μικρό ή όταν οι μεταβλητές δεν κατανέμονται κανονικά (ανάλογα φυσικά με το μέγεθος του εύρους ή της στρέβλωσης). Σημαντικές παραποιήσεις του δείκτη συντελούνται επίσης όταν καταπατείται η προϋπόθεση της γραμμικότητας.

Για την αξιολόγηση της ύπαρξης της κανονικότητας κάποιος θα πρέπει να συμβουλευθεί επίσης το στατιστικό δείκτη Kolmogorov-Smirnov που βρίσκεται στο μενού των μη παραμετρικών τεστ στο SPSS και εξηγείται περισσότερο στο Παράρτημα Α. Στην παρούσα φάση, μια απλή οπτική αξιολόγηση των πινάκων συχνοτήτων μπορεί να μας δώσει μια σχετικά σαφή ιδέα για το πόσο κανονικά κατανέμεται το φαινόμενο που μετρούμε.

2.4. Υπολογισμός και ερμηνεία του συντελεστή συσχέτισης Pearson

Τι πρέπει να γνωρίζουμε για το συντελεστή συσχέτισης Pearson r ; Πρώτον ότι ανήκει στην κατηγορία των σταθμισμένων δεικτών, δηλαδή είναι στατιστικός δείκτης που έχει συγκεκριμένο εύρος με αποτέλεσμα να έχει νόημα από μόνος του και χωρίς τη σύνδεση του με επίπεδα σημαντικότητας ή άλλες πληροφορίες. Το εύρος που μπορεί να πάρει ο r είναι από -1 έως και $+1$. Το πρώτο μέγεθος μας δείχνει μια τέλεια αρνητική σχέση (αντιστρόφως ανάλογη σχέση), ενώ το δεύτερο μια τέλεια θετική (ευθέως ανάλογη) σχέση μεταξύ των δύο μεταβλητών. Το μηδέν αντιπροσωπεύει μηδενική σχέση μεταξύ δύο μεταβλητών. Είναι φυσικό ότι μια σχέση του τύπου 0.6 είναι ταυτόσημη σε μέγεθος με μια σχέση -0.6 , με τη διαφορά ότι η μια είναι θετική ενώ η άλλη αρνητική. Επίσης η παρουσία σχέσεων μεγαλύτερων του 1.0 είναι κάτι αδύνατο/άτοπο (και άρα λάθος). Μεταξύ των 2 ορίων (-1 έως $+1$), έχουν προταθεί διάφορες κατηγοριοποιήσεις που αναφέρονται στην αξιολόγηση του μεγέθους του δείκτη. Μια από αυτές είναι και η παρακάτω:

0.00-0.20	Μηδενική σχέση
0.21-0.40	Μικρή σχέση
0.41-0.60	Μέτρια σχέση
0.61-0.80	Δυνατή σχέση
> 0.81	Εξαιρετικά δυνατή σχέση

Αν και στην πραγματικότητα σχέσεις της τάξης του 0.80 είναι δύσκολο να παρατηρηθούν, όταν συμβαίνει αυτό δημιουργείται και ένας θεωρητικός προβληματισμός: Μήπως αυτές οι δύο μεταβλητές που συσχετίζονται κατά 0.80 μετρούν το ίδιο θεωρητικά αντικείμενο/φαινόμενο; Επομένως, οι πολύ δυνατές σχέσεις θέλουν προσοχή στην αξιολόγησή τους και στην ερμηνεία τους.

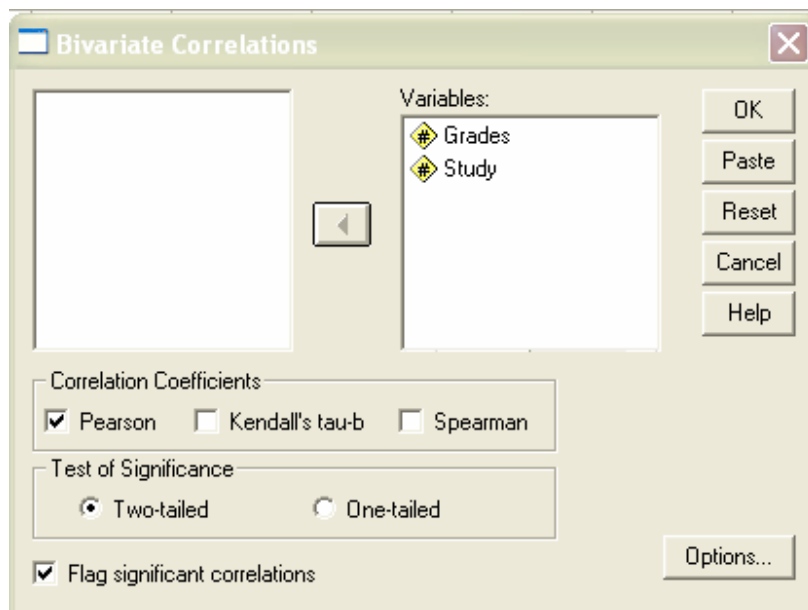
Πώς μπορούμε να υπολογίσουμε το συντελεστή συσχέτισης Pearson r

Πριν την εποχή των υπολογιστών όλες οι πράξεις γίνονταν με το χέρι με αποτέλεσμα να απαιτείται και η κατανόηση των μαθηματικών μοντέλων. Σήμερα αυτό δεν είναι απαραίτητο αφού η χρήση στατιστικών πακέτων έχει αντικαταστήσει αυτή την ανάγκη. Ο τύπος για τον υπολογισμό του συντελεστή συσχέτισης Pearson r έχει ως εξής:

$$r = \frac{\sum (X - \mu_X)(Y - \mu_Y)}{N \sigma_X \sigma_Y}$$

Όπου οι τιμές X και Y αναφέρονται στις τιμές κάθε συμμετέχοντα ενώ οι τιμές μ_X και μ_Y με τους μέσους όρους στις μεταβλητές X και Y . Επομένως η τιμή κάθε συμμετέχοντα στην μεταβλητή X αφαιρείται από το μέσο όρο αυτής της μεταβλητής και όλες αυτές οι τιμές προστίθενται. Το ίδιο γίνεται και για τη μεταβλητή Y . Το γεγονός ότι αυτά τα ζευγάρια τιμών πολλαπλασιάζονται μεταξύ τους εκφράζει τη «συνδιακύμανση» των δύο μεταβλητών. Έτσι για παράδειγμα αν υπάρχει αρνητική σχέση μεταξύ κατάθλιψης και επίδοσης στο σχολείο, τότε ένα άτομο που θα έχει τιμή μεγαλύτερη από το μέσο όρο στην κατάθλιψη, το πιθανότερο είναι ότι θα έχει μικρότερη τιμή από το μέσο όρο στη μεταβλητή επίδοση. Το γινόμενο αυτό θα είναι αρνητικό, και αν αυτό συμβαίνει για τους περισσότερους μαθητές τότε θα έχουμε αρνητική σχέση μεταξύ κατάθλιψης και επίδοσης. Ο παρανομαστής του κλάσματος εκφράζει το γινόμενο των δύο σφαλμάτων το οποίο διορθώνεται για το μέγεθος του δείγματος. Υπάρχουν πολύ πιο εύκολες τυπολογίες για τον υπολογισμό του συντελεστή συσχέτισης με το χέρι, όμως η παραπάνω είναι καλύτερη για την εξήγηση του τι ακριβώς υπολογίζει ο συντελεστής συσχέτισης που δεν είναι τίποτε άλλο από την ύπαρξη της συνδιακύμανσης μεταξύ δύο μεταβλητών.

Για τον υπολογισμό του συντελεστή συσχέτισης με το SPSS επιλέγουμε Analyze → Correlate → Bivariate. Στο παράθυρο τους Σχήματος 12 επιλέγουμε τις μεταβλητές που θέλουμε να συσχετίσουμε. Ο συντελεστής συσχέτισης Pearson r είναι προ-επιλεγμένος, ενώ όπως θα δούμε αργότερα, για την διερεύνηση μη-παραμετρικών σχέσεων μπορούν να επιλεγθούν οι δείκτες των Spearman και Kendall. Επίσης βλέπουμε ότι προ-επιλεγμένο είναι και το είδος τεστ: δίπλευρο¹².



Σχήμα 12: Παράθυρο για τον υπολογισμό του συντελεστή συσχέτισης Pearson r στο SPSS.

¹² Το μονόπλευρο τεστ χρησιμοποιείται στην περίπτωση που ο ερευνητής αξιολογεί μια συγκεκριμένη υπόθεση η οποία έχει κατεύθυνση. Για παράδειγμα η υπόθεση: «Υπάρχει σχέση μεταξύ προσκόλλησης και επίδοσης» δεν δηλώνει κάποια κατεύθυνση και ο ερευνητής οφείλει να διερευνήσει το είδος της σχέσης και στις δύο πλευρές της κατανομής (αν δηλ. η σχέση είναι θετική ή αρνητική). Στην διερεύνηση όμως για παράδειγμα της σχέσης άγχους και κατάθλιψης ο ερευνητής είναι σχεδόν σίγουρος ότι η σχέση είναι θετική και αυτό που τον απασχολεί είναι το μέγεθος αυτής της σχέσης. Έτσι, μπορεί να επιλέξει τη χρήση του μονόπλευρου τεστ. Καλό είναι να χρησιμοποιείται το δίπλευρο τεστ αφού το μονόπλευρο εκφράζει κάποια προκατάληψη (από την πλευρά του ερευνητή) και επίσης δημιουργεί και τις προϋποθέσεις για να ξεπεράσουν τα ευρήματα όρια στατιστικής σημαντικότητας.

		Study	Grades
Study	Pearson Correlation	1	.964**
	Sig. (2-tailed)		.000
	N	10	10
Grades	Pearson Correlation	.964**	1
	Sig. (2-tailed)	.000	
	N	10	10

** . Correlation is significant at the 0.01 level

Σχήμα 13: Αποτελέσματα από τη χρήση του συντελεστή συσχέτισης Pearson r στο SPSS.

Από τα παραπάνω φαίνεται πως η σχέση μεταξύ ωρών διαβάσματος και επίδοσης για τους 10 μαθητές που αξιολογήθηκαν ήταν 0.964 η οποία είναι στατιστικά σημαντική όπως φαίνεται και από την παρατηρούμενη πιθανότητα: $p = .000$. Επειδή πιθανότητα ίση με το μηδέν δεν υφίσταται, η παραπάνω πιθανότητα είναι απλά μικρότερη από 1 τοις 1000. Αφού η πιθανότητα είναι μικρότερη από το συνηθισμένο επίπεδο στατιστικής σημαντικότητας $\alpha = 5\%$ (ή όποιο άλλο επίπεδο ο ερευνητής χρησιμοποιήσει), δεχόμαστε την εναλλακτική υπόθεση σύμφωνα με την οποία υπάρχει στατιστικά σημαντική συνάφεια μεταξύ «ωρών διαβάσματος» και της «επίδοσης».

Η παρατηρούμενη σχέση στο παράδειγμα ήταν ιδιαίτερος δυνατή ώστε εντοπίστηκε παρόλο το μικρό μέγεθος του δείγματος. Χωρίς να μπορούμε σε μαθηματικές λεπτομέρειες τονίζουμε ότι τόσο τα πολύ μικρά (μικρότερα του 10 περίπου) όσο και τα πολύ μεγάλα δείγματα (μεγαλύτερα του 500 περίπου) δημιουργούν πολλά προβλήματα στην αξιοπιστία της στατιστικής σημαντικότητας του συντελεστή συσχέτισης. Σε παρόμοιες περιπτώσεις (οι πιο πολλές εκ των οποίων εμφανίζονται σε πολύ μεγάλα δείγματα) καλό είναι να περιοριζόμαστε σε μία ποιοτική ερμηνεία της τιμής του συντελεστή συσχέτισης (όπως αυτή που περιγράφεται στην αρχή της παραγράφου) και να αγνοείται η πιθανότητα. Πάντως, για την σωστή διερεύνηση σχέσεων ο ερευνητής θα πρέπει να επιλέξει έναν ιδανικό αριθμό συμμετεχόντων προκειμένου να μπορούν να

επαληθευτούν εναλλακτικές υποθέσεις που είναι αληθείς. Πληροφορίες και σχετικοί πίνακες μπορούν να αναζητηθούν στις εξαιρετικές δουλειές του Jacob Cohen (1988).

2.5. Μερικές παρατηρήσεις για το συντελεστή συσχέτισης Pearson

Τι πρέπει να έχετε υπόψη σας/να προσέχετε όταν χρησιμοποιείτε τον συντελεστή συσχέτισης Pearson r :

1. Επηρεάζεται από ακραίες παρατηρήσεις. Αν έχετε άτομα που έχουν πολύ μεγάλες ή πολύ μικρές τιμές είναι πολύ πιθανό αυτές να διαστρεβλώνουν το μέγεθος του συντελεστή συσχέτισης, σε εξαιρετικά μεγάλο βαθμό.
2. Επηρεάζεται, όπως και όλα τα τεστ, από το μέγεθος του δείγματος. Πολύ μικρά ή πολύ μεγάλα δείγματα δημιουργούν προβλήματα στην ερμηνεία της πιθανότητας και της στατιστικής σημαντικότητας.
3. Η αξιοπιστία του «πλήττεται» από την καταπάτηση της προϋπόθεσης της γραμμικότητας οπότε εξετάστε μέσα από το διάγραμμα σκεδασμού την ύπαρξη μη γραμμικών σχέσεων.
4. Επηρεάζεται, όπως και όλα τα τεστ, από την αξιοπιστία των μετρήσεων. Αν οι μετρήσεις έχουν μεγάλες τιμές σε στατιστικό σφάλμα, το πιθανότερο είναι ότι το σφάλμα αυτό θα «φορτώσει» και στον συντελεστή συσχέτισης, με απρόβλεπτες συνέπειες (να τον επηρεάσει δηλαδή και προς τα πάνω, αν και συνήθως θα πρέπει να τον επηρεάζει προς τα κάτω¹³).
5. Είναι ακατάλληλος όταν οι μεταβλητές δεν είναι συνεχείς αλλά διακριτές. Στην δεύτερη περίπτωση συντελεστές όπως ο Φ ή ο point-biserial είναι πιο κατάλληλοι ανάλογα με το αν η μία ή και οι δύο μεταβλητές είναι κατηγορικές.
6. Επηρεάζεται από την ύπαρξη στρεβλών κατανομών, ειδικά όταν αυτές αποκλίνουν κατά πολύ από την κανονικότητα. Αν καταπατείται κάποια από τις

¹³ Υπάρχουν πολλοί λόγοι που αυξάνουν το σφάλμα της μέτρησης, άλλοι στατιστικοί και άλλοι μεθοδολογικοί. Το πιθανότερο είναι ότι αύξηση στο μέγεθος τυχαίων παραγόντων και αναξιοπιστίας του τεστ θα μειώσουν τον συντελεστή συσχέτισης ενώ αύξηση του μεγέθους συστηματικού σφάλματος (π.χ., όταν οι μαθητές «κλέβουν») μπορεί να αυξήσει τον συντελεστή εξαιτίας της μείωσης των ατομικών διαφορών (τυπικής απόκλισης).

προϋποθέσεις εξετάστε¹⁴ την χρήση μη παραμετρικών κριτηρίων, τα οποία συζητώνται περιεκτικά παρακάτω.

7. Δεν μπορεί να χρησιμοποιηθεί για τη διερεύνηση αιτιωδών σχέσεων. Αυτό δεν σημαίνει ότι δεν υπάρχουν αιτιακές σχέσεις μεταξύ των μεταβλητών που μελετήθηκαν. Απλά ότι η χρήση του συντελεστή συσχέτισης δεν επιτρέπει αυτό το συμπέρασμα. Το μόνο συμπέρασμα που μπορεί να προκύψει από τη χρήση του συντελεστή συσχέτισης είναι ότι οι 2 μεταβλητές συνδιακυμαίνονται. Μόνο η μελλοντική χρήση πειραματικών σχεδίων μπορεί να επιβεβαιώσει αν οι αρχικές αυτές συνάφειες έχουν αιτιακό χαρακτήρα.

2.5. Μη-παραμετρικοί συντελεστές συσχέτισης (Spearman)

Όταν οι παραμετρικές προϋποθέσεις δεν ικανοποιούνται (δηλ. η κανονικότητα και η γραμμικότητα, το εύρος των παρατηρήσεων και η ύπαρξη ισοδιαστημικής κλίμακας), τότε πρέπει να χρησιμοποιηθούν εναλλακτικοί στατιστικοί δείκτες για την ανίχνευση σχέσεων μεταξύ μεταβλητών. Ένας από αυτούς είναι και ο δείκτης συσχέτισης του Spearman όπου μετατρέπει τα δεδομένα σε «σειρές» με βάση το μέγεθός τους (δηλ. οι αρχικές τιμές μπαίνουν σε σειρά με βάση το μέγεθός τους: πρώτος, δεύτερος, τρίτος, κλπ.). Αυτό έχει σαν αποτέλεσμα οι αποστάσεις μεταξύ των παρατηρήσεων να χάνουν τη σημασία τους και να αξιολογείται η σειρά των συμμετεχόντων στην πρώτη μεταβλητή σε σχέση με τη σειρά που αυτοί έχουν στην δεύτερη μεταβλητή. Το μέγεθος της συμφωνίας ή όχι της σειράς στις δύο μεταβλητές εκφράζει και το πρόσημο αλλά και το μέγεθος της σχέσης. Για παράδειγμα αν κάποιος που είναι πρώτος στις επιδόσεις είναι και πρώτος στη δημοτικότητα και ακολουθείται από άτομα μικρότερων επιδόσεων και μικρότερης δημοτικότητας θα αναδείξει μια θετική συνάφεια μεταξύ επίδοσης και δημοτικότητας.

Η εξίσωση για τον υπολογισμό του συντελεστή συσχέτισης Spearman είναι ως εξής:

$$1 - \frac{6 \sum D^2}{N(N^2 - 1)}$$

¹⁴ Η καταπάτηση των προϋποθέσεων δεν επιβάλλει την χρήση μη-παραμετρικών κριτηρίων. Η απόφαση αυτή θα πρέπει να είναι αποτέλεσμα προσεκτικής σκέψης μιας και η καταπάτηση των προϋποθέσεων

όπου N είναι ο αριθμός των ζευγαριών και D η διαφορά στη σειρά μεταξύ πρώτης και δεύτερης μέτρησης (δηλ. των ζευγών). Στο λογισμικό SPSS, η χρήση του δείκτη Spearman είναι πολύ απλή αφού το μόνο που χρειάζεται είναι η επιλογή «Spearman» στο μενού «correlate→bivariate». Άλλος σχετικός μη παραμετρικός δείκτης είναι ο «τ» του Kendall ο οποίος δε θα αναλυθεί.

Στο παραπάνω παράδειγμα η εφαρμογή του συντελεστή Spearman έδωσε τα εξής αποτελέσματα:

			Study	Grades
Spearman's rho	Study	Correlation Coefficient	1.000	.985**
		Sig. (2-tailed)	.	.000
		N	10	10
	Grades	Correlation Coefficient	.985**	1.000
		Sig. (2-tailed)	.000	.
		N	10	10

** . Correlation is significant at the 0.01 level (2-tailed).

Σχήμα 14: *Συσχέτιση μεταξύ ωρών διαβάσματος και επίδοσης με το δείκτη Spearman.*

Όπως φαίνεται στο Σχήμα 14, η συνάφεια μεταξύ των δύο μεταβλητών ήταν επίσης πολύ δυνατή $r = .985$ και στατιστικά σημαντική ($p < .001$). Δεν περιμένουμε να δούμε πολύ μεγάλες διαφοροποιήσεις μεταξύ των 2 δεικτών, ειδικά αν η καταπάτηση των προϋποθέσεων δεν έχει γίνει σε μεγάλο βαθμό. Στην παρούσα περίπτωση ο δείκτης Spearman είναι και λίγο μεγαλύτερος από τον δείκτη Pearson, αν και συνήθως περιμένουμε πιο «συντηρητικά» ευρήματα από τη χρήση μη παραμετρικών δοκιμασιών¹⁵.

μπορεί επίσης να οφείλεται και σε άλλους λόγους.

¹⁵ Συνήθως οι μη-παραμετρικοί δείκτες είναι πιο «αυστηροί» στα αποτελέσματά τους αφού στόχο έχουν να διορθώσουν για πιθανά προβλήματα που προκαλούνται από τις καταπατήσεις των προϋποθέσεων. Αυτό βέβαια δεν συμβαίνει συνέχεια όπως φάνηκε και παραπάνω.

3. ΑΝΑΛΥΣΗ ΠΑΛΙΝΔΡΟΜΗΣΗΣ

Μπορεί το οικογενειακό περιβάλλον να προβλέψει τις επιδόσεις, συμπεριφορές ή τις διαπροσωπικές σχέσεις των παιδιών στο σχολείο; Μπορεί ο χρόνος διαβάσματος να προβλέψει/εξηγήσει την επίδοση των μαθητών σε ένα συγκεκριμένο γνωστικό αντικείμενο; Η απάντηση των παραπάνω ερωτημάτων είναι αυτονόητα πολύ σημαντική για την σχολική εμπειρία των παιδιών και των εκπαιδευτικών και για το λόγο αυτό τα προβλεπτικά μοντέλα έχουν τόσο μεγάλη αξία και διάδοση. Η ανάλυση συσχέτισης για την οποία έγινε λόγος παραπάνω, δεν μπορεί να απαντήσει σε ερωτήματα αναφορικά με την πρόβλεψη, επισημαίνει απλώς τη συσχέτιση μεταξύ δύο μεταβλητών.

Το πιο απλό προβλεπτικό μοντέλο είναι η **ανάλυση παλινδρόμησης**, η οποία αποτελεί φυσική επέκταση του συντελεστή συσχέτισης. Οι διαφορές με το συντελεστή συσχέτισης είναι οι εξής:

1. Χρησιμοποιείται για την πρόβλεψη των τιμών ορισμένων μεταβλητών και
2. Μπορεί να χρησιμοποιηθούν περισσότερες από 2 ανεξάρτητες μεταβλητές

Ας εξετάσουμε το ενδεχόμενο μιας απλής πρόβλεψης: θέλουμε να εξετάσουμε αν οι ώρες διαβάσματος είναι σε θέση να προβλέψουν την επίδοση των μαθητών στο σχολείο. Ως εργαλείο μέτρησης θα χρησιμοποιηθεί ένα κριτήριο αξιολόγησης μαθηματικών γνώσεων, το οποίο αποτελείται από 30 ερωτήσεις - ασκήσεις. Τα δεδομένα προέρχονται από 19 μαθητές μιας τάξης και οι υποθέσεις έχουν ως εξής¹⁶:

H_0 : Οι ώρες διαβάσματος **δεν** προβλέπουν την επίδοση

$H_ε$: Οι ώρες διαβάσματος προβλέπουν την επίδοση

Τα δεδομένα από τους 19 συμμετέχοντες και έχουν ως εξής:

Πίνακας 2: Δεδομένα από μια υποθετική μελέτη αναφορικά με την πρόβλεψη της επίδοσης στα μαθηματικά ως συνάρτηση των ωρών διαβάσματος

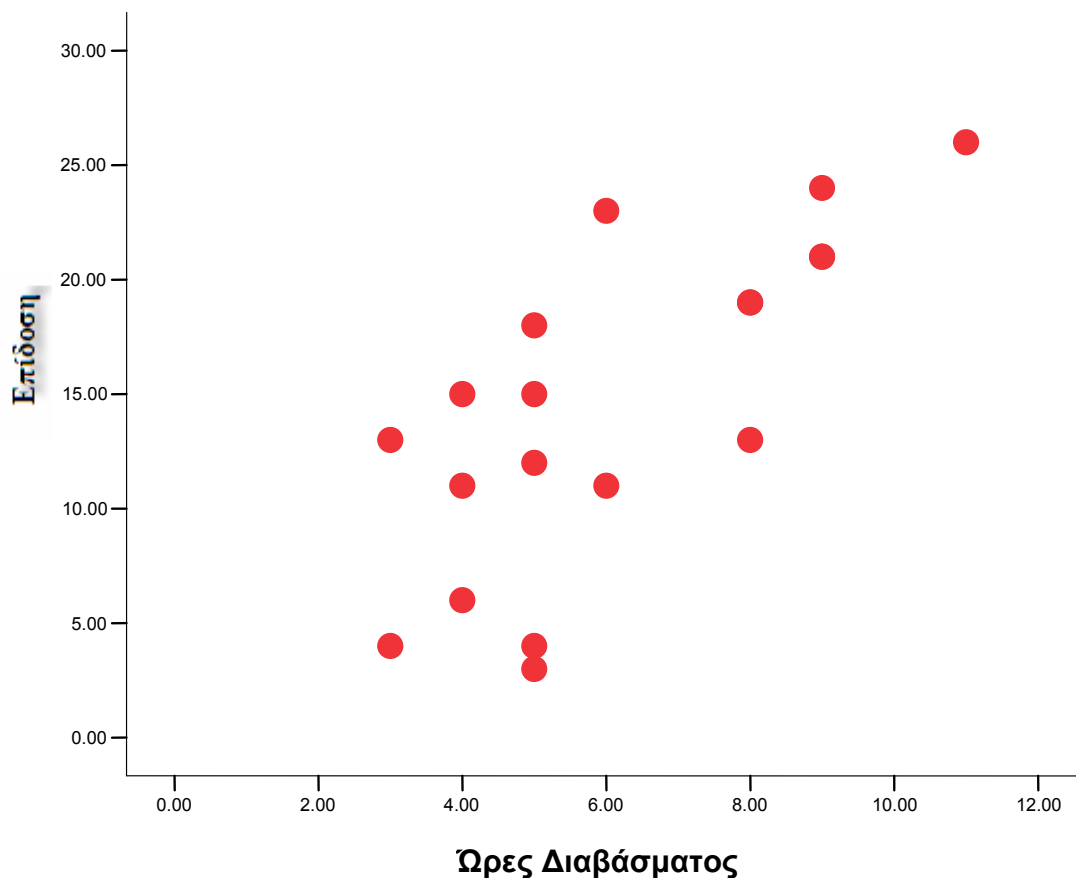
Συμμετέχοντες	1	2	3	4	5	6	7	8	9	10
Ώρες διαβάσματος	11	9	9	9	8	8	8	6	6	5
Επίδοση	26	21	24	21	19	13	19	11	23	15

¹⁶ Θα πρέπει οι χρήστες στατιστικών αναλύσεων να συνηθίσουν στην ιδέα ότι η στατιστική είναι επικουρική και μόνο της ερευνητικής διαδικασίας και έχει νόημα μόνο για την απάντηση συγκεκριμένων ερευνητικών ερωτήσεων. Επομένως πρώτα θα πρέπει να γίνει η διατύπωση των υποθέσεων και μετά η επιλογή του κατάλληλου στατιστικού κριτηρίου που θα δώσει απάντηση στη (συγκεκριμένη) υπόθεση.

Το πρώτο βήμα, όπως και στην ανάλυση συσχέτισης είναι η δημιουργία του διαγράμματος σκεδασμού το οποίο, πέρα από την οπτική ανίχνευση της ύπαρξης σχέσης, θα μας δώσει πληροφορίες και για την ύπαρξη ή όχι γραμμικότητας. Ακολουθώντας τα βήματα που περιγράφηκαν στην ανάλυση συσχέτισης κατά Pearson, δημιουργούμε το παρακάτω διάγραμμα σκεδασμού. Προσοχή: ότι οι παραπάνω υποθέσεις υποδηλώνουν ποια είναι η εξαρτημένη μεταβλητή (αποτέλεσμα) και ποια η ανεξάρτητη¹⁷ (προβλεπτικός παράγοντας). Επομένως στη δημιουργία του διαγράμματος σκεδασμού τοποθετούμε την εξαρτημένη μεταβλητή στον κάθετο άξονα (Y) και την ανεξάρτητη στον οριζόντιο (X).

Από το Σχήμα 15 είναι φανερό ότι υπάρχει θετική συνάφεια μεταξύ των δύο μεταβλητών από τη στιγμή που οι υψηλές τιμές της μιας φαίνεται να σχετίζονται με υψηλές τιμές της άλλης. Επομένως η προϋπόθεση της γραμμικότητας φαίνεται να ικανοποιείται. Σχετικά με την προϋπόθεση της κανονικότητας (και χρησιμοποιώντας το τεστ Kolmogorov-Smirnov), διαπιστώνουμε ότι σε καμιά από τις δύο μεταβλητές (επίδοση ή ώρες διαβάσματος) δεν ακυρώνεται η συγκεκριμένη προϋπόθεση. Ο δείκτης K-S για τις ώρες διαβάσματος είναι 0.941 ενώ για τη επίδοση 0.457 και κανένας τους δεν ξεπερνά τα επίπεδα στατιστικής σημαντικότητας (πιθανότητες 0.34 και 0.99, αντίστοιχα). Από τη στιγμή που καμία πιθανότητα δεν ήταν μικρότερη από το 5%, τότε αυτόματα δεχόμαστε τη μηδενική υπόθεση που υποστηρίζει ότι οι υπό εξέταση κατανομές (της μεταβλητής “επίδοση” και της μεταβλητή “ώρες διαβάσματος”) δεν διαφέρουν από την υποθετική κανονική κατανομή. Έτσι λοιπόν μπορούμε να προχωρήσουμε στην ανάλυση παλινδρόμησης αφού ικανοποιούνται και οι δύο προϋποθέσεις.

¹⁷ Όπως ειπώθηκε πριν, στην ανάλυση συσχέτισης η συνάφεια A-B είναι ίδιας “αξίας” με τη συνάφεια B-A από τη στιγμή που το ενδιαφέρον δεν είναι στο αν η μία μεταβλητή επιδρά στην άλλη, αλλά απλά στο αν οι δύο μεταβλητές “κινούνται” παράλληλα.



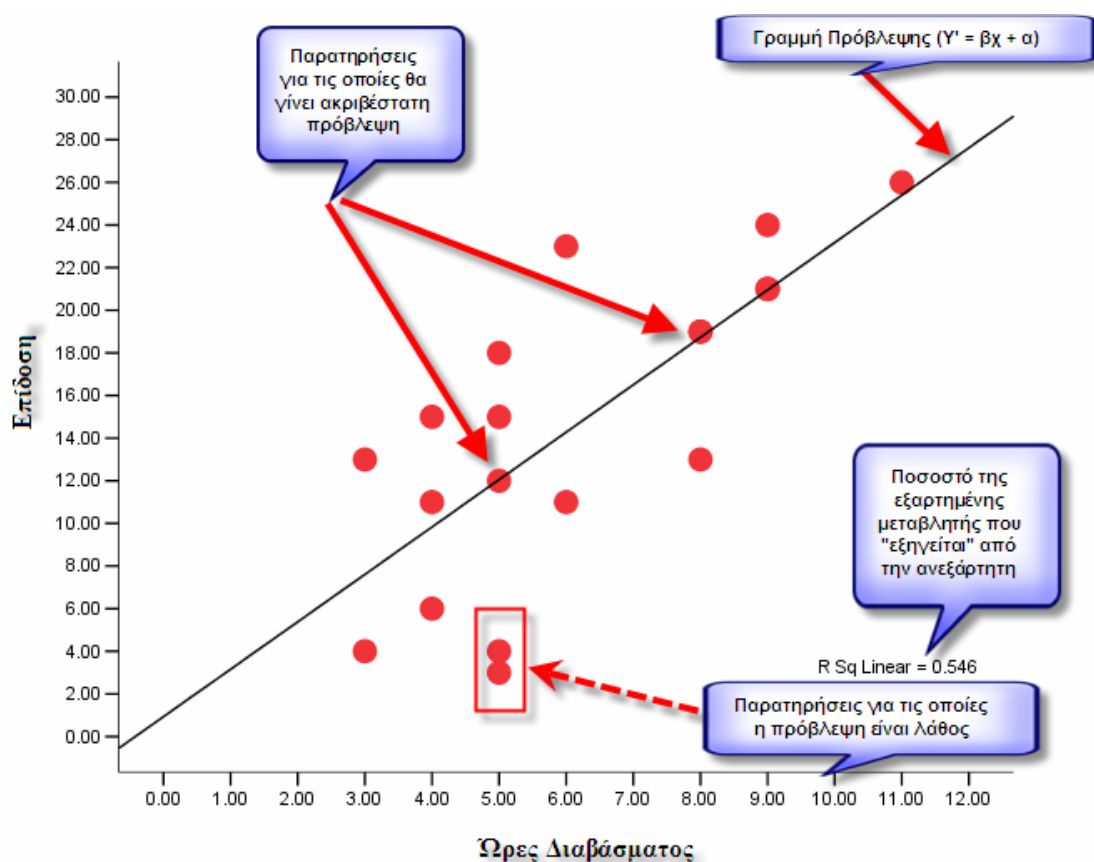
Σχήμα 15: Διάγραμμα σκεδασμού για την πρόβλεψη της σχολικής επίδοσης από τις ώρες διαβάσματος.

Το επόμενο βήμα είναι να εφαρμόσουμε το μοντέλο της ευθείας γραμμής με πρόθεση την πρόβλεψη της σχέσης ανάμεσα στην επίδοση και τις ώρες διαβάσματος. Το μοντέλο αυτό συμβολίζεται με τη συνάρτηση:

$$\hat{Y} = bX + \alpha$$

όπου \hat{Y} είναι η προβλεπτική τιμή επίδοσης και η οποία είναι συνάρτηση ενός συντελεστή κλίσης (b) και μιας σταθεράς (α). Τα σημεία α και b αποτελούν τα δύο αναγκαία σημεία προκειμένου να χαραχθεί μια και μοναδική ευθεία γραμμή η οποία

ταιριάζει όσο το δυνατόν καλύτερα στα δεδομένα μας. Στην ιδεατή περίπτωση θα έπρεπε όλες οι παρατηρήσεις να είναι πάνω στην ευθεία γραμμή, αν όμως αυτό δε συμβαίνει, τότε θα έπρεπε (προκειμένου να ισχύει η ευθύγραμμη σχέση) όσο το δυνατόν περισσότερες παρατηρήσεις να είναι πάνω στη γραμμή και οι υπόλοιπες να βρίσκονται όσο το δυνατόν πιο κοντά της (προς τα πάνω ή προς τα κάτω). Το a (σταθερά) της συνάρτησης είναι το σημείο όπου η ευθεία γραμμή της πρόβλεψης τέμνει τον κάθετο άξονα (αναφέρεται δηλαδή στην τιμή του Y όταν η τιμή του X είναι ίση με μηδέν).



Σχήμα 16: Πρόβλεψη της επίδοσης από τις ώρες διαβάσματος.

Όπως φαίνεται στο Σχήμα 16, τέσσερις περίπου παρατηρήσεις εφάπτονται με τη γραμμή της πρόβλεψης, που σημαίνει ότι για αυτά τα άτομα η πρόβλεψη θα είναι εξαιρετικά ακριβής. Όσο μεγαλύτερη είναι η απόσταση κάποιων παρατηρήσεων από την ευθεία γραμμή, τόσο περισσότερο αποτυγχάνει το μοντέλο πρόβλεψης για τις

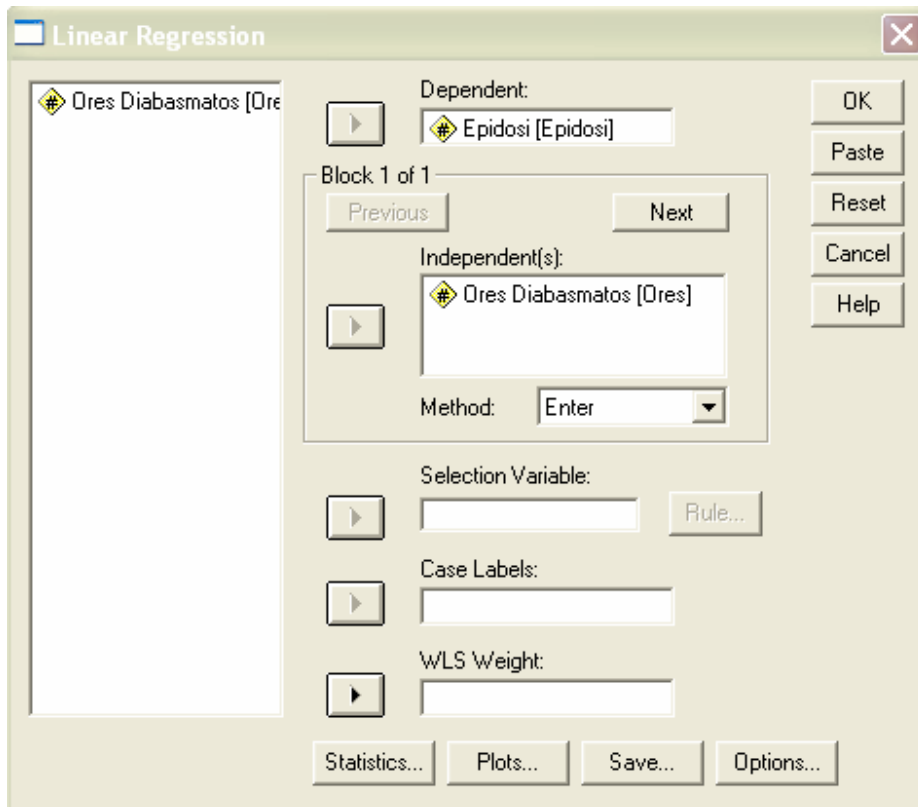
(συγκεκριμένες) παρατηρήσεις (η πρόβλεψη βασίζεται στην “επιτυχία” της συγκεκριμένης ευθείας γραμμής να περιγράψει τα δεδομένα¹⁸). Οι αποστάσεις των παρατηρήσεων από την ευθεία γραμμή αναφέρονται επίσης ως υπόλοιπα (residuals), αφού δεν αποτελούν τίποτε άλλο από την απόκλιση μεταξύ πραγματικών και αναμενόμενων (προβλεπόμενων) τιμών. Όσο μεγαλύτερες είναι οι αποστάσεις, τόσο μεγαλύτερη και η απόκλιση μεταξύ πραγματικότητας και πρόβλεψης. Για παράδειγμα για τον μαθητή-τρια που έχει διαβάσει 5 ώρες (όπως δείχνει και το διακεκομμένο βέλος), προβλέπουμε ότι θα έχει επίδοση ίση με το 12 (αν υπολογίσουμε τη νοητή ευθεία της συγκεκριμένης παρατήρησης προς την γραμμή πρόβλεψης και από τη γραμμή πρόβλεψης στον άξονα Y). Είναι φανερό ότι για τον συγκεκριμένο μαθητή-τρια η πρόβλεψη θα είναι εξαιρετικά φτωχή.

Το παραπάνω παράδειγμα εξετάζει μία μόνο ανεξάρτητη μεταβλητή X (ώρες διαβάσματος) σε σχέση με την εξαρτημένη μεταβλητή Y (επίδοση). Το μοντέλο αυτό ονομάζεται «απλή» γραμμική παλινδρόμηση (simple linear regression). Πιο ρεαλιστικά μοντέλα χρησιμοποιούν περισσότερες ανεξάρτητες μεταβλητές και ονομάζονται μοντέλα «πολλαπλής» γραμμικής παλινδρόμησης (multiple linear regression). Στις σημειώσεις θα επικεντρωθούμε στην απλή παλινδρόμηση αλλά η επέκτασή της στην πολλαπλή περίπτωση είναι άμεση.

3.1. Υπολογισμός της γραμμής παλινδρόμησης με τη χρήση του SPSS

Για τον υπολογισμό της ανάλυσης παλινδρόμησης επιλέξτε: analyze → Regression → Linear. Στην οθόνη σας εμφανίζεται το παράθυρο του σχήματος 17. Επιλέξτε και τοποθετήστε την εξαρτημένη μεταβλητή (επίδοση) στο παράθυρο με τίτλο: «Dependent» και την ανεξάρτητη μεταβλητή (ώρες διαβάσματος) στο παράθυρο με τίτλο «Independent(s)». Επιλέξτε το OK. Για λόγους απλοποίησης δεν απαιτούνται άλλες επιλογές για την βασική κατανόηση της απλής ανάλυσης παλινδρόμησης.

¹⁸ Κάποιος θα μπορούσε πολύ λογικά να ισχυριστεί ότι το πρώτο βήμα πριν την αξιολόγηση της πρόβλεψης είναι να διερευνηθεί αν υπάρχει σχέση μεταξύ των δύο μεταβλητών αφού η έλλειψη σχέσης δηλώνει και απουσία πρόβλεψης. Στο συγκεκριμένο παράδειγμα, η σχέση μεταξύ επίδοσης και ωρών διαβάσματος με τη χρήση του Pearson r ήταν 0.739. Το μέγεθος του συντελεστή συσχέτισης είναι αρκετά



Σχήμα 17: Υπολογισμός της πρόβλεψης της σχολικής επίδοσης από τις ώρες διαβάσματος.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.739(a)	.546	.519	4.88196

a Predictors: (Constant), Ores Diabasmatos

Σχήμα 18: Αποτελέσματα από την ανάλυση παλινδρόμησης (1^{ος} Πίνακας).

Όπως φαίνεται στο Σχήμα 18, ο πρώτος πίνακας που βλέπουμε περιγράφει κάποια συγκεντρωτικά δεδομένα για την ύπαρξη πρόβλεψης. Αρχικά παρουσιάζεται ο δείκτης συνάφειας των δύο μεταβλητών Pearson r , ο οποίος είναι 0.739 και δείχνει μια

μεγάλο, και στατιστικά σημαντικό, ώστε να μπορεί να υπάρχει (και ουσιαστική) πρόβλεψη για τη σχέση των δύο μεταβλητών.

δυνατή θετική συνάφεια μεταξύ ωρών διαβάσματος και επίδοσης. Ο συντελεστής συσχέτισης υπολογισμένος στη δεύτερη δύναμη (τετράγωνο), ορίζει τον «συντελεστή προσδιορισμού» και εξηγεί το ποσοστό της διακύμανσης της εξαρτημένης μεταβλητής που οφείλεται στην επίδραση της ανεξάρτητης. Το ποσοστό αυτό στο συγκεκριμένο παράδειγμα είναι 0.546, δηλ. ~55%. Με απλά λόγια το συγκεκριμένο εύρημα υποδηλώνει ότι αν γνωρίζουμε πόσες ώρες διαβάζει κάποιος μαθητής-τρια μπορούμε να κατανοήσουμε κατά 55% το φαινόμενο της σχολικής επίδοσης¹⁹. Δηλαδή, η γνώση του χρόνου μελέτης φαίνεται να είναι καθοριστικός παράγοντας εκτίμησης της επίδοσης, με την επιφύλαξη (ασφαλώς) ότι η αιτία της επίδοσης δεν είναι απαραίτητα ο χρόνος μελέτης κάποιου μαθητής-τριας. Οι ώρες μελέτης εκτιμάται ότι συμμετέχουν στην εξέταση και ερμηνεία της επίδοσης, αλλά θα πρέπει να δοθεί προσοχή σε δύο παράγοντες: (α) ότι η διερεύνηση αιτιακών σχέσεων (causal relationships) απαιτεί τη χρήση πειραματικών ερευνητικών σχεδίων, και (β) το ότι μια μεταβλητή εξηγεί το 55% της διακύμανσης μιας άλλης μεταβλητής, δεν σημαίνει ότι αποτελεί και την αιτία της αλλαγής της. Στο συγκεκριμένο παράδειγμα οι πιθανές αιτιολογήσεις μπορεί να είναι πολλές. Επομένως, αν διερευνούσαμε την προβλεπτική ικανότητα και άλλων μεταβλητών θα ήταν πιθανό να αξιολογηθεί σε μεγαλύτερο βάθος η σημαντικότητα και η αξία της μεταβλητής «ώρες διαβάσματος» ως προβλεπτικός παράγοντας της επίδοσης, αφού οι ερευνητές θα μπορούσαν να αξιολογήσουν αυτή τη συνεισφορά σαν συνάρτηση της «ταυτόχρονης» συνεισφοράς και άλλων μεταβλητών. Για παράδειγμα, η εισαγωγή μιας δεύτερης προβλεπτικής μεταβλητής (π.χ. κίνητρα) θα μπορούσε να μειώσει τα επίπεδα πρόβλεψης της μεταβλητής επίδοση (π.χ., από 55% σε 45%) αφού ένα 10% της διακύμανσης του φαινομένου της επίδοσης θα μπορούσε να οφείλεται στα κίνητρα για μάθηση ή επίτευξη.

¹⁹ Αυτό δεν είναι ισοδύναμο του γεγονότος ότι για το 55% των συμμετεχόντων θα έχουμε ακριβή πρόβλεψη, ενώ για τους υπόλοιπους συμμετέχοντες η πρόβλεψη θα είναι ατυχή.

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	487.252	1	487.252	20.444	.000 ^a
	Residual	405.169	17	23.833		
	Total	892.421	18			

a. Predictors: (Constant), Ores Diabasmatos

b. Dependent Variable: Epidosi

Σχήμα 19: Αποτελέσματα από την ανάλυση παλινδρόμησης: Τεστ σημαντικότητας (2^{ος} Πίνακας).

Το Σχήμα 19 αξιολογεί το μέγεθος των δεικτών του Μοντέλου (Σχήμα 18) για το αν ξεπερνούν ή όχι επίπεδα στατιστικής σημαντικότητας. Με άλλα λόγια αξιολογεί την υπόθεση ότι η πρόβλεψη είναι σημαντική ή τυχαία. Στον πίνακα της “ανάλυσης διακύμανσης” (ANOVA) αξιολογείται η πιθανότητα ο συντελεστής συσχέτισης r ή ο συντελεστής R^2 να είναι στατιστικώς σημαντικά διαφορετικοί από το μηδέν. Όπως φαίνεται στο Σχήμα 19, το μέγεθος τους τεστ F^{20} είναι 20.444 το οποίο για 1 και 17 βαθμούς ελευθερίας²¹ (στον αριθμητή και στον παρανομαστή, αντίστοιχα), σχετίζεται με μια πιθανότητα που είναι μικρότερη από 1 τοις 1000). Γνωρίζοντας ότι ένα στατιστικά σημαντικό εύρημα σημαίνει αυτόματα την υιοθέτηση της εναλλακτικής υπόθεσης, συμπεραίνει κανείς ότι υπάρχει στατιστικά σημαντική συσχέτιση/πρόβλεψη της επίδοσης από τις ώρες διαβάσματος.

²⁰ Το ίδιο τεστ (F) χρησιμοποιείται συνήθως για την ανίχνευση διαφορών 2 ή και περισσότερων μέσων όρων. Το ίδιο ουσιαστικά μαθηματικό μοντέλο χρησιμοποιείται και για τον υπολογισμό της πρόβλεψης.

²¹ Σε μια προσπάθεια απλοποίησης ισχύει ότι οι βαθμοί ελευθερίας του αριθμητή αντιστοιχούν στον αριθμό των προβλεπτικών μεταβλητών ενώ αυτοί του παρανομαστή περίπου στον αριθμό των παρατηρήσεων (συμμετεχόντων). Φυσικά το παραπάνω αποτελεί υπεραπλούστευση του τι ακριβώς είναι και τι ρόλο παίζουν οι βαθμοί ελευθερίας (βλ. και τη σχετική ενότητα στη σ. ...)

Coefficients^a

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
1 (Constant)	.938	3.229		.290	.775
Ores Diabasmatos	2.224	.492	.739	4.522	.000

a. Dependent Variable: Epidosi

Σχήμα 20: Αποτελέσματα από την ανάλυση παλινδρόμησης: Συντελεστές και προβλεπτική εξίσωση (3^{ος} Πίνακας από το συνολικό output του SPSS).

Στον Πίνακα 3 του Σχήματος 20 εντοπίζουμε τους συντελεστές που ορίζουν την εξίσωση της πρόβλεψης. Αξιοποιούμε την εξίσωση για να προβλέψουμε τις τιμές συγκεκριμένων συμμετεχόντων ή γενικότερα για να κάνουμε προβλέψεις για δεδομένες τιμές X (βλέπε παρακάτω «εφαρμογή της προβλεπτικής εξίσωσης»). Από το Σχήμα 20 εντοπίζουμε δύο στατιστικά τεστ που αναφέρονται στη χρήση του t-test²². Έχουμε ήδη επισημάνει ότι ένα στατιστικό τεστ θα πρέπει να εξετάζει μια συγκεκριμένη ερευνητική υπόθεση (μηδενική και εναλλακτική). Το πρώτο τεστ αξιολογεί τις παρακάτω υποθέσεις αναλογικά με την παράμετρο της σταθεράς (constant, a στην εξίσωση), η οποία όπως είπαμε παραπάνω αναφέρεται στην πρόβλεψη της τιμής του Y όταν η τιμή του X είναι μηδέν. Οι υποθέσεις λοιπόν που αξιολογούνται είναι ως εξής:

H_0 : Η τιμή της σταθεράς είναι ίση με το μηδέν

H_e : Η τιμή της σταθεράς είναι διαφορετική από το μηδέν

Επί τις ουσίας αξιολογείται το αν η τιμή 0.938 είναι ίση με το μηδέν. Από την αξιολόγηση της πιθανότητας, η οποία είναι 77.5%, φαίνεται ότι πρέπει να δεχτούμε τη μηδενική υπόθεση (θα δεχόμασταν την εναλλακτική αν η πιθανότητα ήταν μικρότερη από 5%), η οποία μας λέει ότι η σταθερά δεν είναι στατιστικά σημαντικά διαφορετική από το μηδέν. Ποια είναι λοιπόν η συνεισφορά του τεστ; Σχεδόν καμία αφού η σταθερά, αν και συνήθως είναι στατιστικά σημαντική, δεν συνεισφέρει ουσιαστικές πληροφορίες για το μέγεθος ή την κατεύθυνση της πρόβλεψης. Αντίθετα, η κλίση b μας πληροφορεί

²² Το t-test, όπως και το F-test ανήκει στα τεστ διαφορών όπου αξιολογείται αν ένας μέσος όρος είναι διαφορετικό από κάποια παράμετρο του πληθυσμού ή αν διαφέρουν μεταξύ τους δύο μέσοι όροι.

για το μέγεθος της πρόβλεψης και το είδος της σχέσης μεταξύ των δύο μεταβλητών. Η κλίση δίνεται από τη στήλη B του Σχήματος 20 μετά τη σταθερά (2,224). Επομένως το δεύτερο τεστ αναφέρεται στις εξής υποθέσεις:

H_0 : Η κλίση b είναι ίση με το μηδέν

H_ϵ : Η κλίση b είναι στατιστικώς σημαντικά διαφορετική από το μηδέν

Επειδή ο στατιστικός δείκτης t με μέγεθος 4.522 σχετίζεται με μια πιθανότητα που είναι μικρότερη από το 5%, δεχόμαστε αυτόματα την εναλλακτική υπόθεση που μας λέει ότι η κλίση είναι στατιστικώς σημαντικά διαφορετική από το μηδέν. Σημαντική κλίση σημαίνει φυσικά και σημαντική πρόβλεψη. Επομένως συμπεραίνουμε και από το συγκεκριμένο τεστ (όπως και από το F-test παραπάνω) ότι υπάρχει μια στατιστικά σημαντική πρόβλεψη της επίδοσης σε συνάρτηση με τις ώρες διαβάσματος.

Τι σημαίνει ο δείκτης b ; Ο συντελεστής b μας δείχνει πόσο μεταβάλλεται το αποτέλεσμα (στο παράδειγμά μας η επίδοση) για κάθε μια μονάδα αλλαγής της ανεξάρτητης μεταβλητής (δηλαδή των ωρών διαβάσματος). Η τιμή του b είναι 2.224 που σημαίνει ότι η αύξηση κατά μιας και μόνο ώρας διαβάσματος την ημέρα σχετίζεται με αύξηση της επίδοσης στα μαθηματικά κατά 2 περίπου μονάδες.

3.2. Εφαρμογή της προβλεπτικής εξίσωσης

Η έννοια της πρόβλεψης έχει αξία αν με πληροφορίες που έχουμε σήμερα μπορούμε να προβλέψουμε, με σχετικά μεγάλη ακρίβεια, ότι κάτι θα συμβεί στο μέλλον. Με τη χρήση λοιπόν της εξίσωσης πρόβλεψης μπορούμε να προβλέψουμε τον αριθμό των ασκήσεων στα μαθηματικά που θα λυθούν αν π.χ., κάποιος μαθητής-τρια διαβάσει 5 ώρες την ημέρα. Με βάση την παρακάτω εξίσωση και εφαρμόζοντας όπου $x = 5$, έχουμε:

$$\hat{Y} = \beta X + \alpha \Rightarrow$$

$$\hat{Y} = .938X + 2.224 \Rightarrow$$

$$\hat{Y} = .938(5) + 2.224 \Rightarrow$$

$$\hat{Y} = 6.914, \sim 7 \text{ ασκήσεις}$$

Δηλαδή, περίπου 7 ασκήσεις στα μαθηματικά είναι συνάρτηση διαβάσματος 5 ωρών την ημέρα. Μια και η αξιολόγηση στα μαθηματικά δεν ήταν σταθμισμένη, είναι λίγο δύσκολο να αξιολογηθεί αν το 7 είναι μικρός ή μεγάλος αριθμός. Ας δούμε λοιπόν πόσες

μαθηματικές ασκήσεις θα λύσει κάποιος μαθητής-τρια που διαβάζει 0 ώρες την ημέρα για τα μαθηματικά. Εφαρμόζοντας την παραπάνω προβλεπτική εξίσωση για $X = 0$, βλέπουμε ότι οι μαθητές αυτοί θα λύσουν μόνο 2 μαθηματικές ασκήσεις. Η διαφορά λοιπόν των 2 από τις 7 ασκήσεις είναι εμφανής. Από την πρόβλεψη αυτή φαίνεται πόσο σημαντικό παράγοντα της μαθηματικής επίδοσης αποτελεί το διάβασμα. Όπως αναφέρθηκε παραπάνω, η εικόνα σχετικά με την προβλεπτική ικανότητα των ωρών διαβάσματος μπορεί να αλλάξει δραματικά από την αξιολόγηση (και) άλλων ανεξάρτητων μεταβλητών.

4. ΕΛΕΓΧΟΣ ΓΙΑ ΤΟ ΜΕΣΟ ΟΡΟ

4.1. Εισαγωγή

Ένα από τα πιο ενδιαφέροντα μέτρα θέσης που χαρακτηρίζει ένα πληθυσμό ή ένα δείγμα είναι, χωρίς αμφιβολία, ο μέσος όρος. Στη μεγάλη πλειοψηφία των ποσοτικών εμπειρικών ερευνών όμως δεν υπάρχουν στοιχεία από όλο τον πληθυσμό (εκτός εάν έχουμε στοιχεία απογραφής). Αν λοιπόν μας ενδιαφέρει να εξετάσουμε την τιμή του μέσου όρου ορισμένων χαρακτηριστικών σε έναν ή περισσότερους πληθυσμούς αυτό μπορεί να γίνει μόνο με βάση τα στοιχεία του δείγματος. Ο σκοπός της παρούσας στατιστικής μεθοδολογίας είναι να παρουσιάσει τις τεχνικές μέσω των οποίων ο ερευνητής θα μπορέσει να βγάλει από το δείγμα ασφαλή συμπεράσματα για τη μέση τιμή ορισμένων χαρακτηριστικών του πληθυσμού. Προτού προχωρήσουμε θα παρουσιάσουμε ορισμένα παραδείγματα από το χώρο της Εκπαιδευτικής Έρευνας (και όχι μόνο) όπου η παραπάνω μεθοδολογία είναι χρήσιμη.

Παράδειγμα 1: Γνωρίζουμε ότι το μέσο ποσό που ξοδεύει μηνιαίως για φροντιστήρια μία οικογένεια που έχει ένα παιδί στην Γ' Λυκείου είναι 2000 ευρώ (το νούμερο είναι υποθετικό). Επιθυμούμε να διαπιστώσουμε εάν οι οικογένειες των τελειόφοιτων μαθητών μίας συγκεκριμένης πόλης ξοδεύουν κατά μέσο όρο το ίδιο ή διαφορετικό ποσό για φροντιστήρια. Η σύγκριση αφορά το μηνιαίο μέσο όρο του ποσού που ξοδεύει μία οικογένεια που κατοικεί στη συγκεκριμένη πόλη και έχει ένα παιδί στην Γ' Λυκείου *με την προκαθορισμένη γνωστή τιμή των 2000 ευρώ.*

Παράδειγμα 2: Μας ενδιαφέρει να εξετάσουμε την επίδραση που έχει η χρήση των Νέων Τεχνολογιών στην εκπαιδευτική διαδικασία. Πιο συγκεκριμένα επιθυμούμε να διαπιστώσουμε εάν οι μαθητές που διδάσκονται ορισμένα μαθήματα με τη βοήθεια των Νέων Τεχνολογιών (χρήση εκπαιδευτικού λογισμικού, κλπ) παρουσιάζουν διαφορές ως προς τη βαθμολογία τους με τους μαθητές που διδάσκονται τα ίδια μαθήματα με τον «παραδοσιακό» τρόπο. Στην προκειμένη περίπτωση, η σύγκριση γίνεται *μεταξύ των μέσων όρων βαθμολογίας δύο διαφορετικών μαθητικών πληθυσμών* (ο πληθυσμός που χρησιμοποιεί νέες τεχνολογίες σε σχέση με τον πληθυσμό που διδάσκεται κάποιο γνωστικό αντικείμενο με τον παραδοσιακό τρόπο διδασκαλίας).

Παράδειγμα 3: Τα φαινόμενα σχολικής βίας αρχίζουν να εμφανίζονται σε ολοένα και μικρότερη ηλικία. Οι ερευνητές θεωρούν ότι η προβολή βίαιων εικόνων στα παιδιά ηλικίας 12 ετών συμβάλλει σημαντικά σε αυτό. Για το σκοπό αυτό αναπτύχθηκε οπτικοακουστικό υλικό όπου υπήρχαν ποικίλα ερεθίσματα και εικόνες βίας. Πριν την έκθεση παιδιών ηλικίας 12 ετών στο συγκεκριμένο υλικό καταγράφηκαν τα κρούσματα σχολικής βίας των υπό εξέταση παιδιών κατά το τελευταίο δίμηνο. Στη συνέχεια μέσω μίας ειδικής εκδήλωσης (πχ. προβολή της ταινίας) οι μαθητές ήρθαν σε επαφή με το υλικό βίας και καταγράφηκαν τα κρούσματα βίας των ίδιων μαθητών για το επόμενο δίμηνο. Οι ερευνητές ενδιαφέρονται να εξετάσουν εάν ο μέσος όρος των κρουσμάτων βίας αυξήθηκε μετά την προβολή της ταινίας. Η σύγκριση αφορά το μέσο όρο κρουσμάτων βίας του ίδιου πληθυσμού σε δύο διαφορετικές χρονικές στιγμές.

Παρατήρηση για το τελευταίο παράδειγμα: Ο αναγνώστης μπορεί εύκολα και σωστά να διατυπώσει αμφιβολίες σχετικά με την «ηθική» βάση της διαδικασίας. Πιο συγκεκριμένα με ποιο σκεπτικό οδηγούμε μικρά παιδιά πιθανώς σε αύξηση της βίας συμπεριφοράς τους; Η ένσταση είναι απόλυτα σωστή και για το λόγο αυτό τέτοια πειράματα δεν γίνονται (τουλάχιστον όχι πια). Πάντως κατά τη δεκαετία του 1960 και πριν τα πράγματα ήταν διαφορετικά σχετικά με τις ηθικές αναστολές των ερευνητών τουλάχιστον σε επίπεδο ενηλίκων. Για περισσότερες πληροφορίες ο αναγνώστης μπορεί να διαβάσει για το πείραμα Milgram (βάλτε τις λέξεις *Milgram experiment* στη μηχανή αναζήτησης www.google.com και θα πάρετε αρκετά σχετικά αποτελέσματα. Ένα καλό site είναι πάντως και το εξής: <http://www.cba.uri.edu/Faculty/dellabitta/mr415s98/EthicEtcLinks/Milgram.htm>).

Τονίζουμε ότι όλη η διαδικασία που θα περιγραφεί στις επόμενες ενότητες αφορά μέσους όρους που προέρχονται από συνεχή αριθμητικά δεδομένα (πχ. Βάρος, Εισόδημα στις αρχικές τους μονάδες μέτρησης). Όταν έχουμε κατηγορικές μεταβλητές όπου δεν έχει νόημα η διαβάθμιση μεταξύ των τιμών (πχ. φύλο 1=άνδρας 2=γυναίκα) ο μέσος όρος δεν έχει επίσης νόημα. Αυτά τα δεδομένα ονομάζονται δεδομένα ονομαστικής κλίμακας ή *nominal data*. Σε διακριτά δεδομένα όπου η διαβάθμιση μεταξύ των τιμών έχει νόημα (δεδομένα τακτικής κλίμακας ή *ordinal data* πχ. το εισόδημα στη μορφή 1=Φτωχός, 2=Μέσο εισόδημα 3=Πλούσιος), ο μέσος όρος χρησιμοποιείται αλλά υπάρχουν αρκετές επιφυλάξεις για την ερμηνεία του και τη χρήση του συγκεκριμένου

ελέγχου. Εκτός από το αρχικό κεφάλαιο, για μία καλή περιγραφή των βασικών τύπων δεδομένων ο αναγνώστης αν ανατρέξει στο site <http://www.cmh.edu/stats/definitions.asp> και στο σύνδεσμο Types of data.

4.2. Στατιστική μεθοδολογία

Στα προηγούμενα παραδείγματα η σύγκριση ήταν κατά σειρά

- (i) μεταξύ του μέσου όρου και μίας προκαθορισμένης τιμής
- (ii) μεταξύ δύο διαφορετικών μέσων όρων και
- (iii) μεταξύ του ίδιου μέσου όρου σε δύο διαφορετικές χρονικές στιγμές.

Η πρώτη περίπτωση αφορά την εξέταση του μέσου όρου από έναν πληθυσμό ενώ οι άλλες δύο θα εξετασθούν στην περίπτωση των δύο πληθυσμών (η τελευταία περίπτωση εντάσσεται σε αυτή την κατηγορία γιατί έχουμε δύο δείγματα έστω και αν τυπικά αυτά προέρχονται από τον ίδιο πληθυσμό). Σε όλες τις περιπτώσεις η πιο συνηθισμένη στατιστική μεθοδολογία είναι ο έλεγχος t (t-test). Στη συνέχεια θα περιγράψουμε (χωρίς να αναφερθούμε σε ιδιαίτερες μαθηματικές εξισώσεις) την παραπάνω τεχνική αλλά και διάφορους εναλλακτικούς τρόπους ελέγχου στη περίπτωση όπου το t-test δεν είναι κατάλληλο.

4.2.1. Έλεγχος για το μέσο όρο ενός πληθυσμού

Το σκεπτικό του ελέγχου

Προτού προχωρήσουμε στους τρόπους ελέγχου, θεωρούμε χρήσιμο να παραθέσουμε τον τρόπο που η Στατιστική θεωρία προσεγγίζει το πρόβλημα του ελέγχου για το μέσο όρο ενός πληθυσμού. Πολλές από τις παρακάτω ιδέες θα χρησιμοποιηθούν (με την αναγκαία προσαρμογή) και στην περίπτωση των δύο πληθυσμών.

Όταν θέλουμε να συγκρίνουμε το μέσο όρο μίας μεταβλητής του πληθυσμού με μία προκαθορισμένη τιμή και επειδή δεν έχουμε δεδομένα από όλο τον πληθυσμό (αν υπήρχαν, τότε ο έλεγχος θα ήταν περιττός-θα μπορούσαμε να υπολογίσουμε άμεσα τον πληθυσμιακό μέσο όρο) χρησιμοποιούμε τα αντίστοιχα δεδομένα του δείγματος. Πιο συγκεκριμένα αυτό το οποίο ουσιαστικά κάνουμε είναι να δούμε πόσο «κοντά» είναι ο μέσος όρος του δείγματος για τη συγκεκριμένη μεταβλητή με την τιμή που έχουμε καθορίσει. Εφόσον είναι «αρκετά κοντά» τότε θεωρούμε ότι δεν υπάρχουν διαφορές και

λέμε ότι κατ' επέκταση ο μέσος όρος του πληθυσμού για τη συγκεκριμένη μεταβλητή είναι ίσος με την τιμή που είχαμε προκαθορίσει. Σε αντίθετη περίπτωση εντοπίζουμε τυχόν διαφορές και λέμε ότι ο μέσος όρος του πληθυσμού για τη συγκεκριμένη μεταβλητή είναι διαφορετικός από την προκαθορισμένη τιμή. Ακόμα καλύτερα για την τελευταία περίπτωση μπορούμε να δούμε εάν ο μέσος όρος είναι μικρότερος/μεγαλύτερος της συγκεκριμένης τιμής.

Το κριτήριο ελέγχου

Ένα ενδιαφέρον ζήτημα είναι πως καθορίζεται το «κοντά» που αναφέρθηκε προηγουμένως. Ως μέτρο απόστασης λαμβάνεται αρχικά η διαφορά της προκαθορισμένης τιμής από το μέσο όρο του δείγματος. Όμως αυτή η διαφορά εξαρτάται από τις μονάδες του προβλήματος κάθε φορά. Για παράδειγμα διαφορά 5 μονάδων στην εικοσαβάθμια μαθητική βαθμολογική κλίμακα ή 50 μονάδων στο δείκτη ευφυΐας σημαίνει τεράστια απόκλιση μεταξύ των δύο ποσοτήτων ενώ η διαφορά των 500 ευρώ στο ετήσιο οικογενειακό εισόδημα δεν αντιστοιχεί στην ίδια *ποιοτική* διαφοροποίηση. Επίσης η ερμηνεία αλλάζει και μεταξύ ομοειδών διαφορών. Ως παράδειγμα παίρνουμε τη διαφορά των 500 ευρώ στο ετήσιο οικογενειακό εισόδημα. Αν ο υπό εξέταση πληθυσμός είναι μία κοινωνική ομάδα χαμηλών εισοδημάτων, η παραπάνω διαφοροποίηση είναι πολύ πιο έντονη από το αν εξετάζαμε έναν πληθυσμό με πιο ισόρροπη οικονομική κατανομή. Για το λόγο αυτό ως τελικό μέτρο απόστασης του ελέγχου παίρνουμε τη διαφορά της προκαθορισμένης τιμής από το μέσο όρο του δείγματος διαιρεμένης με την τυπική απόκλιση του μέσου όρου του δείγματος που αποτελεί ένα μέτρο της «ποικιλίας»/διασποράς των τιμών του δείγματος γύρω από το μέσο όρο. Έχουμε δηλαδή περιληπτικά:

$$\text{Κριτήριο ελέγχου} = (\text{Μέσος όρος δείγμ.} - \text{καθορισμένη τιμή}) / \text{Τυπική απόκλιση μέσου}$$

Το κριτήριο αυτό έχει, υπό προϋποθέσεις, κάποιες γνωστές στατιστικές ιδιότητες. Πιο συγκεκριμένα ακολουθεί μία γνωστή στατιστική κατανομή την κατανομή *t* (ή αλλιώς κατανομή Student) για αυτό και ονομάζεται *έλεγχος t* (*t-test*). Η κατανομή αυτή μας βοηθά στον υπολογισμό της τιμής *p* (*p-value*), τη σύγκριση με το *επίπεδο στατιστικής σημαντικότητας* (*statistical significance level*) και τέλος στην εξαγωγή των τελικών συμπερασμάτων όπως θα περιγράψουμε παρακάτω.

Ενδιαφέρον παρουσιάζει η προϋπόθεση που πρέπει να ισχύει για την εφαρμογή του παραπάνω κριτηρίου ελέγχου. Αυτή αναφέρεται στο γεγονός ότι τα δεδομένα πρέπει να προέρχονται από πληθυσμό που ακολουθεί κανονική κατανομή. Η κανονική κατανομή είναι η πιο γνωστή κατανομή της Στατιστικής και έχει τη μορφή συμμετρικής «καμπανοειδούς» καμπύλης. Όπως αναφέρθηκε και σε προηγούμενα κεφάλαια, για να εξακριβώσουμε την ισχύ της παραπάνω προϋπόθεσης καλό είναι να κάνουμε ένα ιστόγραμμα και να δούμε αν υπάρχουν σοβαρές αποκλίσεις από την κανονικότητα. Στις εφαρμογές αλλά και στο Παράρτημα Α θα δούμε επίσης έναν στατιστικό έλεγχο για τη διαπίστωση της κανονικότητας. Σε περίπτωση όπου διαπιστώσουμε ότι υπάρχουν σοβαρές αποκλίσεις από την κανονικότητα ο στατιστικός έλεγχος t δεν είναι πλέον κατάλληλος. Ο πιο αποδεκτός και καθιερωμένος εναλλακτικός έλεγχος στην περίπτωση αυτή είναι το *Wilcoxon signed rank* που βασίζεται στην κατάταξη των παρατηρήσεων.

Εξαγωγή συμπερασμάτων

Σε όλες τις περιπτώσεις το τελικό συμπέρασμα για τον πληθυσμό βασίζεται στα δειγματικά δεδομένα που αποτελούν ένα μικρό υποσύνολο του πληθυσμού. Υπάρχει λοιπόν η περίπτωση *λάθους συμπεράσματος*, δηλ. το τυχαίο δείγμα να αποτελείται από τέτοια δεδομένα που να δίνουν ένα λανθασμένο συμπέρασμα σχετικά με τον πληθυσμό. (Προσοχή: Αυτό δεν σημαίνει αναγκαστικά ότι το δείγμα δεν είναι αντιπροσωπευτικό. Υπάρχει πάντα (έστω και μικρή) η πιθανότητα ένα τυχαίο δείγμα που συλλέχθηκε με τη σωστή τεχνική δειγματοληψίας να δίνει αποτελέσματα που δεν συμβαδίζουν με τη γενική εικόνα του πληθυσμού. Αυτή είναι εξάλλου και η έννοια του «τυχαίου» στη δειγματοληψία. Ως παράδειγμα αναφέρουμε την περίπτωση δειγματοληψίας εισοδήματος σε ένα οικοδομικό τετράγωνο όπου με τυχαίο τρόπο επιλέγεται κάποιος/α που μόλις έχει κερδίσει ένα σημαντικό ποσό στο λαχείο!). Δυστυχώς περιπτώσεις σαν την προηγούμενη δεν μπορούν να αποκλεισθούν. Είναι δυνατόν όμως να φροντίσουμε ώστε να έχουν πολύ μικρή πιθανότητα πραγματοποίησης. Πιο αναλυτικά έχουμε τα εξής:

- Αρχικά υποθέτουμε ότι δεν υπάρχει (ουσιαστική) διαφορά μεταξύ του άγνωστου μέσου όρου του *πληθυσμού* και της προκαθορισμένης τιμής. Είναι δηλαδή αρκετά κοντά αυτές οι δύο τιμές. Αυτή ονομάζεται αρχική ή μηδενική υπόθεση (*null*

hypothesis και συμβολίζεται με H_0) και εξετάζεται σε σχέση με την εναλλακτική υπόθεση ότι υπάρχει διαφορά.

- Στη συνέχεια με τη βοήθεια του κριτηρίου ελέγχου υπολογίζουμε την πιθανότητα η οποιαδήποτε διαφορά του μέσου όρου να οφείλεται σε τυχαίους παράγοντες ή όχι.
- Αν η παραπάνω πιθανότητα είναι μικρότερη από ένα επιτρεπτό όριο τότε απορρίπτουμε την αρχική υπόθεση και θεωρούμε ότι υπάρχει στατιστικώς σημαντική διαφορά (statistically significant difference) μεταξύ του μέσου όρου του πληθυσμού και της καθορισμένης τιμής. Στην αντίθετη περίπτωση θεωρούμε ότι δεν υπάρχει στατιστικώς σημαντική διαφορά.

Η πιθανότητα που περιγράψαμε παραπάνω υπολογίζεται από το στατιστικό έλεγχο t και ονομάζεται τιμή p (*p-value*). Το επιτρεπτό όριο με το οποίο συγκρίνεται η τιμή p ονομάζεται *επίπεδο στατιστικής σημαντικότητας* (statistical significance level) και καθορίζεται από τον ερευνητή. Το επίπεδο σημαντικότητας είναι ουσιαστικά η πιθανότητα λανθασμένης απόρριψης της αρχικής υπόθεσης. Η πιο συνηθισμένη τιμή είναι 5% (ή 0,05) ενώ άλλες τιμές που χρησιμοποιούνται είναι 1% (ή 0,01) και 10% (ή 0,10) για πιο «αυστηρούς» και «χαλαρούς» ελέγχους αντιστοίχως. Είναι λοιπόν φανερό ότι σε κάθε έλεγχο είναι απαραίτητο για τον ερευνητή να έχει καθορίσει ένα επίπεδο σημαντικότητας. Τότε σε περίπτωση όπου η τιμή p είναι μικρότερη ή ίση από το επίπεδο σημαντικότητας δηλώνουμε ότι υπάρχει στατιστικά σημαντική διαφορά μεταξύ του μέσου όρου του πληθυσμού και της προκαθορισμένης τιμής στο συγκεκριμένο επίπεδο σημαντικότητας.

Σχηματικά έχουμε τα ακόλουθα βήματα για την εξαγωγή συμπερασμάτων μέσω ενός ελέγχου t :

1. Έλεγχος της κανονικότητας των δεδομένων (απαραίτητη προϋπόθεση για κάθε έλεγχο t)
2. Καθορισμός της αρχικής υπόθεσης ότι δεν υπάρχει διαφορά μεταξύ πληθυσμιακού μέσου όρου και προκαθορισμένης τιμής. Αντιστοίχως καθορίζεται και η εναλλακτική υπόθεση.
3. Ορισμός του επιπέδου σημαντικότητας
4. Πραγματοποίηση του ελέγχου t
5. Εύρεση της τιμής p

6. Σύγκριση της τιμής p με το επίπεδο σημαντικότητας

6A. Αν η τιμή $p \leq$ επιπ. σημαντικότητας τότε απορρίπτουμε την αρχική υπόθεση δηλαδή υπάρχει στατιστικά σημαντική διαφορά μεταξύ του μέσου όρου του πληθυσμού και της προκαθορισμένης τιμής στο συγκεκριμένο επίπεδο σημαντικότητας, αλλιώς

6B. Αν τιμή $p >$ επιπ. σημαντικότητας τότε δεν απορρίπτουμε την αρχική υπόθεση, δηλαδή πράγματι δεν υπάρχει στατιστικά σημαντική διαφορά μεταξύ του μέσου όρου του πληθυσμού και της προκαθορισμένης τιμής στο συγκεκριμένο επίπεδο σημαντικότητας.

4.3. Σχόλια για τον έλεγχο του μέσου όρου ενός πληθυσμού

Είναι φανερό ότι (σχεδόν) όλα τα παραπάνω μπορεί να υλοποιηθούν με τη χρήση H/Y για μεγαλύτερη ταχύτητα και ευκολία. Το πιο διαδεδομένο στατιστικό λογισμικό είναι το Statistical Package for the Social Sciences (SPSS), παραδείγματα μέσω του οποίου θα εμφανισθούν παρακάτω. Το SPSS μας δίνει την τιμή p για τον παραπάνω έλεγχο και ο ερευνητής, αφού έχει θέσει το επίπεδο σημαντικότητας που είναι κατάλληλο για την έρευνα, καταλήγει στο σχετικό συμπέρασμα.

Αξίζει επίσης να αναφερθεί ότι η παραπάνω ανάλυση εξέτασε την περίπτωση ύπαρξης διαφοράς μεταξύ του μέσου όρου ενός πληθυσμού και μίας προκαθορισμένης τιμής. Δεν έχουμε ασχοληθεί με την πιο πρακτική περίπτωση όπου μας ενδιαφέρει αν ο μέσος όρος είναι μόνο μεγαλύτερος ή μόνο μικρότερος από την τιμή και όχι απλά διαφορετικός. Η αντιμετώπιση είναι ακριβώς η ίδια όπως και προηγουμένως με την ιδιαιτερότητα ότι το στατιστικό λογισμικό που θα χρησιμοποιήσουμε δεν μας δίνει απευθείας την τιμή p αλλά θα πρέπει να την υπολογίσουμε με βάση τις τιμές που μας δίνει το SPSS.

Ο έλεγχος όπου εξετάζεται απλά η διαφορά μεταξύ του μέσου όρου και της καθορισμένης τιμής λέγεται «*δίπλευρος*» έλεγχος (two-sided test) ενώ οι άλλοι λέγονται «*μονόπλευροι*» έλεγχοι (one-sided test).

Σε περιπτώσεις όπου το δείγμα είναι πολύ μεγάλο τότε κατά πάσα πιθανότητα ο έλεγχος θα καταδείξει στατιστικά σημαντικές διαφορές έστω και αν οι αριθμητικές διαφορές είναι μικρές. Αυτό είναι ένα χαρακτηριστικό πολλών τέτοιων ελέγχων.

Παρόμοιος προβληματισμός υπάρχει και στην ανάλυση συσχέτισης αλλά και στο κριτήριο X^2 . Χωρίς να μπορούμε σε περισσότερες λεπτομέρειες (αφού αυτό άλλωστε είναι ένα ιδιαίτερος «πλατύ» θέμα) καλό είναι να αποφεύγονται τα πολύ μεγάλα ή τα πολύ μικρά δείγματα όταν έχουμε στατιστικούς ελέγχους.

Τέλος η παραπάνω συζήτηση συμβολίζει σε κάποιο βαθμό τη διαφορά μεταξύ Στατιστικής και Μαθηματικών. Στη Στατιστική αρκεί να είναι δύο ποσότητες «κοντά» ώστε να πούμε ότι δεν υπάρχει διαφορά. Αντιθέτως στα Μαθηματικά πρέπει να είναι αριθμητικά ίσες μέχρι το τελευταίο δεκαδικό ψηφίο ώστε να πούμε ότι δεν υπάρχει διαφορά.

4.4. Εφαρμογές του ελέγχου για το μέσο όρο ενός πληθυσμού

Οι παρακάτω εφαρμογές θα αναλύσουν τη μεθοδολογία που περιγράψαμε προηγουμένως. Στο πρώτο παράδειγμα η ανάλυση θα είναι λεπτομερής ενώ στη συνέχεια κάποια από τα βήματα θα εξετάζονται μαζί όπως άλλωστε συμβαίνει στην πράξη.

A. Θεωρούμε ότι οι μαθητές ηλικίας 12 ετών ασχολούνται 8 ώρες κατά μέσο όρο εβδομαδιαίως με εξωσχολικές δραστηριότητες (ξένες γλώσσες, αθλητισμός, κλπ). Τα παρακάτω στοιχεία αφορούν τις ώρες που αφιερώνουν την εβδομάδα δέκα τυχαίως επιλεγμένοι μαθητές αυτής της ηλικίας για εξωσχολικές δραστηριότητες. Μας ενδιαφέρει να εξετάσουμε εάν ο αρχικός συλλογισμός είναι σωστός.

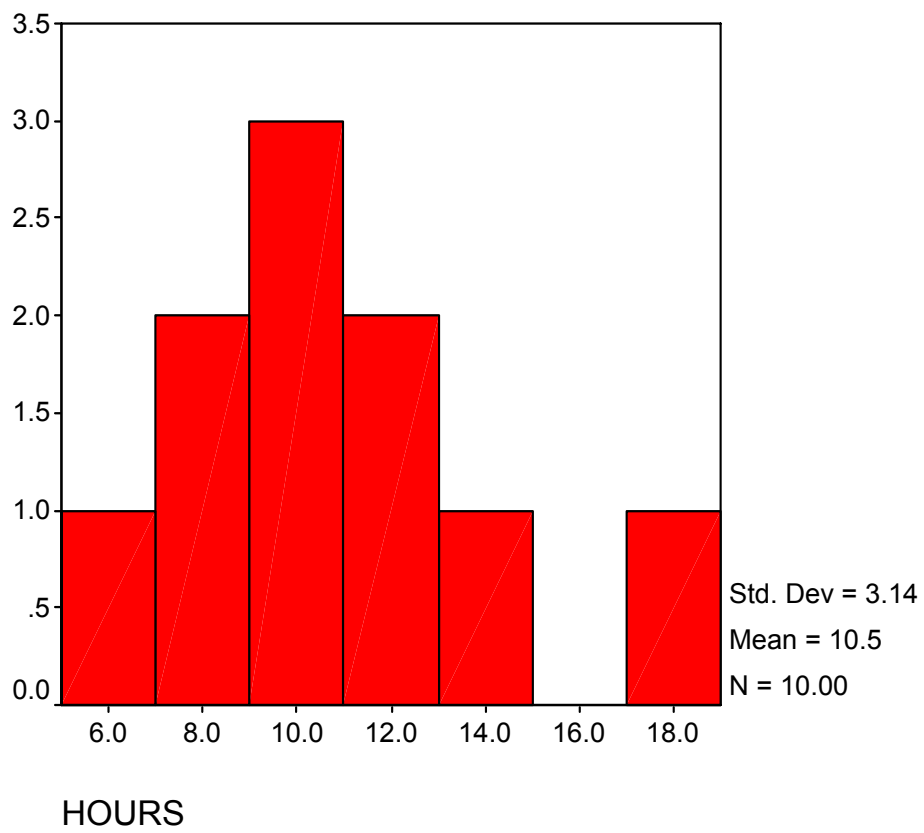
Ώρες για εξωσχολικές δραστηριότητες την εβδομάδα 10 μαθητών									
12	17	9,5	7,5	6,5	9,5	11	14	8,5	9,5

Θα ακολουθήσουμε τα βήματα που αναφέρθηκαν προηγουμένως για την εξαγωγή συμπερασμάτων.

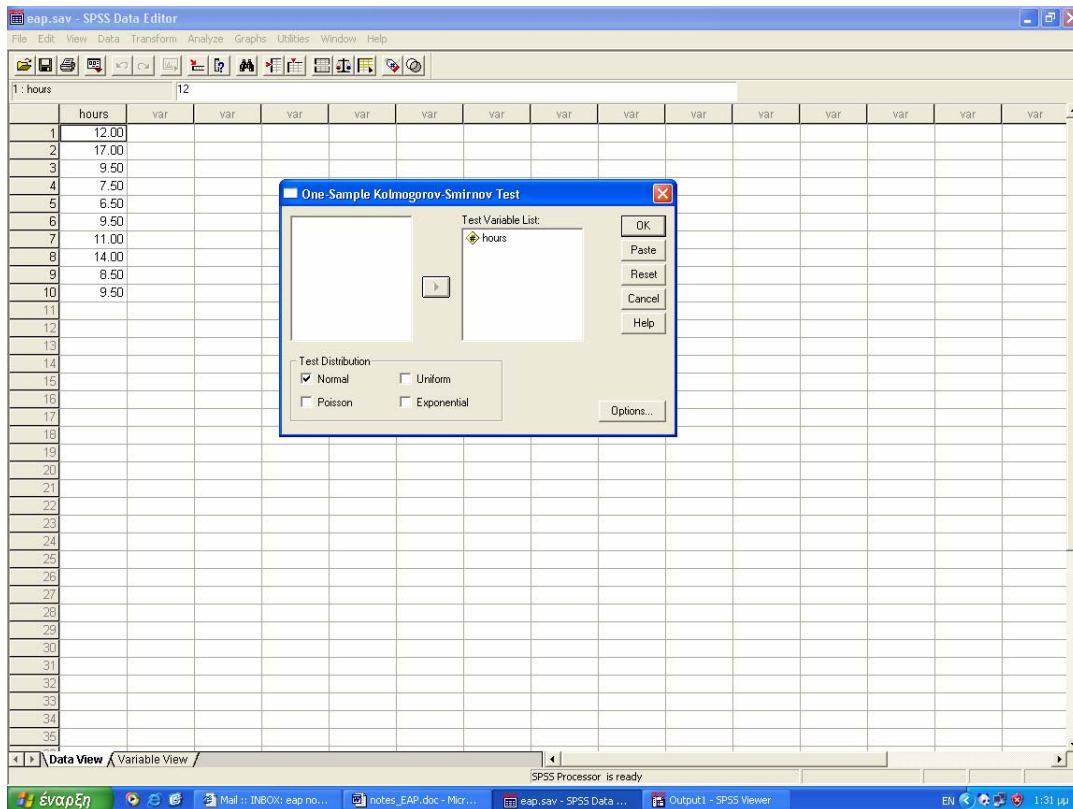
Βήμα 1: Έλεγχος κανονικότητας

Αφού εισάγουμε τα δεδομένα στο SPSS (στη στήλη hours) εξετάζουμε αν προέρχονται από κανονική κατανομή.

Για το λόγο αυτό κάνουμε πρώτα ένα Ιστόγραμμα (εντολές **Graphs** και μετά **Histogram** βάζοντας το όνομα της μεταβλητής εκεί που λέει **Variables** και πατάμε OK). Επίσης μπορούμε να κάνουμε έναν έλεγχο κανονικότητας χρησιμοποιώντας τις εντολές **Analyze-Nonparametric tests-1 Sample K-S**. Στο μενού που θα εμφανισθεί εισάγουμε το όνομα της μεταβλητής στο μέρος που λέει **Test Variable List** φροντίζοντας να υπάρχει κλικ στην επιλογή **Test Distribution Normal**. Τα αποτελέσματα εμφανίζονται αφού πατήσουμε OK (βλέπε Σχήματα 21 και 22).



Σχήμα 21: Ιστόγραμμα για τη μεταβλητή HOURS



Σχήμα 22: Διαδικασία εντολών για την κανονικότητα

Από το ιστόγραμμα παρατηρούμε ότι δεν υπάρχουν σοβαρές αποκλίσεις από την κανονικότητα. Αυτό επιβεβαιώνεται από τα αποτελέσματα του ελέγχου κανονικότητας που παρουσιάζονται παρακάτω.

NPar Tests

One-Sample Kolmogorov-Smirnov Test

		HOURS
N		10
Normal Parameters	Mean	10.5000
	Std. Deviation	3.1447
Most Extreme Differences	Absolute	.225
	Positive	.225
	Negative	-.102
Kolmogorov-Smirnov Z		.711
Asymp. Sig. (2-tailed)		.693
a Test distribution is Normal.		

b Calculated from data.

Επικεντρώνουμε την προσοχή μας στο μέρος των αποτελεσμάτων που λέει: **Asymp. Sig. (2-tailed)**. Αυτό δίνει την τιμή p του ελέγχου για την κανονικότητα. Η αρχική ή μηδενική υπόθεση αναφέρει ότι τα δεδομένα προέρχονται από κανονική κατανομή. Αν η τιμή p είναι μικρότερη ενός επιπέδου σημαντικότητας, έστω 5%, τότε απορρίπτουμε την υπόθεση αυτή. Στην περίπτωσή μας έχουμε ότι $0,693 > 0,05$ και άρα εξακολουθεί να ισχύει η υπόθεση της κανονικότητας και μπορούμε να προχωρήσουμε στον έλεγχο t (t-test).

Βήμα 2: Καθορισμός της αρχικής ή μηδενικής υπόθεσης (H_0) ότι δεν υπάρχει διαφορά μεταξύ πληθυσμιακού μέσου όρου και προκαθορισμένης τιμής. Αντιστοίχως καθορίζεται και η εναλλακτική υπόθεση (H_1)

H_0 : Οι μαθητές απασχολούνται κατά μέσο όρο 8 ώρες την εβδομάδα σε εξωσχολικές δραστηριότητες.

H_1 : Οι μαθητές δεν απασχολούνται κατά μέσο όρο 8 ώρες την εβδομάδα σε εξωσχολικές δραστηριότητες.

Βήμα 3: Ορισμός του επιπέδου σημαντικότητας

Ως επίπεδο σημαντικότητας ορίζεται το 5% (ή 0,05).

Βήμα 4: Πραγματοποίηση του ελέγχου t

Για το σκοπό αυτό επιλέγουμε κατά σειρά τις εξής εντολές (Σχήμα 23) :

Analyze-Compare Means–One Sample T Test

Στη συνέχεια εμφανίζεται η οθόνη με το την εντολή ελέγχου και εισάγουμε στο μέρος που λέει **Test Variable(s)** τη μεταβλητή της οποίας θα ελέγξουμε το μέσο όρο (hours στο SPSS). Η προκαθορισμένη τιμή (στην προκειμένη περίπτωση 8) εισάγεται εκεί που λέει **Test Value** (Σχήμα 24) και πατάμε OK. Τα αποτελέσματα της ανάλυσης εμφανίζονται παρακάτω:

T-Test

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
HOURS	10	10.5000	3.1447	.9944

One-Sample Test

Test Value = 8

t	df	Sig. (2-tailed)	(2-Mean Difference	95% Confidence Interval of the Difference	
				Lower	Upper
2.514	9	.033	2.5000	.2504	4.7496

Βήμα 5: Εύρεση της τιμής p

Μετά τον τίτλο εμφανίζονται κάποια περιγραφικά στοιχεία για τη μεταβλητή Hours όπου βλέπουμε ότι ο δειγματικός μέσος όρος είναι 10.50 αρκετά διαφορετικός από το 8. Η στήλη Std. Error Mean δίνει το μέτρο της διασποράς των δεδομένων που χρησιμοποιείται στον παρονομαστή του κριτηρίου ελέγχου. Πράγματι μπορούμε να διαπιστώσουμε ότι : $t = (10,5-8)/0,944=2,514$ που εμφανίζεται στο δεύτερο πίνακα. Η στήλη Sig. (2-tailed) δίνει την τιμή p για ένα δίπλευρο έλεγχο που είναι 0,033.

Βήμα 6: Σύγκριση της τιμής p με το επίπεδο σημαντικότητας

Παρατηρούμε ότι $0,033 < 0,05$. Άρα υπάρχει στατιστικά σημαντική διαφορά σε σχέση με το 8. Το συμπέρασμα είναι ότι δεν ισχύει η αρχική υπόθεση, δηλαδή οι μαθητές ηλικίας 12 ετών δεν αφιερώνουν 8 ώρες εβδομαδιαίως για εξωσχολικές δραστηριότητες.

Αξίζει να προσέξουμε ότι εάν ήμασταν πιο αυστηροί στον καθορισμό του επιπέδου σημαντικότητας και επιλέγαμε το 1% τότε το συμπέρασμα θα ήταν το ακριβώς αντίθετο. Αυτό δεν είναι παράδοξο αφού το στατιστικό συμπέρασμα για τον πληθυσμό

με βάση το δείγμα εμπεριέχει πάντα την πιθανότητα λάθους. Αν θέλουμε να είμαστε ιδιαίτερος αυστηροί για την ύπαρξη λάθους θα πρέπει τα δειγματικά δεδομένα να είναι πολύ διαφορετικά από την προκαθορισμένη τιμή κάτι το οποίο δεν συμβαίνει σε τόσο μεγάλο βαθμό εδώ.

B. Σε συνέχεια του παραδείγματος A, ας υποθέσουμε ότι μας ενδιαφέρει να εξετάσουμε την εναλλακτική υπόθεση εάν οι μαθητές απασχολούνται υπερβολικά με εξωσχολικές δραστηριότητες δηλαδή μας ενδιαφέρει μόνο εάν ο μέσος όρος είναι μεγαλύτερος του 8. Σε αυτή την περίπτωση οι υποθέσεις παρουσιάζονται ως εξής:

H_0 : Οι μαθητές απασχολούνται κατά μέσο όρο 8 ώρες την εβδομάδα σε εξωσχολικές δραστηριότητες. (δεν αλλάζει)

H_1 : Οι μαθητές απασχολούνται κατά μέσο όρο *περισσότερο* από 8 ώρες την εβδομάδα σε εξωσχολικές δραστηριότητες.

Για να απαντήσουμε στο ερώτημα αυτό επαναλαμβάνουμε τη διαδικασία του SPSS που περιγράψαμε προηγουμένως και στη συνέχεια αφού επιλέξουμε και πάλι το 5% ως επίπεδο σημαντικότητας θα πρέπει να βρούμε την τιμή p. Αυτό γίνεται ως εξής:

1. Βρίσκουμε την τιμή p του δίπλευρου ελέγχου (0,033)

2. Βρίσκουμε την τιμή του κριτηρίου ελέγχου ($t=2,514$).

2A. Αν η τιμή του κριτηρίου είναι θετική και μας ενδιαφέρει ο μονόπλευρος έλεγχος για τιμές μεγαλύτερες της προκαθορισμένης τιμής τότε η τιμή p του μονόπλευρου ελέγχου είναι το μισό της τιμής p του δίπλευρου ελέγχου.

2B. Αν η τιμή του κριτηρίου είναι θετική και μας ενδιαφέρει ο μονόπλευρος έλεγχος για τιμές μικρότερες της προκαθορισμένης τιμής τότε η τιμή p του μονόπλευρου ελέγχου είναι:

1-το μισό της τιμής p του δίπλευρου ελέγχου.

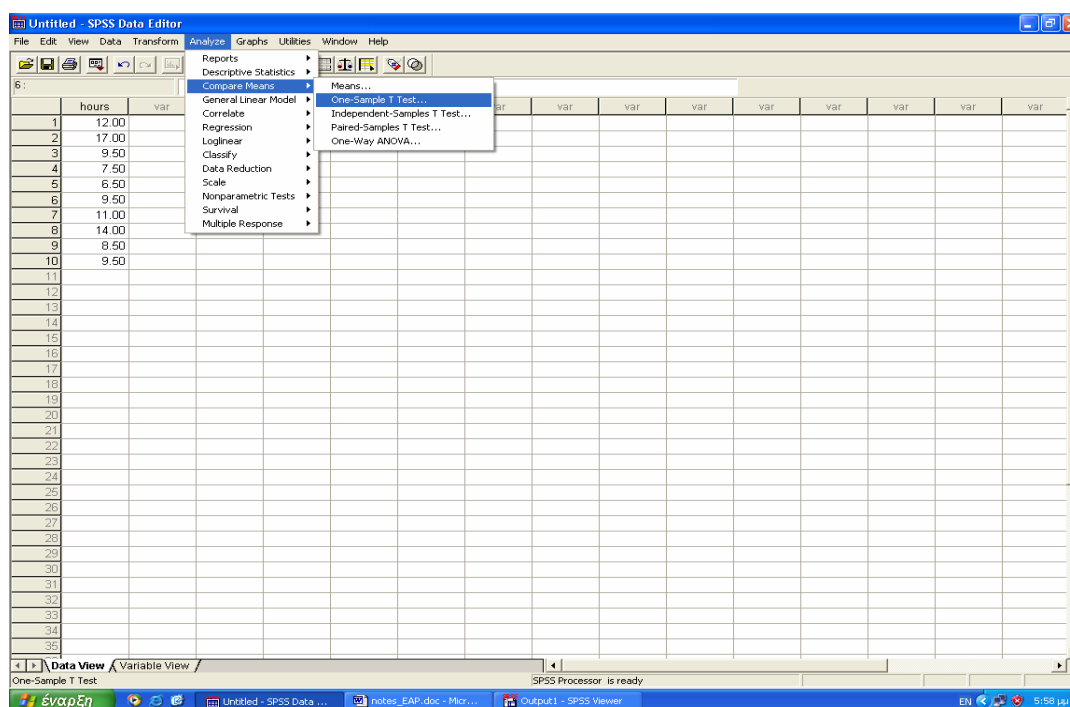
2Γ. Αν η τιμή του κριτηρίου είναι αρνητική και μας ενδιαφέρει ο μονόπλευρος έλεγχος για τιμές μικρότερες της προκαθορισμένης τιμής τότε η τιμή p του μονόπλευρου ελέγχου είναι το μισό της τιμής p του δίπλευρου ελέγχου.

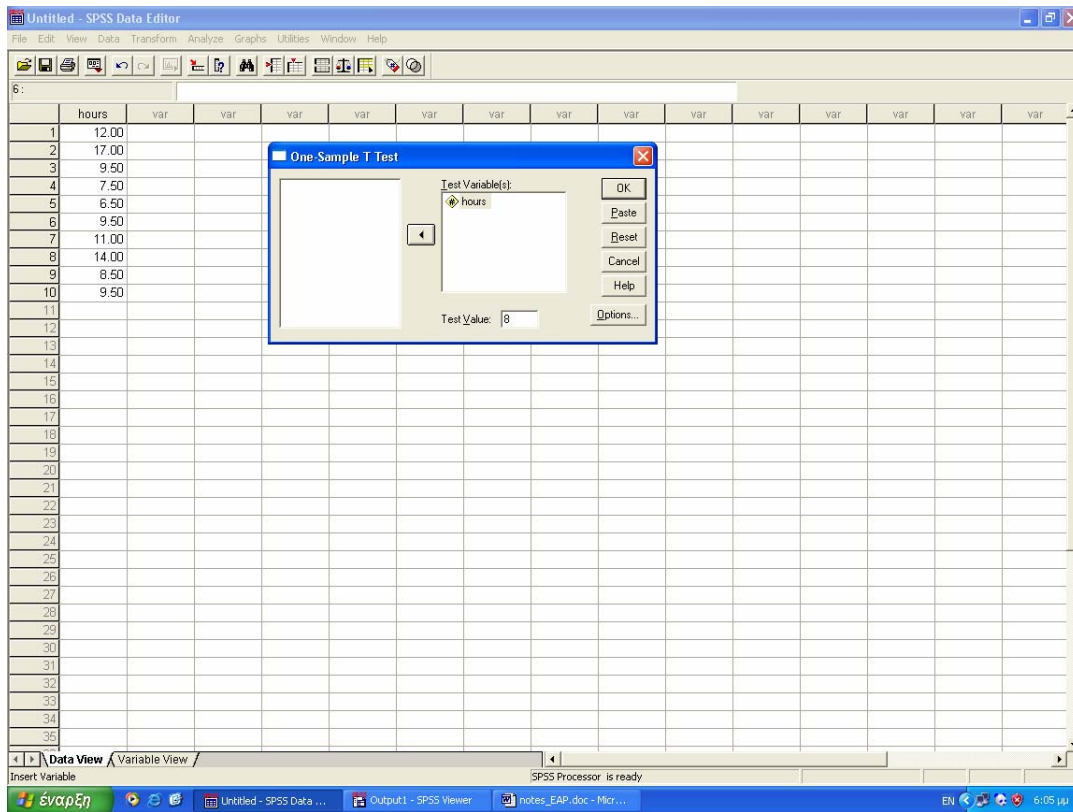
2Δ. Αν η τιμή του κριτηρίου είναι αρνητική και μας ενδιαφέρει ο μονόπλευρος έλεγχος για τιμές μεγαλύτερες της προκαθορισμένης τιμής τότε η τιμή p του μονόπλευρου ελέγχου είναι:

1-το μισό της τιμής p του δίπλευρου ελέγχου.

Στο παράδειγμά μας ισχύει η περίπτωση 2Α και η τιμή p του μονόπλευρου ελέγχου είναι 0.016. Άρα το συμπέρασμα είναι ότι μέσος όρος εξωσχολικών δραστηριοτήτων είναι μεγαλύτερος του 8 σε επίπεδο στατιστικής σημαντικότητας 5%. Αξίζει να τονίσουμε ότι όπως κι στην περίπτωση του δίπλευρου ελέγχου, το συμπέρασμα αντιστρέφεται σε επίπεδο σημαντικότητας 1%.

Τέλος αν στο παράδειγμά μας ενδιέφερε μόνο εάν ο μέσος όρος είναι μικρότερος του 8, η τιμή p δίνεται από την περίπτωση 2Β και είναι $1-0,016=0,984>0,05$. Το συμπέρασμα είναι ότι κατά μέσο όρο οι εξωσχολικές δραστηριότητες δεν είναι λιγότερες των 8 ωρών την εβδομάδα σε επίπεδο στατιστικής σημαντικότητας 5% (ή και οποιοδήποτε άλλο επίπεδο που χρησιμοποιείται στην πράξη).





Σχήματα 23-24: Οθόνη του SPSS με τις εντολές για τον έλεγχο της μέσης τιμής ενός πληθυσμού

4.5. Έλεγχος για τη σύγκριση μέσων όρων δύο πληθυσμών

Θα επεκτείνουμε την παραπάνω διαδικασία για την περίπτωση όπου έχουμε δείγματα από δύο πληθυσμούς. Όπως είδαμε και στην αρχική συζήτηση της ενότητας «Στατιστική μεθοδολογία» στην κατηγορία των δύο πληθυσμών βάζουμε περιπτώσεις όπως του Παραδείγματος 2 (όπου πράγματι έχουμε δύο διαφορετικούς πληθυσμούς) αλλά και του Παραδείγματος 3 (όπου τα δύο δείγματα προέρχονται από τον ίδιο πληθυσμό). Στην πρώτη περίπτωση τα δείγματα είναι ανεξάρτητα ενώ στη δεύτερη τα δείγματα είναι εξαρτημένα αφού οι παρατηρήσεις τους προέρχονται από τα ίδια άτομα. Θα ακολουθήσουμε αυτή τη διαφοροποίηση και εδώ.

4.5.1. Ανεξάρτητα δείγματα

Ο έλεγχος t που περιγράφηκε προηγουμένως μπορεί άνετα να επεκταθεί και στους δύο πληθυσμούς. Το σκεπτικό ανάπτυξης του κριτηρίου ελέγχου παραμένει το ίδιο με την αλλαγή ότι στον αριθμητή έχουμε τη διαφορά μεταξύ των δειγματικών μέσων όρων και στον παρονομαστή έχουμε μία συνολική τυπική απόκλιση μεταξύ των μέσων όρων των δύο δειγμάτων. Πιο συγκεκριμένα έχουμε:

$$\text{Κριτήριο ελέγχου} = (\text{Διαφορά μέσων όρων}) / \text{Συνολική τυπική απόκλιση}$$

Ισχύει η ίδια προϋπόθεση όπως και πριν για την εφαρμογή του παραπάνω κριτηρίου ελέγχου. Αυτή αναφέρεται στο γεγονός ότι τα δεδομένα από τις δύο ομάδες πρέπει να προέρχονται από πληθυσμούς με κανονικές κατανομές. Για να εξακριβώσουμε την ισχύ της παραπάνω προϋπόθεσης καλό είναι να κάνουμε δύο ιστογράμματα για τα αντίστοιχα δείγματα και να δούμε αν υπάρχουν σοβαρές αποκλίσεις από την κανονικότητα ή να χρησιμοποιήσουμε τον έλεγχο που είδαμε προηγουμένως. Σε περίπτωση όπου διαπιστώσουμε ότι υπάρχουν σοβαρές αποκλίσεις από την κανονικότητα ο στατιστικός έλεγχος t δεν είναι πλέον κατάλληλος και για το λόγο αυτό χρησιμοποιούμε το κριτήριο Mann-Whitney το οποίο δεν βασίζεται στη υπόθεση της κανονικότητας. Το κριτήριο αυτό ταξινομεί τα δεδομένα και χρησιμοποιεί αυτή την κατάταξη για την εξαγωγή του τελικού συμπεράσματος. Για να αποφύγουμε την

περαιτέρω μαθηματική εξειδίκευση δεν θα παρουσιάσουμε λεπτομερώς το κριτήριο αυτό αλλά θα δούμε πως εφαρμόζεται στο SPSS.

Εφαρμογές για ανεξάρτητα δείγματα

Τα παρακάτω στοιχεία αφορούν τη βαθμολογία (κλίμακα 0-100) σε μία μαθηματική εξέταση μίας σχολικής τάξης Γ΄ Γυμνασίου σε σχέση με το φύλο. (1-Αγόρι, 2-Κορίτσι). Ερευνήστε αν τα αγόρια έχουν υψηλότερο μέσο όρο βαθμολογίας στην εξέταση από τα κορίτσια.

Βαθμολογία	Φύλο	Βαθμολογία	Φύλο
83	1	73	2
75	1	81	2
67	1	49	2
44	1	36	2
28	1	65	2
56	1	56	2
91	1	73	2
39	1	51	2
71	1	44	2
54	1	65	2
38	1	51	2
57	2		

Αρχικά εισάγουμε τα δεδομένα σε δύο στήλες (Grades –Sex) όπως είναι και στον Πίνακα. Μπορούμε να διαπιστώσουμε ότι τα δεδομένα για τον κάθε πληθυσμό ακολουθούν την κανονική κατανομή. Αυτό γίνεται με την ίδια μεθοδολογία όπως και πριν με τη διαφορά ότι πρέπει να γίνει ξεχωριστά για Αγόρια-Κορίτσια. Για να το κάνετε αυτό στο SPSS θα χρησιμοποιήσετε την ακολουθία εντολών

Data-Select cases

Στη συνέχεια επιλέγουμε **if condition is satisfied** τοποθετώντας μέσα στο διαλογικό κουτί που εμφανίζεται **SEX=1**. Με τη διαδικασία αυτή επιλέγουμε αρχικά μόνο τα αγόρια και εφαρμόζουμε τους ελέγχους κανονικότητας που έχουν αναπτυχθεί

προηγουμένως. Στη συνέχεια επαναλαμβάνουμε τη διαδικασία μόνο για τα κορίτσια (**SEX=2**). Αφού έχουμε τελειώσει με τους ελέγχους κανονικότητας επαναλαμβάνουμε για μία τελευταία φορά τη διαδικασία επιλέγοντας **All cases** ώστε να προχωρήσουμε στον έλεγχο των μέσων όρων και για τα δύο δείγματα.

Οι υποθέσεις διαμορφώνονται ως εξής:

H_0 : ΔΕΝ υπάρχει διαφορά στο μέσο όρο βαθμολογίας μεταξύ των 2 φύλων

H_1 : Τα αγόρια έχουν υψηλότερο μέσο όρο βαθμολογίας από τα κορίτσια

Ο έλεγχος t γίνεται με την παρακάτω ακολουθία εντολών:

Analyze-Compare means-Independent Samples T Test

Στη συνέχεια εισάγουμε το όνομα της μεταβλητής Grades στο χώρο **Test Variable(s)** και το όνομα της μεταβλητής Sex στο χώρο **Grouping Variable**. Κάνουμε κλικ στην επιλογή **Define Groups** βάζοντας τις τιμές που αντιστοιχούν στα αγόρια και στα κορίτσια (1 και 2 στα κουτιά **Group 1** και **Group 2** αντιστοίχως). Επιλέγουμε Continue και OK για να πάρουμε τα παρακάτω αποτελέσματα (στον πίνακα Independent Samples test) έχουν διαγραφεί για λόγους χώρου οι 3 τελευταίες στήλες οι οποίες δεν είχαν κάποιο στοιχείο προς ανάλυση):

T-Test

Group Statistics

	SEX	N	Mean	Std. Deviation	Std. Error Mean
GRADES	1.00	11	58.7273	20.3081	6.1231
	2.00	12	58.4167	13.2765	3.8326

Independent Samples Test

		Levene's Test for Equality of Variances	F	Sig.	t-test for Equality of Means	Sig. (2-tailed)	Mean Difference
					T		
GRADES	Equal variances	3.068	.094	.044	21	.965	0.3106

assumed				
Equal	.043	16.999	.966	0.3106
variances				
not				
assumed				

Τα αποτελέσματα Group Statistics δείχνουν μία πολύ μικρή διαφορά μεταξύ των δύο μέσων όρων (58,73 για τα αγόρια έναντι 58,42 για τα κορίτσια). Αυτό καταρχήν αποτελεί μία ένδειξη ότι δεν πρέπει να υπάρχει στατιστική διαφορά. Επιλέγουμε το 5% ως επίπεδο σημαντικότητας. Κοιτώντας τα αποτελέσματα του ελέγχου t στη στήλη Sig (2-tailed) παρατηρούμε την ύπαρξη 2 τιμών p. Όπως και στην περίπτωση του ενός πληθυσμού αυτές αντιστοιχούν στην τιμή p του δίπλευρου ελέγχου. Ποια όμως από τις δύο θα επιλέξουμε (αν και είναι σχεδόν ίδιες) ;

Η απάντηση θα δοθεί από τον έλεγχο των διασπορών που γίνεται στις αρχικές στήλες. Πιο συγκεκριμένα ο έλεγχος του Levene που εμφανίζεται στην τρίτη στήλη θεωρεί αρχικά ότι οι διασπορές μεταξύ των δύο πληθυσμών είναι ίδιες. Η τιμή p που δίνεται από τη στήλη Sig συγκρίνεται με το επίπεδο σημαντικότητας που έχουμε επιλέξει. Εάν η τιμή p (0,094) είναι μικρότερη του 5% τότε ο έλεγχος θα συμπεράνει ότι οι διασπορές δεν είναι ίδιες και θα επιλεγόταν η δεύτερη τιμή p του δίπλευρου ελέγχου (0,966). Κάτι τέτοιο όμως δε συμβαίνει αφού $0,094 > 0,05$ και άρα οι διασπορές είναι στατιστικά ίδιες και επιλέγουμε την τιμή 0.965 ως τιμή του δίπλευρου ελέγχου.

Όμως ο έλεγχος μεταξύ των μέσων όρων είναι μονόπλευρος αφού μας ενδιαφέρει αν τα αγόρια έχουν μεγαλύτερο μέσο όρο από τα κορίτσια. Για να βρούμε την τιμή p του μονόπλευρου ελέγχου ακολουθούμε το παρακάτω σκεπτικό (παρόμοιο όπως και στην περίπτωση ενός πληθυσμού):

1. Βρίσκουμε την τιμή p του δίπλευρου ελέγχου (0,965)
 2. Βρίσκουμε την τιμή του κριτηρίου ελέγχου ($t=0,044$ που αντιστοιχεί στην παραπάνω τιμή p).
- 2Α. Αν η τιμή του κριτηρίου είναι θετική και μας ενδιαφέρει ο μονόπλευρος έλεγχος όπου ο πρώτος πληθυσμός (Group 1=Αγόρια) έχει μέσο όρο μεγαλύτερο από το δεύτερο

πληθυσμό (Group 2=Κορίτσια) τότε η τιμή p του μονόπλευρου ελέγχου είναι το μισό της τιμής p του δίπλευρου ελέγχου.

2B. Αν η τιμή του κριτηρίου είναι θετική και μας ενδιαφέρει ο μονόπλευρος έλεγχος όπου ο πρώτος πληθυσμός (Group 1=Αγόρια) έχει μέσο όρο μικρότερο από το δεύτερο πληθυσμό (Group 2=Κορίτσια) τότε η τιμή p του μονόπλευρου ελέγχου είναι:

1-το μισό της τιμής p του δίπλευρου ελέγχου.

2Γ. Αν η τιμή του κριτηρίου είναι αρνητική και μας ενδιαφέρει ο μονόπλευρος έλεγχος όπου ο πρώτος πληθυσμός (Group 1=Αγόρια) έχει μέσο όρο μικρότερο από το δεύτερο πληθυσμό (Group 2=Κορίτσια) τότε η τιμή p του μονόπλευρου ελέγχου είναι το μισό της τιμής p του δίπλευρου ελέγχου.

2Δ. Αν η τιμή του κριτηρίου είναι αρνητική και μας ενδιαφέρει ο μονόπλευρος έλεγχος όπου ο πρώτος πληθυσμός (Group 1=Αγόρια) έχει μέσο όρο μεγαλύτερο από το δεύτερο πληθυσμό (Group 2=Κορίτσια) τότε η τιμή p του μονόπλευρου ελέγχου είναι:

1-το μισό της τιμής p του δίπλευρου ελέγχου.

Στις παραπάνω περιπτώσεις με τον όρο «ο πρώτος πληθυσμός» εννοούμε αυτόν ο οποίος ορίστηκε πρώτα (δηλαδή Αγόρι στο παράδειγμα) και όχι κάποιου άλλου είδους διαβάθμιση. Στο παράδειγμα ισχύει η περίπτωση 2A και η τιμή p του μονόπλευρου ελέγχου είναι 0.47. Άρα το συμπέρασμα είναι ότι μέσος όρος βαθμολογίας δεν είναι μεγαλύτερος στα αγόρια απ' ότι στα κορίτσια 8 σε επίπεδο στατιστικής σημαντικότητας 5%.

4.5.2. Εξαρτημένα δείγματα: Κριτήριο ελέγχου και εφαρμογή

Ο έλεγχος αυτός εξετάζει για κάθε στοιχείο του δείγματος τις διαφορές μεταξύ των μέσων όρων δύο διαφορετικών χρονικών στιγμών. Αν αυτή η διαφορά (καταλλήλως διαιρεμένη) είναι αρκετά κοντά στο 0 τότε λέμε ότι δεν υπάρχει στατιστική διαφορά μεταξύ των μέσων όρων για τις δύο διαφορετικές χρονικές στιγμές. Έχουμε λοιπόν το εξής κριτήριο ελέγχου:

Κριτήριο ελέγχου = Μέσος όρος διαφορών / Τυπική απόκλιση του μέσου

Ο έλεγχος αυτός ονομάζεται *έλεγχος t για ζευγαρωτά δείγμα (paired samples t-test)*. Για να γίνει αυτός ο έλεγχος πρέπει να ισχύει η προϋπόθεση της κανονικότητας για τις τιμές των διαφορών μεταξύ των δύο διαφορετικών χρονικών στιγμών. Η διαπίστωση γίνεται με τις μεθόδους που αναφέρθηκαν και προηγουμένως. Σε περίπτωση που υπάρχει σοβαρή απόκλιση από την κανονικότητα τότε χρησιμοποιείται το κριτήριο *Wilcoxon signed rank*.

Εφαρμογή: Μας ενδιαφέρει να εξετάσουμε εάν μία συγκεκριμένη ψυχολογική θεραπεία συμβάλλει στην αύξηση της αυτο-εκτίμησης παιδιών που είχαν διαγνωσθεί ότι βρίσκονται στα αρχικά στάδια κατάθλιψης. Στα παιδιά αυτά δόθηκαν κατάλληλα ερωτηματολόγια πριν και μετά τη θεραπεία. Για κάθε φάση δημιουργήθηκε ο σχετικός δείκτης αυτο-εκτίμησης (κλίμακα 0-100). Τα στοιχεία παρουσιάζονται παρακάτω:

Αυτο-εκτίμηση πριν τη θεραπεία	Αυτο-εκτίμηση μετά τη θεραπεία	Διαφορά (Πριν-Μετά)
12	23	-11
34	45	-11.00
67	73	-6
43	54	-11
81	76	5
54	56	-2
56	66	-10
76	78	-2
65	79	-14
56	63	-7

Αφού περάσετε τα στοιχεία των 2 πρώτων στηλών (before και after) για να δημιουργήσετε στο SPSS τη στήλη της διαφοράς (Diff=before-after) ακολουθείτε τις

εντολές: **Transform-Compute**. Στη συνέχεια στο κουτί **Target variable** βάζετε το όνομα της μεταβλητής της διαφοράς και στο κουτί **Numeric expression** δημιουργείτε τη διαφορά. Είναι εύκολο να δούμε ότι τα στοιχεία της διαφοράς ακολουθούν την κανονική κατανομή.

Ο έλεγχος ανάγεται στην αρχική περίπτωση μέσου όρου ενός πληθυσμού για τη μεταβλητή diff σε σχέση με την προκαθορισμένη τιμή 0. Πιο συγκεκριμένα οι υποθέσεις είναι οι εξής:

H_0 : ΔΕΝ υπάρχει διαφορά κατά μέσο όρο στην αυτο-εκτίμηση των παιδιών πριν και μετά τη θεραπεία

H_1 : Μετά η θεραπεία, η αυτο-εκτίμηση των παιδιών έχει αυξηθεί κατά μέσο όρο

T-Test

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	BEFORE	54.4000	10	20.5329	6.4931
	AFTER	61.3000	10	17.5882	5.5619

Paired Samples Correlations

N	Correlation	Sig.
10	.966	.000

Paired Samples Test

	Paired Differences	Std.	Std. Error Mean	t	df	Sig. (2-tailed)
Pair 1	BEFORE -6.9000	5.7822	1.8285	-3.774	9	.004
	AFTER					

Εναλλακτικά, κάποιος μπορεί να χρησιμοποιήσει τις εντολές:

Analyze-Compare means-Paired samples T Test

Στη συνέχεια επιλέγουμε τις δύο μεταβλητές, εισάγουμε τη σχηματιζόμενη διαφορά στο κουτί **Paired variables** και πατάμε OK. Τα αποτελέσματα με τη δεύτερη προσέγγιση φαίνονται παρακάτω (προφανώς το συμπέρασμα θα είναι το ίδιο με οποιοδήποτε τρόπο και να γίνει η ανάλυση).

Από το πρώτο μέρος των αποτελεσμάτων παρατηρούμε ότι ο μέσος όρος της αυτο-εκτίμησης των μαθητών έχει αυξηθεί αρκετά (54,4 πριν τη θεραπεία έναντι 61,3 μετά). Αυτό είναι μία πρώτη ένδειξη ότι πράγματι υπάρχει στατιστικά σημαντική αύξηση. Η εξαγωγή των συμπερασμάτων γίνεται με βάση τα αποτελέσματα στο μέρος Paired Sample Test. Η στήλη Sig (2-tailed) δίνει την τιμή p για τον δίπλευρο έλεγχο. Για να πάρουμε την τιμή p του μονόπλευρου ελέγχου ακολουθούμε τη διαδικασία που είχαμε αναφέρει και στην περίπτωση του ενός μέσου όρου για τη μεταβλητή diff. Σύμφωνα με την εφαρμογή μας ενδιαφέρει να εξετάσουμε το μονόπλευρο έλεγχο για τιμές μικρότερες της προκαθορισμένης τιμής, αφού η βελτίωση της αυτο-εκτίμησης σημαίνει $\text{diff} < 0$ (περίπτωση 2Γ). Άρα η τιμή p του μονόπλευρου ελέγχου είναι 0,002 και μπορούμε να πούμε ότι σε επίπεδο σημαντικότητας 5% η συγκεκριμένη θεραπεία βελτιώνει την αυτο-εκτίμηση.

4.6. Σχόλια για τον έλεγχο μεταξύ των μέσων όρων 2 πληθυσμών

Σε όλα τα παραπάνω παραδείγματα είδαμε ότι ισχύει η υπόθεση της κανονικότητας. Για το λόγο αυτό χρησιμοποιήσαμε τους κατάλληλους ελέγχους t . Όταν δεν ισχύει η υπόθεση της κανονικότητας τότε χρησιμοποιούμε μία σειρά άλλων ελέγχων που δεν έχουν ως προϋπόθεση τα δεδομένα να ακολουθούν μία κανονική κατανομή. Αυτοί οι έλεγχοι ονομάζονται μη παραμετρικοί (non-parametric tests). Προφανώς η κατηγορία ελέγχων t στους οποίους αναφερθήκαμε προηγουμένως αναφέρονται και ως παραμετρικοί έλεγχοι (parametric tests).

Οι εντολές για κάθε κατηγορία ελέγχου είναι οι ακόλουθες.

- Για 2 ανεξάρτητα δείγματα έχουμε τον έλεγχο Mann Whitney:

Analyze-Non-Parametric Tests-2 Independent Samples

και μετά Grades στο test variable list, sex στο group variable και επέλεξε Mann Whitney U στο test type

- Για εξαρτημένα δείγματα έχουμε τον έλεγχο Wilcoxon signed rank:

Analyze-Non-Parametric Tests-2 related Samples

και μετά before και after στο test pair list και επέλεξε Wilcoxon στο test type.

- Το SPSS δεν έχει ενσωματωμένους μη παραμετρικούς ελέγχους για την περίπτωση του ενός πληθυσμού.

Μπορούμε όμως να δημιουργήσουμε τον έλεγχο αυτό βασιζόμενοι στον έλεγχο Wilcoxon signed rank για εξαρτημένα δείγματα στην περίπτωση δύο πληθυσμών. Πιο συγκεκριμένα εκτός από τη μεταβλητή το μέσο όρο της οποίας εξετάζουμε (Hours για το παράδειγμά μας προηγουμένως) δημιουργούμε μία νέα μεταβλητή (έστω New το όνομα της) η οποία έχει τιμή 0 παντού. Στη συνέχεια εφαρμόζουμε τον έλεγχο Wilcoxon signed rank για εξαρτημένα δείγματα βάζοντας hours και new στο test pair list και επιλέγοντας Wilcoxon στο test type. Η δημιουργία της new γίνεται με την ακόλουθη διαδικασία: **Transform-Compute** και μετά new στο Target Variable και 0 στο Numeric Expression.

Όπως αναφέραμε και στην ανάπτυξη των ελέγχων οι αρχικές υποθέσεις είναι ότι δεν υπάρχει διαφορά μεταξύ

- (i) του μέσου όρου και της προκαθορισμένης τιμής
- (ii) μεταξύ των μέσων όρων από 2 πληθυσμούς και
- (iii) μεταξύ των μέσων όρων του ίδιου πληθυσμού σε 2 διαφορετικές χρονικές στιγμές.

Η αρχική ή μηδενική υπόθεση για κάθε περίπτωση εξετάζεται με βάση το κριτήριο ελέγχου και το επίπεδο στατιστικής σημαντικότητας ώστε να διαπιστώσουμε αν τα δεδομένα την απορρίπτουν ή όχι. Όπως είδαμε η μηδενική υπόθεση εξετάζεται για να διαπιστώσουμε αν από τα δεδομένα μπορούμε να πούμε ότι ισχύει τελικά αυτή η υπόθεση ή απορρίπτεται προς όφελος μίας άλλης υπόθεσης (συνήθως κάποιας συγκεκριμένης μορφής ανισότητα μονόπλευρου ή δίπλευρου ελέγχου). Δηλαδή, η αρχική υπόθεση εξετάζεται σε σχέση με μία άλλη υπόθεση που λέει ότι ο μέσος όρος είναι μεγαλύτερος/μικρότερος/διαφορετικός από την προκαθορισμένη τιμή ή από τον άλλο

μέσο όρο. Αυτή η δεύτερη υπόθεση όπως έχουμε πει ονομάζεται εναλλακτική υπόθεση ή ερευνητική υπόθεση ή υπόθεση εργασίας (alternative hypothesis-research hypothesis). Στην ουσία η υπόθεση εργασίας είναι αυτό που προσπαθούμε να αποδείξουμε και με βάση τα δεδομένα, το κριτήριο ελέγχου και το επίπεδο σημαντικότητας εξετάζουμε αν υπάρχουν επαρκή στοιχεία για να απορρίψουμε την αρχικώς θεωρούμενη ως σωστή μηδενική υπόθεση προς όφελος της εναλλακτικής υπόθεσης ή υπόθεσης εργασίας.

Ένα παρόμοιο σκεπτικό υπάρχει και στο σύστημα απονομής δικαιοσύνης. Αρχικώς το δικαστήριο θεωρεί ότι ο κατηγορούμενος είναι αθώος (μηδενική υπόθεση). Στη συνέχεια μέσα από την ακροαματική διαδικασία (τα δεδομένα του στατιστικού προβλήματος) και το σύστημα απονομής δικαστικής απόφασης (πλειοψηφία ή ομοφωνία των ενόρκων-δικαστών, δηλαδή το ανάλογο του επιπέδου στατιστικής σημαντικότητας) απορρίπτουμε ή όχι την αρχική υπόθεση προς όφελος της καταδικαστικής (εναλλακτική υπόθεση).

5. ΑΝΑΛΥΣΗ ΔΙΑΚΥΜΑΝΣΗΣ

5.1. Εισαγωγή

Η ανάλυση διακύμανσης (Analysis of Variance – ANOVA) αποτελεί τη φυσική επέκταση της διαδικασίας ελέγχου του μέσου όρου ενός ή δύο πληθυσμών που περιγράφηκε προηγουμένως. Προσφέρει όμως περισσότερες δυνατότητες από τον έλεγχο t. Πιο συγκεκριμένα σε αυτή τη μεθοδολογία δεν ισχύει ο περιορισμός στη σύγκριση ενός ή δύο μέσων όρων. Ο ερευνητής μπορεί να συγκρίνει τους μέσους όρους που προέρχονται από τρεις, τέσσερις ή και περισσότερους πληθυσμούς. Επίσης η τεχνική αυτή επιτρέπει στον ερευνητή να εξετάσει την επίδραση περισσότερων από μίας μεταβλητής στους υπό εξέταση πληθυσμούς. Για την καλύτερη κατανόηση των διαφορών με τον έλεγχο t αναφέρουμε το σχετικό παράδειγμα:

Παράδειγμα: Ένας ερευνητής επιθυμεί να εξετάσει τις διαφοροποιήσεις στο μέσο όρο του εισοδήματος των αποφοίτων 3 διαφορετικών οικονομικών σχολών. Είναι επίσης πιθανό, εκτός από την τυχόν διαφοροποίηση μεταξύ των σχολών, ο ερευνητής να ενδιαφέρεται να αναλύσει και τυχόν μεταβολές στο εισόδημα μεταξύ του φύλου των αποφοίτων.

Ο έλεγχος με το κριτήριο t προσφέρει πολύ λίγες δυνατότητες για το σκοπό αυτού του παραδείγματος. Πιο συγκεκριμένα ο οποιοσδήποτε έλεγχος μπορεί να γίνει μόνο μεταξύ 2 πληθυσμών. Στην προκειμένη περίπτωση έχουμε 3 διαφορετικές σχολές. Μία αρχική σκέψη θα ήταν να κάνει διαδοχικούς ελέγχους t για όλους τους πιθανούς συνδυασμούς ζευγαριών (είναι 3 συνολικά) μεταξύ των 3 σχολών. Αυτό εκτός από ασύμφορο από πρακτική άποψη (σε περίπτωση 8 σχολών θα έπρεπε να γίνουν συνολικά 28 έλεγχοι t) είναι και στατιστικά μεμπτό. Ο λόγος είναι ότι οι πολλαπλοί έλεγχοι t αυξάνουν κατά πολύ την πιθανότητα λάθους συμπεράσματος.

Υπάρχει όμως και ένας άλλος περιορισμός. Το κριτήριο t ελέγχει την επίδραση στο εισόδημα μόνο ενός παράγοντα κάθε φορά, (πχ. σχολή ή φύλο) και όχι ταυτόχρονα και των δύο. Στο παράδειγμά μας έχουμε δημιουργήσει 6 διαφορετικά επίπεδα εξέτασης του μέσου όρου (3 σχολές όπου σε κάθε σχολή εξετάζουμε άνδρες και γυναίκες

ξεχωριστά). Μία προσέγγιση θα ήταν να κάνουμε διαδοχικούς ελέγχους t μεταξύ των 6 επιπέδων αλλά τότε επανερχόμαστε στον αρχικό προβληματισμό.

Η μεταβλητή της οποίας τον μέσο όρο εξετάζουμε (όπως εισόδημα στο παράδειγμα) ονομάζεται *εξαρτημένη μεταβλητή (dependent variable)* ενώ οι παράγοντες βάσει των οποίων δημιουργούμε τα επίπεδα εξέτασης (εδώ σχολή και φύλο) ονομάζονται *ανεξάρτητες μεταβλητές (independent variables)*. Σε περίπτωση όπου έχουμε μόνο μία ανεξάρτητη μεταβλητή η μεθοδολογία ονομάζεται *ανάλυση διακύμανσης μίας (ή μονής) κατεύθυνσης (one-way analysis of variance, one-way ANOVA)*. Ομοίως σε περίπτωση δύο ανεξάρτητων μεταβλητών έχουμε *ανάλυση διακύμανσης διπλής κατεύθυνσης*. Είναι προφανές ότι η one-way ANOVA με 2 επίπεδα εξέτασης είναι ισοδύναμη διαδικασία με το γνωστό μας έλεγχο t.

Η διαδικασία ANOVA στηρίζεται (όπως και ο έλεγχος t) στην κανονικότητα των δεδομένων για κάθε επίπεδο πληθυσμού. Θα πρέπει λοιπόν προτού εφαρμοσθεί η συγκεκριμένη τεχνική να εξετάζεται αν υπάρχουν σοβαρές αποκλίσεις από την κανονικότητα. Οι τρόποι για την παραπάνω διαπίστωση είναι οι ίδιοι με του ελέγχου t. Σε περίπτωση σοβαρών αποκλίσεων από την κανονικότητα υπάρχουν διάφορες μη-παραμετρικές στατιστικές διαδικασίες (όπως ο έλεγχος Kruskal-Wallis). Επίσης τα δεδομένα της εξαρτημένης μεταβλητής πρέπει να είναι συνεχή.

Τέλος τονίζουμε την ύπαρξης μεγάλης ποικιλίας μεθόδων ανάλυσης διακύμανσης. Οι σημειώσεις αυτές θα επικεντρωθούν στην περίπτωση της one-way ANOVA με κάποιες αναφορές για την περίπτωση διπλής κατεύθυνσης. Σε όλες τις περιπτώσεις θα έχουμε ανεξάρτητα δείγματα.

5.2. Μεθοδολογία ANOVA

Στην ενότητα αυτή θα παρουσιάσουμε το σκεπτικό ανάπτυξης του κριτηρίου ANOVA, τη μεθοδολογία που ακολουθείται για να διαπιστώσουμε που υπάρχουν οι συγκεκριμένες διαφοροποιήσεις και τέλος κάποιες σκέψεις για την επέκταση στη διπλή κατεύθυνση.

Το σκεπτικό για τη διαδικασία ANOVA

Η ANOVA αναλύει τη συνολική μεταβλητότητα όλων των τιμών. Δηλαδή αναλύει πόσο «διάσπαρτες» είναι όλες οι τιμές σε σχέση με ένα συνολικό μέσο όρο όλων

των τιμών ανεξάρτητα σε ποιο δείγμα ανήκουν. Αυτή η συνολική μεταβλητότητα παραμένει σταθερή και μπορεί να επιμερισθεί σε δύο (όχι κατ' ανάγκη ίσα) μέρη:

1. Τη μεταβλητότητα *μεταξύ των δειγμάτων* δηλαδή πόσο διαφέρει ο μέσος όρος κάθε δείγματος από το συνολικό μέσο όρο και
2. Τη μεταβλητότητα *μέσα στα δείγματα*, δηλαδή σε κάθε δείγμα πόσο διαφέρει κάθε παρατήρηση από το μέσο όρο του συγκεκριμένου δείγματος.

Σε περίπτωση όπου υπάρχουν μεγάλες διαφορές μεταξύ των μέσων όρων τότε η μεταβλητότητα μεταξύ των δειγμάτων είναι ιδιαίτερα μεγάλη και άρα η μεταβλητότητα μέσα στα δείγματα είναι μικρή (αφού η συνολική μεταβλητότητα παραμένει σταθερή). Έτσι το κριτήριο ελέγχου στην ANOVA είναι ο λόγος των δύο πηγών μεταβλητότητας που αναφέρθηκαν παραπάνω, δηλαδή

$$\text{Κριτήριο ελέγχου} = (\text{μεταβλ. μεταξύ των δειγμάτων}) / \text{μεταβλ. μέσα στα δείγματα}$$

Όσο μεγαλύτερες τιμές λαμβάνει αυτό το κλάσμα τόσο πιο πιθανό είναι να υπάρχουν διαφορές μεταξύ των επιπέδων. Ο λόγος αυτός ακολουθεί μία γνωστή στατιστική κατανομή F και ισχύουν όλα είχαμε αναφέρει και στον έλεγχο t για την πιθανότητα λάθους συμπεράσματος. Έτσι και εδώ υπάρχει η τιμή p και το επίπεδο στατιστικής σημαντικότητας.

Post hoc κριτήρια

Σε περίπτωση όπου διαπιστώσουμε την ύπαρξη στατιστικά σημαντικών διαφορών μεταξύ των μέσων όρων, το επόμενο βήμα είναι να εντοπίσουμε τα επίπεδα μεταξύ των οποίων υπάρχουν διαφορές. Για παράδειγμα στο εισαγωγικό παράδειγμα με τον παράγοντα των διαφορετικών σχολών, μας ενδιαφέρει να εξετάσουμε μεταξύ ποιών ακριβώς σχολών υπάρχει διαφορά ως προς το μέσο εισόδημα. Πιο σε συγκεκριμένα, σε ποια ζεύγη μεταξύ των 3 (1^η με 2^η σχολή, 1^η με 3^η σχολή και 2^η με 3^η σχολή) εντοπίζονται διαφορές;

Επειδή αυτή η ανάλυση γίνεται μετά τη διαπίστωση στατιστικά σημαντικών διαφορών μεταξύ των μέσων όρων ονομάζεται *post hoc* ανάλυση (εκ των υστέρων ανάλυση ή και *a posteriori* ανάλυση) και τα κριτήρια *post hoc* κριτήρια. Υπάρχουν πάρα

πολλά σχετικά κριτήρια. Θα αναφερθούμε περιληπτικά στα τέσσερα από τα πιο σημαντικά:

- Το κριτήριο LSD (Least Squares Differences). Ουσιαστικά πρόκειται για πολλαπλούς ελέγχους t και δεν αντιμετωπίζεται το πρόβλημα της αύξησης της πιθανότητας λάθους συμπεράσματος. Λιγότερο επικίνδυνο για μικρό αριθμό ελέγχων, καλό είναι πάντως να αποφεύγεται.
- Το κριτήριο Bonferroni. Είναι ένα «αυστηρό» κριτήριο όσον αφορά το επίπεδο σημαντικότητας και αντιμετωπίζει ικανοποιητικά το πρόβλημα της αύξησης της πιθανότητας λάθους συμπεράσματος.
- Το κριτήριο Tukey HSD (Honestly Significant Difference). Θεωρείται το πιο ασφαλές κριτήριο. Είναι ένα «συντηρητικό» κριτήριο με την έννοια ότι κάποιες στατιστικά σημαντικές διαφορές με άλλα κριτήρια είναι πιθανό να μην εμφανίζονται εδώ.
- Το κριτήριο Scheffe. Θεωρείται συντηρητικό κριτήριο. Εφαρμόζεται συχνά και καλύπτει πλήθος περιπτώσεων. Μαζί με το Tukey είναι από τα πιο χρήσιμα στην πράξη.

5.3. Σχόλια

Η μεθοδολογία ANOVA μπορεί να επεκταθεί και σε παραπάνω από μία ανεξάρτητη μεταβλητή. Το σκεπτικό δεν αλλάζει απλά οι υπολογισμοί γίνονται πιο δύσκολοι, κάτι το οποίο δεν αποτελεί πρόβλημα στο SPSS (ή και σε οποιοδήποτε άλλο στατιστικό λογισμικό). Ένα σημαντικό βήμα που δεν υπήρχε προηγουμένως είναι ότι τώρα μπορεί να εμφανισθεί και αλληλεπίδραση μεταξύ των ανεξάρτητων μεταβλητών. Αυτή η αλληλεπίδραση υπολογίζεται ως μία επιπλέον ανεξάρτητη μεταβλητή.

Τα κριτήρια της post hoc ανάλυσης δίνουν συνήθως (αλλά όχι πάντα) ταυτόσημα συμπεράσματα. Καλό είναι να επιλέγουμε κριτήρια με βάση τις στατιστικές τους ιδιότητες και όχι με βάση το ποιο δίνει στατιστικά σημαντικές διαφορές.

Όπως και στον έλεγχο t , ισχύει το φαινόμενο της «συχνής» και όχι αναγκαστικά πραγματικής στατιστικά σημαντικής διαφοράς όταν τα δείγματα είναι μεγάλα σε μέγεθος.

Στην εφαρμογή θα γίνει επίσης μνεία για τις εντολές στο μη παραμετρικό έλεγχο Kruskal-Wallis. Στην πράξη δεν είναι αρκετά συχνή η εφαρμογή τους. Θα πρέπει να υπάρχουν σοβαρές αποκλίσεις από την κανονικότητα για να τις χρησιμοποιήσουμε.

5.4. Εφαρμογή

Μας ενδιαφέρει να συγκρίνουμε την απόδοση τριών σχολικών τάξεων στη μαθηματική ικανότητα με βάση τις εξετάσεις που διεξάγει κάθε 3 χρόνια σε παιδιά ηλικίας 15 ετών ο ΟΟΣΑ μέσω του προγράμματος PISA (Program for the International Student Assessment, δείτε και το σχετικό site www.pisa.oecd.org). Τα αποτελέσματα παρουσιάζονται παρακάτω (οι αριθμοί είναι υποθετικοί και όχι πραγματικοί):

Σχολείο 1:

469, 474, 478, 465, 459, 489, 478

Σχολείο 2:

462, 465, 447, 431, 453, 467, 466

Σχολείο 3:

467, 453, 472, 451, 457, 443, 456

Εισάγουμε τα αποτελέσματα στο SPSS σε δύο στήλες (pisa και school). Με τους τρόπους που αναλύθηκαν στο προηγούμενο κεφάλαιο παρατηρούμε ότι δεν παραβιάζεται η προϋπόθεση της κανονικότητας και για αυτό το λόγο θα εφαρμόσουμε one-way ANOVA. Θέτουμε το 5% ως επίπεδο στατιστικής σημαντικότητας.

Χρησιμοποιούμε τις εξής εντολές:

Analyze-General Linear Model-Univariate

Ως **Dependent Variable** βάζουμε τη μεταβλητή pisa και ως **Fixed Factor(s)** τη μεταβλητή school. (Σημείωση: εάν είχαμε two-way ANOVA, θα έβαζα και τις δύο ανεξάρτητες μεταβλητές στο σημείο αυτό). Κάνω κλικ στο **Post hoc** και αφού βάλω τη μεταβλητή school στο μενού **Post hoc tests for** επιλέγω τα τεστ Scheffe, Tukey. Πατώ continue και OK και παίρνω τα παρακάτω αποτελέσματα.

Univariate Analysis of Variance

Between-Subjects Factors

		N
SCHOOL	1.00	7
	2.00	7
	3.00	7

Tests of Between-Subjects Effects

Dependent Variable: PISA

Source	Type	III Df	Mean Square	F	Sig.
	Sum of Squares				
Corrected Model	1308.286	2	654.143	5.329	.015
Intercept	4482324.000	1	4482324.000	36512.337	.000
SCHOOL	1308.286	2	654.143	5.329	.015
Error	2209.714	18	122.762		
Total	4485842.000	21			
Corrected Total	3518.000	20			

a. R Squared = .372 (Adjusted R Squared = .302)

Multiple Comparisons

Dependent Variable: PISA

	(I)	(J)	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
	SCHOOL	SCHOOL				Lower Bound	Upper Bound
Tukey HSD	1.00	2.00	17.2857	5.9224	.024	2.1707	32.4007
		3.00	16.1429	5.9224	.035	1.0279	31.2579

	2.00	1.00	-17.2857	5.9224	.024	-32.4007	-2.1707
		3.00	-1.1429	5.9224	.980	-16.2579	13.9721
	3.00	1.00	-16.1429	5.9224	.035	-31.2579	-1.0279
		2.00	1.1429	5.9224	.980	-13.9721	16.2579
Scheffe	1.00	2.00	17.2857	5.9224	.031	1.4949	33.0766
		3.00	16.1429	5.9224	.045	.3520	31.9337
	2.00	1.00	-17.2857	5.9224	.031	-33.0766	-1.4949
		3.00	-1.1429	5.9224	.982	-16.9337	14.6480
	3.00	1.00	-16.1429	5.9224	.045	-31.9337	-.3520
		2.00	1.1429	5.9224	.982	-14.6480	16.9337

Based on observed means.

* The mean difference is significant at the .05 level.

Το κριτήριο ελέγχου F και η αντίστοιχη τιμή p δίνονται από το δεύτερο μέρος των αποτελεσμάτων και ειδικότερα στις στήλες F και Sig. Οι τιμές αυτές είναι 5,329 και 0,015 αντιστοίχως. Άρα σε επίπεδο στατιστικής σημαντικότητας 5% (όχι όμως και αν είχαμε επιλέξει το 1%) υπάρχει στατιστικά σημαντική διαφορά μεταξύ των τριών σχολικών μονάδων. Αυτό είναι μάλλον αναμενόμενο αν κοιτάξουμε και τους 3 μέσους όρους (σχολείο 1=473, 1 σχολείο 2=455,9 και σχολείο 3=457) ενώ οι τυπικές αποκλίσεις τους δεν διαφέρουν σημαντικά (γύρω στο 10).

Η post hoc ανάλυση συγκρίνει με τα κριτήρια Tukey και Scheffe τις διαφορές ανά σχολείο και παρουσιάζεται στο κομμάτι με τον τίτλο Multiple Comparisons. Για κάθε μέθοδο, κάθε σχολείο (I) συγκρίνεται με τα άλλα 2 σχολεία (J). Έτσι με τη μέθοδο Tukey ο δίπλευρος έλεγχος του σχολείου 1 με το σχολείο 2 έδωσε μία τιμή p 0,024. Αν μας ενδιαφέρει ο μονόπλευρος έλεγχος για το αν το σχολείο 1 έχει μέσο όρο μεγαλύτερο από το σχολείο 2 (όπως φαίνεται και από τις αριθμητικές τους τιμές), τότε σύμφωνα με την τεχνική που αναπτύχθηκε στην περίπτωση ελέγχου t για το μέσο όρο 2 πληθυσμών για ανεξάρτητα δείγματα η τιμή p είναι 0,012. Αυτό σημαίνει ότι το σχολείο 1 υπερτερεί κατά μέσο όρο του σχολείου 2 στη μέση μαθηματική επίδοση στο PISA σε επίπεδο στατιστικής σημαντικότητας 5% (όχι όμως και για 1%). Παρόμοια συμπεράσματα βγαίνουν από τη σύγκριση μεταξύ των σχολείων 1 και 3, ενώ δεν υπάρχουν διαφορές μεταξύ των σχολείων 2 και 3. Τέλος τα συμπεράσματα είναι ταυτόσημα και με τις 2 μεθόδους.

Οι εντολές για το μη-παραμετρικό έλεγχο Kruskal Wallis είναι οι ακόλουθες: **Analyze-Non-Parametric Tests-K Independent Samples** και μετά βάζω pisa στο test variable list και school στο group variable.

Πίνακας 3: *Σύνοψη των κυριότερων ελέγχων μέσων όρων*

Περίπτωση	Παραμετρικοί έλεγχοι	Μη-παραμετρικοί
Ένας μέσος όρος	t-test για ένα πληθυσμό	Wicoxon signed rank
Δύο μέσοι όροι (ανεξάρτητα δείγματα)	t-test για δύο πληθυσμούς	Mann-Whitney
Δύο μέσοι όροι (εξαρτημένα δείγματα)	Paired t-test για δύο πληθυσμούς	Wicoxon signed rank
Περισσότεροι από δύο μέσοι όροι (ανεξάρτητα δείγματα)	ANOVA	Kruskal-Wallis (για one-way)

6. ΤΟ ΚΡΙΤΗΡΙΟ X^2

Το κριτήριο X^2 είναι γνωστό ως κριτήριο ελέγχου ανεξαρτησίας (chi square test of independence) ή κριτήριο ελέγχου πινάκων συνάφειας (contingency tables). Οι πίνακες συνάφειας ή πίνακες διπλής εισόδου αναπτύσσονται σε (έστω k) γραμμές και (έστω l) στήλες. Το κριτήριο εξετάζει αν οι δύο μεταβλητές που απαρτίζουν τον πίνακα διπλής εισόδου είναι ανεξάρτητες ή όχι. Πρόκειται ουσιαστικά για στατιστικό έλεγχο ο οποίος βασίζεται στη χρήση της στατιστικής κατανομής X^2 σε επίπεδο σημαντικότητας α . Το σκεπτικό του ελέγχου είναι το εξής: Αρχικά υποθέτουμε ότι δεν υπάρχει σχέση μεταξύ των μεταβλητών (δηλαδή είναι ανεξάρτητες οι μεταβλητές μεταξύ τους). Μπορούμε να υπολογίσουμε τότε τις αναμενόμενες συχνότητες (δηλαδή τις συχνότητες που αναμένουμε αν ισχύει η αρχική υπόθεση ότι δεν υπάρχει σχέση). Τις συγκρίνουμε με τις πραγματικές συχνότητες. Αν διαφέρουν πολύ τότε απορρίπτουμε την υπόθεση της ανεξαρτησίας και λέμε ότι υπάρχει σχέση μεταξύ των 2 μεταβλητών.

Οι προϋποθέσεις εφαρμογής του κριτηρίου X^2 είναι οι ακόλουθες:

- Το εξεταζόμενο δείγμα θα πρέπει να έχει επιλεγεί τυχαία και έχει πλήθος >25
- Οι παρατηρήσεις πρέπει να είναι ανεξάρτητες
- Στις περιπτώσεις των κατηγοριοποιημένων μεταβλητών οι επιμέρους κατηγορίες θα πρέπει να μην επικαλύπτονται
- Δεν πρέπει πολλές (συνήθως πάνω από το 25%) από τις αναμενόμενες συχνότητες να είναι μικρότερες του πέντε
- Το κριτήριο δεν είναι αξιόπιστο όταν το συνολικό δείγμα είναι πολύ μεγάλο γιατί τότε μπορεί να παρουσιάσει ότι υπάρχει στατιστικά σημαντική σχέση μεταξύ των 2 μεταβλητών ενώ αυτό δε θα ισχύει στην πραγματικότητα.

6.1. Εφαρμογή

Ο ερευνητής Α επιχειρεί να ελέγξει κατά πόσο το φύλο επηρεάζει τη δήλωση των υποκειμένων αναφορικά με το βαθμό πολιτικής τους ενημέρωσης (ή διαφορετικά εάν ο βαθμός πολιτικής ενημέρωσης επηρεάζεται από το φύλο)²³.

Προκειμένου να ελεγχθεί η υπόθεση εργασίας, διατυπώνεται η μηδενική της υπόθεση και εξετάζεται η ανεξαρτησία των δύο μεταβλητών με εφαρμογή του ελέγχου ανεξαρτησίας (chi square test of independence). Οι υποθέσεις διαμορφώνονται ως εξής:

H_0 : Ο βαθμός πολιτικής ενημέρωσης ΔΕΝ επηρεάζεται από το φύλο

H_1 : Ο βαθμός πολιτικής ενημέρωσης επηρεάζεται από το φύλο

Επιλέγεται το στατιστικό κριτήριο X^2 , προκειμένου ο ερευνητής να συγκρίνει (υπό όρους) τις παρατηρούμενες συχνότητες με τις συχνότητες που θα εμφανίζονταν εάν η μηδενική υπόθεση ήταν αληθής (αναμενόμενες συχνότητες). Αρχικά υπολογίζονται οι συχνότητες για τις τιμές κάθε μεταβλητής. Ο υπολογισμός των συχνοτήτων με τη βοήθεια του SPSS έχει ως εξής:

Πίνακας 4: Υπολογισμός συχνοτήτων για τις τιμές της μεταβλητής του

Φύλο					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Κορίτσι	83	41.5	41.5	41.5
	Αγόρι	117	58.5	58.5	100.0
	Total	200	100.0	100.0	

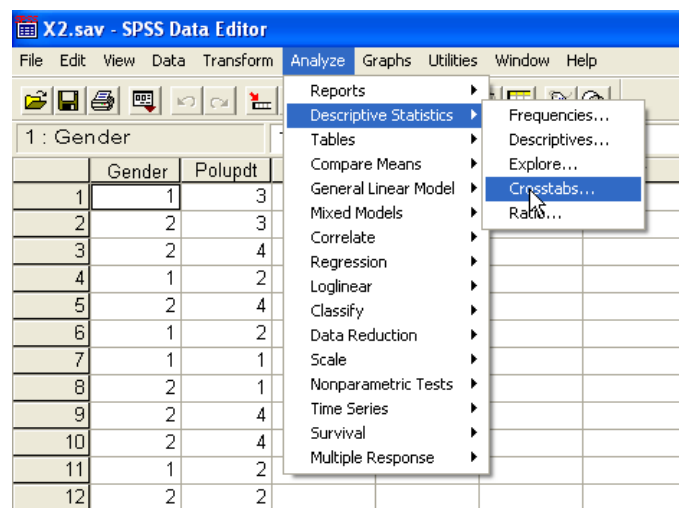
²³ Ενδεικτική σχετική ερώτηση: Πόσο νομίζεις ότι είσαι ενημερωμένη/ος-η για τα πολιτικά ζητήματα; [Πάρα πολύ=1, Πολύ=2, Αρκετά=3, Λίγο=4, Πολύ λίγο=5, Καθόλου=6]

Πίνακας 5: Υπολογισμός συχνοτήτων για τις τιμές της μεταβλητής ΒΑΘΜΟΣ ΠΟΛΙΤΙΚΗΣ ΕΝΗΜΕΡΩΣΗΣ

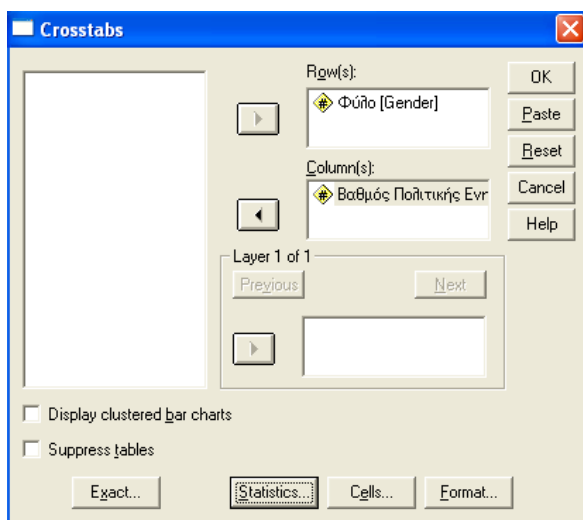
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Πάρα πολύ	37	18.5	18.8	18.8
	Πολύ	40	20.0	20.3	39.1
	Αρκετά	31	15.5	15.7	54.8
	Λίγο	69	34.5	35.0	89.8
	Πολύ λίγο	11	5.5	5.6	95.4
	Καθόλου	9	4.5	4.6	100.0
	Total		197	98.5	100.0
Missing	Missing	3	1.5		
Total		200	100.0		

Αυτό βέβαια που ενδιαφέρει είναι η συνεξέταση των μεταβλητών, γεγονός που επιτυγχάνεται με τη διασταύρωσή τους. Στο SPSS η διασταύρωση των μεταβλητών προκύπτει με τη βοήθεια της επιλογής Crosstabs ως εξής:

- από το Menu Analyze επιλέγεται η εντολή Descriptive Statistics και στη συνέχεια Crosstabs.



Στο πλαίσιο διαλόγου που εμφανίζεται οι μεταβλητές που ενδιαφέρουν τοποθετούνται σε γραμμές (rows) και στήλες (columns).



Σχήματα 25-26: Εντολές του SPSS για το κριτήριο χ^2

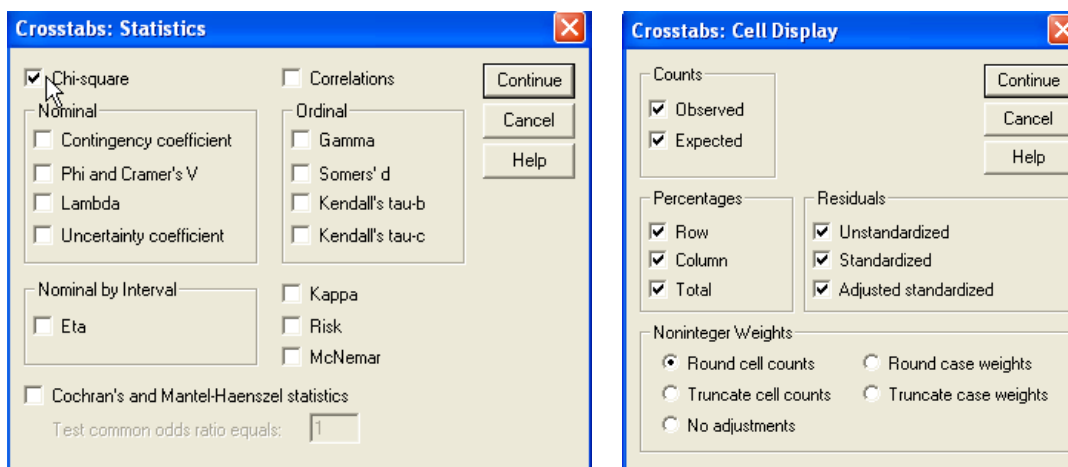
Πίνακας 6: Διασταύρωση της μεταβλητής ΦΥΛΟ με τη μεταβλητή ΒΑΘΜΟΣ ΠΟΛΙΤΙΚΗΣ ΕΝΗΜΕΡΩΣΗΣ

Φύλο * Βαθμός Πολιτικής Ενημέρωσης Crosstabulation

Count		Βαθμός Πολιτικής Ενημέρωσης						Total
		Πάρα πολύ	Πολύ	Αρκετά	Λίγο	Πολύ λίγο	Καθόλου	
Φύλο	Κορίτσι	24	19	17	17	2	2	81
	Αγόρι	13	21	14	52	9	7	116
Total		37	40	31	69	11	9	197

Στον παραπάνω πίνακα συνάφειας οι τιμές στα κελιά του πίνακα είναι οι παρατηρούμενες συχνότητες. Για το στατιστικό έλεγχο είναι απαραίτητος ο προσδιορισμός των αναμενόμενων συχνοτήτων. Οι αναμενόμενες συχνότητες αναφέρονται στις συχνότητες που ο ερευνητής θα διαπίστωνε εάν ίσχυε η μηδενική υπόθεση. Για το SPSS η σχετική διαδικασία είναι ιδιαίτερα απλή. Ως συνέχεια της διαδικασίας που προτάθηκε παραπάνω (Analyse - Descriptive Statistics - Crosstabs), στην επιλογή Statistics (βλ. σχετική εικόνα) να επιλέξουμε το στατιστικό κριτήριο που μας ενδιαφέρει (στη συγκεκριμένη περίπτωση Chi-square)²⁴ και να ζητηθεί η εμφάνιση των παρατηρούμενων και αναμενόμενων συχνοτήτων σε απόλυτες τιμές και ποσοστά.

²⁴ Οι δυνατές επιλογές είναι όσες εμφανίζονται στο σχετικό πλαίσιο διαλόγου. Για την επιλογή είναι απαραίτητο να εξετάζεται η τήρηση των σχετικών προδιαγραφών.



Σχήμα 27: Επιλογές του SPSS για το κριτήριο χ^2

Τα αποτελέσματα έχουν ως εξής:

Πίνακας 7: Διασταύρωση της μεταβλητής Φύλο με τη μεταβλητή Βαθμός Πολιτικής Ενημέρωσης με εμφάνιση των παρατηρούμενων και αναμενόμενων συχνοτήτων

Φύλο * Βαθμός Πολιτικής Ενημέρωσης Crosstabulation

		Βαθμός Πολιτικής Ενημέρωσης						Total	
		Πάρα-πολύ	Πολύ	Αρκετά	Λίγο	Πολύ λίγο	Καθόλου		
Φύλο	Κορίτσι	Count	24	19	17	17	2	81	
	Expected Count	15.2	16.4	12.7	28.4	4.5	3.7	81.0	
	% within Φύλο	29.6%	23.5%	21.0%	21.0%	2.5%	2.5%	100.0%	
	% within Βαθμός Πολιτικής Ενημέρωσης	64.9%	47.5%	54.8%	24.6%	18.2%	22.2%	41.1%	
	% of Total	12.2%	9.6%	8.6%	8.6%	1.0%	1.0%	41.1%	
	Residual	8.8	2.6	4.3	-11.4	-2.5	-1.7		
	Std. Residual	2.3	.6	1.2	-2.1	-1.2	-.9		
	Adjusted Residual	3.3	.9	1.7	-3.5	-1.6	-1.2		
	Αγόρι	Count	13	21	14	52	9	7	116
	Expected Count	21.8	23.6	18.3	40.6	6.5	5.3	116.0	
% within Φύλο	11.2%	18.1%	12.1%	44.8%	7.8%	6.0%	100.0%		
% within Βαθμός Πολιτικής Ενημέρωσης	35.1%	52.5%	45.2%	75.4%	81.8%	77.8%	58.9%		
% of Total	6.6%	10.7%	7.1%	26.4%	4.6%	3.6%	58.9%		
Residual	-8.8	-2.6	-4.3	11.4	2.5	1.7			
Std. Residual	-1.9	-.5	-1.0	1.8	1.0	.7			
Adjusted Residual	-3.3	-.9	-1.7	3.5	1.6	1.2			
Total	Count	37	40	31	69	11	9	197	
Expected Count	37.0	40.0	31.0	69.0	11.0	9.0	197.0		
% within Φύλο	18.8%	20.3%	15.7%	35.0%	5.6%	4.6%	100.0%		
% within Βαθμός Πολιτικής Ενημέρωσης	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%		
% of Total	18.8%	20.3%	15.7%	35.0%	5.6%	4.6%	100.0%		

Annotations in the image:
 - Blue arrow: Παρατηρούμενες Συχνότητες (Observed Frequencies) pointing to the Count row.
 - Red arrow: Αναμενόμενες Συχνότητες (Expected Frequencies) pointing to the Expected Count row.
 - Green arrow: Υπόλοιπο (Residual) pointing to the Residual row.

Πίνακας 8: Διασταύρωση της μεταβλητής Φύλο με τη μεταβλητή Βαθμός Πολιτικής Ενημέρωσης με εμφάνιση των παρατηρούμενων και αναμενόμενων συχνοτήτων

			Βαθμός Πολιτικής Ενημέρωσης					Total	
			Πάρα πολύ	Πολύ	Αρκετά	Λίγο	Πολύ λίγο		Καθόλου
Φύλο	Κορίτσι	Count	24	19	17	17	2	2	81
		Expected Count	15.2	16.4	12.7	28.4	4.5	3.7	81.0
		Residual	8.8	2.6	4.3	-11.4	-2.5	-1.7	
Αγόρι	Count	Count	13	21	14	52	9	7	116
		Expected Count	21.8	23.6	18.3	40.6	6.5	5.3	116.0
		Residual	-8.8	-2.6	-4.3	11.4	2.5	1.7	
Total	Count	Count	37	40	31	69	11	9	197
		Expected Count	37.0	40.0	31.0	69.0	11.0	9.0	197.0

Στον πίνακα εμφανίζονται οι παρατηρούμενες συχνότητες, οι αναμενόμενες συχνότητες, οι ποσοστιαίες αναγωγές και τα υπόλοιπα. Οι ενδείξεις υπολοίπων προκύπτουν από τη διαφορά παρατηρούμενων – αναμενόμενων συχνοτήτων. Θετική τιμή υπολοίπου σημαίνει ότι το ισοζύγιο παρατηρούμενων – αναμενόμενων συχνοτήτων βαρύνει υπέρ των παρατηρούμενων (οι παρατηρούμενες συχνότητες είναι περισσότερες από τις αναμενόμενες). Όσο μεγαλύτερη είναι η διαφορά (θετική ή αρνητική), τόσο πιθανότερο είναι να προκύπτει στατιστικά σημαντικό αποτέλεσμα. Με τις κατάλληλες επιλογές η εφαρμογή επιστρέφει και τον πίνακα

	Value	df	Asymp. Sig. (2-sided)	
Δείκτης Pearson χ^2	Pearson Chi-Square	23.159 ^a	5	.000
Λόγος πιθανοφάνειας	Likelihood Ratio	23.826	5	.000
Γρομμάτι προς γραμμάτι συσχέτιση	Linear-by-Linear Association	18.661	1	.000
Αριθμός έγκυρων περιπτώσεων	N of Valid Cases	197		

^a. 2 cells (16.7%) have expected count less than 5. The minimum expected count is 3.70.

Αριθμός κελιών (από 2) που έχουν αναμενόμενη συχνότητα μικρότερη του 5. Η ελάχιστη αναμενόμενη συχνότητα είναι 3.7

Σχήμα 28: Αναλυτική παρουσίαση των αποτελεσμάτων του SPSS

- Ο δείκτης χ^2 είναι ο στατιστικός δείκτης που χρησιμοποιείται για την εκτίμηση της στατιστικής σημαντικότητας.

- Ο λόγος πιθανοφάνειας στα μεγάλα δείγματα συμπίπτει με την τιμή του X^2 .
- Ο δείκτης γραμμικής προς γραμμική συσχέτιση παρέχει μια ένδειξη γραμμικής συσχέτισης των μεταβλητών.
- Τα δύο κελιά με μικρό αριθμό παρατηρήσεων συνιστούν πρόβλημα (περισσότερο εάν το ποσοστό υπερβαίνει το 25%). Στην τελευταία περίπτωση προτείνεται η επανακωδικοποίηση (recode) της μεταβλητής.

Από τον πίνακα συνάφειας των δύο μεταβλητών διαπιστώνουμε ότι οι παρατηρούμενες τιμές διαφέρουν συστηματικά από τις αναμενόμενες²⁵. Είναι όμως και στατιστικά σημαντική η διαφορά; Η τιμή του Pearson Chi-Square είναι ενδεικτική της εξάρτησης. Η βασική παραδοχή είναι ότι όσο η τιμή αυξάνει, τόσο αυξάνει και η πιθανότητα οι μεταβλητές να μην είναι ανεξάρτητες. Έτσι τιμές της πιθανότητας κοντά στο 1 (υπό)δηλώνουν ανεξαρτησία αφού απορρίπτεται η αρχική υπόθεση.

Για X^2 (Pearson Chi-Square) 564.786, p-value (Significance level) 0.000 και επίπεδο σημαντικότητας 5% (σύνηθες στις έρευνες των Επιστημών της Αγωγής) η μηδενική υπόθεση απορρίπτεται, άρα οι δύο μεταβλητές είναι εξαρτημένες (βλ. ενδεικτικά και <http://home.clara.net/sisa/two2hlp.htm>, <http://www.graphpad.com/quickcalcs/contingency1.cfm>)

6.2. Τυπολόγιο ανάγνωσης ενδείξεων για τους πίνακες του SPSS

Στη συνέχεια, παραθέτουμε μία μικρή εξήγηση για τις πιο σημαντικές ενδείξεις που «βγάζει» το SPSS σε αυτό το κριτήριο. Αρκετές από αυτές τις ενδείξεις συνήθως προϋποθέτουν αρκετά καλή γνώση της Στατιστικής. Για το λόγο αυτό ο αναγνώστης-κοινωνικός ερευνητής ας τις διαβάσει «ενημερωτικά» χωρίς να θεωρεί απαραίτητη την πλήρη κατανόηση όλων των όρων.

²⁵ Σημειώστε τις διαφορές ανάμεσα στις παρατηρούμενες και τις αναμενόμενες συχνότητες στη γραμμική των υπολοίπων.

- **Count** (καταμέτρηση περιπτώσεων) = πρόκειται για τον αριθμό που δηλώνει το πλήθος των παρατηρήσεων για κάθε κατηγορία
- **Exp.count** (αναμενόμενη καταμέτρηση περιπτώσεων) = πρόκειται για το πλήθος που θα εμφανίζονταν εάν οι δύο μεταβλητές ήταν εντελώς ανεξάρτητες μεταξύ τους.
- **Residual** (υπόλοιπο περιπτώσεων) = αριθμός ο οποίος προκύπτει από τη διαφορά του πλήθους των παρατηρούμενων τιμών από τις αναμενόμενες.
- **Row Total** (σύνολο γραμμής) = το συνολικό πλήθος τιμών για κάθε γραμμή
- **Column Total** (σύνολο στήλης) = το συνολικό πλήθος τιμών για κάθε στήλη
- **Chi Square Pearson & Likelihood Ratio** (X^2 και λόγος πιθανοφάνειας)
- **Degrees of Freedom** (βαθμοί ελευθερίας) = πρόκειται για τον αριθμό που προκύπτει ως γινόμενο του αριθμού των κατηγοριών της μιας μεταβλητής μείον 1 επί των αριθμό των κατηγοριών της δεύτερης μεταβλητής μείον -1, δηλαδή $(κ-1)(λ-1)$
- **Significance** (σημαντικότητα) = ο αριθμός δηλώνει την πιθανότητα τα αποτελέσματα που προέκυψαν να είναι τυχαία. Τιμές μεγαλύτερες ($>$) του 0.05 δηλώνουν ότι οι παρατηρούμενες τιμές δεν διαφέρουν με τρόπο στατιστικά σημαντικό από τις αναμενόμενες τιμές.
- **Linear by Linear association** (γραμμική συσχέτιση) = με τη χρήση του συγκεκριμένου στατιστικού μέτρου εξετάζεται εάν υφίσταται (γραμμική) συσχέτιση ανάμεσα στις δύο μεταβλητές.
- **Minimum Expected Count** (ελάχιστη αναμενόμενη εκτίμηση)
- **Cells (.x%) have expected count less than 5** = Τιμές άνω του 25% καθιστούν την τιμή του X^2 προβληματική και πιθανώς μη-έγκυρη.
- **Phi (ϕ)** = πρόκειται για τιμή η οποία αναφέρεται στην εκτίμηση της έντασης στη σχέση δύο ονομαστικών μεταβλητών. Ο συντελεστής Phi είναι συνάρτηση του X^2 , του μεγέθους του δείγματος και ανεξάρτητος από την ταξινόμηση των τιμών των μεταβλητών και το πλήθος των τιμών. Αν οι μεταβλητές είναι ανεξάρτητες, τότε το ϕ τείνει στο μηδέν.
- **Cramer's V** = η τιμή του συντελεστή Cramer's V αναφέρεται (επίσης) στην ισχύ της σχέσης ανάμεσα στις δύο (ονομαστικές) μεταβλητές. Εμπεριέχει τον

συντελεστή ϕ και η διαφορά ως προς τον ϕ έγκειται στο γεγονός ότι η ϕ μπορεί να πάρει τιμές >1 ενώ το εύρος τιμών του Cramer's V είναι ανάμεσα στο 0 και το 1.

Τέλος να σημειώσουμε ότι η δημοσίευση των αποτελεσμάτων σύμφωνα με το APA Style θα μπορούσε να έχει ως εξής:

$$\chi^2(5, N=200) = 23.159, p < 0.05^{26}$$

²⁶ American Psychological Association. (2001). *Publication manual of the American Psychological Association* (5th ed.). Washington, DC: American Psychological Association, p.113

ΠΑΡΑΡΤΗΜΑΤΑ

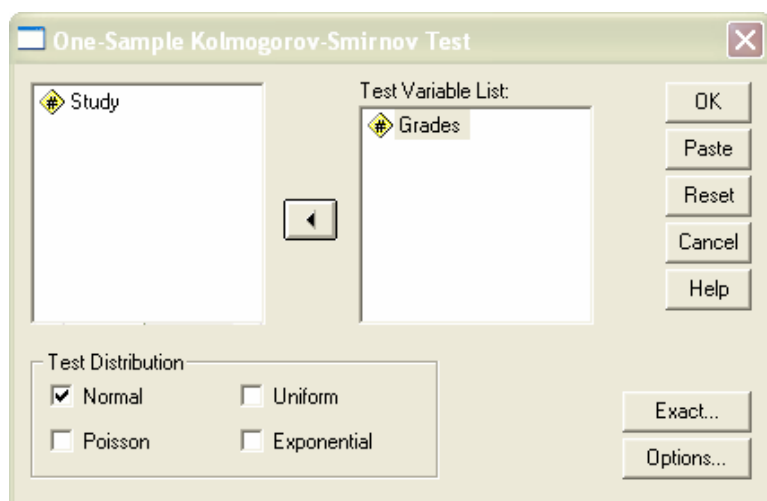
A: Έλεγχος κανονικότητας

Συνήθως η στατιστική σημαντικότητα είναι ένα επιθυμητό αποτέλεσμα μιας και οι ερευνητές προσπαθούν να επαληθεύσουν εναλλακτικές υποθέσεις (δεν έχει νόημα η επαλήθευση μηδενικών υποθέσεων—δεν ψάχνει ο ερευνητής τι ΔΕΝ ισχύει αλλά τι ισχύει). Εξαιρέση αποτελούν τα περισσότερα τεστ προϋποθέσεων (όχι όλα) για τα οποία επιθυμητό αποτέλεσμα είναι η υιοθέτηση της μηδενικής υπόθεσης. Αυτό φαίνεται ξεκάθαρα μόλις κάποιος διατυπώσει τις 2 υποθέσεις.

H_0 : Η κατανομή ΔΕΝ είναι στατιστικά σημαντικά διαφορετική από την κανονική

$H_ε$: Η κατανομή είναι στατιστικά σημαντικά διαφορετική από την κανονική

Από τις παραπάνω υποθέσεις γίνεται φανερό ότι για να ικανοποιηθεί η προϋπόθεση της κανονικής κατανομής πρέπει να ισχύει η μηδενική υπόθεση. Για την διερεύνηση της προϋπόθεσης στη μεταβλητή επίδοση (grades) πηγαίνουμε στο μενού του SPSS και επιλέγουμε: Statistics → NonParametric tests → 1-Sample K-S.



Σχήμα 29: Εντολές του SPSS για το τεστ Kolmogorov-Smirnov

Με επιλογή της κανονικής κατανομής (Normal) ως υποθετική κατανομή, βλέπουμε ότι σχετικά με τη μεταβλητή επίδοση, η πιθανότητα να κατανέμεται αυτό το

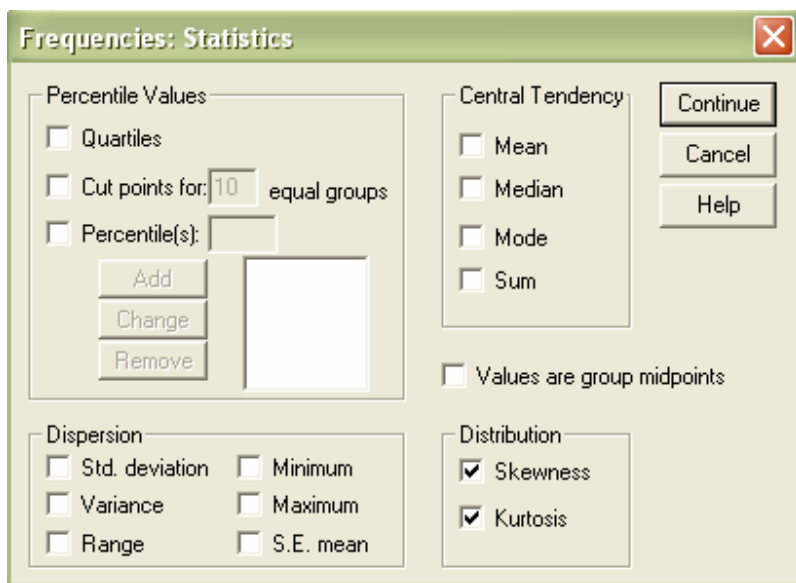
φαινόμενο κανονικά είναι 97%. Επομένως γίνεται δεκτή η μηδενική υπόθεση που επιβεβαιώνει την ύπαρξη της προϋπόθεσης της κανονικότητας.

		Grades	
N		10	Μέγεθος Δείγματος
Normal Parameters ^{a,b}	Mean	14.4000	Μέσος όρος
	Std. Deviation	4.45222	Τυπική Απόκλιση
Most Extreme Differences	Absolute	.154	
	Positive	.104	
	Negative	-.154	
Kolmogorov-Smirnov Z		.486	Μέγεθος Τεστ K-S
Asymp. Sig. (2-tailed)		.972	Πιθανότητα

a. Test distribution is Normal.
b. Calculated from data.

Σχήμα 30: Αποτελέσματα του SPSS για το τεστ Kolmogorov-Smirnov

Ένας εναλλακτικός τρόπος αξιολόγησης της κανονικότητας είναι με τη χρήση της περιγραφικής στατιστικής και των δεικτών κυρτότητας και στρέβλωσης. Όπως πρότεινε και ο Δαφέρμος (2005), ο λόγος λ της κύρτωσης/στρέβλωσης προς το στατιστικό τους σφάλμα εκφράζει την πιθανότητα να κατανέμεται κανονικά το φαινόμενο που μελετάμε. Οι δείκτες αυτοί υπολογίζονται στο μενού συχνότητων. Έτσι επιλέγουμε: Statistics → Descriptive Statistics → Frequencies → Βάζετε στο παράθυρο “Variables” τη/τις μεταβλητή(ες) που θέλετε να αξιολογήσετε και κάντε κλικ στο “statistics”.



Σχήμα 31: Εντολές του SPSS για εναλλακτικό τρόπο

Αφού επιλέξετε Skewness/Kurtosis, επιλέξτε «continue» → OK. Τα αποτελέσματα φαίνονται παρακάτω:

Statistics		
Grades		
N	Valid	10
	Missing	0
Skewness		-.690
Std. Error of Skewness		.687
Kurtosis		-.185
Std. Error of Kurtosis		1.334

Σχήμα 32: Αποτελέσματα του SPSS για εναλλακτικό τρόπο

Όταν ο λόγος λ ξεπερνά το ± 2 τότε υπάρχει ένδειξη ότι μια δεδομένη κατανομή αποκλίνει από την κανονικότητα είτε σε επίπεδο στρέβλωσης είτε σε επίπεδο κύρτωσης.

Σχετικά με τα παραπάνω δεδομένα, ο λόγος για την στρέβλωση ήταν $\lambda = \frac{Skew}{Skew_{Error}}$ είναι

$-.690/.687 = -1.004$, ενώ για την κύρτωση $-.185/1.334 = -.1386$. Μιας και κανένας λόγος λ δεν είναι μεγαλύτερος από ± 2 τότε δεχόμαστε την υπόθεση ότι η μεταβλητή «επίδοση» κατανέμεται κανονικά και επομένως ικανοποιείται αυτή η προϋπόθεση.

Αξιολόγηση της κανονικότητας μπορεί να γίνει και για υπο-ομάδες με το μενού «Explore».

B. Επιλογή κριτηρίου ελέγχου

1. Σύγκριση της ίδιας μεταβλητής σε δύο δείγματα

- 1.1 Όταν η μεταβλητή είναι **κατηγορική** και τα δείγματα είναι **ανεξάρτητα** **Test X^2 ομοιογένειας**.

Προϋποθέσεις:

- (α) ο έλεγχος να γίνεται σε δείγματα μεγέθους >30 ,
- (β) οι θεωρητικές τιμές που υπολογίζονται στον πίνακα συνάφειας να είναι > 1 και
- (γ) το 80% να είναι $>$ του 5

- 1.2 Όταν η μεταβλητή είναι **κατηγορική** και τα δείγματα είναι **εξαρτημένα** **Test Mc Nemar**.

Το test ελέγχει αν είναι στατιστικά σημαντικές οι αλλαγές που επήλθαν από την επίδραση κάποιου εξωγενούς παράγοντα.

Προϋποθέσεις:

- (α) οι μεταβλητές να είναι δίτιμες (dichotomous)
- (b) οι μεταβλητές να έχουν τις ίδιες τιμές (value labels)

- 1.3 Όταν η μεταβλητή είναι **διατακτική** και τα δείγματα είναι **ανεξάρτητα**. **Test Mann & Whitney**

Το test ελέγχει εάν οι μέσες τιμές των δύο πληθυσμών μπορούν να θεωρηθούν ίσες συγκρίνοντας τις διαμέσους των δύο δειγμάτων

- 1.4 Όταν η μεταβλητή είναι **διατακτική** και τα δείγματα είναι **εξαρτημένα**.

Test προσήμου (**Sign test**) ή **test Wilcoxon**. Το Sign test προτείνεται για τη διερεύνηση του ποιοτικού μέρους (ποιο δείγμα υπερέχει) ενώ το Wilcoxon test μετρά πόσο υπερέχει. Η βιβλιογραφία εκτιμά ότι το Sign test είναι αυστηρότερο του test Wilcoxon.

- 1.5 Όταν η μεταβλητή είναι **ποσοτική** και τα δείγματα είναι **ανεξάρτητα**.

T – test

Προϋπόθεση:

(α) τα δύο δείγματα θα πρέπει να έχουν κανονικές κατανομές. Αυτό σημαίνει ότι θα πρέπει να προηγηθεί έλεγχος κανονικότητας. Στην περίπτωση που οι κατανομές δεν ακολουθούν κανονική κατανομή προτείνεται η χρήση του **Mann – Whitney test**. Ανάλογα με το οι διασπορές των δύο δειγμάτων είναι ίσες ή όχι επιλέγεται και ειδική εκδοχή του t-test. Ο έλεγχος ισότητας των διασπορών γίνεται με χρήση της F – κατανομής (**F-test**).

- 1.6 Όταν η μεταβλητή είναι **ποσοτική** και τα δείγματα είναι **εξαρτημένα**.

T – test ζευγαρωτών παρατηρήσεων

Προϋπόθεση:

(α) τα δύο δείγματα θα πρέπει να έχουν κανονικές κατανομές. Αυτό σημαίνει ότι θα πρέπει να προηγηθεί έλεγχος κανονικότητας. Στην περίπτωση που οι κατανομές δεν ακολουθούν κανονική κατανομή προτείνεται η χρήση του **test Wilcoxon**.

2. Σύγκριση της **ίδιας** μεταβλητής σε **περισσότερα από δύο** δείγματα.

- 2.1 Όταν η μεταβλητή είναι **κατηγορική** και τα δείγματα είναι **ανεξάρτητα**.

Πίνακας Συνάφειας. Ο πίνακας έχει n γραμμές (όπου n = ο αριθμός των δειγμάτων) και k στήλες (όπου k = οι τιμές της μεταβλητής). Μπορεί να ισχύει και το αντίστροφο.

X^2 test ομοιογένειας. Προσοχή το X^2 εφαρμόζεται στις συχνότητες και όχι στα ποσοστά.

Προϋπόθεση:

(α) σε όλα τα κελιά του πίνακα συνάφειας να έχουμε θεωρητικές συχνότητες >1 και για το 80% >5 .

- 2.2 Όταν η μεταβλητή είναι **κατηγορική** και τα δείγματα είναι **εξαρτημένα**.

Πίνακας Συνάφειας. Ο πίνακας έχει n γραμμές (όπου n = ο αριθμός των περιπτώσεων) και k στήλες (όπου k = ο αριθμός των δειγμάτων).

Προϋποθέσεις:

(α) οι μεταβλητές να είναι δίτιμες (dichotomous)

Test Cochran (Q – test).

- 2.3 Όταν η μεταβλητή είναι **διατακτική** και τα δείγματα είναι **ανεξάρτητα**.
Kruskal & Wallis test (μη παραμετρικό test). Βασική παραδοχή αποτελεί η υπόθεση ότι οι διάμεσοι των πληθυσμών είναι ίσες.
- 2.4 Όταν η μεταβλητή είναι **διατακτική** και τα δείγματα είναι **εξαρτημένα**.
Friedman test. Βασική παραδοχή αποτελεί η υπόθεση ότι οι διάμεσοι των πληθυσμών είναι ίσες.
- 2.5 Όταν η μεταβλητή είναι **συνεχής** και τα δείγματα είναι **ανεξάρτητα**.
Ανάλυση Διασποράς **ANOVA**.
Προϋποθέσεις:
(α) η κατανομή των παρατηρήσεων θα πρέπει να ακολουθεί την κανονική κατανομή,
(β) η μεταβλητή να παρουσιάζει την ίδια διασπορά σε όλους τους πληθυσμούς.
Προσοχή στον υπολογισμό και την εκτίμηση των υπολοίπων.
- 2.6 Όταν η μεταβλητή είναι **συνεχής** και τα δείγματα είναι **εξαρτημένα**.
Σε γενικές γραμμές υιοθετείται η Ανάλυση Διασποράς.
3. Σύγκριση δύο μεταβλητών στο ίδιο δείγμα
- 3.1 Όταν η **εξαρτημένη μεταβλητή** είναι **συνεχής** και η **ανεξάρτητη** επίσης **συνεχής**.
Γραμμική παλινδρόμηση και υπολογισμός του **συντελεστή γραμμικής συσχέτισης Pearson**.
Προϋπόθεση:
Η κατανομή των παρατηρήσεων θα πρέπει να ακολουθεί την κανονική κατανομή,
- 3.2 Όταν η **εξαρτημένη μεταβλητή** είναι **συνεχής** και η **ανεξάρτητη** είναι **διατακτική**.
Μη παραμετρική παλινδρόμηση και υπολογισμός του **συντελεστή Spearman**.
- 3.3 Όταν η **εξαρτημένη μεταβλητή** είναι **συνεχής** και η **ανεξάρτητη** είναι **κατηγορική**.

Ανάλυση Διασποράς με έναν παράγοντα και υπολογισμός του συντελεστή η^2 .

Παρατήρηση: ανάλογα με την κατανομή της εξαρτημένης μεταβλητής χρησιμοποιείται παραμετρικό ή μη-παραμετρικό test.

- 3.4 Όταν η εξαρτημένη μεταβλητή είναι διατακτική και η ανεξάρτητη είναι συνεχής.

Μη παραμετρική παλινδρόμηση και υπολογισμός του συντελεστή Spearman.

- 3.5 Όταν η εξαρτημένη μεταβλητή είναι διατακτική και η ανεξάρτητη είναι επίσης διατακτική.

Μη παραμετρική παλινδρόμηση και υπολογισμός του συντελεστή Spearman.

- 3.6 Όταν η εξαρτημένη μεταβλητή είναι διατακτική και η ανεξάρτητη είναι κατηγορική.

χ^2 test.

- 3.7 Όταν η εξαρτημένη μεταβλητή είναι κατηγορική και η ανεξάρτητη είναι συνεχής.

χ^2 test.

- 3.8 Όταν η εξαρτημένη μεταβλητή είναι κατηγορική και η ανεξάρτητη είναι διατακτική.

χ^2 test.

- 3.9 Όταν η εξαρτημένη μεταβλητή είναι κατηγορική και η ανεξάρτητη είναι επίσης κατηγορική.

χ^2 test.

ΒΙΒΛΙΟΓΡΑΦΙΑ-ΔΙΑΔΙΚΤΥΟ

- Stevens, S. (1946) On the theory of scales and measurement. *Science*, 103, 667-680
- Μιχαλοπούλου, Κ. (1992). *Κλίμακες Μέτρησης Στάσεων*. Αθήνα: Οδυσσέας
- Τσάντας, Ν., Μουσιάδης, Χ., Μπαγιάτης, Ν., & Χατζηπαντελής, Θ. (1999). *Ανάλυση δεδομένων με τη βοήθεια στατιστικών πακέτων*. Θεσσαλονίκη: Ζήτη.
- Norusis, M. (2005). *Οδηγός Ανάλυσης Δεδομένων με το SPSS 12.0* (Κ.Περάκης, Trans.). Αθήνα: Κλειδάριθμος.
- Yerkes, R. M., & Dowdson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative and Neurological Psychology*, 18, 459-482.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- Ρούσσο, Π., & Τσαούσης, Ι. (2003). *Στατιστική εφαρμοσμένη στις κοινωνικές επιστήμες*. Αθήνα: Ελληνικά Γράμματα

Χρήσιμες διευθύνσεις στο Διαδίκτυο (εκτός αυτών που αναφέρονται στο κείμενο)

1. <http://www.ats.ucla.edu/stat/spss/>
(ιδιαίτερα τους συνδέσμους, [SPSS Starter Kit](#), [What statistical analysis should I use?](#), [Annotated Output](#) και την κατηγορία Links by Topic). Επίσης ο σχετικός σύνδεσμος <http://www.ats.ucla.edu/stat/spss/notes2/analyze.htm> προσφέρει πολύ καλή ανάλυση για τις εντολές που πρέπει να ακολουθούνται για κάθε τύπο ανάλυσης.
2. T-TEST
http://www.une.edu.au/WebStat/unit_materials/c6_common_statistical_tests/index.html
3. ANOVA
<http://www.physics.csbsju.edu/stats/anova.html>
<http://web.umn.edu/~psyworld/virtualstat/anova/anovacalc.html>
<http://faculty.vassar.edu/lowry/ank3.html>
<http://faculty.vassar.edu/lowry/ank4.html>
http://www.une.edu.au/WebStat/unit_materials/c7_anova/index.html
<http://www.graphpad.com/quickcalcs/posttest1.cfm>
<http://home.ubalt.edu/ntsbarsh/Business-stat/otherapplets/ANOVA2Rep.htm>
<http://faculty.vassar.edu/lowry/corr3.html>
<http://home.ubalt.edu/ntsbarsh/Business-stat/otherapplets/ANOVA2Dep.htm>
<http://faculty.vassar.edu/lowry/anova2x2.html>

- <http://faculty.vassar.edu/lowry/anova2x3.html>
<http://home.ubalt.edu/ntsbarsh/Business-stat/otherapplets/ANOVATwo.htm>
<http://faculty.vassar.edu/lowry/corr4>.
http://www.une.edu.au/WebStat/unit_materials/c7_anova/twoway_anova.htm
http://www.une.edu.au/WebStat/unit_materials/c7_anova/twoway_anova.htm
4. <http://www.allenandunwin.com/spss/exercises5.htm> Ερωτήσεις αλλά και απαντήσεις στο σύνδεσμο στο τέλος της σελίδας σχετικά με τις διάφορες στατιστικές τεχνικές.