

Τραυλός, Α. Κ. (2008). Κλινική στατιστική. Στους Γ. Ι. Μπαλτόπουλος & Π. Λ. Ευαγγελοπούλου (Επιμ.), *11^ο Θεματικό Συνέδριο - Εντατική θεραπεία και επείγουσα ιατρική: Κλινικές περιπτώσεις και σπάνια νοσήματα* (σελ. 153-172). Αθήνα, Αττικής: Π. Χ. Πασχαλίδης.

ΠΕΡΙΛΗΨΗ

Η ερευνητική δραστηριότητα στο χώρο της υγείας απαιτεί σωστό σχεδιασμό, έγκυρη στατιστική ανάλυση και τεκμηριωμένη εξαγωγή συμπερασμάτων. Πληρώνοντας τις προϋποθέσεις της ερευνητικής διαδικασίας και της αξιολόγησης, οι επαγγελματίες υγείας θα διαπιστώσουν, ότι οι στατιστικές έννοιες και δεξιότητες είναι ουσιαστικοί παράγοντες για την επικοινωνία και τη λήψη των αποφάσεων στην επιστημονική έρευνα. Κύριος σκοπός της εργασίας είναι να παρουσιαστούν μη-παραμετρικές και παραμετρικές στατιστικές μέθοδοι επεξεργασίας των δεδομένων, καθώς και ο τρόπος παρουσίασης των αποτελεσμάτων για την εποικοδομητικότερη επικοινωνία των επαγγελματιών υγείας. Στο άρθρο παρουσιάζονται (α) ο έλεγχος χ^2 και ο συντελεστής συσχέτισης Spearman για την μη-παραμετρική στατιστική και (β) ο συντελεστής συσχέτισης Pearson, ο έλεγχος t για εξαρτημένα και ανεξάρτητα δείγματα, η απλή ανάλυση διασποράς (ANOVA) για εξαρτημένα και ανεξάρτητα δείγματα, και η παραγοντική ανάλυση (factor analysis). Ο τρόπος επεξήγησης των στατιστικών μεθόδων και των αποτελεσμάτων των στατιστικών ανάλύσεων γίνεται σύμφωνα με τις επιστημονικές και τις απαιτήσεις της σύγχρονης επιστημονικής βιβλιογραφίας.

Λέξεις κλειδιά: ερευνητική διαδικασία, αξιολόγηση, συντελεστής συσχέτισης, έλεγχος χ^2 , έλεγχος -t, απλή ανάλυση διασποράς, παραγοντική ανάλυση.

ΕΙΣΑΓΩΓΗ

Ο όρος *Στατιστική* δημιουργεί τις περισσότερες φορές ανησυχία και προβληματισμό σε φοιτητές και σε νέους επιστήμονες που δεν έδωσαν την απαιτούμενη προσοχή κατά την περίοδο διδασκαλίας του αντίστοιχου μαθήματος. Υπάρχει η "ψευδής εντύπωση" ότι η στατιστική είναι ένα μάθημα με προχωρημένα μαθηματικά. Αποτελεί όμως κοινή διαπίστωση, ότι η στατιστική βασίζεται σε απλές αριθμητικές πράξεις, όπως πρόσθεση, αφαίρεση, πολλαπλασιασμός και διαίρεση. Η κατανόηση των στατιστικών όρων και αλγορίθμων, η οργανωμένη μελέτη και εξάσκηση στις στατιστικές μεθόδους, η λειτουργική αντιπαράθεση των ερευνητικών ερωτήσεων με την υιοθέτηση της κατάλληλης στατιστικής ανάλυσης, η συλλογή και η επεξεργασία των δεδομένων, καθώς και η επεξήγηση των αποτελεσμάτων της στατιστικής ανάλυσης αποτελούν μονόδρομο στη διαδικασία της μάθησης για τον απαιτητικό ερευνητή-επιστήμονα. Οι νέοι ερευνητές στο χώρο της υγείας γνωρίζουν ότι χωρίς στατιστικές γνώσεις και δεξιότητες δεν θα είναι σε θέση να καταξιωθούν ως επαγγελματίες και επιστήμονες στο χώρο της αρμοδιότητάς τους. Στο ελάχιστο, οι επιστήμονες υγείας, ιατροί και νοσηλευτές χρειάζονται την οικειότητα με τη "λογική" των στατιστικών για τις ερευνητικές και κλινικές δραστηριότητές τους για να είναι σε θέση να επικοινωνούν αποτελεσματικά με άλλους επιστήμονες του χώρου τους.

Σκοπός της παρούσας εργασίας είναι να αναπτυχθεί περιληπτικά η σημαντικότητα της ερευνητικής διαδικασίας και της αξιολόγησης και να παρουσιαστούν οι σημαντικότερες στατιστικές μέθοδοι επεξεργασίας δεδομένων μη-παραμετρικής (απαραμετρικής) και παραμετρικής στατιστικής, κα-

* Το άρθρο έχει γραφτεί σύμφωνα με το Διεθνές πρότυπο σύνταξης άρθρων του American Psychological Association (2003).

θώς και ο τρόπος επεξήγησης των αποτελεσμάτων των στατιστικών αναλύσεων όπως παρουσιάζονται στην επιστημονική βιβλιογραφία. Στη μη-παραμετρική επαγωγική στατιστική αναπτύσσονται ο έλεγχος χ^2 και ο συντελεστής συσχέτισης Spearman. Στην παραμετρική επαγωγική στατιστική παρουσιάζονται ο συντελεστής συσχέτισης Pearson, ο έλεγχος -t για εξαρτημένα και ανεξάρτητα δείγματα, η απλή ανάλυση διασποράς (ANOVA) για εξαρτημένα και ανεξάρτητα δείγματα και η παραγοντική ανάλυση (factor analysis).

ΕΡΕΥΝΗΤΙΚΗ ΔΙΑΔΙΚΑΣΙΑ ΚΑΙ ΑΞΙΟΛΟΓΗΣΗ

Στη μελέτη της στατιστικής δεν πρέπει να παραλειφθεί η σημαντικότητα της σωστής μεθοδολογικής προσέγγισης. Ακολουθώντας την ερευνητική διαδικασία θα διαπιστώσουμε, ότι οι στατιστικές έννοιες και δεξιότητες είναι ουσιαστικοί παράγοντες για την επικοινωνία και τη λήψη των αποφάσεων στην επιστημονική έρευνα. Η ακόλουθη επισκόπηση των διαδικασιών έρευνας και αξιολόγησης επεξηγεί το ρόλο των στατιστικών μεθόδων και τεχνικών.

Ερευνητική διαδικασία

Όπως συμβαίνει σε οποιαδήποτε επιστημονική προσπάθεια, η έρευνα στις βιολογικές επιστήμες ή στις επιστήμες υγείας στοχεύει να **περιγράψει**, να **εξηγήσει**, να **προβλέψει**, ή/και να **καταλάβει** τα φαινόμενα. Μέρος αυτής της διαδικασίας περιλαμβάνει την απόκτηση της γνώσης, καθώς και, την ανάπτυξη και την εφαρμογή των θεωριών που μας επιτρέπουν να καταλάβουμε τις χαρακτηριστικά σύνθετες σχέσεις μεταξύ των φαινομένων. Σήμερα, το μεγαλύτερο μέρος της έρευνας ακολουθεί την επιστημονική μέθοδο, χρησιμοποιώντας μια συστηματική, ελεγχόμενη ερευνητική διαδικασία, που παράγει τα αντικειμενικά συμπεράσματα σχετικά με το πρόβλημα. Σύμφωνα με τους Glasnapp & Roggio (1985), τα βήματα σε αυτήν τη διαδικασία είναι:

Βήμα 1: Προσδιορίζουμε το πρόβλημα.

Βήμα 2: Διευκρινίζουμε το πρόβλημα με τον προσδιορισμό των ερευνητικών υποθέσεων ή ερωτήσεων.

Βήμα 3: Αναπτύσσουμε το ερευνητικό σχέδιο για τη συλλογή των δεδομένων.

Βήμα 4: Αναλύουμε τα δεδομένα με την εφαρμογή της στατιστικής.

Βήμα 5: Βγάζουμε τα συμπεράσματα/τις ερμηνείες από τα αποτελέσματα.

Τα δύο πρώτα βήματα είναι συχνά τα δυσκολότερα στη ερευνητική διαδικασία. Η επάρκεια των διαδικασιών που συνεπάγονται στο Βήμα 3 καθορίζει την ισχύ των συμπερασμάτων ενός ερευνητή. Το ερευνητικό σχέδιο και οι διαδικασίες συλλογής δεδομένων περιλαμβάνουν τον προσδιορισμό (α) των ατόμων - συμμετεχόντων (δειγματοληψία) που θα επιτρέψει τη σωστή ερμηνεία των αποτελεσμάτων και τη γενίκευση των αποτελεσμάτων, (β) των οργάνων συλλογής δεδομένων που μετρούν αυτό ακριβώς που ισχυρίζονται ότι μετρούν (validity - εγκυρότητα) με συνέπεια και ακρίβεια (reliability - αξιοπιστία), (γ) των διαδικασιών συλλογής δεδομένων που δεν αποκρύπτουν την έννοια των δεδομένων, και (δ) ενός γενικού ερευνητικού σχεδίου που αναδεικνύει τα κατάλληλα δεδομένα για την απάντηση των ερευνητικών ερωτήσεων ή υποθέσεων που προσδιορίστηκαν στο Βήμα 2, συμπεριλαμβανομένου του προσδιορισμού της κατάλληλης στατιστικής ή της στατιστικής τεχνικής που θα χρησιμοποιηθεί για την ανάλυση των δεδομένων.

Χωρίς κατάλληλη δειγματοληψία, οι τεχνικές μέτρησης, το ερευνητικό σχέδιο, και οι διαδικασίες συλλογής δεδομένων, καθώς και η αντιπροσωπευτικότητα και η δυνατότητα γενίκευσης των αποτελεσμάτων είναι παράγοντες αμφισβητήσιμοι. Συχνά αγνοούνται ή υποβαθμίζονται οι τεχνικές μέτρησης που υιοθετούνται. Είναι μεγάλο λάθος να θεωρηθεί ότι οι στατιστικές μέθοδοι παρέχουν τη θεραπεία για οποιοδήποτε ανεπάρκειες στις διαδικασίες στο βήμα 3. Τίποτα δεν μπορεί να είναι παραπάνω από την αλήθεια! Η συλλογή δεδομένων με λανθασμένες διαδικασίες "γεννούν" άκυρα αποτελέσματα και συμπεράσματα, ανεξάρτητα από το πόσο περίπλοκη είναι η χρησιμοποιούμενη στατιστική τεχνική. Το σχέδιο συλλογής δεδομένων πρέπει να προβληματίζει τον ερευνητή σε οποιαδήποτε ερευνητική προσπάθεια. Οι διαδικασίες για τα βήματα 4 και 5 εξαρτώνται α-

πλώς από τις εκβάσεις των βημάτων 2 και 3. Οι στατιστικές διαδικασίες που υιοθετούνται πρέπει να είναι συμβατές με τις ερευνητικές υποθέσεις ή τις ερωτήσεις και το ερευνητικό σχέδιο. Είναι εμφανές ότι η εφαρμογή των κατάλληλων στατιστικών τεχνικών είναι ένα απαραίτητο συστατικό της ερευνητικής διαδικασίας, αλλιώς όχι και ο μόνος παράγοντας που αναφέρεται στην εγκυρότητα των αποτελεσμάτων.

Στο τελικό βήμα, τα συμπεράσματα συνάγονται βασισμένα στην ερμηνεία των δεδομένων που παρέχονται από τους στατιστικούς δείκτες και τα στατιστικά τεστ. Πάρα πολύ συχνά, τα στατιστικά αποτελέσματα ερμηνεύονται μόνο σε ένα επίπεδο "παπαγαλίας". Είναι σημαντικό να αναφερθεί ότι οι αριθμοί προήλθαν από τα στοιχεία και ότι αυτά τα στοιχεία αντιπροσωπεύουν τις μετρήσεις συγκεκριμένων μεταβλητών (δηλ., γνώρισμα, ιδιότητες, κατασκευάσμα ή χαρακτηριστικό). Ο σημαντικός στόχος είναι η ερμηνεία των στατιστικών αποτελεσμάτων όπως αφορούν τις ερευνητικές ερωτήσεις και τις υποθέσεις από την άποψη των πραγματικών μεταβλητών που εξετάστηκαν (Glasnapp & Poggio, 1985).

Διαδικασία αξιολόγησης

Οι ερευνητές διαφοροποιούνται συνήθως μεταξύ της αξιολόγησης και της έρευνας. Οι διακρίσεις μπορούν να γίνουν μεταξύ των στόχων, των αρχικών σταδίων, και του σχεδίου για τη συλλογή δεδομένων και του τύπου και της χρήσης των δεδομένων που συγκεντρώνονται (Glasnapp & Poggio, 1985; Lockhart, 1998). Η στατιστική γνώση και δεξιότητα είναι τόσο σημαντική στην αξιολόγηση των αποτελεσμάτων (διαδικασία αξιολόγησης) όσο και στη διεξαγωγή της έρευνας (ερευνητική διαδικασία). Οι ίδιες στατιστικές μέθοδοι χρησιμεύουν ως σημαντικά εργαλεία και στις δύο διαδικασίες.

Ο σκοπός της αξιολόγησης αναφέρεται στην άμεση χρησιμότητα των αποτελεσμάτων για τη λήψη των αποφάσεων σε μια συγκεκριμένη πρακτική ή παρεμβατική διαδικασία. Η αξιολόγηση είναι μια διαδικασία που παρέχει τις πληροφορίες σχετικά με την επάρκεια ή την αποτελεσματικότητα της πρακτικής ή της παρεμβατικής διαδικασίας για να επιτευχθούν οι στόχοι της έρευνας. Στην αξιολόγηση, οι συστηματικά συλλεγόμενες πληροφορίες καθοδηγούν τις αποφάσεις, σχετικά με την αποτελεσματικότητα και τη χρησιμότητα της πρακτικής ή της παρεμβατικής διαδικασίας. Τα βήματα στη διαδικασία αξιολόγησης περιλαμβάνουν τα παρακάτω χαρακτηριστικά: Πρώτον, αξιολογώ τις ανάγκες, διευκρινίζω τους στόχους και προσδιορίζω τους αντικειμενικούς σκοπούς της πρακτικής ή της παρεμβατικής διαδικασίας. Δεύτερον, επιλέγω, αναπτύσσω, και εκτελώ την πρακτική ή την παρεμβατική διαδικασία. Τρίτον, εφαρμόζω το σχέδιο αξιολόγησης και συλλέγω τα δεδομένα. Τέταρτον, αναλύω τα δεδομένα εφαρμόζοντας τις στατιστικές τεχνικές. Τέλος, ερμηνεύω τα ευρήματα και λαμβάνω τις αποφάσεις σχετικά με την αποτελεσματικότητα της πρακτικής ή της παρεμβατικής διαδικασίας (Glasnapp & Poggio, 1985; Heiman, 2001).

Ενώ η ερευνητική διαδικασία συνδέεται με τη συναγωγή των συμπερασμάτων για τις υποθέσεις ή τις ερωτήσεις που αφορούν τις συγκεκριμένες θεωρίες, η διαδικασία αξιολόγησης οδηγεί συνήθως σε μια ερμηνεία των στατιστικών δεικτών εξετάζοντας την επίτευξη των συγκεκριμένων στόχων της πρακτικής ή της παρεμβατικής διαδικασίας. Οι δραστηριότητες αξιολόγησης είναι συχνά ποικίλες, κατευθύνοντας την προσοχή και στη διαδικασία και στο αποτέλεσμα. Ακόμα, τα βήματα 3 και 4 στην ερευνητική διαδικασία πρέπει να ακολουθηθούν και στην αξιολόγηση. Πάλι, προσοχή πρέπει να δοθεί στο ερευνητικό - πειραματικό σχέδιο για τη συλλογή των δεδομένων, τον προσδιορισμό της μέτρησης των μεταβλητών και την πραγματική συλλογή των δεδομένων. Η επάρκεια αυτών των διαδικασιών καθορίζει την ισχύ των αποτελεσμάτων αξιολόγησης. Η στατιστική ανάλυση άπλη μια φορά χρησιμεύει ως ένα "μοναδικό και απαραίτητο" εργαλείο επιστημονικής και αντικειμενικής τεκμηρίωσης. Τα αποτελέσματα είναι τόσο καλά, όσο τα δεδομένα πάνω στα οποία βασίζονται (Glasnapp & Poggio, 1985).

ΑΤΟΜΙΚΕΣ ΔΙΑΦΟΡΕΣ ΚΑΙ ΜΕΤΑΒΛΗΤΕΣ

Οι **ατομικές διαφορές** (individual differences) "ευθύνονται" για κάθε συμπεριφορά ή χαρακτηριστικό των ατόμων που δεν είναι το ίδιο για όλα τα άτομα που συμμετέχουν σε μία ερευνητική δια-

δικασία (ή έχουν τις ίδιες πιθανότητες να συμπεριληφθούν στην έρευνα). Τα χαρακτηριστικά ή οι συμπεριφορές που διαφέρουν οι άνθρωποι μεταξύ τους ονομάζονται **μεταβλητές** (variables). Τα χαρακτηριστικά ή οι συμπεριφορές που δεν διαφέρουν οι άνθρωποι μεταξύ τους ονομάζονται **σταθερές** (constants).

Οι **Ανεξάρτητες Μεταβλητές** (Independent Variables) -ΑΜ είναι μεταβλητές που ο ερευνητής ελέγχει ή χειρίζεται σύμφωνα με το σκοπό της έρευνας. Οι ΑΜ μπορεί να είναι μεταβλητές χειρισμού (-ελέγχου) ή μεταβλητές ταξινόμησης. Παραδείγματος χάριν, σε μια μελέτη για να καθοριστεί η επίδραση του επιπέδου της δόσης φαρμάκων στην απόδοση κάποιας δεξιότητας, ο ερευνητής μπορεί να χειριστεί το επίπεδο δόσης. Κατά συνέπεια η δόση είναι μια μεταβλητή χειρισμού στη μελέτη. Εάν η ανεξάρτητη μεταβλητή είναι μια μεταβλητή ταξινόμησης, ταξινομεί απλά τα άτομα που συμμετέχουν στη μελέτη. Παραδείγματος χάριν, εάν ο ερευνητής στην υποθετική μελέτη φαρμάκων μας ταξινομεί τα άτομα κατά φύλο ή κατά ηλικία και συγκρίνει έπειτα την απόδοση αυτών των ομάδων μετά από τη χορήγηση της δόσης, τότε το φύλο ή η ηλικία θα είναι η μεταβλητή ταξινόμησης στη μελέτη.

Η **Εξαρτημένη Μεταβλητή** (Dependent Variable) - ΕΜ είναι ένα μέτρο της επίδρασης της ανεξάρτητης μεταβλητής. Για παράδειγμα, η απόδοση στη αξιοθιγούμενη δεξιότητα είναι η εξαρτημένη μεταβλητή. Δεδομένου ότι η ανεξάρτητη μεταβλητή αλληλάζει ή μεταβάλλεται (λόγω του χειρισμού της από τον ερευνητή), ο ερευνητής παρατηρεί τις αλληλαγές στην εξαρτημένη μεταβλητή προκειμένου να καθορίσει πως συνδέεται με τις αλληλαγές στην ανεξάρτητη μεταβλητή.

ΚΛΙΜΑΚΕΣ ΜΕΤΡΗΣΗΣ

Όλες οι μετρήσεις δεν είναι ίδιες. Μερικές μετρήσεις είναι ακριβέστερες από άλλες. Εάν λέμε ότι ένα άτομο είναι ψηλότερο, ή ότι ο Α είναι ψηλότερος από τον Β, αυτό δεν είναι τόσο ακριβές όπως λέγοντας ότι το άτομο Α έχει ύψος 1 μέτρο και 80 εκατοστά και το άτομο Β, 1 μέτρο και 78 εκατοστά. Η ακρίβεια της μέτρησης μιας μεταβλητής είναι σημαντική στον καθορισμό της στατιστικής μεθόδου που θα χρησιμοποιήσουμε για τη στατιστική ανάλυση. Η ιεράρχηση των κλιμάκων μέτρησης περιλαμβάνει την ονομαστική κλίμακα (nominal scale), την τακτική κλίμακα (ordinal scale), την κλίμακα διαστήματος (interval scale) και την κλίμακα λόγου ή αναλογίας (ratio scale). Για αναλυτική περιγραφή των κλιμάκων μέτρησης ο αναγνώστης μπορεί να συμβουλευτεί τους Heiman (2001) και Pagano & Gauvreau, (2002). Τα δεδομένα των κλιμάκων διαστήματος και λόγου αναλύονται με παραμετρικές στατιστικές μεθόδους (π.χ. συντελεστή συσχέτισης Pearson, έλεγχος t, ανάλυση διασποράς, ανάλυση συνδιακύμανσης κτλ), της τακτικής κλίμακας για εξαρτημένα δείγματα με Wilcoxon T και συντελεστή συσχέτισης Spearman και για ανεξάρτητα δείγματα με Mann-Whitney U, και της ονομαστικής κλίμακας για εξαρτημένα δείγματα με διωνυμικό έλεγχο προσήμου (binomial sign test) και για συχνότητες ανεξάρτητων δειγμάτων με χ^2 (chi square) (βλ. Πίνακα 1) (Coolican, 1996).

ΠΕΡΙΓΡΑΦΙΚΗ ΣΤΑΤΙΣΤΙΚΗ

Η περιγραφική στατιστική περιλαμβάνει συνοπτικά στατιστικά των δεδομένων όπως είναι τα

ΠΙΝΑΚΑΣ 1. Επιλογή στατιστικών ανάλογα με την κλίμακα μέτρησης		
Κλίμακα Μέτρησης	Εξαρτημένα Δείγματα	Ανεξάρτητα Δείγματα
Διαστήματος/Λόγου	Παραμετρική Στατιστική	Παραμετρική Στατιστική
Τακτική	Wilcoxon T Spearman r	Mann-Whitney U
Ονομαστική	Διωνυμικός Έλεγχος Προσήμου (Binomial Sign Test)	χ^2 (chi square)

μέτρα κεντρικής τάσης, (κυρίως, η μέση τιμή - mean, ο διάμεσος - median και η επικρατούσα τιμή - mode) και τα μέτρα μεταβλητότητας ή διασποράς (κυρίως, εύρος -range, διασπορά - variance, τυπική απόκλιση - standard deviation, τυπικό σφάλμα της μέσης τιμής -standard error of the mean). Φυσικά, υπάρχουν και άλλα μέτρα κεντρικής τάσης, όπως γεωμετρικός μέσος - (geometric mean), αρμονικός μέσος (harmonic mean), τετραγωνικός μέσος (quadratic mean), κινητός μέσος (moving average) και μέτρα μεταβλητότητας, όπως ημι-ενδοτεταρτημοριακό εύρος (semi-interquartile range), ενδοεκατοστημοριακό εύρος (interpercentile range), μέση απόλυτη απόκλιση (mean absolute deviation), συντελεστή μεταβλητότητας (coefficient of variation) κ.ά.

ΠΙΝΑΚΑΣ 2. Μέσες τιμές (M) και τυπικές αποκλίσεις (TA) της ορμόνης (mg/dl) στην πειραματική και ομάδα ελέγχου πριν και μετά την άσκηση

Ομάδες		Άσκηση		
		Πριν	Μετά	Σύνολο
Πειραματική n = 10	M	298.40	396.50	347.45
	TA	31.80	55.18	41.41
Ελέγχου n = 10	M	142.00	128.80	135.40
	TA	29.08	28.51	26.16
Σύνολο n = 20	M	220.20	262.65	241.43
	TA	85.54	143.83	113.18

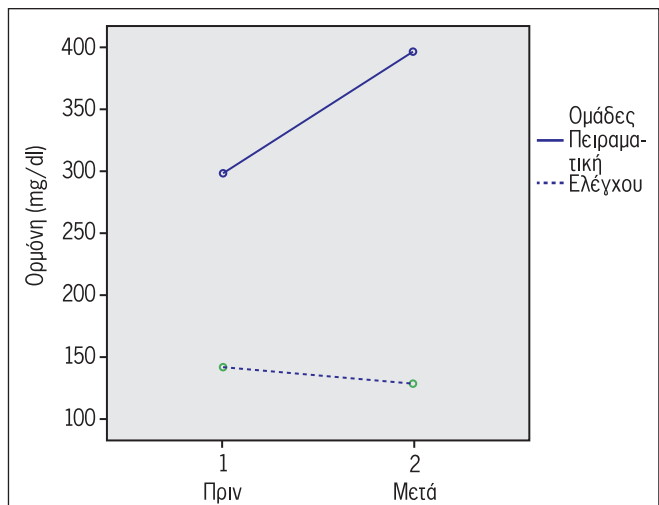
Ο ερευνητής παρουσιάζει σε πίνακες τα κατάλληλα μέτρα κεντρικής τάσης και μεταβλητότητας ανάλογα με την ερευνητική του ερώτηση και τις στατιστικές υποθέσεις. Για παράδειγμα, όταν ένας ερευνητής μελετά τις μεταβολές μιας ορμόνης, πριν και μετά την άσκηση, σε μία πειραματική και σε μία ομάδα ελέγχου (η οποία δεν αθλείται), τότε πρέπει να παρουσιάσει σε πίνακα τις μέσες τιμές και τις τυπικές αποκλίσεις της ορμόνης, πριν και μετά την άσκηση για την πειραματική και την ομάδα ελέγχου (βλ. Πίνακα 2). Στον Πίνακα φαίνεται ο αριθμός των ατόμων που συμμετέχουν στην έρευνα και ο αναγνώστης μπορεί να προσδιορίσει το πειραματικό σχέδιο που χρησιμοποίησε ο ερευνητής. Στο συγκεκριμένο παράδειγμα το πειραματικό σχέδιο είναι διπλής κατεύθυνσης (2-way) με επαναλαμβανόμενες μετρήσεις στον δεύτερο παράγοντα.

Υπάρχει, επίσης, περίπτωση όπου οι περισσότεροι εκδότες προτιμούν την παρουσίαση των μέσων τιμών σε Σχήμα, όπου προβάλονται οι τυπικές αποκλίσεις ή το τυπικό σφάλμα της μέσης τιμής ή τα διαστήματα εμπιστοσύνης. Σε περίπτωση που οι ερευνητικές ερωτήσεις έχουν συγκεκριμένες αντιπαραθέσεις ζευγών, μπορούμε να τοποθετήσουμε και το επίπεδο σημαντικότητας. Ένας απλός τρόπος απεικόνισης των μέσων τιμών παρουσιάζεται στο Σχήμα 1.

ΜΗ-ΠΑΡΑΜΕΤΡΙΚΗ - ΑΠΑΡΑΜΕΤΡΙΚΗ ΣΤΑΤΙΣΤΙΚΗ

Η μη-παραμετρική ή απαραμετρική στατιστική αναλύει δεδομένα ονομαστικής και τακτικής κλίμακας. Ως αντιπροσωπευτικές αναλύσεις της μη-παραμετρικής στατιστικής θα παρουσιαστούν: το χ^2 (chi square) και ο συντελεστής συσχέτισης Spearman. Για αναλυτική παρουσίαση

ΣΧΗΜΑ 1. Μέσες τιμές της ορμόνης (mg/dl) στην πειραματική και ομάδα ελέγχου πριν και μετά την άσκηση.



ΠΙΝΑΚΑΣ 3. Συχνότητες αιτιών θανάτου Α, Β, και Γ ανδρών και γυναικών που απεβίωσαν το ημερολογιακό έτος 2006-2007 (N =1190)

Φύλο	Αιτία Θανάτου		
	A	B	Γ
Άνδρες	270	140	190
Γυναίκες	170	330	90

ση του Διωνυμικού Ελέγχου Προσήμου (Binomial Sign Test), του ελέγχου Wilcoxon T και Mann-Whitney U, ο αναγνώστης μπορεί να ανατρέξει στους Lockhart, (1998) και Pagano & Gauvreau (2002).

Έλεγχος χ^2 (chi square)

Το χ^2 (chi square) χρησιμοποιείται για το έλεγχο της σημαντικότητας μεταξύ **συχνοτήτων** (ονομαστική κλίμακα) δύο ή περισσότερων δειγμάτων. Πρέπει πάντα να έχουμε στο νου μας ότι **σε καμία περίπτωση δεν χρ**

μοιούμε ποσοστιαίες αναλογίες, γιατί θα οδηγηθούμε σε εντελώς λανθασμένα αποτελέσματα. Επίσης, τα δεδομένα για την ανάλυση του χ^2 περιλαμβάνουν τα δεδομένα του κάθε ατόμου **μία και μόνο φορά**. Το άθροισμα των συχνοτήτων είναι ίσο με το σύνολο των ατόμων που συμμετέχουν στην έρευνα (για αναλυτική παρουσίαση βλ. Bland, 2000).

Ας αναλύσουμε τα δεδομένα μιας υποθετικής μελέτης που εξετάζει τις τρεις κυριότερες αιτίες θανάτου (Α, Β και Γ) σε σύνολο 1190 ανδρών και γυναικών που απεβίωσαν το τελευταίο ημερολογιακό έτος σε ένα Νομό της χώρας (βλ. Πίνακα 3). Η ερευνητική μας υπόθεση είναι ότι υπάρχει σχέση μεταξύ των τριών κύριων αιτιών θανάτου και του φύλου.

Εκτελώντας τη στατιστική ανάλυση με έγκυρο στατιστικό πακέτο (π.χ. SPSS, SAS, mini Tab, κτλ)*, έχουμε τον παρακάτω Πίνακα 4, όπου παρουσιάζονται οι πραγματικές και οι αναμενόμενες συχνότητες. Ο αναγνώστης πρέπει να γνωρίζει ότι ο υπολογισμός των αναμενόμενων συχνοτήτων είναι εύκολος και τις περισσότερες φορές δεν χρειάζεται να ανατρέξει στη χρησιμοποίηση στατιστικού πακέτου.

Η τιμή χ^2 , οι βαθμοί ελευθερίας και το επίπεδο σημαντικότητας παρουσιάζονται στον Πίνακα 5. Στην πρώτη γραμμή παρουσιάζεται η τιμή του ελέγχου $\chi^2 = 135.18$. Οι βαθμοί ελευθερίας (degrees of freedom - df) είναι 2 και το επίπεδο σημαντικότητας είναι μικρότερο από 0.001.

Έχοντας αυτά τα αποτελέσματα μπορούμε να προχωρήσουμε στην παρουσίαση και την επεξήγηση των αποτελεσμάτων. Οι παρακάτω προτάσεις είναι οι ενδεικνυόμενες για την παρουσίαση των στατιστικών αποτελεσμάτων του ελέγχου χ^2 .

Μπορούμε να περιγράψουμε τα αποτελέσματα όπως παρακάτω: "Υπάρχει στατιστικά σημαντική διαφορά μεταξύ της πραγματικής και της αναμενόμενης συχνότητας μεταξύ των ανδρών και γυναικών σε ότι αφορά τις τρεις αιτίες θανάτου ($\chi^2 = 135.18$, $df = 2$, $p < .001$)".

Άλλος τρόπος και περισσότερο ακριβής είναι: "Υπάρχει

ΠΙΝΑΚΑΣ 4. Πραγματικές (Π.Σ.) και αναμενόμενες συχνότητες (Α.Σ.) για τις αιτίες θανάτου Α, Β, και Γ σε άνδρες και γυναίκες

Φύλο		Φύλο x Αιτία Θανάτου			Σύνολο
		Αιτία θανάτου			
		A	B	Γ	
Άνδρες	Π.Σ.	270	140	190	600
	Α.Σ.	221.8	237.0	141.2	600.0
	Υπόλοιπο	48.2	-97.0	48.8	
Γυναίκες	Π.Σ.	170	330	90	590
	Α.Σ.	218.2	233.0	138.8	590.0
	Υπόλοιπο	-48.2	97.0	-48.8	
Σύνολο	Π.Σ.	440	470	280	1190
	Α.Σ.	440.0	470.0	280.0	1190.0

* Από το σημείο αυτό και μετά δε θα αναφέρουμε ότι πρέπει να χρησιμοποιηθεί στατιστικό πακέτο ή οποιαδήποτε άλλη στατιστική διαδικασία. Ο αναγνώστης πρέπει να γνωρίζει ότι η ανάλυση δεδομένων μπορεί να γίνει με τη χρήση έγκυρου στατιστικού πακέτου ή ακολουθώντας τη διαδικασία "paper and pencil" από βιβλία στατιστικής που προτείνονται στο άρθρο ή που έχει στη διάθεσή του.

μια σημαντική σχέση μεταξύ φύλου και των τριών κύριων αιτιών θανάτου ($\chi^2 = 135.18, df = 2, p < .001$).

Επιπροσθέτως, πρέπει να προσδιορίσουμε την κατεύθυνση των αποτελεσμάτων. Ένας σωστός τρόπος είναι: “Οι γυναίκες σε σχέση με τους άνδρες είναι πιθανότερο να πεθάνουν από την αιτία Β και λιγότερο πιθανό από τις αιτίες Α και Γ” ή “Συγκρινόμενοι με τις γυναίκες, οι άνδρες είναι πιθανότερο να έχουν ως αιτίες θανάτου την Α και Γ και λιγότερο πιθανό την αιτία Β”.

Σε καμία περίπτωση, ο επαγγελματίας υγείας δεν πρέπει να προχωρήσει σε προσδιορισμούς αιτίας και αποτελέσματος (cause and effect). Δεν είναι λίγοι εκείνοι που πολλές φορές θέλουν να προσδιορίσουν τις αιτίες, όπως π.χ. πιθανές μοιήνσεις του περιβάλλοντος, ύπαρξη βαριάς βιομηχανίας στην περιοχή, κτλ. Για να αποδώσει κάποιος αιτία στο αποτέλεσμα μιας έρευνας θα πρέπει το πειραματικό σχέδιο να το επιτρέπει. Τέτοια πειραματικά σχέδια θα παρουσιαστούν παρακάτω στην παραμετρική επαγωγική στατιστική.

Συντελεστής Συσχέτισης Spearman

Ο συντελεστής συσχέτισης Spearman (r_s) ονομάζεται και συντελεστής συσχέτισης σειρών, γιατί χρησιμοποιείται για τον προσδιορισμό της συνάφειας δύο τακτικών (ordinal scale) μεταβλητών που εκφράζουν τη σειρά κατάταξης σε δύο αντίστοιχες μεταβλητές (ή ιδιότητες) (Βαγενάς, 2003). Ο συντελεστής συσχέτισης Spearman κυμαίνεται από +1 έως -1. Τέλεια θετική συσχέτιση ($r_s = +1.00$) υπάρχει όταν οι δύο μεταβλητές συμπίπτουν ανά ζεύγη, δηλ το κάθε άτομο διατηρεί τη σειρά κατάταξης από την πρώτη στη δεύτερη μεταβλητή. Τέλεια αρνητική συσχέτιση ($r_s = -1.00$) έχουμε όταν οι σειρές των δύο μεταβλητών συμπίπτουν αντίστροφα, δηλ. το άτομο που είναι πρώτο στη σειρά στη μεταβλητή Α είναι τελευταίο στη σειρά στη μεταβλητή Β.

Ας εξετάσουμε τα δεδομένα μιας υποθετικής μελέτης όπου εξετάζεται η σχέση μεταξύ του ποσοστού των εμβολιασμένων παιδιών για DPT (διφθερίτιδα, κοκίτη και τέτανο) (Μεταβλητή Α) και του ρυθμού θνησιμότητας (ανά 1000 παιδιά) σε παιδιά κάτω των 5 χρόνων (Μεταβλητή Β) σε 10 χώρες. Τα υποθετικά δεδομένα παρουσιάζονται στον Πίνακα 6. Τα δεδομένα παρουσιάζονται με αυτόν τον τρόπο (δηλ. πραγματικά ποσοστά (μεταβλητή Α) και αριθμός θανάτων (μεταβλητή Β) και σειρές κατάταξης ανά μεταβλητή) για να δείξουμε τη διαφορά του συντελεστή συσχέτισης Spearman (r_s) με το συντελεστή συσχέτισης Pearson που θα αναπτυχθεί παρακάτω.

Από τη στατιστική ανάλυση των σειρών Α και Β προκύπτει ο συντελεστής συσχέτισης Spearman (r_s) που ισούται με $r_s = -.71$, με βαθμούς ελευθερίας 8 ($df = N-2 = 8$) και επίπεδο σημαντικότητας $p < .05$.

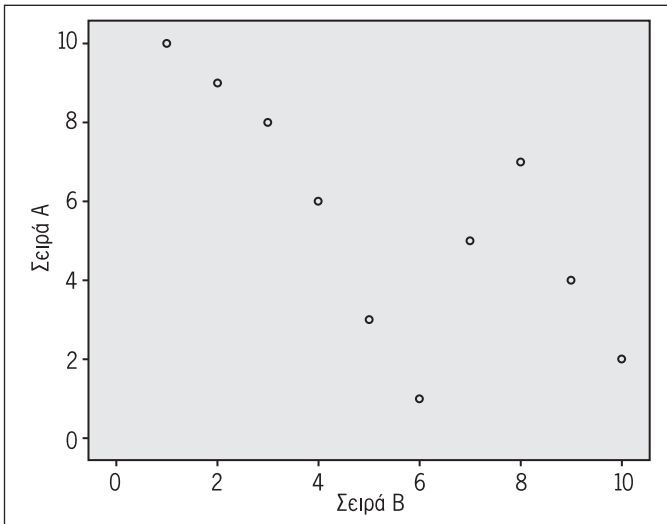
Στην επεξήγηση των δεδομένων, ο ερευνητής μπορεί να επισημάνει: “Υπάρχει στατιστικά σημαντική αρνητική συσχέτιση μεταξύ των σειρών κατάταξης των ποσοστών εμβολιασμού και του ρυθμού θνησιμότητας των παιδιών κάτω των 5 χρόνων στις 10 χώρες (ρ ή $r_s = -.71, df = 8, p < .05$). Οι χώρες που

ΠΙΝΑΚΑΣ 5. Πίνακας ελέγχου χ^2

Έλεγχος χ^2			
	Τιμή	df	Sig. (2-sided)
Pearson Chi-Square	135.18	2	.000
N of Valid Cases	1190		

ΠΙΝΑΚΑΣ 6. Ποσοστά παιδιών εμβολιασμένων (Μεταβλητή Α) και ρυθμός θνησιμότητας (Μεταβλητή Β) σε 10 χώρες, καθώς και οι σειρές κατάταξης των δύο μεταβλητών

Χώρες (N = 10)	Μεταβλητή Α	Μεταβλητή Β	Σειρά Α	Σειρά Β
1	54	16	7	8
2	84	12	4	9
3	21	230	10	1
4	95	50	1	6
5	87	54	3	5
6	50	140	8	3
7	25	200	9	2
8	74	80	6	4
9	80	20	5	7
10	90	10	2	10



ΣΧΗΜΑ 2. Διάγραμμα διασποράς μεταξύ των σειρών κατάταξης Α και Β (υπολογισμός Spearman rs).

είναι πρώτες στην κατάταξη στο ποσοστό εμβολιασμού τείνουν να έχουν τη μικρότερη σειρά κατάταξης στο ρυθμό θνησιμότητας και το αντίστροφο". Επίσης, στα περισσότερα δημοσιευμένα άρθρα υπάρχει και το διάγραμμα διασποράς (scatterplot) (Σχήμα 2).

Με το συντελεστή συσχέτισης Spearman (r_s), ο επαγγελματίας υγείας δεν μπορεί να μιλήσει για τη σχέση των ποσοστών των εμβολιασμένων παιδιών και του ρυθ-

μού θνησιμότητας των παιδιών. Εάν θέλει να μιλήσει γι αυτή τη σχέση, θα πρέπει να χρησιμοποιήσει το συντελεστή συσχέτισης Pearson που είναι κατάλληλος για δεδομένα που έχουν καταγραφή σε κλίμακα διαστήματος.

ΠΑΡΑΜΕΤΡΙΚΗ ΣΤΑΤΙΣΤΙΚΗ

Η παραμετρική στατιστική αναλύει δεδομένα της κλίμακας διαστήματος και λόγου. Ως απλές αντιπροσωπευτικές αναλύσεις παραμετρικής επαγωγικής στατιστικής θα παρουσιαστούν: ο συντελεστής συσχέτισης Pearson (r), ο έλεγχος t για εξαρτημένα και ανεξάρτητα δείγματα, και απλή ανάλυση διασποράς (**Analysis of Variance** - ANOVA) για ανεξάρτητα και εξαρτημένα δείγματα. Θα παρουσιαστεί, επίσης, η στατιστική μέθοδος της παραγοντικής ανάλυσης (factor analysis) με τη μέθοδο εκτίμησης των Κύριων Συνιστωσών και την ορθογώνια περιστροφή Varimax.

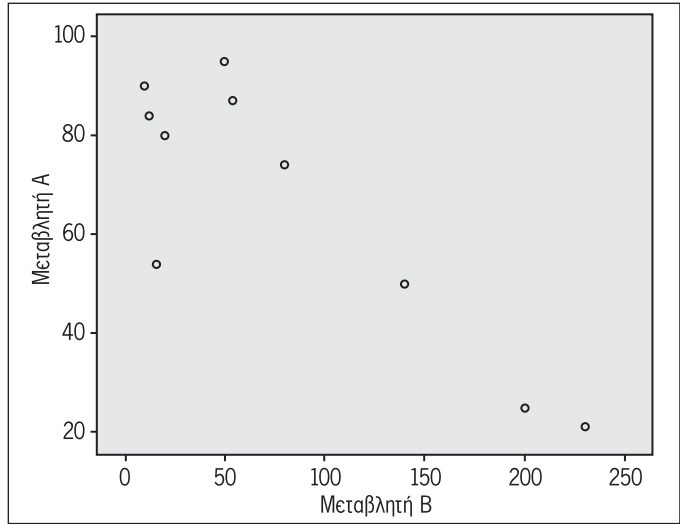
Συντελεστής συσχέτισης Pearson (r)

Ο συντελεστής συσχέτισης Pearson (r) χρησιμοποιείται για τον προσδιορισμό της σχέσης μεταξύ δύο συνεχών μεταβλητών (κλίμακα μέτρησης διαστήματος/αναλογίας) της ίδιας ομάδας ατόμων ή περιπτώσεων. Ο συντελεστής συσχέτισης Pearson (r) κυμαίνεται από +1 έως -1 (όπως και ο συντελεστής συσχέτισης Spearman). Ως παράδειγμα για τον υπολογισμό του συντελεστή συσχέτισης Pearson, θα χρησιμοποιήσουμε τα δεδομένα των μεταβλητών Α και Β του Πίνακα 6. Αναλύοντας τα πραγματικά δεδομένα των δύο συνεχών μεταβλητών και όχι τις σειρές κατάταξης των μεταβλητών, θα διαπιστώσουμε ότι ο συντελεστής συσχέτισης Pearson έχει διαφορετικό μέγεθος από το συντελεστή συσχέτισης Spearman. Πράγματι, η ανάλυση των δεδομένων μας δίνει $r = -.86$, $df = 8$, και $p < .001$.

Ερμηνεύοντας τα αποτελέσματα ο ερευνητής μπορεί να αναφέρει ότι: "Υπάρχει στατιστικά σημαντική αρνητική συσχέτιση μεταξύ του ποσοστού των εμβολιασμένων παιδιών και του ρυθμού θνησιμότητας των παιδιών κάτω των 5 χρόνων ($r = -.86$, $df = 8$, και $p < .001$). Οι χώρες με υψηλά ποσοστά εμβολιασμού των παιδιών για DPT έχουν μικρότερο ρυθμό θνησιμότητας". Επιπροσθέτως, ο ερευνητής μπορεί να παρουσιάσει και το διάγραμμα διασποράς (Σχήμα 3).

Συγκρίνοντας το διάγραμμα διασποράς μεταξύ των μεταβλητών Α και Β (βλ. Σχήμα 3) με το διάγραμμα διασποράς των σειρών κατάταξης των μεταβλητών (Σχήμα 2), διαπιστώνουμε ότι έχουν την ίδια κατεύθυνση, αλλά διαφορετικό μέγεθος (ρ ή $r_s = -.71$, $df = 8$, $p < .05$ έναντι $r = -.86$, $df = 8$, και $p < .001$) και διαφορετικές τοποθετήσεις των σημείων. Αυτό οφείλεται στη μεγαλύτερη ακρίβεια της κλίμακας διαστήματος από την τακτική κλίμακα, καθώς και στο διαφορετικό τρόπο υ-

ΣΧΗΜΑ 3. Διάγραμμα διασποράς μεταξύ των μεταβλητών A και B (υπολογισμός Pearson r).



πολογισμού των συσχετίσεων. Ο αναγνώστης πρέπει να γνωρίζει ότι ανάλογα με την ερευνητική του ερώτηση, θα ακολουθήσει παραμετρική ή μη-παραμετρική προσέγγιση του θέματος.

Κρίνεται απαραίτητο να τονιστεί ότι ο συντελεστής συσχέτισης αποτελεί έναν αριθμό και δεν παρέχει καμιά πληροφορία σχετικά με τις αιτίες που προκάλεσαν τη σχέση μεταξύ των μεταβλητών. Η συσχέτιση μπορεί (α) να αποτελεί απλή σύμπτωση ή τυχαίο γεγονός που προέκυψε από διάφορους παράγοντες (π.χ. λάθος δειγματοληψίας, επιλογή μεταβλητών, λανθασμένη ερευνητική διαδικασία, σφάλματα μέτρησης) και (β) να είναι απόρροια επίδρασης άλλων μεταβλητών.

Επίσης, ο ερευνητής δεν πρέπει να δίνει μεγάλη έμφαση στη σημαντικότητα του r όταν το δείγμα που χρησιμοποιείται είναι μεγάλο ($N > 100$), γιατί η τιμή r επιτυγχάνει στατιστική σημαντικότητα ($p < .05$) σε μέγεθος $r = .17$. Το μέγεθος μιας τέτοιας συσχέτισης, αν και στατιστικά σημαντικό, είναι αρκετά μικρό για να μπορέσουμε να κάνουμε επισημάνσεις για τη σχέση των μεταβλητών. Ο επαγγελματίας υγείας ελέγχει και το **συντελεστή προσδιορισμού r^2** που αποτελεί το δείκτη της κοινής διασποράς των δύο συσχετιζόμενων μεταβλητών και υπολογίζεται από το γινόμενο του πολλαπλασιαζόμενο με το 100. Στο παράδειγμά μας, ο r^2 είναι ίσος με 73.96% [$(-.86) * (-.86) * 100 = 73.96\%$]. Το αποτέλεσμα αυτό σημαίνει ότι η συνδιασπορά των μεταβλητών A και B επεξηγεί το 73.96% της συσχέτισης, ενώ το υπόλοιπο 26.04% μπορεί να οφείλεται σε άλλους παράγοντες.

Έλεγχος t (t-test) για εξαρτημένα δείγματα

Ο έλεγχος t για εξαρτημένα δείγματα χρησιμοποιείται, για να υπολογίσουμε, εάν οι μέσες τιμές δύο μετρήσεων (π.χ. πρώτη μέτρηση και δεύτερη μέτρηση) που έγιναν στο ίδιο δείγμα ανθρώπων διαφέρουν μεταξύ τους. Για παράδειγμα, μας ενδιαφέρει να δούμε τις μεταβολές μιας ορμόνης X μετά από άσκηση συγκεκριμένης έντασης σε κυλιόμενο τάπητα. Γίνονται οι απαραίτητες μετρήσεις πριν την άσκηση και αμέσως μετά την άσκηση. Ας υποθέσουμε ότι στην έρευνα συμμετείχαν 10 ασθενείς. Οι τιμές της ορμόνης X παρουσιάζονται στον Πίνακα 7.

Από την ανάλυση των δεδομένων παρουσιάζουμε τις μέσες τιμές, τις τυπικές αποκλίσεις και τα τυπικά σφάλματα των μέσων τιμών (Πίνακας 8). Σε περίπτωση που κρίνεται αναγκαίο, παρουσιάζουμε και σχήμα, όπου απεικονίζονται οι μέσες τιμές με τις τυπικές αποκλίσεις ή τα τυπικά σφάλματα των μέσων τιμών ή τα διαστήματα εμπιστοσύνης. Επίσης, πρέπει να παρουσιάσουμε και το συντελεστή συσχέτισης των δύο μεταβλητών. Δεν ξεχνάμε ότι στα ε-

ΠΙΝΑΚΑΣ 7. Τιμές ορμόνης X πριν και μετά τη άσκηση σε 10 ασθενείς

ID	Πριν	Μετά
1	154	170
2	184	190
3	121	125
4	195	189
5	187	195
6	150	140
7	125	145
8	174	180
9	180	195
10	190	200

ΠΙΝΑΚΑΣ 8. Μέσες τιμές, τυπικές αποκλίσεις και τυπικά σφάλματα της μέσης τιμής της ορμόνης Χ πριν και μετά την άσκηση

ΟρμόνηΧ	Μέση τιμή	N	Τυπική απόκλιση	Τυπικό σφάλμα μέσης τιμής
Πριν	166.00	10	26.97	8.53
Μετά	172.90	10	26.83	8.48

ξαρτημένα δείγματα ο έλεγχος t γίνεται με τον συνυπολογισμό του Pearson r που στην περίπτωση μας είναι ίσος με 0.94 ($r = .94$, $df = 8$, $p < .001$).

Στη συνέχεια, από τον υπολογισμό των διαφορών των μέσων τιμών (διαφορά - d = -6.90) και του υπολογισμού του τυπικού σφάλματος της διαφοράς (2.968) υπολογίζουμε και την τιμή $t = -2.33$, με $df = 9$ και p

< .05. Στον Πίνακα 9 παρουσιάζεται η εκτύπωση από το στατιστικό πακέτο SPSS. Ο αναγνώστης έχει όλες τις απαραίτητες πληροφορίες όπως διαφορά μεταξύ των μέσων τιμών (-6.90), τυπική απόκλιση της διαφοράς των μέσων τιμών (9.368) τυπικό σφάλμα της διαφοράς των μέσων τιμών (2.968), διάστημα εμπιστοσύνης 95% της διαφοράς των μέσων τιμών, τον έλεγχο t (-2.325), τους βαθμούς ελευθερίας ($df = 9$) και τη σημαντικότητα του t μετά από δίπλευρο έλεγχο της σημαντικότητας ($p = .045$).

Για την επεξήγηση των αποτελεσμάτων μπορούμε να πούμε: "Οι μέσες τιμές της Ορμόνης Χ πριν ($M = 166.00 \pm 26.97$) και μετά την άσκηση ($M = 172.90 \pm 26.83$) παρουσίασαν στατιστικά σημαντική διαφορά ($t = -2.33$, $df = 9$, $p < .05$). Από τα ευρήματα της μελέτης διαφαίνεται ότι η άσκηση επηρεάζει σημαντικά τη συγκέντρωση της ορμόνης Χ."

Υπάρχει, επίσης η περίπτωση όπου οι ερευνητές προτιμούν να παρουσιάσουν τα διαστήματα εμπιστοσύνης. Οπότε μπορούμε να παρουσιάσουμε τα ευρήματά μας ως εξής: "Η μέση τιμή της Ορμόνης Χ, πριν την άσκηση ήταν $M = 166.00 (\pm 26.97)$ και μετά την άσκηση ήταν $M = 172.90 (\pm 26.83)$. Η διαφορά των μέσων τιμών της Ορμόνης πριν και μετά την άσκηση ήταν -6.90. Το 95% διάστημα εμπιστοσύνης για αυτή τη διαφορά κυμάνθηκε από -13.614 έως -.186. Εφόσον, το διάστημα εμπιστοσύνης δεν καλύπτει την τιμή 0.00, η διαφορά των μέσων τιμών είναι στατιστικά σημαντική σύμφωνα με το δίπλευρο έλεγχο της σημαντικότητας στο επίπεδο $p < .05$."

Έλεγχος t (t-test) για ανεξάρτητα δείγματα

Ο Έλεγχος t (t-test) για ανεξάρτητα δείγματα (ή μη συσχετιζόμενα δείγματα) χρησιμοποιείται, για να υπολογίσουμε, εάν οι μέσες τιμές δύο διαφορετικών ομάδων διαφέρουν μεταξύ τους. Ας χρησιμοποιήσουμε, ως παράδειγμα, τα δεδομένα του Πίνακα 7 με τη διαφορά ότι η πρώτη στήλη ανήκει σε μία ομάδα ανδρών και η δεύτερη στήλη σε μία ομάδα γυναικών. Οι δύο αυτές ομάδες μετρήθηκαν σε κατάσταση ηρεμίας στην εξαρτημένη μεταβλητή Ορμόνη Χ. Από την ανάλυση των δεδομένων παρατηρούμε ότι οι μέσες τιμές, οι τυπικές αποκλίσεις και τα τυπικά σφάλματα των μέσων τιμών παραμένουν οι ίδιες. Ωστόσο, τα αποτελέσματα του ελέγχου t είναι εντελώς διαφορετικά (βλ. Πίνακα 10). Εκείνο, όμως, που αλλιάζει είναι η τιμή του t και οι διάφοροι έλεγχοι που πρέπει να γίνουν για την εγκυρότητα των αποτελεσμάτων.

Αναλυτικότερα, σε πρώτη φάση ελέγχουμε εάν οι διασπορές είναι ίσες μεταξύ τους με το Levene's Test for Equality of Variances. Το Levene's Test for Equality of Variances, στην περίπτωση μας, μας υποδηλώνει ότι οι διασπορές δεν είναι ίσες ($F = .007$, $p = .933$) και μπορούμε να χρησι-

ΠΙΝΑΚΑΣ 9. Αποτέλεσμα της στατιστικής ανάλυσης για τον έλεγχο t σε εξαρτημένα δείγματα.

Paired Samples Test									
		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error	95% Confidence Interval of the Difference				
						Lower	Upper		
Pair 1	Pre-post	-6,900	9,386	2,968	-13,614	-,186	-2,325	9	,045

ΠΙΝΑΚΑΣ 10. Αποτέλεσμα της στατιστικής ανάλυσης για τον έλεγχο t σε ανεξάρτητα δείγματα										
Independent Samples Test										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Pre	Equal variances assumed	,007	,933	-,574	18	,573	-6,900	12,030	-32,174	18,374
	Equal variances not assumed			-,574	17,999	,573	-6,900	12,030	-32,174	18,374

μποιήσουμε την τιμή $t = .57$, με βαθμούς ελευθερίας $df = 18$ και επίπεδο σημαντικότητας $p = .57$. Σε περίπτωση που οι διασπορές είναι άνισες, τότε ελέγχουμε τη δεύτερη σειρά (Equal variances not assumed) και αποδίδουμε τη σημαντικότητα σύμφωνα με τους διορθωμένους βαθμούς ελευθερίας. Ο αναγνώστης μπορεί να παρατηρήσει ότι εκείνο που αλλάζει είναι οι βαθμοί ελευθερίας και το επίπεδο σημαντικότητας (στο παράδειγμά μας δεν φαίνεται η αλλαγή στους βαθμούς ελευθερίας, λόγω της υψηλής ομοιογένειας των διασπορών). Όλοι οι άλλοι υπολογισμοί παραμένουν οι ίδιοι.

Ερμηνεύοντας τα αποτελέσματα ο ερευνητής μπορεί να αναφέρει ότι: "Η μέση τιμή της Ορμόνης X των ανδρών ($M = 166.00 \pm 26.97$) δεν παρουσίασε στατιστικά σημαντική διαφορά ($t = -.57$, $df = 18$, $p > .05$) από τη μέση τιμή των γυναικών ($M = 172.90 \pm 26.83$). Από τα ευρήματα της μελέτης διαφαίνεται ότι δεν υπάρχουν σημαντικές διαφορές στη συγκέντρωση της ορμόνης X μεταξύ των ανδρών και γυναικών." Σε περίπτωση, όμως, που υπάρχει στατιστική σημαντικότητα στις διασπορές, τότε αυτό θα πρέπει να αναφερθεί και να επισημανθεί, εάν ο έλεγχος t παραμένει σημαντικός λαμβάνοντας υπόψη την ανισότητα των διασπορών.

Εάν ο ερευνητής προτιμά να χρησιμοποιήσει τα διαστήματα εμπιστοσύνης, τότε μπορούμε να γράψουμε: "Η διαφορά μεταξύ των μέσων τιμών της Ορμόνης X για τους άνδρες ($M = 166.00 \pm 26.97$) και τις γυναίκες ($M = 172.90 \pm 26.83$) είναι -6.90 . Το 95% διάστημα εμπιστοσύνης για αυτή τη διαφορά κυμάνθηκε από -32.174 έως 18.374 . Εφόσον, το διάστημα εμπιστοσύνης συμπεριλαμβάνει την τιμή 0.00, η διαφορά των μέσων τιμών δεν είναι στατιστικά σημαντική, σύμφωνα με το δίπλευρο έλεγχο της σημαντικότητας στο επίπεδο $p < .05$."

Συγκρίνοντας τους δύο ελέγχους t για εξαρτημένα και ανεξάρτητα δείγματα, ο αναγνώστης μπορεί να καταλάβει το λάθος που μπορεί να οδηγηθεί σε περίπτωση που αντιμετωπίσει τα δεδομένα ως ανεξάρτητα, τη στιγμή που είναι εξαρτημένα και αντίστροφα.

Απλή ανάλυση διασποράς για ανεξάρτητα δείγματα

Η Απλή Ανάλυση Διασποράς (*Analysis of Variance* - ANOVA) για ανεξάρτητα ή μη συσχετιζόμενα δείγματα, μας χρησιμεύει να διαπιστώσουμε, εάν υπάρχουν σημαντικές διαφορές μεταξύ τριών ή περισσότερων αριθμητικών μέσων (μέσων τιμών). Θα αναπτύξουμε όλη τη διαδικασία με τα υποθετικά δεδομένα που παρουσιάζονται στο Πίνακα 11.

Σε αυτή τη μελέτη εξετάζουμε τη επίδραση ορμονών στην αντιμετώπιση της κατάθλιψης. Ελέγχουμε κατά πόσο η χορήγηση συγκεκριμένης φαρμακευτικής αγωγής συμβάλλει σημαντικά στην αντιμετώπιση της κατάθλιψης. Χορηγούμε δύο Ορμόνες, την Ορμόνη Α και Β και αντί για ομάδα ελέγχου χρησιμοποιούμε ομάδα Placebo (στη ομάδα αυτή χορηγούμε μια ουσία που είναι πανομοιότυπη με την ερευνητική ουσία, χωρίς φυσικά να έχει καμία φαρμακευτική επίδραση στους συμμετέχοντες).*

* Ο αναγνώστης πρέπει να γνωρίζει ότι σε έρευνες αυτού του είδους ακολουθεί σωστές μεθοδολογικές διαδικασίες (π.χ. double blind experiments), όπου αυτοί που δίνουν τις ορμόνες (Α, Β και Placebo) και οι ασθενείς δεν γνωρίζουν σε ποια ομάδα ανήκουν.

ΠΙΝΑΚΑΣ 11. Δεδομένα για μια υποθετική μελέτη επίδρασης ορμονών στην κατάθλιψη

ID	Ομάδες	Θεραπεία	Κατάθλιψη
1	1	Ορμόνη Α	10
2	1	Ορμόνη Α	11
3	1	Ορμόνη Α	9
4	2	Ορμόνη Β	5
5	2	Ορμόνη Β	5
6	2	Ορμόνη Β	4.5
7	3	Placebo	4
8	3	Placebo	3
9	3	Placebo	4

Δεν είναι λίγες οι φορές, όπου πολλοί νέοι ερευνητές, ενώ χορηγούν τρεις διαφορετικές ορμόνες αρκούνται στο να εξετάσουν τις διαφορές μεταξύ των ορμονών κάνοντας πολλαπλούς ελέγχους t . Σε περίπτωση που καταλήγουν σε τέτοιες τεχνικές και δεν ακολουθούν τη μέθοδο της ANOVA, τότε οδηγούνται σε λανθασμένα συμπεράσματα που οφείλονται κυρίως στη διόγκωση του σφάλματος τύπου I (Type I Error), όπου υπάρχει η πιθανότητα να απορρίψουν τη μηδενική υπόθεση (H_0) και να υιοθετήσουν την εναλλακτική (H_A), χωρίς να υπάρχουν στατιστικά σημαντικές διαφορές.

Ο επαγγελματίας υγείας πρέπει να γνωρίζει πως, όταν καταλήγει σε πολλαπλούς ελέγχους t μπορεί να μιλάει **“αποκλειστικά και μόνο”** για τις δύο συγκρινόμενες μέσες τιμές, χωρίς να συνδυάζει το εύρημα αυτό με τις άλλες συγκρίσεις των μέσων τιμών. Αυτό μπορεί να το κάνει μόνο, όταν ακολουθήσει τη μέθοδο ANOVA και προχωρήσει σε κατάλληλες post hoc αναλύσεις, μετά την εύρεση στατιστικά σημαντικής διαφοράς στα επίπεδα της εξεταζόμενης ανεξάρτητης μεταβλητής (βλ. παρακάτω).

Επίσης, με τις πολλαπλές συγκρίσεις το επίπεδο σημαντικότητας αυξάνεται, ανάλογα με τον αριθμό των δυνατών συνδυασμών των μέσων τιμών. Ο αριθμός των δυνατών συγκρίσεων υπολογίζεται από τον τύπο $x(x-1)/2$, όπου x ο αριθμός των μέσων τιμών. Στην περίπτωσή μας έχουμε $x = 3$ μέσες τιμές, άρα $(3(3-1)/2) = 3$ δυνατούς συνδυασμούς. Η ολική πιθανότητα του σφάλματος τύπου I (α) σε σχέση με τον αριθμό των συγκρίσεων (c) και του προκαθορισμένου επιπέδου σημαντικότητας α (π.χ. $\alpha = 0.05$) για κάθε σύγκριση δίνεται από τη σχέση $\alpha_i = 1 - (1-\alpha)^c = 1 - (1-0.05)^3 = 0.14$. Η ολική πιθανότητα απόρριψης της μηδενικής υπόθεσης ανέρχεται στο επίπεδο $p = .14$ και ο ερευνητής πιστεύει ότι είναι στο επίπεδο $p \leq .05$. Σε περίπτωση που έχουμε 10 συγκρίσεις μέσων τιμών, τότε καταλήγουμε στο επίπεδο πιθανότητας $p = .40$.

Η σωστή διαδικασία για μελέτες όπου συγκρίνουν από τρεις και πάνω μέσες τιμές είναι η ANOVA. Στο παράδειγμά μας, η απλή ανάλυση διασποράς για ανεξάρτητα δείγματα μας έδωσε τα παρακάτω αποτελέσματα. Οι μέσες τιμές και οι τυπικές αποκλίσεις των τριών ομάδων στην εξαρτημένη μεταβλητή της κατάθλιψης παρουσιάζονται στον Πίνακα 12.

Ένας σημαντικός έλεγχος που πρέπει να ληφθεί υπόψη είναι ο έλεγχος της ομοιογένειας της διασποράς που συνήθως γίνεται με τον έλεγχο Levene (Levene statistic). Όταν το επίπεδο σημαντικότητας του ελέγχου Levene δεν είναι στατιστικά σημαντικό, τότε οι διασπορές είναι ομοιογενείς (δεν διαφέρουν μεταξύ τους) και μπορούμε να προχωρήσουμε στην επεξήγηση του πηλίκου F . Στο παράδειγμά μας, ο έλεγχος Levene δεν παρουσίασε στατιστικά σημαντικές διαφορές ($F_{2,6} = 1.17, p = .372$) στην ομοιογένεια των διασπορών (Πίνακας 13). Αυτό αποδεικνύει ότι οι διασπορές

των τριών ομάδων είναι ομοιογενείς σε πληθυσμιακό επίπεδο.

Στη συνέχεια δημιουργούμε τον Πίνακα με τα περιληπτικά αποτελέσματα της ANOVA (Πίνακας 14) και παρουσιάζουμε, τις Πηγές Διασποράς (Source of Variance), τα Αθροίσματα Τετραγώνων (Sums of Squares), τους βαθμούς ελευθερίας, τα Μέσα Τετράγωνα (Mean Squares), το Πηλίκο F (F ratio) και το επίπεδο σημαντικότητας (significance - p).

ΠΙΝΑΚΑΣ 12. Μέσες τιμές και τυπικές αποκλίσεις των ομάδων Ορμόνη Α, Ορμόνη Β, και Placebo στην κατάθλιψη

Ομάδες	N	Μέση τιμή	Τυπική απόκλιση	Τυπικό σφάλμα μέσης τιμής
Ορμόνη Α	3	10.00	1.00	.58
Ορμόνη Β	3	4.83	.29	.17
Placebo	3	3.67	.58	.33
Σύνολο	9	6.17	2.98	.99

Από την ANOVA φαίνεται ότι υπάρχει στατιστικά σημαντική διαφορά μεταξύ των Ομάδων ($F_{2,6} = 72.18, p < 0.001$).^{*} Το αποτέλεσμα αυτό σημαίνει ότι υπάρχει σημαντική διαφορά μεταξύ των ομάδων. Ωστόσο, δεν μας λέει ότι όλες οι ομάδες διαφέρουν μεταξύ τους. Για να εντοπίσουμε τις στατιστικά σημαντικές διαφορές πρέπει να προχωρήσουμε σε μετά-ANOVA πολλαπλές συγκρίσεις (post hoc analyses). Στη στατική υπάρχουν αρκετές *post hoc* αναλύσεις. Η πιο φιλελεύθερη *post hoc* ανάλυση είναι ο έλεγχος ελάχιστης διαφοράς LSD και ο πιο συντηρητικός είναι ο έλεγχος Scheffè (Scheffè, 1953). Στο ενδιάμεσο βρίσκεται ο έλεγχος Tukey και ο έλεγχος διόρθωσης του Bonferroni.

Στο παράδειγμά μας, η Tukey *post hoc* ανάλυση έδειξε ότι η Ομάδα της Ορμόνης A ($M = 10.00 \pm 1.00$) βελτιώθηκε σημαντικά σε σύγκριση με την Ομάδα της Ορμόνης B ($M = 4.83 \pm .29$) και την Ομάδα Placebo ($M = 3.67 \pm .58$), ενώ η Ομάδα της Ορμόνης B δεν διέφερε σημαντικά από την Ομάδα Placebo. Οι ίδιες στατιστικά σημαντικές διαφορές παρουσιάστηκαν με τους ελέγχους, Scheffè και Bonferroni.

Επανερχόμενοι στην επιλογή των πολλαπλών ελέγχων t, χωρίς να ελέγχουμε τις διαφορές των Ομάδων με ANOVA, θα διαπιστώσουμε ότι όλοι οι έλεγχοι t αποδίδουν στατιστικά σημαντικές διαφορές. Αναλυτικότερα, η σύγκριση της Ομάδας A με την Ομάδα B μας έδωσε $t_4 = 8.60$ με $p < .001$, η σύγκριση της Ομάδας A με την Ομάδα Placebo με έδωσε $t_4 = 9.50, p < .001$, και η σύγκριση της Ομάδας B με την Ομάδα Placebo $t_4 = 3.13, p < .05$. Ο επαγγελματίας υγείας που θα ακολουθήσει αυτή τη διαδικασία έχει κάνει το σφάλμα τύπου I και θα οδηγηθεί σε λανθασμένα συμπεράσματα, τονίζοντας την ευεργετική επίδραση της Ορμόνης B στην κατάθλιψη, ενώ η επίδραση αυτή μπορεί να οφείλεται καθαρά σε τυχαίους παράγοντες. Επίσης, ο επαγγελματίας υγείας σε καμία περίπτωση δεν μπορεί να επεξηγήσει την επίδραση της μίας ορμόνης (π.χ. Ορμόνη A) σε σύγκριση με τις άλλες δύο ορμόνες (π.χ. Ορμόνη B και Placebo).

Ο ερευνητής θα μπορούσε να αναφέρει τα αποτελέσματα της έρευνας ως εξής: "Το αποτέλεσμα της επίδρασης των ορμονών στην κατάθλιψη ήταν στατιστικά σημαντικό ($F_{2,6} = 72.18, p < 0.001$). Ακολουθώντας τη μέθοδο διόρθωσης του Bonferroni για τον αριθμό των δυνατών συγκρίσεων, η μόνη στατιστικά σημαντική διαφορά παρουσιάστηκε μεταξύ των μέσων τιμών της Ομάδας Ορμόνη A σε σύγκριση με την Ομάδα Ορμόνη B και Placebo ($p < .05$). Η μέση τιμή στην κατάθλιψη της Ομάδας που χορηγήθηκε η Ορμόνη A ($M = 10.00 \pm 1.00$) ήταν σημαντικά μεγαλύτερη από αυτή της Ομάδας που χορηγήθηκε η Ορμόνη B ($M = 4.83 \pm .29$) και Placebo ($M = 3.67 \pm .58$). Δεν παρουσιάστηκε στατιστικά σημαντική διαφορά μεταξύ των μέσων τιμών της Ομάδας που χορηγήθηκε η Ορμόνη B και Placebo."

Απλή ανάλυση διασποράς για εξαρτημένα δείγματα

Η Απλή Ανάλυση Διασποράς για εξαρτημένα ή συσχετιζόμενα δείγματα μας χρησιμεύει να διαπιστώσουμε εάν υπάρχουν σημαντικές διαφορές μεταξύ τριών ή περισσότερων επαναλαμβανόμενων μετρήσεων της ίδιας εξαρτημένης μεταβλητής πάνω στα ίδια άτομα. Θα χρησιμοποιήσουμε τα ίδια δεδομένα του Πίνακα 11, αλλά με διαφορετική κωδικοποίηση. Η Ομάδα Placebo μετονομάζεται σε θεραπεία Placebo, η ομάδα της Ορμόνης B σε θεραπεία Ασπιρίνης και η Ομάδα A σε θεραπεία Φαρμάκου X. Σκοπός της υποθετικής έρευνας είναι να εξετάσουμε, εάν το φάρμακο X έχει καλύτερη επίδραση από την Ασπιρίνη και την Ομάδα Ελέγχου (Placebo)** στην εξαρτημένη μεταβλη-

ΠΙΝΑΚΑΣ 13. Έλεγχος ομοιογένειας της διασποράς (Levene Statistic) για την κατάθλιψη στις τρεις πειραματικές ομάδες

Levene Statistic	df1	df2	Sig.
1.171	2	6	.372

* Το ηλίκο F υπολογίζεται από τα μέσα τετράγωνα των ομάδων διαιρούμενα με τα μέσα τετράγωνα του σφάλματος και μας δίνει την τιμή $72.18 (34.08/.47 = 18)$.

** Για τη δοσολογία και τον τρόπο χορήγησης του φαρμάκου υποτίθεται ότι ακολουθήθηκε η σωστή μεθοδολογία.

ΠΙΝΑΚΑΣ 14. Περιληπτικά αποτελέσματα της απλής ANOVA για ανεξάρτητα δείγματα των πειραματικών ομάδων στην εξαρτημένη μεταβλητή της κατάθλιψης					
Πηγή διασποράς	Άθροισμα τετραγώνων	df	Μέσα τετράγωνα	Πηλίκιο F	p
Ομάδες	68.17	2	34.08	72.17	.000
Σφάλμα	2.83	6	.47		
Ολική	71.00	8			

τή ανακούφισης από τον πόνο. Τα αρχικά δεδομένα για τα τρία άτομα που συμμετείχαν στην έρευνα παρουσιάζονται στον Πίνακα 15.

Η διαφορά της απλής ANOVA για ανεξάρτητα δείγματα και της ANOVA για εξαρτημένα δείγματα βρίσκεται, κυρίως: (α) στο διαμερισμό της διασποράς, όπου εκτός της διασποράς **μεταξύ των μετρήσεων - θεραπειών** και του **σφάλματος** προστίθεται και η διασπορά **εντός των υποκειμένων** και (β) στον έλεγχο ομοιογένειας των διασπορών και συνδιασπορών των δεδομένων. Οπότε, η ολική διασπορά βρίσκεται από το άθροισμα των 3 προαναφερθεισών διασπορών.

Η απλή ANOVA για εξαρτημένα δείγματα των δεδομένων του Πίνακα 15, μας δίνει ακριβώς τις ίδιες μέσες τιμές, τυπικές αποκλίσεις και τυπικά σφάλματα των μέσων τιμών (βλ. Πίνακα 12). Ωστόσο, παρατηρώντας των Πίνακα 16 με τα περιληπτικά αποτελέσματα της απλής ANOVA για εξαρτημένα δείγματα και συγκρίνοντάς τον με τα περιληπτικά αποτελέσματα του Πίνακα 14 (απλή ANOVA για ανεξάρτητα δείγματα), θα διαπιστώσουμε ότι το Άθροισμα Τετραγώνων της πηγής διασποράς Σφάλμα (2.83) του Πίνακα 14, έχει διαμεριστεί σε διασπορά Εντός των Υποκειμένων και Σφάλματος με τα αντίστοιχα Άθροισμα Τετραγώνων (.50 και 2.33, αντίστοιχα) στον Πίνακα 16. Ο διαμερισμός αυτός της διασποράς οδήγησε σε διαφορετικό Πηλίκιο $F_{2,4} = 58.43$ ($34.08/.47 = 58.43$) και φυσικά με διαφορετικούς βαθμούς ελευθερίας και διαφορετικά μέσα τετράγωνα για την πηγή διασποράς σφάλμα.

Για να έχει τη δυνατότητα να χρησιμοποιήσει τα αποτελέσματα της ANOVA, ο επαγγελματίας υγείας πρέπει να κάνει τους απαραίτητους ελέγχους εγκυρότητας των αποτελεσμάτων. Μία από τις πιο σημαντικές παραδοχές (προϋποθέσεις) της ANOVA με επαναλαμβανόμενες μετρήσεις είναι η παραδοχή της σφαιρικότητας (sphericity assumption) η οποία αναφέρεται στην ομοιογένεια της διασποράς και της συνδιασποράς των δεδομένων. Ο έλεγχος της ομοιογένειας της διασποράς και της συνδιασποράς γίνεται με τη μέθοδο ελέγχου της σφαιρικότητας του Mauchly. Στο παράδειγμά μας, δεν παρατηρήθηκε στατιστικά σημαντική διαφορά ($p > .05$) στον έλεγχο της σφαιρικότητας και αυτό συνεπάγεται ότι μπορούμε να χρησιμοποιήσουμε τα αποτελέσματα της ANOVA. Επιπλέον, όταν υπάρχει παραβίαση της σφαιρικότητας, τότε προχωρούμε στη διόρθωση των βαθμών ελευθερίας με τη μέθοδο των Geisser and Greenhouse (1958). Όλα τα στατιστικά πακέτα υπολογίζουν το Greenhouse-Geisser epsilon (ϵ), το οποίο πολλαπλασιαζόμενο με τους βαθμούς ελευθερίας των πηγών διασποράς που συνθέτουν το πηλίκιο F δημιουργούν ένα πιο συντηρητικό επίπεδο σημαντικότητας. Στο παράδειγμά μας, το Greenhouse-Geisser epsilon είναι ίσο με.50. Διορθώνοντας τώρα τα αποτελέσματα στον Πίνακα 16, δημιουργούμε τον Πίνακα 17 με τους βαθμούς ελευθερίας και το επίπεδο σημαντικότητας διορθωμένα. Όπως διαπιστώσαμε το πηλίκιο F παραμένει το ίδιο, αλλά το

επίπεδο σημαντικότητας έφτασε στο $p = .017$. Εάν επιλέξουμε ως επίπεδο στατιστικής σημαντικότητας το $p < .01$ και διορθώσουμε τους βαθμούς ελευθερίας, λόγω της παραβίασης της παραδοχής της σφαιρικότητας (πράγμα που δεν συμβαίνει στο παράδειγμά μας), τότε δεν μπορούμε να προχωρήσουμε στις post hoc αναλύσεις και δηλώνουμε την ύπαρξη μη-στατιστικά σημαντικής διαφοράς μεταξύ των μετρήσεων.

ΠΙΝΑΚΑΣ 15. Δεδομένα για μια υποθετική μελέτη επίδρασης φαρμάκου στην ανακούφιση από τον πόνο			
ID	Placebo	Ασπιρίνη	Φάρμακο X
1	4	5	10
2	3	5	11
3	4	4.5	9

ΠΙΝΑΚΑΣ 16. Περιληπτικά Αποτελέσματα της απλής ANOVA για εξαρτημένα δείγματα των τριών θεραπειών στην εξαρτημένη μεταβλητή ανακούφισης από το πόνο

Πηγή διασποράς	Άθροισμα τετραγώνων	df	Μέσα τετράγωνα	Πηλίκιο F	p
Ομάδες	68.17	2	34.08	72.17	.000
Μεταξύ Μετρήσεων	68.17	2	34.08	58.43	.001
Εντός των Υποκειμένων	.50	2	.25		
Σφάλμα	2.33	4	.58		
Ολική	71.00	8			

Οι περισσότεροι ερευνητές δεν αναφέρουν τις περισσότερες φορές τον έλεγχο της σφαιρικότητας και ούτε κάνουν τις απαραίτητες διορθώσεις στους βαθμούς ελευθερίας. Όταν χρησιμοποιούμε πάνω από τρεις επαναλαμβανόμενες μετρήσεις πρέπει πάντα να αναφέρουμε, εάν η παραδοχή της σφαιρικότητας έχει παραβιαστεί και εάν έχει γίνει ο έλεγχος της σημαντικότητας σύμφωνα με τη διόρθωση του Greenhouse-Geisser epsilon. Όταν δεν υπάρχει πρόβλημα στο επίπεδο σημαντικότητας μετά τη διόρθωση Greenhouse-Geisser, ή δεν υπάρχει παραβίαση της προϋπόθεσης της σφαιρικότητας, τότε μπορούμε να προχωρήσουμε στις *post hoc* αναλύσεις.*

Μπορούμε λοιπόν να περιγράψουμε τα αποτελέσματα της ANOVA με τον παρακάτω τρόπο. "Η απλή ANOVA για εξαρτημένα δείγματα έδειξε στατιστικά σημαντική διαφορά για το κύριο αποτέλεσμα των μετρήσεων ($F_{2,4} = 58.43, p < .001$). Η στατιστικά σημαντική διαφορά διατήρησε τη σημαντικότητά της και μετά τη διόρθωση των βαθμών ελευθερίας με τη μέθοδο των Greenhouse και Geisser ($p < .017$), λόγω παραβίασης της παραδοχής της σφαιρικότητας.** Η *post hoc* ανάλυση με τον έλεγχο Tukey για τον αριθμό των δυνατών συγκρίσεων, έδειξε ότι η θεραπεία με το Φάρμακο Χ ($M = 10.00 \pm 1.00$) ανακούφισε περισσότερο από τον πόνο σε σύγκριση με την Ασπιρίνη ($M = 4.83 \pm .29$) και τη θεραπεία Placebo ($M = 3.67 \pm .58$), ενώ η θεραπεία με Ασπιρίνη δεν διέφερε σημαντικά από τη θεραπεία Placebo στην εξαρτημένη μεταβλητή ανακούφισης από τον πόνο."

Σχετικά με τη χρησιμοποίηση των πολλαπλών ελέγχων *t*, ισχύει ότι και στην απλή ANOVA για ανεξάρτητα δείγματα. Επίσης, όταν οι επαγγελματίες υγείας χρησιμοποιούν περίπλοκα πειραματικά σχέδια (π.χ. ANOVA δύο κατευθύνσεων, τριών κατευθύνσεων για ανεξάρτητα και εξαρτημένα δείγματα ή μεικτά πειραματικά σχέδια δύο, τριών κατευθύνσεων, κτλ), πρέπει να γνωρίζουν ότι ακολουθούν παρόμοιες διαδικασίες. Ελέγχουν τα επίπεδα σημαντικότητας των κυρίων αποτελεσμάτων (main effects) και των αλληλεπιδράσεων (interactions) και προχωρούν στην επιλογή των καταλληλότερων *post hoc* αναλύσεων χρησιμοποιώντας ως μέσο υπολογισμού τα Μέσα Τετράγωνα του Σφάλματος (Mean Square error) που έδωσε τη στατιστικά σημαντική διαφορά στο πηλίκιο F. Για αναλυτική παρουσίαση του θέματος, ο αναγνώστης μπορεί να ανατρέξει σε βιβλία στατιστικής που ασχολούνται με ανάπτυξη προχωρημένων πειραματικών σχεδίων (π.χ. Kirk, 1995, Winer, 1971).

Παραγοντική ανάλυση (Factor Analysis)

Η παραγοντική ανάλυση (factor analysis) είναι μια στατιστική μέθοδος που μας επιτρέπει να εξηγήσουμε τη δομή και τη μεταβλητότητα - διασπορά που παρατηρείται μεταξύ των αρχικών μεταβλητών (εξαρτημένες μεταβλητές) και να δημιουργήσουμε νέες μεταβλητές που ονομάζονται

* Στο παράδειγμά μας, η *post hoc* ανάλυση έγινε με τη μέθοδο Tukey και τα αποτελέσματά της παρουσιάζονται στην παράγραφο περιγραφής των αποτελεσμάτων.

** Η πρόταση αυτή δεν πρέπει να προστεθεί όταν δεν έχουμε παραβίαση της παραδοχής της σφαιρικότητας. Στη συγκεκριμένη περίπτωση χρησιμοποιήθηκε για λόγους εξηγηματικούς και μόνο.

ΠΙΝΑΚΑΣ 17. Περιληπτικά Αποτελέσματα της απλής ANOVA για εξαρτημένα δείγματα των τριών θεραπειών στην ε-ξαρτημένη μεταβλητή ανακούφισης από το πόνο με τη διόρθωση του Greenhouse-Geisser epsilon ($\epsilon = .50$)

Πηγή διασποράς	Άθροισμα τετραγώνων	df	Μέσα τετράγωνα	Πηλίκιο F	p
Μεταξύ μετρήσεων Διόρθωση ϵ	68.17	22*(.50) = 1	34.0868.17	58.4358.43	.001.017
Εντός των υποκειμένων	.50	2	.25		
Σφάλμα Διόρθωση ϵ	2.33	44*(.50) = 2	.581.17		
Ολική	71.00	8			

παράγοντες (factors). Οι πληροφορίες που αποκομίζουμε από τις στήλη-εξαρτήσεις των μεταβλητών χρησιμοποιούνται, για να μειώσουμε τον αρχικό αριθμό των μεταβλητών σε ένα μικρότερο αριθμό παραγόντων (Καρλής, 2005).

Η παραγοντική ανάλυση χρησιμοποιείται κυρίως, όταν προσπαθούμε να καταλάβουμε το πρότυπο των απαντήσεων των ανθρώπων, όταν συμμετέχουν σε έρευνες που χρησιμοποιούν ως μέτρο αξιολόγησης, κυρίως, ερωτηματολόγια κλειστού-τύπου (closed-ended questionnaires).^{*} Οι ερωτήσεις που αξιολογούν παρόμοιες έννοιες ή χαρακτηριστικά, μπορούν να προσδιοριστούν με την παραγοντική ανάλυση και να δημιουργηθούν οι νέοι παράγοντες. Ο αναγνώστης πρέπει να γνωρίζει ότι οι γνώσεις που απαιτούνται για να κατανοήσει κάποιος την παραγοντική ανάλυση είναι ο συντελεστής συσχέτισης και ο πίνακας συντελεστών συσχέτισης (correlation matrix) μεταξύ των μεταβλητών (Howitt & Cramer, 2001).

Ας υποθέσουμε ότι έχουμε ένα ερωτηματολόγιο κλειστού-τύπου με 10 ερωτήσεις στην κλίμακα Likert (διαφωνώ απόλυτα = 1, διαφωνώ = 2, ουδέτερο = 3, συμφωνώ = 4, συμφωνώ απόλυτα = 5). Η ανάλυση της κάθε ερώτησης ξεχωριστά δεν μας παρέχει χρήσιμες πληροφορίες. Εφόσον, οι ερωτήσεις γίνονται στα ίδια άτομα, επόμενο είναι να θέλουμε να μάθουμε εάν υπάρχει συσχέτιση μεταξύ των ερωτήσεων. Υπολογίζοντας τις συσχετίσεις μεταξύ των 10 ερωτήσεων (μεταβλητών) θα διαπιστώσουμε ότι θα δημιουργήσουμε έναν πίνακα συντελεστών συσχέτισης με 100 συσχετίσεις (10^2). Ο πίνακας των συντελεστών συσχέτισης είναι συμμετρικός και στη διαγώνιο από την πάνω αριστερή προς την κάτω δεξιά γωνία υπάρχουν οι συσχετίσεις της κάθε μεταβλητής με τον εαυτό της ($r = 1.00$). Οι υπολειπόμενες 45 διαφορετικές συσχετίσεις είναι οι συσχετίσεις που μας αφορούν και εξετάζουμε (για παρόμοιο παράδειγμα βλ. Πίνακα 18).

Ας χρησιμοποιήσουμε ένα υποθετικό παράδειγμα, όπου 10 άτομα αξιολογήθηκαν σε 6 μεταβλητές. Το παράδειγμα χρησιμοποιείται για τους σκοπούς του άρθρου. Ο αναγνώστης πρέπει να γνωρίζει ότι για να μπορέσει να εκτελέσει μια έγκυρη παραγοντική ανάλυση θα πρέπει να έχει αρκετά μεγαλύτερο αριθμό ατόμων. Για να γίνουν σωστά οι επεξεργασίες του μοντέλου, ο ερευνητής πρέπει να έχει στο μυαλό του την αντιστοιχία των 10 ατόμων ανά ερώτηση ή μέτρηση. Στο παράδειγμά μας, θα έπρεπε να έχουμε τουλάχιστον 60 άτομα.

Η παραγοντική ανάλυση στο συγκεκριμένο παράδειγμα έγινε με Ανάλυση Κύριων Συνιστωσών (Principal Components Analysis) και την ορθογώνια περιστροφή Varimax. Για τη δημιουργία των νέων παραγόντων λάβαμε υπόψη τις ιδιοτιμές (eigenvalues) που είχαν τιμή μεγαλύτερη από 1.

Για να καταλάβουμε την παραγοντική ανάλυση πρέπει να αρχίσουμε με τον πίνακα των συντελεστών συσχέτισης (Πίνακας 18). Θα διαπιστώσουμε ότι υπάρχουν δύο διαφορετικές ομάδες μεταβλητών. Οι μεταβλητές Α, Γ και Ε συσχετίζονται μεταξύ τους (υπογραμμισμένες συσχετίσεις, π.χ.

^{*} Σε καμία περίπτωση, ο αναγνώστης δεν πρέπει να θεωρήσει ότι η παραγοντική ανάλυση γίνεται μόνο σε ερωτηματολόγια.

$r = .868$) και οι μεταβλητές Β, Δ, και ΣΤ συσχετίζονται επίσης μεταξύ τους (πλήγιες συσχετίσεις, π.χ. $r = .760$). Μπορούμε να παρατηρήσουμε επίσης ότι οι μεταβλητές Α, Γ, και Ε δεν επιτυγχάνουν υψηλές συσχετίσεις με τις μεταβλητές Β, Δ, και ΣΤ και φαίνεται να αποτελούν διαφορετικές ομάδες (ή συστάδες) μεταβλητών.

ΠΙΝΑΚΑΣ 18. Πίνακας συντελεστών συσχέτισης μεταξύ των μεταβλητών Α, Β, Γ, Δ, Ε, και ΣΤ

		A	B	Γ	Δ	Ε	ΣΤ
Συσχέτιση	A	1.000	-.200	<u>.868</u>	-.371	<u>.894</u>	-.180
	B	-.200	1.000	-.161	.760	-.154	<u>.731</u>
	Γ	.868	-.161	1.000	-.359	<u>.772</u>	-.075
	Δ	-.371	.760	-.359	1.000	-.225	<u>.779</u>
	Ε	.894	-.154	.772	-.225	1.000	-.188
	ΣΤ	-.180	.731	-.075	.779	-.188	1.000

Εφαρμόζοντας την παραγοντική ανάλυση, μειώνουμε τον πίνακα των συντελεστών συσχέτισης που περιέχει όλες τις μεταβλητές σε ένα μικρότερο αριθμό παραγόντων.

Η ανάλυση σε Κύριες Συνιστώσες (Principal Components Analysis)* έδωσε δύο παράγοντες. Ο Πίνακας 19 περιέχει τις ιδιοτιμές (eigenvalues) και το ποσοστό της διασποράς ή διακύμανσης που κάθε ιδιοτιμή ερμηνεύει. Από το παράδειγμά μας διαφαίνεται ότι οι δύο παράγοντες εξηγούν το 87.12% της διασποράς των μεταβλητών.

Στη συνέχεια η παραγοντική ανάλυση μας παρέχει έναν Πίνακα με τις επιβαρύνσεις των παραγόντων (Component Matrix). Στον Πίνακα 20 παρουσιάζονται οι επιβαρύνσεις των παραγόντων που προκύπτουν για το μοντέλο με δύο παράγοντες. Ο ερευνητής μπορεί να δει πως εκφράζεται κάθε μεταβλητή με τη χρήση των 2 παραγόντων. Παρατηρώντας τις σχέσεις των μεταβλητών και παραγόντων διαπιστώνουμε ότι για τον πρώτο παράγοντα οι μεταβλητές Γ, Ε, και ΣΤ έχουν αρνητικά πρόσημα, ενώ οι υπόλοιπες έχουν θετικά πρόσημα.

Συνδυάζοντας τα ευρήματα αυτά με τον Πίνακα 21, μπορούμε να δούμε τις εταιρικότητες (communalities) ή τις διασπορές που εξηγούν οι παράγοντες που προσαρμόσαμε. Παρατηρούμε ότι οι αριθμοί κυμαίνονται από το 0 έως το 1 και αναφέρονται στο ποσοστό διασποράς της κάθε μεταβλητής που εξηγείται από τον αριθμό των παραγόντων που προσαρμόστηκαν.** Παρατηρούμε επομένως ότι με το μοντέλο που εφαρμόσαμε ερμηνεύουμε τα 95% του παράγοντα Α. Το ποσοστό αυτό προκύπτει από το άθροισμα των τετραγώνων των επιβαρύνσεων των παραγόντων στη συγκεκριμένη μεταβλητή. Χρησιμοποιώντας τον Πίνακα 20 μπορούμε να επιβεβαιώσουμε τα αποτελέσματα της στήλης Εκτίμηση (Extraction) του Πίνακα 21. Για τη μεταβλητή Α $(.831)^2 + (.509)^2 = 0.949$, για τη μεταβλητή Β, $(.769)^2 + (.529)^2 = 0.825$, κοκ.

Αφαιρώντας τη δεύτερη στήλη από τη μονάδα (Πίνακας 21) έχουμε το ποσοστό εκείνο της διασποράς που δεν μπορεί να εξηγήσει το παραγοντικό μοντέλο για τη συγκεκριμένη μεταβλητή. Για τη μεταβλητή Α έχουμε $.051$ ($1.00 - .949 = .051$).

Από την μέχρι τώρα ανάλυση αξιολογήσαμε (α) τις συσχετίσεις που υπάρχουν μεταξύ των μεταβλητών, (β) τους 2

ΠΙΝΑΚΑΣ 19. Ιδιοτιμές του Πίνακα συσχέτισης των δεδομένων και ποσοστά διασποράς των ιδιοτιμών

Παράγοντες	Initial Eigenvalues		
	Total	% of Variance	Cumulative%
1	3.258	54.303	54.303
2	1.969	32.814	87.117
3	.343	5.718	92.836
4	.263	4.382	97.218
5	.106	1.775	98.993
6	.060	1.007	100.000

* Υπάρχουν και άλλες μέθοδοι εκτίμησης των παραγόντων, όπως η μέθοδος των ελάχιστων τετραγώνων, η γενικευμένη μέθοδος ελάχιστων τετραγώνων, και η μέθοδος των κύριων αξόνων (principal axis model).

** Όταν χρησιμοποιούμε τη μέθοδο των κύριων συνιστωσών η πρώτη στήλη (initial) έχει την τιμή 1. Χρησιμοποιώντας τις άλλες μεθόδους θα διαπιστώσουμε ότι στην πρώτη στήλη θα υπάρχουν τιμές που θα κυμαίνονται από το 0 έως το 1.

ΠΙΝΑΚΑΣ 20. Επιβαρύνσεις επιλεγμένων παραγόντων

Component Matrix ^a		
	Component	
	1	2
A	.831	.509
B	.769	.529
Γ	-.767	.529
Δ	.764	.536
E	-.627	.671
ΣΤ	-.642	.643

Extraction Method: Principal Component Analysis.
a 2 components extracted.

ΠΙΝΑΚΑΣ 21. Οι εταιρικότητες των μεταβλητών για το παραγοντικό μας μοντέλο

Communalities		
Μεταβλητές	Initial	Extraction
A	1.000	.949
B	1.000	.825
Γ	1.000	.871
Δ	1.000	.868
E	1.000	.871
ΣΤ	1.000	.843

Extraction Method: Principal Component Analysis

νέους παράγοντες που δημιουργήθηκαν με την μέθοδο εκτίμησης των Κύριων Συνιστωσών, και (γ) τις διασπορές των μεταβλητών που εξηγούν (και δεν μπορούν να εξηγήσουν) οι δύο παράγοντες.

Για να γίνουν πιο ερμηνεύσιμοι οι παράγοντες πρέπει να προχωρήσουμε στην περιστροφή (rotation) των παραγόντων. Με την περιστροφή δεν αλλιάζουν η προσαρμοστικότητα και το ποσοστό της διασποράς (και συνδιασποράς) που ερμηνεύει το μοντέλο. Αλλιάζουν μόνο οι τιμές των επιβαρύνσεων (loadings). Οι βασικές μέθοδοι περιστροφής είναι η Varimax, η Quartimax, η Equimax, και η Obllique (μη ορθογώνια περιστροφή). Για το παράδειγμά μας, θα χρησιμοποιήσουμε τη μέθοδο περιστροφής Varimax, η οποία προσπαθεί να ελαχιστοποιήσει τον αριθμό των μεταβλητών που έχουν μεγάλες επιβαρύνσεις για κάθε παράγοντα (Πίνακας 22).

Παρατηρούμε ότι οι μεταβλητές Α μέχρι ΣΤ τοποθετούνται με σειρά σύμφωνα με την τιμή της επιβάρυνσης (από τη μεγαλύτερη προς τη μικρότερη). Η τοποθέτηση αυτή είναι χρήσιμη, γιατί μας βοηθά να αποφασίσουμε πως θα ονομάσουμε τον κάθε παράγοντα. Στον πρώτο παράγοντα, η μεγαλύτερη επιβάρυνση ανήκει στη μεταβλητή Α (.962), η αμέσως επόμενη στη μεταβλητή Ε (.929) και Γ (.928). Στο δεύτερο παράγοντα, τη μεγαλύτερη επιβάρυνση έχει η μεταβλητή ΣΤ (.918) και ακολουθούν η Β (.905) και η Δ (.900). Η σειρά και η τιμή της επιβάρυνσης μαζί με το εννοιολογικό περιεχόμενο των μεταβλητών συμβάλλουν στον ονομαστικό προσδιορισμό του κάθε παράγοντα.

Όπως αναφέραμε και προηγουμένως εκείνο που αλλιάζει με την περιστροφή είναι οι τιμές των επιβαρύνσεων. Παρατηρείται, ωστόσο και μια "εσωτερική" διαφοροποίηση στα ποσοστά της διασποράς των δύο παραγόντων μετά από την περιστροφή Varimax, χωρίς να υπάρχει καμιά μεταβολή στην ολική επεξηγούμενη διασπορά (Πίνακας 23). Το ποσοστό της διασποράς για τον καθένα παράγοντα μετά την περιστροφή Varimax, αλλιάζει για τον παράγοντα 1, από 54.303 σε 45.228 και για τον παράγοντα 2, από 32.814 σε 41.889. Ωστόσο, το αθροιστικό ποσοστό διασποράς (cumulative%) των δύο παραγόντων παραμένει το ίδιο (87.117). Η διαφοροποίηση αυτή δεν επηρεάζει την εγκυ-

ΠΙΝΑΚΑΣ 22. Επιβαρύνσεις των παραγόντων με την περιστροφή Varimax

Rotated Component Matrix ^a		
Μεταβλητές	Παράγοντες	
	1	2
A	.962	-.153
E	.929	-.089
Γ	.928	-.098
ΣΤ	-.041	.918
B	-.070	.905
Δ	-.239	.900

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization
^a Rotation converged in 3 iterations

ρότητα του μοντέλου και οφείλεται στη δημιουργία της νέας δομής των παραγόντων μετά την ορθογώνια περιστροφή.

Ο ερευνητής θα μπορούσε να αναφέρει τα αποτελέσματα της παραγοντικής ανάλυσης ως εξής: "Η παραγοντική ανάλυση έγινε με τη μέθοδο

των Κύριων Συνιστωσών στις συσχετίσεις των 6 μεταβλητών. Από την ανάλυση προέκυψαν δύο παράγοντες που οι ιδιοτιμές τους ήταν πάνω από 1. Η ορθογώνια περιστροφή Varimax δημιούργησε τη δομή των παραγόντων που παρουσιάζεται στον Πίνακα 24. Ο πρώτος παράγοντας εξηγεί το 45.23% και ο δεύτερος το 41.89% της διασποράς των αρχικών μεταβλητών. Ο πρώτος παράγοντας φαίνεται να αξιολογεί την ιδιότητα Χ και ο δεύτερος την Υ.*

ΣΥΜΠΕΡΑΣΜΑΤΑ - ΓΕΝΙΚΕΣ ΕΠΙΣΗΜΑΝΣΕΙΣ

Το κύριο μήλημα αυτής της εργασίας είναι η επεξήγηση τρόπων και μέσων, που υπηρετούν με τον καλύτερο τρόπο τη στατιστική ανάλυση των δεδομένων που προέρχονται από το χώρο της υγείας. Η επιθετική επικέντρωση σε συγκεκριμένες στατιστικές μεθόδους και τεχνικές αποβλήπει στην αποτελεσματικότερη συλλογή, ανάλυση και ερμηνεία των δεδομένων. Η κατανόηση και η σωστή παρουσίαση των βασικών εννοιών και δεικτών της περιγραφικής στατιστικής (μέτρα κεντρικής τάσης και διασποράς) αποτελούν προϋπόθεση και γνωστικό προαπαιτούμενο για την κατανόηση των απλών και σύνθετων στατιστικών αναλύσεων. Η παρουσίαση και ανάπτυξη των ενδεικτικών στατιστικών αναλύσεων της μη παραμετρικής στατιστικής (έλεγχος χ^2 και συντελεστής συσχέτισης Spearman) και της συσχέτισης μεταξύ δύο συνεχών μεταβλητών (συντελεστής συσχέτισης Pearson) προδικάζει τα ερμηνευτικά όρια των στατιστικών ευρημάτων και προστατεύει τον επαγγελματία υγείας από επεξηγηματικούς προσδιορισμούς αιτίας και αποτελέσματος.

Ο προσεκτικός σχεδιασμός της ερευνητικής διαδικασίας, η συλλογή των δεδομένων σε κλίμακα διαστήματος ή αναλογίας και η επιλογή της αρμόζουσας στατιστικής ανάλυσης προδιαγράφουν την αναμενόμενη (τις περισσότερες φορές) ερμηνεία των στατιστικών αποτελεσμάτων. Ο έλεγχος της ομοιογένειας των διασπορών και των συνδιασπορών, η επισήμανση των στατιστικά σημαντικών ευρημάτων και η επιλογή των κατάλληλων μετά-ANOVA (post hoc) αναλύσεων διασφαλίζουν την εγκυρότητα των αποτελεσμάτων και επιτρέπουν την τεκμηρίωση της σχέσης "αιτίας και αιτιατού".

Η δημιουργία κοινών παραγόντων ανάμεσα σε

ΠΙΝΑΚΑΣ 23. Ιδιοτιμές του Πίνακα συσχέτισης των δεδομένων και ποσοστά διασποράς των ιδιοτιμών πριν και μετά την περιστροφή

Παράγοντες	Initial Eigenvalues			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3.258	54.303	54.303	2.714	45.228	45.228
2	1.969	32.814	87.117	2.513	41.889	87.117
3	.343	5.718	92.836			
4	.263	4.382	97.218			
5	.106	1.775	98.993			
6	.060	1.007	100.00			

ΠΙΝΑΚΑΣ 24. Επιβαρύνσεις παραγόντων μετά την ορθογώνια περιστροφή των 6 παραγόντων

Μεταβλητές	Παράγοντας 1	Παράγοντας 2
A	.962	-.153
B	-.070	.905
Γ	.928	-.098
Δ	-.239	.900
E	.929	-.089
ΣΤ	-.041	.918

* Στο σημείο αυτό ο ερευνητής θα πρέπει να ονομάσει τους παράγοντες 1 και 2.

μια ομάδα μεταβλητών με τη στατιστική μέθοδο της παραγοντικής ανάλυσης απαιτεί ιδιαίτερη προσοχή. Για τους σκοπούς του άρθρου, παρουσιάστηκε “περιληπτικά” η μέθοδος εκτίμησης των Κύριων Συνιστωσών με περιστροφή Varimax. Ωστόσο, υπάρχουν και άλλες μέθοδοι εκτίμησης, όπως της Μέγιστης Πιθανοφάνειας, των Ελάχιστων Τετραγώνων και των Κύριων Αξόνων. Διαφορετικές μέθοδοι εκτίμησης ή ακόμα και η ίδια μέθοδος εκτίμησης μπορούν να δώσουν ένα μεγάλο αριθμό ισοδύναμων εκτιμήσεων, οι οποίες επηρεάζονται και από τον τρόπο περιστροφής των μεταβλητών (π.χ. Quartimax, ή Equimax ή Oblique). Ο αριθμός των νέων παραγόντων εξαρτάται, κυρίως, από τη μέθοδο εκτίμησης, η οποία επιτρέπει στον ερευνητή να “μεροληπτεί” προς τη θεμιτή κατεύθυνση των αποτελεσμάτων. Ωστόσο, η παραγοντική ανάλυση αποτελεί πρόκληση για κάθε επιστήμονα, η οποία συνδυαζόμενη (α) με αξιόπιστους και έγκυρους τρόπους μέτρησης των μεταβλητών, και (β) την ενδελεχή αιτιολόγηση και εφαρμογή των θεωριών και ερευνητικών ευρημάτων στο χώρο της επιστήμης του, μπορεί να γίνει αφετηρία δημιουργίας σημαντικών μοντέλων πρόβλεψης και επεξήγησης συμπεριφορών στο χώρο της υγείας.

ΕΠΙΛΕΓΜΕΝΗ ΒΙΒΛΙΟΓΡΑΦΙΑ

1. American Psychological Association (2003). *Publication Manual of the American Psychological Association* (5th ed.). Washington, DC: American Psychological Association.
2. Βαγενάς, Γ. Κ. (2003). *Στατιστικές εφαρμογές στην αθλητική επιστήμη*. Αθήνα, Αττικής: αυτοέκδοση.
3. Bland, M. (2000). *An introduction to medical statistics*. Oxford, UK: Oxford University Press.
4. Coolican, H. (1996). *Introduction to research methods and statistics in psychology* (2nd ed.). London, UK: Hodder & Stoughton.
5. Geisser, S., & Greenhouse, S. W. (1958). An extension of Box's results on the use of the F distribution in multivariate analysis. *Annals of Mathematical Statistics*, 29, 885-891.
6. Glasnapp, D. R., & Poggio, J. P. (1985). *Essentials of statistical analysis for the behavioral sciences*. Columbus, OH: Merrill.
7. Heiman, G. W. (2001). *Understanding research methods and statistics* (2nd ed.). Boston, NY: Houghton Mifflin Company.
8. Howitt, D., & Cramer, D. (2001). *A guide to computing statistics with SPSS for windows*. Essex, UK: Pearson Education.
9. Καρλής, Δ. (2005). *Πολυμεταβλητή στατιστική Ανάλυση*. Αθήνα, Αττικής: Σταμούλης.
10. Kirk, R. E. (1995). *Experimental Design: Procedures for the behavioral sciences* (3rd ed.). Belmont, CA: Brooks/Cole.
11. Lockhart, R.S. (1998). *Introduction to statistics and data analysis for the behavioral sciences*. New York, NY: W. H. Freeman.
12. Pagano, M., & Gauvreau, K. (2002). *Αρχές βιοστατιστικής*. Αθήνα, Αττικής: Έλλην.
13. Scheffe, H. (1953). A Method for judging all contrasts in the analysis of variance. *Biometrika*, 40, 87-104.
14. Winer, B. J. (1971). *Statistical principles in experimental design*. New York, NY: McGraw-Hill.