

ΕΙΣΑΓΩΓΗ

«Μια κατά προσέγγιση απάντηση στο σωστό πρόβλημα αξίζει πολύ περισσότερο απ' ό τι μια ακριβής απάντηση σε ένα κατά προσέγγιση πρόβλημα».

John Tukey, διάσημος μαθηματικός

Το βιβλίο αυτό γράφτηκε με παραδείγματα από τον αθλητισμό και τις επιστήμες υγείας, πράγμα που βοηθάει πρωτίστως τους φοιτητές και ερευνητές. Πιθανόν, όμως, κι άλλοι νεοεισερχόμενοι στη στατιστική να ωφεληθούν στην προσπάθειά τους να την κατανοήσουν μέσα από αυτές τις εφαρμογές. Όλοι έχουμε εικόνες από τον αθλητισμό, οι οποίες εμπεριέχουν οικείες σε όλους μας μετρήσεις, όπως για παράδειγμα άλμα 7,5 μέτρων ή 85 πόντους στο αγώνα μπάσκετ ή κέρδισε-έχασε τον αγώνα, κ.ά. Οι οικείες εικόνες με τα στατιστικά στον αθλητισμό μπορούν να βοηθήσουν τον καθένα να καταλάβει πως κάποια μέτρηση μπορεί να πάρει τη μορφή μεταβλητής και, στη συνέχεια, πως μια μέτρηση-μεταβλητή μπορεί να συσχετισθεί με άλλες μετρήσεις-μεταβλητές που επίσης συναντάμε στον αθλητισμό. Όπως εξηγείται σ' αυτό το βιβλίο, ανάλογα με το είδος μετρήσεων-μεταβλητών χρησιμοποιούμε και τις ανάλογες στατιστικές αναλύσεις. Μ' αυτόν τον τρόπο, μπορεί κανείς πολύ εύκολα να καταλάβει ποιός είναι ο σκοπός της κάθε στατιστικής ανάλυσης και πώς να ερμηνεύει αποτελέσματα από προσφιλή δεδομένα, όπως τα σκορ και οι μετρήσεις στον αθλητισμό.

Για τους ερευνητές εφαρμοσμένων επιστημών, όπως είναι οι αθλητικές επιστήμες, η στατιστική είναι κάτι σαν το όχημα για τον οδηγό του. Χρειάζεται να ξέρει πώς να το οδηγεί σωστά για να φτάσει με ασφάλεια στον προ-



2 ΕΦΑΡΜΟΓΕΣ ΤΗΣ ΣΤΑΤΙΣΤΙΚΗΣ ΣΤΙΣ ΕΠΙΣΤΗΜΕΣ ΤΟΥ ΑΘΛΗΤΙΣΜΟΥ

ορισμό του. Αντίστοιχα, οι ερευνητές εφαρμοσμένων επιστημών χρειάζεται να γνωρίζουν ποιες στατιστικές αναλύσεις είναι κατάλληλες για την επίλυση ερευνητικών προβλημάτων του κλάδου τους, ή, όπως αναφέρεται σε επόμενο κεφάλαιο, για τη διερεύνηση συγκεκριμένων ερευνητικών υποθέσεων στον επιστημονικό τους τομέα, πώς να χρησιμοποιούν τις στατιστικές αναλύσεις για να διερευνήσουν τις ερευνητικές τους υποθέσεις και, τέλος, πώς να ερμηνεύουν και να περιγράψουν τα αποτελέσματα των αναλύσεων. Ο οδηγός ενός οχήματος δεν είναι ούτε ο κατασκευαστής του ούτε καν ο μηχανικός που θα το επιδιορθώσει σε περίπτωση βλάβης. Αντίστοιχα, ο ερευνητής στις αθλητικές επιστήμες δεν καλείται να κατασκευάσει μια στατιστική ανάλυση για να επιλύσει ένα πρόβλημα που έχει, απλώς διαλέγει την κατάλληλη ανάλυση για το σκοπό του και τη χρησιμοποιεί με τον τρόπο που περιγράφει στον οδηγό (manual) ο κατασκευαστής, δηλαδή, ο επιστήμονας της στατιστικής.

Βεβαίως, όσο καλύτερα ο οδηγός γνωρίζει το πώς δουλεύει η μηχανή ή το σύστημα πέδησης και μάλιστα σε διαφορετικές συνθήκες οδήγησης, τόσο καλύτερα θα ανταποκριθεί στην οδήγηση. Αντίστοιχα, όσο βαθύτερη γνώση στατιστικής και συγκεκριμένων στατιστικών αναλύσεων έχει κάποιος, τόσο καλύτερα θα ανταποκριθεί στην επίλυση ερευνητικών προβλημάτων. Ωστόσο, αυτό δε σημαίνει ότι κάποιος χρειάζεται να αναπτύξει κορυφαίου επιπέδου γνώσεων στατιστικής για να ξεκινήσει να τη χρησιμοποιεί στην έρευνά του. Όπως η οδήγηση, έτσι και η στατιστική μαθαίνονται σιγά-σιγά μέσα από την πράξη. Σε σχέση με έμπειρους οδηγούς κάποιος που έχει μόλις πάρει δίπλωμα οδήγησης, κινδυνεύει περισσότερο να κάνει λάθος και να υποπέσει σε ατύχημα. Ωστόσο αν δεν τολμήσει να οδηγήσει, δε θα μάθει ποτέ. Αν ο ερευνητής εφαρμοσμένων επιστημών δεν προσπαθήσει ποτέ να κάνει στατιστικές αναλύσεις για να επιλύσει ερευνητικά προβλήματα, δε θα μάθει ποτέ πώς να χρησιμοποιεί την στατιστική.

Άρα το πρώτο που χρειάζεται για να μάθει κάποιος στατιστική, είναι να έχει ένα ερευνητικό πρόβλημα που θέλει να επιλύσει. Προβλήματα έχουν όλοι και παντού. Σε σχέση με κάποιον εμπειρικό, ο επιστήμονας μπορεί να διατυπώσει κάποιο πρόβλημα με τέτοιο τρόπο ώστε να μπορέσει να το διερευνήσει με επιστημονική μεθοδολογία και, στη συνέχεια, να αναλύσει τα δεδομένα που θα συλλέξει με την κατάλληλη στατιστική ανάλυση. Για το λόγο αυτό, στο παρόν βιβλίο επιλέξαμε να ξεκινήσουμε με παραδείγματα ερευνητικών προβλημάτων που συναντώνται στο χώρο των αθλητικών επιστημών και να τα διατυπώσουμε με τέτοιο τρόπο ώστε στη συνέχεια να επιλέξουμε την κατάλληλη στατιστική ανάλυση για την επίλυσή τους.



Η στατιστική δεν είναι εντελώς ανεξάρτητη από την ερευνητική μεθοδολογία. Όσο καλύτερες γνώσεις στατιστικής έχει κάποιος επιστήμονας, τόσο πιο πιθανό είναι να διατυπώσει με αποτελεσματικό τρόπο τις ερευνητικές του υποθέσεις και να σχεδιάσει καλύτερα την ερευνητική του μεθοδολογία. Όπως εξηγείται παρακάτω, οι στατιστικές αναλύσεις που επιλέγουμε εξαρτώνται από το είδος των μεταβλητών που θα χρησιμοποιηθούν. Επειδή το είδος των μεταβλητών αντανακλά την κλίμακα μέτρησης, κάποιος που γνωρίζει ποιες στατιστικές αναλύσεις είναι κατάλληλες για να αναλυθούν δεδομένα που συλλέχθηκαν με διαφορετικές κλίμακες μέτρησης, οι οποίες χρησιμοποιούνται στην επιστήμη του, έχει και την ευχέρεια να σχεδιάσει την έρευνά του επιλέγοντας τις καταλληλότερες κλίμακες μέτρησης. Αυτός είναι άλλος ένας πολύ σημαντικός λόγος για τον οποίο επιλέξαμε να διδάξουμε τη στατιστική με τη χρήση ερευνητικών προβλημάτων που συναντώνται στον αθλητισμό. Ως παραδείγματα χρησιμοποιούμε οικείες εικόνες από το χώρο του αθλητισμού, εξηγώντας τον τρόπο με τον οποίο τις εικόνες αυτές τις αποδίδουμε ως μεταβλητές. Το κάθε κεφάλαιο ξεκινάει με ένα παράδειγμα που διευκολύνει τον αναγνώστη να καταλάβει ποιος είναι ο σκοπός του κεφαλαίου. Για το παρόν κεφάλαιο της εισαγωγής επιλέξαμε το παρακάτω παράδειγμα.

Παράδειγμα. Δύο μεταπτυχιακοί φοιτητές συναντούν στη βιβλιογραφία μια στατιστική ανάλυση που δεν έχουν διδαχθεί. Συζητώντας μεταξύ τους συμφωνούν ότι η συγκεκριμένη ανάλυση φαίνεται κατάλληλη για την ανάλυση των δεδομένων τους, ωστόσο τους φαίνεται εντελώς άγνωστη και ιδιαίτερα περίπλοκη. Ο πρώτος φοιτητής αποφασίζει να ζητήσει βοήθεια από ένα φίλο του ειδικό στη στατιστική. Ο φίλος του κάνει τις αναλύσεις και στη συνέχεια του εξηγεί τί σημαίνουν τα δεδομένα και πώς πρέπει να περιγραφούν. Σε εύλογο διάστημα ο φοιτητής γράφει ένα άψογο κείμενο περιγραφής των αποτελεσμάτων του. Ο δεύτερος φοιτητής αποφασίζει να κάνει μόνος του την ανάλυση διαβάζοντας πληροφορίες σε σχετικά βιβλία και στο internet, προμηθεύεται ένα στατιστικό πρόγραμμα και μελετά το manual του στατιστικού προγράμματος. Μετά από πολλές δοκιμές και ώρες δουλειάς, γράφει τα αποτελέσματα που θεωρεί ότι είναι κατάλληλα και τα δίνει σ' έναν ειδικό στη στατιστική να του πει αν είναι σωστά. Αυτός του λέει ότι αρκετά απ' αυτά που έγραψε έχουν λογική αλλά του επισημαίνει ότι θα έπρεπε προηγουμένως να είχε εξετάσει κάποιες προϋποθέσεις που δεν εξέτασε, να είχε κάνει κάποια



τροποποίηση σε κάποια απ' τις μεταβλητές του –πράγμα που δεν έκανε–, και να είχε χρησιμοποιήσει και κάποια επιπλέον ανάλυση που ομοίως δεν την έκανε. Κατά τη γνώμη σας ποιος από τους δύο παραπάνω φοιτητές είναι πιο αποτελεσματικός; Τεκμηριώστε την άποψή σας προτού δείτε την απάντηση στο τέλος του κεφαλαίου¹.

Μετά από κάθε παράδειγμα που χρησιμοποιείται ως εισαγωγή σε κάθε κεφάλαιο, το οποίο εστιάζει σε κάποια στατιστική ανάλυση, ακολουθεί ο σκοπός της στατιστικής ανάλυσης που περιγράφεται στο συγκεκριμένο κεφάλαιο, τα χαρακτηριστικά των μεταβλητών που χρησιμοποιούνται στην συγκεκριμένη στατιστική ανάλυση, καθώς και οι προϋποθέσεις για να εκτελεσθεί η ανάλυση. Συγκεκριμένα, ποια είναι τα σημεία κλειδιά που πρέπει να προσέξει ο αναγνώστης ιδιαίτερα όταν προσπαθεί να διαβάσει τα αποτελέσματα του SPSS και τί ακριβώς σημαίνουν αυτά τα σημεία-κλειδιά. Στη συνέχεια, ακολουθεί ένα παράδειγμα ερευνητικής υπόθεσης στον αθλητισμό με μεταβλητές και δεδομένα τα οποία χρησιμοποιούνται για να εκτελεσθεί η στατιστική ανάλυση με τη χρήση του SPSS-PASW. Ακολούθως, ο αναγνώστης έχει τη δυνατότητα να διαβάσει και να κατανοήσει τον Πίνακα των αποτελεσμάτων που προκύπτουν από το SPSS-PASW. Στο τέλος περιγράφονται τα αποτελέσματα με τον τρόπο που θα περιγραφόταν σε μία επιστημονική διατριβή ή σε ένα επιστημονικό άρθρο. Η παρουσίαση των αριθμών και συμβόλων στη συγγραφή των αποτελεσμάτων γίνεται με βάση το στιλ συγγραφής που προτείνεται από την American Psychological Association (2010) που χρησιμοποιείται ευρέως στις επιστήμες της συμπεριφοράς. Για να κατανοήσει ο αναγνώστης τη λειτουργία της στατιστικής ανάλυσης, στις πιο απλές αναλύσεις η περιγραφή της ανάλυσης γίνεται βήμα-βήμα και χωρίς τη χρήση SPSS, δηλαδή με τον παραδοσιακό τρόπο όπως θα γινόταν πριν ανακαλυφθούν τα προγράμματα στατιστικής. Στις πιο πολύπλοκες αναλύσεις, όπως οι πολυμεταβλητές αναλύσεις, δεν έχει ιδιαίτερο νόημα για τον συνηθισμένο αναγνώστη ή ερευνητή να παρακολουθήσει μια πολυσέλιδη ανάλυση με το χέρι αφού μπορεί πολύ πιο εύκολα να γίνει με κάποιο στατιστικό πρόγραμμα. Αναγνώστες με καλές γνώσεις μαθηματικών που θέλουν να πάρουν μια μικρή γεύση απ' αυτού του είδους τις αναλύσεις μπορούν να συμβουλευτούν βιβλία όπως των Tabachnick και Fidell (2012).

Το κάθε κεφάλαιο κλείνει με ερωτήσεις προς τους αναγνώστες που τους βοηθούν να μελετήσουν ξανά τα κύρια σημεία του κεφαλαίου, με προτάσεις επιστημονικών άρθρων όπου χρησιμοποιείται η συγκεκριμένη στατιστική ανάλυση και με προτεινόμενες ασκήσεις για επίλυση, οι οποίες συνοδεύονται



από πραγματικά δεδομένα τα οποία δίνονται σε CD που συνοδεύει το βιβλίο. Το CD περιλαμβάνει επίσης και το αρχείο δεδομένων που χρησιμοποιείται στα παραδείγματα του συγκεκριμένου κεφαλαίου που περιγράφει κάποια στατιστική ανάλυση. Με τον τρόπο αυτό, ο αναγνώστης μπορεί να επαναλάβει μόνος του την στατιστική ανάλυση που περιγράφεται στο κεφάλαιο ώστε αντιπαραβάλλοντας τα αποτελέσματα που βρίσκει με τα αποτελέσματα που περιγράφονται στο βιβλίο, να επιβεβαιώσει βήμα-βήμα την ορθή εκτέλεση της ανάλυσης και τη σωστή ανάγνωση και κατανόηση των αποτελεσμάτων.

Παρακάτω δίνονται οδηγίες για το πώς να χρησιμοποιηθεί το παρόν βιβλίο. Οι οδηγίες έχουν σχέση με το είδος των μεταβλητών που χρησιμοποιείται σε κάθε στατιστική ανάλυση. Για το λόγο αυτό προηγείται μια σύντομη αναφορά στο τί είναι μεταβλητές και ποιά είναι τα χαρακτηριστικά τους που τις κατηγοριοποιούν σε δύο γενικές ομάδες: (1) ποσοτικού ή συνεχούς τύπου, και (2) ποιοτικού ή διακριτού τύπου. Εξηγείται, επίσης, τί είναι δείγμα και τί πληθυσμός, δύο κρίσιμες έννοιες στις επιστήμες της στατιστικής και της μεθοδολογίας της έρευνας, οι οποίες αναφέρονται συνεχώς σ' αυτό το βιβλίο.

Μεταβλητές

Κάθε μέτρηση γίνεται με κάποιο όργανο που έχει συγκεκριμένη **κλίμακα**, π.χ., το άλμα σε μήκος μετριέται με μετροταινία σε μέτρα, η επίδοση στον αγώνα ταχύτητας στα 400 μέτρα μετριέται με χρονόμετρο σε δευτερόλεπτα, το φύλο των παιδιών μπορεί να μετρηθεί με την απάντηση σε μια ερώτηση που επιτρέπει δύο πιθανές απαντήσεις Αγόρι-Κορίτσι, κλπ. Αφού συλλεχθούν τα δεδομένα, πρέπει να καταγραφούν σε μια βάση δεδομένων, που συνήθως είναι ένας Πίνακας με στήλες και γραμμές. Τα δεδομένα που συλλέγονται από κάποια κλίμακα καταγράφονται σε κάποια στήλη του Πίνακα που στην κορυφή της φέρει μια ονομασία, π.χ., άλμα σε μήκος. Στη συγκεκριμένη στήλη σε κάθε γραμμή του Πίνακα περνάμε την επίδοση του κάθε αθλητή ή αθλήτριας, δηλαδή η κάθε γραμμή αναπαριστά ένα διαφορετικό άτομο και γι' αυτό η κάθε γραμμή θα έχει μια συγκριμένη τιμή (π.χ., 4,50 μέτρα) που πιθανά θα διαφέρει από την τιμή ενός άλλου ατόμου (π.χ., 4,75 μέτρα). Δηλαδή στη στήλη «άλμα σε μήκος», σε κάθε γραμμή η τιμή μεταβάλλεται ανάλογα με την επίδοση του κάθε ατόμου. Το πεδίο, στο συγκεκριμένο παράδειγμα η στήλη που χρησιμοποιούμε για να καταγράψουμε τις τιμές των ατόμων σε κάποια μέτρηση, ονομάζεται «**μεταβλητή**»².

Όπως δηλώνει η ετυμολογία της λέξης, η μεταβλητή περιγράφει κάτι που μεταβάλλεται. Για παράδειγμα, η επίδοση στο άλμα σε μήκος είναι ένας αριθμός που μεταβάλλεται ανάλογα με την επίδοση του κάθε αθλητή ή αθλήτριας. Σ' αυτό το παράδειγμα η μεταβολή του αριθμού έχει μια συνεχή ροή όπως ο δείκτης του ρολογιού, για παράδειγμα κάποιος αθλητής μπορεί να πηδήξει 6,51 μέτρα, κάποιος άλλος 6,52 μέτρα, ένας τρίτος 6,53 μέτρα, κλπ. Για το λόγο αυτό τις μεταβλητές που παίρνουν τιμές που έχουν μια συνεχή ροή τις ονομάζουμε «**συνεχούς τύπου**».

Διαφορετικό παράδειγμα μεταβλητής είναι το φύλο των αθλητών που μεταβάλλεται ανάλογα με το αν αθλητής είναι άνδρας ή γυναίκα. Σ' αυτό το παράδειγμα η μεταβλητή διακρίνει τα υποκείμενα σε ομάδες ή κατηγορίες, για το λόγο αυτό ονομάζεται «**διακριτού τύπου**». Στο συγκεκριμένο παράδειγμα που αφορά στο φύλο των αθλητών, η διάκριση είναι μεταξύ δύο ομάδων (1 = άνδρες, 2 = γυναίκες), γι' αυτό το είδος αυτής της μεταβλητής εκτός από διακριτού τύπου μπορεί να ονομαστεί και «**διχοτόμος**». Άλλο παράδειγμα διακριτής μεταβλητής είναι το αποτέλεσμα ενός αγώνα για μια ομάδα, που μπορεί να πάρει τρεις τιμές: νίκη, ισοπαλία ή ήττα. Άλλο παράδειγμα διακριτής μεταβλητής είναι η θρησκεία, που περιγράφει χριστιανούς ορθόδοξους (ομάδα 1), χριστιανούς καθολικούς (ομάδα 2), χριστιανούς προτεστάντες (ομάδα 3), βουδιστές (ομάδα 4), μουσουλμάνους σουνίτες (ομάδα 5), κλπ.

Μεταβλητές διακριτού τύπου όπως το φύλο, η θρησκεία ή το αποτέλεσμα του αγώνα περιγράφουν ομάδες που διαφοροποιούνται μεταξύ τους ως προς διαφορετικά ποιοτικά χαρακτηριστικά, για το λόγο αυτό οι μεταβλητές αυτές ονομάζονται και «**ποιοτικές**», σε αντίθεση με τις μεταβλητές συνεχούς τύπου, όπως το άλμα σε μήκος, που ονομάζονται «**ποσοτικές**» επειδή ακριβώς περιγράφουν ποσότητες.

Μεταξύ των δύο παραπάνω γενικών κατηγοριών, μπορεί να βρει κανείς και μεταβλητές που μπορεί να έχουν χαρακτηριστικά διάκρισης-ποιότητας και συνέχειας-ποσότητας ταυτόχρονα, για παράδειγμα η μεταβλητή «ηλικιακή ομάδα» μπορεί να έχει ομάδες παιδιών, νέων, μεσήλικων και ηλικιωμένων, όπου οι ομάδες διαφοροποιούνται μεταξύ τους εντός μιας συνέχειας (από τις ομάδες μικρότερης ηλικίας προς τις ομάδες μεγαλύτερης ηλικίας). Συνήθως αυτές τις μεταβλητές τις χρησιμοποιούμε ως διακριτού τύπου. Κάποιες φορές, όμως, οι μεταβλητές αυτές προσεγγίζουν περισσότερο τα χαρακτηριστικά των μεταβλητών συνεχούς-ποσοτικού τύπου κι αυτό συμβαίνει όταν τα διαστήματα μεταξύ των ομάδων είναι ίσα μεταξύ τους, για παράδειγμα η συχνότητα άσκησης ενός ασκούμενου στο γυμναστήριο το μήνα περιγράφεται ως εξής:

1-5 φορές το μήνα, 6-10 φορές το μήνα, 11-15 φορές το μήνα, 16-20 φορές το μήνα, 21-25 φορές το μήνα, 26-30 φορές το μήνα. Οι μεταβλητές αυτές ονομάζονται «ίσων διαστημάτων» και περιγράφουν **ομαδοποιημένα ποσοτικά δεδομένα**.

Στις αναλύσεις που παρουσιάζονται σ' αυτό το βιβλίο χρησιμοποιούνται μόνο μεταβλητές συνεχούς ή διακριτού τύπου οι οποίες συναντώνται και πιο συχνά στην έρευνα. Για τις ειδικές περιπτώσεις όπου χρησιμοποιούνται μεταβλητές «ίσων διαστημάτων» επιφυλασσόμαστε να συμπεριλάβουμε σχετικά παραδείγματα σε μια επόμενη έκδοση.

Στο SPSS/PASW και στη βιβλιογραφία συναντάμε συνήθως και δύο ακόμη ονομασίες για τις μεταβλητές: **Εξαρτημένη (dependent) και ανεξάρτητη (independent)**. Σε γενικές γραμμές, ανεξάρτητη μεταβλητή (σε αναλύσεις όπως η παλινδρόμηση ονομάζεται και ερμηνευτική ή προβλεπτική μεταβλητή) ορίζεται αυτή που θεωρείται ότι προκαλεί κάποια επίδραση (π.χ., σε ένα πείραμα) σε άλλη ή άλλες μεταβλητές (π.χ., την εξαρτημένη μεταβλητή). Εξαρτημένη είναι η μεταβλητή (σε αναλύσεις όπως η παλινδρόμηση συναντάται και με την ονομασία μεταβλητή κριτήριο) της οποίας η τιμή καθορίζεται μερικώς από την επίδραση κάποιας άλλης μεταβλητής ή μεταβλητών (π.χ. της ανεξάρτητης μεταβλητής). Για παράδειγμα, κάποιος θέλει να εξετάσει την επίδραση δύο διαφορετικών προπονητικών προγραμμάτων στην επίδοση των αθλητών στο άλμα σε μήκος. Ανεξάρτητη μεταβλητή είναι η πειραματική συνθήκη που παίρνει δύο τιμές, την τιμή 1 αν κάποιος προπονήθηκε με το πρόγραμμα 1 και την τιμή 2 αν κάποιος προπονήθηκε με το πρόγραμμα 2. Εξαρτημένη μεταβλητή είναι η επίδοση στο άλμα σε μήκος που ως συνεχούς τύπου μεταβλητή παίρνει πάρα πολλές τιμές (θεωρητικά άπειρες, αν συμπεριληφθούν οι δεκαδικές τιμές) μεταξύ της μέγιστης και ελάχιστης δυνατής επίδοσης.

Επειδή στις έρευνες δεν έχουμε πάντοτε σχέση αιτίου-αιτιατού το οποίο συνήθως συναντάται σε πειραματικές διαδικασίες με αυστηρή μεθοδολογία, στο παρόν βιβλίο χρησιμοποιούνται συχνότερα οι όροι συνεχής/ποσοτική και διακριτή/ποιοτική διότι δεν εμπεριέχουν την προϋπόθεση ότι κάποια μεταβλητή επηρεάζει την άλλη.

Πληθυσμός, δείγμα

Όταν κάνουμε μια έρευνα, συνήθως κάνουμε μετρήσεις σε ένα συγκεκριμένο αριθμό ατόμων, και σχεδόν ποτέ δεν περιλαμβάνουμε απολύτως όλα τα άτομα

του πληθυσμού που μελετάμε. **Πληθυσμός** είναι η πλήρης συλλογή όλων των ατόμων που μοιράζονται ένα συγκεκριμένο χαρακτηριστικό, π.χ., ο πληθυσμός των Ελλήνων είναι όλοι οι πολίτες που μιλάνε Ελληνικά ως μητρική γλώσσα και πιθανά έχουν μια σχετικά κοινή γεννητική κληρονομιά. Τον ακριβή αριθμό του πληθυσμού των Ελλήνων τον μετράμε πολύ σπάνια, παραδείγματος χάριν σε μια απογραφή, στην οποία πολύ λίγες επιπλέον μεταβλητές είναι δυνατόν να αξιολογήσουμε (π.χ. πέρα από το φύλο, ηλικία, τόπο διαμονής κλπ.). Λόγω του κόστους της έρευνας, για όλες τις υπόλοιπες μεταβλητές, π.χ. βάρος, συχνότητα άθλησης των Ελλήνων, κλπ., συνήθως επιλέγουμε με τυχαίο τρόπο υπο-ομάδες του πληθυσμού για να κάνουμε τις συγκεκριμένες μετρήσεις που επιθυμούμε. Η επιλογή μιας συγκεκριμένης υπο-ομάδας του πληθυσμού ονομάζεται **δείγμα**.

Στην επιστημονική έρευνα κάνοντας μετρήσεις σε ένα δείγμα συνήθως θέλουμε να βγάλουμε συμπεράσματα για τον πληθυσμό που μελετάμε. Επειδή δεν έχουμε όλο τον πληθυσμό αλλά μόνο δείγμα αυτού, με βάση τις μετρήσεις μας στο δείγμα θέλουμε να βγάλουμε συμπεράσματα για το τί ισχύει σε όλο τον πληθυσμό. Αυτό είναι αντικείμενο της **επαγωγικής στατιστικής** που αποτελεί και τη θεματολογία του κύριου όγκου αυτού του βιβλίου.

Για να περιγράψουμε τα δεδομένα που συλλέγουμε σε κάποια μέτρηση σε μια μεταβλητή χρησιμοποιούμε την περιγραφική στατιστική, για παράδειγμα δίνουμε το μικρότερο και χαμηλότερο σκορ στο άλμα σε μήκος, τη μέση τιμή των επιδόσεων στο άλμα σε μήκος, το ποσοστό των αγοριών και κοριτσιών που συμμετείχαν στην έρευνα ή που είχαν επιδόσεις πάνω από ένα κρίσιμο σκορ που μας ενδιαφέρει κλπ. Στο παρόν βιβλίο, η ύλη ξεκινά με τον τρόπο που υπολογίζουμε δύο από τα βασικότερα στοιχεία της περιγραφικής στατιστικής, τη μέση τιμή και τη διάμεσο και, ευθύς αμέσως, εξηγούνται με παραδείγματα οι βασικότερες έννοιες της επαγωγικής στατιστικής τις οποίες χρειάζεται να γνωρίζει κανείς για να κάνει κάποια στατιστική ανάλυση. Στο υπόλοιπο του βιβλίου και σε κάθε κεφάλαιο περιγράφεται μία συγκεκριμένη στατιστική ανάλυση.

Το ευρύτερο γραμμικό μοντέλο

Για να γίνει εύκολη η κατανόηση του σκοπού της κάθε στατιστικής ανάλυσης, οι στατιστικές αναλύσεις με τη χρήση του στατιστικού προγράμματος SPSS/PASW που περιγράφονται σ' αυτό το βιβλίο μπορούν να εκληφθούν ως μέρος

ενός Ευρύτερου Γραμμικού Μοντέλου. Στον παρακάτω Πίνακα Α φαίνεται η θέση όλων των αναλύσεων σ' αυτό το ευρύτερο γραμμικό μοντέλο. Σε κάθε ανάλυση ουσιαστικά εξετάζουμε τη σχέση μεταξύ μεταβλητών. Στο ευρύτερο γραμμικό μοντέλο, οι σχέσεις αφορούν μεταβλητές που κατατάσσονται σε δύο διαστάσεις, στον άξονα X και στον άξονα Y. Σαν μεταβλητές X ονομάζονται οι μεταβλητές που διαμορφώνουν τις Y.

Για παράδειγμα, στην απλή μορφή παλινδρόμησης ο τύπος της σχέσης δύο μεταβλητών Y και X είναι ο εξής:

$$Y = A + BX + e,$$

όπου

- B είναι ο συντελεστής παλινδρόμησης που περιγράφει πόσο μεγάλη είναι η σχέση της Y με την X,
- A είναι η σταθερή τιμή (που ισούται με τη Y όταν X=0),
- και e είναι το σφάλμα πρόβλεψης.

Όλες οι στατιστικές αναλύσεις που χρησιμοποιούμε γίνονται με βάση τη μορφή των μεταβλητών, αν δηλαδή είναι κατηγορικές/διακριτού τύπου/ποιοτικές ή αν είναι συνεχούς τύπου/ποσοτικές και ανάλογα με τη θέση που παίρνουν στους άξονες X και Y. Η συγκεκριμένη στατιστική ανάλυση που χρησιμοποιούμε κάθε φορά για να εξετάσουμε τη σχέση των μεταβλητών X και Y ανάλογα με τα χαρακτηριστικά τους περιγράφεται στον παρακάτω Πίνακα Α.

Στο παρόν βιβλίο παρουσιάζονται όλες οι αναλύσεις που εξετάζουν τη σχέση δύο μεταβλητών καθώς και οι περισσότερες από τις πιο πολύπλοκες αναλύσεις οι οποίες και συναντώνται συχνότερα στη βιβλιογραφία. Ανάλογα με την υποδοχή που θα επιφυλάξουν οι αναγνώστες σ' αυτό το βιβλίο, αναλύσεις που χρησιμοποιούνται σπανιότερα στη βιβλιογραφία στις αθλητικές επιστήμες, όπως η Ανάλυση Διάκρισης (Discriminant Analysis) και η Κανονική Ανάλυση Συσχέτισης (Canonical Correlation Analysis) μπορεί να συμπεριληφθούν σε μια επόμενη έκδοση. Σας ευχόμαστε καλή ανάγνωση και καλές στατιστικές αναλύσεις!

Πίνακας Α. Σύνοψη των στατιστικών αναλύσεων του Γενικότερου Γραμμικού Μοντέλου

<u>ΜΟΡΦΗ ΑΝΑΛΥΣΗΣ</u>	<u>ΜΟΡΦΗ ΜΕΤΑΒΛΗΤΩΝ Χ ΚΑΙ Υ</u>
A. Αναλύσεις με δύο μεταβλητές	
1. Συσχέτιση Pearson product moment correlation	Η Χ συνεχής, η Υ συνεχής
2. χ^2 Chi square	Η Χ διακριτή, η Υ διακριτή
3. Τ-τεστ σε ανεξάρτητα δείγματα (Independent samples T-test)	Η Χ διχοτόμος, η Υ συνεχής
4. Ανάλυση διακύμανσης προς έναν παράγοντα/μιας κατεύθυνσης (One-way ANOVA)	Η Χ διακριτή, η Υ συνεχής
B. Αναλύσεις με μία Υ (εξαρτημένη) και πολλές Χ (ανεξάρτητες)	
5. Ανάλυση Παλινδρόμησης (Regression)	Οι Χ συνεχείς, η Υ συνεχής
6. Ανάλυση διακύμανσης προς δύο ή περισσότερους παράγοντες (Factorial ANOVA)	Οι Χ διακριτές, η Υ συνεχής
7. Ανάλυση Συνδιακύμανσης (ANCOVA)	Κάποιοι Χ διακριτές και κάποιοι συνεχείς, η Υ συνεχής
8. Discriminant Analysis	Όλες οι Χ συνεχείς, η Υ διακριτή
9. Loglinear analysis ή Multiway Frequency analysis	Όλες οι Χ διακριτές, η Υ διακριτή
Γ. Αναλύσεις με πολλές Υ και πολλές Χ	
10. Canonical correlation	Όλες οι Χ συνεχείς, όλες οι Υ συνεχείς
11. Πολυμεταβλητή Ανάλυση Διακύμανσης (MANOVA)	Όλες οι Χ διακριτές, όλες οι Υ συνεχείς
12. Πολυμεταβλητή Ανάλυση Συνδιακύμανσης (MANCOVA)	Κάποιοι Χ συνεχείς και κάποιοι διακριτές, όλες οι Υ συνεχείς
13. Discriminant Analysis	Όλες οι Χ συνεχείς, όλες οι Υ διακριτές
14. Παραγοντική Ανάλυση (Factor Analysis)	Όλες οι Χ συνεχείς, οι Υ αφανείς-λανθάνουσες (latent)
15. Loglinear analysis ή Multiway Frequency analysis	Όλες οι Χ διακριτές, όλες οι Υ διακριτές

Αναφορές στη βιβλιογραφία

- American Psychological Association. (2010). *Publication manual of the American Psychological Association* (6th ed.). Washington, DC.
- Tabachnick, B.G., & Fidell, L.S. (2012). *Using Multivariate Statistics* (6th Edition). N.J.: Pearson.

Υποσημείωση

1. Εξαρτάται πώς ορίζεται η έννοια «αποτελεσματικός». Αν αποτελεσματικός σημαίνει το καλύτερο αποτέλεσμα στην συντομότερο χρόνο τότε πλεονεκτεί ο πρώτος φοιτητής. Αν αποτελεσματικός σημαίνει να μάθει κανείς στατιστική, τότε ο δεύτερος φοιτητής ακολούθησε αποτελεσματικότερη διαδικασία.
2. Το πεδίο μιας μεταβλητής μπορεί να αποδοθεί με διάφορες μορφές και όχι μόνο με μορφή στήλης. Π.χ. σε ένα Πίνακα μπορεί μια μεταβλητή να αποδοθεί ως γραμμή όπου οι στήλες περιγράφουν τα άτομα. Στο SPSS και στις περισσότερες βάσεις δεδομένων οι μεταβλητές έχουν τη μορφή στήλης ενός Πίνακα δύο διαστάσεων, όπου η κάθε γραμμή αφιερώνεται για κάθε άτομο.