

# ΕΦΑΡΜΟΣΜΕΝΗ ΙΑΤΡΙΚΗ ΕΡΕΥΝΑ APPLIED MEDICAL RESEARCH

ΑΡΧΕΙΑ ΕΛΛΗΝΙΚΗΣ ΙΑΤΡΙΚΗΣ 2009, 26(5):699-711  
ARCHIVES OF HELLENIC MEDICINE 2009, 26(5):699-711

## Στατιστικές μέθοδοι ανάλυσης δεδομένων

Π. Γαλάνης

Εργαστήριο Οργάνωσης  
και Αξιολόγησης Υπηρεσιών Υγείας,  
Τμήμα Νοσηλευτικής, Πανεπιστήμιο  
Αθηνών, Αθήνα

Statistical methods of data analysis

Abstract at the end of the article

1. Εισαγωγή
2. Επίπεδο στατιστικής σημαντικότητας
3. Είδη δεδομένων
  - 3.1. Ονομαστικά δεδομένα
  - 3.2. Διατάξιμα δεδομένα
  - 3.3. Δεδομένα διαστηματικής κλίμακας
  - 3.4. Δεδομένα κλίμακας λόγου
  - 3.5. Συνεχή και ασυνεχή δεδομένα
4. Αριθμητικά περιληπτικά μέτρα
5. Στατιστικές μέθοδοι ανάλυσης
  - 5.1. Ποσοτικές μεταβλητές
    - 5.1.1. Παραμετρικές μέθοδοι
    - 5.1.2. Μη παραμετρικές μέθοδοι
  - 5.2. Διατάξιμες μεταβλητές
  - 5.3. Ονομαστικές μεταβλητές
6. Συσχέτιση
7. Γραμμική παλινδρόμηση
8. Λογιστική παλινδρόμηση
9. Ανάλυση επιβίωσης
10. Συμπεράσματα

### Λέξεις ευρητηρίου

Ανάλυση δεδομένων  
Ανάλυση επιβίωσης  
Γραμμική παλινδρόμηση  
Λογιστική παλινδρόμηση  
Παραμετρικές μέθοδοι

Υποβλήθηκε 24.10.2008

Εγκρίθηκε 22.11.2008

### 1. ΕΙΣΑΓΩΓΗ

Το αντικείμενο των επιδημιολογικών μελετών είναι μια συνάρτηση συχνότητας (occurrence function) και πιο συγκεκριμένα ένα μέτρο σχέσης\* (measure of association) που ποσοτικοποιεί τη σχέση μεταξύ του μελετώμενου προσδιοριστή\*\* και της έκβασης.\*\*\* Προκειμένου να διαπιστωθεί αν η σχέση αυτή είναι στατιστικά σημαντική χρησιμοποιούνται οι έλεγχοι της στατιστικής σημαντικότητας (statistical significance tests) ή στατιστικοί έλεγχοι ή στατιστικές δοκιμασίες.<sup>1-8</sup>

Οι έλεγχοι της στατιστικής σημαντικότητας (ΕΣΣ) χρησιμοποιούνται για τον έλεγχο της επιστημονικής υπόθεσης, που περιλαμβάνει δύο είδη υποθέσεων, τη μηδενική υπόθεση (null hypothesis) και την εναλλακτική (alternative).<sup>1,3,4,8</sup> Σημειώνεται ότι η επιστημονική υπόθεση αποτελεί καταφατική αποφαντική πρόταση και εισάγει μια νέα πληροφορία, ενώ η μηδενική ή η άκυρη υπόθεση δεν αποτελεί καν επιστημονική πρόταση, καθώς δεν υποστηρίζει το διαψεύσιμο

ισχυρισμό για την ύπαρξη σχέσης μεταξύ προσδιοριστή και έκβασης –και γι' αυτό καλείται συχνά και υπόθεση της

\*\* Παράγοντας κινδύνου (risk factor) ή έκθεση (exposure) ή προσδιοριστής (determinant), όπως τελικά επικράτησε να λέγεται σήμερα, είναι το χαρακτηριστικό –συγγενές, περιβαλλοντικό ή συμπεριφορικό – των ατόμων από το οποίο εξαρτάται (σχετίζεται ή συναρτάται) η συχνότητα εμφάνισης της μελετώμενης έκβασης (η οποία στην αιτιογνωστική επιδημιολογία είναι η εμφάνιση της πάθησης).<sup>1,4,8</sup> Ο προσδιοριστής της συχνότητας εμφάνισης μιας έκβασης περιλαμβάνει δύο κατηγορίες, την ενδεικτική κατηγορία (index category) και την κατηγορία αναφοράς (reference category). Η επιλογή της κατηγορίας αναφοράς του προσδιοριστή είναι μείζονος σημασίας στο σχεδιασμό της μελέτης. Είναι κρίσιμης σημασίας να γίνει αντιληπτό ότι η κατηγορία αναφοράς δεν είναι συμφυές χαρακτηριστικό του προσδιοριστή, αλλά αποτέλεσμα επιλογής του ερευνητή, άρα συνέπεια του σχεδιασμού της μελέτης. Από την επιλογή της κατηγορίας αναφοράς εξαρτάται η ερμηνεία των αποτελεσμάτων μιας μελέτης. Για παράδειγμα, προσδιοριστής της συχνότητας εμφάνισης της νεφρικής πάθησης δεν είναι η αρτηριακή υπέρταση, αλλά η αρτηριακή πίεση. Η αρτηριακή υπέρταση είναι μια κατηγορία και συνήθως η ενδεικτική κατηγορία του προσδιοριστή, στην οποία μελετάται η συχνότητα εμφάνισης της νεφρικής πάθησης σε σχέση πάντοτε με τη συχνότητα εμφάνισής της στην κατηγορία αναφοράς, στην προκειμένη περίπτωση στην κατηγορία των ατόμων που δεν έχουν αρτηριακή υπέρταση.

\* Τα μέτρα σχέσης που μπορούν να υπολογιστούν σε μια επιδημιολογική μελέτη είναι η διαφορά ή ο λόγος των μέτρων συχνότητας, το αποδοτέο κλάσμα στους εκτεθειμένους ή στο σύνολο του πληθυσμού, ο λόγος των οτς και οι συντελεστές παλινδρόμησης.<sup>1</sup>

\*\*\* Ο όρος έκβαση στην επιδημιολογία χρησιμοποιείται για να δηλώσει άλλοτε την εμφάνιση μιας πάθησης (health outcome) και άλλοτε την ίαση ή το θάνατο ή την εμφάνιση καταλοίπων από μια πάθηση (illness outcome).

μη διαφοράς– και διατυπώνεται με σκοπό να απορριφθεί. Η συμπληρωματική έννοια της μηδενικής υπόθεσης καλείται εναλλακτική υπόθεση και υποστηρίζει ότι υπάρχει σχέση μεταξύ προσδιοριστή και έκβασης. Η εναλλακτική υπόθεση μπορεί να είναι «μονής κατεύθυνσης» (one-sided), υποστηρίζοντας ότι υπάρχει συγκεκριμένη σχέση –είτε θετική είτε αρνητική– μεταξύ προσδιοριστή και έκβασης, ή «διπλής κατεύθυνσης» (two-sided), υποστηρίζοντας την ύπαρξη σχέσης χωρίς όμως να καθορίζεται εάν είναι θετική ή αρνητική.

Εάν είναι σαφές, πριν από τη διεξαγωγή μιας μελέτης, ότι υπάρχει θετική ή αρνητική σχέση μεταξύ προσδιοριστή και έκβασης, τότε η εναλλακτική υπόθεση μπορεί να είναι «μονής κατεύθυνσης». Αν μια μελέτη, π.χ., διερευνά τη σχέση μεταξύ καπνισματικής συνήθειας (προσδιοριστή) και συχνότητας εμφάνισης καρκίνου του πνεύμονα (έκβασης), τότε η εναλλακτική υπόθεση μπορεί να είναι μονής κατεύθυνσης, δηλαδή ότι το κάπνισμα αυξάνει τη συχνότητα εμφάνισης του καρκίνου του πνεύμονα. Αντιθέτως, εάν πριν από τη διεξαγωγή μιας μελέτης δεν είναι σαφής η σχέση μεταξύ προσδιοριστή και έκβασης, τότε η εναλλακτική υπόθεση είναι «διπλής κατεύθυνσης», γεγονός που συμβαίνει στην πλειονότητα των επιδημιολογικών μελετών.

Με τους ελέγχους στατιστικής σημαντικότητας ελέγχεται η μηδενική υπόθεση και είτε απορρίπτεται είτε όχι. Σε καμία περίπτωση όμως τα δεδομένα μιας μελέτης δεν μπορούν να προσφέρουν απόδειξη ότι η μηδενική υπόθεση είναι αληθής. Αν δεν απορριφθεί η μηδενική υπόθεση, τότε υποστηρίζεται ότι τα δεδομένα στα οποία στηρίζεται ο έλεγχος, δηλαδή τα δεδομένα της μελέτης, δεν επαρκούν για την απόρριψή της. Τονίζεται και πάλι ότι ο στατιστικός έλεγχος των υποθέσεων δεν οδηγεί στην απόδειξη της υπόθεσης, αλλά απλά παρέχει την πληροφορία εάν τα διαθέσιμα δεδομένα στηρίζουν την υπόθεση. Όταν η μηδενική υπόθεση δεν απορρίπτεται, δεν σημαίνει ότι είναι και αληθινή. Αντιθέτως, η απόρριψη της μηδενικής υπόθεσης σημαίνει ότι, σύμφωνα με τα δεδομένα μιας συγκεκριμένης μελέτης, υπάρχει σχέση μεταξύ προσδιοριστή και συχνότητας εμφάνισης της έκβασης.

Σε μια μελέτη που διερευνά τη σχέση μεταξύ της χρήσης κινητών τηλεφώνων και της συχνότητας εμφάνισης όγκου στον εγκέφαλο, η μηδενική υπόθεση υποστηρίζει ότι δεν υπάρχει σχέση μεταξύ προσδιοριστή και έκβασης, ενώ η εναλλακτική υπόθεση «διπλής κατεύθυνσης» υποστηρίζει την ύπαρξη σχέσης. Ανάλογα με το είδος των δεδομένων της μελέτης πραγματοποιείται ο κατάλληλος ΕΣΣ και απορρίπτεται ή όχι η μηδενική υπόθεση. Αν απορριφθεί η μηδενική υπόθεση, τότε σύμφωνα με τα δεδομένα της μελέτης αυτής υπάρχει σχέση μεταξύ της χρήσης κινητών και της

συχνότητας εμφάνισης όγκου στον εγκέφαλο. Αντιθέτως, αν δεν απορριφθεί η μηδενική υπόθεση, τότε δεν σημαίνει ότι δεν υπάρχει σχέση μεταξύ της χρήσης κινητών και της συχνότητας εμφάνισης όγκου στον εγκέφαλο. Απλά σημαίνει ότι τα δεδομένα της μελέτης αυτής δεν προσέφεραν ένδειξη για την απόρριψη της μηδενικής υπόθεσης. Είναι πολύ πιθανό μια άλλη μελέτη να καταλήξει στο συμπέρασμα ότι υπάρχει σχέση. Προφανώς, όσο περισσότερες μελέτες απορρίψουν τη συγκεκριμένη μηδενική υπόθεση τόσο μεγαλύτερη ένδειξη υπάρχει ότι η χρήση κινητών τηλεφώνων σχετίζεται με τη συχνότητα εμφάνισης όγκου στον εγκέφαλο. Δεν είναι όμως δυνατόν να επαληθευτεί\* η αιτιακή σχέση μεταξύ της χρήσης κινητών τηλεφώνων και της συχνότητας εμφάνισης όγκου στον εγκέφαλο. Η μη απόρριψη, εξάλλου, της μηδενικής υπόθεσης σε μια μελέτη δεν σημαίνει ότι είναι αληθής και ότι δεν υπάρχει σχέση ανάμεσα στο μελετώμενο προσδιοριστή και τη συχνότητα εμφάνισης της έκβασης.

## 2. ΕΠΙΠΕΔΟ ΣΤΑΤΙΣΤΙΚΗΣ ΣΗΜΑΝΤΙΚΟΤΗΤΑΣ

Το ερώτημα που γεννιέται είναι πώς μπορεί να απορριφθεί η μηδενική υπόθεση με τη χρήση του κατάλληλου στατιστικού ελέγχου της υπόθεσης. Για το σκοπό αυτό χρησιμοποιείται το παρατηρούμενο επίπεδο στατιστικής σημαντικότητας (τιμή  $p$ ), που υπολογίζεται με τη χρήση του κατάλληλου στατιστικού μοντέλου ανάλογα με τα δεδομένα μιας μελέτης και το οποίο συγκρίνεται με ένα προκαθορισμένο επίπεδο στατιστικής σημαντικότητας (τιμή  $\alpha$ ). Το προκαθορισμένο επίπεδο σημαντικότητας καθορίζεται αυθαίρετα από τον ερευνητή και συνήθως λαμβάνει την τιμή 0,05 ή σπανιότερα την τιμή 0,10 ή την τιμή 0,01.<sup>1,3,4,8-10</sup>

Η τιμή  $p$  υπολογίζεται σε σχέση με τη μηδενική υπόθεση, σύμφωνα με την οποία δεν υπάρχει σχέση μεταξύ προσδιοριστή και συχνότητας εμφάνισης της έκβασης.<sup>1,3,4,8</sup> Η τιμή  $p$  είναι η πιθανότητα, με δεδομένο ότι η μηδενική υπόθεση είναι αληθής, τα δεδομένα μιας μελέτης να παρουσιάσουν μια σχέση μεταξύ προσδιοριστή και συχνότητας εμφάνισης της έκβασης τόσο ακραία ή πιο ακραία από αυτήν που πραγματικά λαμβάνεται. Παραδείγματος χάρη, μια μελέτη διερευνά τη σχέση μεταξύ της χρήσης κινητών τηλεφώνων και της συχνότητας εμφάνισης όγκου στον εγκέφαλο και υπολογίζεται ότι ο λόγος των επιπτώσεων-πυκνοτήτων είναι ίσος με 2,5. Η μηδενική υπόθεση είναι ότι δεν υπάρχει σχέση μεταξύ της χρήσης κινητών τηλεφώνων και της

\* Ο όρος verification αποδίδεται με τον όρο επαλήθευση, ενώ ο όρος confirmation αποδίδεται με τον όρο επιβεβαίωση ή επικύρωση. Τέλος, ο όρος support αποδίδεται με τον όρο στήριξη.

συχνότητας εμφάνισης όγκου στον εγκέφαλο, οπότε ο λόγος των επιπτώσεων-πυκνοτήτων σε μια τέτοια περίπτωση είναι ίσος με τη μονάδα. Ανάλογα με τα δεδομένα της μελέτης χρησιμοποιείται το κατάλληλο στατιστικό μοντέλο για τον έλεγχο της μηδενικής υπόθεσης, οπότε προκύπτει μια τιμή  $p$ . Αυτή η τιμή  $p$  είναι η πιθανότητα, με δεδομένο ότι ο λόγος των επιπτώσεων-πυκνοτήτων είναι ίσος με τη μονάδα, μια μελέτη να δώσει λόγο των επιπτώσεων-πυκνοτήτων  $\geq 2,5$ . Πρέπει να τονιστεί ότι η τιμή  $p$  δεν είναι η πιθανότητα ότι η μηδενική υπόθεση είναι αληθής. Η τιμή  $p$  υπολογίζεται με την προϋπόθεση ότι η μηδενική υπόθεση είναι αληθής.

Όπως προαναφέρθηκε, για τον έλεγχο της μηδενικής υπόθεσης συγκρίνεται η τιμή  $p$  με την τιμή  $\alpha$ , η οποία είναι αυθαίρετη και καθορίζεται από τον ερευνητή. Εάν η τιμή  $p$ , που προκύπτει από τα δεδομένα μιας μελέτης, είναι μικρότερη από την τιμή  $\alpha$ , τότε απορρίπτεται η μηδενική υπόθεση, ενώ αν η τιμή  $p$  είναι μεγαλύτερη από την τιμή  $\alpha$ , τότε δεν απορρίπτεται η μηδενική υπόθεση, αλλά και ούτε επιβεβαιώνεται. Εάν η τιμή  $\alpha$  καθοριστεί ως 0,05, τότε οποιαδήποτε τιμή  $p$  η οποία είναι  $< 0,05$  οδηγεί στην απόρριψη της μηδενικής υπόθεσης. Μια τιμή  $p$  ίση με 0,07 οδηγεί στην απόρριψη της μηδενικής υπόθεσης όταν η τιμή  $\alpha$  είναι 0,10, ενώ οδηγεί στη μη απόρριψη της μηδενικής όταν η τιμή  $\alpha$  είναι 0,05. Έτσι, η ίδια τιμή  $p$  μπορεί να οδηγήσει σε ακριβώς αντίθετα συμπεράσματα ανάλογα με την τιμή  $\alpha$ , οπότε γίνεται αντιληπτό ότι η χρήση των ελέγχων της στατιστικής σημαντικότητας και των παρατηρούμενων επιπέδων σημαντικότητας (τιμές  $p$ ) για την εξαγωγή συμπερασμάτων είναι εξαιρετικά επισφαλής. Ακόμη και αν η τιμή  $p$  είναι μικρότερη από την τιμή  $\alpha$ , αυτό δεν σημαίνει ότι η μηδενική υπόθεση δεν αποτελεί την πλέον λογική εξήγηση για τα δεδομένα μιας μελέτης.

Συχνά, δεν αναφέρεται καν η ακριβής τιμή  $p$ , παρά μόνο αν η τιμή αυτή είναι μεγαλύτερη ή μικρότερη από την τιμή  $\alpha$ . Δεν υπάρχει κανένας λόγος βέβαια ένα συνεχές μέτρο, όπως είναι η τιμή  $p$ , να μετατρέπεται σε διχοτόμο, επειδή έτσι χάνεται πολύτιμη πληροφορία. Όταν η τιμή  $\alpha$  καθορίζεται ως 0,05, τότε τιμές  $p$  ίσες με 0,001 ή 0,049 οδηγούν σε απόρριψη της μηδενικής υπόθεσης. Είναι όμως πολύ μεγάλη η διαφορά ανάμεσα σε μια τιμή  $p$  ίση με 0,001 και μια τιμή  $p$  ίση με 0,049. Για το λόγο αυτόν πρέπει να αναφέρεται η ακριβής τιμή  $p$  και όχι απλά αν είναι μεγαλύτερη ή μικρότερη από την τιμή  $\alpha$ .

Μολονότι ακόμη και σήμερα πολλοί χρησιμοποιούν τους ελέγχους της στατιστικής σημαντικότητας και τις τιμές  $p$  για την εξαγωγή συμπερασμάτων,\* είναι φανερό ότι οι

τιμές  $p$  αποτελούν απλά ένα συνεχές μέτρο της συμβατότητας μεταξύ της μηδενικής υπόθεσης και των δεδομένων μιας μελέτης. Αυθαίρετο κριτήριο της συμβατότητας αυτής αποτελεί η τιμή  $\alpha$  που καθορίζεται από τον ερευνητή. Έτσι, αν η τιμή  $\alpha$  καθοριστεί ως 0,05, τότε οποιαδήποτε τιμή  $p < 0,05$  υποδηλώνει μικρή συμβατότητα της μηδενικής υπόθεσης με τα δεδομένα μιας μελέτης. Όσο μικρότερη δηλαδή είναι η τιμή  $p$  τόσο μικρότερη είναι η συμβατότητα μεταξύ της μηδενικής υπόθεσης και των δεδομένων μιας μελέτης.

Συμπερασματικά, η χρήση των τιμών  $p$  για την εξαγωγή ασφαλών συμπερασμάτων σχετικά με τον έλεγχο των επιστημονικών υποθέσεων είναι εξαιρετικά επισφαλής. Η τιμή  $p$  που προκύπτει σε μια μελέτη δεν πρέπει να αποτελεί κριτήριο για το αν μια μηδενική υπόθεση είναι αληθής ή όχι, καθώς απαιτείται η διεξαγωγή πολλών μελετών και η σύγκριση των αποτελεσμάτων που προκύπτουν με την εφαρμογή της μετα-ανάλυσης. Η τιμή  $p$  είναι απλά ένα μέτρο της συμβατότητας μεταξύ της μηδενικής υπόθεσης και των δεδομένων μιας μελέτης και χρησιμοποιείται για την απόρριψη ή όχι της μηδενικής υπόθεσης.<sup>17</sup> Βέβαια, η μη απόρριψη της μηδενικής υπόθεσης δεν σημαίνει ότι είναι αληθής, αλλά ότι τα δεδομένα μιας μελέτης δεν προσφέρουν αρκετή απόδειξη για την απόρριψή της. Η τιμή  $p$  παρουσιάζει ένα σημαντικό μειονέκτημα, καθώς δεν εκφράζει το μέγεθος της σχέσης μεταξύ του προσδιοριστή και της συχνότητας εμφάνισης της έκβασης και την ακρίβεια –απουσία τυχαίου σφάλματος– με την οποία μετράται η σχέση αυτή. Για την καλύτερη ερμηνεία των δεδομένων μιας μελέτης είναι εξαιρετικά σημαντικό να διαχωρίζεται η πληροφορία για το μέγεθος της σχέσης αναφέροντας την τιμή του μέτρου σχέσης και η πληροφορία για την ακρίβεια της μελέτης αναφέροντας το κατάλληλο διάστημα εμπιστοσύνης.

Αξίζει να αναφερθεί ότι στην εφαρμογή της τεχνικής του ελέγχου των υποθέσεων υπάρχουν δύο τύποι σφάλματος.<sup>4-7</sup> Το πρώτο ονομάζεται *σφάλμα τύπου I* (type I error) και είναι επίσης γνωστό ως σφάλμα απόρριψης (rejection error) ή, αλλιώς, σφάλμα  $\alpha$  ( $\alpha$  error). Ένα σφάλμα τύπου I συμβαίνει όταν απορρίπτεται η μηδενική υπόθεση, ενώ είναι αληθής. Η πιθανότητα να διαπραχθεί ένα σφάλμα τύπου I καθορίζεται από το προκαθορισμένο επίπεδο σημαντικότητας του ελέγχου (τιμή  $\alpha$ ). Εάν πραγματοποιηθούν κατ'επανάληψη έλεγχοι της υπόθεσης διατηρώντας το επίπεδο σημαντικότητας στο 0,05, θα απορριφθεί λανθασμένα η μηδενική υπόθεση, ενώ είναι αληθής, 5 φορές στους 100 ΕΣΣ. Σε ορισμένες περιπτώσεις επιλέγεται η πιθανότητα 0,01, οπότε με τον τρόπο αυτόν απορρίπτεται λανθασμένα η μηδενική υπόθεση, ενώ είναι αληθής, μόνο μία φορά στους 100 ΕΣΣ.

Ο δεύτερος τύπος σφάλματος που μπορεί να διαπραχθεί κατά την εφαρμογή ενός ελέγχου υπόθεσης είναι το *σφάλμα τύπου II* (type II error), που είναι επίσης γνωστό ως σφάλμα αποδοχής (acceptance error) ή, αλλιώς, σφάλ-

\* Μια μελέτη δεν επιτρέπει την εξαγωγή συμπερασμάτων, παρέχοντας μόνο ένδειξη (evidence) ενάντια της υπόθεσης.

μα  $\beta$  ( $\beta$  error). Ένα σφάλμα τύπου II συμβαίνει όταν δεν απορρίπτεται η μηδενική υπόθεση, ενώ δεν είναι αληθής. Παραδείγματος χάρη, εάν  $\beta=0,10$ , τότε η πιθανότητα μη απόρριψης της μηδενικής υπόθεσης, ενώ δεν είναι αληθής, είναι 0,10 ή 10%.

Εάν το  $\beta$  είναι η πιθανότητα να διαπραχθεί ένα σφάλμα τύπου II, τότε  $(1-\beta)$  ονομάζεται *στατιστική ισχύς* (statistical power) του ελέγχου της υπόθεσης. Η ισχύς είναι η πιθανότητα να απορριφθεί η μηδενική υπόθεση, ενώ δεν είναι αληθής. Με άλλη διατύπωση, είναι η πιθανότητα να αποφευχθεί ένα σφάλμα τύπου II. Επίσης, η ισχύς μπορεί να θεωρηθεί ως η πιθανότητα μια συγκεκριμένη μελέτη να διακρίνει μια απόκλιση από τη μηδενική υπόθεση δεδομένου ότι αυτή υπάρχει.

Το συνηθέστερο σφάλμα είναι να θεωρείται μια θεωρία (υπόθεση) ως αληθής, ενώ δεν είναι, και σπανιότερα το αντίθετο.

Γενικά, οι ερευνητές προσπαθούν να σχεδιάσουν ΕΣΣ που έχουν υψηλή ισχύ. Δεν αρκεί να υπάρχει μικρή πιθανότητα να απορριφθεί η μηδενική υπόθεση όταν αληθεύει. Πρέπει να υπάρχει μεγάλη πιθανότητα να απορριφθεί η μηδενική υπόθεση όταν δεν είναι αληθής. Ένας τρόπος αύξησης της ισχύος ενός ελέγχου είναι η αύξηση της τιμής  $\alpha$ . Αύξηση της τιμής  $\alpha$  προκαλεί αύξηση του σφάλματος τύπου I, αλλά συγχρόνως μείωση του σφάλματος τύπου II. Αντιθέτως, μείωση της τιμής  $\alpha$  προκαλεί μείωση του σφάλματος τύπου I και αύξηση του σφάλματος τύπου II. Για το λόγο αυτόν, ο σχεδιασμός μιας μελέτης απαιτεί ιδιαίτερη προσοχή και ιδιαίτερα όσον αφορά στον προκαθορισμό της τιμής  $\alpha$ .

### 3. ΕΙΔΗ ΔΕΔΟΜΕΝΩΝ

Ο εμπειρικός καθορισμός των τιμών μιας αριθμητικής συνάρτησης ή, αλλιώς, μιας ποσοτικής έννοιας ονομάζεται μέτρηση. Ο όρος μέτρηση όμως χρησιμοποιείται συχνά και με ευρύτερη έννοια, περιλαμβάνοντας και την κατανομή των αντικειμένων μιας κατηγορίας σε τάξεις (ποιοτικές έννοιες). Ουσιαστικά, με τη διαδικασία της μέτρησης επιτυγχάνεται η συστηματική απόδοση αριθμών στα αντικείμενα και τις ιδιότητές τους, με αποτέλεσμα να διευκολύνεται σημαντικά η χρήση των μαθηματικών μεθόδων στη μελέτη και την περιγραφή των αντικειμένων και των σχέσεών τους.

Διακρίνονται διάφορα είδη ή επίπεδα μέτρησης με τη χρήση των ίδιων όρων και για τα δεδομένα\* που αντιστοι-

\* Δεδομένα (data) ή εμπειρικά δεδομένα (empirical data), με την έννοια ότι υπάρχουν ανεξάρτητα από τη νόηση, δεν υφίστανται. Ουσιαστικά, αποτελούν παρατηρήσεις (observations) νοητικά φορτισμένες. Αυτό εξάλλου διακρίνει τις παρατηρήσεις από τις εντυπώσεις (impressions).

χούν σε κάθε επίπεδο. Τα διάφορα αυτά είδη δεδομένων διαφέρουν τόσο στην ερμηνεία των αριθμητικών τιμών που χρησιμοποιούνται όσο και στον ΕΣΣ που επιλέγεται ως ο κατάλληλος για τη στατιστική ανάλυση. Οι μεταβλητές –και κατ' επέκταση και τα δεδομένα– ανάλογα με τα μαθηματικά τους χαρακτηριστικά διακρίνονται σε *ποιοτικές* (qualitative) και *ποσοτικές* (quantitative), με τις πρώτες να διακρίνονται σε ονομαστικές και διατάξιμες και τις δεύτερες σε μεταβλητές διαστηματικής κλίμακας και μεταβλητές κλίμακας λόγου.<sup>4-7,12</sup>

#### 3.1. Ονομαστικά δεδομένα

Τα ονομαστικά δεδομένα (nominal data) αποτελούν την απλούστερη και περισσότερο συνηθισμένη μορφή δεδομένων. Στην περίπτωση αυτή, τα αντικείμενα μιας ορισμένης κατηγορίας διασπώνται και εντάσσονται σε διάφορες ομάδες με τους αριθμούς να αποτελούν ουσιαστικά ονόματα ή χαρακτηρισμούς τάξεων και χωρίς να έχουν αριθμητικό νόημα. Για να είναι επιστημονικώς χρήσιμα τα δεδομένα ονομαστικής κλίμακας θα πρέπει (α) οι τάξεις να μην έχουν κοινά στοιχεία, δηλαδή να έχουν σαφώς προσδιορισμένο πλάτος, ώστε να αποκλείεται να ανήκει ένα στοιχείο σε περισσότερες από μία τάξεις και (β) η κατανομή των τάξεων να εξαντλεί την περιοχή όπου αναλύεται.

Σημειώνεται ότι οι μεταβλητές (ο αγγλικός όρος είναι variate και όχι variable) δεν υπάρχουν στη φύση και αποτελούν στατιστικές έννοιες που σχεδιάζονται από τους ερευνητές, ενώ ως δεδομένα νοούνται οι τιμές που λαμβάνει μια μεταβλητή κατά τη μέτρησή της. Το φύλο, για παράδειγμα, δεν είναι μεταβλητή. Στην περίπτωση δημιουργίας μιας μεταβλητής για το φύλο, οι άνδρες μπορούν να λάβουν την τιμή 0 και οι γυναίκες την τιμή 1. Οι τιμές αυτές 0 και 1 δεν έχουν αριθμητικό νόημα και βεβαίως δεν αποτελούν τη μοναδική επιλογή. Είναι δυνατόν οι άνδρες να λάβουν την τιμή 1 και οι γυναίκες την τιμή 2 χωρίς να επηρεαστεί η ανάλυση\*\* των δεδομένων. Παρόμοιο παράδειγμα αποτελεί η ομάδα αίματος, με τα άτομα, π.χ., της ομάδας

\*\* Ο όρος «ανάλυση δεδομένων» περιλαμβάνει (α) τα περιγραφικά στατιστικά μέτρα (descriptive statistics) που συνοψίζουν τα δεδομένα και (β) τα διαλογισμικά στατιστικά μέτρα (inferential statistics) που χρησιμοποιούνται στα στατιστικά υποδείγματα ή, αλλιώς, μοντέλα αναφορικά με την εξαγωγή συμπερασμάτων για τα αντικείμενα μιας μελέτης, που είναι, ουσιαστικά, η παράμετρος που ενδιαφέρει τους ερευνητές. Πρέπει να σημειωθεί ότι ο όρος «ανάλυση δεδομένων» (data analysis) είναι εσφαλμένος εννοιολογικά, καθώς οι παρατηρήσεις δεν είναι δεδομένα και η επεξεργασία των παρατηρήσεων, οι οποίες θεωρητικά είναι φορτισμένες, αποτελεί σύνθεση και όχι ανάλυση. Στην ουσία, η «ανάλυση των δεδομένων» είναι το σύνολο της ένδειξης (evidence) που η εμπειρική έρευνα παράγει για τον έλεγχο της υπόθεσης. Τα «δεδομένα» αποτελούν το μέσο επιλογής του κατάλληλου μαθηματικού υποδείγματος, δηλαδή του υποδείγματος που ταιριάζει καλύτερα στα δεδομένα.

Α να λαμβάνουν την τιμή 1, τα άτομα της ομάδας Β την τιμή 2, τα άτομα της ομάδας ΑΒ την τιμή 3 και τα άτομα της ομάδας Ο την τιμή 4. Οι συμβολισμοί των διαφόρων ομάδων με αριθμούς διευκολύνει σημαντικά τη διαχείρισή τους με τη χρήση των ηλεκτρονικών υπολογιστών κατά την ανάλυση των δεδομένων. Τα δεδομένα ονομαστικής κλίμακας είναι γνωστά και ως *κατηγορικά δεδομένα* (categorical data), καθώς με τις μετρήσεις ένα ορισμένο αντικείμενο συνδέεται με μια ορισμένη κατηγορία ή τάξη (π.χ. άνδρας ή γυναίκα).

Όταν τα δεδομένα ονομαστικής κλίμακας μπορούν να λάβουν μόνο μία από δύο συγκεκριμένες τιμές, όπως άνδρες και γυναίκες, τότε ονομάζονται *διχότομα* ή *δυαδικά* (dichotomous, binary). Τα διχότομα δεδομένα –για παράδειγμα, η εμφάνιση ή όχι μιας πάθησης, θετικό ή αρνητικό αποτέλεσμα μιας εργαστηριακής δοκιμασίας κ.ά.– χρησιμοποιούνται συχνά στις επιστήμες υγείας και γι' αυτό έχουν αναπτυχθεί ιδιαίτερες μέθοδοι ανάλυσης, όπως π.χ. η λογιστική παλινδρόμηση.

### 3.2. Διατάξιμα δεδομένα

Τα διατάξιμα δεδομένα (ordinal data) είναι εκείνα στα οποία η σειρά ή, αλλιώς, η διάταξη μεταξύ των διαφόρων κατηγοριών έχει σημασία, έτσι ώστε οι μεγαλύτερες αριθμητικές τιμές να αντιπροσωπεύουν την παρουσία ενός χαρακτηριστικού σε μεγαλύτερο βαθμό και οι μικρότερες την παρουσία του ίδιου χαρακτηριστικού σε μικρότερο βαθμό. Στην περίπτωση αυτή, τα αντικείμενα μιας ορισμένης κατηγορίας όχι μόνο διασπώνται και εντάσσονται σε διάφορες κατηγορίες ή τάξεις, αλλά είναι δυνατή και η διάταξη των τάξεων αυτών με τρόπο που επιτρέπει τις μεταξύ τους συγκρίσεις.

Χαρακτηριστικό παράδειγμα διατάξιμης μεταβλητής στις επιστήμες υγείας αποτελεί ο βαθμός εγκαύματος, που λαμβάνει συνήθως τιμές 1–4, με τις υψηλότερες τιμές να αντιπροσωπεύουν σοβαρότερη μορφή εγκαύματος. Ένα άλλο παράδειγμα αποτελεί η ταξινόμηση των τραυματισμών σύμφωνα με το επίπεδο σοβαρότητάς τους, με τη μεταβλητή αυτή να λαμβάνει, π.χ., τιμές 1–4, όπου 1 αντιστοιχεί σε ελαφρύ τραυματισμό, 2 σε μέτριο, 3 σε σοβαρό και 4 σε θανατηφόρο. Και στις δύο περιπτώσεις η διάταξη των τάξεων ή κατηγοριών πραγματοποιείται λογικά, αλλά δεν είναι δυνατόν να ποσοτικοποιηθεί η διαφορά μεταξύ των κατηγοριών και να καθοριστεί αν η διαφορά, για παράδειγμα, μεταξύ εγκαυμάτων πρώτου και δεύτερου βαθμού είναι ίδια με τη διαφορά μεταξύ εγκαυμάτων τρίτου και τέταρτου βαθμού.

Ιδιαίτερη κατηγορία διατάξιμων δεδομένων αποτελούν

τα *διατεταγμένα δεδομένα* (ranked data), στα οποία οι παρατηρήσεις τοποθετούνται στη σειρά από τη μεγαλύτερη στη μικρότερη –ή και αντίστροφα– και στη συνέχεια αντικαθίστανται από αριθμούς που αντιστοιχούν στη θέση κάθε παρατήρησης στην ακολουθία. Ως παραδείγματα αναφέρονται η διάταξη των πολιτειών των ΗΠΑ με βάση τον πληθυσμό ή την έκτασή τους, καθώς και η διάταξη των αιτιών θανάτου με βάση τον αριθμό θανάτων που προκαλούν. Σημειώνεται ότι στην περίπτωση των διατάξιμων δεδομένων ως μέτρο κεντρικής τάσης χρησιμοποιείται η διάμεσος και όχι ο μέσος.

### 3.3. Δεδομένα διαστηματικής κλίμακας

Στα δεδομένα διαστηματικής κλίμακας (interval scale data) έχει σημασία τόσο η διάταξη όσο και το μέγεθος, με τους αριθμούς να αντιπροσωπεύουν πραγματικές μετρήσιμες ποσότητες και όχι ονόματα ή χαρακτηρισμούς τάξεων. Τα διαστήματα που χωρίζουν τη μια τιμή της μεταβλητής από την άλλη έχουν διαφορετική βαρύτητα στα διάφορα σημεία της κλίμακας. Χαρακτηριστικό παράδειγμα της κατηγορίας αυτής αποτελεί η κλίμακα μέτρησης της έντασης των σεισμών ή, αλλιώς, κλίμακα Ρίχτερ, στην οποία οι τιμές της μεταβλητής μπορεί να απέχουν μεταξύ τους κατά ίση αριθμητική διαφορά, αλλά αυτό δεν συνεπάγεται κίολας ότι οι διαφορές αυτές είναι ισοδύναμες. Στην περίπτωση της κλίμακας Ρίχτερ, η διαφορά της έντασης μεταξύ δύο σεισμών 4 και 5 βαθμών της κλίμακας Ρίχτερ δεν είναι ισοδύναμη με τη διαφορά μεταξύ δύο σεισμών 7 και 8 βαθμών της κλίμακας Ρίχτερ. Τα δεδομένα διαστηματικής κλίμακας είναι μαθηματικά αξιοποιήσιμα, αλλά δεν υφίσταται η σχέση του πολλαπλασιασμού μεταξύ των τιμών τους. Έτσι, έχει νόημα ότι ένας σεισμός 8 βαθμών της κλίμακας Ρίχτερ είναι ισχυρότερος από ένα σεισμό 4 βαθμών της κλίμακας Ρίχτερ, αλλά όχι ότι ο πρώτος είναι δύο φορές ισχυρότερος από το δεύτερο. Σημειώνεται ότι στην πράξη τα δεδομένα διαστηματικής κλίμακας χρησιμοποιούνται ελάχιστα.

### 3.4. Δεδομένα κλίμακας λόγου

Τα δεδομένα κλίμακας λόγου (ratio scale data) υπερέρχουν έναντι των δεδομένων διαστηματικής κλίμακας στο γεγονός ότι διαθέτουν ένα φυσικό μηδενικό σημείο έναρξης της μέτρησης. Οι περισσότερες μεταβλητές –όπως το βάρος, το ύψος, η ηλικία, το εισόδημα, οι δαπάνες υγείας κ.ά.– ανήκουν στην κατηγορία αυτή. Με τα δεδομένα κλίμακας λόγου είναι δυνατή η πρόσθεση, η αφαίρεση, ο πολλαπλασιασμός και η διαίρεση των δεδομένων. Παράδειγματος χάρη, ένα άτομο με ετήσιο εισόδημα ίσο με

30.000 ευρώ έχει διπλάσιο εισόδημα από ένα άτομο με ετήσιο εισόδημα ίσο με 15.000 ευρώ. Επιπλέον, ένα άτομο βάρους 100 kg είναι δύο φορές βαρύτερο από ένα άτομο βάρους 50 kg. Σημειώνεται ότι τα δεδομένα των ευρύτατα χρησιμοποιούμενων κλιμάκων Likert, μολονότι είναι διατάξιμα, σε αρκετές περιπτώσεις αντιμετωπίζονται ως δεδομένα κλίμακας λόγου διευκολύνοντας σημαντικά τη στατιστική ανάλυση. Από στατιστική άποψη, η προσέγγιση αυτή δεν είναι η πλέον ενδεδειγμένη, αλλά εφόσον επιλεγεί πρέπει να αναφέρονται με σαφήνεια τα κριτήρια και οι προϋποθέσεις εφαρμογής της.

### 3.5. Συνεχή και ασυνεχή δεδομένα

Σε αρκετές περιπτώσεις, τα ποσοτικά δεδομένα διακρίνονται σε συνεχή (continuous) και διακριτά ή ασυνεχή (discrete). Η διαφορά τους έγκειται στο γεγονός ότι τα ασυνεχή δεδομένα περιορίζονται στο να λαμβάνουν μόνο ακέραιες τιμές, οι οποίες διαφέρουν μεταξύ τους κατά συγκεκριμένες ποσότητες, ενώ τα συνεχή δεδομένα μπορούν να λάβουν οποιαδήποτε τιμή, ακόμη και δεκαδική. Τα περισσότερα δεδομένα διαστηματικής κλίμακας και κλίμακας λόγου αποτελούν συνεχή δεδομένα, γεγονός που διευκολύνει σημαντικά τη στατιστική ανάλυση. Παραδείγματα ασυνεχών μεταβλητών αποτελούν ο αριθμός των τροχαίων ατυχημάτων, ο αριθμός των γεννήσεων μιας γυναίκας και ο αριθμός των νέων περιπτώσεων μιας πάθησης που καταγράφηκαν σε μια χώρα σ' ένα συγκεκριμένο χρονικό διάστημα, ενώ παραδείγματα συνεχών μεταβλητών αποτελούν το βάρος, το ύψος, η ηλικία, το εισόδημα, ο χρόνος, η θερμοκρασία, η αρτηριακή πίεση κ.ά. Σημειώνεται ότι σε ορισμένες περιπτώσεις οι συνεχείς μεταβλητές μετατρέπονται σε διατάξιμες εφόσον απαιτείται μικρότερος βαθμός λεπτομέρειας. Ανάλογα με το δείκτη μάζας σώματος, π.χ., τα άτομα μπορούν να ταξινομηθούν σε αδύνατα, φυσιολογικά, υπέρβαρα ή παχύσαρκα.

## 4. ΑΡΙΘΜΗΤΙΚΑ ΠΕΡΙΛΗΠΤΙΚΑ ΜΕΤΡΑ

Τα αριθμητικά περιληπτικά μέτρα διακρίνονται στα μέτρα θέσης ή κεντρικής τάσης, που περιλαμβάνουν (α) τη μέση τιμή ή μέσο (mean), (β) τη διάμεσο (median) και (γ) την επικρατούσα τιμή (mode), και στα μέτρα μεταβλητότητας ή διασποράς, που περιλαμβάνουν (α) το εύρος (range), (β) το ενδοτεταρτημοριακό εύρος (interquartile range), (γ) τη διασπορά (variance), (δ) την τυπική απόκλιση (standard deviation) και (ε) το συντελεστή μεταβλητότητας (coefficient of variance).<sup>5,6</sup>

Ο μέσος χρησιμοποιείται ως περιληπτικό μέτρο για συνεχή και ασυνεχή δεδομένα, ενώ επηρεάζεται σημαντικά

από την ύπαρξη ακραίων τιμών. Η διάμεσος επηρεάζεται σε πολύ μικρότερο βαθμό από τις ακραίες τιμές σε σχέση με το μέσο και χρησιμοποιείται τόσο για συνεχή και ασυνεχή δεδομένα όσο και για διατάξιμα. Όταν η κατανομή μιας ομάδας δεδομένων είναι συμμετρική (ή, αλλιώς, κανονική), τότε ο μέσος αποτελεί το καλύτερο μέτρο κεντρικής τάσης, ενώ όταν η κατανομή παρουσιάζει ασυμμετρία ενδείκνυται η χρήση της διαμέσου. Στην περίπτωση δεδομένων που ακολουθούν την κανονική κατανομή, ως το πλέον κατάλληλο μέτρο εκτίμησης της μεταβλητότητας θεωρείται η τυπική απόκλιση, ενώ το ενδοτεταρτημοριακό εύρος αποτελεί το αντίστοιχο μέτρο στην περίπτωση δεδομένων που δεν ακολουθούν την κανονική κατανομή.

## 5. ΣΤΑΤΙΣΤΙΚΕΣ ΜΕΘΟΔΟΙ ΑΝΑΛΥΣΗΣ

Τα τελευταία χρόνια, η ευρεία χρήση των ηλεκτρονικών υπολογιστών και των στατιστικών προγραμμάτων, όπως SPSS (SPSS Inc, Chicago, Illinois, USA), STATA (Stata Corp LP, Texas, USA), SAS (SAS Institute Inc, North Carolina, USA), Statgraphics (Statpoint Inc, Virginia, USA), Statistica (Statsoft Inc, Tulsa, USA) κ.λπ., έχει διευκολύνει σημαντικά την ανάλυση δεδομένων ακόμη και για εκείνους που δεν είναι ιδιαίτερα εξοικειωμένοι με τις έννοιες της στατιστικής. Για το λόγο αυτόν κρίθηκε σκόπιμο να μην αναλυθούν οι μαθηματικές ιδιότητες που χρησιμοποιούνται στις διάφορες στατιστικές μεθόδους παρά μόνο να δημιουργηθεί ένας συνοπτικός και χρηστικός οδηγός ανάλυσης δεδομένων, χωρίς βέβαια να εξαντλείται το πολυδιάστατο αυτό θέμα. Στους πίνακες 1 και 2 παρουσιάζονται συνοπτικά οι στατιστικές μέθοδοι που χρησιμοποιούνται συχνότερα στις επιστήμες υγείας ανάλογα με το είδος και τον αριθμό των μεταβλητών.<sup>2,5-7</sup>

### 5.1. Ποσοτικές μεταβλητές

Στην περίπτωση των ποσοτικών μεταβλητών απαιτείται καταρχήν η διεξαγωγή ελέγχου κανονικότητας, όπου στην πλειονότητα των περιπτώσεων είναι ο έλεγχος *Kolmogorov-Smirnov*.<sup>2,5-7</sup> Σημειώνεται ότι στην περίπτωση όπου ο αριθμός των παρατηρήσεων είναι μικρός, συνιστάται η εφαρμογή του ελέγχου κανονικότητας *Shapiro-Wilk*. Με τους ελέγχους κανονικότητας εκτιμάται εάν η κατανομή των πληθυσμών από τους οποίους προέρχονται τα δεδομένα των «δειγμάτων» είναι η κανονική κατανομή ή αν την προσεγγίζει. Παραδείγματος χάρη, εάν το αντικείμενο μιας μελέτης είναι η σύγκριση του αναστήματος των αγοριών σε σχέση με εκείνο των κοριτσιών, τότε το πρώτο βήμα είναι να ελεγχθεί αν η κατανομή των πληθυσμών των αγοριών και των κοριτσιών από τους οποίους προέρχονται τα μελετώμενα

**Πίνακας 1.** Στατιστικές μέθοδοι ανάλυσης δεδομένων στην περίπτωση μιας εξαρτημένης και μιας ανεξάρτητης μεταβλητής.

| Εξαρτημένη μεταβλητή        | Ανεξάρτητη μεταβλητή       | Κατανομή πληθυσμών | «Δείγματα» | Μέθοδος   |
|-----------------------------|----------------------------|--------------------|------------|---|
| Διχοτόμος                   | Ποιοτική, ποσοτική         | Κανονική           | Ανεξάρτητα | Απλή λογιστική παλινδρόμηση                                   |
| Ονομαστική                  | Ονομαστική                 |                    | Ανεξάρτητα | Έλεγχος $\chi^2$ , ακριβής έλεγχος του Fisher                 |
| Ονομαστική                  | Ονομαστική                 |                    | Ανά ζεύγη  | Έλεγχος του McNemar   |
| Ποσοτική                    | Διχοτόμος                  | Κανονική           | Ανεξάρτητα | Έλεγχος t για ανεξάρτητα «δείγματα»                           |
| Ποσοτική                    | Διχοτόμος                  | Κανονική           | Ανά ζεύγη  | Έλεγχος t για «δείγματα» ανά ζεύγη                            |
| Ποσοτική                    | Ονομαστική (>2 κατηγορίες) | Κανονική           | Ανεξάρτητα | Ανάλυση διασποράς μίας κατεύθυνσης                            |
| Ποσοτική                    | Ονομαστική (>2 κατηγορίες) | Κανονική           | Ανά ζεύγη  | Ανάλυση διασποράς για επαναλαμβανόμενες μετρήσεις             |
| Ποσοτική, διατάξιμη         | Διχοτόμος                  | Μη κανονική        | Ανεξάρτητα | Έλεγχος Mann-Whitney U  |
| Ποσοτική, διατάξιμη         | Διχοτόμος                  | Μη κανονική        | Ανά ζεύγη  | Έλεγχος προσημασμένης διάταξης του Wilcoxon, έλεγχος προσήμου |
| Ποσοτική, διατάξιμη         | Ονομαστική (>2 κατηγορίες) | Μη κανονική        | Ανεξάρτητα | Έλεγχος Kruskal-Wallis, έλεγχος διαμέσου                      |
| Ποσοτική, διατάξιμη         | Ονομαστική (>2 κατηγορίες) | Μη κανονική        | Ανά ζεύγη  | Έλεγχος Friedman, έλεγχος Kendall's W                         |
| Ποσοτική                    | Ποσοτική                   | Κανονική           | Ανεξάρτητα | Απλή γραμμική παλινδρόμηση                                    |
| Ποσοτική (χρόνος επιβίωσης) | Διχοτόμος                  | Μη κανονική        | Ανεξάρτητα | Έλεγχος log-rank  |

**Πίνακας 2.** Στατιστικές μέθοδοι ανάλυσης δεδομένων στην περίπτωση μιας εξαρτημένης και περισσότερων από μία ανεξάρτητων μεταβλητών.

| Εξαρτημένη μεταβλητή        | Ανεξάρτητη μεταβλητή | Κατανομή πληθυσμών | «Δείγματα» | Μέθοδος                             |
|-----------------------------|----------------------|--------------------|------------|-------------------------------------|
| Διχοτόμος                   | Ποιοτικές, ποσοτικές | Κανονική           | Ανεξάρτητα | Πολλαπλή λογιστική παλινδρόμηση     |
| Ποσοτική                    | Ονομαστικές          | Κανονική           | Ανεξάρτητα | Πολυμεταβλητή ανάλυση διασποράς     |
| Ποσοτική                    | Ποσοτικές            | Κανονική           | Ανεξάρτητα | Ανάλυση συνδιασποράς                |
| Ποσοτική                    | Ποιοτικές, ποσοτικές | Κανονική           | Ανεξάρτητα | Πολυμεταβλητή ανάλυση συνδιασποράς  |
| Ποσοτική                    | Ποιοτικές, ποσοτικές | Κανονική           | Ανεξάρτητα | Πολλαπλή γραμμική παλινδρόμηση      |
| Ποσοτική (χρόνος επιβίωσης) | Ποιοτικές, ποσοτικές | Μη κανονική        | Ανεξάρτητα | Μοντέλο αναλογικών κινδύνων του Cox |

«δείγματα» είναι η κανονική κατανομή ή όχι. Στην περίπτωση αυτή, απαιτείται η διεξαγωγή δύο ελέγχων κανονικότητας, τόσο δηλαδή για τον πληθυσμό των αγοριών όσο και για τον πληθυσμό των κοριτσιών. Εάν η σύγκριση, εξάλλου, αφορά στο βάρος κοριτσιών που γυμνάζονται 0 ώρες/εβδομάδα, 1–7 ώρες/εβδομάδα ή >7 ώρες/εβδομάδα, τότε απαιτούνται τρεις έλεγχοι κανονικότητας για τους τρεις αυτούς διαφορετικούς πληθυσμούς κοριτσιών.

Στην περίπτωση όπου η κατανομή των πληθυσμών από τους οποίους προέρχονται τα μελετώμενα «δείγματα» είναι η κανονική κατανομή (normal distribution), τότε χρησιμοποιούνται οι παραμετρικές μέθοδοι (parametric methods), ενώ στην περίπτωση που δεν είναι η κανονική κατανομή χρησιμοποιούνται οι μη παραμετρικές μέθοδοι.<sup>2,5–7</sup> Το σημαντικότερο πλεονέκτημα των μη παραμετρικών μεθόδων είναι ότι μπορούν να εφαρμοστούν στην περίπτωση όπου η κατανομή των πληθυσμών δεν είναι η κανονική κατανομή.

Επιπλέον, στις μη παραμετρικές μεθόδους χρησιμοποιούνται οι διατάξεις των παρατηρήσεων και όχι οι πραγματικές τιμές των παρατηρήσεων, με αποτέλεσμα να είναι δυνατή η εφαρμογή τους και στην περίπτωση «δειγμάτων» με μικρό αριθμό παρατηρήσεων. Η χρήση, εξάλλου, των διατάξεων δίνει τη δυνατότητα εφαρμογής των μη παραμετρικών μεθόδων και στην περίπτωση διατάξιμων δεδομένων και όχι μόνο ποσοτικών. Επίσης, οι μη παραμετρικές μέθοδοι είναι λιγότερο ευαίσθητες στο σφάλμα μέτρησης εξαιτίας της χρήσης των διατάξεων. Οι μη παραμετρικές μέθοδοι, ωστόσο, παρουσιάζουν και ορισμένα μειονεκτήματα έναντι των αντίστοιχων παραμετρικών, όπως μικρότερη στατιστική ισχύ, λιγότερο ειδικές υποθέσεις και μη χρήση όλης της πληροφορίας που είναι γνωστή για μια κατανομή, καθώς χρησιμοποιούνται οι διατάξεις των παρατηρήσεων και όχι οι πραγματικές τους τιμές. Τέλος, στην περίπτωση των παραμετρικών μεθόδων, για τον έλεγχο της μηδενικής υπόθεσης χρησιμοποιούνται οι μέσες τιμές και η μεταβλητότητα των

παρατηρήσεων εκτιμάται με τη χρήση της τυπικής απόκλισης, ενώ στην περίπτωση των μη παραμετρικών μεθόδων χρησιμοποιούνται οι διάμεσοι και το ενδοτεταρτημοριακό εύρος, αντίστοιχα.

**5.1.1. Παραμετρικές μέθοδοι.** Στην περίπτωση όπου τα «δείγματα» είναι δύο και ανεξάρτητα μεταξύ τους –για παράδειγμα, σύγκριση του αναστήματος μεταξύ αγοριών και κοριτσιών– χρησιμοποιείται ο *στατιστικός έλεγχος t για ανεξάρτητα «δείγματα»* (independent samples t-test).<sup>2,5-7</sup> Στην περίπτωση όπου τα «δείγματα» είναι δύο, αλλά ανά ζεύγη, οπότε για κάθε παρατήρηση στην πρώτη ομάδα υπάρχει μια αντίστοιχη παρατήρηση στη δεύτερη ομάδα, χρησιμοποιείται ο *έλεγχος t για «δείγματα» ανά ζεύγη* (paired samples t-test). Ως παραδείγματα στην κατηγορία αυτή αναφέρονται η σύγκριση της συστολικής πίεσης πριν και μετά από τη χορήγηση ενός φαρμάκου σε μια ομάδα πασχόντων και η σύγκριση των γνώσεων σχετικά με τις μεταμοσχεύσεις ιστών και οργάνων πριν και μετά από την εφαρμογή ενός εκπαιδευτικού προγράμματος σε μια ομάδα φοιτητών. Στους δύο αυτούς ελέγχους t, η μηδενική υπόθεση που ελέγχεται είναι ότι οι μέσες τιμές της ποσοτικής μεταβλητής στους δύο πληθυσμούς από τους οποίους προέρχονται τα μελετώμενα «δείγματα» είναι ίσες. Έτσι, υπολογίζονται η τιμή p, η διαφορά μεταξύ των μέσων των δύο «δειγμάτων» και το 95% διάστημα εμπιστοσύνης για τη διαφορά αυτή. Όπως προαναφέρθηκε, όταν η τιμή p είναι μικρότερη από την τιμή  $\alpha$ , η οποία συνήθως ορίζεται στο επίπεδο του 0,05, τότε απορρίπτεται η μηδενική υπόθεση και η διαφορά μεταξύ των μέσων των δύο πληθυσμών είναι στατιστικώς σημαντική. Σημειώνεται ότι όταν το διάστημα εμπιστοσύνης δεν περιλαμβάνει την τιμή 0, τότε η διαφορά μεταξύ των δύο μέσων είναι στατιστικώς σημαντική. Το διάστημα εμπιστοσύνης όμως δεν πρέπει να χρησιμοποιείται μόνο για τον έλεγχο της μηδενικής υπόθεσης, αλλά κυρίως για την εκτίμηση του τυχαίου σφάλματος σε μια μέτρηση. Έτσι, τα μικρά διαστήματα εμπιστοσύνης υποδηλώνουν μικρότερο τυχαίο σφάλμα και μεγαλύτερη ακρίβεια, ενώ τα μεγάλα διαστήματα εμπιστοσύνης υποδηλώνουν μεγαλύτερο τυχαίο σφάλμα και μικρότερη ακρίβεια. Σε μια μελέτη, π.χ., βρέθηκε ότι το μέσο ανάστημα του «δείγματος» των κοριτσιών ήταν 164 cm, το μέσο ανάστημα του «δείγματος» των αγοριών ήταν 172 cm, η διαφορά μεταξύ των μέσων αναστημάτων ήταν 8 cm, το 95% διάστημα εμπιστοσύνης για τη διαφορά αυτή ήταν 6,6–9,3 cm και η τιμή p του ελέγχου t για ανεξάρτητα «δείγματα» ήταν <0,001. Με βάση τα αποτελέσματα της ανάλυσης των δεδομένων των δύο «δειγμάτων» προέκυψε ότι το μέσο ανάστημα του πληθυσμού των αγοριών ήταν μεγαλύτερο από εκείνο του πληθυσμού των κοριτσιών και μάλιστα η σχέση αυτή ήταν και στατιστικώς σημαντική.

Για τη σύγκριση των μέσων τιμών περισσότερων από δύο πληθυσμών χρησιμοποιείται η *ανάλυση διασποράς μίας κατεύθυνσης* (one-way analysis of variance, ANOVA). Στην περίπτωση, παραδείγματος χάρι, όπου συγκρίνονται οι μέσοι τριών πληθυσμών, η μηδενική υπόθεση είναι ότι οι μέσοι των τριών πληθυσμών από τους οποίους προέρχονται οι παρατηρήσεις των αντίστοιχων τριών «δειγμάτων» είναι ίσοι μεταξύ τους. Η ανάλυση διασποράς μπορεί να ελέγξει τη μηδενική αυτή υπόθεση, αλλά για να πραγματοποιηθούν ανά δύο συγκρίσεις μεταξύ των μέσων χωρίς να αυξηθεί το σφάλμα τύπου I απαιτείται επιπλέον η εφαρμογή μιας post hoc μεθόδου πολλαπλής σύγκρισης. Συνήθως εφαρμόζεται το κριτήριο του *Bonferroni* (για μέχρι 10 συγκρίσεις) και σπανιότερα το κριτήριο του *Tukey* ή του *Scheffe*. Με τον τρόπο αυτόν επιτυγχάνεται η διεξαγωγή πολλαπλών συγκρίσεων χωρίς αύξηση του σφάλματος τύπου I.

Όταν η εξαρτημένη μεταβλητή είναι ποσοτική μεταβλητή που ακολουθεί την κανονική κατανομή –π.χ. δείκτης μάζας σώματος– και οι ανεξάρτητες μεταβλητές είναι δύο ή περισσότερες ποιοτικές μεταβλητές (για παράδειγμα, φύλο, φυσική άσκηση ή όχι, κατανάλωση οινοπνεύματος ή όχι κ.λπ.), τότε χρησιμοποιείται η *πολυμεταβλητή ανάλυση διασποράς* (multivariate analysis of variance, MANOVA), ενώ όταν οι ανεξάρτητες μεταβλητές είναι δύο ή περισσότερες ποσοτικές μεταβλητές (π.χ. ηλικία, ώρες άθλησης, αριθμός σιγάρων κ.ά.), τότε χρησιμοποιείται η *ανάλυση συνδιασποράς* (analysis of covariance, ANCOVA). Η πολυμεταβλητή ανάλυση διασποράς χρησιμοποιείται λιγότερο συχνά και στην περίπτωση περισσότερων από μίας συσχετιζόμενων εξαρτημένων μεταβλητών αποβλέποντας στην πραγματοποίηση ενός μόνο στατιστικού ελέγχου για το σύνολο αυτό των μεταβλητών αντί της διεξαγωγής πολλαπλών ξεχωριστών ελέγχων. Τέλος, όταν οι ανεξάρτητες μεταβλητές (ποιοτικές και ποσοτικές) είναι δύο ή περισσότερες, τότε χρησιμοποιείται η *πολυμεταβλητή ανάλυση συνδιασποράς* (multivariate analysis of covariance, MANCOVA).

Στην περίπτωση της σύγκρισης μέσων τιμών περισσότερων από δύο πληθυσμών ανά ζεύγη χρησιμοποιείται η *ανάλυση διασποράς μίας κατεύθυνσης για επαναλαμβανόμενες μετρήσεις* (repeated measures one-way analysis of variance). Χαρακτηριστικό παράδειγμα της κατηγορίας αυτής αποτελεί η σύγκριση των μέσων τιμών μιας ποσοτικής μεταβλητής σε περισσότερες από δύο χρονικές στιγμές. Η εκτίμηση, για παράδειγμα, της αποτελεσματικότητας μιας δίαιτας σ' ένα μελετώμενο πληθυσμό 50 γυναικών πραγματοποιείται με τη σύγκριση του μέσου βάρους αυτών των γυναικών ένα μήνα, 2 μήνες και 3 μήνες μετά από την έναρξη της δίαιτας. Δηλαδή, πρόκειται για τρία «δείγματα» με επαναλαμβανόμενες μετρήσεις του βάρους. Τα «δείγματα» αυτά αποτελούνται από τις ίδιες, 50 γυναίκες, σε διαφορετικές όμως χρονικές στιγμές.

**5.1.2. Μη παραμετρικές μέθοδοι.** Όπως προαναφέρθηκε,

στην περίπτωση όπου η κατανομή των πληθυσμών από τους οποίους προέρχονται τα μελετώμενα «δείγματα» δεν είναι η κανονική κατανομή, τότε χρησιμοποιούνται οι μη παραμετρικές μέθοδοι, στις οποίες λαμβάνονται υπόψη οι διατάξεις των παρατηρήσεων και όχι οι πραγματικές τιμές των παρατηρήσεων, με αποτέλεσμα να είναι δυνατή η εφαρμογή τους και στην περίπτωση «δειγμάτων» με μικρό αριθμό παρατηρήσεων.<sup>2,5-7</sup> Στις μη παραμετρικές μεθόδους, η μηδενική υπόθεση δεν αφορά στη σύγκριση των μέσων τιμών των πληθυσμών από τους οποίους προέρχονται τα «δείγματα», αλλά στη σύγκριση των διαμέσων.

Στην περίπτωση όπου τα «δείγματα» είναι δύο και ανεξάρτητα μεταξύ τους χρησιμοποιείται ο στατιστικός έλεγχος *Mann-Whitney U* ή, αλλιώς, έλεγχος *αθροίσματος διατάξεων του Wilcoxon* (*Wilcoxon rank-sum test*) ή, αλλιώς, έλεγχος *Mann-Whitney-Wilcoxon*. Στην περίπτωση όπου τα «δείγματα» είναι δύο, αλλά ανά ζεύγη, οπότε για κάθε παρατήρηση στην πρώτη ομάδα υπάρχει μια αντίστοιχη παρατήρηση στη δεύτερη ομάδα, χρησιμοποιείται ο έλεγχος *προσημασμένης διάταξης του Wilcoxon* (*Wilcoxon signed-ranked test*) ή ο έλεγχος *προσήμου* (*sign test*). Στην περίπτωση όπου τα «δείγματα» είναι περισσότερα από δύο και ανεξάρτητα μεταξύ τους χρησιμοποιείται ο έλεγχος *Kruskal-Wallis* ή ο έλεγχος της *διαμέσου* (*median test*), ενώ στην περίπτωση όπου τα «δείγματα» είναι περισσότερα από δύο και ανά ζεύγη χρησιμοποιείται ο έλεγχος *Friedman* ή ο έλεγχος *Kendall's W* ή ο έλεγχος *Cochran*.

## 5.2. Διατάξιμες μεταβλητές

Όπως προαναφέρθηκε, η χρήση των διατάξεων στις μη παραμετρικές μεθόδους παρέχει τη δυνατότητα εφαρμογής τους και στην περίπτωση διατάξιμων δεδομένων και όχι μόνο ποσοτικών. Έτσι, οι στατιστικές μέθοδοι για την ανάλυση διατάξιμων δεδομένων είναι οι μη παραμετρικές μέθοδοι που χρησιμοποιούνται στην περίπτωση των ποσοτικών δεδομένων.

## 5.3. Ονομαστικές μεταβλητές

Στην περίπτωση όπου τα «δείγματα» είναι ανεξάρτητα μεταξύ τους χρησιμοποιείται ο έλεγχος  $\chi^2$  (*chi-square test*), ενώ όταν οι παρατηρήσεις των «δειγμάτων» είναι ανά ζεύγη χρησιμοποιείται ο έλεγχος του *McNemar*.<sup>2,5-7</sup> Με τους ελέγχους  $\chi^2$  και *McNemar* συγκρίνονται οι συχνότητες (ή τα ποσοστά) στις διάφορες κατηγορίες των ονομαστικών μεταβλητών. Παραδείγματος χάρη, η διερεύνηση της σχέσης μεταξύ του φύλου και της συχνότητας εμφάνισης καρδιαγγειακών παθήσεων μπορεί να πραγματοποιηθεί με τον έλεγχο  $\chi^2$ , χωρίς όμως ο έλεγχος αυτός να αποτελεί και τη μοναδική

επιλογή. Έτσι, στο συγκεκριμένο παράδειγμα, όπου η εξαρτημένη μεταβλητή –εμφάνιση ή όχι καρδιαγγειακών παθήσεων– είναι διχοτόμος, προτιμάται η εφαρμογή της λογιστικής παλινδρόμησης και όχι του ελέγχου  $\chi^2$ . Ορισμένοι προτείνουν να χρησιμοποιείται ο έλεγχος  $\chi^2$  με διόρθωση συνέχειας (*continuity correction*) ή, αλλιώς, με διόρθωση του Yates (*Yates' correction*), επιτυγχάνοντας έτσι τη μείωση του σφάλματος τύπου I, αυξάνοντας όμως ταυτόχρονα την πιθανότητα του σφάλματος τύπου II.

Όταν διερευνάται η σχέση μεταξύ δύο διχοτόμων μεταβλητών και ο συνολικός αριθμός των παρατηρήσεων είναι μικρός (<20 παρατηρήσεις), τότε αντί του ελέγχου  $\chi^2$  χρησιμοποιείται ο *ακριβής έλεγχος του Fisher* (*Fisher's exact test*). Ένα δεύτερο κριτήριο εφαρμογής του ελέγχου Fisher είναι η ύπαρξη έστω και μίας αναμενόμενης τιμής (*expected value*) <10 σε ένα από τα τέσσερα κελιά του 2x2 πίνακα συνάφειας (*contingency table*) που δημιουργείται με βάση τις παρατηρούμενες τιμές (*observed values*).

## 6. ΣΥΣΧΕΤΙΣΗ

Το είδος και το μέγεθος της γραμμικής σχέσης μεταξύ δύο ποσοτικών μεταβλητών που ακολουθούν την κανονική κατανομή εκτιμάται με το συντελεστή συσχέτισης του Pearson (*Pearson's coefficient of correlation*), ενώ όταν έστω και μία από τις δύο ποσοτικές μεταβλητές δεν ακολουθεί την κανονική κατανομή, τότε χρησιμοποιείται ο συντελεστής συσχέτισης διατάξεων του Spearman (*Spearman's rank correlation coefficient*).<sup>2,5-7</sup> Επιπλέον, ο συντελεστής συσχέτισης διατάξεων του Spearman χρησιμοποιείται και στην περίπτωση όπου η μία ή και οι δύο μεταβλητές είναι σε διατεταγμένη κλίμακα.

Ο συντελεστής συσχέτισης του Pearson συμβολίζεται με  $r$  όταν αφορά σε ολόκληρο τον πληθυσμό και με  $r$  όταν υπολογίζεται για ένα συγκεκριμένο «δείγμα» (αποτελώντας έτσι εκτίμηση για τον πληθυσμό από τον οποίο προέρχεται το «δείγμα»). Ο συντελεστής συσχέτισης του Pearson δεν έχει μονάδες μέτρησης και λαμβάνει τιμές από -1 έως +1. Οι τιμές -1 έως +1 συμβαίνουν όταν υπάρχει μια τέλεια (αρνητική ή θετική, αντίστοιχα) γραμμική σχέση μεταξύ των δύο μεταβλητών. Όταν ο συντελεστής συσχέτισης λαμβάνει την τιμή 0 σημαίνει ότι δεν υπάρχει γραμμική σχέση μεταξύ των δύο μεταβλητών, αλλά δεν αποκλείεται η ύπαρξη μιας άλλης μη γραμμικής σχέσης. Εάν το μέγεθος της μιας μεταβλητής τείνει να αυξάνεται όπως αυξάνεται το μέγεθος και της άλλης μεταβλητής, τότε υπάρχει θετική συσχέτιση (*positive correlation*) μεταξύ των δύο μεταβλητών και ο συντελεστής συσχέτισης είναι >0. Αντίθετα, εάν το μέγεθος της μιας μεταβλητής τείνει να αυξάνεται

όπως ελαττώνεται το μέγεθος της άλλης μεταβλητής, τότε υπάρχει αρνητική συσχέτιση (negative correlation) μεταξύ των δύο μεταβλητών και ο συντελεστής συσχέτισης είναι  $<0$ . Ο συντελεστής συσχέτισης του Pearson παρουσιάζει ορισμένα μειονεκτήματα, καθώς (α) ποσοτικοποιεί τη σχέση μεταξύ δύο μεταβλητών μόνο στην περίπτωση κατά την οποία η σχέση αυτή είναι γραμμική, (β) είναι πολύ ευαίσθητος σε ακραίες τιμές και η ύπαρξή τους σε αρκετές περιπτώσεις οδηγεί σε λανθασμένα συμπεράσματα, (γ) η εκτιμηθείσα συσχέτιση δεν μπορεί να επεκταθεί πέρα από το παρατηρούμενο εύρος των μεταβλητών και (δ) η ύπαρξη μιας ισχυρής συσχέτισης μεταξύ δύο μεταβλητών δεν υποδηλώνει απαραίτητα και την ύπαρξη μιας σχέσης μεταξύ αιτίας και έκβασης.

Ο συντελεστής συσχέτισης διατάξεων του Spearman όταν υπολογίζεται σε «δείγματα» συμβολίζεται με  $r_s$  και λαμβάνει τιμές από  $-1$  έως  $+1$ , με την ερμηνεία να είναι ίδια με εκείνη στην περίπτωση του συντελεστή συσχέτισης του Pearson. Ο συντελεστής συσχέτισης διατάξεων του Spearman είναι πολύ λιγότερο ευαίσθητος σε ακραίες τιμές από το συντελεστή συσχέτισης του Pearson και μπορεί να χρησιμοποιηθεί όταν η μία ή και οι δύο μεταβλητές είναι σε διατεταγμένη κλίμακα, αλλά επειδή ως μη παραμετρική μέθοδος χρησιμοποιεί τις διατάξεις και όχι τις παρατηρούμενες τιμές, δεν λαμβάνει υπόψη όλη την πληροφορία που είναι γνωστή για μια κατανομή.

## 7. ΓΡΑΜΜΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ

Η απλή γραμμική παλινδρόμηση (simple linear regression), καθώς και η ανάλυση συσχέτισης, αποτελεί μια μέθοδο διερεύνησης της φύσης της σχέσης μεταξύ δύο ποσοτικών μεταβλητών που ακολουθούν την κανονική κατανομή.<sup>4,8,13,14</sup> Η σημαντικότερη διαφορά ανάμεσα στην ανάλυση παλινδρόμησης και την ανάλυση συσχέτισης είναι ότι η πρώτη παρέχει τη δυνατότητα να διερευνηθεί η αλλαγή σε μια ποσοτική μεταβλητή, που ονομάζεται *εξαρτημένη μεταβλητή* (dependent variate) ή, αλλιώς, *απόκριση* (response), η οποία οφείλεται σε μια δεδομένη αλλαγή στη δεύτερη ποσοτική μεταβλητή, γνωστή ως *ανεξάρτητη μεταβλητή* (independent variate) ή, αλλιώς, *επεξηγηματική μεταβλητή* (explanatory variate). Η ανάλυση συσχέτισης δεν έχει τη δυνατότητα να πραγματοποιήσει τη διαφοροποίηση αυτή αντιμετωπίζοντας συμμετρικά τις δύο ποσοτικές μεταβλητές, όπως προαναφέρθηκε εκτενώς. Η ανάλυση παλινδρόμησης αποβλέπει στην πρόβλεψη ή την εκτίμηση της τιμής της εξαρτημένης μεταβλητής που αντιστοιχεί σε μια συγκεκριμένη τιμή της επεξηγηματικής μεταβλητής. Για παράδειγμα, η απλή γραμμική παλινδρόμηση μπορεί να χρησιμοποιηθεί προκειμένου να διερευνηθεί εάν

η συχνότητα εμφάνισης καρκίνου του πνεύμονα εξαρτάται από τον ημερήσιο αριθμό τσιγάρων.

Στην απλή γραμμική παλινδρόμηση, η σχέση μεταξύ των δύο μεταβλητών περιγράφεται από μια ευθεία γραμμή σύμφωνα με την ακόλουθη ισότητα:

$$Y=A_0+A_1X_1+\varepsilon \quad (1)$$

Στην ισότητα 1, το  $Y$  είναι η εξαρτημένη μεταβλητή της απλής γραμμικής παλινδρόμησης, ενώ το  $X_1$  είναι η ανεξάρτητη μεταβλητή. Ουσιαστικά, το  $Y$  αντιστοιχεί στη μελετώμενη έκβαση, ενώ το  $X_1$  αντιστοιχεί στο μελετώμενο προσδιοριστή. Το  $A_0$  είναι η σταθερά της απλής γραμμικής παλινδρόμησης και είναι η μέση τιμή που λαμβάνει η μεταβλητή  $Y$  όταν η μεταβλητή  $X_1$  ισούται με 0. Το  $A_1$  περιγράφει την κλίση της ευθείας γραμμής που συσχετίζει το  $X_1$  με το  $Y$ . Το  $A_1$  είναι ο αριθμός των μονάδων που μεταβάλλεται το  $Y$  κάθε φορά που η τιμή του  $X_1$  μεταβάλλεται κατά μία μονάδα. Το  $\varepsilon$  είναι το τυχαίο σφάλμα που αντιπροσωπεύει την τυχαία απόκλιση από την αναμενόμενη τιμή της εξαρτημένης μεταβλητής  $Y$ . Η μέση τιμή του τυχαίου σφάλματος γενικά θεωρείται ίση με 0. Οι τιμές των  $A_0$  και  $A_1$  κυμαίνονται θεωρητικά από  $-\infty$  έως  $+\infty$ .

Η ισότητα 1 χρησιμοποιείται στην περίπτωση της απλής γραμμικής παλινδρόμησης, επειδή υπάρχει μόνο μία ανεξάρτητη μεταβλητή. Το μοντέλο, ωστόσο, της γραμμικής παλινδρόμησης μπορεί να επεκταθεί, περιλαμβάνοντας περισσότερες από μία ανεξάρτητες μεταβλητές (ποσοτικές ή ποιοτικές), οπότε προκύπτει το μοντέλο της πολλαπλής γραμμικής παλινδρόμησης:

$$Y=A_0+A_1X_1+A_2X_2+A_3X_3+\dots \quad (2)$$

Το μοντέλο της πολλαπλής γραμμικής παλινδρόμησης (multiple linear regression) και γενικότερα τα μοντέλα που περιλαμβάνουν περισσότερες από μία ανεξάρτητες μεταβλητές ταυτόχρονα χρησιμοποιούνται για την εξουδετέρωση των συγχυτών. Στα πολυμεταβλητά μαθηματικά μοντέλα περιλαμβάνονται περισσότεροι από δύο προσδιοριστές, με το αποτέλεσμα του καθενός να μη συγχέεται από τη δράση των υπολοίπων, καθιστώντας έτσι τα μοντέλα αυτά ιδανικά για την εξουδετέρωση των συγχυτών μιας μελέτης. Παραδείγματος χάρη, εάν διερευνάται η σχέση μεταξύ της συχνότητας εμφάνισης καρκίνου του πνεύμονα και του ημερήσιου αριθμού τσιγάρων, της ημερήσιας κατανάλωσης οιοπνεύματος και του φύλου, τότε το μοντέλο της πολλαπλής γραμμικής παλινδρόμησης θα περιλαμβάνει τρεις ανεξάρτητες μεταβλητές. Οι τρεις αυτές μεταβλητές συσχετίζονται μεταξύ τους, οπότε αναμένεται καθένα από τα τρία αυτά χαρακτηριστικά να είναι συγχυτής της σχέσης του άλλου με την εξαρτημένη μεταβλητή, που στη συγκεκριμένη περίπτωση είναι η συχνότητα εμφάνισης

καρκίνου του πνεύμονα. Στο συγκεκριμένο παράδειγμα, η εφαρμογή της ισότητας 2 περιλαμβάνει τρεις προσδιοριστές, τον ημερήσιο αριθμό τσιγάρων ( $X_1$ ), την ημερήσια κατανάλωση οιοπνεύματος ( $X_2$ ) και το φύλο ( $X_3$ ), οπότε υπολογίζονται τρεις συντελεστές,  $A_1$ ,  $A_2$  και  $A_3$ , αντίστοιχα. Οι συντελεστές  $A_1$ ,  $A_2$  και  $A_3$  εκτιμούν το αποτέλεσμα της δράσης του καπνίσματος, του οιοπνεύματος και του φύλου, αντίστοιχα, εξουδετερώνοντας ταυτόχρονα τη σύγχυση που ενδεχομένως υπάρχει στη σχέση μεταξύ της ενδεικτικής κατηγορίας του κάθε προσδιοριστή ξεχωριστά και της συχνότητας εμφάνισης του καρκίνου του πνεύμονα.

Μαθηματικά, τουλάχιστον, δεν υπάρχει περιορισμός στον αριθμό των ανεξάρτητων μεταβλητών που μπορούν να χρησιμοποιηθούν στην πολλαπλή γραμμική παλινδρόμηση, αν και μικρός αριθμός δεδομένων επιτρέπει να χρησιμοποιούνται λίγες μόνο ανεξάρτητες μεταβλητές. Ορισμένοι ερευνητές θεωρούν ότι για κάθε ανεξάρτητη μεταβλητή πρέπει να υπάρχουν τουλάχιστον 10 παρατηρήσεις στο μελετώμενο «δείγμα», δηλαδή στην περίπτωση πέντε ανεξάρτητων μεταβλητών θα πρέπει να υπάρχουν τουλάχιστον 50 παρατηρήσεις, ενώ ορισμένοι άλλοι θεωρούν ότι για κάθε ανεξάρτητη μεταβλητή θα πρέπει να υπάρχουν τουλάχιστον 40 παρατηρήσεις.

**8. ΛΟΓΙΣΤΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ**

Όταν η εξαρτημένη μεταβλητή δεν είναι ποσοτική αλλά διχοτόμος (όπως π.χ. η εμφάνιση ή όχι μιας πάθησης), οπότε μπορεί να λάβει μόνο δύο πιθανές τιμές, τότε χρησιμοποιείται η λογιστική παλινδρόμηση (logistic regression).<sup>1,4,8,15,16</sup> Στην περίπτωση της απλής λογιστικής παλινδρόμησης ισχύει η εξής ισότητα:

$$\ln \left[ \frac{Y}{1-Y} \right] = A_0 + A_1 X_1 \tag{3}$$

Προφανώς, όπως και σε άλλα γραμμικά μοντέλα, είναι δυνατόν να συμπεριληφθούν περισσότερες από μία ανεξάρτητες μεταβλητές (ποσοτικές ή ποιοτικές), οπότε προκύπτει το μοντέλο της πολλαπλής λογιστικής παλινδρόμησης:

$$\ln \left[ \frac{Y}{1-Y} \right] = A_0 + A_1 X_1 + A_2 X_2 + A_3 X_3 + \dots \tag{3}$$

Στις ισότητες 3 και 4, το  $Y$  είναι η πιθανότητα εμφάνισης της μελετώμενης έκβασης. Έτσι, η λογιστική παλινδρόμηση χρησιμοποιείται για την εκτίμηση της πιθανότητας εμφάνισης μιας έκβασης με βάση την ύπαρξη διαφόρων προσδιοριστών. Σημειώνεται ότι στη λογιστική παλινδρόμηση ο αντιλογάριθμος ( $e^{A_i}$ ) του συντελεστή παλινδρόμησης ( $A_i$ ) μιας διχοτόμου ή, αλλιώς, ενδεικτικής ανεξάρτητης μεταβλητής ( $X_i$ ) αποτελεί εκτίμηση του λόγου των odds στους εκτεθειμένους σε σχέση με τους μη εκτεθειμένους.

**9. ΑΝΑΛΥΣΗ ΕΠΙΒΙΩΣΗΣ**

Σε ορισμένες περιπτώσεις, η εξαρτημένη μεταβλητή είναι το χρονικό διάστημα ανάμεσα σε ένα συμβατικό χρονικό σημείο έναρξης της μέτρησης –για παράδειγμα, η στιγμή της διάγνωσης, η πραγματοποίηση μιας χειρουργικής επέμβασης, η έναρξη μιας φαρμακευτικής αγωγής, η μεταμόσχευση ενός μοσχεύματος κ.ά.– και την εμφάνιση της μελετώμενης έκβασης. Στις μελέτες όπου διεξάγεται ανάλυση επιβίωσης, οι συμμετέχοντες είναι ήδη πάσχοντες, οπότε η μελετώμενη έκβαση δεν είναι η εμφάνιση μιας πάθησης, αλλά ο θάνατος, η ίαση, η εμφάνιση καταλοιπών, η διασπορά ενός καρκίνου, η απόρριψη ενός μοσχεύματος, η εμφάνιση μιας λοίμωξης κ.ά. Το χρονικό διάστημα ανάμεσα στο συμβατικό χρονικό σημείο έναρξης της μέτρησης και την εμφάνιση της μελετώμενης έκβασης ονομάζεται *χρόνος επιβίωσης* (survival time). Μολονότι τα δεδομένα είναι ποσοτικά, οι κατανομές των χρόνων επιβίωσης σε ελάχιστες μόνο περιπτώσεις ακολουθούν την κανονική κατανομή, τείνοντας να έχουν δεξιά ασυμμετρία.

Στην πλειοψηφία των περιπτώσεων όπου διεξάγεται ανάλυση επιβίωσης, η μελετώμενη έκβαση δεν εμφανίζεται σε όλους τους συμμετέχοντες. Έτσι, η ανάλυση των δεδομένων πραγματοποιείται χωρίς να έχει συμβεί η μελετώμενη έκβαση σε όλους τους πάσχοντες. Υπάρχουν, εξάλλου, και πάσχοντες που αποτελούν τις απώλειες μιας μελέτης γιατί χάθηκε η επικοινωνία μαζί τους είτε λόγω μετακίνησής τους πριν από την ολοκλήρωση της μελέτης είτε λόγω άρνησής τους να συμμετάσχουν έως την ολοκλήρωση της μελέτης. Η ελλιπής παρατήρηση ενός συμμετέχοντα έως την εμφάνιση της μελετώμενης έκβασης είναι γνωστή ως «περικεκομμένη» (censored). Η ύπαρξη «περικεκομμένων» παρατηρήσεων διαφοροποιεί σημαντικά την ανάλυση δεδομένων που αφορά στην επιβίωση.

Η ανάλυση επιβίωσης (survival analysis)<sup>1,4-8,17</sup> αποβλέπει κυρίως στη δημιουργία των κλινικών πινάκων επιβίωσης, που χρησιμοποιούνται για την εκτίμηση της πρόγνωσης σε παθήσεις οι οποίες δυνητικά είναι θανατηφόρες και έχουν σχετικά μεγάλη χρονική διάρκεια, ιδιαίτερα στις κακοήθεις νεοπλασίες. Οι κλινικοί πίνακες επιβίωσης χρησιμοποιούνται ευρύτατα στις κλινικές δοκιμές (clinical trials), οι οποίες έχουν ως σκοπό να διαπιστώσουν αν μια θεραπευτική παρέμβαση είναι αποδοτικότερη έναντι μιας άλλης. Δύο ομάδες πασχόντων, π.χ., που προέκυψαν έπειτα από τυχαιοποίηση, δέχονται μόσχευμα νεφρού και η μια ομάδα λαμβάνει τη φαρμακευτική αγωγή Α, ενώ η άλλη ομάδα λαμβάνει τη φαρμακευτική αγωγή Β. Σκοπός της μελέτης αυτής είναι η διαπίστωση της αποδοτικότερης φαρμακευτικής αγωγής. Στην περίπτωση αυτή, η μελετώμενη έκβαση είναι η απόρριψη του νεφρικού μοσχεύ-

ματος. Για την πραγματοποίηση της σύγκρισης απαιτείται να κατασκευαστούν οι κλινικοί πίνακες επιβίωσης για τις δύο ομάδες πασχόντων και να συγκριθούν οι αντίστοιχες καμπύλες επιβίωσης (survival curves).

Για τη σύγκριση δύο καμπυλών επιβίωσης, και κατ'επέκταση δύο θεραπευτικών παρεμβάσεων, χρησιμοποιείται, συνήθως, ο στατιστικός έλεγχος *log-rank*.

Οι κλινικοί πίνακες επιβίωσης κατασκευάζονται είτε με την εφαρμογή της ασφαλιστικής μεθόδου (actuarial method)<sup>18</sup> είτε με την εφαρμογή της μεθόδου Kaplan-Meier.<sup>19</sup>

Η χρήση της ασφαλιστικής μεθόδου δεν είναι δυνατή όταν ο αριθμός των συμμετεχόντων σε μια μελέτη είναι μικρός, όταν τα επιμέρους χρονικά διαστήματα στα οποία χωρίζεται ο συνολικός χρόνος διεξαγωγής της μελέτης είναι σχετικά μεγάλα, όταν οι απώλειες είναι πολλές και όταν δεν είναι λογικό να θεωρηθεί ότι οι απώλειες κατανέμονται ομοιόμορφα σε ένα δεδομένο χρονικό διάστημα. Στις περιπτώσεις αυτές είναι απαραίτητη η εφαρμογή της μεθόδου Kaplan-Meier.

Η μέθοδος Kaplan-Meier είναι παρόμοια με την ασφαλιστική, αλλά διαφέρει στο ότι δεν απαιτείται να χωριστεί ο συνολικός χρόνος σε ενδιάμεσα χρονικά διαστήματα. Για το λόγο αυτόν, η μέθοδος Kaplan-Meier είναι κατάλληλη, κυρίως, για μελέτες με μικρό αριθμό πασχόντων. Η μέθοδος Kaplan-Meier απαιτεί λιγότερο πολύπλοκους υπολογισμούς από την ασφαλιστική, κυρίως επειδή η εκτίμηση της επιβίωσης πραγματοποιείται κάθε φορά που ένας πάσχοντας εμφανίζει τη μελετώμενη έκβαση και όχι σε συγκεκριμένα χρονικά διαστήματα, όπως συμβαίνει στην ασφαλιστική. Η μέθοδος Kaplan-Meier, σε αντίθεση με την ασφαλιστική, δεν λαμβάνει υπόψη της κατά τους υπολογισμούς τις απώλειες της μελέτης.

Το μοντέλο αναλογικών κινδύνων του Cox (Cox's proportional hazards model) ή απλούστερα το μοντέλο του Cox αποτελεί μια γενικότερη μέθοδο ανάλυσης της επιβίωσης, στην οποία ενδεχομένως να συμπεριλαμβάνονται περισ-

σότεροι από ένας προσδιοριστές και όχι μόνο το είδος της θεραπευτικής παρέμβασης.

## 10. ΣΥΜΠΕΡΑΣΜΑΤΑ

Ο κατάλληλος μεθοδολογικός σχεδιασμός και ο περιορισμός στο μεγαλύτερο δυνατό βαθμό του τυχαίου και του συστηματικού σφάλματος –και ιδιαίτερα των συγχυτών– αποτελούν το πρώτο και πλέον καθοριστικό βήμα στη διεξαγωγή μιας μελέτης και την εξαγωγή ασφαλών συμπερασμάτων. Εξίσου σημαντική βεβαίως είναι και η στατιστική επεξεργασία των δεδομένων με την επιλογή της κατάλληλης μεθόδου ανάλογα με το είδος τόσο της ερευνητικής υπόθεσης όσο και των δεδομένων. Σημειώνεται ότι η ύπαρξη των πολυμεταβλητών μαθηματικών μοντέλων επιτρέπει την εξουδετέρωση έως ένα βαθμό της σύγχυσης που προκαλείται από διάφορα χαρακτηριστικά, αρκεί ο αριθμός των παρατηρήσεων να είναι επαρκώς μεγάλος. Η γνώση του σχεδιασμού μιας μελέτης και δευτερευόντως της ανάλυσης των δεδομένων είναι ζητήματα πρωταρχικής σημασίας για τους επιστήμονες υγείας. Και αν σωστά θεωρηθεί ότι η ανάλυση των δεδομένων μπορεί να αποτελέσει αντικείμενο και των στατιστικών, τότε σίγουρα ο ερευνητικός σχεδιασμός αποτελεί αντικείμενο των επιστημόνων υγείας.

Σε αρκετές περιπτώσεις, η παρουσίαση των αποτελεσμάτων περιορίζεται στην αναφορά των παρατηρούμενων επιπέδων στατιστικής σημαντικότητας (τιμές *p*) αγνοώντας τον καθοριστικό ρόλο που παίζει η παράθεση των διαστημάτων εμπιστοσύνης στην ερμηνεία του τυχαίου σφάλματος. Επιπλέον, οι τιμές *p* δεν προσδιορίζουν το μέγεθος της σχέσης ανάμεσα στο μελετώμενο προσδιοριστή και την έκβαση, αποτελώντας απλά ένα μέτρο της συμβατότητας μεταξύ της μηδενικής υπόθεσης και των δεδομένων μιας μελέτης. Τα μέτρα σχέσης είναι εκείνα που καθορίζουν το μέγεθος της σχέσης μεταξύ προσδιοριστή και έκβασης και γι' αυτό πρέπει να καθορίζεται με σαφήνεια ο τρόπος υπολογισμού τους.

## ABSTRACT

### Statistical methods of data analysis

P. GALANIS

Center for Health Services, Management and Evaluation, Department of Nursing, University of Athens, Athens, Greece

Archives of Hellenic Medicine 2009, 26(5):699–711

Measures of association are the object of epidemiologic studies and are used to quantify the relation between the determinant under study and the outcome, while statistical significance tests are used to identify statistically sig-

nificant relationships between these two variates. Statistical analysis of data is achieved using a variety of methods. Choice of the appropriate method depends on the scientific hypothesis and the kind of data. Data commonly are distinguished into qualitative (nominal and ordinal data) and quantitative (interval scale and ratio scale data). Concerning quantitative variates which follow a normal distribution, parametric methods should be applied. The independent samples t-test is used to determine whether two population means are significantly different, but for this to apply, the distribution of both populations must be normal and the two samples must not be related to each other. One-way analysis of variance is employed to determine whether more than two population means are different. Paired samples t-test and repeated measures of one-way analysis of variance are the appropriate tests when the samples (two or more than two respectively) are related to each other. Multivariate mathematical models are used to evaluate the effect of one or more characteristics while simultaneously controlling for possible confounding effects of other characteristics. In a multivariate model, the inclusion of several variates results in each term being unconfounded by the other terms. In epidemiology, the most frequently used multivariate models are the multiple linear and the logistic regression models. The outcome (or dependent variate) in linear regression is a quantitative variate, while in logistic regression it is a dichotomous variate.

**Key words:** Data analysis, Linear regression, Logistic regression, Parametric methods, Survival analysis

## Βιβλιογραφία

1. ΣΠΑΡΟΣ Λ, ΓΑΛΑΝΗΣ Π. *Δοκίμια επιδημιολογίας*. Εκδόσεις Παρισιάνου, Αθήνα, 2006
2. ΤΡΙΧΟΠΟΥΛΟΣ Δ, ΤΖΩΝΟΥ Α, ΚΑΤΣΟΥΓΙΑΝΝΗ Κ. *Βιοστατιστική*. Εκδόσεις Παρισιάνου, Αθήνα, 2000
3. ΓΑΛΑΝΗΣ Π, ΣΠΑΡΟΣ Λ. Ανάλυση δεδομένων: Μη μπαγιασιανή προσέγγιση. *Αρχ Ελλ Ιατρ* 2005, 22:377–391
4. ROTHMAN KJ, GREENLAND S, LASH TL. *Modern epidemiology*. 3rd ed. Lippincott Williams & Wilkins, Philadelphia, 2008
5. BOSLAUGH S, WATTERS P. *Statistics in a nutshell*. O'Reilly Media Inc, Cambridge, 2008
6. DAWSON B, TRAPP R. *Basic and clinical biostatistics*. 4th ed. McGraw-Hill, New York, 2004
7. FOSTER J, BARKUS E, YAVORSKY C. *Understanding and using advanced statistics*. SAGE Publications, London, 2006
8. ROTHMAN KJ. *Epidemiology. An introduction*. Oxford University Press, New York, 2002
9. GOODMAN S. P values, hypothesis tests, and likelihood: Implications for epidemiology of a neglected historical debate. *Am J Epidemiol* 1993, 137:485–496
10. GOODMAN S. A comment on replication, p-values and evidence. *Stat Med* 1992, 11:875–879
11. GOODMAN S, ROYALL R. Evidence and scientific research. *Am J Public Health* 1988, 78:1568–1574
12. ΓΕΜΤΟΣ Π. *Μεθοδολογία των κοινωνικών επιστημών. Μεταθεωρία και ιδεολογική κριτική των επιστημών του ανθρώπου*. 3η έκδοση, 2ος τόμος. Εκδόσεις Παπαζήση, Αθήνα, 1987:180–197
13. STRIKE P. *Statistical methods in laboratory medicine*. Butterworth Heinemann, Cambridge, 1991
14. MIETTINEN OS. *Theoretical epidemiology. Principles of occurrence research in medicine*. John Wiley & Sons, New York, 1985
15. BLAND M. *An introduction to medical statistics*. 2nd ed. Oxford Medical Publications, Oxford, 1996
16. HOSMER D, LEMESHOW S. *Applied logistic regression*. John Wiley & Sons, New York, 1989
17. ΓΑΛΑΝΗΣ Π, ΣΠΑΡΟΣ Λ. Κλινικοί πίνακες επιβίωσης. *Αρχ Ελλ Ιατρ* 2006, 23:393–400
18. ANDERSEN P, KEIDING N. Survival analysis, overview. In: Armitage P, Colton T (eds) *Encyclopaedia of biostatistics*. John Wiley & Sons, New York, 1998:4452–4461
19. BORGAN Ø. Kaplan-Meier estimator. In: Armitage P, Colton T (eds) *Encyclopaedia of biostatistics*. John Wiley & Sons, New York, 1998:2154–2160

*Corresponding author:*

P. Galanis, 14 Dikis street, GR-157 73 Athens, Greece  
e-mail: pegalan@nurs.uoa.gr