ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

ΣΧΟΛΗ ΕΠΙΣΤΗΜΩΝ & ΤΕΧΝΟΛΟΓΙΑΣ ΤΗΣ ΠΛΗΡΟΦΟΡΙΑΣ

ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ

ΜΕΤΑΠΤΥΧΙΑΚΟ

ΣΥΓΚΡΙΣΗ ΜΕΛΕΤΩΝ ΜΕΣΩ ΜΕΤΡΗΣΗΣ ΣΥΜΦΩΝΙΑΣ ΜΕΘΟΔΩΝ

Σοφία Νικολάου Ζαφείρη

ΕΡΓΑΣΙΑ

Που υποβλήθηκε στο Τμήμα Στατιστικής
του Οικονομικού Πανεπιστημίου Αθηνών
ως μέρος των απαιτήσεων για την απόκτηση
Μεταπτυχιακού Διπλώματος
Συμπληρωματικής Ειδίκευσης στη Στατιστική
Μερικής Παρακολούθησης (Part-time)

Αθήνα
Ιανουάριος 2017

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ


ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΕΥΧΑΡΙΣΤΙΕΣ

Αισθάνομαι την ανάγκη να εκφράσω τις ειλικρινείς ευχαριστίες μου στον επιβλέποντα της μεταπτυχιακής μου εργασίας Καθηγητή κύριο Βασίλειο Βασδέκη για την καθοδήγησή του, αλλά και την ανεξάντλητη υπομονή και κατανόηση που υπέδειξε κατά τη διάρκεια εκπόνησής της. Επίσης, θέλω να ευχαριστήσω ιδιαιτέρως την οικογένειά μου και τους ανθρώπους που στάθηκαν δίπλα μου όλα αυτά τα χρόνια.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ


ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΒΙΟΓΡΑΦΙΚΟ ΣΗΜΕΙΩΜΑ

Είμαι απόφοιτος της Σχολής Εφαρμοσμένων Μαθηματικών και Φυσικών Επιστημών του Εθνικού Μετσόβειου Πολυτεχνείου με κατεύθυνση Μαθηματικά, τελική βαθμολογία 7.0/10 και έτος αποφοίτησης το 2011. Τον Οκτώβριο του 2012 έγινα δεκτή στο μεταπτυχιακό πρόγραμμα του Τμήματος Στατιστικής του Οικονομικού Πανεπιστημίου Αθηνών και συγκεκριμένα στο πρόγραμμα «Συμπληρωματικής Ειδίκευσης στη Στατιστική» και με την παρούσα διπλωματική ολοκληρώνω τις σπουδές μου σε αυτό.

Από το Φεβρουάριο του 2012 εργάζομαι σε φαρμακευτική βιομηχανία στο τμήμα του Market Access.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ


ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ABSTRACT

Sophia Zafeiri

Comparison studies through measuring agreement between methods

January 2017

In clinical measurements comparison between two methods can be based on the differences arisen from observations made on the same subjects. By creating an interval within which the 95% of differences between measurements by the two methods are expected to lie, estimated by mean difference ±1.96 standard deviation of these differences, we reach 95% limits of agreement. Using graphical methods we explore the way in which they can be used to further investigate the assumptions of each method and we provide confidence intervals. Moreover, we examine data where there is a relationship between differences and magnitude of measurements using the method of logarithmic transformation and regression method, in order to extend our basic approach. We also estimate the 95% limits of agreement between the two methods expanding on repeated measurements with equal numbers of replicates by each method on each subject, unequal numbers of replicates and replicated data in pairs where the underlying value of the quantity being measured is changing. Finally, we examine a non-parametric approach to comparing methods.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ


ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΠΕΡΙΛΗΨΗ

Σοφία Ζαφείρη

ΣΥΓΚΡΙΣΗ ΜΕΛΕΤΩΝ ΜΕΣΩ ΜΕΤΡΗΣΗΣ ΣΥΜΦΩΝΙΑΣ ΜΕΘΟΔΩΝ

Ιανουάριος 2017

Η συμφωνία μεταξύ δύο μεθόδων κλινικών μετρήσεων, μπορεί να ποσοτικοποιηθεί χρησιμοποιώντας τις διαφορές μεταξύ των μετρήσεων που πραγματοποιήθηκαν από τις δύο υπό σύγκριση μεθόδους στα ίδια υποκείμενα. Το 95% όριο συμφωνίας, εκτιμάται από τη μέση τιμή της διαφοράς ± 1.96 της τυπικής τους απόκλισης, παρέχοντας έτσι ένα διάστημα εντός του οποίου αναμένεται να βρίσκονται το 95% των διαφορών μεταξύ των μετρήσεων που πραγματοποιήθηκαν από τις δύο υπό σύγκριση μεθόδους. Περιγράφονται οι τρόποι με τους οποίους οι γραφικές παραστάσεις μπορούν να χρησιμοποιηθούν ώστε να ελεγχθούν οι υποθέσεις των μεθόδων, και επίσης υπολογίζονται τα διαστήματα εμπιστοσύνης. Παρουσιάζεται η ανάλυση των δεδομένων όταν υπάρχει σχέση μεταξύ των διαφορών και του μεγέθους των μετρήσεων με τη χρήση λογαριθμικού μετασχηματισμού αλλά και με την προσέγγιση μέσω παλινδρόμησης. Εξετάζεται η σημαντικότητα της επαναληψιμότητας των μετρήσεων για κάθε μέθοδο ξεχωριστά και συγκρίνονται οι εκτιμήσεις τους με τα όρια συμφωνίας. Επιπλέον, μελετώνται τα όρια συμφωνίας για δεδομένα με επαναλαμβανόμενες μετρήσεις, αναλύοντας ξεχωριστά τις περιπτώσεις με ίσο αλλά και άνισο πλήθος επαναλήψεων για κάθε μέθοδο σε κάθε υποκείμενο και επίσης εξετάζεται η περίπτωση επαναλαμβανόμενων μετρήσεων σε ζεύγη για τα οποία η πραγματική τιμή των μετρήσεων μεταβάλλεται. Τέλος, παρουσιάζεται μια μη παραμετρική μέθοδος σύγκρισης μεθόδων μέτρησης.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ


ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΚΑΤΑΛΟΓΟΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

Σελίδα

ΕΥΧΑΡΙΣΤΙΕΣ I

ΒΙΟΓΡΑΦΙΚΟ ΣΗΜΕΙΩΜΑ III

ABSTRACT V

ΠΕΡΙΛΗΨΗ VII

ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ XI

ΚΑΤΑΛΟΓΟΣ ΓΡΑΦΗΜΑΤΩΝ XIII

Κεφάλαιο 1: Το πρόβλημα της σύγκρισης των μεθόδων 1

Κεφάλαιο 2: Τα όρια συμφωνίας 3

2.1. Γραφική Απεικόνιση της συμφωνίας 9

2.2. Ακρίβεια των εκτιμώμενων ορίων συμφωνίας 12

Κεφάλαιο 3. Σχέση διαφοράς και μεγέθους 15

3.1. Λογαριθμικός μετασχηματισμός 17

3.2. Προσέγγιση παλινδρόμησης για μη ομοιόμορφες διαφορές 21

Κεφάλαιο 4. Η σημαντικότητα της επαναληψιμότητας 29

4.1. Εκτιμώντας την επαναληψιμότητα 30

Κεφάλαιο 5. Μέτρηση συμφωνίας μεθόδων με τη χρήση επαναλαμβανόμενων μετρήσεων 33

5.1. Ισάριθμο πλήθος επαναλήψεων 33

5.2. Άνισο πλήθος επαναλήψεων 38

5.3. Επαναλαμβανόμενες μετρήσεις σε ζεύγη 42

Κεφάλαιο 6. Μη παραμετρική προσέγγιση στη σύγκριση μεθόδων 45

Αναφορές 49

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ


ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ

Πίνακας Σελίδα

Πίνακας 2.1. Μετρήσεις συστολικής αρτηριακής πίεσης οι οποίες πραγματοποιήθηκαν ταυτόχρονα από δύο παρατηρητές (J και R) και από ένα ημιαυτόματο μηχάνημα καταγραφής αρτηριακής συστολικής πίεσης (S), πραγματοποιώντας ο καθένας 3 διαδοχικές μετρήσεις (τα δεδομένα παρέχονται από τον Dr E O' Brien, του Altman και Bland) 4

Πίνακας 2.2. Μέση τιμή και τυπική απόκλιση των διαφορών των δύο μεθόδων J1 και S1. 6

Πίνακας 2.3. Μέση τιμή και τυπική απόκλιση των διαφορών των δύο μεθόδων J1 και S1,έχοντας εξαιρέσει τα αντικείμενα 78 και 80. 8

Πίνακας 2.4. Spearman's rank correlations, μεταξύ των απόλυτων τιμών των διαφορών και της μέσης τιμής των μετρήσεων των δύο μεθόδων. 10

Πίνακας 3.1. Μετρήσεις όγκου πλάσματος εκφρασμένες ως ποσοστό των κανονικών τιμών 99 υποκειμένων, χρησιμοποιώντας δύο διαφορετικές μεθόδους Nadler και Hurley. (τα δεδομένα προέρχονται από C Dore). 15

Πίνακας 3.2. Μέση τιμή και τυπική απόκλιση των διαφορών (Nadler-Hurley) μετά το λογαριθμικό μετασχηματισμό. 18

Πίνακας 3.3. Περιεκτικότητες λίπους στο ανθρώπινο γάλα με την ενζυμική διαδικασία καθορισμού των τριγλυκεριδίων, και με την τυπική μέθοδο Gerber (g/100ml). 23

Πίνακας 3.4. Ανάλυση παλινδρόμησης του μοντέλου . $\widehat{D} = b_0 + b_1 A$ για τα δεδομένα του Πίνακα 3.3. 25

Πίνακας 3.5. Ανάλυση παλινδρόμησης του μοντέλου . $\widehat{R} = c_0 + c_1 A$ για τα δεδομένα του Πίνακα 3.3. 27

Πίνακας 4.1. Ανάλυση διακύμανσης των μεθόδων μέτρησης συστολικής αρτηριακής πίεσης από τους παρατηρητές J, R και το μηχάνημα S. 31

Πίνακας 5.1. Μέση τιμή και διακύμανση των διαφορών των μέσων τιμών των υποκειμένων με τις μεθόδους J και S. 35

Πίνακας 5.2. Δεδομένα καρδιακής παροχής 12 ασθενών, τα οποία μετρήθηκαν με δύο μεθόδους RV και IS (Dr LS Bowling) 39

Πίνακας 5.3. Εκτίμηση των εντός των υποκειμένων διακυμάνσεων για τις μεθόδους RV και IC. 41

Πίνακας 5.4. Μέση τιμή και διακύμανση των διαφορών των μέσων τιμών των υποκειμένων (RV – IC). 42

Πίνακας 6.1 Αξιολόγηση των μηχανημάτων μέτρησης αρτηριακής πίεσης, βασιζόμενοι στις διαφορές μεταξύ των μετρήσεων (σε mmHg) των μηχανημάτων και σφυγμομανόμετρου. 46

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ


ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΚΑΤΑΛΟΓΟΣ ΓΡΑΦΗΜΑΤΩΝ

Γράφημα Σελίδα

Γράφημα 1. Συστολική αρτηριακή πίεση, καταμετρημένη από τον Παρατηρητή J χρησιμοποιώντας συγμομανόμετρο και από το μηχάνημα S. 4

Γράφημα 2. Διαφορά των μετρήσεων της συστολικής αρτηριακής πίεσης, του Παρατηρητή J και του μηχανήματος S, σε σχέση με τη μέση τιμή τους. 11

Γράφημα 3. Διαφορά των μετρήσεων της συστολικής αρτηριακής πίεσης, του Παρατηρητή J και του μηχανήματος S σε σχέση με τη μέση τιμή του, με το 95% όριο συμφωνίας. 12

Γράφημα 4. Μετρήσεις όγκου πλάσματος σύμφωνα με τα στοιχεία του Πίνακα 3.1 16

Γράφημα 5. Διαφορές και αναμενόμενες τιμές των μετρήσεων των δύο μεθόδων. 17

Γράφημα 6-α. Λογαριθμοποιημένες μετρήσεις όγκου πλάσματος σύμφωνα με τα στοιχεία του Πίνακα 3.1 19

Γράφημα 6-β. Διαφορές και αναμενόμενες τιμές των λογαριθμοποιημένων μετρήσεων των δύο μεθόδων με το 95% όριο συμφωνίας. 19

Γράφημα 7. Λόγος όγκου πλάσματος (N/H) σε σχέση με τις αναμενόμενες τιμές των μετρήσεων των δύο μεθόδων με το 95% όριο συμφωνίας. 21

Γράφημα 8. Περιεκτικότητα λίπους στο ανθρώπινο γάλα με τη μέθοδο που καθορίζεται από την ενζυμική διαδικασία (Trig) και με την τυπική μέθοδο Gerber (g/100ml) 24

Γράφημα 9. Διαφορές των μετρήσεων των δύο μεθόδων (Trig – Gerber) σε σχέση με την αναμενόμενη τιμή τους. 24

Γράφημα 10. 95% όριο συμφωνίας βασισμένο στην παλινδρόμηση των δύο μεθόδων Trig και Gerber. 27

Γράφημα 11. Τυπική απόκλιση και μέση τιμή των μετρήσεων κάθε υποκειμένου με τις μεθόδους RV και IC ξεχωριστά. 40

Γράφημα 12. Διαφορές των μέσων τιμών των μετρήσεων σε σχέση με τη μέση τιμή για κάθε υποκείμενο. 41

Γράφημα 13. Διαφορά των μετρήσεων της συστολικής αρτηριακής πίεσης, του Παρατηρητή J και του μηχανήματος S σε σχέση με τη μέση τιμή τους με την προσθήκη του διαστήματος ±15mmHg. 47

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ


ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

Κεφάλαιο 1: Το πρόβλημα της σύγκρισης των μεθόδων

Στην κλινική ιατρική καλούμαστε συχνά να μετρήσουμε ποσότητες που αφορούν το ανθρώπινο σώμα, όπως την ένταση του καρδιακού παλμού ή την πίεση ενός ασθενούς. Μια τέτοια μέτρηση όμως, θα ήταν εξαιρετικά δύσκολο έως απίθανο να συμβεί χωρίς να επηρεαστεί ο ίδιος ο ασθενής από τη διαδικασία της μέτρησης, με αποτέλεσμα οι πραγματικές τιμές να παραμένουν άγνωστες. Χρησιμοποιούμε λοιπόν έμμεσες μεθόδους μέτρησης, και όταν προτείνεται μια νέα μέθοδος είμαστε σε θέση να εκτιμήσουμε την εγκυρότητά της, μόνο μέσω σύγκρισης με άλλες ήδη καθιερωμένες μεθόδους και όχι μέσω σύγκρισης με τις πραγματικές τιμές των ποσοτήτων που μετράμε. Αποτέλεσμα του γεγονότος αυτού, είναι να μην είμαστε σε θέση να γνωρίζουμε με σιγουριά αν οι μετρήσεις των μεθόδων είναι αναμφισβήτητα σωστές, οπότε προσπαθούμε να εκτιμήσουμε το βαθμό συμφωνίας μεταξύ τους. Η καθιερωμένη μέθοδος που χρησιμοποιείται είναι κάποιες φορές γνωστή και ως «χρυσός κανόνας», χωρίς όμως αυτό να υποδηλώνει -τουλάχιστον δε θα έπρεπε- ότι οι μετρήσεις της δεν περιέχουν σφάλματα.

Αξίζει να σημειωθεί ότι, αναπόφευκτα, πάντα θα υπάρχει κάποια έλλειψη συμφωνίας μεταξύ των μετρήσεων διαφορετικών μεθόδων και αυτό που έχει σημασία είναι το μέγεθος αυτής της ασυμφωνίας. Θέλουμε να γνωρίζουμε κατά πόσο οι μετρήσεις της νέας προς εξέταση μεθόδου είναι πιθανό να διαφέρουν από της παλιάς, ώστε αν η διαφορά που προκύπτει δεν οδηγεί σε λανθασμένα κλινικά συμπεράσματα να αντικαταστήσουμε την παλιά με τη νέα ή να έχουμε τη δυνατότητα να τις χρησιμοποιούμε και τις δύο εναλλακτικά[1]. Για παράδειγμα, εάν ένα νέο μηχάνημα μέτρησης αρτηριακής πίεσης είναι απίθανο να δώσει μετρήσεις για ένα υποκείμενο, οι οποίες να διαφέρουν περισσότερο από, ας πούμε, 10mmHg σε σχέση με τις μετρήσεις που λαμβάνουμε χρησιμοποιώντας ένα σφυγμομανόμετρο, θα μπορούσαμε να βασιστούμε στις μετρήσεις του νέου μηχανήματος, καθώς διαφορές μικρότερες από 10mmHg στις μετρήσεις, δεν θα επηρέαζαν ουσιαστικά τη λήψη αποφάσεων. Από την άλλη μεριά, διαφορές ίσες ή μεγαλύτερες των 30mmHg δεν θα μπορούσαν να θεωρηθούν ως σφάλμα αυτού του μεγέθους

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

και θα οδηγούσαν σε πιθανώς αλλαγές στην αντιμετώπιση του υποκειμένου. Το πόσο μπορούν να αποκλίνουν οι μετρήσεις, χωρίς να οδηγούν σε πρόβλημα, εξαρτάται από το είδος των μετρήσεων που κάνουμε και είναι ένα ερώτημα κλινικής κρίσεως. Οι στατιστικές μέθοδοι δεν είναι σε θέση να απαντήσουν ένα τέτοιο ερώτημα. Μέθοδοι οι οποίες συμφωνούν σε ικανοποιητικό βαθμό για κάποιο συγκεκριμένο σκοπό ίσως να μην ικανοποιούν τις απαιτήσεις για κάποιον άλλο σκοπό. Ιδανικά θα πρέπει να καθορίζουμε εξ αρχής την συμφωνία που μας ικανοποιεί σε κάθε περίπτωση.

Στη συγκεκριμένη εργασία περιγράφουμε μια προσέγγιση ανάλυσης τέτοιων δεδομένων, χρησιμοποιώντας απλές γραφικές μεθόδους και βασικούς στατιστικούς υπολογισμούς[2,3]. Η προσέγγισή μας βασίζεται στην ποσοτικοποίηση της διακύμανσης μεταξύ των διαφορών των μετρήσεων των μεθόδων για κάθε ξεχωριστό υποκείμενο. Παρέχουμε μια μέθοδο εύχρηστη για τον κλινικό ερευνητή και η οποία απαιτεί μόνο τη χρήση βασικών στατιστικών πακέτων. Οι εκτιμήσεις που προκύπτουν μπορούν να ερμηνευτούν με ευκολία και δίνονται πάντα στις ίδιες μονάδες μέτρησης με τις αρχικές μετρήσεις. Επιπλέον εστιάζουμε στην ερμηνεία μεμονωμένων μετρήσεων για κάθε υποκείμενο ξεχωριστά.

Επεκτείνουμε την προσέγγισή μας και στην περίπτωση όπου οι διαφορές μεταξύ των μετρήσεων των μεθόδων εξαρτώνται από το μέγεθος της μέτρησης, και παρουσιάζουμε τον τρόπο να συγκρίνουμε μεθόδους όταν έχουμε διαθέσιμες επαναλαμβανόμενες μετρήσεις. Τέλος, περιγράφουμε μια μη παραμετρική προσέγγιση για να χρησιμοποιείται όταν ενδεχομένως υπάρχουν έντονες αποκλίσεις μεταξύ των προς σύγκριση μεθόδων.

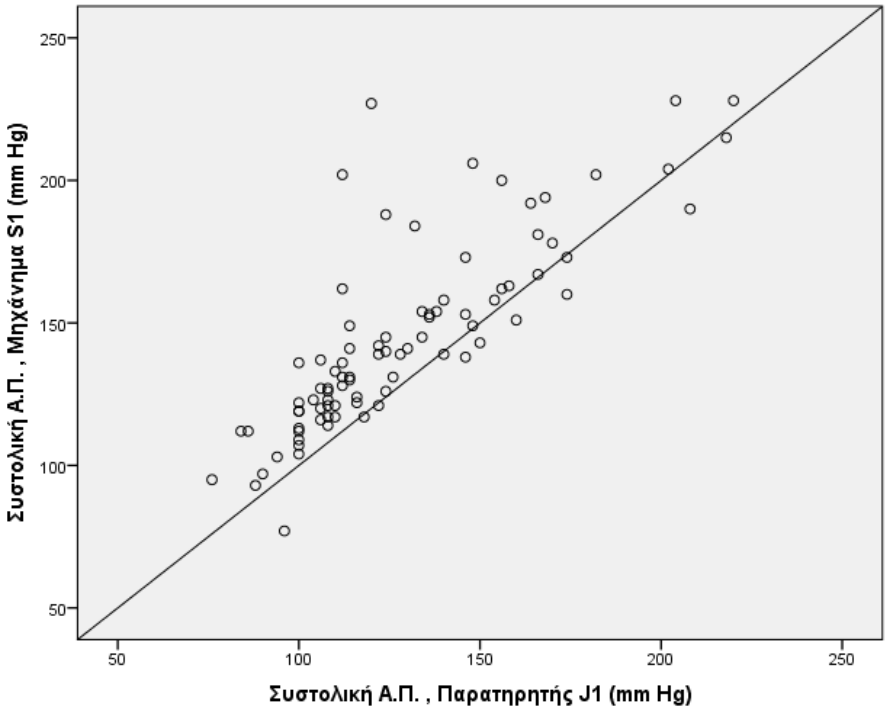
ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

Κεφάλαιο 2: Τα όρια συμφωνίας

Στόχος μας είναι να βρούμε έναν τρόπο να υπολογίζουμε το κατά πόσο συμφωνούν ή αντίστοιχα αποκλίνουν δύο μέθοδοι. Χρειαζόμαστε λοιπόν ένα μέγεθος το οποίο να εκφράζει αυτή τη συμφωνία, να είναι όμως εύκολα υπολογίσιμο καθώς και να μπορεί να ερμηνευτεί χωρίς δυσκολία. Μια πρώτη σκέψη είναι να χρησιμοποιήσουμε τις διαφορές μεταξύ των μετρήσεων διαφορετικών μεθόδων στα ίδια υποκείμενα. Αν από τις διαφορές αυτές διαπιστωθεί το φαινόμενο, μία από τις δύο μεθόδους να έχει την τάση να δίνει σταθερά υψηλότερα ή σταθερά χαμηλότερα αποτελέσματα, τότε το ονομάζουμε *κλίση (bias)* και μπορούμε να το εκτιμήσουμε από τη μέση διαφορά των μετρήσεων[1]. Είναι αναμενόμενο να υπάρχει διακύμανση της μέσης αυτής διαφοράς η οποία εκτιμάται μέσω της τυπικής απόκλισης των διαφορών. Για να έχουν όμως νόημα οι εκτιμήσεις αυτές, θα πρέπει να ελεγχθεί αρχικά (μέσω γραφημάτων) η ομοιομορφία στο μέγεθος του δείγματος τόσο της κλίσης όσο και της διακύμανσης, έλεγχος που παρουσιάζεται αναλυτικά στην Παράγραφο 2.1.

Στον Πίνακα 2.1 που ακολουθεί παρουσιάζονται δεδομένα που αφορούν στις μετρήσεις της αρτηριακής συστολικής πίεσης 85 ασθενών. Δύο έμπειροι παρατηρητές (J και R) χρησιμοποιώντας σφυγμομανόμετρο και ένα ημιαυτόματο μηχάνημα καταγραφής αρτηριακής συστολικής πίεσης (S) πραγματοποιούν 3 ταυτόχρονες μετρήσεις,οι οποίες και καταγράφονται. Αρχικά θα παρουσιάσουμε την ανάλυση για μη επαναλαμβανόμενες μετρήσεις, οπότε θα χρησιμοποιήσουμε τα δεδομένα της πρώτης μέτρησης του παρατηρητή J (J_1) και της πρώτης μέτρησης του μηχανήματος S (S_1). (Γράφημα 1).

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ



Γράφημα 1. Συστολική αρτηριακή πίεση, καταμετρημένη από τον Παρατηρητή J χρησιμοποιώντας συγμομανόμετρο και από το μηχάνημα S.

Subject	J1	J2	J3	R1	R2	R3	S1	S2	S3
1	100	106	107	98	98	111	122	128	124
2	108	110	108	108	112	110	121	127	128
3	76	84	82	76	88	82	95	94	98
4	108	104	104	110	100	106	127	127	135
5	124	112	112	128	112	114	140	131	124
6	122	140	124	124	140	126	139	142	136
7	116	108	102	118	110	102	122	112	112
8	114	110	112	112	108	112	130	129	135
9	100	108	112	100	106	112	119	122	122
10	108	92	100	108	98	100	126	113	111
11	100	106	104	102	108	106	107	113	111
12	108	112	122	108	116	120	123	125	125
13	112	112	110	114	112	110	131	129	122
14	104	108	104	104	108	104	123	126	114
15	106	108	102	104	106	102	127	119	126
16	122	122	114	118	122	114	142	133	137
17	100	102	102	102	102	100	104	116	115
18	118	118	120	116	118	118	117	113	112
19	140	134	138	138	136	134	139	127	113
20	150	148	144	148	146	144	143	155	133
21	166	154	154	164	154	148	181	170	166
22	148	156	134	136	154	132	149	156	140
23	174	172	166	170	170	164	173	170	154
24	174	166	150	174	166	154	160	155	170

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

25	140	144	144	140	144	144	158	152	154
26	128	134	130	128	134	130	139	144	141
27	146	138	140	146	138	138	153	150	154
28	146	152	148	146	152	148	138	144	131
29	220	218	220	220	218	220	228	228	226
30	208	200	192	204	200	190	190	183	184
31	94	84	86	94	84	88	103	99	106
32	114	124	116	112	126	118	131	131	124
33	126	120	122	124	120	120	131	123	124
34	124	124	132	126	126	120	126	129	125
35	110	120	128	110	122	126	121	114	125
36	90	90	94	88	88	94	97	94	96
37	106	106	110	106	108	110	116	121	127
38	218	202	208	218	200	206	215	201	207
39	130	128	130	128	126	128	141	133	146
40	136	136	130	136	138	128	153	143	138
41	100	96	88	100	96	86	113	107	102
42	100	98	88	100	98	88	109	105	97
43	124	116	122	126	116	122	145	102	137
44	164	168	154	164	168	154	192	178	171
45	100	102	100	100	104	102	112	116	116
46	136	126	122	136	124	122	152	144	147
47	114	108	122	114	108	122	141	141	137
48	148	120	132	146	130	132	206	188	166
49	160	150	148	160	152	146	151	147	136
50	84	92	98	86	92	98	112	125	124
51	156	162	152	156	158	152	162	165	189
52	110	98	98	108	100	98	117	118	109
53	100	106	106	100	108	108	119	131	124
54	100	102	94	100	102	96	136	116	113
55	86	74	76	88	76	76	112	115	104
56	106	100	110	106	100	108	120	118	132
57	108	110	106	106	118	106	117	118	115
58	168	188	178	170	188	182	194	191	196
59	166	150	154	164	150	154	167	160	161
60	146	142	132	144	142	130	173	161	154
61	204	198	188	206	198	188	228	218	189
62	96	94	86	96	94	84	77	89	101
63	134	126	124	132	126	124	154	156	141
64	138	144	140	140	142	138	154	155	148
65	134	136	142	136	134	140	145	154	166
66	156	160	154	156	162	156	200	180	179
67	124	138	138	122	140	136	188	147	139
68	114	110	114	112	114	114	149	217	192
69	112	116	122	112	114	124	136	132	133
70	112	116	134	114	114	136	128	125	142

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

71	202	220	228	200	220	226	204	222	224
72	132	136	134	134	136	132	184	187	192
73	158	162	152	158	164	150	163	160	152
74	88	76	88	90	76	86	93	88	88
75	170	174	176	172	174	178	178	181	181
76	182	176	180	184	174	178	202	199	195
77	112	114	124	112	112	126	162	166	148
78	120	118	120	118	116	120	227	227	219
79	110	108	106	110	108	106	133	127	126
80	112	112	106	112	110	106	202	190	213
81	154	134	130	156	136	132	158	121	134
82	116	112	94	118	114	96	124	149	137
83	108	110	114	106	110	114	114	118	126
84	106	98	100	104	100	100	137	135	134
85	122	112	112	122	114	114	121	123	128

Πίνακας 2.1. Μετρήσεις συστολικής αρτηριακής πίεσης οι οποίες πραγματοποιήθηκαν ταυτόχρονα από δύο παρατηρητές (J και R) και από ένα ημιαυτόματο μηχάνημα καταγραφής αρτηριακής συστολικής πίεσης (S), πραγματοποιώντας ο καθένας 3 διαδοχικές μετρήσεις (τα δεδομένα παρέχονται από τον Dr E O' Brien, του Altman και Bland[21]).

Από τα στοιχεία του Πίνακα 2.1, υπολογίζουμε τη μέση διαφορά μεταξύ των μετρήσεων J1 και S1 και όπως φαίνεται στον Πίνακα 2.2 που ακολουθεί, είναι ίση με $\bar{d} = -16.29 mmHg$ και η τυπική απόκλιση των διαφορών ισούται με $s_d = 19.61 mmHg$.

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
d	85	-107,00	19,00	**-16,2941**	**19,61099**
Valid N (listwise)	85				

Πίνακας 2.2. Μέση τιμή και τυπική απόκλιση των διαφορών των δύο μεθόδων J1 και S1.

Θεωρώντας ότι οι διαφορές ακολουθούν κανονική κατανομή, τότε αναμένουμε το 95% αυτών να βρίσκεται εντός του διαστήματος $(\bar{d} \pm 1.96 s_d)$, (θα μπορούσαμε να χρησιμοποιήσουμε και το διάστημα $(\bar{d} \pm 2 s_d)$ χωρίς σημαντικές απώλειες στην ακρίβεια του αποτελέσματος). Στην περίπτωσή μας το διάστημα αυτό υπολογίζεται ως εξής:

$$(\bar{d} \pm 1.96 s_d) = (-16.29 \pm 1.96 \cdot 19.61) = (-16.29 \pm 38.44) = (-54.73, 22.15)$$

Αναμένουμε λοιπόν το 95% των διαφορών κάθε ζεύγους μετρήσεων(J_{1i},S_{1i}) με i=1…85, να βρίσκεται εντός του συγκεκριμένου διαστήματος,το οποίο και

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ονομάζουμε 95% όριο συμφωνίας. Και στην περίπτωση που θεωρήσουμε ότι οι διαφορές εντός των συγκεκριμένων ορίων συμφωνίας δε κρίνονται κλινικά σημαντικές, τότε έχουμε τη δυνατότητα να χρησιμοποιούμε τις δύο μεθόδους εναλλάξ χωρίς να επηρεάζει η μέθοδος που ακολουθείται το αποτέλεσμα της μέτρησης.

Αξίζει να σημειωθεί ότι παρά την ομοιότητα τους, το όριο συμφωνίας δεν πρέπει να συγχέεται με το διάστημα εμπιστοσύνης, γεγονός που θα ήταν πιο ξεκάθαρο αν χρησιμοποιούσαμε διαφορετικό ποσοστό από αυτό του 95%, προτιμήθηκε όμως η χρήση του συγκεκριμένου ποσοστού καθώς αποτελεί μια από πιο συνήθης επιλογές ποσοστών.

Όπως αναφέρθηκε,ο υπολογισμός του 95% ορίου συμφωνίας βασίστηκε στην υπόθεση ότι οι διαφορές των μετρήσεων των δύο μεθόδων, ακολουθούν κανονική κατανομή, γεγονός που μπορεί να θεωρηθεί αρκετά πιθανό, καθώς έχει αφαιρεθεί σημαντική διακύμανση μεταξύ των ασθενών και το σφάλμα των μετρήσεων που έχει απομείνει είναι πιθανό να ακολουθεί κανονική κατανομή. Συνεπώς ενώνοντας τα δύο είδη σφάλματος, η τάση για κανονικότητα αυξάνεται. Για τον έλεγχο της κανονικότητας της κατανομής των διαφορών μπορούν να χρησιμοποιηθούν γραφικές μέθοδοι όπως ιστόγραμμα ή διάγραμμα κανονικότητας. Στην περίπτωση που η κατανομή δεν είναι συμμετρική ή παρουσιάζει ιδιαίτερα «μακριά» άκρα, τότε ενδεχομένως να μην ισχύει η υπόθεση της κανονικότητας. Κάτι τέτοιο είναι πιθανό να συμβαίνει όταν υπάρχει σχέση μεταξύ των διαφορών και της μέσης τιμής τους, σε μια τέτοια περίπτωση η αντιμετώπιση που προτείνεται αναλύεται στο Κεφάλαιο 3. Ακόμη όμως και αν δεν ισχύει η υπόθεση κανονικότητας,είναι πιθανό να μην αποτελεί τόσο σοβαρό πρόβλημα στη συγκεκριμένη περίπτωση όσο σε κάποια άλλα στατιστικά πλαίσια, εξάλλου ακόμη και σε μια μη κανονική κατανομή υπάρχει πιθανότητα το 5% των παρατηρήσεων να βρίσκεται μεταξύ δύο τυπικών αποκλίσεων από το μέσο αν και οι περισσότερες τιμές εκτός ορίων, πιθανότατα είναι διαφορές προς την ίδια κατεύθυνση. Το σημαντικό άλλωστε δεν είναι να ακολουθούν οι μετρήσεις κανονική κατανομή, αλλά οι διαφορές τους, κάτι που παρατηρείται συχνά στην πράξη καθώς τα υποκείμενα επιλέγονται με τέτοιο τρόπο ώστε να

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

παρέχουν μια ευρεία και ομοιόμορφη κατανομή των ποσοτήτων που μας ενδιαφέρουν, παρά να αποτελούν ένα τυχαίο δείγμα.

Στην περίπτωση που παρατηρηθεί σταθερή κλίση θα πρέπει απλώς,αν κριθεί απαραίτητο, να γίνει μια προσαρμογή αφαιρώντας τη μέση διαφορά από τις μετρήσεις της νέας μεθόδου[1]. Γενικώς, μεγαλύτερο και σοβαρότερο πρόβλημα αποτελεί μια μεγάλη τυπική απόκλιση διαφορών και κατά συνέπεια ένα μεγάλο διάστημα για το όριο συμφωνίας.

Στο παράδειγμα της μετρήσεως της αρτηριακής συστολικής πίεσης ασθενών, παρατηρούνται κάποιες ιδιαίτερα μεγάλες διαφορές, στις οποίες το μηχάνημα καταγραφής δίνει τιμές μεγαλύτερες από το σφυγμομανόμετρο. Διαπιστώνουμε ότι υπάρχουν 4 μετρήσεις στο σύνολο των 85, δηλαδή το 5% του συνόλου, στις οποίες οι διαφορές των μετρήσεων βρίσκονται εκτός του ορίου συμφωνίας, και στις 4 όμως το μηχάνημα έχει καταγράψει μεγαλύτερες τιμές. Πιο συγκεκριμένα, στα υποκείμενα 48, 67, 78 και 80 καταγράφονται διαφορές -58, -64, -107 και -90 αντίστοιχα. Μπορούμε να εκτιμήσουμε την επίδραση των δύο μεγαλύτερων διαφορών (-107 και -90 των αντικειμένων 78 και 80 αντίστοιχα), θεωρώντας ότι πρόκειται προφανώς για ακραίες τιμές, αφαιρώντας τες από το δείγμα των μετρήσεων και υπολογίζοντας ξανά το 95% όριο συμφωνίας. Όπως φαίνεται στον Πίνακα 2.3 η νέα μέση διαφορά είναι -14.31mmHg και η τυπική απόκλιση 14,95mmHg και συνεπώς το νέο όριο συμφωνίας που προκύπτει είναι (-43.61, 15.00).

$$\left(\bar{d} \pm 1.96 s_d\right) = (-14.31 \pm 1.96 \cdot 14.95) = (14.31 \pm 29.30) = (-43.61, 15.00)$$

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
d	83	-64,00	19,00	**-14,3133**	**14,94700**
Valid N (listwise)	83				

Πίνακας 2.3. Μέση τιμή και τυπική απόκλιση των διαφορών των δύο μεθόδων J_1 και S_1,έχοντας εξαιρέσει τα αντικείμενα 78 και 80.

Συμπεραίνουμε ότι ενώ το εύρος του διαστήματος μειώθηκε από 76.88 mmHg σε 58.60mmHg η μείωση αυτή δεν είναι ιδιαίτερα μεγάλη, άρα η μέθοδος της ανάλυσης που ακολουθούμε δεν παρουσιάζει ευαισθησία σε μια ή δυο

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ακραίες τιμές, συνεπώς δεν προτείνεται η εξαίρεσή τους από το δείγμα, παρά μόνο η εκτίμηση της επίδρασής τους.

Παρατηρώντας τα στοιχεία του Πίνακα 2.1 διαπιστώνουμε ότι μεγάλες διαφορές όπως εκείνες που αναφέρθηκαν παραπάνω δεν παρατηρούνται μόνο στην πρώτη μέτρηση των δύο μεθόδων αλλά και στις τρεις μετρήσεις κάθε μεθόδου, φαινόμενο καθόλου σπάνιο στην περίπτωση μετρήσεως της αρτηριακής συστολικής πίεσης με τη χρήση μηχανήματος. Για αυτό ακριβώς το λόγο αναπτύχθηκε και παρουσιάζεται σε επόμενο κεφάλαιο, μια μη παραμετρική προσέγγιση αυτών των δεδομένων.

Όπως αναφέρθηκε και στο Κεφάλαιο 1, η απόφαση για την αποδοχή ή μη της συμφωνίας μεταξύ των διαφόρων μεθόδων, λαμβάνεται με βάση κλινικά χαρακτηριστικά και είναι κάτι που ξεπερνά τα όρια του στατιστικού αναλυτή. Η μόνη διαπίστωση που μπορεί να γίνει με βάση την ανωτέρω ανάλυση, είναι ότι το ημιαυτόματο μηχάνημα καταγραφής αρτηριακής συστολικής πίεσης μπορεί να δώσει μετρήσεις μεγαλύτερες κατά 55mmHg και μικρότερες κατά 22mmHg συγκριτικά με το σφυγμομανόμετρο, το αν οι συγκεκριμένες διαφοροποιήσεις στις μετρήσεις θεωρούνται αποδεκτές ή όχι αποτελεί ερώτημα κλινικής κρίσης.

2.1. Γραφική Απεικόνιση της συμφωνίας

Στη μέθοδο συγκρίσεως μελετών, οι τεχνικές γραφικών απεικονίσεων είναι ιδιαίτερα χρήσιμες. Στο Γράφημα 1 απεικονίστηκαν οι μετρήσεις του Παρατηρητή J_1 ο οποίος χρησιμοποίησε σφυγμομανόμετρο σε σχέση με τις μετρήσεις S_1 του μηχανήματος. Στο γράφημα σχεδιάστηκε η γραμμή ισότητας, η γραμμή δηλαδή, που θα προέκυπτε αν οι δύο παρατηρητές έβρισκαν ακριβώς το ίδιο αποτέλεσμα σε κάθε μέτρηση. Δεν υπολογίστηκε, ούτε και σχεδιάστηκε η γραμμή παλινδρόμησης καθώς δεν ενδιαφερόμαστε για την εκτιμώμενη πρόβλεψη της μιας μεθόδου σε σχέση με την άλλη, αλλά με την θεωρητική σχέση ισότητας και των αποκλίσεων από αυτή. Είναι ιδιαίτερα χρήσιμο, οι δύο άξονες του γραφήματος να είναι στην ίδια κλίμακα

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ώστε η γραμμή της ισότητας να σχηματίζει γωνία 45^{o} και με τους δύο άξονες. Αποκτούμε έτσι με μεγαλύτερη ευκολία οπτική εικόνα της συμφωνίας ή όχι των δύο μεθόδων. Ωστόσο, στην περίπτωση που το εύρος της διακύμανσης των μετρήσεων είναι μεγάλο συγκρινόμενο με τις διαφορές μεταξύ των μεθόδων, το συγκεκριμένο είδος γραφήματος ενδέχεται να αποκρύπτει σημαντικές πληροφορίες.

Ένας καλύτερος τρόπος παρουσίασης των δεδομένων είναι παραστήσουμε γραφικά τις διαφορές μεταξύ των μετρήσεων των δύο μεθόδων για κάθε υποκείμενο, σε σχέση με τη μέση τιμή τους. Η ανάλογη γραφική απεικόνιση για τα δεδομένα της συστολικής αρτηριακής πίεσης παρουσιάζεται στο Γράφημα 2, από το οποίο συμπεραίνουμε ξεκάθαρα την έλλειψη συμφωνίας μεταξύ των δύο μεθόδων, γεγονός που δεν ήταν ιδιαίτερα φανερό από το Γράφημα 1. Επιπλέον, ένα τέτοιο γράφημα μας δίνει τη δυνατότητα να ερευνήσουμε κάθε πιθανή σχέση μεταξύ των αποκλίσεων και των πραγματικών τιμών. Μπορούμε να εξετάσουμε την πιθανή αυτή σχέση με τη βοήθεια του συντελεστή συσχέτισης Spearman, ο οποίος στη συγκεκριμένη περίπτωση έχει τιμή r_s=0.07.(Πίνακας 2.4)

Correlations

			meanJ1_S1	abs_d
Spearman's rho	Μέση τιμή μετρήσεων	Correlation Coefficient	1,000	**,068**
		Sig. (2-tailed)	.	,539
		N	85	85
	Απόλυτες τιμές διαφορών	Correlation Coefficient	,068	1,000
		Sig. (2-tailed)	,539	.
		N	85	85

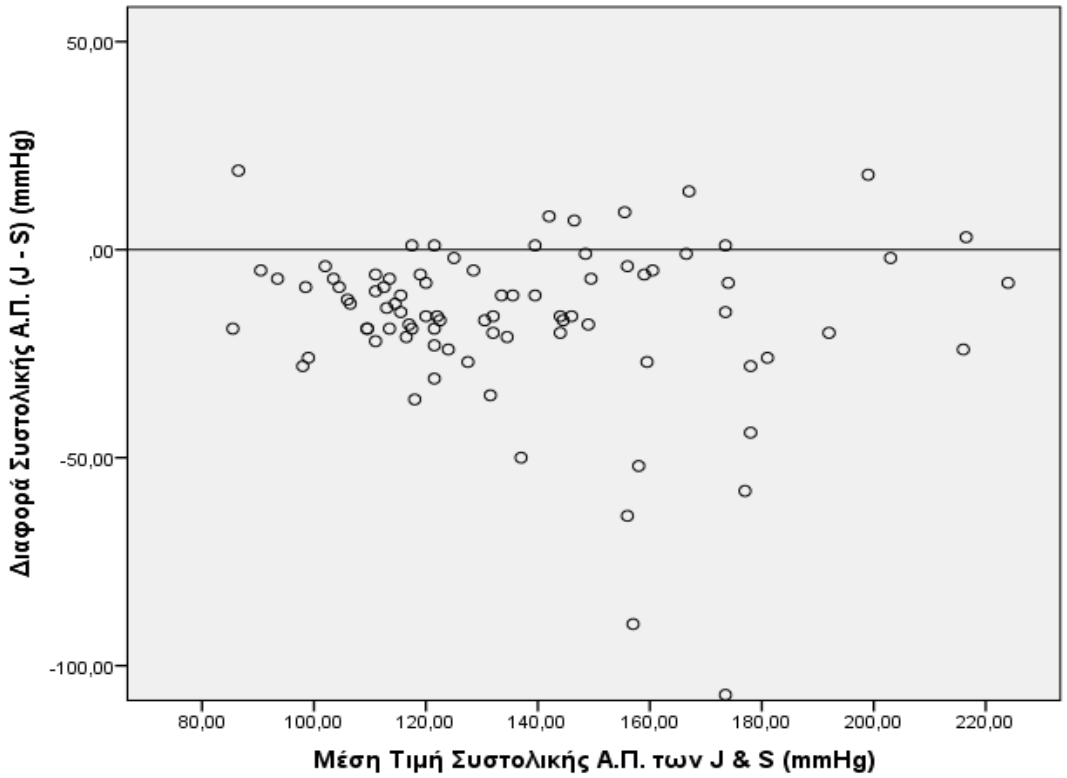
Πίνακας 2.4. Spearman's rank correlations, μεταξύ των απόλυτων τιμών των διαφορών και της μέσης τιμής των μετρήσεων των δύο μεθόδων.

Επιπλέον, από το Γράφημα 2 μπορούμε να παρατηρήσουμε πιθανών ακραίες τιμές ή τιμές εκτός των ορίων συμφωνίας. Και πάλι είναι χρήσιμο να χρησιμοποιούμε την ίδια κλίμακα και στους δύο άξονες, ώστε να είμαστε σε θέση να διαπιστώσουμε αν υπάρχει σχέση μεταξύ των αποκλίσεων και του μεγέθους των μετρήσεων. Στην περίπτωση που οι δύο μέθοδοι συμφωνούν αρκετά είναι αδύνατον πρακτικά να υπάρχει ίδια κλίση, συνεπώς υπάρχει πιθανότητα να χάνεται και πάλι πολύτιμη πληροφορία, και για να το

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

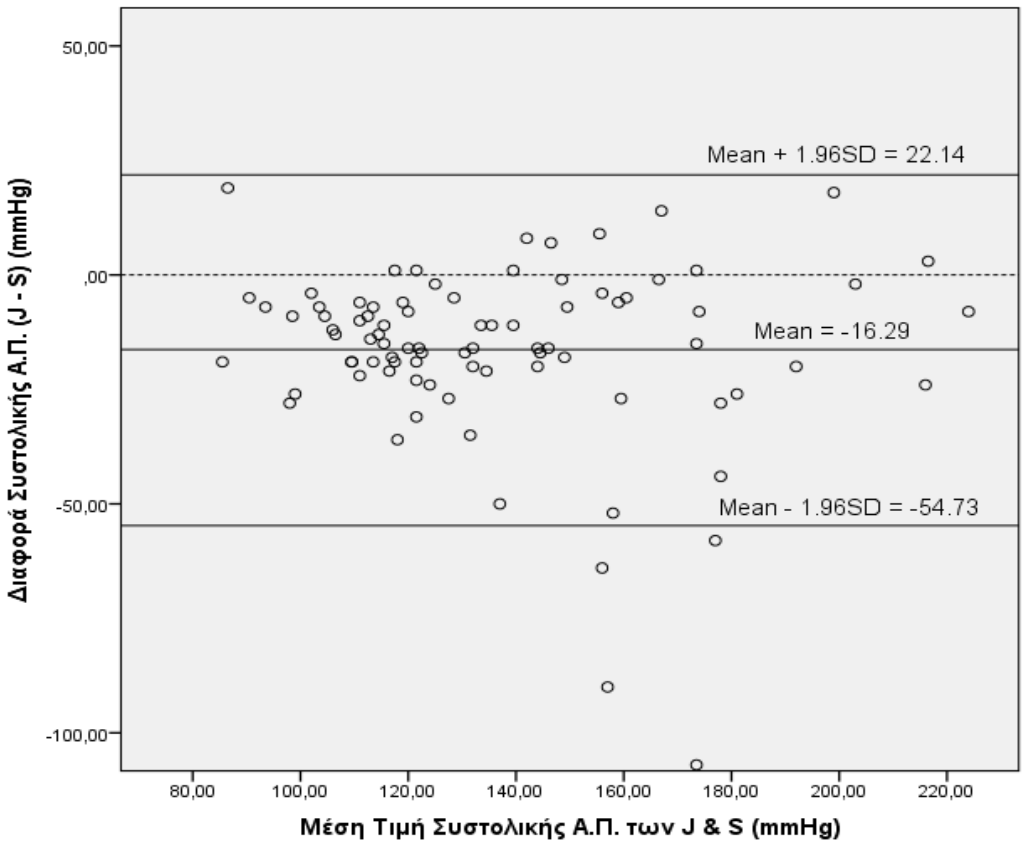
αποφύγουμε προσθέτουμε το 95% όριο συμφωνίας (Γράφημα 3) ώστε να έχουμε καλύτερη απεικόνιση της ύπαρξης συμφωνίας μεταξύ των δύο μεθόδων ή όχι.

Σε τέτοιου είδους μεθόδους, επειδή δεν γνωρίζουμε τις πραγματικές τιμές των ποσοτήτων τις οποίες μετράμε, χρησιμοποιούμε τη μέση τιμή αυτών και για τις δύο μεθόδους σαν την καλύτερη εκτίμηση. Το να παραστήσουμε γραφικά τις διαφορές σε σχέση με τις τιμές θα οδηγούσε σε λάθος συμπεράσματα καθώς υπάρχει συσχέτιση μεταξύ τους[4].



Γράφημα 2. Διαφορά των μετρήσεων της συστολικής αρτηριακής πίεσης, του Παρατηρητή J και του μηχανήματος S, σε σχέση με τη μέση τιμή τους.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ



Γράφημα 3. Διαφορά των μετρήσεων της συστολικής αρτηριακής πίεσης, του Παρατηρητή J και του μηχανήματος S σε σχέση με τη μέση τιμή του, με το 95% όριο συμφωνίας.

2.2. Ακρίβεια των εκτιμώμενων ορίων συμφωνίας

Αν υποθέσουμε ότι οι διαφορές ακολουθούν προσεγγιστικά κανονική κατανομή, τότε μπορούμε να υπολογίσουμε τα τυπικά σφάλματα και τα διαστήματα εμπιστοσύνης των ορίων συμφωνίας. Η διακύμανση της αναμενόμενης μέσης τιμής των διαφορών $\bar{d}$, εκτιμάται από τη σχέση s_d^2/n, όπου n το μέγεθος του δείγματος. Δεδομένου ότι οι διαφορές ακολουθούν κανονική κατανομή, η διακύμανση της τυπικής απόκλισης των διαφορών εκτιμάται προσεγγιστικά από τη σχέση $s_d^2/2(n-1)$, και αυτό γιατί s_d^2 ακολουθεί κατανομή $\chi^2 \times \sigma_d^2/(n-1)$ και η διακύμανση του $\sqrt{\chi^2}$ είναι προσεγγιστικά ½. Η μέση τιμή των διαφορών και η τυπική απόκλισή τους είναι ανεξάρτητες, συνεπώς η διακύμανση των ορίων συμφωνίας εκτιμάται από :

$$Var\left(\bar{d} \pm 1.96 s_d\right) = Var\left(\bar{d}\right) + 1.96^2 Var(s_d) = \frac{s_d^2}{n} + 1.96^2 \frac{s_d^2}{2(n-1)}$$

$$= \left(\frac{1}{n} + \frac{1.96^2}{2(n-1)}\right) s_d^2$$

Και εάν το n δεν είναι μικρό, τότε προσεγγιστικά η προηγούμενη σχέση γίνεται

$$Var\left(\bar{d} \pm 1.96 s_d\right) = \left(1 + \frac{1.96^2}{2}\right) \frac{s_d^2}{n} = 2.92 \frac{s_d^2}{n} = 1.71^2 \frac{s_d^2}{n}$$

Επομένως τα τυπικά σφάλματα του $\bar{d} - 1.96 s_d$ και $\bar{d} + 1.96 s_d$ ισούνται προσεγγιστικά με $1.71 s_d/\sqrt{n} = 1.71 SE\left(\bar{d}\right)$. Τα 95% διαστήματα εμπιστοσύνης μπορούμε να τα υπολογίσουμε βρίσκοντας το κατάλληλο ποσοστιαίο σημείο της t κατανομής, με n-1 βαθμούς ελευθερίας, οπότε το διάστημα εμπιστοσύνης θα προκύπτει από τα t τυπικά σφάλματα κάθε πλευράς της παρατηρούμενης τιμής.

Για τα συγκεκριμένα δεδομένα της συστολικής αρτηριακής πίεσης η τυπική απόκλιση είναι $s_d = 19.61\, mmHg$, και συνεπώς το τυπικό σφάλμα της κλίσης $\bar{d}$ είναι $s_d/\sqrt{n} = 19.61/\sqrt{85} = 2.13 mmHg$. Και για το αντίστοιχο 95% διάστημα εμπιστοσύνης με 84 βαθμούς ελευθερίας και t=1.99 έχουμε : $(-16.29 - 1.99 \times 2.13, -16.29 + 1.99 \times 2.13) = (-20.53, -12.05)$mmHg. Το τυπικό σφάλμα του 95% ορίου συμφωνίας όπως δείξαμε προηγουμένως ισούται με $1.71 SE\left(\bar{d}\right) = 3.64 mmHg$, επομένως το 95% διάστημα εμπιστοσύνης για το κάτω άκρο του ορίου συμφωνίας είναι $(-54.7 - 1.99 \times 3.64, -54.7 + 1.99 \times 3.64) = (-61.9, -47.5) mmHg$ και για το άνω άκρο αντίστοιχα $(22.1 - 1.99 \times 3.64, 22.1 + 1.99 \times 3.64) = (14.9, 29.3) mmHg$. Λόγω του μεγάλου δείγματος τα συγκεκριμένα διαστήματα είναι λογικό να είναι στενά. Και πάλι όμως μας δείχνουν ότι ακόμα και αν επιλέγαμε τις ευνοϊκότερες τιμές που θα μπορούσαν να έχουν τα άκρα του ορίου συμφωνίας θα υπήρχαν αξιοσημείωτες διαφορές μεταξύ των δύο μεθόδων μέτρησης,και συνεπώς θα μπορούσαμε να συμπεράνουμε ότι δεν μπορεί να είναι αποδεκτός ο βαθμός της συμφωνίας των δύο αυτών μεθόδων.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

Τα εν λόγω διαστήματα εμπιστοσύνης βασίζονται στην υπόθεση ότι η μόνη αβεβαιότητα προκύπτει από τα δειγματοληπτικά σφάλματα, επίσης θεωρούμε ότι το δείγμα των υποκειμένων είναι αντιπροσωπευτικό. Επιπλέον, όλες οι μετρήσεις με το σφυγμομανόμετρο πραγματοποιήθηκαν από τον ίδιο έμπειρο παρατηρητή. Συνεπώς, η υπολογισμένη αβεβαιότητα η οποία σχετίζεται με τα όρια συμφωνίας, είναι πιθανόν να είναι κάπως αισιόδοξη[1].

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

Κεφάλαιο 3. Σχέση διαφοράς και μεγέθους

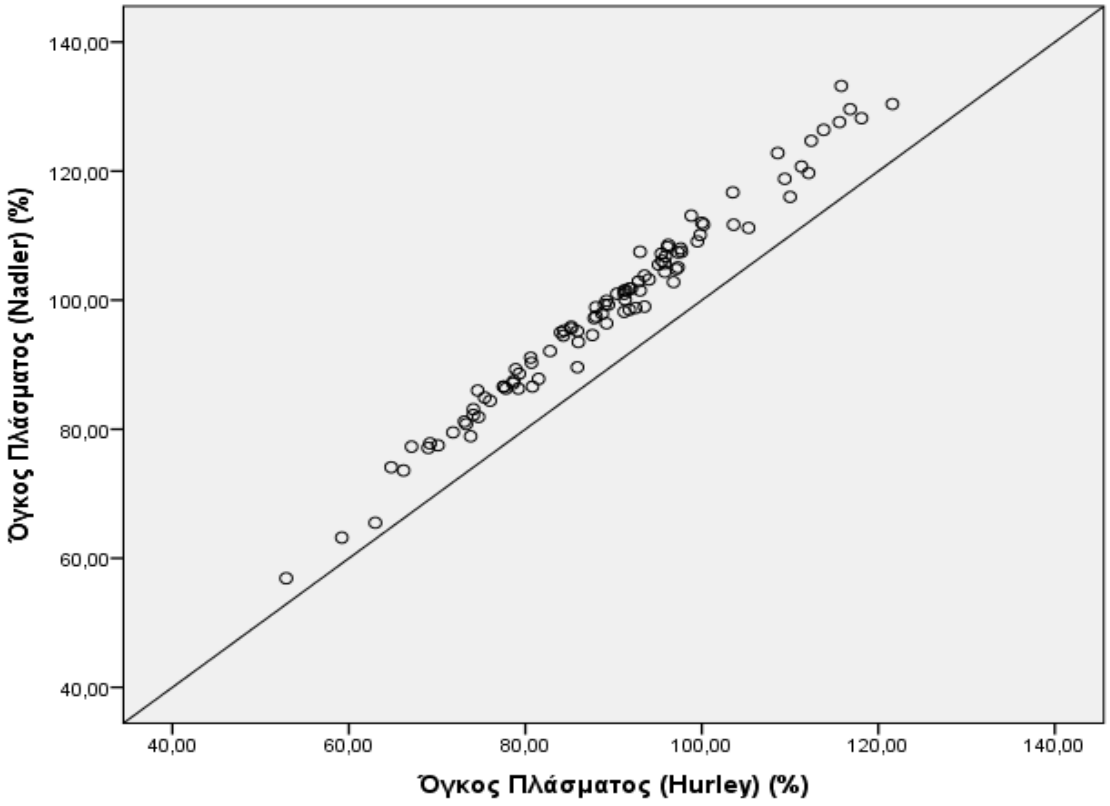
Στο Κεφάλαιο 2 υποθέσαμε ότι η μέση τιμή και η τυπική απόκλιση των διαφορών παραμένουν σταθερές σε όλο το εύρος των μετρήσεων. Η πιο συνηθισμένη απόκλιση από τις υποθέσεις αυτές είναι η αύξηση της μεταβλητότητας των διαφορών καθώς το μέγεθος των μετρήσεων αυξάνεται. Σε αυτές τις περιπτώσεις η γραφική απεικόνιση της μιας μεθόδου σε σχέση με την άλλη παρουσιάζει μια διασπορά των δεδομένων στις μεγαλύτερες μετρήσεις. Και η αναμενόμενη μέση τιμή της διαφοράς $(\bar{d})$ ενδεχομένως να είναι προσεγγιστικά ανάλογη του μεγέθους της μέτρησης, γεγονός που είναι ιδιαίτερα εμφανές στη γραφική απεικόνιση μεταξύ των διαφορών και της αναμενόμενης τιμής των μετρήσεων. Στον Πίνακα 3.1 που ακολουθεί, παρουσιάζονται μετρήσεις του όγκου πλάσματος εκφρασμένες σαν ποσοστό της αναμενόμενης τιμής για φυσιολογικά υποκείμενα, οι οποίες έγιναν με δύο εναλλακτικές μεθόδους Nadler και Hurley. Τα συγκεκριμένα δεδομένα απεικονίζονται στο Γράφημα 4, από όπου μπορούμε να διαπιστώσουμε άμεσα, ότι οι δύο μέθοδοι δίνουν συστηματικά διαφορετικές μετρήσεις και ότι όλες βρίσκονται πάνω από τη γραμμή ισότητας. Το γεγονός ότι οι διαφορές μεγαλώνουν για μεγάλες τιμές του όγκου πλάσματος είναι περισσότερο εμφανές στο Γράφημα 5, όπου παρουσιάζεται η διαφορά σε σχέση με τη αναμενόμενη τιμή των υποκειμένων.

Sub	Nadler	Hurley	Sub	Nadler	Hurley	Sub	Nadler	Hurley
1	56,9	52,9	34	93,5	86	67	104,8	97,1
2	63,2	59,2	35	94,5	84,3	68	105,1	97,3
3	65,5	63	36	94,6	87,6	69	105,5	95,1
4	73,6	66,2	37	95	84	70	105,7	95,8
5	74,1	64,8	38	95,2	85,9	71	106,1	95,5
6	77,1	69	39	95,3	84,4	72	106,8	95,9
7	77,3	67,1	40	95,6	85,2	73	107,2	95,4
8	77,5	70,1	41	95,9	85,2	74	107,4	97,3
9	77,8	69,2	42	96,4	89,2	75	107,5	97,7
10	78,9	73,8	43	97,2	87,8	76	107,5	93
11	79,5	71,8	44	97,5	88	77	108	97,6
12	80,8	73,3	45	97,9	88,7	78	108,2	96,1
13	81,2	73,1	46	98,2	91,2	79	108,6	96,2
14	81,9	74,7	47	98,5	91,8	80	109,1	99,5

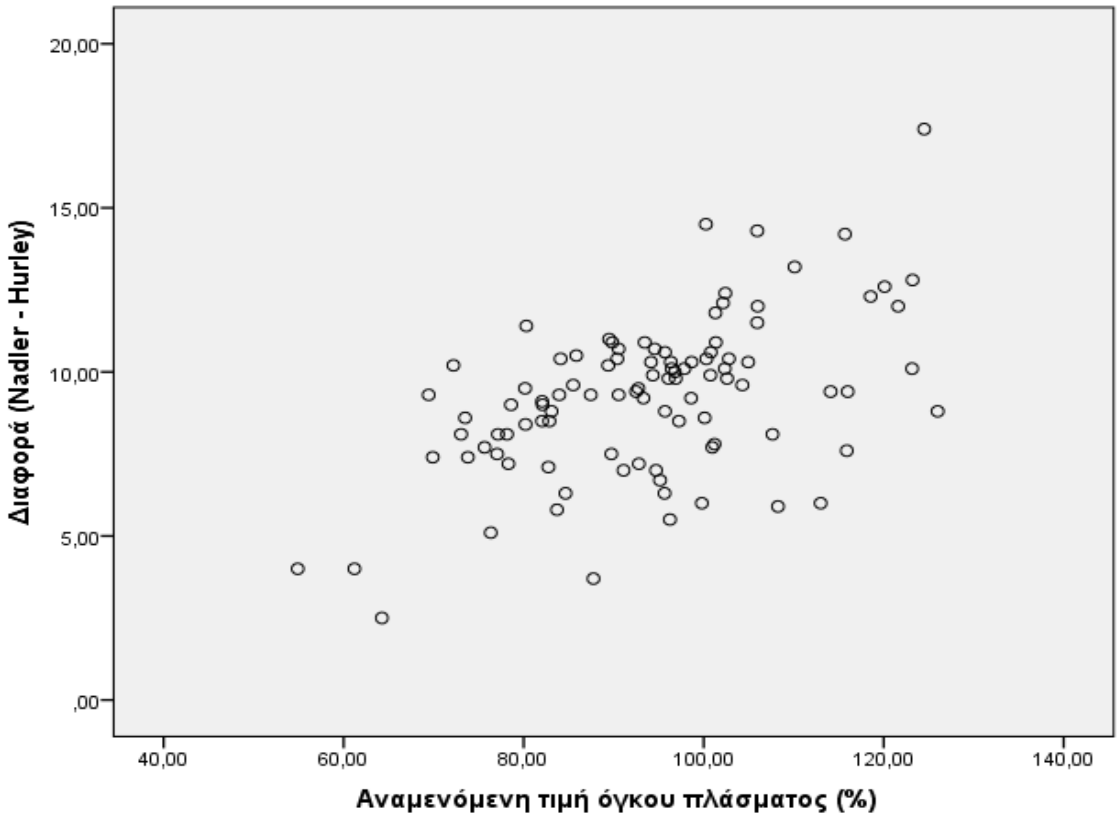
ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

15	82,2	74,1	48	98,8	92,5	81	110,1	99,8
16	83,1	74,1	49	98,9	88	82	111,2	105,3
17	84,4	76	50	99	93,5	83	111,7	103,6
18	84,9	75,4	51	99,3	89	84	111,7	100,2
19	86	74,6	52	99,3	89,4	85	112	100
20	86,3	79,2	53	99,9	89,2	86	113,1	98,8
21	86,3	77,8	54	100,1	91,3	87	116	110
22	86,6	80,8	55	101	90,4	88	116,7	103,5
23	86,6	77,6	56	101	91,2	89	118,8	109,4
24	86,6	77,5	57	101,5	91,4	90	119,7	112,1
25	87,1	78,6	58	101,5	93	91	120,7	111,3
26	87,5	78,7	59	101,5	91,2	92	122,8	108,6
27	87,8	81,5	60	101,8	92	93	124,7	112,4
28	88,6	79,3	61	101,8	91,8	94	126,4	113,8
29	89,3	78,9	62	102,8	96,8	95	127,6	115,6
30	89,6	85,9	63	102,9	92,8	96	128,2	118,1
31	90,3	80,7	64	103,2	94	97	129,6	116,8
32	91,1	80,6	65	103,8	93,5	98	130,4	121,6
33	92,1	82,8	66	104,4	95,8	99	133,2	115,8

Πίνακας 3.1. Μετρήσεις όγκου πλάσματος εκφρασμένες ως ποσοστό των κανονικών τιμών 99 υποκειμένων, χρησιμοποιώντας δύο διαφορετικές μεθόδους Nadler και Hurley. (τα δεδομένα προέρχονται από C Dore[22]).



Γράφημα 4. Μετρήσεις όγκου πλάσματος σύμφωνα με τα στοιχεία του Πίνακα 3.1



Γράφημα 5. Διαφορές και αναμενόμενες τιμές των μετρήσεων των δύο μεθόδων.

Αν αγνοούσαμε τη σχέση μεταξύ διαφορών και του μεγέθους των μετρήσεων και προχωρούσαμε στην ανάλυσή μας όπως στο Κεφάλαιο 2, θα καταλήγαμε σε όρια συμφωνίας που θα περιελάμβαναν τις περισσότερες διαφορές, αλλά για μικρές τιμές του όγκου πλάσματος θα είχαμε μεγαλύτερο διάστημα από αυτό που πραγματικά θα χρειαζόμασταν και αντίστοιχα στενότερο από το απαιτούμενο για μεγάλες τιμές. Για το λόγο αυτό είναι προτιμότερο να προσπαθήσουμε να αφαιρέσουμε τη σχέση αυτή είτε μετασχηματίζοντας τις μετρήσεις είτε με ακολουθώντας μια πιο γενική μέθοδο εάν ο μετασχηματισμός δεν είχε επιτυχία[1].

3.1. Λογαριθμικός μετασχηματισμός

Υπό αυτές τις συνθήκες ο λογαριθμικός (log) μετασχηματισμός των μετρήσεων και των δύο μεθόδων πριν την έναρξη της ανάλυσης, θα επιτρέψει τη χρήση της τυπικής προσέγγισης. Τα όρια συμφωνίας που θα προκύψουν από τα λογαριθμοποιημένα δεδομένα θα πρέπει να τα

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

απολογαριθμοποιήσουμε ώστε να καταλήξουμε στα όρια που θα αφορούν το λόγο των πραγματικών μετρήσεων[3]. Ενώ θα μπορούσαν να χρησιμοποιηθούν και άλλου είδους μετασχηματισμοί, (όπως τετραγωνικές ρίζες ή αντίστροφες τιμές), μόνο ο λογαριθμικός μετασχηματισμός μας επιτρέπει να ερμηνεύσουμε τα αποτελέσματα σε σχέση με τα αρχικά δεδομένα. Και δεδομένου ότι υπάρχει ξεκάθαρη ανάγκη, ειδικά σε τέτοιου είδους μελέτες, να είμαστε σε θέση να εκφράσουμε τα αποτελέσματα σε σχέση με τις πραγματικές μετρήσεις, δεν μπορούμε παρά να επιλέξουμε τον συγκεκριμένο μετασχηματισμό.

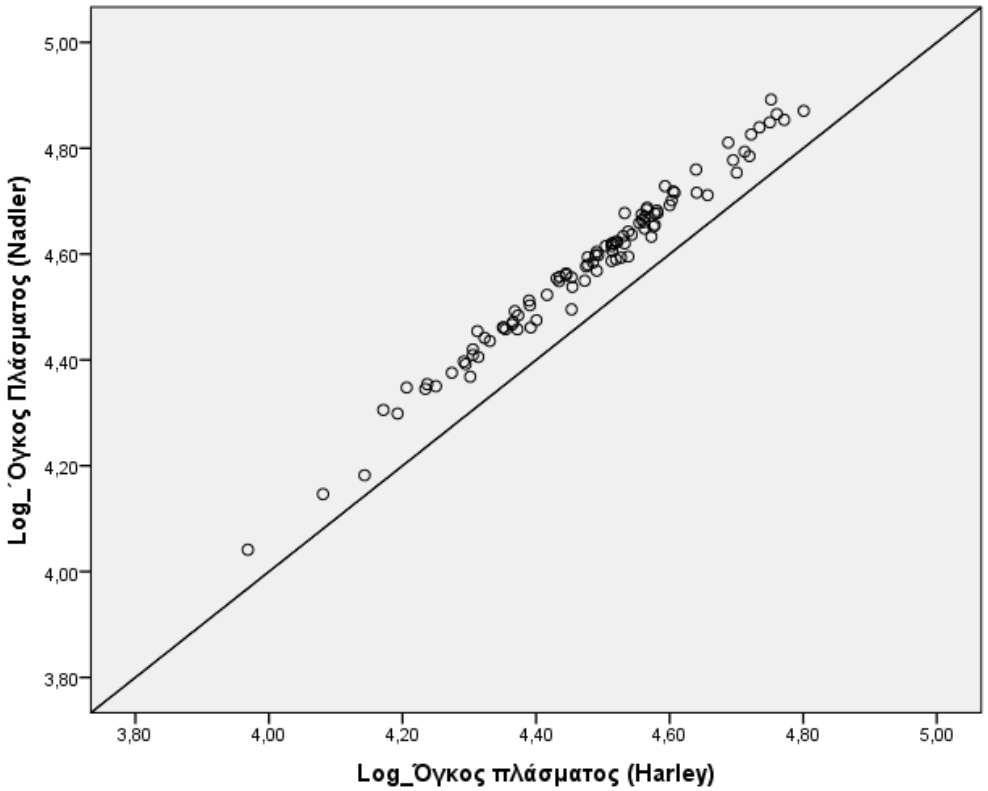
Για τα δεδομένα του Πίνακα 3.1. ο λογαριθμικός μετασχηματισμός έχει πραγματικά μεγάλη επιτυχία στην εξάλειψη της σχέσης μεταξύ των διαφορών και της αναμενόμενης τιμής. Στο Γράφημα 6(α) παρουσιάζονται τα λογαριθμοποιημένα δεδομένα του Πίνακα 3.1. και στο Γράφημα 6(β) οι αντίστοιχες διαφορές σε σχέση με την αναμενόμενη τιμή του όγκου πλάσματος έχοντας συμπεριληφθεί και το 95% όριο συμφωνίας. Όπως φαίνεται στον Πίνακα 3.2. η μέση τιμή των διαφορών (LogNadler – LogHurley) ισούται με 0.098 και η τυπική τους απόκλιση ισούται με 0.022, συνεπώς το 95% όριο συμφωνίας προκύπτει από τη σχέση : $(0.099 - 1.96 \cdot 0.022,\ 0.099 + 1.96 \cdot 0.022) = (0.056,\ 0.141)$. Λόγω της υψηλής συσχέτισης και του μεγάλου μεγέθους του δείγματος τα διαστήματα εμπιστοσύνης των ορίων συμφωνίας θα είναι στενά. Για παράδειγμα το 95% διάστημα εμπιστοσύνης με 98 βαθμούς ελευθερίας και t=1.984 για το κάτω άκρο είναι $\left(0.056 \pm 1.984 \times 1.71 \frac{0.022}{\sqrt{99}}\right) = (0.049,0.064)$.

Descriptive Statistics

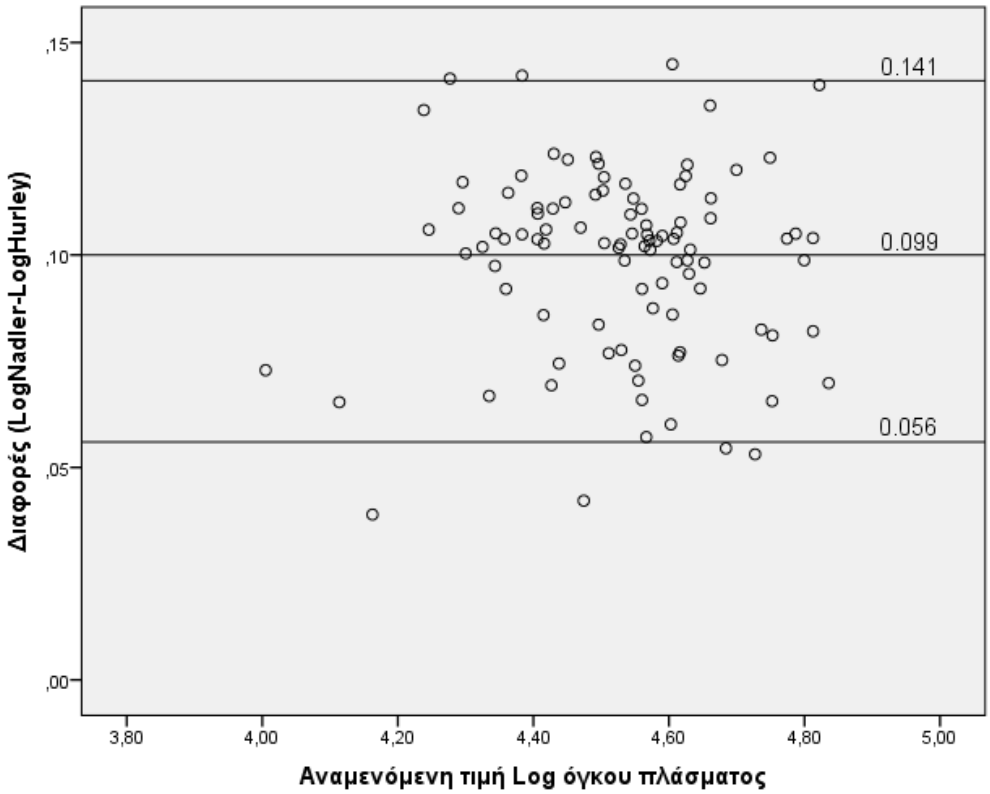
	N	Minimum	Maximum	Mean	Std. Deviation
d_log	99	,04	,14	**,0989**	**,02170**
Valid N (listwise)	99				

Πίνακας 3.2. Μέση τιμή και τυπική απόκλιση των διαφορών (Nadler-Hurley) μετά το λογαριθμικό μετασχηματισμό.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ



Γράφημα 6(α). Λογαριθμοποιημένες μετρήσεις όγκου πλάσματος σύμφωνα με τα στοιχεία του Πίνακα 3.1



Γράφημα 6(β). Διαφορές και αναμενόμενες τιμές των λογαριθμοποιημένων μετρήσεων των δύο μεθόδων με το 95% όριο συμφωνίας.

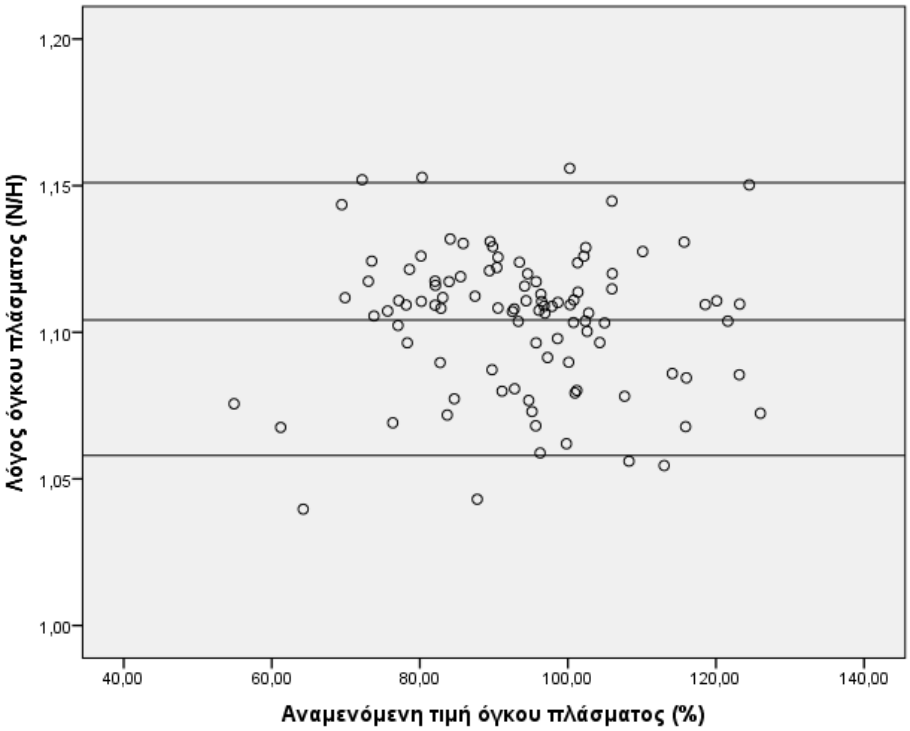
ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

Καθώς τα αποτελέσματα αυτά αφορούν τα λογαριθμοποιημένα ποσοστά είναι δύσκολη η ερμηνεία τους. Όπως αναφέρθηκε προηγουμένως, θα πρέπει να τα απολογαριθμοποιήσουμε ώστε να προκύψουν τιμές σχετικές με το λόγο των μετρήσεων των δύο μεθόδων. Ο λόγος του γεωμετρικού μέσου της μεθόδου Nadler σε σχέση με τη μέθοδο Hurley ισούται με $e^{0.099} = 1.11$ και το αντίστοιχο όριο συμφωνίας ισούται με $(e^{0.056}, e^{0.141}) = (1.06, 1.15)$. Συνεπώς, η μέθοδος Nadler δίνει μεγαλύτερες τιμές σε σχέση με τη μέθοδο Hurley από 6% έως 15% στο 95% των μετρήσεων.

Το παράδειγμα αυτό απεικονίζει τη σημαντικότητα τόσο της κλίσης $(\bar{d})$, όσο και της τυπικής απόκλισης των διαφορών μεταξύ των δύο μεθόδων (s_d).Όπως φαίνεται και από το Γράφημα 4 τα δεδομένα παρουσιάζουν μια ξεκάθαρη κλίση με όλες τις παρατηρήσεις να βρίσκονται πάνω από τη γραμμή ισότητας. Επιπλέον, επειδή οι διαφορές στις μετρήσεις των μεθόδων είναι σχετικά μικρές και συγκεντρωμένες γύρω από τη μέση διαφορά, εάν θέλουμε να εμφανίσουμε μια τέλεια συμφωνία μεταξύ των μεθόδων μπορούμε να εφαρμόσουμε συντελεστή μετατροπής πολλαπλασιάζοντας τη μέθοδο Hurley ή διαιρώντας τη μέθοδο Nadler με το λόγο του γεωμετρικού μέσου 1.11.

Θα μπορούσαμε να κάνουμε τη διαδικασία μετασχηματισμού πιο ξεκάθαρη εάν υπολογίζαμε και δουλεύαμε εξαρχής με το λόγο των δύο μεθόδων. Αντί δηλαδή, να υπολογίσουμε τους λογάριθμους και τις αντίστοιχες διαφορές των δύο μεθόδων, να υπολογίσουμε το λόγο των δύο μετρήσεων για κάθε υποκείμενο και να βρούμε την αναμενόμενη τιμή καθώς και την τυπική απόκλισή τους και στη συνέχεια τα όρια συμφωνίας. Στο Γράφημα 7 παρουσιάζεται η απεικόνιση των παραπάνω υπολογισμών και παρατηρούμε ότι είναι πανομοιότυπο με το Γράφημα 6(β).

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ



Γράφημα 7. Λόγος όγκου πλάσματος (N/H) σε σχέση με τις αναμενόμενες τιμές των μετρήσεων των δύο μεθόδων με το 95% όριο συμφωνίας.

Το Γράφημα 7 προτείνεται σαν μια γενική μέθοδος προσέγγισης για τη σύγκριση δύο μεθόδων[5], χωρίς ωστόσο να γίνεται αναφορά για την ποσοτικοποίηση των διαφορών μεταξύ των μεθόδων. Μια άλλη εναλλακτική είναι να δημιουργία γραφήματος όπου στον κάθετο άξονα να παρουσιάζονται οι διαφορές μεταξύ των δύο μεθόδων ως ποσοστό σε σχέση με την μέση τιμή τους[6].

3.2. Προσέγγιση παλινδρόμησης για μη ομοιόμορφες διαφορές

Κάποιες φορές η σχέση μεταξύ των διαφορών και των αναμενόμενων τιμών τους είναι περίπλοκη και ο λογαριθμικός μετασχηματισμός δεν επιλύει το πρόβλημα. Για παράδειγμα, οι διαφορές ενδέχεται να τείνουν προς τη μία κατεύθυνση για τις μικρές μετρήσεις και προς την αντίθετη για τις μεγάλες. Για τα δεδομένα στα οποία η σχέση μεταξύ διαφορών και μεγέθους δείγματος δεν εξαλείφεται με το λογαριθμικό μετασχηματισμό, ένα γράφημα όπως το Γράφημα 6, παραμένει εξαιρετικά χρήσιμο για τη σύγκριση των δύο μεθόδων[1]. Σε αυτές τις περιπτώσεις, η τυπική ανάλυση που περιγράφεται στο

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

Κεφάλαιο 2 έχει την τάση να δίνει όρια συμφωνίας τα οποία έχουν μεγάλη απόσταση μεταξύ τους, με αποτέλεσμα να οδηγούν στη αποδοχή της έλλειψης συμφωνίας μεταξύ των μεθόδων. Συνεπώς, είναι πιο χρήσιμο να χρησιμοποιούμε καλύτερες προσεγγίσεις όταν πρέπει να εξετάσουμε τέτοια δεδομένα.

Μια προτεινόμενη μέθοδος, είναι η μοντελοποίηση της διακύμανσης των τυπικών αποκλίσεων (SD) των διαφορών (d_i) ως συνάρτηση του επιπέδου των μετρήσεων, χρησιμοποιώντας μια μέθοδο που βασίζεται στα απόλυτα κατάλοιπα που προκύπτουν από την προσαρμογή της κατάλληλης γραμμής παλινδρόμησης. Εάν η αναμενόμενη τιμή αλλάζει και πάλι σε συνάρτηση με το επίπεδο των μετρήσεων, μπορούμε πρώτα να μοντελοποιήσουμε τη σχέση αυτή και στη συνέχεια να μοντελοποιήσουμε τις τυπικές αποκλίσεις (SD) των καταλοίπων. Η προσέγγιση αυτή βασίζεται στην αντίστοιχη που χρησιμοποιείται για τη δημιουργία διαστημάτων αναφοράς που σχετίζονται με την ηλικία[7].

Αρχικά εξετάζουμε την αναμενόμενη διαφορά μεταξύ των δύο μεθόδων σε σχέση με το μέγεθος των μετρήσεων. Είναι γνωστό[2] ότι σε μεθόδους σύγκρισης μελετών όταν η συμφωνία των δύο μεθόδων ποικίλει καθώς οι μετρήσεις ποικίλουν, να εφαρμόζεται παλινδρόμηση μεταξύ των διαφορών των μεθόδων (D) και της αναμενόμενης τιμής αυτών (A). Υπάρχει βέβαια και η περίπτωση και η τυπική απόκλιση των διαφορών (s_d) να μην παραμένει σταθερή σε σχέση με την αναμενόμενη τιμή των διαφορών (A), παρ' όλα αυτά δεν θα εξετάσουμε την συγκεκριμένη περίπτωση. Ανάλογη ιδέα προτάθηκε από τον Marshall et al.[8], ο οποίος χρησιμοποίησε πιο πολύπλοκη μέθοδο από αυτή που προτείνουμε.

Εκτός και αν το γράφημα των δεδομένων δείχνει ξεκάθαρη καμπυλότητα (γεγονός που είναι απίθανο να συμβεί), το μοντέλο της απλής γραμμικής παλινδρόμησης είναι το καταλληλότερο

$$\widehat{D} = b_0 + b_1 A \quad (3.1)$$

Στην περίπτωση που η κλίση b_1 δεν είναι σημαντική, τότε $\widehat{D} = \bar{d}$ η αναμενόμενη τιμή της διαφοράς. Να σημειωθεί ότι η σημαντικότητα ή όχι δεν

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

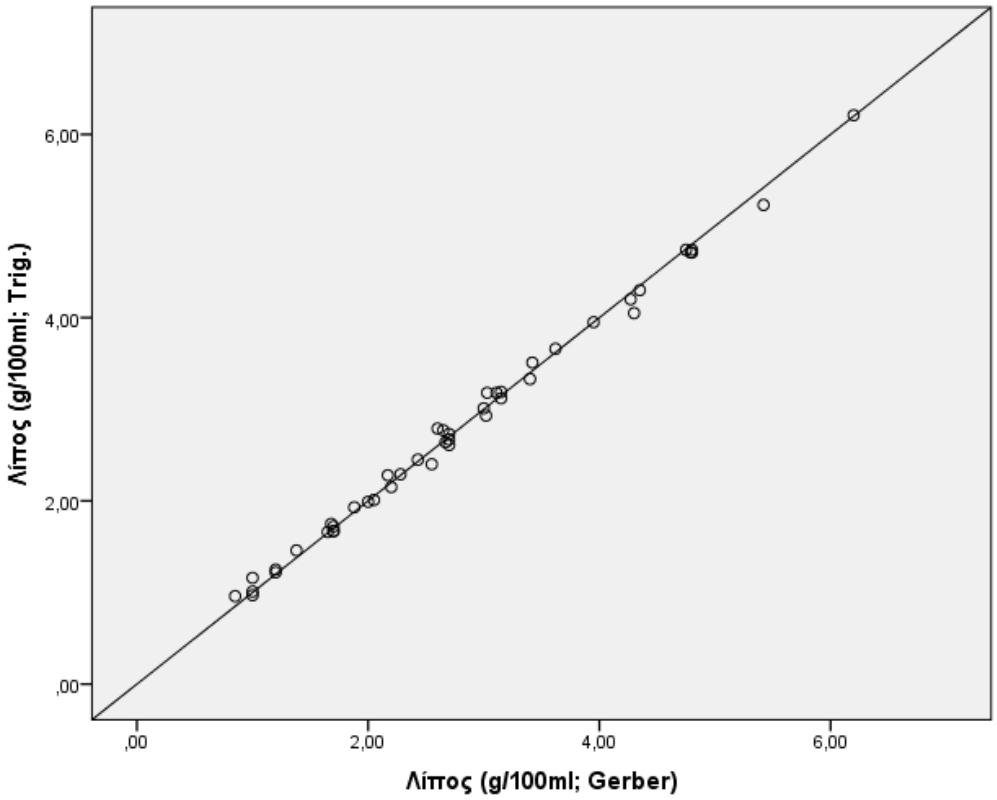
εξαρτάται μόνο από τη στατιστική ερμηνεία αλλά και από την κλινική κρίση. Εάν τώρα, η b_1 διαφέρει σημαντικά από το μηδέν, η εκτίμηση των διαφορών των δύο μεθόδων υπολογίζεται βάσει της σχέσης (3.1) για κάθε πραγματική τιμή των μετρήσεων, εκτιμώμενες από τη μεταβλητή A.

Στον Πίνακα 3.3. που ακολουθεί περιέχονται οι εκτιμώμενες περιεκτικότητες λίπους στο ανθρώπινο γάλα (g/100ml). Οι μετρήσεις καθορίζονται από τη γλυκερίνη, η οποία απελευθερώνεται από την ενζυμική υδρόλυση των τριγλυκεριδίων και από τη τυπική μέθοδο Gerber[9]. Από το Γράφημα 8 φαίνεται ότι οι δύο μέθοδοι συμφωνούν αρκετά, όμως στο Γράφημα 9 παρατηρούμε την τάση των διαφορών να βρίσκονται σε αντίθετες κατευθύνσεις για χαμηλές και υψηλές περιεκτικότητες. Η διακύμανση (SD) των διαφορών φαίνεται να παραμένει σταθερή σε όλα τα επίπεδα της περιεκτικότητας του λίπους. Το γεγονός αυτό επιβεβαιώνεται και από την ανάλυση παλινδρόμησης που ακολουθεί.

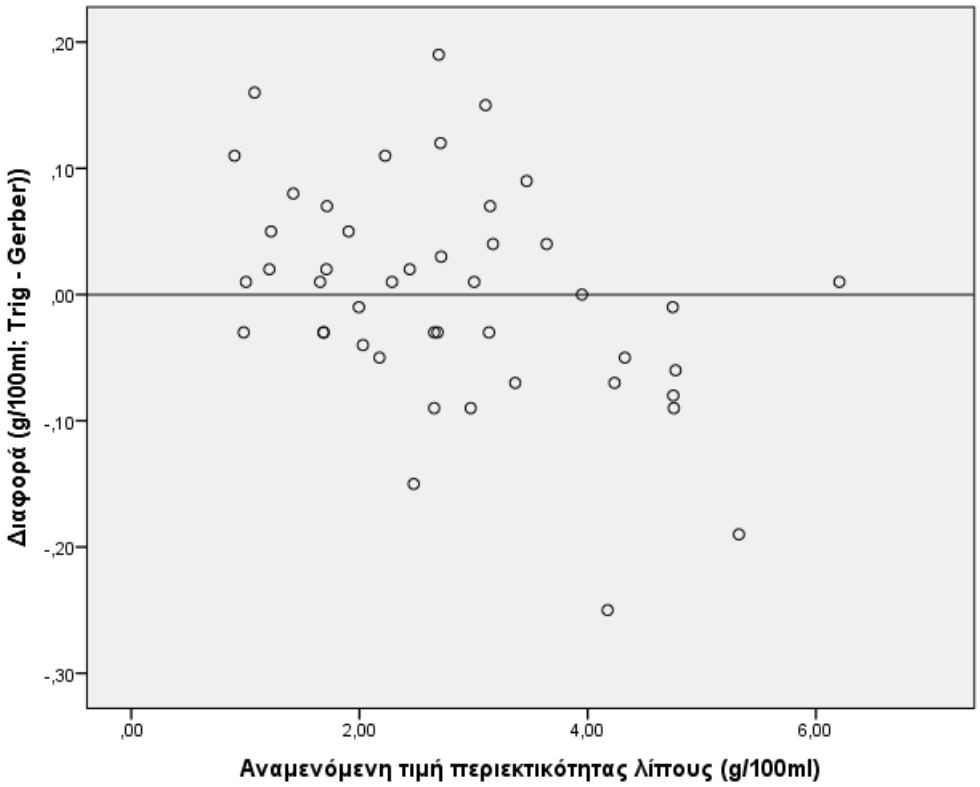
Trig	Gerber	Trig	Gerber	Trig	Gerber
0,96	0,85	2,28	2,17	3,19	3,15
1,16	1,00	2,15	2,20	3,12	3,15
0,97	1,00	2,29	2,28	3,33	3,40
1,01	1,00	2,45	2,43	3,51	3,42
1,25	1,20	2,40	2,55	3,66	3,62
1,22	1,20	2,79	2,60	3,95	3,95
1,46	1,38	2,77	2,65	4,20	4,27
1,66	1,65	2,64	2,67	4,05	4,30
1,75	1,68	2,73	2,70	4,30	4,35
1,72	1,70	2,67	2,70	4,74	4,75
1,67	1,70	2,61	2,70	4,71	4,79
1,67	1,70	3,01	3,00	4,71	4,80
1,93	1,88	2,93	3,02	4,74	4,80
1,99	2,00	3,18	3,03	5,23	5,42
2,01	2,05	3,18	3,11	6,21	6,20

Πίνακας 3.3.Περιεκτικότητες λίπους στο ανθρώπινο γάλα με την ενζυμική διαδικασία καθορισμού των τριγλυκεριδίων, και με την τυπική μέθοδο Gerber (g/100ml)[9].

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ



Γράφημα 8.Περιεκτικότητα λίπους στο ανθρώπινο γάλα με τη μέθοδο που καθορίζεται από την ενζυμική διαδικασία (Trig) και με την τυπική μέθοδο Gerber (g/100ml)



Γράφημα 9.Διαφορές των μετρήσεων των δύο μεθόδων (Trig – Gerber) σε σχέση με την αναμενόμενη τιμή τους.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

Το μοντέλο παλινδρόμησης (το οποίο συμφωνεί και με τον Lucas et al[9]) είναι :

$$\widehat{D} = 0.079 - 0.028\, A\ g/100ml$$

Όπως προκύπτει από τον Πίνακα 3.4. που ακολουθεί, και η τυπική απόκλιση των καταλοίπων ισούται με SD = 0.08033.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	**,079**	,029		2,720	,009
	Average	**-,028**	,009	-,415	-2,993	,005

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,415[a]	,172	,153	**,08033**

Πίνακας 3.4. Ανάλυση παλινδρόμησης του μοντέλου . $\widehat{D} = b_0 + b_1 A$ *για τα δεδομένα του Πίνακα 3.3.*

Στη δεύτερη φάση της ανάλυσης, θεωρούμε τη διακύμανση των μετρήσεων γύρω από τη καλύτερη γραμμή συμφωνίας, δηλαδή τη γραμμή παλινδρόμησης του μοντέλου 3.1. Θα πρέπει να μοντελοποιήσουμε τη διασπορά των καταλοίπων του μοντέλου 3.1. ως συνάρτηση του μεγέθους των μετρήσεων, το οποίο εκτιμάται από τη μεταβλητή A. Η μοντελοποίηση αυτή απλοποιείται σημαντικά από την παραδοχή ότι τα κατάλοιπα ακολουθούν κανονική κατανομή ανεξαρτήτως του μεγέθους των μετρήσεων, γεγονός το οποίο αποτελεί μια φυσική επέκταση της υπόθεσης που έχουμε ήδη κάνει σε τέτοιου τύπου αναλύσεις. Η διαδικασία αυτή, θα γίνει με βάση τις απόλυτες τιμές των καταλοίπων R στις μετρήσεις A, πιο συγκεκριμένα:

$$\widehat{\mathrm{R}} = \mathrm{c}_0 + \mathrm{c}_1\mathrm{A} \quad (3.2)$$

Εάν θεωρήσουμε μια κανονική κατανομή με μέση τιμή μηδέν και διακύμανση σ^2, εύκολα αποδεικνύεται ότι η αναμενόμενη τιμή των απόλυτων τιμών αυτής, οι οποίες θα ακολουθούν half-normal κατανομή, ισούται με $\sigma\sqrt{2/\pi}$.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

Για να υπολογίσουμε τις εκτιμήσεις των τυπικών αποκλίσεων των καταλοίπων, πολλαπλασιάζουμε τις εκτιμώμενες τιμές των απόλυτων καταλοίπων με $\sqrt{\pi/2}$. Τα όρια συμφωνίας υπολογίζονται συνδυάζοντας και τις δύο ισότητες παλινδρόμησης, (Altman[7]).

Αν και οποιοδήποτε μοντέλο παλινδρόμησης μπορεί να εφαρμοστεί στα δεδομένα, αν το SD δεν είναι σταθερό, θεωρούμε ότι το μοντέλο της γραμμικής παλινδρόμησης είναι επαρκές[1]. Στη περίπτωση που δεν υπάρχει «σημαντική» σχέση μεταξύ R και A, η εκτίμηση της τυπικής απόκλισης είναι απλά η τυπική απόκλιση των διαφορών, δηλαδή η τυπική απόκλιση των καταλοίπων του μοντέλου 3.1.

Στη γενική περίπτωση όπου και τα δύο μοντέλα (3.1) και (3.2) χρησιμοποιούνται, η αναμενόμενη τιμή των διαφορών μεταξύ των δύο μεθόδων, δίνεται από τη σχέση $\widehat{D} = b_0 + b_1 A$ (3.1.) και το 95% όριο συμφωνίας υπολογίζεται ως:

$$\widehat{D} \pm 1.96\sqrt{\pi/2}\,\hat{R} = \widehat{D} \pm 2.46\hat{R}$$

ή ως:

$$\mathrm{b_0} + \mathrm{b_1 A} \pm 2.46\{\mathrm{c_0} + \mathrm{c_1 A}\} \quad (3.3)$$

Επιστρέφοντας στο παράδειγμά μας, όπως διαπιστώνουμε από τον Πίνακα 3.5 που ακολουθεί, δεν υπάρχει σχέση μεταξύ των απόλυτων καταλοίπων R και των μετρήσεων A, συνεπώς ο υπολογισμός του SD των διαφορών, θα γίνει από το SD των καταλοίπων του μοντέλου 3.1 και θα ισούται με $s_d = 0.08033$. Το 95% όριο συμφωνίας υπολογίζεται ως $0.079 - 0.0283\alpha \pm 1.96 \times 0.08033\ g/100ml$, όπου α είναι το μέγεθος (μέσος όρος μετρήσεων των δύο μεθόδων) της περιεκτικότητας λίπους. Οι υπολογισμοί αυτοί απεικονίζονται στο Γράφημα 10. Φυσικά, όταν χρησιμοποιείται μόνο μια μέθοδος, το α αντιστοιχεί στην παρατηρούμενη τιμή της μεθόδου.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,133[a]	,018	-,005	,04987

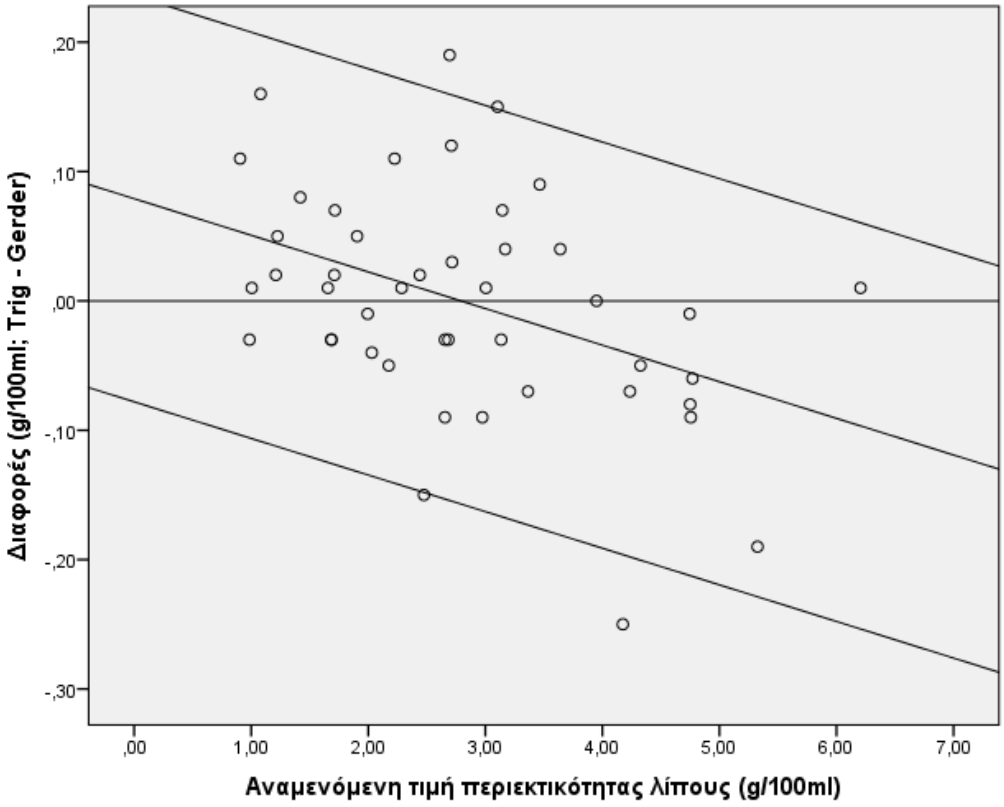
ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	,002	1	,002	,776	**,383[b]**
	Residual	,107	43	,002		
	Total	,109	44			

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	,047	,018		2,590	,013
	Average	,005	,006	,133	,881	,383

Πίνακας 3.5. Ανάλυση παλινδρόμησης του μοντέλου . $\hat{R} = c_0 + c_1 A$ *για τα δεδομένα του Πίνακα 3.3.*



Γράφημα 10. 95% όριο συμφωνίας βασισμένο στην παλινδρόμηση των δύο μεθόδων Trig και Gerber.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ


ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

Κεφάλαιο 4. Η σημαντικότητα της επαναληψιμότητας

Οι επαναλαμβανόμενες μετρήσεις σε μεθόδους μέτρησης, μπορεί να επηρεάσουν τη συμφωνία μεταξύ των δύο συγκρινόμενων μεθόδων, συνεπώς θα πρέπει να εξετάζεται και η επαναληψιμότητα των μετρήσεων. Περιέργως, οι επαναλαμβανόμενες μετρήσεις δεν μελετώνται, με αποτέλεσμα μια σημαντική πτυχή της σύγκρισης των μελετών να παραβλέπεται. Στην περίπτωση που πραγματοποιείται μία μέτρηση από κάθε μέθοδο σε κάθε υποκείμενο, δεν είμαστε σε θέση να αξιολογήσουμε ποια μέθοδος είναι περισσότερο επαναληπτική (ακριβής). Η έλλειψη επαναληψιμότητας μπορεί να επέμβει στη σύγκριση των δύο μεθόδων, καθώς αν μια μέθοδος έχει χαμηλή επαναληψιμότητα, με την έννοια της ύπαρξης σημαντικής διακύμανσης στις επαναλαμβανόμενες μετρήσεις πάνω στο ίδιο υποκείμενο, η συμφωνία των δύο μεθόδων θα είναι αναγκαστικά μικρή. Ακόμη και αν οι μετρήσεις των δύο μεθόδων συμφωνούν κατά μέσο όρο, η χαμηλή επαναληψιμότητας της μιας μπορεί να οδηγήσει σε ασθενή συμφωνία των δύο μεθόδων για τα υποκείμενα. Όταν μια παλιά μέθοδος έχει χαμηλή επαναληψιμότητα, ακόμη και μια νέα μέθοδος η οποία θα ήταν ιδανική, δεν θα συμφωνούσε. Η έλλειψη συμφωνίας σε μελέτες κατά τις οποίες δεν έγιναν επαναλαμβανόμενες μετρήσεις, υποδηλώνει ότι η νέα μέθοδος δεν μπορεί να χρησιμοποιηθεί, γεγονός όμως που μπορεί να οφείλεται στην έλλειψη επαναληψιμότητας της κύριας μεθόδου. Αν και οι δύο μέθοδοι έχουν χαμηλή επαναληψιμότητα, τότε η έλλειψη της μεταξύ τους συμφωνίας είναι πολύ πιθανή. Για τους λόγους αυτούς, προτείνεται η ταυτόχρονη εκτίμηση της επαναληψιμότητας και της συμφωνίας, συγκεντρώνοντας επαναλαμβανόμενα δεδομένα.

Είναι σημαντικό να διευκρινίσουμε τι εννοούμε λέγοντας επαναλαμβανόμενες μετρήσεις, στη πράξη πρόκειται για δύο ή περισσότερες μετρήσεις που πραγματοποιούνται στο ίδιο υποκείμενο κάτω από τις ίδιες συνθήκες. Γενικεύοντας, η απαίτηση αυτή προϋποθέτει μετρήσεις οι οποίες γίνονται σε γρήγορη διαδοχή.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

Επιπλέον είναι σημαντικό οι επαναλαμβανόμενες μετρήσεις να είναι ανεξάρτητες μεταξύ τους. Στην ουσία, ανεξαρτησία μεταξύ των μετρήσεων επιτυγχάνεται όταν διασφαλίζεται ότι ο παρατηρητής κάνει κάθε μέτρηση μη λαμβάνοντας υπόψη την τιμή της προηγούμενης, γεγονός που ενδεχομένως να είναι δύσκολα εφαρμόσιμο στην πράξη.

4.1. Εκτιμώντας την επαναληψιμότητα

Η ανάλυση που εφαρμόζεται για την ποσοτικοποίηση της επαναληψιμότητας μια μεθόδου λαμβάνοντας υπόψη τις επαναλαμβανόμενες μετρήσεις, είναι σχεδόν παρόμοια με εκείνη που χρησιμοποιούμε για τον καθορισμό των ορίων συμφωνίας. Εφαρμόζοντας ανάλυση διακύμανσης (one-way anova), με παράγοντα το υποκείμενο, μπορούμε να εκτιμήσουμε την τυπική απόκλιση εντός των υποκειμένων, s_w, από την τετραγωνική ρίζα του μέσου τετραγωνικού σφάλματος. Συγκρίνοντας τις τυπικές αποκλίσεις διαφορετικών μεθόδων μπορούμε να διαπιστώσουμε ποια μέθοδος εμφανίζει μεγαλύτερη επαναληψιμότητα. Κάθε τυπική απόκλιση μπορεί να χρησιμοποιηθεί για τη δημιουργία ορίων, εντός των οποίων αναμένεται να κυμαίνονται οι διαφορές μεταξύ δύο μετρήσεων της ίδιας μεθόδου. Επιπλέον, η επαναληψιμότητα υποδεικνύει μια βάση, σύμφωνα με την οποία μπορούμε να κρίνουμε τη μεταβλητότητα δύο μεθόδων.

Η ανάλυση θεωρείται απλή, καθώς αναμένουμε η μέση διαφορά μεταξύ των επαναλήψεων να ισούται με το μηδέν, άλλωστε δεν αποτελεί σύνηθες φαινόμενο οι δεύτερες μετρήσεις στο ίδιο δείγμα να διαφέρουν συστηματικά από τις πρώτες. Στη πράξη μια τέτοια συστηματική διαφορά θα υποδήλωνε ότι οι μετρήσεις δεν μπορούν να θεωρηθούν επαναλαμβανόμενες[1]. Με τη χρήση ενός γραφήματος θα μπορούσαμε να διαπιστώσουμε αν οι υποθέσεις αυτές ισχύουν και εάν οι διαφορές των μετρήσεων κάθε υποκειμένου, είναι ανεξάρτητες της μέσης τιμής τους. Στην περίπτωση που η επαναληψιμότητα χειροτερεύει καθώς αυξάνεται ο αριθμός των μετρήσεων, ίσως χρειαστεί να εφαρμόσουμε λογαριθμικό μετασχηματισμό στα δεδομένα, ανάλογο με αυτό που εφαρμόστηκε στη σύγκριση των μεθόδων.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

Μελετώντας και πάλι τα δεδομένα του Πίνακα 1 για τη συστολική αρτηριακή πίεση, και όπως φαίνεται και στον Πίνακα 4.1 που ακολουθεί, για τον παρατηρητή J,η διακύμανση εντός των υποκειμένων είναι $s_{wJ}^2 = 37.408$, για τον παρατηρητή R $s_{wR}^2 = 37.980$, και για το μηχάνημα S $s_{wS}^2 = 83.141$. Διαπιστώνουμε ότι οι δύο παρατηρητές έχουν καλύτερη επαναληψιμότητα σε σχέση με το μηχάνημα, και σχεδόν ίδιες επιδόσεις.

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
	Between Groups	238796,251	84	2842,813	75,995	,000
J	Within Groups	6359,333	170	**37,408**		
	Total	245155,584	254			
	Between Groups	234290,965	84	2789,178	73,437	,000
R	Within Groups	6456,667	170	**37,980**		
	Total	240747,631	254			
	Between Groups	254748,808	84	3032,724	36,477	,000
S	Within Groups	14134,000	170	**83,141**		
	Total	268882,808	254			

Πίνακας 4.1. Ανάλυση διακύμανσης των μεθόδων μέτρησης συστολικής αρτηριακής πίεσης από τους παρατηρητές J, R και το μηχάνημα S.

Οι απόλυτες διαφορές δύο μετρήσεων από την ίδια μέθοδο, αναμένουμε να κυμαίνονται γύρω από τη τιμή $1.96\sqrt{2}s_w$ ή $2.77s_w$ για το 95% των υποκειμένων. Η τιμή αυτή ονομάζεται *συντελεστής επαναληψιμότητας*. Για τον παρατηρητή J, ο οποίος χρησιμοποιεί σφυγμομανόμετρο έχουμε ότι $s_{wJ} = \sqrt{37.408} = 6.116$ και συνεπώς ο συντελεστής επαναληψιμότητάς του είναι $2.77 \times 6.116 = 16.95mmHg$. Για το μηχάνημα S αντίστοιχα, $s_{ws} = \sqrt{83.141} = 9.118$ και ο συντελεστής επαναληψιμότητας ισούται με $2.77 \times 9.118 = 25.27mmHg$. Συνεπώς η επαναληψιμότητα του μηχανήματος είναι σχεδόν 50% μεγαλύτερη από του παρατηρητή J. Μπορούμε να συγκρίνουμε αυτούς τους 95% συντελεστές επαναληψιμότητας με το 95% όριο συμφωνίας. Τα 95% όρια συμφωνίας αντιστοιχούν στο διάστημα $(-2.77s_w, 2.77s_w)$. Αν διαπιστώσουμε ότι είναι παρόμοια, τότε η έλλειψη συμφωνίας μεταξύ των μεθόδων ερμηνεύεται από την έλλειψη επαναληψιμότητας. Εάν τα όρια συμφωνίας είναι αξιοσημείωτα ευρύτερα από ότι θα υποδείκνυε η

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

επαναληψιμότητα, τότε θα πρέπει να υπάρχει κάποιος άλλος παράγοντας στον οποίο οφείλεται η μείωση της συμφωνίας των μεθόδων.

Η χρήση της τυπικής απόκλισης εντός των υποκειμένων, δεν σημαίνει ότι άλλες προσεγγίσεις για την επαναληψιμότητα δεν είναι κατάλληλες. Η χρήση όμως της s_w διευκολύνει τη σύγκριση με τα όρια συμφωνίας και επιπλέον βοηθά στην ερμηνεία των ατομικών μετρήσεων οι οποίες πραγματοποιούνται στις ίδιες μονάδες μέτρησης.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

Κεφάλαιο 5. Μέτρηση συμφωνίας μεθόδων με τη χρήση επαναλαμβανόμενων μετρήσεων

Όταν διαθέτουμε επαναλαμβανόμενες μετρήσεις για δύο μεθόδους στα ίδια υποκείμενα, είναι ιδιαιτέρως επιθυμητό να χρησιμοποιούμε όλα τα δεδομένα για τη σύγκριση των μεθόδων αυτών. Αρχικά υπολογίζουμε τη μέση τιμή των επαναλαμβανόμενων μετρήσεων για κάθε υποκείμενο και στις δύο μεθόδους. Στη συνέχεια, μπορούμε να χρησιμοποιήσουμε τα νέα πλέον ζεύγη τιμών που προκύπτουν από τις μέσες τιμές για να κάνουμε τη σύγκριση χρησιμοποιώντας τη μέθοδο των ορίων συμφωνίας. Η εκτίμηση της κλίσης (bias) θα μείνει ανεπηρέαστη από τη χρήση των μέσων όρων, όμως η εκτίμηση της τυπικής απόκλισης των διαφορών θα είναι πολύ μικρή, δεδομένου ότι μέρος των σφαλμάτων μέτρησης έχει ήδη αφαιρεθεί. Χρειαζόμαστε την τυπική απόκλιση των διαφορών μεταξύ των μονών μετρήσεων των δύο μεθόδων, και όχι την τυπική απόκλιση των διαφορών μεταξύ των μέσων στις επαναλαμβανόμενες μετρήσεις. Θα περιγράψουμε κάποιες μεθόδους για τον χειρισμό τέτοιων δεδομένων, αρχικά για την περίπτωση που διαθέτουμε ισάριθμο πλήθος επαναλήψεων και στη συνέχεια για άνισο αριθμό επαναλήψεων.

Υποθέτουμε ότι ακόμα και αν διαθέτουμε πολλαπλές μετρήσεις, για την κύρια μέθοδο υπάρχει μία τιμή μέτρησης. Στις περιπτώσεις στις οποίες συνηθίζεται να χρησιμοποιείται για την κύρια μέθοδο ο μέσος όρος δύο ή περισσότερων μετρήσεων σε κλινική χρήση, η προσέγγιση που θα παρουσιαστεί παρακάτω δεν συνίσταται, είναι προτιμότερο τα όρια συμφωνίας να αναφέρονται στις μέσες τιμές.

5.1. Ισάριθμο πλήθος επαναλήψεων

Όταν κάνουμε επαναλαμβανόμενες μετρήσεις στο ίδιο υποκείμενο για κάθε μία από τις δύο μεθόδους που θέλουμε να συγκρίνουμε, οι μετρήσεις κάθε μεθόδου κατανέμονται γύρω από την αναμενόμενη τιμή της αντίστοιχης

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

μεθόδου για το κάθε υποκείμενο. Οι αναμενόμενες τιμές αυτές, δεν είναι αναγκαίο να ταυτίζονται για τις δύο μεθόδους. Η διακύμανση αυτή υποδηλώνει τη χρήση μεθόδου αλληλεπίδρασης ομάδας χρόνου. Ορίζουμε ως Χ και Υ τις μετρήσεις των δύο μεθόδων. Χρειαζόμαστε τη διακύμανση των διαφορών των μονών μετρήσεων κάθε μεθόδου, δηλαδή $D = X - Y$. Η διακύμανση για κάθε μέθοδο ξεχωριστά είναι :

$$Var(X) = \sigma_t^2 + \sigma_{xI}^2 + \sigma_{xw}^2$$

$$Var(X) = \sigma_t^2 + \sigma_{yI}^2 + \sigma_{yw}^2$$

όπου σ_t^2 η διακύμανση των πραγματικών τιμών,σ_{xI}^2 και σ_{yI}^2 η αλληλεπίδραση χρόνου και ομάδας στο υποκείμενο και σ_{xw}^2 και σ_{yw}^2 οι εντός των υποκειμένων διακυμάνσεις των μετρήσεων ίδιας μεθόδου, για τις μεθόδους Χ και Υ αντίστοιχα. Συνεπώς, η διακύμανση των διαφορών μεταξύ των υποκειμένων, για τις μονές μετρήσεις κάθε μεθόδου είναι[1] :

$$\mathrm{Var(X - Y)} = \sigma_{\mathrm{D}}^2 = \sigma_{\mathrm{xI}}^2 + \sigma_{\mathrm{yI}}^2 + \sigma_{\mathrm{xw}}^2 + \sigma_{\mathrm{yw}}^2 \qquad (5.1)$$

Θα θέλαμε να εκτιμήσουμε τη διακύμανση των διαφορών ($Var(\bar{X} - \bar{Y})$) από την ανάλυση των μέσων μετρήσεων για κάθε υποκείμενο, $\bar{D} = \bar{X} + \bar{Y}$. Με το μοντέλο αυτό, η χρήση της μέση τιμής των επαναλήψεων θα μειώσει τη διακύμανση εντός των υποκειμένων, αλλά δεν θα επηρεάσει τους όρους της αλληλεπίδρασης που αντιπροσωπεύουν τις εξατομικευμένες διαφορές των ασθενών. Συνεπώς :

$$Var(\bar{X}) = \sigma_t^2 + \sigma_{xI}^2 + \frac{\sigma_{xw}^2}{m_x}$$

Όπου m_x είναι το πλήθος των παρατηρήσεων σε κάθε υποκείμενο με τη μέθοδο Χ, γιατί υπολογίζεται η μέση τιμή του σφάλματος μόνο εντός των υποκειμένων και εντός της ίδια μεθόδου. Αντίστοιχα:

$$Var(\bar{Y}) = \sigma_t^2 + \sigma_{yI}^2 + \frac{\sigma_{yw}^2}{m_y}$$

$$Var(\bar{X} - \bar{Y}) = Var(\bar{D}) = \sigma_{xI}^2 + \frac{\sigma_{xw}^2}{m_x} + \sigma_{yI}^2 + \frac{\sigma_{yw}^2}{m_y}$$

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

Η κατανομή της $\overline{D}$ εξαρτάται μόνο από τα σφάλματα και τις αλληλεπιδράσεις, διότι η πραγματική της τιμή συμπεριλαμβάνεται και στη X και στη Y, οι οποίες διαφοροποιούνται. Από τη εξίσωση (5.1) προκύπτει ότι[1] :

$$\mathrm{Var}(\mathrm{X}-\mathrm{Y}) = \mathrm{Var}(\overline{\mathrm{D}}) + \left(1-\frac{1}{\mathrm{m_x}}\right)\sigma^2_{\mathrm{xw}} + \left(1-\frac{1}{\mathrm{m_y}}\right)\sigma^2_{\mathrm{yw}} \qquad (5.2)$$

Εάν $s_{\bar{d}}^2$ είναι η παρατηρούμενη διακύμανση των διαφορών των εντός των υποκειμένων μέσων η $Var(X-Y) = \sigma_{\bar{d}}^2$ εκτιμάται από τη σχέση :

$$\widehat{\sigma}^2_{\mathrm{d}} = \mathrm{s}^2_{\bar{\mathrm{d}}} + \left(1-\frac{1}{\mathrm{m_x}}\right)\mathrm{s}^2_{\mathrm{xw}} + \left(1-\frac{1}{\mathrm{m_y}}\right)\mathrm{s}^2_{\mathrm{yw}} \qquad (5.3)$$

Και στη συνηθισμένη περίπτωση των δύο επαναλήψεων για κάθε μέθοδο, έχουμε :

$$\widehat{\sigma}_{\bar{d}}^2 = s_{\bar{d}}^2 + \frac{s_{xw}^2}{2} + \frac{s_{yw}^2}{2}$$

όπως δίνεται από τους Bland και Altman[3]. Εύκολα διαπιστώνουμε ότι η μέθοδος αυτή δουλεύει ακόμα και στην περίπτωση όπου η μία μέθοδος έχει επαναλαμβανόμενες μετρήσεις και η άλλη όχι.

Ως παράδειγμα θα εξετάσουμε και πάλι τα δεδομένα του Πίνακα 1. Όπως δείξαμε στην Παράγραφο 4.1 οι εντός των υποκειμένων διακυμάνσεις των δύο μεθόδων για την μέτρηση της συστολικής αρτηριακής πίεσης είναι 37.408 για το σφυγμομανόμετρο (παρατηρητής J) και 38.141 για τη μηχανή (S). Από τον Πίνακα 5.1 βλέπουμε ότι η μέση τιμή των διαφορών μεταξύ των μέσων τιμών των υποκειμένων είναι ίση με -15.62mmHg, η διακύμανσή τους ισούται με $s_{\bar{d}}^2 = 358.492$ και εφόσον έχουμε τρεις παρατηρήσεις και για τις δύο μεθόδους $m_J = m_S = 3$.

Descriptive Statistics

	N	Mean	Std. Deviation	Variance
d_mean	85	**-15,6196**	18,93390	**358,492**
Valid N (listwise)	85			

Πίνακας 5.1. Μέση τιμή και διακύμανση των διαφορών των μέσων τιμών των υποκειμένων με τις μεθόδους J και S.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

Από την Εξίσωση (5.3) η προσαρμοσμένη διακύμανση των διαφορών δίνεται από τη σχέση:

$$\hat{\sigma}_d^2 = 358.492 + \left(1 - \frac{1}{3}\right) \times 37.408 + \left(1 - \frac{1}{3}\right) \times 83.141 = 438.858$$

Συνεπώς $s_d = \sqrt{438.858} = 20.95 mmHg$ και το 95% όριο συμφωνίας είναι $(-15.62 - 1.96 \times 20.95, -15.62 + 1.96 \times 20.95) = (-56.68, 25.44)$.
Διαπιστώνουμε ότι η εκτίμηση αυτή είναι παρόμοια με εκείνη του Κεφαλαίου 2 (-54.73,22.15).

Το τυπικό σφάλμα και τα διαστήματα εμπιστοσύνης για αυτό το όριο συμφωνίας υπολογίζονται ως εξής[1]:

Θεωρώντας ότι τα σφάλματα ακολουθούν κανονική κατανομή και είναι ανεξάρτητα, για n υποκείμενα η $n(m_x - 1)\, s_{xw}^2/\sigma_{xw}^2$ ακολουθεί X^2 κατανομή με $n(m_x - 1)$ βαθμούς ελευθερίας και διακύμανση είναι $2n(m_x - 1)$. Συνεπώς:

$$\mathrm{Var}(s_{xw}^2) = \frac{2\sigma_{xw}^4}{n(m_x - 1)} \quad \text{και} \quad \mathrm{Var}(s_{yw}^2) = \frac{2\sigma_{yw}^4}{n(m_y - 1)} \tag{5.4}$$

Από τις Εξισώσεις (5.3) και (5.4) προκύπτει ότι:

$$\mathrm{Var}\left(\left(1 - \frac{1}{m_x}\right) s_{xw}^2 + \left(1 - \frac{1}{m_y}\right) s_{yw}^2\right) =$$

$$\left(1 - \frac{1}{m_x}\right)^2 \frac{2\sigma_{xw}^4}{n(m_x - 1)} + \left(1 - \frac{1}{m_y}\right)^2 \frac{2\sigma_{yw}^4}{n(m_y - 1)} = \tag{5.5}$$

$$\frac{2(m_x - 1)\sigma_{xw}^4}{n m_x^2} + \frac{2(m_y - 1)\sigma_{yw}^4}{n m_y^2} \tag{5.6}$$

Αντιστοίχως η διακύμανση του s_d^2 δίνεται από τη σχέση:

$$\mathrm{Var}(s_d^2) = \frac{2\sigma_d^4}{n - 1} \tag{5.7}$$

Εφαρμόζοντας τις Εξισώσεις (5.6) και (5.7) στην (5.3) προκύπτει:

$$\mathrm{Var}(\hat{\sigma}_d^2) = \frac{2\sigma_d^4}{n - 1} + \frac{2(m_x - 1)\sigma_{xw}^4}{n m_x^2} + \frac{2(m_y - 1)\sigma_{yw}^4}{n m_y^2} \tag{5.8}$$

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

Γνωρίζοντας ότι

$$Var(f(z)) \approx \left(\frac{df(z)}{dz}\right)^2_{z=E(z)} Var(z)$$

Συμπεραίνουμε

$$Var(\sqrt{z}) \approx \left(\frac{1}{2\sqrt{z}}\right)^2_{z=E(z)} Var(z) = \frac{1}{4E(z)} Var(z)$$

Προκύπτει για τη $Var(\hat{\sigma}_d)$

$$Var(\hat{\sigma}_d) = \frac{1}{4\sigma_d^2}\left(\frac{2\sigma_{\bar{d}}^4}{n-1} + \frac{2(m_x-1)\sigma_{xw}^4}{nm_x^2} + \frac{2(m_y-1)\sigma_{yw}^4}{nm_y^2}\right)$$

$$= \frac{1}{2\sigma_{\bar{d}}^2}\left(\frac{\sigma_{\bar{d}}^4}{n-1} + \frac{(m_x-1)\sigma_{xw}^4}{nm_x^2} + \frac{(m_y-1)\sigma_{yw}^4}{nm_y^2}\right) \quad (5.9)$$

Η διακύμανση της μέσης τιμής των διαφορών εκτιμάται ως $\hat{\sigma}_d^2/n$, η μέση τιμή και η τυπική απόκλιση των διαφορών είναι ανεξάρτητες. Αντικαθιστώντας την εκτίμηση της διακύμανσης στην Εξίσωση (5.9) η διακύμανση του ορίου συμφωνίας $\bar{d} \pm 1.96\hat{\sigma}_d$ εκτιμάται από τη σχέση:

$$\mathrm{Var}\left(\bar{d} \pm 1.96\hat{\sigma}_d\right) = \frac{\hat{\sigma}_{\bar{d}}^2}{n} + \frac{1.96^2}{2\hat{\sigma}_{\bar{d}}^2}\left(\frac{s_{\bar{d}}^4}{n-1} + \frac{(m_x-1)s_{xw}^4}{nm_x^2} + \frac{(m_y-1)s_{yw}^4}{nm_y^2}\right) \quad (5.10)$$

Και για $m_x = m_y = 2$, η σχέση αυτή γίνεται

$$Var\left(\bar{d} \pm 1.96\hat{\sigma}_d\right) = \frac{\hat{\sigma}_d^2}{n} + \frac{1.96^2}{2\hat{\sigma}_d^2}\left(\frac{s_{\bar{d}}^4}{n-1} + \frac{s_{xw}^4}{4n} + \frac{s_{yw}^4}{4n}\right)$$

και για μη επαναλαμβανόμενες μετρήσεις με $m_x = m_y = 1$, η $\hat{\sigma}_d$ αντικαθίσταται από την s_d και προκύπτει :

$$Var\left(\bar{d} \pm 1.96s_d\right) = \frac{s_d^2}{n} + \frac{1.96^2}{2s_d^2}\frac{s_d^4}{n-1} = s_d^2\left(\frac{1}{n} + \frac{1.96^2}{2(n-1)}\right)$$

όπως στην Παράγραφο 2.2. Οι τιμές αυτές μπορούν να χρησιμοποιηθούν για την εκτίμηση του 95% διαστήματος εμπιστοσύνης για τα όρια συμφωνίας.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

Για τα συγκεκριμένα δεδομένα της συστολικής αρτηριακής πίεσης, η διακύμανση του ορίου συμφωνίας είναι :

$$Var(\bar{D} \pm 1.96 s_D) = \frac{438.858}{85} + \frac{1.96^2}{2 \times 438.858}\left(\frac{358.492^2}{84} + \frac{2 \times 37.408^2}{9 \times 85} + \frac{2 \times 83.141^2}{9 \times 85}\right) = 11.956$$

Με τυπικό σφάλμα $\sqrt{11.956} = 3.458 mmHg$ και το 95% διάστημα εμπιστοσύνης για το κάτω όριο συμφωνίας ισούται με (-56.68-1.96x3.458, -56.68+1.96x3.458)=(-63.46, -49.90) και αντίστοιχα το άνω όριο είναι (25.44-1.96x3.458, 25.44+1.96x3.458)=(18.66, 32.22).

Το τυπικό σφάλμα είναι πολύ κοντά σε αυτό που υπολογίστηκε για μία μόνο μέτρηση (3.64mmHg) στην Παράγραφο (2.2). Η χρήση των επαναλήψεων μειώνει μόνο εκείνο το κομμάτι της διακύμανσης που οφείλεται στην έλλειψη ακρίβειας μέτρησης κάθε μεθόδου ενώ η αλληλεπίδραση ομάδας χρόνου παραμένει ανεπηρέαστη. Εάν οι αποκλίσεις είναι μεγάλες σε όλες τις επαναλήψεις του υποκειμένου οι επαναλαμβανόμενες μετρήσεις δε βελτιώνουν ιδιαίτερα την ακρίβεια του ορίου συμφωνίας. Παρόλα αυτά δύο επαναλήψεις είναι προτιμότερες έτσι ώστε να είναι δυνατή η εξέταση της επαναληψιμότητας της μεθόδου.

5.2. Άνισο πλήθος επαναλήψεων

Θεωρούμε την περίπτωση στην οποία υπάρχει άνισος αριθμός παρατηρήσεων για κάθε υποκείμενο, m_{xi} *και* m_{yi}, για τις μεθόδους X και Y αντίστοιχα για κάθε υποκείμενο i. Τέτοια δεδομένα προκύπτουν εάν πραγματοποιούνται μετρήσεις σε ασθενείς ανά τακτά χρονικά διαστήματα καθ όλη τη διάρκεια διαδικασιών οι οποίες διαφέρουν ως προς τη διάρκειά τους, όπως παραδείγματος χάριν χειρουργεία. Στον Πίνακα 5.2 παρουσιάζονται δεδομένα καρδιακής παροχής τα οποία μετρήθηκαν με δύο μεθόδους, καρδιογράφημα εμπέδησης (IC) και ραδιοϊσοτοπική κοιλιογραφία (RV), σε δώδεκα ασθενείς[10].

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

Sub	RV	IC	Sub	RV	IC	Sub	RV	IC
1	7,83	6,57	5	3,13	3,03	9	4,48	3,17
1	7,42	5,62	5	2,98	2,86	9	4,92	3,12
1	7,89	6,9	5	2,85	2,77	9	3,97	2,96
1	7,12	6,57	5	3,17	2,46	10	4,22	4,35
1	7,88	6,35	5	3,09	2,32	10	4,65	4,62
2	6,16	4,06	5	3,12	2,43	10	4,74	3,16
2	7,26	4,29	6	5,92	5,9	10	4,44	3,53
2	6,71	4,26	6	6,42	5,81	10	4,5	3,53
2	6,54	4,09	6	5,92	5,7	11	6,78	7,2
3	4,75	4,71	6	6,27	5,76	11	6,07	6,09
3	5,24	5,5	7	7,13	5,09	11	6,52	7
3	4,86	5,08	7	6,62	4,63	11	6,42	7,1
3	4,78	5,02	7	6,58	4,61	11	6,41	7,4
3	6,05	6,01	7	6,93	5,09	11	5,76	6,8
3	5,42	5,67	8	4,54	4,72	12	5,06	4,5
4	4,21	4,14	8	4,81	4,61	12	4,72	4,2
4	3,61	4,2	8	5,11	4,36	12	4,9	3,8
4	3,72	4,61	8	5,29	4,2	12	4,8	3,8
4	3,87	4,68	8	5,39	4,36	12	4,9	4,2
4	3,92	5,04	8	5,57	4,2	12	5,1	4,5

Πίνακας 5.2. Δεδομένα καρδιακής παροχής 12 ασθενών, τα οποία μετρήθηκαν με δύο μεθόδους RV και IS (Dr LS Bowling[10])

Στην περίπτωση των ισάριθμων επαναλήψεων η επίλυση της Εξίσωση (5.2) βασίζεται στο ότι η διακύμανση της αναμενόμενης τιμής n ανεξάρτητων μεταβλητών με την ίδια μέση τιμή και διακύμανση σ^2 είναι σ^2/n. Εάν W_i είναι η μέση τιμή των m_i παρατηρήσεων οι οποίες έχουν αντίστοιχα μέση τιμή μ και διακύμανση σ^2, η διακύμανσή της θα ισούται με σ^2/m_i, και συνεπώς η αναμενόμενη διακύμανση των μέσων θα είναι

$$\mathrm{Var}(\mathrm{W_i}) = \frac{1}{\mathrm{n}}\left(\sum \frac{1}{\mathrm{m_i}}\right)\sigma^2 \qquad (5.11)$$

Για κάθε υποκείμενο i, θεωρούμε ως m_{xi} τις παρατηρήσεις από τη μέθοδο X και m_{yi} τις παρατηρήσεις της μεθόδου Y αντίστοιχα. Υπολογίζουμε τις διαφορές των μέσων τιμών των δύο μεθόδων για κάθε υποκείμενο και στη συνέχεια την εκτίμηση της διακύμανσης των μέσων διαφορών, από τη σχέση :

$$\mathrm{Var}(\overline{\mathrm{D}}) = \sigma^2_{\mathrm{xI}} + \frac{1}{\mathrm{n}}\left(\sum \frac{1}{\mathrm{m_{xi}}}\right)\sigma^2_{\mathrm{xw}} + \sigma^2_{\mathrm{yI}} + \frac{1}{\mathrm{n}}\left(\sum \frac{1}{\mathrm{m_{yi}}}\right)\sigma^2_{\mathrm{yw}} \qquad (5.12)$$

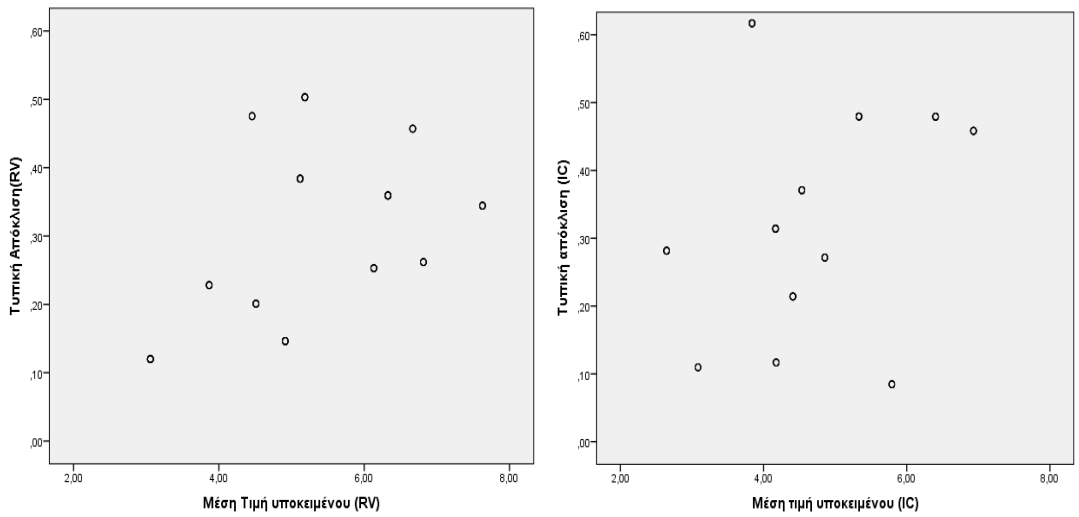
ATHENS UNIVERSITY OF ECONOMICS & BUSINESS ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

και καταλήγουμε:

$$\mathrm{Var(D)} = \mathrm{Var}(\overline{\mathrm{D}}) + \left(1 - \frac{1}{\mathrm{n}}\left(\sum \frac{1}{\mathrm{m_{xi}}}\right)\right)\sigma^2_{\mathrm{xw}} + \left(1 - \frac{1}{\mathrm{n}}\left(\sum \frac{1}{\mathrm{m_{yi}}}\right)\right)\sigma^2_{\mathrm{yw}} \quad (5.13)$$

η οποία ταυτίζεται με την Εξίσωση (5.2) όταν $m_{xi} = m_x$ και $m_{yi} = m_y$.

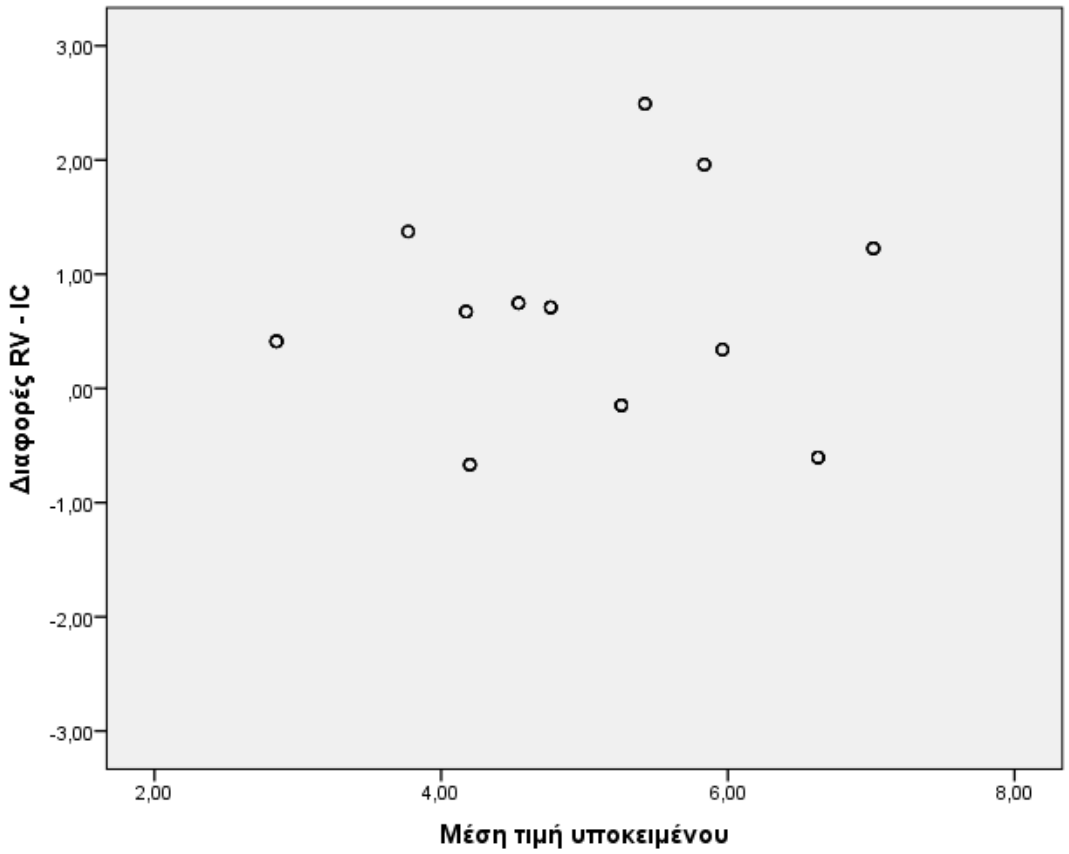
Για τα δεδομένα του Πίνακα 5.2, αρχικά θα πρέπει να εξετάσουμε την υπόθεση ανεξαρτησίας των διακυμάνσεων των μέσων τιμών των υποκειμένων. Για κάθε μέθοδο ξεχωριστά παριστάνουμε γραφικά τις εντός των υποκειμένων τυπικές αποκλίσεις σε σχέση με τις μέσες τιμές. Και όπως φαίνεται στο Γράφημα 11 που ακολουθεί, η υπόθεση της ανεξαρτησίας αυτής επιβεβαιώνεται.



Γράφημα 11 Τυπική απόκλιση και μέση τιμή των μετρήσεων κάθε υποκειμένου με τις μεθόδους RV και IC ξεχωριστά.

Στο Γράφημα 12 παριστάνουμε γραφικά για κάθε υποκείμενο τις διαφορές των μέσων τιμών των δύο μεθόδων σε σχέση με τις αναμενόμενες τιμές, και πάλι η υπόθεση ανεξαρτησίας επιβεβαιώνεται.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ



Γράφημα 12 Διαφορές των μέσων τιμών των μετρήσεων σε σχέση με τη μέση τιμή για κάθε υποκείμενο.

Στη συνέχεια υπολογίζουμε τις εκτιμήσεις των εντός των υποκειμένων διακυμάνσεις σ_{xw}^2 και σ_{yw}^2 των μεθόδων RV και IC, και όπως φαίνεται και στον Πίνακα 5.3. έχουν τιμή $s_{xw}^2 = 0.107$ *και* $s_{yw}^2 = 0.138$.

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
RV	Between Groups	99,729	11	9,066	84,551	,000
	Within Groups	5,147	48	**,107**		
	Total	104,876	59			
IC	Between Groups	91,953	11	8,359	60,631	,000
	Within Groups	6,618	48	**,138**		
	Total	98,571	59			

Πίνακας 5.3. Εκτίμηση των εντός των υποκειμένων διακυμάνσεων για τις μεθόδους RV και IC.

Στο παράδειγμά μας και για τις δύο μεθόδους έχουμε τον ίδιο αριθμό μετρήσεων και επαναλήψεων δηλαδή $m_{xi} = m_{yi}$ και κατά συνέπεια

$$\frac{1}{n}\left(\sum \frac{1}{m_{xi}}\right) = \frac{1}{n}\left(\sum \frac{1}{m_{yi}}\right) = \frac{1}{12}\left(\frac{1}{3} + \frac{3}{4} + \frac{3}{5} + \frac{5}{6}\right) = 0.2097$$

Η διακύμανση των διαφορών των μέσων των δύο μεθόδων για κάθε υποκείμενο ισούται με $Var(\bar{D}) = 0.912$ και η μέση τιμή τους $\bar{d} = 0.7092$ όπως φαίνεται στον Πίνακα 5.4. που ακολουθεί :

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation	Variance
Διαφορές μέσων RV-IC	12	-,67	2,49	**,7092**	,95475	**,912**
Valid N (listwise)	12					

Πίνακας 5.4. Μέση τιμή και διακύμανση των διαφορών των μέσων τιμών των υποκειμένων (RV – IC).

Εφαρμόζοντας τις τιμές αυτές στην Εξίσωση (5.13) υπολογίζουμε τη διακύμανση :

$$Var(RV - IC) = 0.912 + (1 - 0.2097)0.107 + (1 - 0.2097)0.138 = 1.106$$

Και η τυπική απόκλιση των διαφορών μεταξύ των μονών μετρήσεων ισούται με $\hat{\sigma}_d = \sqrt{1.106} = 1.052$. Συνεπώς το 95% όριο συμφωνίας των μεθόδων RV και IC θα είναι $(0.7092 - 1.96 \times 1.052, 0.7092 + 1.96 \times 1.052) = (-1.35, 2.77)$, από το οποίο συμπεραίνουμε ότι μια μέτρηση με τη μέθοδο RV αναμένεται να μην ξεπερνά την αντίστοιχη μέτρηση με την IC περισσότερο από 2.77 ή να μην είναι μικρότερη αυτής κατά 1.35.

5.3. Επαναλαμβανόμενες μετρήσεις σε ζεύγη

Οι μέθοδοι που παρουσιάστηκαν στις Παραγράφους 5.1. και 5.2. βασίστηκαν στην υπόθεση ότι οι πραγματικές τιμές των μετρήσεων των υποκειμένων δεν μεταβάλλονται στις επαναλαμβανόμενες μετρήσεις. Σε κάποιες περιπτώσεις όμως, ενδιαφερόμαστε για την στιγμιαία μέτρηση μιας τιμής η οποία μεταβάλλεται συνεχώς. Θα μπορούσαμε να δημιουργήσουμε ζεύγη μετρήσεων από τις δύο μεθόδους που μας ενδιαφέρουν, στο ίδιο υποκείμενο, στα οποία η πραγματική τιμή θα αλλάζει από ζευγάρι σε ζευγάρι. Για να μπορέσουμε να εκτιμήσουμε τα όρια συμφωνίας, θα εφαρμόσουμε την τεχνική των συνιστωσών διακύμανσης, χρησιμοποιώντας τις διαφορές που προκύπτουν από τα ζεύγη μετρήσεων. Το μοντέλο που θα εφαρμόσουμε για τις διαφορές για κάθε ζευγάρι j μετρήσεων, στο i υποκείμενο έχει τη μορφή[1] :

$$D_{ij} = B + I_i + E_{ij}$$

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

όπου Β είναι η σταθερά του μοντέλου, I_i η αλληλεπίδραση μεθόδου χρόνου στο υποκείμενο i, και E_{ij} τα τυχαία σφάλματα εντός του υποκειμένου i για κάθε ζεύγος μέτρησης j. Η διακύμανση των D_{ij} δίνεται από τη σχέση :

$$\sigma_d^2 = \sigma_{dI}^2 + \sigma_{dw}^2$$

Μπορούμε να εκτιμήσουμε τις σ_{dI}^2 και σ_{dw}^2 από τις συνιστώσες που θα προκύψουν από την ανάλυση διακυμάνσεων του μοντέλου[11]. Εάν υποθέσουμε ότι στο σύνολο n υποκειμένων, για κάθε υποκείμενο i διαθέτουμε m_i ζεύγη μετρήσεων,μέσω της ανάλυσης διακυμάνσεων θα προκύψει το τετραγωνικό σφάλμα εντός των υποκειμένων MS_{sw} και το τετραγωνικό σφάλμα μεταξύ των υποκειμένων MS_b. Τότε η διακύμανση μπορεί να εκτιμηθεί από τις σχέσεις $\hat{\sigma}_{dw}^2 = MS_w$ και

$$\hat{\sigma}_{dI}^2 = \frac{(\sum m_i)^2 - \sum m_i^2}{(n-1)\sum m_i}(MS_b - MS_w)$$

Και συνεπώς από το άθροισμα των εκτιμήσεων αυτών μας προκύπτει η εκτίμηση της διακύμανσης των διαφορών $\hat{\sigma}_d^2$. Η αναμενόμενη τιμή των διαφορών εκτιμάται από τη σχέση $\left(\sum m_i \bar{d}_i\right)/(\sum m_i)$, όπου $\bar{d}_i$ είναι η μέση τιμή των διαφορών του υποκειμένου i. Κατά συνέπεια μπορούμε να εκτιμήσουμε το 95% όριο συμφωνίας.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ


ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

Κεφάλαιο 6. Μη παραμετρική προσέγγιση στη σύγκριση μεθόδων

Οι μεταξύ των μεθόδων διαφορές δεν ακολουθούν πάντα κανονική κατανομή, όπως όμως αναφέρθηκε και στο Κεφάλαιο 2, κάτι τέτοιο σε γενικές γραμμές δεν θα επηρέαζε ιδιαίτερα τα όρια συμφωνίας. Ωστόσο, στην περίπτωση που παρατηρήσουμε μία ή περισσότερες ακραίες αποκλίσεις μεταξύ των δύο μεθόδων, η μη παραμετρική προσέγγιση θα ήταν ιδανική. Μια τέτοια περίπτωση προκύπτει αναμφισβήτητα από τα δεδομένα του Πίνακα 2.1, όπως παρουσιάζονται και στο Γράφημα 1, για την εκτίμηση της συστολικής αρτηριακής πίεσης με τη χρήση ημι- αυτόματων μηχανημάτων μέτρησης. Για το λόγο αυτό το πρωτόκολλο που προτείνεται από τη Βρετανική Κοινότητα Υπερτασικών (British Hypertension Society BHS), για την αξιολόγηση τέτοιων μηχανημάτων, προτείνει μια απλή μη παραμετρική μέθοδο[13].

Μπορούμε να διατηρήσουμε τη βασική προσέγγιση που παρουσιάστηκε στην Παράγραφο 2.1. η οποία περιλαμβάνει το διάγραμμα των διαφορών σε σχέση με τις μέσες τιμές των δύο μεθόδων. Υπάρχουν δύο παρόμοιοι μέθοδοι περιγραφής τέτοιων δεδομένων χωρίς την προϋπόθεση της κανονικότητας των διαφορών. Σύμφωνα με την πρώτη μέθοδο, μπορούμε να υπολογίσουμε το ποσοστό των διαφορών οι οποίες είναι μεγαλύτερες από κάποιες τιμές αναφοράς (για παράδειγμα $\pm 10 mmHg$). Οι τιμές αυτές αναφοράς, θα απεικονίζονται στο γράφημα των διαφορών σε σχέση με τις αναμενόμενες τιμές. Στη δεύτερη μέθοδο, μπορούμε να υπολογίσουμε τις τιμές, οι οποίες καθορίζουν ένα συγκεκριμένο ποσοστό (για παράδειγμα 10%), των πιο ακραίων παρατηρήσεων. Για να το εφαρμόσουμε αυτό, απλώς ταξινομούμε ανά αύξουσα σειρά τις παρατηρήσεις και αφαιρούμε από κάθε άκρο (αρχή και τέλος) το ανάλογο ποσοστό (για παράδειγμα 5% από κάθε άκρο). Τα αντίστοιχα ποσοστιαία σημεία που θα χρησιμοποιήσουμε ορίζονται και αυτά στο γράφημα. Η δεύτερη μέθοδος παρουσιάζει μεγαλύτερη αποτελεσματικότητα στη μη παραμετρική προσέγγιση για τη δημιουργία ορίων συμφωνίας. Και οι δύο όμως μη παραμετρικές μέθοδοι, είναι γενικώς λιγότερο αξιόπιστες σε σχέση με εκείνες που υποθέτουν κανονικότητα, ειδικά

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

για την περίπτωση μικρών δειγμάτων μετρήσεων. Τα διαστήματα εμπιστοσύνης μπορούν να δημιουργηθούν με τη χρήση των τυπικών μεθόδων για τα διωνυμικά ποσοστά ή το τυπικό σφάλμα των ποσοστιαίων σημείων.

Το BHS πρωτόκολλο για την αξιολόγηση μηχανημάτων μέτρησης αρτηριακής πίεσης, προτείνει τη χρήση της πρώτης μεθόδου, από αυτές που αναφέρθηκαν παραπάνω, χρησιμοποιώντας τα ποσοστά των διαφορών εντός καθορισμένων ορίων[13]. Πιο συγκεκριμένα, έχουν δημιουργηθεί τρεις αξιολογήσεις οι οποίες αναφέρονται στο ποσοστό των διαφορών που προτείνεται να υπάρχει, μεταξύ των ορίων 5,10 και 15mmHg. Στον Πίνακα 6.1 παρουσιάζονται οι συνθήκες, τις οποίες πρέπει να ικανοποιούν τα δεδομένα ώστε να αξιολογηθούν με A, B ή C αντίστοιχα. Οι συνθήκες αυτές βασίζονται στα ποσοστά των τιμών των μετρήσεων που θα επιτυγχάναμε χρησιμοποιώντας σφυγμομανόμετρο.

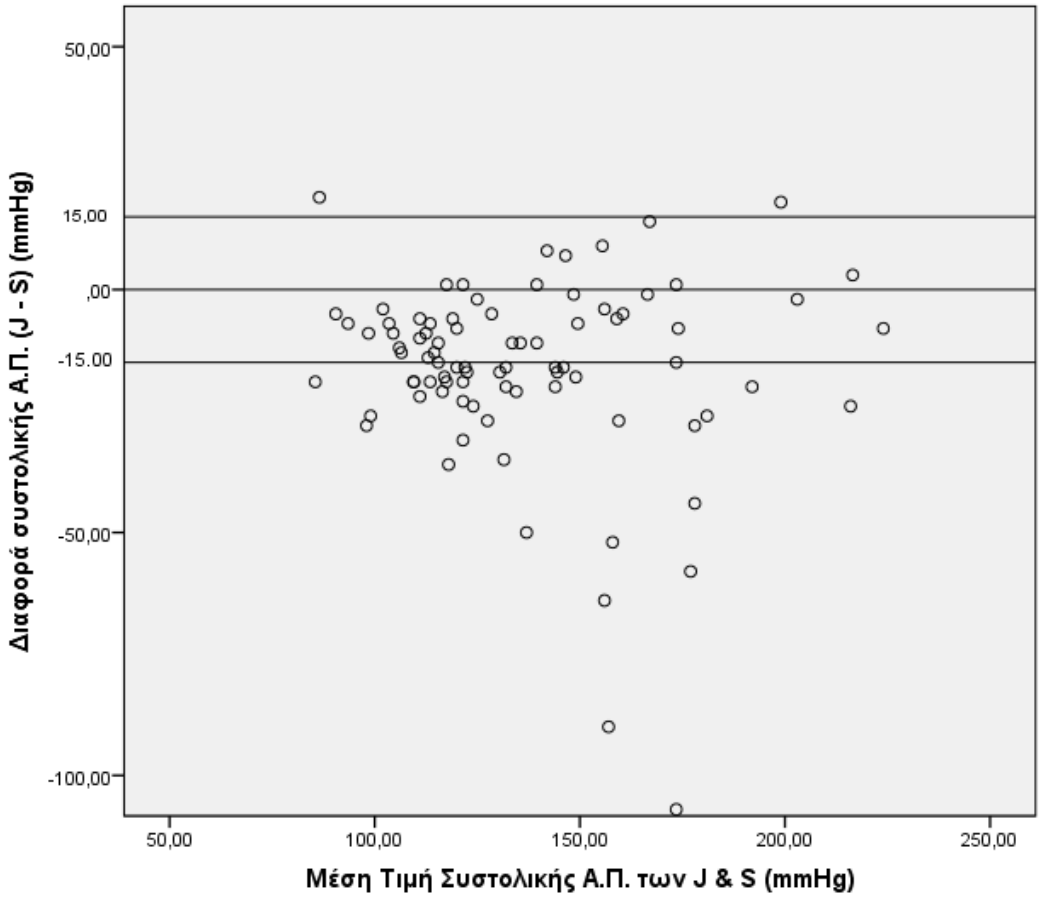
	Difference (mmHg)		
Grade	**≤ 5**	**≤ 10**	**≤ 15**
A	60	85	95
B	50	75	90
C	40	65	85
D		fails to achieve C	

Πίνακας 6.1 Αξιολόγηση των μηχανημάτων μέτρησης αρτηριακής πίεσης, βασιζόμενοι στις διαφορές μεταξύ των μετρήσεων (σε mmHg) των μηχανημάτων και σφυγμομανόμετρου[13].

Χρησιμοποιώντας τα δεδομένα του Πίνακα 2.1, παρουσιάζουμε στο Γράφημα 13 τις διαφορές των δύο μεθόδων (J-S) σε σχέση με τη μέση τιμή τους, έχοντας προσθέσει τα όρια μόνο τις τιμές ± 15 καθώς αυτές περιλαμβάνουν το μεγαλύτερο ποσοστό των διαφορών. Υπολογίζοντας τα ποσοστά των αντίστοιχων διαφορών εντός των διαστημάτων για 5, 10 και 15mmHg διαπιστώνουμε ότι είναι ίσα με 16%, 35% και 49% αντίστοιχα, συμπεραίνουμε λοιπόν ότι βάσει του Πίνακα 6.1 το εν λόγω μηχάνημα αξιολογείται με βαθμό D.

Η μη παραμετρική μέθοδος παρέχει αφοπλιστικά απλά ερμηνεύσιμα αποτελέσματα, και έχει ήδη χρησιμοποιηθεί σε παλαιότερες αναλύσεις[14,15] αλλά όχι ιδιαίτερα συχνά. Ενδεχομένως, αυτή ακριβώς η απλότητά της έχει οδηγήσει στην πεποίθηση ότι δεν είναι η κατάλληλη για την ανάλυση τέτοιων δεδομένων.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ



Γράφημα 13. Διαφορά των μετρήσεων της συστολικής αρτηριακής πίεσης, του Παρατηρητή J και του μηχανήματος S σε σχέση με τη μέση τιμή τους με την προσθήκη του διαστήματος ±15mmHg.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ


ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

Αναφορές

1. **Bland, J.M. and Altman, D.G. (1999).** Measuring agreement in method comparison studies, *Statistical Methods in Medical Research*, 8, 135-160

2. **Altman, D.G. and Bland, J.M. (1983).** Measurement in medicine: the analysis of method comparison studies, *Statistician*, 32, 307-17

3. **Bland, J.M. and Altman, D.G. (1986).** Statistical methods for assessing agreement between two methods of clinical measurement, *Lancet*, i, 307-10

4. **Bland, J.M. and Altman, D.G. (1995).** Comparing methods of measurement: why plotting difference against standard method is misleading, *Lancet*, 346, 1085-87

5. **Eksborg, S. (1981).** Evaluation of method-comparison data, *Clinical Chemistry*, 27,1311-12

6. **Linnet, K. and Bruunshuus, I. (1991).** HPLC with enzymatic detection as a candidate reference method for serum creatinine, *Clinical Chemistry*, 37, 1669-75

7. **Altman, D.G. (1993).** Calculating age-related reference centiles using absolute residuals, *Statistics in Medicine*, 12, 917-24

8. **Marshall, G.N. and Hays, R.D. and Nicholas, R. (1994).** Evaluating agreement between clinical assessment methods, *International Journal of Methods in Psychiatric Research*, 4, 249-57

9. **Lucas, A. and Hudson, G.J. and Simpson, P. and Cole, T.J. and Baker, B.A. (1987).** An automated enzymic micromethod for the measurement of fat in human milk, *Journal of Dairy Research*, 54, 487-92

10. **Bowling, L.S. and Sageman, W.S. and O'Connor, S.M. and Cole, R. and Amundson, D.E. (1993).** Lack of agreement between measurements of ejection fraction by impedance cardiography versus radionuclide ventriculography, *Critical Care Medicine*, 21, 1523-27

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

11. **Searle, S.R. and Cassela, G. and McCulloch, C.E. (1992).** *Variance components*, New York, New York

12. **Burdick, R.K. and Graybill, F.A. (1992).** *Confidence intervals of variance components*, New York, Dekker

13. **O'Brien, E. and Petrie, J. and Littler, W. and de Swiet M. and Padfield, P.L. and Altman, D.G. and Bland, M. and Coats, A. and Atkins, N. (1993).**The British Hypertension Society protocol for the evaluation of blood pressure measuring devices, *Journal of Hypertension*, 11(suppl. 2), S43-S62

14. **Polk, B.F. and Rosner, B. and Feudo, F. and Vandenburgh, M. (1980).** An evaluation of the Vita-Stat automatic blood pressure measuring device, *Hypertension*, 2, 221-27

15. **Latis, G.O. and Simionato, L. and Ferraris, G. (1981).** Clinical assessment of gestational age in the newborn infant, *Early Human Development*, 5, 29-37

16. **Altman, D.G. and Bland, J.M. (1987).** Comparing methods of measurement, *Applied Statistics*, 36, 224-25

17. **Bland, J.M. and Altman, D.G. (1991).** A note on the use of the intraclass correlation coefficient in the evaluation of agreement between two methods of measurement, *Computers in Biology and Medicine*, 20, 337-40

18. **Schoolman, H.M. and Becktel, J.M. and Best, W.R. and Johnson, A.F. (1968).** Statistics in medical research: principles versus practices, *Journal of Laboratory and Clinical Medicine*, 71, 357-67

19. **Westgard, J.O. and Hunt, M.R. (1973).** Use and interpretation of common statistical tests in method-comparison studies, *Clinical Chemistry*, 19, 49-57

20. **Feinstein, A.R. (1976).** Clinical biostatistics. XXXVII. Demeaned errors, confidence games, nonplussed minuses, inefficient coefficients, and other

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

statistical disruptions of scientific communication, *Clinical Pharmacology & Therapeutic*, 20,617-31

21. **Altman, D.G. and Bland, J.M. (1991).** The analysis of blood pressure data, In O'Brien E, O'Malley K eds. *Blood pressure measurement*, Amsterdam, Elsevier, 287-314

22. **Cotes, P.M. and Dore, C.J. and Liu Yin, J.A. and Lewis, S.M. and Messinezy, M. and Pearson, T.C. and Reid, C. (1986).** Determination of serum immunoreactive erythropoietin in the investigation of erythrocytosis, *New England Journal of Medicine*, 315, 283-87

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ