

**Πανεπιστήμιο Θεσσαλίας  
Τμήμα Ιατρικής  
Εργαστήριο Βιομαθηματικών**

***ΒΙΟΣΤΑΤΙΣΤΙΚΗ***  
***ΜΕ ΣΤΟΙΧΕΙΑ ΥΠΟΛΟΓΙΣΤΙΚΗΣ ΣΤΑΤΙΣΤΙΚΗΣ***

**Ανάλυση Δεδομένων με τη βοήθεια  
Στατιστικών Πακέτων (SPSS)**

**Εργαστηριακές Σημειώσεις**

**Ζιντζαράς Ηλίας, M.Sc., Ph.D.**  
**Επίκουρος Καθηγητής Βιομετρίας**

**Λάρισα 2008**

## ΠΕΡΙΕΧΟΜΕΝΑ

Βασιζόμενη σε ενδείξεις ιατρική, στατιστική και επιστημονική μέθοδολογία	5
Προβλήματα στην ιατρική πρακτική και έρευνα	6
Πιθανότητες (Probabilities)	7
Συνδυασμός πιθανοτήτων	8
Πιθανότητα για ένα από δύο γεγονότα	8
Πιθανότητα για δύο γεγονότα μαζί	8
Δεσμευμένη πιθανότητα	8
Βασικοί στατιστικοί όροι	10
Είδη μεταβλητών	10
Γραφικές μέθοδοι περιγραφής δεδομένων	10
Ραβδόγραμμα	10
Ιστόγραμμα	11
Καμπύλη συχνοτήτων	12
Κοιλότητα	12
Διάγραμμα πίτας ή ράβδου	13
Διάγραμμα συσχέτισης	14
Ποσοτικές μέθοδοι περιγραφής δεδομένων	15
Μέση τιμή	15
Διάμεσος (median)-εκατοστημόρια (percentiles)	15
Διάγραμμα κουτιού (Box-plot)	16
Τρόποι μέτρησης της διασποράς	17
Εύρος	18
Διακύμανση( $s^2$ ) (variance)	18
Τυπική απόκλιση(s) (standard deviation)	19
Τυπικό σφάλμα μέσης τιμής (standard error)	20
Κανονική κατανομή	21
Τυπική κανονική κατανομή SND	21
Διάστημα εμπιστοσύνης (δ.ε.) – κανονική κατανομή	23
t-κατανομή	25
Διάστημα εμπιστοσύνης (δ.ε.) – t-κατανομή	25
Έλεγχος υποθέσεων	27
t-test για ζευγαρωτές παρατηρήσεις	28
Στάθμη σημαντικότητας (P-value)	29
Διάστημα εμπιστοσύνης (δ.ε.) μέσης τιμής διαφοράς δύο δειγμάτων	30
t-test για δύο ανεξάρτητα δείγματα	31
Διάστημα εμπιστοσύνης διαφοράς δύο μέσων τιμών	32
Ανάλυση διασποράς με ένα παράγοντα (One-way ANOVA)	34
Πολλαπλές συγκρίσεις - Διόρθωση Bonferroni	37
Ανάλυση διασποράς με δύο παράγοντες (Two-way ANOVA) χωρίς αλληλεπίδραση	38
Ανάλυση διασποράς με δύο παράγοντες (Two-way ANOVA) με αλληλεπίδραση	40
Αλληλεπίδραση	41
Μετατροπή των δεδομένων	43
Έλεγχος για ένα ποσοστό	44
Διάστημα εμπιστοσύνης για ένα ποσοστό	45
Έλεγχος ποσοστού με διόρθωση	46
Έλεγχος για τη σύγκριση δύο ποσοστών	47
Διάστημα εμπιστοσύνης για τη διαφορά δύο ποσοστών	48
Πίνακες συνάφειας - $\chi^2$ test	49

Λόγος αναλογιών (Odds Ratio) σε μελέτες ασθενών-μαρτύρων	51
Διαγνωστικοί έλεγχοι	53
Έλεγχος συμφωνίας δύο διαγνωστικών μεθόδων- McNemar's-test	54
Συντελεστής συσχέτισης (correlation coefficient) r	56
Έλεγχος σημαντικότητας της r	58
Γραμμή γραμμικής παλινδρόμησης (regression line)	59
Έλεγχος γραμμικής σχέσης	61
Πρόβλεψη	62
Έλεγχος εγκυρότητας	62
Πολλαπλή γραμμική παλινδρόμηση (multiple linear regression)	64
Λογαριθμιστική παλινδρόμηση	65
Μη-παραμετρικοί έλεγχοι	66
Wilcoxon-test ζευγαρωτές παρατηρήσεις	66
Mann-Whitney-Wilcoxon test για τη σύγκριση δύο ανεξάρτητων δειγμάτων	67
Ανάλυση επιβίωσης	69
Κλινικές μελέτες	71
Φάσεις κλινικών μελετών για φάρμακα	71
Οργάνωση μιας κλινικής μελέτης – Πρωτόκολλο	72
Εισαγωγή και περιγραφή του γενικού σκοπού της μελέτης	73
Περιγραφή συγκεκριμένων αντικειμενικών σκοπών	73
Καθορισμός της θεραπείας που θα αξιολογηθεί	73
Επιλογή των ασθενών	73
Εγγραφή και κατανομή των ασθενών στις συγκρινόμενες ομάδες	74
Αξιολόγηση της ανταπόκρισης των ασθενών	74
Σχεδιασμός της μελέτης	75
Παρακολούθηση των ασθενών	75
Αναγκαίος αριθμός ασθενών σε μία κλινική μελέτη	76
Ποιοτικό θεραπευτικό αποτέλεσμα	77
Ποσοτικό θεραπευτικό αποτέλεσμα	78
Πιλοτική μελέτη	78
Σχεδιασμοί κλινικών μελετών	79
Διαστρωματικός σχεδιασμός	79
Πολυπαραγοντικός σχεδιασμός	79
Σχεδιασμός αυτοσύγκρισης	80
Διαδοχικές κλινικές μελέτες	80
Πολυκεντρικές κλινικές μελέτες	80
Στατιστική ανάλυση μιας κλινικής μελέτης	81
Ποιοτικό θεραπευτικό αποτέλεσμα	81
Ποσοτικό θεραπευτικό αποτέλεσμα	82
Ενδιάμεση ανάλυση (interim analysis) μιας κλινικής μελέτης	82
Intent to Treat (ITT) ανάλυση μιας κλινικής μελέτης	83
Μελέτες ισοδυναμίας – Κλινική σημαντικότητα vs στατιστική σημαντικότητα	83
Tables 1-5	84
Αναφορές	85

<b>Πρακτικός οδηγός χρήσης του SPSS release 10.0.1 με παραδείγματα</b>	92
Εισαγωγή στο Πρόγραμμα SPSS-Διαχείριση Αρχείων και δεδομένων	92
Περιγραφική Στατιστική	95
Αποθήκευση αρχείων	102
t-test για δύο ανεξάρτητα δείγματα	103
t-test για ζευγαρωτές παρατηρήσεις	108
Έλεγχος για τη σύγκριση δύο ποσοστών -Πίνακες συνάφειας - $\chi^2$ test	110
Mann-Whitney-Wilcoxon test για τη σύγκριση δύο ανεξάρτητων δειγμάτων	112
Ανάλυση διασποράς με ένα παράγοντα (One-way ANOVA)	115
Ανάλυση διασποράς με δύο παράγοντες (Two-way ANOVA) χωρίς αλληλεπίδραση	119
Συντελεστής συσχέτισης - Γραμμή γραμμικής παλινδρόμησης	124
Πολλαπλή γραμμική παλινδρόμηση (multiple linear regression)	128

## **Βασιζόμενη σε ενδείξεις ιατρική, στατιστική και επιστημονική μεθοδολογία**

Η στατιστική στην ιατρική πρακτική και έρευνα περιλαμβάνει:

- ◆ Τον σχεδιασμό πειραμάτων και μελετών
- ◆ Τη συλλογή δεδομένων
- ◆ Την επεξεργασία δεδομένων
- ◆ Την ανάλυση δεδομένων
- ◆ Την παρουσίαση δεδομένων και αποτελεσμάτων
- ◆ Την ερμηνεία αποτελεσμάτων

Η σύγχρονη Ιατρική χρησιμοποιεί την παρατήρηση και το πείραμα για την αξιολόγηση των διαφόρων απόψεων. Για να μπορούμε να αξιολογήσουμε και να κρίνουμε τις διάφορες θεωρίες και συμπεράσματα είναι απαραίτητο να γνωρίζουμε τη φύση των αποδεικτικών στοιχείων (data), τον τρόπο συλλογής, ανάλυσης και παρουσιάσής τους.

Για παράδειγμα αν ένας γνωστός επιστήμονας ισχυρίζεται ότι ένα φάρμακο θεραπεύει ένα συγκεκριμένο νόσημα, δεν μπορούμε να δεχτούμε την αποψη του χωρίς προβληματισμό. Θα θέλαμε να γνωρίζουμε αν έχουν γίνει μελέτες, πόσες και με ποιόν τρόπο ώστε να αποδειχθεί αυτός ο ισχυρισμός. Συνεπώς ο ρόλος της στατιστικής είναι ο πλέον σημαντικός στην ιατρική έρευνα και πράξη. Με τη γνώση της είμαστε σε θέση να:

- ◆ Διεξάγουμε επιστημονική έρευνα
- ◆ Κατανοούμε και εκτιμούμε σωστά τα πορίσματα άλλων μελετών
- ◆ Καταλαβαίνουμε την αιτιολογηση σε απόψεις άλλων επιστημόνων
- ◆ Αντιλαμβανόμαστε πιθανά μεθοδολογικά «κόλπα»
- ◆ Κρίνουμε πόση αξιοπιστία υπάρχει σε κάθε πόρισμα και να
- ◆ Μπορούμε να κρίνουμε την κριτική και τα επιχειρήματα των άλλων.

Η Στατιστική αποτελεί έναν από τους πυρήνες της σύγχρονης επιστημονικής μεθοδολογίας. Η γνώση της είναι απαραίτητη για το σχεδιασμό πειραμάτων και ερευνών, τη μεθοδολογία σχεδιασμού, ανάλυσης και παρουσίασης αυτών και την αξιοποίησή τους στην εξαγωγή συμπερασμάτων. Επίσης, είναι απαραίτητη στους γιατρούς για να είναι σε θέση να αξιολογήσουν κριτικά την ιατρική βιβλιογραφία.

## Προβλήματα στην ιατρική πρακτική και έρευνα

Τα παρακάτω πέντε ερωτήματα αποτελούν ζητήματα κλειδιά τόσο στην ιατρική έρευνα όσο και στην κλινική ιατρική που η στατιστική είναι αναγκαία για την απάντησή τους:

- |                          |                            |
|--------------------------|----------------------------|
| 1. Περι τίνος πρόκειται? | Το πρόβλημα της διάγνωσης  |
| 2. Πόσο σοβαρό είναι?    | Η αξιολόγηση της βαρύτητας |
| 3. Πόσο συχνό είναι?     | Η εκτίμηση της επίπτωσης   |
| 4. Ποιες οι αιτίες?      | Η αναζήτηση συσχετίσεων    |
| 5. Μπορεί να θεραπευθεί? | Η αξιολόγηση των θεραπειών |

### 1. Περι τίνος πρόκειται και πόσο σοβαρό είναι?

Εδώ περιλαμβάνεται η αναζήτηση του ιστορικού και η φυσική εξέταση. Τα δύο αυτά στοιχεία συμπληρώνουν οι διάφορες μετρήσεις και εργαστηριακές εξετάσεις ώστε ο κλινικός ιατρός να καταλήξει από τι πάσχει ο ασθενής και πόσο σοβαρά είναι. Για να απαντηθούν σωστά τα δύο αυτά ερωτήματα θα πρέπει να έχουν ελεγχθεί και οι διάφορες μέθοδοι και συσκευές και που χρησιμοποιούνται για τις μετρήσεις, δηλαδή η αξιοπιστία τους.

Η διάγνωση αποτελεί ένα δύσκολο πρόβλημα καθώς συνήθως παρόμοια συμπτώματα μπορεί να εμφανίζονται σε πολλές διαφορετικές παθολογικές καταστάσεις. Ο στόχος στην περίπτωση αυτή είναι να επιλεγεί εκείνη η διαγνωστική κατηγορία που είναι περισσότερο πιθανή βάσει των χαρακτηριστικών του αρρώστου. Αυτό περιλαμβάνει ιδιαίτερα τη στατιστική έννοια της πιθανότητας

### 2. Πόσο συχνό είναι?

Η απάντηση στην ερώτηση αυτή δεν είναι πάντα όσο απλή φαίνεται. Υπάρχουν δύο τρόποι μέτρησης της συχνότητας μιας ασθένειας, ο επιπολασμός (ο αριθμός των νέων περιστατικών μιας νόσου κατά τη διάρκεια μιας συγκεκριμένης χρονικής περιόδου) και η επίπτωση (ο αριθμός των ασθενών είτε κάποια συγκεκριμένη χρονική στιγμή είτε σε ένα δεδομένο χρονικό διάστημα. Εκτιμά το ρυθμό εμφάνισης μιας νόσου και για αυτό αποτελεί μέτρο του κινδύνου της νόσου).

Τόσο η επίπτωση όσο και ο επιπολασμός αποτελούν δείκτες της νοσηρότητας. Αν η ασθένεια είναι θανατηφόρος, υπολογίζουμε τον αριθμό των ατόμων που πεθαίνουν. Έτσι προκύπτουν εκτιμήσεις για τη θνητότητα.

### 3. Ποιες είναι οι αιτίες?

Η αναζήτηση των πιθανών αιτιών μιας ασθένειας είναι μια δύσκολη διαδικασία με πολλές φορές αβέβαια αποτελέσματα. Αποτελεί αντικείμενο της επιδημιολογίας. Πολλές επιδημιολογικές μελέτες αποκτούν μεγάλη δημοσιότητα καθώς αφορούν ζητήματα ευρέος ενδιαφέροντος, όπως είναι η σχέση καπνίσματος και καρκίνου των πνευμόνων, η πιθανή επίδραση της πυρηνικής ακτινοβολίας στη εμφάνιση λευχαιμίας κοκ.

### 4. Υπάρχει θεραπεία?

Στο παρελθόν τα στοιχεία για την αποτελεσματικότητα μιας θεραπείας προερχόταν από την προσωπική εμπειρία του ιατρού. Ο ενδεδειγμένος τρόπος για την αξιολόγηση διαφορετικών θεραπειών είναι με τη χρήση προσεκτικά σχεδιασμένων πειραμάτων, γνωστών ως κλινικών μελετών (controlled clinical trial).

### **Πιθανότητες (Probabilities)**

Δεν υπάρχει ολοκληρωμένη και ικανοποιητική εξήγηση της έννοια της πιθανότητας ( $P$ ) όμως στην καθημερινή μας ζωή όπως και στην ιατρική τον όρο των χρησιμοποιούμε συχνά. Για παράδειγμα, το αποτέλεσμα μίας θεραπείας με ένα αντιβιοτικό είναι ότι η λοίμωξη έχει θεραπευθεί ή δεν έχει θεραπευθεί σε διάστημα 5 ημερών. Ένας παθολόγος παρήγγειλε ότι 3 στους 4 ασθενείς έχουν θεραπευθεί. Τότε, για τον παθολόγο αυτόν η πιθανότητα ότι ένας ασθενής θα θεραπευτεί όταν του χορηγηθεί το αντιβιοτικό είναι 0.75 ή 75%. Όμως, ένας μεγαλύτερος αριθμός ασθενών θα μπορούσε να δώσει μία καλύτερη εκτίμηση της πιθανότητας θεραπείας.

Η τιμή μιας πιθανότητας κυμαίνεται από 0 έως 1. Ένα γεγονός με πιθανότητα 0 δεν μπορεί να συμβεί, ενώ ένα γεγονός με πιθανότητα 1 θα συμβεί σίγουρα. Όσο πιο μεγάλη είναι η πιθανότητα ενός γεγονότος τόσο πιο δυνατόν είναι να συμβεί αυτό. Ένα γεγονός με πιθανότητα  $1/5$  ή 0.20 σημαίνει ότι είναι δυνατόν να συμβεί 1 στις 5.

Η πιθανότητα ενός γεγονότος ορίζεται ως η αναλογία ή το ποσοστό που συμβαίνει το γεγονός αυτό σε μια μεγάλη σειρά παρατηρήσεων. Για παράδειγμα, αν μεταξύ 100000 νεογέννητων 51000 ήταν αγόρια, η πιθανότητα ένα νεογέννητο να είναι αγόρι παριστάνεται από το κλάσμα  $51000/100000 = 0.51$ , δηλ.  $P(\text{αγόρι})=0.51$ . Όμοια, η πιθανότητα να είναι κορίτσι είναι 0.49, δηλ.  $P(\text{κορίτσι})=0.49$ .

### Συνδυασμός πιθανοτήτων

Συχνά χρειάζεται να υπολογίσουμε την πιθανότητα με κάποιο συνδυασμό γεγονότων. Για παράδειγμα ποια είναι η πιθανότητα μία οικογένεια με δύο παιδιά, να είναι ένα αγόρι και ένα κορίτσι?

### Πιθανότητα για ένα από δύο γεγονότα

Η πιθανότητα να συμβεί ένα γεγονός ή ένα άλλο όταν αυτά δεν μπορούν να συμβούν ταυτόχρονα δίνεται από το άθροισμα των επιμέρους πιθανοτήτων. Για παράδειγμα, αν  $P(\text{αγόρι})=0.51$  και  $P(\text{κορίτσι})=0.49$  τότε η πιθανότητα να έχουμε ένα αγόρι ή ένα κορίτσι είναι  $P(\text{αγόρι ή κορίτσι})=0.51+0.49=1.00$ .

### Πιθανότητα για δύο γεγονότα μαζί

Η πιθανότητα να συμβούν μαζί δύο γεγονότα όταν αυτά είναι ανεξάρτητα (δηλ. το αποτέλεσμα του ενός γεγονότος δεν επηρεάζει το άλλο γεγονός) δίνεται από το γινόμενο των επιμέρους πιθανοτήτων. Για παράδειγμα, η πιθανότητα δύο άτομα σε μία οικογένεια να είναι και τα δύο αγόρια είναι  $P(\text{αγόρι και αγόρι})=0.51*0.51=0.26$ .

Αν τα δύο γεγονότα δεν είναι ανεξάρτητα τότε ο κανόνας του γινομένου δεν ισχύει.

### Δεσμευμένη πιθανότητα

Η πιθανότητα να συμβεί ένα γεγονός A με τη συνθήκη ότι ένα άλλο γεγονός B έχει ήδη συμβεί δίνεται από τον τύπο  $P(A|B)=P(A \text{ και } B)/P(B)$ .

Για παράδειγμα, σε ένα σύνολο 1.000.000 δισκίων: 50.000 έχουν κόκκους βρωμιάς (η πιθανότητα τυχαία ένα δισκίο με κόκκους είναι  $50.000/1.000.000=0.05$ , δηλ.  $P(\text{κόκοι})=0.05$ ), 30.000 έχουν χτυπημένες άκρες ( $P(\text{χτυπημένο})=0.03$ ), 40.000 είναι ξεθωριασμένα



( $P(\text{ξεθωριασμένο})=0.04$  και 20.000 έχουν κόκους και χτυπήματα μαζί ( $P(\text{κόκοι \& χτυπημένο})=0.02$ ). Τότε η πιθανότητα ένα δισκίο να είναι με κόκους δοθέντος ότι το δισκίο είναι ήδη κτυπημένο είναι:

$$P(\text{κόκοι}|\text{χτυπημένο})=P(\text{κόκοι \& χτυπημένο})/P(\text{χτυπημένο})=0.02/0.03 = 2/3.$$

Η πιθανότητα ενά δισκίο να είναι με κόκους δοθέντος ότι είναι επίσης χτυπημένο είναι  $2/3$ .

## **Βασικοί στατιστικοί όροι**

Στην Ιατρική έρευνα και κλινική πρακτική συλλέγοντας δεδομένα από ένα δείγμα ατόμων εξάγουμε συμπεράσματα για έναν πληθυσμό στον οποίο ανήκει το δείγμα.

*Παράδειγμα:* Αν θέλουμε να ερευνήσουμε τη σχέση μεταξύ της αύξησης του βάρους της εγκύου κατά τη διάρκεια της κύησης και του βάρους του νεογέννητου πρέπει να μελετήσουμε ένα δείγμα από εγκύους. Ποτέ δεν μπορούμε να μελετήσουμε όλες τις εγκύους.

Μερικοί από τους πιο βασικούς όρους στη στατιστική μεθοδολογία είναι οι παρακάτω:

Πληθυσμός: όλο το σύνολο των αντικειμένων που ερευνούμε, π.χ. όλες οι έγκυοι

Δείγμα: κάθε υποσύνολο του πληθυσμού που ερευνάται, π.χ. 30 έγκυοι έχουν συλλεχθεί τυχαία από ένα νοσοκομείο.

Μεταβλητή: κάθε χαρακτηριστικό που μπορεί να μετρηθεί ή να παρατηρηθεί, π.χ. αύξηση βάρους σώματος.

Παρατήρηση: η τιμή της μεταβλητής για ένα συγκεκριμένο αντικείμενο που μελετάται, π.χ. η αύξηση βάρους της 12<sup>ης</sup> εγκύου κατά τη διάρκεια της κύησης ήταν 7Kg.

## **Είδη μεταβλητών**

Μια μεταβλητή είναι ποιοτική όταν παίρνει ορισμένες διακριτές τιμές ή ποσοτική όταν παίρνει τιμές σε μία συνεχή κλίμακα.

*Παράδειγμα:* Ποιοτικές μεταβλητές είναι: το φύλο (άνδρας/γυναίκα), η εγκυμοσύνη (έγκυος/μη έγκυος), η επιβίωση (επιβιώνει/πεθαίνει). Ποσοτικές μεταβλητές (αριθμητικές) είναι: το ύψος, το βάρος, η πίεση αίματος, η ηλικία.

## **Γραφικές μέθοδοι περιγραφής δεδομένων**

Είναι σημαντικό να αρχίζουμε πάντα την ανάλυσή μας εξετάζοντας τα δεδομένα μας με γραφικές μεθόδους.

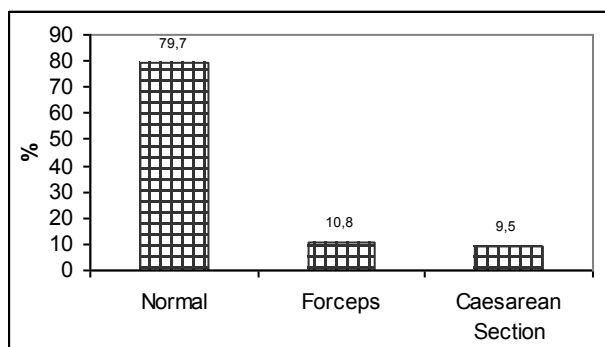
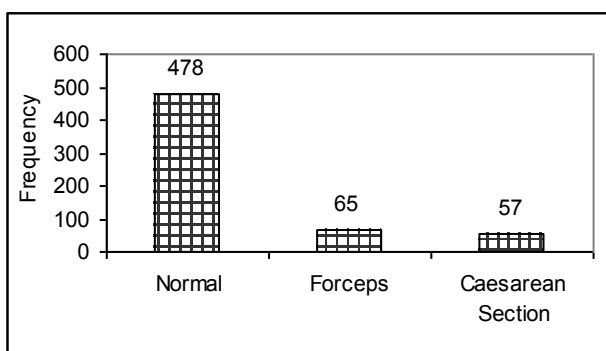
### **Ραβδόγραμμα**

Για να δείξουμε τη συχνότητα μιας ποιοτικής μεταβλητής χρησιμοποιούμε το ραβδόγραμμα.

*Παράδειγμα:* σε ένα νοσοκομείο υπήρξαν 600 γεννήσεις με διαφορετικούς τρόπους τοκετού. Η ποιοτική μεταβλητή που ερευνάται είναι η μέθοδος τοκετού:

Μέθοδος τοκετού	γεννήσεις	%
Κανονικός	478	79.7
Εμβρυουλκία	65	10.8
Καισαρική τομή	57	9.5

Οπότε στο ραβδόγραμμα το μήκος των ράβδων είναι ανάλογο των συχνοτήτων:



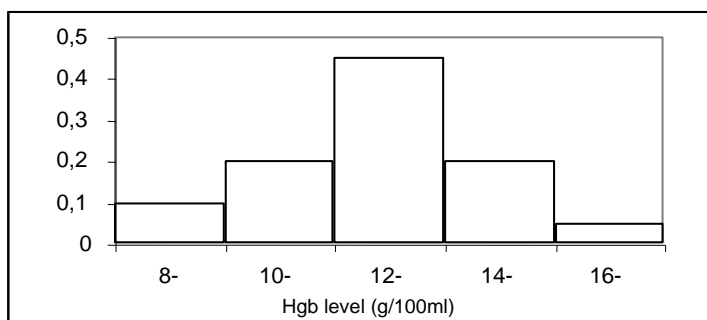
### Ιστόγραμμα

Όταν η μεταβλητή που μελετάμε είναι ποσοτική τότε κατασκευάζουμε μία κατανομή των συχνοτήτων που περιγράφεται με ένα ιστόγραμμα. Αν οι τιμές είναι πολλές τότε τις ομαδοποιούμε σε 5-8 ομάδες.

*Παράδειγμα:* Τα επίπεδα αιμοσφαιρίνης (g/100ml) 20 γυναικών έχουν μετρηθεί και είναι:

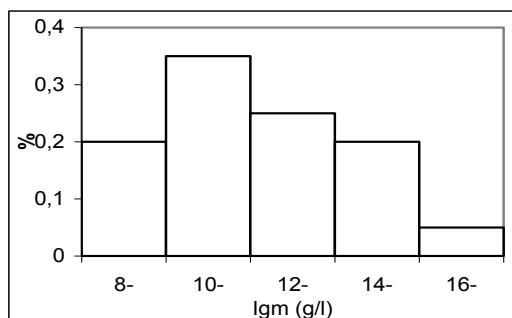
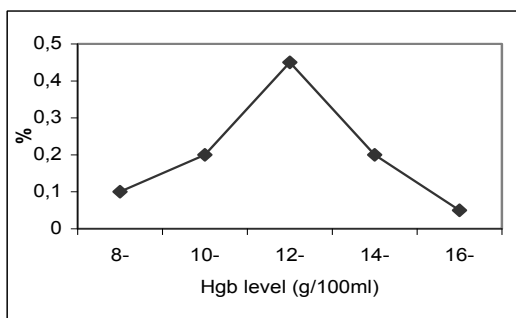
Hgb levels	
8,8	12,9
9,3	12,9
10,5	12,9
10,6	13,3
11,1	13,4
11,4	14,5
12	14,6
12	14,6
12,1	15,1
12,1	16,1

Hgb	Frequency	
8-	2	0,1
10-	4	0,2
12-	9	0,45
14-	4	0,2
16-	1	0,05
Σύνολο	20	



### Καμπύλη συχνοτήτων

Το ιστόγραμμα για πρακτικούς λόγους παριστάνεται με μία καμπύλη που κατασκευάζεται αν ενώσουμε τα διαδοχικά μέσα των πάνω βάσεων των ορθογώνιων του ιστογράμματος (καμπύλη συχνοτήτων). Όσο μεγαλώνει το μέγεθος του δείγματος ο αριθμός των ομάδων μπορεί να αυξηθεί και τότε η καμπύλη προσεγγίζεται από μία λεία καμπύλη.



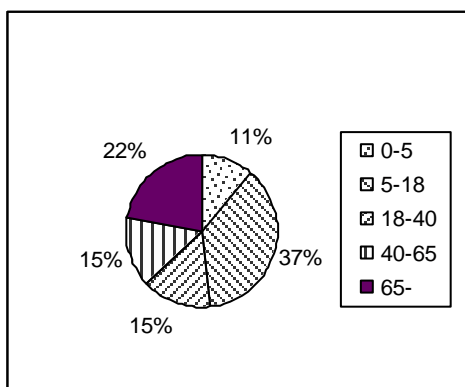
### Κοιλότητα

Το ιστόγραμμα μας υποδεικνύει την κοιλότητα ή την συμμετρία των παρατηρήσεων. Η κατανομή της ανοσοσφαιρίνης σε 20 παιδιά είναι κοίλη προς τα δεξιά ενώ η κατανομή της αιμοσφαιρίνης Hgb των 20 γυναικών είναι συμμετρική.

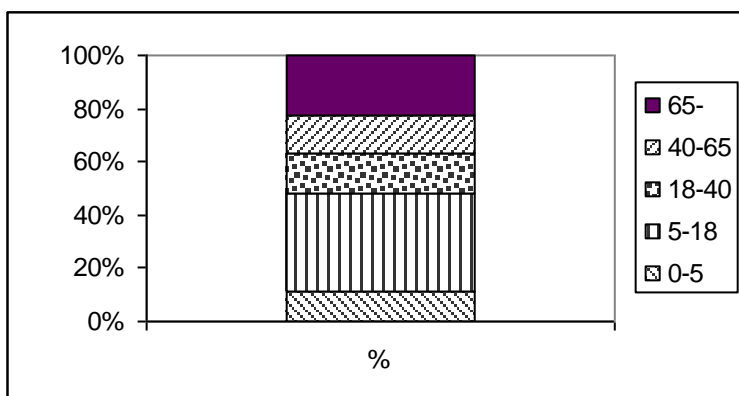
### Διάγραμμα πίτας ή ράβδου

Η πίτα είναι ένας κύκλος που είναι χωρισμένος σε κομμάτια και κάθε κομμάτι αντιπροσωπεύει τιμές μιας μεταβλητής για διάφορες κατηγορίες.

Παράδειγμα: Η παρακάτω πίτα δείχνει τη κατανομή του πληθυσμού μιας πόλης:



Το αντίστοιχο διάγραμμα σε ράβδο είναι:



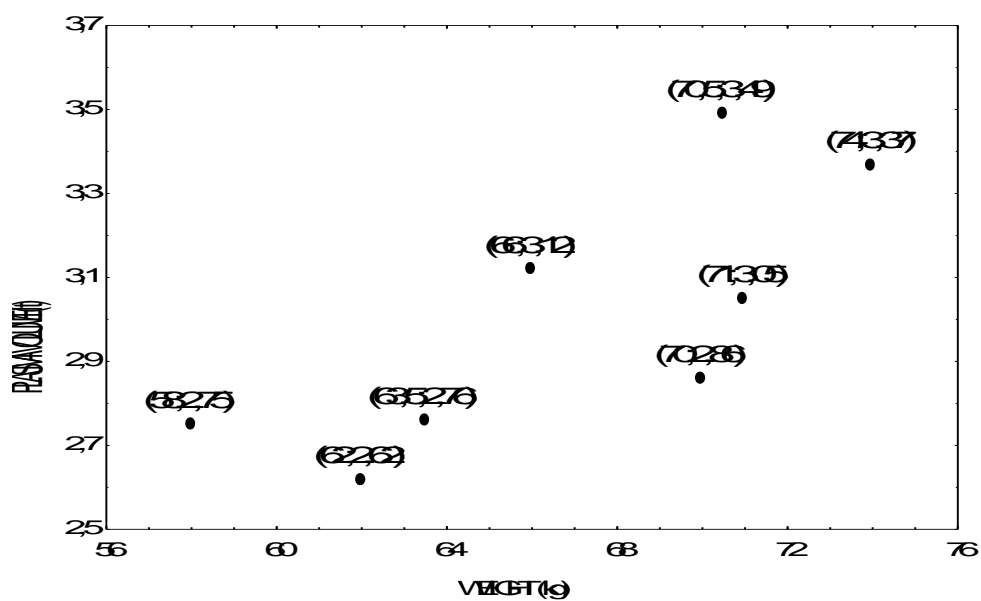
### Διάγραμμα συσχέτισης

Όταν υπάρχουν παρατηρήσεις από δύο ποσοτικές μεταβλητές και μας ενδιαφέρει η σχέση που έχουν μεταξύ τους τότε τα δεδομένα παρουσιάζονται με ένα διάγραμμα συσχέτισης.

*Παράδειγμα:* Το βάρος του σώματος και ο όγκος πλάσματος από 8 υγιείς άνδρες είναι:

Άτομο	βάρος σε Kg (x)	όγκος πλάσματος σε lt (y)
1	58.0	2.75
2	70.0	2.86
3	74.0	3.37
4	63.5	2.76
5	62.0	2.62
6	70.5	3.49
7	71.0	3.05
8	66.0	3.12

Τότε το διάγραμμα συχνοτήτων είναι:



## Ποσótικες μέθοδοι περιγραφής δεδομένων

### Μέση τιμή

Ο πιο απλός τρόπος για να περιγραφεί ένα σύνολο παρατηρήσεων από μία συνεχή μεταβλητή είναι η μέση τιμή: το άθροισμα όλων των παρατηρήσεων δια τον αριθμό των παρατηρήσεων.

*Παράδειγμα:* Οι όγκοι πλάσματος (x) από 8 υγιείς ενήλικες άνδρες είναι:

2.75, 2.86, 3.37, 2.76, 2.62, 3.49, 3.05, 3.12 lt

$x_1=2.75, x_2=2.86, x_3=3.37, x_4=2.76, x_5=2.62, x_6=3.49, x_7=3.05, x_8=3.12$

Το άθροισμα των τιμών είναι:

$$\Sigma x = x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 =$$

$$= 2.75 + 2.86 + 3.37 + 2.76 + 2.62 + 3.49 + 3.05 + 3.12 = 24.02,$$

ο αριθμός των παρατηρήσεων είναι:  $n=8$

οπότε η μέση τιμή είναι:

$$\bar{x} = \Sigma x / n = 2.75 + 2.86 + 3.37 + 2.76 + 2.62 + 3.49 + 3.05 + 3.12 / 8 = 24.02 / 8 = 3.00$$

### Διάμεσος (median)-εκατοστημόρια (percentiles)

Όταν υπάρχουν μία ή δύο υπερβολικά μικρές ή μεγάλες παρατηρήσεις, η μέση τιμή δεν αντιπροσωπεύει τα δεδομένα. Τότε οι παρατηρήσεις περιγράφονται καλύτερα από τη διάμεσο ή 50ο εκατοστημόριο.

Οι παρατηρήσεις ταξινομούνται κατά αύξουσα σειρά και η κεντρική παρατήρηση είναι η διάμεσος. Αν ο αριθμός των παρατηρήσεων είναι ζυγός τότε η διάμεσος είναι η μέση τιμή των δύο κεντρικών τιμών.

*Παράδειγμα:* Η μέγιστη εισπνευστική πίεση (PImax) 9 ασθενών με κυστική ίνωση είναι:

80, 85, 110, 95, 95, 100, 45 95, 130 (cm H<sub>2</sub>O)

Ταξινομούμε τις τιμές κατά αύξουσα σειρά::

45, 80, 85, 95, 95, 95, 100, 110, 130

Τότε η διάμεσος είναι η κεντρική παρατήρηση που είναι η 5η τιμή, δηλ.

διάμεσος = 95.

Οι όγκοι πλάσματος από 8 υγιείς ενήλικες άνδρες είναι:

2.75, 2.86, 3.37, 2.76, 2.62, 3.49, 7.32, 3.05 lt

Ταξινομούμε τις τιμές κατά αύξουσα σειρά::

2.62, 2.75, 2.76, 2.86, 3.05, 3.37, 3.49, 7.32

τότε η διάμεσος είναι η μέση τιμή της 4ης και 5ης τιμής, δηλ.

διάμεσος =  $(2.86+3.05)/2=2.96$

Άλλα εκατοστημόρια που συνήθως χρησιμοποιούνται είναι το 25ο και το 75ο που είναι οι τιμές κάτω από τις οποίες βρίσκεται το 25% και 75% των δεδομένων, αντίστοιχα.

Στο παράδειγμά μας, το 25ο εκατοστημόριο είναι ίσο με 2.76 ( $8*0.25=2$ , δηλ. είναι η τρίτη παρατήρηση) και 75ο εκατοστημόριο είναι ίσο με 3.49 ( $8*0.75=6$ , δηλ. είναι η έβδομη παρατήρηση).

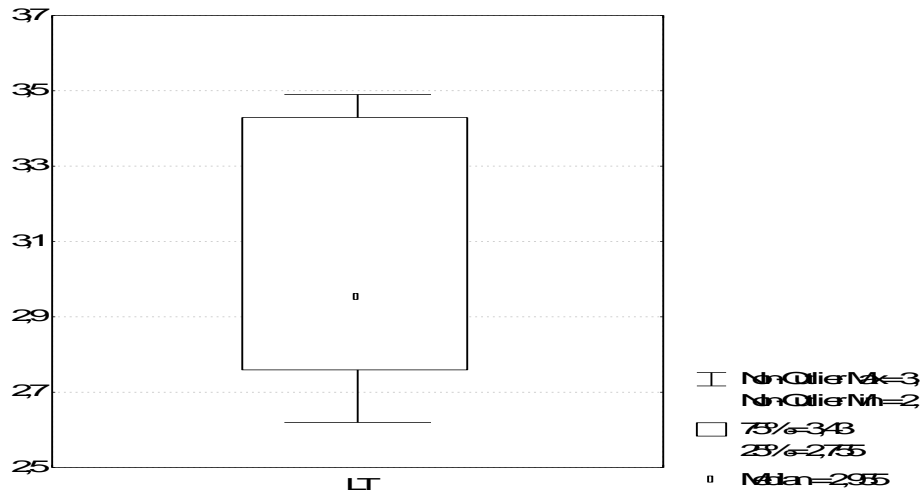
### Διάγραμμα κουτιού (Box-plot)

Το διάγραμμα κουτιού παρουσιάζει τα δεδομένα σαν ένα κουτί με ένα «μουστάκι». Το πάνω μέρος του κουτιού δείχνει το 75ο εκατοστημόριο και το κάτω μέρος του κουτιού το 25ο



εκατοστημόριο. Η διάμεσος δίνεται με μία οριζόντια γραμμή στο κουτί. Οι άκρες του μουστακιού δείχνουν την μέγιστη και την ελάχιστη τιμή (συνήθως ακραίες τιμές αγνοούνται).

*Παράδειγμα:* Με τα δεδομένα από τους όγκους πλάσματος παράγεται το παρακάτω διάγραμμα κουτιού (η τιμή 7.32 θεωρείται ακραία και αγνοείται):

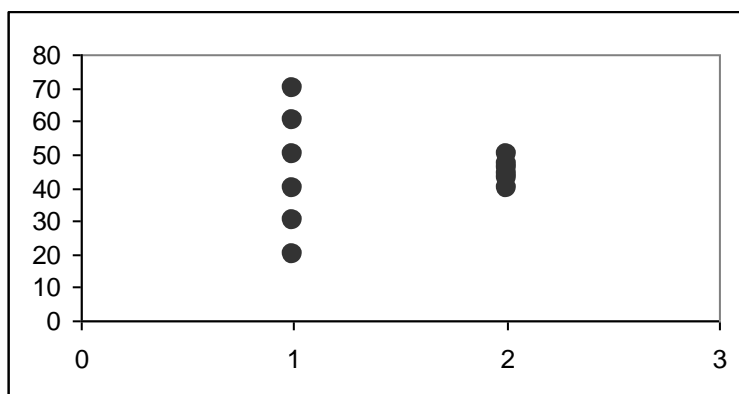


### Τρόποι μέτρησης της διασποράς

Όμως χρειαζόμαστε και ένα μέτρο της διασποράς των δεδομένων. Από μόνη της η μέση τιμή δεν μας επιτρέπει να διαφοροποιήσουμε δείγματα:

Δείγμα 1: 20, 30, 40, 50, 60, 70,  $\bar{x} = 45$  αλλά μεγάλη διασπορά

Δείγμα 2: 40, 43, 44, 46, 47, 50  $\bar{x} = 45$  αλλά μικρή διασπορά



## Εύρος

Τα εύρος είναι η διαφορά μεταξύ της μεγαλύτερης και της μικρότερης παρατήρησης. Δεν δείχνει όμως πώς κατανέμονται οι υπόλοιπες παρατηρήσεις μεταξύ αυτών των δύο.

*Παράδειγμα:* Οι όγκοι πλάσματος x από 8 υγιείς ενήλικες άνδρες είναι:

2.75, 2.86, 3.37, 2.76, 2.62, 3.49, 3.05, 3.12 lt

$$\text{Εύρος} = \max - \min = 3.49 - 2.62 = 0.87$$

## Διακύμανση( $s^2$ ) (variance)

Έτσι ο πιο κατάλληλος τρόπος να μετρηθεί η διασπορά είναι η διακύμανση, που δείχνει πώς κατανέμονται η τιμές γύρω από τη μέση τιμή: η μέση των τετραγωνικών αποστάσεων των παρατηρήσεων από τη μέση τιμή.

Το άθροισμα των τετραγωνικών αποστάσεων διαιρείται με (n-1) αντί με n γιατί έτσι εκτιμάται καλύτερα η διακύμανση του συνολικού πληθυσμού.

Το (n-1) ονομάζεται βαθμοί ελευθερίας (β.ε.).

*Παράδειγμα:* Οι όγκοι πλάσματος x από 8 υγιείς ενήλικες άνδρες είναι:

2.75, 2.86, 3.37, 2.76, 2.62, 3.49, 3.05, 3.12 lt ή

$$x_1=2.75, x_2=2.86, x_3=3.37, x_4=2.76, x_5=2.62, x_6=3.49, x_7=3.05, x_8=3.12$$

$$\text{β.ε.} = n-1=8-1=7$$

$$\text{Διακύμανση} = s^2 = \Sigma(x - \bar{x})^2 / (n-1) = [\Sigma x^2 - (\Sigma x)^2 / n] / (n-1)$$

$$\Sigma x = 2.75 + 2.86 + 3.37 + 2.76 + 2.62 + 3.49 + 3.05 + 3.12 = 24.02$$

$$\Sigma x^2 = 2.75^2 + 2.86^2 + 3.37^2 + 2.76^2 + 2.62^2 + 3.49^2 + 3.05^2 + 3.12^2 = 72.798$$

$$(\Sigma x)^2 = (24.02)^2$$

$$\text{οπότε } s^2 = [72.798 - (24.02)^2 / 8] / 7 = 0.0961$$

### Τυπική απόκλιση(s) (standard deviation)

Η διακύμανση έχει το μειονέκτημα ότι είναι το τετράγωνο των τιμών των παρατηρήσεων. Γι' αυτό εκφράζουμε τη διασπορά με τη τυπική απόκλιση που είναι η τετραγωνική ρίζα της διακύμανσης

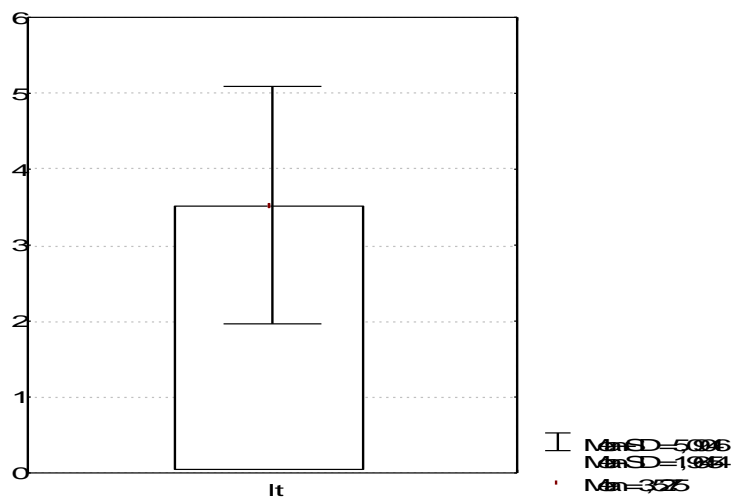
*Παράδειγμα:* Οι όγκοι πλάσματος x από 8 υγιείς ενήλικες άνδρες είναι:

2.75, 2.86, 3.37, 2.76, 2.62, 3.49, 3.05, 3.12 lt

Είναι  $s^2 = 0.0961$

Τυπική απόκλιση =  $\sqrt{\text{διακύμανση}} = s = 0.311$

Συνήθως τα δεδομένα περιγράφονται με ένα ραβδόγραμμα που δείχνει τη μέση τιμή και την τυπική απόκλιση:



### Τυπικό σφάλμα μέσης τιμής (standard error)

Όπως είπαμε στην αρχή το δείγμα μας ενδιαφέρει για να εξάγουμε συμπεράσματα για τον πληθυσμό που αντιπροσωπεύει.

Η μέση τιμή του δείγματος,  $\bar{X}$ , και η τυπική απόκλιση,  $s$ , χρησιμοποιούνται για να εκτιμήσουμε τη μέση τιμή και την τυπική απόκλιση του πληθυσμού που συμβολίζονται με  $\mu$  και  $\sigma$ , αντίστοιχα.

Η μέση τιμή ενός δείγματος είναι απίθανο να είναι ακριβώς ίδια με αυτή του πληθυσμού όπου προέρχεται.

Ένα διαφορετικό δείγμα θα έδινε διαφορετική μέση τιμή και η διαφορά οφείλεται στη δειγματοληπτική διακύμανση.

Αν συλλέξουμε πολλά ανεξάρτητα δείγματα του ίδιου μεγέθους και υπολογίσουμε τη μέση τιμή του καθενός και κατασκευάσουμε την κατανομή συχνότητας των μέσων τιμών τότε η μέση τιμή αυτής της κατανομής είναι η μέση τιμή του πληθυσμού και η τυπική της απόκλιση είναι ίση με το τυπικό σφάλμα μέσης τιμής:

$$\text{s.e.}(\mu) = \sigma/\sqrt{n}.$$

Αυτή η ποσότητα λέγεται τυπικό σφάλμα της μέσης τιμής του δείγματος και μετρά πόσο καλά η μέση τιμή του πληθυσμού εκτιμάται από τη μέση τιμή του δείγματος.

*Παράδειγμα:* Οι όγκοι πλάσματος ( $x$ ) από 8 υγιείς ενήλικες άνδρες είναι:

2.75, 2.86, 3.37, 2.76, 2.62, 3.49, 3.05, 3,12 lt

μέση τιμή:  $\bar{X} = 3.001$

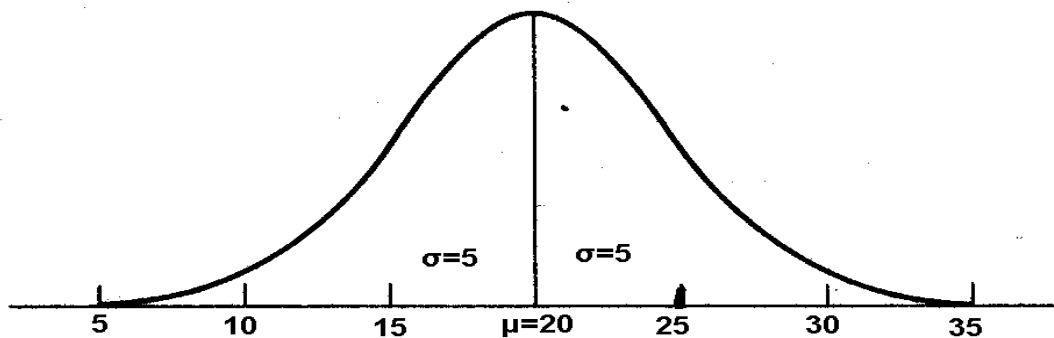
τυπική απόκλιση:  $s = 0.311$

το τυπικό σφάλμα μέσης τιμής:  $\text{s.e.}(\bar{X}) = s/\sqrt{n} = 0.311/\sqrt{8} = 0.111$

Αν το μέγεθος του δείγματος προσεγγίζει το μέγεθος του πληθυσμού τότε το s.e. τείνει στο μηδέν.

## Κανονική κατανομή

Το ιστόγραμμα περιγράφει την κατανομή μιας μεταβλητής. Μία ιδιαίτερη μορφή κατανομής είναι η κανονική κατανομή (N), η οποία έχει μορφή καμπάνας και είναι συμμετρική γύρω από τη μέση τιμή  $\mu$ .



Η κανονική κατανομή είναι σημαντική γιατί: 1) οι περισσότερες μεταβλητές στη φύση την ακολουθούν, όπως η πίεση αίματος, θερμοκρασία σώματος, ύψος, επίπεδα αιμοσφαρίνης και 2) ακόμη και αν μία μεταβλητή δεν ακολουθεί την κανονική μεταβλητή, η μέση τιμή της από ένα δείγμα θα την ακολουθεί και η τυπική απόκλιση της θα είναι  $\sigma/\sqrt{n}$ , το τυπικό σφάλμα.

## Τυπική κανονική κατανομή SND

Η κανονική κατανομή μας επιτρέπει να υπολογίσουμε το ποσοστό του πληθυσμού που οι τιμές του κυμαίνονται σε ένα συγκεκριμένο διάστημα (π.χ. ύψος από 159cm έως 184cm). Το διάστημα αυτό δίνεται από το εμβαδό της καμπύλης που περιλαμβάνεται στο διάστημα αυτό.

Όμως, είναι δύσκολο μαθηματικά να υπολογίζεται το εμβαδό αυτό για οποιαδήποτε μεταβλητή και πληθυσμο και συνεπώς υπάρχει ένας απεριόριστος αριθμός κανονικών κατανομών που εξαρτάται από το  $\mu$  και  $\sigma$ .

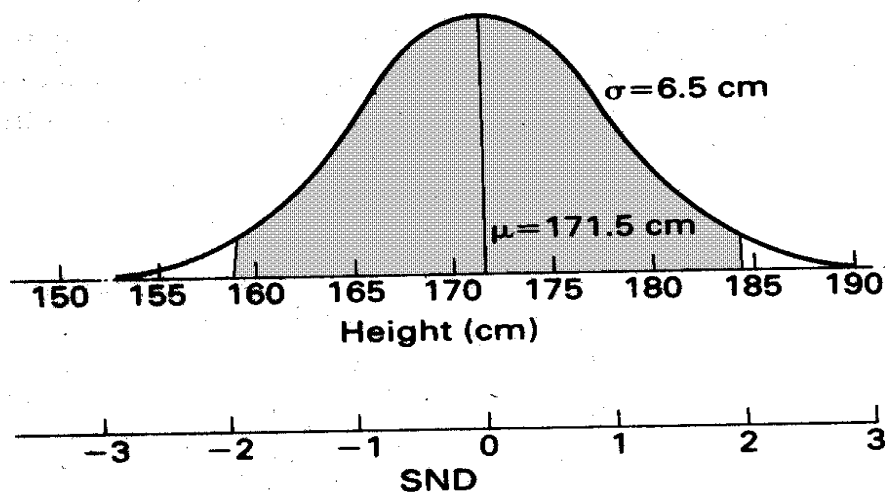
Για το λόγο αυτό η κάθε κανονική κατανομή ανάγεται στην τυπική κανονική SND κατανομή με την μετατροπή

$$z = (x - \mu) / \sigma.$$

Οπότε το εμβαδόν για οποιοδήποτε διάστημα έχει υπολογισθεί για αυτή την κατανομή και δίνεται από έναν πίνακα.

Το συνολικό εμβαδό της κατανομής αυτής είναι 100% ή 1.

*Παράδειγμα:* Η κατανομή του ύψους των ενηλίκων ανδρών στην Ελλάδα περιγράφεται από την ακόλουθη κανονική κατανομή

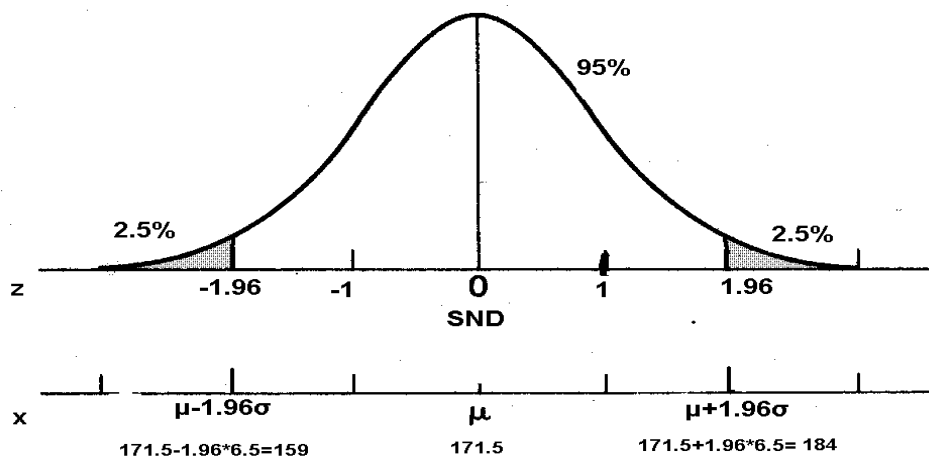


Για να υπολογίσουμε το ποσοστό των ανδρών με ύψος μεταξύ 159 και 184 μετατρέπουμε τις τιμές αυτές σε τιμές τις SND, οπότε

$$z = (159 - 171.5) / 6.5 = -1.96 \text{ και}$$

$$z = (184 - 171.5) / 6.5 = 1.96$$

Από τον πίνακα της SND είναι



Τότε το ποσοστό δίνεται από το εμβαδόν μεταξύ του διαστήματος  $-1.96$  έως  $1.96$  που είναι 95%.

Το 95% της κατανομής του πληθυσμού περιτοιχίζεται μεταξύ των τιμών  $-1.96$  και  $1.96$  της SND.

Δηλαδή, το 95% του πληθυσμού των ανδρών έχει ύψος μεταξύ  $171.5 - 1.96 * 6.5 = 159$  και  $171.5 + 1.96 * 6.5 = 184$ .

Το  $1.96$  λέγεται 5% σημείο της SND.

### **Διάστημα εμπιστοσύνης (δ.ε.) – κανονική κατανομή**

Η μέση τιμή ενός δείγματος ακολουθεί κανονική κατανομή που έχει μέση τιμή  $\mu$  και τυπική απόκλιση  $s.e.(\mu) = \sigma/\sqrt{n}$ .

Οπότε το 95% των δειγματικών μέσων τιμών βρίσκεται στο διάστημα από

$$(\mu - 1.96 * s.e.(\mu), \mu + 1.96 * s.e.(\mu)).$$

Στην πράξη όμως η  $\mu$  και το  $s.e.$  εκτιμούνται από το ίδιο το δείγμα, δηλαδή

$$\mu = \bar{X} \text{ και } s.e. = s/\sqrt{n}.$$

Οπότε, οι πιθανές τιμές της μέσης τιμής του πληθυσμού βρίσκονται στο διάστημα από

$$(\bar{X} - 1.96 * s/\sqrt{n}, \bar{X} + 1.96 * s/\sqrt{n}).$$

Το διάστημα αυτό καλείται 95% διάστημα εμπιστοσύνης για τη μέση τιμή του πληθυσμού.

*Παράδειγμα:* Οι όγκοι πλάσματος ( $x$ ) από 8 υγιείς ενήλικες άνδρες είναι:

2.75, 2.86, 3.37, 2.76, 2.62, 3.49, 3.05, 3.12 lt

Είναι:  $\bar{x} = 3.001$ ,  $s = 0.311$ ,  $s.e.(\bar{x}) = s/\sqrt{n} = 0.311/\sqrt{8} = 0.111$  και

$$(\bar{x} - 1.96 * s/\sqrt{n}, \bar{x} + 1.96 * s/\sqrt{n}),$$

Οπότε  $(3.001 - 1.96 * 0.111, 3.001 + 1.96 * 0.111) = (2.783, 3.219)$ .

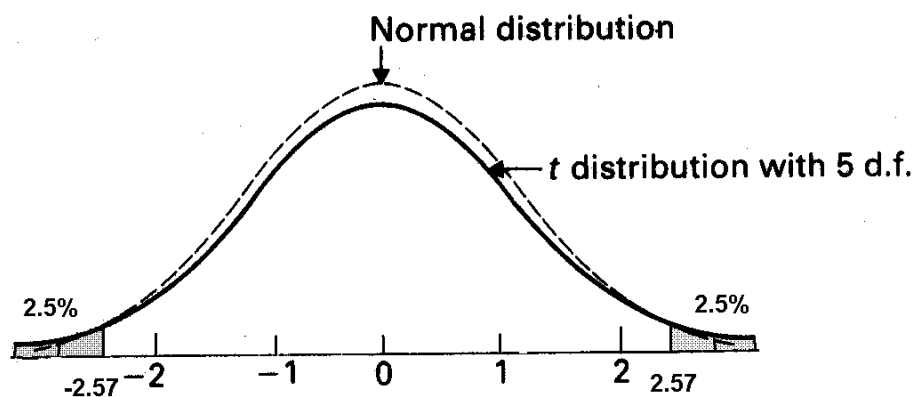


## t-κατανομή

Όταν το μέγεθος του δείγματος είναι μικρό (<60) αντί για την κανονική κατανομή χρησιμοποιούμε την t-κατανομή.

Η t-κατανομή μοιάζει με την κανονική και το σχήμα της εξαρτάται από τους β.ε. (το μέγεθος του δείγματος) για μεγάλο δείγμα έχει το ίδιο σχήμα ακριβώς με τη κανονική κατανομή.

Η τιμές της κατανομής (εμβαδόν) για διάφορα μεγέθη δείγματος δίνονται στο πίνακα 1.



## Διάστημα εμπιστοσύνης (δ.ε.) – t-κατανομή

Οπότε, για μικρό δείγμα η ποσότητα  $(\bar{x} - \mu)/(s/\sqrt{n})$  ακολουθεί όχι την SND αλλά τη t-κατανομή. Σε αυτή την περίπτωση το 95% δ.ε. για την μέση τιμή δίνεται από

**(μέση τιμή)  $\pm$  (5% σημείο t-κατανομής) \* (τυπικό σφάλμα της μέσης τιμής) ή**

$$(\bar{x} - t * \text{s.e.}(\bar{x}), \bar{x} + t * \text{s.e.}(\bar{x})) \text{ ή } (\bar{x} - t * s/\sqrt{n}, \bar{x} + t * s/\sqrt{n}),$$

όπου t είναι η τιμή της t-κατανομής για n-1 β.ε. και δίνεται στον πίνακα 1.

*Παράδειγμα:* Η διάρκεια ανακούφισης 6 ασθενών με αρθρίτιδα μετά τη χορήγηση ενός νέου φαρμάκου είναι: 2.2, 2.4, 4.9, 2.5, 3.7, 4.3 hrs

$$\bar{x}=3.3, s=1.13, \quad n=6, \quad \beta.ε.=6-1=5,$$

$$s.e.(\bar{x}) = s/\sqrt{n} = 0.46$$

από τον πίνακα 1, η τιμή της t-κατανομής για 5 β.ε. στο 5% σημείο είναι 2.57.

Τότε το 95% δ.ε. για το μέσο χρόνο ανακούφισης ασθενών με αρθριτικά είναι

$$(3.3-2.57*0.46, 3.3+2.57*0.46) \text{ ή } (2.1, 4.5\text{hrs})$$

Συνεπώς, υπάρχει 95% πιθανότητα η μέση τιμή του πραγματικού χρόνου ανακούφισης να βρίσκεται στο διάστημα αυτό.

## Έλεγχος υποθέσεων

Με τον έλεγχο υποθέσεων ελέγχουμε αν μια υπόθεση ή συνθήκη είναι αληθής χρησιμοποιώντας μία στατιστική παράμετρο η οποία ακολουθεί μία συγκεκριμένη κατανομή όταν η υπόθεση είναι αληθής. Τότε ανάλογα με την τιμή της παραμέτρου η υπόθεση απορρίπτεται με μία συγκεκριμένη πιθανότητα που προσδιορίζεται από τη στάθμη σημαντικότητας ή από το % σημείο της κατανομής.

Συνήθως ελέγχουμε τις εξής υποθέσεις:

1) Αν η μέση τιμή ενός δείγματος συμφωνεί με μία υποθετική τιμή της μέσης τιμής του πληθυσμού ( $\mu$ ), δηλ.  $\bar{X} - \mu = 0$ , ή αν οι μέσες τιμές από δύο (εξαρτημένα) δείγματα με ζευγαρωτές παρατηρήσεις είναι ίδια. Ο έλεγχος βασίζεται στην t-κατανομή και γίνεται με το t-test.

2) Αν οι μέσες τιμές δύο ανεξάρτητων δειγμάτων είναι ίδιες, δηλ.  $\bar{X}_1 - \bar{X}_2 = 0$ . Ο έλεγχος βασίζεται στην t-κατανομή και γίνεται με το t-test.

3) Αν ένα ποσοστό από ένα δείγμα συμφωνεί με ένα υποθετικό ποσοστό ενός πληθυσμού ( $\pi$ ), δηλ.  $p - \pi = 0$ . Ο έλεγχος βασίζεται στην κανονική-κατανομή και γίνεται με το z-test.

4) Αν δύο ποσοστά από δύο ανεξάρτητα δείγματα είναι ίδια, δηλ.  $p_1 - p_2 = 0$ . Ο έλεγχος βασίζεται στην κανονική-κατανομή και γίνεται με το z-test ή στην  $\chi^2$ -κατανομή και γίνεται με το  $\chi^2$ -test.

5) αν οι μέσες τιμές περισσότερων από δύο ανεξάρτητων δειγμάτων (π.χ. κ δείγματα) είναι ίδιες, δηλ.  $\bar{X}_1 = \bar{X}_2 = \bar{X}_3 = \dots = \bar{X}_k$ . Ο έλεγχος βασίζεται στην F-κατανομή και γίνεται με ανάλυση διασποράς (analysis of variance-ANOVA) και το F-test.

### t-test για ζευγαρωτές παρατηρήσεις

Όταν οι παρατηρήσεις από ένα δείγμα είναι ζευγάρια με τις παρατηρήσεις ενός άλλου δείγματος τότε για να ελέγξουμε αν υπάρχει διαφορά μεταξύ των δύο δειγμάτων χρησιμοποιείται το t-test για ζευγαρωτές παρατηρήσεις.

*Παράδειγμα:* Σε μία κλινική μελέτη για να ελέγξουμε την αποτελεσματικότητα ενός υπνωτικού φαρμάκου, παρατηρήθηκε η διάρκεια του ύπνου (hrs) σε 10 ασθενείς σε ένα βράδυ μετά τη χορήγηση του φαρμάκου και σε ένα άλλο βράδυ μετά τη χορήγηση εικονικού φαρμάκου (placebo). Τα ποτελέσματα ήταν:

Ασθενής	φάρμακο	placebo	διαφορά
1	6.1	5.2	0.9
2	7.0	7.9	-0.9
3	8.2	3.9	4.3
4	7.6	4.7	2.9
5	6.5	5.3	1.2
6	8.4	5.4	3.0
7	6.9	4.2	2.7
8	6.7	6.1	0.6
9	7.4	3.8	3.6
10	5.8	6.3	-0.5
Μέση τιμή	7.06	5.28	$\bar{x} = 1.78$

Τότε για να ελέγξουμε την υπόθεση ότι το φάρμακο δεν είναι αποτελεσματικό, δηλ  $\bar{x} - \mu = 0$  όπου  $\mu = 0$ , χρησιμοποιούμε το t-test:

**$t = (\text{μέση τιμή των διαφορών των δύο δειγμάτων}) / (\text{τυπικό σφάλμα της μέσης τιμής})$  ή**

$$t = (\bar{x} - \mu) / \text{s.e.}(\bar{x}), \text{ όπου } \text{s.e.}(\bar{x}) = s/\sqrt{n}.$$

Τότε  $\bar{x} = 1.78$ ,  $\mu = 0$ ,  $s = 1.77$ ,  $n = 10$  και  $\text{s.e.}(\bar{x}) = s/\sqrt{n} = 0.56$ .

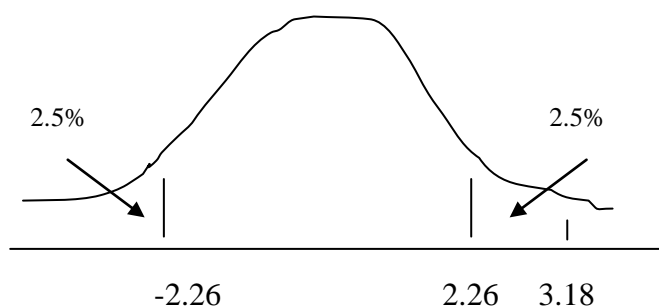
Οπότε  $t = 1.78 / 0.56 = 3.18$  (αν το πρόσημο είναι αρνητικό τότε αγνοείται).

Αν η διαφορά μεταξύ των φαρμάκων είναι πράγματι 0 τότε η ποσότητα  $t = (\bar{x} - \mu) / (s/\sqrt{n})$  είναι μια τυπική τιμή της t-κατανομής με  $(n-1)$  β.ε.

### Στάθμη σημαντικότητας (P-value)

Από τον πίνακα 1 μπορούμε να βρούμε διάφορες στάθμες σημαντικότητας ή τα % σημεία της t-κατανομής.

Η τιμή της t-κατανομής για  $10-1=9$  β.ε. με στάθμη σημαντικότητας 5% είναι 2.26. Αυτό σημαίνει ότι κατά μέσο όρο 5% των παρατηρήσεων από μια τέτοια κατανομή θα είναι μεγαλύτερο από 2.26 ή μικρότερο από -2.26 (αν το φάρμακο δεν είναι αποτελεσματικό-τα δύο δείγματα δεν διαφέρουν)



Συνεπώς, η τιμή  $t=3.18$  είναι μεγαλύτερη από 2.26 και συνεπώς είναι μια ακραία τιμή της t-κατανομής. Οπότε η πιθανότητα να παρατηρήσουμε μία τέτοια τιμή  $t=3.18$  κατά τύχη είναι λιγότερη από 5%, δηλαδή  $P<0.05$ . Δηλαδή η διαφορά που βρέθηκε ( $\bar{X}=1.78$ ) δεν είναι τυχαία, είναι μια πραγματική, οπότε τα φάρμακα διαφέρουν.

Επειδή η πιθανότητα είναι μικρή απορρίπτουμε την υπόθεση ότι τα φάρμακα δεν διαφέρουν και συμπεραίνουμε ότι υπάρχει ένδειξη ( $P<0.05$ ) ότι το υπνωτικό φάρμακο είναι αποτελεσματικό σε σχέση με το εικονικό φάρμακο.

Είναι το φάρμακο αποτελεσματικό σε στάθμη σημαντικότητας 1% ; Από τον πίνακα 1, η τιμή της t-κατανομής με στάθμη σημαντικότητας 1% είναι 3.25. Οπότε  $t=3.18$  είναι μικρότερο από 3.25 συνεπώς δεν υπάρχει ισχυρή ένδειξη ότι το φάρμακο είναι διαφορετικό, απλώς ένδειξη. Συνεπώς το φάρμακο δεν είναι αποτελεσματικό ( $P>0.01$ ).

### Πόσο ισχυρή είναι η ένδειξη διαφοράς?

Πιθανότητα (P)

ισχύς της ένδειξης ενάντια στην υπόθεση:  $\bar{x} - \mu = 0$

να παρατηρήσουμε το t

---

Μεγαλύτερη από 5% ( $P > 0.05$ )	όχι ένδειξη (δεν απορρίπτουμε την υπόθεση)
Μικρότερη από 5% ( $P \leq 0.05$ )	ένδειξη (απορρίπτουμε την υπόθεση)
Μικρότερη από 1% ( $P \leq 0.01$ )	ισχυρή ένδειξη (απορρίπτουμε την υπόθεση)

### Διάστημα εμπιστοσύνης (δ.ε.) μέσης τιμής διαφοράς δύο δειγμάτων

Το 95% δ.ε. για την μέση τιμή  $\bar{x}$  των διαφορών δύο δειγμάτων με ζευγαρωτές παρατηρήσεις είναι ακριβώς το ίδιο με το δ.ε. για την t-κατανομή που εξετάσαμε πιο πάνω και είναι

**(μέση τιμή)  $\pm$  (5% σημείο t-κατανομής)\*(τυπικό σφάλμα της μέσης τιμής) ή**

$(\bar{x} - t^* \text{s.e.}(\bar{x}), \bar{x} + t^* \text{s.e.}(\bar{x}))$ , όπου  $\text{s.e.}(\bar{x}) = s/\sqrt{n}$ .

t είναι η τιμή της t-κατανομής για n-1 β.ε. με 5% και δίνεται στον πίνακα 1.

*Παράδειγμα:* Στο παράδειγμά μας είναι:  $\bar{x} = 1.78$ ,  $s = 1.77$ ,  $n = 10$ , β.ε. =  $10 - 1 = 9$  και  $\text{s.e.}(\bar{x}) = s/\sqrt{n} = 0.56$ .

Από τον πίνακα 1, η τιμή της t-κατανομής για 9 β.ε. στο 5% σημείο είναι 2.26.

Τότε το 95% δ.ε. για τη μέση τιμή της διαφοράς των δύο φαρμάκων είναι

$(1.78 - 2.26 * 0.56, 1.78 + 2.26 * 0.56)$  ή  $(0.51, 3.05 \text{hrs})$ .

Συνεπώς, υπάρχει 95% πιθανότητα η μέση τιμή της διαφοράς των δύο φαρμάκων να βρίσκεται στο διάστημα αυτό. Επειδή το 0 δεν συμπεριλαμβάνεται στο 95% δ.ε. συμπεραίνουμε ότι το πραγματικό φάρμακο είναι αποτελεσματικό και βελτιώνει την διάρκεια του ύπνου από μισή έως 3 ώρες.

### t-test για δύο ανεξάρτητα δείγματα

Όταν οι παρατηρήσεις από δύο δείγματα είναι ανεξάρτητες (δηλ. από διαφορετικά άτομα) τότε μπορούμε να συγκρίνουμε τις μέσες τιμές των δύο δειγμάτων χρησιμοποιώντας το t-test για ανεξάρτητα δείγματα.

*Παράδειγμα:* Το βάρος των νεογέννητων παιδιών που γεννήθηκαν από 15 μη-καπνίστριες και από 14 καπνίστριες είναι:

	Δείγμα 1 Μη-καπνίστριες	Δείγμα 2 Καπνίστριες
	-----	
	3.99	3.18
	3.79	2.84
	3.60	2.90
	3.73	3.27
	3.21	3.85
	3.60	3.52
	4.08	3.23
	3.61	2.76
	3.83	3.60
	3.31	3.75
	4.13	3.59
	3.26	3.63
	3.54	2.38
	3.51	2.34
	2.71	
$\bar{x}$	3.593	3.203
s	0.3707	0.4927
n	15	14

Υπάρχει διαφορά στο βάρος των παιδιών μεταξύ των δύο ομάδων;

Υποθέτουμε ότι και οι δύο ομάδες προέρχονται από έναν πληθυσμό που έχει ίδια τυπική απόκλιση  $\sigma$ . Τότε  $s_1$  και  $s_2$  χρησιμοποιούνται για να εκτιμήσουν την  $\sigma$  με την κοινή τυπική απόκλιση  $s_p$ :

$$s_p = \sqrt{[(n_1-1)s_1^2 + (n_2-1)s_2^2] / (n_1+n_2-2)} \text{ ή } s_p = \sqrt{[(14 \cdot 0.3707^2 + 13 \cdot 0.4927^2) / (15+14-2)]} = 0.4337$$

Ελέγχουμε την υπόθεση ότι οι δύο μέσες τιμές είναι ίδιες,  $\bar{X}_1 - \bar{X}_2 = 0$ , χρησιμοποιώντας την ποσότητα:

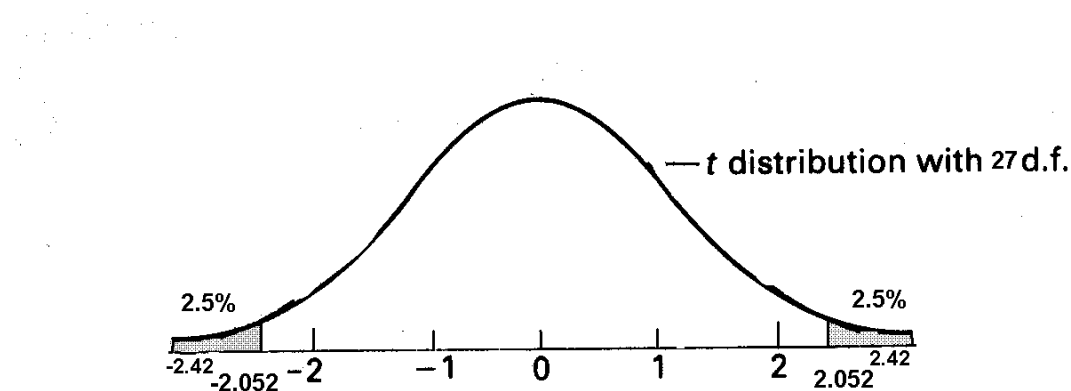
$t = (\text{διαφορά των δύο μέσων τιμών}) / (\text{τυπικό σφάλμα της διαφοράς})$  ή

$t = (\bar{X}_1 - \bar{X}_2) / \text{s.e.}(\bar{X}_1 - \bar{X}_2)$ , όπου  $\text{s.e.}(\bar{X}_1 - \bar{X}_2) = s_p \cdot \sqrt{(1/n_1 + 1/n_2)}$ .

Τότε  $t = (3.593 - 3.203) / [0.4337 \cdot \sqrt{(1/15 + 1/14)}] = 2.42$

Αν οι δύο μέσες τιμές είναι πραγματικά ίδιες, τότε η ποσότητα  $t$  ακολουθεί τη  $t$ -κατανομή με  $(n_1 - 1) + (n_2 - 1) = (15 - 1) + (14 - 1) = 27$  β.ε.

Από τον πίνακα 1, η τιμή της  $t$ -κατανομής με 27 βε με στάθμη σημαντικότητας 5% είναι 2.052. Επειδή η τιμή  $t = 2.42$  είναι μεγαλύτερη από 2.052 απορρίπτουμε την υπόθεση ότι οι μέσες τιμές είναι ίδιες. Οπότε υπάρχει ένδειξη ( $P < 0.05$ ) ότι το βάρος των παιδιών από μη-καπνίστριες διαφέρει από αυτό των καπνιστριών, δηλ. οι καπνίστριες γεννούν μωρά με μικρότερο βάρος.



### Διάστημα εμπιστοσύνης διαφοράς δύο μέσων τιμών

Το 95% δ.ε. της διαφοράς των δύο μέσων τιμών είναι:

$(\text{διαφορά των δύο μέσων τιμών}) \pm (5\% \text{ σημείο } t\text{-κατανομής}) \cdot (\text{τυπ. σφάλμα διαφοράς})$

$((\bar{X}_1 - \bar{X}_2) - t \cdot \text{s.e.}(\bar{X}_1 - \bar{X}_2), (\bar{X}_1 - \bar{X}_2) + t \cdot \text{s.e.}(\bar{X}_1 - \bar{X}_2))$ ,



όπου  $s.e. (\bar{X}_1 - \bar{X}_2) = s_p \sqrt{(1/n_1 + 1/n_2)}$

t είναι η τιμή της t-κατανομής για  $(n_1-1)+(n_2-1)$  β.ε. σε 5% στάθμη σημαντικότητας και δίνεται από τον πίνακα 1.

*Παράδειγμα:* Στο παράδειγμά μας το 5% σημείο της t-κατανομής με  $(15-1)+(14-1)=27$  β.ε. είναι 2.052 (πίνακας 1).

Τότε το 95% δ.ε. είναι

$(0.39 - 2.052 * [0.4337 * \sqrt{(1/15 + 1/14)}], 0.39 + 2.052 * [(0.4337 * \sqrt{(1/15 + 1/14)})])$  ή

(0.0593, 0.721)

Συνεπώς, με πιθανότητα 95% μπορούμε να ισχυριστούμε ότι οι μη-καπνίστριες γεννούν παιδιά με περισσότερο βάρος από 0.0593 έως 0.721 Kg από ότι οι καπνίστριες

Αν το δ.ε. συμπεριλάμβανε την τιμή 0 τότε θα ισχυριζόμασταν ότι μεταξύ των δύο ομάδων δεν υπάρχει διαφορά.

### Ανάλυση διασποράς με ένα παράγοντα (One-way ANOVA)

Η ανάλυση διασποράς με ένα παράγοντα χρησιμοποιείται όταν θέλουμε να συγκρίνουμε τις μέσες τιμές πολλών ομάδων (περισσότερες από δύο). Η ανάλυση καλείται με ένα παράγοντα επειδή τα δεδομένα ταξινομούνται σύμφωνα με ένα παράγοντα ή ομάδα.

*Παράδειγμα:* Σε μία μελέτη, έχει καταγραφεί το βάρος του ήπατος (εκφρασμένο ως ποσοστό του βάρους του σώματος) ποντικών που ανήκουν σε  $k=4$  ομάδες που τράφηκαν με 4 δίαιτες. Θέλουμε να ερευνήσουμε αν υπάρχουν συστηματικές διαφορές μεταξύ των 4 ομάδων, δηλ. να θέλουμε να συγκρίνουμε τις μέσες τιμές των 4 ομάδων:  $\mu_a, \mu_b, \mu_c, \mu_d$ .

	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>
	3.42	3.17	3.34	3.64
	3.96	3.63	3.72	3.93
	3.87	3.38	3.81	3.77
	4.19	3.47	3.66	4.18
	3.58	3.39	3.55	4.21
	3.76	3.41	3.51	3.88
$\bar{x}$	3.80	3.41	3.60	3.94
Sum T	22.78	20.45	21.59	23.61
$s^2$	0.252	0.142	0.182	0.192

Η μέθοδος βασίζεται στο να προσδιορίσουμε πόσο από την συνολική διακύμανση των δεδομένων οφείλεται στις διαφορές μεταξύ των μέσων τιμών των ομάδων και να συγκρίνουμε τη ποσότητα αυτή με το ποσό της συνολικής διακύμανσης που οφείλεται στις διαφορές μεταξύ των ατόμων στην ίδια ομάδα.

Αν η συνολική διακύμανση είναι:

$$s^2 = (\text{άθροισμα τετραγώνων SS}) / (d.f.) \text{ ή } s^2 = \sum (x - \bar{x})^2 / (n-1),$$

τότε η ANOVA διαιρεί το παραπάνω άθροισμα τετραγώνων SS σε δύο μέρη:

- 1) Το SS που οφείλεται στις διαφορές μεταξύ των μέσων τιμών ομάδων,
- 2) Το SS που οφείλεται στις διαφορές μεταξύ των παρατηρήσεων μέσα στη κάθε ομάδα, δηλ. το Υπόλοιπο SS (Residual SS ή Within groups SS).

Με όμοιο τρόπο διαιρείται το σύνολο  $(n-1)$  των d.f.

Η υπόθεση που ελέγχεται εδώ είναι ότι δεν υπάρχουν διαφορές μεταξύ των ομάδων, δηλ.  
 $\mu_a = \mu_b = \mu_c = \mu_d$ .

Για τον προσδιορισμό της ANOVA χρειάζονται οι παρακάτω μαθηματικοί υπολογισμοί:

$$n=24, \Sigma x = 3.42 + 3.96 + \dots + 4.21 + 3.88 = 88.43, \Sigma x^2 = 3.42^2 + 3.96^2 + \dots + 4.21^2 + 3.88^2 = 327.66$$

$$1) \text{Total SS} = \Sigma (x - \bar{x})^2 = \Sigma x^2 - (\Sigma x)^2/n = 327.66 - 88.43^2/24 = 1.83$$

$$\text{Total d.f.} = n - 1 = 24 - 1 = 23$$

$$2) \text{Between groups SS} = T_a^2/n_a + T_b^2/n_b + T_c^2/n_c + T_d^2/n_d - (\Sigma x)^2/n = 22.78^2/6 + 20.45^2/6 + 21.59^2/6 + 23.61^2/6 - 88.43^2/24 = 0.954$$

$$\text{Between groups d.f.} = 4 - 1 = 3$$

$$3) \text{Within groups (residual ή error) SS} = \text{Total SS} - \text{Between groups SS} = 1.83 - 0.954 = 0.876$$

$$\text{Within groups (residual) d.f.} = \text{Total d.f.} - \text{Between groups d.f.} = 23 - 3 = 20$$

Μετά υπολογίζεται, το ποσό της διακύμανσης ανά βαθμό ελευθερίας (d.f.) που είναι η μέση μεταβολή (Mean Square, MS), δηλ.  $MS = SS/df$ .

Ο έλεγχος για τις διαφορές μεταξύ των ομάδων βασίζεται στη σύγκριση μεταξύ του Between groups MS με το Within groups MS. Τα MS συγκρίνονται χρησιμοποιώντας το F-test:

$$F = (\text{Between groups MS}) / (\text{Within groups MS})$$

Η ANOVA παρουσιάζεται με τον εξής πίνακα:

Source of variation	df	SS	MS=SS/df	F
Between groups	k-1=4-1=3	0.954	0.318	7.23
Within groups (residual)	23-3=20	0.876	0.044=s <sup>2</sup>	
Total	n-1=24-1=23	1.83		

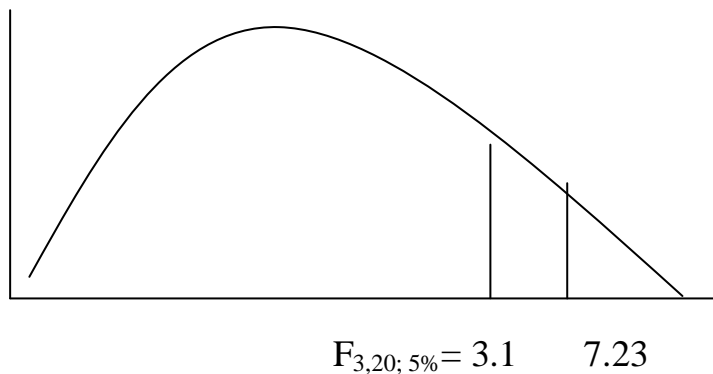
Το Residual MS δίνει μία εκτίμηση της s<sup>2</sup> της κοινής εντός των ομάδων διακύμανσης.

Ελέγχουμε την υπόθεση ότι δεν υπάρχουν διαφορές μεταξύ των ομάδων, δηλ.  $\mu_a = \mu_b = \mu_c = \mu_d$ , χρησιμοποιώντας την παραπάνω τιμή της F.

Αν η υπόθεση είναι αληθινή τότε η F ακολουθεί την F-κατανομή με 3 και 20 df, δηλ. οι Between groups df και οι Within groups df. Τότε το F προσεγγίζει την τιμή 1.

Αν υπάρχουν μεγάλες διαφορές μεταξύ των ομάδων τότε η τιμή της F γίνεται πολύ μεγαλύτερη του 1.

Οι τιμές της F-καταν. με στάθμη σημαντικότητας 5% δίνονται στον πίνακα 3. Στο παράδειγμά ,  $F=7.23$  και η τιμή της F-κατανομής με 3 και 20 df είναι 3.1.



Η τιμή της F είναι 7.23 που είναι μεγαλύτερη από τη τιμή της F-κατανομής με 3 και 20 df σε στάθμη σημαντικότητας 5%  $F_{3,20; 5\%} = 3.1$  ( $P < 0.05$ ), δηλ. απορρίπτουμε τη αρχική υπόθεση. Συνεπώς υπάρχει ένδειξη ( $P < 0.05$ ) για διαφορές μεταξύ των ομάδων.

Αν η ANOVA δείξει ότι υπάρχουν διαφορές τότε μπορούμε να κάνουμε επιμέρους συγκρίσεις μεταξύ των ομάδων, π.χ. να συγκρίνουμε την ομάδα a με την ομάδα b, χρησιμοποιώντας το t-test ή βρίσκοντας το 95% δ.ε. για τη διαφορά των μέσων τιμών:

είναι  $(\bar{X}_1 - \bar{X}_2) = 0.39$  και  $s.e. (\bar{X}_1 - \bar{X}_2) = \sqrt{\{s^2 * (1/n_a + 1/n_b)\}} = \sqrt{\{0.044 * (1/6 + 1/6)\}} = 0.12$

Η τιμή της t-κατανομής με 20 df σε 5% στάθμη σημαντικότητας είναι 2.09.

Το 95% δ.ε. είναι:  $((\bar{X}_1 - \bar{X}_2) - 2.09 * s.e. (\bar{X}_1 - \bar{X}_2), (\bar{X}_1 - \bar{X}_2) + 2.09 * s.e. (\bar{X}_1 - \bar{X}_2))$  ή  $(0.39 - 2.09 * 0.12, 0.39 + 2.09 * 0.12) = (0.14, 0.64)$ .

Με 95% βεβαιότητα η ομάδα a έχει υψηλότερο βάρος μεταξύ 0.14 και 0.64 από ότι η ομάδα b.

Το t-test είναι:  $t = (\bar{X}_a - \bar{X}_b) / s.e.(\bar{X}_1 - \bar{X}_2) = 0.39 / 0.12 = 3.25$

τιμή που είναι μεγαλύτερη από την αντίστοιχη τιμή της t-κατανομής (2.09), δηλ. υπάρχει διαφορά μεταξύ των ομάδων a και b ( $P < 0.05$ ).

Όταν οι ομάδες έχουν διαφορετικό αριθμό ατόμων ο τρόπος προσδιορισμού της ANOVA είναι λίγο διαφορετικός.

### **Πολλαπλές συγκρίσεις - Διόρθωση Bonferroni**

Όλες οι πιθανές συγκρίσεις μεταξύ των ομάδων δεν είναι ανεξάρτητες. Για το λόγο αυτό όταν γίνονται πολλαπλές συγκρίσεις (k) η στάθμη σημαντικότητας (P) πρέπει να διορθώνεται σε  $P' = kP$ .

### Ανάλυση διασποράς με δύο παράγοντες (Two-way ANOVA) χωρίς αλληλεπίδραση

Όταν τα δεδομένα ταξινομούνται σε δύο παράγοντες και υπάρχει μία παρατήρηση για κάθε συνδυασμό των δύο παραγόντων τότε ο έλεγχος της επίδρασης του κάθε παράγοντα (δηλ της διαφοράς μεταξύ των επιπέδων του παράγοντα) γίνεται με την ανάλυση διασποράς με δύο παράγοντες. Η ανάλυση γίνεται συνήθως με τη χρήση στατιστικού προγράμματος.

*Παράδειγμα:* Σε ένα πείραμα για να συγκρίνουμε την επίδραση  $k=3$  φαρμάκων στον αριθμό λεμφοκυττάρων σε ποντίκια, χρησιμοποιήθηκε ένας σχεδιασμός με 3 ποντίκια από  $b=4$  διαφορετικά κλουβιά. Ο αριθμός λεμφοκυττάρων ήταν:

<i>Litters</i>	1	2	3	4	μέση τιμή
<i>Drugs</i>					
<i>a</i>	7.1	6.1	6.9	5.6	6.43
<i>b</i>	6.7	5.0	5.9	5.1	5.68
<i>c</i>	6.6	5.4	5.8	5.2	5.75

Τα δεδομένα ταξινομούνται σε δύο παράγοντες, κατά κλουβί και κατά φάρμακο. Το σύνολο των παρατηρήσεων είναι  $n=12$ .

Η διαίρεση της συνολικής διακύμανσης ακολουθεί την ίδια φιλοσοφία με την One-way ANOVA.

Η ANOVA παρουσιάζεται με τον εξής πίνακα:

Source of variation	df	SS	MS=SS/df	F=MS/Resid. MS
Between litters	$b-1=4-1=3$	4.23	1.40	
Between drugs	$k-1=3-1=2$	1.37	0.683	14.89
Residual	$11-3-2=6$	0.275	$s^2=0.046$	
Total	$n-1=12-1=11$			

Ελέγχουμε τη υπόθεση ότι τα φάρμακα δε διαφέρουν, δηλ.  $\mu_a=\mu_b=\mu_c$ , συγκρίνοντας την τιμή  $F=(\text{Between drugs MS})/(\text{Residual MS})$  με την τιμή της **F-κατανομής** με 2 και 6 df (Between drugs df και Residual df) σε 5% στάθμη σημαντικότητας.

Επειδή η τιμή της  $F=14.89$  είναι μεγαλύτερη από τη τιμή της F-κατανομής που είναι 5.14 τότε υπάρχει διαφορά μεταξύ των φαρμάκων με  $P<0.05$ .

Το τυπικό σφάλμα της διαφοράς μεταξύ των μέσων τιμών δύο φαρμάκων είναι  $\sqrt{\{s^2(1/k + 1/k)\}} = \sqrt{\{0.046(1/4 + 1/4)\}} = 0.15$ .

Μπορούμε να κάνουμε συγκρίση των μέσων τιμών δύο φαρμάκων χρησιμοποιώντας το t-test, π.χ. για να συγκρίνουμε τις μέσες τιμές του a και c το t-test είναι:

$$t = (\bar{X}_x - \bar{X}_c) / \text{s.e.}(\bar{X}_a - \bar{X}_c) = (6.43 - 5.75) / 0.15 = 4.53,$$

τιμή που είναι μεγαλύτερη από την αντίστοιχη τιμή της t-κατανομής με 6 df σε 5% στάθμη σημαντικότητας (2.45), δηλ. υπάρχει διαφορά μεταξύ των δύο φαρμάκων ( $P < 0.05$ ).

### Ανάλυση διασποράς με δύο παράγοντες (Two-way ANOVA) με αλληλεπίδραση

Όταν τα δεδομένα ταξινομούνται σε δύο παράγοντες και υπάρχουν πολλαπλές παρατηρήσεις για κάθε συνδυασμό των δύο παραγόντων τότε ο έλεγχος της επίδρασης του κάθε παράγοντα (δηλ της διαφοράς μεταξύ των επιπέδων του παράγοντα) ή της αλληλεπίδρασης μεταξύ των παραγόντων γίνεται με την ανάλυση διασποράς με δύο παράγοντες και αλληλεπίδραση.

*Παράδειγμα:* Για να ερευνήσουμε την επίδραση του σορβικού οξέος (sa) και του pH του νερού στην επιβίωση της σαλμονέλας, χρησιμοποιήσαμε  $w=3$  επίπεδα pH (5.0, 5.5, 6.0) και  $s=2$  επίπεδα σορβικού οξέος (0, 100 p.p.m.). Για τον κάθε συνδυασμό sa και pH υπάρχουν  $k=3$  παρατηρήσεις. Μία εβδομάδα μετά μετρήθηκε η ποσότητα σαλμονέλας που επιβίωσε ( $\log(\text{πυκνότητα/ml})$ ). Τα δεδομένα ήταν:

Sa	pH	repli cations		
		1	2	3
0	6.0	8.2	8.4	8.3
	5.5	5.9	6.0	6.1
	5.0	4.3	4.3	4.2
100	6.0	7.6	7.8	7.6
	5.5	5.0	5.3	5.8
	5.0	4.1	4.4	4.2

Η ANOVA παρουσιάζεται με τον εξής πίνακα:

Source of variation	df	SS	MS=SS/df	F=MS/s <sup>2</sup>
pH	$w-1=2$	42.46	21.23	590
sa	$s-1=1$	0.86	0.86	23.89
Interaction pHXsa	$(w-1)*(s-1)=2$	0.34	0.17	4.72
Residual	$17-2-1-2=12$	0.43	$s^2=0.036$	
Total	$n=18-1=17$			

Ελέγχουμε τη υπόθεση ότι τα επίπεδα pH δε διαφέρουν, δηλ.  $\mu_{5.0}=\mu_{5.5}=\mu_{6.0}$ , συγκρίνοντας την τιμή  $F=(\text{pH MS})/(\text{Residual MS})$  με την τιμή της **F-κατανομής** με 2 και 12 df (pH df και Residual df) σε 5% στάθμη σημαντικότητας. Επειδή η τιμή της  $F=590$  είναι μεγαλύτερη από τη τιμή της F-κατανομής (3.89) τότε υπάρχει διαφορά μεταξύ των επιπέδων pH με  $P<0.05$ .

Ελέγχουμε την υπόθεση ότι τα επίπεδα sa δε διαφέρουν, δηλ.  $\mu_0=\mu_{100}$ , συγκρίνοντας την τιμή  $F=(\text{sa MS})/(\text{Residual MS})$  με την τιμή της **F-κατανομής** με 1 και 12 df (sa df και



Residual df) σε 5% στάθμη σημαντικότητας. Επειδή η τιμή της  $F=23.89$  είναι μεγαλύτερη από τη τιμή της F-κατανομής (4.75) τότε υπάρχει διαφορά μεταξύ των επιπέδων pH με  $P<0.05$ .

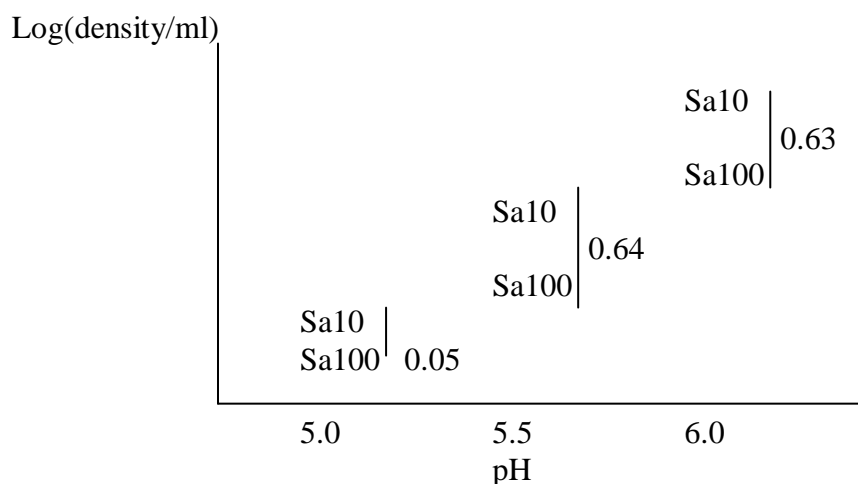
Ελέγχουμε τη υπόθεση ότι δεν υπάρχει αλληλεπίδραση μεταξύ pH και sa συγκρίνοντας την τιμή  $F=(\text{Interaction MS})/(\text{Residual MS})$  με την τιμή της F-κατανομής με 2 και 12 df (Interaction df και Residual df) σε 5% στάθμη σημαντικότητας. Επειδή η τιμή της  $F=4.72$  είναι μεγαλύτερη από τη τιμή της F-κατανομής (3.89) τότε υπάρχει αλληλεπίδραση μεταξύ pH και sa με  $P<0.05$ .

### Αλληλοεπίδραση

Η ύπαρξη της αλληλεπίδρασης σημαίνει ότι η διαφορά sa0-sa100 δεν είναι σταθερή για τα διαφορετικά επίπεδα pH. Τότε τα αποτελέσματα παρουσιάζονται με ένα 2\*3 πίνακα μέσω των τιμών για τον κάθε συνδυασμό sa και pH:

Sa	pH			sa means
	6.0	5.5	5.0	
0	8.3	6.1	4.3	6.2
100	7.7	5.4	4.2	5.8
pH means	8.0	5.7	4.3	6.0

Τα αποτελέσματα μπορούν να παρουσιασθούν και με το αντίστοιχο διάγραμμα:



Το τυπικό σφάλμα της διαφοράς μεταξύ των μέσων τιμών δύο επιπέδων pH είναι  $\sqrt{\{s^2(1/(s*k) + 1/(s*k))\}} = \sqrt{\{0.036(1/6 + 1/6)\}} = 0.11$ .

Το τυπικό σφάλμα της διαφοράς μεταξύ των μέσων τιμών δύο επιπέδων sa είναι  $\sqrt{\{s^2(1/(w*k) + 1/(w*k))\}} = \sqrt{\{0.036(1/9 + 1/9)\}} = 0.09$ .

Το τυπικό σφάλμα της διαφοράς μεταξύ των μέσων τιμών δύο επιπέδων pH για ένα συγκεκριμένο επίπεδο sa είναι  $\sqrt{\{s^2(1/k + 1/k)\}} = \sqrt{\{0.036(1/3 + 1/3)\}} = 0.15$ .

Μπορούμε να κάνουμε σύγκριση των μέσων τιμών δύο επιπέδων pH ή δύο επιπέδων sa χρησιμοποιώντας το t-test, π.χ. για να συγκρίνουμε τις μέσες τιμές του pH6.0 και pH5.0 το t-test είναι:

$$t = (\bar{X}_{\text{pH6.0}} - \bar{X}_{\text{pH5.0}}) / \text{s.e.}(\bar{X}_{\text{pH6.0}} - \bar{X}_{\text{pH5.0}}) = (8.0 - 4.3) / 0.11 = 37.3,$$

τιμή που είναι μεγαλύτερη από την αντίστοιχη τιμή της t-κατανομής με 12 df σε 5% στάθμη σημαντικότητας (2.18), δηλ. υπάρχει διαφορά μεταξύ των δύο επιπέδων pH ( $P < 0.05$ ).

Επίσης, μπορούμε να κάνουμε σύγκριση των μέσων τιμών δύο επιπέδων pH για ένα επίπεδο του sa χρησιμοποιώντας το t-test, π.χ. για να συγκρίνουμε τις μέσες τιμές του pH6.0 και pH5.0 για sa100 το t-test είναι:

$$t = (\bar{X}_{\text{pH6.0,sa100}} - \bar{X}_{\text{pH5.0,sa100}}) / \text{s.e.}(\bar{X}_{\text{pH6.0,sa100}} - \bar{X}_{\text{pH5.0,sa100}}) = (7.7 - 4.2) / 0.15 = 25.6,$$

τιμή που είναι μεγαλύτερη από την αντίστοιχη τιμή της t-κατανομής με 12 df σε 5% στάθμη σημαντικότητας (2.18), δηλ. υπάρχει διαφορά μεταξύ των δύο επιπέδων pH για sa100 ( $P < 0.05$ ).

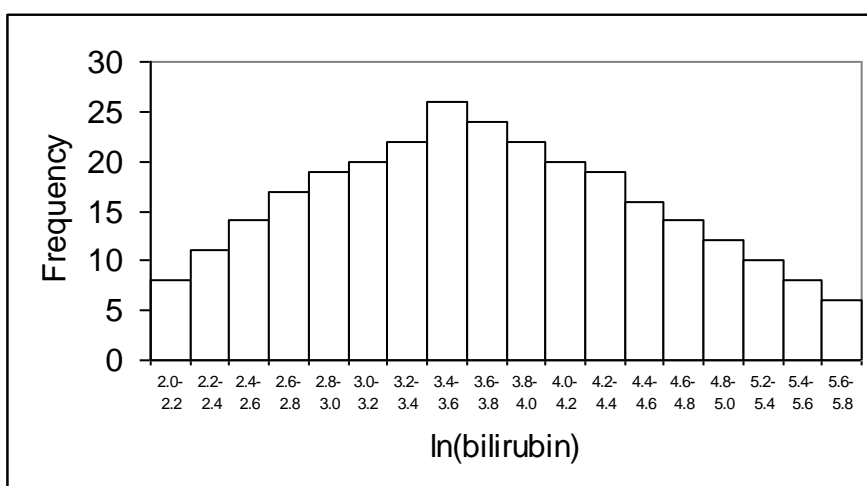
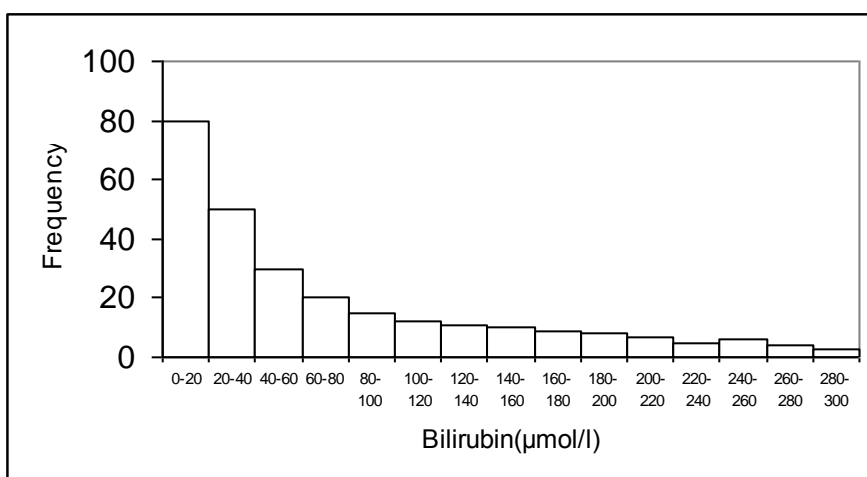
## Μετατροπή των δεδομένων

Για να είναι έγκυρος ένας στατιστικός έλεγχος όπως t-test ή ANOVA πρέπει να υποθέσουμε ότι: 1) οι παρατηρήσεις έχουν κανονική κατανομή και 2) οι διακυμάνσεις των ομάδων που συγκρίνονται είναι ίσες.

Αν δεν πληρούνται οι παραπάνω υποθέσεις και τα δείγματα είναι μικρά σε μέγεθος τότε τα δεδομένα χρειάζεται να μετασχηματισθούν στους λογαρίθμους τους. Η ανάλυση τότε θα βασίζεται στα λογαριθμοποιημένα δεδομένα. Αν τα δεδομένα είναι συχνότητες τότε ίσως χρειάζεται να μετασχηματισθούν στις τετραγωνικές τους ρίζες.

Ο λογάριθμος (log) μίας ποσότητας  $x$  είναι η ποσότητα  $y$ ,  $y=\log(x)$ , έτσι ώστε  $x=e^y$ , όπου  $e=2.718$ . Ο λογάριθμος του 1 είναι 0 και του 0 είναι άπειρο. Μπορούμε να βρούμε τον λογάριθμο μόνο θετικών τιμών.

*Παράδειγμα:* Η κατανομή των τιμών χολερυθρίνης ορού μίας ομάδας ατόμων δεν ακολουθεί την κανονική κατανομή. Τότε με έναν λογαριθμικό μετασχηματισμό τα δεδομένα ακολουθούν κανονική κατανομή. (σελ 144 altman)



### Έλεγχος για ένα ποσοστό

Η ανάλυση των ποσοστών βασίζεται στη δυωνυμική κατανομή (που εξετάζεται παρακάτω) η οποία για δείγματα μεγάλου μεγέθους προσεγγίζεται από την κανονική κατανομή. Ένα δείγμα θεωρείται μεγάλου μεγέθους όταν  $np$  και  $n(1-p)$  είναι και τα δύο μεγαλύτερα από 5, όπου  $p$  είναι ένα παρατηρούμενο ποσοστό και  $\pi$  είναι το αναμενόμενο ποσοστό (ή το ποσοστό ενός πληθυσμού).

Για τον έλεγχο ενός ποσοστού χρησιμοποιείται το z-test και βασίζεται στην κανονική κατανομή.

*Παράδειγμα:* Σε ένα νοσοκομείο συλλέχθηκε ένα τυχαίο δείγμα από 215 γυναίκες από τους καταλόγους νοσηλευομένων και βρέθηκε ότι 39 από αυτές έχουν ιστορικό άσθματος. Είναι γνωστό ότι ο επιπολασμός της νόσου είναι 15%. Το ποσοστό των γυναικών που βρέθηκε με άσθμα συμφωνεί με αυτό του γενικού πληθυσμού;

Τότε για να ελέγξουμε την υπόθεση ότι το ποσοστό των περιστατικών δεν διαφέρει από τον επιπολασμό της νόσου, δηλ.  $p-\pi=0$ , χρησιμοποιούμε το z-test:

**$z = (\text{διαφορά παρατηρούμενου} - \text{θεωρητικού ποσοστού}) / (\text{τυπ. σφάλμα θεωρ. ποσοστού})$  ή**

$$z = (p - \pi) / \text{s.e.}(\pi), \text{ όπου } \text{s.e.}(\pi) = \sqrt{\{\pi(1-\pi)/n\}}$$

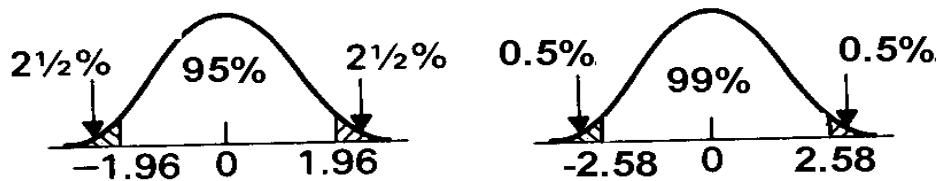
Τότε  $n=215$ ,  $r=39$ ,  $p = r / n = 39 / 215 = 0.18$  το ποσοστό των περιστατικών,  $\pi=0.15$ , ο επιπολασμός της νόσου και

$$\text{s.e.}(\pi) = \sqrt{\{\pi(1-\pi)/n\}} = \sqrt{\{0.15(1-0.15)/215\}} = 0.024$$

Οπότε  $z = (0.18 - 0.15) / \sqrt{\{0.15(1-0.15)/215\}} = 0.03 / 0.024 = 1.23$  (το πρόσημο αγνοείται)

Αν το ποσοστό των περιστατικών είναι πράγματι ίδιο με τον επιπολασμό της νόσου τότε η ποσότητα  $z$  είναι μια τυπική τιμή της τυπικής κανονικής κατανομής.

Το 5% σημείο της τυπικής κανονικής κατανομής είναι 1.96 και το 1% σημείο της κανονικής κατανομής είναι 2.58.



Επειδή  $z = 1.23$  είναι μικρότερο από 1.96 συμπεραίνουμε ότι το ποσοστό γυναικών με άσθμα στο νοσοκομείο δε διαφέρει από αυτό του γενικού πληθυσμού ( $P > 0.05$ ). Προφανές είναι ότι δε διαφέρουν και για  $P > 0.01$ .

### Διάστημα εμπιστοσύνης για ένα ποσοστό

Το 95% δ.ε. ενός παρατηρούμενου ποσοστού δίνεται προσεγγιστικά από:

**(παρατηρούμενο ποσοστό)  $\pm$  (1.96) \* (τυπικό σφάλμα ποσοστού) ή**

$(p - 1.96 * s.e.(p), p + 1.96 * s.e.(p))$ , όπου  $s.e.(p) = \sqrt{\{p(1-p)/n\}}$ .

Το 99% δ.ε. ενός ποσοστού δίνεται προσεγγιστικά από:

$(p - 2.58 * s.e.(p), p + 2.58 * s.e.(p))$ , όπου  $s.e.(p) = \sqrt{\{p(1-p)/n\}}$

*Παράδειγμα:* Στο παράδειγμά μας, το 95% δ.ε. του ποσοστού των γυναικών με άσθμα στον πληθυσμό είναι προσεγγιστικά:

$(0.18 - 1.96 * \sqrt{\{0.18(1-0.18)/215\}}, 0.18 + 1.96 * \sqrt{\{0.18(1-0.18)/215\}})$  ή  $(0.13, 0.23)$

Αν υποθέσουμε ότι το δείγμα των γυναικών από το νοσοκομείο είναι αντιπροσωπευτικό όλων των γυναικών στη χώρα τότε μπορούμε να ισχυριστούμε με βεβαιότητα ότι ο επιπολασμός του άσθματός στη χώρα είναι μεταξύ 13 και 23%.

Το 99% δ.ε. του ποσοστού είναι προσεγγιστικά:

$$(0.18 - 2.58 \cdot \sqrt{\{0.18(1-0.18)/215\}}, 0.18 + 2.58 \cdot \sqrt{\{0.18(1-0.18)/215\}}) \text{ ή } (0.11, 0.25)$$

Δηλαδή, είμαστε 99% βέβαιοι ότι ο επιπολασμός είναι μεταξύ 11 και 25%.

### **Έλεγχος ποσοστού με διόρθωση**

Για σχετικά μικρά δείγματα η εξής ποσότητα δίνει πιο ακριβή αποτελέσματα:

$$z_c = \{ |p - \pi| - 1/(2n) \} / \sqrt{\{\pi(1-\pi)/n\}}$$

*Παράδειγμα:* Για το παράδειγμά μας  $z_c = 1.14$  που είναι πολύ λίγο χαμηλότερο από το  $z = 1.23$  (που βρήκαμε χωρίς τη διόρθωση) επειδή το μέγεθος του δείγματος είναι μεγάλο.

### Έλεγχος για τη σύγκριση δύο ποσοστών

Για τον έλεγχο της σύγκρισης δύο ποσοστών χρησιμοποιείται το z-test και βασίζεται στην κανονική κατανομή.

*Παράδειγμα:* Είκοσι από τα 240 άτομα που εμβολιάστηκαν με πραγματικό εμβόλιο προσβλήθηκαν από γρίπη σε σύγκριση με τα 80 από τα 220 που εμβολιάστηκαν με placebo. Υπάρχει ένδειξη ότι το εμβόλιο ήταν αποτελεσματικό;

Τότε για να ελέγχσουμε την υπόθεση ότι το ποσοστό προσβολής από γρίπη όταν γίνεται εμβολιασμός δεν είναι διαφορετικό από το ποσοστό με το placebo, δηλ.  $p_1 - p_2 = 0$ , χρησιμοποιείται το z-test:

**$z = (\text{διαφορά δύο ποσοστών}) / (\text{τυπικό σφάλμα της διαφοράς})$  ή**

$$z = (p_1 - p_2) / \text{s.e.}(p_1 - p_2), \text{όπου } \text{s.e.}(p_1 - p_2) = \sqrt{\{p(1-p)(1/n_1 + 1/n_2)\}}$$

$$\text{και } p_1 = r_1/n_1, p_2 = r_2/n_2 \text{ και } p = (r_1 + r_2)/(n_1 + n_2)$$

$$\text{Τότε } r_1 = 20, n_1 = 240, r_2 = 80, n_2 = 220, p_1 = 20/240 = 0.083, p_2 = 80/220 = 0.364,$$

$$\text{Το συνολικό ποσοστό με γρίπη ήταν: } p = (20 + 80)/(240 + 220) = 0.217 \text{ ή } 21.7\% \text{ και}$$

$$\text{s.e.}(p_1 - p_2) = \sqrt{\{p(1-p)(1/n_1 + 1/n_2)\}} = \sqrt{\{0.217(1-0.217)(1/240 + 1/220)\}} = 0.038.$$

$$\text{Οπότε } z = (0.083 - 0.364) / \sqrt{\{0.217(1-0.217)(1/240 + 1/220)\}} = -0.281/0.038 = -7.30 \text{ (το πρόσημο αγνοείται)}$$

Αν δεν υπάρχει διαφορά μεταξύ των δύο ποσοστών τότε η ποσότητα της z είναι μία τυπική τιμή της τυπικής κανονικής κατανομής. Η τιμή  $z = 7.30$  είναι μεγαλύτερη από την τιμή της κανονικής κατανομής στο 1% σημείο που είναι 2.58.

Συνεπώς, υπάρχει σημαντική ένδειξη ( $P < 0.01$ ) ότι ο πραγματικός εμβολιασμός ελαττώνει τον κίνδυνο προσβολής από γρίπη.

### Διάστημα εμπιστοσύνης για τη διαφορά δύο ποσοστών

Το 95% δ.ε. για τη διαφορά μεταξύ των δύο ποσοστών δίνεται προσεγγιστικά από:

**(διαφορά δύο ποσοστών) $\pm$ (1.96)\*(τυπικό σφάλμα διαφοράς) ή**

**(( $p_1 - p_2$ ) $-1.96*s.e.(p_1 - p_2)$ , ( $p_1 - p_2$ ) $+1.96*s.e.(p_1 - p_2)$ ),**

**όπου  $s.e. (p_1 - p_2) = \sqrt{\{p_1(1-p_1)/n_1 + p_2(1-p_2)/n_2\}}$ .**

*Παράδειγμα:* Στο παράδειγμά μας, το 95% δ.ε. αυτής της ελάττωσης του ποσοστού των ατόμων με γρίπη λόγω του εμβολίου είναι:

**( $-0.281 - 1.96*0.037$ ,  $-0.281 + 1.96*0.037$ ) ή ( $-0.354$ ,  $-0.208$ ).**

Συνεπώς, η πιθανότητα προσβολής από γρίπη όταν γίνεται πραγματικός εμβολιασμός είναι χαμηλότερος μεταξύ 20.8% και 35.4% σε σχέση με τη χορήγηση placebo.



### Πίνακες συνάφειας - $\chi^2$ test

Για να ελέγχσουμε τη σχέση μεταξύ δύο ποιοτικών δυαδικών μεταβλητών που μπορούν να διαταχθούν σε έναν  $2 \times 2$  πίνακα συχνοτήτων (συνάφειας) χρησιμοποιούμε το  $\chi^2$ -test. Ο έλεγχος αυτός είναι ισοδύναμος με τον έλεγχο της υπόθεσης ότι δύο ποσοστά δε διαφέρουν.

*Παράδειγμα:* Σε μία κλινική μελέτη για να αξιολογήσουμε την αποτελεσματικότητα μίας νέας αντικαρκινικής μεθόδου (Α) σε σχέση με μία παλιά (Β), οι ασθενείς χωρίστηκαν τυχαία σε δύο ομάδες που αντιστοιχούν στις δύο μεθόδους. Από τους 257 ασθενείς που εφαρμόστηκε η Α, οι 41 απεβίωσαν και από τους 244 ασθενείς που εφαρμόστηκε η μέθοδος Β, οι 64 απεβίωσαν. Υπάρχει καμία διαφορά στους δείκτες θνητότητας για τις δύο μεθόδους; ή, υπάρχει σχέση μεταξύ μεθόδου και αποτελέσματος;

Μέθοδος	Αποτέλεσμα		σύνολο
	θάνατος	επιβίωση	
A	41 (53.86)	216 (203.14)	257
B	64 (51.14)	180 (192.86)	244
Σύνολο	105	396	501

Για να ελέγχσουμε την υπόθεση ότι δεν υπάρχει σχέση μεταξύ αντικαρκινικής μεθόδου και αποτελέσματος ή ότι το αποτέλεσμα είναι ανεξάρτητο της μεθόδου ή ότι οι θνητότητες των δύο μεθόδων δε διαφέρουν, χρησιμοποιούμε το  $\chi^2$ -test.

Το  $\chi^2$  test: συγκρίνει της παρατηρούμενες συχνότητες (Ο) σε κάθε μία από τις 4 κατηγορίες του πίνακα με τις αναμενόμενες συχνότητες (Ε) αν δεν υπήρχαν διαφορές μεταξύ των θνητοτήτων:

$$\chi^2 = (\mathbf{O_1 - E_1})^2/\mathbf{E_1} + (\mathbf{O_2 - E_2})^2/\mathbf{E_2} + (\mathbf{O_3 - E_3})^2/\mathbf{E_3} + (\mathbf{O_4 - E_4})^2/\mathbf{E_4}$$

Ο συνολικός δείκτης θνητότητας είναι 105/501, αν οι δείκτες θνητότητας για την Α και Β ήταν ίδιοι, η καλύτερη εκτίμηση της κοινής θνητότητας θα ήταν 105/501. Αφού έχουμε συνολικά 257 ασθενείς στην Α, ο αναμενόμενος αριθμός θανάτων για την Α (αν οι θνητότητες ήταν ίδιες) είναι:  $E_1 = 105/501 \times 257 = (105 \times 257)/501 = 53.86$ .

$$\text{Όμοια } E_2 = (105 \cdot 244) / 501 = 51.14,$$

$$E_3 = (396 \cdot 257) / 501 = 203.14,$$

$$E_4 = (396 \cdot 244) / 501 = 192.86.$$

Τότε είναι

$$\chi^2 = \{(41-53.86)^2/53.86\} + \{(64-51.14)^2/51.14\} + \{(216-203.14)^2/203.14\} + \{(180-192.86)^2/192.8\} = 7.98$$

Αν η υπόθεση ότι δεν υπάρχει σχέση μεταξύ μεθόδου και αποτελέσματος είναι αληθινή τότε η ποσότητα  $\chi^2$  ακολουθεί την  $\chi^2$ -κατανομή με 1 β.ε (πίνακας 2).

Επειδή το 7.98 είναι μεγαλύτερο από την τιμή της  $\chi^2$ -κατανομής με 1 β.ε. με στάθμη σημαντικότητας 5% που είναι 3.84, συμπεραίνουμε ότι η υπόθεση που κάναμε δεν είναι σωστή. Συνεπώς υπάρχει ένδειξη ( $P < 0.05$ ) ότι το αποτέλεσμα σχετίζεται με τη μέθοδο ή ότι οι θνητότητες των δύο μεθόδων διαφέρουν.

### Λόγος αναλογιών (Odds Ratio) σε μελέτες ασθενών-μαρτύρων –

Οι μελέτες ασθενών-μαρτύρων χρησιμοποιούνται στην επιδημιολογία για να ερευνήσουμε τη σχέση μεταξύ ενός πιθανού παράγοντα κινδύνου (π.χ. κάπνισμα) και μίας νόσου (π.χ. καρκίνο του πνεύμονα).

*Παράδειγμα:* Μία ομάδα από άτομα που νοσούν (ασθενείς) συγκρίνεται με μία ομάδα μαρτύρων από άτομα που δεν νοσούν. Τότε υπολογίζεται η συχνότητα για την κάθε ομάδα αυτών που έχουν εκτεθεί και αυτών που δεν έχουν εκτεθεί στο παράγοντα κινδύνου. Τα αποτελέσματα μίας μελέτης παρουσιάζονται σε έναν 2x2 πίνακα:

	<i>Cases</i>	<i>controls</i>	<i>total</i>
<i>Exposed</i>	a	c	a+c
<i>Not Exposed</i>	b	d	b+d
<i>Total</i>	a+b	c+d	n

Τότε οι δύο ομάδες μπορούν να συγκριθούν χρησιμοποιώντας το λόγο αναλογιών (odds ratio, OR) που εκτιμά τον κίνδυνο να νοσήσει κάποιος αν είναι εκτεθειμένος σε σχέση με τον κίνδυνο αν δεν είναι εκτεθειμένος.

Το OR ορίζεται ως  $OR = (a/c)/(b/d) = (ad/bc)$

Το s.e. του  $\log(OR)$  είναι  $s.e.(\log OR) = \sqrt{\{ 1/a + 1/b + 1/c + 1/d \}}$

Το 95% δ.ε. για το  $\log(OR)$  είναι  $(\log(OR) \pm 1.96 * s.e.\log(OR))$

*Παράδειγμα:* Σε μία μελέτη ασθενών-μαρτύρων για να ερευνήσουμε τη σχέση μεταξύ κατανάλωσης καφέ και καρκίνου του παγκρέατος βρέθηκαν τα παρακάτω αποτελέσματα:

	diseased		
Coffee drinking	cases	controls	total
Yes	32	118	150
No	17	127	144
Total	49	245	294

Θέλουμε να ερευνήσουμε τη σχέση μεταξύ του πιθανού παράγοντα κινδύνου που είναι η κατανάλωση καφέ και της εμφάνισης της νόσου που είναι καρκίνος του παγκρέατος.

Τότε ο κίνδυνος να νοσήσει κάποιος αν καταναλώνει καφέ σε σχέση με τον κίνδυνο αν δεν καταναλώνει είναι:

$$OR=(32*127)/(118*17)=2.026 \text{ και } \log(OR)=0.706.$$

$$\text{Επίσης } s.e.\log(OR)=\sqrt{\{ 1/32+1/17 +1/118 + 1/127 \}} = 0.326$$

Το 95% δ.ε. για το  $\log(OR)$  είναι  $(0.706-1.96*0.326, 0.706+1.96*0.326) = (0.067, 1.345)$ , οπότε το 95% δ.ε. για το OR είναι  $(e^{0.067}, e^{1.345}) = (1.069, 3.840)$ .

Επειδή το 1 δεν συμπεριλαμβάνεται στο δ.ε., συμπεραίνουμε ότι υπάρχει αυξανόμενος κίνδυνος (διπλάσιος) για καρκίνο σε αυτούς που πίνουν καφέ σε σχέση με αυτούς που δεν πίνουν.

Όταν ο κίνδυνος να νοσήσει εξαρτάται την ίδια στιγμή και από άλλες μεταβλητές (ή/και παράγοντες κινδύνου) τότε ο υπολογισμός των OR γίνεται εφαρμόζοντας την ανάλυση λογιστικής παλινδρόμησης (logistic regression):

## Διαγνωστικοί έλεγχοι

Η διάγνωση είναι ένα σημαντικό μέρος της κλινικής πρακτικής και αρκετές ιατρικές μελέτες γίνονται με σκοπό την βελτίωση των μεθόδων διάγνωσης.

Για την διαγνωστική αξιολόγηση των μελετών αυτών, συνήθως οι ασθενείς ταξινομούνται σε δύο ομάδες ανάλογα με τα αποτελέσματα μίας διαγνωστικής μεθόδου που θεωρείται σημείο αναφοράς και η οποία δίδει τη πραγματική νοσολογική κατάσταση του ασθενή, π.χ. βιοψία. Μετά εφαρμόζεται η προς αξιολόγηση διαγνωστική μέθοδος στις δύο ομάδες και συγκρίνονται τα αποτελέσματα.

*Παράδειγμα:* Σε μία μελέτη για τη σχέση μεταξύ των αποτελεσμάτων του σπινθηρογραφήματος του ήπατος(liver scanning) και της διάγνωσης των ιστολογικών αλλοιώσεων του ήπατος που βασίζεται σε βιοψία, χρησιμοποιήθηκαν 344 ασθενείς και βρέθηκαν τα παρακάτω αποτελέσματα:

Liver scan	Βιοψία		σύνολο
	κακοήθες (+)	καλοήθες (-)	
Κακοήθες (+)	a=231	b=32	263
Καλοήθες (-)	c=27	d=54	81
Σύνολο	258	86	n=344

Το ερώτημα είναι, κατά πόσο είναι καλό το σπινθηρογράφημα στη διάγνωση της ιστοπαθολογίας και αν υπάρχει συμφωνία μεταξύ σπινθηρογραφήματος και βιοψίας.

Για να περιγράψουμε την ικανότητα του σπινθηρογραφήματος να διαγνώσει την πραγματική κατάσταση ενός ασθενή χρησιμοποιούμε την ευαισθησία (sensitivity), την ειδικότητα (specificity), την θετική διαγνωστική αξία (positive predictive value, PPV) και την αρνητική διαγνωστική αξία (negative predictive value, NPV).

Η ευαισθησία είναι το ποσοστό των ορθά θετικών ευρημάτων από σπινθηρογράφημα στο σύνολο των ασθενών με κακοήθεια ή η πιθανότητα ενός θετικού ευρήματος όταν η παρουσία της κακοήθειας είναι δεδομένη.

Η ειδικότητα είναι το ποσοστό των ορθά αρνητικών ευρημάτων στο σύνολο των εξετασθέντων με καλοήθεια ή η πιθανότητα ενός αρνητικού ευρήματος όταν η απουσία της κακοήθειας είναι δεδομένη.

Η θετική διαγνωστική αξία είναι το ποσοστό των «ορθά θετικών» στο σύνολο των θετικών ευρημάτων.

Η αρνητική διαγνωστική αξία είναι το ποσοστό των «ορθά αρνητικών» στο σύνολο των αρνητικών ευρημάτων.

Οπότε η ευαισθησία είναι:  $Se=a/(a+c)=231 / 258 = 0.90$ .

Το 90% των θετικών στη βιοψία ανιχνεύθηκε σωστά από το σπινθηρογράφημα (% ορθά θετικών).

Η ειδικότητα είναι:  $Sp=d/(b+d)=54 / 86 = 0.63$ .

Το 63% των αρνητικών στη βιοψία ανιχνεύθηκε σωστά από το σπινθηρογράφημα (% ορθά αρνητικών).

Η θετική διαγνωστική αξία είναι:  $PPV=a/(a+b) = 231 / 263 = 0.88$ .

Το 88% των θετικών ασθενών από το σπινθηρογράφημα διαγνώσθηκαν σωστά.

Η αρνητική διαγνωστική αξία είναι:  $NPV=d/(c+d)= 54 / 81 = 0.67$

Το 67% των αρνητικών ασθενών από το σπινθηρογράφημα διαγνώσθηκαν σωστά.

### **Έλεγχος συμφωνίας δύο διαγνωστικών μεθόδων- McNemar's-test**

Για να ερευνήσουμε την συμφωνία μεταξύ των δύο διαγνωστικών μεθόδων, ελέγχουμε την υπόθεση ότι οι δύο μέθοδοι συμφωνούν μεταξύ τους χρησιμοποιώντας το McNemar's-test:

$$X^2=\{|b-c|-1\}^2/(b+c)$$

Οπότε  $X^2=\{|32-27|-1\}^2/(32+27) = 16/59=0.27$ .

Αν οι δύο μέθοδοι συμφωνούν τότε η ποσότητα  $X^2$  είναι μια τυπική τιμή της  $\chi^2$ -κατανομής με 1 β.ε. Η ποσότητα  $X^2=0.27$  είναι μικρότερη από τη τιμή της  $\chi^2$ -κατανομής με 1 β.ε. με στάθμη σημαντικότητας 5% (που είναι 3.84). Συνεπώς δεν απορρίπτουμε την υπόθεση ότι οι δύο διαγνωστικές μέθοδοι συμφωνούν ( $P<0.05$ ).

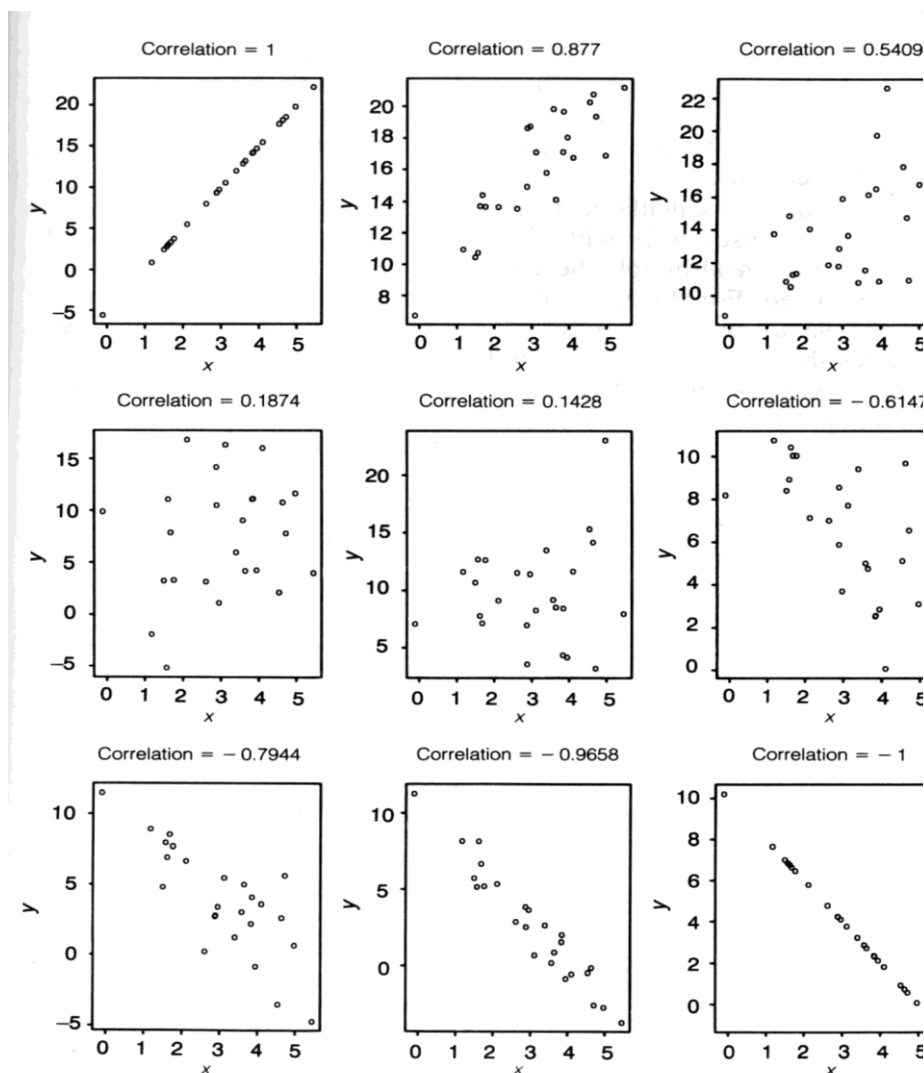
### Συντελεστής συσχέτισης (correlation coefficient) r

Η συσχέτιση δείχνει πόσο καλά συσχετίζονται μεταξύ τους δύο ποσοτικές μεταβλητές x και y. Το μέτρο της συσχέτισης δίνεται από:

$$r = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{[\sum (x - \bar{x})^2 \sum (y - \bar{y})^2]}}$$

Ο συντελεστής συσχέτισης είναι πάντα μεταξύ -1 και +1, και είναι 0 αν οι μεταβλητές δε συσχετίζονται. Είναι θετικός αν το x και το y τείνουν σε μεγάλες ή χαμηλές τιμές μαζί, και όσο πιο μεγάλη είναι η τιμή του r τόσο πιο μεγάλη είναι και η συσχέτισή τους. Η μέγιστη τιμή 1 αποκτάται αν όλα τα σημεία βρίσκονται σε μία γραμμή.

Τα παρακάτω διαγράμματα δείχνουν τη δομή των δεδομένων για διάφορες τιμές του συντελεστή συσχέτισης.

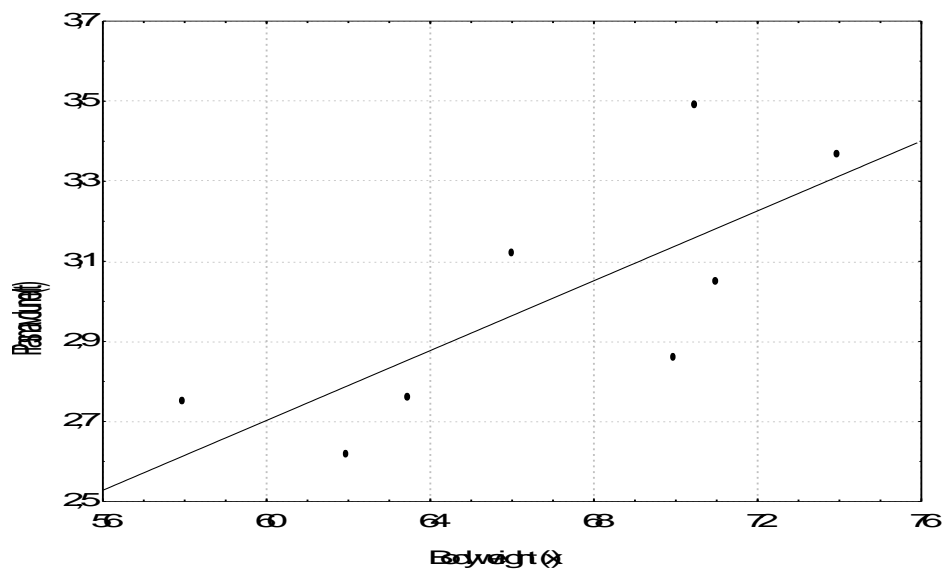




Παράδειγμα: Το βάρος του σώματος και ο όγκος πλάσματος από 8 υγιείς άνδρες είναι:

Άτομο	βάρος σε Kg (x)	όγκος πλάσματος σε lt (y)
1	58.0	2.75
2	70.0	2.86
3	74.0	3.37
4	63.5	2.76
5	62.0	2.62
6	70.5	3.49
7	71.0	3.05
8	66.0	3.12

Η σχέση μεταξύ βάρους σώματος και όγκου πλάσματος παρουσιάζεται με το παρακάτω διάγραμμα συσχέτισης το οποίο δείχνει ότι ο αυξημένος όγκος πλάσματος τείνει να σχετίζεται με υψηλό βάρος και χαμηλός όγκος πλάσματος τείνει να σχετίζεται με χαμηλό βάρος:



Είναι  $\sum (x - \bar{x})(y - \bar{y}) = \sum xy - (\sum x)(\sum y)/n$ ,  $\sum (x - \bar{x})^2 = \sum x^2 - (\sum x)^2/n$  και  $\sum (y - \bar{y})^2 = \sum y^2 - (\sum y)^2/n$ .

Οπότε  $n = 8$ ,  $\sum x = 535$ ,  $\bar{x} = 66.875$ ,  $\sum x^2 = 35983.5$ ,

$\sum y = 24.02$ ,  $\bar{y} = 3.0025$ ,  $\sum y^2 = 72.7980$

$\sum xy = 58 \times 2.75 + 70.0 \times 2.86 + \dots + 66.0 \times 3.12 = 1615.295$ , οπότε

$$r = 8.96 / \sqrt{(205.38 * 0.6780)} = 0.76$$

### **Έλεγχος σημαντικότητας της r**

Για να ελέγξουμε την υπόθεση ότι η τιμή της r δεν είναι διαφορετική από το 0 χρησιμοποιούμε το t-test:

$$t = r \sqrt{[(n-2)/(1-r^2)]}$$

*Παράδειγμα:* Για να ελέγξουμε την υπόθεση ότι η τιμή  $r=0.76$  δεν είναι διαφορετική από το 0, (δηλ. οφείλεται απλά στη τύχη) υπολογίζουμε την ποσότητα:

$$t = r \sqrt{[(n-2)/(1-r^2)]} = 0.76 \sqrt{[6/(1-0.76^2)]} = 2.86$$

Αν η υπόθεση είναι πραγματική η τιμή  $t=2.86$  είναι μια τυπική τιμή της t-κατανομής με  $n-2=8-2=6$  β.ε. Η τιμή  $t=2.86$  είναι μεγαλύτερη από τη τιμή της t-κατανομής με 6 β.ε. σε στάθμη σημαντικότητας 5% που είναι (πίνακας 1) 2.45. Συνεπώς, υπάρχει σημαντική συσχέτιση ( $P<0.05$ ) μεταξύ όγκου πλάσματος και βάρους σώματος.

### Γραμμή γραμμικής παλινδρόμησης (regression line)

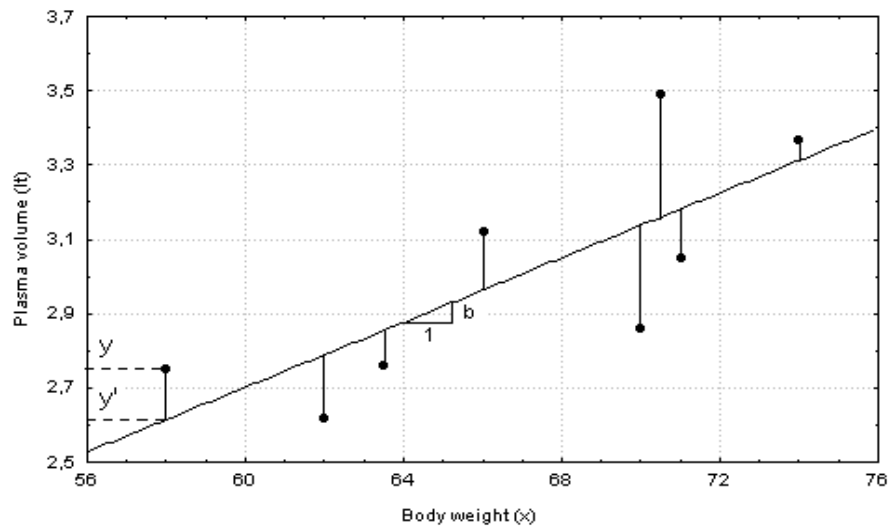
Συχνά ενδιαφερόμαστε για τη σχέση μεταξύ δύο μεταβλητών και πως η μεταβολή της τιμής μιας μεταβλητής επηρεάζει την τιμή μιας άλλης μεταβλητής, π.χ. η σχέση μεταξύ πίεσης του αίματος και ηλικίας, και πως η μεταβολή της ηλικίας επηρεάζει την πίεση του αίματος ή σχέση μεταξύ όγκου πλάσματος και βάρους σώματος, και πως η αύξηση του βάρους σώματος επηρεάζει τον όγκο του πλάσματος. Τότε η σχέση μεταξύ δύο μεταβλητών ελέγχεται με τη γραμμή γραμμικής παλινδρόμησης που είναι η ευθεία γραμμή που περιγράφει πώς μία μεταβλητή εξαρτάται από μία άλλη μεταβλητή.

*Παράδειγμα:* Το βάρος του σώματος και ο όγκος πλάσματος από 8 υγιείς άνδρες είναι:

Άτομο	βάρος σε Kg (x)	όγκος πλάσματος σε lt (y)
1	58.0	2.75
2	70.0	2.86
3	74.0	3.37
4	63.5	2.76
5	62.0	2.62
6	70.5	3.49
7	71.0	3.05
8	66.0	3.12

Η σχέση μεταξύ βάρους σώματος και όγκου πλάσματος παρουσιάζεται με το παρακάτω διάγραμμα συσχέτισης το οποίο δείχνει ότι ο αυξημένος όγκος πλάσματος τείνει να σχετίζεται με υψηλό βάρος και χαμηλός όγκος πλάσματος τείνει να σχετίζεται με χαμηλό βάρος:

Η σχέση μεταξύ δύο μεταβλητών μπορεί να περιγραφεί από μία ευθεία γραμμή που δείχνει πώς η μεταβλητή  $y$  αυξάνεται ή μειώνεται, όταν αυξάνεται η μεταβλητή  $x$ . Η μεταβλητή  $y$  λέγεται εξαρτημένη και η μεταβλητή  $x$  λέγεται ανεξάρτητη. Για παράδειγμα, ο όγκος πλάσματος θεωρείται εξαρτημένη μεταβλητή και το βάρος σώματος ανεξάρτητη μεταβλητή, η πίεση αίματος θεωρείται εξαρτημένη μεταβλητή και η ηλικία ανεξάρτητη μεταβλητή.



Επομένως πρέπει να βρούμε μία γραμμή η οποία αντιπροσωπεύει πολύ στενά τα ζεύγη τιμών  $x$  και  $y$ , δηλαδή μία γραμμή που περνάει ανάμεσα από τα ζεύγη  $x$  και  $y$ . Η κλίση  $b$  και η τομή  $a$  (η τιμή της  $y$  όταν το  $x$  είναι 0) βρίσκονται με τη μέθοδο των ελαχίστων τετραγώνων. Με τη μέθοδο αυτή εκτιμούμε τα  $a$  και  $b$  έτσι ώστε το άθροισμα των τετραγώνων των κάθετων αποστάσεων των σημείων, που ορίζονται από τα ζεύγη  $x$  και  $y$ , από την ευθεία να είναι ελάχιστο. Η γραμμή αυτή λέγεται γραμμή γραμμικής παλινδρόμησης.

Εφαρμόζοντας τη μέθοδο των ελαχίστων τετραγώνων τα  $a$  και  $b$  εκτιμώνται από τις παρακάτω ποσότητες:

$$b = \frac{\sum xy - (\sum x \sum y) / n}{\sum x^2 - (\sum x)^2 / n} \text{ και}$$

$$a = \bar{y} - b \bar{x}, \text{ όπου } \bar{y} \text{ η μέση τιμή της μεταβλητής } y \text{ και } \bar{x} \text{ η μέση τιμή της μεταβλητής } x.$$

Τότε οι προβλεπόμενες τιμές (predicted-fitted values) της  $y$ ,  $y'$ , βρίσκονται πάνω στη γραμμή παλινδρόμησης δίνονται από τον τύπο

$$y' = a + bx$$

Η γραμμή παλινδρόμησης μπορεί να χρησιμοποιηθεί όχι μόνο για να περιγράψει τη σχέση δύο μεταβλητών αλλά και να προβλέψει τιμές της  $y$  για τιμές της  $x$  που είναι μεγαλύτερες από τις παρτηρούμενες..

Για να υπολογίσουμε τη γραμμή παλινδρόμησης υποθέτουμε ότι η ανεξάρτητη μεταβλητή  $x$  δεν έχει σφάλμα στη μέτρησή του και ότι η εξαρτημένη μεταβλητή μετράται με σφάλμα που η κατανομή του είναι η τυπική κανονική κατανομή.

Στο παραδειγμά μας το διάγραμμα δείχνει ότι όσο αυξάνει το βάρος του σώματος  $x$  ανάλογα αυξάνει ο όγκος του πλάσματος  $y$ .

Για να υπολογίσουμε τα  $a$  και  $b$  χρειάζονται η παρακάτω βασικοί υπολογισμοί

$$n=8, \Sigma y=2.75+\dots+3.12=24.02, \Sigma y^2=2.75^2+\dots+3.12^2=72.8, \bar{y}=24.02/8=3,$$

$$\Sigma x=58.0+\dots+66.0=535, \Sigma x^2=58.0^2+\dots+66.0^2=35983.5, \bar{x}=66.88,$$

$$\Sigma xy=(2.75*58.0)+\dots+(3.12*66.0)=1615.3,$$

Τότε είναι

$$b=\{\Sigma xy-(\Sigma x \Sigma y)/n\} / (\Sigma x^2-(\Sigma x)^2/n) = 8.96 / 205.38 = 0.044.$$

$$a=\bar{y}-b*\bar{x}=3-0.044*66.88=0.086$$

Τότε η εξάρτηση του όγκου του πλάσματος από το βάρος του σώματος περιγράφεται από τη σχέση:

$$\{\text{όγκος πλάσματος (y)}\}=0.086+0.044*\{\text{βάρος σώματος (x)}\}$$

### Έλεγχος γραμμικής σχέσης

Για να ελέγξουμε ότι υπάρχει γραμμική σχέση μεταξύ  $y$  και  $x$ , δηλαδή ότι το  $b$  είναι σημαντικά διαφορετικό από το 0, αρκεί να ελέγξουμε την υπόθεση ότι  $r=0$ . Ο έλεγχος της υπόθεσης αυτής βασίζεται όπως είδαμε στην ποσότητα:

$$t = r \sqrt{(n-2)/(1-r^2)}$$

Στο παράδειγμά μας, είναι  $r=0.76$ ,  $n=6$  και  $t = 0.76 \sqrt{[6/(1-0.76^2)]} = 2.86$ , ποσότητα που είναι στατιστικά σημαντική ( $P<0.05$ ).

Το πόσο καλά η γραμμή παλινδρόμησης αντιπροσωπεύει τα δεδομένα δίνεται από το  $r^2$  ή  $R^2$ . Στο παράδειγμά μας είναι  $R^2=0.762=0.58$ , αυτό σημαίνει ότι 58% της διακύμανσης του όγκου πλάσματος εξηγείται από τη γραμμική του σχέση με το βάρος σώματος. Αν το  $R^2$  ήταν 100% τότε τα σημεία θα ήταν πάνω στη γραμμή παλινδρόμησης και συνεπώς ο όγκος πλάσματος δεν θα είχε καμία διακύμανση γύρω από τη γραμμή.

### Πρόβλεψη

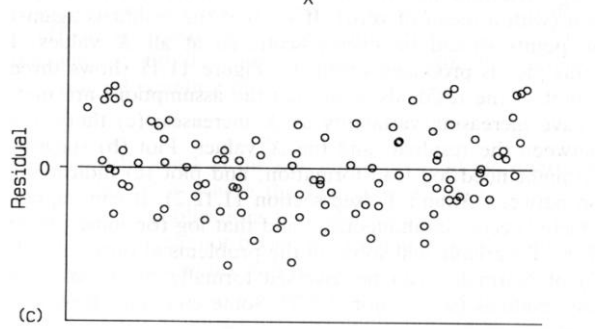
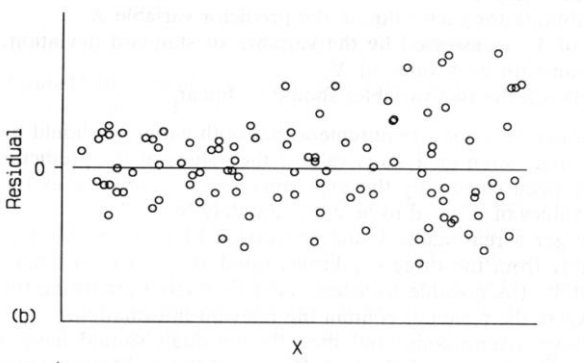
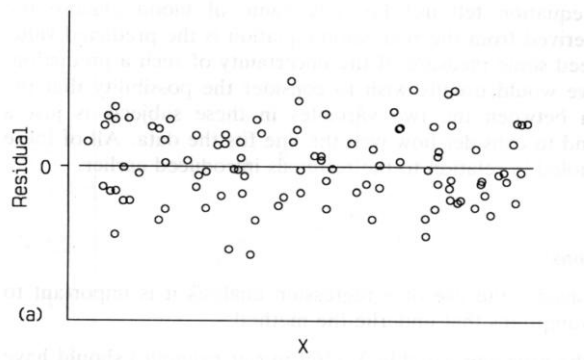
Η εξίσωση της γραμμικής παλινδρόμησης μπορεί να χρησιμοποιηθεί για να προβλέψουμε μία τιμή της εξαρτημένης μεταβλητής  $y$  για μία συγκεκριμένη τιμή της ανεξάρτητης μεταβλητής  $x$ .

*Παράδειγμα:* Για έναν άνδρα που ζυγίζει  $x=66\text{kg}$  ο προβλεπόμενος όγκος πλάσματος είναι

$$y'=a+b*x=0.086+0.044*66=2.96$$

### Έλεγχος εγκυρότητας

Η εγκυρότητα της γραμμής παλινδρόμησης εξετάζεται με την ανάλυση των υπολοίπων. Το υπόλοιπο ορίζεται ως την διαφορά της πραγματική τιμής μείον την αντίστοιχη προβλεπόμενη τιμή. Η ανάλυση των υπολοίπων περιλαμβάνει ένα διάγραμμα των  $x$  vs τα αντίστοιχα υπόλοιπα. Αν δεν υπάρχει κάποια συστηματική δομή των δεδομένων τότε η γραμμή θεωρείται έγκυρη. Έγκυρη θεωρείται η γραμμή παλινδρόμησης που αντιστοιχεί στο πρώτο από τα παρακάτω διαγράμματα:



### Πολλαπλή γραμμική παλινδρόμηση (multiple linear regression)

Η πολλαπλή γραμμική παλινδρόμηση είναι μία προέκταση της απλής γραμμικής παλινδρόμησης, στην οποία θέλουμε να συσχετίσουμε μία εξαρτημένη μεταβλητή  $y$  με περισσότερες από μία ανεξάρτητες μεταβλητές, π.χ.  $x_1$  και  $x_2$ .

Οι ανεξάρτητες μεταβλητές  $x_1$  και  $x_2$  αντιπροσωπεύουν παράγοντες που μπορεί να επηρεάζουν την εξαρτημένη μεταβλητή. Σκοπός της πολλαπλής γραμμικής παλινδρόμησης είναι να βρούμε την εξίσωση που περιγράφει τη σχέση μεταξύ της  $y$  και των  $x_1$  και  $x_2$ , δηλαδή

$$y' = b_0 + b_1 x_1 + b_2 x_2,$$

όπου  $y'$  είναι η προβλεπόμενη τιμή της  $y$ .

Οι συντελεστές  $b_0$ ,  $b_1$  και  $b_2$  εκτιμούνται με τη μέθοδο των ελαχίστων τετραγώνων και υπολογίζονται χρησιμοποιώντας στατιστικά προγράμματα. Ο  $b_0$  είναι ο αντίστοιχος συντελεστής της τομής  $a$  της γραμμής παλινδρόμησης και οι  $b_1$ ,  $b_2$  αντιστοιχούν στην κλίση  $b$  της γραμμής παλινδρόμησης. Η εξίσωση αυτή χρησιμοποιείται για να προβλέψουμε την τιμή της  $y$  για διάφορους συνδυασμούς των  $x_1$  και  $x_2$  ή για να ελέγξουμε πως η μεταβλητές  $x_1$  και  $x_2$  επηρεάζουν την  $y$ .

*Παράδειγμα:* Για να δούμε πως εξαρτάται ο δείκτης ισχύος των αναπνευστικών μυών  $PE_{\max}$  (cmH<sub>2</sub>O) από το ύψος (cm) και το βάρος του σώματος (kg) βρήκαμε την ακόλουθη εξίσωση πολλαπλής γραμμικής παλινδρόμησης:

$$PE_{\max} = 47.35 + 0.147 * \text{ύψος} + 1.024 * \text{βάρος}$$

Οι συντελεστές  $\beta_1=0.147$  και  $\beta_2=1.024$  υποδεικνύουν την προβλεπόμενη αύξηση στον  $PE_{\max}$  για κάθε μονάδα αύξησης στις ανεξάρτητες μεταβλητές, 1 cm ύψους και 1 kg βάρους. Η τιμή  $\beta_0=47.35$  είναι μία σταθερά που αντιστοιχεί στον  $PE_{\max}$  όταν το ύψος και το βάρος μαζί είναι μηδέν.

Επίσης, μπορούμε να αποκτήσουμε το τυπικό σφάλμα για τους συντελεστές  $b_1$  και  $b_2$  και να ελέγξουμε την στατιστική σημαντικότητά τους με το  $t$ -test:  $t=b_1/s.e.(b_1)$  και  $t=b_2/s.e.(b_2)$ , δηλαδή να ελέγξουμε την υπόθεση ότι είναι 0.



## Λογαριθμιστική παλινδρόμηση

Σε πολλές μελέτες το αποτέλεσμα που μας ενδιαφέρει είναι η παρουσία ή όχι κάποιας συνθήκης, π.χ. η ανταπόκριση σε μία θεραπεία, η ύπαρξη εμφράγματος του μυοκαρδίου ή η εμφάνιση καρκίνου, και πως αυτή συνθήκη επηρεάζεται από ορισμένες ανεξάρτητες μεταβλητές, π.χ. πώς η πιθανότητα εμφάνισης καρκίνου του πάγκρεας επηρεάζεται από την κατανάλωση καφέ και την κατανάλωση οινόπνευματος.

Σε αυτές τις περιπτώσεις δεν μπορούμε να χρησιμοποιήσουμε την συνηθισμένη πολλαπλή γραμμική παλινδρόμηση αλλά μπορούμε να χρησιμοποιήσουμε την λογαριθμιστική παλινδρόμηση. Η κύρια διαφορά της λογαριθμιστικής παλινδρόμησης είναι ότι εδώ το ρόλο της εξαρτημένης μεταβλητής τον έχει ένας μετασχηματισμός, ο  $\text{logit}$  μετασχηματισμός:

$$\text{logit}(p) = \log(p/(1-p))$$

όπου  $p$  είναι η πιθανότητα να παρουσιασθεί η συνθήκη.

*Παράδειγμα:* Σε μία μελέτη ασθενών-μαρτύρων για να ερευνήσουμε τη σχέση μεταξύ κατανάλωσης καφέ, κατανάλωσης οινόπνευματος και ηλικίας με την εμφάνιση καρκίνου του πάγκρεας εφαρμόσαμε την λογαριθμιστική παλινδρόμηση και βρέθηκε η παρακάτω εξίσωση:

$$\text{logit}(p = \text{Πιθανότητα καρκίνου}) = b_0 + b_1(\text{ηλικία}) + b_2(\text{καφές}) + b_3 * (\text{οινόπνευμα}) \text{ ή}$$

$$\text{logit}(p = \text{Πιθανότητα καρκίνου}) = 0.02 + 0.47 * (\text{ηλικία}) + 0.69 * (\text{καφές}) + 1.39 * (\text{οινόπνευμα}).$$

Οι συντελεστές  $b_1$ ,  $b_2$  και  $b_3$  μας υποδεικνύουν κατά πόσο οι μεταβλητές ηλικία, καφές και οινόπνευμα επηρεάζουν την πιθανότητα εμφάνισης καρκίνου, αντίστοιχα. Με βάση αυτούς τους συντελεστές μπορούμε να βρούμε τους αντίστοιχους λόγους αναλογιών (odds ratio). Για παράδειγμα, ο λόγος αναλογιών για τον καφέ είναι:

$$\text{OR}_{\text{coffee}} = e^{b_2} = e^{0.693} = 2.00.$$

***Οπότε, συμπεραίνουμε ότι υπάρχει αυξανόμενος κίνδυνος (διπλάσιος) για καρκίνο σε αυτούς που πίνουν καφέ σε σχέση με αυτούς που δεν πίνουν για άτομα με ίδια ηλικία που καταναλώνουν οινόπνευμα***

## Μη-παραμετρικοί έλεγχοι

Οι μη-παραμετρικοί έλεγχοι (tests) χρησιμοποιούνται όταν το μέγεθος των δειγμάτων είναι πολύ μικρό ( $<6$ ) ή όταν δεν μπορούμε να υποθέσουμε ότι τα δεδομένα ακολουθούν μία κανονική κατανομή και ότι τα δείγματα έχουν παρόμοια διακύμανση.

Οι μη-παραμετρικοί έλεγχοι δεν κάνουν καμία υπόθεση για την κατανομή των δεδομένων και είναι λιγότεροι ισχυροί από ότι οι αντίστοιχοι παραμετρικοί έλεγχοι (t-test, z-test, F-test,  $\chi^2$ -test). Οι έλεγχοι αυτοί βασίζονται στη τάξη των τιμών (ranks) των δειγμάτων, δηλ. δεν αναλύουν τα πραγματικά δεδομένα.

### Wilcoxon-test ζευγαρωτές παρατηρήσεις

Ο αντίστοιχος μη-παραμετρικός έλεγχος του t-test για ζευγαρωτές παρατηρήσεις Wilcoxon test για ζευγαρωτές παρατηρήσεις.

*Παράδειγμα:* Θέλουμε να συγκρίνουμε τη βιοδιαθεσιμότητα δύο φαρμάκων T και R με βάση την παράμετρο T<sub>max</sub> που είναι γνωστό ότι η κατανομή της δεν είναι κανονική. Στη μελέτη παίρνουν μέρος δώδεκα άτομα που λαμβάνουν πρώτα το T και μετά το R, οι τιμές του T<sub>max</sub> είναι:

Subject	T	R	T-R	T-R		Sign	T-R  rank	Rank ( R)
1	0.76	0.76	0.00	0.00			0.00	1
2	0.56	0.75	-0.19	0.19		Neg	0.18	2
3	0.78	0.55	0.23	0.23		Neg	0.19	3
4	0.75	1.25	-0.50	0.50		Neg	0.20	4,5
5	0.72	0.50	0.22	0.22			0.20	4,5
6	0.53	0.75	-0.22	0.22			0.22	6,5
7	0.50	0.70	-0.20	0.20		Neg	0.22	6,5
8	0.50	1.20	-0.70	0.70			0.23	8
9	1.00	0.55	0.45	0.45		Neg	0.25	9
10	1.20	1.00	0.20	0.20			0.45	10
11	0.75	1.00	-0.25	0.25		Neg	0.50	11
12	0.52	0.70	-0.18	0.18		Neg	0.70	12

Ελέγχουμε την υπόθεση ότι T και R έχουν την ίδια βιοδιαθεσιμότητα συγκρίνοντάς τα σύμφωνα με την τάξη των τιμών τους (ranks) και χρησιμοποιώντας το άθροισμα του λιγότερου εμφανιζόμενου πρόσημου  $\Sigma R_i^+$ .

Ο έλεγχος έχει τα ακόλουθα στάδια:

- 1) Υπολογίζουμε τις διαφορές μεταξύ T και R ( $T_i - R_i$ ) για κάθε άτομο i.
- 2) Παίρνουμε τις απόλυτες τιμές των διαφορών  $|T_i - R_i|$  και τις ταξινομούμε ( $R_i$ ) κατά αύξουσα σειρά.
- 3) Υπολογίζουμε το άθροισμα των ranks που σχετίζεται με το λιγότερο εμφανιζόμενο πρόσημο (θετικό ή αρνητικό,  $\Sigma R_i^+$  ή  $\Sigma R_i^-$ ).
- 4) Ορίζουμε τη στάθμη σημαντικότητας (P-value) συγκρίνοντας το άθροισμα με τις τιμές του Πίνακα 4. Αν η τιμή του  $\Sigma R_i^+$  ή  $\Sigma R_i^-$  είναι μικρότερη από τη τιμή του πίνακα με μέγεθος δείγματος n τότε T και R είναι διαφορετικά στην υποδεικνυόμενη από τον πίνακα στάθμη σημαντικότητας (π.χ.  $P=0.05$ ).

Το άθροισμα του λιγότερου εμφανιζόμενου πρόσημου είναι:  $\Sigma R_i^+ = 1+4.5+6.5+8+10=30$

Η τιμή 30 για  $n=12$  συγκρίνεται με τις τιμές του Πίνακα 4.. Οπότε δεν υπάρχει διαφορά μεταξύ T και R επειδή  $P>0.05$ .

### **Mann-Whitney-Wilcoxon test για τη σύγκριση δύο ανεξάρτητων δειγμάτων**

Ο αντίστοιχος μη-παραμετρικός έλεγχος του t-test για να συγκρίνουμε δύο ανεξάρτητα δείγματα είναι το Mann-Whitney-Wilcoxon test.

*Παράδειγμα:* Η κατανάλωση ενέργειας μέσα σε 24 ώρες μίας ομάδας λεπτών γυναικών και μίας ομάδας παχύσαρκων γυναικών δίνεται παρακάτω. Υπάρχει διαφορά στην κατανάλωση ενέργειας μεταξύ των δύο ομάδων ;

<u>Λεπτές (n=13)</u>	<u>παχύσαρκες (n=9)</u>
6.13	8.79
7.05	9.19
7.48	9.21
7.48	9.68
7.53	9.69
7.58	9.97
7.90	11.51
8.08	11.85
8.09	12.79
8.11	
8.40	
10.15	
10.88	

Στο Mann-Whitney-Wilcoxon test βρίσκουμε τις τάξεις (ranks) όλων των δεδομένων θεωρώντας ότι είναι ένα δείγμα. Κατόπιν βρίσκουμε το άθροισμα T των τάξεων τις πιο μικρής ομάδας (παχύσαρκες).

Λεπτές rank	ενέργεια	παχύσαρκες ενέργεια	rank
-----			
1	6.13		
2	7.05		
3.5	7.48		
3.5	7.48		
5	7.53		
6	7.58		
7	7.90		
8	8.08		
9	8.09		
10	8.11		
11	8.40		
		8.79	12
		9.19	13
		9.21	14
		9.68	15
		9.69	16
		9.97	17
18	10.15		
19	10.88		
		11.51	20
		11.85	21
		12.79	22
-----			
T=103		T=150	

Ελέγχουμε την υπόθεση ότι οι δύο ομάδες δε διαφέρουν μεταξύ τους χρησιμοποιώντας τη μικρότερη ποσότητα απο τα δύο T, δηλ. T=150.

Αυτή η ποσότητα T συγκρίνεται με τις τιμές του πίνακα 5 για  $n_1=9$  και  $n_2=13$  (η τιμή του πίνακα είναι ίδια όταν  $n_1=13$  και  $n_2=9$ ). Επειδή η τιμή 150 βρίσκεται εκτός του διαστήματος 65-142 για  $P=0.01$  αλλά εντός του διαστήματος 56-151 για  $P=0.001$  συμπεραίνουμε ότι η υπόθεση της ισότητας των δύο ομάδων απορρίπτεται σε  $P<0.01$ .

## Ανάλυση επιβίωσης

Στις μελέτες που ερευνούν χρόνια νοσήματα που είναι ενδεχομένως και θανατηφόρα, π.χ. καρκίνος, συνήθως τα δεδομένα που καταγράφονται είναι ο χρόνος από την εισαγωγή ενός ασθενή στη μελέτη μέχρι κάποιο συγκεκριμένο ανεπιθύμητο αποτέλεσμα, π.χ. από την στιγμή που έγινε μία συγκεκριμένη χειρουργική επέμβαση μέχρι το θάνατο του ασθενή. Τότε πρέπει να ορίζεται με ακρίβεια τόσο η αρχή της μέτρησης του χρόνου (π.χ. επέμβαση ή ημερομηνία διάγνωσης) όσο το τελικό αποτέλεσμα (π.χ. θάνατος ή υποτροπή).

Στις μελέτες αυτές, το αποτέλεσμα που εξετάζουμε συνήθως δε θα πραγματοποιηθεί για όλους τους ασθενείς. Για παράδειγμα, αν μελετάται ο χρόνος μέχρι το θάνατο σε ασθενείς με καρκίνο που υποβλήθηκαν σε μία θεραπεία, είναι δυνατόν σε μερικά άτομα να μην υπάρχει το αποτέλεσμα του θανάτου κατά το διάστημα παρακολούθησης ή κατά τη διάρκεια της μελέτης. Συνεπώς ο χρόνος μέχρι το θάνατο δεν είναι δυνατό να καθορισθεί σε όλους τους ασθενείς και οι ασθενείς αυτοί ονομάζονται αποκομμένοι (censored).

Οπότε στην ανάλυση επιβίωσης για κάθε ασθενή καταγράφουμε τα εξής:

- 1) αν είναι σε ζωή ή θάνατο
- 2) το χρόνο επιβίωσης: το χρόνο από την εισαγωγή στη μελέτη μέχρι το θάνατο ή το χρόνο από την εισαγωγή στη μελέτη μέχρι πότε ήταν γνωστό ότι ήταν ζωντανός.

Η ανάλυση επιβίωσης συνεκτιμά τόσο τη συχνότητα όσο και το χρόνο επέλευσης του θανάτου ή κάποιου άλλου ανεπιθύμητου γεγονότος (π.χ. υποτροπή). Σκοπός της είναι να εκτιμήσει την πιθανότητα ή το ποσοστό επιβίωσης σε ένα συγκεκριμένο χρονικό διάστημα.

*Παράδειγμα:* Σε 11 ασθενείς με ΣΕΛ χορηγήθηκε κυκλοφωσφαμίδα και παρακολούθηθηκαν για 120 μέρες και ένας αριθμός από αυτούς υποτροπίασαν. Δύο ασθενείς αποχώρησαν από τη μελέτη, ο ένας μετά 50 μέρες και ο άλλος μετά από 66 μέρες (αποκομμένοι). Τα δεδομένα επιβίωσης τους, δηλ. ο χρόνος σε μήνες από την έναρξη της θεραπείας μέχρι το ανεπιθύμητο αποτέλεσμα-υποτροπή, ήταν:

30, 50, 50\*, 51, 66\*, 82, 92, 120\*, 120\*, 120\*, 120\* μέρες,

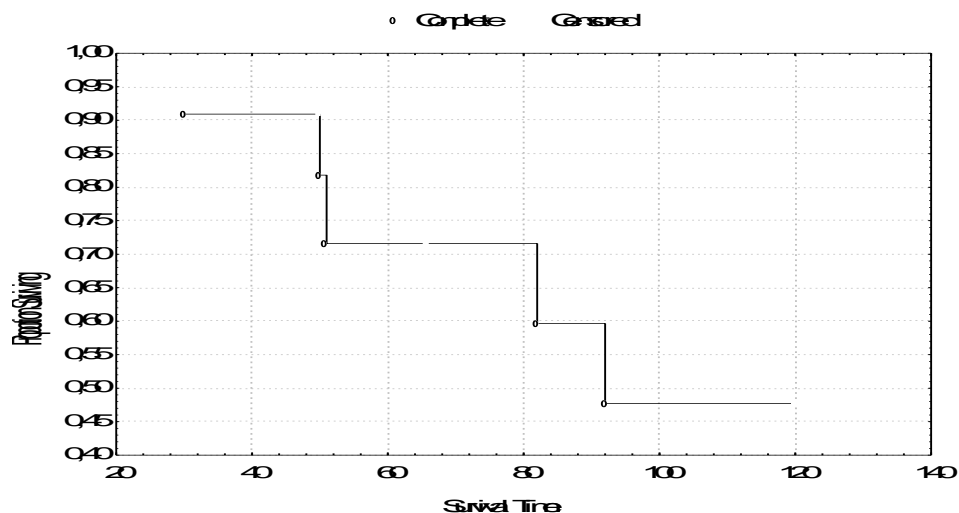
όπου \* σημαίνει ότι δεν υποτροπίασαν μέχρι το τέλος της μελέτης (censored).

Η πιθανότητα επιβίωσης (μη-υποτροπής) για κάθε χρονική στιγμή υπολογίζεται σύμφωνα με τον παρακάτω πίνακα:

patients k	Time in Days	Time Interval in days	Number at risk ( $r_k$ )	Failures ( $f_k$ )	$Q =$ ( $r_k - f_k$ ) / $r_k$	Survival proport. $p_k =$ $P_{k-1} * Q$	s.e.( $p_k$ )	95% c.i.
1	30	(1-29)	11	1	0.91	0.91	0.08	(0.75, 1.00)
2	50	(1-49)	10	1	0.90	0.82	0.11	(0.60, 1.00)
3	50*							
4	51	(1-50)	8	1	0.88	0.72	0.13	(0.45, 0.98)
5	66*							
6	82	(1-81)	6	1	0.83	0.60	0.15	(0.29, 0.90)
7	92	(1-91)	5	1	0.80	0.48	0.15	(0.18, 0.78)
8	120*							
9	120*							
10	120*							
11	120*							

Όπου  $s.e.(p_k) = p_k \sqrt{\{(1-p_k)/r_{p_k}\}}$ , και το 95% δ.ε. είναι  $(p_k - 1.96 * s.e.(p_k), p_k + 1.96 * s.e.(p_k))$ .

Ο άνω πίνακας παρουσιάζεται με μία καμπύλη που λέγεται καμπύλη επιβίωσης ή Kaplan-Meier καμπύλη. Η καμπύλη αυτή σχεδιάζεται σαν τμηματική συνάρτηση (δηλ. όχι λεία καμπύλη): το ποσοστό επιβίωσης παραμένει αμετάβλητο μεταξύ των υποτροπών ακόμη και αν υπάρχουν στο ενδιάμεσο κάποιες αποκομμένες παρατηρήσεις.



Η πιθανότητα μη-υποτροπής (επιβίωσης) σε 90 μέρες είναι 0.60 με 95% δ.ε. (0.29, 0.90)

Αν θέλουμε να συγκρίνουμε δύο καμπύλες επιβίωσης, π.χ. η μία να αντιστοιχεί στην κυκλοφωσφαμίδη και η άλλη στην ομάδα ελέγχου, τότε η υπόθεση ότι οι δύο ομάδες δε διαφέρουν στην πιθανότητα επιβίωσης ελέγχεται με το logrank test που είναι ένας μη-παραμετρικός έλεγχος.

## **Κλινικές μελέτες**

Στην ιατρική πράξη, κάθε δραστηριότητα που σκοπεύει στην αναστολή μιας νοσηρής διαδικασίας ή στην εξάλειψη των συμπτωμάτων ενός νοσήματος λέγεται θεραπεία. Η θεραπεία πραγματοποιείται με θεραπευτικά μέτρα που μπορεί να είναι φαρμακευτικά, ή χειρουργικά.

Η αξιολόγηση της αποτελεσματικότητας των θεραπευτικών μέτρων γίνεται σε ομάδες ασθενών με τυχαιοποιημένες και ελεγχόμενες έρευνες θεραπευτικής παρεμβάσης ή κλινικές μελέτες (clinical trials).

Ένα θεραπευτικό μέτρο, π.χ. φάρμακο, θεωρείται αποτελεσματικό αν μπορεί να επηρεάσει ευνοϊκά την πρόγνωση ενός νοσήματος.

Η σχετική (ή απόλυτη) αποτελεσματικότητα ενός θεραπευτικού μέτρου ορίζεται ως το ποσοστό ασθενών που επιβιώνουν ή θεραπεύονται με αυτό το μέτρο δια (ή μείον) το ποσοστό των ασθενών που επιβιώνουν ή βελτιώνονται με μη εφαρμογή του μέτρου.

## **Φάσεις κλινικών μελετών για φάρμακα**

Οι κλινικές μελέτες για φάρμακα ταξινομούνται σε 4 διαδοχικές κατηγορίες:

*Φάση I:* Αφορά ασφάλεια φαρμάκων και σκοπός της είναι να ορισθεί η αποδεκτή δόση ενός φαρμάκου. Συνήθως περιλαμβάνει 20-80 ασθενείς.

*Φάση II:* Είναι μικρής κλίμακας έρευνα για την αποτελεσματικότητα και την ασφάλεια. Μερικές φορές είναι και η διαδικασία συλλογής των λιγότερο τοξικών φαρμάκων που θα προχωρήσουν στην επόμενη φάση. Συνήθως περιλαμβάνει 100-200 ασθενείς σε κάθε φάρμακο.

*Φάση III:* Είναι μεγάλης κλίμακας έρευνα που ουσιαστικά συγκρίνει ένα φάρμακο με ένα τυπικό θεραπευτικό μέτρο με σκοπό να αποδείξει την αποτελεσματικότητα του φαρμάκου. Εδώ βασίζονται οι οργανισμοί για να εγκρίνουν την κυκλοφορία του φαρμάκου.

*Φάση IV:* Παρακολούθηση του φαρμάκου όταν ήδη κυκλοφορεί στην αγορά.

## Οργάνωση μιας κλινικής μελέτης - Πρωτόκολλο

Η οργάνωση μιας κλινικής μελέτης βασίζεται σε ένα ερευνητικό πρωτόκολλο: ένα κείμενο που προσδιορίζει πως θα διεξαχθεί και θα οργανωθεί η μελέτη.

Μερικά από τα κύρια χαρακτηριστικά ενός πρωτοκόλλου και της οργάνωσης μιας μελέτης είναι τα ακόλουθα:

- ◆ Εισαγωγή και περιγραφή του γενικού σκοπού της μελέτης
- ◆ Περιγραφή συγκεκριμένων αντικειμενικών σκοπών
- ◆ Καθορισμός της θεραπείας που θα αξιολογηθεί
- ◆ Επιλογή των ασθενών
- ◆ Εγγραφή και κατανομή των ασθενών στις συγκρινόμενες ομάδες
- ◆ Αξιολόγηση της ανταπόκρισης των ασθενών
  - ◆ *Αξιολόγηση στην έναρξη*
  - ◆ *Θεραπευτικό αποτέλεσμα*
  - ◆ *Παρενέργειες ή ανεπιθύμητα αποτελέσματα:*
- ◆ Σχεδιασμός της μελέτης:
  - ◆ *Επιλογή του συγκριτικού δείγματος*
  - ◆ *Τυφλός έλεγχος (blindness)*
  - ◆ *Σχήματα κατανομής των ασθενών-μέθοδοι & σχεδιασμοί*
- ◆ Παρακολούθηση των ασθενών
- ◆ Αναγκαίος αριθμός ασθενών
- ◆ Τρόπος στατιστικής ανάλυσης



### **Εισαγωγή και περιγραφή του γενικού σκοπού της μελέτης**

Εδώ εξηγείται γιατί η μελέτη αξίζει να γίνει και με ποιο τρόπο βασίσθηκε στην εμπειρία προηγούμενων ερευνών.

### **Περιγραφή συγκεκριμένων αντικειμενικών σκοπών**

Εδώ περιγράφονται οι σκοποί πιο συγκεκριμένα και αναφέρονται ακριβώς οι υποθέσεις σχετικά με την αποτελεσματικότητα και την ασφάλεια του θεραπευτικού μέτρου που θα εξετασθούν από τη μελέτη.

### **Καθορισμός της θεραπείας που θα αξιολογηθεί**

Τα χαρακτηριστικά και ο τρόπος θεραπείας που θα δοκιμασθεί πρέπει να καθορισθούν προσεκτικά, να καταγραφούν στο πρωτόκολλο και να τηρηθούν με συνέπεια.

Για θεραπείες με φάρμακα τα ακόλουθα στοιχεία πρέπει να περιγραφούν στο πρωτόκολλο: 1) μορφή του φαρμάκου, 2) οδός χορήγησης, 3) δοσολογία, 4) διάρκεια θεραπείας, 5) αντιμετώπιση παρενεργειών, αλλαγή δοσολογίας, 6) αντιμετώπιση μη-συμμόρφωσης με τη θεραπεία, 7) συσκευασία των φαρμάκων

### **Επιλογή των ασθενών**

Εδώ περιγράφεται ο καθορισμός των κριτηρίων και της διαδικασίας επιλογής των ασθενών.

Το δείγμα των ασθενών πρέπει να είναι αντιπροσωπευτικό της νόσου.

Το κριτήριο επιλογής ασθενών σε μία μελέτη για ήπια υπέρταση που έγινε από το MRC ήταν:

- ♦ 90 < mean diastolic BP of four readings < 110 mm Hg and
- ♦ mean systolic BP < 200 mm Hg

### **Εγγραφή και κατανομή των ασθενών στις συγκρινόμενες ομάδες**

Η κατανομή των ασθενών στις δύο ομάδες πρέπει να γίνει με τρόπο τυχαίο. Η τυχειότητα εξουδετερώνει ενδεχόμενους συγχυτικούς παράγοντες, κατανέμοντάς τους ισομερώς στις συγκρινόμενες ομάδες. Αν είναι αδύνατη η τυχειοποίηση, η συγκρισιμότητα των δύο ομάδων θα γίνει μέσα από άλλου τύπου σχεδιασμό ή μέσα από τη στατιστική ανάλυση.

### **Αξιολόγηση της ανταπόκρισης των ασθενών**

Οι μέθοδοι που αξιολογούν και καταγράφουν την εξέλιξη ενός ασθενή πρέπει να ορίζονται με ακρίβεια στο πρωτόκολλο της μελέτης. Η αξιολόγηση των ασθενών διακρίνονται σε τρεις κατηγορίες:

*Αξιολόγηση στην έναρξη (baseline):* Καταγράφεται η αρχική κλινική εικόνα του κάθε ασθενή μαζί με τα προσωπικά του χαρακτηριστικά και το κλινικό ιστορικό του. Αυτά είναι χρήσιμα σε περίπτωση που μερικά μπορούν να επηρεάσουν την ανταπόκριση του ασθενή στο θεραπευτικό μέτρο (π.χ. πολλαπλές μετρήσεις της πίεσης του αίματος ενός ασθενή πριν την εισαγωγή του σε μελέτη για αντιυπερτασικό φάρμακο).

*Θεραπευτικό αποτέλεσμα:* Το θεραπευτικό αποτέλεσμα πρέπει να καθορίζεται στους συγκεκριμένους αντικειμενικούς σκοπούς. Εκφράζεται συνήθως με βάση τη συχνότητα και το χρόνο εμφάνισης κάποιου ανεπιθύμητου γεγονότος (π.χ. υποτροπή) Η αξιολόγηση του δοκιμαζόμενου θεραπευτικού μέτρου βασίζεται συνήθως σε περισσότερα από ένα θεραπευτικά αποτελέσματα (π.χ. βελτίωση κλινική εικόνας και μείωση εργαστηριακών δεικτών) τότε πρέπει να αποφασισθεί ποιο από τα επιδιωκόμενα θεραπευτικά αποτελέσματα είναι πρωτεύον και ποιά είναι δευτερεύοντα. Η μελέτη πρέπει να επικεντρωθεί στο πρωτεύον.

*Παρενέργειες ή ανεπιθύμητα αποτελέσματα:* Σε κάθε μελέτη είναι απαραίτητο να εκτιμήσουμε εκτός από την αποτελεσματικότητα και την ασφάλεια. Τα θεραπευτικά μέτρα συγκρίνονται και ως προς τις παρενέργειες. Τότε πρέπει να ετοιμαστεί ένας κατάλογος με τις πιθανές παρενέργειες για τον κάθε ασθενή.

## **Σχεδιασμός της μελέτης**

*Επιλογή του συγκριτικού δείγματος:* Στην ομάδα ελέγχου εφαρμόζεται κάποιο θεραπευτικό μέτρο με τεκμηριωμένη αποτελεσματικότητα. Αυτό γίνεται για ηθικούς λόγους και γιατί το δοκιμαζόμενο νέο θεραπευτικό μέτρο πρέπει να κριθεί καλύτερο από τα ήδη υπάρχοντα. Αν δεν υπάρχει άλλο θεραπευτικό μέτρο με τεκμηριωμένη αποτελεσματικότητα τότε στα άτομα της ομάδας ελέγχου θα πρέπει να εφαρμοσθεί κάποιο ανενεργό μέτρο (placebo).

*Τυφλός έλεγχος (blindness):* Για να αποφευχθούν οι συνέπειες από τις υποκειμενικές προκαταλήψεις ή αντιδράσεις (αυθυποβολή) των ασθενών ή των ερευνητών οι έρευνες γίνονται κατά κανόνα με τρόπο τυφλό ή διπλά τυφλό. Με τον διπλά τυφλό τρόπο ούτε ο ασθενής, ούτε ο γιατρός γνωρίζουν ποιος πήρε το νέο φάρμακο. Με τον τυφλό τρόπο ο ασθενής δεν γνωρίζει εάν πήρε το νέο φάρμακο.

*Σχήματα κατανομής των ασθενών:* παράλληλος, διαστρωματικός, διαδοχικός, πολυπαραγοντικός, αυτοσύγκρισης.

## **Παρακολούθηση των ασθενών**

Είναι ανάγκη να περιορίζεται ο αριθμός των ασθενών που διακόπτουν τη συνεργασία τους (withdrawals) ή αδυνατούν να παρακολουθήσουν τη μελέτη. Ο αριθμός των απωλειών πρέπει να περιορίζεται στο 10% του συνόλου των ασθενών που περιλήφθηκαν αρχικά στην έρευνα. Η διακοπή είναι συνηθέστερη όταν η θεραπεία δεν είναι αποτελεσματική. Όταν το ποσοστό των ασθενών που διακόπτουν είναι μεγάλο τα ευρήματα της έρευνας οδηγούν σε πλασματική υπερεκτίμηση της αποτελεσματικότητας των δύο συγκρινόμενων μέτρων.

### Αναγκαίος αριθμός ασθενών σε μία κλινική μελέτη

Προσδιορίζεται από στατιστικές μεθόδους. Πρακτικά θέματα όπως: η διαθεσιμότητα των ασθενών, οι οικονομικοί πόροι και θέματα ηθικής (ο ασθενής να μη λάβει ένα λιγότερο αποτελεσματικό φάρμακο) πρέπει επίσης να ληφθούν υπόψη.

*Παράδειγμα:* Στο πρωτόκολλο της μελέτης για τη σύγκριση του anturan με το placebo για να προσδιορισθεί το μέγεθος του δείγματος έλαβαν υπόψη τα ακόλουθα:

(1) τον κύριο σκοπό της μελέτης: να ελέγξουν αν το anturan αξίζει για να εμποδίσει την θνητότητα μετά από έμφραγμα το μυοκαρδίου. Η περαιτέρω προφύλαξη από μη-θανατηφόρο έμφραγμα ήταν σχετικό αλλά όχι ο κύριος σκοπός.

(2) Την κύρια μέτρηση του θεραπευτικού αποτελέσματος στον κάθε ασθενή: Ο θάνατος από οποιαδήποτε αιτία μέσα σε ένα χρόνο από την πρώτη χορήγηση της θεραπείας θεωρήθηκε κύρια ένδειξη αποτυχίας του θεραπευτικού μέτρου.

(3) Τον τρόπο ανάλυσης των δεδομένων ώστε να ανιχνεύσουν διαφορές μεταξύ των θεραπευτικών μέτρων: Η ανάλυση συμπεριλάμβανε την σύγκριση των ποσοστών των ασθενών που πέθαναν μέσα σε ένα χρόνο με anturan και placebo. Χρησιμοποιήθηκε το chi-squared test και η διαφορά ελέγχθηκε σε στάθμη σημαντικότητας 5%.

(4) Το αναμενόμενο αποτέλεσμα με ένα τυπικό θεραπευτικό μέτρο: Με τη χρήση του placebo αναμένεται να πεθάνει το 10% των ασθενών.

(5) Τη διαφορά των θεραπευτικών μέτρων που πρέπει να ανιχνευθεί και το βαθμό βεβαιότητας ότι υπάρχει αυτή η διαφορά: Αποφασίσθηκε ότι είναι σημαντικό το anturan να ελαττώσει κατά το ήμισυ την διαφορά σε θνητότητα (5% κάθε χρόνο) και έπρεπε να υπάρχει 90% βεβαιότητα ότι η διαφορά αυτή είναι στατιστικά σημαντική.

### Ποιοτικό θεραπευτικό αποτέλεσμα

Το θεραπευτικό αποτέλεσμα σε κάθε ασθενή ήταν επιτυχία (επιβίωση) ή αποτυχία (θάνατος σε ένα χρόνο). Σύμφωνα με τα προηγούμενα πρέπει να διαλέξουμε 4 πράγματα:

$p_1$  = ποσοστό των επιτυχιών που αναμένεται από ένα τυπικό θεραπευτικό μέτρο.

$p_2$  = ποσοστό των επιτυχιών στο εξεταζόμενο θεραπευτικό μέτρο που θέλουμε να δείξουμε ότι είναι διαφορετικό από το  $p_1$ .

$\alpha$  = τη στάθμη σημαντικότητας του chi-squared test που χρησιμοποιήθηκε για να ανιχνεύσουμε διαφορά μεταξύ των θεραπευτικών μέτρων

$1-\beta$  = το βαθμό βεβαιότητας ότι αν υπάρχει πράγματι η διαφορά  $p_1 - p_2$  τότε θα ανιχνευθεί.

Στη μελέτη για το anturan επέλεξαν τα ακόλουθα:

$p_1 = 90\%$  (αναμένονταν να επιβιώσουν ένα χρόνο με το placebo)

$p_2 = 95\%$

$\alpha = 0.05$

$\beta = 0.1$

Τότε ο απαιτούμενος αριθμός ασθενών σε κάθε θεραπευτικό μέτρο δίνεται από τον ακόλουθο τύπο:

$$n = [p_1 \times (1 - p_1) + p_2 \times (1 - p_2)] \times (z_\alpha + z_\beta)^2 / (p_2 - p_1)^2$$

$z_\alpha$  και  $z_\beta$  είναι τα  $\alpha$  και  $\beta$  σημεία της SND, δηλ.  $z_{0.05}=1.96$ ,  $z_{0.1}=2.32$  και  $(z_\alpha + z_\beta)^2 = 18.3$ .

$$n = [0.90 \times (1 - 0.90) + 0.95 \times (1 - 0.95)] \times 18.3 / (0.95 - 0.90)^2 = 1006$$

απαιτούνται για κάθε θεραπευτικό μέτρο.

Αν η μελέτη είχε γίνει με μικρότερο αριθμό ασθενών από ότι υποδεικνύουν αυτοί οι υπολογισμοί τότε ελαττώνεται η δύναμη να ανιχνεύσουμε σημαντικές διαφορές μεταξύ των θεραπευτικών μέτρων και ο κίνδυνος για λανθασμένα αρνητικό συμπέρασμα αυξάνεται.

### Ποσοτικό θεραπευτικό αποτέλεσμα

Σε μία μελέτη για να αξιολογήσουμε την αποτελεσματικότητα του συμπληρώματος βιταμίνης D σε εγκύους στην προφύλαξη από νεογνική υπασβεσταιμία ένας αριθμός εγκύων πρέπει να κατανεμηθεί τυχαία σε δύο ομάδες: με βιταμίνη D και placebo. Το θεραπευτικό αποτέλεσμα μετράται με το επίπεδο ασβεστίου στο αίμα του νεογνού μία εβδομάδα μετά τη γέννηση.

Πρέπει να ορίσουμε τα ακόλουθα:

Τη μέση τιμή  $\mu_1$  ασβεστίου και την τυπική απόκλιση  $\sigma$  που αναμένεται από ένα τυπικό θεραπευτικό μέτρο.

Τη διαφορά μεταξύ των δύο θεραπευτικών μέτρων  $\delta = \mu_2 - \mu_1$  που είναι σημαντική για να ανιχνεύσουμε χρησιμοποιώντας t-test σε στάθμη σημαντικότητας  $\alpha$ .

Το βαθμό βεβαιότητας  $(1 - \beta)$  για να ανιχνεύσουμε μία τέτοια διαφορά αν πράγματι υπάρχει.

Από προηγούμενες εξετάσεις σε γυναίκες υποθέτουμε ότι είναι  $\mu_1 = 9.0$  και  $\sigma = 1.8\text{mg}/100\text{ml}$  για το ασβέστιο στο αίμα.

Θεωρείται ότι μία αύξηση σε  $9.5\text{mg}/100\text{ml}$  είναι κλινικά σημαντική, δηλ  $\delta = \mu_2 - \mu_1 = 9.5 - 9.0 = 0.5$ .  
 $\alpha = 0.05$  και  $\beta = 0.1$

Τότε ο απαιτούμενος αριθμός ασθενών είναι:

$$N = 2\sigma^2 \times (z_\alpha + z_\beta)^2 / (\mu_2 - \mu_1)^2$$

$$N = 2 \times 1.8^2 \times 18.3 / 0.5^2 = 474 \text{ σε κάθε θεραπευτικό μέτρο}$$

### Πιλοτική μελέτη

Αν δεν υπάρχουν προηγούμενες μελέτες ή ενδείξεις για τις τιμές των  $\mu_1$ ,  $\mu_2$  και  $\sigma$  τότε συνήθως γίνεται μία πιλοτική μελέτη σε περιορισμένο αριθμό ασθενών ώστε να εκτιμηθούν οι ποσότητες αυτές.

## Σχεδιασμοί κλινικών μελετών

Εκτός από τις κλασικές συγκριτικές θεραπευτικές έρευνες στις οποίες κάθε ασθενής κατανέμεται τυχαία σε ένα από τα συγκρινόμενα μέτρα (σχεδιασμός παράλληλων ομάδων) και στο τέλος της μελέτης τα θεραπευτικά αποτελέσματα αξιολογούνται με απλές στατιστικές αναλύσεις, υπάρχουν και άλλοι πολυπλοκότεροι σχεδιασμοί:

- ◆ Διαστρωματικός
- ◆ Πολυπαραγοντικός
- ◆ Αυτοσύγκρισης
- ◆ Διαδοχικές
- ◆ Πολυκεντρικές

### Διαστρωματικός σχεδιασμός

Όταν η πρόγνωση της νόσου που μελετάται επηρεάζεται από προγνωστικούς παράγοντες (π.χ. φύλο, ηλικία, στάδιο νόσου, ιστολογική εικόνα) τότε η τυχαιοποίηση των ασθενών γίνεται χωριστά σε κάθε προγνωστική υποκατηγορία, ώστε να υπάρχει εξομοίωση των συγκρινόμενων ομάδων. Αν όμως το δείγμα είναι πολύ μεγάλο τα αποτελέσματα θα είναι αδιάβλητα.

### Πολυπαραγοντικός σχεδιασμός

Ο πολυπαραγοντικός σχεδιασμός επιτρέπει τον ταυτόχρονο έλεγχο της αποτελεσματικότητας πολλών θεραπευτικών μέτρων με τον ίδιο αριθμό ασθενών και είναι ο μόνος σχεδιασμός που επιτρέπει τον έλεγχο για αλληλεπίδραση μεταξύ των θεραπευτικών μέτρων A και B.

		A		
		+	-	
B	+	(1)	(3)	(1)+(3)
	-	(2)	(4)	(2)+(4)
-----				
		(1)+(2)	(3)+(4)	

Αν  $(1)-(2) > (3)-(4)$  τότε υπάρχει συνέργεια μεταξύ Α και Β: η δράση του Β ισχυροποιείται από την παρουσία του Α.

### Σχεδιασμός αυτοσύγκρισης

Όταν δύο ή περισσότερα θεραπευτικά μέτρα (φάρμακα) εφαρμόζονται διαδοχικά στους ίδιους ασθενείς. Με το σχεδιασμό αυτό μειώνεται ο αριθμός των ατόμων σε σχέση με άλλους σχεδιασμούς πετυχαίνοντας ταυτόχρονα αύξηση της ισχύος των στατιστικών ελέγχων. Μεταξύ των περιόδων χορήγησης των φαρμάκων πρέπει να υπάρχει μία περίοδος απόπλυσης. Χρησιμοποιείται συνήθως γαι τον έλεγχο της βιοϊσοδυναμίας φαρμάκων.

### Διαδοχικές κλινικές μελέτες

Στο σχεδιασμό αυτό τα μέτρα (φάρμακα) εφαρμόζονται ανά ένα σε διαδοχικά ζεύγη ασθενών και γίνεται αξιολόγηση της επίδρασης των μέτρων σε κάθε ζεύγος ασθενών. Στις διαδοχικές μελέτες δεν υπάρχει προκαθορισμένος αριθμός ασθενών που πρέπει να μελετηθούν, τα θεραπευτικά αποτελέσματα των συγκρινόμενων μέτρων αξιολογούνται συνεχώς και η έρευνα διακόπτεται αμέσως μόλις τεκμηριωθεί η υπεροχή του ενός από τα συγκρινόμενα μέτρα ή διαπιστωθεί ότι δεν υπάρχει ουσιαστική διαφορά αποτελεσματικότητας μεταξύ τους. Δεν είναι χρήσιμος για νοσήματα με μεγάλη χρονική διάρκεια.

### Πολυκεντρικές κλινικές μελέτες

Η εκτέλεση μιας μελέτης γίνεται συνήθως από πολλά κέντρα ώστε να αποφεύγεται η διαμόρφωση μιας προκατάληψης υπέρ ή κατά ενός θεραπευτικού μέτρου (πηγή συστηματικού σφάλματος) και να είναι πιο σύντομη η έρευνα αφού ο απαιτούμενος αριθμός ασθενών μπορεί να συμπληρωθεί γρηγορότερα.



## Στατιστική ανάλυση μιας κλινικής μελέτης

Η μέθοδος ανάλυσης των αποτελεσμάτων μίας μελέτης εξαρτάται από τον σχεδιασμό της μελέτης και από το αν το θεραπευτικό αποτέλεσμα είναι ποιοτικό ή ποσοτικό.

### Ποιοτικό θεραπευτικό αποτέλεσμα

Η αξιολόγηση της αποτελεσματικότητας των θεραπευτικών μέτρων γίνεται με βάση το δείκτη θνητότητας ή κάποιο άλλο δείκτη προσβολής που εκφράζει τη συχνότητα μιας ανεπιθύμητης κατάληξης χωρίς αναφορά στο χρόνο του γεγονότος.

Η σύγκριση των δεικτών γίνεται με ένα chi-squared test. Όταν οι συγκρινόμενες ομάδες διαφέρουν ως προς τη κατανομή προγνωστικών-συγχυτικών παραγόντων η σύγκριση γίνεται με τη μέθοδο των Mantel-Haenszel  $\chi^2$ MH.

Τα αποτελέσματα μίας μελέτης για να συγκρίνουμε τον επιπολασμό των αντισωμάτων κατά της λεπτόσπειρας σε αγροτικές και αστικές περιοχές μίας χώρας ήταν:

Antibodies	rural	urban	total
Yes	60 (30%)	60 (30%)	120
No	140	140	280
Total	200	200	400

Το φύλο όμως θεωρείται συγχυτικός παράγοντας που διαφοροποιεί τον επιπολασμό:

#### Άνδρες

Antibodies	rural	urban	total
Yes	a=36 (72%)	e=50 (50%)	e=86
No	f=14	f=50	f=64
Total	g=50	h=100	n=150

$$\chi^2 = 5.73 \text{ (P<0.05)}$$

#### Γυναίκες

Antibodies	rural	urban	total
Yes	a=24 (16%)	e=10 (10%)	e=34
No	f=126	f=90	f=216
Total	g=150	h=100	n=250

$$\chi^2 = 1.36 \text{ (P>0.05)}$$

$$\chi^2_{MH} = (\Sigma a - \Sigma E)^2 / \Sigma V, \text{ όπου } E = eg/n \text{ και } V = efgh/[n^2(n-1)].$$

$$\Sigma a = 36 + 24 = 60,$$

$$\Sigma E = [86 \times 50 / 150] + [34 \times 150 / 250] = 49.1$$

$$\Sigma V = [86 \times 64 \times 50 \times 100 / 150^2 (150 - 1)] + [34 \times 216 \times 150 \times 100 / 250^2 (250 - 1)] = 15.29$$

$$\chi^2_{MH} = 7.08 \text{ (} P < 0.05 \text{), (η } \chi^2_{MH} \text{ ακολουθεί την } \chi^2 \text{ –κατανομή με 1 β.ε.)}$$

### Ποσοτικό θεραπευτικό αποτέλεσμα

Όταν η αξιολόγηση της αποτελεσματικότητας των θεραπευτικών μέτρων γίνεται με τη χρήση ενός ποσοτικού δείκτη (π.χ. διαφορά της διαστολικής πίεσης DBP από baseline) και ο σχεδιασμός της μελέτης είναι παράλληλος σε δύο ομάδες (αντιυπερτασικό φάρμακο vs placebo) τότε η σύγκριση των δεικτών γίνεται με ένα t-test ή κατασκευάζοντας το 95% δ.ε. για τη διαφορά των ποσοτικών δεικτών.

### Ενδιάμεση ανάλυση (interim analysis) μιας κλινικής μελέτης

Για να είναι ηθικά αποδεκτή μία κλινική μελέτη είναι σημαντικό να ελέγχεται η στατιστική σημαντικότητα της διαφοράς των θεραπευτικών μέτρων με ενδιάμεσες αναλύσεις. Έτσι αποφεύγεται η χρονική παράταση της μελέτης όταν η διαφορά των μέτρων είναι ήδη διαπιστωμένη.

Αν δε διακοπεί η έρευνα τότε οι ασθενείς θα εκτεθούν χωρίς λόγο σε μία θεραπευτική αγωγή που είναι κατώτερη από μία άλλη. Όμως, ο επαναλαμβανόμενος στατιστικός έλεγχος για την τεκμηρίωση μιας ενδεχόμενης διαφοράς κατά τη πρόοδο μιας κλινικής μελέτης αυξάνει σημαντικά την πιθανότητα να προκύψει συμπτωματικά μία διαφορά στατιστικά σημαντική χωρίς να υπάρχει στην πραγματικότητα τέτοια διαφορά. Τότε πρέπει να ορισθεί ο αριθμός των ενδιάμεσων αναλύσεων στο πρωτόκολλο όπως και η στάθμη σημαντικότητας της διαφοράς (συνήθως  $P < 0.01$  για λιγότερες από 10 ενδιάμεσες αναλύσεις).

### Intent to Treat (ITT) ανάλυση μιας κλινικής μελέτης

Η αξιολόγηση της αποτελεσματικότητας ενός θεραπευτικού μέτρου (φαρμάκου) βασίζεται στους ασθενείς που παρακολούθησαν το πρωτόκολλο (per protocol analysis, PPA).

Όμως μία τέτοια ανάλυση εξαιρεί τους ασθενείς που εγκατέλειψαν τη μελέτη για διάφορους λόγους (π.χ. μη-συμμόρφωση με τη δοσολογία, χορήγηση θεραπείας εκτός πρωτοκόλλου κ.ά.). Η ITT ανάλυση τότε περιλαμβάνει όλους τους ασθενείς που κατανεμήθηκαν στα θεραπευτικά μέτρα.

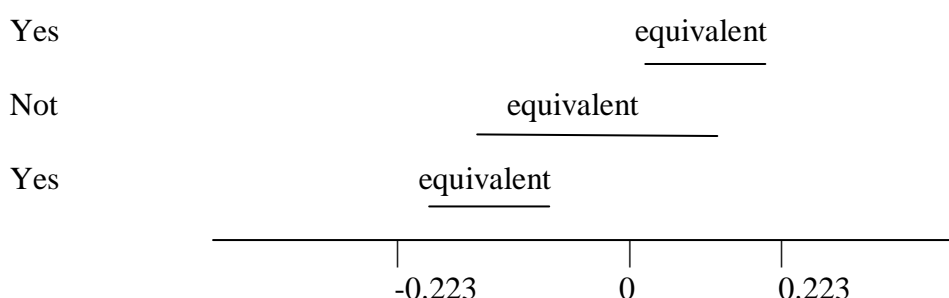
Αν η PPA και η ITT αναλύσεις παράγουν διαφορετικά αποτελέσματα τότε περισσότερη έμφαση δίνεται στην ITT.

### Μελέτες ισοδυναμίας – Κλινική σημαντικότητα vs στατιστική σημαντικότητα

Σκοπός μίας μελέτης ισοδυναμίας είναι να αποδείξει τη θεραπευτική ισοδυναμία μεταξύ δύο θεραπευτικών μέτρων, συνήθως ενός νέου φαρμάκου και ενός ήδη υπάρχοντος φαρμάκου για το ίδιο νόσημα (active comparator). Η ανάλυση μιας τέτοιας μελέτης πρέπει να βασίζεται στο 95% δ.ε. της διαφοράς μεταξύ των φαρμάκων. Σε αυτές τις μελέτες προσδιορίζεται εκ των προτέρων τα όρια της κλινικής τους ισοδυναμίας ( $-\Delta$ ,  $+\Delta$ ). Αν το 95% δ.ε. βρίσκεται μέσα σε αυτά τα όρια τότε τα δύο φάρμακα θεωρούνται θεραπευτικά ισοδύναμα.

Ισοδυναμία μεταξύ T και R δεν σημαίνει απαραίτητα και μη στατιστική διαφορά.

Στατιστική  
Διαφορά (T vs R)



Ένα καινούργιο δισκίο με sodium clodronate έχει αναπτυχθεί και είναι πιο εύχρηστο από την κάψουλα. Οι δύο μορφές θα θεωρηθούν ισοδύναμες αν η διαφορά των  $C_{max}$  βρίσκεται μεταξύ -0.223 και 0.223.

Το 95% δ.ε. για την διαφορά στο  $C_{max}$  μεταξύ δισκίου και κάψουλας είναι: (0.120, 0.210) άρα οι δύο μορφές θεωρούνται θεραπευτικά ισοδύναμες.

**Table 1 Percentage points of the  $t$  distribution.**

d.f.	P			
	0.1	0.05	0.02	0.01
1	6.31	12.71	31.82	63.66
2	2.92	4.30	6.96	9.92
3	2.35	3.18	4.54	5.84
4	2.13	2.78	3.75	4.60
5	2.02	2.57	3.36	4.03
6	1.94	2.45	3.14	3.71
7	1.90	2.36	3.00	3.50
8	1.86	2.31	2.90	3.36
9	1.83	2.26	2.82	3.25
10	1.81	2.23	2.76	3.17
11	1.80	2.20	2.72	3.11
12	1.78	2.18	2.68	3.06
13	1.77	2.16	2.65	3.01
14	1.76	2.14	2.62	2.98
15	1.75	2.13	2.60	2.95
16	1.75	2.12	2.58	2.92
17	1.74	2.11	2.57	2.90
18	1.73	2.10	2.55	2.88
19	1.73	2.09	2.54	2.86
20	1.72	2.09	2.53	2.84
21	1.72	2.08	2.52	2.83
22	1.72	2.07	2.51	2.82
23	1.71	2.07	2.50	2.81
24	1.71	2.06	2.49	2.80
25	1.71	2.06	2.48	2.79
26	1.71	2.06	2.48	2.78
27	1.70	2.05	2.47	2.77
28	1.70	2.05	2.47	2.76
29	1.70	2.04	2.46	2.76
30	1.70	2.04	2.46	2.75
40	1.68	2.02	2.42	2.70
60	1.67	2.00	2.39	2.66
120	1.66	1.98	2.36	2.62
$\infty$	1.65	1.96	2.33	2.58

**Table 2** Percentage points of the  $\chi^2$  distribution.

d.f.	<i>P</i> value			
	0.1	0.05	0.025	0.01
1	2.71	3.84	5.02	6.63
2	4.61	5.99	7.38	9.21
3	6.25	7.81	9.35	11.34
4	7.78	9.49	11.14	13.28
5	9.24	11.07	12.83	15.09
6	10.64	12.59	14.45	16.81
7	12.02	14.07	16.01	18.48
8	13.36	15.51	17.53	20.09
9	14.68	16.92	19.02	21.67
10	15.99	18.31	20.48	23.21
11	17.28	19.68	21.92	24.73
12	18.55	21.03	23.34	26.22
13	19.81	22.36	24.74	27.69
14	21.06	23.68	26.12	29.14
15	22.31	25.00	27.49	30.58
16	23.54	26.30	28.85	32.00
17	24.77	27.59	30.19	33.41
18	25.99	28.87	31.53	34.81
19	27.20	30.14	32.85	36.19
20	28.41	31.41	34.17	37.57
21	29.62	32.67	35.48	38.93
22	30.81	33.92	36.78	40.29
23	32.01	35.17	38.08	41.64
24	33.20	36.42	39.36	42.98
25	34.38	37.65	40.65	44.31
26	35.56	38.89	41.92	45.64
27	36.74	40.11	43.19	46.96
28	37.92	41.34	44.46	48.28
29	39.09	42.56	45.72	49.59
30	40.26	43.77	46.98	50.89

**Table 3. 5% points of the F-distribution**

Degrees of freedom in denominator	Degrees of freedom in numerator															
	1	2	3	4	5	6	7	8	9	10	20	40	∞	1	2	3
1	161	200	216	225	230	234	237	239	241	242	248	251	254	161	200	216
2	18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.5	19.5	18.5	19.0	19.2
3	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.66	8.59	8.53	10.1	9.55	9.28
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.80	5.72	5.63	7.71	6.94	6.59
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.56	4.46	4.37	6.61	5.79	5.41
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	3.87	3.77	3.67	5.99	5.14	4.76
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.44	3.34	3.23	5.59	4.74	4.35
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.15	3.04	2.93	5.32	4.46	4.07
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	2.94	2.83	2.71	5.12	4.26	3.86
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.77	2.66	2.54	4.96	4.10	3.71
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.65	2.53	2.40	4.84	3.98	3.59
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.54	2.43	2.30	4.75	3.89	3.49
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.46	2.34	2.21	4.67	3.81	3.41
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.39	2.27	2.13	4.60	3.74	3.34
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.33	2.20	2.07	4.54	3.68	3.29
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.28	2.15	2.01	4.49	3.63	3.24
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.23	2.10	1.96	4.45	3.59	3.20
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.19	2.06	1.92	4.41	3.55	3.16
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.16	2.03	1.88	4.38	3.52	3.13
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.12	1.99	1.84	4.35	3.49	3.10
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.01	1.87	1.71	4.24	3.39	2.99
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	1.93	1.79	1.62	4.17	3.32	2.92
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	1.84	1.69	1.51	4.08	3.23	2.84
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.66	1.50	1.25	3.92	3.07	2.68
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.57	1.39	1.00	3.84	3.00	2.60

**Table 4.**

Sample size, $N$	5% Level <sup>a</sup>	1% Level
6	0	—
7	2	—
8	3	0
9	5	1
10	8	3
11	10	5
12	13	7
13	17	10
14	21	13
15	25	16
16	30	19
17	35	23
18	40	28
19	46	32
20	52	37

Table 5. Mann-Whitney-Wilcoxon test.

$n_1$	$n_2$	Two-tailed probability (P)				
		0.1	0.05	0.02	0.01	0.001
3	3	6–15	–	–	–	–
3	4	6–18	–	–	–	–
4	4	11–25	10–26	–	–	–
2	5	3–13	–	–	–	–
3	5	7–20	6–21	–	–	–
4	5	12–28	11–29	10–30	–	–
5	5	19–36	17–38	16–39	15–40	–
2	6	3–15	–	–	–	–
3	6	8–22	7–23	–	–	–
4	6	13–31	12–32	11–33	10–34	–
5	6	20–40	18–42	17–43	16–44	–
6	6	28–50	26–52	24–54	23–55	–
2	7	3–17	–	–	–	–
3	7	8–25	7–26	6–27	–	–
4	7	14–34	13–35	11–37	10–38	–
5	7	21–44	20–45	18–47	16–49	–
6	7	29–55	27–57	25–59	24–60	–
7	7	39–66	36–69	34–71	32–73	28–77
2	8	4–18	3–19	–	–	–
3	8	9–27	8–28	6–30	–	–
4	8	15–37	14–38	12–40	11–41	–
5	8	23–47	21–49	19–51	17–53	–
6	8	31–59	29–61	27–63	25–65	21–69
7	8	41–71	38–74	35–77	34–78	29–83
8	8	51–85	49–87	45–91	43–93	38–98



$n_1$	$n_2$	Two-tailed probability (P)				
		0.1	0.05	0.02	0.01	0.001
2	9	4-20	3-21	-	-	-
3	9	10-29	8-31	7-32	6-33	-
4	9	16-40	14-42	13-43	11-45	-
5	9	24-51	22-53	20-55	18-57	15-60
6	9	33-63	31-65	28-68	26-70	22-74
7	9	43-76	40-79	37-82	35-84	30-89
8	9	54-90	51-93	47-97	45-99	40-104
9	9	66-105	62-109	59-112	56-115	50-121
2	10	4-22	3-23	-	-	-
3	10	10-32	9-33	7-35	6-36	-
4	10	17-43	15-45	13-47	12-48	-
5	10	26-54	23-57	21-59	19-61	15-65
6	10	35-67	32-70	29-73	27-75	23-79
7	10	45-81	42-84	39-87	37-89	31-95
8	10	56-96	53-99	49-103	47-105	41-111
9	10	69-111	65-115	61-119	58-122	52-128
10	10	82-128	78-132	74-136	71-139	63-147
2	11	4-24	3-25	-	-	-
3	11	11-34	9-36	7-38	6-39	-
4	11	18-46	16-48	14-50	12-52	-
5	11	27-58	24-61	22-63	20-65	16-69
6	11	37-71	34-74	30-78	28-80	23-85
7	11	47-86	44-89	40-93	38-95	32-101
8	11	59-101	55-105	51-109	49-111	42-118
9	11	72-117	68-121	63-126	61-128	53-136
10	11	86-134	81-139	77-143	73-147	65-155
11	11	100-153	96-157	91-162	87-166	78-175
2	12	5-25	4-26	-	-	-
3	12	11-37	10-38	8-40	7-41	-
4	12	19-49	17-51	15-53	13-55	-
5	12	28-62	26-64	23-67	21-69	16-74
6	12	38-76	35-79	32-82	30-84	24-90
7	12	49-91	46-94	42-98	40-100	33-107
8	12	62-106	58-110	53-115	51-117	43-125
9	12	75-123	71-127	66-132	63-135	55-143
10	12	89-141	84-146	79-151	76-154	67-163
11	12	104-160	99-165	94-170	90-174	81-183
12	12	120-180	115-185	109-191	105-195	95-205

$n_1$	$n_2$	Two-tailed probability (P)				
		0.1	0.05	0.02	0.01	0.001
2	13	5-27	4-28	3-29	-	-
3	13	12-39	10-41	8-43	7-44	-
4	13	20-52	18-54	15-57	13-59	10-62
5	13	30-65	27-68	24-71	22-73	17-78
6	13	40-80	37-83	33-87	31-89	25-95
7	13	52-95	48-99	44-103	41-106	34-113
8	13	64-112	60-116	56-120	53-123	45-131
9	13	78-129	73-134	68-139	65-142	56-151
10	13	92-148	88-152	82-158	79-161	69-171
11	13	108-167	103-172	97-178	93-182	83-192
12	13	125-187	119-193	113-199	109-203	98-214
13	13	142-209	136-215	130-221	125-226	114-237
2	14	6-28	4-30	3-31	-	-
3	14	13-41	11-43	8-46	7-47	-
4	14	21-55	19-57	16-60	14-62	10-66
5	14	31-69	28-72	25-75	22-78	17-83
6	14	42-84	38-88	34-92	32-94	26-100
7	14	54-100	50-104	45-109	43-111	35-119
8	14	67-117	62-122	58-126	54-130	46-138
9	14	81-135	76-140	71-145	67-149	58-158
10	14	96-154	91-159	85-165	81-169	71-179
11	14	112-174	106-180	100-186	96-190	85-201
12	14	129-195	123-201	116-208	112-212	100-224
13	14	147-217	141-223	134-230	129-235	116-248
14	14	166-240	160-246	152-254	147-259	134-272
2	15	6-30	4-32	3-33	-	-
3	15	13-44	11-46	9-48	8-49	-
4	15	22-58	20-60	17-63	15-65	10-70
5	15	33-72	29-76	26-79	23-82	18-87
6	15	44-88	40-92	36-96	33-99	26-106
7	15	56-105	52-109	47-114	44-117	36-125
8	15	69-123	65-127	60-132	56-136	47-145
9	15	84-141	79-146	73-152	69-156	60-165
10	15	99-161	94-166	88-172	84-176	73-187
11	15	116-181	110-187	103-194	99-198	87-210
12	15	133-203	127-209	120-216	115-221	103-233
13	15	152-225	145-232	138-239	133-244	119-258
14	15	171-249	164-256	156-264	151-269	137-283
15	15	192-273	184-281	176-289	171-294	156-309

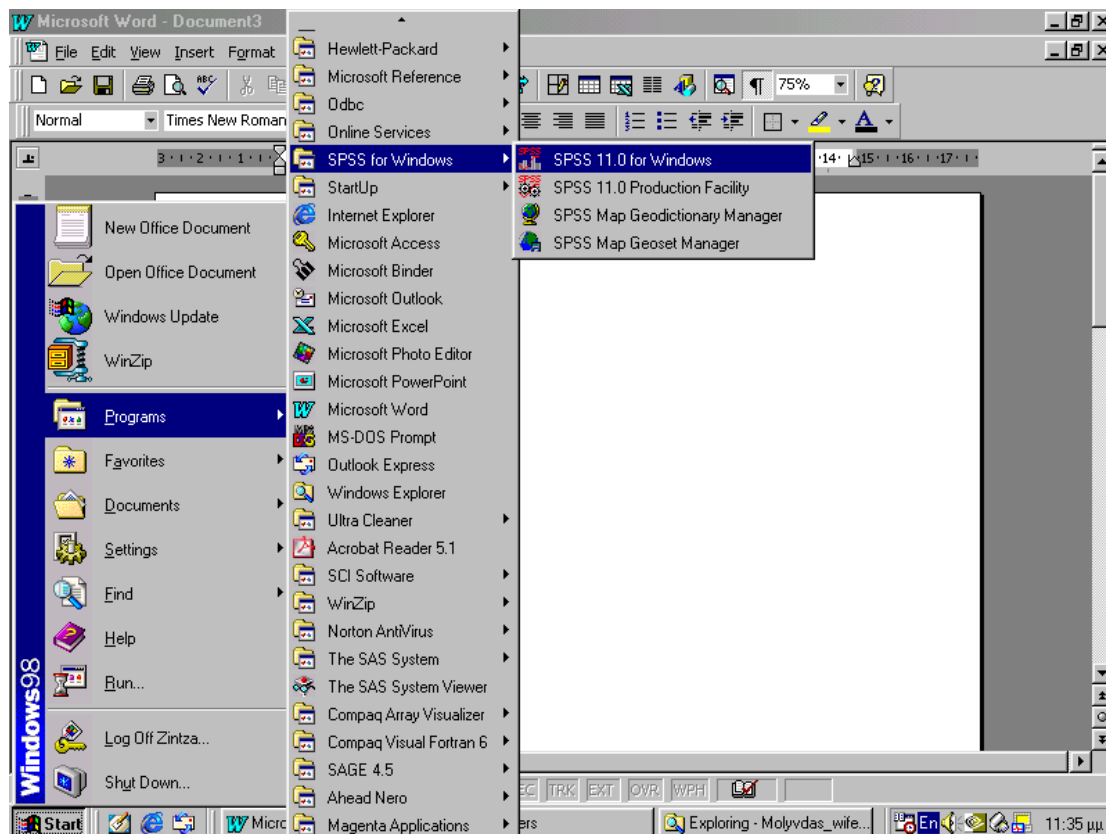
## ΑΝΑΦΟΡΕΣ

- Δ. Τριχόπουλος, Α. Τζώνου & Κ. Κατσουγιάννη (2000) Βιοστατιστική. Παρισιάνος, Αθήνα.
- Δ. Τριχόπουλος (1982) Επιδημιολογία. Παρισιάνος. Αθήνα.
- D.G. Altman (1991) Practical Statistics for Medical Research. Chapman & Hall, London.
- S. Bolton (1997) Pharmaceutical Statistics. Marcel Dekker, New York.
- G. Dunn and B. Everitt (1995) Clinical Biostatistics. Edward Arnold, London.
- B.R. Kirkwood (1988) Medical Statistics. Blackwell, Oxford.
- R. Mead and R.N. Curnow (1983) Statistical Methods in Agriculture and Experimental Biology. Chapman & Hall, London.
- S.J. Pocock (1989) Clinical Trials. Wiley, Chichester.

## Πρακτικός οδηγός χρήσης του SPSS release 10.0.1 με παραδείγματα

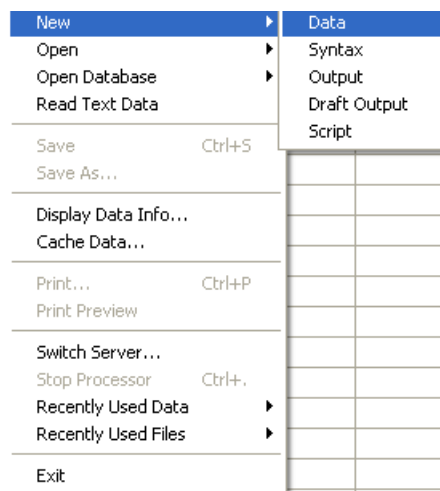
### Εισαγωγή στο Πρόγραμμα SPSS-Διαχείριση Αρχείων και δεδομένων

Για να αρχίσετε τη λειτουργία του SPSS επιλέξτε στο Desktop: “Start”-“Programs”-“SPSS for Windows”-“SPSS 11.0 for Windows”.



Το SPSS δίνει την δυνατότητα επεξεργασίας δεδομένων που είτε εισάγει ο χρήστης από το πληκτρολόγιο είτε είναι αποθηκευμένα σε Excel, Βάσεις Δεδομένων κ.τ.λ.

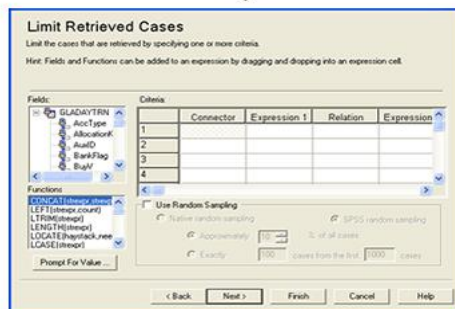
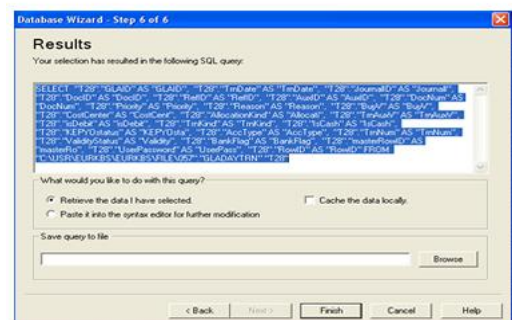
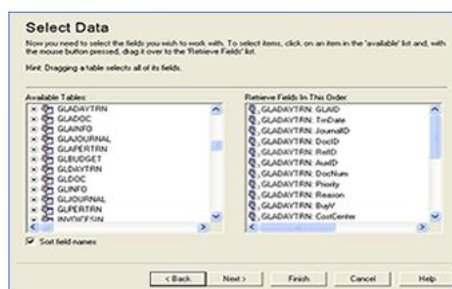
Έτσι επιλέγοντας από το μενού επιλογών File έχουμε τις εξής επιλογές



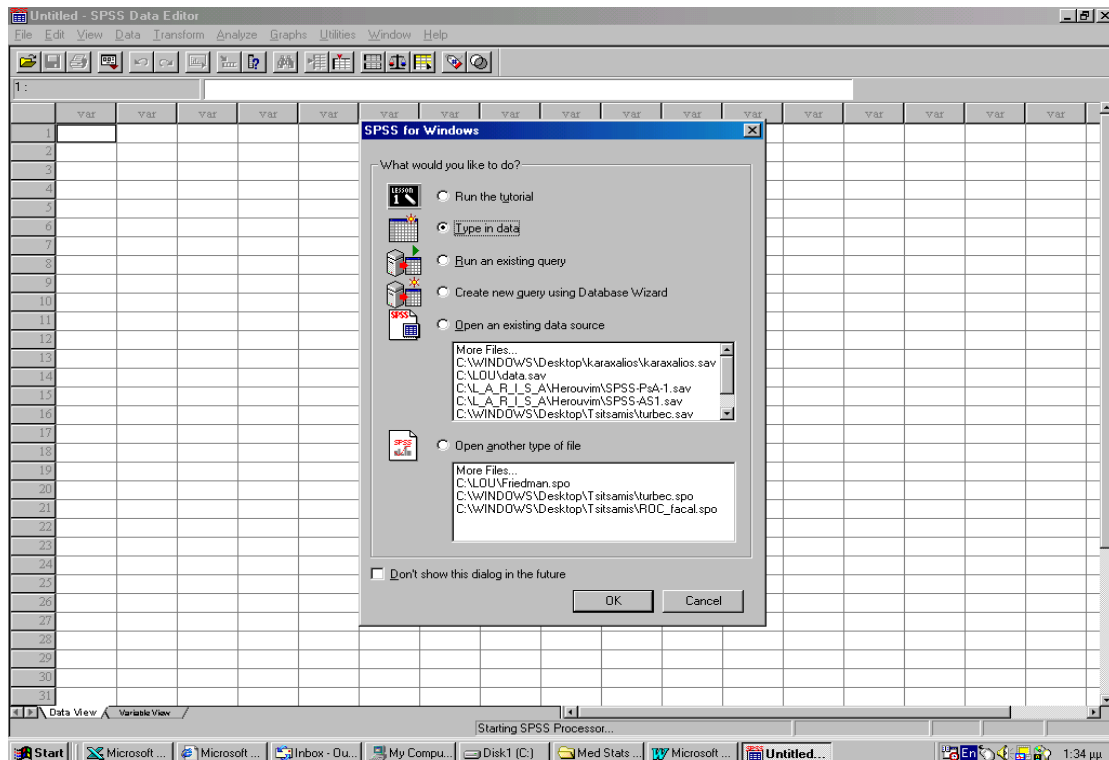
Επιλέγοντας New, μπορούμε να εισάγουμε ένα καινούργιο data set σε ένα πίνακα δεδομένων (Βλ. **Error! Reference source not found.**). Μπορούμε επίσης να δημιουργήσουμε ένα νέο παράθυρο εξόδου δεδομένων, όπου το SPSS θα βάλει τα αποτελέσματα μιας ανάλυσης.

Εάν τώρα τα δεδομένα που θέλουμε να επεξεργαστούμε, είναι αποθηκευμένα σε μια Βάση Δεδομένων (ΒΔ), μπορούμε να τα ανακτήσουμε, επιλέγοντας Open Database. Εδώ έχουμε την δυνατότητα είτε να δημιουργήσουμε μια νέα αναζήτηση, δηλαδή να καθορίσουμε τους πίνακες της ΒΔ, από τους οποίους το SPSS θα ανακτήσει τα δεδομένα μας, να τροποποιήσουμε μια αναζήτηση που έχουμε ήδη δημιουργήσει ή τέλος να ανακτήσουμε μια αναζήτηση που έχουμε αποθηκεύσει σε κάποια άλλη στιγμή.

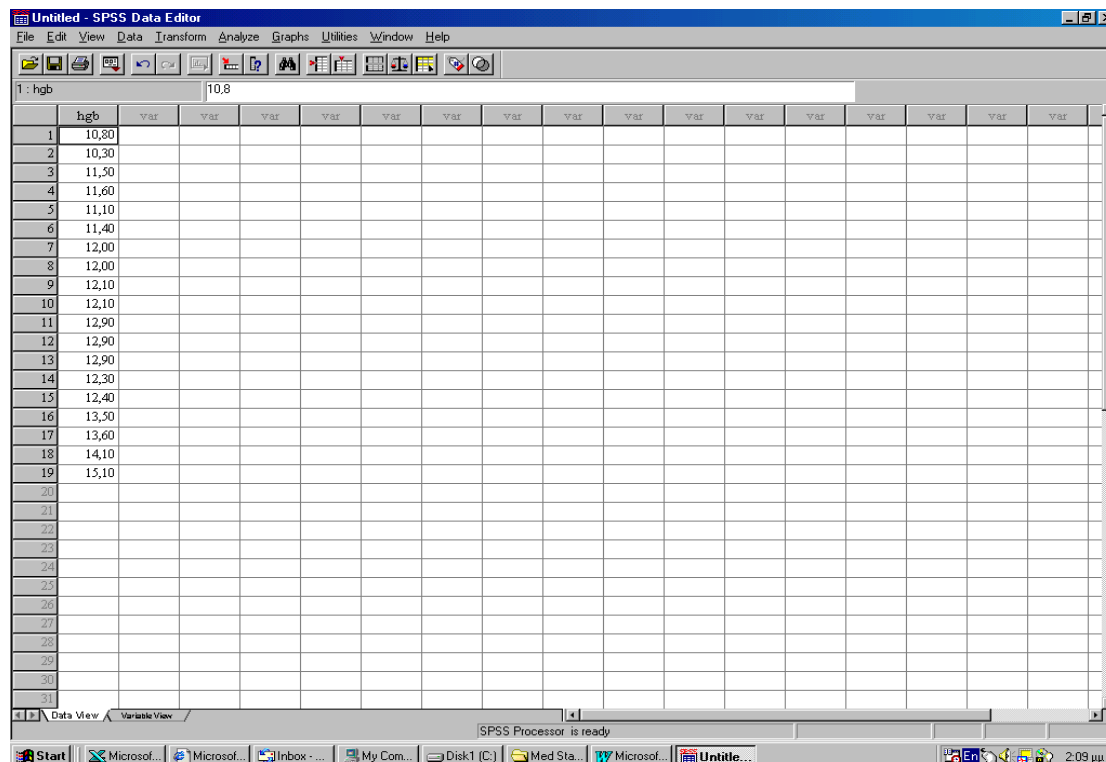
Όταν καθορίζουμε μια αναζήτηση, το SPSS μας παρουσιάζει έναν οδηγό βημάτων (wizard) που με την βοήθεια επιλογών, μας βοηθά στο να καθορίσουμε τους πίνακες της ΒΔ αλλά και τον τρόπο που αυτοί οι πίνακες συνδέονται μεταξύ τους για να ανακτήσουμε τα δεδομένα.



Έστω ότι θέλουμε να εισάγουμε ένα dataset . Όταν ανοίξουμε για πρώτη φορά το SPSS μας παρουσιάζεται η οθόνη με επιλογές για το είδος του dataset (εάν το εισάγουμε εμείς από το πληκτρολόγιο ή εάν ανακτήσουμε τα δεδομένα μας από μια ΒΔ). Επιλέγουμε “Type in data” για να πληκτρολογήσουμε τα δεδομένα.

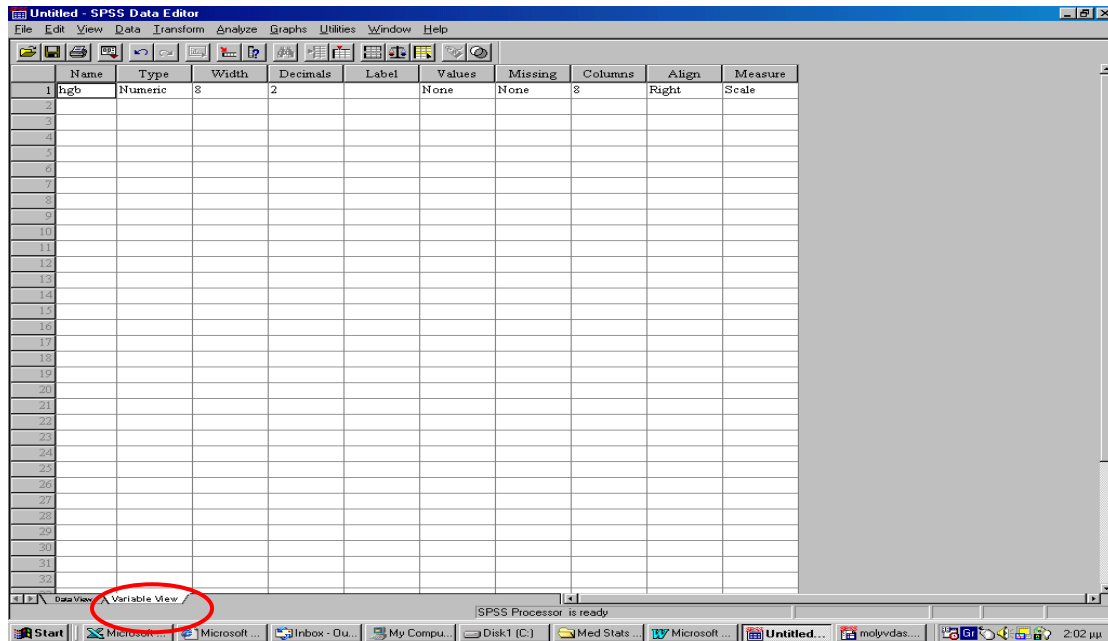


Καταχωρήστε τα επίπεδα αιμοσφαιρίνης 20 ατόμων.



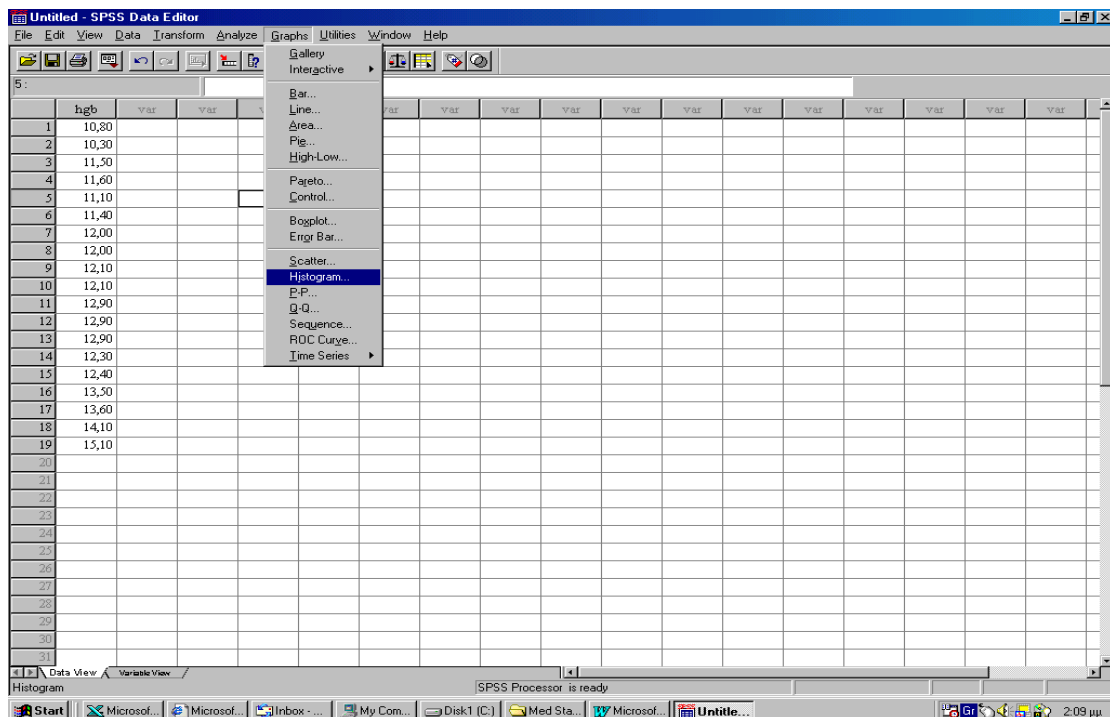
Η αλλαγή του ονόματος (Name) της μεταβλητής γίνεται επιλέγοντας το tab ‘Variable View’, στο κάτω αριστερό τμήμα του πλαισίου, που γίνεται η καταχώρηση των δεδομένων είτε επιλέγοντας από το μενού επιλογών ‘View-

Variables'. Εδώ μπορούμε, εκτός του ονόματος, να καθορίσουμε και άλλα στοιχεία της μεταβλητής, όπως τον τύπο της, το μήκος, τον αριθμό των δεκαδικών ψηφίων κτλ



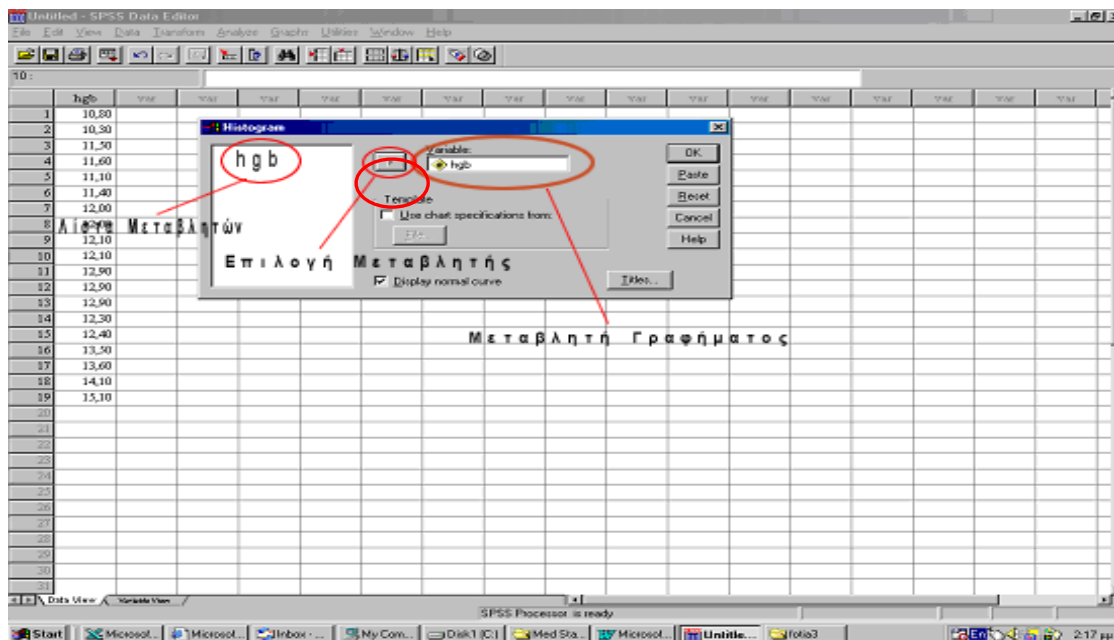
## Περιγραφική Στατιστική

Η επιλογή “Graph” από το μενού επιλογών, μας βοηθάει να κατασκευάσουμε διάφορα γραφήματα των δεδομένων όπως “Bar”, “Pie”, “Histogram” κτλ. Εάν π.χ. θέλουμε ένα ιστόγραμμα που να περιγράφει τις τιμές της Hgb που καταχωρήσαμε στο **Error! Reference source not found.**, επιλέγουμε από το μενού επιλογών, επιλέγουμε “Graphs” -> “Histogram”.

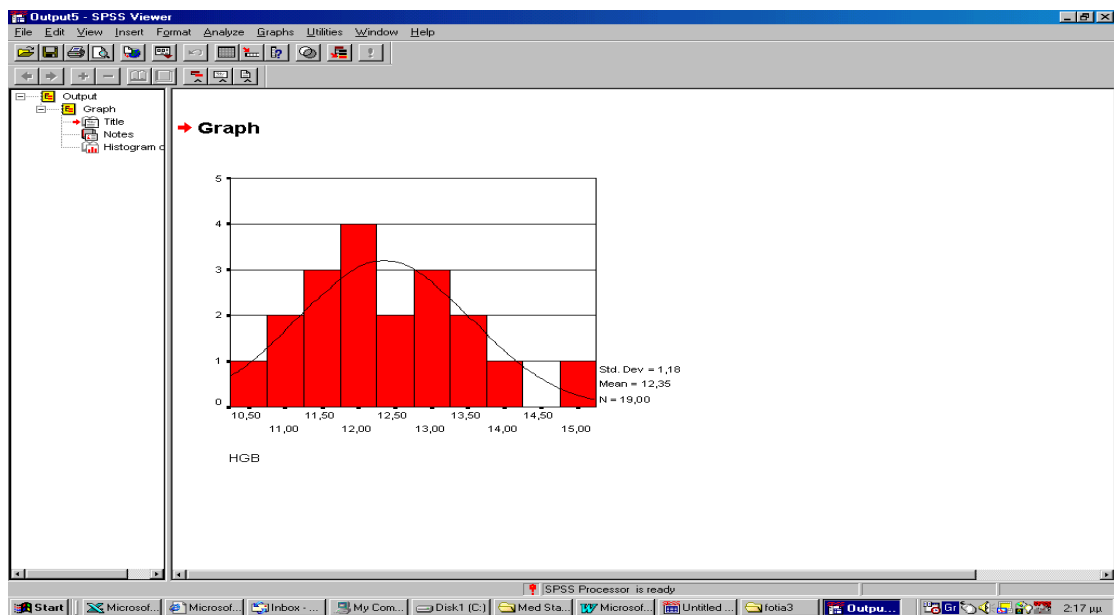


Στην συνέχεια εμφανίζεται η φόρμα επιλογών. Αφού επιλέξουμε την μεταβλητή hgb από την Λίστα Μεταβλητών και κάνοντας κλικ στο κουμπί Επιλογής Μεταβλητής, τοποθετούμε την επιλεγμένη μεταβλητή στο

πεδίο Μεταβλητή Γραφήματος (Variable). Επιλέγουμε επίσης “Display normal curve” για να συμπεριλάβουμε στο ιστόγραμμα μας και την καμπύλη της κανονικής κατανομής. Στην συνέχεια και κάντε κλικ στο “OK”.

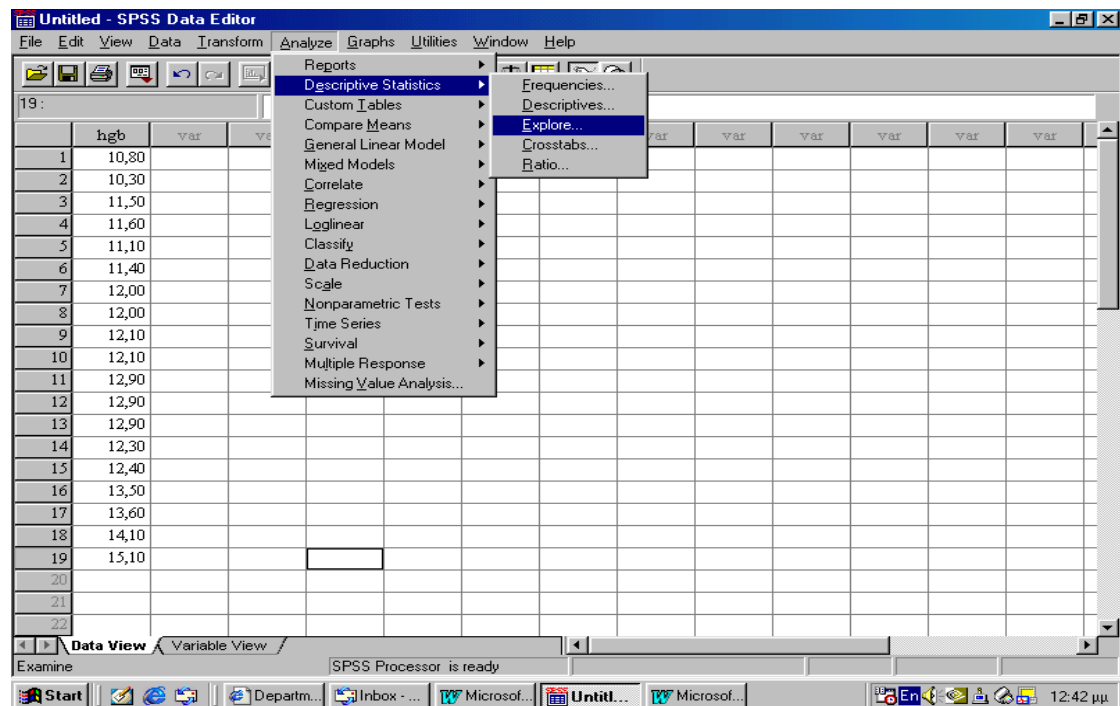


Τότε παράγεται το ιστόγραμμα με την καμπύλη της κανονικής κατανομής σε “output” αρχείο του SPSS.

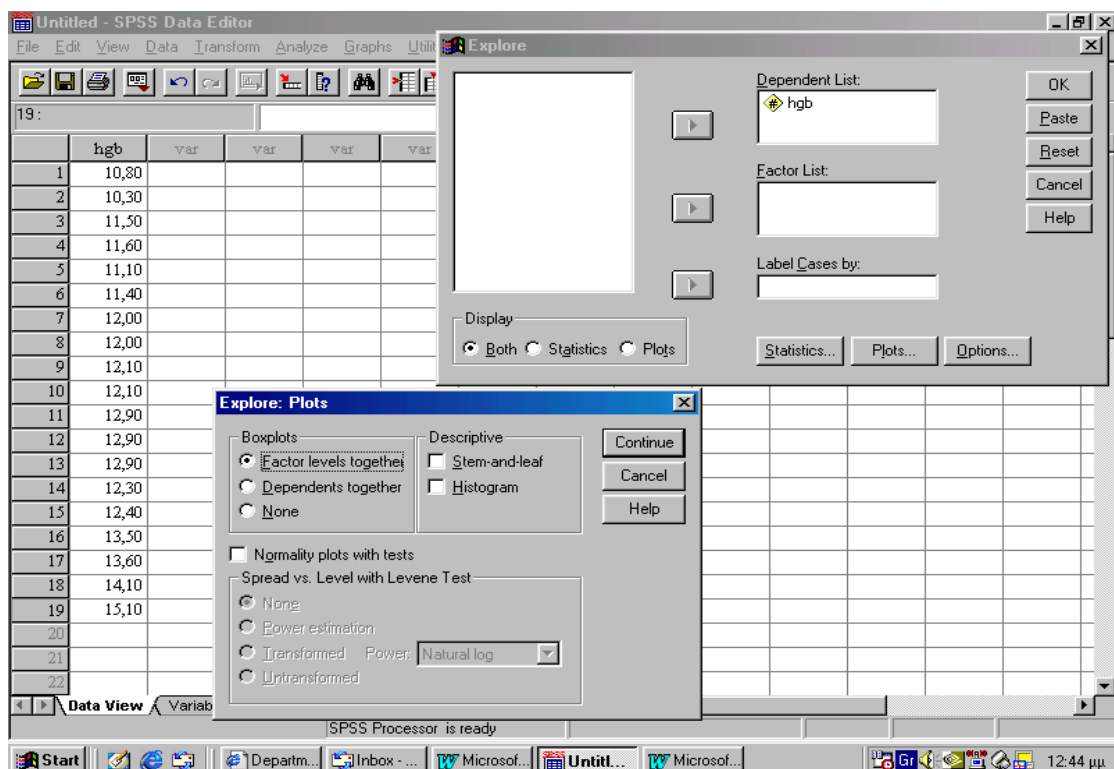


Για να δημιουργήσουμε ένα box plot, επιλέγουμε από το μενού επιλογών Analyze->Descriptive Statistics->Explore.

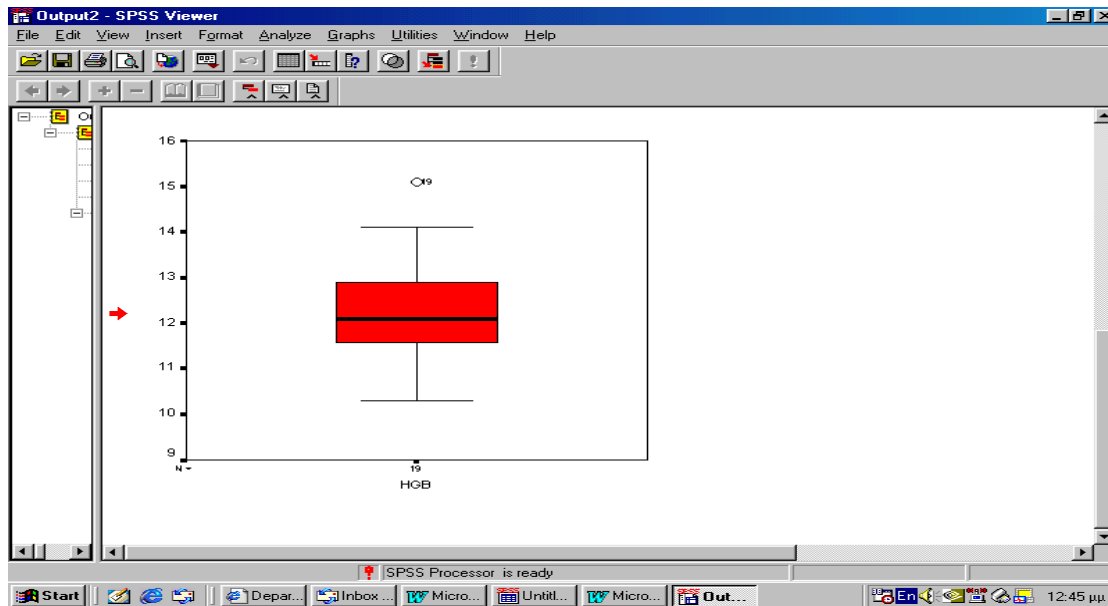




Στην συνέχεια από την Λίστα Μεταβλητών επιλέγουμε την μεταβλητή, για την οποία θέλουμε να δημιουργήσουμε το Box Plot. Στην συνέχεια κάνοντας κλικ στο κουμπί Plots, εμφανίζεται το παράθυρο με τίτλο “Explore:Plots” και στην ομάδα “Box Plots” επιλέγουμε Factor levels Together.

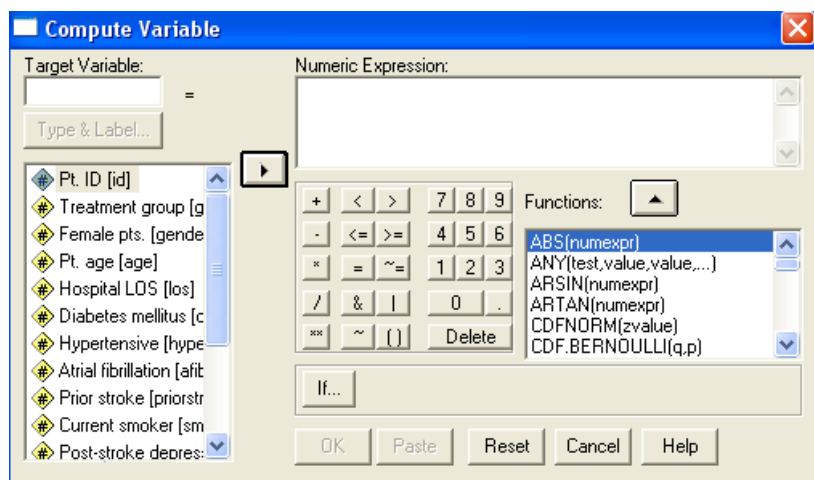


Το output αρχείο δείχνει τα αποτελέσματα της ανάλυσης



Η μεσαία γραμμή στο κόκκινο κουτί, παριστάνει την διάμεσο, ενώ οι άνω κα κάτω πλευρά του κουτιού, παριστάνει τα 25 και 75 Quantiles. Οι κάθετες γραμμές, που καταλήγουν στις δύο οριζόντιες (μουστάκια), παριστάνουν την μικρότερη και μεγαλύτερη τιμή του δείγματός μας.

Εάν θέλουμε να τροποποιήσουμε μια μεταβλητή, είτε υπολογίζοντας τον λογάριθμο της είτε διαιρώντας την με ένα αριθμό, αυτό μπορούμε να το πετύχουμε, επιλέγοντας από το μενού επιλογών Transform-> Compute.

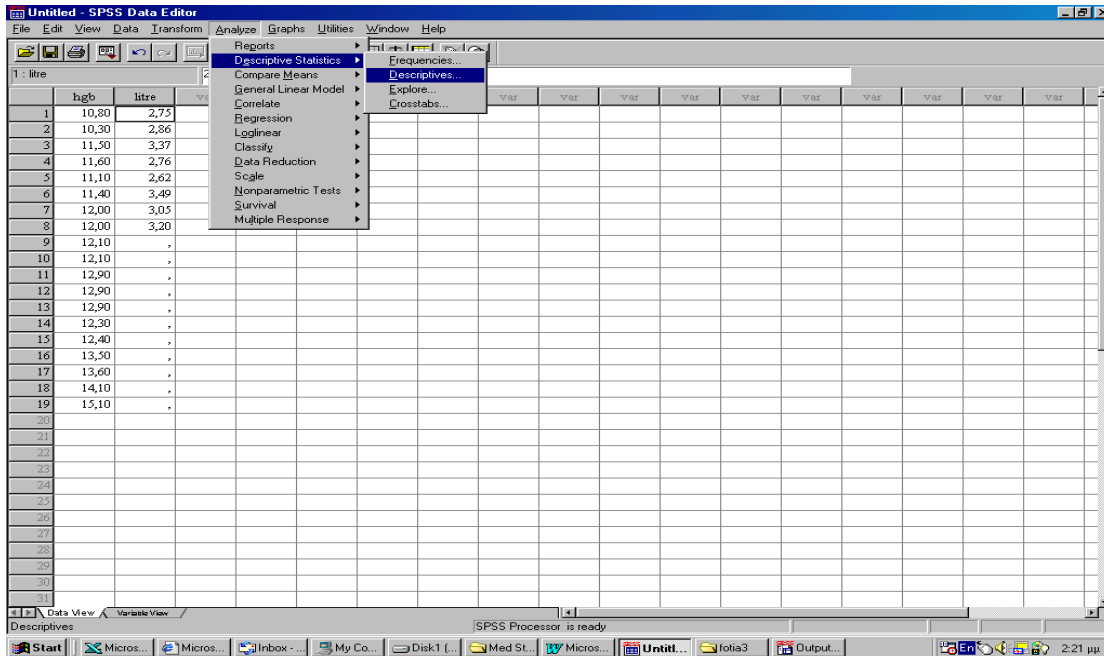


Το SPSS έρχεται με μια μεγάλη λίστα συναρτήσεων που μπορούμε να εφαρμόσουμε σε μια μεταβλητή. Μας δίνει επίσης την δυνατότητα να δημιουργήσουμε και νέες κάνοντας διάφορες συνδυασμούς.

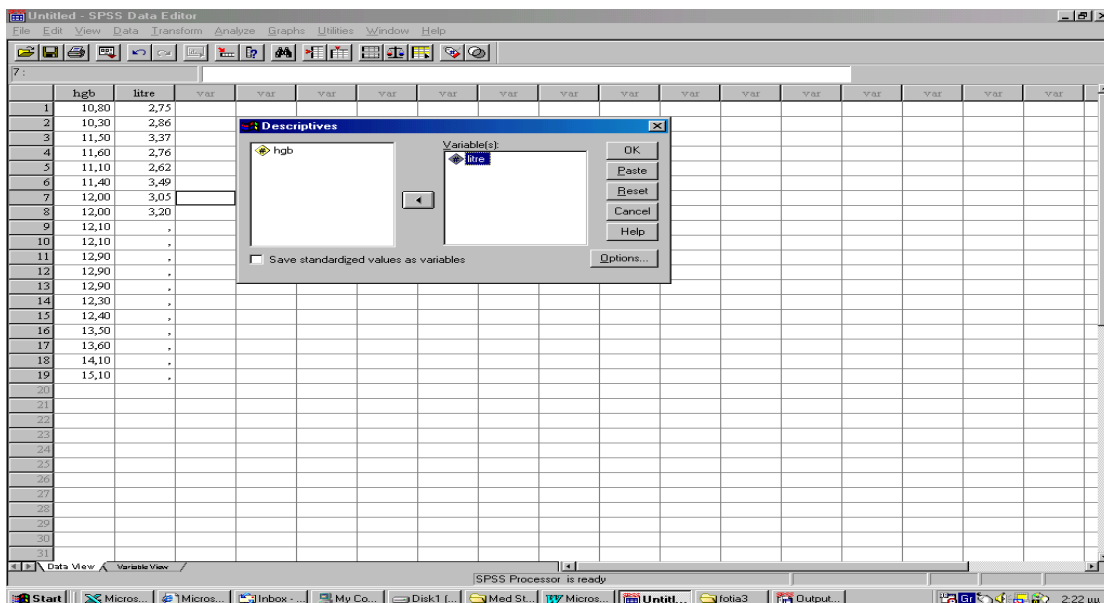
Καταχωρήστε τους όγκους πλάσματος από 8 άτομα, και δώστε στη μεταβλητή το όνομα “litre” – ακολουθώντας την διαδικασία που περιγράφηκε παραπάνω.

Για να υπολογίσουμε διάφορα περιγραφικά δεδομένα του δείγματός μας όπως μέση τιμή, τυπική απόκλιση, μέγιστη τιμή, ελάχιστη τιμή κ.λ. από το μενού επιλέγουμε Analyze -> Descriptive Statistics. Στην συνέχεια εμφανίζεται ένα άλλο μενού επιλογών, όπου διαλέγουμε το είδος των περιγραφικών στοιχείων που επιθυμούμε να βρούμε.

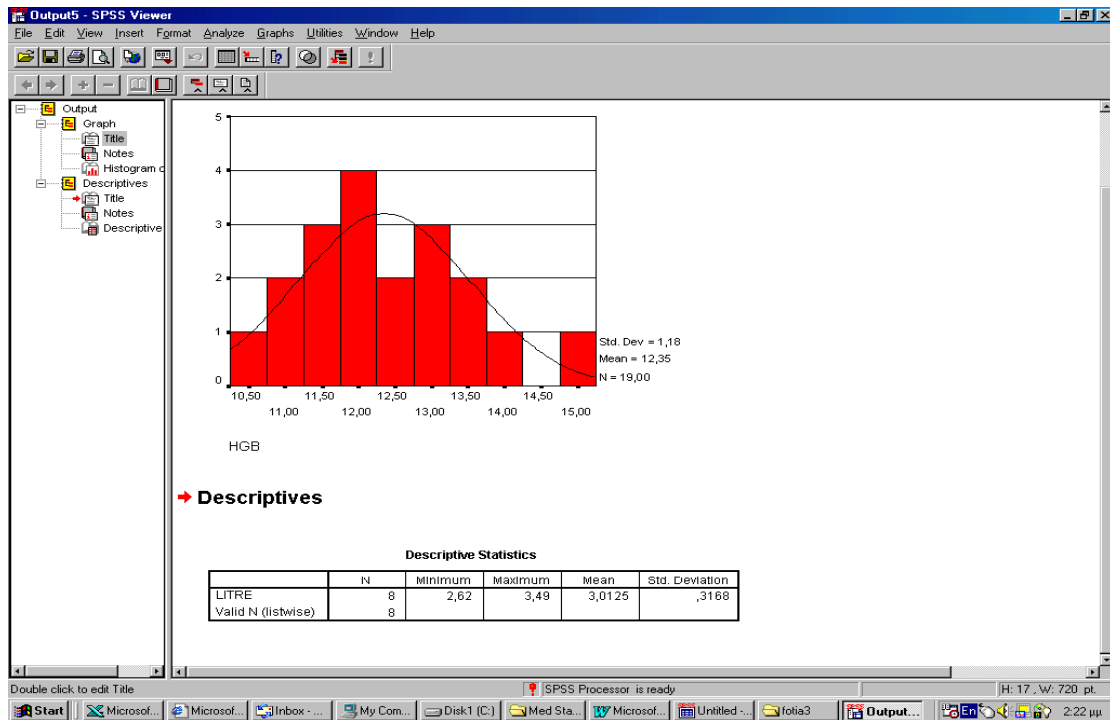
Έτσι για να υπολογίσουμε την μέση τιμή, τη τυπική απόκλιση, το ελάχιστο και το μέγιστο από τις επιλέγουμε “Descriptives”.



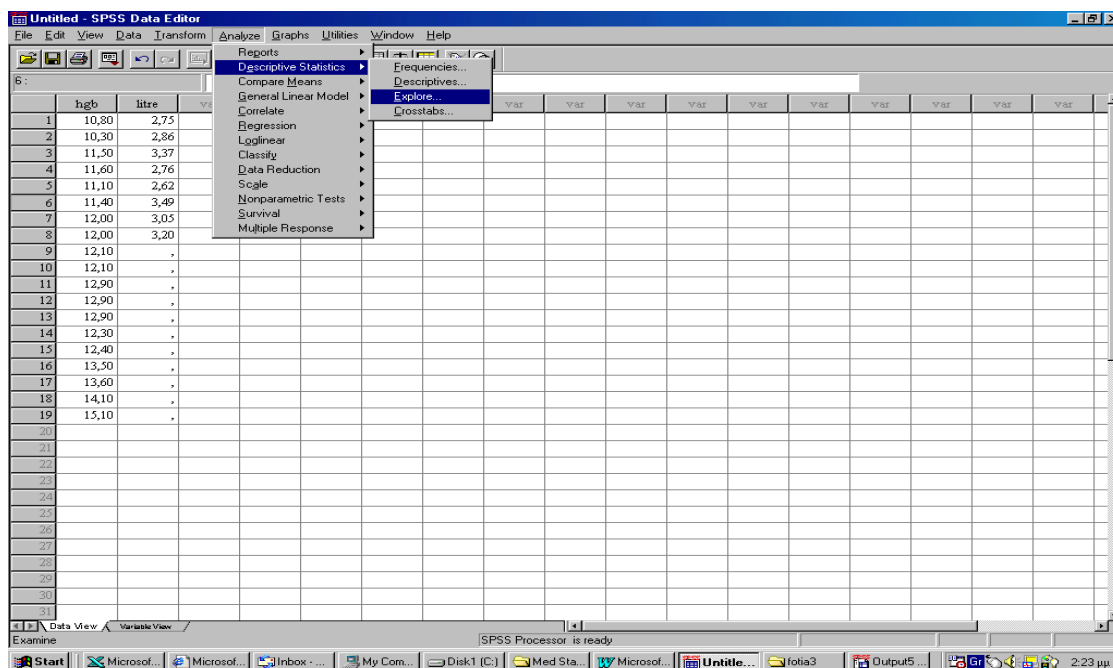
Στην συνέχεια επιλέγουμε από την λίστα μεταβλητών, την μεταβλητή που αναφέρεται στ δεδομένα μας, δηλ. litre (Βλ. **Error! Reference source not found.**).



Το output αρχείο δεξιά, κάτω από τον τίτλο **Descriptives**, δείχνει τα αποτελέσματα της ανάλυσης.

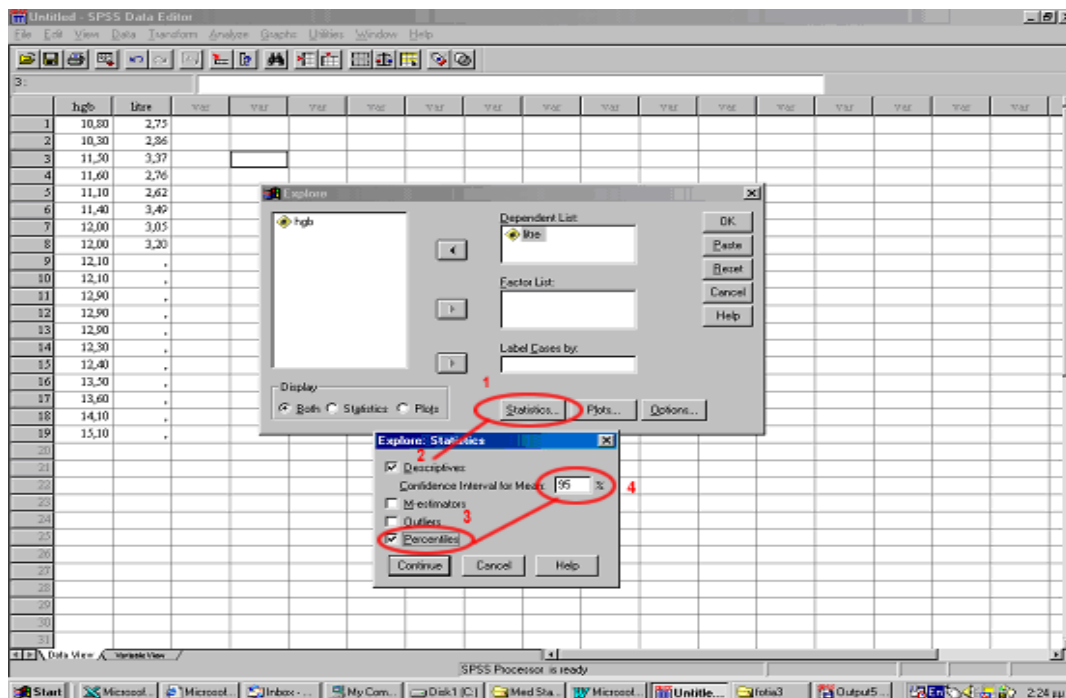


Στο ίδιο δείγμα, για να υπολογίσουμε το 95% δ.ε. για τη μέση τιμή της μεταβλητής “litre” και διάφορα περιγραφικά στατιστικά της μεταβλητής, επιλέγουμε Explore.



Μπορούμε επίσης να προσδιορίσουμε οι ίδιοι το % του δ.ε.

Έτσι στην επιλογή “Explore”, και αφού επιλέξουμε την μεταβλητή από την Λίστα Μεταβλητών, κάνοντας κλικ στο κουμπί Statistics, εμφανίζεται η μικρή οθόνη με τίτλο “Explore : Statistics”. Εδώ μαρκάροντας την επιλογή Descriptives, καθορίζουμε το δ.ε. που μας ενδιαφέρει. Επίσης μαρκάροντας την επιλογή Percentiles, μπορούμε να βρούμε κα τη διάμεσο.



Τα αποτελέσματα των παραπάνω βρίσκονται στο output. Στην προκειμένη περίπτωση, αγνοούμε το Case Processing Summary.

**Case Processing Summary**

	Cases				Total	
	Valid		Missing		N	Percent
LITRE	N	Percent	N	Percent	N	Percent
	8	42,1%	11	57,9%	19	100,0%

**Descriptives**

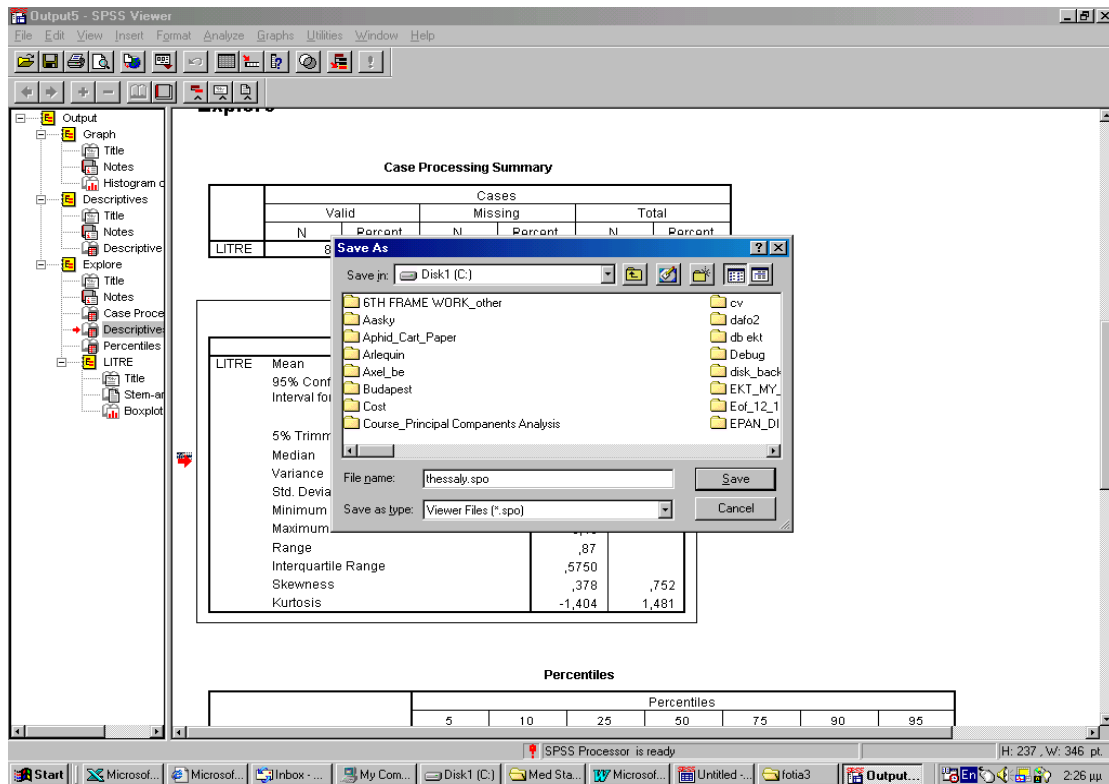
			Statistic	Std. Error
LITRE	Mean		3,0125	,1120
	95% Confidence Interval for Mean	Lower Bound	2,7477	
		Upper Bound	3,2773	
	5% Trimmed Mean		3,0078	
	Median		2,9550	
	Variance		,100	
	Std. Deviation		,3168	
	Minimum		2,62	
	Maximum		3,49	
	Range		,87	
	Interquartile Range		,5750	
	Skewness		,378	,752
	Kurtosis		-1,404	1,481

**Percentiles**

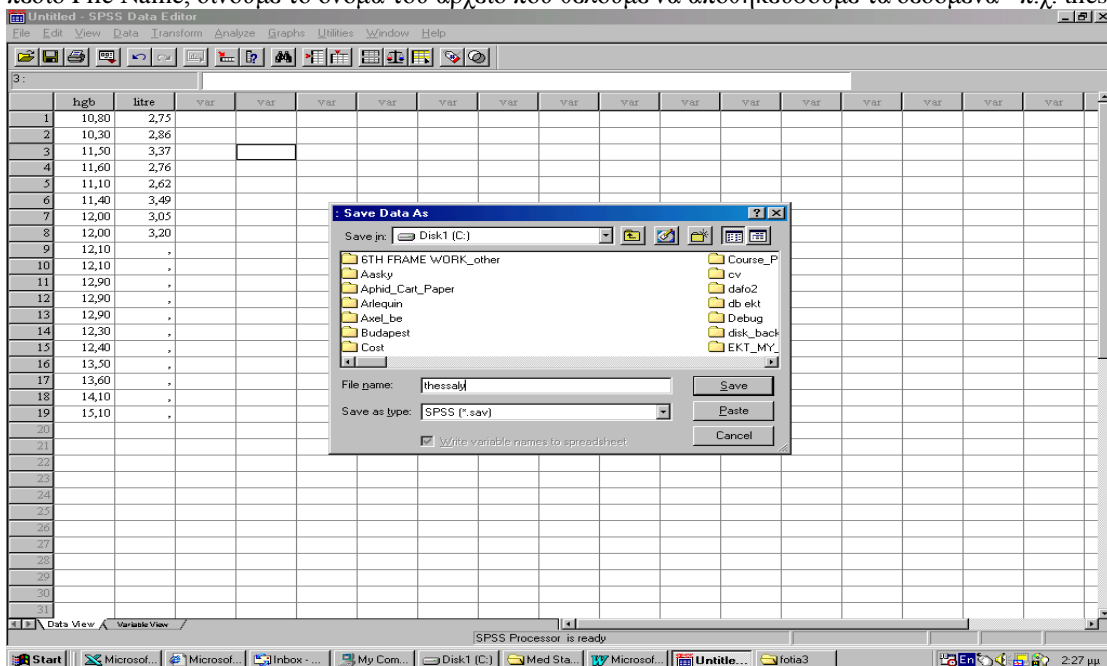
## Αποθήκευση αρχείων

Μπορούμε να αποθηκεύσουμε στο δίσκο (είτε στο floppy drive είτε στο hard drive) το output.

Από το μενού επιλογών διαλέγουμε File->Save As και στο πεδίο File Name, δίνουμε το όνομα του αρχείου – π.χ. thessaly.spo. Έτσι θα μπορέσουμε αργότερα να ανακτήσουμε τα αποτελέσματα της ανάλυσης που έχουμε κάνει.

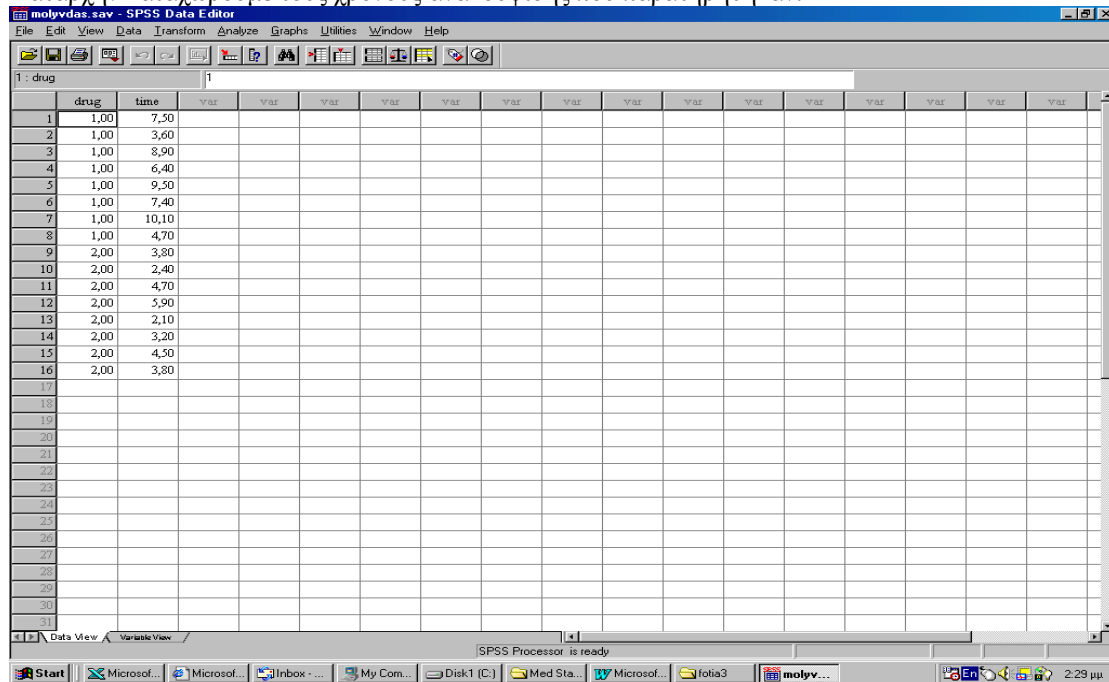


Μπορούμε επίσης να αποθηκεύσουμε και τα δεδομένα σε κάποιο αρχείο, για μελλοντική χρήση. Έτσι από το μενού επιλογών, διαλέγουμε File->Save Data As (Βλ. **Error! Reference source not found.**) Στο πεδίο File Name, δίνουμε το όνομα του αρχείου που θέλουμε να αποθηκεύσουμε τα δεδομένα –π.χ. thessaly.sav.



## t-test για δύο ανεξάρτητα δείγματα

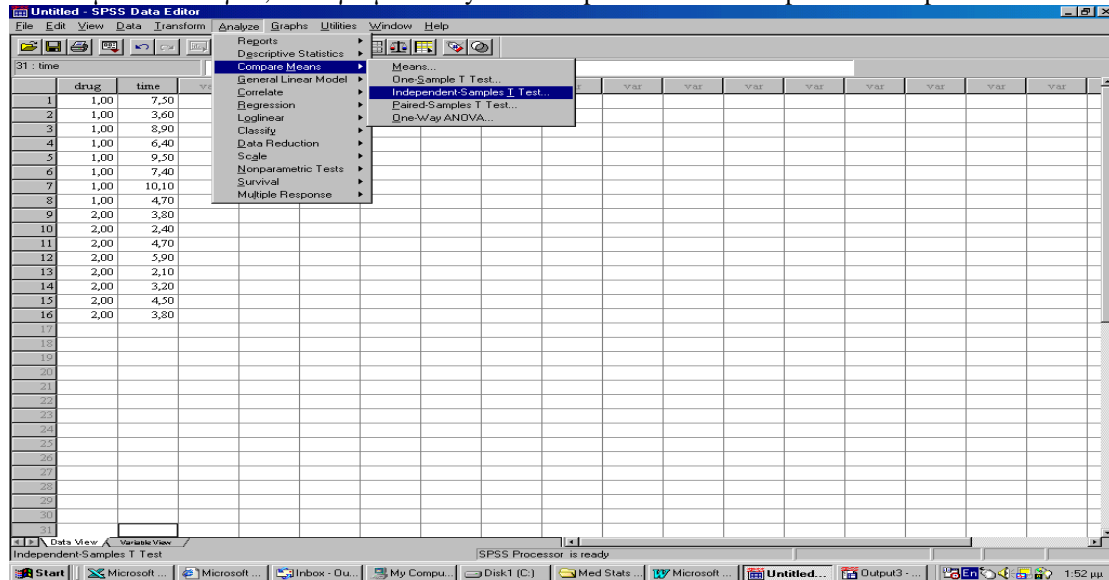
Έστω ότι θέλουμε να ελέγξουμε εάν δύο αναλγητικά φάρμακα διαφέρουν ως προς τους χρόνους ανακούφισης. Για τη σύγκριση δύο αναλγητικών φαρμάκων “1” και “2”, 16 ασθενείς διαιρέθηκαν τυχαία στις δύο ομάδες. Καταρχήν καταχωρούμε τους χρόνους ανακούφισης που παρατηρήθηκαν.



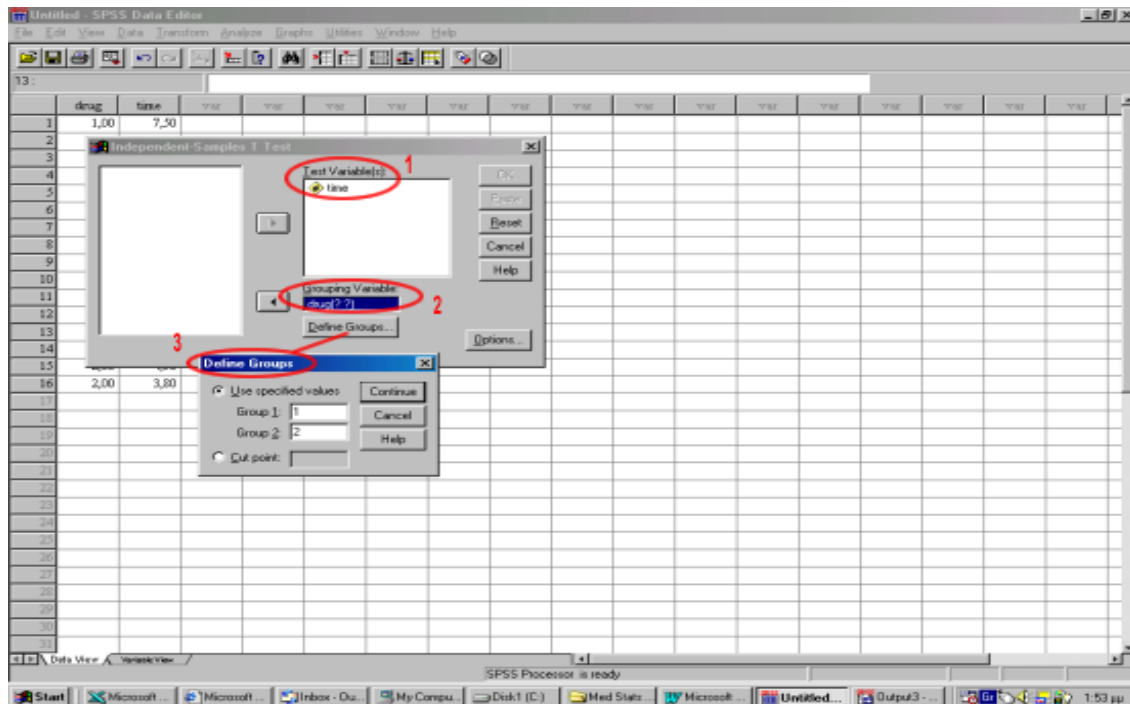
Για να ελέγξουμε εάν υπάρχει διαφορά μεταξύ των δύο φαρμάκων, θα εφαρμόσουμε το t-test καθορίζοντας την στάθμη σημαντικότητας (significance level ή P-value).

Έστω ότι θέλουμε να υπολογίσουμε το 95% δ.ε. της διαφοράς των μέσων τιμών του “1” και “2”.

Από το μενού επιλογών, διαλέγουμε Analyze-> Compare Means-> Independent Samples T-test.



Στην συνέχεια επιλέγουμε από την Λίστα Μεταβλητών την μεταβλητή (time) για την οποία θα εφαρμόσουμε το t-test (Test Variable), και στην συνέχεια την μεταβλητή (drug) σύμφωνα με την οποία ομαδοποιούμε τα δεδομένα μας (Grouping Variable). Με την επιλογή Define Variables, εμφανίζεται η οθόνη Define Groups, στα πεδία της οποία καθορίζουμε τις τιμές που διαφοροποιούν τις διάφορες ομάδες – τα φάρμακα στην δικιά μας περίπτωση.



Εφόσον εφαρμόσουμε το t-test. Τα αποτελέσματα εμφανίζονται στο output. Υποθέτουμε ότι τα δύο φάρμακα έχουν ίδια διακύμανση. Τότε  $P=0.002$  δηλ.  $P<0.05$  οπότε διαφέρουν.

**Output4 - SPSS Viewer**

File Edit View Insert Format Analyze Graphs Utilities Window Help

Output

- T-Test
  - Title
  - Notes
  - Group Statistics
  - Independent

→ **T-Test**

**Group Statistics**

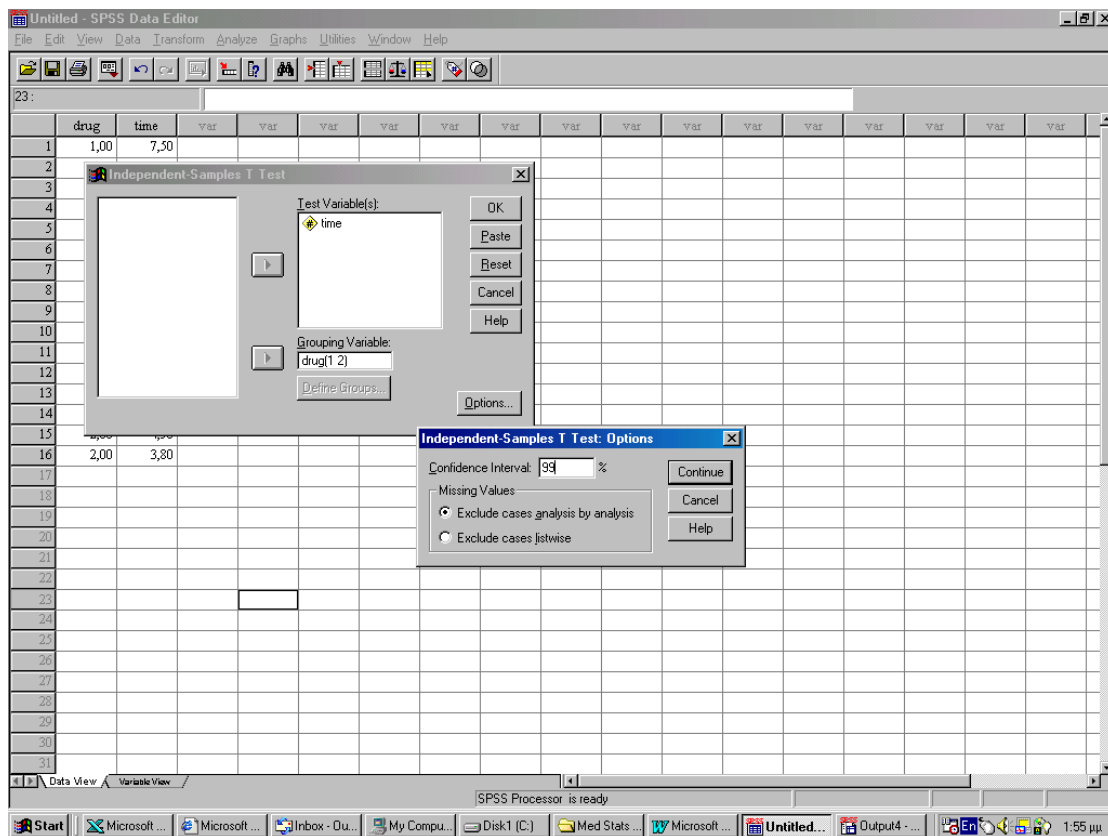
	DRUG	N	Mean	Std. Deviation	Std. Error Mean
TIME	1,00	8	7,2625	2,2847	,8078
	2,00	8	3,8000	1,2490	,4416

**Independent Samples Test**

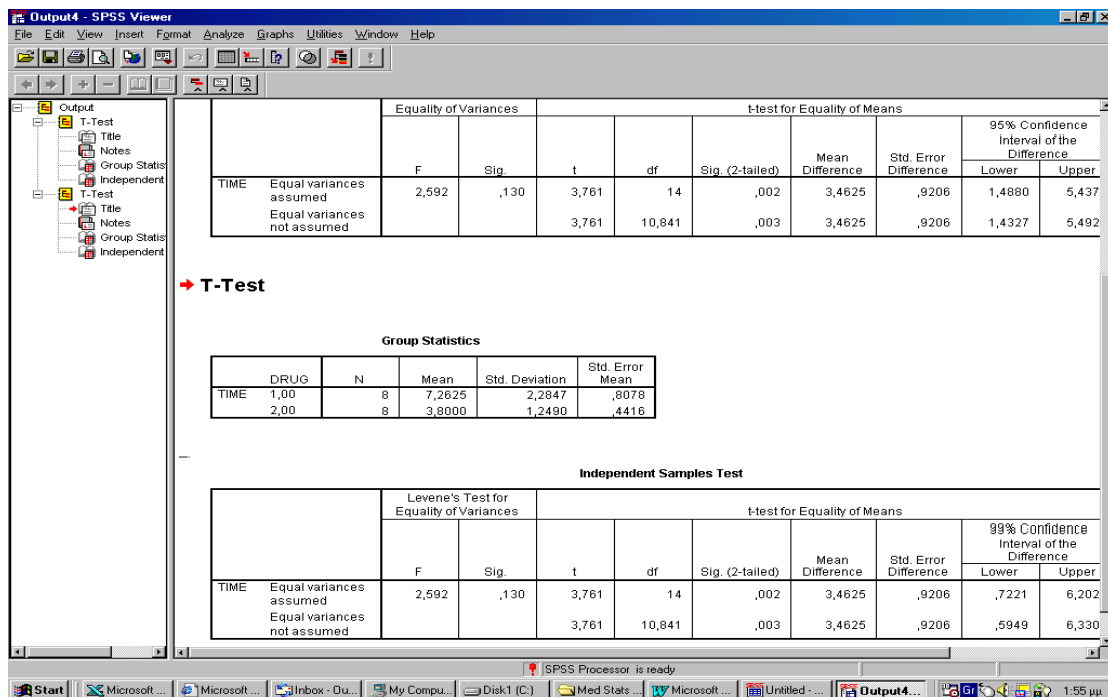
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
TIME	Equal variances assumed	2,592	,130	3,761	14	,002	3,4625	,9206	1,4880	5,4370
	Equal variances not assumed			3,761	10,841	,003	3,4625	,9206	1,4327	5,4923

Για να υπολογίσουμε το 99% δ.ε. της διαφοράς των μέσων τιμών, επιλέγουμε το "Options" και στην συνέχεια καθορίζουμε το δ.ε. που μας ενδιαφέρει.

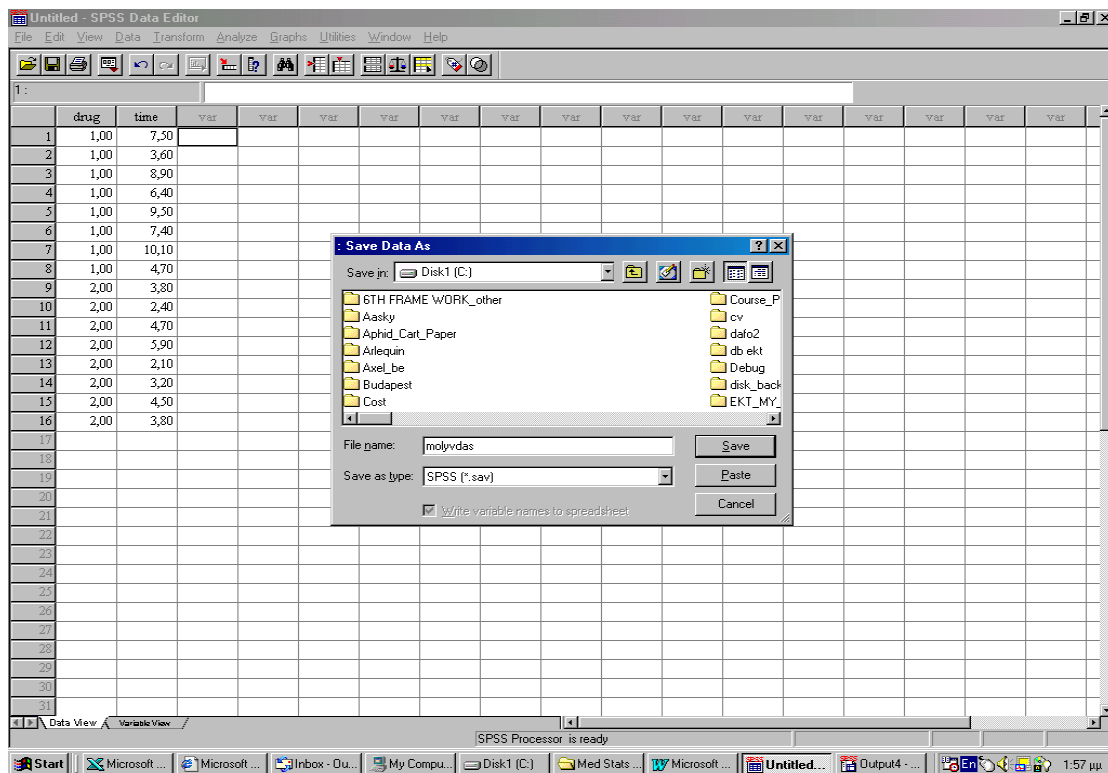




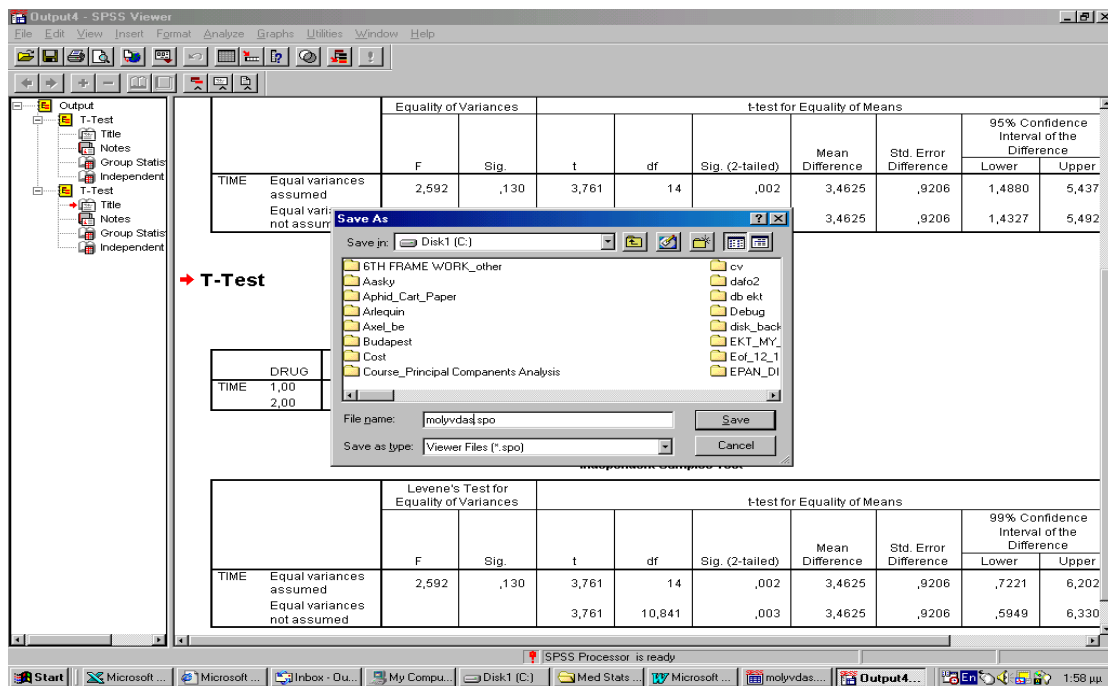
Τα αποτελέσματα εμφανίζονται στο output.



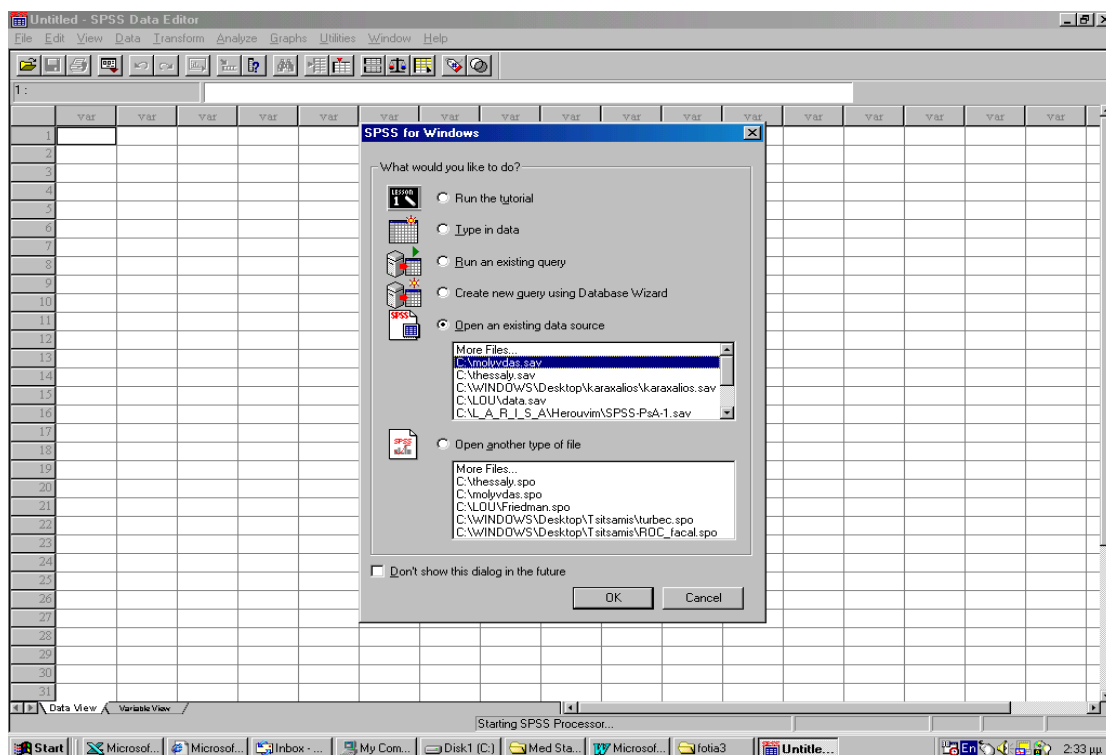
Αποθηκεύουμε ξανά τα δεδομένα σε ένα αρχείο για μελλοντική χρήση



Αποθηκεύουμε το output με τα αποτελέσματα της ανάλυσης σε κάποιο άλλο αρχείο για μελλοντική χρήση.



Για να ανακτήσουμε δεδομένα που έχουμε αποθηκεύσει, από το μενού επιλογών, διαλέγουμε File->Open. Στην συνέχεια επιλέγουμε το όνομα του αρχείου που θέλουμε- π.χ. “molyvdas.sav”.



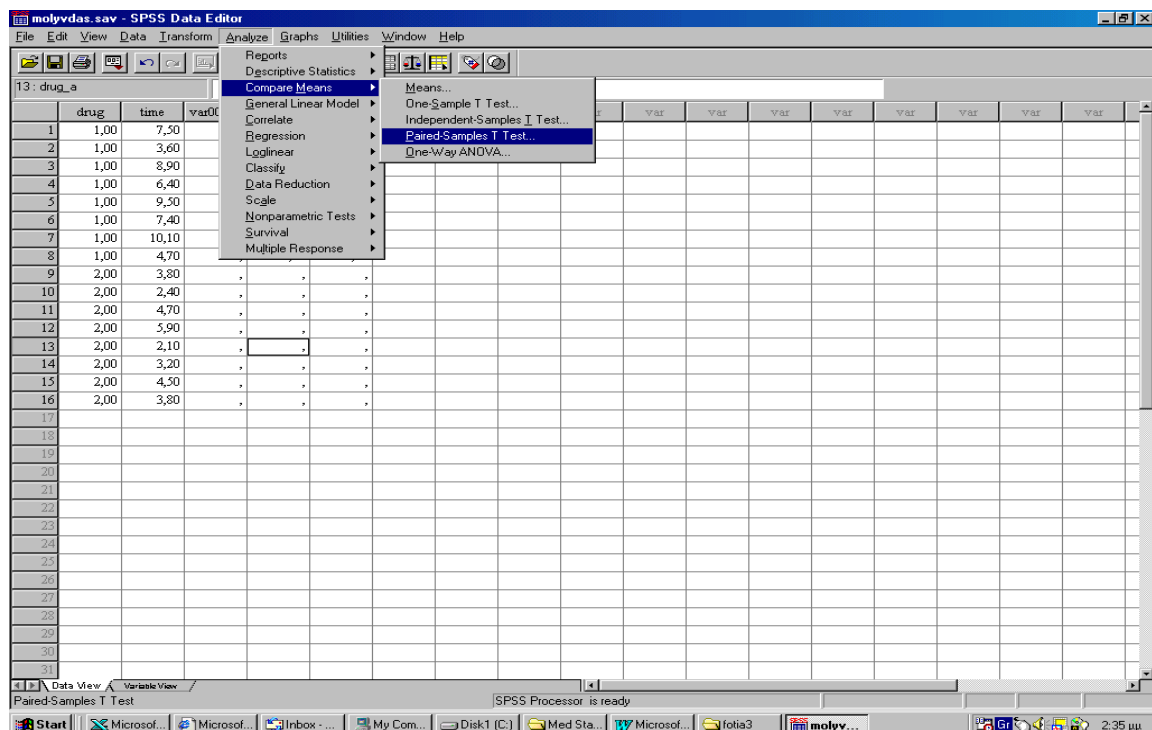
### t-test για ζευγαρωτές παρατηρήσεις

Έστω ότι θέλουμε να συγκρίνουμε τους χρόνους ανακούφισης δύο αναλγητικών φαρμάκων Α και Β. Έστω επίσης ότι σε 8 ασθενείς χορηγήθηκε πρώτα το φάρμακο Α και μετά το Β. Παρατηρήθηκαν οι εξής χρόνοι ανακούφισης:

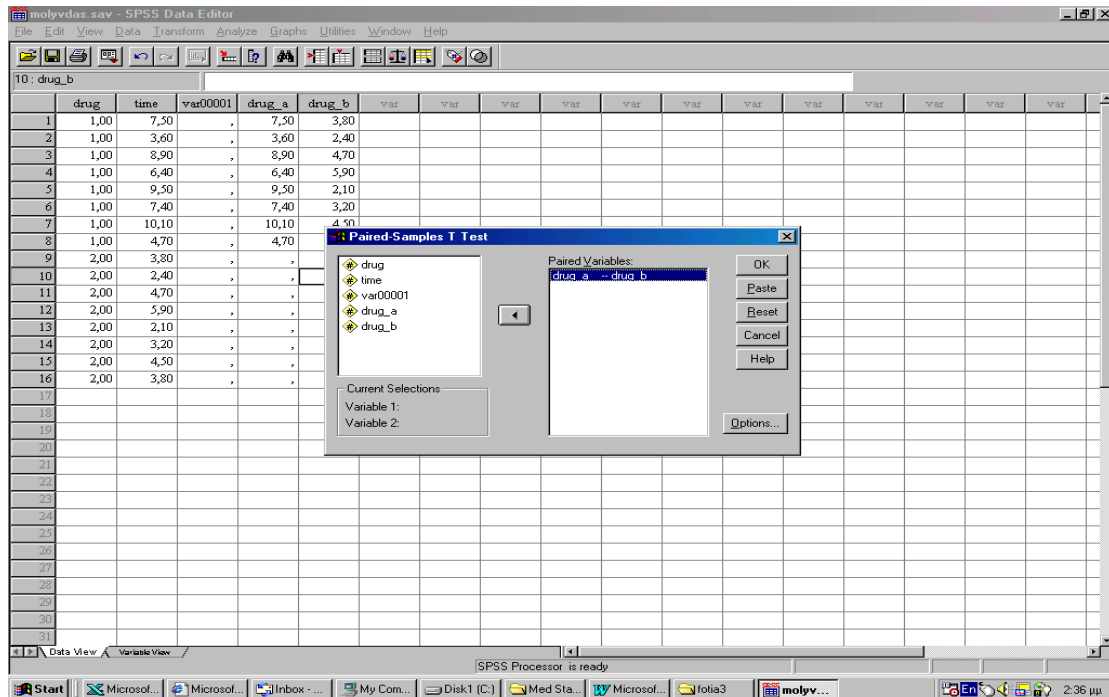
	drug	time	var00001	drug_a	drug_b	var00002	var00003	var00004	var00005	var00006	var00007	var00008	var00009	var00010	var00011	var00012	var00013
1	1,00	7,50	,	7,50	3,80												
2	1,00	3,60	,	3,60	2,40												
3	1,00	8,90	,	8,90	4,70												
4	1,00	6,40	,	6,40	5,90												
5	1,00	9,50	,	9,50	2,10												
6	1,00	7,40	,	7,40	3,20												
7	1,00	10,10	,	10,10	4,50												
8	1,00	4,70	,	4,70	3,80												
9	2,00	3,80	,	,	,												
10	2,00	2,40	,	,	,												
11	2,00	4,70	,	,	,												
12	2,00	5,90	,	,	,												
13	2,00	2,10	,	,	,												
14	2,00	3,20	,	,	,												
15	2,00	4,50	,	,	,												
16	2,00	3,80	,	,	,												
17																	
18																	
19																	
20																	
21																	
22																	
23																	
24																	
25																	
26																	
27																	
28																	
29																	
30																	
31																	

Θέλουμε να ελέγξουμε αν υπάρχει διαφορά στην αποτελεσματικότητα των Α και Β με το t-test για ζευγαρωτές παρατηρήσεις.

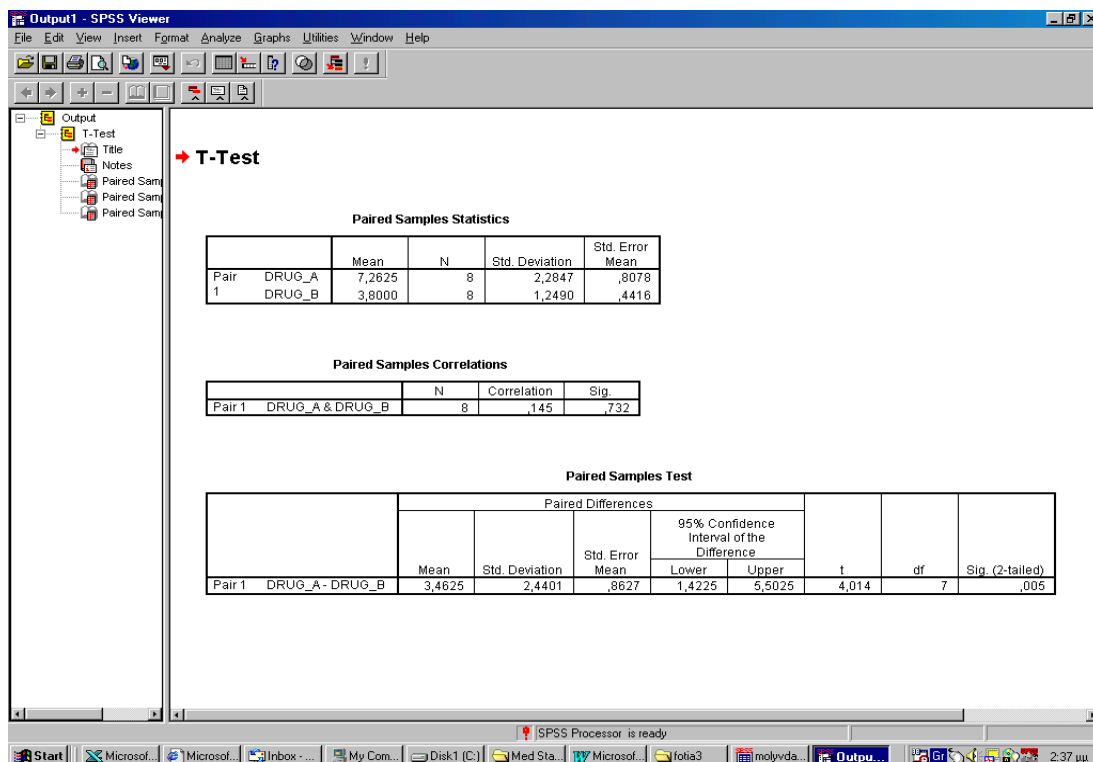
Από το μενού επιλογών διαλέγουμε Analyze->“Compare Means-> Paired-Samples T-Test.



Στην συνέχεια στην οθόνη Paired Samples T-Test, διαλέγουμε τις μεταβλητές που παριστάνουν τις ζευγαρωτές παρατηρήσεις.



Τα αποτελέσματα εμφανίζονται στο output. Παρατηρούμε ότι τα φάρμακα διαφέρουν με  $P=0.005$  ή  $P<0.05$ . Το 95% δ.ε. είναι (1.42, 5.50) που δεν περιέχει το 1.



### Έλεγχος για τη σύγκριση δύο ποσοστών -Πίνακες συνάφειας - $\chi^2$ test

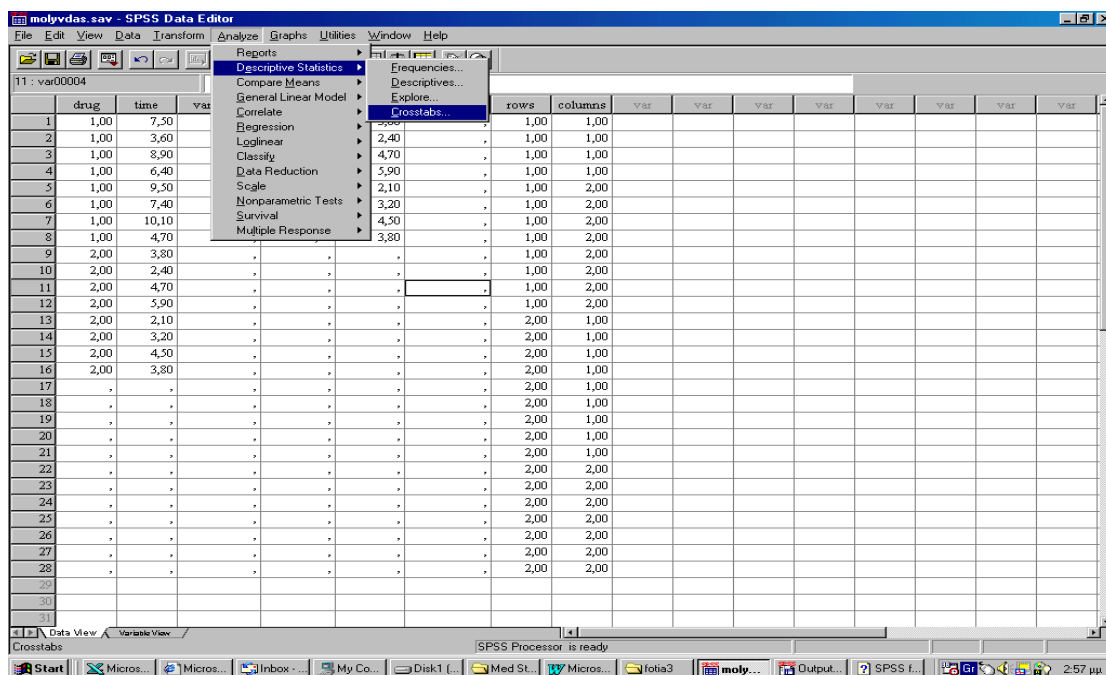
Ας υποθέσουμε ότι σε μια μελέτη 28 άτομα έλαβαν το φάρμακο Α ή Β. Από τα 12 άτομα που έλαβαν το φάρμακο Α, θεραπεύτηκαν 4 και από τα 16 άτομα που έλαβαν το φάρμακο Β, θεραπεύτηκαν 9. Θέλουμε να ελέγξουμε εάν υπάρχει διαφορά μεταξύ των δύο φαρμάκων και εάν υπάρχει σχέση μεταξύ φαρμάκου και θεραπευτικού αποτελέσματος.

Τα δεδομένα έχουν τη μορφή ενός 2x2 τραπεζίου

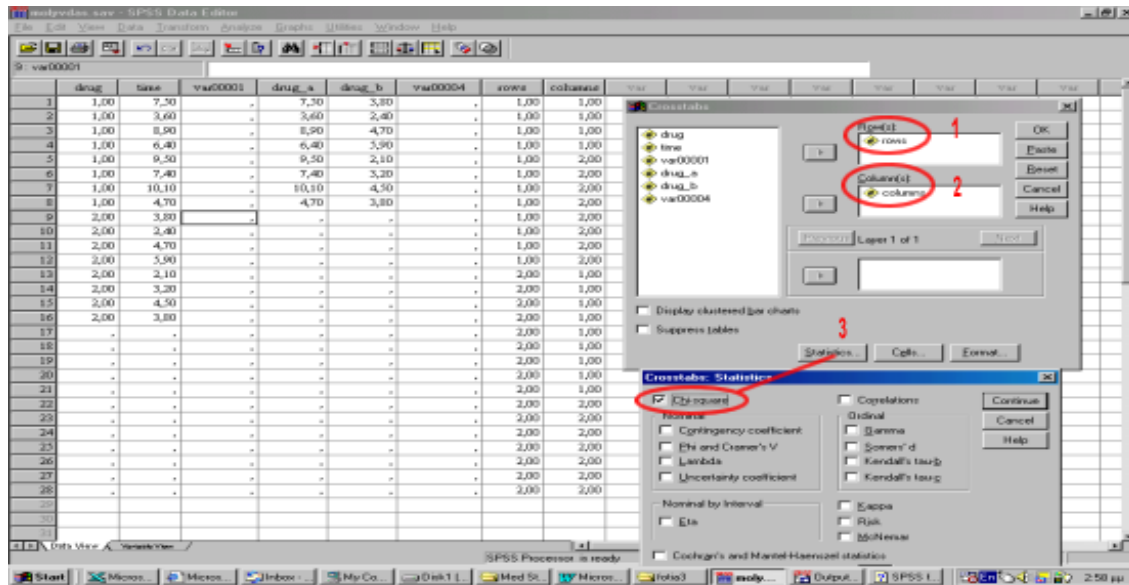
Φάρμακο		<u>Cured=1 not cured=2</u>	
	A=1	4	8
	B=2	9	7

και καταχωρούνται στις μεταβλητές με τα ονόματα “rows” και “columns” όπως παρακάτω.

Για να ελέγξουμε την παραπάνω υπόθεση θα εφαρμόσουμε ένα  $\chi^2$ -test. Από το μενού επιλογών, διαλέγουμε Analyze-> Descriptive Statistics-> Crosstabs .



Από την Λίστα Μεταβλητών, διαλέγουμε τις μεταβλητές για τις γραμμές και τις στήλες και στην συνέχεια από την επιλογή Statistics, διαλέγουμε Chi-Square.



Τα αποτελέσματα εμφανίζονται στο output. Παρατηρούμε ότι το  $\chi^2$  είναι 1.448 που είναι σημαντικό σε  $P=0.229$ , δηλ.  $P>0.05$ , οπότε δεν υπάρχει σχέση.

**Output1 - SPSS Viewer**

File Edit View Insert Format Analyze Graphs Utilities Window Help

Output

- T-Test
- Title
- Notes
- Paired Samples Statistics
- Paired Samples Correlations
- Paired Samples Test
- Crosstabs
  - Title
  - Notes
  - Case Processing Summary
  - ROWS \* COLUMNS Crosstab
  - Chi-Square Tests

**Crosstabs**

**Case Processing Summary**

	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
ROWS * COLUMNS	28	100,0%	0	,0%	28	100,0%

**ROWS \* COLUMNS Crosstabulation**

Count		COLUMNS		Total
		1,00	2,00	
ROWS	1,00	4	8	12
	2,00	9	7	16
Total		13	15	28

**Chi-Square Tests**

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	1,448 <sup>a</sup>	1	,229		
Continuity Correction <sup>a</sup>	,673	1	,412		
Likelihood Ratio	1,467	1	,226		
Fisher's Exact Test				,276	,207
Linear-by-Linear Association	1,396	1	,237		
N of Valid Cases	28				

a. Computed only for a 2x2 table

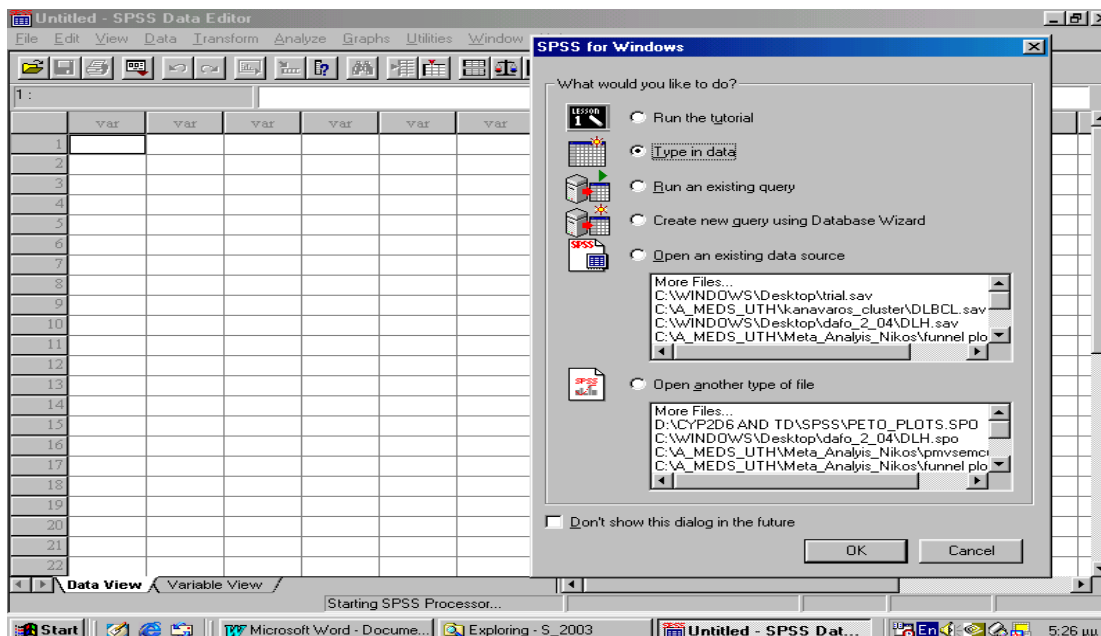
1 items selected (1 hidden/collapsed)

SPSS Processor is ready

Start | Micros... | Micros... | Inbox... | My Co... | Disk1 (... | Med St... | Micros... | fotia3 | molyvd... | Outp... | SPSS f... | 3:00 μμ

## Mann-Whithey-Wilcoxon test για τη σύγκριση δύο ανεξάρτητων δειγμάτων

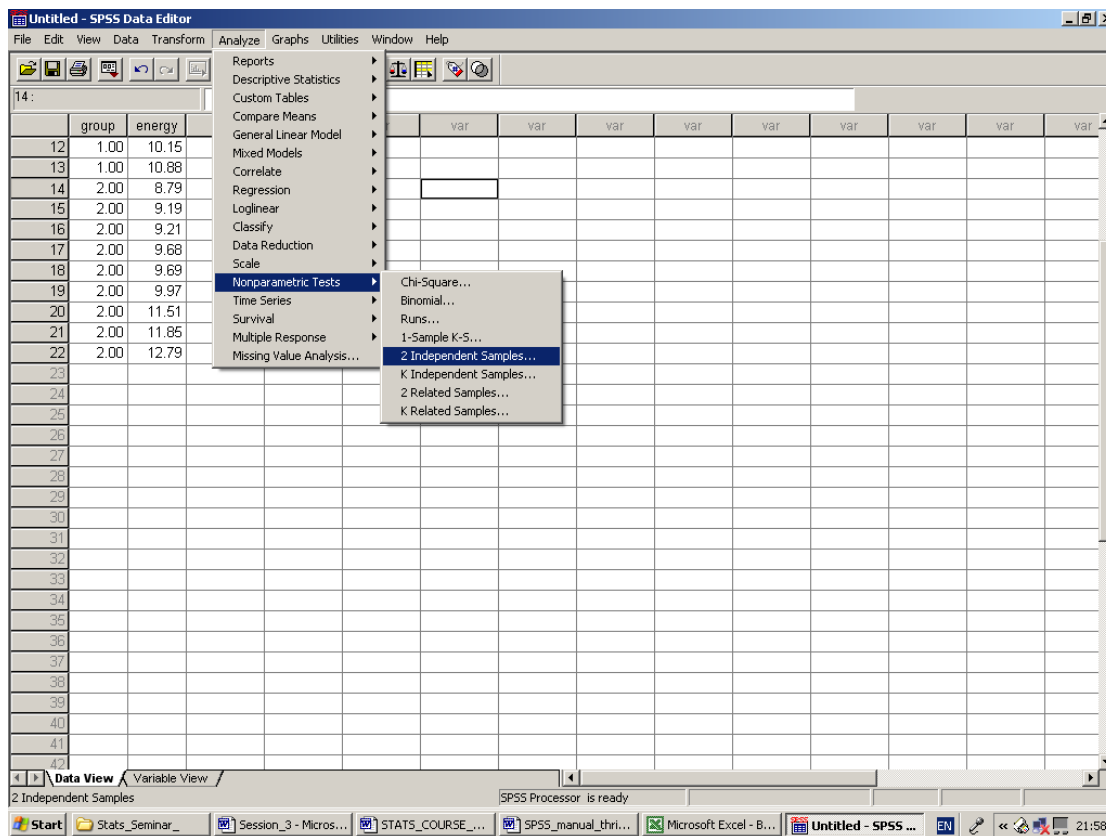
Ανοίξτε το “SPSS” από το “Start” στο Desktop του H/Y και επιλέξτε “Type in data”. Τώρα μπορείτε να εισάγεται τα δεδομένα στο “Data View”.



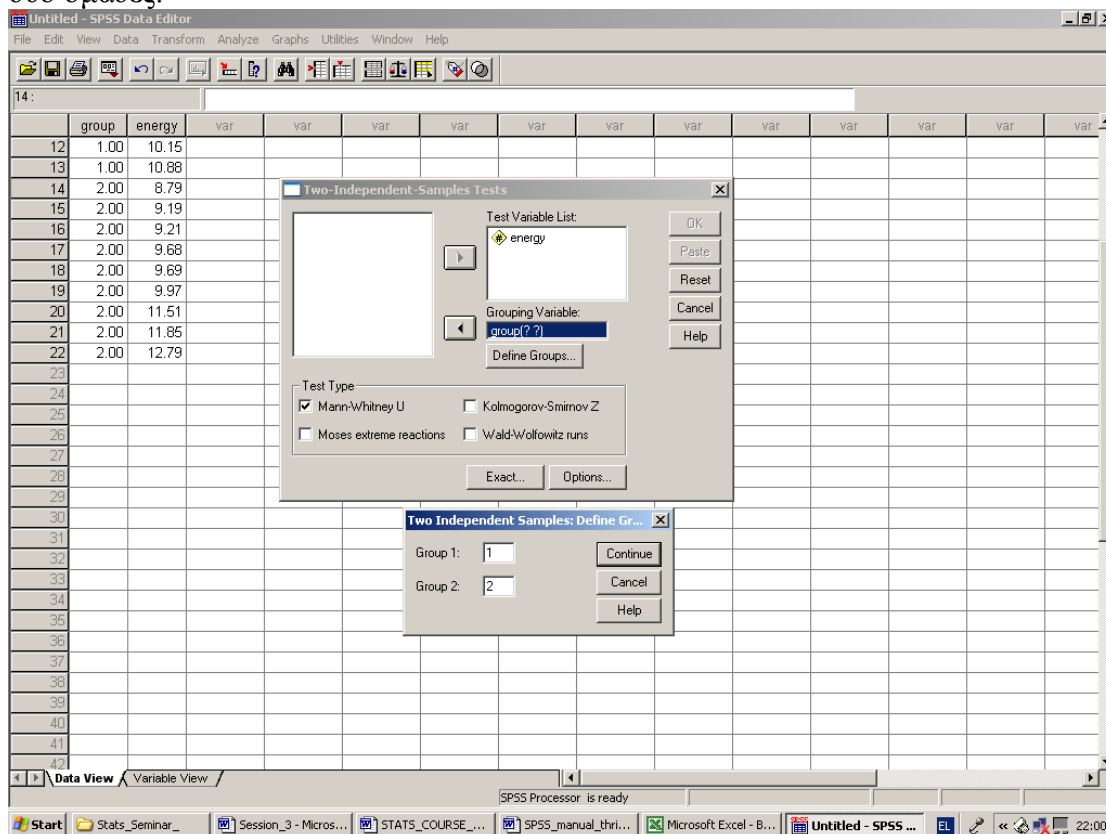
Εισάγετε τα δεδομένα της άσκησης στο “Data View” όπως φαίνεται παρακάτω. Μία μεταβλητή αντιπροσωπεύει τις ομάδες και μία άλλη την αντίστοιχη κατανάλωση ενέργειας, μετά ονομάστε τις μεταβλητές “group” και “energy”, αντίστοιχα, μέσα από το “Variable View”.

Για να αναλύσετε τα δεδομένα επιλέξτε “Nonparametric Tests” και “2 Independent Samples”.





Τοποθετείστε στο πεδίο “Test Variable List:” την μεταβλητή “group” και στο πεδίο “Grouping Variable” τη μεταβλητή “group”. Επιλέξτε και εισάγετε 1 και 2 αντίστοιχα για τις δύο ομάδες.



Τότε εμφανίζεται το παράθυρο με τα αποτελέσματα. Μπορείτε να αποθηκεύσετε τα αποτελέσματα σε ένα αρχείο για μελλοντική χρήση.

The screenshot shows the SPSS Output Viewer window titled "Output4 - SPSS Viewer". The left pane displays a tree view of the output structure: Output > NPar Tests > Mann-Whitney Test > Test Statistics. The main pane shows the "Mann-Whitney Test" results.

**Ranks**

	GROUP	N	Mean Rank	Sum of Ranks
ENERGY	1.00	13	7.92	103.00
	2.00	9	16.67	150.00
Total		22		

**Test Statistics<sup>b</sup>**

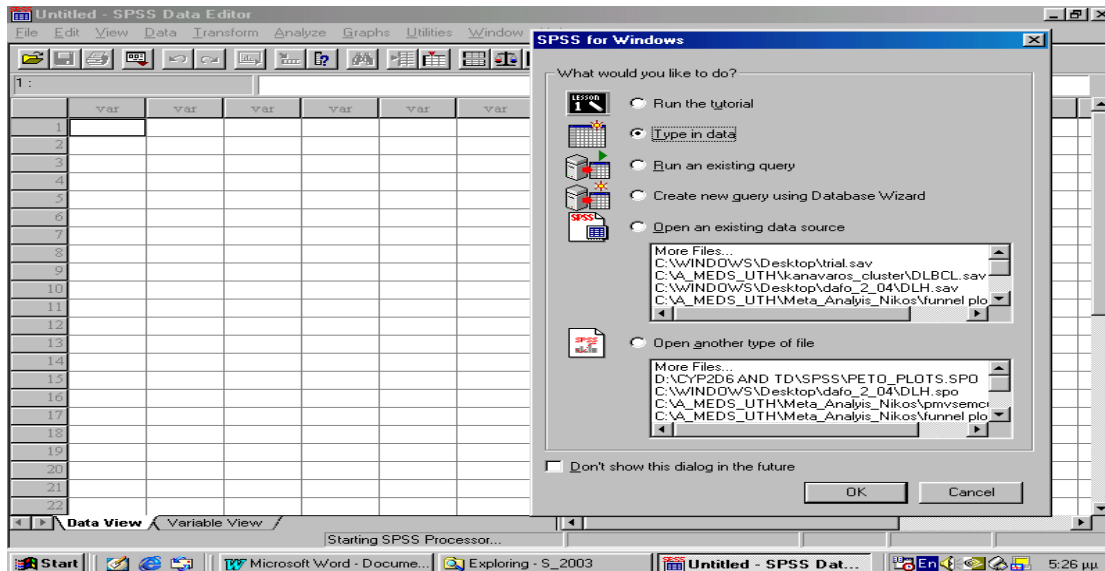
	ENERGY
Mann-Whitney U	12.000
Wilcoxon W	103.000
Z	-3.106
Asymp. Sig. (2-tailed)	.002
Exact Sig. [Z*(1-tailed Sig.)]	.001 <sup>a</sup>

a. Not corrected for ties.  
b. Grouping Variable: GROUP

A "Save As" dialog box is open over the bottom right of the main window. It shows the "Save in:" location as "Desktop". The file name is "Output4" and the "Save as type" is "Viewer Files (\*.spo)".

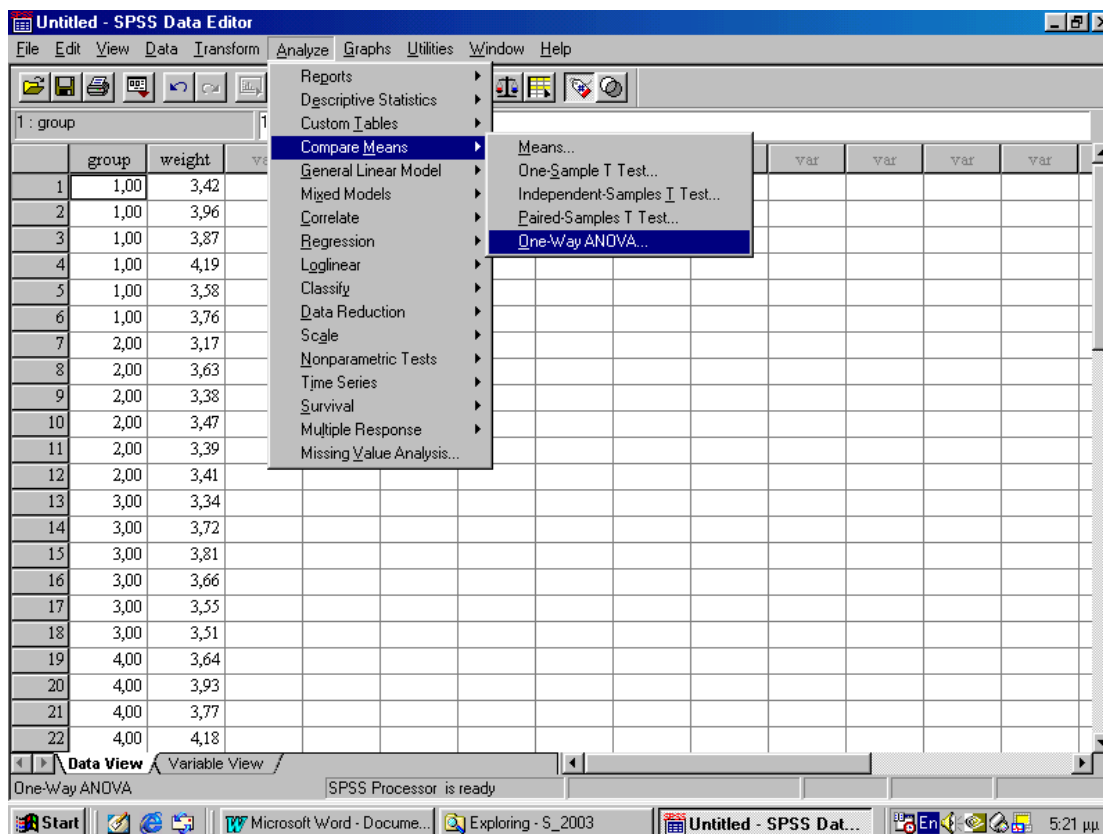
## Ανάλυση διασποράς με ένα παράγοντα (One-way ANOVA)

Ανοίξτε το “SPSS” από το “Start” στο Desktop του Η/Υ και επιλέξτε “Type in data”. Τώρα μπορείτε να εισάγεται τα δεδομένα στο “Data View”.

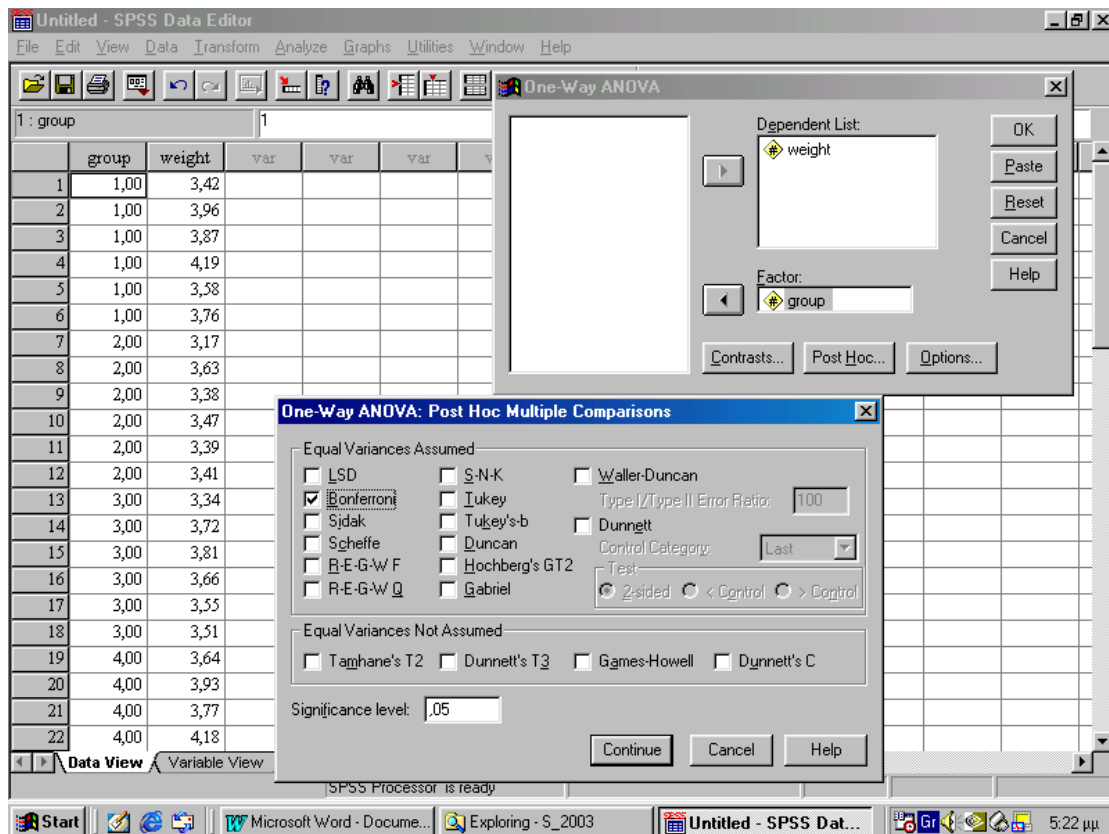


Εισάγετε τα δεδομένα της άσκησης στο “Data View” όπως φαίνεται παρακάτω. Μία μεταβλητή αντιπροσωπεύει τις ομάδες και μία άλλη τα αντίστοιχα βάρη, μετά ονομάστε τις μεταβλητές “group” και “weight”, αντίστοιχα, μέσα από το “Variable View”.

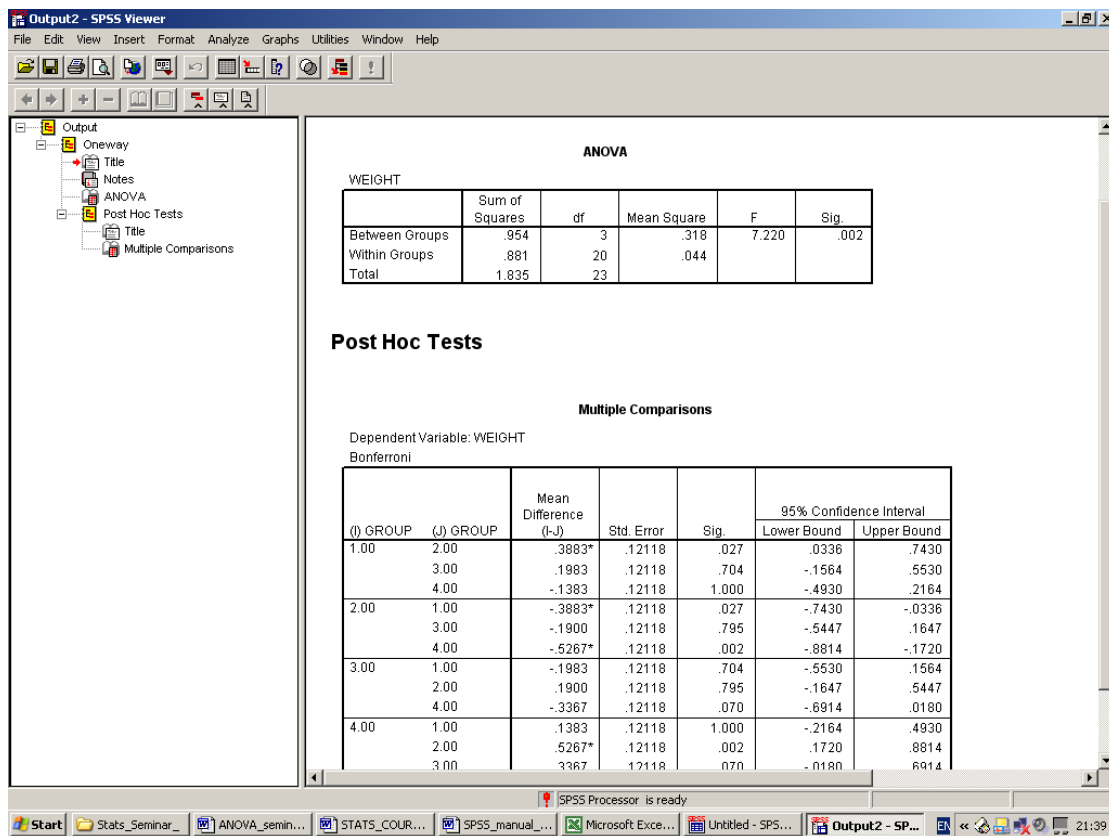
Για να αναλύσετε τα δεδομένα επιλέξτε “Analyse” “Compare Means” “One Way ANOVA”



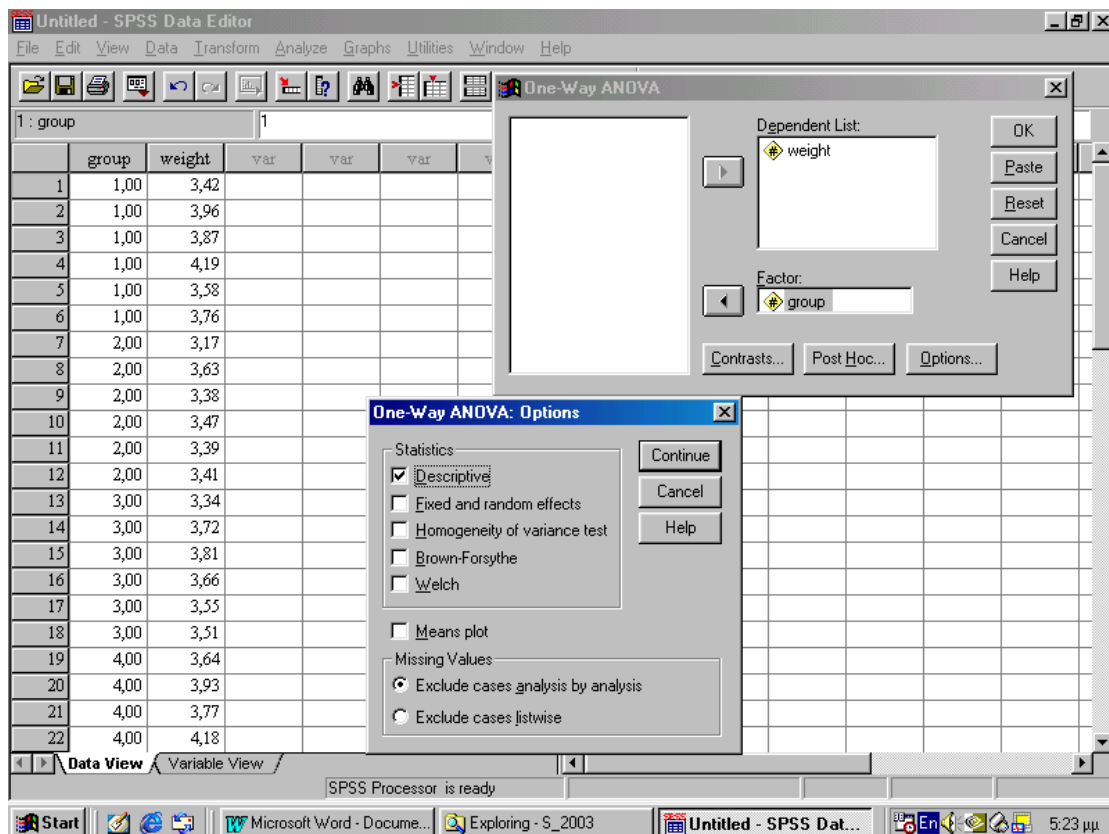
Στο παράθυρο “One-Way ANOVA” επιλέξτε τη μεταβλητή “weight” στο “Dependent List” και τη μεταβλητή “group” στο “Factor”. Μετά επιλέξτε την επιλογή “Post Hoc”. Τότε εμφανίζεται το παράθυρο “One-Way ANOVA Post Hoc Multiple Comparisons”, επιλέξτε την επιλογή “Bonferroni” και μετά “Continue”. Τέλος επιλέξτε “Ok” για να δείτε τα αποτελέσματα της ανάλυσης.



Το “Output” παράθυρο εμφανίζει την ANOVA και τα αποτελέσματα των πιθανών συγκρίσεων μεταξύ των ομάδων.



Μετά επιλέξετε “Options” “Descriptives” για να υπολογίσετε περιγραφικά στατιστικά για την κάθε ομάδα.



Τότε εμφανίζεται το παράθυρο με τα αποτελέσματα. Μπορείτε να αποθηκεύσετε τα αποτελέσματα σε ένα αρχείο για μελλοντική χρήση.

**Output3 - SPSS Viewer**

File Edit View Insert Format Analyze Graphs Utilities Window Help

**Output**

- Oneway
  - Title
  - Notes
  - ANOVA
  - Post Hoc Tests
    - Title
    - Multiple Comparisons
- Oneway
  - Title
  - Notes
  - Descriptives
  - ANOVA
  - Post Hoc Tests
    - Title
    - Multiple Comparisons

**Oneway**

**Descriptives**

WEIGHT

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
1.00	6	3.7967	.27457	.11209	3.5085	4.0848	3.42	4.19
2.00	6	3.4083	.14892	.06080	3.2521	3.5646	3.17	3.63
3.00	6	3.5983	.16750	.06838	3.4226	3.7741	3.34	3.81
4.00	6	3.9350	.22492	.09182	3.6990	4.1710	3.64	4.21
Total	24	3.6846	.28247	.05766	3.5653	3.8039	3.17	4.21

**ANOVA**

Source	df	Mean Square	F	Sig.
Between Groups	3	.318	7.220	.002
Within Groups	20	.044		
Total	23			

**Multiple Comparisons**

**Save As**

Save in: Desktop

My Documents, My Computer, My Network Places, Cornelis\_Zintzaras, Debug, DOCUMENTS, Endocrine Reviews -- Wu and von Eckardstein 24 (2) 183.files, kanavaros\_cluster, kukulis\_e, Med\_application\_DEP, Release, Sage, Stats\_Seminar, Output1, Shortcut to H...

File name: Output3

Save as type: Viewer Files (\*.spo)

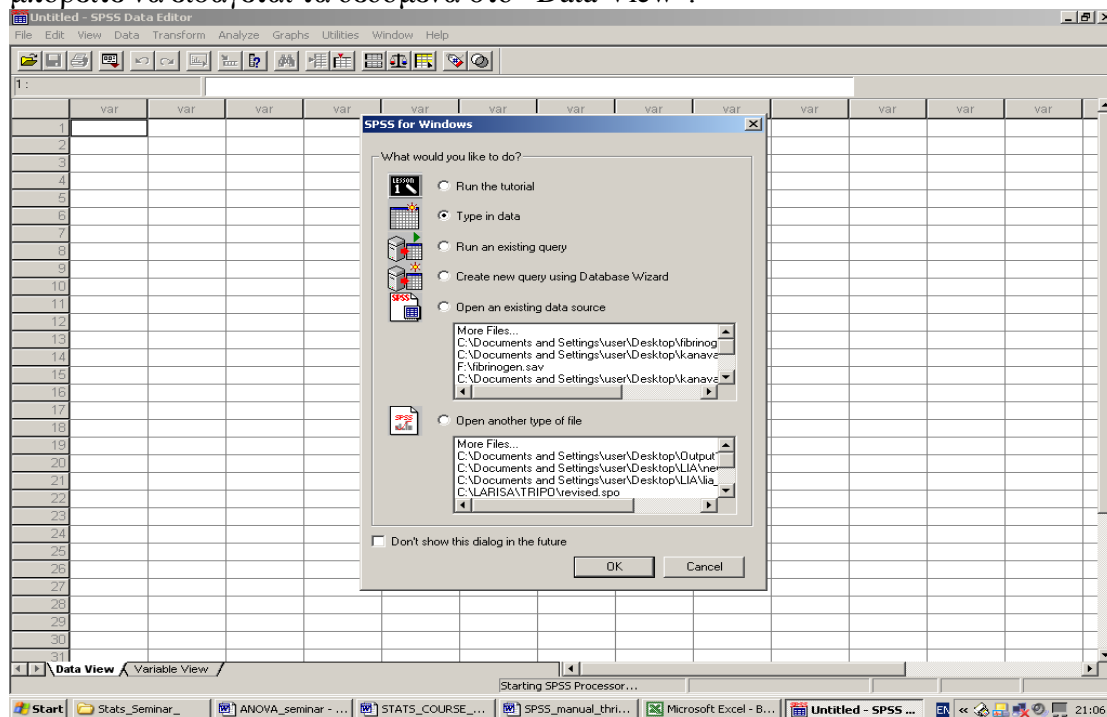
Save, Cancel

SPSS Processor is ready

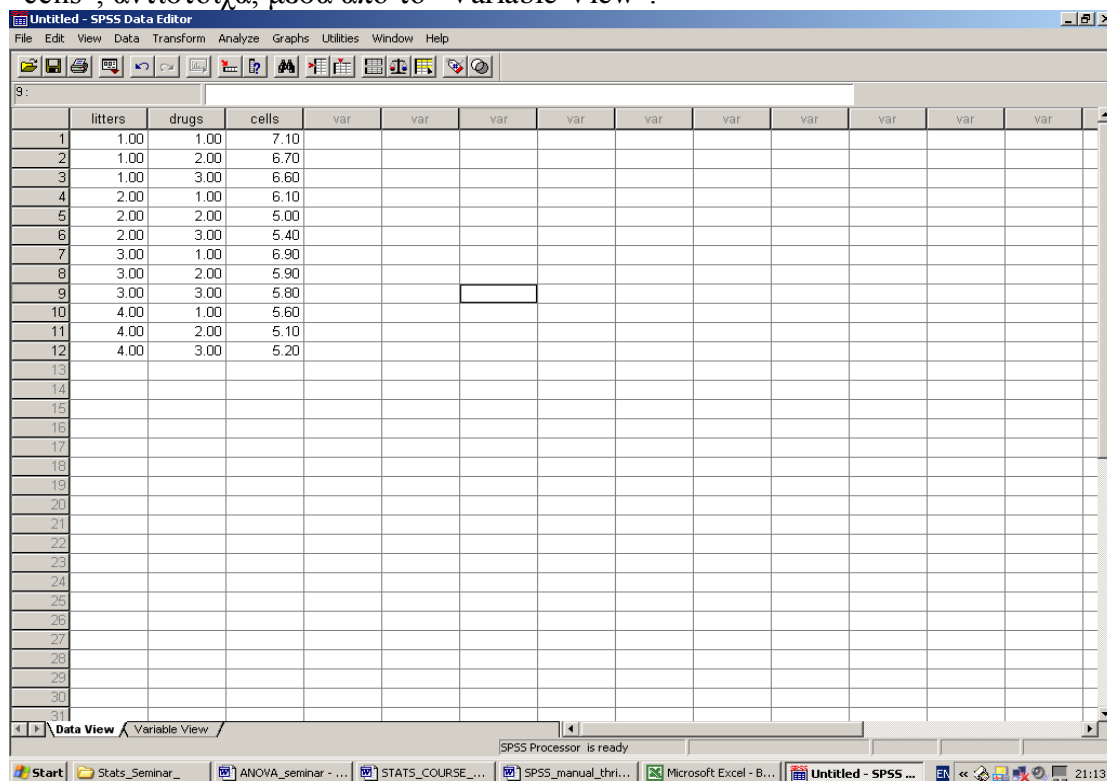
Start, Stats\_Seminar..., ANOVA\_semin..., STATS\_COUR..., SPSS\_manual..., Microsoft Exce..., Untitled - SPS..., Output3 - SP..., 21:45

## Ανάλυση διασποράς με δύο παράγοντες (Two-way ANOVA) χωρίς αλληλεπίδραση

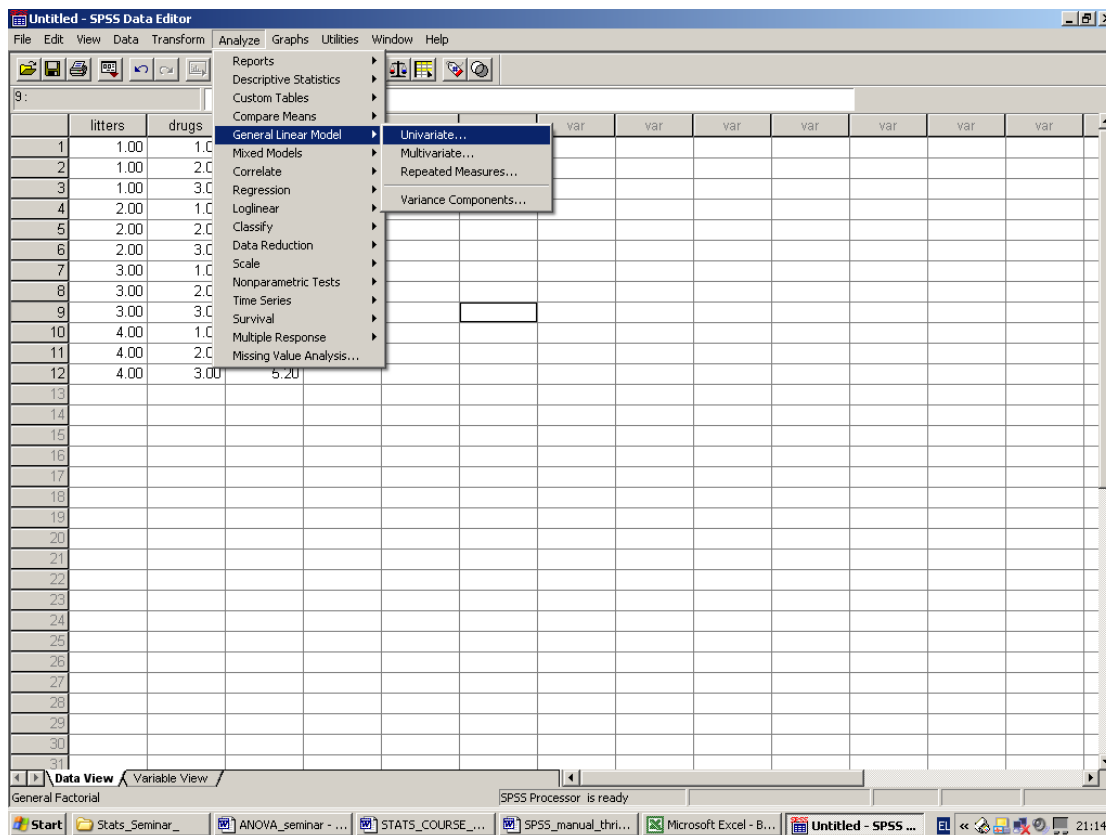
Ανοίξτε το “SPSS” από το “Start” στο Desktop του H/Y και επιλέξτε “Type in data”. Τώρα μπορείτε να εισάγετε τα δεδομένα στο “Data View”.



Εισάγετε τα δεδομένα της άσκησης στο “Data View” όπως φαίνεται παρακάτω. Μία μεταβλητή αντιπροσωπεύει τα κλουβιά (litters), μία άλλη τα φάρμακα και μία άλλη τον αντίστοιχο αριθμό λεμφοκυττάρων, μετά ονομάστε τις μεταβλητές “litters”, “drugs” και “cells”, αντίστοιχα, μέσα από το “Variable View”.

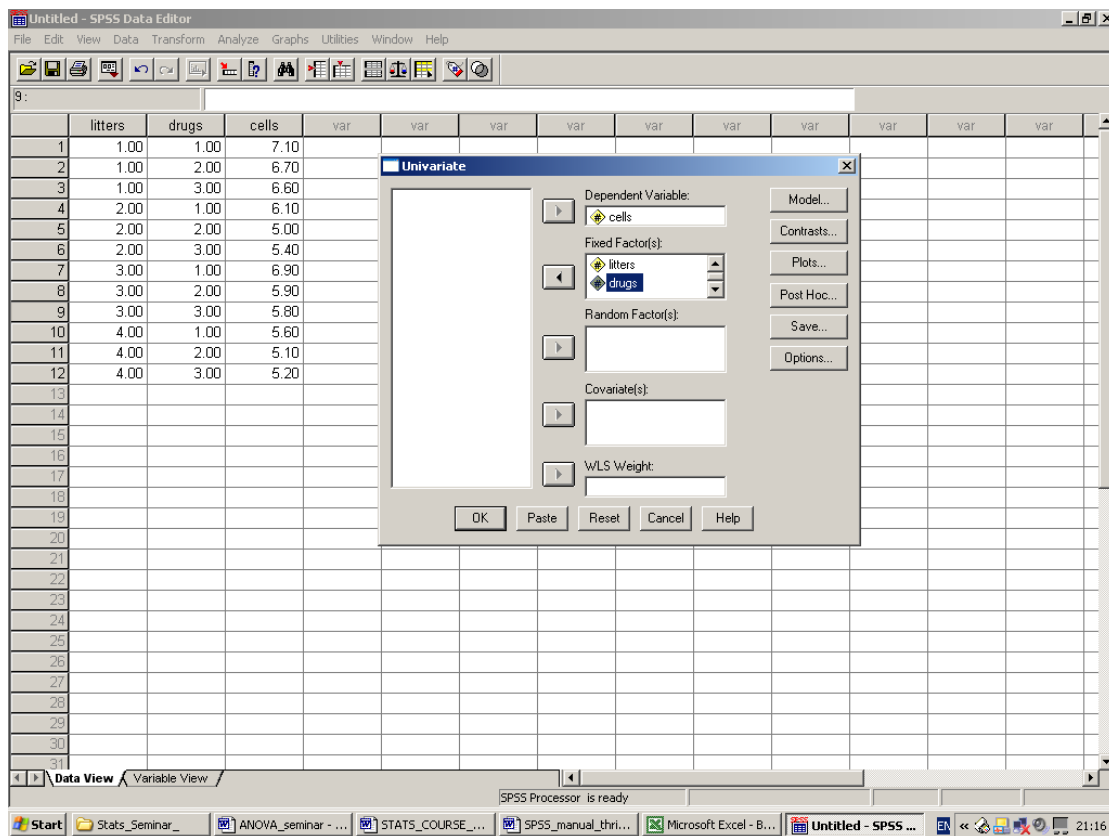


Για να αναλύσετε τα δεδομένα επιλέξτε “Analyse”, “General Linear Model”, “Univariate”

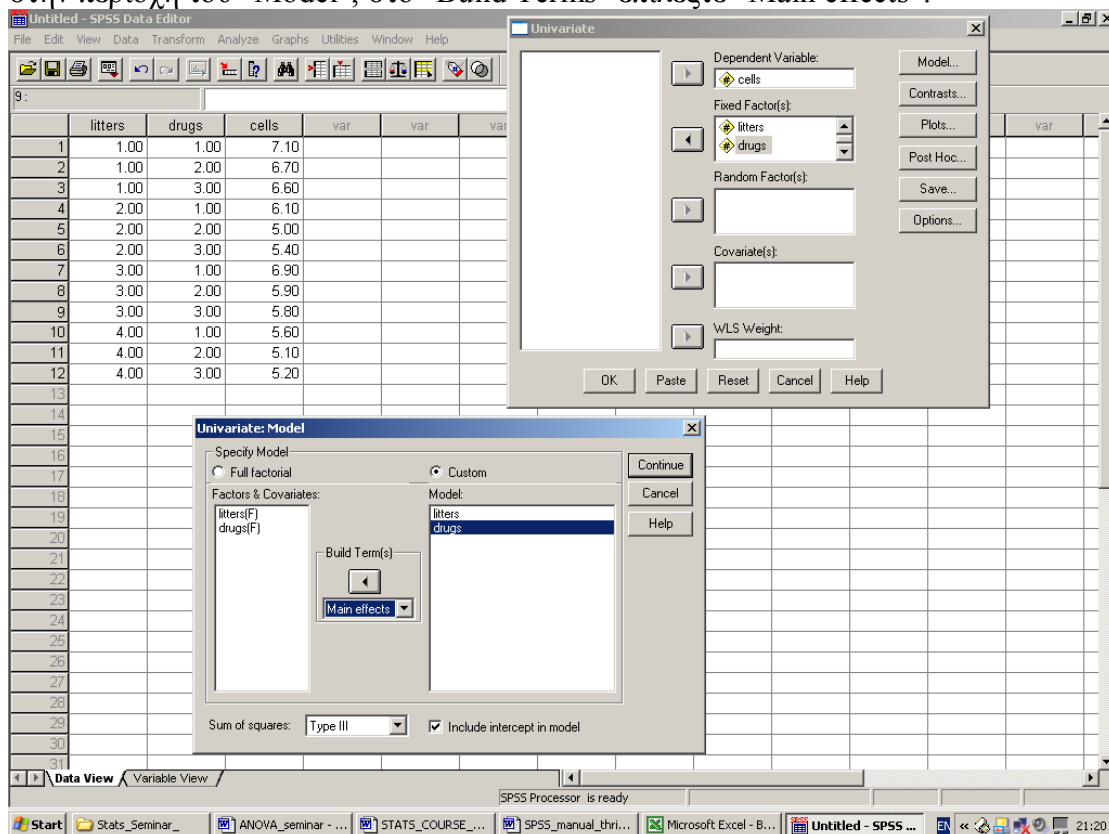


Δηλώστε ως “Dependent Variable” τη μεταβλητή “cells” και σαν “Fixed Factor(s)” τις μεταβλητές “litters” και “drugs”.

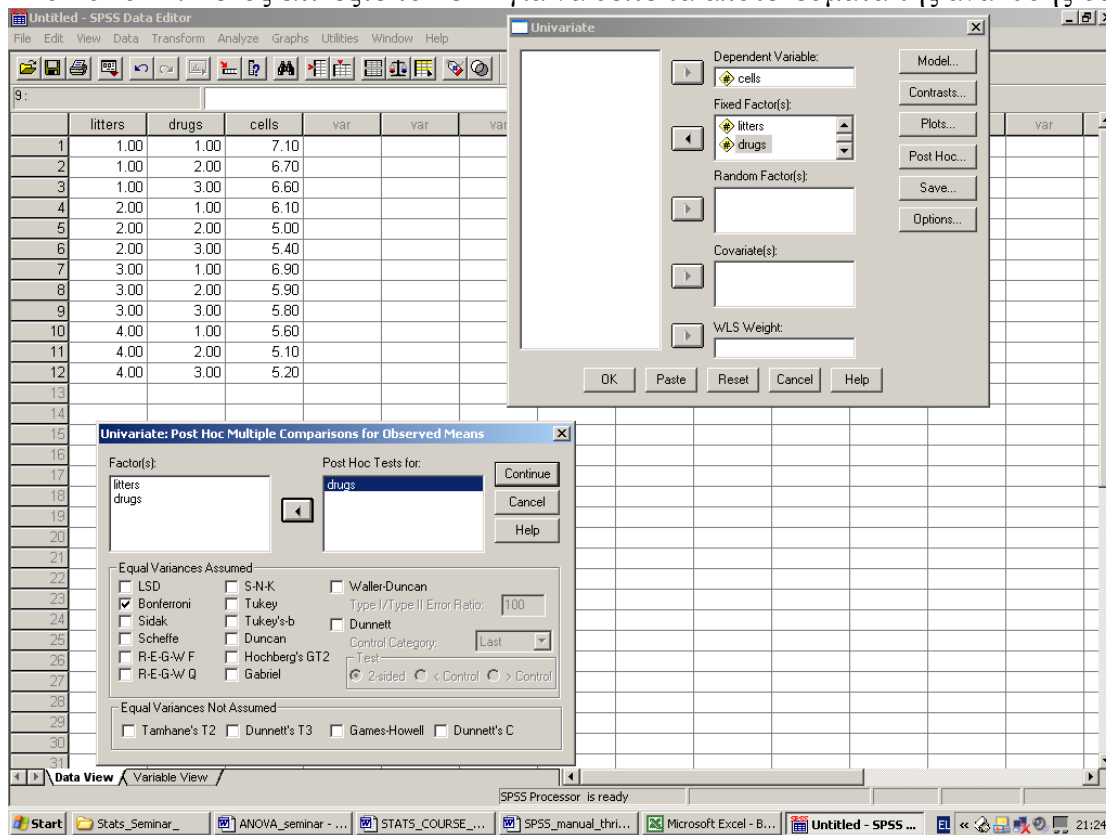




Επιλέξτε “Model” και μετά “Custom”. Τοποθετείστε τις μεταβλητές “litters” και “drugs” μέσα στην περιοχή του “Model”, στο “Build Terms” επιλέξτε “Main effects”.



Επιλέξτε την επιλογή “Post Hoc” και τοποθετείστε στο παράθυρο που θα εμφανισθεί μέσα στο πεδίο “Post Hoc Tests for:” την μεταβλητή “drugs”. Κατόπιν επέλεξε την επιλογή “Bonferroni”. Τέλος επιλέξτε το “Ok” για να δείτε τα αποτελέσματα της ανάλυσής σας.



Τα αποτελέσματα της ανάλυσης μπορείτε να τα αποθηκεύσετε σε ένα αρχείο για μελλοντική χρήση.

Output1 - SPSS Viewer

File Edit View Insert Format Analyze Graphs Utilities Window Help

Output

- Univariate Analysis of Variance
  - Title
  - Notes
  - Between-Subjects Factors
  - Tests of Between-Subjects Effects
    - Post Hoc Tests
      - Title
      - DRUGS
        - Title
        - Multiple Comparisons

Tests of Between-Subjects Effects

Dependent Variable: CELLS

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	5.595 <sup>a</sup>	5	1.119	24.415	.001
Intercept	424.830	1	424.830	9269.018	.000
LITTERS	4.230	3	1.410	30.764	.000
DRUGS	1.365	2	.683	14.891	.005
Error	.275	6	4.583E-02		
Total	430.700	12			
Corrected Total	5.870	11			

(Adjusted R Squared = .914)

Save As

Save in: Desktop

File name: Output1

Save as type: Viewer Files (\*.spo)

Multiple Comparisons

CELLS

(I) DRUGS	(J) DRUGS	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	.7500*	.15138	.008	.2523	1.2477
	3.00	.6750*	.15138	.013	.1773	1.1727
2.00	1.00	-.7500*	.15138	.008	-1.2477	-.2523
	3.00	-.0750	.15138	1.000	-.5727	.4227

SPSS Processor is ready

Τις μεταβλητές με τα δεδομένα τους μπορείτε να τις αποθηκεύσετε σε ένα αρχείο για μελλοντική χρήση.

Untitled - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

9:

	litters	drugs	cells	var	var	var	var	var	var	var	var	var	var
1	1.00	1.00	7.10										
2	1.00	2.00	6.70										
3	1.00	3.00	6.60										
4	2.00	1.00	6.10										
5	2.00	2.00	5.00										
6	2.00	3.00	5.40										
7	3.00	1.00	6.90										
8	3.00	2.00	5.90										
9	3.00	3.00	5.80										
10	4.00	1.00	5.60										
11	4.00	2.00	5.10										
12	4.00	3.00	5.20										
13													
14													
15													
16													
17													
18													
19													
20													
21													
22													
23													
24													
25													
26													
27													
28													
29													
30													
31													

Save Data As

Save in: Desktop

File name: elkad

Save as type: SPSS (\*.sav)

☒ Write variable names to spreadsheet

SPSS Processor is ready

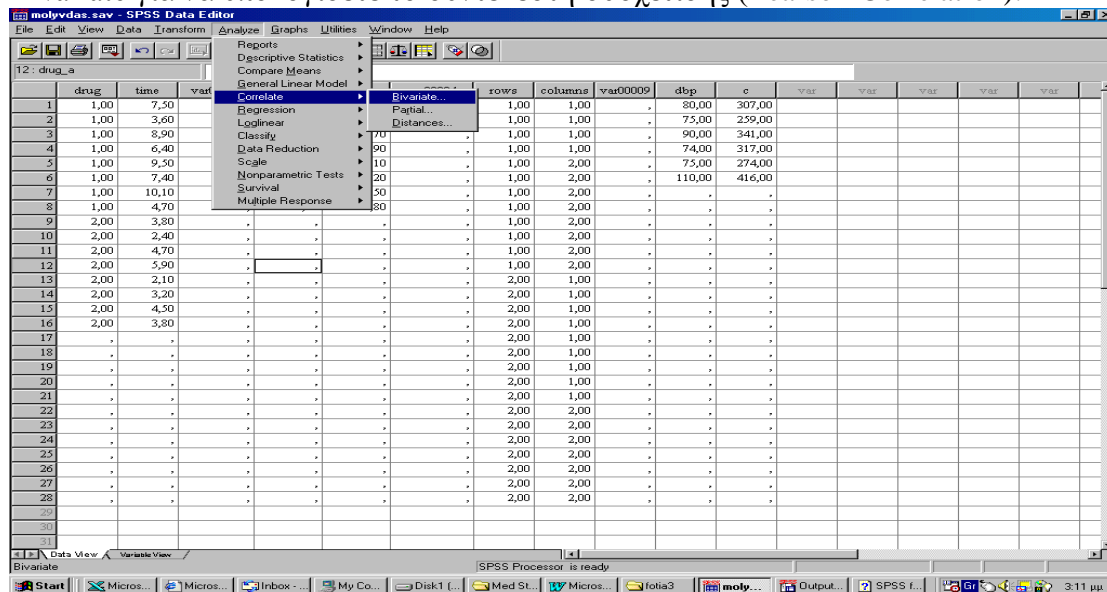
## Συντελεστής συσχέτισης (correlation coefficient) $r$ - Γραμμή γραμμικής παλινδρόμησης (regression line)

Ανοίξτε ένα από τα αρχεία δεδομένων που έχετε ήδη αποθηκεύσει (πχ molyndas.sav) και στο ίδιο αρχείο καταχωρήστε τα νέα δεδομένα.

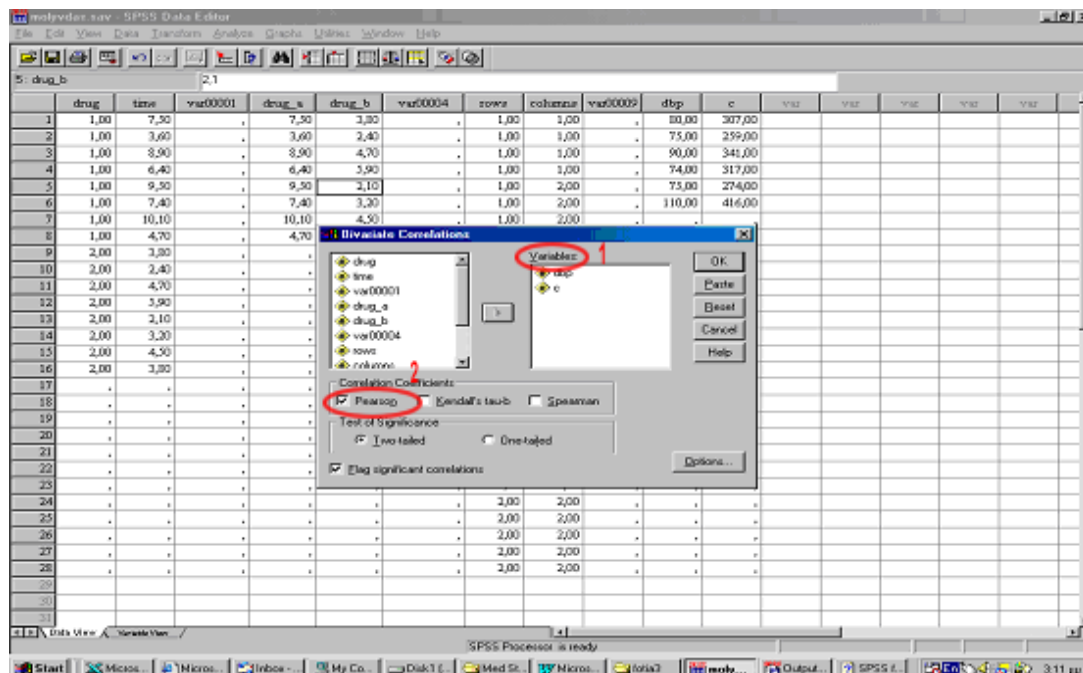
Έστω ότι σε 6 άτομα μετρήθηκε η πίεση στο αίμα (DBP) και τα επίπεδα χοληστερόλης (C) .

Θέλουμε να ελέγξουμε την υπόθεση εάν υπάρχει σχέση μεταξύ DBP και C και ποια προκαλεί τη μεταβολή ποιας και κατά πόσο?

Καταχωρούμε τα δεδομένα και από το μενού επιλογών διαλέγουμε Analyze->Correlate->Bivariate για να υπολογίσουμε το συντελεστή συσχέτισης (Pearson Correlation).



Στην συνέχεια από την Λίστα Μεταβλητών, επιλέγουμε τις μεταβλητές (DBP και C), για τις οποίες θέλουμε να υπολογίσουμε τον συντελεστή συσχέτισης, και κατόπιν την επιλογή “Pearson Correlation”.



Στα αποτελέσματα του output βρίσκουμε ότι  $r=0.931$  που είναι σημαντικό σε  $P=0.007$ .

**Chi-Square Tests**

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	1,448 <sup>a</sup>	1	,229		
Continuity Correction <sup>a</sup>	,673	1	,412		
Likelihood Ratio	1,467	1	,226		
Fisher's Exact Test				,276	,207
Linear-by-Linear Association	1,396	1	,237		
N of Valid Cases	28				

a. Computed only for a 2x2 table  
b. 0 cells (.0%) have expected count less than 5. The minimum expected count is 5,57.

**Correlations**

		DBP	C
DBP	Pearson Correlation	1,000	,931**
	Sig. (2-tailed)		,007
	N	6	6
C	Pearson Correlation	,931**	1,000
	Sig. (2-tailed)	,007	
	N	6	6

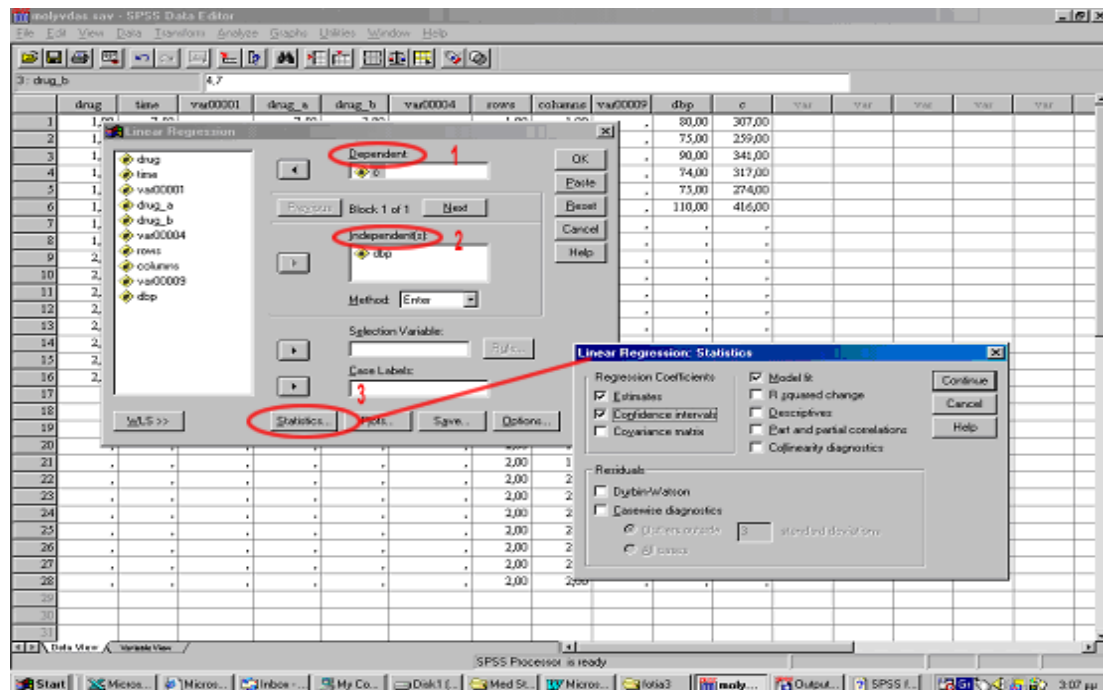
\*\* . Correlation is significant at the 0.01 level

Η σχέση μεταξύ DBP και C περιγράφεται με τη γραμμή γραμμικής παλινδρόμησης. Για να την βρούμε από το μενού επιλογών, διαλέγουμε “Regression”, “Linear”.

**Linear Regression**

SPSS Processor is ready

Στην συνέχεια από την Λίστα Μεταβλητών, διαλέγουμε την εξαρτώμενη μεταβλητή (C) και την ανεξάρτητη μεταβλητή (DBP). Η επιλογή “Statistics”, εμφανίζει το παράθυρο “Linear Regression Statistics”, όπου μπορούμε να καθορίσουμε διάφορες παραμέτρους, όπως διαστήματα εμπιστοσύνης κτλ.



Τα αποτελέσματα εμφανίζονται στο output. Το a είναι 7,861 και το b είναι 3,704

**Output1 - SPSS Viewer**

File Edit View Insert Format Analyze Graphs Utilities Window Help

Output

- T-Test
- Title
- Notes
- Paired Samples Statistics
- Paired Samples Correlations
- Paired Samples Test
- Crosstabs
- Title
- Notes
- Case Processing Summary
- ROWS \* COLUMNS Crosstab
- Chi-Square Tests
- Regression
  - Model Summary
  - ANOVA
  - Coefficients

**Model Summary**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,931 <sup>a</sup>	,867	,834	22,8215

a. Predictors: (Constant), DBP

**ANOVA<sup>a</sup>**

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	13582,716	1	13582,716	26,079	,007 <sup>a</sup>
	Residual	2083,284	4	520,821		
	Total	15666,000	5			

a. Predictors: (Constant), DBP  
b. Dependent Variable: C

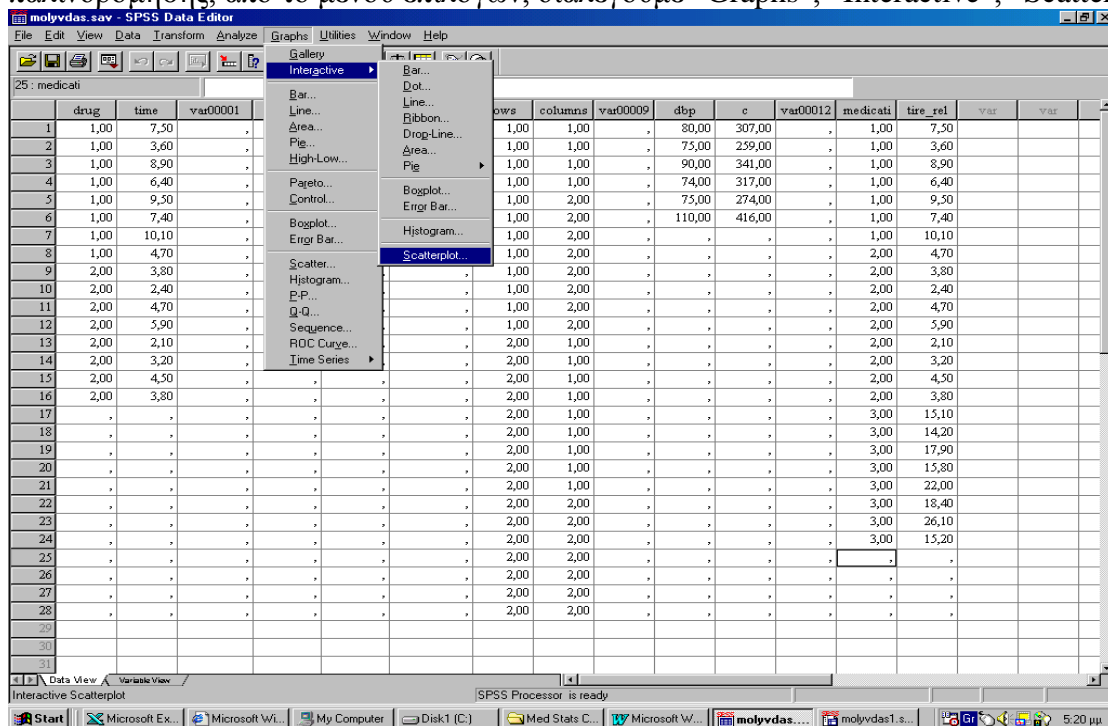
**Coefficients<sup>a</sup>**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95% Confidence Interval for B	
		B	Std. Error				Lower Bound	Upper Bound
1	(Constant)	7,861	61,635		,128	,905	-163,262	178,983
	DBP	3,704	,725	,931	5,107	,007	1,690	5,718

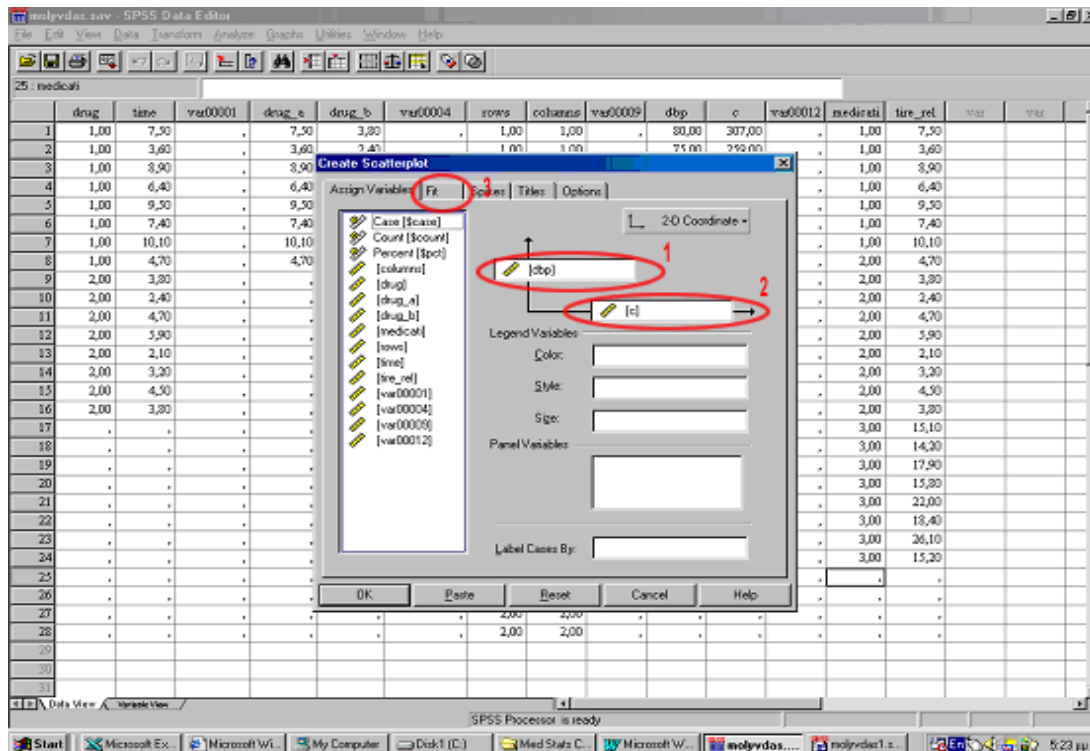
a. Dependent Variable: C

SPSS Processor is ready

**Reactive**, **Popular**, **GIS**, **Web**, **to**, **Access**, **Online**, **CSV**, **Storage**, **Cloud**, **Graphs**, **Interactive**, **Scatterplot**.

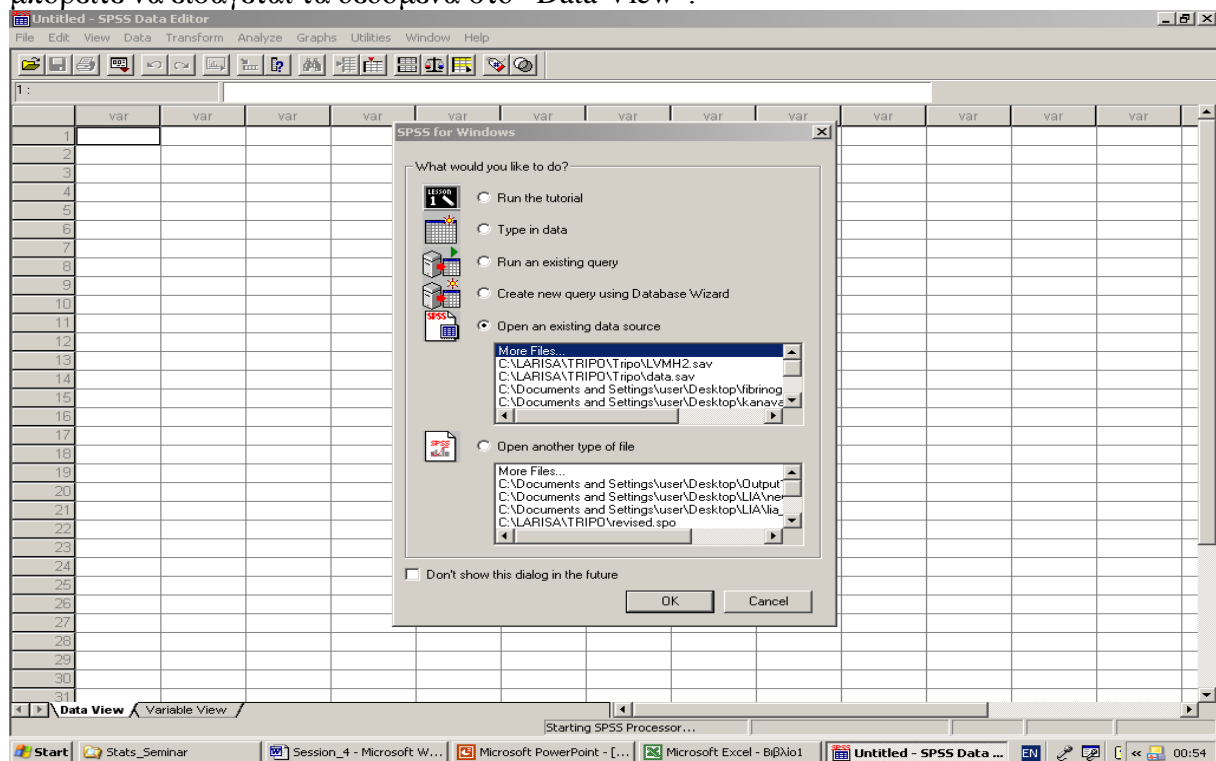


Από την Λίστα Μεταβλητών επιλέγουμε την εξαρτημένη μεταβλητή (DBP) και την ανεξάρτητη μεταβλητή όπως παρακάτω επιλέξτε στο tab “Fit” Regression.

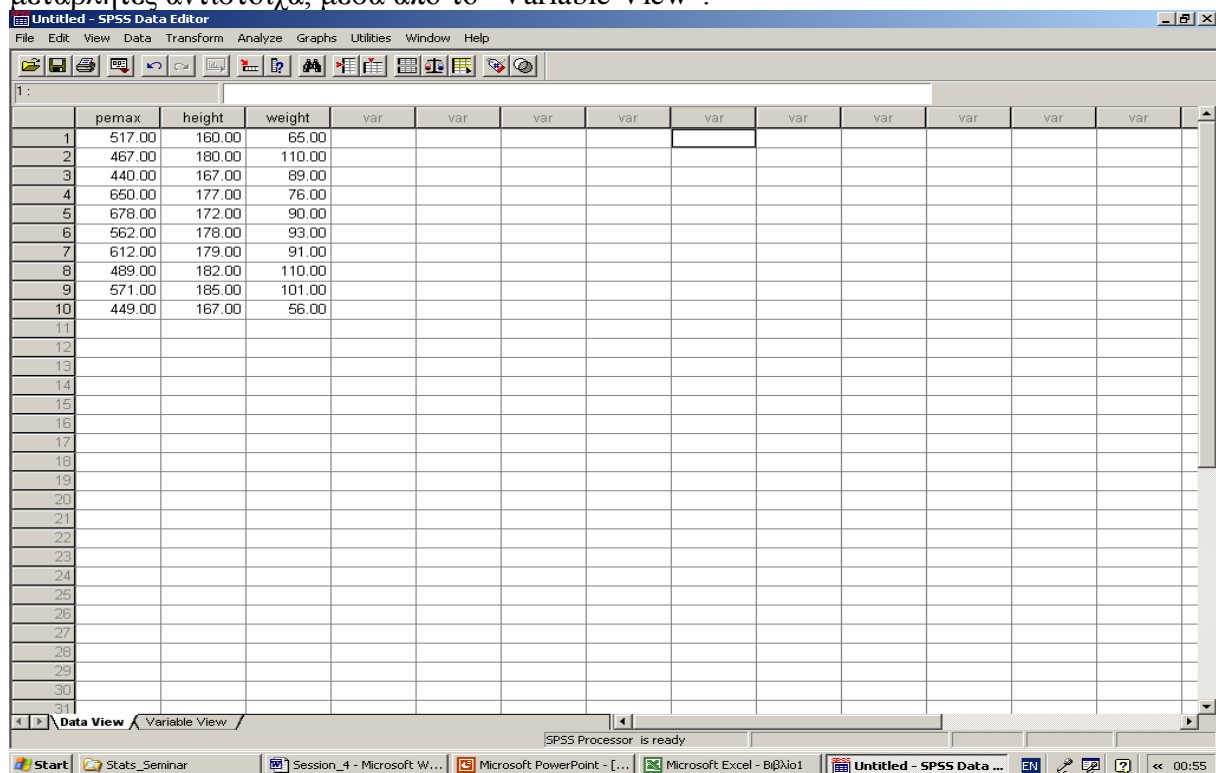


## Πολλαπλή γραμμική παλινδρόμηση (multiple linear regression)

Ανοίξτε το “SPSS” από το “Start” στο Desktop του H/Y και επιλέξτε “Type in data”. Τώρα μπορείτε να εισάγεται τα δεδομένα στο “Data View”.

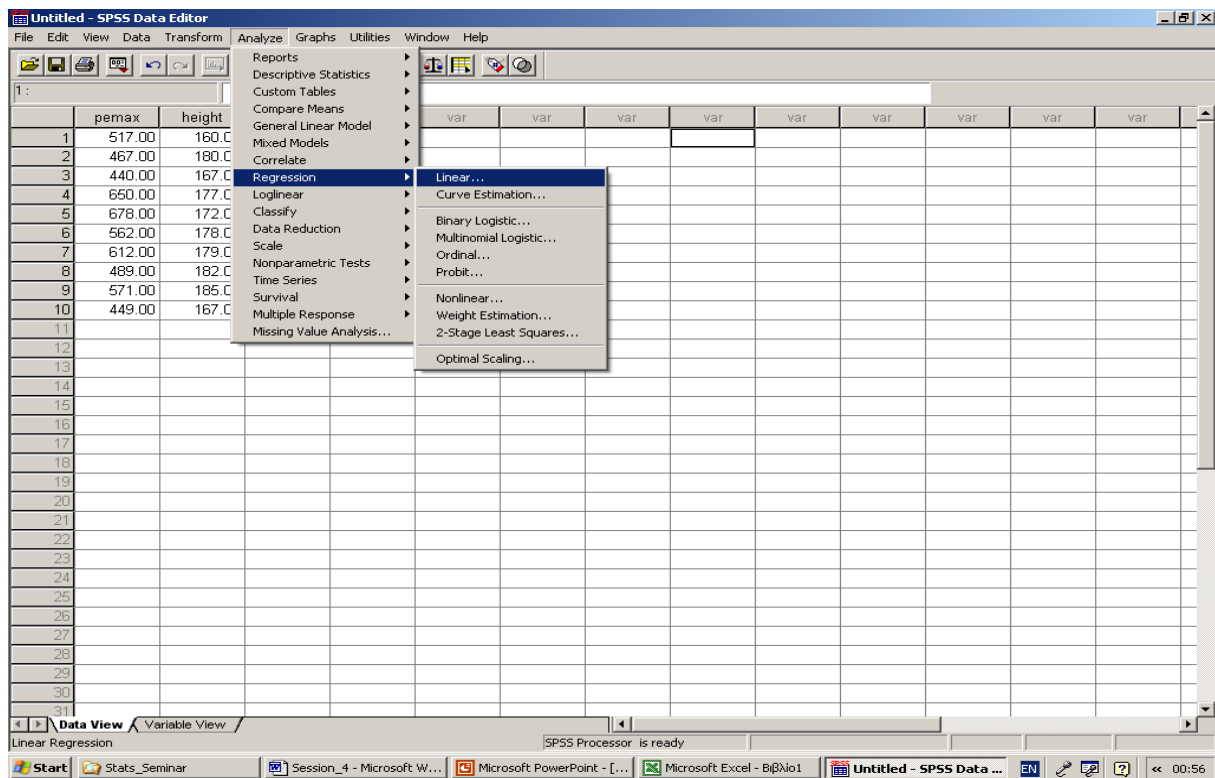


Εισάγετε τα δεδομένα της άσκησης στο “Data View” όπως φαίνεται παρακάτω. Μία μεταβλητή αντιπροσωπεύει τη PEmax, μία το Height και μία Weight, μετά ονομάστε τις μεταβλητές αντίστοιχα, μέσα από το “Variable View”.

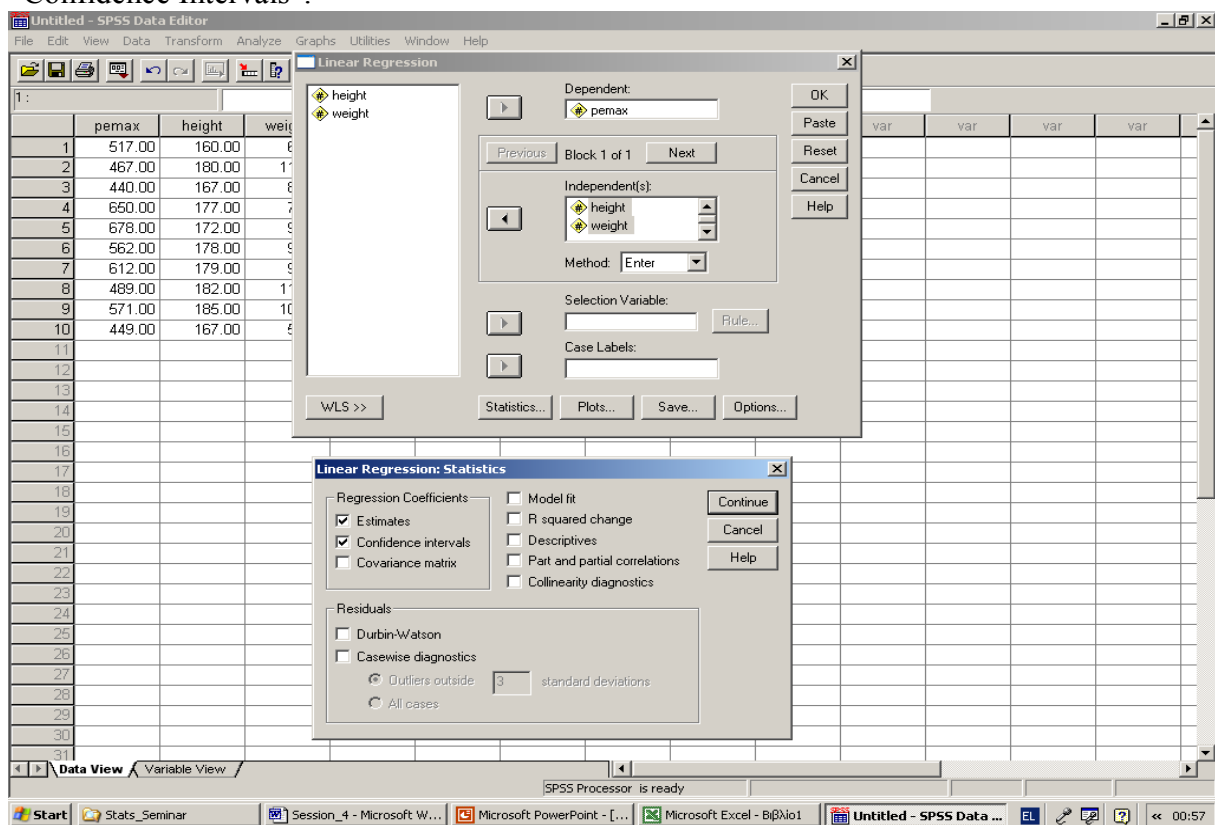




Για να αναλύσετε τα δεδομένα επιλέξτε “Regression”, “Linear”.



Δηλώστε ως “Dependent” τη μεταβλητή “pemax” και σαν “Independent(s)” τις μεταβλητές “height” και “weight”. Διαλέξτε την επιλογή “Statistics” και μετά επιλέξτε “Estimates” και “Confidence Intervals”.



Τα αποτελέσματα της ανάλυσης φαίνονται από το παράθυρο “Output1”: Τα P-value για τις σημαντικότητες των μεταλητών “height” και “weight” είναι 0.285 και 0.421, αντίστοιχα, δηλ. μη σημαντικά στην ανταπόκριση της PEmax. Τα αποτελέσματα μπορεί να αποθηκεύσετε για μελλοντική χρήση από την επιλογή “File”, “Save as” όπως προηγουμένως.

**Regression**

**Variables Entered/Removed<sup>a</sup>**

Model	Variables Entered	Variables Removed	Method
1	WEIGHT <sup>a</sup> HEIGHT	.	Enter

a. All requested variables entered.  
b. Dependent Variable: PEMAX

**Coefficients<sup>a</sup>**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	-440.547	855.708		-.515	.623	-2463.974	1582.881
	HEIGHT	6.747	5.833	.631	1.157	.285	-7.046	20.540
	WEIGHT	-2.209	2.585	-.467	-.855	.421	-8.321	3.903

a. Dependent Variable: PEMAX