ΑΝΑΛΥΣΗ ΔΕΔΟΜΕΝΩΝ

Δρ. Βασίλης Π. Αγγελίδης
Τμήμα Μηχανικών Παραγωγής & Διοίκησης
Δημοκρίτειο Πανεπιστήμιο Θράκης

Περιεχόμενα

- **Έλεγχος κανονικότητας**
  - P-P Plot και Q-Q Plot
  - Τεστ Κανονικότητας
  - Τεστ Κανονικότητας – Διαστήματα Εμπιστοσύνης
  - Τεστ ανεξαρτησίας
- Βασικές έννοιες θεωρίας ελέγχων υποθέσεων
  - Στατιστική Επαγωγή
  - Διαστήματα Εμπιστοσύνης
- **Έλεγχοι**
  - **T-Test**
    - Έλεγχος της μέσης τιμής – Παραμετρικό/Μη Παραμετρικό Τεστ
    - Σύγκριση των μέσων τιμών μίας μεταβλητής - Παραμετρικό /Μη Παραμετρικό Τεστ
    - Σύγκριση των μέσων τιμών δύο εξαρτημένων δειγμάτων. Παραμετρικό /Μη Παραμετρικό Τεστ
    - Συντελεστές γραμμικής συσχέτισης
  - **Έλεγχος F**
    - Ανάλυση διακύμανσης (διασποράς) κατά ένα παράγοντα (One-way Anova)
  - **Έλεγχος χ^2**
    - Σύγκριση δύο ποιοτικών μεταβλητών (ανεξάρτητα / εξαρτημένα δείγματα)
- Διαστήματα Εμπιστοσύνης

Έλεγχος κανονικότητας και διαστήματα εμπιστοσύνης

- Αναφέρθηκε στο προηγούμενο μάθημα ότι όταν το ιστόγραμμα συχνοτήτων των ποσοτικών μεταβλητών έχει το σχήμα "**καμπάνας**", τότε λέμε ότι τα δεδομένα ακολουθούν την κανονική κατανομή ή κατανέμονται κανονικά.
- Το ιστόγραμμα όμως δεν είναι "**ικανό**" να μας απαντήσει στη ερώτηση αν είναι κανονικά τα δεδομένα ή αν προέρχονται από μία κανονική κατανομή με ένα μέσο και μία διακύμανση.
- Ένα από τα πιο σημαντικά προβλήματα στην στατιστική είναι η διερεύνηση της πληροφορίας η οποία σχετίζεται με την μορφή της **κατανομής** από την οποία προέρχεται ένα τυχαίο δείγμα.
- Οι περισσότεροι έλεγχοι γίνονται με την προϋπόθεση ότι (**υπό την H_0**) το τυχαίο δείγμα προέρχεται από μια συγκεκριμένη κατανομή. Μια τέτοια περίπτωση είναι τα **t-tests** που θα εξετάσουμε στη συνέχεια και τα οποία, ιδιαίτερα για μικρά δείγματα, προϋποθέτουν ότι το δείγμα προέρχεται από κανονικής κατανομής πληθυσμό.
- Εάν το τυχαίο δείγμα δεν προέρχεται από την κατανομή κάτω από την οποία έχει κατασκευασθεί κάποιος έλεγχος τότε προφανώς το αντίστοιχο **p-value** που λαμβάνεται δεν είναι ακριβές.

Έλεγχος κανονικότητας και διαστήματα εμπιστοσύνης

- Συνεπώς είναι αρκετά χρήσιμη η δυνατότητα να ελέγχουμε αν κάποια δεδομένα προέρχονται από μια συγκεκριμένη κατανομή ή όχι.
- Έλεγχοι αυτής της μορφής καλούνται «**έλεγχοι καλής προσαρμογής**» των δεδομένων σε μια συγκεκριμένη κατανομή και έχουν προταθεί αρκετοί.
- Για τον έλεγχο αν τα δεδομένα ακολουθούν την κανονική κατανομή αρχικά μπορούμε να κατασκευάσουμε δύο γραφήματα με το **SPSS**, το **P-P Plot** και το **Q-Q Plot** (Επιλέγοντας **Analyze→Descriptive Statistics→P-P Plots** ή **Q-Q Plots**).
- Με αυτά τα γραφήματα ελέγχουμε οπτικά την ύπαρξη κανονικότητας στα δεδομένα. Όσο πιο κοντά στην ευθεία είναι τα σημεία του σχήματος τόσο πιο πολλές είναι οι ενδείξεις ότι τα δεδομένα ακολουθούν την κανονική κατανομή.
- Το μάτι όμως πάλι μπορεί να “**πέσει έξω**” και να **ξεγελαστούμε**. Για αυτό το λόγο καταφεύγουμε σε τεστ κανονικότητας για να απαντήσουμε στην προηγούμενη ερώτηση.

Έλεγχος κανονικότητας και διαστήματα εμπιστοσύνης

Ο έλεγχος κανονικότητας υπάγεται σε μία ευρύτερη οικογένεια ελέγχων, τη λεγόμενη «**έλεγχοι υποθέσεων**». Όταν ακούμε για ελέγχους υποθέσεων μας έρχονται πολλά πράγματα στο μυαλό. Κάποια από αυτά είναι η μηδενική υπόθεση (**Null Hypothesis** ή H_0), η εναλλακτική υπόθεση (**Alternative Hypothesis** ή H_1), το **επίπεδο στατιστικής σημαντικότητας** (α) και το παρατηρηθέν επίπεδο στατιστικής σημαντικότητας (**p-value** ή **Significance**). Οι υποθέσεις αυτές είναι της ακόλουθης μορφής:

H_0: Η κατανομή των δεδομένων δε διαφέρει από την κανονική κατανομή

H_1: Η κατανομή των δεδομένων διαφέρει από την κανονική κατανομή

Για τη διεξαγωγή των ελέγχων υποθέσεων χρησιμοποιούνται κάποιοι μαθηματικοί τύποι, που καλούνται ελεγχοσυναρτήσεις. Με βάση το αποτέλεσμα τους οδηγούμαστε στο συμπέρασμα ότι η μηδενική υπόθεση απορρίπτεται ή όχι.

Έλεγχος κανονικότητας και διαστήματα εμπιστοσύνης

- Η μηδενική υπόθεση στην συγκεκριμένη περίπτωση την οποία θέλουμε να ελέγξουμε είναι ότι τα δεδομένα ακολουθούν την κανονική ή ότι προέρχονται από ένα πληθυσμό που ακολουθεί την κανονική κατανομή.
- Η εναλλακτική είναι ότι τα δεδομένα δεν ακολουθούν την κανονική κατανομή.
- Το επίπεδο στατιστικής σημαντικότητας ορίζεται συνήθως ίσο με 0,05 ή 5%.
- Το παρατηρηθέν επίπεδο στατιστικής σημαντικότητας ορίζεται ως η πιθανότητα η τιμή του ελέγχου (ελεγχοσυνάρτησης) να πάρει μία τιμή τόσο ακραία ή περισσότερο ακραία από αυτή που πήρε στο συγκεκριμένο δείγμα κάτω από τη μηδενική υπόθεση.
- Αν η **p-value** είναι μικρότερη του 0,05, τότε λέμε ότι η μηδενική υπόθεση απορρίπτεται.
- Αν η **p-value** είναι μεγαλύτερη ή ίση του 0,05, τότε λέμε ότι η μηδενική υπόθεση δεν απορρίπτεται.

Έλεγχος κανονικότητας και διαστήματα εμπιστοσύνης

Για τους ελέγχους καλής προσαρμογής αρχικά θα εξετάσουμε:

- κάποιους «εμπειρικούς» ελέγχους οι οποίοι γίνονται μέσω κάποιων γραφημάτων (**P-P** και **Q-Q plots**) έτσι ώστε να πάρουμε μια πρώτη εποπτική εικόνα για τα δεδομένα (τα γραφήματα αυτά δεν οδηγούν με σχετική «**ασφάλεια**» σε κάποια απόφαση)

ενώ στη συνέχεια θα περάσουμε στους πιο σημαντικούς ελέγχους καλής προσαρμογής:

- το χι-τετράγωνο τεστ καλής προσαρμογής
- και το Kolmogorv- Smirnov τεστ.

Τέλος, ένα ενδιαφέρον παρεμφερές πρόβλημα αφορά δύο δείγματα και τον έλεγχο της υπόθεσης ότι τα δείγματα αυτά προέρχονται από τον ίδιο πληθυσμό (δηλαδή από την ίδια κατανομή). Για τον έλεγχο αυτό θα παρουσιαστούν τρία μη παραμετρικά τεστ:

- το Kolmogorov- Smirnov για δυο δείγματα,
- το το Wald-Wolfowitz τέστ των ροών
- και το Mann-Whitney U τέστ

P-P Plot και Q-Q Plot

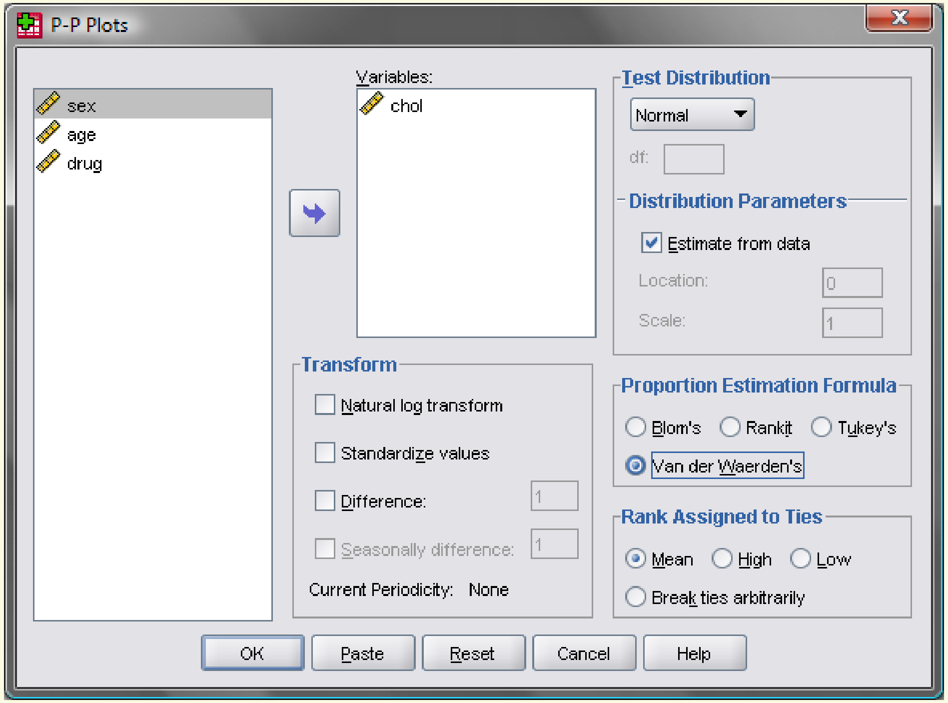
- Τα **P-P Plot** και **Q-Q plot** (probability-probability plot και Quantile-Quantile plot) είναι δύο γραφήματα τα οποία μας βοηθούν να ελέγξουμε αν κάποια δεδομένα προέρχονται από κάποια συγκεκριμένη κατανομή (π.χ. κανονική).
- Και στα δύο γραφήματα, αν τα σημεία που εμφανίζονται βρίσκονται «**κοντά**» στη διαγώνιο (και «**τυχαία**» γύρω από αυτήν) τότε μπορεί να θεωρηθεί ότι τα δεδομένα προέρχονται από μία κατανομή (π.χ. κανονική)
- ο έλεγχος μέσω των παραπάνω γραφημάτων δεν μπορεί να είναι **αξιόπιστος** διότι δεν βασίζεται σε κάποιο **στατιστικό κριτήριο** που μας οδηγεί σε σωστή απόφαση π.χ. στο $1-a$ % των περιπτώσεων.
- Συνήθως γίνεται για να πάρουμε μια **πρώτη εποπτική εικόνα** και για να δούμε αν υπάρχουν κάποιες έκτροπες, σε σχέση με τις αναμενόμενες υπό την F_0, παρατηρήσεις.
- Είναι προφανές ότι, εκτός της κανονικής, μπορούμε γραφικά να ελέγξουμε την καλή προσαρμογή των δεδομένων και σε άλλες κατανομές (αλλάζουμε την Test distribution)..

Παράδειγμα P-P Plot

Να ελεγχθεί γραφικά (μέσω **P-P plot** ή **Q-Q plot**) αν οι παρατηρήσεις του δείκτη χοληστερίνης (μεταβλητή chol) προέρχονται από την κανονική κατανομή (αρχείο test1)

Στο **SPSS** από το μενού **analyze/descriptive statistics** επιλέγουμε **P-P plot**

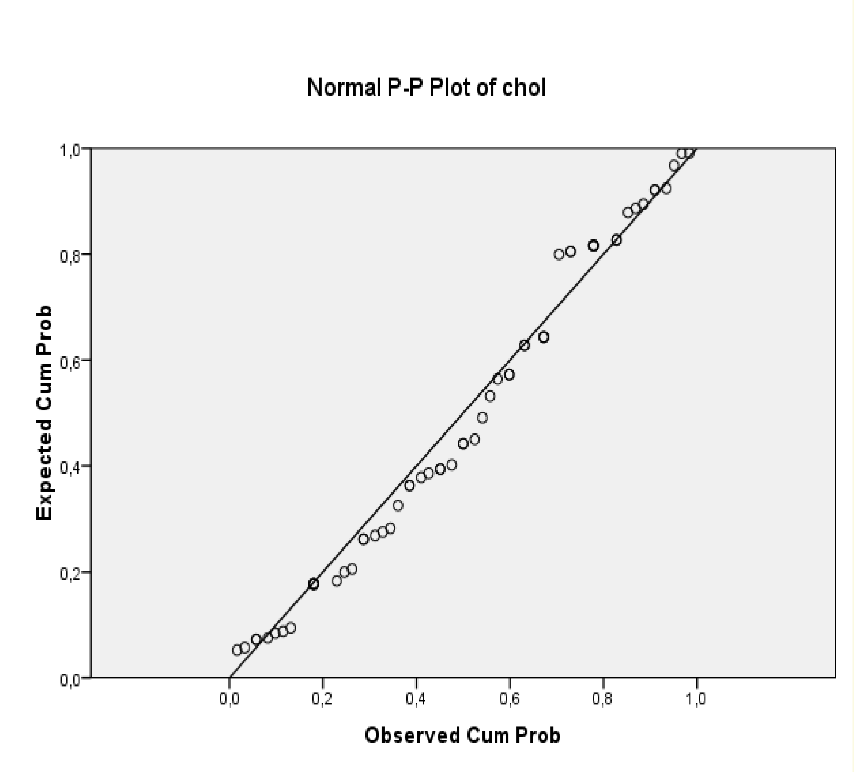
Για μέτρια ή μεγάλα δείγματα δεν υπάρχει ουσιαστική διαφορά οποιαδήποτε proportion estimation formula και αν επιλέξουμε για μικρά όμως δείγματα καλύτερα να επιλέξουμε την Van der Waerden's κάποιοι άλλοι ερευνητές παραθέτοντας κάποια δικαιολόγηση έχουν προτείνει την (Blom's) ή κάποια άλλη



Αποτελέσματα P-P Plot

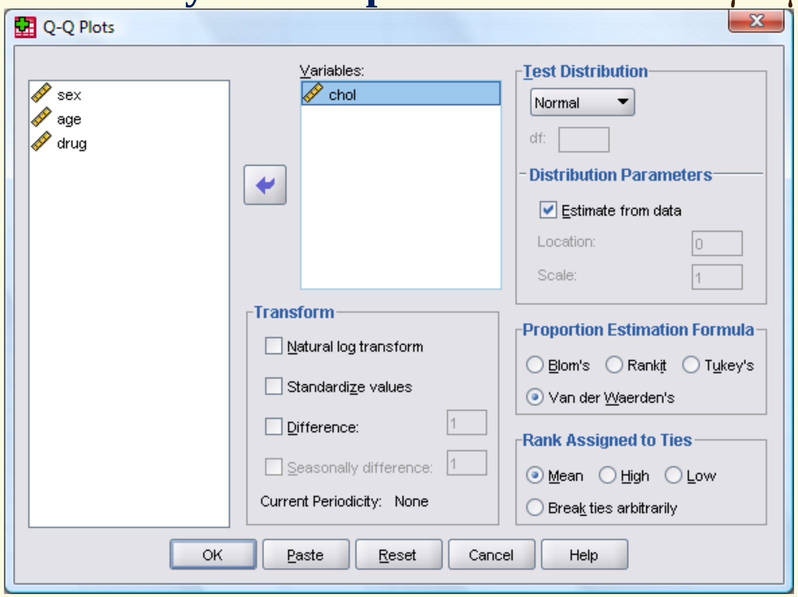
Φαίνεται ότι τα 60 σημεία του επιπέδου δεν «απέχουν» πολύ από την διαγώνιο, ούτε φαίνονται κάποιες «έκτροπες» παρατηρήσεις. Επομένως, τουλάχιστον γραφικά, δεν φαίνεται να υπάρχει επαρκής λόγος ώστε να μην θεωρήσουμε τα δεδομένα ως κανονικά (για να είμαστε πιο ακριβείς θα πρέπει να προχωρήσουμε και σε έλεγχο με δεδομένο ε.σ. a, π.χ. χ^2 ή **K-S** που θα εξετάσουμε παρακάτω).

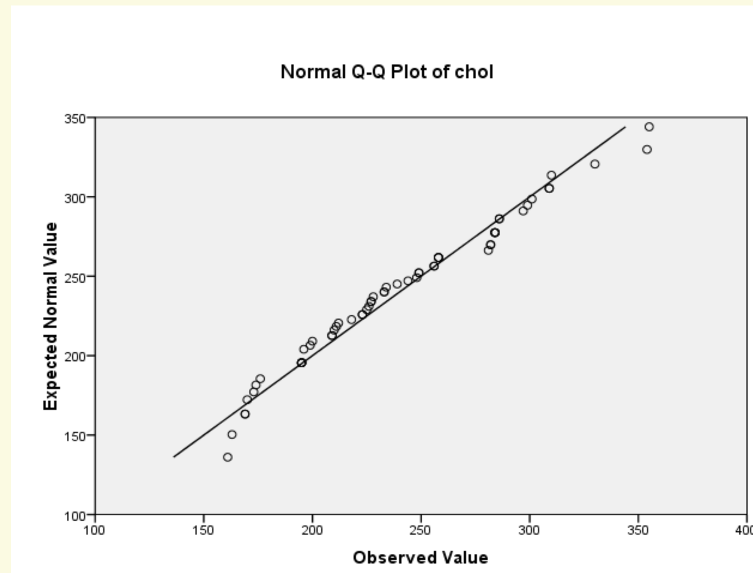
Αξίζει να παρατηρήσουμε ότι αν υπήρχαν κάποιες έκτροπες παρατηρήσεις (παρατηρήσεις αρκετά «απομακρυσμένες» από την διαγώνιο) τότε θα έπρεπε αυτές να επανεξεταστούν λεπτομερέστερα ώστε να βεβαιωθούμε ότι δεν ισχύουν κάποιες ειδικές συνθήκες για αυτές ή ότι δεν έχουν περαστεί λάθος στο **SPSS**.



Παράδειγμα P-P Plot

Να ελεγχθεί γραφικά (μέσω **Q-Q plot**) αν οι παρατηρήσεις του δείκτη χοληστερίνης (μεταβλητή chol) προέρχονται από την κανονική κατανομή (αρχείο test1). Στο **SPSS** από το μενού **analyze/descriptive statistics** επιλέγουμε **Q-Q**



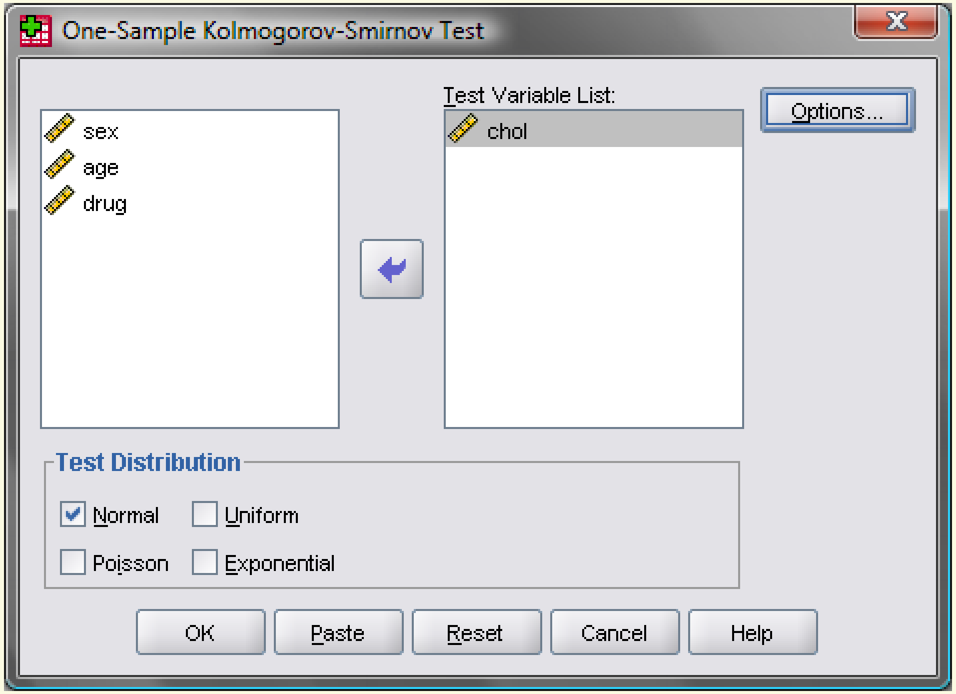


Αυτό το εμπειρικό τεστ είναι ισοδύναμο με το προηγούμενο και επομένως τα αποτελέσματα είναι τα ίδια. Και εδώ φαίνεται ότι όλα τα σημεία του επιπέδου δεν «**απέχουν**» πολύ από την διαγώνιο και επομένως δεν υπάρχει επαρκής λόγος ώστε απορρίψουμε ότι τα δεδομένα είναι κανονικά.

Τεστ Κανονικότητας

Για να ελέγξουμε αν η **κατανομή** μιας μεταβλητής είναι συμβατή με την **κανονική** εφαρμόζουμε το test **Kolmogorov-Smirnov**.

- **H_0(Μηδενική υπόθεση):** Η υπό έλεγχο κατανομή, δε διαφέρει από την κανονική κατανομή έναντι της
- **H_1(Εναλλακτικής υπόθεσης):** Η υπό έλεγχο κατανομή διαφέρει από την κανονική κατανομή.



Για την εκτέλεση του τεστ κανονικότητας με το SPSS :

- πατάμε **Analyze → Non parametric tests → One sample K-S**
- Βάζουμε στο **test variable list** τις μεταβλητές που θέλουμε να ελέγξουμε την κανονικότητα τους,
- Τσεκάρουμε **Normal** και **OK**

Τεστ Κανονικότητας - Αποτελέσματα

One-Sample Kolmogorov-Smirnov Test

		chol
N		60
Normal Parameters[a]	Mean	240,0833
	Std. Deviation	48,72729
Most Extreme Differences	Absolute	,099
	Positive	,083
	Negative	-,099
Kolmogorov-Smirnov Z		,770
Asymp. Sig. (2-tailed)		,593

a. Test distribution is Normal.

Η υπόθεση την οποία θέλουμε να ελέγξουμε είναι ότι οι μεταβλητές ακολουθούν την κανονική κατανομή. Από ότι φαίνεται για δεν έχουμε εκλιπούσες τιμές. Εμφανίζονται επίσης ο μέσος και η τυπική απόκλιση για κάθε μεταβλητή. Για τον έλεγχο της κανονικότητας μας ενδιαφέρουν η τιμή της Asymp. Sig. (2-tailed).

Το τεστ των Kolmogorov-Smirnov είναι ένα απλά μία τιμή p-value που υπολογίζεται με μία ελαχιστοσυνάρτηση. Παρατηρούμε ότι η τιμή p-value είναι 0,593 και όπως αναφέραμε προηγουμένως αν η p-value είναι μεγαλύτερη από το 0,05, τότε δεν απορρίπτουμε την υπόθεση της κανονικότητας των δεδομένων. Άρα η υπόθεση ότι οι μετρήσεις που αφορούν την μέτρηση της χοληστερίνης κατανέμονται κανονικά δεν απορρίπτεται σε επίπεδο στατιστικής σημαντικότητας α=0,05 ή α=5%. Ειδάλλως, μπορούμε να πούμε ότι υπάρχουν ενδείξεις ότι αυτές οι μετρήσεις ακολουθούν την κανονική κατανομή.

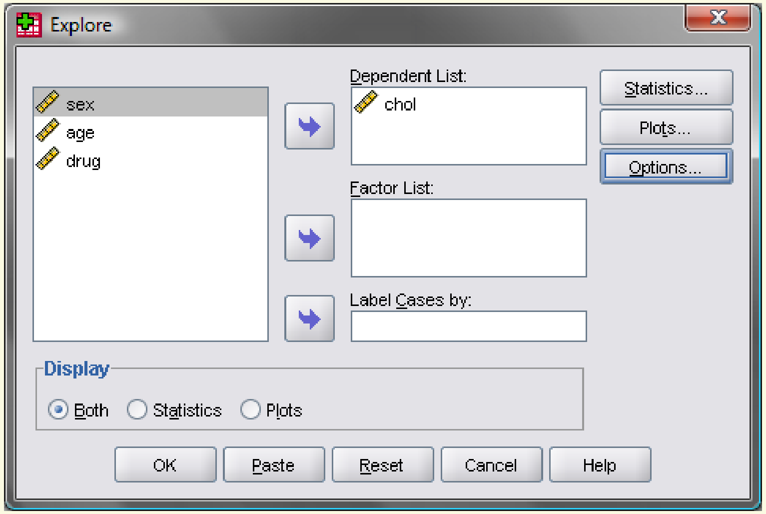
Τεστ Κανονικότητας – Διαστήματα Εμπιστοσύνης

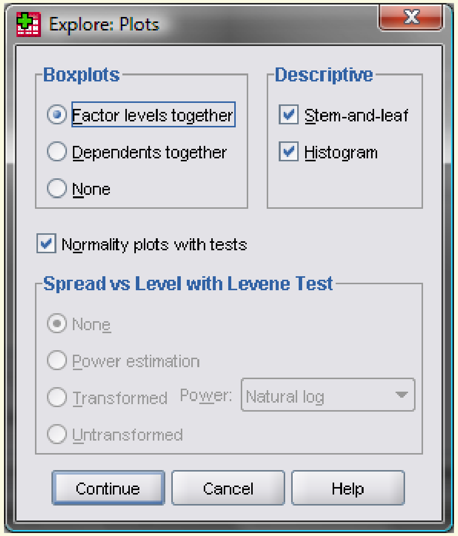
Στην προσπάθεια να εκτιμήσουμε την πραγματική τιμή του μέσου, χρησιμοποιούμε το μέσο ενός δείγματος. Στη συνέχεια κατασκευάζουμε ένα 95% διάστημα εμπιστοσύνης με βάση ένα μαθηματικό τύπο, ο οποίος είναι στην ουσία 2 τυπικά σφάλματα αριστερά και δεξιά της τιμής του μέσου που βρήκαμε για το δείγμα. Αν επαναλάβουμε τη δειγματοληψία ν φορές θα εκτιμήσουμε ν διαφορετικούς μέσους και προφανώς ν διαφορετικά (πολλά θα είναι αλληλοεπικαλυπτόμενα) διαστήματα εμπιστοσύνης.

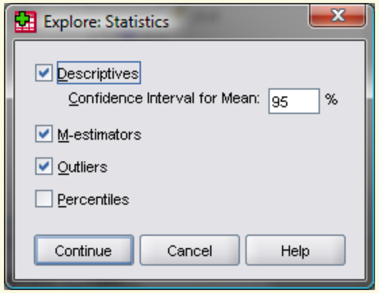
Αυτό που θέλουμε είναι ότι στο 95% των ν περιπτώσεων τα διαστήματα εμπιστοσύνης που υπολογίσαμε θα έχουν περικλείσει, ή “πιάσει”, ή “χτυπήσει” την τιμή του πραγματικού μέσου. Επομένως, αν για παράδειγμα “παίρναμε” κάθε φορά 100 δείγματα από έναν πληθυσμό και κατασκευάζαμε 100 διαστήματα εμπιστοσύνης για τη μέση τιμή μίας μεταβλητής και επαναλαμβάναμε τη διαδικασία άπειρες τιμές, κατά μέσο όρο στο 95% των περιπτώσεων θα είχαμε φτιάξει διαστήματα εμπιστοσύνης που θα είχαν “πιάσει” τον πραγματικό μέσο του πληθυσμού. Το 95% θα το ονομάζουμε **βαθμό ή επίπεδο εμπιστοσύνης**. Το υπόλοιπο 5% είναι αυτό που έχουμε ήδη ορίσει επίπεδο στατιστικής σημαντικότητας.

Τεστ Κανονικότητας – Διαστήματα Εμπιστοσύνης

Για να κατασκευάσουμε ένα 95% διάστημα εμπιστοσύνης για τη μέση τιμή μίας μεταβλητής εργαζόμαστε ως εξής: πατάμε **Analyze→Descriptive Statistics→Explore**







Από την επιλογή Plots μπορούμε να επιλέξουμε αν θέλουμε να εμφανιστεί ένα ιστόγραμμα των ή της μεταβλητής που θα περάσουμε στο άνω λευκό κουτάκι (Dependent List:) ακόμη και ένα τεστ κανονικότητας. Από την επιλογή Statistics μπορούμε να επιλέξουμε διάφορα περιγραφικά στατιστικά καθώς και το διάστημα εμπιστοσύνης

Τεστ Κανονικότητας – Διαστήματα Εμπιστοσύνης

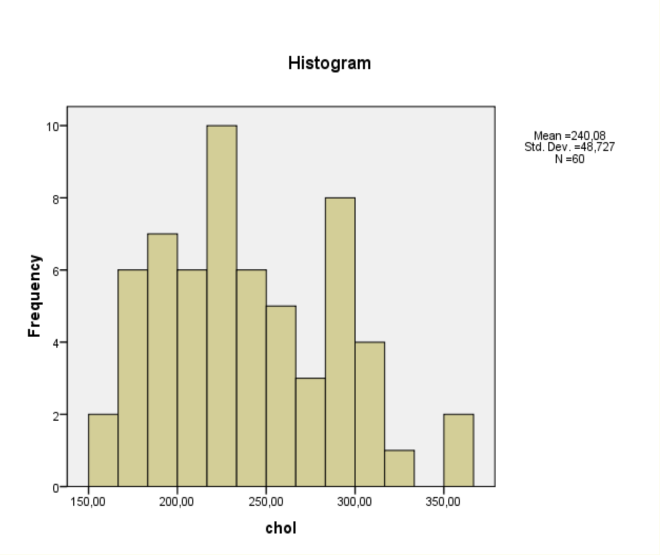
Descriptives

		Statistic	Std. Error
Mean		240,0833	6,29067
95% Confidence Interval for Mean	Lower Bound	227,4957	
	Upper Bound	252,6709	
5% Trimmed Mean		238,3889	
Median		233,0000	
Variance		2374,349	
Std. Deviation		48,72729	
Minimum		161,00	
Maximum		355,00	
Range		194,00	
Interquartile Range		84,25	
Skewness		,363	,309
Kurtosis		-,539	,608

Ο πίνακας περιέχει τα περιγραφικά μέτρα στα οποία έχουμε ήδη αναφερθεί μαζί με λίγα ακόμα για τα οποία δεν έχουμε μιλήσει. Η μεταβλητή επιλέχθηκε τυχαία. Η πρώτη γραμμή του πίνακα περιέχει τη μέση τιμή για τις τιμές αυτής της μεταβλητής. Οι επόμενες δύο τιμές είναι το κάτω και το άνω άκρο του 95% διαστήματος εμπιστοσύνης για τον πραγματικό μέσο του πληθυσμού. Η επόμενη γραμμή είναι ο μέσος των τιμών της μεταβλητής από την οποία έχουμε αφαιρέσει το 5% των μεγαλύτερων και μικρότερων τιμών.

Το Interquartile range (ενδοτεταρτημοριακό εύρος) είναι η διαφορά μεταξύ τρίτου και πρώτου τεταρτημόριου. Σε αυτό το εύρος βρίσκεται το 50% των κεντρικών παρατηρήσεων της μεταβλητής.

Τεστ Κανονικότητας – Διαστήματα Εμπιστοσύνης

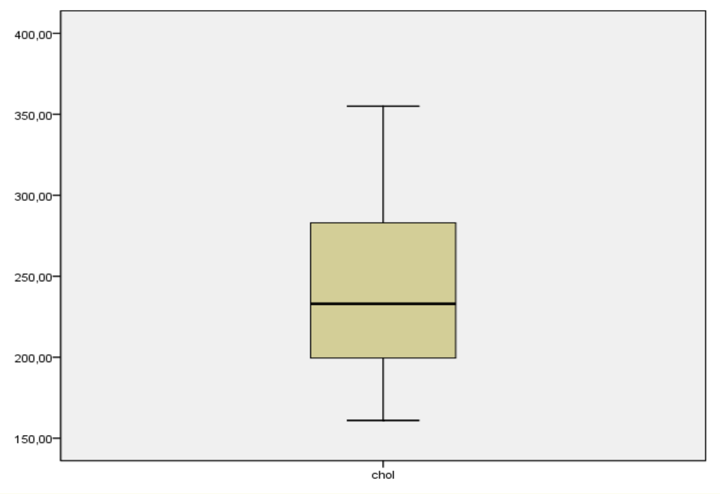


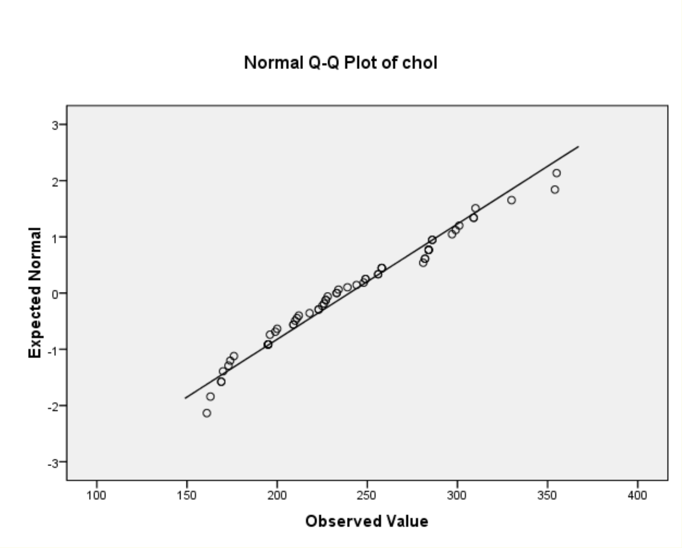
Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
chol	,099	60	,200*	,966	60	,091

a. Lilliefors Significance Correction

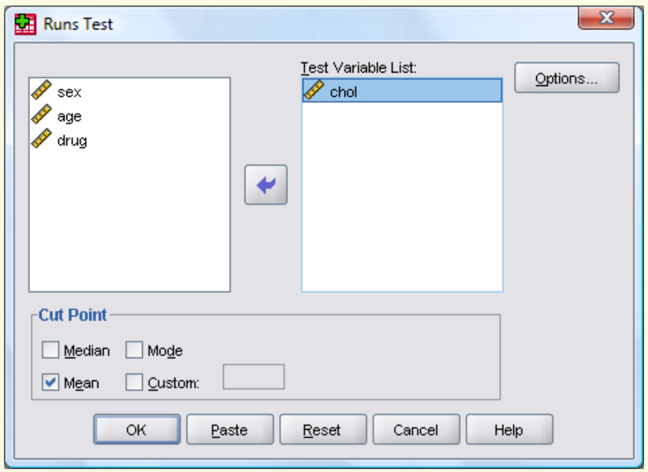
*. This is a lower bound of the true significance.





Έλεγχος ανεξαρτησίας

Ένας άλλος χρήσιμος έλεγχος είναι ο έλεγχος που βασίζεται στα λεγόμενα **runs**, δηλαδή τις μεταβολές προσήμου της σειράς. Ο έλεγχος είναι διαθέσιμος στο μενού **Analyze** → **Non parametric tests** → **Runs**. Για να ελέγξουμε τις τιμές της χολυστερίνης για τυχαιότητα χρησιμοποιούμε τις επιλογές αυτές και έχουμε την ακόλουθη οθόνη.



Runs Test

	chol
Test Value[a]	240,0833
Cases < Test Value	33
Cases >= Test Value	27
Total Cases	60
Number of Runs	4
Z	-7,025
Asymp. Sig. (2-tailed)	**,000**

a. Mean

Ο έλεγχος αυτός είναι έλεγχος Z . Η τιμή της στατιστικής είναι –7,025 και έχει πιθανότητα 0,000. Αυτό σημαίνει ότι σε επίπεδο σημαντικότητας 5% (αλλά και 1%) μπορούμε να απορρίψουμε την υπόθεση ότι η σειρά είναι τυχαία. Και αυτό μπορούμε να το δούμε και στα δεδομένα τα οποία είναι ταξινομξμένα κατά αύξουσα σειρά.

Ξέρω στατιστική σημαίνει ότι γνωρίζω ……..

- ποιο στατιστικό είναι κατάλληλο για κάθε ερευνητική ερώτηση,
- πώς να υπολογίσω το στατιστικό αυτό και
- πώς να το ερμηνεύσω

Η επιλογή του κατάλληλου στατιστικού είναι ένα από τα σημαντικότερα βήματα στην διαδικασία της εκπαιδευτικής έρευνας και της στατιστικής ανάλυσης

Για την επιλογή του κατάλληλου στατιστικού χρησιμοποιούμε δύο γενικά κριτήρια:

- Το λόγο (σκοπό) για τον οποίο χρειαζόμαστε το στατιστικό
  - η περιγραφή μεταβλητών ή σχέσεων μεταξύ μεταβλητών, με τα οποία ασχολείται η Περιγραφική Στατιστική και
  - η γενίκευση από το δείγμα στον πληθυσμό, με την οποία ασχολείται η επαγωγική
- Την κλίμακα (επίπεδο) μέτρησης των μεταβλητών
  - Ονομαστική, Τακτική, Ισοδιαστημική, Αναλογική

Στατιστική Επαγωγή

- Ο απώτερος σκοπός της στατιστικής ανάλυσης είναι να μας βοηθήσει να περιγράψομε τον πληθυσμό.
- Όταν όμως τα μόνα διαθέσιμα στοιχεία είναι από ένα δείγμα, ο πληθυσμός δεν μπορεί να περιγραφεί άμεσα.
- Αυτό που μπορούμε να κάνομε σε τέτοιες περιπτώσεις είναι να υπολογίσομε τα στατιστικά του δείγματος (περιγραφικά ή εξηγητικά) και από αυτά δι' επαγωγής να βγάλομε συμπεράσματα για τον πληθυσμό ή να προσεγγίσομε τις παραμέτρους του.
- Τα στατιστικά και οι τεχνικές που χρησιμοποιούνται για να κάνομε γενικεύσεις από το δείγμα στον πληθυσμό λέγονται Επαγωγικά Στατιστικά.
- Οι επαγωγές που ενδιαφέρουν τους ερευνητές στις κοινωνικές επιστήμες και την εκπαίδευση είναι δύο ειδών.
  - Τα **Διαστήματα Εμπιστοσύνης** (Confidence Intervals) και
  - Ο **Έλεγχος Υποθέσεων** (Hypothesis Testing)

Τα Διαστήματα Εμπιστοσύνης

Έχουμε τα παρακάτω ερωτήματα:

- Ποιο είναι το εκπαιδευτικό επίπεδο των Ελλήνων;
- Ποια είναι η διαφορά μεταξύ αγοριών και κοριτσιών στην επίδοση τους στα μαθήματα;
- Ποια είναι η σχέση μεταξύ του κοινωνικοοικονομικού επιπέδου της οικογένειας και της επίδοσης των μαθητών;

Η λογική των διαστημάτων εμπιστοσύνης είναι αρκετά απλή:

- βασισμένοι στα στατιστικά του δείγματος και στη θεωρία των πιθανοτήτων φτιάχνομε διαστήματα τιμών μέσα στα οποία θα πέφτουν οι αντίστοιχες παράμετροι με κάποια δεδομένη πιθανότητα (συνήθως 95%). Αυτή την πιθανότητα ονομάζομε «εμπιστοσύνη» (confidence).

Για παράδειγμα, αν στο δείγμα μιας δημοσκόπησης 32% των ερωτηθέντων έχει απαντήσει ότι σκοπεύει να ψηφίσει ΠΑΣΟΚ στις επόμενες εκλογές, μπορούμε να πούμε ότι με πιθανότητα 95% το ποσοστό του πληθυσμού που σκοπεύει να ψηφίσει ΠΑΣΟΚ είναι 32% ± 3% ή 29-35%.

Ο έλεγχος υποθέσεων

Έχουμε τα παρακάτω ερωτήματα:

- Είναι το εκπαιδευτικό επίπεδο των Ελλήνων (δηλ. ο μέσος όρος) διαφορετικό από το μέσο όρο των κρατών της Ευρωπαϊκής Ένωσης;
- Υπάρχει καμιά διαφορά μεταξύ αγοριών και κοριτσιών στην επίδοση τους στα γλωσσικά μαθήματα;
- Υπάρχει καμιά διαφορά στη μαθητική επίδοση στο δημοτικό μεταξύ αυτών που ζουν στις διάφορες γεωγραφικές περιοχές της χώρας;

Οι υποθέσεις :

- Το εκπαιδευτικό επίπεδο των Ελλήνων δεν είναι διαφορετικό από το μέσο όρο των κρατών της Ευρωπαϊκής Ένωσης.
- Δεν υπάρχει καμιά διαφορά μεταξύ αγοριών και κοριτσιών στην επίδοση τους στα γλωσσικά μαθήματα.
- Δεν υπάρχει καμιά διαφορά στη μαθητική επίδοση στο δημοτικό μεταξύ αυτών που ζουν στις διάφορες γεωγραφικές περιοχές της χώρας.

Ο έλεγχος υποθέσεων

Οι έλεγχοι υποθέσεων προσπαθούν να απαντήσουν σε ερωτήσεις του τύπου «υπάρχει διαφορά;», «υπάρχει σχέση;», «υπάρχει επίδραση;» κ.τ.λ..

Όλοι οι έλεγχοι ακολουθούν περίπου την ίδια διαδικασία και λογική. Όλοι αρχίζουν με μιαν υπόθεση, ότι στον **πληθυσμό δεν υπάρχει** σχέση, διαφορά, επίδραση κ.τ.λ..

Η υπόθεση αυτή είναι γνωστή ως **μηδενική υπόθεση**, γιατί συνήθως υποθέτομε ότι οι σχέσεις, οι διαφορές και επιδράσεις είναι μηδενικές.

Για να ελέγξομε τις υποθέσεις αυτές, συλλέγομε δεδομένα από ένα δείγμα πιθανοτήτων και υπολογίζομε τη διαφορά, σχέση επίδραση κ.τ.λ.. Αν στο δείγμα βρούμε ότι μας υπάρχει διαφορά, επίδραση ή σχέση (και συνήθως κάτι υπάρχει), δύο τινα μπορεί να συμβαίνουν:

- Είτε ισχύει η μηδενική μας υπόθεση και βρήκαμε τη διαφορά κατά τύχη (κατά σύμπτωση, κατά λάθος),
- Είτε δεν τη βρήκαμε κατά τύχη και άρα δεν ισχύει η μηδενική μας υπόθεση.

Ο έλεγχος υποθέσεων

- Για να αποφανθούμε περί του τι ισχύει εξετάζομε την πιθανότητα να έχομε βρει τη διαφορά αυτή κατά τύχη.
- Υπολογίζομε, δηλαδή, την πιθανότητα να προέρχεται το δείγμα από έναν πληθυσμό όπου ισχύει η μηδενική υπόθεση (η πιθανότητα αυτή συμβολίζεται με ένα p).
- Αν η πιθανότητα αυτή είναι μικρή, συμπεραίνομε ότι το δείγμα μας δεν προέρχεται από τέτοιο πληθυσμό, αλλά από έναν άλλο, στον οποίο δεν ισχύει η μηδενική υπόθεση.
- Αν, με άλλα λόγια, με βάση τα δεδομένα, η μηδενική υπόθεση φαντάζει απίθανη, την απορρίπτομε και συμπεραίνομε ότι υπάρχει σχέση, διαφορά ή επίδραση στον πληθυσμό.

Ο έλεγχος υποθέσεων

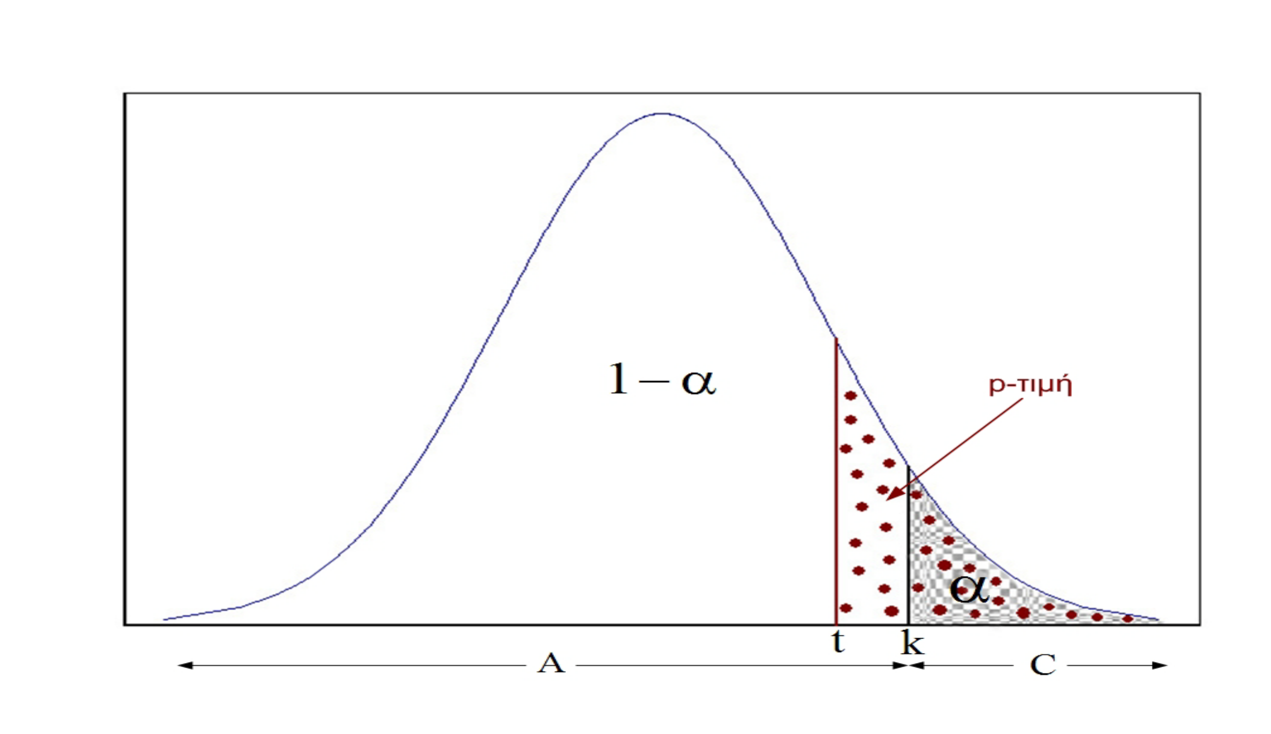
Για να αποφασίσομε αν απορρίπτεται η μηδενική υπόθεση πρέπει :

- να γνωρίζομε πώς να υπολογίσομε το p (την πιθανότητα να έχομε βρει κάτι κατά λάθος) και
- να αποφασίσομε πότε θα λέμε η πιθανότητα αυτή είναι πολύ μικρή (ώστε να λέμε ότι αυτό που βρήκαμε δεν είναι κατά λάθος και άρα δεν ισχύει η μηδενική υπόθεση).

Το πρώτο θα το δούμε παρακάτω για διάφορες περιπτώσεις. Το δεύτερο καθορίζεται a priori, μάλλον «αυθαίρετα», και είναι γνωστό ως επίπεδο στατιστικής σημαντικότητας (συμβολίζεται με ένα α - άλφα). Η λογική τελειώνει κάπως έτσι:

- Αν $p \leq \alpha$ απορρίπτομε τη μηδενική υπόθεση και αποφαινόμαστε ότι υπάρχει στατιστικά σημαντική διαφορά.
- Αν $p > \alpha$ αποτυγχάνομε να απορρίψομε τη μηδενική υπόθεση και αποφαινόμαστε ότι δεν υπάρχει στατιστικά σημαντική διαφορά.

Ο έλεγχος υποθέσεων



- Απορρίπτουμε την H_0 όταν p – τιμή < α .
  - *Για να είμαστε σε θέση να αξιολογήσουμε πόσο αξιόπιστη είναι η p-τιμή που θα βρούμε, πρώτα πρέπει να ελέγξουμε αν ο πληθυσμός μας είναι κανονικός.*
  - *Σε περίπτωση που δεν είναι, θα πρέπει να αξιολογήσουμε αν απέχει πολύ ή όχι.*
  - *Επίσης, θα λάβουμε υπ' όψιν και το μέγεθος του δείγματος.*

Ο έλεγχος υποθέσεων

Γενικά, τις πιθανότητες p μπορούμε να τις βρούμε, αν γνωρίζομε:

- Τη δειγματοληπτική κατανομή (sampling distribution) ενός στατιστικού (ενός οποιουδήποτε στατιστικού, και
- τις πιθανότητες που αντιστοιχούν στις διάφορες τιμές αυτής της κατανομής.

Κάτω από ορισμένες προϋποθέσεις, οι πληροφορίες αυτές υπάρχουν σε πίνακες για τα περισσότερα στατιστικά.

Η κατανομή και ο πίνακας που χρησιμοποιούμε εξαρτάται από το στατιστικό που θέλομε να ελέγξομε. Έτσι, χρησιμοποιούμε άλλη κατανομή (και άλλο πίνακα) όταν θέλομε να ελέγξομε τη διαφορά μεταξύ πολλών μέσων όρων και άλλη όταν θέλομε να ελέγξομε το συντελεστή παλινδρόμησης. Τα κυριότερα στατιστικά που χρησιμοποιούνται στην εκπαίδευση και τις κοινωνικές επιστήμες ακολουθούν, εκτός από την κανονική (**z**), και τις παρακάτω κατανομές:

- Κατανομή t (t Distribution)
- Κατανομή F (F Distribution)
- Κατανομή χ^2 (CHI SQUARE Distribution)

Από την κατανομή που ακολουθούν τα διάφορα στατιστικά, παίρνουν την ονομασία τους και οι έλεγχοι στατιστικής σημαντικότητας που χρησιμοποιούνται για τα στατιστικά αυτά.

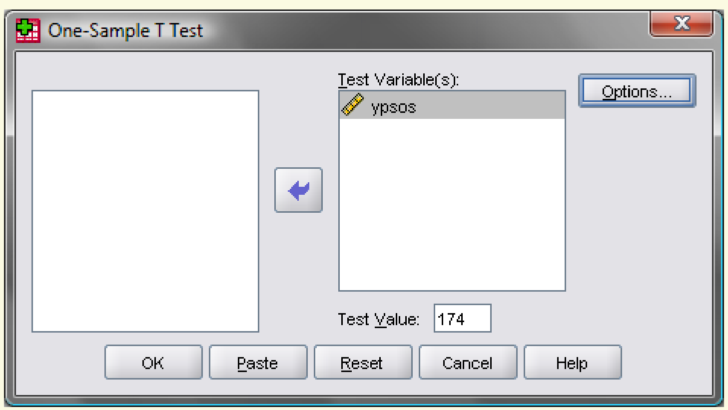
Έλεγχος της μέσης τιμής μίας ποσοτικής μεταβλητής (έλεγχος $\mu=\mu_0$)

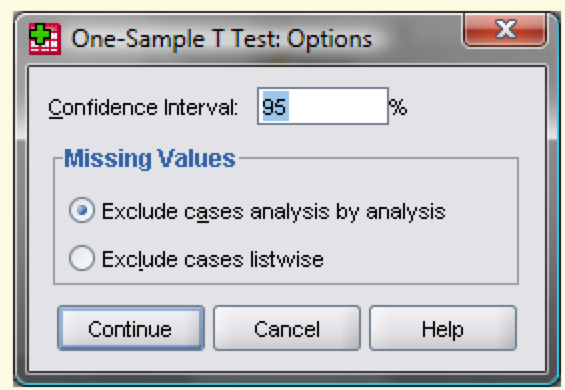
Πολλές φορές θέλουμε να ελέγξουμε, αν η μέση τιμή μίας μεταβλητής είναι (στατιστικά) ίση με έναν συγκεκριμένο αριθμό. Για παράδειγμα, αν ο μέσος όρος του ύψους ενός τυχαίου δείγματος 60 ανδρών μ είναι ίσος με μ_0=174 εκατοστά (test2)

H_0: Μηδενική υπόθεση $\mu=\mu_0$ έναντι της H_α: Εναλλακτικής υπόθεσης $\mu\neq\mu_0$

Η κατάλληλη δοκιμασία σε αυτή την περίπτωση είναι το One sample t-test, όταν **ισχύει** η προϋπόθεση της **κανονικότητας** της μεταβλητής μας.

SPSS: Analyze → Compare Means → One sample t test και δηλώνουμε την τιμή ελέγχου μ_0 **(test value).**





Αποτελέσματα

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
ypsos	60	174,6683	3,01749	,38956

Ο πρώτος πίνακας περιέχει την τιμή του μέσου, της τυπικής απόκλισης, της τυπικής απόκλισης του μέσου (τυπικό σφάλμα του μέσου) και το μέγεθος του δείγματος

One-Sample Test

	Test Value = 174					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
ypsos	1,716	59	,091	,66833	-,1112	1,4478

Η πρώτη στήλη του πίνακα περιέχει την ονομασία της μεταβλητής,

Η δεύτερη στήλη περιέχει την τιμή του t (1,716) τεστ για αυτή τη μεταβλητή, η οποία χρησιμοποιήθηκε για να υπολογιστεί η p-value (0,091) (Sig. (2-tailed)). Η στήλη δίπλα από την p-value περιέχει τη διαφορά ανάμεσα στην τιμή της μηδενικής υπόθεσης (174) και στη μέση τιμή της μεταβλητής. Οι επόμενες δύο στήλες περιέχουν ένα 95% διάστημα εμπιστοσύνης για αυτή τη διαφορά. Το επίπεδο στατιστικής σημαντικότητας είναι ίσο με **0,091**. Εφόσον είναι **μεγαλύτερο** του **0,05**, καταλήγουμε στο συμπέρασμα ότι δεν μπορούμε να απορρίψουμε τη μηδενική υπόθεση ότι ο πραγματικός μέσος είναι ίσος με 174.

Σταδιακή ανάλυση

Ερευνητικό ερώτημα:

«Ο μέσος όρος του ύψους ενός δείγματος 60 ανδρών διαφέρει από 174 εκατοστά ($\mu_{0)}$»

ΒΗΜΑ 1ο: Γράφουμε τη μηδενική υπόθεση, η οποία είναι αντίθετη με το ερευνητικό ερώτημα:

$H_0: \mu = \mu_0$ ή $H_0: \mu - \mu_0 = 0$

Όπου ***μ***, ο μέσος όρος του δείγματος και μ_0 ο αριθμός που θέλουμε να ερευνήσουμε αν διαφέρει από τον μέσο όρο.

ΒΗΜΑ 2ο: Γράφουμε την εναλλακτική υπόθεση, η οποία είναι αντίθετη με τη μηδενική και περίπου ίδια με το ερευνητικό ερώτημα:

$H_a: \mu \neq \mu_0$ ή $H_a: \mu - \mu_0 \neq 0$

ΒΗΜΑ 3ο: Ορίζουμε το επίπεδο στατιστικής σημαντικότητας (α):

$\alpha = 0{,}05$

ΒΗΜΑ 4ο: Βρίσκουμε την πιθανότητα (**p**) να είναι ο μέσος του δείγματος να είναι διαφορετικός από την ελεγχόμενη τιμή, «**κατά λάθος**»

ΒΗΜΑ 5ο: Συγκρίνομε την πιθανότητα **p** με το **α**. Αν **p** > **α** αποτυγχάνομε να απορρίψομε τη μηδενική υπόθεση και συμπεραίνομε ότι δεν υπάρχει στατιστικά σημαντική διαφορά μεταξύ του μέσου όρου και της συγκρινόμενης τιμής

Έλεγχος της μέσης τιμής - Μη παραμετρικό Τεστ

Το αντίστοιχο μη παραμετρικό τεστ καλείται έλεγχος των προσημασμένων τάξεων μεγέθους του **Wilcoxon** για τη διάμεσο ενός πληθυσμού.

Επειδή βασίζεται στις τάξεις μεγέθους των παρατηρήσεων και όχι στις παρατηρήσεις αυτές κάθε αυτές δε χρειάζεται **καμία προϋπόθεση** ως προς την κατανομή των παρατηρήσεων. Αυτό **ισχύει** για όλα τα μη παραμετρικά τεστ που θα δούμε.

Ο έλεγχος όμως εδώ βασίζεται στη **διάμεσο** και όχι στο μέσο του δείγματος. Αυτό βέβαια δε σημαίνει ότι αυτού του είδους τα τεστ είναι χαμηλότερης αξιοπιστίας σε σύγκριση με τα παραμετρικά. Απεναντίας, σε πολλές περιπτώσεις είναι **πιο ισχυρά**.

Γενικά όμως η ισχύς τους πλησιάζει πάρα πολύ των ισχύ των κλασικών παραμετρικών τεστ. Αυτός είναι και ο λόγος που η μόνη προϋπόθεση που χρειάζεται είναι η συμμετρία της κατανομής.

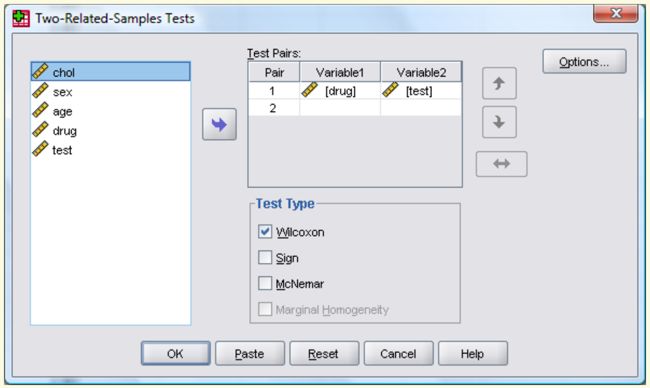
Ένα **πλεονέκτημα** όμως των μη παραμετρικών ελέγχων που θα εξεταστούν είναι ότι εφαρμόζονται και στις περιπτώσεις που οι μεταβλητές είναι **ποιοτικές διατεταγμένης κλίμακας** (καλό, καλύτερο, πολύ καλό). Σε αυτήν την περίπτωση οι μέθοδοι ελέγχων υποθέσεων που θα δούμε είναι από **τις πιο ισχυρές** μεθόδους που υπάρχουν

Έλεγχος της μέσης τιμής - Μη παραμετρικό Τεστ

Ο έλεγχος όμως αυτός δεν προσφέρεται από το **SPSS** μέσω του μενού επιλογών, για αυτό θα πρέπει να γίνει μία διεργασία πρώτα. Η διεργασία έχει ως εξής:

- θα περάσουμε σε μία νέα στήλη την τιμή την οποία θέλουμε να ελέγξουμε τόσες φορές, όσες και οι παρατηρήσεις της μεταβλητής στην οποία θέλουμε να κάνουμε τον έλεγχο υπόθεσης.
- Επομένως θέλοντας να ελέγξουμε με ένα ***μη παραμετρικό τεστ*** αν ο μέσος όρος του ύψους ενός τυχαίου δείγματος 60 ανδρών μ είναι ίσος με μ_0=174 εκατοστά στο αρχείο test2 θα προσθέσουμε μια νέα μεταβλητή (π.χ. Diam) που θα περιέχει την τιμή 174, 60 φορές, όσες δηλαδή είναι και οι παρατηρήσεις του δείγματος υπό εξέταση
- Επιλέγουμε από το μενού εντολών τα εξής: **Analyze→Nonparamteric Tests→2 Related Samples**

Αποτελέσματα



Η επιλογή **Options** μας επιτρέπει να εμφανίσουμε και κάποια περιγραφικά μέτρα. Η μία μεταβλητή θα είναι η στήλη της οποίας τη διάμεσο ελέγχουμε και η άλλη μεταβλητή η στήλη με την τιμή την οποία ελέγχουμε (174 για το παράδειγμα)

Test Statistics[b]

	diam - ypsos
Z	-1,684[a]
Asymp. Sig. (2-tailed)	,092

a. Based on positive ranks.

b. Wilcoxon Signed Ranks Test

Από τα αποτελέσματα του ελέγχου μας ενδιαφέρει κυρίως η γραμμή της Asymp. Sig. (2-tailed). Η τιμή p-value που υπολογίζεται μέσω του τεστ είναι 0,092. Αφού είναι μεγαλύτερη του 0,05, οδηγούμαστε στο συμπέρασμα ότι δεν μπορούμε να απορρίψουμε την μηδενική υπόθεση. Δηλαδή η διάμεσος του δείγματος δεν διαφέρει στατιστικά σημαντικά από την τιμή που ελέγξαμε (174).

Σύγκριση των μέσων τιμών μίας μεταβλητής σε δύο ανεξάρτητα δείγματα (έλεγχος $\mu_1=\mu_2$)

Πολλές φορές θέλουμε να ελέγξουμε, αν η μέση τιμή μιας ποσοτικής μεταβλητής διαφέρει σε δύο ανεξάρτητα δείγματα. Για παράδειγμα, αν το μέσο ύψος του δείγματος των 60 ανδρών (test2) του προηγούμενου παραδείγματος διαφέρει από το μέσο ύψος ενός νέου δείγματος 40 γυναικών (test 3).

Μηδενική υπόθεση: $\mu_1=\mu_2$ έναντι της Εναλλακτικής υπόθεσης: $\mu_1\neq\mu_2$

Η κατάλληλη δοκιμασία σε αυτή την περίπτωση είναι το **Independent Samples t-test**, αρκεί να ισχύουν οι προϋποθέσεις:

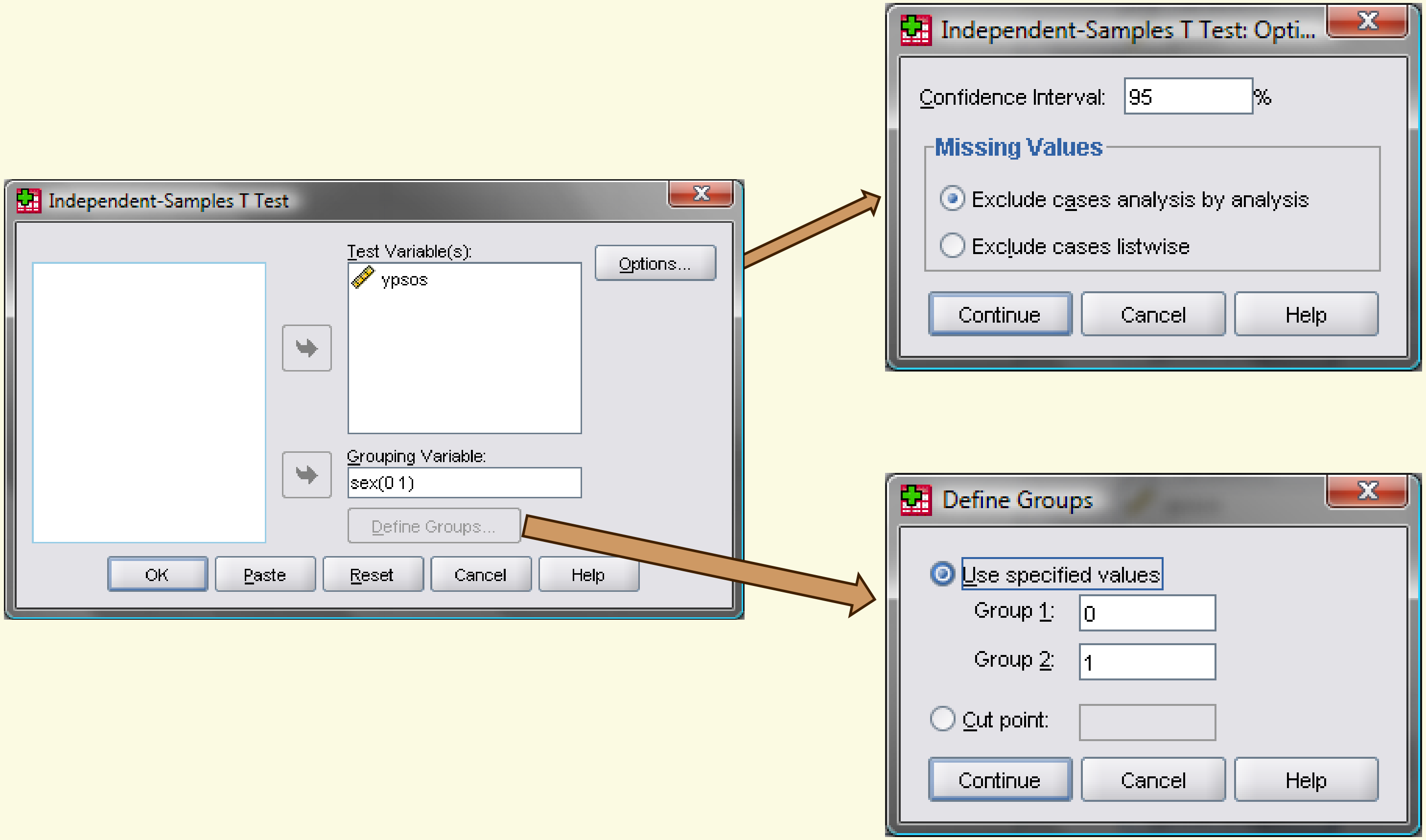
A) και οι δυο να κατανέμονται κανονικά

B) οι διασπορές τους να μην απέχουν πολύ.

Αν οι προϋποθέσεις αυτές δεν ισχύουν, τότε ανατρέχουμε σε μη παραμετρικό τεστ.

SPSS: Analyze → Compare Means → Independent Samples t-test και δηλώνουμε την μεταβλητή που ξεχωρίζει τα δύο ανεξάρτητα δείγματα (grouping variable)

Διαδικασία


Independent-Samples T Test
Test Variable(s):
ypsos
Options...
Grouping Variable:
sex(0 1)
Define Groups...
OK
Paste
Reset
Cancel
Help
Independent-Samples T Test: Opti...
Confidence Interval: 95 %
Missing Values
Exclude cases analysis by analysis
Exclude cases listwise
Continue
Cancel
Help
Define Groups
Use specified values
Group 1: 0
Group 2: 1
Cut point:
Continue
Cancel
Help

Αποτελέσματα

Group Statistics

	sex	N	Mean	Std. Deviation	Std. Error Mean
ypsos	man	60	174,6683	3,01749	,38956
	woman	40	176,3200	3,38179	,53471

Ο πρώτος πίνακας περιέχει κάποια περιγραφικά μέτρα για τα δύο δείγματα. Ο δεύτερος πίνακας είναι αυτός που μας ενδιαφέρει. Το t τεστ έχει δύο "**κατευθύνσεις**". Η μία κατεύθυνση είναι αυτή που δεν μπορούμε να υποθέσουμε ότι οι διακυμάνσεις των δύο δειγμάτων είναι περίπου ίσες και αυτή που μπορούμε να υποθέσουμε ότι είναι ίσες.

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
ypsos	Equal variances assumed	,320	,573	-2,555	98	,012	-1,65167	,64656	-2,93475	-,36859
	Equal variances not assumed			-2,497	77,041	,015	-1,65167	,66156	-2,96900	-,33434

Αποτελέσματα

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means					95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
ypsos	Equal variances assumed	,320	,573	-2,555	98	,012	-1,65167	,64656	-2,93475	-,36859
	Equal variances not assumed			-2,497	77,041	,015	-1,65167	,66156	-2,96900	-,33434

Το **τεστ του Levene** ελέγχει την υπόθεση της ισότητας των δύο διακυμάνσεων και υπολογίζει μία **p-value**. Αν η p-value είναι **μικρότερη** του 0,05, απορρίπτεται η υπόθεση της ισότητας των διακυμάνσεων. Στην αντίθετη περίπτωση δεν απορρίπτεται. Επομένως, ανάλογα με την p-value (Sig.) του τεστ του Levene, κοιτάζουμε την πρώτη ή τη δεύτερη γραμμή αποτελεσμάτων. Στην προκειμένη περίπτωση η p-value είναι μεγαλύτερη του 0,05, άρα μπορούμε να υποθέσουμε ισότητα των δύο διακυμάνσεων. Επομένως θα κοιτάξω τη πρώτη γραμμή αποτελεσμάτων του πίνακα.

Συνεχίζουμε λοιπόν με την πρώτη γραμμή της ανάλυσης που αφορά **equal variances** και έχουμε σύμφωνα με το p-value την απόρριψη της $\mathbf{H_0 : \mu 1 - \mu 2 = 0}$ σε ε.σ. a = 0,05 (0,012 < 0,05). Δηλαδή φαίνεται ότι **υπάρχει** στατιστικά σημαντική διαφορά στα μέσα ύψη των δυο πληθυσμών.

Σταδιακή ανάλυση

Ερευνητικό ερώτημα:

«Υπάρχει διαφορά στο ύψος μεταξύ των ανδρών και των γυναικών του δείγματος»

ΒΗΜΑ 1ο: Γράφουμε τη μηδενική υπόθεση, η οποία είναι αντίθετη με το ερευνητικό ερώτημα:

$H_0: \mu_\alpha = \mu_\gamma$ ή $H_0: \mu_\alpha - \mu_\gamma = 0$

Όπου μ_α, ο μέσος όρος του ύψους των ανδρών και μ_γ ο μέσος όρος του ύψους των γυναικών

ΒΗΜΑ 2ο: Γράφουμε την εναλλακτική υπόθεση, η οποία είναι αντίθετη με τη μηδενική και περίπου ίδια με το ερευνητικό ερώτημα:

$H_a: \mu_\alpha \neq \mu_\gamma$ ή $H_a: \mu_\alpha - \mu_\gamma \neq 0$

ΒΗΜΑ 3ο: Ορίζουμε το επίπεδο στατιστικής σημαντικότητας (α):

$\alpha = 0,05$

ΒΗΜΑ 4ο: Βρίσκουμε την πιθανότητα **(p)** να είναι ο μέσος του δείγματος να είναι διαφορετικός από την ελεγχόμενη τιμή, «**κατά λάθος**»

ΒΗΜΑ 5ο: Συγκρίνομε την πιθανότητα **p** με το **α.** Αν **p** > **α** αποτυγχάνομε να απορρίψομε τη μηδενική υπόθεση και συμπεραίνομε ότι δεν υπάρχει στατιστικά σημαντική διαφορά μεταξύ των συγκρινόμενων μέσων όρων.

Σύγκριση των μέσων τιμών μίας μεταβλητής – Μη Παραμετρικό Τεστ

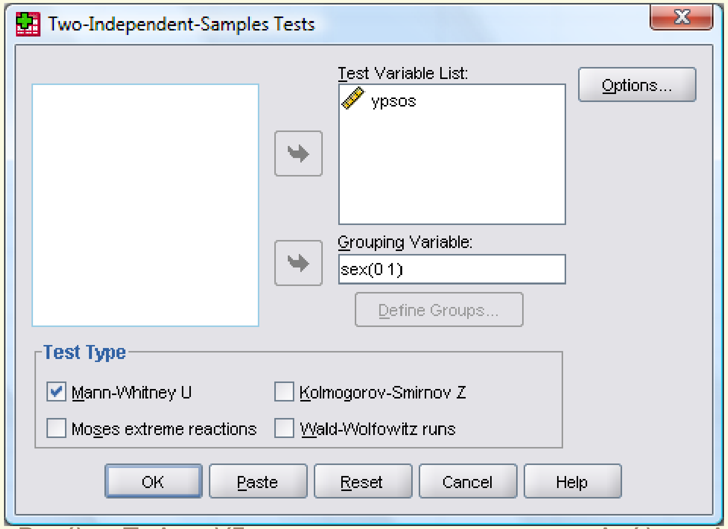
Το αντίστοιχο μη παραμετρικό ανάλογο του **t τεστ** είναι το τεστ των Mann-Whitney-Wilcoxon. Η διαδικασία την οποία πρέπει να κάνουμε για εκτελέσουμε αυτό το μη παραμετρικό τεστ είναι ίδια με προηγουμένως (συνένωση των δύο μεταβλητών σε μία στήλη κ.λ.π.).

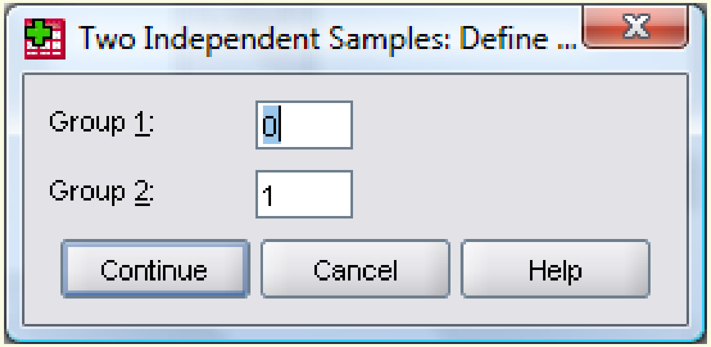
Οι υποθέσεις σε αυτήν την περίπτωση ορίζονται ως εξής:

– **H_0:** Οι άνδρες και οι γυναίκες δεν διαφέρουν ως προς το ύψος τους

– **H_1:** Οι άνδρες και οι γυναίκες διαφέρουν ως προς το ύψος τους

Επιλέγουμε **Analyze→Nonparametric Tests→2 Independent Samples**





Αποτελέσματα

Test Statistics[a]

	ypsos
Mann-Whitney U	818,500
Wilcoxon W	2648,500
Z	-2,685
Asymp. Sig. (2-tailed)	,007

a. Grouping Variable: sex

Μας ενδιαφέρει κυρίως η γραμμή της Asymp. Sig. (2-tailed). Η τιμή **p-value** που υπολογίζεται μέσω του τεστ είναι 0,007. Αφού είναι μικρότερη του 0,05 οδηγούμαστε στο συμπέρασμα ότι μπορούμε να απορρίψουμε την μηδενική υπόθεση.

Δηλαδή ότι ο μέσος όρος του ύψους των ανδρών διαφέρει με στατιστικά σημαντικό τρόπο από τον μέσο όρο του ύψους των γυναικών

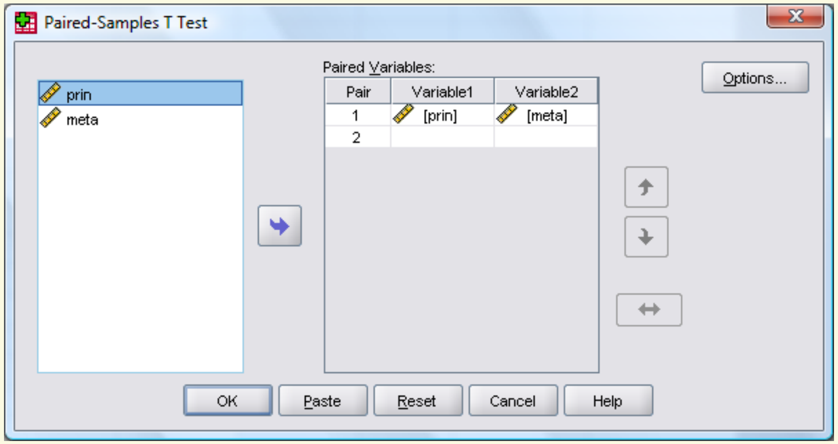
Σύγκριση των μέσων τιμών δύο εξαρτημένων δειγμάτων

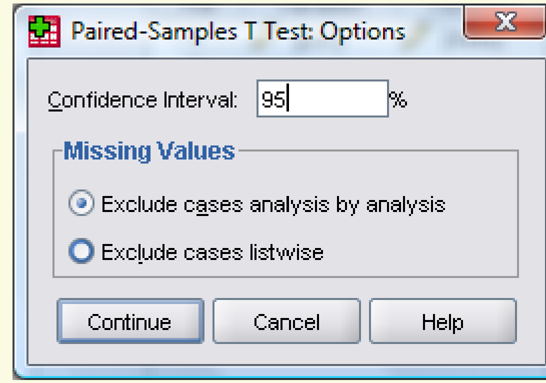
Παραμετρικό Τεστ

- Είδαμε πως να κάνουμε έλεγχο υποθέσεων για δύο ανεξάρτητα δείγματα. Τι γίνεται όμως όταν τα δύο δείγματα δεν είναι, ή δεν μπορούμε να υποθέσουμε ότι είναι ανεξάρτητα;
- Την απάντηση τη δίνουν το **t τεστ** για ζεύγη παρατηρήσεων και ο έλεγχος των προσημασμένων τάξεων μεγέθους του **Wilcoxon** για δείγμα ζευγών παρατηρήσεων.
- Κλασικό παράδειγμα εφαρμογής των δύο αυτών τεστ είναι στην ιατρική, όταν έχουμε μετρήσεις για κάποια άτομα πριν και μετά από μία δίαιτα και ενδιαφερόμαστε να δούμε κατά πόσο ήταν αποτελεσματική η δίαιτα ή όχι. Και τις δύο φορές μετρήσαμε το βάρος των ίδιων ατόμων. Άρα γίνεται σαφές ότι τα μεγέθη των δύο δειγμάτων πρέπει να είναι απαραίτητα ίσα.
- **Παράδειγμα :** Για τον έλεγχο της αποτελεσματικότητας ενός σκευάσματος που καταπολεμά την παχυσαρκία, χορηγήθηκε συγκεκριμένη ποσότητά του σε 20 κατάλληλα πειραματόζωα. Σε καθένα από αυτά καταγράφηκε το βάρος του αμέσως πριν και μια εβδομάδα μετά την χορήγηση του σκευάσματος (test5).

Διαδικασία – Αποτελέσματα

Για το t τεστ εργαζόμαστε ως εξής: επιλέγουμε **Analyze → Compare Means → Paired - Samples T Test**





Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	prin	80,9700	20	5,50551	1,23107
	meta	79,1850	20	5,01779	1,12201

Ο πρώτος πίνακας (Paired Samples Statistics) περιέχει το μέγεθος των δειγμάτων, τους μέσους, τις τυπικές αποκλίσεις των μέσων και τις τυπικές αποκλίσεις κάθε δείγματος.

Διαδικασία – Αποτελέσματα

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	prin & meta	20	,824	,000

Τα δεδομένα που χρησιμοποιήθηκαν στο παράδειγμα προέρχονται από μέτρηση του βάρους 20 πειραματόζωων που χρησιμοποιήθηκαν για τον έλεγχο της αποτελεσματικότητας ενός σκευάσματος που καταπολεμά την παχυσαρκία πριν και μετά την χορήγηση του. Πρόκειται δηλαδή για τα ίδια πειραματόζωα γεγονός που είναι εύλογο να υποθέσουμε εξάρτηση μεταξύ των δύο μετρήσεων του βάρους.

Ο πίνακας (Paired Samples Correlations) επιβεβαιώνει την υπόθεση αυτή. Ο συντελεστής γραμμικής συσχέτισης που υπολογίστηκε έχει υψηλή τιμή, φανερώνοντας υψηλή γραμμική συσχέτιση μεταξύ των δύο μετρήσεων. Το παρατηρηθέν επίπεδο στατιστικής σημαντικότητας για τον έλεγχο της υπόθεσης ότι δεν υπάρχει γραμμική συσχέτιση μεταξύ των δύο μετρήσεων βάρους είναι ίσο με 0, γεγονός που σημαίνει ότι ο συντελεστής γραμμικής συσχέτισης είναι στατιστικά σημαντικός

Διαδικασία – Αποτελέσματα

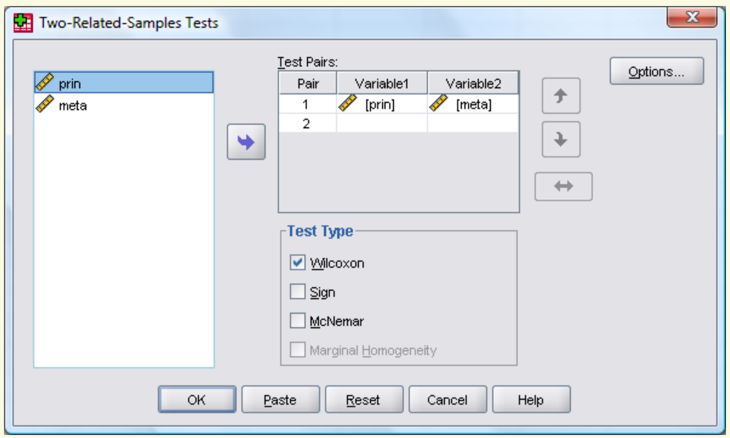
Paired Samples Test

		Paired Differences						
				95% Confidence Interval of the Difference				
	Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	Sig. (2-tailed)
Pair 1 prin - meta	1,78	3,15	,70548	,30842	3,26158	2,530	19	,020

Η τελευταία στήλη του παραπάνω πίνακα περιέχει το παρατηρηθέν επίπεδο στατιστικής σημαντικότητας, το οποίο είναι ίσο με 0,02 Αφού είναι μικρότερο του 0,05 συμπεραίνουμε ότι οι μέσοι των δύο μετρήσεων του βάρους πριν και μετά την χορήγηση του σκευάσματος διαφέρουν με στατιστικά σημαντικό τρόπο. Επομένων το σκεύασμα έφερε τα επιδιωκόμενα αποτελέσματα.

Σύγκριση των μέσων τιμών δύο εξαρτημένων δειγμάτων
Μη Παραμετρικό Τεστ

Το t τεστ όμως προϋποθέτει κανονικότητα των δεδομένων αν και αυτό το πρόβλημα μπορεί να ξεπεραστεί. Το αντίστοιχο μη παραμετρικό τεστ του Wilcoxon εκτελείται στο SPSS επιλέγοντας τα εξής: **Analyze→Nonparametric Tests→2 Related Samples**



Test Statistics[b]

	meta - prin
Z	-2,278[a]
Asymp. Sig. (2-tailed)	,023

a. Based on positive ranks.

b. Wilcoxon Signed Ranks Test

Μας ενδιαφέρει κυρίως η γραμμή της Asymp. Sig. (2-tailed). Η τιμή p-value που υπολογίζεται μέσω του τεστ είναι 0,023. Αφού είναι μικρότερητου 0,05, οδηγούμαστε στο συμπέρασμα ότι μπορούμε να απορρίψουμε την μηδενική υπόθεση. Δηλαδή συμπεραίνουμε ότι οι μέσοι των δύο μετρήσεων του βάρους πριν και μετά την χορήγηση του σκευάσματος διαφέρουν με στατιστικά σημαντικό τρόπο

Συντελεστές γραμμικής συσχέτισης

- Οι συντελεστές που θα εξετάσουμε αναφέρονται στη γραμμικής φύσεως σχέση που μπορεί να συνδέει τις δύο μεταβλητές.
- Οι τιμές που μπορεί να πάρει ένας συντελεστής συσχέτισης είναι από -1 έως +1.
- Αρνητικές τιμές του συντελεστή γραμμικής συσχέτισης δύο μεταβλητών σημαίνει ότι έχουμε την ύπαρξη αρνητικής γραμμικής συσχέτισης. Δηλαδή, οι μεγαλύτερες τιμές της μίας μεταβλητής τείνουν να αντιστοιχούν στις μικρότερες τιμές της άλλης μεταβλητής.
- Θετικές τιμές του συντελεστή γραμμικής συσχέτισης είναι ένδειξη θετικής γραμμικής συσχέτισης μεταξύ των δύο μεταβλητών. Δηλαδή, οι μεγαλύτερες τιμές της μίας μεταβλητής τείνουν να αντιστοιχούν στις μεγαλύτερες τιμές της άλλη μεταβλητής.
- Τιμές κοντά στο μηδέν αποτελούν ένδειξη ότι δεν υπάρχει στατιστικά σημαντική γραμμική συσχέτιση μεταξύ των δύο μεταβλητών.
- Όσο πιο μεγάλες είναι οι τιμές του συντελεστή, ή όσο πιο κοντά βρίσκονται στη μονάδα (σε απόλυτη τιμή πάντα), τόσο πιο ισχυρή είναι η γραμμική συσχέτιση μεταξύ τους.
- Οι πιο γνωστοί συντελεστές γραμμικής συσχέτισης είναι οι συντελεστές του **Pearson**, του **Spearman** και του **Kendall**.

Συντελεστές γραμμικής συσχέτισης

Η μηδενική και η εναλλακτική υπόθεση εδώ είναι οι εξής:

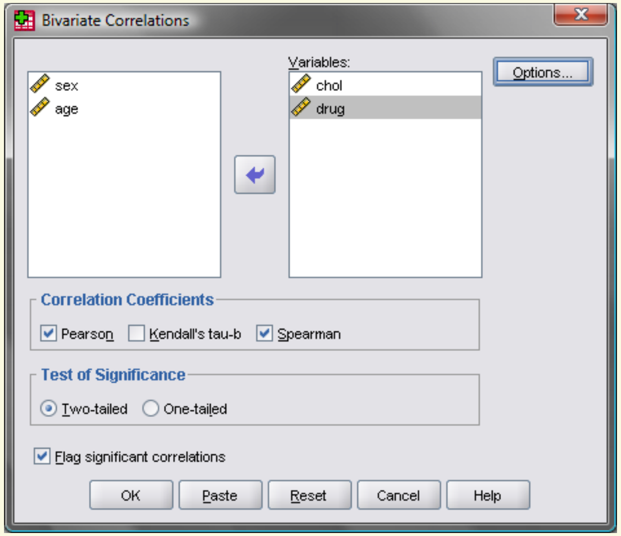
H_0: ρ=0 ή δεν υπάρχει γραμμική συσχέτιση μεταξύ των δύο μεταβλητών

H1: ρ≠0 ή υπάρχει γραμμική συσχέτιση μεταξύ των δύο μεταβλητών

- Ο συντελεστής συσχέτισης του Pearson “χρειάζεται” την υπόθεση της κανονικότητας των δεδομένων, σε αντίθεση με τους άλλους δύο που δεν “χρειάζονται” την υπόθεση της κανονικότητας των δεδομένων.
- Βέβαια, για μεγάλα δείγματα, μεγέθους 30 παρατηρήσεων και πάνω και όσο το μέγεθος του δείγματος μεγαλώνει η θεωρία μας λέει ότι οι τιμές των συντελεστών “πλησιάζουν” η μία την άλλη.
- Η κύρια διαφορά των συντελεστών είναι ότι ο συντελεστής του Pearson υπολογίζεται με βάση τα δεδομένα, ενώ οι άλλοι δύο υπολογίζονται με βάση τις τάξεις μεγέθους των δεδομένων. Ειδικότερα, ο συντελεστής του Spearman είναι ο συντελεστής του Pearson στην ουσία υπολογισμένος για τις τάξεις μεγέθους των δεδομένων.
- Το γεγονός λοιπόν ότι οι συντελεστές του Spearman και του Kendall υπολογίζονται με βάση τις τάξεις μεγέθους των δεδομένων είναι που επιτρέπει την ελευθερία ως προς τη μη ικανοποίηση της κανονικότητας των μεταβλητών.

Συντελεστές γραμμικής συσχέτισης

Για να υπολογίσουμε τους τρεις αυτούς συντελεστές συσχέτισης στο SPSS επιλέγουμε τα εξής: **Analyze→Correlate→Bivariate**



Στο κουτάκι “variables” πρέπει να περάσουμε τουλάχιστον δύο μεταβλητές, διότι οι συντελεστές συσχέτιση υπολογίζονται για ζεύγη μεταβλητών. Οπότε αν περάσουμε περισσότερες από δύο μεταβλητές, θα υπολογιστούν οι συντελεστές γραμμικής συσχέτισης για όλα τα ζεύγη των μεταβλητών.

Βλέπουμε ότι μόνο ο συντελεστής του Pearson είναι επιλεγμένος. Αν θέλουμε να εμφανιστούν και οι άλλοι δύο συντελεστές απλά τους επιλέγουμε.

Η επιλογή Options μας δίνει τη δυνατότητα εμφάνισης των μέσων, των τυπικών αποκλίσεων και των πληθών των τιμών για κάθε μεταβλητή. Παρατηρήστε ότι στο κάτω αριστερό μέρος του παραθύρου είναι επιλεγμένη μία επιλογή (Flag significant correlations).

Συντελεστές γραμμικής συσχέτισης - Αποτελέσματα

Correlations

		chol	drug
chol	Pearson Correlation	1,000	,354**
	Sig. (2-tailed)		,006
	N	60	60
drug	Pearson Correlation	,354**	1,000
	Sig. (2-tailed)	,006	
	N	60	60

**. Correlation is significant at the 0.01 level (2-tailed).

Correlations

			chol	drug
Spearman's rho	chol	Correlation Coefficient	1,000	,357**
		Sig. (2-tailed)	.	,005
		N	60	60
	drug	Correlation Coefficient	,357**	1,000
		Sig. (2-tailed)	,005	.
		N	60	60

**. Correlation is significant at the 0.01 level (2-tailed).

Βλέπουμε ότι για όλες τις τιμές των συντελεστών γραμμικής συσχέτισης υπάρχουν δύο αστεράκια. Αυτό γίνεται μέσω της επιλογής Flag significant correlations.

Κάτω από κάθε τιμή του συντελεστή συσχέτισης εμφανίζεται μία p-value (Sig. (2-tailed)). Η p-value που έχει υπολογιστεί για κάθε συντελεστή ξεχωριστά και αναφέρεται στον έλεγχο της υπόθεσης ότι στο συγκεκριμένο ζεύγος μεταβλητών δεν υπάρχει γραμμική συσχέτιση.

Αφού το παρατηρηθέν επίπεδο στατιστικής σημαντικότητας είναι μικρότερο του 0,05, συμπεραίνουμε ότι αυτή η υπόθεση απορρίπτεται σε α=0,05.

Άρα υπάρχει στατιστικά σημαντική γραμμική συσχέτιση μεταξύ του ζεύγους. Στην περίπτωση που η p-value είναι μικρότερη του 0,01, τότε ο συντελεστής συσχέτισης εμφανίζεται με δύο αστεράκια αντί για μόνο ένα.

Συντελεστές γραμμικής συσχέτισης - Αποτελέσματα

- Σε αυτό το σημείο καλό θα ήταν να αναφέρουμε ότι ο συντελεστής του Kendall μπορεί να χρησιμοποιηθεί και στην περίπτωση που έχουμε κατηγορικές μεταβλητές οι οποίες όμως είναι υποχρεωτικά σε κλίμακα διάταξης. Είναι δηλαδή διατακτικές κατηγορικές μεταβλητές.
- Ακόμα να αναφέρουμε ότι με το συντελεστή γραμμικής συσχέτισης ελέγχουμε αν σε ένα ζεύγος μεταβλητών υπάρχει γραμμική συσχέτιση μόνο. Δηλαδή μπορεί να υπάρχει συσχέτιση μεταξύ των δύο μεταβλητών, αλλά όχι γραμμικής φύσεως. Σε αυτήν την περίπτωση αυτή η σχέση που συνδέει τις δύο μεταβλητές δεν μπορεί να ανιχνευτεί με το συντελεστή γραμμικής συσχέτισης.
- Οπότε προσοχή στην ερμηνεία που δίνουμε στο συντελεστή συσχέτισης. Να υπενθυμίσουμε επίσης ότι η λογική με την οποία απορρίπτουμε ή όχι μία υπόθεση είναι πάντα η ίδια. Αν το παρατηρηθέν επίπεδο στατιστικής σημαντικότητας είναι μικρότερο του 0,05 η υπόθεση απορρίπτεται. Στην αντίθετη περίπτωση δεν απορρίπτεται

Έλεγχος F

Ο έλεγχος F χρησιμοποιείται για τα στατιστικά των οποίων η δειγματοληπτική κατανομή ακολουθεί την κατανομή F.

Η ανάλυση διακύμανσης χρησιμοποιείται επίσης για να εξεταστεί η υπόθεση ότι οι μέσοι όροι μιας μέτρησης (εξαρτημένη μεταβλητή, ποσοτική) δεν διαφέρουν μεταξύ ομάδων που δημιουργεί μια ανεξάρτητη μεταβλητή (ποιοτική). Η ανάλυση αυτή είναι κατ' ουσία μία επέκταση του τεστ ανεξαρτησίας αφού συμπεριλαμβάνει στην ανάλυση περισσότερες των δύο ομάδες τις τιμές της ανεξάρτητης μεταβλητής.

Τα στατιστικά τα οποία ακολουθούν την κατανομή F και θα παρουσιασθούν εδώ είναι:

- Η διαφορά μεταξύ περισσοτέρων των δύο αριθμητικών μέσων
- Η διαφορά μεταξύ πολλαπλών μετρήσεων

Η Διαδικασία του Ελέγχου Στατιστικής Σημαντικότητας της Διαφοράς Μεταξύ Πολλών Αριθμητικών Μέσων (Ομάδων)

Έλεγχος F (ANOVA)

Το ερευνητικό ερώτημα:

«Υπάρχει διαφορά στην μέτρηση της χοληστερίνης μεταξύ του δείγματος που λαμβάνουν διαφορετική ποσότητα φαρμάκου (η δοσολογία φαρμάκου χωρίζει το δείγμα σε πέντε κατηγορίες).

ΒΗΜΑ 1ο: Γράφουμε τη μηδενική υπόθεση η οποία είναι αντίθετη με το ερευνητικό ερώτημα: $\boldsymbol{H_0}$**:** $\boldsymbol{\mu_0 = \mu_1 = \mu_2 = \mu_3 = \mu_4}$ (δεν υπάρχει διαφορά)

όπου $\boldsymbol{\mu_0}$: ο μέσος όρος των ανθρώπων που λαμβάνουν καθόλου φάρμακο, $\boldsymbol{\mu_1}$: ο μέσος όρος που λαμβάνουν λίγο φάρμακο, $\boldsymbol{\mu_2}$ **,** $\boldsymbol{\mu_3}$, $\boldsymbol{\mu_4}$ ………

ΒΗΜΑ 2ο: Γράφομε την εναλλακτική υπόθεση, η οποία είναι αντίθετη με τη μηδενική και περίπου ίδια με το ερευνητικό ερώτημα: $\boldsymbol{H_a}$**:** $\boldsymbol{\mu_0 = \mu_1 = \mu_2 = \mu_3 = \mu_4}$ ***δεν ισχύει*** (υπάρχει διαφορά)

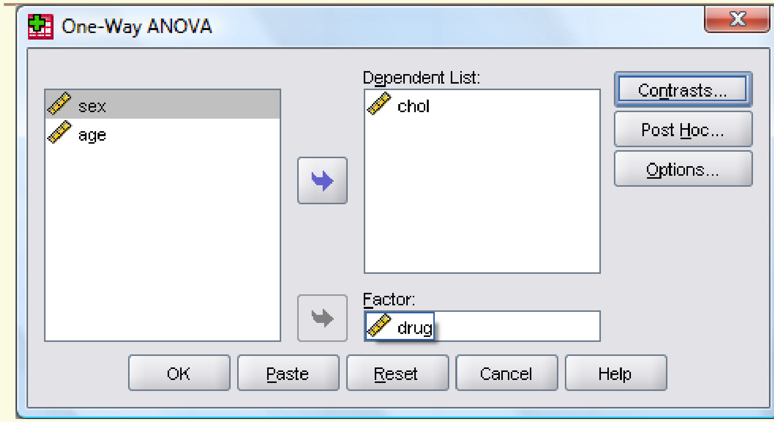
ΒΗΜΑ 3ο: Ορίζομε το επίπεδο στατιστικής σημαντικότητας (α):

$\alpha = 0{,}05$

ΒΗΜΑ 4ο: Βρίσκομε την πιθανότητα (**p**) να υπάρχει διαφορά, «κατά λάθος».

ΒΗΜΑ 5ο: Συγκρίνομε την πιθανότητα **p** με το **α.** Αν **p** < **α** απορρίπτουμε τη μηδενική υπόθεση και συμπεραίνομε ότι υπάρχει στατιστικά σημαντική διαφορά μεταξύ των μέσων ορών

Ανάλυση διακύμανσης (διασποράς) κατά ένα παράγοντα (One-way Anova).



Από τα μενού **analyze**, *compare means* και One-Way ANOVA. Στη θέση **factor** ο χρήστης τοποθετεί την ανεξάρτητη χωρίς να χρειάζεται ο ορισμός των τιμών των ομάδων.

ANOVA

chol

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	17624,056	3	5874,685	2,686	,055
Within Groups	122462,527	56	2186,831		
Total	140086,583	59			

Η πρώτη στήλη του πίνακα (Sum of Squares) αναφέρεται στο άθροισμα τετραγώνων μεταξύ των ομάδων, στο άθροισμα τετραγώνων εντός των ομάδων και τέλος στο σύνολο των ανωτέρω.

Στη δεύτερη στήλη εμφανίζονται βαθμοί ελευθερίας μεταξύ των ομάδων (k-1) δηλαδή εδώ 2, οι βαθμοί ελευθερίας για το σύνολο των παρατηρήσεων (cases) και τέλος οι βαθμοί ελευθερίας που αντιστοιχούν στην μεταβλητότητα εντός των ομάδων.

Στη τρίτη στήλη εμφανίζονται τα μέσα τετράγωνα (mean square) που προέρχονται από διαιρέσεις των αθροισμάτων τετραγώνων με τους βαθμούς ελευθερίας και η διαίρεση των οποίων δίνει το στατιστικό F με το επίπεδο σημαντικότητας στην τελευταία στήλη. Στην προκειμένη περίπτωση δεν υπάρχει στατιστικά σημαντική διαφορά μεταξύ των ομάδων ($p>0,05$).

One-way Anova

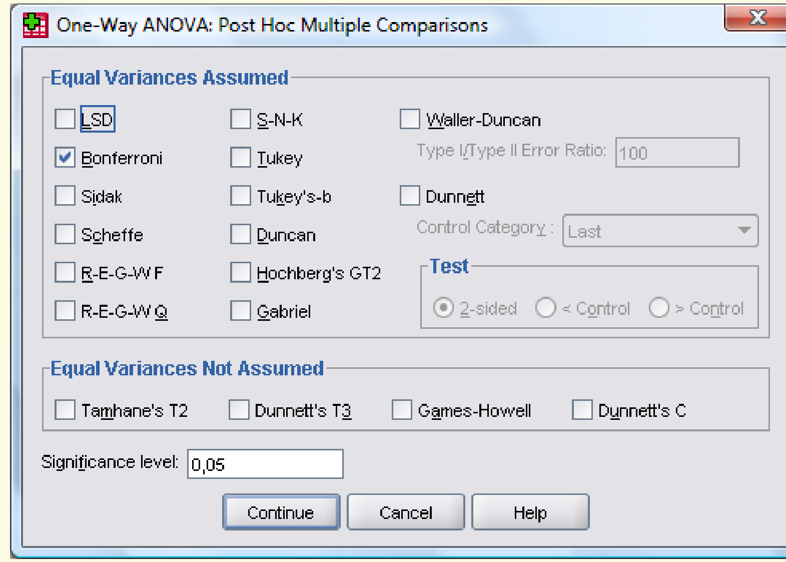
Η ανάλυση αυτή μας δίνει μια γενική αλλά όχι πλήρη απάντηση.

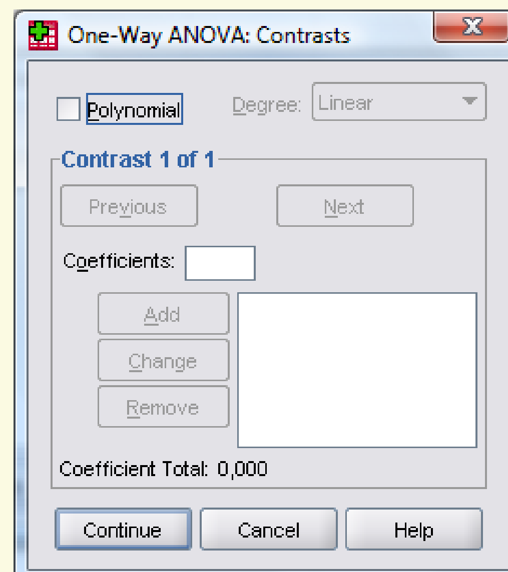
Η στατιστική σημαντικότητα μέχρι εδώ υποδηλώνει ότι δεν υπάρχουν στατιστικά σημαντικές διαφορές αλλά δεν μας λέει ούτε πόσες ούτε ποιες διαφορές δεν είναι στατιστικά σημαντικές.

Για να πάρομε τις λεπτομερείς αυτές πληροφορίες πρέπει να κάνουμε ένα ακόμη βήμα.

Πριν πατήσομε “OK” στο προηγούμενο πλαίσιο διαλόγου, πατούμε το κουμπί “Post Hoc…” και στο πλαίσιο που ανοίγει επιλέγομε έναν από τους ελέγχους Πολλαπλών Συγκρίσεων (Multiple Comparisons) που παρουσιάζονται, όπως φαίνεται στην παρακάτω εικόνα.

Εδώ επιλέγουμε τον έλεγχο “Bonferonni





One-way Anova - Αποτελέσματα

Multiple Comparisons

Chol Bonferroni

(I) drug	(J) drug	Mean Dif	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
καθόλου	ελάχιστη	-17,35294	16,03977	1,000	-61,2252	26,5194
	μικρή	-29,37433	18,09530	,661	-78,8690	20,1203
	μέτρια	-45,71373*	16,56580	,047	-91,0248	-,4026
ελάχιστη	καθόλου	17,35294	16,03977	1,000	-26,5194	61,2252
	μικρή	-12,02139	18,09530	1,000	-61,5160	37,4732
	μέτρια	-28,36078	16,56580	,555	-73,6719	16,9503
μικρή	καθόλου	29,37433	18,09530	,661	-20,1203	78,8690
	ελάχιστη	12,02139	18,09530	1,000	-37,4732	61,5160
	μέτρια	-16,33939	18,56318	1,000	-67,1138	34,4350
μέτρια	καθόλου	45,71373*	16,56580	,047	,4026	91,0248
	ελάχιστη	28,36078	16,56580	,555	-16,9503	73,6719
	μικρή	16,33939	18,56318	1,000	-34,4350	67,1138

*. The mean difference is significant at the 0.05 level.

Όπως φαίνεται στον πίνακα (από τις τιμές των p) υπάρχουν στατιστικά σημαντικές διαφορές μόνο μεταξύ της ομάδας που παίρνει καθόλου δόση φαρμάκου και της ομάδας που παίρνει μέτρια δόση φαρμάκου ενώ μεταξύ των υπόλοιπων ομάδων δεν υπάρχουν στατιστικά σημαντικές διαφορές.

Έλεγχος χ2

- Ο έλεγχος που χρησιμοποιείται για τις σχέσεις μεταξύ δύο ή περισσότερων ονομαστικών ή και τακτικών μεταβλητών είναι γνωστός ως έλεγχος **χ^2** (**Chi Square**). Αν και υπάρχουν επί μέρους στατιστικοί έλεγχοι για τα διάφορα μέτρα συνάφειας, ο έλεγχος χ^2 είναι ένας *γενικός* έλεγχος που μπορεί να χρησιμοποιηθεί για όλες τις περιπτώσεις που μπορούμε να φτιάξομε ένα πίνακα διμεταβλητών συχνοτήτων (CROSS-TABS), δηλαδή για όλες σχεδόν τις συνάφειες μη ισοδιαστημικών μεταβλητών.
- Η απαιτούμενη κλίμακα μέτρησης των μεταβλητών είναι η ονομαστική, παρόλο που και μεταβλητές με διατακτική κλίμακα μπορούν να χρησιμοποιηθούν.
- Ο **X^2 έλεγχος** ανεξαρτησίας χρησιμοποιείται για τον έλεγχο της υπόθεσης ότι δύο κατηγορικές μεταβλητές είναι ανεξάρτητες μεταξύ τους.
- Οι κατηγορικές μεταβλητές μπορούν να έχουν οσαδήποτε επίπεδα (ή κατηγορίες), αρκεί βέβαια η κάθε μία να έχει τουλάχιστον δύο επίπεδα.
- Το SPSS, όταν υπολογίζομε τους συντελεστές συσχετισμού (φ, λ κ.ά.), δίνει και το p που αντιστοιχεί στο καθένα στην τελευταία στήλη κάτω από την ένδειξη "Approximate Significance". Για τον έλεγχο στατιστικής σημαντικότητας του κάθε συντελεστή, δεν έχομε παρά να συγκρίνομε το p με το α.

Παράδειγμα

- Έχουμε ταξινομήσει ένα δείγμα 419 γυναικών ανάλογα με το αν πάσχουν από κατάθλιψη και αν είχαν κάποια τραυματική εμπειρία στη ζωή τους.
- Το ερώτημα είναι αν υπάρχει εξάρτηση μεταξύ της τραυματικής εμπειρίας και της κατάθλιψης.

	Κατάθλιψη		
Τραυματική Εμπειρία	Όχι (0)	Ναι (1)	Σύνολο
Όχι (0)	251	4	255
Ναι (1)	131	33	164
Σύνολο	382	37	419

- Η μηδενική υπόθεση είναι πάντα αυτή που **δεν υποθέτει** εξάρτηση (υποθέτει ανεξαρτησία μεταξύ των μεταβλητών). Η προϋπόθεση που απαιτείται από τον X^2 έλεγχο ανεξαρτησίας είναι οι συχνότητες των κελιών να είναι τουλάχιστον ίσες με 5.
- Το SPSS χρησιμοποιεί το άλλο είδος υπόθεσης που θέλει τις αναμενόμενες συχνότητες των κελιών να είναι τουλάχιστον ίσες με 5. Ένα αποδεκτό ποσοστό κελιών που θα έχουν συχνότητες μικρότερες του 5 είναι το 25%, δηλαδή το πολύ ένα στα τέσσερα κελιά να έχει μία τιμή μικρότερη του 5 χωρίς να μειώνεται σημαντικά η αποτελεσματικότητα του τεστ. Αυτό ισχύει βέβαια και για πίνακες που έχουν περισσότερα κελιά. Αν αυτή η υπόθεση δεν ικανοποιείται, τότε κοιτάζουμε την p-value που υπολογίζεται με βάση το ακριβές τεστ του Fisher

Σταδιακή ανάλυση

Το ερευνητικό ερώτημα: «Το ερώτημα είναι αν υπάρχει εξάρτηση μεταξύ της τραυματικής εμπειρίας και της κατάθλιψης»;

ΒΗΜΑ 1ο: Γράφουμε τη μηδενική υπόθεση:

H_0: $\chi^2 = 0$ Δεν υπάρχει Σχέση

ΒΗΜΑ 2ο: Γράφουμε την εναλλακτική υπόθεση:

H_a: $\chi^2 \neq 0$ Υπάρχει Σχέση

ΒΗΜΑ 3ο: Ορίζουμε το επίπεδο στατιστικής σημαντικότητας (α):

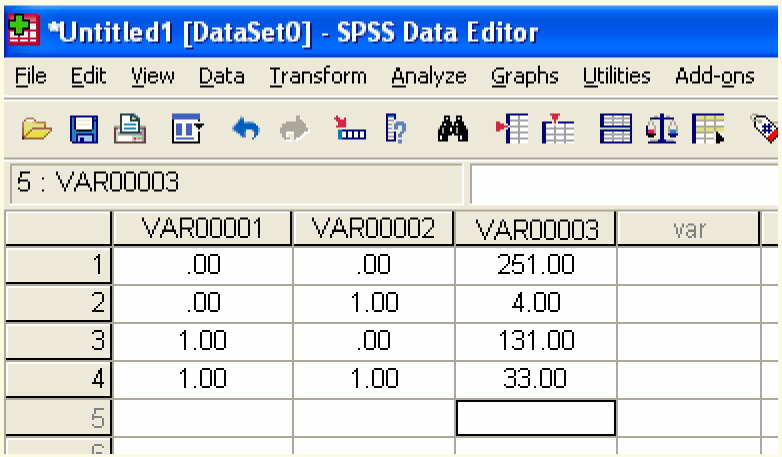
α = 0,05

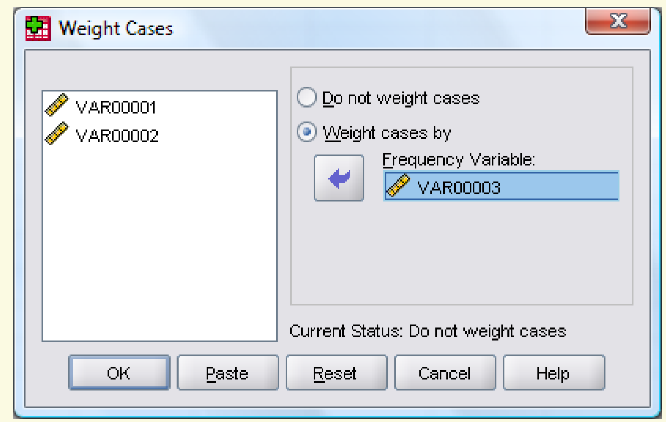
ΒΗΜΑ 4ο: Βρίσκομε την πιθανότητα (**p**) να είναι το χ2 διαφορετικό το 0, «κατά λάθος».

ΒΗΜΑ 5ο: Συγκρίνουμε την πιθανότητα **p** με το **α.** Αν **p** < **α** (< *0,05)* απορρίπτομε τη μηδενική υπόθεση και συμπεραίνομε ότι υπάρχει στατιστικά σημαντική σχέση μεταξύ των δύο μεταβλητών.

Εισαγωγή δεδομένων (πίνακα συχνοτήτων)

- Μέσα στις παρενθέσεις έχουν τοποθετηθεί το (0 και 1) για να διευκολύνουν στο να περαστούν τα δεδομένα στο SPSS. Πρέπει να προσέξουμε ώστε ο κάθε συνδυασμός γραμμής και στήλης να περιέχει τον αριθμό του κελιού που πρέπει. Περνώντας τα δεδομένα στο SPSS θα έχουν την εξής μορφή (σημασία έχει ο κάθε συνδυασμός γραμμής και στήλης να περιέχει το σωστό αριθμό):

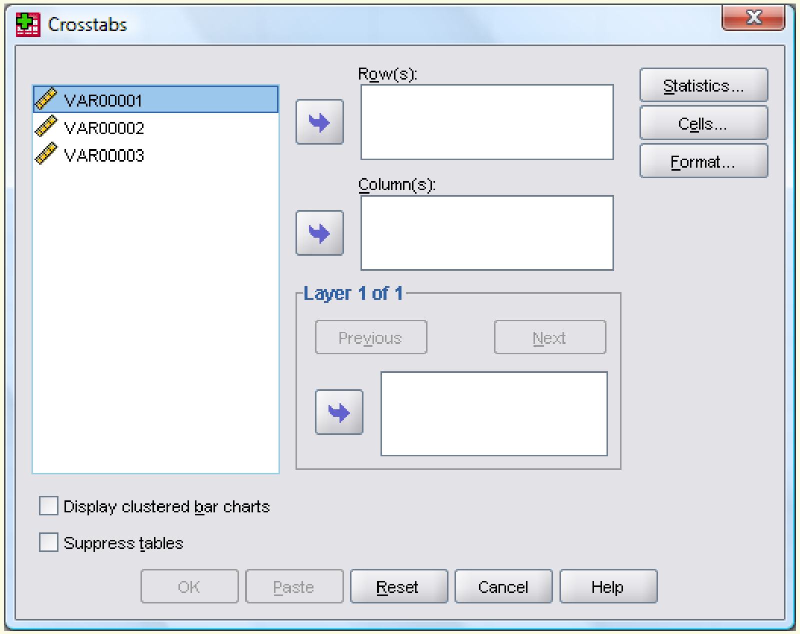




- Για να "δώσουμε" στο SPSS να "καταλάβει" ότι η τρίτη στήλη περιέχει τις συχνότητες των κελιών θα επιλέξουμε τα εξής: Data→ Weight Cases
- Θα επιλέξουμε Weight cases by και στο λευκό κουτάκι από κάτω θα περάσουμε την τρίτη στήλη (VAR00003 για την περίπτωση μας) ή τη στήλη που περιέχει τις συχνότητες των κελιών.

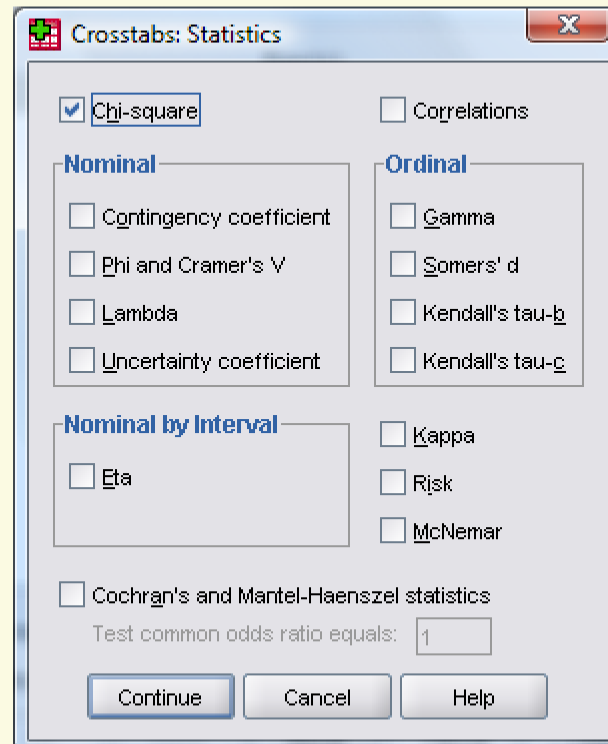
Διαδικασία

Επιλέγουμε **Analyze→Descriptive Statistics→Crosstabs**



Επιλέγοντας Display clustered bar charts θα εμφανιστεί ένα ραβδόγραμμα .

Πατώντας Statistics θα εμφανιστεί το παράθυρο, στο οποίο θα επιλέξουμε Chi-square



Αποτελέσματα

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	42,675[a]	1	**,000**		
Continuity Correction[b]	40,402	1	**,000**		
Likelihood Ratio	44,365	1	**,000**		
Fisher's Exact Test				**,000**	**,000**
Linear-by-Linear Association	42,573	1	**,000**		
N of Valid Cases	419				

a. 0 cells (,0%) have expected count less than 5. The minimum expected count is 14,48.

b. Computed only for a 2x2 table

Αυτό που κοιτάζουμε είναι οι p-value για κάθε τεστ. Κοιτάζουμε το Asymp. Sig. (2 sided) και το Exact Sig. (2-sided) για τα Pearson Chi-Square και Fisher's Exact Test. Μπορούμε φυσικά να κοιτάξουμε και τις p-value για το Likelihood Ratio.

Το πρώτο μήνυμα μας πληροφορεί για το αν ικανοποιείται η προϋπόθεση ισχύος του χι-τετράγωνο τεστ. Θέλουμε το πολύ το 25% των κελιών να έχουν τιμές μικρότερες από 5. Αν δεν ισχύει αυτό τότε δεν εμπιστευόμαστε τα αποτελέσματα του χι-τετράγωνο ελέγχου, παρά μόνο του Fisher για την περίπτωση δισδιάστατων πινάκων Στη βιβλιογραφία αναφέρεται και μία πιο αυστηρή προϋπόθεση όσον αφορά τα κελιά με αριθμούς μικρότερους του 5 και η οποία θέλει όλα τα κελιά να έχουν τιμές μεγαλύτερες του 5.

Πίνακας ανακεφαλαίωσης Ι

ΣΤΟΧΟΣ	**Παραμετρικός Έλεγχος**	**Προϋποθέσεις**	**Εντολές SPSS**	**Μη Παραμετρικός Έλεγχος**
Σύγκριση ενός Μέσου Όρου- με μια τιμή	*One Sample* **t-test**	**Ανεξάρτητες** Παρατηρήσεις **Τυχαίο** δείγμα **Κανονική** κατανομή	**Statistics** … Compare Means … One-Sample t Test...	
Σύγκριση δύο Ομάδων (Μέσων Όρων)	*Independent Samples* **test**	**Ανεξάρτητα Τυχαία** δείγματα **Κανονικές** κατανομές **Ομοιογένεια διασποράς**	**Statistics** …. Compare Means … Independent-Samples t Test...	*Mann-Whitney U test*
Σύγκριση δύο Μετρήσεων ή δύο Μεταβλητών (Μέση Διαφορά)	*Paired Samples* **t-test**	**Ανεξάρτητες Παρατηρήσεις** **Τυχαίο** δείγμα **Κανονική** κατανομή	Statistics Compare Means Paired - Samples t Test..	*Wilcoxon Signed-rank test*
Σχέση μεταξύ δύο Ισοδιαστημικών Μεταβλητών	**t-test**	**Τυχαίο** δείγμα **Γραμμική** Σχέση **Κανονική** κατανομή	Statistics Correlate Bivariate...	
Επίδραση μιας Ανεξάρτητης Μεταβλητής σε μια Εξαρτημένη	**t-test**	**Τυχαίο** δείγμα **Γραμμική** Σχέση **Κανονική** κατανομή **Ομοιογένεια** διασποράς	Statistics Regression Linear…	

Πίνακας ανακεφαλαίωσης ΙΙ

ΣΤΟΧΟΣ	Παραμετρικός Έλεγχος	Προϋποθέσεις	Εντολές SPSS	Μη Παραμετρικός Έλεγχος
Έλεγχος Ενός Ποσοστού	***z-test***	***Ανεξάρτητες Παρατηρήσεις*** ***Τυχαίο** δείγμα* ***Κανονική** κατανομή*	***Statistics*** ***Compare Means*** ***Paired-Samples t Test..***	***Wilcoxon Signed-rank test***
Σύγκριση Δύο Ποσοστών	***z-test***	***Ανεξάρτητες Παρατηρήσεις*** ***Τυχαίο** δείγμα* ***Κανονική** κατανομή*	***Statistics*** ***Compare Means*** ***Paired-Samples t Test..***	***Wilcoxon Signed-rank test***
Σύγκριση περισσοτέρων των δύο Ομάδων (Μέσων Όρων)	*ANOVA* ***F-test***	***Ανεξάρτητα Τυχαία** δείγματα* ***Κανονικές** κατανομές* *Ομοιογένεια διασποράς*	***Statistics*** ***Compare Means*** ***One-Way ANOVA..***	***Kruskal-Wallis H-test***
Σύγκριση Επαναλαμβανόμενων Μετρήσεων	*ANOVA* ***F-test***	***Ανεξάρτητες Παρατηρήσεις*** ***Τυχαίο** δείγμα* ***Κανονικές** κατανομές* ***Ομοιογένεια διασποράς***		***Friedman Test***
Επίδραση πολλών Ανεξάρτητων Μεταβλητών σε μια Εξαρτημένη	***F-test***	***Τυχαίο** δείγμα* ***Γραμμική** Σχέση* ***Ανεξάρτητες** Παρατηρήσεις* ***Κανονική** κατανομή* ***Ομοιογένεια** διασποράς*	***Statistics*** ***Regression*** ***Linear...***	
Σχέση μεταξύ δύο Ονομαστικών ή Τακτικών Μεταβλητών		***Ανεξάρτητες Παρατηρήσεις*** ***Τυχαίο** δείγμα*	***Statistics*** ***Summarize*** ***Crosstabs...*** ***Chi-square***	χ^2

Διαστήματα Εμπιστοσύνης: Υπολογισμός

Για να φτιάξομε τα διαστήματα εμπιστοσύνης πρέπει να γνωρίζομε τα εξής:

1. *Πάντοτε* φτιάχνομε τα διαστήματα γύρω από τις **παραμέτρους** του πληθυσμού και πάντοτε χρησιμοποιούμε το αντίστοιχο περιγραφικό (ή εξηγητικό) **στατιστικό** για να τα φτιάξομε. Έτσι, αν θέλομε να φτιάξομε διαστήματα εμπιστοσύνης γύρω από τον αριθμητικό μέσο του πληθυσμού (μ), χρησιμοποιούμε τον αριθμητικό μέσο του δείγματος κ.τ.λ.
2. Για την κατασκευή διαστημάτων εμπιστοσύνης πρέπει να γνωρίζομε το τυπικό σφάλμα της παραμέτρου (ή του συνδυασμού των παραμέτρων) που θέλομε να προσεγγίσομε.
3. Πρέπει επίσης να γνωρίζομε την κατάλληλη δειγματοληπτική κατανομή, δηλαδή, το κατάλληλο στατιστικό ελέγχου. Εδώ θα παρουσιάσομε τα πιο συνηθισμένα διαστήματα εμπιστοσύνης, δηλαδή τα διαστήματα εμπιστοσύνης για τα στατιστικά για τα οποία χρησιμοποιήσαμε τους ελέγχους t και z. Αφού καθορίσομε το στατιστικό, βρίσκομε την τιμή του στατιστικού ($t_{\alpha/2}$ $z_{\alpha/2}$) που αντιστοιχεί στο επίπεδο α/2 που επιθυμούμε και στους βαθμούς ελευθερίας από τον αντίστοιχο πίνακα,.
4. Αξίζει να το επαναλάβομε: Για τα διαστήματα εμπιστοσύνης *χρησιμοποιούμε πάντοτε το **κρίσιμο** στατιστικό (την τιμή του πίνακα).*

Διαστήματα Εμπιστοσύνης: Υπολογισμός

- Όταν έχομε αυτές τις πληροφορίες, η κατασκευή των διαστημάτων εμπιστοσύνης ή CI (Confidence Intervals) είναι πολύ απλή:
- Βρίσκομε το γινόμενο του τυπικού σφάλματος επί το κρίσιμο στατιστικό. $t_{a/2}s_b$ Την ποσότητα αυτή την προσθέτομε στο περιγραφικό στατιστικό για να βρούμε το ανώτατο όριο και την αφαιρούμε για να βρούμε το κατώτατο όριο. Επειδή τα διαστήματα εμπιστοσύνης εκτείνονται και από τις δύο πλευρές του στατιστικού, το επίπεδο α διαιρείται δια του 2. Για παράδειγμα το κρίσιμο t για διαστήματα εμπιστοσύνης 95% είναι το $t_{a/2}$ ή $t_{0,025}$.
- Τα διαστήματα εμπιστοσύνης μπορούν να υπολογιστούν σχετικά εύκολα για όλα σχεδόν τα στατιστικά για τα οποία χρησιμοποιήσαμε τους ελέγχους t και z. Για τα περισσότερα από τα στατιστικά αυτά το SPSS υπολογίζει αυτόματα (ή με την κατάλληλη εντολή) τα διαστήματα εμπιστοσύνης. Εδώ θα παρουσιάσομε τον υπολογισμό τους για τους συντελεστές παλινδρόμησης. Να σημειώσομε ότι για τα διαστήματα εμπιστοσύνης πρέπει να ισχύουν οι ίδιες προϋποθέσεις που ισχύουν για τους αντίστοιχους ελέγχους.

ΤΕΛΟΣ