

Κεφάλαιο 6

Παραμετρικοί και μη παραμετρικοί έλεγχοι υποθέσεων

Σύνοψη

Το κεφάλαιο αυτό παρουσιάζει τους γνωστότερους παραμετρικούς και μη παραμετρικούς ελέγχους θεωρητικά και στο πρόγραμμα. Συγκεκριμένα του ελέγχους: T-Test για τον έλεγχο Μέσης Τιμής Πληθυσμού, T-test για τον έλεγχο της διαφοράς δύο μέσων από ανεξάρτητα δείγματα, Mann-Whitney U, T-test για τον έλεγχο της διαφοράς δύο μέσων από συσχετιζόμενα δείγματα, Wilcoxon, Means, One-Way ANOVA για ανεξάρτητα δείγματα, Kruskal-Wallis, Two-Way ANOVA for Independent samples Friedman, χ^2 test, Mc Nemar test.

Σημειώνεται ότι τα δεδομένα των εφαρμογών είναι κατά πλειοψηφία δεδομένα που χρησιμοποιήθηκαν κατά τις παραδόσεις των αντίστοιχων μαθημάτων του ΠΜΣ «Βιοστατιστικής» των πανεπιστημίων Αθηνών και Ιωαννίνων. Για τις εφαρμογές των 6.2, 6.6. και 6.11 χρησιμοποιήθηκαν δεδομένα από την αντίστοιχη σελίδα του πανεπιστημίου του Princeton (βλ. βιβλιογραφικές αναφορές). Πολύτιμες πληροφορίες για τη συγκεκριμένη ύλη υπάρχουν στον αντίστοιχο ιστότοπο του προγράμματος (βλ. βιβλιογραφικές αναφορές). Τέλος για την παράγραφο 6.14 αξιολογή πηγή ήταν η διπλωματική εργασία του Γ. Κουτσοχέρα (βλ. βιβλιογραφικές αναφορές).

Προαπαιτούμενη γνώση

Απαιτείται πολύ καλή γνώση του προηγούμενου κεφαλαίου καθώς και θεωρητική γνώση των κατανομών κατανομής και χ^2 , και εξοικείωση με τη διαδικασία ελέγχου υποθέσεων (έννοιες της ελεγχουσυνάρτησης και της περιοχής απόρριψης της μηδενικής υπόθεσης).

6.1 T-Test για τον έλεγχο Μέσης Τιμής Πληθυσμού

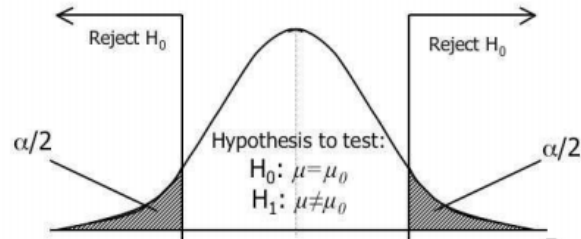
Η δυνατότητα προσφέρεται από την επιλογή *Analyze > Compare Means > One-Sample T-test*, με σκοπό τον έλεγχο της ισότητας της μέσης τιμής ενός πληθυσμού με μια συγκεκριμένη τιμή μ_0 . Δηλαδή οι υποθέσεις που εξετάζονται είναι:

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu \neq \mu_0$$

Εξίσωση 6.1

Η στατιστική συνάρτηση του ελέγχου διαμορφώνεται ανάλογα με το εάν το δείγμα μας προέρχεται από κανονικό πληθυσμό ή όχι, με γνωστή ή άγνωστη διασπορά. Η υπόθεση απορρίπτεται όταν η τιμή του κριτηρίου είναι κατά απόλυτη τιμή μεγαλύτερη από την τιμή που θα λάμβανε εάν ίσχυε η μηδενική υπόθεση και δίνεται από τους αντίστοιχους πίνακες της κανονικής, ή της t-student κατανομής.



Εικόνα 6.1: Περιοχές απόρριψης σε δίπλευρο έλεγχο

Μορφές της ελεγχοσυνάρτησης

1^η Περίπτωση. Διασπορά γνωστή:

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

Εξίσωση 6.2

Απορρίπτουμε την H_0 εάν η περιοχή απόρριψης είναι:

$$R = \{|z| > z_{\alpha/2}\}$$

Εξίσωση 6.3

2^η Περίπτωση. Διασπορά άγνωστη, $n \geq 30$:

$$z = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$$

Εξίσωση 6.4

Απορρίπτουμε την H_0 εάν η περιοχή απόρριψης είναι:

$$R = \{|z| > z_{\alpha/2}\}$$

Εξίσωση 6.5

3^η Περίπτωση. Διασπορά άγνωστη, $n < 30$

$$t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$$

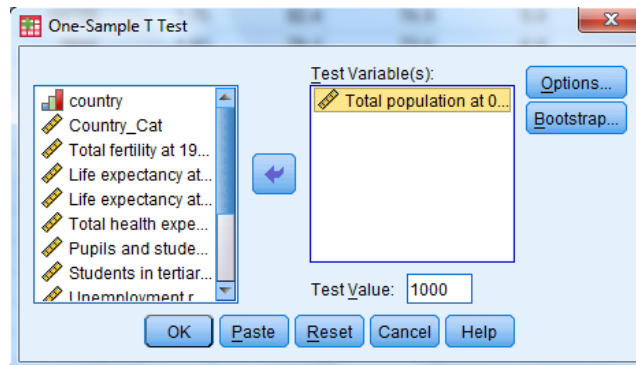
Εξίσωση 6.6

Απορρίπτουμε την H_0 εάν η περιοχή απόρριψης είναι:

$$R = \{|t| > t_{n-1; \alpha/2}\}$$

Εξίσωση 6.7

Στο περιβάλλον του SPSS, στο πεδίο *Test Variable(s)* εισάγουμε την ή τις συνεχείς μεταβλητές που θα εξεταστούν και στο πεδίο *Test Value* εισάγουμε την υπό εξέταση τιμή μ_0 .



Εικόνα 6.2: Το menu One-Sample T-test

Από το πλήκτρο *Options* ορίζουμε το επίπεδο σημαντικότητας για τη δημιουργία των διαστημάτων εμπιστοσύνης της διαφοράς των δύο μέσων (της μέσης τιμής και της υπό εξέτασης τιμής).

Η διαδικασία ολοκληρώνεται με το πλήκτρο *OK* όπου και λαμβάνουμε:

One-Sample Statistics				
	N	Mean	Std. Deviation	Std. Error Mean
Total population at 01/01/1998.1000s	15	24372,20	27121,193	7002,662

One-Sample Test						
	Test Value = 1000					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
Total population at 01/01/1998.1000s	3,338	14	,005	23372,200	8352,98	38391,42

Εικόνα 6.3: Περιγραφικά μέτρα της μεταβλητής και αποτελέσματα ελέγχου της ισότητας της μέσης τιμής με την τιμή 1000

Στον πρώτο πίνακα παρουσιάζονται το πλήθος παρατηρήσεων, η μέση τιμή, η τυπική απόκλιση και το τυπικό σφάλμα μέσης τιμής. Διαπιστώνεται ότι η μέση τιμή 24.372,2 διαφέρει από την υπό εξέταση τιμή $\mu_0 = 1.000$. Η διαφορά αυτή είναι στατιστικά σημαντική σε επίπεδο σημαντικότητας 5%, όπως επιβεβαιώνεται από τα αποτελέσματα του ελέγχου ($p\text{-value} = 0,005 < 0,05$).

6.2 T-test για τον έλεγχο της διαφοράς δύο μέσων από ανεξάρτητα δείγματα

Η δυνατότητα προσφέρεται από την επιλογή *Analyze > Compare Means > Independent-Samples T-test*, με σκοπό τον έλεγχο της συσχέτισης μιας συνεχούς μεταβλητής με μια δίτιμη κατηγορική μεταβλητή, στην οποία οι κατηγορίες αφορούν μετρήσεις που γίνονται σε διαφορετικά άτομα. Για παράδειγμα, εξετάζουμε το πλήθος θανάτων σε σχέση με το φύλο. Η μεταβλητή «Φύλο» είναι δίτιμη και οι μετρήσεις διενεργούνται σε διαφορετικά άτομα: άνδρες / γυναίκες.

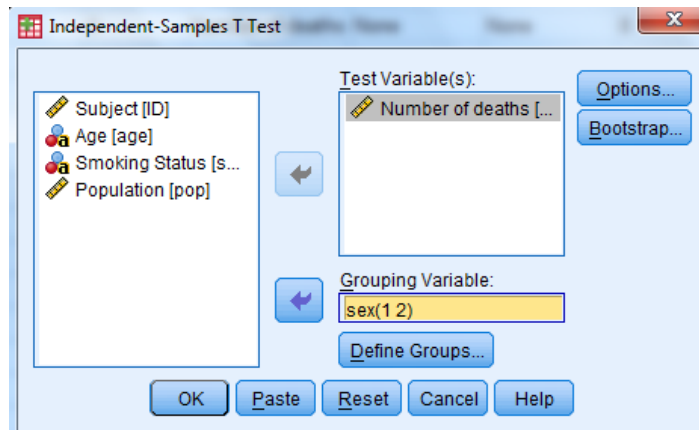
Στην πραγματικότητα η υπόθεση που εξετάζεται είναι η ισότητα των μέσων των δύο ανεξάρτητων ομάδων.

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2$$

Εξίσωση 6.8

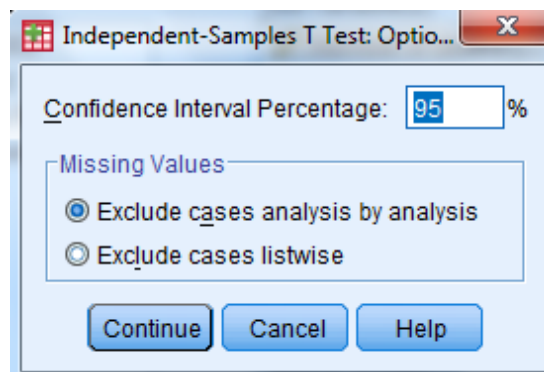
Εάν η υπόθεση της ισότητας δεν απορριφθεί τότε θα συνεπάγεται ότι οι δύο μέσοι είναι ίδιοι στις δύο ομάδες, συνεπώς η κατηγοριοποίηση δεν οδηγεί σε διαφορετικά αποτελέσματα, ή με άλλα λόγια η κατηγορική μεταβλητή δεν σχετίζεται με τη συνεχή μεταβλητή.



Εικόνα 6.4: Το menu Independent-Samples T-test

Στο πεδίο *Test Variable(s)* εισάγουμε την ή τις συνεχείς μεταβλητές που θα εξεταστούν, ενώ στο πεδίο *Grouping Variable* εισάγουμε την κατηγορική μεταβλητή. Καθώς ενδέχεται η κατηγορική που θα επιλεγεί να έχει περισσότερες από δύο κατηγορίες. Μέσω του πλήκτρου *Define Groups* ορίζουμε τους κωδικούς των δύο κατηγοριών που θα χρησιμοποιηθούν για την εφαρμογή του T-test.

Από το πλήκτρο *Options* ορίζουμε το επίπεδο σημαντικότητας για τη δημιουργία των διαστημάτων εμπιστοσύνης της διαφοράς των δύο μέσων $\mu_1-\mu_2$.



Εικόνα 6.5: Το menu Independent-Samples T-test > Options

Η διαδικασία ολοκληρώνεται με το πλήκτρο *OK* όπου και λαμβάνουμε δύο πίνακες.

Στον πρώτο πίνακα παρουσιάζονται τα περιγραφικά μέτρα (στο συγκεκριμένο παράδειγμα του πλήθους θανάτων) στις δύο ομάδες. Απ' αυτά φαίνεται ότι οι άνδρες έχουν κατά μέσο όρο μεγαλύτερο πλήθος θανάτων σε σχέση με τις γυναίκες.

Group Statistics					
Sex		N	Mean	Std. Deviation	Std. Error Mean
Number of deaths	Male	19	266,05	308,119	70,687
	Female	17	239,71	208,927	50,672

Εικόνα 6.6: Περιγραφικά μέτρα του πλήθους θανάτων στις δύο ομάδες

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
Number of deaths	Equal variances assumed	2,962	,094	,297	34	,769	26,347	88,833	-154,184	206,877
	Equal variances not assumed			,303	31,804	,764	26,347	86,973	-150,855	203,549

Εικόνα 6.7: Αποτελέσματα του T-test για ανεξάρτητα δείγματα

Στον δεύτερο πίνακα λαμβάνουμε τα αποτελέσματα του ελέγχου. Στον πίνακα αυτόν εμφανίζονται δύο εγγραφές για τον έλεγχο T-test της ισότητας των δύο μέσων. Αυτό οφείλεται στο ότι για τον υπολογισμό χρησιμοποιούνται διαφορετικοί τύποι, αναλόγως εάν ισχύει ή όχι η ισότητα των διασπορών στις δύο ομάδες. Συνεπώς για να αποφανθούμε ποια αποτελέσματα θα χρησιμοποιήσουμε, ελέγχουμε αρχικά τον έλεγχο Levene για την ισότητα των διασπορών. Ο έλεγχος υποδεικνύει ότι η υπόθεση της ισότητας δεν μπορεί να απορριφθεί σε επίπεδο 5% ($p\text{-value} = 0,094 > 0,05$), συνεπώς θα χρησιμοποιηθούν τα αποτελέσματα του ελέγχου T-test από την πρώτη γραμμή, από όπου και διαπιστώνουμε ότι η υπόθεση της ισότητας δεν απορρίπτεται ($p\text{-value} = 0,769 > 0,05$), άρα το μέσο πλήθος θανάτων μεταξύ ανδρών και γυναικών είναι σχεδόν το ίδιο.

6.3 Mann-Whitney U

Πρόκειται για τον αντίστοιχο μη παραμετρικό έλεγχο της ισότητας δύο μέσων μεταξύ ανεξάρτητων δειγμάτων, τα οποία όμως δεν ακολουθούν κανονική κατανομή. Η δυνατότητα προσφέρεται από την επιλογή *Analyze > Non Parametric Tests > Independent-Samples*.

Το κριτήριο Mann-Whitney είναι ένας βαθμολογικός έλεγχος που εξετάζει την υπόθεση ότι το άθροισμα των βαθμών των θέσεων των παρατηρήσεων του ενός δείγματος είναι ίσο με το άθροισμα των βαθμών των θέσεων των παρατηρήσεων του άλλου δείγματος.

$$H_0 : W_A = W_B$$

$$H_1 : W_A \neq W_B$$

Εξίσωση 6.9

Εάν η υπόθεση ότι τα αθροίσματα είναι ίσα δεν απορριφθεί, τότε συνεπάγεται ότι η κατανομή των παρατηρήσεων της μιας ομάδας είναι σχεδόν ίδια με την κατανομή της άλλης και ότι οι παρατηρήσεις εμφανίζονται με τυχαίο τρόπο τόσο στη μια όσο και στην άλλη ομάδα. Συνεπώς, η διαφοροποίηση λόγω της κατηγορικής μεταβλητής δεν σχετίζεται με τη συνεχή μεταβλητή.

Η ελεγχοσυνάρτηση βασίζεται στη σχέση:

$$U = \min(U_A, U_B)$$

Εξίσωση 6.10

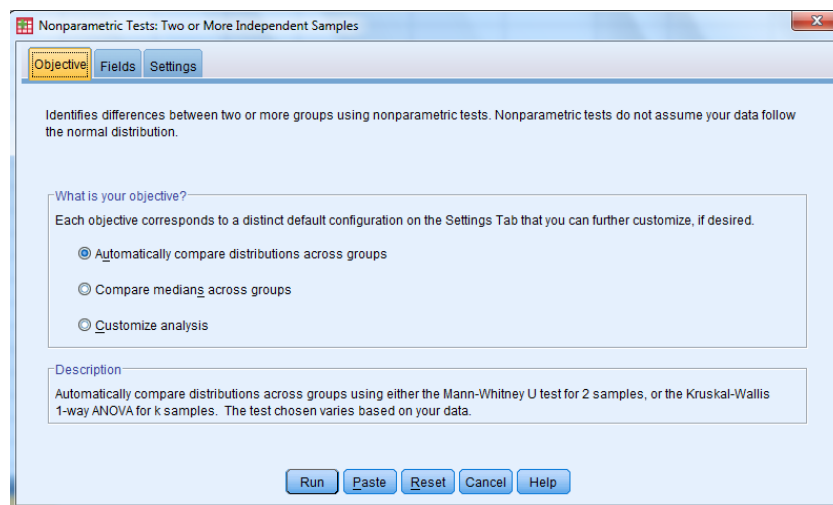
Όπου:

$$U_A = n_A n_B + \frac{n_A(n_A + 1)}{2} - W_A$$

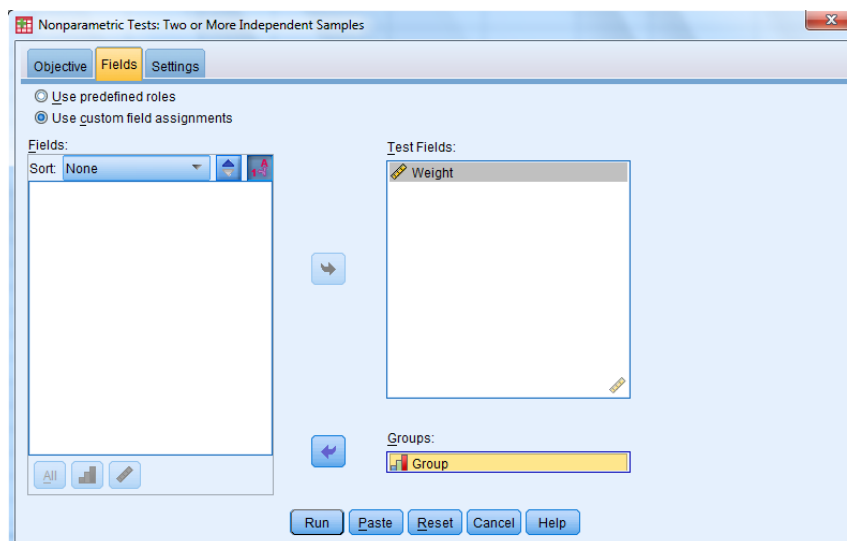
$$U_B = n_A n_B + \frac{n_B(n_B + 1)}{2} - W_B$$

Εξίσωση 6.11

Έστω ότι έχουμε δύο ομάδες ποντικών (ίδιας ηλικίας και φύλου) στα οποία μετά από τυχαιοποίηση χορηγήθηκε ένα πρόγραμμα διατροφής υψηλής (ομάδα Α) και χαμηλής (ομάδα Β) περιεκτικότητας πρωτεϊνών. Στο τέλος του προγράμματος μετρήθηκε το βάρος των ποντικών ώστε να εξεταστεί ποιο είδος διατροφής επιδρά στο βάρος των ποντικών.

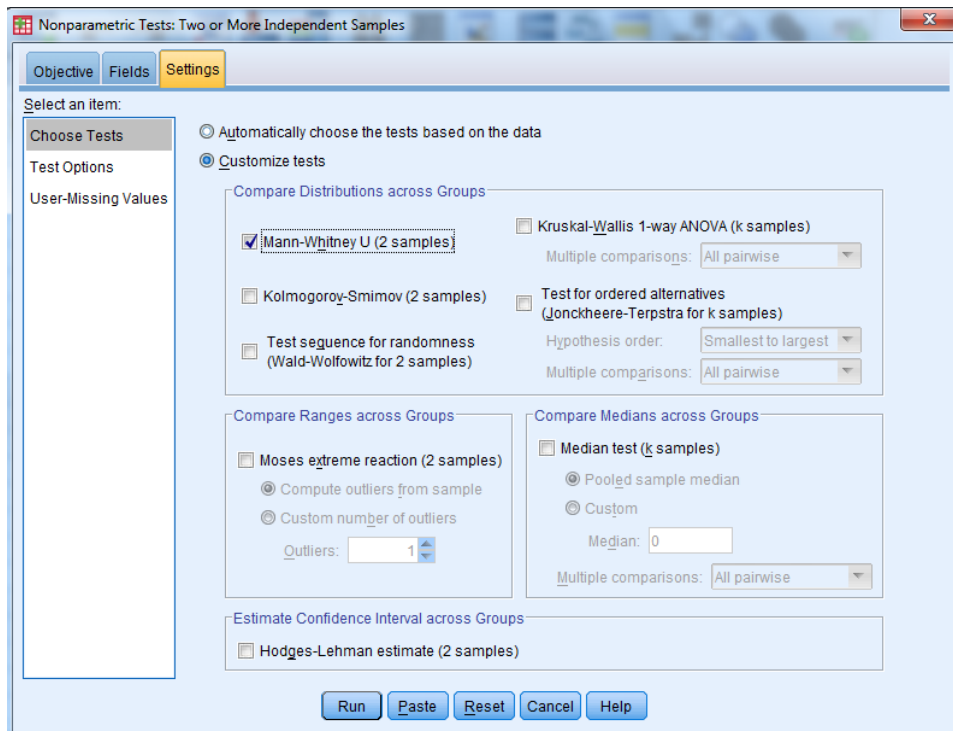


Εικόνα 6.8: To menu Nonparametric Tests: Two or More Independent-Samples



Εικόνα 6.9: To menu Nonparametric Tests: Two or More Independent-Samples > Fields

Στο πεδίο *Test Fields* ορίζουμε τη συνεχή μεταβλητή, και στο πεδίο *Groups* την κατηγορική μεταβλητή.



Εικόνα 6.10: Το menu *Nonparametric Tests: Two or More Independent-Samples* > *Settings*

Στη συνέχεια στην ενότητα *Settings* επιλέγουμε το κριτήριο σύγκρισης των δύο ομάδων *Mann-Whitney U (2 samples)*.

Η διαδικασία ολοκληρώνεται με το πλήκτρο *RUN* όπου και λαμβάνουμε:

Ranks				
	Group	N	Mean Rank	Sum of Ranks
Weight	Group A	12	12.13	145.50
	Group B	8	8.06	64.50
	Total	20		

Εικόνα 6.11: Περιγραφικά μέτρα των βαθμών των θέσεων των παρατηρήσεων

Το SPSS μας δίνει το μέσο της ιεράρχησης ή το μέσο της κατάταξης (Mean Rank) και όχι περιγραφικά μέτρα για τις δύο μεταβλητές ή για τις μετρήσεις της ίδιας μεταβλητής. Στις περιπτώσεις που η κανονικότητα παραβιάζεται συνηθίζουμε να δίνουμε τη διάμεσο ή το εύρος. Καθώς αυτά τα περιγραφικά δεν παρέχονται αυτόματα από το SPSS τα εξάγουμε από την εντολή Means, βλ. ομόνομη ενότητα.

Test Statistics ^b	
	Weight
Mann-Whitney U	28.500
Wilcoxon W	64.500
Z	-1.506
Asymp. Sig. (2-tailed)	.132
Exact Sig. [2*(1-tailed Sig.)]	.135 ^a

a. Not corrected for ties.

b. Grouping Variable: Group

Εικόνα 6.12: Αποτελέσματα του ελέγχου Mann-Whitney U

Ο έλεγχος υποδεικνύει ότι η υπόθεση της ισότητας δεν μπορεί να απορριφθεί σε επίπεδο 5% ($p\text{-value} = 0,132 > 0,05$), άρα δεν μπορούμε να ισχυριστούμε ότι οι δύο μέσοι διαφέρουν.

6.4 T-test για τον έλεγχο της διαφοράς δύο μέσων από συσχετιζόμενα δείγματα

Η δυνατότητα προσφέρεται από την επιλογή *Analyze > Compare Means > Paired-Samples T-test*, με σκοπό τον έλεγχο της συσχέτισης μιας συνεχούς μεταβλητής με μια δίτιμη κατηγορική μεταβλητή, στην οποία οι κατηγορίες αφορούν μετρήσεις που γίνονται στα ίδια άτομα πριν και μετά από τη μεσολάβηση ενός γεγονότος (π.χ. χορήγηση φαρμάκου). Για παράδειγμα, εξετάζουμε εάν οι άνθρωποι χάνουν βάρος, εάν υποβληθούν σε μια διαίτα υψηλή σε υδατάνθρακες και χαμηλά λιπαρά. Τα δύο δείγματα καλούνται εξαρτημένα (correlated) και οι παρατηρήσεις ζευγαρωτές παρατηρήσεις.

Καθώς οι μετρήσεις εφαρμόζονται στο ίδιο άτομο, το δείγμα που προκύπτει προέρχεται από τη διαφορά των δύο μετρήσεων (πριν και μετά το γεγονός). Η υπόθεση που εξετάζεται είναι εάν ο μέσος όρος αυτών των διαφορών ισούται με το μηδέν ή όχι.

$$H_0 : \bar{d} = 0$$

$$H_1 : \bar{d} \neq 0$$

Εξίσωση 6.12

Η ελεγχουσυνάρτηση είναι η:

$$t = \frac{\bar{d}}{\hat{\sigma}_D} \sim t_{N-1}$$

Εξίσωση 6.13

Όπου:

$$\bar{d} = \sum_{i=1}^n \frac{x_i - y_i}{n}$$

Εξίσωση 6.14

και

$$\hat{\sigma}_d = s_d \sqrt{\frac{1}{N}}$$

Εξίσωση 6.15

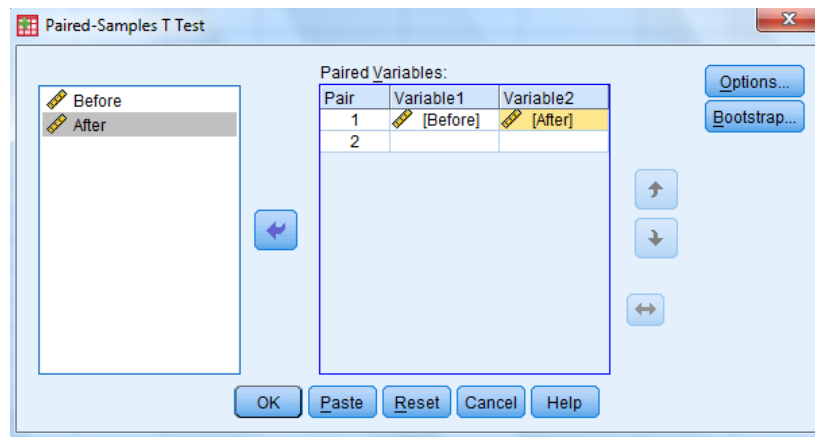
Η υπόθεση απορρίπτεται όταν $t > t_{N-1}$.

Συνεπώς, εάν η υπόθεση της ισότητας δεν απορριφθεί τότε θα συνεπάγεται ότι οι διαφορές είναι κατά μέσο όρο ίσες με το μηδέν, δηλ. ότι το γεγονός δεν επηρέασε την αρχική κατάσταση.

Στο πεδίο *Paired Variables* στις θέσεις *Variable1* & *Variable2* εισάγουμε τις δύο μεταβλητές που αντιπροσωπεύουν τις μετρήσεις πριν και μετά το γεγονός.

Προσοχή! Διαπιστώνουμε ότι η μορφή των δεδομένων για να πραγματοποιηθεί η ανάλυση είναι *Wide Format*. Εάν τα δεδομένα μας δεν είναι σε αυτή την μορφή, τότε βλ. ενότητα *Restructure* για τον μετασχηματισμό.

Από το πλήκτρο *Options* ορίζουμε το επίπεδο σημαντικότητας για τη δημιουργία των διαστημάτων εμπιστοσύνης.



Εικόνα 6.13: Το menu Paired-Samples T-test

Η διαδικασία ολοκληρώνεται με το πλήκτρο *OK* όπου και λαμβάνουμε δύο πίνακες.

Στον πρώτο πίνακα παρουσιάζονται τα περιγραφικά μέτρα στις δύο ομάδες. Απ' αυτά φαίνεται ότι υπάρχει μια μείωση της μέσης τιμής πριν και μετά.

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Before	65,780	15	5,1988	1,3423
	After	64,207	15	5,2380	1,3524

Εικόνα 6.14: Περιγραφικά μέτρα των δύο ομάδων

		Paired Differences				t	df	Sig. (2-tailed)	
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
Pair 1	Before - After							Lower	Upper
		1,5733	,4301	,1110	1,3352	1,8115	14,169	14	,000

Εικόνα 6.15: Αποτελέσματα του T-test για εξαρτημένα δείγματα

Ο δεύτερος πίνακας παρουσιάζει τα αποτελέσματα του ελέγχου από όπου διαπιστώνουμε ότι η διαφορά είναι σε επίπεδο 5% ($p\text{-value} < 0,001$), συνεπώς η διαφορά (μείωση) είναι στατιστικά σημαντική και το γεγονός (δίαιτα) συνετέλεσε στη μείωση της μέσης τιμής (σε μείωση του βάρους).

6.5 Wilcoxon

Ο έλεγχος αυτός είναι ο μη παραμετρικός ισοδύναμος του T-test για εξαρτημένα δείγματα. Η δυνατότητα προσφέρεται από την επιλογή *Analyze > Nonparametric Tests > Related Samples*.

Το κριτήριο Wilcoxon είναι ένας βαθμολογικός έλεγχος που εξετάζει την υπόθεση ότι το άθροισμα των βαθμών των αρνητικών διαφορών είναι ίσο με το άθροισμα των βαθμών των θετικών διαφορών που προκύπτουν από τη διαφορά των τιμών των δύο δειγμάτων.

Η ελεγχουσυνάρτηση βασίζεται στη σχέση:

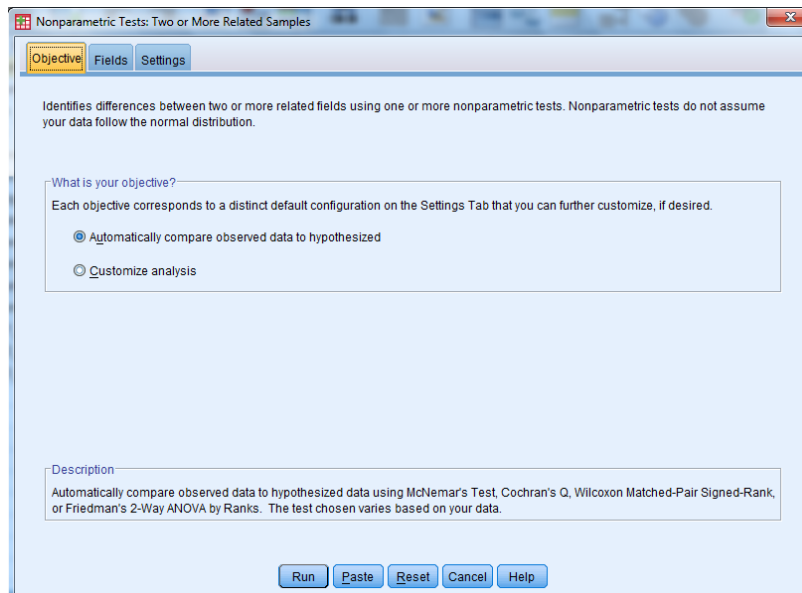
$$T = \min(T^+, T^-)$$

Εξίσωση 6.16

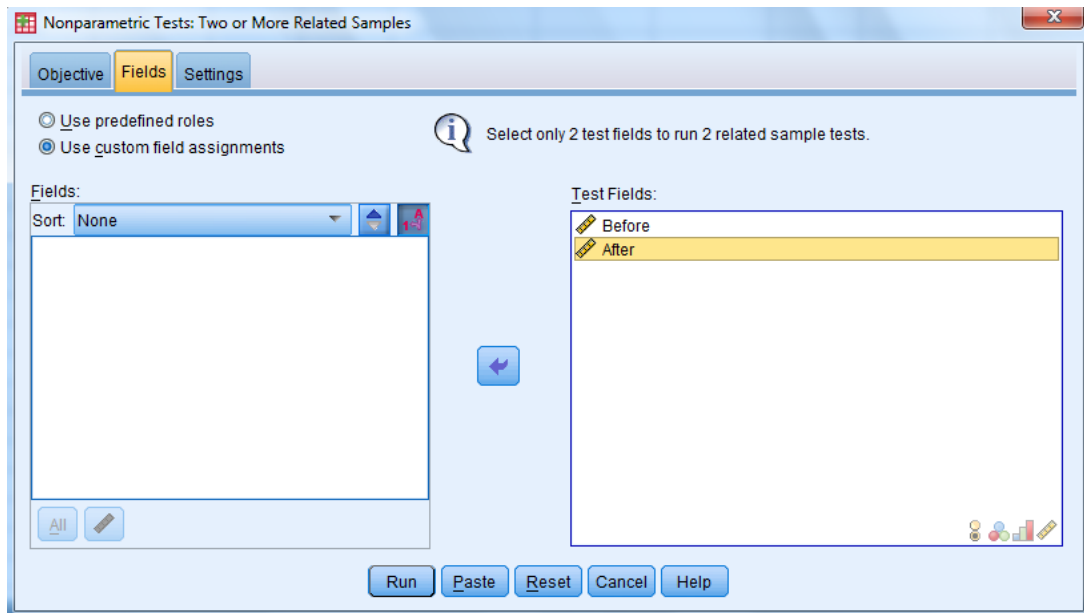
Όπου:

T^+ : το άθροισμα των βαθμών των θετικών διαφορών και

T^- : το άθροισμα των βαθμών των αρνητικών διαφορών μεταξύ των δύο ομάδων (πριν και μετά).



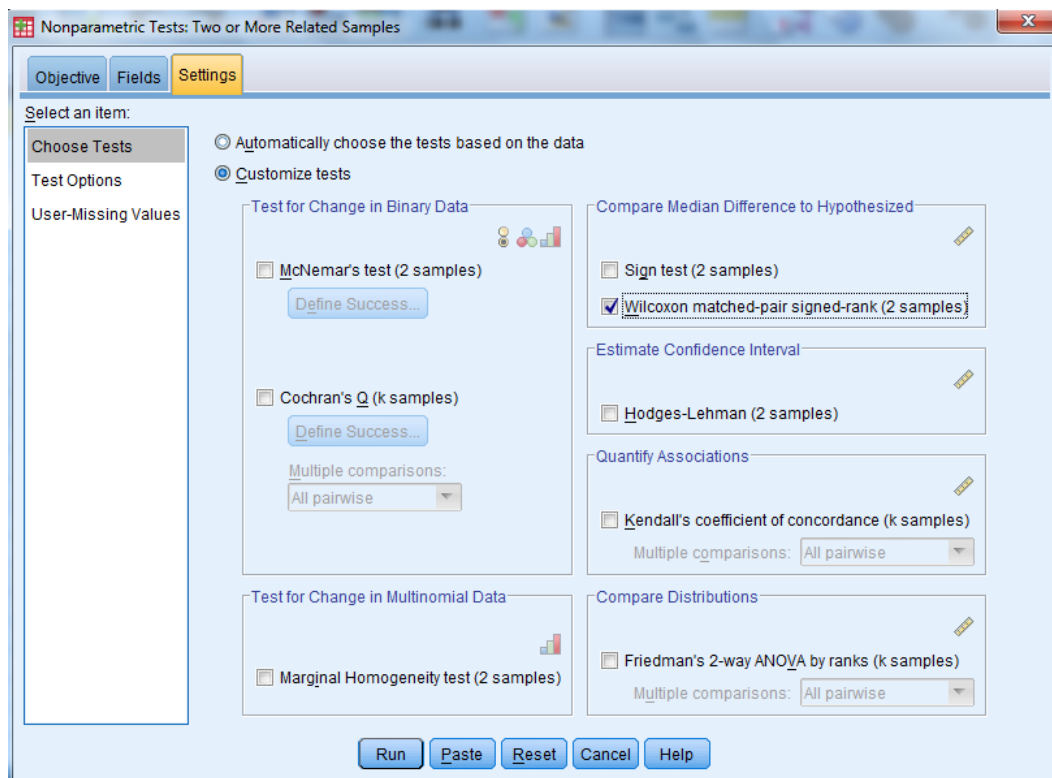
Εικόνα 6.16: Το menu *Nonparametric Tests: Two or More Related-Samples*



Εικόνα 6.17: Το menu *Nonparametric Tests: Two or More Independent-Samples > Fields*

Στο πεδίο *Test Fields* ορίζουμε τις δύο μεταβλητές που αντιπροσωπεύουν τις μετρήσεις πριν και μετά το γεγονός.

Στη συνέχεια στην ενότητα *Settings* επιλέγουμε το κριτήριο σύγκρισης των δύο ομάδων *Wilcoxon matched-pair signed-rank (2 samples)*.



Εικόνα 6.18: Το menu *Nonparametric Tests: Two or More Independent-Samples > Settings*

Η διαδικασία ολοκληρώνεται με το πλήκτρο *RUN* όπου και λαμβάνουμε:

Ranks				
		N	Mean Rank	Sum of Ranks
After - Before	Negative Ranks	15 ^a	8.00	120.00
	Positive Ranks	0 ^b	.00	.00
	Ties	0 ^c		
	Total	15		

a. After < Before

b. After > Before

c. After = Before

Εικόνα 6.19: Περιγραφικά μέτρα των βαθμών των θέσεων των παρατηρήσεων

Test Statistics ^b	
	After - Before
Z	-3.410 ^a
Asymp. Sig. (2-tailed)	.001

a. Based on positive ranks.

b. Wilcoxon Signed Ranks Test

Εικόνα 6.20: Αποτελέσματα του ελέγχου Wilcoxon

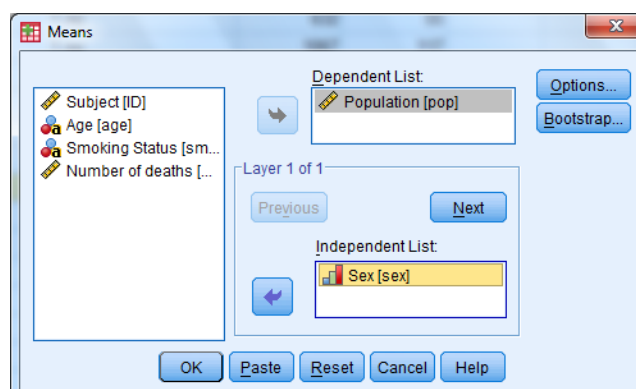
Το SPSS μας δίνει το μέσο της ιεράρχησης ή το μέσο της κατάταξης (Mean Rank) και όχι περιγραφικά μέτρα για τις δύο μεταβλητές ή για τις μετρήσεις της ίδιας μεταβλητής. Στις περιπτώσεις που η κανονικότητα παραβιάζεται συνηθίζουμε να δίνουμε τη διάμεσο ή το εύρος. Καθώς αυτά τα περιγραφικά δεν παρέχονται αυτόματα από το SPSS τα εξαγωγή από την εντολή Means, βλ. ομόνομη ενότητα.

Ο δεύτερος πίνακας παρουσιάζει τα αποτελέσματα του ελέγχου απ' όπου διαπιστώνουμε ότι σε επίπεδο 5% ($p\text{-value} < 0,001$) απορρίπτεται η H_0 , συνεπώς η διαφορά (μείωση) είναι στατιστικά σημαντική και το γεγονός (δίαιτα) συνετέλεσε στη μείωση της μέσης τιμής (σε μείωση του βάρους).

6.6 Means

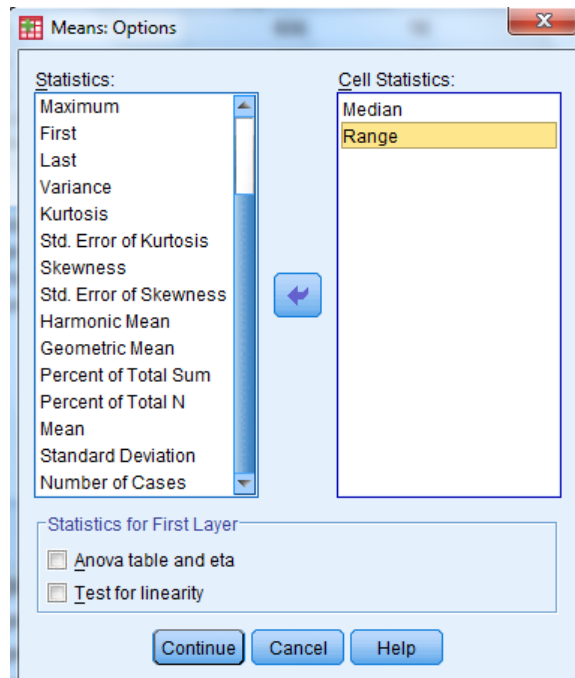
Η δυνατότητα προσφέρεται από την επιλογή *Analyze > Compare Means > Means*, με σκοπό την εξαγωγή περιγραφικών μέτρων σε ομάδες δεδομένων.

Στο πεδίο *Dependent List* εισάγουμε τη συνεχή μεταβλητή για την οποία θα εξαχθούν τα περιγραφικά μέτρα και στο πεδίο *Independent List* την κατηγορική μεταβλητή ομαδοποίησης των δεδομένων.



Εικόνα 6.21: Το menu Means

Από το πλήκτρο *Options* επιλέγουμε τα στατιστικά που επιθυμούμε να εξαχθούν εισάγοντάς τα στο πεδίο *Cell Statistics*.



Εικόνα 6.22: Το menu Means > Options

Η διαδικασία ολοκληρώνεται με το πλήκτρο *OK* όπου και λαμβάνουμε:

Report

Population

Sex	Median	Range	Mean
Male	1195,00	5948	1874,00
Female	667,00	4584	1206,82
Total	858,50	5954	1558,94

Εικόνα 6.23: Αποτελέσματα εντολής Means

Ο πίνακας παρουσιάζει τα περιγραφικά μέτρα τόσο για την κάθε ομάδα, όσο και για το σύνολο του δείγματος.

6.7 One-Way ANOVA για ανεξάρτητα δείγματα

Η δυνατότητα προσφέρεται από την επιλογή *Analyze > Compare Means > One- Way ANOVA*, με σκοπό τον έλεγχο συσχέτισης μεταξύ μιας συνεχούς μεταβλητής με μια κατηγορική η οποία έχει περισσότερες από δύο κατηγορίες. Ανάλογα με τις κατηγορίες- «επίπεδα» της κατηγορικής μεταβλητής η μέθοδος προσπαθεί να εξηγήσει τη μεταβλητότητα που οφείλεται στα διαφορετικά επίπεδα της συνεχούς μεταβλητής. Ακόμα ερευνάται η μεταβλητότητα που υπάρχει εντός της κάθε κατηγορίας «επίπεδο» αποδίδοντας σε αυτήν το ανάλογο «ποσοστό».

Υποθέσεις

Για την εφαρμογή της μεθόδου υποθέτουμε ότι:

- Οι πληθυσμοί / επίπεδα από τους οποίους προέρχονται οι μετρήσεις έχουν την ίδια διασπορά:

$$\sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2$$

Εξίσωση 6.17

- Οι πληθυσμοί κατανέμονται κανονικά.
- Οι μετρήσεις σε κάθε δείγμα (επίπεδο) είναι ανεξάρτητες και κατανέμονται κανονικά.

Εφόσον ισχύουν οι παραπάνω υποθέσεις, αναρωτιόμαστε γιατί να μην γίνει εφαρμογή του ελέγχου T-test λαμβάνοντας ανά δύο τα ζεύγη των επιπέδων της κατηγορικής μεταβλητής, ώστε να συγκριθούν οι ομάδες που διαφέρουν μεταξύ τους. Αυτό δεν μπορεί να εφαρμοστεί, γιατί θα έπρεπε να εφαρμοστούν:

$$\binom{k}{2}$$

Εξίσωση 6.18

Έλεγχοι για τα k ανεξάρτητα δείγματα, γιατί εάν το κάθε T-test απαιτούσε ένα επίπεδο στατιστικής σημαντικότητας α (σφάλμα Τύπου I), τότε το επίπεδο στατιστικής σημαντικότητας για την συνολική ανάλυση θα ήταν $k\alpha$. Αυτό οδηγεί στο ότι το επίπεδο σημαντικότητας της κάθε μηδενικής υπόθεσης θα έπρεπε να περιοριστεί σε: $\alpha = 5\% / k$.

Με αυτόν τον τρόπο η περιοχή απόρριψης της μηδενικής υπόθεσης περιορίζεται πολύ. Ως εκ τούτου αντί των πολλαπλών T-test εφαρμόζουμε ένα F-test.

Ανάλυση της μεθόδου

Η συνολική μεταβλητότητα των δεδομένων που εκφράζεται μέσω τετραγωνικών αποκλίσεων δίνεται από τη σχέση:

$$SST = SS_{bg} + SS_{wg}$$

Εξίσωση 6.19

Όπου:

$$SST = \sum_{i=1}^k \sum_{j=1}^{n_i} (X_{ij} - \bar{X})^2$$

Εξίσωση 6.20

Η διακύμανση όλων των παρατηρήσεων γύρω από τον κοινό δειγματικό μέσο.

$$\bar{X} = \frac{1}{N} \sum_{i=1}^k \sum_{j=1}^{n_i} X_{ij}$$

Εξίσωση 6.21

Ο ολικός δειγματικός μέσος όπου:

$$N = n_1 + \dots + n_k$$

Εξίσωση 6.22

$$SS_{wg} = \sum_{i=1}^k \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2$$

Εξίσωση 6.23

Η διακύμανση των παρατηρήσεων γύρω από τον μέσο κάθε δείγματος, δηλ. «μεταβλητότητα» μέσα στο κάθε δείγμα (within groups).

$$SS_{bg} = \sum_{i=1}^k \sum_{j=1}^{n_i} (\bar{X}_i - \bar{X})^2$$

Εξίσωση 6.24

Η διακύμανση των δειγματικών μέσων γύρω από τον ολικό μέσο, δηλαδή «μεταβλητότητα» μεταξύ των δειγμάτων (between groups).

Η ανάλυση διακύμανσης ερευνά αν οι διαφορές ανάμεσα στους μέσους των διαφόρων «επιπέδων» της τυχαίας μεταβλητής X οφείλονται σε «τυχαίους» παράγοντες:

$$SS_{wg} = X_{ij} - \bar{X}_i$$

Εξίσωση 6.25: Μέτρο μεταβλητότητας στο κάθε δείγμα

ή σε συστηματική διαφορά λόγω των διαφόρων επιπέδων της τυχαίας μεταβλητής X :

$$SS_{bg} = \bar{X}_i - \bar{X}$$

Εξίσωση 6.26: Μέτρο μεταβλητότητας μεταξύ των δειγμάτων

ANOVA Table

Source of variability	Sum of Squares SS	Degrees of freedom $d.f.$	Mean of SS MS	Ratio F
Between Groups	SS_{bg}	$k - 1$	$MSS_{bg} = \frac{SS_{bg}}{k-1}$	$F = \frac{MSS_{bg}}{MSS_{wg}}$
Within Groups	SS_{wg}	$N - k$	$MSS_{wg} = \frac{SS_{wg}}{N-k}$	
Total variability	SST	$N - 1$		

Εικόνα 6.24: One-way ANOVA Table

Η υπόθεση που εξετάζεται είναι εάν οι μέσοι όλων των ομάδων είναι ίσοι μεταξύ τους, ή υπάρχει τουλάχιστον ένας μέσος που διαφέρει.

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_k = \mu$$

$$H_1 : \exists 0 < i < j \leq k : \mu_i \neq \mu_j$$

Εξίσωση 6.27

Η ελεγχουσυνάρτηση δίνεται από τη σχέση:

$$F = \frac{MSS_{bg}}{MSS_{wg}} \sim F_{k-1, N-k}$$

Εξίσωση 6.28

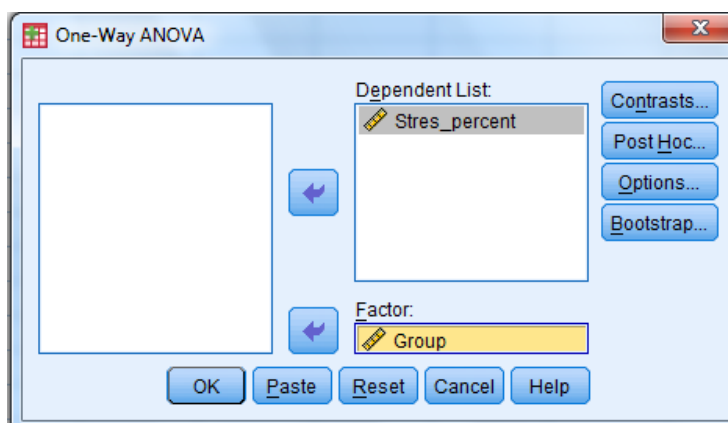
Όπου:

MSS_{bg} , η μέση μεταβλητότητα μεταξύ των ομάδων, και MSS_{wg} , η μέση μεταβλητότητα εντός των ομάδων.

Η υπόθεση απορρίπτεται όταν $F > F_{k-1, k; \alpha}$. Τα αποτελέσματα του ελέγχου παρουσιάζονται στην Εικόνα 6.24.

Έστω ότι έχουμε έναν πληθυσμό από 20 ποντίκια σε καθεστώς άγχους χωρισμένα σε 4 ισοπληθείς ομάδες, στα οποία χορηγείται ενδοφλεβίως ένα φάρμακο σε διαφορετικές δόσεις. Εξετάζουμε εάν η δοσολογία του φαρμάκου επιδρά στη μείωση των ποσοστών ταραχής.

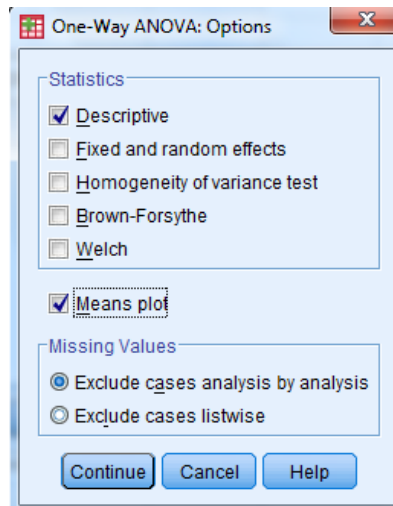
Στο πεδίο *Dependent List* εισάγουμε την/ις υπό εξέταση συνεχή/είς μεταβλητή/ές (π.χ. τη μεταβλητή «ποσοστό άγχους») και στο πεδίο *Factor* την κατηγορική μεταβλητή. Στο παράδειγμά μας είναι η μεταβλητή που καθορίζει σε ποια ομάδα δοσολογίας ανήκει το κάθε ποντίκι.



Εικόνα 6.25: Το menu One-way ANOVA

Από το πλήκτρο *Options* ορίζουμε την εμφάνιση πρόσθετων πληροφοριών και ελέγχων, όπως:

- *Descriptive*, για την εμφάνιση περιγραφικών μέτρων (μέση τιμή, τυπική απόκλιση, τυπικό σφάλμα, 95% Δ.Ε. του μέσου, ελάχιστη και μέγιστη παρατήρηση) της κάθε κατηγορίας.
- *Fixed and Random Effects*, για την εκτίμηση της επίδρασης της κατηγορικής μεταβλητής στη συνεχή μεταβλητή μέσω ενός υποδείγματος σταθερών επιδράσεων και μέσω ενός υποδείγματος τυχαίων επιδράσεων.
- *Homogeneity of variance test*, για τον έλεγχο της ισότητας των διασπορών μεταξύ των ομάδων. Ο έλεγχος που υπολογίζεται είναι αυτός του Levene. Εάν ο έλεγχος απορριφθεί τότε η χρήση της μεθόδου δεν συνιστάται δεδομένου ότι παραβιάζεται η υπόθεση εφαρμογής της μεθόδου. Σε αυτή την περίπτωση θα πρέπει να γίνει χρήση της μη παραμετρικής μεθόδου Kruskal-Wallis, ή των ελέγχων Brown-Forsythe και Welch.
- *Means Plot*, για την εμφάνιση διαγράμματος των μέσων της κάθε κατηγορίας.
- *Missing Values*, για τον χειρισμό των ελλειπουσών τιμών.

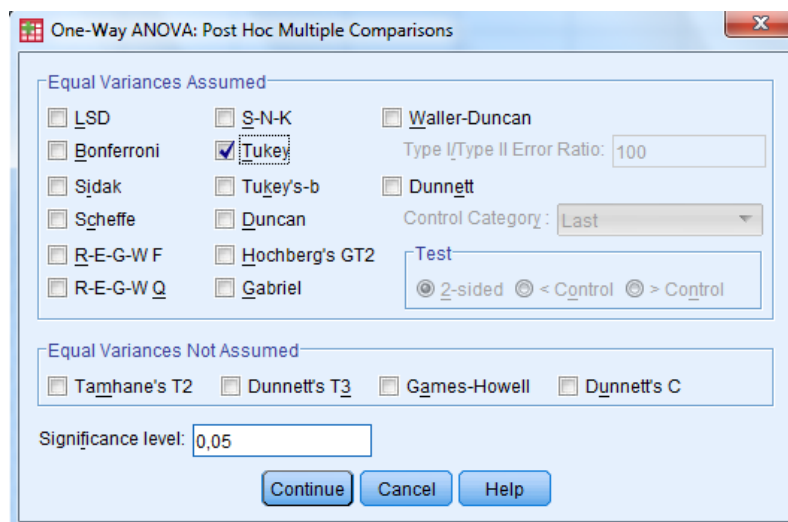


Εικόνα 6.26: Το menu One-way ANOVA > Options

Σε περίπτωση που ο έλεγχος One-way ANOVA απορριφθεί, δεν είναι σε θέση να μας δώσει ποιος ήταν ο μέσος που διέφερε σε σχέση με τους υπόλοιπους. Για να προσδιοριστούν οι μέσοι που διαφέρουν μεταξύ τους, θα πρέπει να γίνουν πολλαπλές συγκρίσεις μεταξύ των ομάδων (ανά δύο). Επειδή όπως προαναφέραμε θα πρέπει να ληφθεί υπόψη η μείωση στο επίπεδο σημαντικότητας λόγω των πολλαπλών συγκρίσεων, έχουν διαμορφωθεί ορισμένες μέθοδοι, όπως:

- Έλεγχοι που βασίζονται σε προσεγγίσεις pairwise multiple comparisons tests ή range tests (Tukey, Hochberg's GT2, Gabriel, Scheffe),
- Έλεγχοι που κάνουν μόνο πολλαπλές συγκρίσεις των μέσων τιμών ανά δύο (Bonferroni, Sidak, Dunnett C, LSD-Least Significant Difference).

Η πιο ευρέως διαδεδομένη είναι ο έλεγχος του Tukey.



Εικόνα 6.27: Οι δυνατές επιλογές για Post-Hoc Multiple Comparisons

Tukey Test

Είναι μια διαδικασία πολλαπλών συγκρίσεων που χρησιμοποιείται κυρίως όταν τα μεγέθη των k δειγμάτων είναι ίσα, και εξετάζει την υπόθεση:

$$H_0 : \mu_L = \mu_S$$

$$H_1 : \mu_L \neq \mu_S$$

Εξίσωση 6.29

Ο έλεγχος βασίζεται στον υπολογισμό της στατιστικής συνάρτησης Q η οποία ορίζεται ως:

$$Q = \frac{\bar{X}_L - \bar{X}_S}{\sqrt{\frac{MSS_{wg}}{N_{p/s}}}}$$

Εξίσωση 6.30

Όπου:

$$\bar{X}_L$$

Ο μέγιστος δειγματικός μέσος

$$\bar{X}_S$$

Ο ελάχιστος δειγματικός μέσος

$$N_{p/s}$$

Το πλήθος των μετρήσεων σε κάθε δείγμα.

Όλοι οι μέσοι συγκρίνονται ανά δύο και απορρίπτουμε τη μηδενική υπόθεση όταν $Q > Q_\alpha$, όπου η στατιστική συνάρτηση $Q \sim$ κατανομή (k, df_{wg}) .

Σημείωση

Αν τα δείγματα έχουν διαφορετικό μέγεθος, τότε το $N_{p/s}$ υπολογίζεται από τον αρμονικό μέσο των μεγεθών των δειγμάτων:

$$N_{p/s} = \frac{k}{\frac{1}{n_1} + \dots + \frac{1}{n_k}}$$

Εξίσωση 6.31

Η διαδικασία ολοκληρώνεται με το πλήκτρο *OK* όπου και λαμβάνουμε:

Descriptives

Stres_percent

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
1	5	28,8600	2,83161	1,26633	25,3441	32,3759	26,20	33,50
2	5	25,0400	2,42343	1,08379	22,0309	28,0491	22,80	27,70
3	5	22,5000	3,36378	1,50433	18,3233	26,6767	19,30	27,80
4	5	22,3000	1,96850	,88034	19,8558	24,7442	19,60	23,90
Total	20	24,6750	3,67364	,82145	22,9557	26,3943	19,30	33,50

Εικόνα 6.28: Περιγραφικά μέτρα των ομάδων

ANOVA

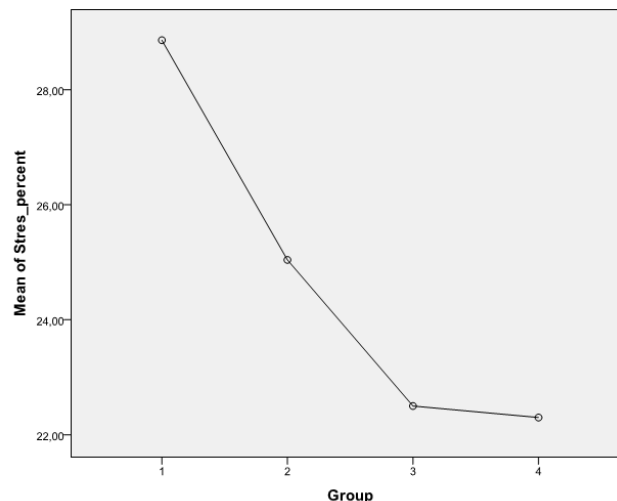
Stres_percent

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	140,094	3	46,698	6,423	,005
Within Groups	116,324	16	7,270		
Total	256,418	19			

Εικόνα 6.29: ANOVA Table με τα αποτελέσματα του F-test

Στα αποτελέσματα λαμβάνουμε τον πίνακα με τα περιγραφικά μέτρα των μέσων των ομάδων, και τον πίνακα ANOVA με τα αποτελέσματα του F-test, από τον οποίο διαπιστώνουμε ότι η μηδενική υπόθεση απορρίπτεται ($p\text{-value} = 0,005$), συνεπώς υπάρχει τουλάχιστον ένας μέσος που διαφέρει σημαντικά από τους υπόλοιπους. Αναφορικά με το παράδειγμά μας μπορεί να θεωρηθεί ότι υπάρχουν σημαντικές διαφορές στους μέσους των τεσσάρων ομάδων, επομένως η αύξηση της δοσολογίας που χορηγήθηκε στα ποντίκια μειώνει τα επίπεδα στρες.

Για το λόγο αυτό θα πρέπει να ληφθεί υπόψη και ο πίνακας πολλαπλών συγκρίσεων μέσω κάποιας από τις μεθόδους της ενότητας Post Hoc Multiple Comparisons.



Εικόνα 6.30: Γραφική απεικόνιση των μέσων τιμών των ομάδων

Multiple Comparisons

		Stres_percent Tukey HSD				
(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1	2	3,82000	1,70532	,155	-1,0589	8,6989
	3	6,36000*	1,70532	,009	1,4811	11,2389
	4	6,56000*	1,70532	,007	1,6811	11,4389
2	1	-3,82000	1,70532	,155	-8,6989	1,0589
	3	2,54000	1,70532	,466	-2,3389	7,4189
	4	2,74000	1,70532	,403	-2,1389	7,6189
3	1	-6,36000*	1,70532	,009	-11,2389	-1,4811
	2	-2,54000	1,70532	,466	-7,4189	2,3389
	4	,20000	1,70532	,999	-4,6789	5,0789
4	1	-6,56000*	1,70532	,007	-11,4389	-1,6811
	2	-2,74000	1,70532	,403	-7,6189	2,1389
	3	-,20000	1,70532	,999	-5,0789	4,6789

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

Εικόνα 6.31: Πίνακας Πολλαπλών συγκρίσεων μέσω της μεθόδου Tukey

Ο πίνακας Multiple Comparisons παρουσιάζει τις διαφορές των μέσων τιμών μεταξύ των ομάδων ανά δύο, σημειώνοντας με αστερίσκο τις στατιστικά σημαντικές διαφορές.

Από την εφαρμογή του Tukey test έχουμε:

$$\sqrt{\frac{MSS_{wg}}{N_{p/s}}} = \sqrt{\frac{7,27}{5}} = 1.2058$$

Εξίσωση 6.32

$Q_{\alpha=0.05} = 4,05$ και $Q_{\alpha=0.01} = 5,2$ ποσοστιαία σημεία της Q .

Οπότε η διαφορά των μέσων των ομάδων 1 & 2:

$$\bar{X}_L - \bar{X}_S < Q_{\alpha} \sqrt{\frac{MSS_{wg}}{N_{p/s}}} \Rightarrow 3.82 < 4.88$$

Εξίσωση 6.33

Άρα η H_0 δεν απορρίπτεται.

Η διαφορά των μέσων των ομάδων 1 & 3:

$$\bar{X}_L - \bar{X}_S < Q_{\alpha} \sqrt{\frac{MSS_{wg}}{N_{p/s}}} \Rightarrow 6.36 > 4.88$$

Εξίσωση 6.34

Άρα η H_0 απορρίπτεται κ.ο.κ. Επιπρόσθετα στα αποτελέσματα εμφανίζεται ο ακόλουθος πίνακας Homogeneous Subsets στον οποίο παρουσιάζονται τα ομοιογενή υποσύνολα που δύναται να δημιουργηθούν από τις υφιστάμενες ομάδες. Στο παράδειγμά μας διαπιστώνουμε ότι μπορούν να δημιουργηθούν δύο

υποσύνολα στα οποία είτε το πρώτο θα περιέχει τις ομάδες 4,3 και 2 και το δεύτερο την ομάδα 1 μόνη της, είτε το πρώτο θα περιέχει τις ομάδες 4 και 3 και το δεύτερο τις ομάδες 2 και 1.

Homogeneous Subsets

Stres_percent

Tukey HSD^a

Group	N	Subset for alpha = 0.05	
		1	2
4	5	22,3000	
3	5	22,5000	
2	5	25,0400	25,0400
1	5		28,8600
Sig.		,403	,155

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 5,000.

Εικόνα 6.32: Έλεγχος ομοιογένειας των ομάδων με σκοπό την ομαδοποίηση τους

6.8 Kruskal-Wallis H

Η δυνατότητα προσφέρεται από την επιλογή *Analyze > Non parametric Tests > Legacy Dialogs > k Independent Samples*, με σκοπό τον έλεγχο συσχέτισης μεταξύ μιας συνεχούς μεταβλητής με μια κατηγορική, η οποία έχει περισσότερες από δύο κατηγορίες με ίσο ή διαφορετικό πλήθος ατόμων ανά κατηγορία. Η μέθοδος χρησιμοποιείται όταν δεν ισχύουν οι υποθέσεις εφαρμογής της One-Way ANOVA, δηλ. όταν οι κατανομές δεν είναι κανονικές και οι διασπορές δεν είναι ίσες, ή όταν η μεταβλητή είναι διατεταγμένη.

Η μέθοδος βασίζεται στην ίδια λογική των ελέγχων εξετάζοντας αν ένα σύνολο k ανεξάρτητων ομάδων προέρχονται από τον ίδιο πληθυσμό. Σε αντίθεση με τον έλεγχο One-way ANOVA που εξετάζει την ισότητα των μέσων, εδώ χρησιμοποιείται το άθροισμα των βαθμών των θέσεων των παρατηρήσεων. Η μέθοδος είναι μια επέκταση του κριτηρίου Mann-Whitney U για k ανεξάρτητες ομάδες. Κατ' επέκταση η υπόθεση που εξετάζεται είναι αν οι ομάδες έχουν ίσο μέσο βαθμών (δηλ. τυχαία διάταξη) ή κάποια ομάδα διαφοροποιεί τη διάταξη των παρατηρήσεων.

Για την εφαρμογή της μεθόδου οι παρατηρήσεις διατάσσονται κατά αύξουσα σειρά και σε κάθε παρατήρηση προσδιορίζεται η σχετική της θέση (rank). Στη συνέχεια υπολογίζονται οι ποσότητες:

$$SS_k = \sum_{j=1}^k \frac{R_j^2}{n_j}$$

Εξίσωση 6.35

Όπου:

$$R_j = \sum_{i=1}^{n_j} r_{ij}$$

Εξίσωση 6.36

Το άθροισμα των θέσεων της κάθε j-ομάδας, το οποίο είναι το αντίστοιχο «άθροισμα τετραγώνων των θέσεων μεταξύ των ομάδων»:

$$SS_r = \sum_{j=1}^k \sum_{i=1}^{n_j} r_{ij}^2$$

Εξίσωση 6.37

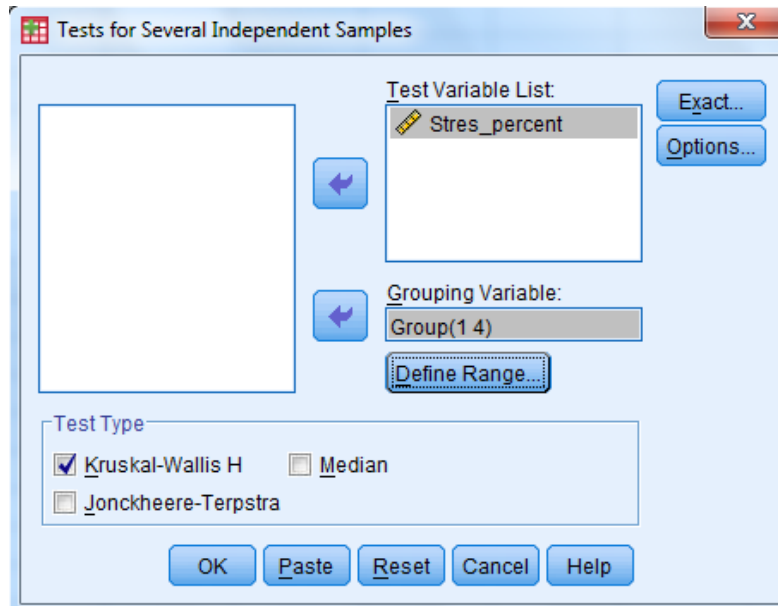
Το «συνολικό άθροισμα τετραγώνων των θέσεων των παρατηρήσεων όλων των ομάδων».

Η ελεγχουσυνάρτηση προκύπτει από τη σχέση:

$$H = \frac{12}{N(N+1)} \sum_{j=1}^k \frac{R_j^2}{n_j} - 3(N+1) \sim X_{k-1}^2$$

Εξίσωση 6.38

Η υπόθεση απορρίπτεται όταν: $H > X_{k-1;\alpha}^2$



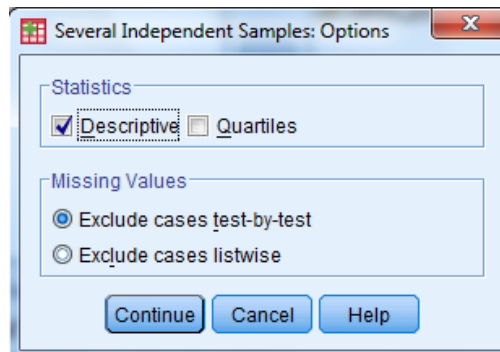
Εικόνα 6.33: Το menu εκτέλεσης Kruskal-Wallis H

Στο πεδίο *Test Variable List* εισάγουμε την/ις υπό εξέταση συνεχή/είς μεταβλητή/ές (π.χ. την μεταβλητή «ποσοστό άγχους») και στο πεδίο *Grouping Variable* την κατηγορική μεταβλητή. Στη συνέχεια από το πλήκτρο *Define Range* ορίζουμε το εύρος των κωδικών των επιπέδων της κατηγορικής μεταβλητής (ελάχιστο και μέγιστο κωδικό) που θα χρησιμοποιηθούν στην ανάλυσή μας.

Σ' αυτή την ενότητα παρέχονται επιπρόσθετα οι έλεγχοι Median και Jonckheere-Terpstra:

- *Median Test*, η οποία είναι μια μη παραμετρική διαδικασία που βασίζεται στον έλεγχο της συνάφειας μεταξύ των ομάδων και του πλήθους των παρατηρήσεων που είναι μικρότερες ή μεγαλύτερες από τη διάμεσο.
- *Jonckheere-Terpstra* που χρησιμοποιείται αντί του Kruskal-Wallis όταν μεταξύ των επιπέδων των ομάδων υπάρχει σχετική διάταξη, ώστε να εξεταστεί το ενδεχόμενο της ύπαρξης γραμμικής σχέσης, μεταξύ της διάταξης και της εξαρτημένης μεταβλητής.

Από το πλήκτρο *Options* επιλέγουμε την εμφάνιση των περιγραφικών μέτρων (Descriptive) και των τεταρτημορίων (Quartiles) για την εξαρτημένη μεταβλητή.



Εικόνα 6.34: Το menu εκτέλεσης Kruska -Wallis H > Options

Η διαδικασία ολοκληρώνεται με το πλήκτρο *OK* όπου και λαμβάνουμε:

Descriptive Statistics								
	N	Mean	Std. Deviation	Minimum	Maximum	Percentiles		
						25th	50th (Median)	75th
Stres_percent	20	24,6750	3,67364	19,30	33,50	22,1250	23,8000	27,6750
Group	20	2,50	1,147	1	4	1,25	2,50	3,75

Εικόνα 6.35: Περιγραφικά μέτρα των ομάδων

Ranks			
	Group	N	Mean Rank
Stres_percent	1	5	16,80
	2	5	11,20
	3	5	6,80
	4	5	7,20
	Total	20	

Εικόνα 6.36: Οι μέσοι των βαθμών των ομάδων. Ο έλεγχος εξετάζει την ισότητά τους

Test Statistics ^{a,b}	
	Stres_percent
Chi-square	9,258
df	3
Asymp. Sig.	,026

a. Kruskal Wallis Test

b. Grouping Variable:
Group

Εικόνα 6.37: Αποτελέσματα του ελέγχου Kruskal-Wallis

Από τα αποτελέσματα του ελέγχου Kruskal-Wallis διαπιστώνουμε ότι η υπόθεση απορρίπτεται ($p\text{-value} = 0,026$).

Σε περίπτωση που ο έλεγχος Kruskal-Wallis απορριφθεί, δεν είναι σε θέση να μας δώσει ποιος ήταν ο μέσος που διέφερε σε σχέση με τους υπόλοιπους. Για να προσδιοριστούν οι μέσοι που διαφέρουν μεταξύ τους, θα πρέπει να γίνουν πολλαπλές συγκρίσεις μεταξύ των ομάδων (ανά δύο).

Σε αυτή την περίπτωση ο ερευνητής μπορεί χειροκίνητα να προβεί σε επί μέρους εφαρμογή του Mann-Whitney U test για όλους τους ανά δύο πιθανούς συνδυασμούς, διορθώνοντας το επίπεδο σημαντικότητας διαιρώντας το με το πλήθος των συγκρίσεων (μέθοδος Bonferroni).

6.9 Two-Way ANOVA for Independent Samples

Η δυνατότητα προσφέρεται από την επιλογή *Analyze > General Linear Model > Univariate*, με σκοπό τον έλεγχο συσχέτισης μεταξύ μιας συνεχούς μεταβλητής με δύο κατηγορικές μεταβλητές. Σε περίπτωση που θέλουμε να εξετάσουμε τη συσχέτιση με περισσότερες από δύο κατηγορικές, τότε εφαρμόζεται μια three way ANOVA, κ.ο.κ.

Ανάλυση της μεθόδου

Στην ανάλυση διακύμανσης άνα δύο παράγοντες έχουμε δύο κατηγορικές μεταβλητές X_1 (παράγοντας A) και X_2 (παράγοντας B), και μια εξαρτημένη ποσοτική μεταβλητή Y . Έτσι, οι υποθέσεις που ελέγχουμε είναι εάν:

- Τα διαφορετικά επίπεδα τιμών της X_1 επηρεάζουν τις τιμές της Y .
- Τα διαφορετικά επίπεδα τιμών της X_2 επηρεάζουν τις τιμές της Y .
- Υπάρχει αλληλεπίδραση (interaction) μεταξύ των μεταβλητών X_1 και X_2 .

Έλλειψη αλληλεπίδρασης σημαίνει ότι η επίδραση της μεταβλητής X_1 στην Y είναι ίδια για κάθε επίπεδο της μεταβλητής X_2 και αντίστροφα.

Για να εξεταστούν οι παραπάνω υποθέσεις, θα καταλήξουμε σε τρία *F-test*.

Υποθέσεις

Έστω ότι η μεταβλητή X_1 έχει r επίπεδα και η μεταβλητή X_2 έχει c επίπεδα. Οι παρατηρήσεις συμβολίζονται ως Y_{ijk} .

Όπου:

$i = 1, \dots, r$: τα επίπεδα της μεταβλητής X_1 ,

$j = 1, \dots, c$: τα επίπεδα της μεταβλητής X_2 ,

$k = 1, \dots, N_T$: το σύνολο των μετρήσεων.

Οι παρατηρήσεις εκφράζονται από τη σχέση:

$$Y_{ijk} = \mu_{ij} + \varepsilon_{ijk}$$

Εξίσωση 6.39

Όπου:

$$\varepsilon_{ijk} \sim N(0, \sigma^2)$$

Εξίσωση 6.40

Για την εφαρμογή της ανάλυσης ως προς δύο παράγοντες θα πρέπει να ισχύουν:

- Οι πληθυσμοί από τους οποίους προέρχονται οι παρατηρήσεις να κατανέμονται κανονικά:

$$Y_{ijk} \sim N(\mu_{ij}, \sigma^2)$$

Εξίσωση 6.41

- Οι πληθυσμοί από τους οποίους προέρχονται οι παρατηρήσεις να έχουν ίσες διασπορές:

$$\sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2$$

Εξίσωση 6.42

Παράδειγμα

Θέλουμε να εξετάσουμε την επίδραση και την αλληλεπίδραση δύο τύπων ναρκωτικών Α και Β, στη φυσιολογική διέγερση. Έχουμε 40 εργαστηριακά ποντίκια που χωρίζονται σε τέσσερις ομάδες των δέκα ατόμων. Κάθε ομάδα έλαβε μια συγκεκριμένη δόσοληψία από τα ναρκωτικά Α και Β (0 μονάδες και 1 μονάδα). Η τυχαία μεταβλητή Y είναι η φυσιολογική διέγερση.

Στόχος μας είναι να κατασκευάσουμε τον ANOVA-Table οπότε εργαζόμαστε ως εξής:

1^{ον}: Υπολογίζουμε τους μέσους όρους των ομάδων.

2^{ον}: Υπολογίζουμε τα τετραγωνικά αθροίσματα για την κάθε ομάδα, όπου:

$$SS = \sum Y_i^2 - \frac{(\sum Y_i)^2}{N}$$

$$SST = SS_{bg} + SS_{wg}$$

$$SS_{wg} = SS_{g1} + SS_{g2} + SS_{g3} + SS_{g4}$$

$$SS_{bg} = SST - SS_{wg}$$

$$SS_{bg} = SS_r + SS_c + SS_{IN}$$

Εξισώσεις 6.43

Δηλαδή η μεταβλητότητα μεταξύ των ομάδων αναλύεται στη μεταβλητότητα μεταξύ των επιπέδων r , των επιπέδων c και στην αλληλεπίδραση.

3^{ον}: Υπολογίζουμε τους βαθμούς ελευθερίας:

$$df_r = r - 1$$

$$df_c = c - 1$$

$$df_{IN} = (r - 1)(c - 1)$$

$$df_T = N_T - 1$$

$$df_{error} = df_{wg} = N_T - rc$$

$$df_{bg} = rc - 1$$

$$df_T = df_{wg} + df_{bg}$$

$$df_{bg} = df_r + df_c + df_{IN}$$

Εξισώσεις 6.44

4^{ον}: Υπολογίζουμε τα:

$MS_{wg}, MS_r, MS_c, MS_{IN}$ (ή αλλιώς MS_{rc})

$$MSS_r = \frac{SS_r}{df_r}$$

$$MSS_c = \frac{SS_c}{df_c}$$

$$MSS_{IN} = \frac{SS_{IN}}{df_{IN}} \text{ και } MSS_{error} = MSS_{wg} = \frac{SS_{wg}}{df_{wg}}.$$

Εξισώσεις 6.45

5^{ον}: Υπολογίζονται τα F -tests:

- 1) F -test για τον έλεγχο της επίδρασης των επιπέδων της X_1 στις τιμές της Y . Η υπόθεση που εξετάζεται είναι:

H_0 : Δεν υπάρχει στατιστικά σημαντική διαφορά στους μέσους μεταξύ των δύο γραμμών.

H_1 : Υπάρχει στατιστικά σημαντική διαφορά στους μέσους μεταξύ των δύο γραμμών.

$$F_r = \frac{MSS_r}{MSSE}$$

Εξίσωση 6.46

Απορρίπτουμε την H_0 όταν:

$$F_r > F_{r-1, N_T - rc; \alpha}$$

Εξίσωση 6.47

- 2) F -test για τον έλεγχο της επίδρασης των επιπέδων της X_2 στις τιμές της Y . Η υπόθεση που εξετάζεται είναι:

H_0 : Δεν υπάρχει στατιστικά σημαντική διαφορά στους μέσους μεταξύ των δύο στηλών.

H_1 : Υπάρχει στατιστικά σημαντική διαφορά στους μέσους μεταξύ των δύο στηλών.

$$F_c = \frac{MSS_c}{MSSE}$$

Εξίσωση 6.48

Απορρίπτουμε την H_0 όταν:

$$F_c > F_{c-1, N_T - rc; \alpha}$$

Εξίσωση 6.49

- 3) F -test για τον έλεγχο ύπαρξης αλληλεπίδρασης των επιπέδων της X_1 στις τιμές της X_2 . Η υπόθεση που εξετάζεται είναι:

H_0 : Δεν υπάρχει αλληλεπίδραση μεταξύ γραμμών και στηλών.

H_1 : Υπάρχει αλληλεπίδραση μεταξύ γραμμών και στηλών.

$$F_{IN} = \frac{MSS_{IN}}{MSSE}$$

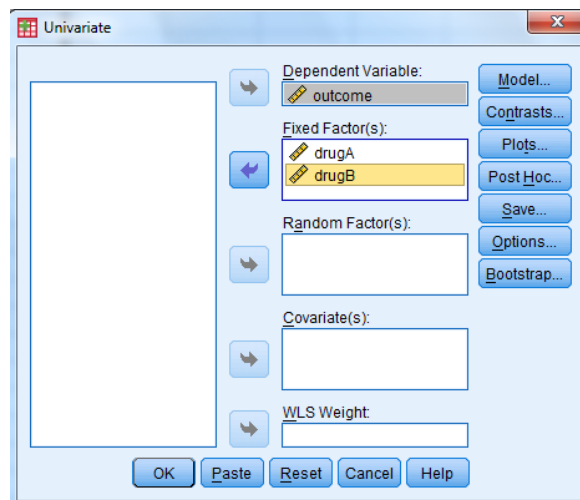
Εξίσωση 6.50

Απορρίπτουμε την H_0 όταν:

$$F_{IN} > F_{(r-1)(c-1), N_T - rc; \alpha}$$

Εξίσωση 6.51

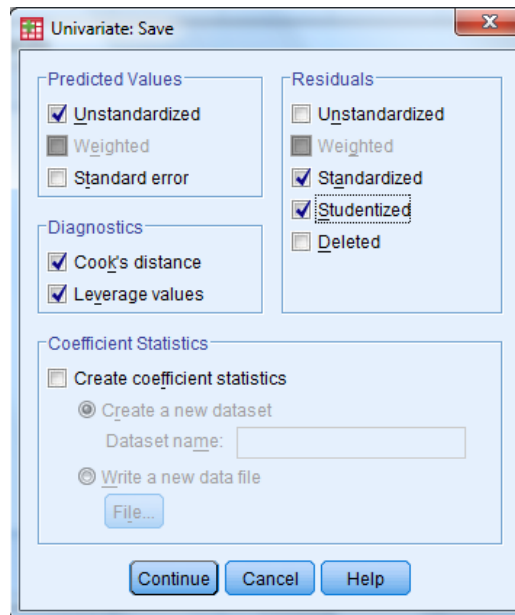
Από την επιλογή *Analyze > General Linear Model > Univariate* ορίζουμε στο πεδίο *Dependent Variable* την εξαρτημένη μεταβλητή και στο πεδίο *Fixed Factor(s)* τις κατηγορικές μεταβλητές.



Εικόνα 6.38: Το menu Univariate

Από το πλήκτρο *Save* ο χρήστης έχει τη δυνατότητα να αποθηκεύσει διάφορες τιμές που προβλέπονται από το μοντέλο, σφάλματα (residuals) ή συναφή μέτρα αξιολόγησης της ανάλυσης, τα οποία αποθηκεύονται ως νέες μεταβλητές στο φύλλο δεδομένων. Αναλυτικότερα:

- *Predicted Values*: Περιέχει επιλογές με σκοπό τον υπολογισμό τιμών που προκύπτουν από προβλέψεις του μοντέλου.
- *Unstandardized*: Η τιμή πρόβλεψης για την εξαρτημένη μεταβλητή.
- *Weighted*: Οι σταθμισμένες μη τυποποιημένες προβλεπόμενες τιμές. Η δυνατότητα προσφέρεται μόνο αν η μεταβλητή έχει προηγουμένως δηλωθεί ως σταθμισμένη.
- *Standard error*: Μια εκτίμηση της τυπικής απόκλισης του μέσου όρου της εξαρτημένης μεταβλητής για τις περιπτώσεις που έχει τις ίδιες τιμές των ανεξάρτητων μεταβλητών.
- *Diagnostics*: Μας παρέχει μέτρα για τον εντοπισμό ασυνήθιστων περιπτώσεων. Π.χ. ακραίων παρατηρήσεων ή παρατηρήσεων επίδρασης στο μοντέλο.
- *Cook's distance*: Είναι ένα μέτρο για τον εντοπισμό ακραίων παρατηρήσεων.
- *Leverage values*: Είναι ένα μέτρο για τον εντοπισμό παρατηρήσεων με μεγάλη επίδραση (μόγλευση) στο μοντέλο.



Εικόνα 6.39: Το menu Univariate > Save

Residuals: Περιέχει τις επιλογές αποθήκευσης διαφόρων μορφών σφαλμάτων του μοντέλου, όπου σφάλμα νοείται ως η διαφορά μεταξύ της προβλεπόμενης τιμής και της πραγματικής παρατήρησης.

Unstandardized: Μη τυποποιημένα κατάλοιπα.

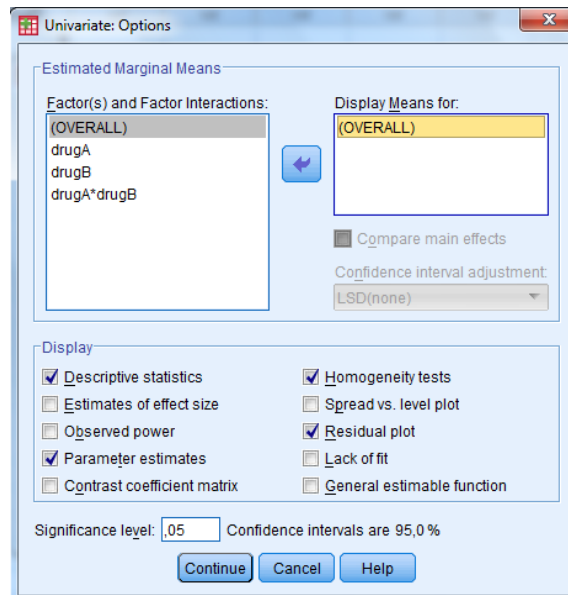
Weighted: Σταθμισμένα μη τυποποιημένα σφάλματα. Η δυνατότητα προσφέρεται μόνο αν η μεταβλητή έχει προηγουμένως δηλωθεί ως σταθμισμένη.

Standardized: Τυποποιημένα σφάλματα. Εάν ισχύει η υπόθεση της κανονικότητας, τότε τα τυποποιημένα σφάλματα έχουν μέση τιμή μηδέν και τυπική απόκλιση 1.

Studentized: Τα σφάλματα διαιρούνται με την εκτιμήτρια της τυπικής απόκλισης η οποία διαφέρει ανάλογα με την κάθε περίπτωση και εξαρτάται από την απόσταση της κάθε παρατηρούμενης τιμής των ανεξάρτητων μεταβλητών από τη μέση τιμή της ανεξάρτητης μεταβλητής.

Deleted: Τα σφάλματα που προκύπτουν όταν η συγκεκριμένη παρατήρηση εξαιρείται από τον υπολογισμό του συντελεστή παλινδρόμησης. Είναι η διαφορά μεταξύ της τιμής της εξαρτημένης μεταβλητής και της προσαρμοσμένης προβλεπόμενης τιμής.

Coefficient Statistics: Δημιουργεί έναν πίνακα διακυμάνσεων - συνδιακυμάνσεων των εκτιμώμενων παραμέτρων του μοντέλου είτε σε μια νέα βάση δεδομένων στην τρέχουσα ενότητα είτε σε νέο εξωτερικό αρχείο. Επίσης για κάθε εξαρτημένη μεταβλητή, θα υπάρξει μια σειρά εκτιμώμενων παραμέτρων, μια σειρά στατιστικά σημαντικών τιμών από ελέγχους T-test των εκτιμήσεων των συντελεστών του μοντέλου και μια σειρά από τους βαθμούς ελευθερίας των σφαλμάτων. Για ένα πολυπαραγοντικό μοντέλο υπάρχουν παρόμοιες σειρές για κάθε εξαρτημένη μεταβλητή. Αυτό το αρχείο δύναται να χρησιμοποιηθεί από άλλες διαδικασίες που έχουν τη δυνατότητα να διαβάσουν αρχεία πινάκων (μήτρας).



Εικόνα 6.40: Το menu Univariate > Options

Από το πλήκτρο *Options* λαμβάνουμε (προαιρετικά) διάφορα στατιστικά στοιχεία τα οποία υπολογίζονται βάσει ενός μοντέλου σταθερών επιδράσεων (fixed effects). Αναλυτικά:

- *Estimated Marginal Means*: Σ' αυτή την ενότητα επιλέγουμε τους παράγοντες και τις αλληλεπιδράσεις για τους οποίους θα υπολογιστούν οι εκτιμώμενοι μέσοι των πληθυσμών των κατηγοριών των παραγόντων. Οι μέσοι αυτοί είναι προσαρμοσμένοι σε σχέση με τις συμμεταβλητές, αν υπάρχουν. Οι αλληλεπιδράσεις είναι διαθέσιμες μόνο αν έχουν οριστεί αλληλεπιδράσεις στο προσαρμοσμένο (custom) μοντέλο.
- *Compare main effects*: Παρέχει μη διορθωμένες κατά ζεύγη συγκρίσεις μεταξύ των εκτιμώμενων μέσων για κάθε κύρια επίδραση του μοντέλου, τόσο για συγκρίσεις μεταξύ των ομάδων όσο και εντός των ομάδων. Η δυνατότητα προσφέρεται μόνο αν έχουν δηλωθεί κύριες επιδράσεις στην ενότητα Display Means For.
- *Confidence interval adjustment*: Επιλέγουμε τη μέθοδο προσαρμογής του επιπέδου στατιστικής σημαντικότητας για τις πολλαπλές συγκρίσεις. Μπορούμε να επιλέξουμε μεταξύ των μεθόδων LSD, Bonferroni, ή Sidak. Η δυνατότητα προσφέρεται μόνο αν έχει επιλεγεί η παράμετρος Compare main effects.

Η επιλογή Display Εμφανίζει πίνακες στατιστικών στοιχείων, όπως:

- *Descriptive statistics*: Παρέχει έναν πίνακα με περιγραφικά μέτρα όπως μέσο, τυπική απόκλιση και μέγεθος δείγματος για όλες τις εξαρτημένες μεταβλητές.
- *Estimates of effect size*: Παρέχει την εκτιμήτρια η^2 για κάθε επίδραση και κάθε εκτιμώμενη παράμετρο. Η εκτιμήτρια η^2 περιγράφει την αναλογία της συνολικής μεταβλητότητας που αναλογεί σε κάθε παράγοντα.
- *Observed power*: Παρέχει την ισχύ του ελέγχου όταν η εναλλακτική υπόθεση βασίζεται στην παρατηρούμενη τιμή.
- *Parameter estimates*: Παρέχει τον πίνακα με τους εκτιμώμενους συντελεστές του μοντέλου, τις τυπικές τους αποκλίσεις, τα t tests, τα διαστήματα εμπιστοσύνης και την ισχύ (p-values) των ελέγχων.
- *Homogeneity tests*: Παρέχει τον έλεγχο του Levene για τον έλεγχο της ομοιογένειας (ισότητας) της διασποράς της εξαρτημένης μεταβλητής σε όλα τα επίπεδα συνδυασμών των ομάδων που προκύπτουν από τις κατηγορικές μεταβλητές. Επίσης, περιλαμβάνονται οι έλεγχοι Box's M test της ομοιογένειας του πίνακα συνδιασποράς της εξαρτημένης μεταβλητής.

- *Spread-versus-level*: Χρησιμοποιείται για τον έλεγχο των υποθέσεων του μοντέλου, και εμφανίζεται όταν δεν υπάρχουν παράγοντες.
- *Residual plots*: Χρησιμοποιείται για τον έλεγχο των υποθέσεων του μοντέλου, π.χ. με τη δημιουργία ενός γραφήματος διασποράς των προβλεπόμενων παρατηρήσεων σε σχέση με τα τυποποιημένα σφάλματα της εξαρτημένης μεταβλητής, ώστε να εξεταστούν οι υποθέσεις κανονικότητας, ομοσκεδαστικότητας και ανεξαρτησίας.
- *Lack of fit*: Για τον έλεγχο της περίπτωσης που η σχέση μεταξύ της εξαρτημένης και των ανεξάρτητων μεταβλητών μπορεί να περιγραφεί επακριβώς από το μοντέλο.
- *General estimable function*: Μας επιτρέπει την κατασκευή ειδικών (custom) υποθέσεων βασισόμενες σε μετασχηματισμούς γενικών συναρτήσεων.
- *Significance level*: Ορίζεται το επίπεδο στατιστικής σημαντικότητας που θα χρησιμοποιείται από τις αναλύσεις (π.χ. post hoc tests, ή διαστήματα εμπιστοσύνης). Η τιμή αυτή χρησιμοποιείται επίσης για τον υπολογισμό της παρατηρούμενης ισχύος του ελέγχου (p-value).

Η διαδικασία ολοκληρώνεται με το πλήκτρο *OK* όπου και λαμβάνουμε:

Descriptive Statistics

Dependent Variable: outcome

drugA	drugB	Mean	Std. Deviation	N
0	0	20,4300	2,41387	10
	1	23,8200	2,73813	10
	Total	22,1250	3,05543	20
1	0	24,2700	2,80913	10
	1	27,8100	3,67225	10
	Total	26,0400	3,66382	20
Total	0	22,3500	3,22156	20
	1	25,8150	3,75882	20
	Total	24,0825	3,87529	40

Εικόνα 6.41: Περιγραφικά μέτρα των ομάδων

Levene's Test of Equality of Error Variances^a

Dependent Variable: outcome

F	df1	df2	Sig.
,587	3	36	,627

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + drugA + drugB + drugA * drugB

Εικόνα 6.42: Έλεγχος της ισότητας των διασπορών

Tests of Between-Subjects Effects

Dependent Variable: outcome

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	273,391 ^a	3	91,130	10,505	,000
Intercept	23198,672	1	23198,672	2674,139	,000
drugA	153,272	1	153,272	17,668	,000
drugB	120,062	1	120,062	13,840	,001
drugA * drugB	,056	1	,056	,006	,936
Error	312,307	36	8,675		
Total	23784,370	40			
Corrected Total	585,698	39			

a. R Squared = ,467 (Adjusted R Squared = ,422)

Εικόνα 6.43: ANOVA Table

Parameter Estimates

Dependent Variable: outcome

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	27,810	,931	29,858	,000	25,921	29,699
[drugA=0]	-3,990	1,317	-3,029	,005	-6,661	-1,319
[drugA=1]	0 ^a
[drugB=0]	-3,540	1,317	-2,688	,011	-6,211	-,869
[drugB=1]	0 ^a
[drugA=0] * [drugB=0]	,150	1,863	,081	,936	-3,628	3,928
[drugA=0] * [drugB=1]	0 ^a
[drugA=1] * [drugB=0]	0 ^a
[drugA=1] * [drugB=1]	0 ^a

a. This parameter is set to zero because it is redundant.

Εικόνα 6.44: Ο πίνακας των συντελεστών (b) του γραμμικού μοντέλου

Από τον πίνακα διαπιστώνουμε ότι δεν υπάρχει αλληλεπίδραση ($p\text{-value} = 0,936 > 0,005$), ενώ οι δύο πρώτες υποθέσεις απορρίπτονται στα επίπεδα στατιστικής σημαντικότητας $\alpha = 0,05$ και $\alpha = 0,01$. Συνεπώς οι διαφορές μεταξύ των μέσων των δύο στηλών και των δύο γραμμών είναι στατιστικά σημαντικές που σημαίνει ότι 1 μονάδα του ναρκωτικού A παράγει μεγαλύτερη διέγερση από 0 μονάδες του ναρκωτικού και παρόμοια 1 μονάδα του ναρκωτικού B παράγει μεγαλύτερη διέγερση από 0 μονάδες του ίδιου ναρκωτικού. Οπότε 1 μονάδα από τα ναρκωτικά A και B γενικά αυξάνει τη σωματική διέγερση χωρίς ταυτόχρονα να υπάρχει αλληλεπίδραση μεταξύ των δύο ναρκωτικών.

6.10 Friedman

Η δυνατότητα προσφέρεται από την επιλογή *Analyze > Non parametric Tests > Legacy Dialogs > Related Samples*, με σκοπό τον έλεγχο συσχέτισης επαναλαμβανόμενων μετρήσεων μιας συνεχούς μεταβλητής. Η μέθοδος χρησιμοποιείται όταν δεν ισχύουν οι υποθέσεις εφαρμογής της Repeated Measures ANOVA, δηλ. όταν οι κατανομές δεν είναι κανονικές και οι διασπορές δεν είναι ίσες. Η μέθοδος είναι μια επέκταση του κριτηρίου sign test για k συσχετιζόμενα δείγματα.

Για την εφαρμογή της μεθόδου οι παρατηρήσεις διατάσσονται κατά αύξουσα σειρά και σε κάθε παρατήρηση προσδιορίζεται η σχετική της θέση (rank) και στη συνέχεια υπολογίζεται το άθροισμα των σχετικών θέσεων.

Η ελεγχουσυνάρτηση προκύπτει από τη σχέση:

$$T = \frac{12}{nk(k+1)} \sum_{j=1}^k \frac{R_j^2}{n_j} - 3h(k+1) \sim X_{k-1}^2$$

Εξίσωση 6.52

Όπου:

$$R_j = \sum_{i=1}^{n_j} r_{ij}$$

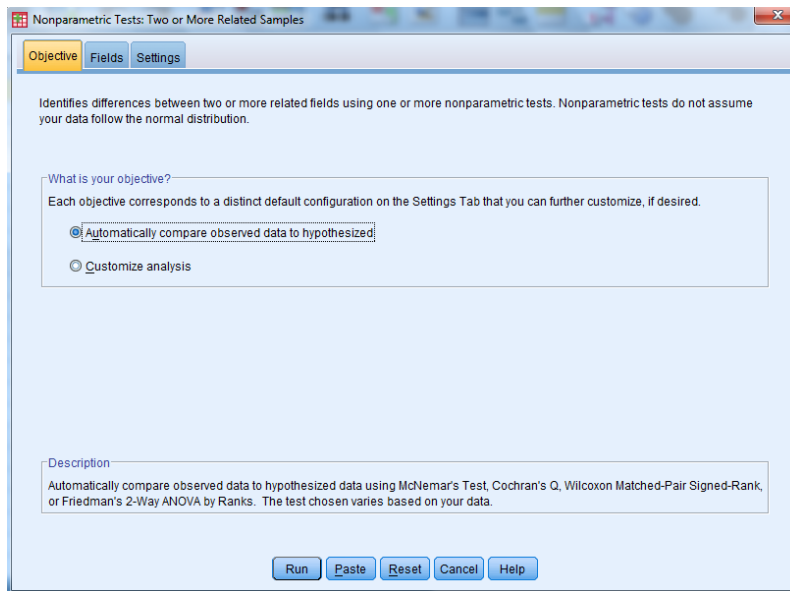
Εξίσωση 6.53

Το άθροισμα των θέσεων της κάθε j -ομάδας.

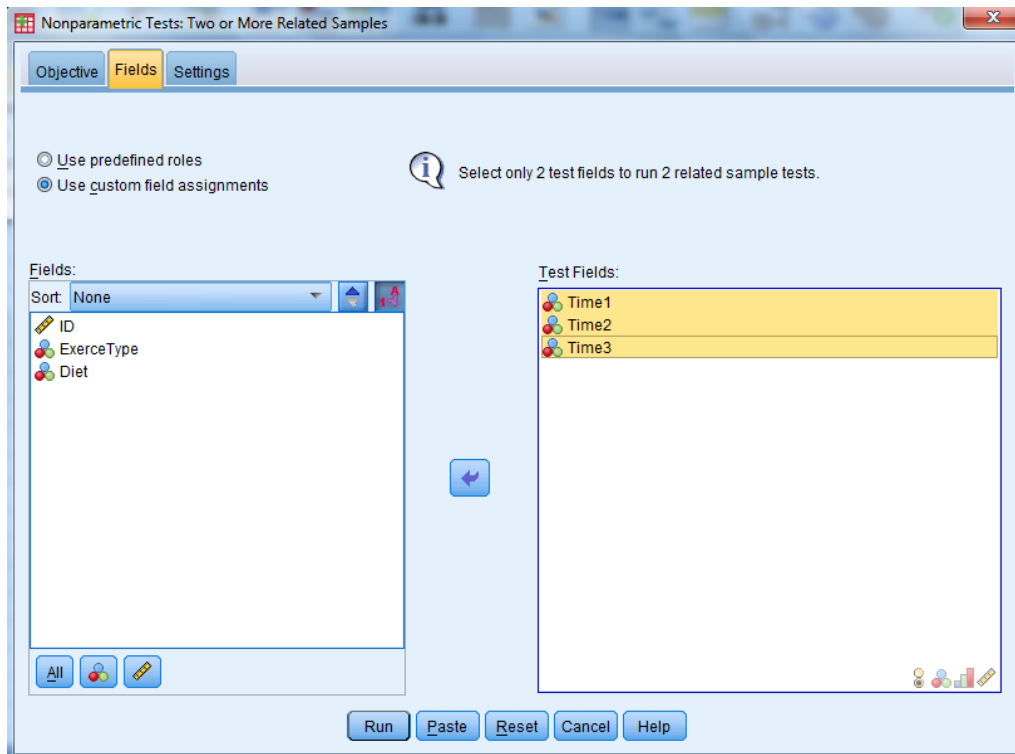
Η υπόθεση απορρίπτεται όταν:

$$T > X_{k-1; \alpha}^2$$

Εξίσωση 6.54

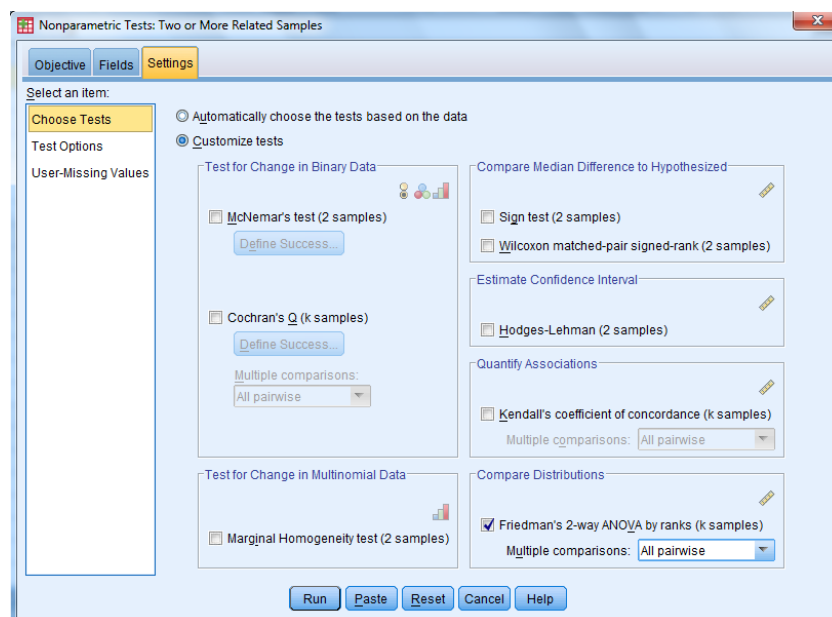


Εικόνα 6.45: Το menu Non parametric Tests: Two or More Related Samples > Step 1 Objective



Εικόνα 6.46: To menu *Non parametric Tests: Two or More Related Samples* > Step 2 Fields

Στο πεδίο *Test Fields* εισάγουμε την ή τις επαναλαμβανόμενες μετρήσεις:



Εικόνα 6.47: To menu *Non parametric Tests: Two or More Related Samples* > Step 3 Settings

Στην ενότητα Settings επιλέγουμε τον έλεγχο Friedman's Two-way ANOVA by ranks (k samples) για τη σύγκριση των κατανομών των επαναλαμβανόμενων μετρήσεων.

Σε αυτή την ενότητα παρέχονται επιπρόσθετα οι έλεγχοι:

- *Mc Nemar Test (2 samples)*: Βλ. επίσης ομώνυμη ενότητα.

- *Cohran's Q* (*k samples*): Είναι μια μη παραμετρική διαδικασία η οποία ελέγχει κατά πόσο *k* δοκιμές (π.χ. θεραπείες) δίτιμων μεταβλητών έχουν ταυτόσημα αποτελέσματα. Ο έλεγχος είναι ισοδύναμος του Mc Nemar Test.

Η διαδικασία ολοκληρώνεται με το πλήκτρο *OK* όπου και λαμβάνουμε:

	Mean	Std. Deviation	N
Time1	93,41	6,062	29
Time2	102,10	14,477	29
Time3	105,00	18,949	29

Εικόνα 6.48: Περιγραφικά μέτρα των μέσων

	Mean Rank
Time1	1.26
Time2	2.26
Time3	2.48

Εικόνα 6.49: Οι μέσοι των βαθμών των ομάδων

N	29
Chi-Square	25.748
df	2
Asymp. Sig.	.000

a. Friedman Test

Εικόνα 6.50: Αποτελέσματα του ελέγχου Friedman

Από τα αποτελέσματα του ελέγχου διαπιστώνουμε ότι η υπόθεση της ισότητας των μέσων απορρίπτεται, συνεπώς υπάρχει διαφορά μεταξύ των επαναλαμβανόμενων μετρήσεων.

6.11 X²-test

Η δυνατότητα προσφέρεται από την επιλογή *Analyze > Descriptive Statistics > Crosstabs*, με σκοπό τον έλεγχο για την ύπαρξη σχέσης μεταξύ δύο ποιοτικών μεταβλητών.

Έχουμε ήδη αναφέρει ότι περιγραφικά οι από κοινού κατανομές συχνότητας δύο ποιοτικών μεταβλητών απεικονίζονται σε έναν πίνακα συνάφειας (contingency table) διπλής εισόδου.

Οι υποθέσεις που εξετάζονται είναι:

H_0 : Δεν υπάρχει συσχέτιση μεταξύ των μεταβλητών A & B, ή αλλιώς οι παρατηρηθείσες συχνότητες είναι ίσες με τις αναμενόμενες.

H_1 : Υπάρχει συσχέτιση μεταξύ των μεταβλητών A & B, ή αλλιώς οι παρατηρηθείσες συχνότητες δεν είναι ίσες με τις αναμενόμενες.

Η στατιστική συνάρτηση δίνεται από τη σχέση:

$$X^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(o_{ij} - e_{ij})^2}{e_{ij}} \sim X^2_{(r-1)(c-1);a}$$

Εξίσωση 6.55

Όπου:

O_{ij} : Οι παρατηρούμενες τιμές από κάθε κελί και e_{ij} : οι αναμενόμενες τιμές κάθε κελιού, οι οποίες υπολογίζονται από το πηλίκο του γινομένου του οριζόντιου και του κάθετου αθροίσματος του κάθε κελιού, προς το γενικό σύνολο.

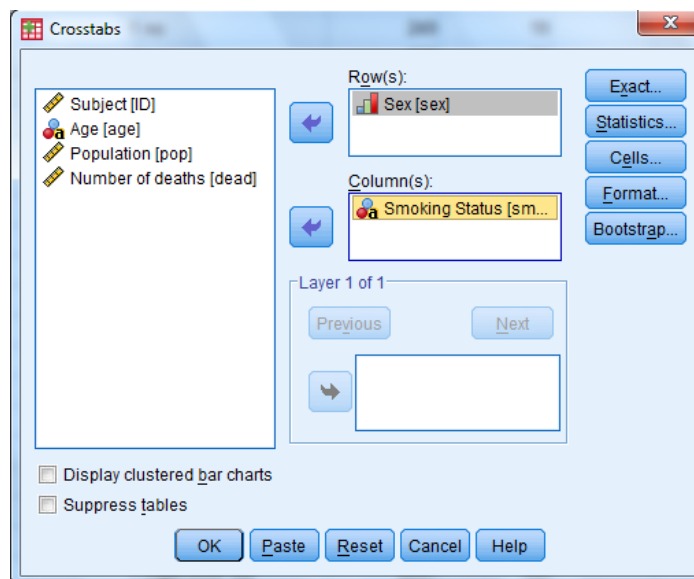
Υποθέσεις του ελέγχου:

- Όλες οι αναμενόμενες τιμές είναι μεγαλύτερες του ένα
- Περισσότερες από το 80% των αναμενόμενων τιμών ξεπερνά το 5

Όταν οι υποθέσεις εφαρμογής του X^2 τεστ δεν ισχύουν, τότε το SPSS εμφανίζει σημείωση που το αναφέρει και τότε λαμβάνουμε υπόψη το Fisher's Exact Test.

Έστω, ότι έχουμε τις μεταβλητές «Φύλο» και «Συνήθεια Καπνίσματος» και θέλουμε να εξετάσουμε την πιθανή συσχέτισή τους από τα αποτελέσματα των μετρήσεων σε μια ομάδα θεραπείας. Συνεπώς:

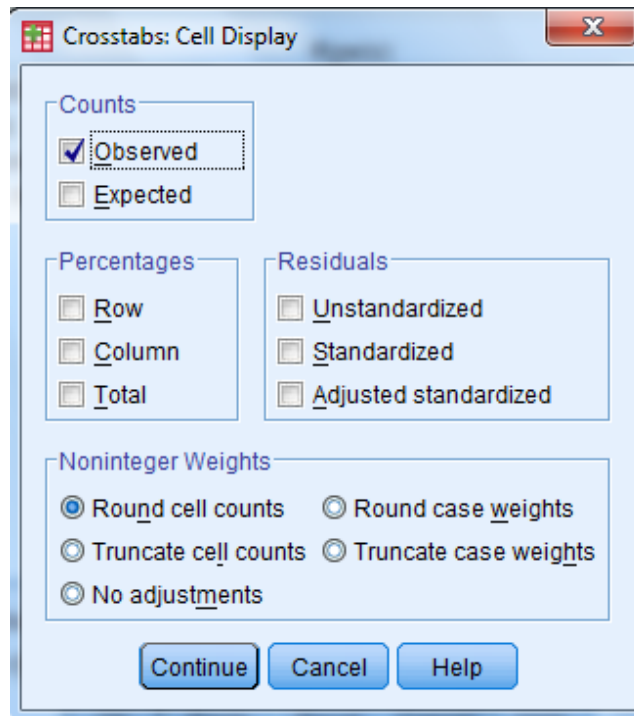
Οι δύο μεταβλητές εισάγονται στα πεδία Row(s) και Column(s). Συνηθίζουμε να εισάγουμε στις γραμμές τη μεταβλητή με τις περισσότερες κατηγορίες για λόγους ομοιομορφίας του πίνακα συνάφειας που θα δημιουργηθεί.



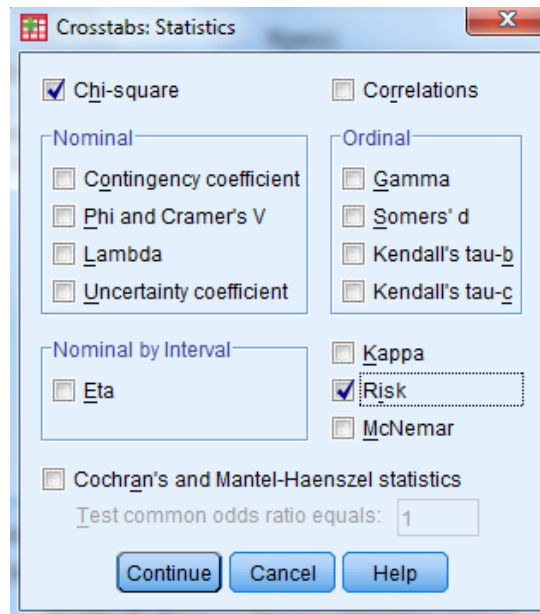
Εικόνα 6.51: Το menu Crosstabs

Από το πλήκτρο *Cell* αν επιθυμούμε επιλέγουμε επιπρόσθετα *Expected* για να εμφανιστούν οι αναμενόμενες συχνότητες. Περισσότερα για τις επιλογές του πλήκτρου *Cell* βλ. Παράρτημα «Crosstabs > Cell».

Από το πλήκτρο *Statistics* επιλέγουμε *Chi-square* για να υπολογιστεί το στατιστικό κριτήριο. Περισσότερα για τις επιλογές του πλήκτρου *Statistics* βλ. Παράρτημα «Crosstabs > Statistics».



Εικόνα 6.52: To menu Crosstabs > Cell



Εικόνα 6.53: To menu Crosstabs > Statistics

Η διαδικασία ολοκληρώνεται με το πλήκτρο *OK* όπου και λαμβάνουμε:

Sex * Smoking Status Crosstabulation

			Smoking Status				Total
			cigarPipeOnly	cigaretteOnly	cigarettePlus	no	
Sex	Male	Count	3	5	6	5	19
		Expected Count	4,8	4,8	4,8	4,8	19,0
	Female	Count	6	4	3	4	17
		Expected Count	4,3	4,3	4,3	4,3	17,0
Total		Count	9	9	9	9	36
		Expected Count	9,0	9,0	9,0	9,0	36,0

Εικόνα 6.54: Πίνακας συνάφειας του Φύλου με την Καπνιστική Συνήθεια

Ο πίνακας συνάφειας εμφανίζει τις παρατηρηθείσες και τις αναμενόμενες συχνότητες. Παράδειγμα, για το πρώτο κελί η αναμενόμενη τιμή δίνεται από τη σχέση $(19 \cdot 9) / 36$.

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	2,118 ^a	3	,548
Likelihood Ratio	2,150	3	,542
N of Valid Cases	36		

a. 8 cells (100,0%) have expected count less than 5. The minimum expected count is 4,25.

Εικόνα 6.55: Αποτελέσματα του χ^2 τεστ

Από τον πίνακα διαπιστώνουμε ότι η μηδενική υπόθεση δεν μπορεί να απορριφθεί ($p\text{-value} = 0,548 > 0,05$) συνεπώς δεν υπάρχει στατιστικά σημαντική συσχέτιση μεταξύ του φύλου και τις καπνιστικής συνήθειας.

Παρακάτω παραθέτουμε μια περίπτωση δεδομένων στα οποία οι υποθέσεις εφαρμογής του χ^2 τεστ δεν ισχύουν, οπότε το SPSS εμφανίζει σχετική σημείωση που το αναφέρει, και σε αυτή την περίπτωση λαμβάνουμε υπόψη το Fisher's Exact Test.

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	,026 ^b	1	,873		
Continuity Correction ^a	,000	1	1,000		
Likelihood Ratio	,026	1	,873		
Fisher's Exact Test				1,000	,598
Linear-by-Linear Association	,025	1	,875		
N of Valid Cases	26				

a. Computed only for a 2x2 table

b. 1 cells (25,0%) have expected count less than 5. The minimum expected count is 3,81.

Εικόνα 6.56: Αποτελέσματα του Fisher Exact τεστ στο παράδειγμα Διακοπής Σχολείου με Διακοπή Θεραπευτικής Αγωγής

Μέσω του χ^2 τεστ ελέγχουμε ανά δύο την ύπαρξη συσχετίσεων (μονοπαραγοντική ανάλυση). Ο έλεγχος αυτός είναι ένα μέτρο που μας δείχνει αν οι μεταβλητές συσχετίζονται ή όχι, δηλαδή δεν μας λέει το πόσο συσχετίζονται, παρά μόνο την ύπαρξη ή όχι κάποιας συσχέτισης.

Λόγος Συμπληρωματικών Πιθανοτήτων

Συνεπώς, για να εξεταστεί ο βαθμός συσχέτισης, στις περιπτώσεις 2x2 πινάκων συνάφειας, όπου ο έλεγχος προκύπτει στατιστικά σημαντικός, υπολογίζουμε τα μέτρα κινδύνου π.χ. το OR (Odds Ratio).

Risk Estimate			
	Value	95% Confidence Interval	
		Lower	Upper
Odds Ratio for Διακοπή Σχολείου (OXI / NAI)	1,143	,224	5,841
For cohort Διακοπή Θεραπευτικής Σχέσης = OXI	1,059	,522	2,147
For cohort Διακοπή Θεραπευτικής Σχέσης = NAI	,926	,367	2,337
N of Valid Cases	26		

Εικόνα 6.57: Εκτίμηση μέτρων του βαθμού κινδύνου, στο παράδειγμα Διακοπής Σχολείου με Διακοπή Θεραπευτικής Αγωγής

Έστω ότι έχουμε τον ακόλουθο γενικό πίνακα, στον οποίο παρουσιάζονται οι συχνότητες ατόμων που εκτέθηκαν σε έναν παράγοντα και εμφάνισαν ή όχι μια ασθένεια. Εάν προκύψει ότι ο παράγοντας σχετίζεται με την ασθένεια θα πρέπει να διερευνηθεί ο βαθμός αυτής της συσχέτισης.

		Ασθένεια		Σύνολο
		Ναι (D^+)	Όχι (D^-)	
Παράγοντας	Ναι (E^+)	a	b	$a + b$
	Όχι (E^-)	c	d	$c + d$
Σύνολο		$a + c$	$b + d$	n

Εικόνα 6.58: 2x2 πίνακας συνάφειας Ασθένειας και Παράγοντα

Το OR είναι ο λόγος των συμπληρωματικών πιθανοτήτων, δηλαδή του λόγου πιθανότητας εμφάνισης της ασθένειας στα άτομα που εκτέθηκαν στον παράγοντα, προς τον λόγο πιθανότητας εμφάνισης της ασθένειας στα άτομα που δεν εκτέθηκαν στον παράγοντα, και δίνεται από τη σχέση:

$$OR = \frac{\frac{P(D^+ | E^+)}{P(D^- | E^+)}}{\frac{P(D^+ | E^-)}{P(D^- | E^-)}} = \frac{\frac{a / (a + c)}{c / (a + c)}}{\frac{b / (b + d)}{d / (b + d)}} = \frac{\frac{a}{c}}{\frac{b}{d}} = \frac{ad}{bc}$$

Εξίσωση 6.56

Ερμηνεία

- Όταν $OR = 1$ τότε δεν υπάρχει συσχέτιση μεταξύ της ασθένειας και της έκθεσης στον παράγοντα.
- Όταν $OR = a > 1$ τότε η έκθεση είναι επιβαρυντική για την ασθένεια, δηλαδή η πιθανότητα εμφάνισης της ασθένειας είναι a φορές μεγαλύτερη από την πιθανότητα μη εμφάνισης της ασθένειας. Δηλαδή υπάρχει μεγαλύτερη πιθανότητα, τόσες φορές όσες είναι το OR ασθένειας αν υπάρχει έκθεση στον παράγοντα, σε σχέση με την πιθανότητα ασθένειας αν δεν υπάρχει έκθεση στον παράγοντα.
- Όταν $OR = a < 1$ τότε η έκθεση στον παράγοντα είναι προστατευτική, δηλαδή εάν εκτεθείς στον παράγοντα έχεις $1-a$ φορές λιγότερη πιθανότητα να ασθηνήσεις.

Στο παράδειγμα Διακοπής Σχολείου με Διακοπή Θεραπευτικής Αγωγής, έχουμε $OR = 1,143$ δηλαδή οι ασθενείς που δεν διακόπτουν τη θεραπευτική αγωγή έχουν 1,143 φορές περισσότερη πιθανότητα να μη διακόψουν το σχολείο σε σχέση με αυτούς που διέκοψαν τη θεραπευτική αγωγή.

6.12 Mc Nemar test

Η δυνατότητα προσφέρεται από την επιλογή *Analyze > Descriptive Statistics > Crosstabs*, επιλέγοντας Mc Nemar από το πλήκτρο *Statistics*.

Σκοπός είναι να εξεταστεί η συμφωνία δύο κατηγορικών μεταβλητών, όταν οι παρατηρήσεις εμφανίζουν αντιστοιχία κατά ζεύγη. Ο έλεγχος αυτός είναι ο μη παραμετρικός ισοδύναμος του *T-test* για εξαρτημένα δείγματα. Σ' αυτή την περίπτωση τα δεδομένα παρουσιάζονται σε έναν πίνακα συνάφειας της μορφής:

		Κατηγορική Β		Σύνολο
		+	-	
Κατηγορική Α	+	c_1	d_1	$c_1 + d_1$
	-	d_2	c_2	$d_2 + c_2$
Σύνολο		$c_1 + d_2$	$d_1 + c_2$	n

Εικόνα 6.59: 2x2 Πίνακας συνάφειας κατά ζεύγη δεδομένων κατηγορικών μεταβλητών

Η υπόθεση που εξετάζεται είναι:

H_0 : Δεν υπάρχει διαφορά μεταξύ των απαντήσεων μεταβλητών A και B

H_1 : Υπάρχει διαφορά μεταξύ των απαντήσεων μεταβλητών A και B

Η ελεγχοσυνάρτηση δίνεται από τη σχέση:

$$X^2 = \frac{(d_1 - d_2)^2}{d_1 + d_2} \sim X^2_{1;a}$$

Εξίσωση 6.57

Ελεγχοσυνάρτηση με διόρθωση κατά Yates:

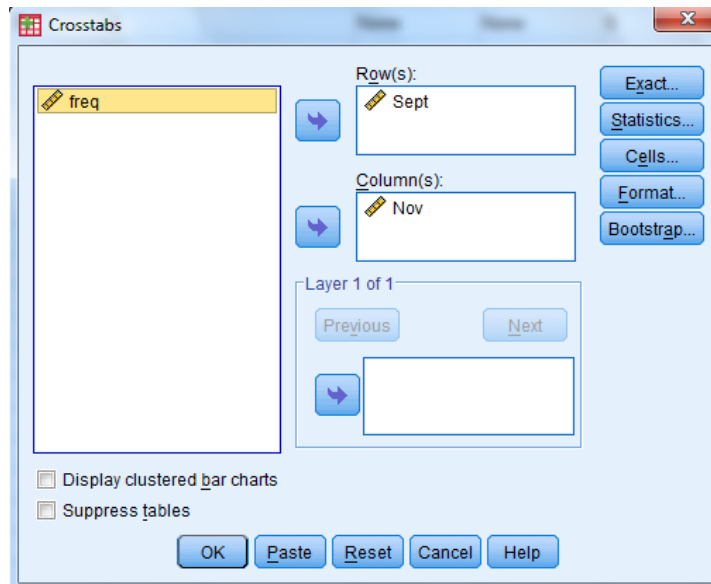
$$X^2 = \frac{(|d_1 - d_2| - 1)^2}{d_1 + d_2} \sim X^2_{1;a}$$

Εξίσωση 6.58

Η υπόθεση απορρίπτεται όταν:

$$X^2 > X^2_{1;a}$$

Εξίσωση 6.59



Εικόνα 6.60: Το menu Crosstabs

Από το πλήκτρο *Statistics* επιλέγουμε Mc Nemar για να υπολογιστεί το στατιστικό κριτήριο. Η διαδικασία ολοκληρώνεται με το πλήκτρο *OK* όπου και λαμβάνουμε:

Sept * Nov Crosstabulation

Count

		Nov		Total
		,00	1,00	
Sept	,00	60	15	75
	1,00	5	20	25
Total		65	35	100

Εικόνα 6.61: Πίνακας συνάφειας των ατόμων που κρυολόγησαν τους μήνες Σεπτέμβριο και Νοέμβριο

Τον πίνακα συνάφειας των συχνοτήτων των παρατηρήσεων για τους συνδυασμούς των κατηγορικών μεταβλητών.

Chi-Square Tests

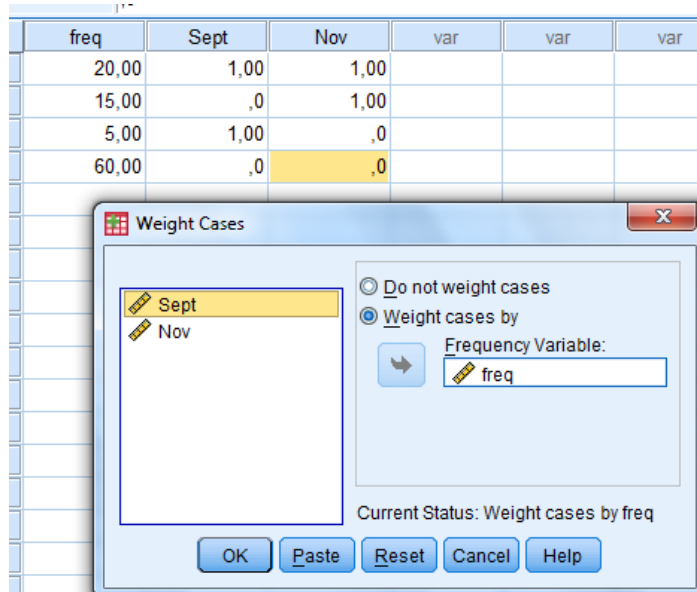
	Value	Exact Sig. (2-sided)
McNemar Test		,041 ^a
N of Valid Cases	100	

a. Binomial distribution used.

Εικόνα 6.62: Αποτελέσματα του ελέγχου

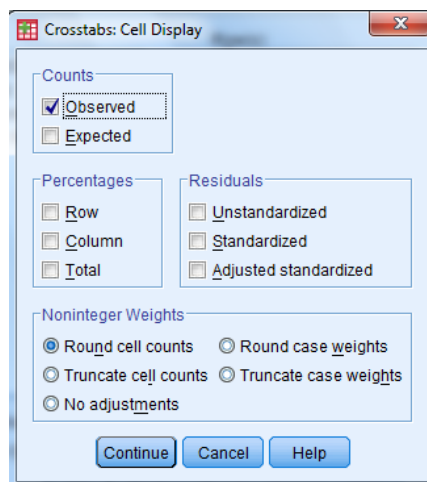
Η υπόθεση που εξετάζεται είναι εάν υπάρχει διαφορά στην πιθανότητα νόσησης μεταξύ των μηνών Σεπτεμβρίου και Νοεμβρίου. Από τα αποτελέσματα του ελέγχου διαπιστώνουμε ότι η μηδενική υπόθεση απορρίπτεται ($p\text{-value} = 0,041 < 0,05$), συνεπώς δεν υπάρχει διαφορά στην πιθανότητα νόσησης μεταξύ των δύο μηνών.

Σημείωση: Στην περίπτωση που τα δεδομένα έχουν γραφεί σε *Wide Format*, τότε θα πρέπει να σταθμιστούν μέσω της εντολής *Weight Cases* (βλ. ομώνυμη ενότητα).



Εικόνα 6.63: Στάθμιση δεδομένων που έχουν καταγραφεί σε Wide Format

6.13 Παράρτημα: Crosstabs > Cell



Εικόνα 6.64: Το menu Crosstabs > Cell

Συχνότητες (Counts)

- *Observed*: Επιλέγεται για να εμφανιστούν οι παρατηρούμενες συχνότητες των συνδυασμών των κατηγοριών της /των ποιοτικής/ών μεταβλητής/τών.
- *Expected*: Επιλέγεται για να εμφανιστούν οι αναμενόμενες συχνότητες των συνδυασμών των κατηγοριών της /των ποιοτικής/ών μεταβλητής/τών.

Ποσοστά (Percentages)

- *Row*: Επιλέγεται για να εμφανιστούν οι σχετικές συχνότητες (%) ως προς τα αθροίσματα στις γραμμές του πίνακα.
- *Column*: Επιλέγεται για να εμφανιστούν οι σχετικές συχνότητες (%) ως προς τα αθροίσματα στις στήλες του πίνακα.

- *Total*: Επιλέγεται για να εμφανιστούν οι σχετικές συχνότητες (%) ως προς το σύνολο των παρατηρήσεων.

Υπόλοιπα (Residuals)

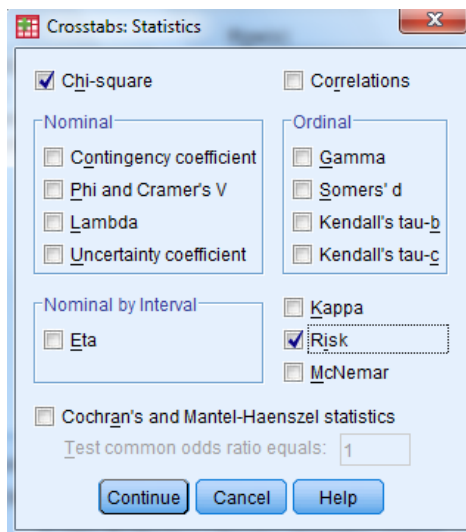
- *Unstandardized*: Επιλέγεται για να εμφανιστούν τα μη τυποποιημένα υπόλοιπα από τη διαφορά παρατηρούμενων και αναμενόμενων συχνοτήτων.
- *Standardized*: Επιλέγεται για να εμφανιστούν τα τυποποιημένα υπόλοιπα από τη διαφορά παρατηρούμενων και αναμενόμενων συχνοτήτων.
- *Adjusted standardized*: Επιλέγεται για να εμφανιστούν τα αναπροσαρμοσμένα υπόλοιπα. Ισχύει ότι με τη μηδενική υπόθεση ότι οι 2 μεταβλητές είναι ανεξάρτητες, τα αναπροσαρμοσμένα υπόλοιπα θα ακολουθούν την τυπική κανονική κατανομή, δηλαδή έχουν μέση τιμή 0 και τυπική απόκλιση 1. Έτσι, ένα προσαρμοσμένο υπόλοιπο που είναι πάνω από 1,96 (κατά συνθήκη χρησιμοποιείται 2,0) δείχνει ότι ο αριθμός των παρατηρήσεων στο εν λόγω κελί είναι σημαντικά μεγαλύτερος από ό,τι θα αναμενόταν εάν η μηδενική υπόθεση ήταν αληθής, με επίπεδο σημαντικότητας 0,05. Ένα αναπροσαρμοσμένο υπόλοιπο μικρότερο από -2,0 υποδεικνύει ότι ο αριθμός των παρατηρήσεων στο εν λόγω κελί είναι σημαντικά μικρότερος από ό,τι θα αναμενόταν εάν η μηδενική υπόθεση ήταν αληθής. Έτσι, ανάλογα με το πρόσημο του αναπροσαρμοσμένου υπολοίπου, διαπιστώνουμε αν σε κάθε κελί υπάρχουν λιγότερες ή περισσότερες παρατηρήσεις από ό,τι θα περίμενε κανείς αν οι δύο μεταβλητές ήταν ανεξάρτητες. Αν ο πίνακας συνάφειας είναι 2x2, τότε όλα τα προσαρμοσμένα σφάλματα θα έχουν την ίδια απόλυτη τιμή, με ακριβώς 2 από αυτά να είναι αρνητικά.
- *Noninteger Weights*: Συνήθως κάθε κελί περιέχει ακέραιους αριθμούς, δεδομένου ότι αποτυπώνουν τις συχνότητες εμφάνισης. Ωστόσο, στις περιπτώσεις που τα δεδομένα σταθμίζονται από μεταβλητή με βαρύτητα κλασματικών τιμών (π.χ. 1,25), τότε επίσης ενδέχεται ο αριθμός που θα προκύψει σε κάθε κελί να είναι κλασματικός. Σ' αυτές τις περιπτώσεις, υπάρχουν επιλογές με σκοπό την περικοπή ή στρογγυλοποίηση πριν ή μετά τον υπολογισμό των συχνοτήτων σε κάθε κελί.

Οι επιλογές που παρέχονται είναι:

- *Round Cell Counts*: Επιλέγεται έτσι ώστε οι σταθμισμένες συχνότητες των κελιών να χρησιμοποιηθούν όπως είναι, αλλά οι αθροιστικές σταθμισμένες τιμές στρογγυλοποιούνται πριν από τον υπολογισμό των στατιστικών μέτρων.
- *Truncate Cell Counts*: Επιλέγεται έτσι ώστε οι σταθμισμένες συχνότητες των κελιών να χρησιμοποιηθούν όπως είναι, αλλά οι αθροιστικές σταθμισμένες τιμές να περικοπούν πριν από τον υπολογισμό των στατιστικών μέτρων. Σ' αυτή την περίπτωση δεν γίνεται στρογγυλοποίηση στις τιμές αλλά αποκοπή από ένα σημείο και μετά. Για παράδειγμα, οι αριθμοί 1,4142 και 1,4987 περικόπτονται και οι δύο στον αριθμό 1,4.
- *Round Case Weights*: Επιλέγεται έτσι ώστε κάθε στάθμιση να στρογγυλοποιείται πριν τη χρήση της.
- *Truncate Case Weights*: Επιλέγεται έτσι ώστε κάθε στάθμιση να περικόπτεται πριν τη χρήση της.
- *No adjustments*: Επιλέγεται έτσι ώστε να μην απαιτούνται προσαρμογές. Οι σταθμίσεις χρησιμοποιούνται όπως είναι, το ίδιο και οι κλασματικές συχνότητες των κελιών. Ωστόσο, όταν πρέπει να υπολογιστούν ακριβή στατιστικά μέτρα (τα οποία είναι διαθέσιμα μόνο μέσω της επιλογής Exact tests), τότε οι αθροιστικές σταθμίσεις των κελιών είτε θα στρογγυλοποιηθούν είτε θα περικοπούν πριν τη χρήση τους στον υπολογισμό των στατιστικών μέτρων.

6.14 Παράρτημα: Crosstabs > Statistics

Από το πλήκτρο *Statistics* γίνεται η επιλογή των ελέγχων υποθέσεων και δεικτών συνάφειας μεταξύ των κατηγορικών μεταβλητών.



Εικόνα 6.65: Το menu *Crosstabs > Statistics*

- *Chi-square*: Επιλέγεται για τον υπολογισμό των ελέγχων Pearson X^2 τεστ, likelihood-ratio X^2 τεστ, Fisher X^2 exact test, και X^2 τεστ κατά Yates (διόρθωση συνέχειας) σε πίνακες 2x2. Το Fisher X^2 exact test υπολογίζεται για πίνακες 2x2 όταν δεν ισχύουν οι υποθέσεις του X^2 test, δηλ. όταν υπάρχει αναμενόμενη τιμή μικρότερη του ένα, ή όταν περισσότερες από το 80% των αναμενόμενων είναι μικρότερες του 5. Για πίνακες ανεξαρτήτων πλήθους γραμμών και στηλών υπολογίζονται τα Pearson X^2 τεστ και likelihood-ratio X^2 τεστ, ενώ όταν και οι δύο μεταβλητές είναι ποσοτικές τότε προσφέρεται και ο έλεγχος γραμμικής τάσης (the linear-by-linear association test).
- *Correlations*: Επιλέγεται για τον υπολογισμό του πίνακα συσχετίσεων, για πίνακες στους οποίους οι γραμμές και οι στήλες περιέχουν διατεταγμένες τιμές. Οι συσχετίσεις αποδίδονται από τον συντελεστή συσχέτισης του Spearman, r (μόνο για αριθμητικά δεδομένα). Περισσότερα για τον συντελεστή συσχέτισης r του Spearman βλ. παράγραφο *Correlations*.
- *Nominal*: Περιέχει επιλογές με σκοπό τον υπολογισμό στατιστικών μέτρων συνάφειας για ονομαστικά δεδομένα (κατηγορικά). Σε αυτή την υποενοότητα μπορούμε να επιλέξουμε:
- *Contingency coefficient*: Ο συντελεστής συνάφειας είναι ένα μέτρο της σύνδεσης συσχέτισης των μεταβλητών. Οι τιμές του κυμαίνονται μεταξύ 0 και 1, με τιμή 0 να μην υποδεικνύει συσχέτιση μεταξύ των γραμμών και των στηλών (δηλ. συσχέτισης των μεταβλητών) και τιμές κοντά στο 1 να υποδεικνύουν υψηλό βαθμό συσχέτισης μεταξύ των μεταβλητών. Η μέγιστη δυνατή τιμή εξαρτάται από τον αριθμό των σειρών και των στηλών σε έναν πίνακα.
- *Phi and Cramer's V*: Ο συντελεστής συνάφειας ϕ (Yule, 1912) είναι ένα μέτρο που βασίζεται στο X^2 του Pearson και προκύπτει από τη ρίζα του πηλίκου του X^2 με το μέγεθος δείγματος,

$$\phi = \sqrt{\frac{X^2}{n}}$$

Εξίσωση 6.60

Στη βιβλιογραφία ο συντελεστής αναφέρεται και ως συντελεστής μέσης τετραγωνικής συνάφειας του Pearson (Pearson's coefficient of mean square contingency). Για πίνακες 2x2 ο συντελεστής ϕ λαμβάνει τιμές στο διάστημα $-1, 1$ χωρίς ωστόσο αυτό να είναι απόλυτο.

Ο συντελεστής V (Crammer, 1946) είναι μια επέκταση του συντελεστή ϕ για $I \times J$ πίνακες συνάφειας με $I, J > 2$ και προκύπτει από τη σχέση:

$$V = \sqrt{\frac{\phi^2}{q-1}}$$

Εξίσωση 6.61

Όπου:

$$q = \min\{I, J\}$$

Εξίσωση 6.62

Εκφράζοντας τη συνάφεια μεταξύ δύο μεταβλητών ως ποσοστό της μέγιστης δυνατής μεταβλητότητάς τους. Ο συντελεστής V λαμβάνει τιμές στο διάστημα $[0,1]$ με $V = 1$ όταν οι δύο μεταβλητές έχουν ίσα περιθώρια αθροίσματα. Καθώς τα περιθώρια αθροίσματα διαφέρουν όλο και περισσότερο η τιμή του V πλησιάζει το 0.

- *Lambda*: Ο συντελεστής συνάφειας λ (Goodman & Kruskal, 1954) ανήκει στα μέτρα προγνωστικής συνάφειας, και αποτελεί ένα μέτρο της αναλογικής μείωσης του σφάλματος, όταν οι τιμές της ανεξάρτητης μεταβλητής χρησιμοποιούνται για να προβλέψουν τις τιμές της εξαρτημένης μεταβλητής. Ο συντελεστής λ λαμβάνει τιμές στο διάστημα $[0,1]$, με $\lambda = 0$, όταν οι μεταβλητές είναι ανεξάρτητες, και $\lambda = 1$, όταν με βάση τη μια μεταβλητή μπορεί να γίνει τέλεια πρόγνωση των τιμών της άλλης.
- *Uncertainty coefficient*: Ο συντελεστής αβεβαιότητας U (Theil's uncertainty coefficient, 1972) ανήκει και αυτός στα μέτρα προγνωστικής συνάφειας και επίσης, είναι γνωστός και ως συντελεστής εντροπίας (entropy coefficient). Διαφοροποιείται από τον συντελεστή λ με την έννοια ότι λαμβάνει υπόψη του ολόκληρη την κατανομή της εξαρτημένης μεταβλητής και όχι μόνο την πιο συχνή κατηγορία της (όπως συμβαίνει στον συντελεστή λ). Λαμβάνει τιμές στο διάστημα $[0,1]$, με $U = 0$ όταν οι μεταβλητές είναι ανεξάρτητες, και $U = 1$, όταν με βάση τη μια μεταβλητή μπορεί να γίνει τέλεια πρόγνωση των τιμών της άλλης. Για παράδειγμα, $U = 0,83$ συνεπάγεται ότι η γνώση της μιας μεταβλητής, μειώνει το σφάλμα πρόβλεψης των τιμών της άλλης κατά 83%. Το πρόγραμμα υπολογίζει και τις συμμετρικές και τις μη συμμετρικές εκδοχές του συντελεστή αβεβαιότητας.
- *Ordinal*: Περιέχει επιλογές με σκοπό τον υπολογισμό στατιστικών μέτρων συνάφειας για διατάξιμα δεδομένα. Σ' αυτή την υποενότητα μπορούμε να επιλέξουμε:
- *Gamma*: Ο συντελεστής γ (Goodman & Kruskal, 1954) είναι ένα μέτρο διατακτικής συνάφειας, που βασίζεται στη διαφορά «συμφωνιών» (concordant) και «ασυμφωνιών» (discordant). Λαμβάνει τιμές στο διάστημα $[-1, 1]$, με $\gamma = 1$ ή $\gamma = -1$ να υποδεικνύει ισχυρή θετική ή αρνητική συσχέτιση, ενώ για $\gamma = 0$ μικρή ή καμία συσχέτιση.
- *Somers' d*: Ο συντελεστής D (Somers, 1962) είναι ένα μέτρο για την πρόβλεψη της εξαρτημένης μεταβλητής γνωρίζοντας την ανεξάρτητη. Είναι παραλλαγή του μέτρου Gamma, λαμβάνοντας υπόψη τυχόν ισοβαθμίες (ties) των ζευγών των παρατηρήσεων. Λαμβάνει τιμές στο διάστημα $[-1, 1]$, με $D = 1$ όταν δεν υπάρχουν «ασύμφωνα» ζεύγη μεταξύ των μεταβλητών X και Y , και η διάταξη των X είναι σε πλήρη συμφωνία με τη διάταξη των τιμών της Y (ισχυρή συσχέτιση μεταξύ των δύο μεταβλητών), ενώ $D = -1$ όταν δεν υπάρχουν «σύμφωνα» ζεύγη μεταξύ των μεταβλητών X και Y , και η διάταξη των τιμών της X είναι σε πλήρη ασυμφωνία με τη διάταξη των τιμών της Y .
- *Kendall's tau-b*: Ο συντελεστής τ (Kendall, 1938) είναι ένα διατακτικό μέτρο συνάφειας, το οποίο έχει τρεις παραλλαγές. Ο συντελεστής τ_b λαμβάνει υπόψη τις περιπτώσεις ισοβαθμιών (ties). Λαμβάνει τιμές στο διάστημα $[-1, 1]$, με το πρόσημο να καθορίζει την κατεύθυνση της σχέσης, όπως και στα προηγούμενα μέτρα.

- *Kendall's tau-c*: Ο συντελεστής τ (Kendall, 1938) είναι ένα διατακτικό μέτρο συνάφειας, το οποίο έχει τρεις παραλλαγές. Ο συντελεστής τ_c , σε αντίθεση με τον προηγούμενο δείκτη, αγνοεί τις περιπτώσεις ισοβαθμιών (ties). Λαμβάνει και αυτός τιμές στο διάστημα $[-1, 1]$, με το πρόσημο ομοίως, να καθορίζει την κατεύθυνση της σχέσης.
- *Nominal by Interval*: Περιέχει επιλογές με σκοπό τον υπολογισμό στατιστικών μέτρων συνάφειας μεταξύ μιας (κωδικοποιημένης) κατηγορικής και μιας ποσοστικής μεταβλητής.
- *Eta*: Ο συντελεστής η είναι ένα μέτρο συσχέτισης μεταξύ δύο μεταβλητών. Λαμβάνει τιμές στο διάστημα $[0, 1]$, με $\eta = 0$ όταν δεν υπάρχει συσχέτιση μεταξύ των μεταβλητών και $\eta = 1$ ή κοντά στο 1, όταν υπάρχει υψηλός βαθμός συσχέτισης. ο συντελεστής είναι κατάλληλος όταν η μια εξαρτημένη μεταβλητή μετράται σε κλίμακα (π.χ. το εισόδημα) και η άλλη ανεξάρτητη μεταβλητή είναι κατηγορική, με περιορισμένο αριθμό κατηγοριών (π.χ. το φύλο).
- *Kappa*: Ο συντελεστής k Cohen είναι ένα στατιστικό μέτρο της συμφωνίας μεταξύ των αξιολογήσεων δύο βαθμολογητών όταν και οι δύο βαθμολογούν το ίδιο αντικείμενο. Λαμβάνει τιμές στο διάστημα $[0, 1]$, με την τιμή 1 να δηλώνει τέλεια συμφωνία. Για να γίνει ο υπολογισμός του συντελεστή θα πρέπει και οι δύο μεταβλητές να έχουν ίδιο τύπο δεδομένων (string ή numeric) και το ίδιο καθορισμένο μήκος.
- *Risk*: Ο συντελεστής κινδύνου, υπολογίζεται για πίνακες 2×2 και είναι ένα μέτρο του βαθμού συσχέτισης δύο παραγόντων. Ορίζεται ως ο λόγος της επίπτωσης (αναλογία πιθανοτήτων) δύο ομάδων με διαφορετική έκθεση σε έναν παράγοντα. Αν το διάστημα εμπιστοσύνης του συντελεστή κινδύνου περιέχει την τιμή 1, τότε δεν είμαστε σε θέση να υποθέσουμε ότι κάποιος παράγοντας συνδέεται με το αποτέλεσμα/συμβάν. Η αναλογία πιθανοτήτων μπορεί να χρησιμοποιηθεί ως μια εκτίμηση του σχετικού κινδύνου ή όταν η εμφάνιση του παράγοντα είναι σπάνια.
- *McNemar*: Πρόκειται για μία μη παραμετρική στατιστική δοκιμασία για τον έλεγχο της συμφωνίας δύο κατηγορικών μεταβλητών, όταν οι παρατηρήσεις εμφανίζουν αντιστοιχία κατά ζεύγη. Ο έλεγχος αυτός είναι ο μη παραμετρικός ισοδύναμος του T-test για εξαρτημένα δείγματα, βλ. ομώνυμη ενότητα.
- *Cochran's and Mantel-Haenszel statistics*: Μ' αυτή την επιλογή λαμβάνονται οι στατιστικές Cochran και Mantel-Haenszel για τον έλεγχο της ανεξαρτησίας μεταξύ μιας δίτιμης μεταβλητής παράγοντα και μιας δίτιμης μεταβλητής απόκρισης, που εξαρτάται από μια μεταβλητή που ορίζει ένα ή περισσότερα στρώματα (ελέγχου). Θα πρέπει να ληφθεί υπόψη ότι ενώ άλλα στατιστικά μέτρα υπολογίζονται ξεχωριστά για κάθε στρώμα, οι στατιστικές Cochran και Mantel-Haenszel υπολογίζονται μία φορά για όλες τις στρώσεις.

Βιβλιογραφικές Αναφορές Κεφαλαίου

Ελληνόγλωσσες

Κουτσογέρας, Γ. (2010). *Μέτρα συνάφειας και μέτρα ασυμμετρίας για πίνακες συνάφειας*. Μεταπτυχιακή εργασία, τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης, Πανεπιστήμιο Πειραιά.

Ξενόγλωσσες

Cramér, H. (1946). *Mathematical Methods of Statistics* (Princeton New Jersey: *Princeton Mathematical Series, NJ*, pp. 367-369.

Goodman, L. A. & Kruskal, W. H. (1954). Measures of association for cross classifications. *Journal of the American Statistical Association*, 49(268), pp. 732-764.

Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika*, pp. 81-93.

Somers, R. H. (1962). A new asymmetric measure of association for ordinal variables. *American Sociological Review*, pp. 799-811.

Theil, H. (1972). *Statistical decomposition analysis; with applications in the social and administrative sciences* Amsterdam: North Holland Publishing Company.

Yule, G. U. (1912). On the methods of measuring association between two attributes. *Journal of the Royal Statistical Society*, (75-76) pp. 579-652.

IBM Knowledge Center. Retrieved August 30, 2015 from http://www-01.ibm.com/support/knowledgecenter/SSLVMB_20.0.0/com.ibm.spss.statistics.help/idh_glm_u_sav.htm