ΚΕΦΑΛΑΙΟ ΕΚΤΟ

Έλεγχος για τις παραμέτρους θέσης δύο πληθυσμών με εξαρτημένα δείγματα

Στο κεφάλαιο αυτό θα ασχοληθούμε με τον έλεγχο της υπόθεσης της ισότητα δύο μέσων τιμών με εξαρτημένα δείγματα. Εξαρτημένα δείγματα εμφανίζονται συνήθως στις ακόλουθες περιπτώσεις:

α) Σε πειράματα, μελέτες των οποίων ο σκοπός είναι η διερεύνηση της αποτελεσματικότητας μίας θεραπείας. Για το λόγο αυτό οι τιμές μίας ή περισσοτέρων μεταβλητών καταγράφονται στην ίδια πειραματική μονάδα πριν και μετά την εφαρμογή της μεθόδου.

β) Στην περίπτωση των διδύμων.

γ) Όταν θεωρούμε πειραματικές μονάδες που μοιάζουν σε όλα τα υπόλοιπα χαρακτηριστικά πλην αυτού που θέλουμε να μελετήσουμε (ταιριαστά δεδομένα).

Στη βάση των εξαρτημένων δειγμάτων θα ασχοληθούμε με το ακόλουθο πρόβλημα:

Έστω ένα τυχαίο δείγμα $X_1,\ldots,X_n$ μεγέθους n από έναν πληθυσμό με μέση τιμή μ_1 και διακύμανση σ_1^2. Επιπλέον έστω ένα τυχαίο δείγμα $Y_1,\ldots,Y_n$ μεγέθους n από έναν πληθυσμό με μέση τιμή μ_2 και διακύμανση σ_2^2. Επιπρόσθετα υποθέτουμε ότι τα δύο δείγματα είναι εξαρτημένα. Ενδιαφερόμαστε για τον έλεγχο, σε επίπεδο σημαντικότητας α, της μηδενικής υπόθεσης

$$H_0 : \mu_1 = \mu_2,$$

ως προς μία εκ των

$$H_a : \mu_1 > \mu_2, \quad H_a : \mu_1 < \mu_2, \; H_a : \mu_1 \neq \mu_2 .$$

Το πρώτο βήμα για τη μελέτη του προβλήματος είναι η δημιουργία των διαφορών $D_i = X_i - Y_i,\; i = 1,\ldots,n$.

Το παραπάνω πρόβλημα ελέγχεται υπό κάποιες υποθέσεις με τον παραμετρικό έλεγχο του t-test. Όταν κάποιες από τις υποθέσεις αυτές δεν ικανοποιείται και δεν υπάρχει τρόπος διόρθωσης του προβλήματος ο έλεγχος ανάγεται σε αυτόν ότι οι πληθυσμιακές διάμεσοι είναι ίσες. Τα αποτελέσματα του τελευταίου ελέγχου γενικεύονται για τον δοθέν έλεγχο όταν τα δεδομένα είναι συμμετρικά.

6.1 Μεθοδολογία-Υλοποίηση στο S.P.S.S.

Η μεθοδολογία που θα χρησιμοποιηθεί για τη στατιστική ανάλυση ενός τέτοιου προβλήματος εξαρτάται από το αν πληρούνται ή όχι κάποιες προϋποθέσεις, τις οποίες και πρέπει αρχικά να ελέγξει ο ερευνητής. Πιο συγκεκριμένα, ελέγχουμε

α) αν το ποσοστό των ακραίων τιμών στις διαθέσιμες παρατηρήσεις $D_i = X_i - Y_i,\ i = 1,...,n$, ξεπερνά το 10% αυτών, και

β) αν ο πληθυσμός από τον οποίο λαμβάνεται το τυχαίο δείγμα $D_i = X_i - Y_i,\ i = 1,...,n$, μπορούμε να ισχυριστούμε ότι περιγράφεται ικανοποιητικά από την κανονική κατανομή.

Ανάλογα με τα αποτελέσματα των παραπάνω ελέγχων προβαίνουμε στον παραμετρικό ή στο μη παραμετρικό έλεγχο. Για το λόγο αυτό στη συνέχεια παρουσιάζονται όλα τα πιθανά αποτελέσματα των α) και β), τα διάφορα βήματα της ανάλυσης και οι αποφάσεις στις οποίες οδηγούμαστε.

1. Αρχικά ελέγχουμε αν υπάρχουν ακραίες τιμές στις διαθέσιμες δειγματικές τιμές D_i . Αν το ποσοστό των ακραίων τιμών δε ξεπερνά το 10%, τότε προχωρούμε στο επόμενο βήμα. Αν το ποσοστό των ακραίων τιμών ξεπερνά το 10%, τότε δοκιμάζουμε μήπως ο μετασχηματισμός του λογαρίθμου διορθώνει το πρόβλημα. Αν το πρόβλημα αυτό διορθώνεται τότε μεταβαίνουμε στο βήμα 2, σε διαφορετική περίπτωση συμπεραίνουμε ότι θα χρησιμοποιηθεί ο μη παραμετρικός έλεγχος (βλέπε βήμα 4).
2. Στο βήμα 2, χρησιμοποιώντας το τεστ των Shapiro-Wilk καθώς και γραφικούς τρόπους, ελέγχουμε αν οι διαθέσιμες δειγματικές παρατηρήσεις D_i (είτε οι αρχικές είτε οι μετασχηματισμένες του βήματος 1) μπορούν να θεωρηθούν ότι προέρχονται από έναν πληθυσμό που περιγράφεται ικανοποιητικά από την κανονική κατανομή. Αν ο έλεγχος της κανονικότητας μας υποδεικνύει ότι η υπόθεση της κανονικότητας δεν απορρίπτεται (p-τιμή >α), τότε η ανάλυση θα συνεχιστεί με τον παραμετρικό έλεγχο (βλέπε βήμα 3). Αν η υπόθεση της κανονικότητας απορρίπτεται (τεστ Shapiro-Wilk, p-τιμή <α), τότε ελέγχουμε αν το πρόβλημα της μη κανονικότητας διορθώνεται μετασχηματίζοντας κατάλληλα τα δεδομένα (Box-Cox μετασχηματισμός) και επανελέγχοντας την ύπαρξη ακραίων τιμών, δηλαδή ξεκινώντας την ανάλυση από το βήμα 1. Αν με κάποιο μετασχηματισμό των δεδομένων

επιτυγχάνεται η κανονικότητα συνεχίζουμε την ανάλυση παραμετρικά (βήμα 3). Σε αντίθετη περίπτωση, αν το πλήθος των δειγματικών παρατηρήσεων D_i, μη λαμβάνοντας υπόψη αυτές που έχουν αφαιρεθεί στο βήμα 1, είναι μεγάλο (συνήθως μεγαλύτερο του 30), κάνοντας χρήση του Κεντρικού Οριακού Θεωρήματος, προβαίνουμε στον παραμετρικό έλεγχο της υπό έλεγχο υπόθεσης (βλέπε βήμα 3), όπου η p-τιμή του ελέγχου και το διάστημα εμπιστοσύνης θα είναι προσεγγιστικά. Στην περίπτωση τώρα που το πρόβλημα της μη κανονικότητας δε διορθώνεται (τεστ Shapiro-Wilk, p-τιμή <α), και ταυτόχρονα το πλήθος των δειγματικών παρατηρήσεων, μη λαμβάνοντας υπόψη αυτές που έχουν αφαιρεθεί στο βήμα 1, είναι μικρό (συνήθως μικρότερο του 30), συνεχίζεται η περαιτέρω ανάλυση μη παραμετρικά (βήμα 4).

3. <u>Παραμετρικός έλεγχος- Τ τεστ συγκρίσεως ζευγών</u>: Χρησιμοποιούμε τη στατιστική συνάρτηση $t = \dfrac{\bar{D}}{S_D / \sqrt{n}} \overset{H_0}{\sim} t_{n-1}$, όπου $\bar{D}$ και S_D η μέση τιμή και τυπική απόκλιση του δείγματος των διαφορών $D_i = X_i - Y_i,\ i = 1,2,...,n$. Οι κρίσιμες περιοχές του ελέγχου είναι αντίστοιχα $t \geq t_{n-1,a}$, $t \leq -t_{n-1,a}$, $|t| \geq t_{n-1,a/2}$ $\left(t \geq t_{n-1,a/2} \text{ ή } t \leq -t_{n-1,a/2}\right)$. Επιπλέον το 100(1-α)% Δ.Ε. για την $\mu_1 - \mu_2$ είναι:

$$\left(\bar{D} - t_{n-1,a/2}\frac{S_D}{\sqrt{n}}, \bar{D} + t_{n-1,a/2}\frac{S_D}{\sqrt{n}}\right).$$

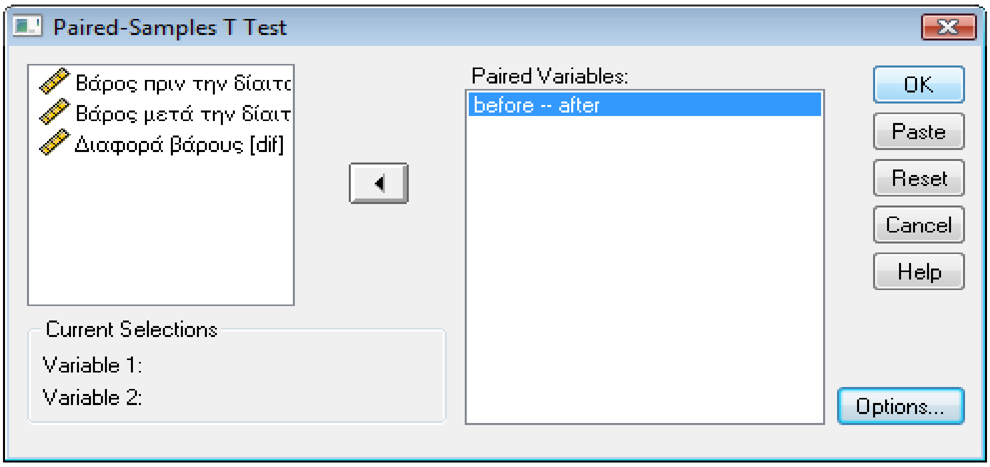
<u>Επισήμανση</u>: Σε περίπτωση που έχει χρησιμοποιηθεί κάποιος μετασχηματισμός διόρθωσης του προβλήματος είτε της ύπαρξης πολλών ακραίων τιμών είτε της μη κανονικότητας, τότε όλα τα παραπάνω αναφέρονται στις μετασχηματισμένες τιμές και στο τροποποιημένο σε μέγεθος δείγμα. Ειδικότερα, αν έχει χρησιμοποιηθεί ο μετασχηματισμός του λογαρίθμου, θα προβούμε στον έλεγχο αν ο μέσος λογάριθμος του ενός πληθυσμού δε διαφέρει στατιστικά σημαντικά από το μέσο λογάριθμο του άλλου.

<u>Υλοποίηση στο S.P.S.S.</u>

Για τη διεξαγωγή του t τεστ συγκρίσεως ζευγών, από το κύριο μενού του λογισμικού επιλέγουμε

i. Analyze→Compare Means→Paired-Samples T Test.

ii. Στο νέο παράθυρο διαλόγου που προκύπτει διαλέγουμε αρχικά τη μεταβλητή που καταγράφει π.χ. τις τιμές του βάρους πριν την εφαρμογή της δίαιτας και η οποία θα χαρακτηρισθεί ως Variable 1. Έπειτα επιλέγουμε εκείνη που καταγράφει τις τιμές του βάρους μετά την εφαρμογή της δίαιτα και η οποία θα πάει στο πλαίσιο Variable2. Στη συνέχεια μετακινούμε το ζεύγος των μεταβλητών στο πλαίσιο Paired Variables.



iii. Από την επιλογή Options έχουμε τη δυνατότητα να καθορίσουμε τον τρόπο χειρισμού των ελλιπών τιμών καθώς και να προσδιορίσουμε το βαθμό εμπιστοσύνης του διαστήματος εμπιστοσύνης που θα κατασκευαστεί για τη μέση απώλεια βάρους.

<u>Σχόλιο</u>: Εναλλακτικά ο έλεγχος της υπόθεσης $H_0 : \mu_D = d_0$ θα μπορούσε να διενεργηθεί μέσω της διαδικασίας Compare Means One Sample για τη μεταβλητή όπου καταγράφονται οι δειγματικές τιμές της διαφοράς και με Test Value την τιμή d_0.

4. <u>Μη παραμετρικός έλεγχος-Wilcoxon</u>: Σύμφωνα με αυτόν θεωρούμε προς έλεγχο την ισοδύναμη μηδενική υπόθεση ότι οι διαφορές $D_i = X_i - Y_i$, $i = 1,\ldots,n$, προέρχονται από μία συμμετρική περί το μηδέν κατανομή. Για την υλοποίηση του αρχικά τοποθετούμε τις διαφορές $D_i = X_i - Y_i$, $i = 1,\ldots,n$, σε αύξουσα τάξη μεγέθους μη λαμβάνοντας υπόψη το πρόσημο. Έπειτα αντικαθιστούμε την j κατά σειρά μεγέθους διαφορά με $+j$ ή $-j$ ανάλογα με το αν η συγκεκριμένη διαφορά είναι θετική ή αρνητική αντίστοιχα. Αν υπάρχουν μηδενικές διαφορές απορρίπτονται από τη μελέτη με ανάλογη μείωση του μεγέθους του δείγματος. Έπειτα υπολογίζουμε το στατιστικό $T = \min\left(T^+, T^-\right)$, όπου T^+ και T^- αντίστοιχα είναι το άθροισμα των θετικών και αρνητικών τάξεων αντίστοιχα. Αποδεικνύεται τότε ότι για $n \geq 8$, $Z = \dfrac{T - \dfrac{n(n+1)}{4}}{\sqrt{\dfrac{n(n+1)(2n+1)}{24}}} \underset{H_0}{\overset{\pi\rho o\sigma.}{\sim}} N\left(0,1\right)$, και ο έλεγχος γίνεται κατά τα γνωστά

χρησιμοποιώντας αυτήν τη στατιστική συνάρτηση, με τις p-τιμές να προσδιορίζονται κατά ανάλογο τρόπο με το Z τεστ.

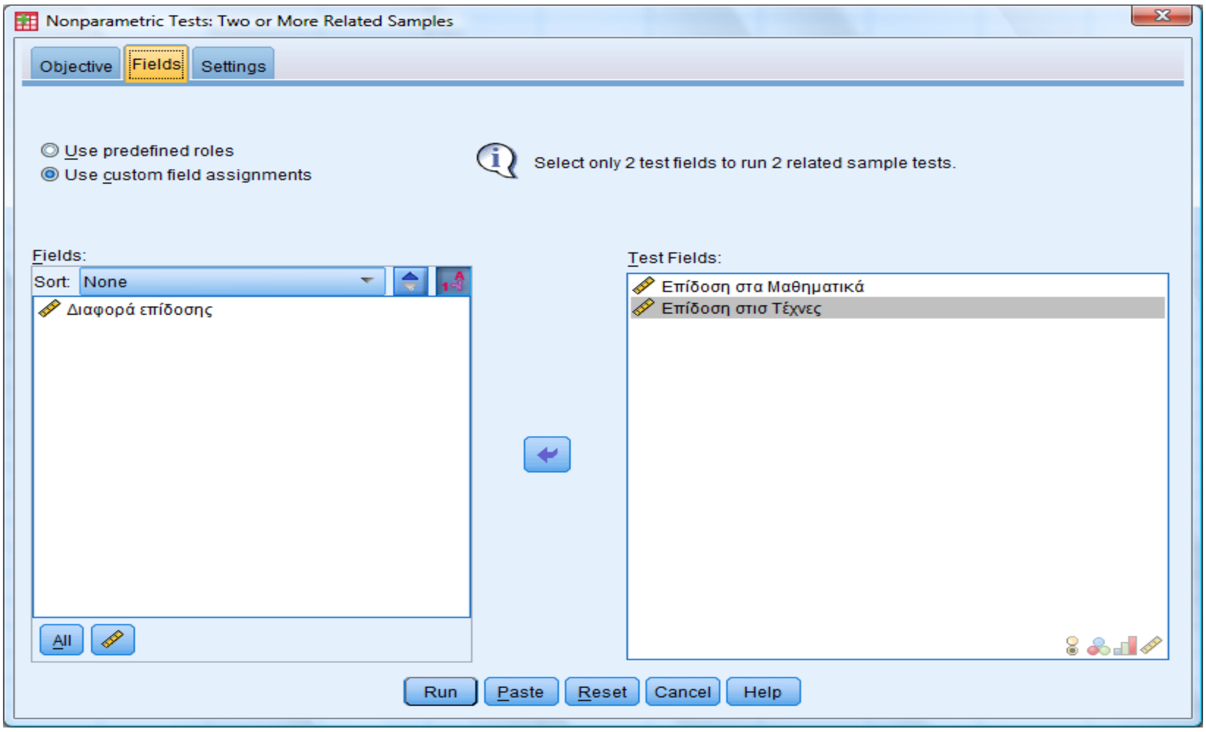
<u>Σχόλιο</u>: Για τον έλεγχο της υπό μελέτης μηδενικής υπόθεσης έχει προταθεί στη βιβλιογραφία και το προσημικό τεστ για συγκρίσεις κατά ζεύγη. Όμως, το τεστ του Wilcoxon είναι αποτελεσματικότερο για αυτό και δεν αναφέρθηκε το προσημικό τεστ.

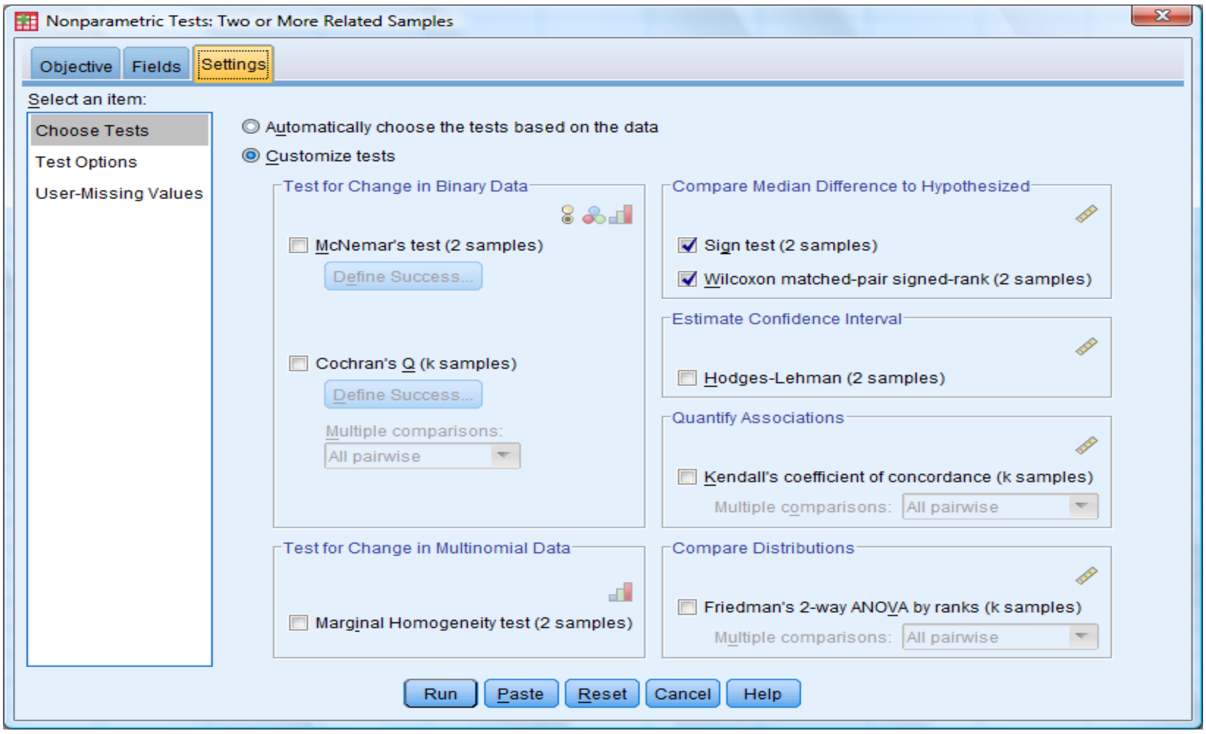
<u>Υλοποίηση στο S.P.S.S.</u>

Από το κύριο μενού επιλέγουμε

i. Analyze→Non Parametric Tests→Related Samples.

ii. Στο νέο παράθυρο διαλόγου που προκύπτει επιλέγουμε στο πλαίσιο Objective την επιλογή Customize analysis, έτσι ώστε στη συνέχεια από τα πλαίσια Fields και Settings να καθορίσουμε τον έλεγχο τον οποίο θέλουμε να διενεργηθεί όπως φαίνεται στα σχήματα που ακολουθούν.





Στο παράθυρο διαλόγου Settings επομένως επιλέγουμε τον προσημικό έλεγχο (Sign test) και τον έλεγχο των Mann-Whiney.

6.2 Παραδείγματα

<u>Παράδειγμα 1°</u>

Στον πίνακα που ακολουθεί (βλέπε Ζωγράφος, 2003, σελ. 206) δίνεται το βάρος σε κιλά 9 γυναικών πριν και μετά την εφαρμογή μίας δίαιτας αδυνατίσματος τεσσάρων εβδομάδων

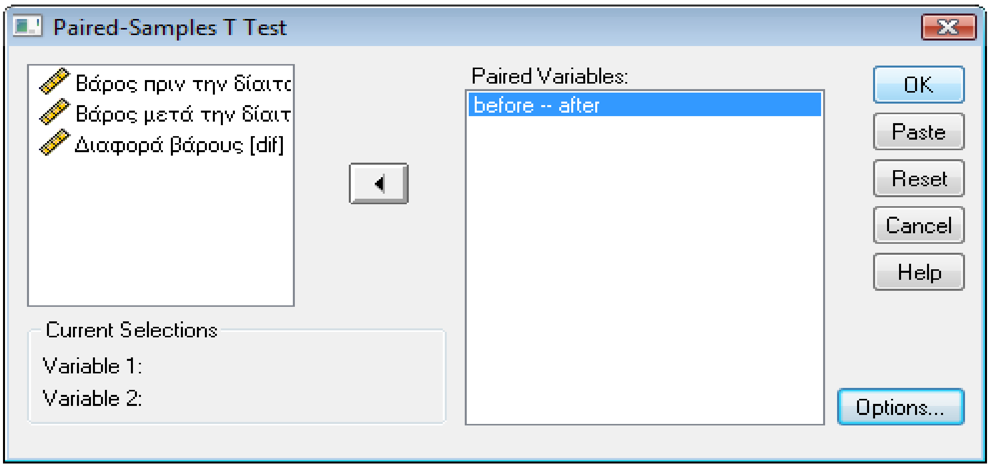
Πριν: 67 81 57 68 67 69 66 77 54

Μετά: 59 69 56 59 63 66 64 71 54.

Με βάση τα δεδομένα αυτά να ελεγχθεί αν είναι εφικτό ο ισχυρισμός ότι η δίαιτα είναι αποτελεσματική.

Επειδή πρόκειται για μετρήσεις του βάρους στα ίδια άτομα πριν και μετά την εφαρμογή μίας δίαιτας τα δύο δείγματα είναι εξαρτημένα. Επομένως, αρχικά ελέγχουμε αν δεν υπάρχουν ακραίες τιμές στις δειγματικές τιμές της απώλεια βάρους και αν αυτές μπορούμε να ισχυριστούμε ότι προέρχονται από κανονικό πληθυσμό (δημιουργία στήλης διαφορών μέσω της διαδικασίας Transform Compute). Προκύπτει από το θηκόγραμμα έχουμε ότι δεν υπάρχουν ακραίες τιμές και χρησιμοποιώντας το στατιστικό τεστ των Shapiro-Wilk προκύπτει ότι η υπόθεση της κανονικότητας δεν μπορεί να απορριφθεί (p-τιμή=0,727>0.05).

Από τη διεξαγωγή του t τεστ συγκρίσεως ζευγών, που επιτυγχάνεται μέσω της διαδικασίας Analyze→Compare Means→Paired-Samples T Test όπως φαίνεται στο σχήμα που ακολουθεί προκύπτουν τα αποτελέσματα που παρατίθενται και ερμηνεύονται



<u>Ερμηνεία αποτελεσμάτων</u>

Από τον πίνακα Paired Samples Statistics έχουμε ότι το μέσο βάρος των 9 ατόμων πριν την δίαιτα ήταν 67.33 κιλά με τυπική απόκλιση 8.44097 και τυπικό σφάλμα για τη μέση τιμή 2.81366. Οι αντίστοιχες ποσότητες μετά τη διενέργεια της δίαιτας είναι 62.33 κιλά, 5.78792 και 1.92931, αντίστοιχα. Παρατηρούμε δηλαδή μία μείωση του βάρους κατά 5 κιλά. Θα εξεταστεί στη συνέχεια αν αυτή είναι στατιστικά σημαντική.

Στον πίνακα Paired Samples Correlations έχουμε ότι το βάρος πριν και το βάρος μετά την δίαιτα είναι συσχετισμένα (υψηλή στατιστικά σημαντική θετική γραμμική συσχέτιση).

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Βάρος πριν την δίαιτα	67,3333	9	8,44097	2,81366
	Βάρος μετά την δίαιτα	62,3333	9	5,78792	1,92931

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Βάρος πριν την δίαιτα & Βάρος μετά την δίαιτα	9	,906	,001

Τα σημαντικότερα όμως αποτελέσματα δίνονται στον πίνακα Paired Samples Test. Προκύπτει ότι η μέση απώλεια βάρους είναι 5 κιλά (δίνεται στη στήλη Mean), ενώ η τυπική απόκλιση των διαφορών και το τυπικό σφάλμα για τη μέση τιμή είναι 4.03113 και 1.34371 κιλά αντίστοιχα. Επιπλέον μας δίνεται το 95% διάστημα εμπιστοσύνης για τη μέση διαφορά (1.90140,8.09860), καθώς και η τιμή του t τεστ για τον έλεγχο ότι δεν υπάρχουν στατιστικά σημαντικές διαφορές στο μέσο βάρος των ατόμων πριν και μετά τη δίαιτα. Από την p-τιμή του ελέγχου συμπεραίνουμε ότι υπάρχουν στατιστικά σημαντικές διαφορές στο μέσο βάρος πριν και μετά τη δίαιτα και αφού η μέση διαφορά είναι 5 καταλαβαίνουμε ότι η δίαιτα επιφέρει στατιστικά σημαντική μείωση του βάρους.

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	Βάρος πριν την δίαιτα - Βάρος μετά την δίαιτα	5	4,03113	1,34371	1,90140	8,09860	3,721	8	,006

Η αναφορά αφήνεται ως άσκηση.

Παράδειγμα 2^o

Παρακάτω παρατίθενται οι επιδόσεις 15 μαθητών στα Μαθηματικά και στις Καλές Τέχνες (βλέπε Παπαΐωάννου και Φερεντίνος, 2000, σελ. 263). Να ελεγχθεί με επίπεδο σημαντικότητας 5% αν υπάρχει στατιστικά σημαντική διαφορά στις επιδόσεις των μαθητών στα δύο γνωστικά αντικείμενα.

Μαθηματικά: 22 37 36 38 42 58 58 60 62 65 66 56 66 67 62

Καλές Τέχνες: 53 68 42 49 51 65 51 71 55 74 68 64 67 73 65.

Πρόκειται για μετρήσεις της επίδοσης σε δύο γνωστικά αντικείμενα και στους ίδιους μαθητές. Επομένως γίνεται εύκολα κατανοητό ότι έχουμε δύο εξαρτημένα δείγματα.

Επειδή πρόκειται για μετρήσεις της επίδοσης στα ίδια άτομα τα δύο δείγματα είναι εξαρτημένα. Επομένως, αρχικά ελέγχουμε αν δεν υπάρχουν ακραίες τιμές στις δειγματικές τιμές της διαφοράς της επίδοσης και αν αυτές μπορούμε να ισχυριστούμε ότι προέρχονται από κανονικό πληθυσμό. Αρχικά σχηματίζουμε στο S.P.S.S. τη στήλη των διαφορών των επιδόσεων στα Μαθηματικά και στις Καλές τέχνες (έστω κωδική ονομασία diafora και Label Διαφορά Επίδοσης). Το πρώτο βήμα της ανάλυσης είναι ο έλεγχος της ύπαρξης ακραίων τιμών και της κανονικότητας των διαφορών. **Διαπιστώνεται (αφήνεται ως άσκηση) ότι υπάρχουν ακραίες τιμές σε ποσοστό μεγαλύτερο του 10% και το πρόβλημα δε διορθώνεται με το μετασχηματισμό του λογαρίθμου, επομένως καταφεύγουμε σε μη παραμετρικούς τρόπους ελέγχου**. Ακολουθώντας τα βήματα που αναλυτικά περιγράφηκαν στην παράγραφο 6.1 προκύπτουν τα ακόλουθα αποτελέσματα

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The median of differences between Επίδοση στισ Τέχνες and Επίδοση στα Μαθηματικά equals 0.	Related-Samples Sign Test	.007[1]	Reject the null hypothesis.
2	The median of differences between Επίδοση στα Μαθηματικά and Επίδοση στισ Τέχνες equals 0.	Related-Samples Wilcoxon Signed Ranks Test	.009	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

[1] Exact significance is displayed for this test.

Από την p-τιμή του στατιστικού τεστ του Wilcoxon συμπεραίνουμε ότι υπάρχουν στατιστικά σημαντικές διαφορές στην επίδοση στα Μαθηματικά και στις Τέχνες (p-τιμή=0.009<0.05). Επιπλέον από τον πίνακα που ακολουθεί (και αποκτήθηκε μέσω της διαδικασίας Analyze Descriptive Statistics Explore) προκύπτει ότι η πληθυσμιακή διάμεσος της επίδοση στις τέχνες είναι στατιστικά σημαντικά μεγαλύτερη και τα αποτελέσματα δεν μπορούν να γενικευτούν στις μέσες τιμές.

Descriptives

			Statistic	Std. Error
Επίδοση στα Μαθηματικά	Mean		53,0000	3,65148
	95% Confidence Interval for Mean	Lower Bound	45,1683	
		Upper Bound	60,8317	
	5% Trimmed Mean		53,9444	
	Median		58,0000	
	Variance		200,000	
	Std. Deviation		14,14214	
	Minimum		22,00	
	Maximum		67,00	
	Range		45,00	
	Interquartile Range		27,00	
	Skewness		-,949	,580
	Kurtosis		-,278	1,121
Επίδοση στισ Τέχνες	Mean		61,0667	2,57546
	95% Confidence Interval for Mean	Lower Bound	55,5428	
		Upper Bound	66,5905	
	5% Trimmed Mean		61,4074	
	Median		65,0000	
	Variance		99,495	
	Std. Deviation		9,97473	
	Minimum		42,00	
	Maximum		74,00	
	Range		32,00	
	Interquartile Range		17,00	
	Skewness		-,466	,580
	Kurtosis		-1,062	1,121

Παράδειγμα 3^{o} Αρχείο Teaching.sav[*]

Στο αρχείο αυτό υπάρχει η βαθμολογία 60 μαθητών οι οποίοι επιλεχθήκαν τυχαία από το σύνολο των μαθητών μια πόλης. Σκοπός της μελέτης αυτής ήταν να εξετάσει, αν μια καινούργια μέθοδος διδασκαλίας ενός μαθήματος βελτιώνει την απόδοση των μαθητών. Pretest είναι η βαθμολογία των μαθητών στο συγκεκριμένο μάθημα πριν την

διδασκαλία με την νέα μέθοδο και Posttest είναι η βαθμολογία των ίδιων μαθητών στο συγκεκριμένο μάθημα μετά την διδασκαλία της νέας μεθόδου. Να εξετασθεί η αποτελεσματικότητα της νέας μεθόδου διδασκαλίας.

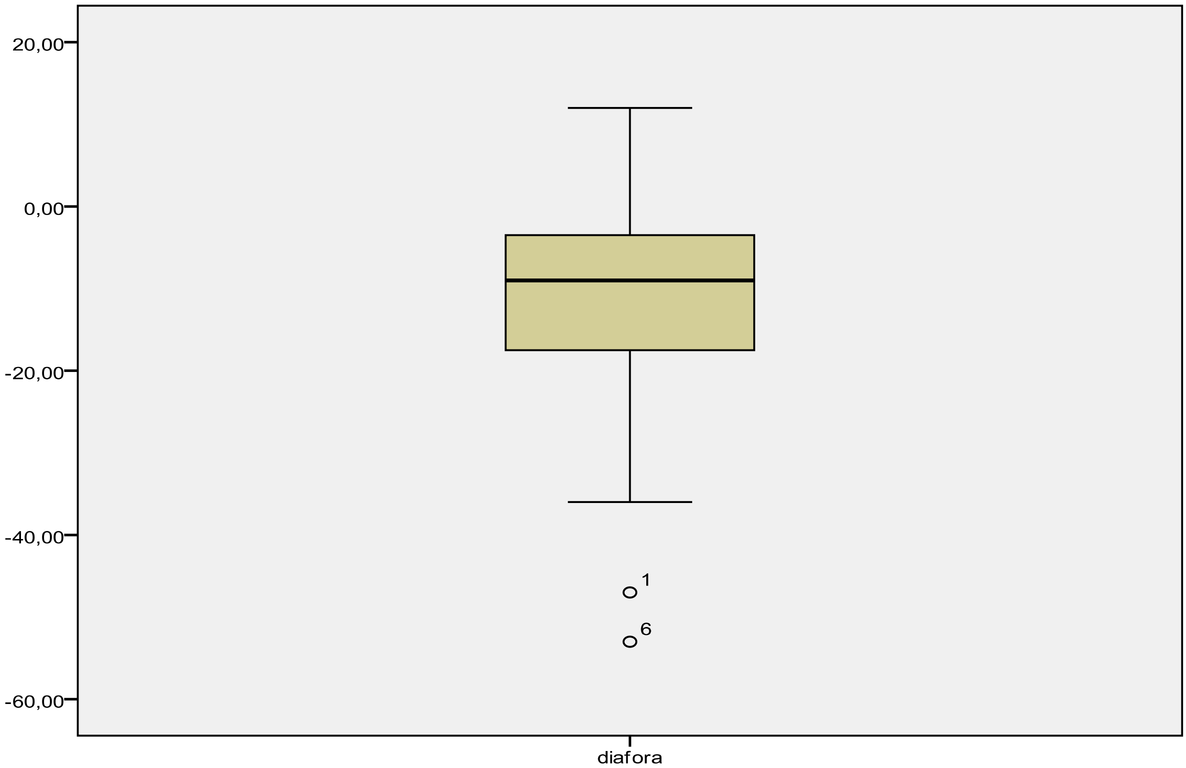
Η υλοποίηση αφήνεται ως άσκηση, ενώ δίνεται η αναφορά.

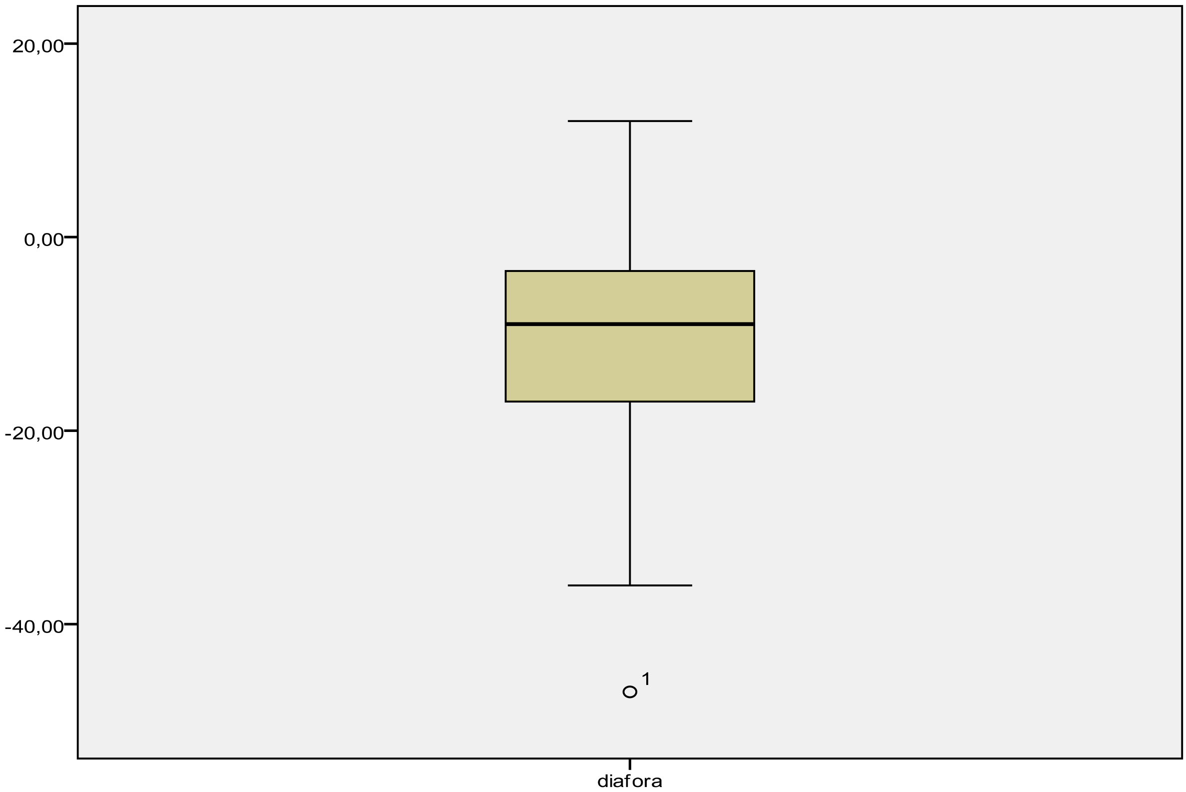
Αναφορά:

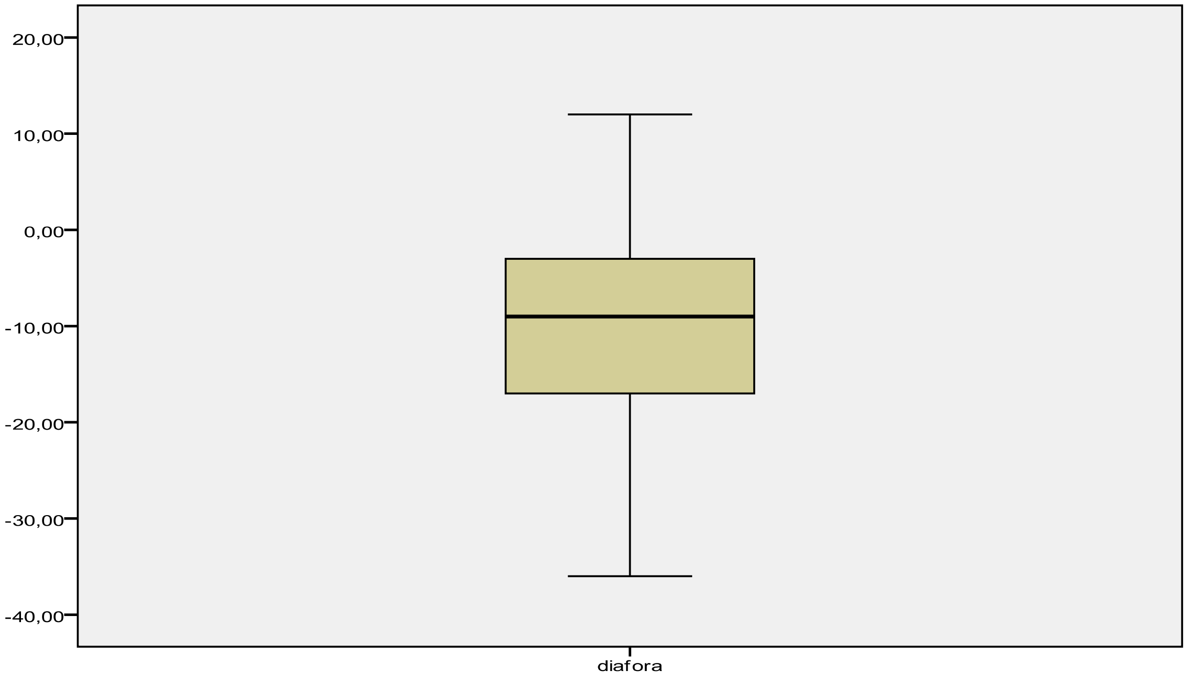
Στο πρόβλημα αυτό θέλουμε να εξετάσουμε την αποτελεσματικότητα της νέας μεθόδου διδασκαλίας σε σχέση με την υπάρχουσα μέθοδο. Το πρόβλημα μας είναι ένα πρόβλημα ελέγχου ισότητας των μέσων τιμών δυο πληθυσμών. Επειδή όμως οι μετρήσεις μας, και για τις δυο μεθόδους, γίνονται στις ίδιες πειραματικές μονάδες συμπεραίνουμε ότι τα δείγματα μας δεν είναι ανεξάρτητα. Το πείραμα αυτό είναι της μορφής ΠΡΙΝ-ΜΕΤΑ. Για να απαντήσουμε στο αρχικό μας ερώτημα θα πρέπει να σχηματίσουμε τις διαφορές π. χ. pretest-posttest και έτσι το πρόβλημά μας μετατρέπετε σε ένα πρόβλημα ελέγχου για την μέση τιμή ενός πληθυσμού. Πιο συγκεκριμένα για το αν μέση τιμή της διαφοράς στη βαθμολογία πριν και μετά τη μέθοδο διδασκαλίας είναι ίση με μηδέν (0). Για να κάνουμε χρήση του t-τεστ για έναν πληθυσμό θα πρέπει, για το δείγμα μας, να ικανοποιούνται οι επόμενες προϋποθέσεις:

1. Να είναι τυχαίο.
2. Να μην έχει ακραίες τιμές σε ποσοστό μεγαλύτερο του 10%.
3. Να προέρχεται από πληθυσμό που περιγράφεται ικανοποιητικά από την κανονική κατανομή..

Από το θηκόγραμμα προέκυψε ότι υπάρχουν δύο ακραίες παρατηρήσεις στις δειγματικές τιμές των διαφορών στην βαθμολογία πριν και μετά την εφαρμογή της μεθόδου διδασκαλίας οι παρατηρήσεις με αύξοντα αριθμό 6 και 1 με τιμές -53 και -47 αντίστοιχα (βλέπε θηκογράμματα 1,2,3). Οι παρατηρήσεις αυτές αποκλείονται από την περαιτέρω ανάλυση, επειδή ο συνολικός τους αριθμός δεν υπερβαίνει το 10% των παρατηρήσεων, 2/60*100%<10%).


20,00
0,00
-20,00
-40,00
-60,00
1
6
diafora


20,00
0,00
-20,00
-40,00
1
diafora



Στη συνέχεια ελέγχουμε αν οι 58 δειγματικές παρατηρήσεις της διαφοράς της βαθμολογίας προέρχονται από κανονικό πληθυσμό. Η κρίσιμη πιθανότητα του τεστ των Shapiro-Wilk είναι p=0,021. Αυτό σημαίνει ότι η υπόθεση της κανονικής κατανομής με επίπεδο σημαντικότητας 5% θα πρέπει να απορριφθεί, ενώ με επίπεδο σημαντικότητας 1% δεν μπορεί να απορριφθεί. Αποτέλεσμα αυτού ήταν να καθορίσουμε το επίπεδο σημαντικότητας στο 1% σε όσα έπονται και θα χρησιμοποιηθεί για τον έλεγχο της υπό μελέτης υπόθεσης ο παραμετρικός έλεγχος του t-τεστ.

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
diafora	,158	58	,001	,951	58	,021

a. Lilliefors Significance Correction

Σαν συμπέρασμα από την υλοποίηση του t-τεστ προέκυψε ότι: οι δύο μέθοδοι διδασκαλίας, η παλαιά και η καινούργια, διαφέρουν στατιστικά σημαντικά μεταξύ τους (p<0,001). Πιο συγκεκριμένα η μέση απόδοση της καινούργιας μεθόδου είναι κατά 10,5172 βαθμούς καλύτερη από την παλαιά. Ένα 99% διάστημα εμπιστοσύνης για την διαφορά protest-posttest είναι το (-14,0858 , -6,9487).

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
diafora	58	-10,5172	10,19845	1,33912

One-Sample Test

	Test Value = 0					
	t	df	Sig. (2-tailed)	Mean Difference	99% Confidence Interval of the Difference	
					Lower	Upper
diafora	-7,854	57	,000	-10,51724	-14,0858	-6,9487