ΠΑΝΕΠΙΣΤΗΜΙΟ ΙΩΑΝΝΙΝΩΝ
ΑΝΟΙΚΤΑ ΑΚΑΔΗΜΑΪΚΑ ΜΑΘΗΜΑΤΑ

Τίτλος Μαθήματος: Στατιστική Ανάλυση Δεδομένων

Ενότητα: Έλεγχος για τις παραμέτρους θέσης δύο πληθυσμών με εξαρτημένα δείγματα

Διδάσκων: Επίκ. Καθ. Απόστολος Μπατσίδης

Τμήμα: Μαθηματικών

Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο
ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ
Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης
ΕΣΠΑ 2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

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 $D_i = X_i - Y_i,\; i = 1,...,n$.

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:

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

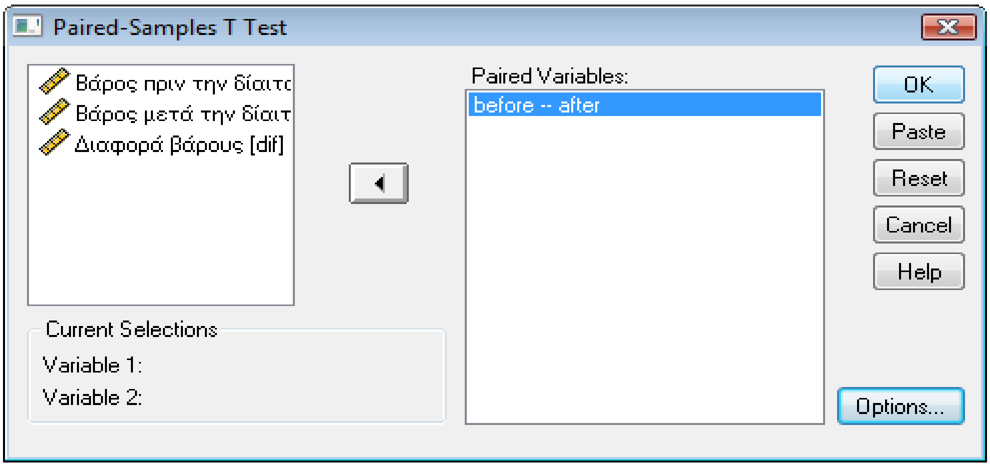
<u>Επισήμανση</u>: Σε περίπτωση που έχει χρησιμοποιηθεί κάποιος μετασχηματισμός διόρθωσης του προβλήματος είτε της ύπαρξης πολλών ακραίων τιμών είτε της μη κανονικότητας, τότε όλα τα παραπάνω αναφέρονται στις μετασχηματισμένες τιμές και στο τροποποιημένο σε μέγεθος δείγμα. Ειδικότερα, αν έχει χρησιμοποιηθεί 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.

Σχόλιο: 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. Μη παραμετρικός έλεγχος-Wilcoxon: Σύμφωνα με αυτόν θεωρούμε προς έλεγχο την ισοδύναμη μηδενική υπόθεση ότι οι διαφορές $D_i = X_i - Y_i$, $i = 1,...,n$, προέρχονται από μία συμμετρική περί το μηδέν κατανομή. Για την υλοποίηση του αρχικά τοποθετούμε τις διαφορές $D_i = X_i - Y_i$, $i = 1,...,n$, 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 $n \geq 8$, $Z = \dfrac{T - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}} \underset{H_0}{\overset{\pi\rho o\sigma.}{\sim}} N\left(0,1\right)$, 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.

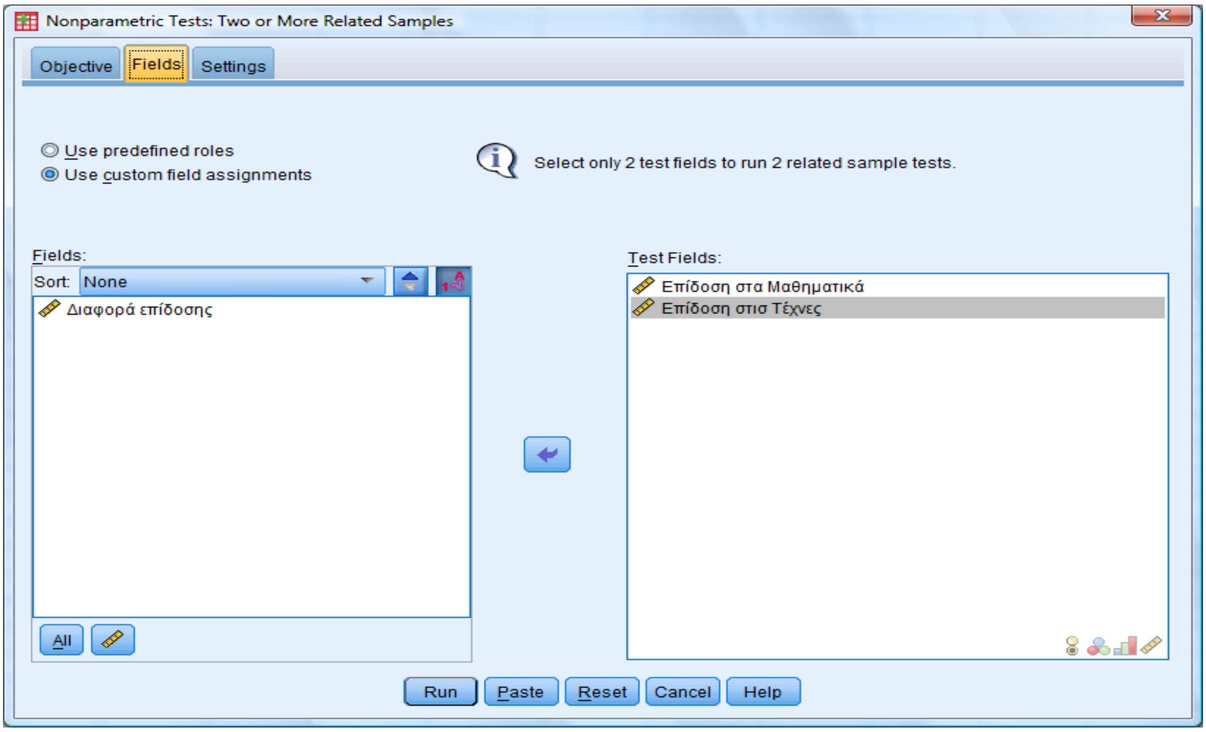
Σχόλιο: Για τον έλεγχο της υπό μελέτης μηδενικής υπόθεσης έχει προταθεί στη βιβλιογραφία και το προσημικό τεστ για συγκρίσεις κατά ζεύγη. Όμως, το τεστ του Wilcoxon είναι αποτελεσματικότερο για αυτό και δεν αναφέρθηκε το προσημικό τεστ.

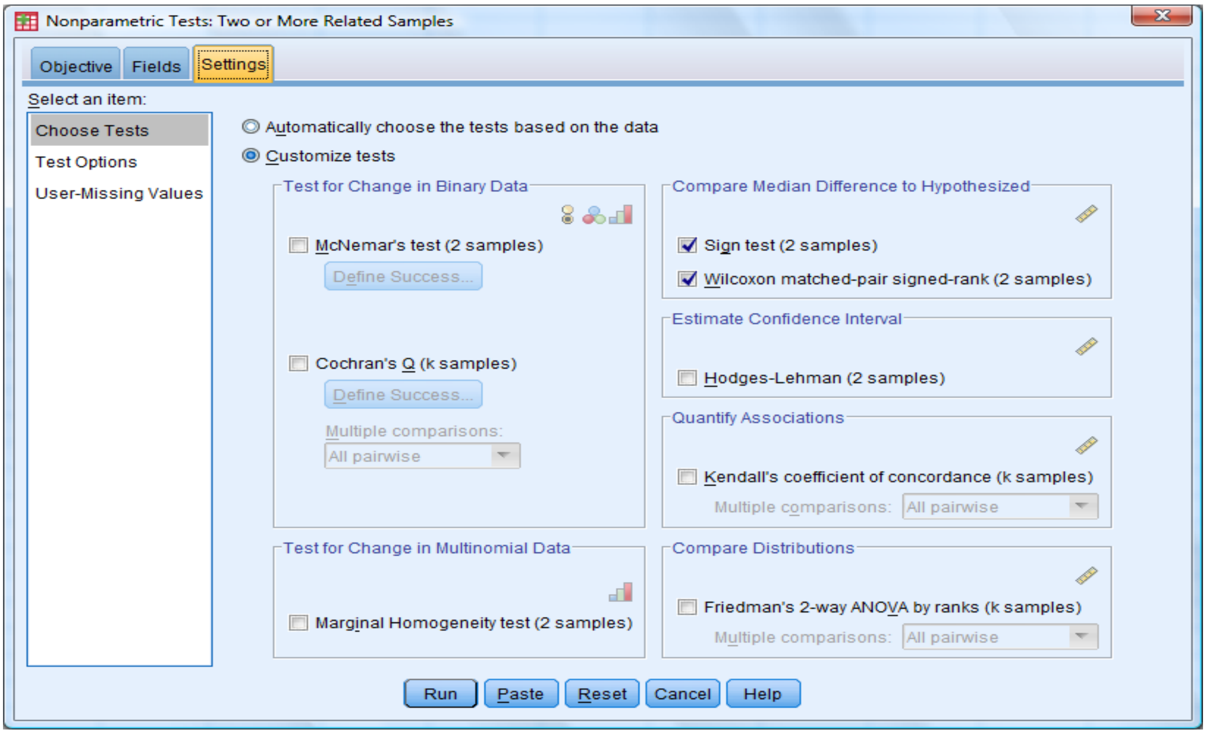
Υλοποίηση στο S.P.S.S.

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 1ο</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

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.

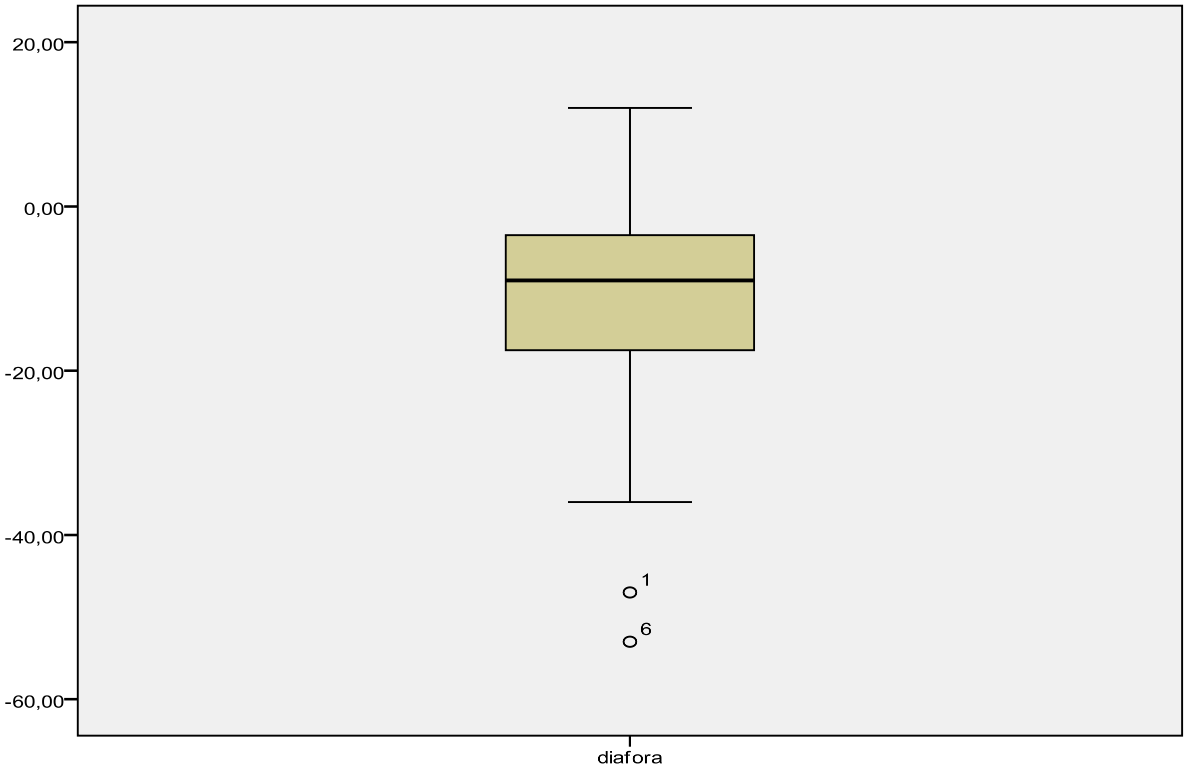
Η υλοποίηση αφήνεται ως άσκηση, ενώ δίνεται η αναφορά.

Αναφορά:

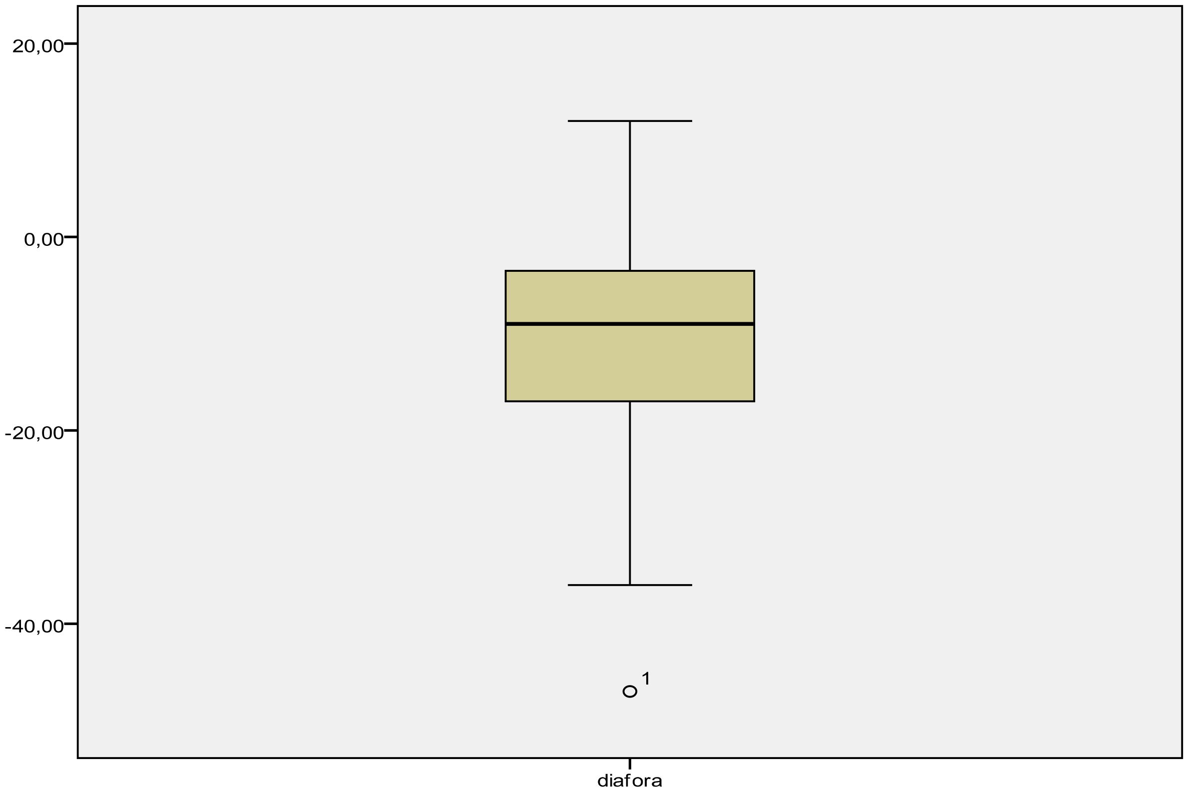
Στο πρόβλημα αυτό θέλουμε να εξετάσουμε την αποτελεσματικότητα της νέας μεθόδου διδασκαλίας σε σχέση με την υπάρχουσα μέθοδο. Το πρόβλημα μας είναι ένα πρόβλημα ελέγχου ισότητας των μέσων τιμών δυο πληθυσμών. Επειδή όμως οι μετρήσεις μας, και για τις δυο μεθόδους, γίνονται στις ίδιες πειραματικές μονάδες 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..

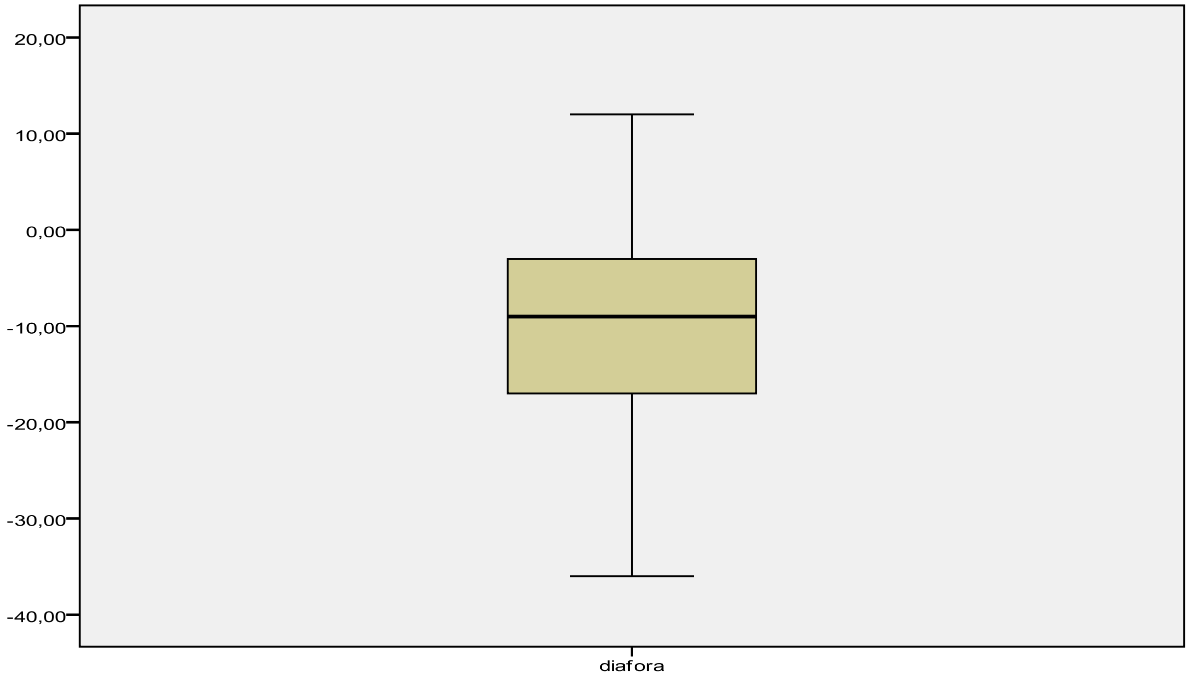
Από το θηκόγραμμα προέκυψε ότι υπάρχουν δύο ακραίες παρατηρήσεις στις δειγματικές τιμές των διαφορών στην βαθμολογία πριν και μετά την εφαρμογή της μεθόδου διδασκαλίας οι παρατηρήσεις με αύξοντα αριθμό 6 και 1 με τιμές -53 και -47 αντίστοιχα (βλέπε θηκογράμματα 1,2,3). Οι παρατηρήσεις αυτές αποκλείονται 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					
					99% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
diafora	-7,854	57	,000	-10,51724	-14,0858	-6,9487

Ανοικτά Ακαδημαϊκά Μαθήματα

Πανεπιστήμιο Ιωαννίνων

Τέλος Ενότητας

Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «**Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Ιωαννίνων**» έχει χρηματοδοτήσει μόνο τη αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.

Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο
ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ
Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης
ΕΣΠΑ 2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

Σημειώματα

Σημείωμα Αναφοράς

Σημείωμα Αδειοδότησης