ΚΕΦΑΛΑΙΟ ΤΕΤΑΡΤΟ

Έλεγχος ότι η παράμετρος θέσης ενός πληθυσμού είναι ίση με δοθείσα γνωστή τιμή

Έστω ένα τυχαίο δείγμα $X_1,\ldots,X_n$ μεγέθους n από έναν πληθυσμό με μέση τιμή μ και διακύμανση σ^2, άγνωστη. Ενδιαφερόμαστε για τον έλεγχο, σε επίπεδο σημαντικότητας α, της μηδενικής υπόθεσης

$$H_0 : \mu = \mu_0,$$

με μ_0 μια σταθερά, ως προς μία εκ των

$$H_a : \mu > \mu_0, \quad H_a : \mu < \mu_0,\ H_0 : \mu \neq \mu_0 .$$

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 η πληθυσμιακή διάμεσος είναι ίση με τη δοθείσα τιμή με μεθόδους της μη παραμετρικής στατιστικής. Τα αποτελέσματα του τελευταίου ελέγχου γενικεύονται για τον δοθέν έλεγχο όταν τα δεδομένα είναι συμμετρικά.

Σχόλιο: Επισημαίνεται ότι οι μη παραμετρικές μέθοδοι είναι λιγότερο ισχυρές στο να «ανακαλύπτουν» στατιστικά σημαντικές σχέσεις συγκριτικά με τις ανάλογες, αντίστοιχες παραμετρικές.

Υπενθύμιση: Σε κάθε στατιστικό έλεγχο αποφασίζουμε στη βάση ενός στατιστικού για την αποδοχή ή την απόρριψη μίας υπόθεσης (της μηδενικής υπόθεσης όπως λέγεται και η οποία συμβολίζεται με H_0). Επομένως υπάρχει ο «κίνδυνος» είτε ο στατιστικός να απορρίπτει την προς έλεγχο μηδενική υπόθεση (να αποδέχεται την λεγόμενη εναλλακτική υπόθεση H_a), ενώ η H_0 είναι αληθής είτε ο στατιστικός να αποδέχεται την H_0, ενώ η H_a είναι αληθής. Στην πρώτη περίπτωση έχουμε το λεγόμενο **σφάλμα τύπου I**, ενώ στη δεύτερη το **σφάλμα τύπου II**. Είναι τότε:

$$a = P(\text{σφάλμα τύπου I}) = P(\text{απορρίπτω } H_0 / H_0 \text{ αληθής})$$

και

$$\beta = P(\text{σφάλμα τύπου II}) = P(\text{αποδέχομαι } H_0 / H_\alpha \text{ αληθής}).$$

Το επιθυμητό θα ήταν να επιτυγχάνεται η ταυτόχρονη ελαχιστοποίηση των α και β. Όμως κάτι τέτοιο είναι αδύνατο. Το πρόβλημα αυτό παρακάμπτεται, προκαθορίζοντας το α και ελαχιστοποιώντας το β ή ισοδύναμα μεγιστοποιώντας την **ισχύ του τεστ** $\gamma = 1 - \beta = P(\text{απορρίπτω } H_0 / H_\alpha \text{ αληθής})$. Το προκαθορισμένο α είναι γνωστό και ως **επίπεδο σημαντικότητας** και συνήθως επιλέγεται να είναι είτε 5% είτε 1%.

Στα στατιστικά πακέτα η απόφαση για την αποδοχή ή απόρριψη της υπόθεσης δεν γίνεται εξετάζοντας αν η τιμή του στατιστικού ανήκει στην **περιοχή απόρριψης** (γνωστή και ως **κρίσιμη περιοχή**), αλλά στη βάση των p-τιμών (p-value ή Sig.) Η **p-τιμή** ενός στατιστικού τεστ είναι η μικρότερη τιμή του επιπέδου σημαντικότητας για την οποία απορρίπτεται η μηδενική υπόθεση. Εύκολα προκύπτει τότε ότι **απορρίπτουμε την προς έλεγχο μηδενική υπόθεση αν η p-τιμή είναι μικρότερη από το προκαθορισμένο επίπεδο σημαντικότητας (συνήθως 0.05).**

4.1 Μεθοδολογία- Υλοποίηση στο S.P.S.S.

Η μεθοδολογία που θα χρησιμοποιηθεί για τη στατιστική ανάλυση του πιο πάνω 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 δειγματικές παρατηρήσεις ξεπερνά το 10% αυτών, και

β) αν ο πληθυσμός από τον οποίο λαμβάνεται το τυχαίο δείγμα μπορούμε να ισχυριστούμε ότι περιγράφεται ικανοποιητικά από την κανονική κατανομή.

Ανάλογα με τα αποτελέσματα των παραπάνω ελέγχων προβαίνουμε στον παραμετρικό έλεγχο του t test ή στο μη παραμετρικό έλεγχο (προσημικό στατιστικό τεστ). Ο μη παραμετρικός έλεγχος ουσιαστικά ελέγχει αν η πληθυσμιακή διάμεσος είναι ίση με την προκαθορισμένη τιμή και τα αποτελέσματα όπως θα δούμε στη συνέχεια μπορούν να γενικευθούν υπό κάποια προϋπόθεση για τον ζητούμενο έλεγχο.

Από τα παραπάνω ίσως έγινε ήδη αντιληπτό ότι κομβικό σημείο για τον τρόπο διεξαγωγής του υπό μελέτη ελέγχου αποτελεί η διενέργεια των προκαταρκτικών ελέγχων α) και β), με βάση τα αποτελέσματα των οποίων θα αποφανθούμε αν θα προχωρήσουμε παραμετρικά ή μη παραμετρικά. 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. Αν το ποσοστό των ακραίων τιμών, που προκύπτει αφαιρώντας μία ακραία κάθε φορά με τη βοήθεια του θηκογράμματος, δε ξεπερνά το 10%, τότε προχωρούμε στο βήμα 2. Αν το ποσοστό των ακραίων τιμών ξεπερνά το 10%, τότε, αφού συμπεριλάβουμε τις δειγματικές παρατηρήσεις που προηγούμενα έχουν αποκλειστεί από την ανάλυση, 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 (είτε οι αρχικές είτε οι μετασχηματισμένες του βήματος 1) 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 του t τεστ (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 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, εννοείται χωρίς να προκύπτει πρόβλημα ύπαρξης ακραίων τιμών, συνεχίζουμε την ανάλυση παραμετρικά (βήμα 3). Σε αντίθετη περίπτωση, αν το πλήθος των δειγματικών παρατηρήσεων (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 η υπόθεση της κανονικότητας απορρίπτεται τόσο για τις αρχικές όσο και για τις μετασχηματισμένες δειγματικές τιμές (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 t τεστ</u>: Οι κρίσιμες περιοχές μεγέθους α για τον έλεγχο της μηδενικής υπόθεσης

$$H_0 : \mu = \mu_0,$$

με μ_0 μια σταθερά, ως προς τις εναλλακτικές

$$H_a : \mu > \mu_0, \quad H_a : \mu < \mu_0, \; H_a : \mu \neq \mu_0,$$

είναι αντίστοιχα

$$t \geq t_{n-1,a} \, , \; t \leq -t_{n-1,a} \, , \; |t| \geq t_{n-1,a/2} \; \left(t \geq t_{n-1,a/2} \text{ ή } t \leq -t_{n-1,a/2} \right),$$

όπου $t = \dfrac{\overline{X} - \mu_0}{s/\sqrt{n}} \overset{H_0}{\sim} t_{n-1}$.

Επιπλέον το 100(1-α)% Δ.Ε. για τη μέση τιμή είναι

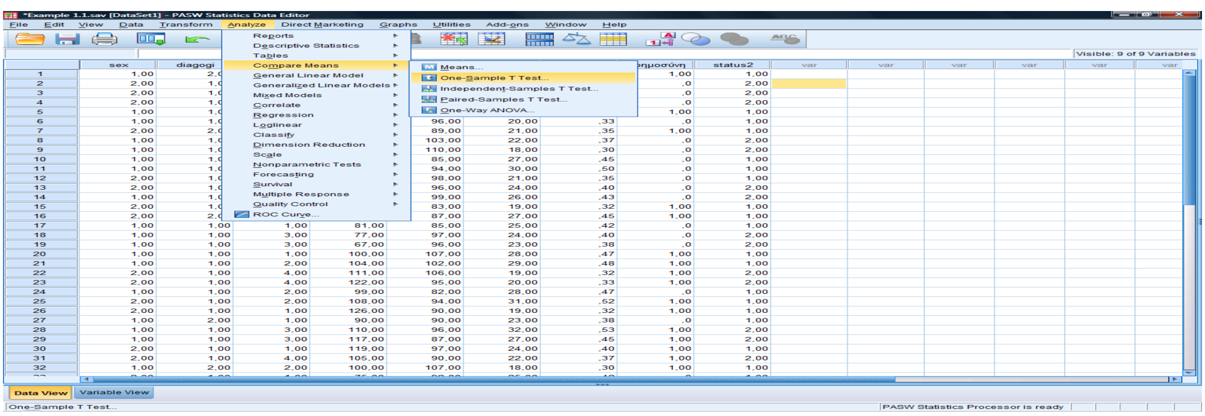
$$\left(\overline{X} - t_{n-1,a/2} \frac{S}{\sqrt{n}}, \overline{X} + t_{n-1,a/2} \frac{S}{\sqrt{n}} \right).$$

<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 δε διαφέρει στατιστικά σημαντικά από το λογάριθμο της δοθείσας γνωστής τιμής.

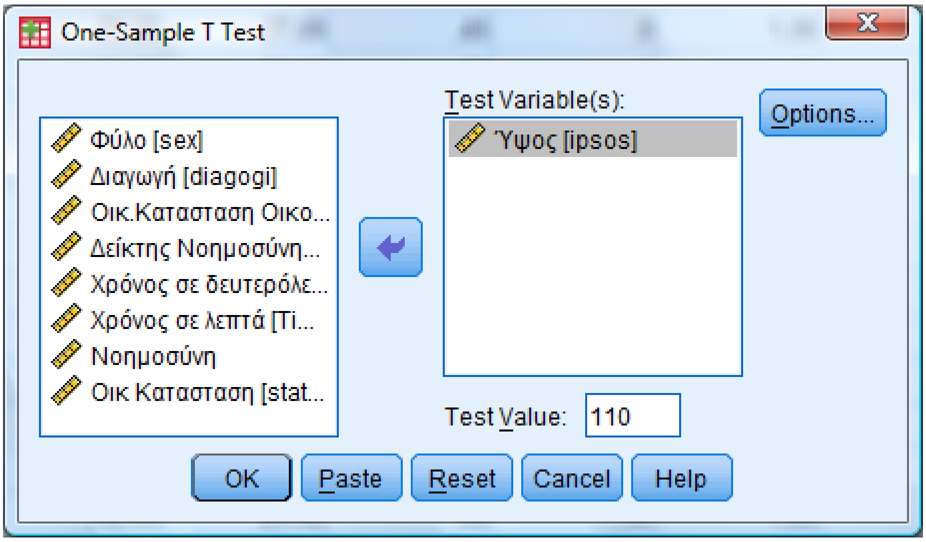
<u>Υλοποίηση στο S.P.S.S.</u>

Ο έλεγχος αυτός υλοποιείται ως εξής:

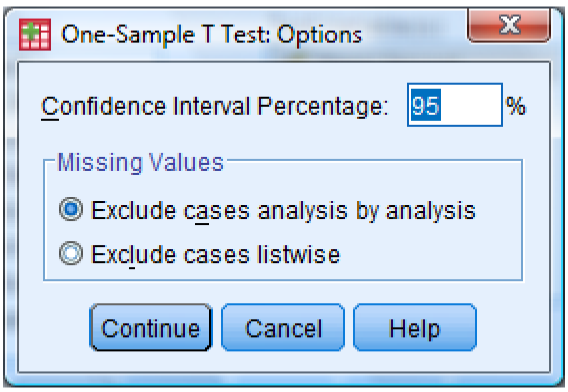
i. Analyze→Compare Means→One-Sample T Test.



ii. Στο παράθυρο διαλόγου που προκύπτει μετακινούμε στο πλαίσιο Test Variable τη μεταβλητή (στήλη του αρχείου δεδομένων, έστω το Ύψος) που καταγράφονται οι δειγματικές τιμές της υπό μελέτη μεταβλητής. Στο πλαίσιο Test Value εισάγουμε την τιμή ως προς την οποία θέλουμε να γίνει ο έλεγχος (έστω 110 cm). Προσοχή: Αν έχει χρησιμοποιηθεί κατάλληλος μετασχηματισμός διόρθωσης του προβλήματος των ακραίων τιμών ή της μη κανονικότητας θα πρέπει να τροποποιείται κατάλληλα και η αρχική τιμή προς έλεγχο.



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 διαφορά της καθορισμένης τιμής από την πληθυσμιακή μέση τιμή, δηλαδή για την ποσότητα $\mu - \mu_0$. Έτσι για παράδειγμα ζητάμε τον υπολογισμό του 95% διαστήματος εμπιστοσύνης για το $\mu - \mu_0$.



<u>Ερμηνεία αποτελεσμάτων του S.P.S.S.</u>

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Ύψος	35	94,8571	7,18308	1,21416

One-Sample Test

	Test Value = 110					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
Ύψος	-12,472	34	,000	-15,14286	-17,6103	-12,6754

Από τον πίνακα One-Sample Statistics πληροφορούμαστε ότι 35 παρατηρήσεις είναι διαθέσιμες στη μεταβλητή Ύψος. Δηλαδή, έχει καταγραφεί το ύψος 35 παιδιών. Το μέσο ύψος αυτών των παιδιών είναι 94,8571 εκατοστά, η δειγματική τυπική απόκλιση του ύψους είναι S =7.18308 εκατοστά και το τυπικό σφάλμα για τη μέση τιμή είναι $\frac{S}{\sqrt{n}}$=1.21416.

Από τον δεύτερο πίνακα (One-Sample Test) το λογισμικό αρχικά στο πλαίσιο Test Value μας υπενθυμίζει ότι η τιμή ελέγχου είναι ίση με 110 (Test Value =110). Επιπλέον μας δίνει την τιμή του t στατιστικού για τον έλεγχο της υπόθεσης ότι το πληθυσμιακό μέσο ύψος είναι ίσο με 110 εκατοστά. Επιπλέον προκύπτει ότι το μέσο ύψος των παιδιών είναι στατιστικά σημαντικά διαφορετικό από 110 εκατοστά ($t =$-12.472, p-τιμή<0.001). Μάλιστα το μέσο ύψος των παιδιών είναι στατιστικά σημαντικά μικρότερο από 110 εκατοστά (καθώς από τη στήλη Mean Difference παρατηρούμε ότι η μέση διαφορά (Μέσο ύψος -110) είναι ίση με -15.14286 εκατοστά). Τέλος, το 95% διάστημα εμπιστοσύνης για τη διαφορά $\mu - 110$ είναι το (-17.6103, -12.6754) και επομένως ένα 95% διάστημα εμπιστοσύνης για το μέσο ύψος μ είναι το (92.3897,97.3246).

4. <u>Τεστ του Wilcoxon (Wilcoxon Signed Rank Test) </u>: Ο έλεγχος της μηδενικής υπόθεσης

$$H_0 : \mu = \mu_0,$$

με μ_0 μια σταθερά, ως προς μία εκ των $H_a : \mu > \mu_0$, $H_a : \mu < \mu_0$, $H_a : \mu \neq \mu_0$, <u>υπό την προϋπόθεση ότι ο πληθυσμός είναι συμμετρικός</u> γύρω από το μ_0, ανάγεται σε έλεγχο της διαμέσου και μεταξύ άλλων έχει προταθεί ο προσημικός έλεγχος καθώς και το τεστ του Wilcoxon για ένα δείγμα (για περισσότερες πληροφορίες βλέπε Μπατσίδης (2010), Παπαϊωάννου και Λουκάς (2002)).

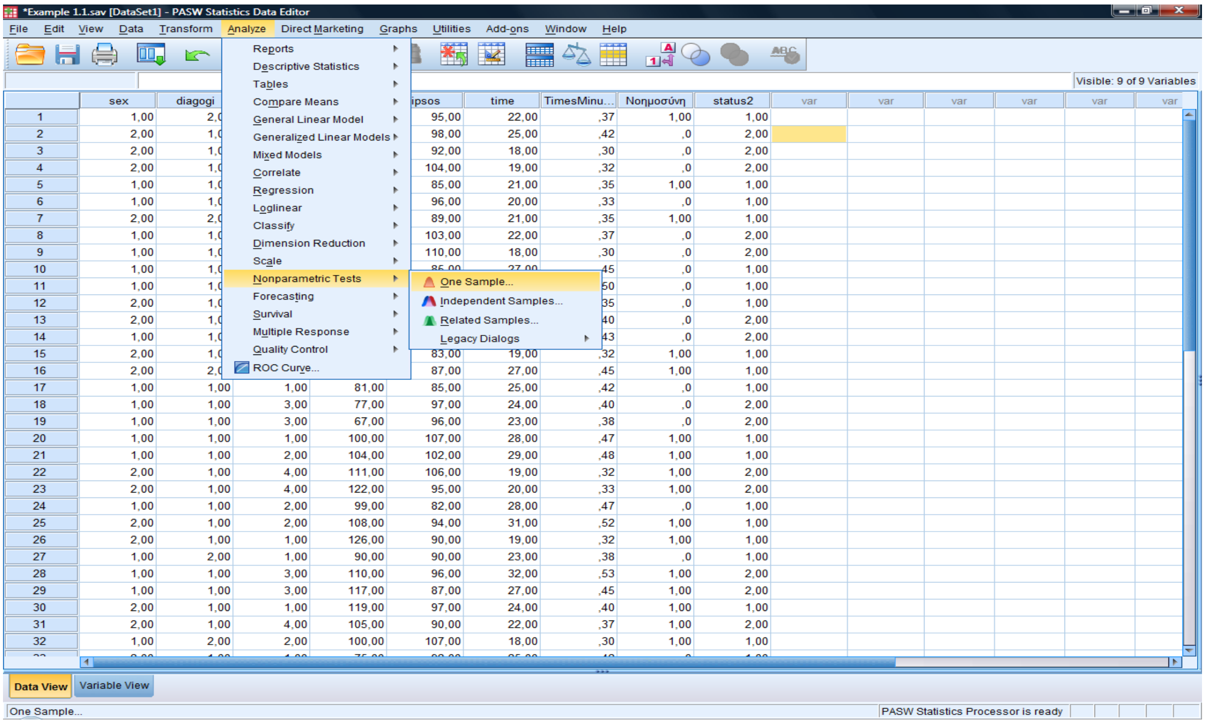
Επιγραμματικά αναφέρουμε ότι αν $D_i = X_i - m_0$, $i = 1,...,n$, είναι οι μη μηδενικές διαφορές και $d_1,...,d_c$, ο αριθμός των παρατηρήσεων σε καθεμία από τις c διαφορετικές απόλυτες διαφορές (σε αύξουσα τάξη μεγέθους), με $d_i \geq 1$, και $\sum_{i=1}^{c} d_i = \frac{n(n+1)}{2}$ τότε η προσεγγιστική κατανομή του στατιστικού T^+ που παριστάνει το άθροισμα των τάξεων που αντιστοιχούν στις θετικές διαφορές, υπό τη μηδενική υπόθεση, είναι

$$W = \frac{T^+ - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24} - \sum_{i=1}^{c} \frac{d_i(d_i^2 - 1)}{48}}} \overset{\pi\rho o\sigma}{\underset{H_0}{\sim}} N(0,1)$$

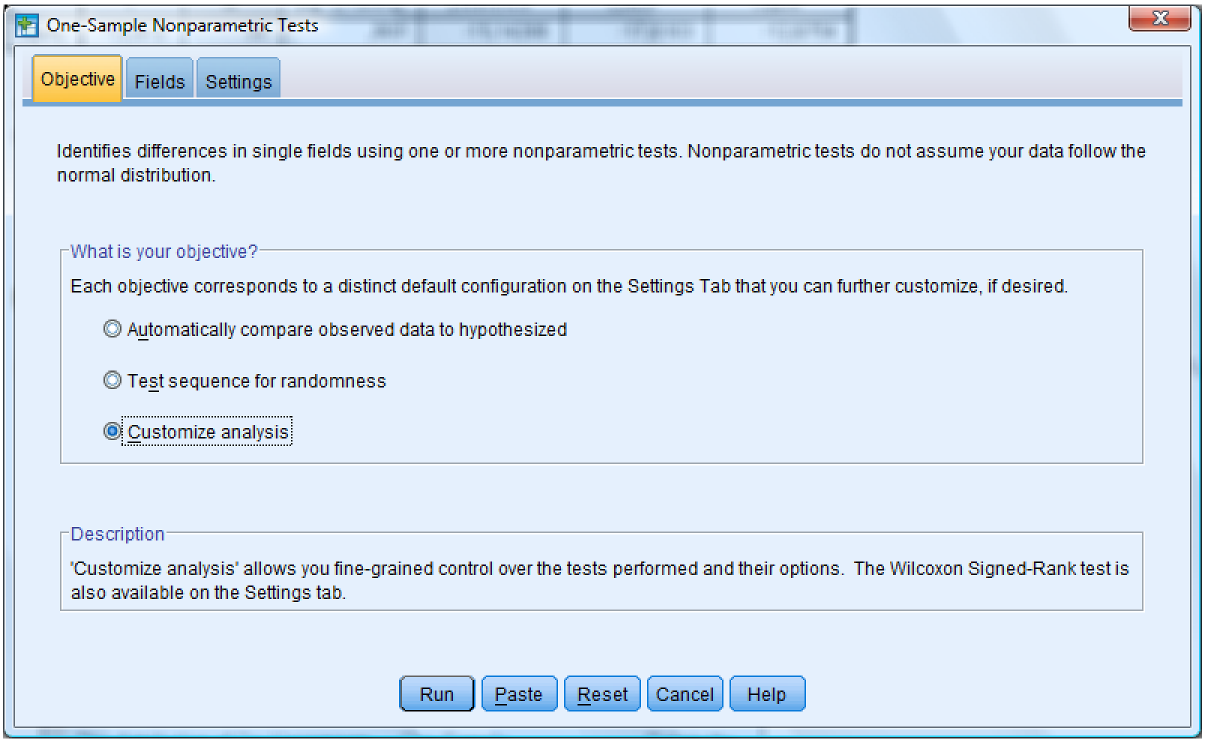
και αυτή η στατιστική συνάρτηση χρησιμοποιείται για τον έλεγχο της υπό μελέτης μηδενικής υπόθεσης.

Ο έλεγχος αυτός υλοποιείται ως εξής:

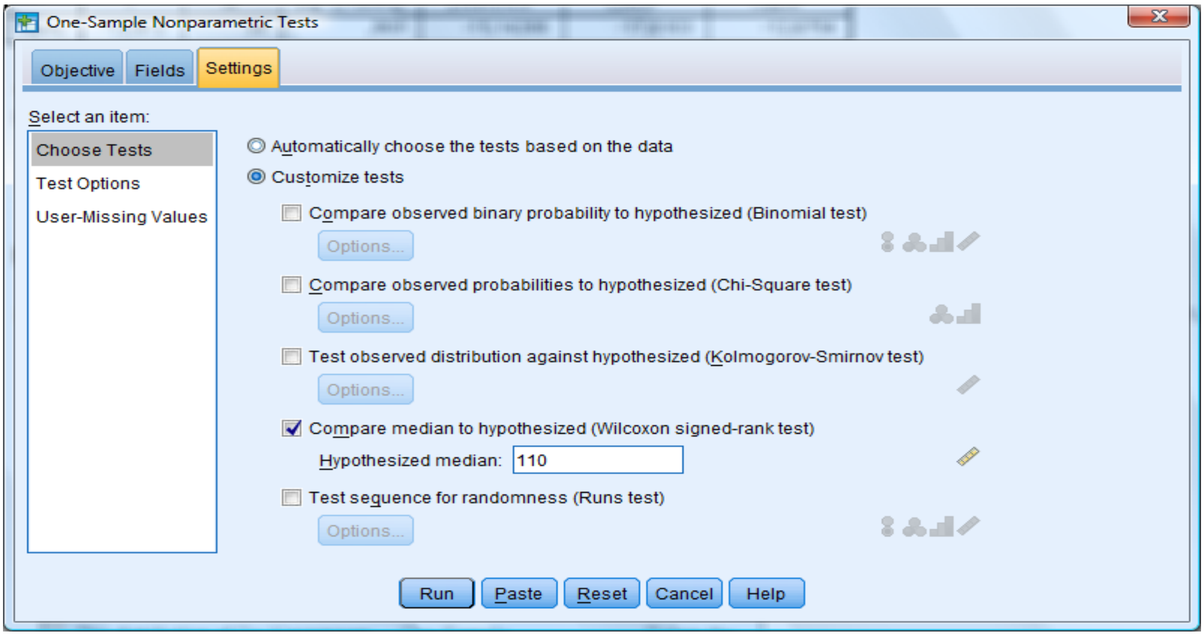
i. Analyze→Non Parametric Tests→One Sample.



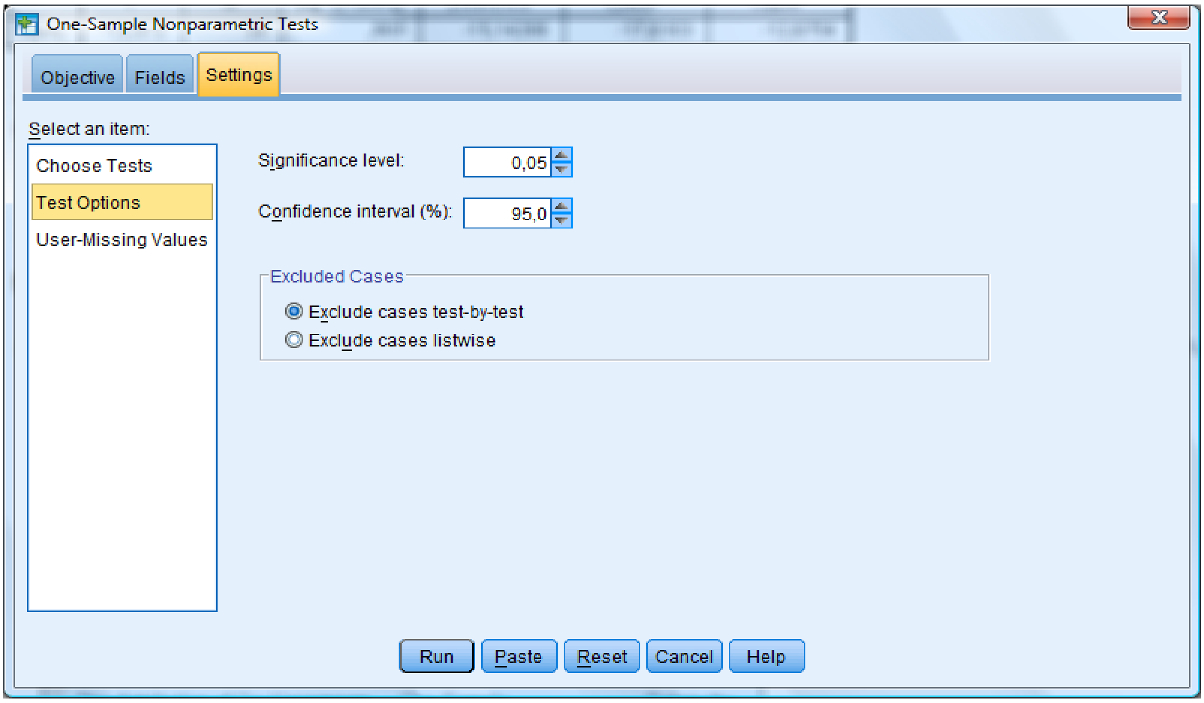
ii. Στο παράθυρο διαλόγου που προκύπτει επιλέγουμε στο πλαίσιο Objective την επιλογή Customize analysis, έτσι ώστε στη συνέχεια από τα πλαίσια Fields και Settings να καθορίσουμε τον έλεγχο τον οποίο θέλουμε να διενεργηθεί.



Στο πλαίσιο Settings και έχοντας ενεργοποιημένη την επιλογή Choose Tests επιλέγουμε τον κατάλληλο μη παραμετρικό έλεγχο, δηλαδή στη συγκεκριμένη περίπτωση την επιλογή Compare median to hypothesized (Wilcoxon signed-rank test) τοποθετώντας στο πλαίσιο Hypothesized median την προκαθορισμένη τιμή π.χ. 110.



Από την επιλογή Test Options μας δίνεται μεταξύ άλλων η δυνατότητα καθορισμού του επιπέδου σημαντικότητας και του βαθμού εμπιστοσύνης των Δ.Ε. Επιλέγοντας Run διενεργούνται οι επιλογές που ζητήσαμε.



<u>Ερμηνεία αποτελεσμάτων του S.P.S.S.</u>

Η υπόθεση ότι η πληθυσμιακή διάμεσος του ύψους ισούται με 110 cm απορρίπτεται καθώς η p-τιμή του One-Sample Wilcoxon Signed Ranks test είναι μικρότερη από 0.001. Το αποτέλεσμα αυτό για να μπορεί να γενικευθεί στην πληθυσμιακή μέση τιμή του ύψους θα πρέπει η δειγματική μέση τιμή να είναι κοντά στην δειγματική διάμεσο, κάτι που μπορεί να ελέγξει κάποιος μέσω της διαδικασίας Descriptive Statistics Explore.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The median of Φύλο equals 110.	One-Sample Wilcoxon Signed Ranks Test	.000	Reject the null hypothesis.
2	The median of Διαγωγή equals 110.	One-Sample Wilcoxon Signed Ranks Test	.000	Reject the null hypothesis.
3	The median of Οικ.Κατασταση Οικογένειας equals 110.	One-Sample Wilcoxon Signed Ranks Test	.000	Reject the null hypothesis.
4	The median of Δείκτης Νοημοσύνης equals 110.	One-Sample Wilcoxon Signed Ranks Test	.000	Reject the null hypothesis.
5	The median of Χρόνος σε δευτερόλεπτα equals 110.	One-Sample Wilcoxon Signed Ranks Test	.000	Reject the null hypothesis.
6	The median of Χρόνος σε λεπτά equals 110.	One-Sample Wilcoxon Signed Ranks Test	.000	Reject the null hypothesis.
7	The median of Νοημοσύνη equals 110.	One-Sample Wilcoxon Signed Ranks Test	.000	Reject the null hypothesis.
8	The median of Οικ Κατασταση equals 110.	One-Sample Wilcoxon Signed Ranks Test	.000	Reject the null hypothesis.
9	The median of Ύψος equals 110.	One-Sample Wilcoxon Signed Ranks Test	.000	Reject the null hypothesis.

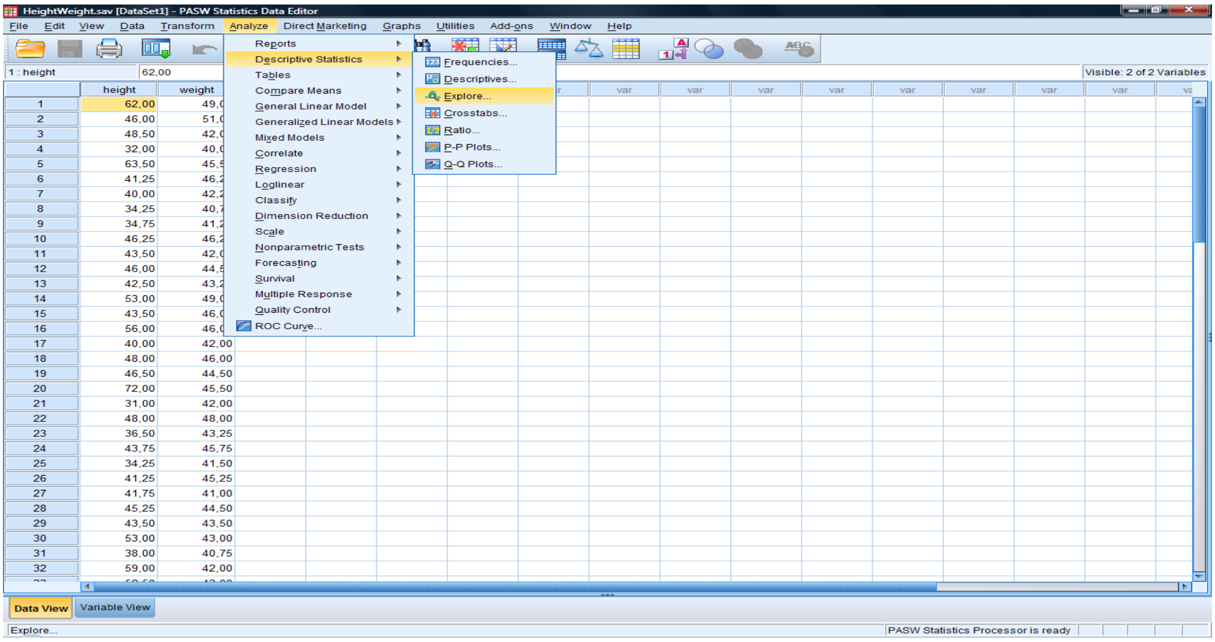
Asymptotic significances are displayed. The significance level is .05.

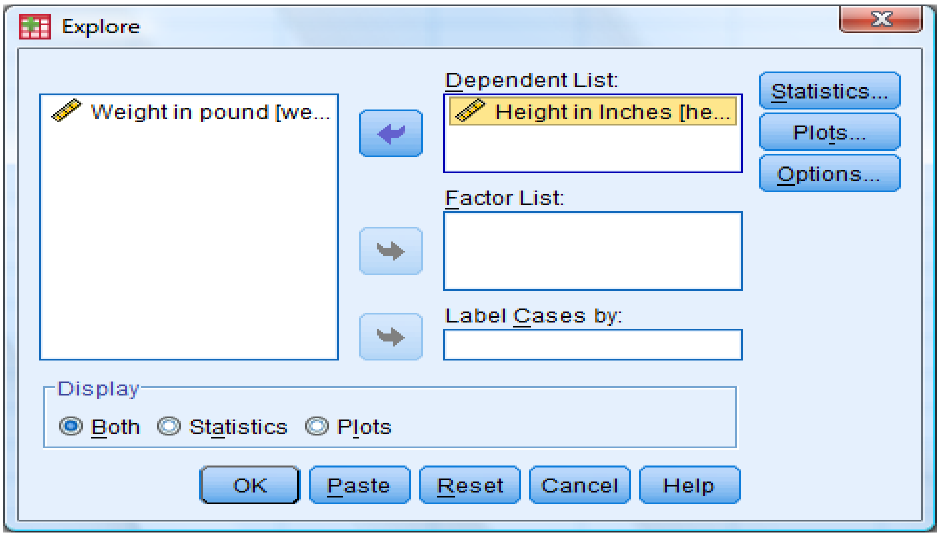
4.2 Παραδείγματα

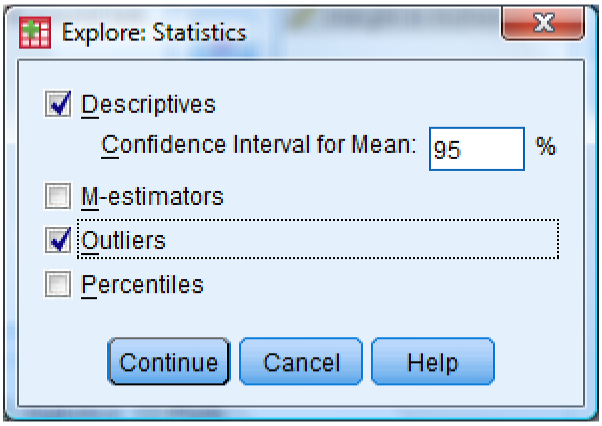
<u>Παράδειγμα 1ο</u> Στο αρχείο HeightWeight.sav* καταγράφονται οι τιμές του βάρους και του ύψους (σε ίντσες) 73 τυχαία επιλεγμένων ατόμων από έναν πληθυσμό. Θέλουμε να ελέγξουμε, αν είναι εφικτό, αν το μέσο ύψος του πληθυσμού είναι στατιστικά σημαντικά διαφορετικό από τις 60 inches.

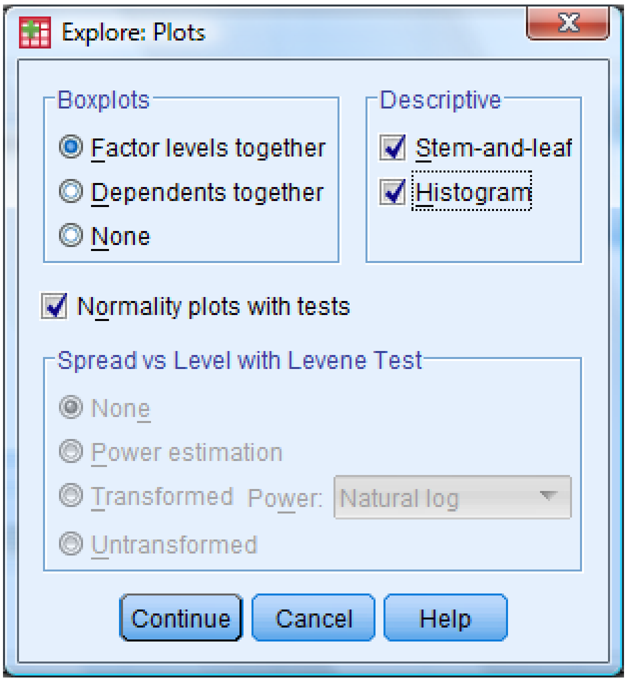
Υλοποίηση: Η υλοποίηση θα γίνει ακολουθώντας τη μεθοδολογία της παραγράφου 4.1.

Έλεγχος ακραίων τιμών. Ο έλεγχος για την ύπαρξη ακραίων τιμών γίνεται με το θηκόγραμμα. Παρατήρηση: ταυτόχρονα ζητάμε και στατιστικούς και γραφικούς τρόπους ελέγχου της κανονικότητας έτσι ώστε να μην επανερχόμαστε σε περίπτωση που δεν υπάρχει πρόβλημα ακραίων τιμών. Ακολουθούμε τα επόμενα βήματα μέσω της διαδικασίας Analyze Descriptive Statistics Explore.

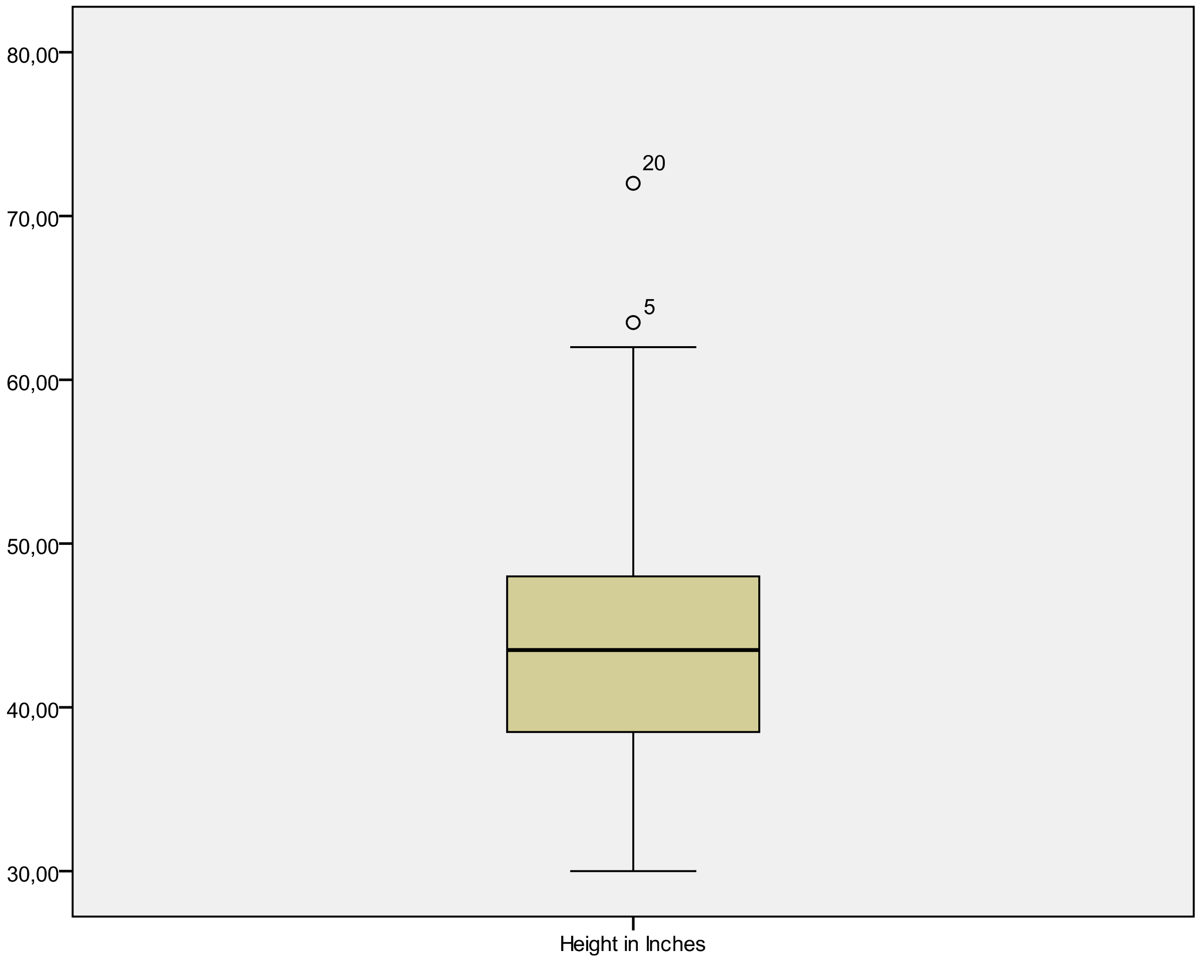




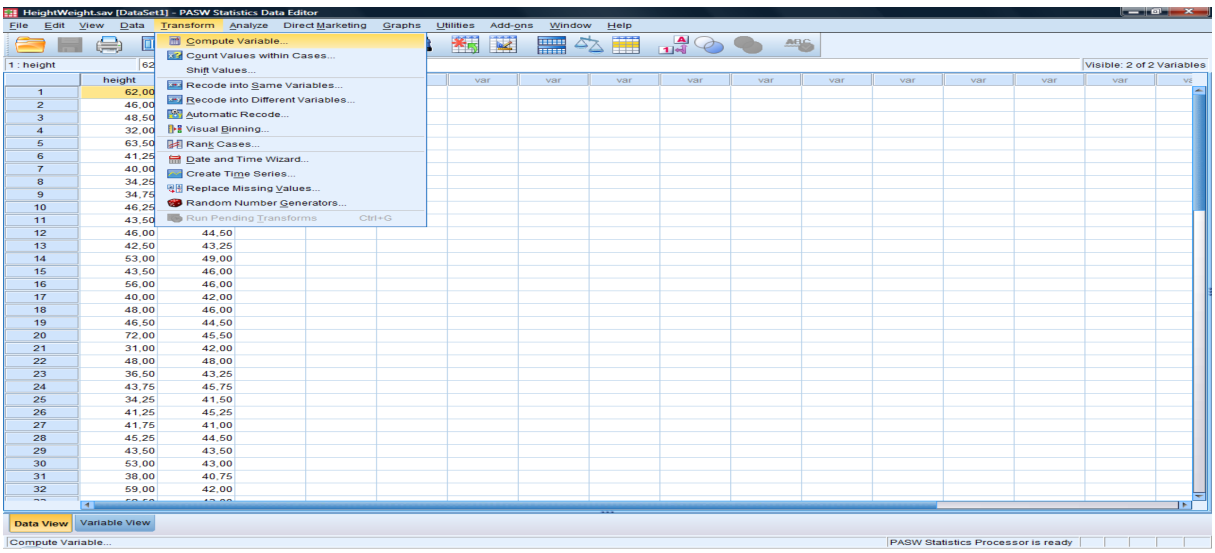


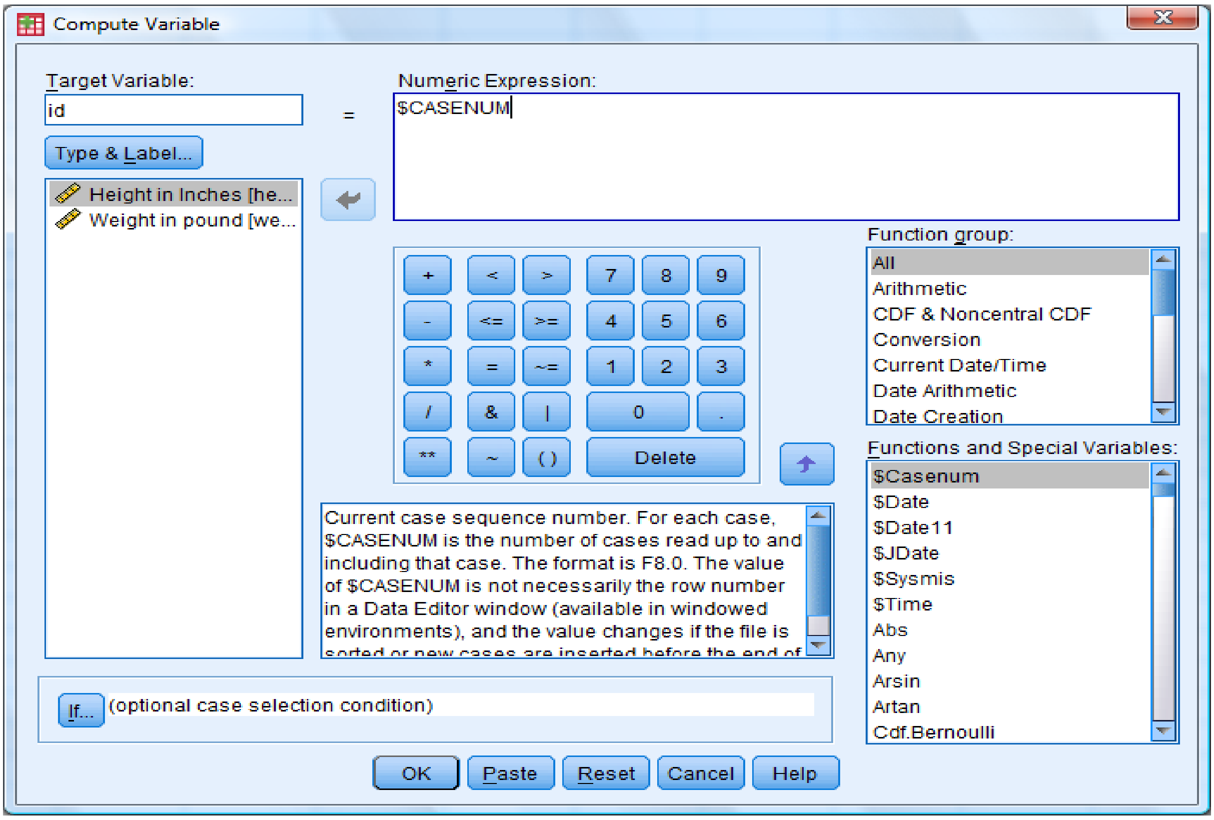


Προκύπτει μεταξύ άλλων το ακόλουθο θηκόγραμμα:

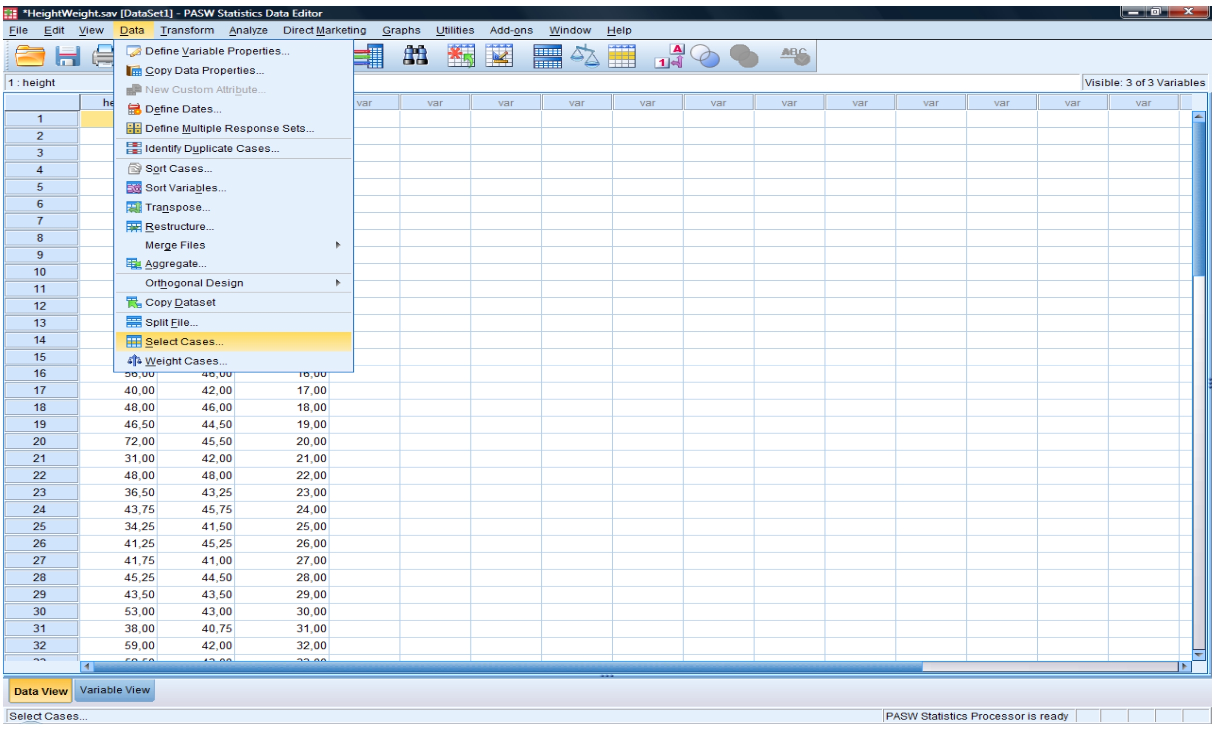


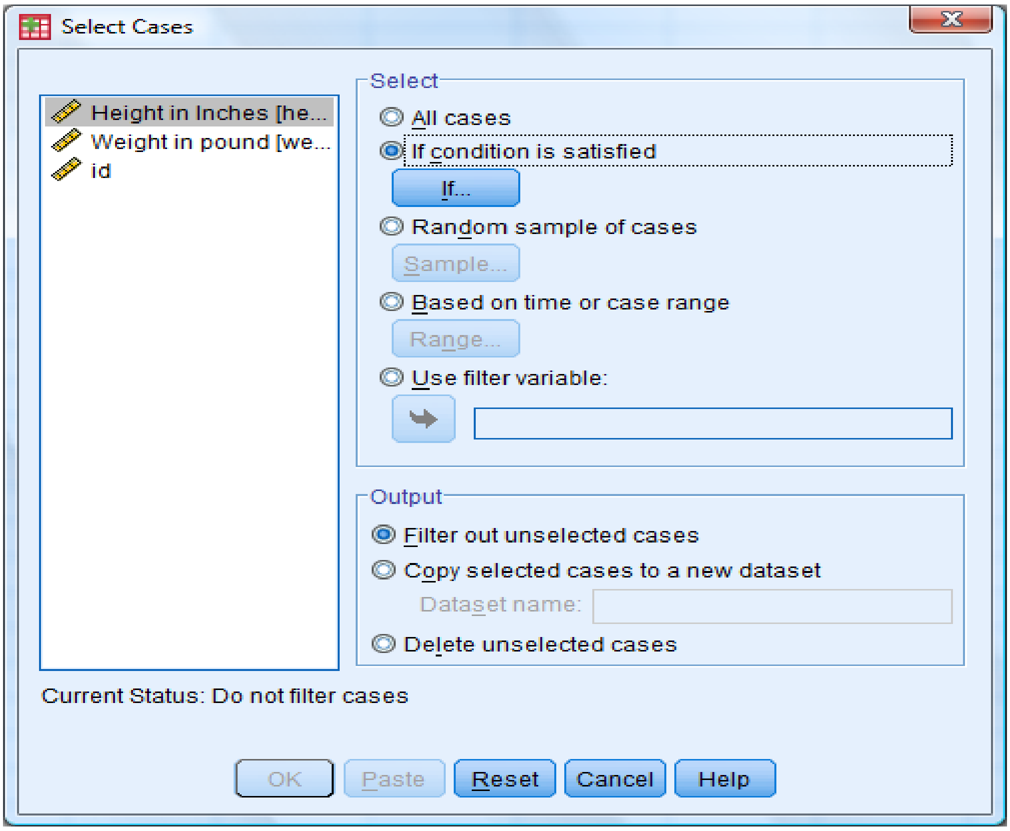
Στο σχήμα αυτό βλέπουμε ότι έχουμε μία τουλάχιστον ακραία τιμή, την παρατήρηση 20 την οποία και θα αποκλείσουμε από την περαιτέρω ανάλυση. Δημιουργούμε για αυτό το σκοπό μία νέα στήλη, μέσω της διαδικασίας Transform Compute Variable, που θα μας δίνει τον αύξοντα αριθμό των παρατηρήσεων,.

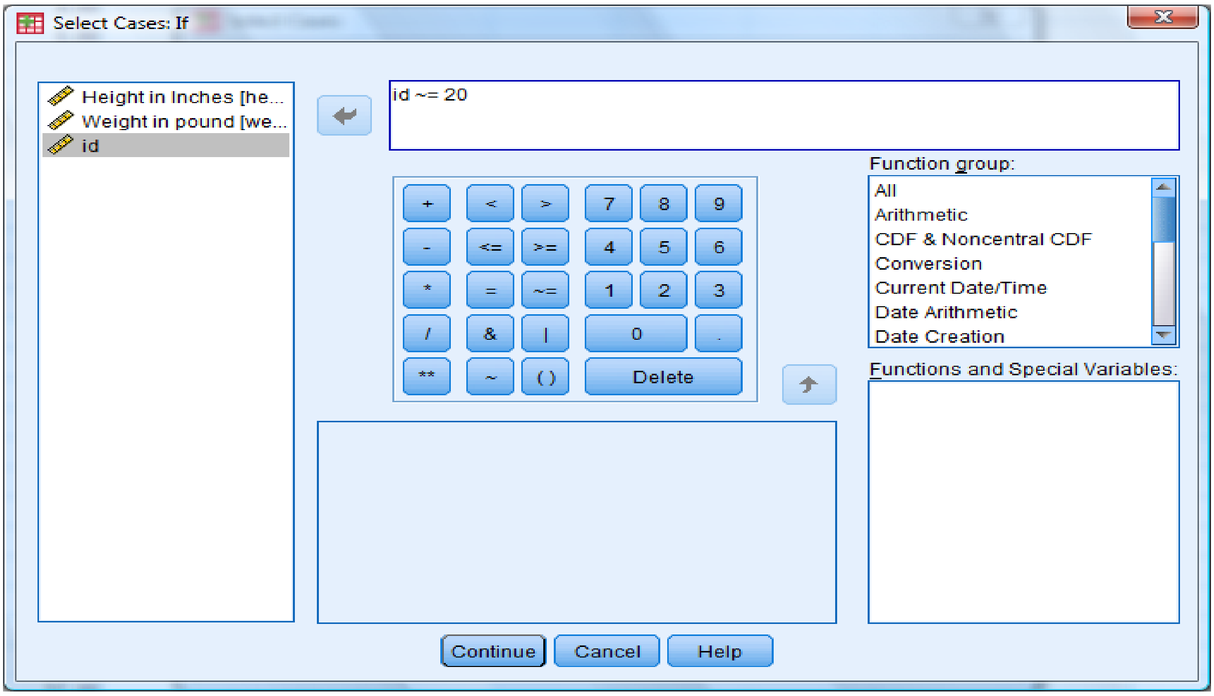




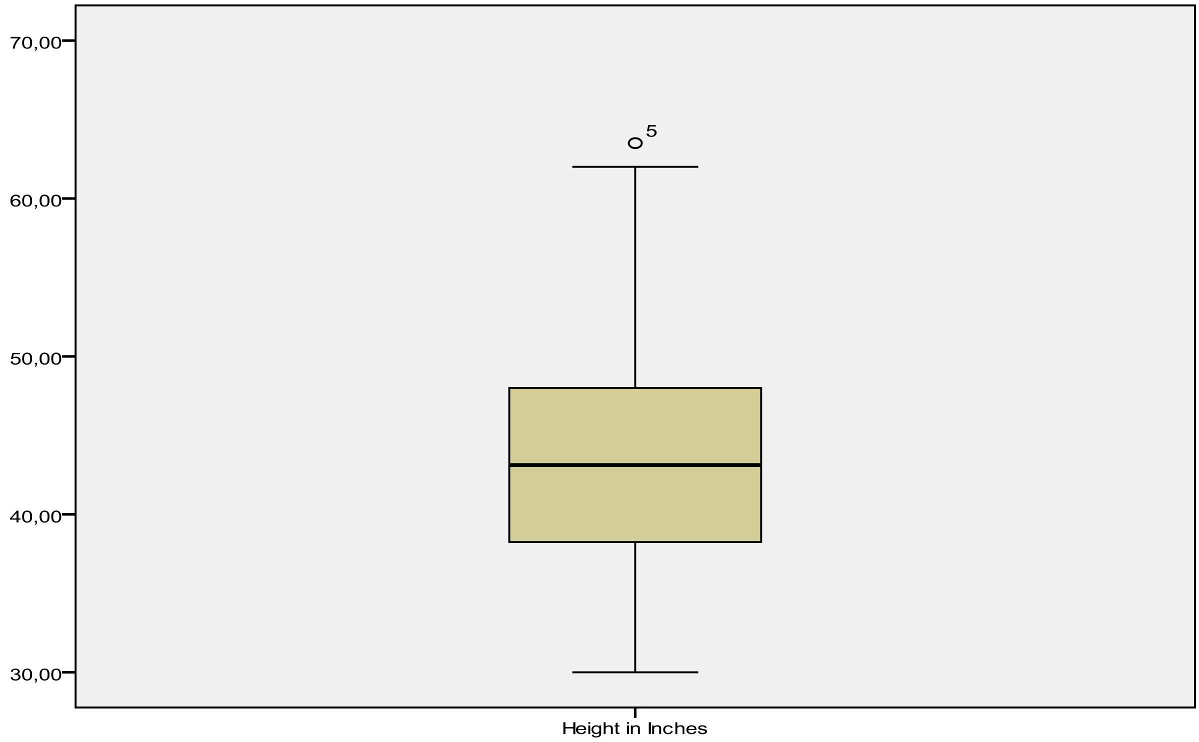
Με τη βοήθεια της στήλης αυτής αποκλείουμε από την περαιτέρω ανάλυση την παρατήρηση 20 μέσω της διαδικασίας Data → Select Cases ως εξής:





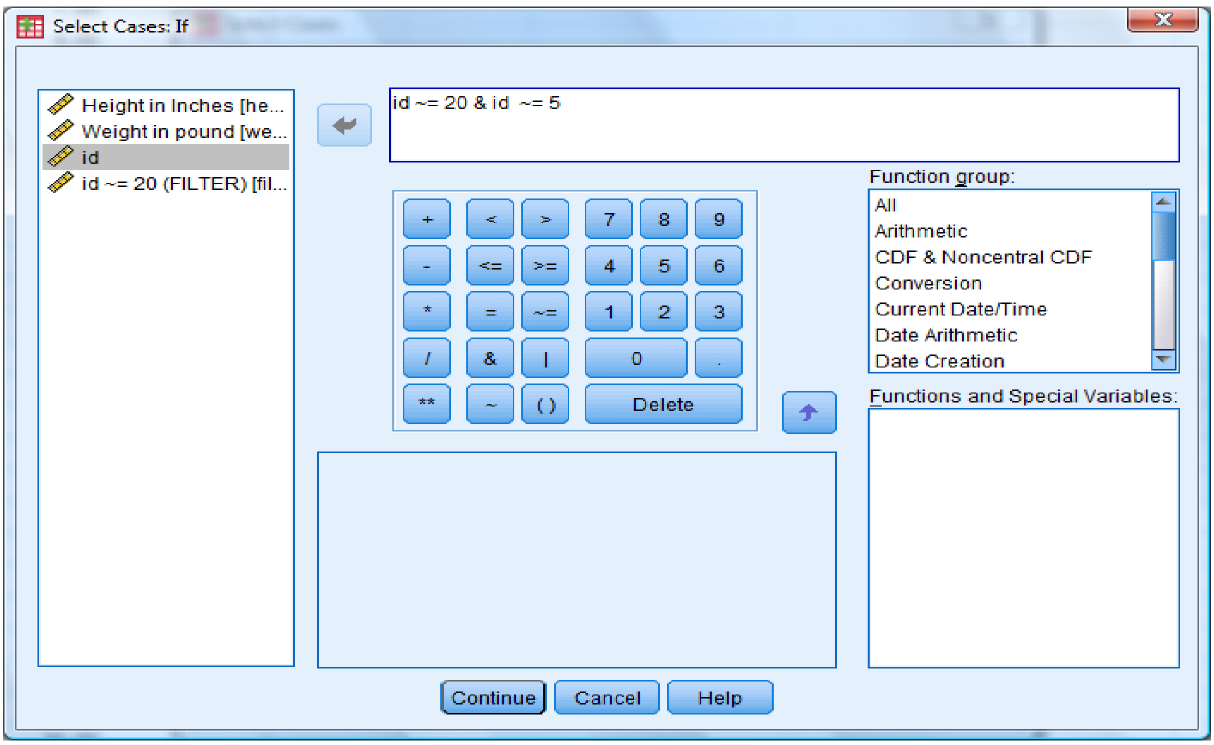


Έπειτα επαναλαμβάνεται ο έλεγχος των ακραίων τιμών και προκύπτει το ακόλουθο θηκόγραμμα:



Επομένως καθώς είναι ακραία τιμή πρέπει να αποκλειστεί και η παρατήρηση 5 ($2^{η}$ ακραία σε σύνολο 73 παρατηρήσεων, ποσοστό μικρότερο του 10%) από την περαιτέρω ανάλυση.

Επιλέγουμε μέσω της διαδικασίας Data → Select Cases τα ακόλουθα:



Έπειτα επαναλαμβάνεται η διαδικασία ελέγχου ακραίων τιμών.

Στην περίπτωση αυτή βλέπουμε ότι δεν υπάρχουν άλλες ακραίες τιμές.

Έλεγχος κανονικής κατανομής: Το αποτέλεσμα αυτού του ελέγχου είναι διαθέσιμο ήδη αφού στο προηγούμενο βήμα ζητάμε ταυτόχρονα και γραφικούς και στατιστικούς τρόπους ελέγχου της υπόθεσης της κανονικότητας.

Tests of Normality

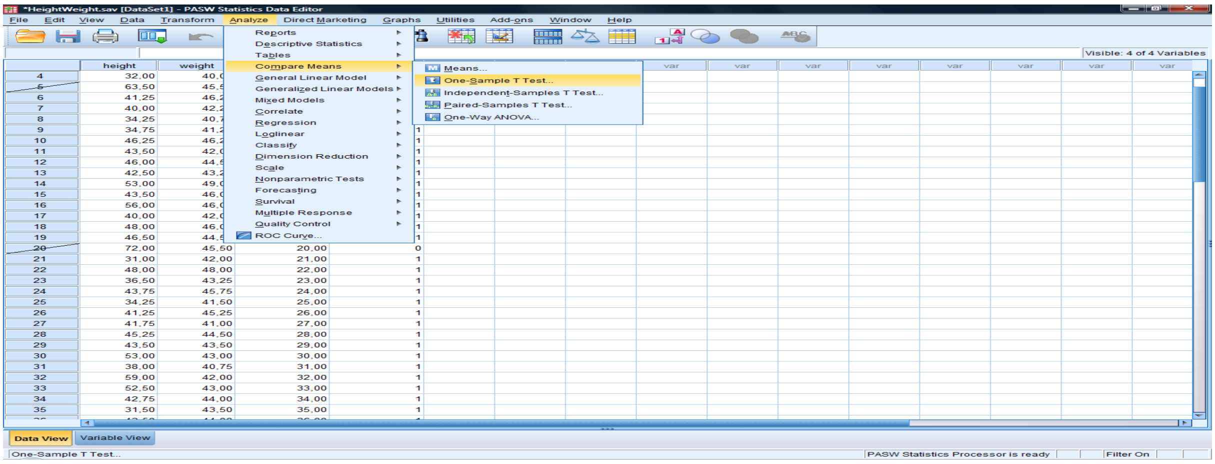
	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Height in Inches	,078	71	,200*	,969	71	,079

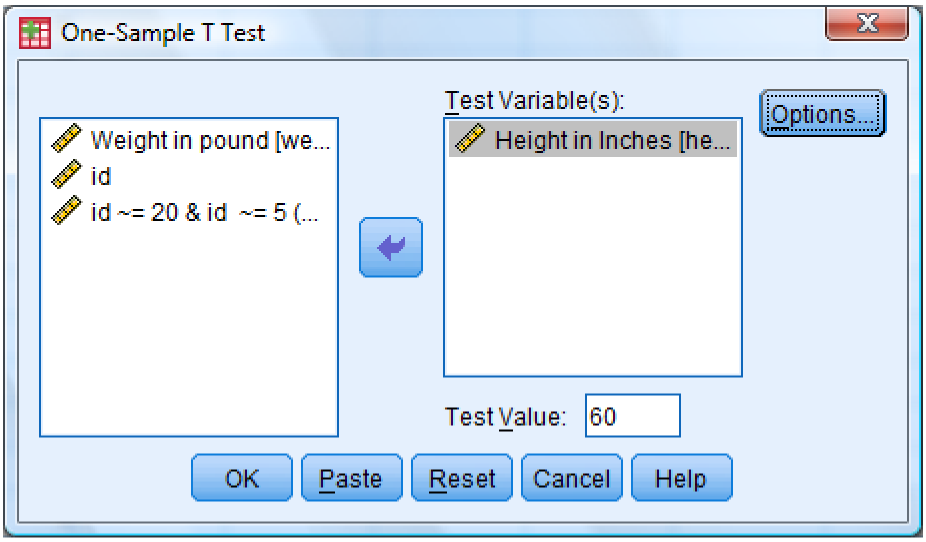
a. Lilliefors Significance Correction

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

Η τιμή που κοιτάζουμε στον πίνακα αυτό, είναι η η κρίσιμη πιθανότητα, (p-value) στη στήλη Sig του Shapiro Wilk. Επειδή η τιμή είναι μεγαλύτερη του 5% (δηλ. του α=0,05), λέμε ότι η υπόθεση ότι οι δειγματικές τιμές του ύψους προέρχονται από έναν πληθυσμό που περιγράφεται ικανοποιητικά από την κανονική κατανομή δεν μπορεί να απορριφθεί. Επομένως συμπεραίνουμε ότι θα ελέγξουμε την υπόθεση ότι το μέσο ύψος του πληθυσμού είναι ίσο με 60 ίντσες παραμετρικά.

<u>Παραμετρικός έλεγχος t-Test</u> Ο έλεγχος διεξάγεται μέσω της διαδικασίας Analyze→Compare Means→One-Sample T Test όπως αναλυτικά περιγράφηκε στην παράγραφο 4.1.





Στο παράθυρο των αποτελεσμάτων προκύπτουν τα ακόλουθα αποτελέσματα:

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Height in Inches	71	43,2359	7,36278	,87380

One-Sample Test

	Test Value = 60					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
Height in Inches	-19,185	70	,000	-16,76408	-18,5068	-15,0213

Κοιτάζουμε την κρίσιμη πιθανότητα, Sig. (2-tailed). Στην περίπτωσή μας p<0,001. Άρα η τιμή αυτή είναι μικρότερη από το 5% (0,000<0,05), συνεπώς το μέσο ύψος είναι στατιστικά σημαντικά διαφορετικό από τις 60 inches και καθώς το πρόσημο της μέσης διαφοράς του μέσου ύψους από τις 60 ίντσες είναι αρνητικό (-16,764) συμπεραίνουμε ότι το μέσο ύψος του πληθυσμού είναι στατιστικά σημαντικά μικρότερο από την τιμή των 60 inches.

Από την παραπάνω ανάλυση προκύπτει η ακόλουθη αναφορά:

Αναφορά Θέλουμε να ελέγξουμε αν η μέση τιμή του ύψους των ατόμων του πληθυσμού από τον οποίο επιλέξαμε το δείγμα μας ισούται με 60 ίντσες. Το πρόβλημα αυτό είναι ένας έλεγχος για τη μέση τιμή ενός πληθυσμού. Θα ελέγξουμε αρχικά αν ικανοποιούνται οι παρακάτω υποθέσεις χρήσης του παραμετρικού αυτού ελέγχου.

1. Το δείγμα μας είναι τυχαίο.
2. Δεν υπάρχουν ακραίες τιμές στα δεδομένα μας που ξεπερνούν σε ποσοστό το 10%.

3. Τα δεδομένα μας ακολουθούν κανονική κατανομή.

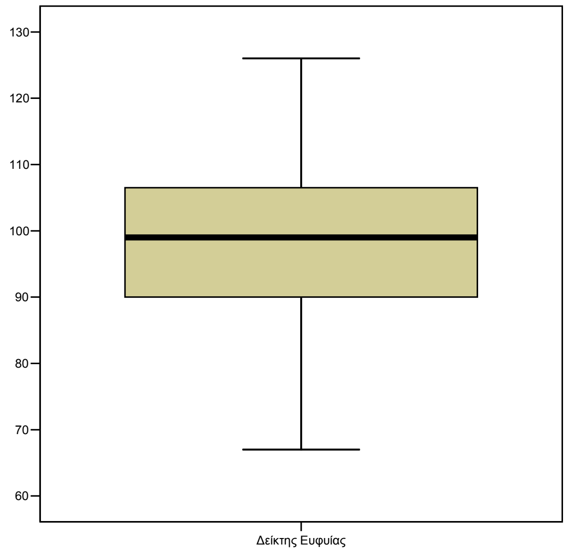
Η πρώτη από τις προϋποθέσεις σχετίζεται με τον τρόπο που επιλέξαμε το δείγμα μας και ικανοποιείται.

Ο έλεγχος των ακραίων τιμών έγινε με το θηκόγραμμα και έδειξε ότι υπάρχουν 2 ακραίες τιμές οι παρατηρήσεις με αύξοντα αριθμό 20 και 5 και τιμές του ύψους 72 και 63,5 ίντσες αντίστοιχα (βλέπε θηκογράμματα 1,2,3). Καθώς το ποσοστό των ακραίων τιμών (2/73*100%) δεν υπερβαίνει το 10% συνεχίζουμε την περαιτέρω ανάλυση έχοντας αποκλείσει τις δύο αυτές παρατηρήσεις. Από το τεστ των Shapiro-Wilk έχουμε ότι η υπόθεση ότι οι δειγματικές τιμές του ύψους προέρχονται από έναν πληθυσμό που περιγράφεται ικανοποιητικά από την υπόθεση της κανονικής κατανομής δεν μπορεί να απορριφθεί (τιμή του τεστ 0.961, β.ε. 71, p=0,079).

Εφόσον ικανοποιούνται όλες οι προϋποθέσεις, μπορούμε να κάνουμε χρήση του παραμετρικού ελέγχου t-Test για τον έλεγχο της υπόθεσης ότι το μέσο ύψος του πληθυσμού είναι ίσο με 60 ίντσες. Από τον έλεγχο αυτό προκύπτει ότι το μέσο ύψος είναι στατιστικά σημαντικά διαφορετικό από τις 60 inches ($p<0,001$) και καθώς το πρόσημο της μέσης διαφοράς του μέσου ύψους από τις 60 ίντσες είναι αρνητικό (-16,764) συμπεραίνουμε ότι το μέσο ύψος του πληθυσμού είναι στατιστικά σημαντικά μικρότερο από την τιμή των 60 inches. Επιπλέον ένα 95% Δ.Ε. για το μέσο ύψος του πληθυσμού είναι το (60-18.5068,60-15.0213).

<u>Παράδειγμα 2ο</u> Στο αρχείο GeneralExample.sav[*] καταγράφονται μεταξύ άλλων ο δείκτης ευφυΐας 35 ατόμων τυχαία επιλεγμένων από τον υπό μελέτη πληθυσμό. Θέλουμε να ελέγξουμε, αν είναι εφικτό, αν ο μέσος δείκτης ευφυΐας του πληθυσμού είναι στατιστικά σημαντικά διαφορετικός από τις 100 μονάδες.

Η υλοποίηση είναι ανάλογη του προηγούμενου παραδείγματος και παραλείπεται. Στη συνέχεια θα δοθεί η αναφορά των ακόλουθων αποτελεσμάτων



Tests of Normality

	Kolmogorov-Smirnov(a)			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
iq	,087	35	,200(*)	,986	35	,932

* This is a lower bound of the true significance.
a Lilliefors Significance Correction

One-Sample Test

	Test Value = 100					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
iq	-1,250	34	,220	-3,057	-8,03	1,91

Αναφορά: Θέλουμε να ελέγξουμε αν η μέση τιμή του δείκτη ευφυΐας για τον πληθυσμό από τον οποίο επιλέξαμε το δείγμα μας ισούται με 100. Το πρόβλημα αυτό είναι ένας έλεγχος για τη μέση τιμή ενός πληθυσμού. Θα ελέγξουμε αρχικά αν ικανοποιούνται οι παρακάτω υποθέσεις χρήσης του παραμετρικού αυτού ελέγχου.

1. Το δείγμα μας είναι τυχαίο.
2. Δεν υπάρχουν ακραίες τιμές στα δεδομένα μας που ξεπερνούν σε ποσοστό το 10%.
3. Τα δεδομένα μας ακολουθούν κανονική κατανομή.

Η πρώτη από τις προϋποθέσεις σχετίζεται με τον τρόπο που επιλέξαμε το δείγμα μας και ικανοποιείται.

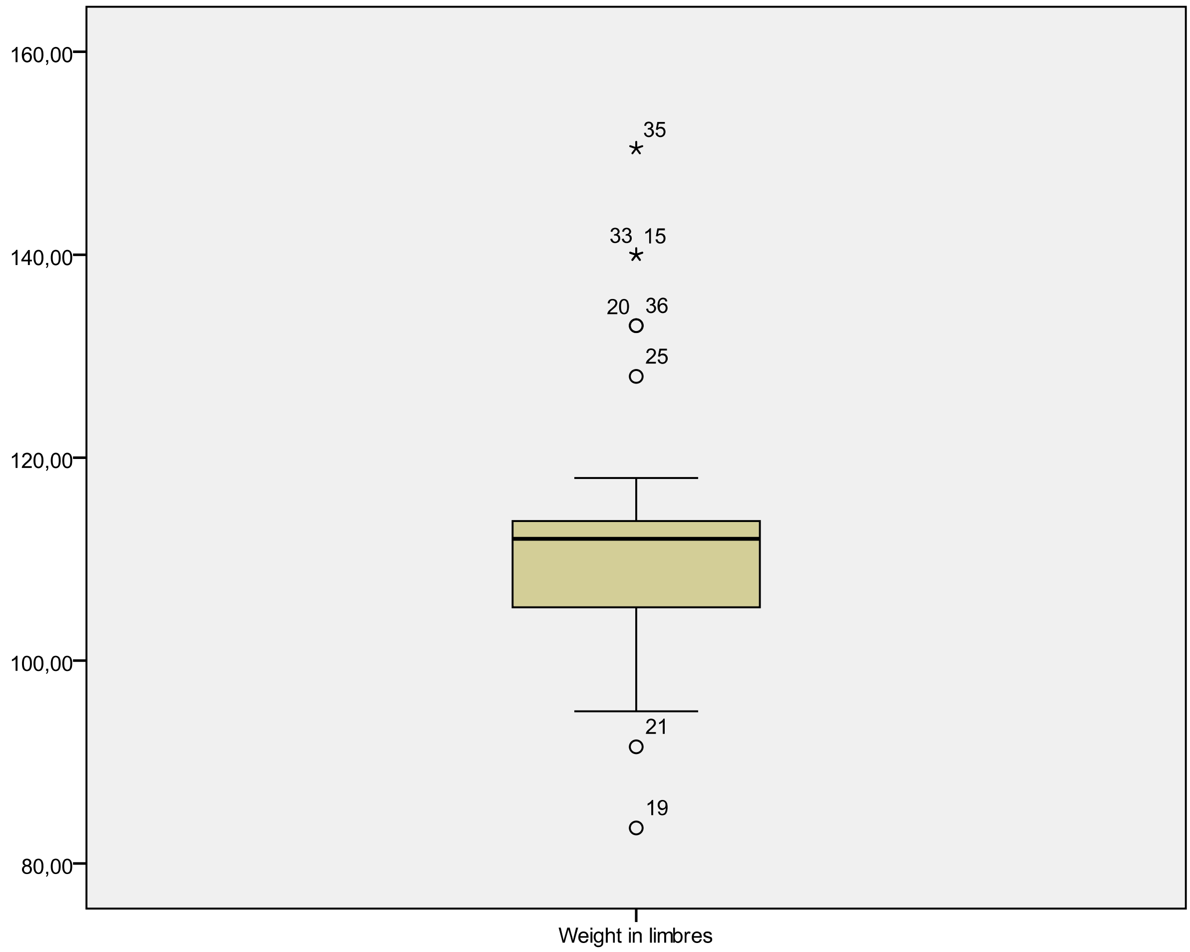
Ο έλεγχος των ακραίων τιμών έγινε με το θηκόγραμμα και έδειξε ότι δεν υπάρχουν (βλέπε θηκόγραμμα 1) τέτοιες στις δειγματικές τιμές που καταγράφεται ο δείκτης ευφυΐας των 35 ατόμων. Καθώς δεν υπάρχουν ακραίες τιμές συνεχίζουμε την περαιτέρω ανάλυση ελέγχοντας την υπόθεση ότι οι διαθέσιμες δειγματικές παρατηρήσεις που καταγράφεται ο δείκτης ευφυΐας προέρχονται από έναν πληθυσμό που περιγράφεται ικανοποιητικά από την κανονική κατανομή. Από το τεστ των Shapiro-Wilk έχουμε ότι η υπόθεση ότι οι δειγματικές τιμές του δείκτη ευφυΐας προέρχονται από έναν πληθυσμό που περιγράφεται ικανοποιητικά από την υπόθεση της κανονικής κατανομής δεν μπορεί να απορριφθεί (τιμή του τεστ 0.986, β.ε. 35, p=0,932).

Εφόσον ικανοποιούνται όλες οι προϋποθέσεις, μπορούμε να κάνουμε χρήση του παραμετρικού ελέγχου t-Test για τον έλεγχο της υπόθεσης ότι ο μέσος δείκτης ευφυΐας του πληθυσμού είναι ίσος με 100. Από τον έλεγχο αυτό προκύπτει ότι ο μέσο δείκτης ευφυΐας δε διαφέρει στατιστικά σημαντικά από το 100 (p=0,220), ενώ ένα 95% Δ.Ε. για το μέσο δείκτη ευφυΐας είναι το (100-8.03,100+1.91).

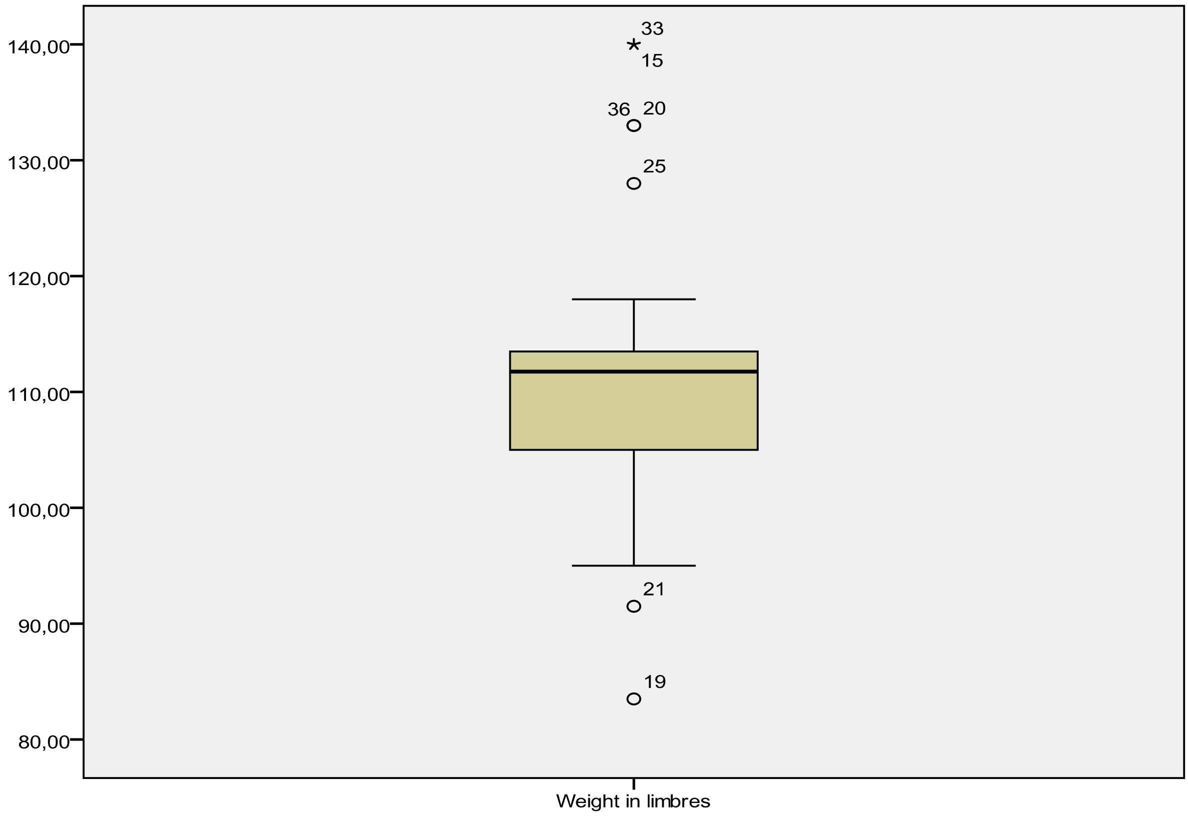
<u>Παράδειγμα 3^o</u> Στο αρχείο HeightWeight.sav[*] καταγράφονται οι τιμές του βάρους (σε λίμπρες) και του ύψους (σε ίντσες) 73 τυχαία επιλεγμένων ατόμων από έναν πληθυσμό. Θέλουμε να ελέγξουμε, αν είναι εφικτό, αν το μέσο βάρος του πληθυσμού είναι στατιστικά σημαντικά διαφορετικό από τις 70 λίμπρες.

Υλοποίηση-Αποτελέσματα:

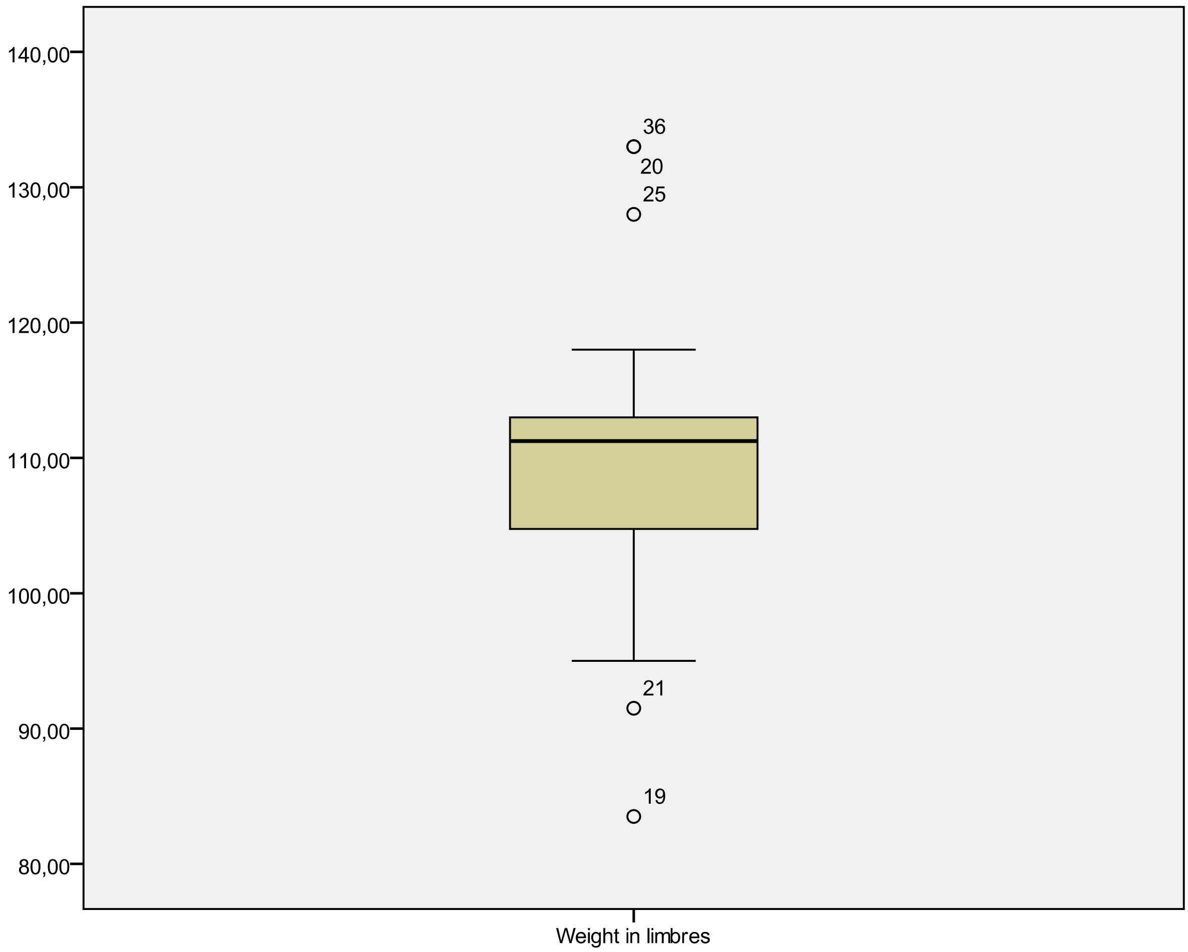
Έλεγχος ακραίων τιμών. Ο έλεγχος για την ύπαρξη ακραίων τιμών γίνεται με το θηκόγραμμα, όπως αναλυτικά περιγράφηκε στο Παράδειγμα 1, μέσω της διαδικασίας Analyze Descriptive Statistics Explore.



Καθώς η παρατήρηση 35 είναι εκείνη που είναι πιο απομακρυσμένη από τους φράκτες (whiskers) θα είναι αυτή που αρχικά αποκλείουμε από την περαιτέρω ανάλυση με τον τρόπο που περιγράφηκε στο Παράδειγμα 1. Έπειτα επαναλαμβάνουμε τον έλεγχο ύπαρξης ακραίων τιμών. Από το νέο θηκόγραμμα προκύπτει ότι οι παρατηρήσεις με αύξοντα αριθμό 33 και 15 είναι επίσης ακραίες. Καθώς το ποσοστό των ακραίων παρατηρήσεων δεν ξεπέρασε ακόμη το 10% των διαθέσιμων παρατηρήσεων τις αποκλείουμε και προβαίνουμε σε επανέλεγχο για την ύπαρξη ακραίων τιμών.

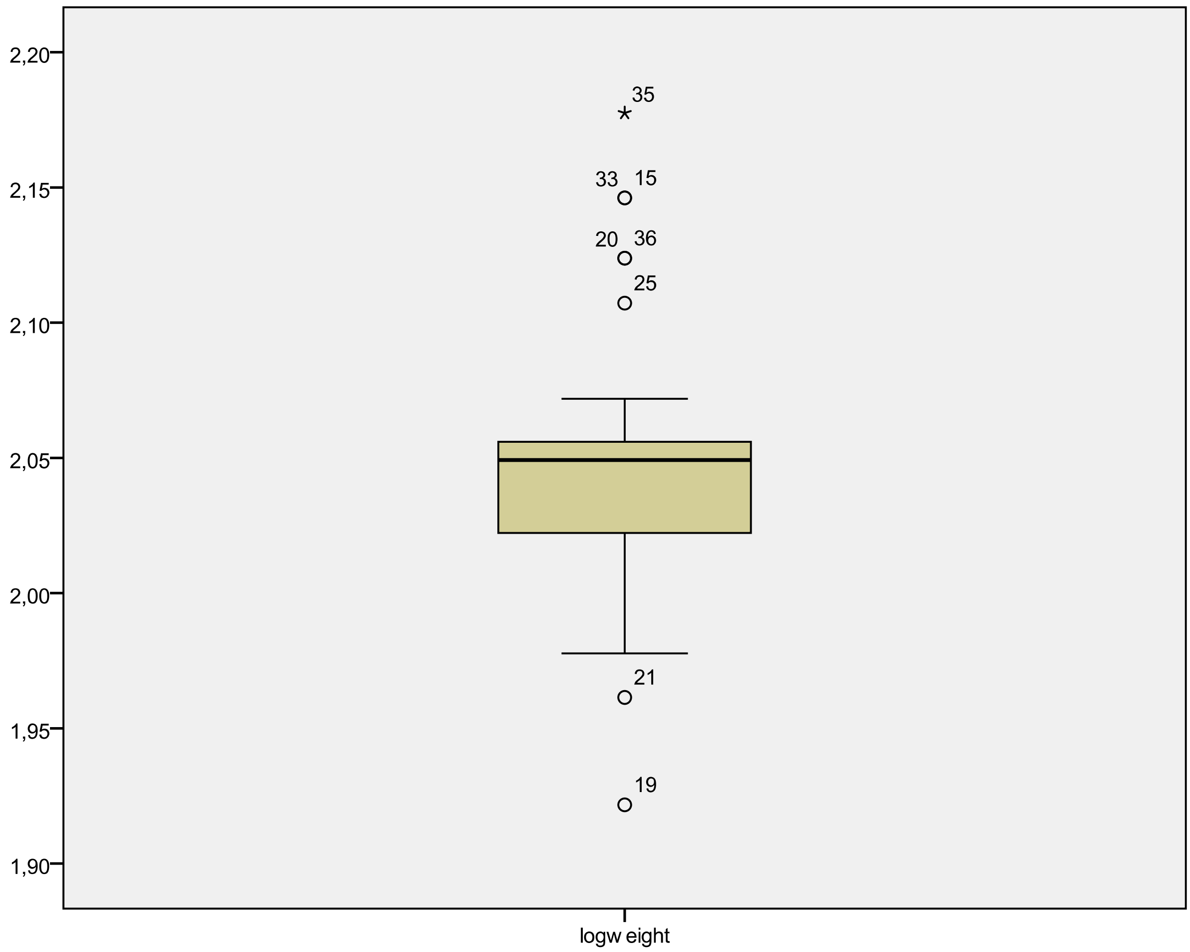


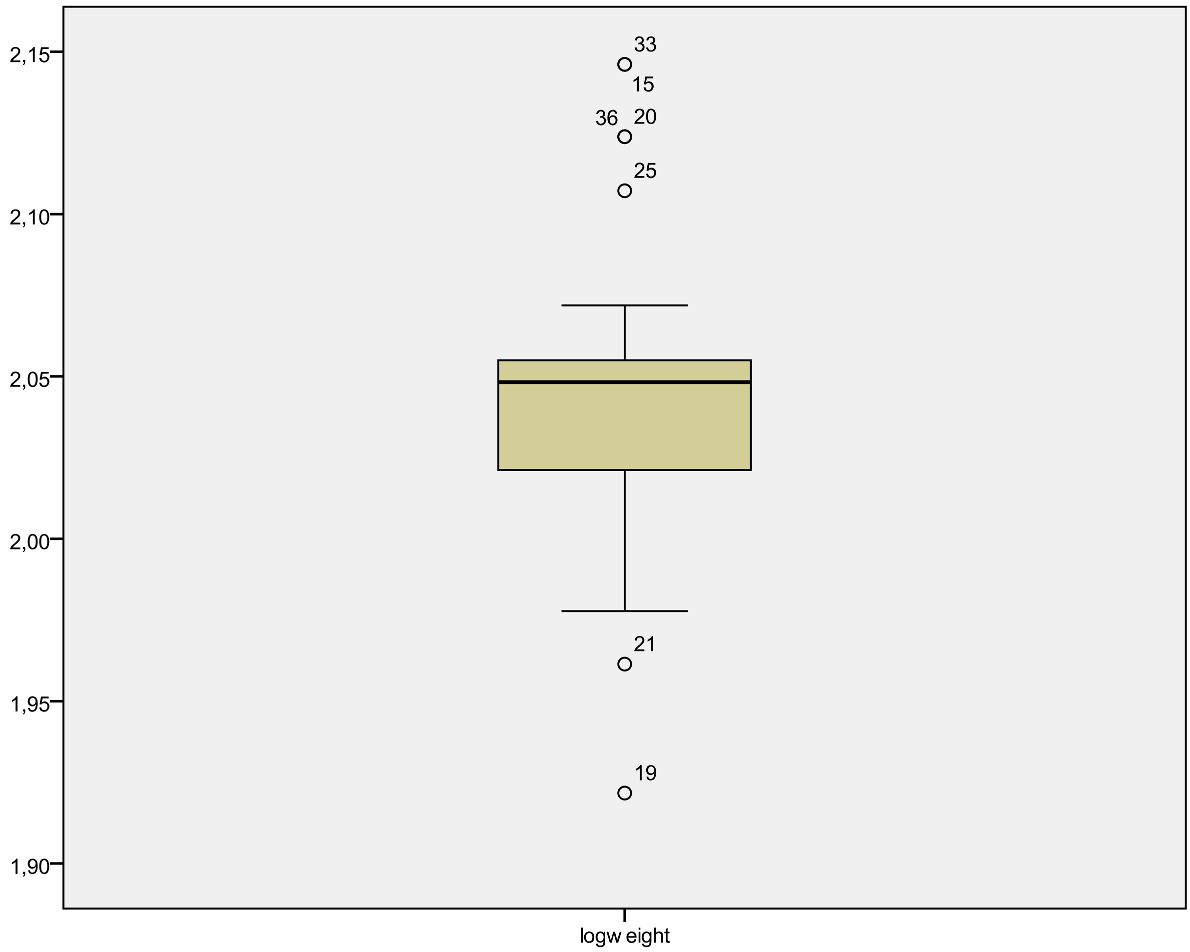
Το νέο θηκόγραμμα που προκύπτει είναι το:

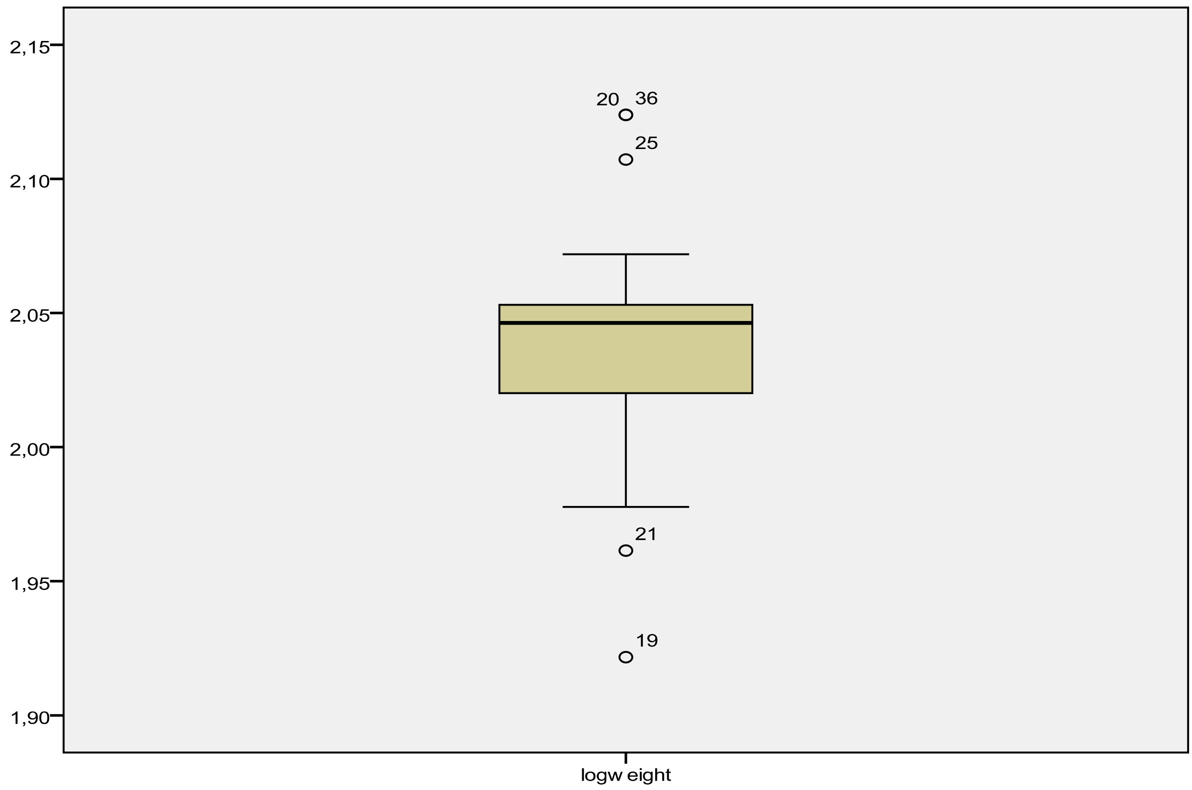


Επομένως προκύπτει ότι οι παρατηρήσεις με αύξοντα αριθμό 36 και 20 είναι ακραίες. Επομένως πλέον έχουμε διαπιστώσει ότι υπάρχουν τουλάχιστον 5 ακραίες τιμές οι παρατηρήσεις με αύξοντα αριθμό 35,33,15,36,20 και το ποσοστό τους ξεπερνά το 10% (καθώς 5/39*100%>10%). Αφού τις επαναφέρουμε όλες τις παρατηρήσεις θα εξετάσουμε αν ο μετασχηματισμός του λογαρίθμου διορθώνει το πρόβλημα. Προσοχή: Πάντοτε πριν μετασχηματίσουμε τα δεδομένα μας κάνοντας χρήση της συνάρτησης του λογαρίθμου θα πρέπει λόγω ορισμού της να ελέγχουμε αν περιέχονται στα δεδομένα μη θετικές τιμές. Σε μία τέτοια περίπτωση αν X είναι η μεταβλητή που θα μετασχηματιστεί και α η μικρότερη μη θετική τιμή τότε προτείνεται ο μετασχηματισμός $\log\left(X+|a|+1\right)$.

Μετασχηματίζοντας τα δεδομένα και μέσω της διαδικασίας Explore προκύπτουν τα ακόλουθα θηκογράμματα με παρόμοιο τρόπο αποκλείοντας μία-μία τις ακραίες τιμές.




2,15
2,10
2,05
2,00
1,95
1,90
33
15
36
20
25
21
19
logw eight


2,15
2,10
2,05
2,00
1,95
1,90
20
36
25
21
19
logw eight

Επομένως ο μετασχηματισμός του λογαρίθμου δε διορθώνει το πρόβλημα καθώς το ποσοστό αυτών στα μετασχηματισμένα δεδομένα είναι μεγαλύτερο του 10%.

Άρα θα προβούμε στον μη παραμετρικό έλεγχο ότι η πληθυσμιακή διάμεσος του βάρους είναι 70 λίμπρες μέσω της διαδικασίας Analyze→Non Parametric Tests→One Sample που αναλυτικά περιγράφηκε προηγούμενα, χρησιμοποιώντας προφανώς όλες τις παρατηρήσεις και τα αρχικά δεδομένα.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The median of Height in inches equals 70.	One-Sample Wilcoxon Signed Ranks Test	.000	Reject the null hypothesis.
2	The median of Weight in limbres equals 70.	One-Sample Wilcoxon Signed Ranks Test	.000	Reject the null hypothesis.
3	The median of id equals 70.	One-Sample Wilcoxon Signed Ranks Test	.000	Reject the null hypothesis.
4	The median of id ~= 35 & id ~= 33 & id ~= 15 (FILTER) equals 70.	One-Sample Wilcoxon Signed Ranks Test	.000	Reject the null hypothesis.
5	The median of logweight equals 70.	One-Sample Wilcoxon Signed Ranks Test	.000	Reject the null hypothesis.

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

Καθώς ο παραπάνω πίνακας μας πληροφορεί αν η πληθυσμιακή διάμεσος του βάρους είναι στατιστικά σημαντικά διαφορετική από 70 λίμπρες, από τη διαδικασία Analyze Descriptive Statistics Explore θα διαπιστώσουμε αν τα αποτελέσματα που αφορούν την πληθυσμιακή διάμεσο μπορούν να γενικευτούν στην πληθυσμιακή μέση τιμή, εξετάζοντας αν τα δεδομένα είναι συμμετρικά.

Descriptives

			Statistic	Std. Error
Weight in limbres	Mean		112,2051	2,10084
	95% Confidence Interval for Mean	Lower Bound	107,9522	
		Upper Bound	116,4581	
	5% Trimmed Mean		111,7400	
	Median		112,0000	
	Variance		172,128	
	Std. Deviation		13,11975	
	Minimum		83,50	
	Maximum		150,50	
	Range		67,00	
	Interquartile Range		9,00	
	Skewness		,933	,378
	Kurtosis		1,754	,741

Αναφορά: Θέλουμε να ελέγξουμε αν το μέσο βάρος του πληθυσμού είναι 70 λίμπρες. Το πρόβλημα αυτό είναι ένας έλεγχος για τη μέση τιμή ενός πληθυσμού. Θα ελέγξουμε αρχικά αν ικανοποιούνται οι παρακάτω υποθέσεις χρήσης του παραμετρικού αυτού ελέγχου.

1. Το δείγμα μας είναι τυχαίο
2. Δεν υπάρχουν ακραίες τιμές στα δεδομένα μας που ξεπερνούν σε ποσοστό το 10%.
3. Τα δεδομένα μας ακολουθούν κανονική κατανομή.

Η πρώτη από τις προϋποθέσεις σχετίζεται με τον τρόπο που επιλέξαμε το δείγμα μας και ικανοποιείται.

Ο έλεγχος των ακραίων τιμών έδειξε ότι έχουμε μεγάλο αριθμό ακραίων παρατηρήσεων που ξεπερνούν σε ποσοστό το 10%. Συγκεκριμένα υπάρχουν τουλάχιστον 5 ακραίες παρατηρήσεις, οι δειγματικές παρατηρήσεις με αύξοντα αριθμό 35,33,15,20,36 (βλέπε θηκογράμματα 1,2,3). Ο μετασχηματισμός του λογαρίθμου (που πραγματοποιείται επαναφέροντας όλες τις παρατηρήσεις που είχαν αποκλειστεί) δε διορθώνει το πρόβλημα των ακραίων τιμών (βλέπε θηκογράμματα 4,5,6). Συγκεκριμένα ακραίες παρατηρήσεις, με την σειρά που εμφανίσθηκαν, ήταν οι δειγματικές παρατηρήσεις με αύξοντα αριθμό 35,33,15,20,36. Για το λόγο αυτό θα καταφύγουμε στον μη παραμετρικό έλεγχο της υπόθεσης ότι η πληθυσμιακή διάμεσος του βάρους ισούται με 70 λίμπρες.

Επειδή η κρίσιμη πιθανότητα είναι p<0,001 συμπεραίνουμε ότι η διάμεσος του βάρος του πληθυσμού είναι στατιστικά σημαντικά διαφορετική από 70 λίμπρες. Καθώς η δειγματική μέση τιμή του βάρους είναι 112,2051 λίμπρες ενώ η δειγματική διάμεσος είναι 112, τα αποτελέσματα γενικεύονται για τη μέση τιμή και ειδικότερα συμπεραίνουμε ότι το μέσο βάρος του πληθυσμού είναι στατιστικά σημαντικά μεγαλύτερο από 70 λίμπρες.

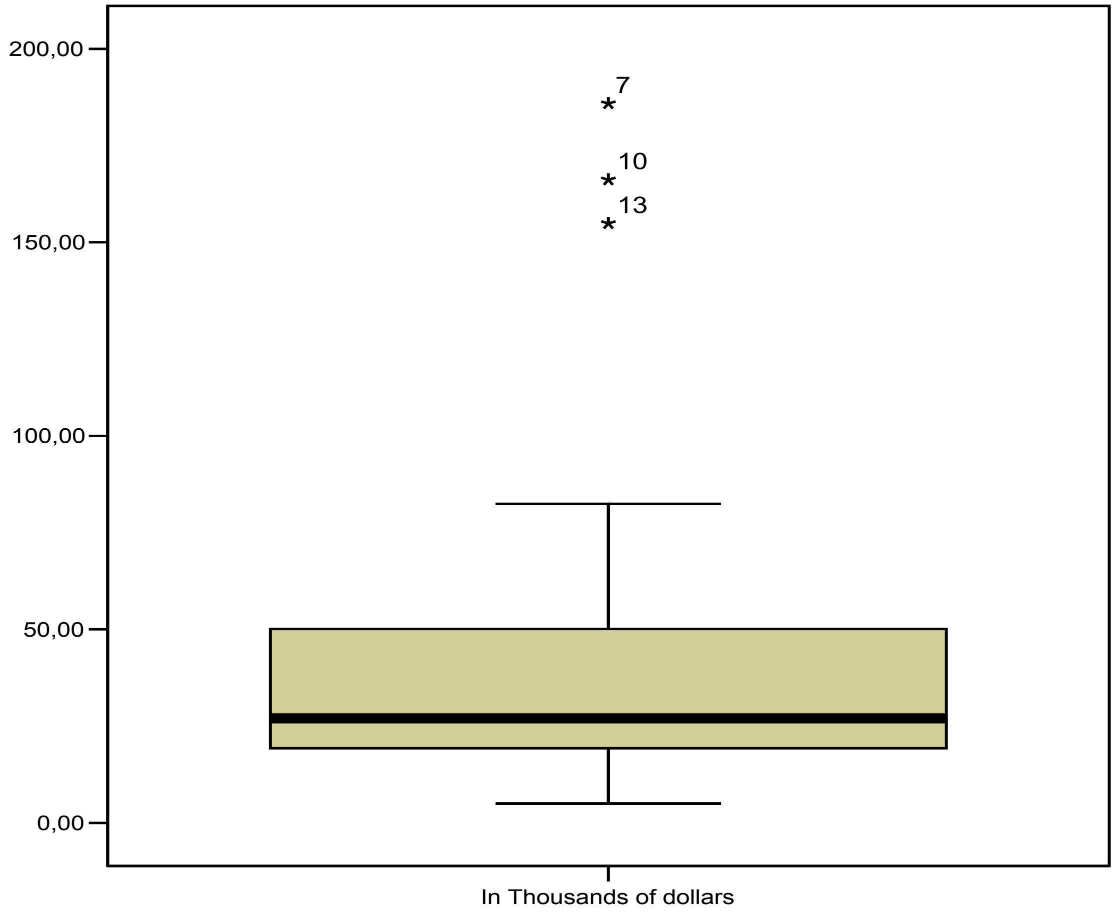
<u>Παράδειγμα 4^o</u> Στο αρχείο TVAdv.sav[*] καταγράφεται το ποσό που δαπανούν 21 τυχαία επιλεγμένες εταιρείες ενός πληθυσμού σε διαφημίσεις. Να ελεγχτεί αν είναι εφικτό, αν το οι μέσες δαπάνες των εταιρειών διαφέρουν στατιστικά σημαντικά από τα 184 εκατομμύρια δολάρια (τα δεδομένα δίνονται σε εκατομμύρια δολάρια).

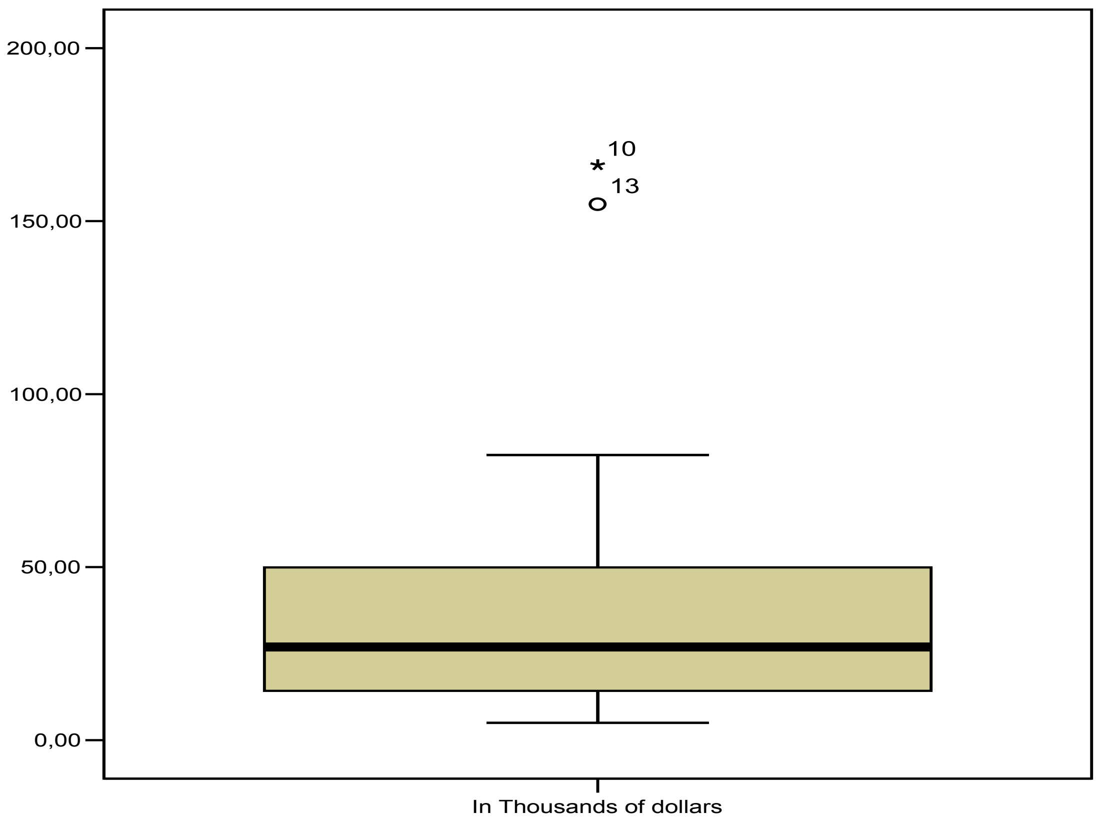
Αναφορά: Θέλουμε να ελέγξουμε αν κατά μέσο όρο το ποσό που δαπανούν οι εταιρείες ισούται με 185 εκατομμύρια δολάρια. Το πρόβλημα αυτό είναι ένας έλεγχος για τη μέση τιμή ενός πληθυσμού. Για να το ελέγξουμε θα χρησιμοποιήσουμε το t-Test για έναν πληθυσμό εφόσον ικανοποιούνται οι εξής προϋποθέσεις:

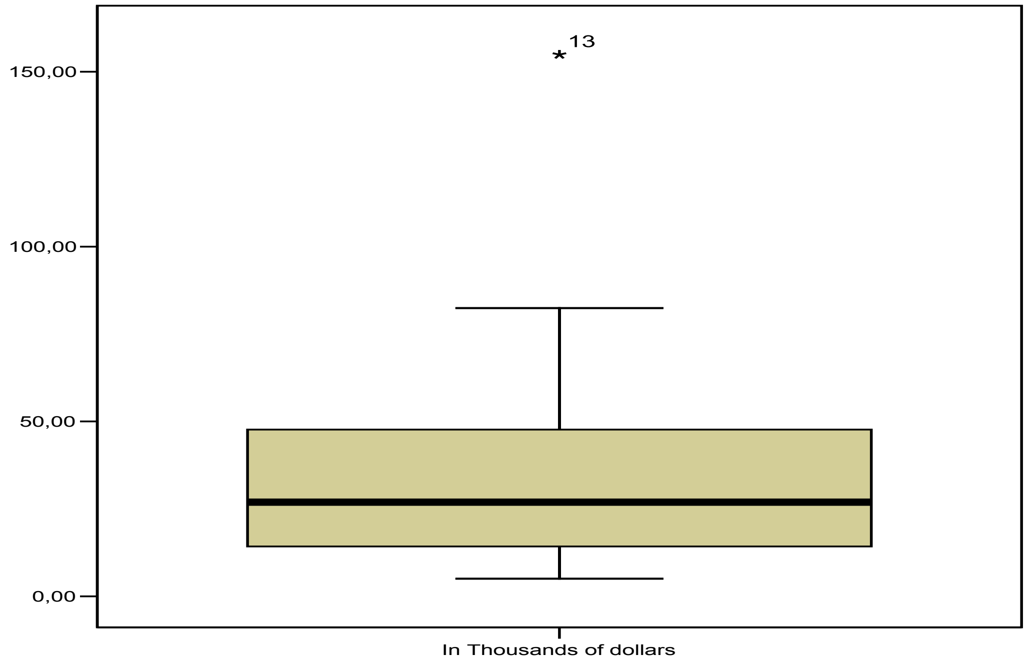
1. Το δείγμα μας είναι τυχαίο
2. Δεν υπάρχουν ακραίες τιμές στα δεδομένα μας.
3. Τα δεδομένα μας ακολουθούν κανονική κατανομή.

Η πρώτη από τις προϋποθέσεις σχετίζεται με τον τρόπο που επιλέξαμε το δείγμα μας και ικανοποιείται.

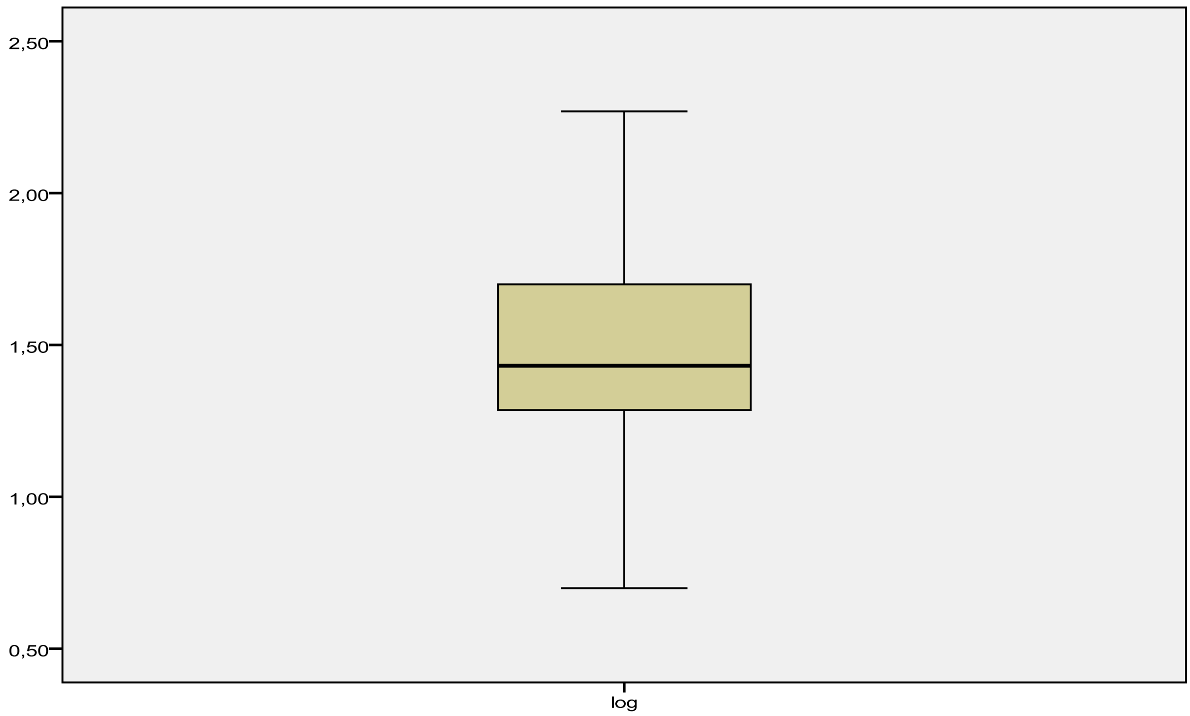
Ο έλεγχος των ακραίων τιμών έδειξε ότι ο αριθμός των ακραίων παρατηρήσεων υπερβαίνει το 10% του μεγέθους του δείγματος (3/21*100%>10%), καθώς υπάρχουν τουλάχιστον 3 ακραίες τιμές οι παρατηρήσεις με αύξοντα αριθμό 7,10,13 και τιμές στις δαπάνες 185.9, 166.20 και 154.90 αντίστοιχα (βλέπε θηκογράμματα 1,2,3).


200,00
150,00
100,00
50,00
0,00
7
10
13
In Thousands of dollars


200,00
150,00
100,00
50,00
0,00
10
13
In Thousands of dollars



Για τον λόγο αυτό εξετάζουμε αν ο μετασχηματισμός του λογαρίθμου θα διορθώσει το πρόβλημα αφού πρώτα επαναφέρουμε τις ακραίες παρατηρήσεις που πρωτύτερα έχουν αποκλειστεί. Ο μετασχηματισμός του λογαρίθμου διορθώνει το πρόβλημα καθώς από το θηκόγραμμα που προέκυψε (βλέπε θηκόγραμμα 4) συμπεραίνουμε ότι δεν υπάρχουν ακραίες παρατηρήσεις στις δειγματικές παρατηρήσεις του λογαρίθμου των δαπανών των εταιρειών.



Καθώς δεν υπάρχουν ακραίες τιμές στις δειγματικές παρατηρήσεις του λογαρίθμου των δαπανών συνεχίζουμε την περαιτέρω ανάλυση ελέγχοντας την υπόθεση ότι οι διαθέσιμες δειγματικές παρατηρήσεις που καταγράφεται ο λογάριθμος των δαπανών προέρχονται από έναν πληθυσμό που περιγράφεται ικανοποιητικά από την κανονική κατανομή. Από το τεστ των Shapiro-Wilk έχουμε ότι η υπόθεση ότι οι δειγματικές τιμές των δαπανών για διαφημίσεις των εταιρειών προέρχονται από έναν πληθυσμό που περιγράφεται ικανοποιητικά από την υπόθεση της κανονικής κατανομής δεν μπορεί να απορριφθεί (p=0,381).

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
log	,107	21	,200*	,953	21	,381

a. Lilliefors Significance Correction

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

Εφόσον ικανοποιούνται όλες οι προϋποθέσεις, μπορούμε να κάνουμε χρήση του παραμετρικού ελέγχου t-Test για τον έλεγχο της υπόθεσης ότι ο μέσος λογάριθμος των δαπανών είναι ίσος με το δεκαδικό λογάριθμο του 185, δηλαδή με 2.2267. Από τον έλεγχο αυτό προκύπτει ότι ο μέσο λογάριθμος των δαπανών διαφέρει στατιστικά σημαντικά από το 2,267 ($p<0.001$), ενώ ένα 95% Δ.Ε. για το μέσο λογάριθμο των δαπανών είναι το (2.267-1,0082, 2.267-0,5768).

<u>Παράδειγμα 5^o</u> Στο αρχείο HeightWeight15.sav* καταγράφονται οι τιμές του βάρους και του ύψους (σε ίντσες) τυχαία επιλεγμένων ατόμων από έναν πληθυσμό. Θέλουμε να ελέγξουμε, αν είναι εφικτό, αν το μέσο ύψος του πληθυσμού είναι στατιστικά σημαντικά διαφορετικό από τις 65 inches.

Η υλοποίηση παραλείπεται και αφήνεται ως άσκηση ενώ δίδεται μία συνοπτική αναφορά.

Συνοπτική Αναφορά: Θέλουμε να ελέγξουμε αν η μέση τιμή του ύψους για τον πληθυσμό, από τον οποίο επιλέξαμε το δείγμα μας, ισούται με 65 ίντσες. Το πρόβλημα αυτό είναι ένας έλεγχος για τη μέση τιμή ενός πληθυσμού. Για να το ελέγξουμε θα χρησιμοποιήσουμε το t-Test για έναν πληθυσμό εφόσον ικανοποιούνται οι εξής προϋποθέσεις:

1. Το δείγμα μας είναι τυχαίο
2. Δεν υπάρχουν ακραίες τιμές στα δεδομένα μας.
3. Τα δεδομένα μας ακολουθούν κανονική κατανομή.

Η πρώτη από τις προϋποθέσεις σχετίζεται με τον τρόπο που επιλέξαμε το δείγμα μας και ικανοποιείται.

Ο έλεγχος των ακραίων τιμών έγινε με το θηκόγραμμα και έδειξε ότι δεν υπάρχουν ακραίες τιμές.

Ο έλεγχος της κανονικής κατανομής με το τεστ των Shapiro-Wilk έδειξε ότι η υπόθεση αυτή απορρίπτεται για επίπεδο σημαντικότητας 5% ενώ δεν μπορεί να απορριφθεί για επίπεδο σημαντικότητας 1% (p=0,037). Συνέπεια αυτού είναι στη συνεχεία να χρησιμοποιούμε το 1% σαν επίπεδο σημαντικότητας, καθώς θέλουμε να αποφύγουμε τόσο το μετασχηματισμό των δεδομένων όσο και τη χρήση του Κ.Ο.Θ.

Εφόσον ικανοποιούνται όλες οι προϋποθέσεις, μπορούμε να κάνουμε χρήση του t-Test. Επειδή p=0,020 το συμπέρασμα από τη χρήση του τεστ αυτού είναι ότι η υπόθεση πως ο μέσο ύψος του πληθυσμού είναι στατιστικά σημαντικά ίσο με 65 ίντσες, δεν μπορεί να απορριφθεί για επίπεδο σημαντικότητας 1%.