ΕΛΕΓΧΟΙ ΥΠΟΘΕΣΕΩΝ

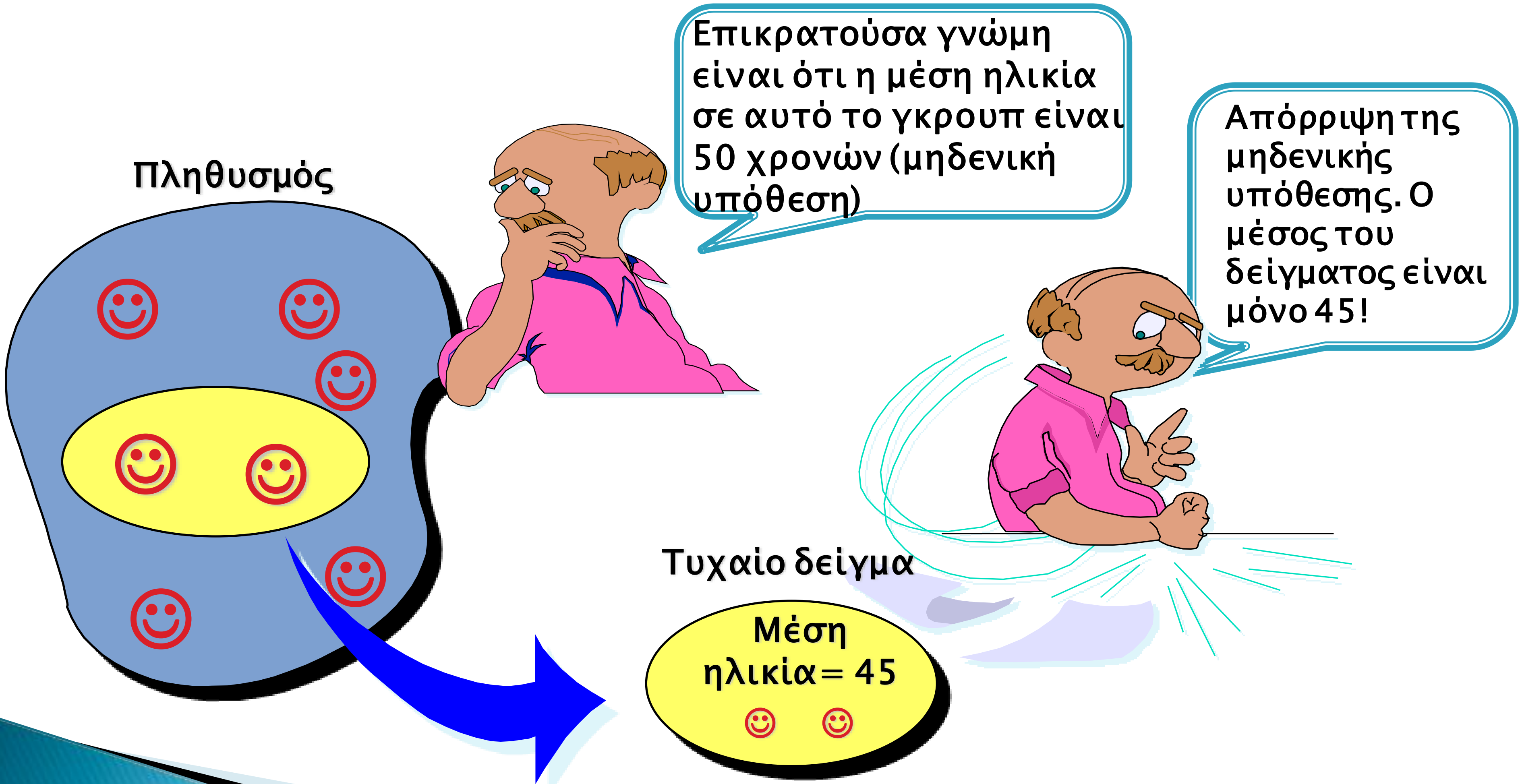
Κλωνάρης Στάθης

ΠΜΣ: Οργάνωση & Διοίκηση Επιχειρήσεων Τροφίμων και Γεωργίας

Έλεγχος Υποθέσεων

- Η Υπόθεση είναι μία πεποίθηση σχετικά με μία παράμετρο
- Παράμετρος μπορεί να είναι ο μέσος ενός πληθυσμού, ένα ποσοστό, ένας συντελεστής συσχέτισης,...

Η Λογική πίσω από τον έλεγχο υποθέσεων



Έλεγχος Υποθέσεων

Έλεγχος Υποθέσεων: Είναι μια μέθοδος της Στατιστικής Συμπερασματολογίας (statistical inference) που μας βοηθά να βγάλουμε συμπεράσματα σε σχέση με την τιμή μιας παραμέτρου του πληθυσμού, συγκρίνοντας τα αποτελέσματα του δείγματος με εκείνα που θα περιμέναμε εάν η **υπόθεση** είναι αληθινή.

Η μέθοδος αυτή μας βοηθά να διαπιστώσουμε εάν τα δεδομένα του δείγματος υποστηρίζουν την υπόθεση ότι η παράμετρος του πληθυσμού έχει μια συγκεκριμένη τιμή.

Με άλλα λόγια προσπαθούμε να προσδιορίσουμε την πιθανότητα να προκύψει ένα δείγμα όπου η τιμή της παραμέτρου είναι π.χ. $\boldsymbol{\mu}$ από έναν πληθυσμό με πραγματική τιμή $\boldsymbol{\mu_0}$. Εάν η πιθανότητα είναι μεγάλη τότε η διαφορά μεταξύ του μ και μ_0 οφείλεται στις τυχαίες κυμάνσεις της δειγματοληψίας. Αντίθετα, εάν διαπιστωθεί ότι η πιθανότητα είναι μικρή, τότε η υπόθεση δεν ισχύει.

Παραδείγματα Ελέγχου Υποθέσεων

H_0: Το μέσο ύψος των ανδρών είναι 174.

H_1: Το μέσο ύψος των ανδρών είναι μεγαλύτερο από174.

H_0: Ο μισός πληθυσμός είναι υπέρ της χρήσης πυρηνικής ενέργειας .

H_1: Λιγότερο από τον μισό πληθυσμό είναι υπέρ της χήρησης πυρηνικής ενέργειας

H_0: Το ποσοστό ανεργίας είναι ίσο μεταξύ ανδρών και γυναικών.

H_1: Το ποσοστό ανεργίας είναι μεγαλύτερο στις γυναίκες.

H_0: Δεν υπάρχει συσχέτιση μεταξύ των επιτοκίων και της τιμής του χρυσού.

H_1: Υπάρχει συσχέτιση μεταξύ των επιτοκίων και της τιμής του χρυσού.

Βασικές έννοιες….

Οι βασικές έννοιες του ελέγχου υποθέσεων είναι οι εξής:

1. Υπάρχουν δύο υποθέσεις. Η πρώτη ονομάζεται μηδενική υπόθεση (null hypothesis, H_0)και η δεύτερη εναλλακτική υπόθεση ή υπόθεση έρευνας (alternative hypothesis, H_1).
2. Ο έλεγχος ξεκινά θωρώντας ότι η μηδενική υπόθεση είναι αληθής.
3. Στόχος του ελέγχου είναι να καθοριστεί αν υπάρχουν αρκετές αποδείξεις που να στηρίζουν την αλήθεια της εναλλακτικής υπόθεσης.
4. Οι δυνατές αποφάσεις είναι δύο:
    - Απόρριψη της μηδενικής υπόθεσης
    - Αποδοχή (μη- απόρριψη) της μηδενικής Υπόθεσης

Βασικές έννοιες….

- **Μηδενική υπόθεση** (Null Hypothesis) H_0, είναι η υπόθεση που ελέγχουμε.
- **Εναλλακτική υπόθεση** (Alternative Hypothesis) H_1, είναι η υπόθεση που θα δεχθούμε εάν η μηδενική υπόθεση βρεθεί εσφαλμένη.
- Η **Εναλλακτική υπόθεση** δεν περιέχει το σύμβολο ίσον (=) όσον αφορά την τιμή της παραμέτρου του πληθυσμού αλλά βασίζεται στο $\neq$, $>$, $<$
- Η αποδοχή ή η απόρριψη της μηδενικής υπόθεσης βασίζεται στη τιμή του **κριτηρίου ελέγχου** (test statistic). Κριτήριο ελέγχου είναι η στατιστική (μεταβλητή) που υπολογίζεται με βάση τις πληροφορίες του δείγματος και ακολουθεί μια γνωστή κατανομή (Z κανονικής κατανομής, t κατανομής student, κτλ). Εφόσον το κριτήριο ακολουθεί μια γνωστή κατανομή, με την χρήση των εμβαδών της αντίστοιχης κατανομής θα προσδιορίσουμε την πιθανότητα με βάση την οποία θα κρίνουμε εάν η H_0 είναι αληθινή ή όχι.
- **Επίπεδο σημαντικότητας** (significant level) και συμβολίζεται με α, είναι η τιμή της πιθανότητας που μικρότερή της θα θεωρείται απίθανο να αποδοθεί στην τύχη ενώ μεγαλύτερη θα είναι φυσικό να ην αποδώσουμε στην τύχη.

Εσφαλμένα Συμπεράσματα από την Διεξαγωγή των Ελέγχων Υποθέσεων

Κατά το έλεγχο υποθέσεων υπάρχουν δύο δυνατοί τύποι σφαλμάτων. Το **σφάλμα τύπου I** συμβαίνει όταν απορρίπτουμε μία αληθινή μηδενική υπόθεση, ενώ το **σφάλμα τύπου II** συμβαίνει όταν δεχόμαστε μία ψευδή μηδενική απόφαση.

Πραγματικότητα	**Απόφαση**	
	Επιλέγω H_0	Επιλέγω H_1
H_0 Αληθινή	Σωστή Απόφαση	Σφάλμα Τύπου I
H_1 Αληθινή	Σφάλμα τύπου II	Σωστή Απόφαση

Παράδειγμα Εσφαλμένων αποφάσεων

Ένα συνηθισμένο παράδειγμα είναι από τον χώρο της δικαιοσύνης. Όταν κάποιος αντιμετωπίζει μια σοβαρή κατηγορία οδηγείται στο δικαστήριο, όπου ο δημόσιος κατήγορο και ο συνήγορος υπεράσπισης παρουσιάζουν αποδείξεις και αναπτύσσουν επιχειρήματα, και τελικά οι ένορκοι αποφασίζουν αν ο κατηγορούμενος είναι αθώος ή ένοχος. Αυτό που κάνουν οι ένορκοι είναι ένας έλεγχος υποθέσεων.

H0: ο κατηγορούμενος είναι αθώος

H1: ο κατηγορούμενος είναι ένοχος

Πραγματικότητα	Απόφαση δικαστηρίου	
	Επιλέγω H_0 (Αθώωση)	Επιλέγω H_1 (Καταδίκη)
H_0 Αληθινή (αθώος)	Σωστή Απόφαση	Σφάλμα Τύπου I
H_1 Αληθινή (ένοχος)	Σφάλμα τύπου II	Σωστή Απόφαση

Ερώτηση ....

1. Μπορούμε να αποφύγουμε τα σφάλματα κατά τον έλεγχο των υποθέσεων;

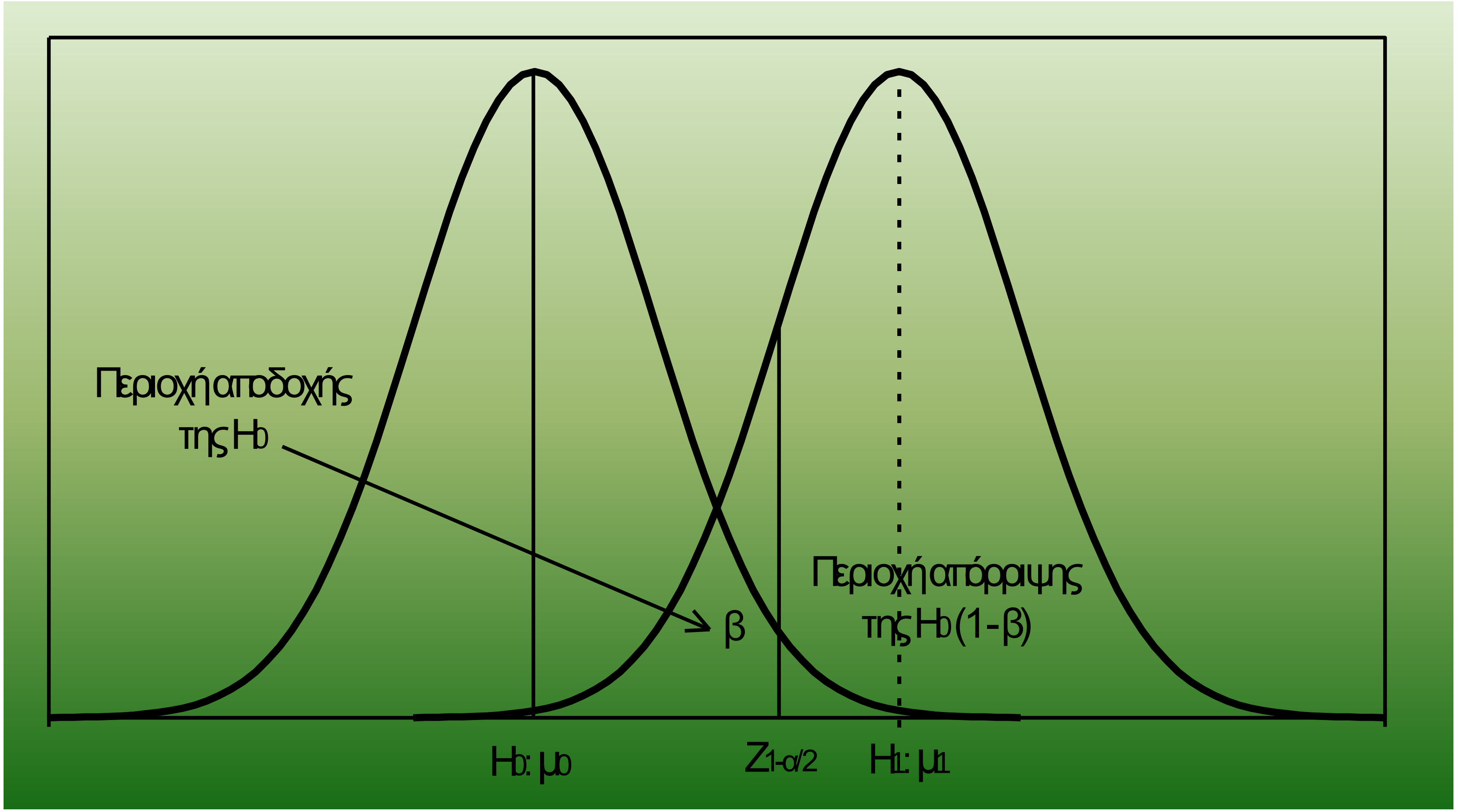
2. Μπορούμε να υπολογίσουμε την πιθανότητα να διαπράξουμε κατά τον έλεγχο των υποθέσεων;

$\alpha = P$(Σφάλμα Τύπου I)$=P$(Απορρίπτω την H_0 | H_0 σωστή)$=$

$\beta = P$(Σφάλμα Τύπου II)$=P$(Δεν απορρίπτω την H_0 | H_1 σωστή)

$1-\beta = P$(απορρίπτω H_0 | H_1 σωστή)
καλείται Δύναμη κριτηρίου ή ***Ισχύς ελέγχου*** *και* εκφράζεται συνήθως ως συνάρτηση των τιμών $\mu_1 - \mu_o$ δηλαδή $1-\beta = f(\mu_1 - \mu_0)$

Κατανομές Δειγματοληψίας και Περιοχές Απόρριψης Αποδοχής της H_o όταν η H_o είναι Λανθασμένη



Ένα δείγμα
1. Έλεγχος υπόθεσης Μέσου (γνωστή σ)

H_0: $\mu=\mu_o$ H_1: $\mu\neq\mu_o$

Κριτήριο ελέγχου: $$Z=\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$$

Εφόσον το επίπεδο σημαντικότητας είναι ***α,*** τότε η περιοχή απόρριψης έχει εμβαδόν (πιθανότητα) ***α,*** και η περιοχή αποδοχής εμβαδόν ίσο με *1-α*

Οι τιμές του κριτηρίου Ζ που αντιστοιχούν στα όρια της περιοχής απόρριψης ονομάζονται **κριτικές τιμές** (critical values). Η κάτω τιμή συμβολίζεται με $Z_{\alpha/2}$ και η άνω τιμή με $Z_{1-\alpha/2}$

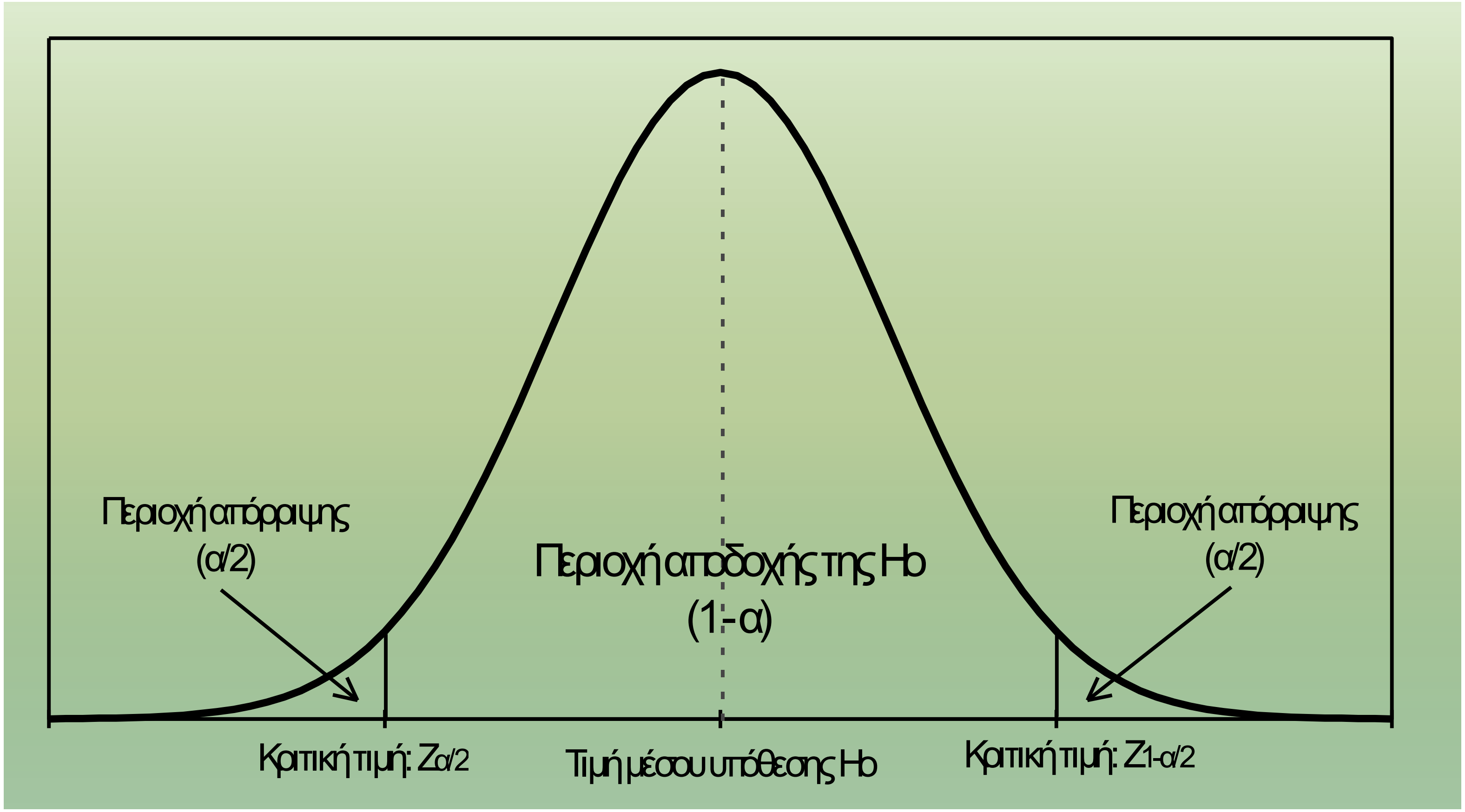
Εάν Ζ είναι η τιμή που προκύπτει από τα δεδομένα

του δείγματος τότε εάν:

$Z<Z_{\alpha/2}$ ή $Z>Z_{1-\alpha/2}$ τότε η H_0 απορρίπτεται

$Z_{\alpha/2}<Z<Z_{1-\alpha/2}$ τότε η H_0 γίνεται δεκτή

Κατανομή Δειγματοληψίας και Περιοχές Απόρριψης και Αποδοχής της H_o



Παράδειγμα

Μία εταιρεία παράγει συνθετικό δέρμα που χρησιμοποιείται για την υποδηματοποιία. Μία από τις βασικές προδιαγραφές που πρέπει να τηρούνται είναι το πάχος. Από τις προδιαγραφές είναι γνωστό ότι ο συγκεκριμένος τύπος συνθετικού δέρματος παράγεται με μέσο πάχος 4 χιλιοστά και σ=0,1 χιλιοστά. Από δείγμα 50 μετρήσεων προέκυψε μέσο παχος=4,02 χιλιοστά. Το ερώτημα που απασχολεί τον υπεύθυνο παραγωγής είναι εάν η παραγωγή εξελίσσεται ομαλά σύμφωνα με τις προδιαγραφές ή κάποιο πρόβλημα στην παραγωγή προκαλεί μεταβολή στο πάχος του συνθετικού δέρματος.

$n=50$

H_0: $\mu = 4$ χιλ.

H_1: $\mu \neq 4$χιλ.

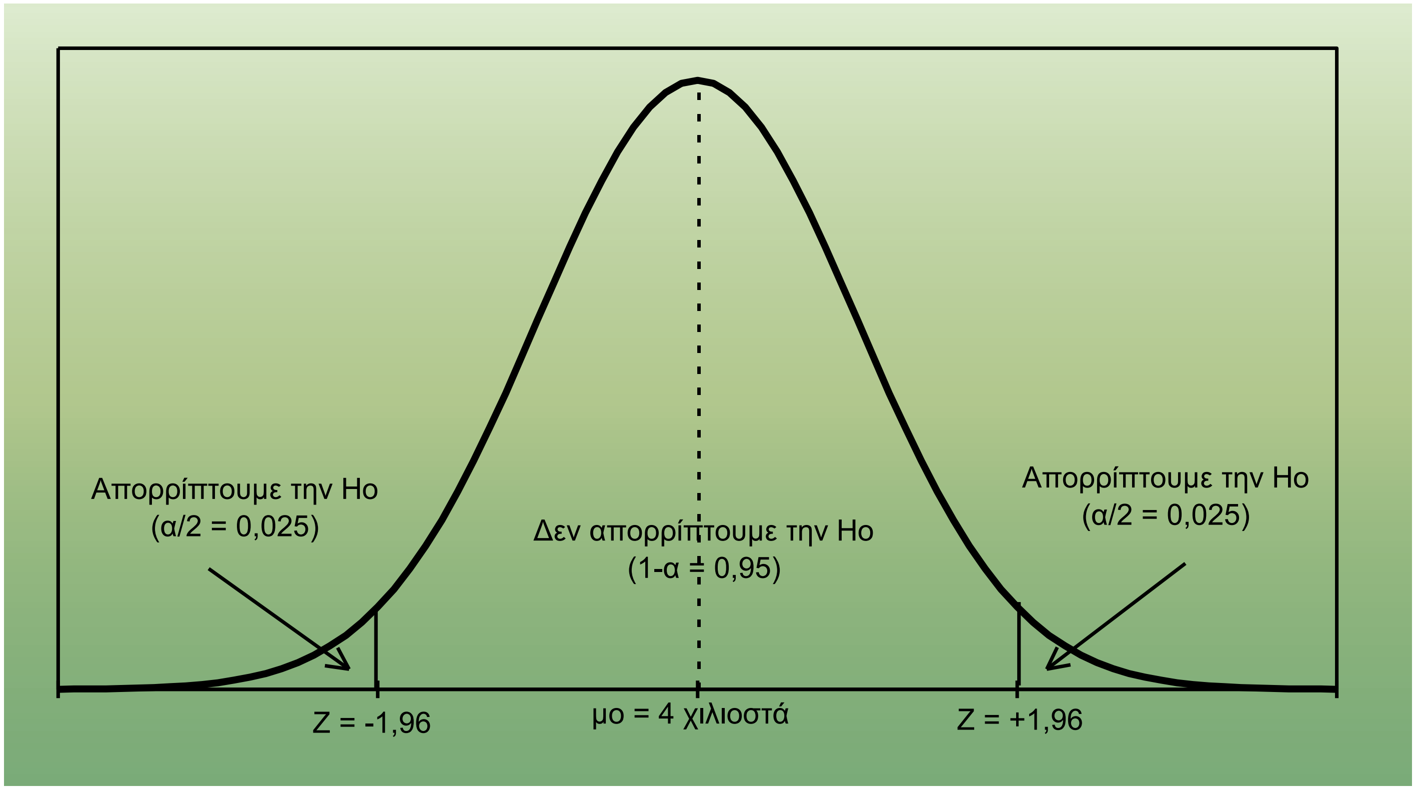
Οι κριτικές τιμές του κριτηρίου Z σύμφωνα με την τυποποιημένη κανονική κατανομή είναι -1,96 και 1,96

$$Z = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}} = \frac{4{,}02 - 4{,}00}{0{,}1 / \sqrt{50}} = +1{,}41$$

Για α=0.05 εύκολα βρίσκουμε από τους Πίνακες της τυποποιημένης Κανονικής Κατανομής ότι $z_{0.025}=1.96$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

Έλεγχος υπόθεσης Μέσου (γνωστή σ) Σε επίπεδο σημαντικότητας α=0,05



Άρα, $z=1.41<|z_{0.025}|=1.96$ οπότε δεν απορρίπτουμε την μηδενική υπόθεση H_0

Έλεγχος υπόθεσης Μέσου (γνωστή σ) Σε επίπεδο σημαντικότητας α=0,05

Ισοδύναμα θα μπορούσαμε να κατασκευάσουμε ένα συμμετρικό 95% Διάστημα Εμπιστοσύνης για το μ, το οποίο εν είναι τίποτε άλλο από την περιοχή αποδοχής του αμφίπλευρου ελέγχου:

$$(\bar{x} - z_{\frac{a}{2}}\sigma/\sqrt{n}, \bar{x} + z_{\frac{a}{2}}\sigma/\sqrt{n}) = (4.02 - 1.96 * \frac{0.1}{\sqrt{50}}, 4.02 + 1.96 * \frac{0.1}{\sqrt{50}}) =$$
$$(3.9922\ , 4.0477)$$

Παρατηρούμε ότι το παραπάνω ΔΕ περιέχει την υποτιθέμενη τιμή με βάση την H_0 (4 mm) και άρα οδηγούμαστε στο συμπέρασμα να δεχτούμε την H_0

p-value....

Στο προηγούμενο παράδειγμα (εταιρία παραγωγής συνθετικού δέρματος) χρησιμοποιήσαμε έναν τυποποιημένο έλεγχο (*standardized test statistic*). Το μειονέκτημα αυτής της μεθόδου είναι ότι η πληροφορία που μας δίνει είναι ένα απλό ναι ή όχι το οποίο οδηγεί στην λήψη της απόφασης.

Η τιμή p (*p-value*) είναι η δεσμευμένη πιθανότητα να πάρει ο έλεγχος μια τιμή σαν αυτή που έχει υπολογιστεί από το δείγμα, με δεδομένη την αλήθεια της μηδενικής υπόθεσης.

p-value….

Στο προηγούμενο παράδειγμα, η τιμή p είναι η πιθανότητα σε ένα πληθυσμό με μέσο 4,0 και τυπική απόκλιση 0,1, ένα δείγμα μεγέθους 50 μετρήσεων να έχει μέσο όρο διαφορετικό του 4,0.

$$P(\overline{X} \neq 4{,}0) = a/2 \Leftrightarrow P\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} > \frac{4.02-4.00}{0.1/\sqrt{50}}\right) = a/2$$

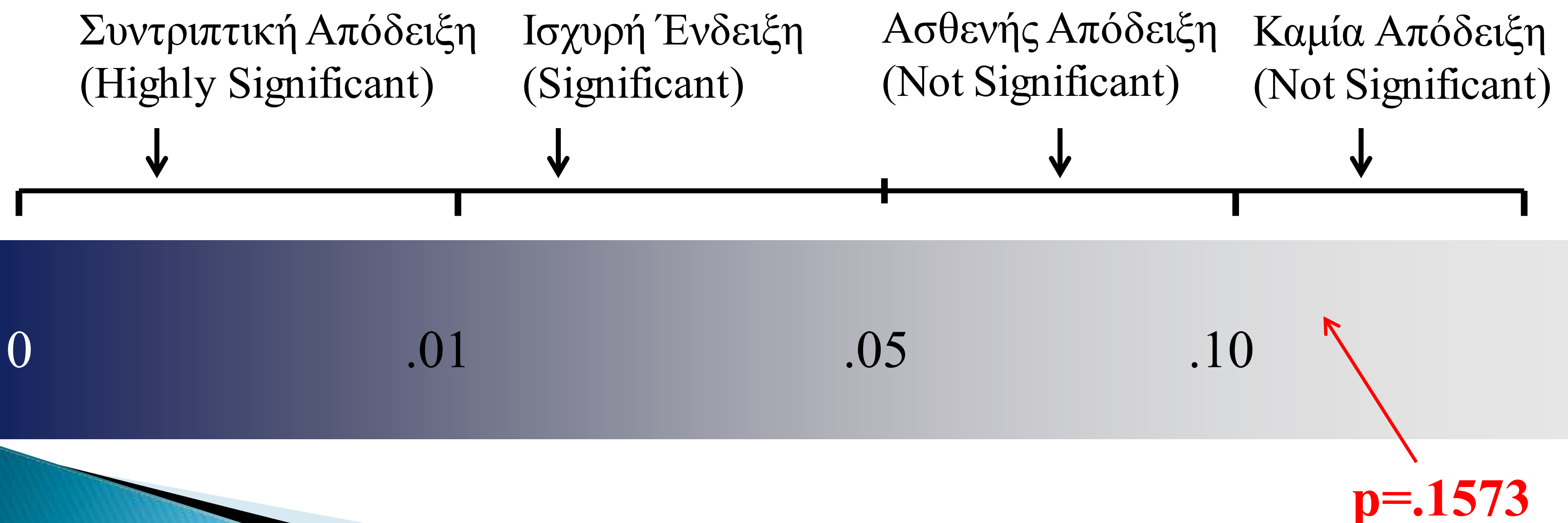
$$P(|Z| > 1{,}414) = 0.0786 \text{ άρα } P = 0.1573$$

Με άλλα λόγια, υπάρχει μία μεγάλη πιθανότητα το δείγμα με μέσο 4,02 να προέρχεται από έναν πληθυσμό με μέσο 4,0 ή αλλιώς δεν έχουμε καμία απόδειξη ώστε να απορρίψουμε την μηδενική υπόθεση

Όταν το p-value είναι μικρότερο του α, τότε η H_0 απορρίπτεται

Ερμηνεύοντας την τιμή *p*

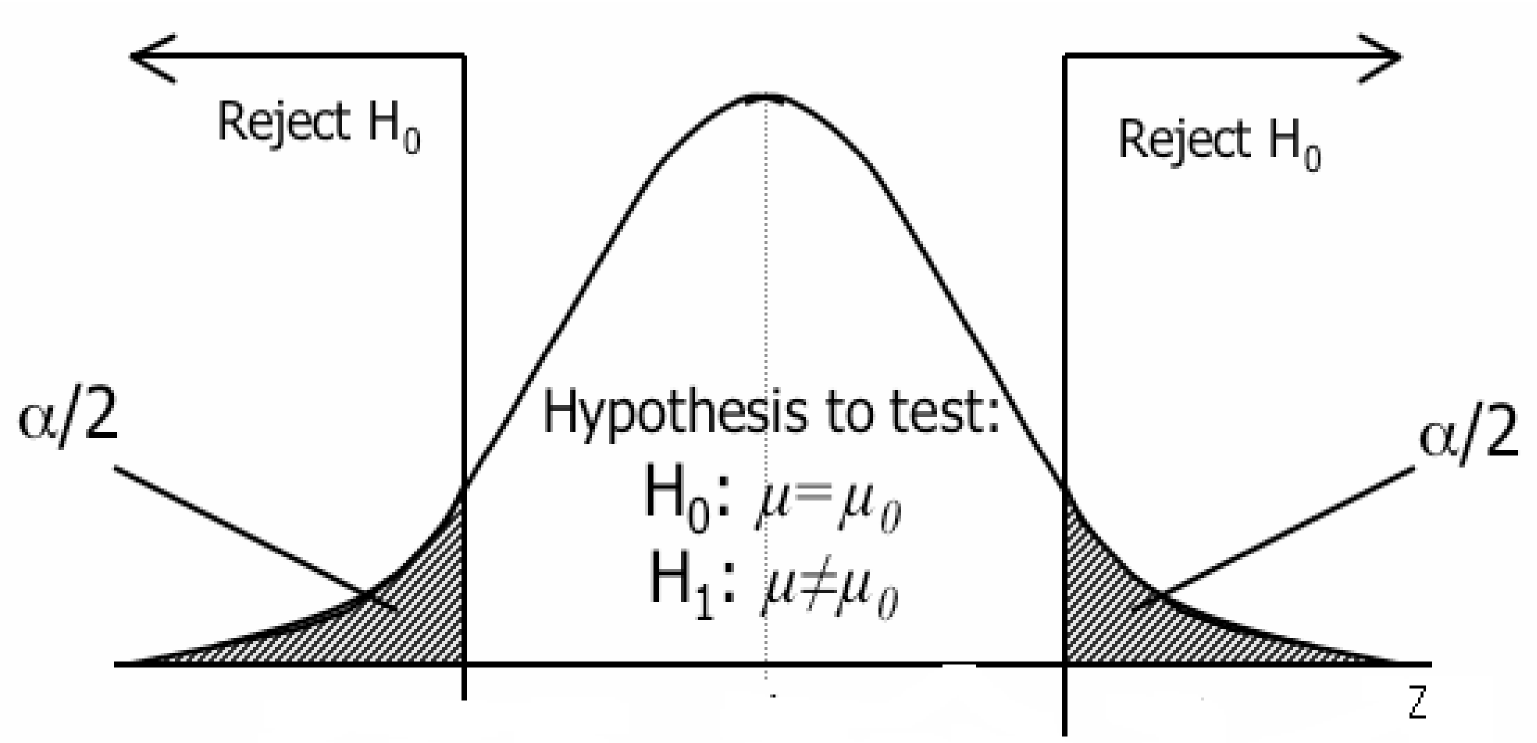
Η τιμή P ενός ελέγχου υπόθεσης αποτελεί μία πολύτιμη πληροφορία επειδή είναι ένα μέτρο στατιστική βαρύτητας των στοιχείων που στηρίζουν την εναλλακτική υπόθεση.



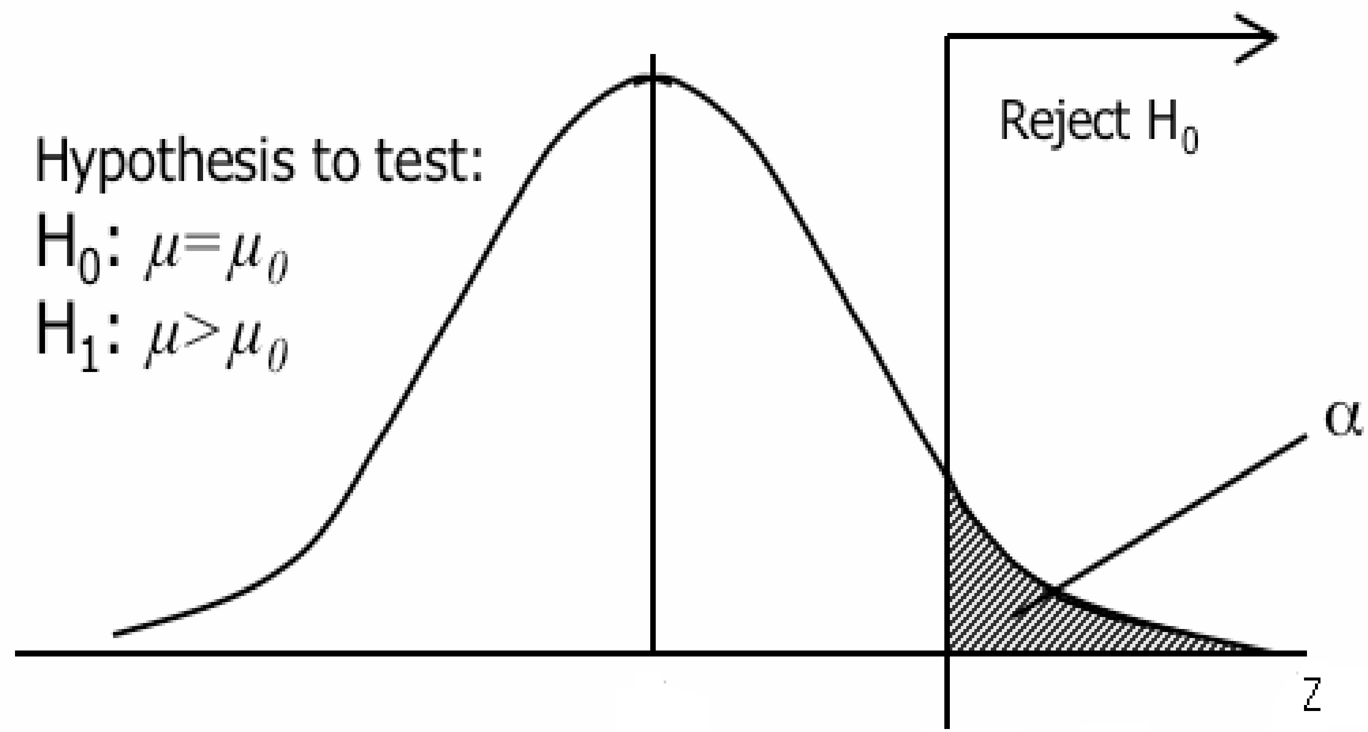
Μονόπλευρα και Αμφίπλευρα κριτήρια ελέγχου

Ανάλογα με το πώς εκφράζεται η εναλλακτική υπόθεση έχουμε μονοκατάληκτα και δικατάληκτα κριτήρια ελέγχου.

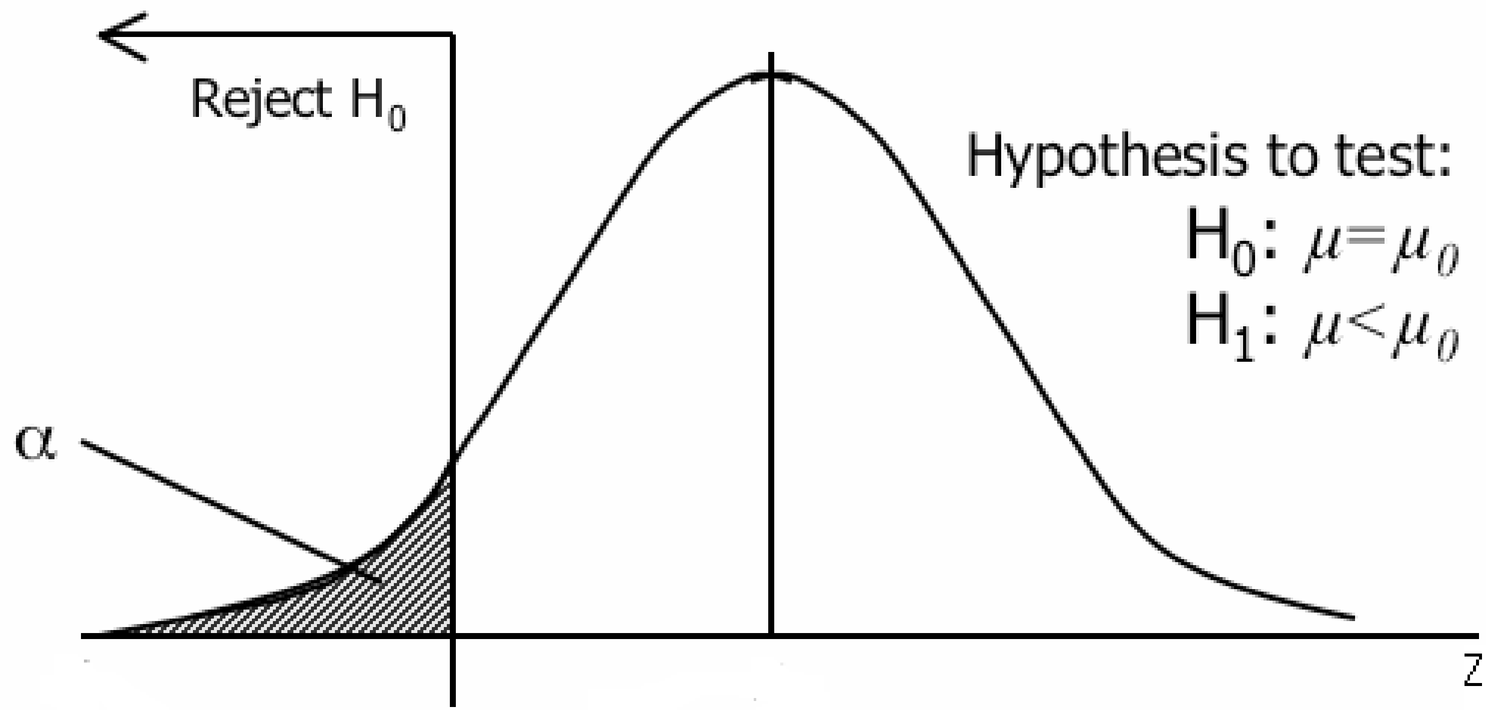
Δικατάληκτο κριτήριο χρησιμοποιείται όταν ελέγχουμε μία υπόθεση όπου η παράμετρος παίρνει μια τιμή διάφορη (≠) από μία άλλη.



Μονόπλευρο προς τα δεξιά (Right tail testing)



Μονόπλευρο αριστερά (Left tail testing)



Μέσος πληθυσμού όταν η τυπική απόκλιση είναι άγνωστη

Μέχρι τώρα εκτιμούσαμε τον μέσο του πληθυσμού γνωρίζοντας την τυπική απόκλιση του πληθυσμού χρησιμοποιώντας τον τύπο

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

Αλλά πόσο συχνά γνωρίζουμε την τυπική απόκλιση του πληθυσμού;

Στην περίπτωση όπου η τυπική απόκλιση του πληθυσμού είναι άγνωστη χρησιμοποιούμε για την εκτίμηση του μέσου το κριτήριο t-stundent

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

Μέσος πληθυσμού όταν η τυπική απόκλιση είναι άγνωστη

Όταν η σ είναι άγνωστη, χρησιμοποιούμε την τυπική απόκλιση του δείγματος ***s***

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} \quad \Longrightarrow \quad t = \frac{\bar{x} - \mu}{s / \sqrt{n}}$$

και η στατιστική z αντικαθίσταται από την στατιστική t, που ακολουθεί κατανομή t-student με βαθμούς ελευθερίας ν=n-1.

Ένα δείγμα:
2. Εκτίμηση του μ όταν σ είναι άγνωστη

Όταν η τυπική απόκλιση σ του πληθυσμού είναι άγνωστη και ο πληθυσμός ακολουθεί κανονική κατανομή, ο έλεγχος υποθέσεων για τον μέσο μ του πληθυσμού υπολογίζεται από τον τύπο:

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

Όταν η τυπική απόκλιση σ του πληθυσμού είναι άγνωστη και ο πληθυσμός ακολουθεί κανονική κατανομή, **ο εκτιμητής διαστήματος εμπιστοσύνης** του μέσου μ του πληθυσμού δίνεται από τον τύπο:

$$\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$$

Ένα δείγμα:
2. Εκτίμηση του μ όταν σ είναι άγνωστη

- Απαραίτητη προϋπόθεση για το παραπάνω είναι τα δεδομένα μας να είναι Κανονικά Κατανεμημένα ή το μέγεθος του δείγματος να είναι μεγάλο (n>50).
- Όταν το δείγμα μας είναι μεγάλο τότε η κατανομή του Student προσεγγίζει την κανονική κατανομή οπότε μπορούμε να θεωρήσουμε ότι το Tα ακολουθεί την N(0,1). Ο συγκεκριμένος έλεγχος καλείται **one sample t-test**
- Αν το μέγεθος του δείγματος δεν είναι μεγάλο και δεν ισχύει η κανονικότητα, τότε χρησιμοποιούμε το αντίστοιχο μη-παραμετρικό έλεγχο που θα εξετάσουμε παρακάτω.

Παράδειγμα

Ένας διαιτολόγος ισχυρίζεται ότι ο μέσος Έλληνας είναι πάνω από 20 κιλά βαρύτερος από το ιδανικό βάρος. Για τον έλεγχο αυτής της υπόθεσης επιλέχθηκε ένα τυχαίο δείγμα 20 ατόμων και καταγράφηκε η διαφορά ανάμεσα στο πραγματικό και το ιδανικό βάρος του καθενός. Μπορούμε από τα δεδομένα να συμπεράνουμε σε επίπεδο σημαντικότητας 5% ότι ο ισχυρισμός είναι αληθής;

16	23	18	41	22	18	23	19	22	15
18	35	16	15	17	19	23	15	16	26

Αναγνωρίστε την παράμετρο

Τα δεδομένα είναι ποσοτικά και το ζητούμενο είναι ο έλεγχος μιας υπόθεσης για το μέσο βάρος πάνω από το ιδανικό βάρος.

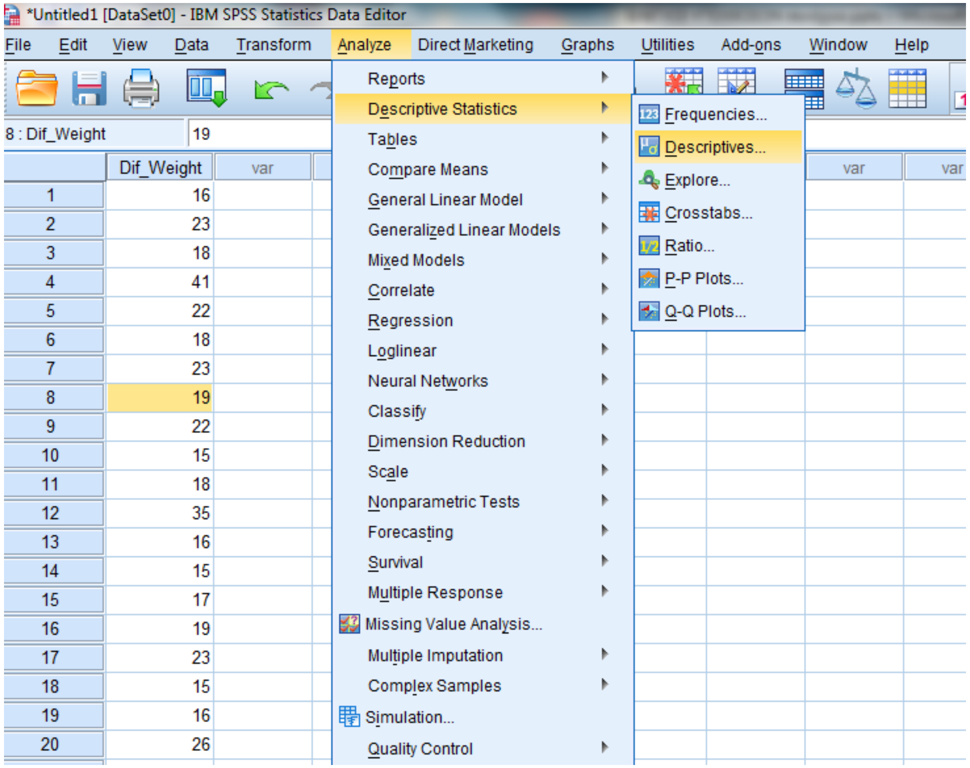
Θέλουμε να γνωρίζουμε εάν υπάρχουν αρκετές αποδείξεις ώστε να συμπεράνουμε ότι ο μέσος Έλληνας είναι πάνω από 20 κιλά υπέρβαρος. Έτσι,

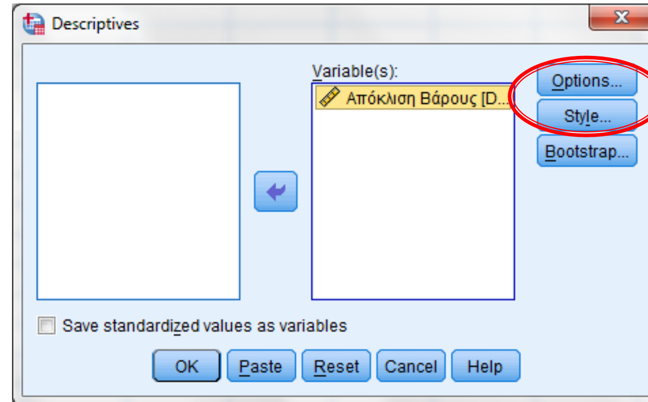
$$H_1: \mu > 20$$

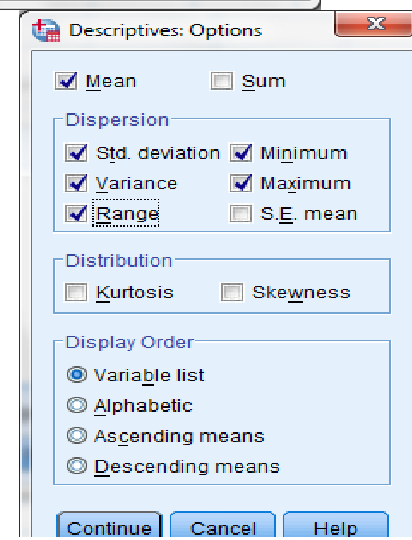
Ως εκ τούτου ορίζουμε την μηδενική υπόθεση ως:

$$H_0: \mu = 20$$

Ένα δείγμα:
2. Εκτίμηση του μ όταν σ είναι άγνωστη



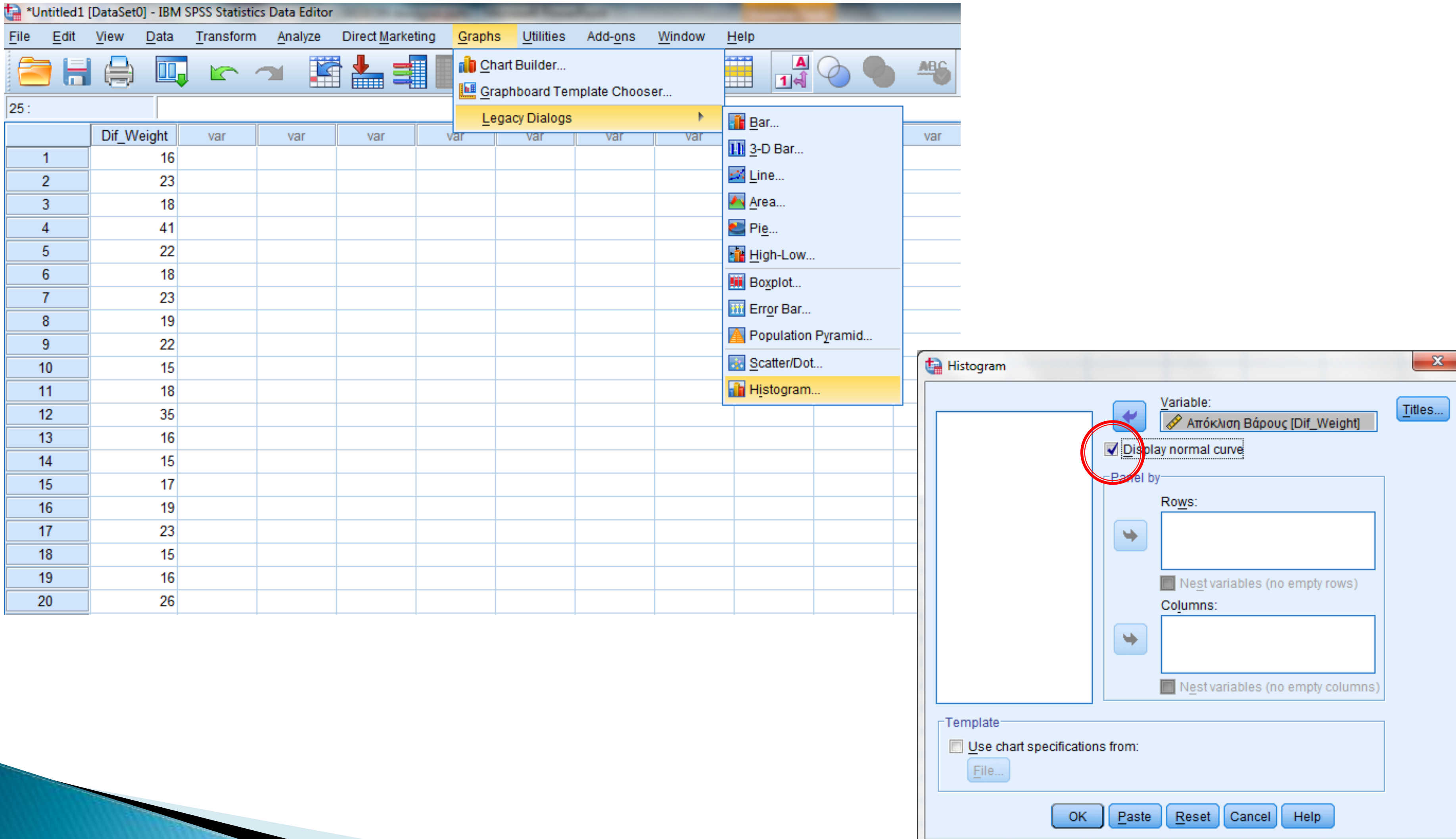




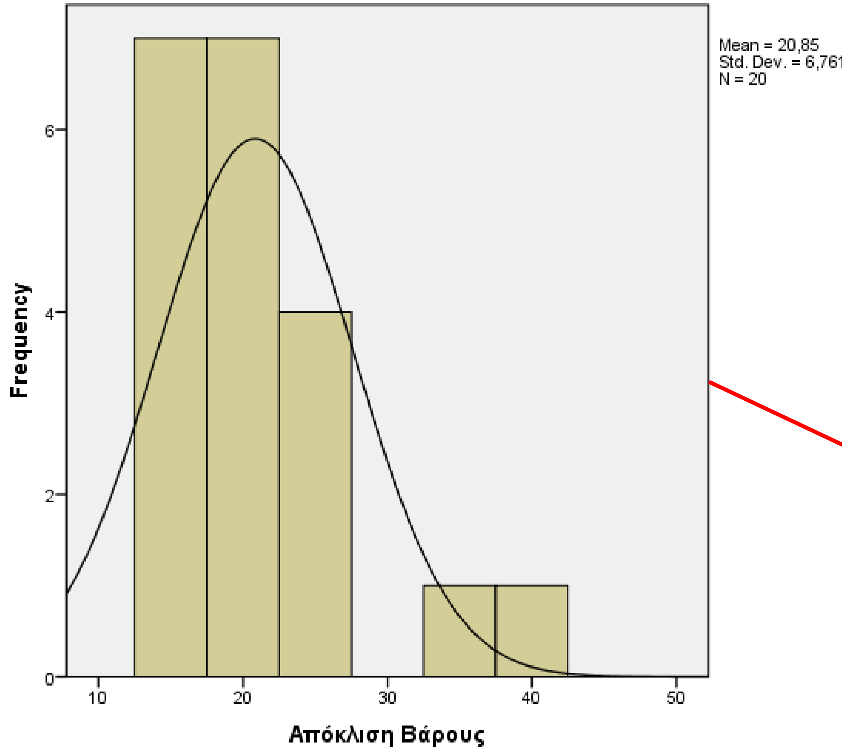
Descriptive Statistics

	N	Range	Minimum	Maximum	Mean	Std. Deviation	Variance
Απόκλιση Βάρους	20	26	15	41	20,85	6,761	45,713
Valid N (listwise)	20						

Ένα δείγμα:
2. Εκτίμηση του μ όταν σ είναι άγνωστη



Ένα δείγμα:
2. Εκτίμηση του μ όταν σ είναι άγνωστη



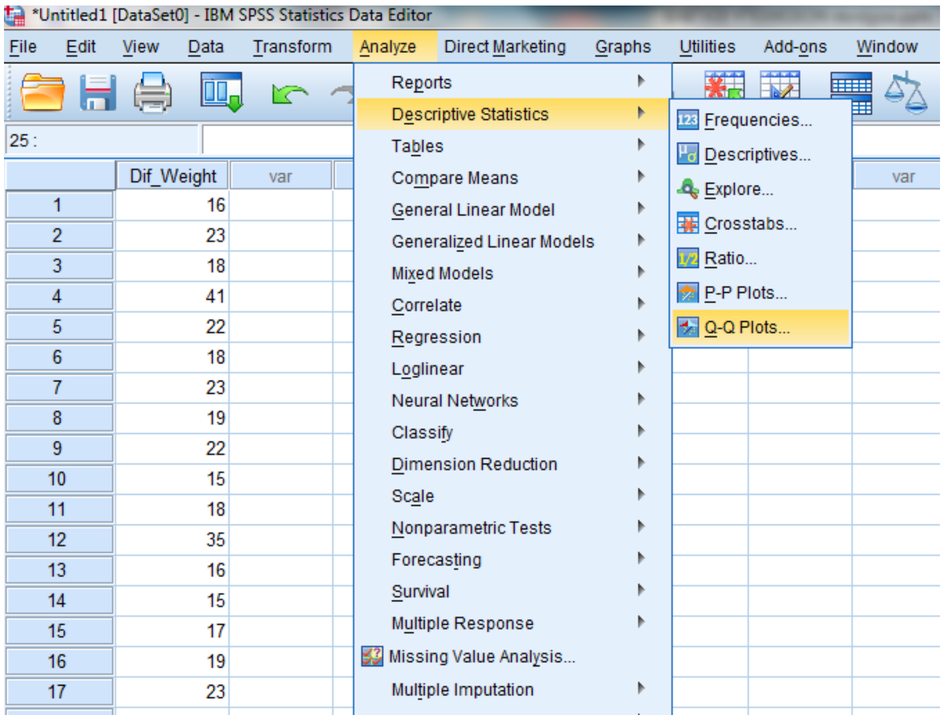
Επειδή το δείγμα μας δεν είναι μεγάλο, αρχικά ελέγχουμε αν η υπόθεση της κανονικότητας είναι λογική.

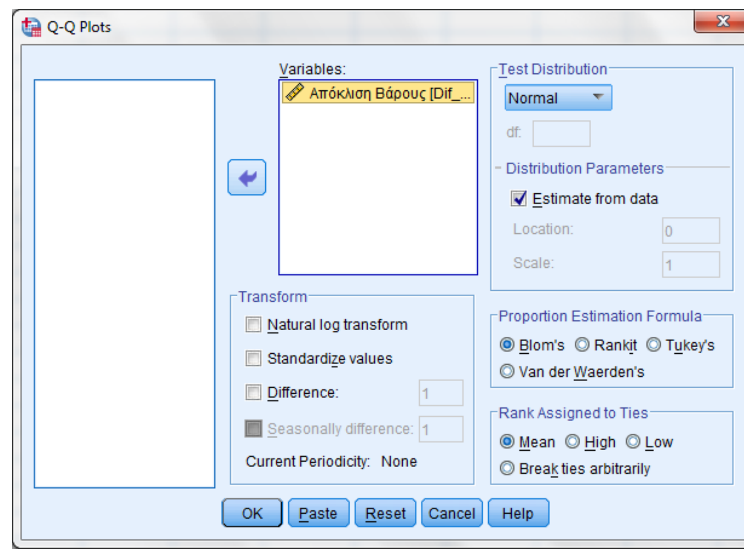
Η υπόθεση της κανονικότητας δεν φαίνεται και πολύ λογική

Ένα δείγμα:
2. Εκτίμηση του μ όταν σ είναι άγνωστη

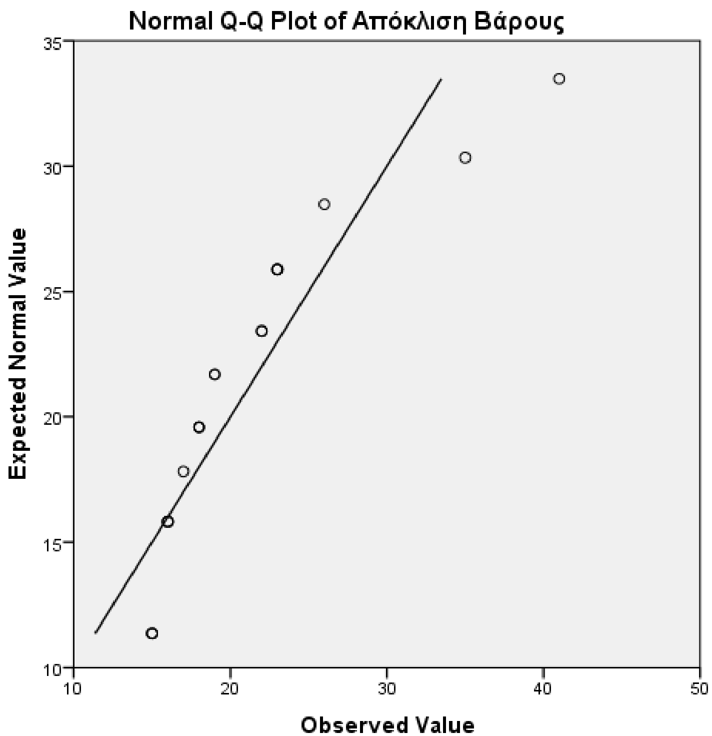
- Για να ελέγξουμε αν τα δεδομένα μας προέρχονται από την Κανονική κατανομή κάνουμε επίσης μία γραφική παράσταση των δειγματικών ποσοστημορίων ως προς τα θεωρητικά ποσοστημόρια της Κανονικής Κατανομής (**QQ – PLOT**). Όσο πιο κοντά στην γραμμή που αναπαριστά τα θεωρητικά ποσοστημόρια είναι τα σημεία που αναπαριστούν τα δειγματικά ποσοστημόρια , τόσο καλύτερη προσαρμογή έχουμε.
- Ισοδύναμα υπάρχει και το PP-PLOT που ελέγχει δειγματικές αθροιστικές πιθανότητες με αναμενόμενες αθροιστικές πιθανότητες με βάση την υπόθεση της κανονικής κατανομής. Και σε αυτή την περίπρτωση, όσο πιο κοντά στην γραμμή βρίσκονται τα σημεία τόσο καλύτερη προσαρμογή έχουμε.

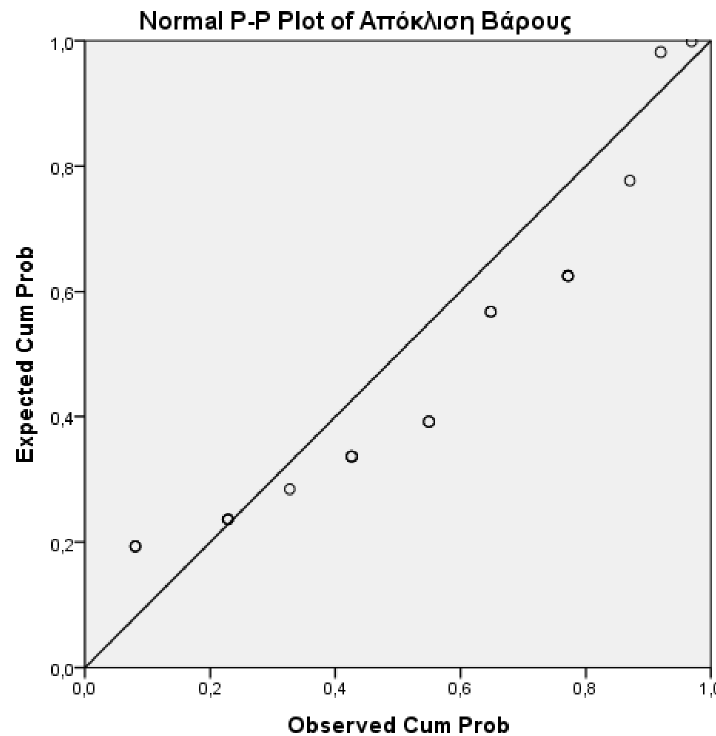
Ένα δείγμα:
2. Εκτίμηση του μ όταν σ είναι άγνωστη





Ένα δείγμα:
2. Εκτίμηση του μ όταν σ είναι άγνωστη





Η υπόθεση της κανονικότητας
δεν φαίνεται και πολύ λογική

Παράδειγμα

Υπολογίστε

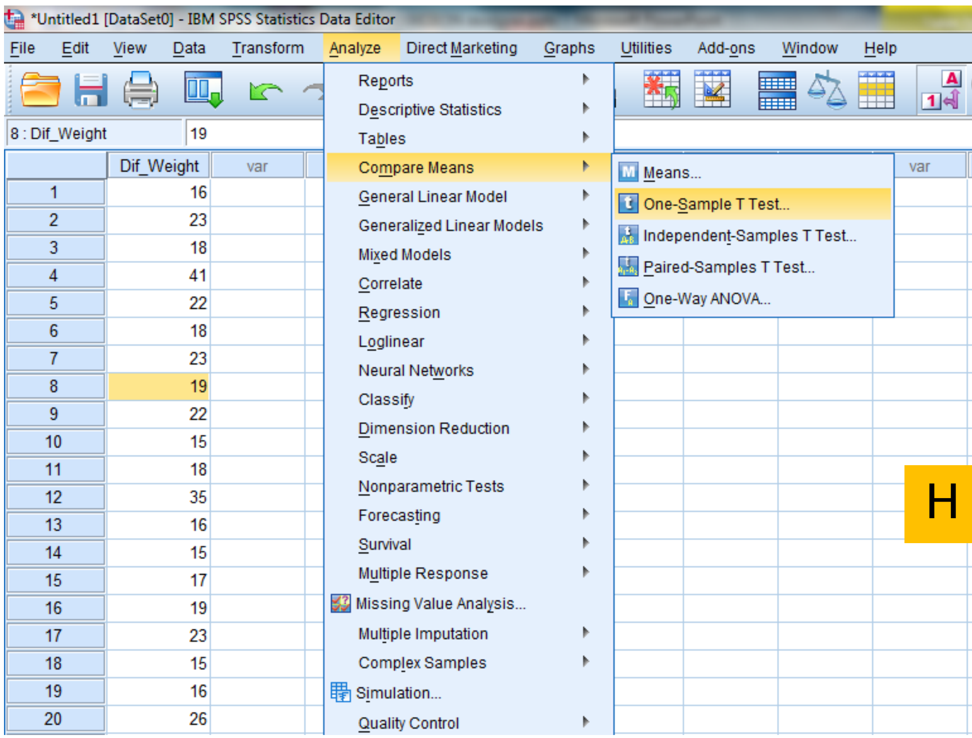
Επειδή η τυπική απόκλιση του πληθυσμού δεν είναι γνωστή, θα υπολογίσουμε τον μέσο χρησιμοποιώντας την στατιστική ***t*** που ακολουθεί *t-student* κατανομή με βαθμούς ελευθερίας ***ν***=η-1=20-1=19.

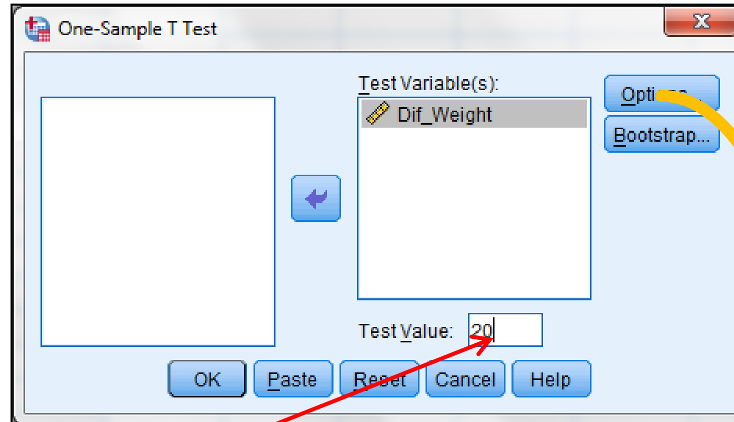
$$\bar{x} = \frac{\sum x_i}{n} = 20.85$$

$$s^2 = \frac{\sum (X_i - \bar{X})}{n-1} = \frac{\sum X_i^2 - \frac{(\sum X_i)^2}{n}}{n-1} = 45.69 \quad s = \sqrt{s^2} = \sqrt{45.69} = 6.76$$

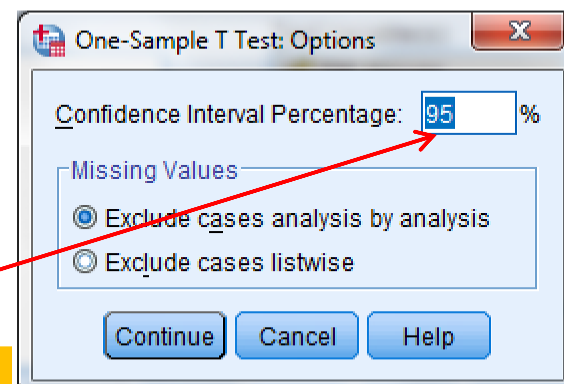
$$t = \frac{\bar{x} - \mu}{\mathrm{s}/\sqrt{\mathrm{n}}} = \frac{20.85 - 20}{6{,}76/\sqrt{20}} = 0.5622 \qquad t_{0.05,19} = 1.729$$

Ένα δείγμα : Με το SPSS





Η Τιμή που ελέγχουμε



Επίπεδο σημαντικότητας

Ένα δείγμα: Με το SPSS

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Dif_Weight	20	20,85	6,761	1,512

One-Sample Test

	Test Value = 20					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
Dif_Weight	,562	19	,581	,850	-2,31	4,01

Η τιμή της στατιστικής t

Βαθμοί ελευθερίας

P –τιμή του αμφίπλευρου ελέγχου

Διαφορά δειγματικού μέσου από την υποτιθέμενη τιμή κάτω από την H_0

95% Δ.Ε. της διαφοράς ου μέσου από την υποτιθέμενη τιμή κάτω από την H_0

Ένα δείγμα

Ερμηνεύστε

Η τιμή της t-statistics είναι t=0.5622 και η τιμή p είναι ίση με 0.29 (=0.58/2).

Δεν υπάρχουν αρκετές αποδείξεις ώστε να συμπεράνουμε ότι ο μέσος Έλληνας είναι υπέρβαρος πάνω από 20 κιλά.

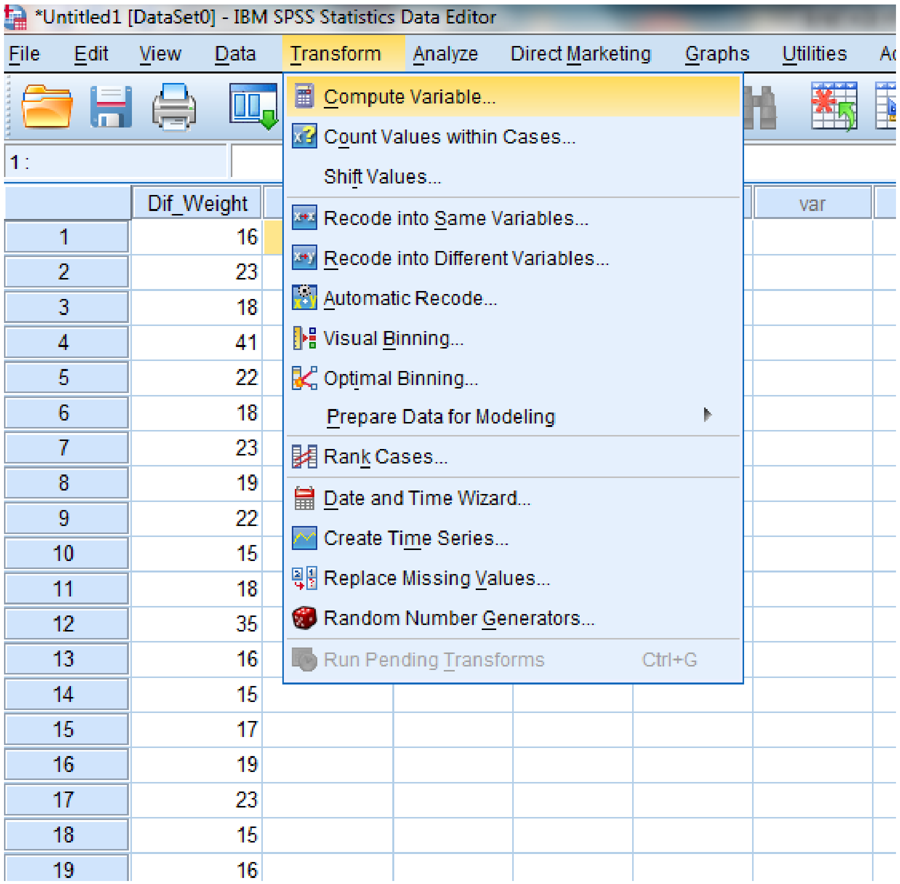
Παρατηρήστε ότι ο έλεγχος βρίσκεται αρκετά μακριά στο όριο περιοχής απόρριψης και η τιμή p-value είναι 0.29. Αυτό σημαίνει ότι η πιθανότητα να ισχύει η μηδενική υπόθεση και να έχουμε ένα δείγμα με τον μέσο που βρήκαμε μεγαλύτερο των 20 κιλών είναι 29%.

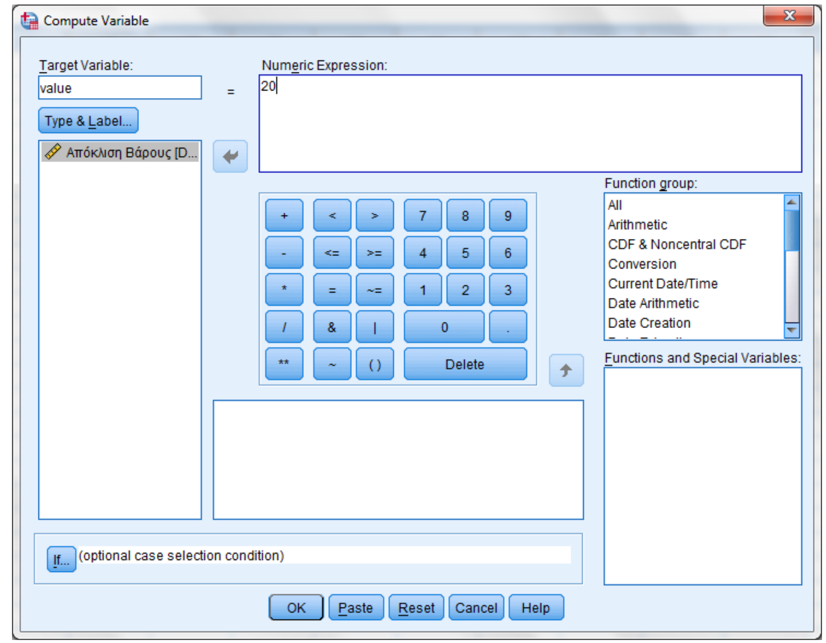
Συνεπώς, δεν μπορούμε να συμπεράνουμε ότι οι Έλληνες είναι υπέρβαροι πάνω από 20 κιλά.

Μη παραμετρικό τεστ

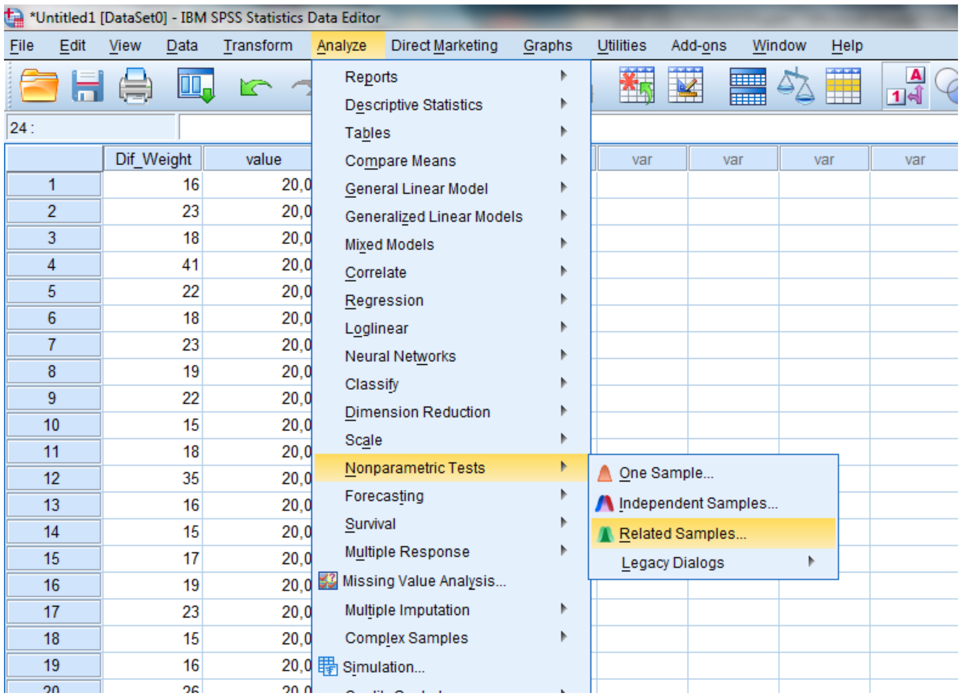
- Όταν το μέγεθος του δείγματος είναι μικρό και δεν ισχύει η υπόθεση της κανονικότητας, μπορούμε να εφαρμόσουμε τον αντίστοιχο μη-παραμετρικό έλεγχο που καλείται **Wilcoxon test.**
- Για την εφαρμογή του παραπάνω ελέγχου στο SPSS πρέπει να κατασκευάσουμε μία καινούργια μεταβλητή (π.χ. Value), στην οποία επαναλαμβάνουμε την τιμή κάτω από την H_0 (20 για το συγκεκριμένο παράδειγμα) για κάθε μονάδα του δείγματος

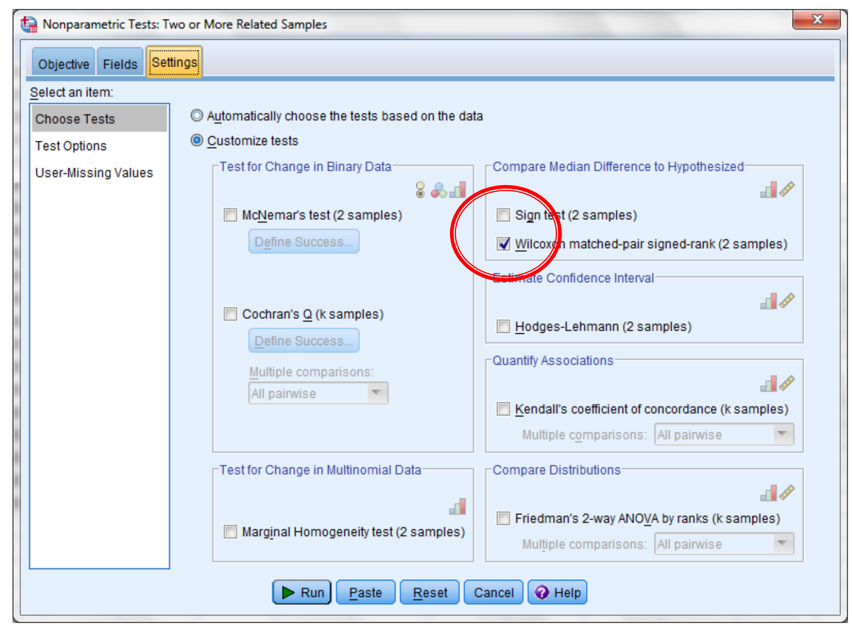
Μη παραμετρικό τεστ





Μη παραμετρικό τεστ



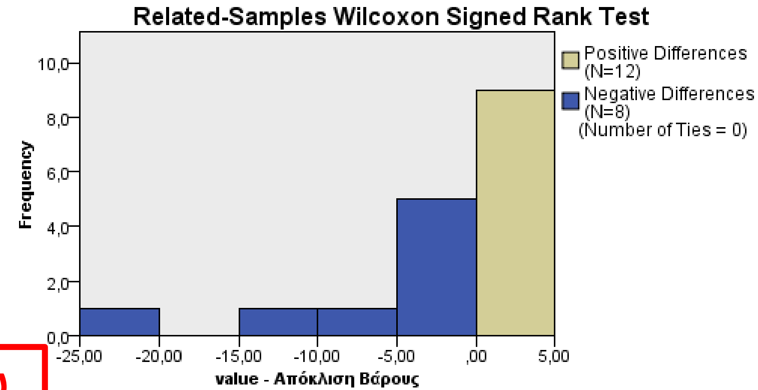


Μη παραμετρικό τεστ

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The median of differences between Απόκλιση Βάρους and value equals 0.	Related-Samples Wilcoxon Signed Rank Test	,722	Retain the null hypothesis.

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



Total N	20
Test Statistic	114,500
Standard Error	26,695
Standardized Test Statistic	,356
Asymptotic Sig. (2-sided test)	,722

Αποτέλεσμα του ελέγχου

Ασυμπτωτική P-τιμή

Η P-τιμή> 0.05 οπότε δεν έχουμε σοβαρές ενδείξεις εναντίον της μηδενικής υπόθεσης

Ένα δείγμα
3.Έλεγχος Ποσοστού *p*…

Αν τα δεδομένα μας είναι ονομαστικά, μπορούμε να μετρήσουμε μόνο σχετικές συχνότητες ή ποσοστά. Ως εκ τούτου, η παράμετρος που περιγράφει έναν πληθυσμό ονομαστικών δεδομένων είναι το ποσοστό του πληθυσμού ***p***.

Μία ονομαστική μεταβλητή μπορεί να έχει περισσότερες από δύο τιμές. Στην πράξη όμως ενδιαφερόμαστε για μία μόνο τιμή (ή ομάδα τιμών) την οποία ονομάζουμε «επιτυχία» ενώ τις άλλες τιμές τις ονομάζουμε «αποτυχία». Με τον τρόπο αυτό κάθε ονομαστική μεταβλητή μπορεί να θεωρηθεί ότι έχει μόνο δύο δυνατά αποτελέσματα. (π.χ. δημοσκοπήσεις πρόθεσης ψήφου, προτιμήσεις καταναλωτών κ.α.)

Ένα δείγμα
3.Έλεγχος Ποσοστού *p*...

Ο έλεγχος της υπόθεσης που χρησιμοποιείται για ένα ποσοστό p ενός πληθυσμού είναι:

$$z = \frac{\hat{p} - p}{\sqrt{p(1-p)/n}}$$

και ακολουθεί κατά προσέγγιση την κανονική κατανομή με την προϋπόθεση ότι **np>5** και **n(1-p)>5** .

$\hat{p} = \frac{x}{n}$ και **x** είναι ο αριθμός των επιτυχιών σε ένα δείγμα και **n** το μέγεθος του δείγματος.

Εκτιμητής διαστηματος Εμπιστοσύνης ποσοστού ενός πληθυσμού

Ο εκτιμητής διαστήματος εμπιστοσύνης για το ποσοστό ***p*** υπολογίζεται από τον τύπο:

$$\hat{p} \pm z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n}$$

(με την προϋπόθεση ότι ***np***$>$5 και ***n(1-p)***$>$5)

Ένα δείγμα
3.Έλεγχος Ποσοστού *p*…

Επειδή στο SPSS δεν υπάρχει η στατιστική z, ισοδύναμα θα μπορούσαμε να είχαμε κατασκευάσει έναν πίνακα συχνοτήτων και να ελέγχαμε αν οι παρατηρηθείσες συχνότητες (δείγμα) διαφέρουν από τις αναμενόμενες (np_0 και $n(1-p_0)$) υπολογίζοντας την παρακάτω στατιστική ελέγχου:

$$x^2 = \sum \frac{(\text{παρατηρηθείσες συχνότητες} - \textit{αναμενόμενες συχνότητες})^2}{(\textit{αναμενόμενες συχνότητες})}$$

Ένα δείγμα
3.Έλεγχος Ποσοστού *p*…

- Η παραπάνω στατιστική ελέγχου κάτω από την μηδενική υπόθεση ακολουθεί μία x^2 κατανομή με ένα βαθμό ελευθερίας, οπότε η p-τιμή του ελέγχου προκύπτει όπως πριν.
- Για να εφαρμοστεί ο προηγούμενος έλεγχος πρέπει οι αναμενόμενες συχνότητες να είναι τουλάχιστον 5

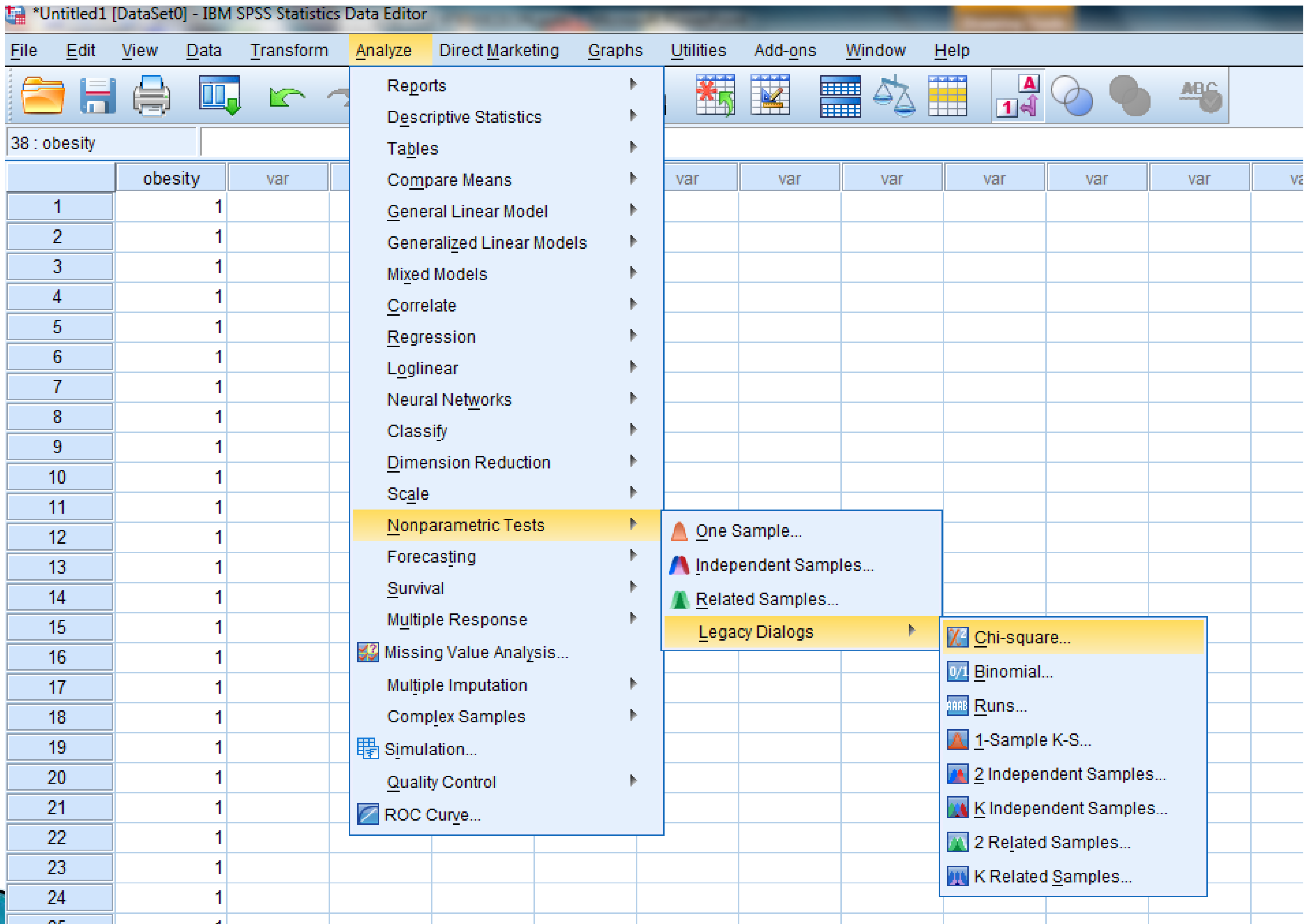
Ένα δείγμα
3.Έλεγχος Ποσοστού *p*…

Παράδειγμα: Από προηγούμενη πανελλαδική έρευνα είναι γνωστό ότι το 25% των Ελλήνων πολιτών είναι υπέρβαροι. Για να ελέγξουμε αυτή την υπόθεση οργανώνουμε μία νέα έρευνα όπου σε ένα τυχαίο δείγμα n=100 Ελλήνων πολιτών βρέθηκαν κ=30 υπέρβαροι ($\hat{p}=30/100$).

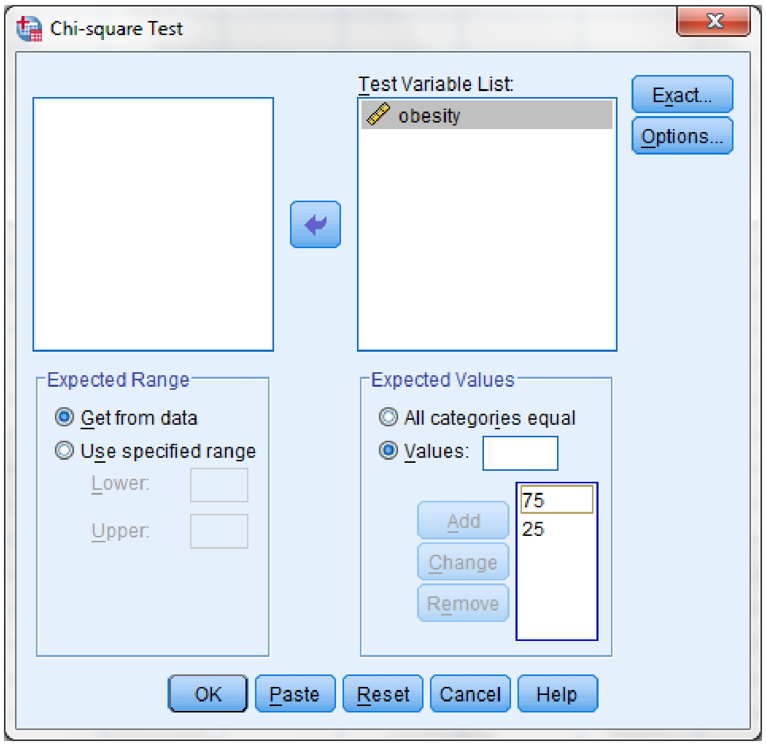
H_0: $p=0.25$ με εναλλακτική H_1: $p\neq 0.25$

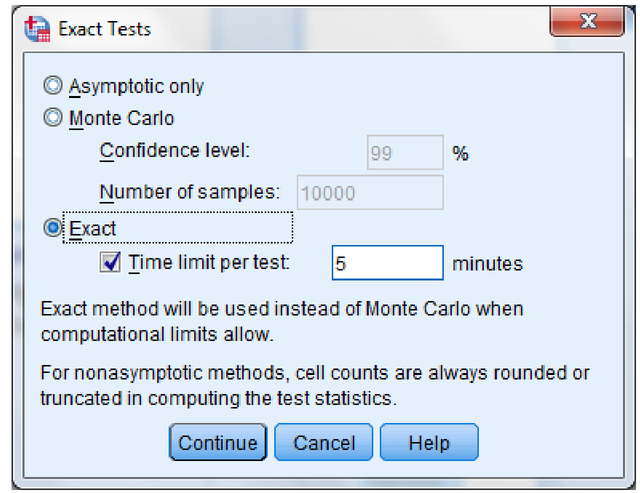
Παρατηρούμε ότι np_0-25 και $n(1-p_0)=75$

Ένα δείγμα
3.Έλεγχος Ποσοστού *p*...



Ένα δείγμα
3.Έλεγχος Ποσοστού *p*…





Ένα δείγμα
3.Έλεγχος Ποσοστού *p*...

obesity

	Observed N	Expected N	Residual
0	70	75,0	-5,0
1	30	25,0	5,0
Total	100		

Test Statistics

	obesity
Chi-Square	1,333[a]
df	1
Asymp. Sig.	,248
Exact Sig.	,298
Point Probability	,095

a. 0 cells (0,0%) have expected frequencies less than 5. The minimum expected cell frequency is 25,0.

Συχνότητες από το δείγμα

Αναμενόμενες συχνότητες

X2 στατιστική

P-τιμή του ακριβή αμφίπλευρου ελέγχου

Έλεγχος των προϋποθέσεων

3. Έλεγχος Ποσοστού *p*… Αντιστοιχία z-test και x^2-test

z-Test: Proportion			
			obesity
Sample Proportion			0.3
Observations			100
Hypothesized Proportion			0.25
z Stat			**1.1547**
P(Z<=z) one-tail			0.1241
z Critical one-tail			1.6449
P(Z<=z) two-tail			**0.2482**
z Critical two-tail			1.96

Test Statistics

	obesity
Chi-Square	1,333[a]
df	1
Asymp. Sig.	,248
Exact Sig.	,298
Point Probability	,095

a. 0 cells (0,0%) have expected frequencies less than 5. The minimum expected cell frequency is 25,0.

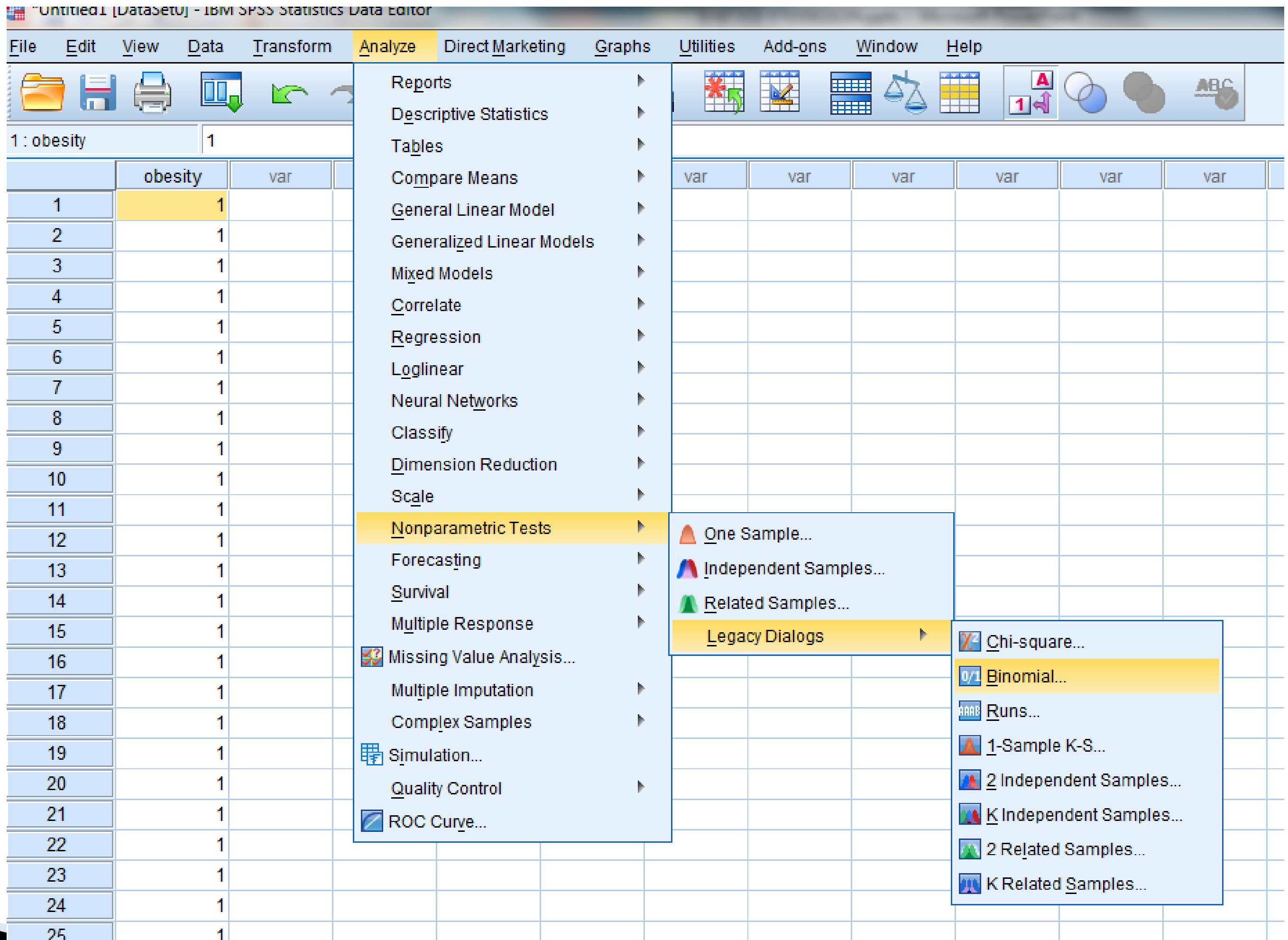
$$\boldsymbol{z = \sqrt{x^2} \Rightarrow z = \sqrt{1,333} \Rightarrow z = 1.154}$$

Ένα δείγμα
3.Έλεγχος Ποσοστού *p*…

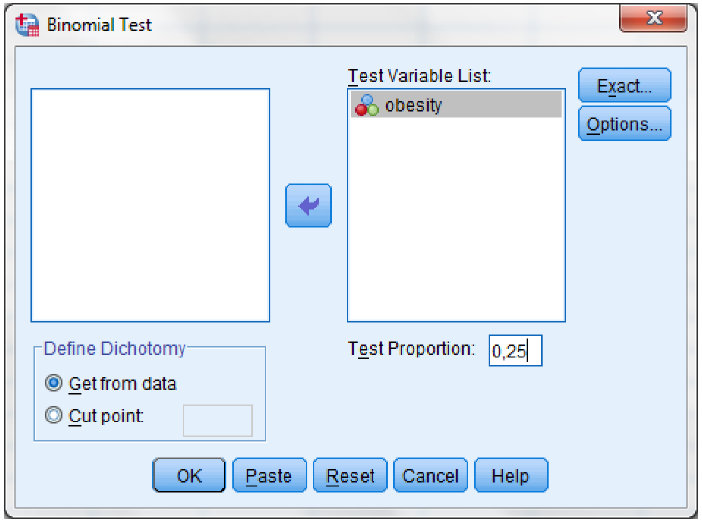
- Από τα αποτελέσματα του παραπάνω ελέγχου καταλήγουμε ότι, σε ε.σ. 5% δεν έχουμε σοβαρές ενδείξεις εναντίον της μηδενικής υπόθεσης, οπότε δεν την απορρίπτουμε. Δηλαδή, δεχόμαστε ότι το ποσοστό των παχύσαρκων Ελλήνων είναι 25%.

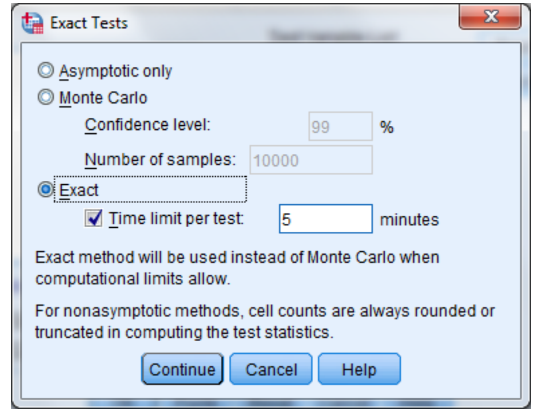
- Όταν δεν ισχύουν οι προϋποθέσεις (οι αναμενόμενες συχνότητες να είναι τουλάχιστον 5) θα πρέπει να εφαρμόσουμε το Διωνυμικό κριτήριο (**Binominal test**)

Ένα δείγμα
3.Έλεγχος Ποσοστού *p*...



Ένα δείγμα
3.Έλεγχος Ποσοστού *p*...





Binomial Test

		Category	N	Observed Prop.	Test Prop.	Exact Sig. (1-tailed)	Exact Sig. (1-tailed)
obesity	Group 1	1	30	,30	,25	,150	,150
	Group 2	0	70	,70			
	Total		100	1,00			

Ακριβής P-τιμή μονόπλευρου ελέγχου
Η ακριβής τιμή αμφίπλευρου ελέγχου είναι ίση με 2 Χ 0.15=0.30

Παράδειγμα με Πίνακα Συχνοτήτων

Η κυβέρνηση πριν πάρει την απόφαση να διεξάγει ένα δημοψήφισμα σχετικά με την υπογραφή ή όχι ενός τρίτου Μνημονίου διεξάγει μία έρευνα σχετικά με το εάν οι Έλληνες ναι υπέρ ή κατά με την ενδεχόμενη υπογραφή ενός Μνημονίου Νο 3. Οι απαντήσεις είναι μόνο δύο 1. ΝΑΙ και 2. ΌΧΙ. Από πανελλαδικό δείγμα 765 ατόμων ο αριθμός των ΟΧΙ είναι 407. Μπορεί η κυβέρνηση σε επίπεδο σημαντικότητας 5% να συμπεράνει ότι ο Ελληνικός λαός είναι κατά της υπογραφής ενός νέου μνημονίου;

Πράδειγμα

Αναγνωρίστε την παράμετρο

Το πρόβλημα είναι η περιγραφή του πληθυσμού των ψήφων σε ολόκληρη την Ελλάδα. Τα δεδομένα είναι ονομαστικά καθώς οι τιμές είναι «ΝΑΙ» και «ΌΧΙ» Η παράμετρος που πρέπει να εξεταστεί είναι το ποσοστό μίας από τις δύο απαντήσεις. Για να πάρει η κυβέρνηση την απόφαση να πάει σε ένα δημοψήφισμα ελέγχει την υπόθεση:

$H_1: p>0.5 \quad H_0: p=0.5$

Θα χρησιμοποιήσει την στατιστική

$$z = \frac{\hat{p} - p}{\sqrt{p(1-p)/n}}$$

Η κατανομή είναι κατά προσέγγιση η τυποποιημένη κανονική κατανομή και για στάθμη σημαντικότητας 5% θα χρησιμοποιηθεί το διάστημα: $z > z_{\alpha} = z_{0.05} = 1.645$

Παράδειγμα

Υπολογισμός

n=765 και ο αριθμός των επιτυχιών (ψήφων υπέρ του ΌΧΙ) είναι x=407. Έτσι το ποσοστό υπέρ του ΌΧΙ στο δείγμα είναι:

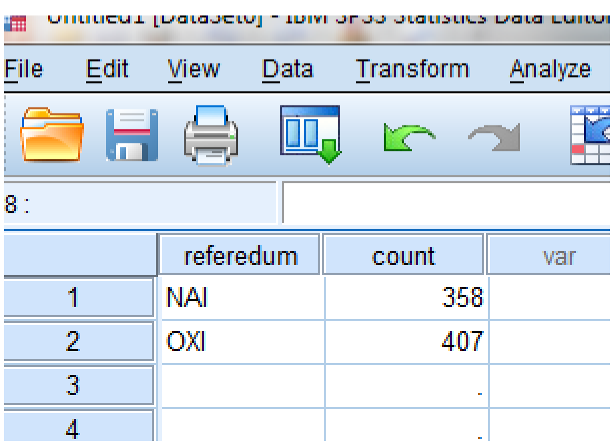
$$\widehat{p} = \frac{x}{n} = \frac{407}{765} = 0.532$$

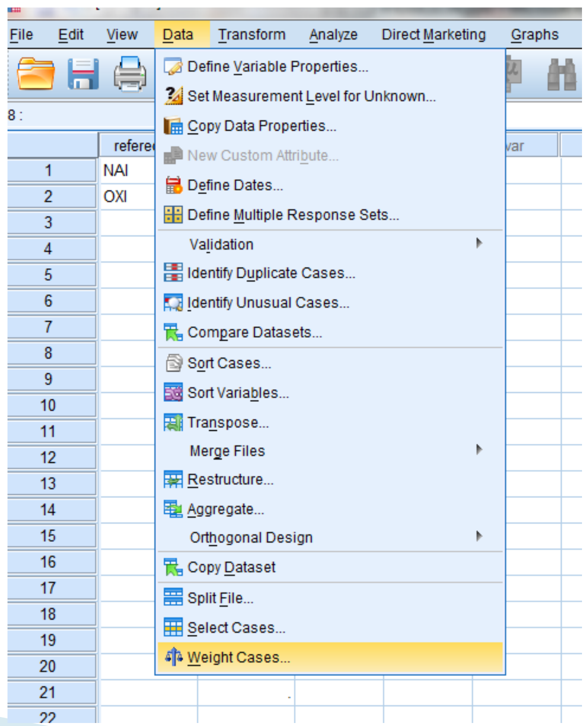
$$z = \frac{\widehat{p} - p}{\sqrt{p(1-p)/n}} = \frac{0.532 - 0.5}{\sqrt{0.532(1-0.5)/765}} = 1.77$$

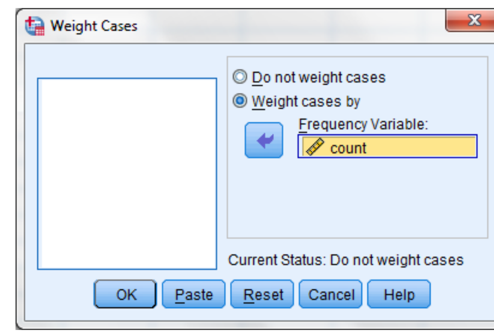
Έλεγχος ποσοστού με x^2 στατιστική στο SPSS

Υπολογισμός

Βάζουμε τις κατηγορίες του δημοψηφίσματος (ΝΑΙ και ΌΧΙ) σε μία μεταβλητή (Referendum) και τις παρατηρήσεις μας σε μία άλλη μεταβλητή (Count) (….αριστερή εικόνα). Κατόπιν σταθμίζουμε τις κατηγορίες της ονομαστικής μεταβλητής με τις παρατηρήσεις.

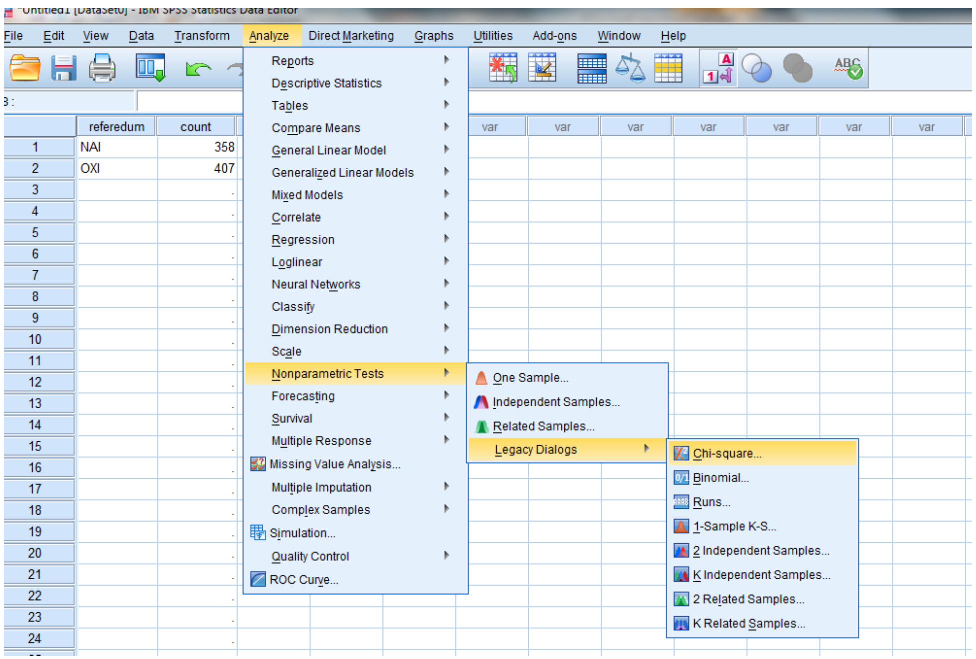


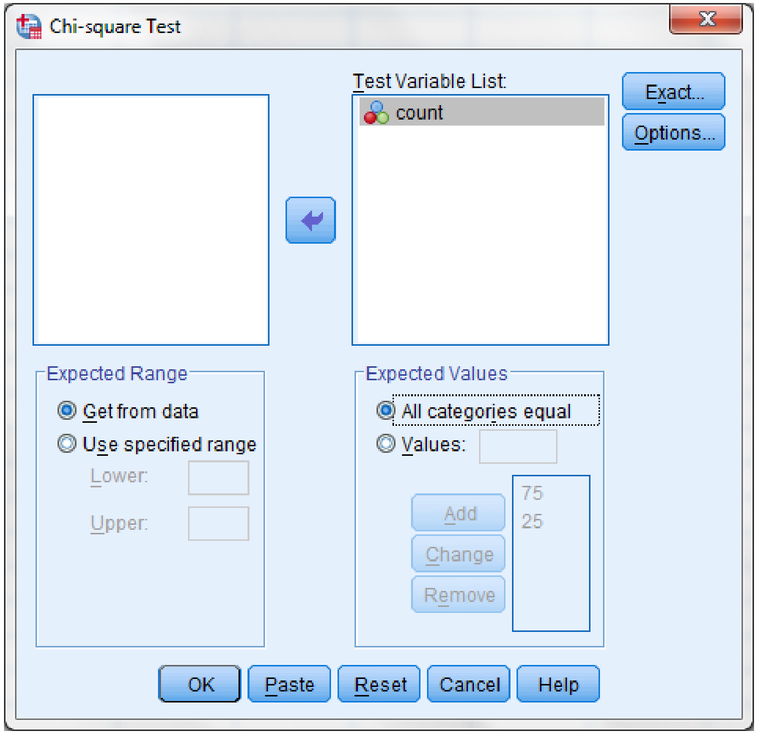




Έλεγχος ποσοστού με x^2 στατιστική στο SPSS

Επιλέγουμε, Analyze>Non parametric tests>Legacy Dialogs>Chi square…





Έλεγχος ποσοστού με x^2 στατιστική στο SPSS

Υπολογισμός

count

	Observed N	Expected N	Residual
358	358	382,5	-24,5
407	407	382,5	24,5
Total	765		

Test Statistics

	count
Chi-Square	3,139[a]
df	1
Asymp. Sig.	,076
Exact Sig.	,083
Point Probability	,012

a. 0 cells (0,0%) have expected frequencies less than 5. The minimum expected cell frequency is 382,5.

Παράδειγμα

Ερμηνεία

Ο έλεγχος είναι z=1.77 ($z = \sqrt{3.139} = 1.77$ και η τιμή p=0.0384. Σε επίπεδο σημαντικότητας 5% μπορούμε να απορρίψουμε την υπόθεση μηδέν και να συμπεράνουμε ότι οι Έλληνες είναι κατά της υπογραφής ενός νέου μνημονίου.

Ένα σημαντικό θέμα σε κάθε έλεγχο υπόθεσης είναι το κόστος των σφαλμάτων I και II. Αυτό που θα ήθελε να αποφύγει η Ελληνική Κυβέρνηση είναι να δεχθεί την υπόθεση μηδέν που είναι λανθασμένη (σφάλμα τύπου II) και να προχωρούσε σε δημοψήφισμα όπου το αποτέλεσμα θα ήταν η μη υπογραφή ενός νέου μνημονίου.

Αν η δειγματοληπτική έρευνα έθετε επίπεδο σημαντικότητα α=1% ($z_{0.01}$=2.575) σε αυτή την περίπτωση θα είχαμε αποδοχή της μηδενικής απόφασης και διεξαγωγή δημοψηφίσματος. Έτσι, όσο μεγαλύτερο επίπεδο σημαντικότητας διάλεγε η Ελληνική Κυβέρνηση για να ελέγξει την υπόθεση, τόσο μικρότερη θα είναι η πιθανότητα να πέσει σε Σφάλμα τύπου II

Επιλογή μεγέθους δείγματος

Αν η παράμετρος που πρέπει να εκτιμηθεί είναι ένα ποσοστό, το όριο σφάλματος εκτίμησης δίνεται από τον τύπο:

$$z_{a/2}\sqrt{p(1-p)/n} = \widehat{p} - p = B$$

Αν επιλύσουμε ως προς n βρίσκουμε την έκφραση για το επιθυμητό μέγεθος του δείγματος

$$n = [\frac{z_{a/2}\sqrt{\widehat{p}(1-\widehat{p}}}{B}]^2$$

Επιλογή μεγέθους δείγματος

Όπως έχουμε δεί σε πολλά παραδείγματα βρίσκουμε διαδοχικά
1−α=0.95, α=0.05, α/2=0.025, $z_{\alpha/2}=z_{0.025}=1.96$ Έτσι:

$$n=[\frac{z_{a/2}\sqrt{\widehat{p}(1-\widehat{p})}}{B}]^2=[\frac{1.96\sqrt{\widehat{p}(1-\widehat{p})}}{B}]$$

Ας υποθέσουμε ότι μία εταιρεία θέλει να εκτιμήσει το ποσοστό των καταναλωτών που προτιμούν το δικό της προϊόν, με επίπεδο εμπιστοσύνης 95% και μέγιστη απόκλιση 0.03. Αν γνωρίζει ότι το μερίδιο αγοράς στην αγορά είναι στο 20%, μπορούμε να θέσουμε p=0. 2 και να βρούμε

$$n==[\frac{1.96\sqrt{0.2(1-0.2)}}{0.03}]^2=683$$