

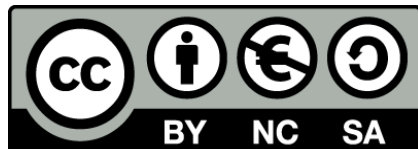


ΤΕΙ Αθήνας

Μεθοδολογία της έρευνας και Ιατρική στατιστική

Ενότητα 3: Έλεγχοι υποθέσεων - Διαστήματα εμπιστοσύνης

Δρ.Ευσταθία Παπαγεωργίου, Αναπληρώτρια Καθηγήτρια





Οι ερευνητικές υποθέσεις

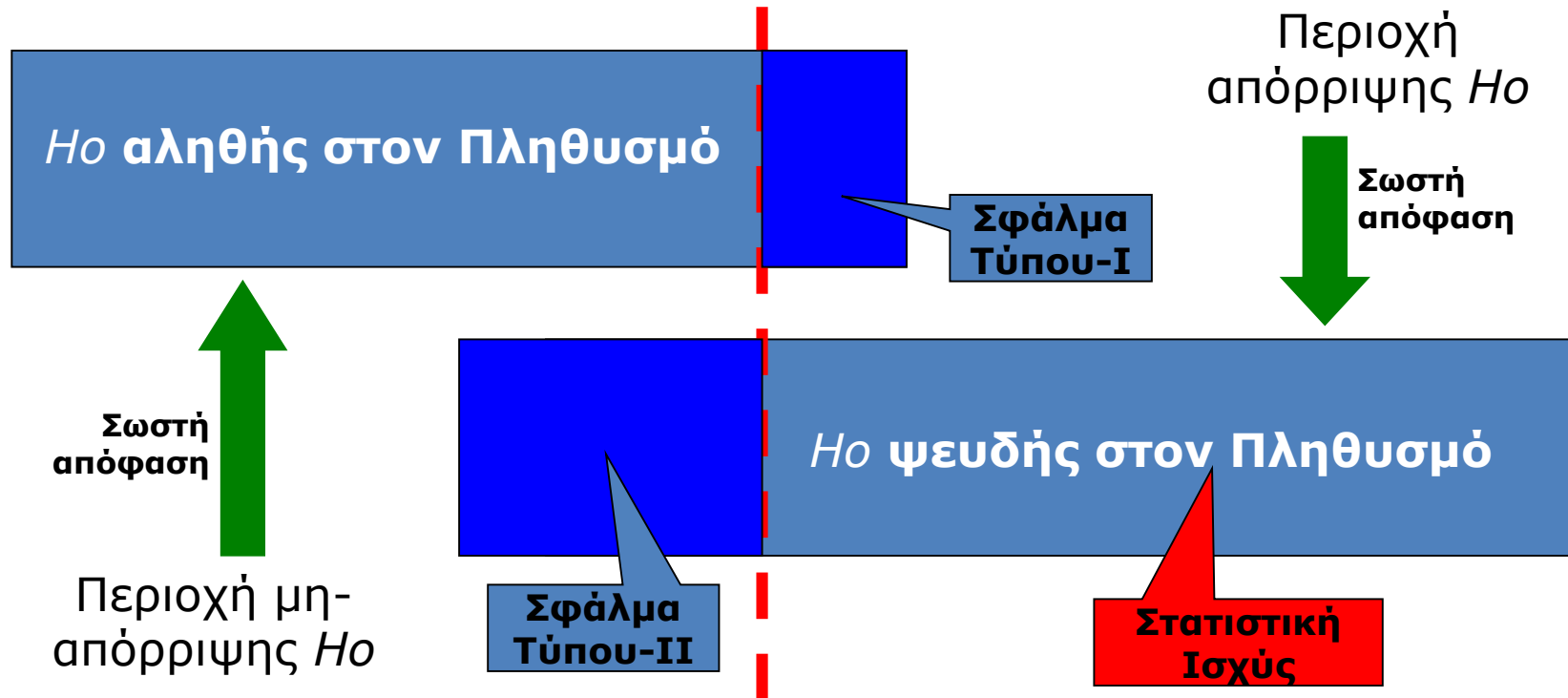
- Στην έρευνα ελέγχουμε υποθέσεις, με βάση τα πραγματικά δεδομένα μας.
 - π.χ. ο μεγάλος χρόνος εισόδου στο Νοσοκομείο από την έναρξη των συμπτωμάτων, συσχετίζεται με αυξημένο κίνδυνο θανάτου;
 - Μια διατροφή πλούσια σε υδατάνθρακες συσχετίζεται με μειωμένο σωματικό βάρος;

- Η διαδικασία που ακολουθείται για την λήψη τέτοιου είδους αποφάσεων ονομάζεται **έλεγχος υποθέσεων**.
- Η υπόθεση που θέλουμε να ελέγξουμε συμβολίζεται με H_0 και ονομάζεται μηδενική υπόθεση ενώ η εναλλακτική της υπόθεση συμβολίζεται με H_1 .
- Σε κάθε έλεγχο είναι δυνατόν να πραγματοποιηθούν δύο ειδών σφάλματα:
 - **Σφάλμα τύπου I:** Απόρριψη της H_0 ενώ στην πραγματικότητα είναι αληθής.
 - **Σφάλμα τύπου II:** Απόρριψη της H_1 (Αποδοχή της H_0) ενώ στην πραγματικότητα η H_1 είναι αληθής.

Σφάλματα στη λήψη απόφασης

	Αποδοχή υπόθεσης H_0 από το δείγμα	Απόρριψη υπόθεσης H_0 από το δείγμα
Υπόθεση H_0 αληθής στον πληθυσμό		Σφάλμα τύπου I
Υπόθεση H_0 ψευδής στον πληθυσμό	Σφάλμα τύπου II	 Στατιστική ισχύς

Έλεγχοι Υποθέσεων



Τιμές στατιστικού κριτηρίου

- $\alpha = P(\text{σφάλμα τύπου I}) = P(\text{Απόρριψη της } H_0 \text{ ενώ στην πραγματικότητα είναι αληθής})$
- $\beta = P(\text{σφάλμα τύπου II}) = P(\text{Αποδοχή της } H_0 \text{ ενώ στην πραγματικότητα η } H_1 \text{ είναι αληθής})$
- Η πιθανότητα $\gamma = 1 - \beta$ ονομάζεται ισχύς του ελέγχου και εκφράζει το ποσοστό των «σωστών» απορρίψεων της H_0 .

Το α ονομάζεται επίπεδο σημαντικότητας.

Ερμηνεία του α :

Εάν για παράδειγμα σε έναν έλεγχο επιλέξουμε επίπεδο σημαντικότητας $\alpha = 0.05$ και απορρίψουμε την υπόθεση, αυτό σημαίνει ότι σε 100 όμοιες περιπτώσεις, είναι δυνατό έχουμε κάνει λάθος και να απορρίψουμε την H_0 ενώ είναι αληθής, μόνο σε 5. Σε μια τέτοια περίπτωση λέμε ότι η υπόθεση απορρίπτεται σε επίπεδο σημαντικότητας 0.05

Statistical Tests – Confidence Intervals

- Κριτήριο για την αποδοχή ή όχι της H_0 είναι το p-value.
- Το μικρότερο επίπεδο σημαντικότητας για το οποίο απορρίπτεται η H_0 ονομάζεται p-value.
- Απορρίπτεται η H_0 αν η τιμή του p-value είναι μικρή.
- Συγκεκριμένα, απορρίπτεται η H_0 αν η τιμή του p-value είναι μικρότερη του α για αυτό το επίπεδο σημαντικότητας.
- Όσο μειώνεται το α τόσο δυσκολεύει η απόφαση της απόρριψης.

Τι δεν είναι το p -value

- Το p -value **δεν** είναι η πιθανότητα να επαληθευθεί η μηδενική υπόθεση
 - και αυτό γιατί οι υποθέσεις δεν εκφράζονται με πιθανότητες στην στατιστική.

Τι δεν είναι το p -value

- Το p -value **δεν** είναι η πιθανότητα να απορριφθεί λανθασμένα η μηδενική υπόθεση.
 - Το να απορριφθεί λανθασμένα η μηδενική υπόθεση είναι το σφάλμα Τύπου I.
 - ✓ Αυτό το σφάλμα είναι μια εκδοχή της καλούμενης «σφάλμα του εισαγγελέα» (“prosecutor's fallacy”) όπου κρίνει αθώο τον κατηγορούμενο ενώ έχει διαπράξει το έγκλημα.
 - Το σφάλμα Τύπου I είναι στενά συνυφασμένο με το p -value, αφού απορρίπτουμε τη μηδενική υπόθεση όταν το p -value είναι μικρότερο από κάποιο προκαθορισμένο όριο α (επίπεδο σημαντικότητας) του σφάλματος τύπου-I.

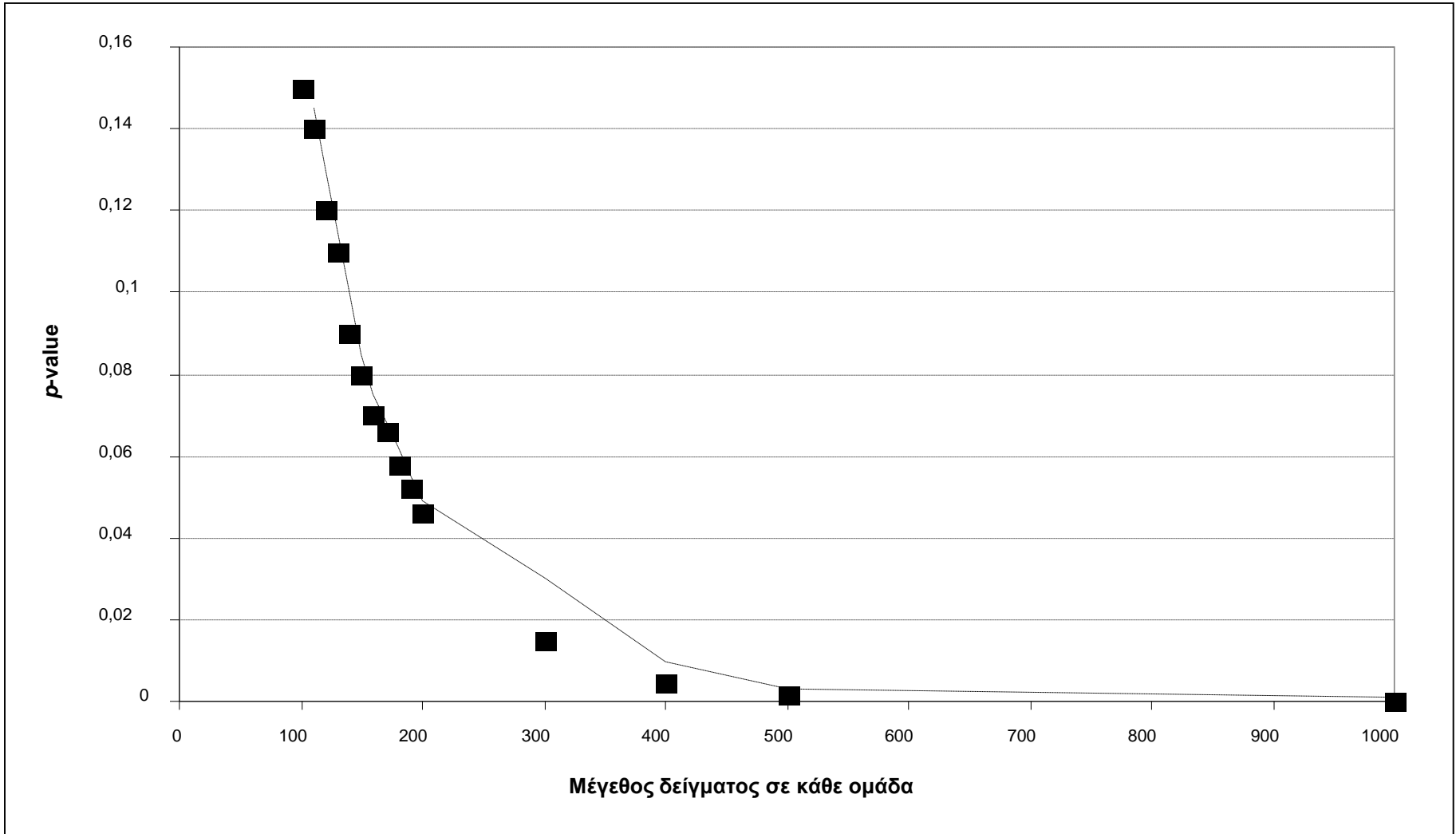
p -value και μέγεθος του δείγματος

- Το p -value επηρεάζεται ισχυρά από το μέγεθος του δείγματος.

Συγκεκριμένα

- Υπάρχει **αντίστροφη συσχέτιση** μεταξύ του μεγέθους δείγματος και του p -value.

p-value και μέγεθος του δείγματος για μια δεδομένη συσχέτιση



Το μέγεθος του δείγματος

- Το **επαρκές** μέγεθος του δείγματος είναι μεγίστης σημασίας για την αξιοπιστία της έρευνας.

Οι «αρχές» της δειγματοληψίας

- Πρέπει όμως να ληφθεί υπόψη ότι **σχετικά μεγάλο δείγμα** συνεπάγεται και μεγάλο κόστος
 - **χωρίς αυτό να σημαίνει και απαραίτητα αξιόπιστα αποτελέσματα,**
- ενώ **πολύ μικρό δείγμα** μπορεί να οδηγήσει σε **συστηματικό σφάλμα** και **μεροληπτικές** αποφάσεις για τον πληθυσμό.

Το μέγεθος του δείγματος καθορίζεται από:

- Το επίπεδο στατιστικής σημαντικότητας των ελέγχων, το οποίο συμβολίζεται με α και στο χώρο των επιστημών έχει καθοριστεί να είναι $< 0,01$ ή $< 0,05$.
- Το μέγεθος της αναζητούμενης σχέσης, π.χ. πόσο μεγάλη θα πρέπει να είναι η διαφορά στα επίπεδα ολικής χοληστερόλης μεταξύ της θεραπευτικής προσέγγισης A και της θεραπευτικής προσέγγισης B έτσι ώστε να θεωρείται κλινικά αξιόλογη.
- Τη στατιστική ισχύ των ελέγχων, η οποία συμβολίζεται με γ και στο χώρο των επιστημών της Υγείας έχει καθοριστεί να είναι $> 0,80$ ή $> 0,90$.
- Το επίπεδο ακρίβειας στις μετρήσεις, το οποίο εξαρτάται και από την συνείδηση των ερευνητών που διεξάγουν την έρευνα.
- Το μέγεθος του πληθυσμού αναφοράς.
- Τη μεταβλητότητα στα χαρακτηριστικά του πληθυσμού, η οποία αν είναι μεγάλη συνεπάγεται και ανάλογη αύξηση του μεγέθους του δείγματος.
- Το διαθέσιμο χρηματικό ποσό για την έρευνα.

- Τα διαστήματα εμπιστοσύνης αποτελούν έναν εναλλακτικό τρόπο εκτίμησης παραμέτρων.
- Εκτιμάμε μία παράμετρο, με ένα διάστημα που έχει άκρα τυχαίες μεταβλητές.
- Το διάστημα θα έχει την μορφή: $P[L \leq \theta \leq U] = \gamma$
- Ένα τέτοιο διάστημα ονομάζεται διάστημα εμπιστοσύνης με βαθμό εμπιστοσύνης γ . Ο αριθμός $\gamma = 1 - \alpha$ εκφράζει την ακρίβεια με την οποία θέλουμε να γίνει η εκτίμηση, ενώ ο α εκφράζει τον βαθμό ανεκτικότητας ώστε το διάστημα να μην περιέχει την πραγματική τιμή της παραμέτρου.
- Για παράδειγμα αν $\gamma = 0.95$ αναμένεται σε 100 δείγματα της μορφής $[L, U]$ τα 95 να περιλαμβάνουν την σωστή τιμή.

Παράδειγμα

- Μετρήθηκε το κάλιο του ορού σε 9 υγιή άτομα και σε 4 άτομα που έπασχαν από μία νόσο. Στα υγιή άτομα βρέθηκε μέση τιμή 4 mEq/L και σταθερή απόκλιση 0.9 mEq/L , ενώ στους ασθενείς βρέθηκε μέση τιμή 5 mEq/L και σταθερή απόκλιση 0.8 mEq/L .

Υπάρχει διαφορά των μέσων τιμών του καλίου του ορού στις δύο αυτές ομάδες;

Έλεγχοι υποθέσεων και δ.ε. για διαφορά μέσων τιμών σε ανεξάρτητους πληθυσμούς σε μικρά δείγματα και με ισότητα διασπορών ($\sigma_1 = \sigma_2 = \sigma$):

$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 > \mu_2$	$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 < \mu_2$	$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 \neq \mu_2$
$R = \{t > t_{n_1+n_2-2; a}\}$	$R = \{t < -t_{n_1+n_2-2; a}\}$	$R = \{ t > t_{n_1+n_2-2; \frac{a}{2}}\}$

$(\bar{x}_1 - \bar{x}_2 - s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} t_{n_1+n_2-2; \frac{a}{2}}, \bar{x}_1 - \bar{x}_2 + s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} t_{n_1+n_2-2; \frac{a}{2}})$, όπου $s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$.

Το κριτήριο t δίνεται από τον τύπο: $t = \frac{\bar{x}_1 - \bar{x}_2}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$

STATGRAPHICS Plus - Untitled StatFolio

File Edit Plot Describe Compare Relate Special SnapStats!! View Window Help

Hypothesis Tests

Lbl: Row:

```

Hypothesis Tests
-----
Sample means = 4,0 and 5,0
Sample standard deviations = 0,9 and 0,8
Sample sizes = 9 and 4

95,0% confidence interval for difference between means: -1,0 +/- 1,1558 [-2,1558;0,155798]

Null Hypothesis: difference between means = 0,0
Alternative: not equal
Computed t statistic = -1,9043
P-Value = 0,0833412
Do not reject the null hypothesis for alpha = 0,05.

(Equal variances assumed).

The StatAdvisor
-----
This analysis shows the results of performing a hypothesis test

```

Όπως διαπιστώνουμε δεχόμαστε την μηδενική υπόθεση $H_0 : \mu_1 = \mu_2$ έναντι της εναλλακτικής $H_1 : \mu_1 \neq \mu_2$, δηλαδή δεχόμαστε ότι δεν υπάρχει διαφορά στις τιμές του καλίου του ορού στις δύο αυτές ομάδες.

Συγκεκριμένα:

Null Hypothesis: difference between means = 0,0

Alternative: not equal

Computed t statistic = -1,9043

P-Value = 0,0833412

Do not reject the null hypothesis for alpha = 0,05

(Equal variances assumed)

Δεχόμαστε την μηδενική υπόθεση H_0 για επίπεδο σημαντικότητας $\alpha = 0.05$, διότι η τιμή του p-value είναι $0.08334 > 0.05$. Επίσης το στατιστικό λογισμικό μας υπολογίζει και την τιμή του t κριτηρίου ίση με -1.9043 .

Σημειώνεται ότι αναφερόμαστε σε κανονικούς πληθυσμούς με άγνωστες και ίσες διασπορές ($\sigma_1 = \sigma_2 = \sigma$).

Statistical Tests –Confidence Intervals

Statistical tests I

- Όπως διαπιστώνουμε επίσης το 95% διάστημα εμπιστοσύνης για την διαφορά των μέσων τιμών $\mu_1 - \mu_2$ του καλίου του ορού στις δύο αυτές ομάδες είναι:

[-2,1558;0,155798]

Statistical Tests –Confidence Intervals

Statistical tests I

Παράδειγμα:

Σε τέσσερα άτομα με αυξημένες τιμές των τριγλυκεριδίων του ορού (mg/dl) χορηγήθηκε για ένα μήνα φάρμακο που πιστεύεται ότι ελαττώνει τα επίπεδα των τριγλυκεριδίων. Οι τιμές των τριγλυκεριδίων στα τέσσερα αυτά άτομα πριν και μετά τη χορήγηση του φαρμάκου ήταν:

Άτομο	Πριν τη χορήγηση	Μετά τη χορήγηση
1 ^ο	180	120
2 ^ο	200	220
3 ^ο	240	130
4 ^ο	230	160

Βρείτε ένα 95% δ.ε. για την διαφορά των μέσων $\mu_1 - \mu_2$ στα επίπεδα των τριγλυκεριδίων πριν και μετά την χορήγηση.

Ελαττώνει τα επίπεδα των τριγλυκεριδίων το φάρμακο αυτό;

(Άσκηση 65 σελ. 16 του Βιβλίου Ασκήσεων Βιοστατιστικής Α. Τζώνου & Κ. Κατσουγιάννη)

Statistical Tests – Confidence Intervals

Statistical tests I

Έλεγχοι υποθέσεων και δ.ε. για παρατηρήσεις κατά ζεύγη:

$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 > \mu_2$	$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 < \mu_2$	$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 \neq \mu_2$
$R = \left\{ \frac{\bar{z}}{s_z} \sqrt{n} > t_{n-1; a} \right\}$	$R = \left\{ \frac{\bar{z}}{s_z} \sqrt{n} < -t_{n-1; a} \right\}$	$R = \left\{ \left \frac{\bar{z}}{s_z} \sqrt{n} \right > t_{n-1; \frac{a}{2}} \right\}$

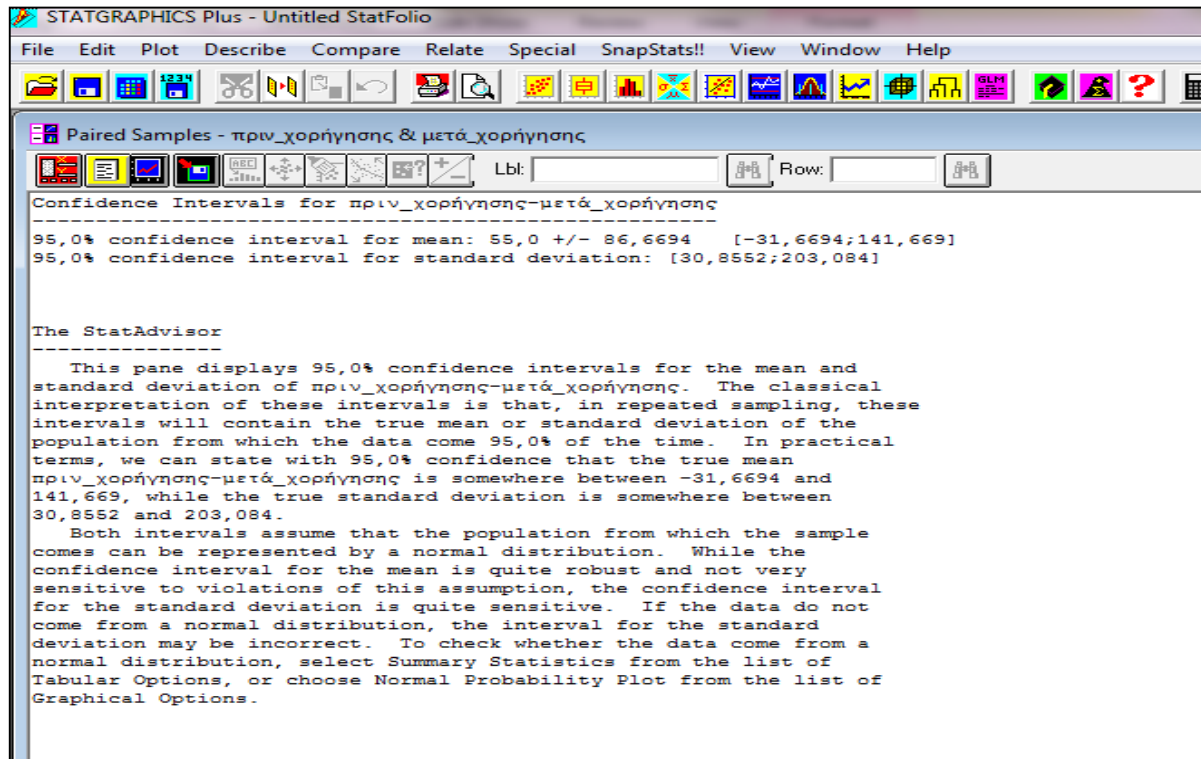
$$\left(\bar{z} - \frac{s_z}{\sqrt{n}} t_{n-1; \frac{a}{2}}, \bar{z} + \frac{s_z}{\sqrt{n}} t_{n-1; \frac{a}{2}} \right), \text{ όπου } z = x_i - y_i.$$

Statistical Tests –Confidence Intervals

Statistical tests I

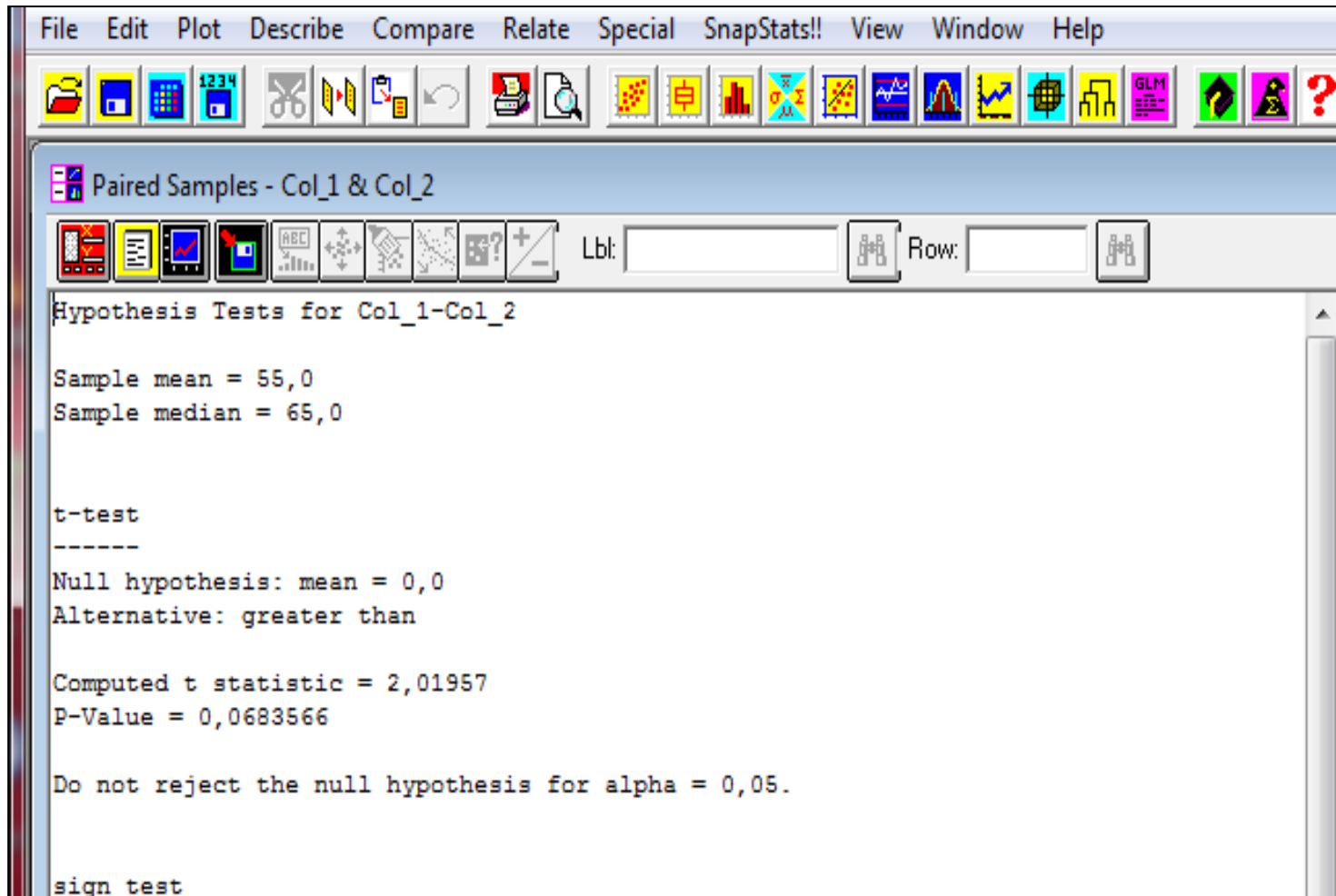
Όπως φαίνεται και στην παρακάτω εικόνα το 95% διάστημα εμπιστοσύνης για την διαφορά των μέσων στα επίπεδα τριγλυκεριδίων πριν και μετά την χορήγηση είναι:

$$55,0 \pm 86,6694 = [-31,6694;141,669]$$



Statistical Tests – Confidence Intervals

Statistical tests I



The screenshot displays a software window titled "Paired Samples - Col_1 & Col_2". The main area shows the following text:

```
Hypothesis Tests for Col_1-Col_2  
  
Sample mean = 55,0  
Sample median = 65,0  
  
t-test  
-----  
Null hypothesis: mean = 0,0  
Alternative: greater than  
  
Computed t statistic = 2,01957  
P-Value = 0,0683566  
  
Do not reject the null hypothesis for alpha = 0,05.  
  
sign test
```


Statistical Tests –Confidence Intervals

Statistical tests I

Όπως παρατηρούμε παράγονται τα εξής συμπεράσματα:

Null hypothesis: mean = 0,0

Alternative: greater than

Computed t statistic = 2,01957

P-Value = 0,0683566

Do not reject the null hypothesis for alpha = 0,05.

Δηλ. δεχόμαστε (δεν απορρίπτουμε) την μηδενική υπόθεση σε επίπεδο σημαντικότητας $\alpha=5\%$ και συνεπώς το φάρμακο δεν ελαττώνει τα επίπεδα των τριγλυκεριδίων.

Αυτό συμβαίνει διότι η τιμή του P είναι $0,068 > 0,05$ και άρα δέχομαι την $H_0: \mu_1 = \mu_2$.

Ταυτόχρονα υπολογίζεται και η τιμή του κριτηρίου t statistic ίση με 2,01957.

Τέλος Ενότητας