



ΓΕΩΠΟΝΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
AGRICULTURAL UNIVERSITY OF ATHENS

Έλεγχος υποθέσεων - Σύγκριση δύο μέσων όρων - Έλεγχος t και F

Κατσιλέρος Αναστάσιος

2017

Έλεγχος υποθέσεων

Ο έλεγχος υποθέσεων είναι μια διαδικασία κατά την οποία ελέγχουμε ισχυρισμούς που αφορούν τις παραμέτρους του πληθυσμού, με την επιλογή δειγμάτων από τον πληθυσμό.

Υπάρχουν δύο είδη στατιστικών υποθέσεων:

Η μηδενική υπόθεση (H_0), είναι η υπόθεση που δηλώνει μη ύπαρξης διαφοράς.

Η εναλλακτική υπόθεση (H_1), είναι η υπόθεση που δηλώνει ύπαρξης διαφοράς (αμφίπλευρες και μονόπλευρες).

Σφάλμα τύπου I

Η απόρριψη της μηδενικής υπόθεσης, όταν στην πραγματικότητα ισχύει λέγεται **Σφάλμα τύπου I**.

Η πιθανότητα διάπραξης Σφάλματος τύπου I καθορίζεται από την εκλογή του επίπεδου σημαντικότητας α και συνήθως είναι 0,05 ή 0,01.

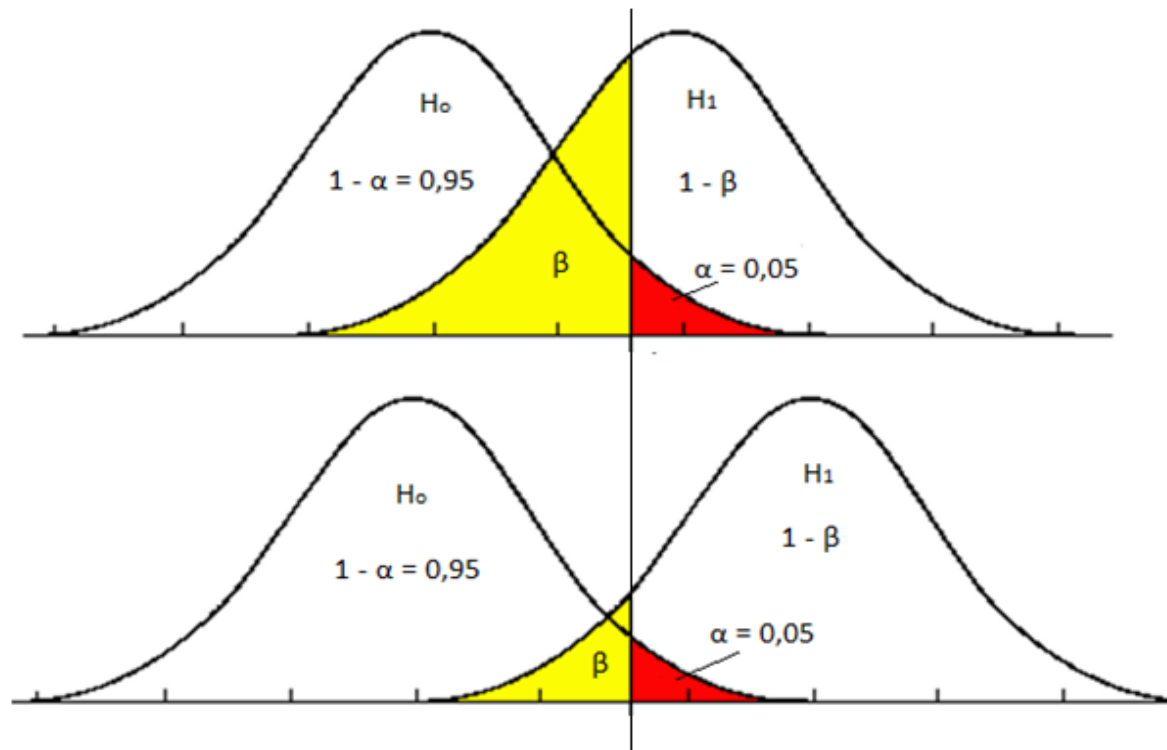
Η πιθανότητα να αποδεχτούμε την H_0 όταν είναι ορθή είναι $1 - \alpha$. Το $1 - \alpha$ είναι ο συντελεστής εμπιστοσύνης του ελέγχου.

Σφάλμα τύπου II

Η μη απόρριψη μιας λανθασμένης μηδενικής υπόθεσης όταν πράγματι ισχύει λέγεται **Σφάλμα τύπου II**

Η πιθανότητα διάπραξης Σφάλματος τύπου II είναι β το οποίο δεν είναι γνωστό από πριν και δεν καθορίζεται. Για ένα ορισμένο αριθμό παρατηρήσεων η τιμή του β είναι αντίστροφα ανάλογη με αυτή του α . Η πιθανότητα να απορρίψουμε την H_0 και να δεχτούμε την H_1 είναι $1 - \beta$ και ονομάζεται **ισχύς του ελέγχου**. Ο μόνος τρόπος να μειώσουμε ταυτόχρονα τις πιθανότητες των δύο σφαλμάτων είναι αυξάνοντας τον αριθμό των παρατηρήσεων.

Πειραματικοί Σχεδιασμοί



α : επίπεδο σημαντικότητας, $1-\alpha$: συντελεστής εμπιστοσύνης, $1-\beta$: ισχύς ελέγχου

	Ισχύει H_0	Δεν Ισχύει H_0
Μη απόρριψη της H_0	☺ Σωστή απόφαση	Σφάλμα τύπου II
Απόρριψη της H_0	Σφάλμα τύπου I	☺ Σωστή απόφαση

Στάδια ελέγχου υποθέσεων

- ✓ Διαμόρφωση της μηδενικής H_0 και εναλλακτικής υπόθεσης H_1 .
- ✓ Επιλογή **επίπεδου σημαντικότητας** α το οποίο ορίζεται η πιθανότητα που χρησιμοποιούμε ως κριτήριο για την απόρριψη της H_0 .
- ✓ Επιλογή της στατιστικής δοκιμασίας. Ορισμός των **κρίσιμων τιμών** της στατιστικής δοκιμασίας που αντιστοιχούν στην πιο πάνω πιθανότητα. Υπολογισμός της τιμής της στατιστικής δοκιμής (t ή F) και σύγκριση με τις κρίσιμες τιμές ή χρήση της **p-value**.

Η **p-value** είναι η ακριβής πιθανότητα P για την εμφάνιση μιας τιμής ίσης ή πιο ακραίας από την τιμή της στατιστικής που υπολογίστηκε, υπό την προϋπόθεση ότι ισχύει η H_0

Η **p-value** είναι η ελάχιστη τιμή του επιπέδου σημαντικότητας για την οποία απορρίπτεται η H_0 και εκφράζει πόσο ισχυρές είναι οι αποδείξεις που προκύπτουν από το δείγμα εναντίον της H_0 .

Έλεγχος υπόθεσης ότι ο μέσος όρος ενός πληθυσμού έχει μια συγκεκριμένη τιμή

Ο έλεγχος εξαρτάται από το αν η παραλλακτικότητα του πληθυσμού είναι γνωστή ή άγνωστη

Στην περίπτωση που είναι γνωστή η παραλλακτικότητα χρησιμοποιείται η σχέση:

$$z = \frac{\bar{Y} - \mu_0}{\sigma / \sqrt{n}}$$

με όρια εμπιστοσύνης: $\bar{Y} \pm z_{\alpha/2, (n-1)} \frac{\sigma}{\sqrt{n}}$

ενώ όταν είναι άγνωστη χρησιμοποιείται η σχέση:

$$t = \frac{\bar{Y} - \mu_0}{s / \sqrt{n}}$$

με όρια εμπιστοσύνης: $\bar{Y} \pm t_{\alpha/2, (n-1)} s_{\bar{Y}}$

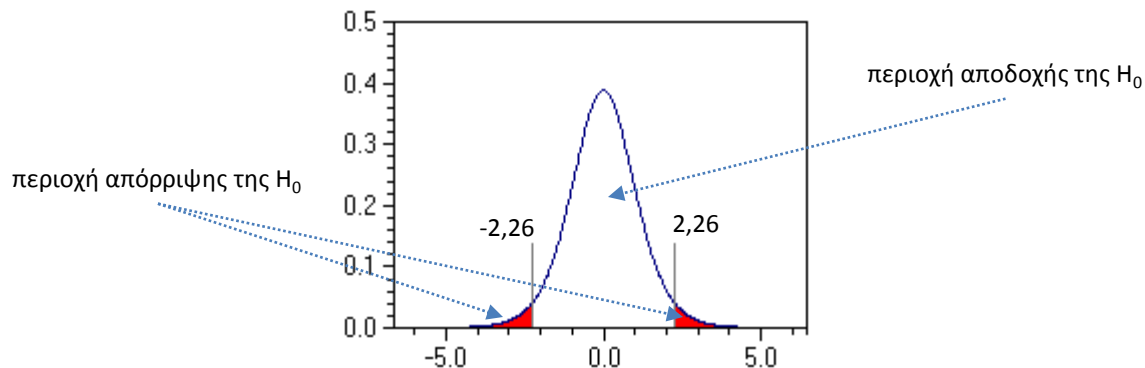
Παράδειγμα 1.

Τα δεδομένα (101, 96, 93, 108, 102, 92, 98, 97, 95, 103) αναφέρονται στο ύψος σε εκατοστά 10 αγοριών ηλικίας 4 χρονών. Ελέγξτε με πιθανότητα 0,95 ότι ο μέσος όρος του πληθυσμού είναι 100.

$$H_0: \mu = 100, \quad H_1: \mu \neq 100,$$

$$t = \frac{\bar{Y} - \mu_0}{s / \sqrt{n}} = \frac{98,5 - 100}{4,97 / \sqrt{10}} = \frac{-1,5}{1,57} = -0,954$$

Επειδή ο έλεγχος είναι αμφίπλευρος η περιοχή απόρριψης της H_0 περιλαμβάνει τιμές μικρότερες από -2,26 και μεγαλύτερες από 2,26 ($t_{0,025,9}$) και η περιοχή αποδοχής της H_0 τιμές μεταξύ -2,26 και 2,26.



Διάστημα εμπιστοσύνης

Ο τύπος του διαστήματος εμπιστοσύνης για τον έλεγχο της υπόθεσης H_0 είναι:

$$\bar{Y} \pm t_{\alpha/2, (n-1)} s_{\bar{Y}} = 98,5 \pm 2,26 * 1,57 = 98,5 \pm 3,548$$

Το κατώτερο όριο είναι: 94,94
και το ανώτερο όριο είναι: 102,05

```
> x=c(101, 96, 93, 108, 102, 92, 98, 97, 95, 103)
> t.test(x, mu=100, alternative = c("two.sided"), conf.level = 0.95)
```

One Sample t-test

```
data: x
t = -0.954, df = 9, p-value = 0.365
alternative hypothesis: true mean is not equal to 100
95 percent confidence interval:
 94.94314 102.05686
sample estimates:
mean of x
98.5
```

Έλεγχος υποθέσεων για τη διαφορά των μέσων όρων δύο πληθυσμών

Έστω ότι μας ενδιαφέρει η μελέτη δύο πληθυσμών, οι οποίοι γνωρίζουμε ότι ακολουθούν την κανονική κατανομή αλλά άγνωστες μέσες τιμές.

Παίρνουμε ένα δείγμα από κάθε πληθυσμό και υπολογίζουμε τη μέση τιμή τους.

Η διαφορά των μέσων τιμών των δύο δειγμάτων ακολουθεί την **κανονική κατανομή** με μέσο όρο και διακύμανση:

$$\mu_{\bar{Y}_1 - \bar{Y}_2} = \mu_1 - \mu_2 \quad \text{και} \quad \sigma_{\bar{Y}_1 - \bar{Y}_2}^2 = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

Αν οι παραλλακτικότητες των δύο πληθυσμών σ^2_1 και σ^2_2 είναι **γνωστές**, η μεταβλητή z ακολουθεί την τυπική κανονική κατανομή.

$$\frac{(\bar{Y}_1 - \bar{Y}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = z \sim N(0,1)$$

Για να ελέγξουμε την υπόθεση ότι η διαφορά των δύο μέσων τιμών των πληθυσμών είναι μηδέν ($H_0: \mu_1 - \mu_2 = 0$), δηλαδή ότι οι μέσες τιμές είναι ίσες, χρησιμοποιούμε την στατιστική συνάρτηση ελέγχου:

$$z = \frac{(\bar{Y}_1 - \bar{Y}_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Όταν οι παραλλακτικότητες των δύο πληθυσμών είναι άγνωστες, το οποίο είναι το πιο σύνηθες, εκτιμώνται από τα δεδομένα των δειγμάτων.

Ο τρόπος σύγκρισης των μέσων τιμών εξαρτάται από το αν οι τιμές είναι κατά ζεύγη ή ανεξάρτητες.

Αν οι τιμές είναι ανεξάρτητες ο τρόπος σύγκρισης εξαρτάται από το αν οι παραλλακτικότητες των πληθυσμών από τους οποίους προήλθαν είναι ίσες ή όχι.

- α) Τα δείγματα είναι ανεξάρτητα και οι παραλλακτικότητες των πληθυσμών από τους οποίους προέρχονται είναι ίσες.
- β) Τα δείγματα είναι ανεξάρτητα και οι παραλλακτικότητες των πληθυσμών από τους οποίους προέρχονται είναι διαφορετικές.
- γ) εάν οι παρατηρήσεις είναι κατά ζεύγη

Έλεγχος ομοιογένειας των δύο διακυμάνσεων

Ο έλεγχος της υπόθεσης ότι παραλλακτικότητες των πληθυσμών είναι ίσες ($H_0: \sigma^2_1 = \sigma^2_2$) γίνεται με τη δοκιμασία Hartley's $F_{\max}$:

$$F_{\max} = \frac{\text{μεγαλύτερη } s_1^2}{\text{μικρότερη } s_2^2}$$

και σύγκριση με κρίσιμη τιμή από πίνακα του F σε επίπεδο σημαντικότητας $\alpha/2$ (αμφίπλευρη δοκιμασία) και για βαθμούς ελευθερίας $n_1 - 1$ για τον αριθμητή και $n_2 - 1$ για τον παραιννομαστή.

Τα δείγματα είναι ανεξάρτητα και οι παραλλακτικότητες ίσες.

Έστω δύο ανεξάρτητα τυχαία δείγματα από δύο κανονικούς πληθυσμούς με άγνωστες αλλά ίσες διακυμάνσεις, ισχύει:

$$t = \frac{(\bar{Y}_1 - \bar{Y}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}} \sim t_{n_1+n_2-2}$$

Όπου $\bar{Y}_1$ και $\bar{Y}_2$ οι μέσοι όροι των δειγμάτων, μ_1 και μ_2 οι μέσοι όροι των πληθυσμών, n_1 και n_2 τα μεγέθη των δειγμάτων και s_p^2 η εκτιμήτρια της κοινής διακύμανσης σ^2 των δύο πληθυσμών με βάση και τα δύο δείγματα.

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 - 1) + (n_2 - 1)} \quad \text{όταν } n_1 = n_2 : s_p^2 = \frac{s_1^2 + s_2^2}{2}$$

Εάν μ_1 και μ_2 είναι οι μέσοι όροι των δύο πληθυσμών, να ελεγχθεί η υπόθεση:

$$H_0: \mu_1 = \mu_2 \quad \text{και} \quad H_1: \mu_1 \neq \mu_2$$

Εφαρμόζεται η δοκιμασία του t :

$$t = \frac{\bar{Y}_1 - \bar{Y}_2}{s_{\bar{Y}_1 - \bar{Y}_2}}$$

με βαθμούς ελευθερίας $n_1 + n_2 - 2$

$$s_{\bar{Y}_1 - \bar{Y}_2} = \sqrt{s_p^2 \frac{n_1 + n_2}{n_1 n_2}}$$

το τυπικό σφάλμα της διαφοράς των δύο μέσων όρων.

Παράδειγμα 2.

Τα δεδομένα αναφέρονται στο ύψος 10 φυτών από δύο ποικιλίες κριθαριού. Ελέγξτε αν οι δυο ποικιλίες διαφέρουν σημαντικά.

Κρόνος	Νίκη
95	104
99	74
103	84
90	81
121	62
98	67
102	64
94	93
87	73
95	87
$n_1 = 10$	$n_2 = 10$
$\bar{Y}_1 = 98,4$	$\bar{Y}_2 = 78,9$
$s^2_1 = 87,6$	$s^2_2 = 181,4$

Έλεγχος ομοιογένειας των δύο διακυμάνσεων

$$H_0: \sigma^2_1 = \sigma^2_2, \quad H_1: \sigma^2_1 \neq \sigma^2_2,$$

$$F_{\max} = \frac{\text{μεγαλύτερη } s^2_2}{\text{μικρότερη } s^2_1} = \frac{181,4}{86,7} = 2,09$$

Η κρίσιμη τιμή F για $\alpha/2 = 0,025$ και βαθμούς ελευθερίας $n_1 - 1 = 9$ για τον αριθμητή και $n_2 - 1 = 9$ για τον παρανομαστή ισούται με 4,02.

Επειδή η τιμή που υπολογίστηκε (2,09) είναι μικρότερη από την κρίσιμη τιμή F , δεν απορρίπτεται η υπόθεση H_0 ισότητας των δύο διακυμάνσεων, σε $\alpha = 0,05$.

Επειδή δεν απορρίφτηκε η H_0 , η συμψηφισμένη διακύμανση υπολογίζεται ως εξής:

$$s_p^2 = \frac{s_1^2 + s_2^2}{2} = \frac{(86,7 + 181,4)}{2} = 134,05$$

Το τυπικό σφάλμα της διαφοράς:

$$s_{\bar{Y}_1 - \bar{Y}_2} = \sqrt{\frac{2s_p^2}{n}} = \sqrt{\frac{2(134,05)}{10}} = \sqrt{26,81} = 5,18$$

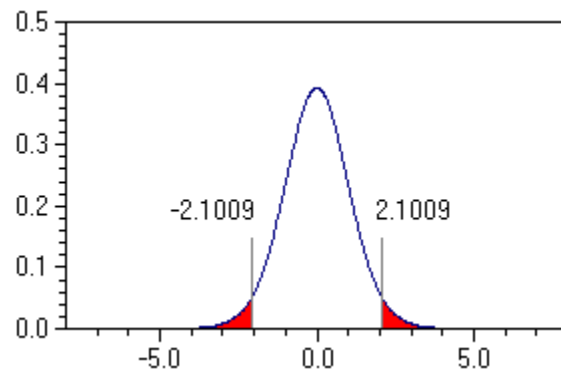
Εφαρμόζεται η δοκιμασία του t :

$$t = \frac{\bar{Y}_1 - \bar{Y}_2}{s_{\bar{Y}_1 - \bar{Y}_2}} = \frac{98,4 - 78,9}{5,18} = \frac{19,5}{5,18} = 3,76$$

Η κρίσιμη τιμή t για $\alpha = 0,05$ και βαθμούς ελευθερίας $n_1 + n_2 - 2 = 18$ ισούται με 2,1.

Επειδή η τιμή που υπολογίστηκε (3,76) είναι μεγαλύτερη από την κρίσιμη τιμή t , απορρίπτεται η μηδενική υπόθεση H_0 , σε $\alpha = 0,05$.

Επομένως οι ποικιλίες διαφέρουν στατιστικά σημαντικά ως προς το ύψος.



Διάστημα εμπιστοσύνης

Ο τύπος του διαστήματος εμπιστοσύνης για τον έλεγχο της υπόθεσης $H_0 \mu_1 = \mu_2$ είναι:

$$(\bar{Y}_1 - \bar{Y}_2) \pm t_{\alpha/2, (n_1+n_2-2)} S_{\bar{Y}_1 - \bar{Y}_2}$$

$$(98,4 - 78,9) \pm 2,1 * 5,18 = 19,5 \pm 10,9$$

Το κατώτερο όριο είναι: $l_1 = 8,6$
και το ανώτερο όριο είναι: $l_2 = 30,4$

Επειδή το διάστημα εμπιστοσύνης δεν περιέχει την τιμή 0, συμπεραίνουμε ότι οι μέσοι όροι διαφέρουν.

Πειραματικοί Σχεδιασμοί

```
> x=c(95, 99, 103, 90, 121, 98, 102, 94, 87, 95)
```

```
> y=c(104, 74, 84, 81, 62, 67, 64, 93, 73, 87)
```

```
> var(x)
```

```
[1] 87.6
```

```
> var(y)
```

```
[1] 181.4333
```

```
> var.test(y, x, ratio = 1, alternative = c("two.sided"), conf.level = 0.95)
```

F test to compare two variances

data: y and x

F = 2.0712, num df = 9, denom df = 9, p-value = 0.2931

alternative hypothesis: true ratio of variances is not equal to 1

95 percent confidence interval:

0.514446 8.338465

sample estimates:

ratio of variances

2.071157

```
> t.test(x, y, alternative = c("two.sided"), mu = 0, paired = FALSE, var.equal = TRUE, conf.level=0.95)
```

Two Sample t-test

data: x and y

t = 3.7595, df = 18, p-value = 0.001435

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

8.602849 30.397151

sample estimates:

mean of x mean of y

98.4 78.9

Τα δείγματα είναι ανεξάρτητα και οι παραλλακτικότητες άνισες.

Όταν οι παραλλακτικότητες των πληθυσμών είναι άνισες δεν μπορούμε να υπολογίσουμε ένα κοινό s^2 από τα δύο δείγματα. Η σημαντικότητα της διαφοράς των μέσων όρων των πληθυσμών ελέγχεται με τον τύπο:

$$t' = \frac{\bar{Y}_1 - \bar{Y}_2}{S_{\bar{Y}_1 - \bar{Y}_2}}$$

Όπου:

$$S_{\bar{Y}_1 - \bar{Y}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \text{ και } BE' = \frac{(s_1^2 / n_1 + s_2^2 / n)^2}{(s_1^2 / n_1)^2 / (n_1 - 1) + (s_2^2 / n_2)^2 / (n_2 - 1)}$$

Παράδειγμα 3.

Τα δεδομένα αναφέρονται σε μελέτη προσθήκης φυτορρυθμιστικής ουσίας (IAA) σε θρεπτικό υπόστρωμα Murasighe & Skoog και μελετήθηκε η ανάπτυξη των εκφύτων.

MS	MS + IAA
11,9	21,4
15,2	35,4
14,1	22,4
13,9	25,7
10,1	31,4
13,2	26,7
12,5	34,2
-	24,5
$n_1 = 7$	$n_2 = 8$
$\bar{Y}_1 = 12,9$	$\bar{Y}_2 = 27,7$
$s^2_1 = 2,8$	$s^2_2 = 28,3$

Έλεγχος ομοιογένειας των δύο διακυμάνσεων

$$H_0: \sigma^2_1 = \sigma^2_2, \quad H_1: \sigma^2_1 \neq \sigma^2_2,$$

$$F_{\max} = \frac{\text{μεγαλύτερη } s^2_2}{\text{μικρότερη } s^2_1} = \frac{27,7}{12,9} = 10,13$$

Η κρίσιμη τιμή F για $\alpha/2 = 0,025$ και βαθμούς ελευθερίας $n_2 - 1 = 7$ για τον αριθμητή και για τον παρανομαστή $n_1 - 1 = 6$ ισούται με 5,69.

Επειδή η τιμή που υπολογίστηκε (10,13) είναι μεγαλύτερη από την κρίσιμη τιμή F , απορρίπτεται η μηδενική υπόθεση H_0 ισότητας των δύο διακυμάνσεων, σε $\alpha = 0,05$.

Επειδή απορρίφτηκε η H_0 , το τυπικό σφάλμα της διαφοράς υπολογίζεται ως εξής:

$$s_{\bar{Y}_1 - \bar{Y}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{2,8^2}{7} + \frac{28,3^2}{8}} = 1,98$$

Ο έλεγχος της σημαντικότητας της διαφοράς :

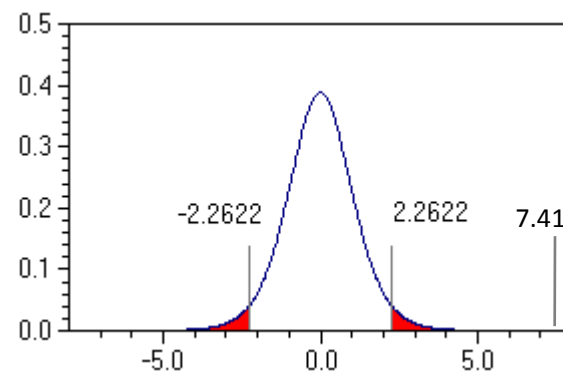
$$t' = \frac{\bar{Y}_1 - \bar{Y}_2}{s_{\bar{Y}_1 - \bar{Y}_2}} = \frac{12,9 - 27,7}{1,98} = -7,41$$

$$BE' = \frac{(2,8^2 / 7 + 28,3^2 / 8)^2}{(2,8^2 / 7)^2 / (7 - 1) + (28,3^2 / 8)^2 / (8 - 1)} = 8,54 \approx 9$$

Η κρίσιμη τιμή t για $\alpha = 0,05$ και βαθμούς ελευθερίας 9 ισούται με 2,26.

Επειδή η τιμή που υπολογίστηκε (7,41) είναι μεγαλύτερη από την κρίσιμη τιμή t , απορρίπτεται η μηδενική υπόθεση H_0 , σε $\alpha = 0,05$.

Επομένως οι ποικιλίες διαφέρουν στατιστικά σημαντικά ως προς το ύψος. Η πιθανότητα αυτό το συμπέρασμα να είναι λάθος, είναι το πολύ 0,05.



Διάστημα εμπιστοσύνης

Ο τύπος του διαστήματος εμπιστοσύνης για τον έλεγχο της υπόθεσης $H_0: \mu_1 = \mu_2$ είναι:

$$(\bar{Y}_1 - \bar{Y}_2) \pm t_{\alpha/2, BE} S_{\bar{Y}_1 - \bar{Y}_2}$$

$$(12,9 - 27,7) \pm 2,26 * 1,98 = -14,7 \pm 4,5$$

Το κατώτερο όριο είναι: $l_1 = -19,2$
και το ανώτερο όριο είναι: $l_2 = -10,2$

Επειδή το διάστημα εμπιστοσύνης δεν περιέχει την τιμή 0, οι μέσοι όροι διαφέρουν.

Πειραματικοί Σχεδιασμοί

```
> x=c(11.9, 15.2, 14.1, 13.9, 10.1, 13.2, 12.5)
> y=c(21.4, 35.4, 22.4, 25.7, 31.4, 26.7, 34.2, 24.5)
> var(x)
[1] 2.794762
```

```
> var(y)
[1] 28.32125
```

```
> var.test(y, x, ratio = 1, alternative = c("two.sided"), conf.level = 0.95)
```

F test to compare two variances

data: y and x

F = 10.134, num df = 7, denom df = 6, p-value = 0.01179

alternative hypothesis: true ratio of variances is not equal to 1

95 percent confidence interval:

1.779254 51.870270

sample estimates:

ratio of variances

10.13369

```
> t.test(x, y, alternative=c("two.sided"), mu = 0, paired=FALSE, var.equal=FALSE, conf.level=0.95)
```

Welch Two Sample t-test

data: x and y

t = -7.4198, df = 8.5412, p-value = 5.31e-05

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

-19.25374 -10.19983

sample estimates:

mean of x mean of y

12.98571 27.71250

Σύγκριση τιμών κατά ζεύγη.

Όταν οι τιμές των δύο δειγμάτων είναι κατά ζεύγη υπολογίζουμε την παραλλακτικότητα ως εξής:

$$s_D^2 = \frac{\sum D_i^2 - (\sum D_i)^2 / n}{n - 1}$$

όπου D_i οι διαφορές μεταξύ των ζευγών

Η δοκιμασία του t :
$$t = \frac{\bar{D}}{s_{\bar{D}}} \quad \text{για } BE = n - 1$$

όπου
$$s_{\bar{D}} = \sqrt{\frac{s_D^2}{n}}$$

Παράδειγμα 4.

Τα δεδομένα αναφέρονται στην περιεκτικότητα σε mg βιταμίνης C ενός τροφίμου πριν και μετά την τοποθέτηση του σε φούρνο μικροκυμάτων.

Ζεύγη	πριν	μετά	D
1	6,28	6,85	-0,57
2	7,05	6,84	0,21
3	8,24	7,07	1,17
4	5,2	5,05	0,15
5	6,16	6,18	-0,02
6	6,67	6,71	-0,04
7	7,01	6,49	0,52
8	5,46	5,24	0,22
Μέσος Όρος	6,51	6,30	0,205
Σύνολο	52,07	50,43	1,64

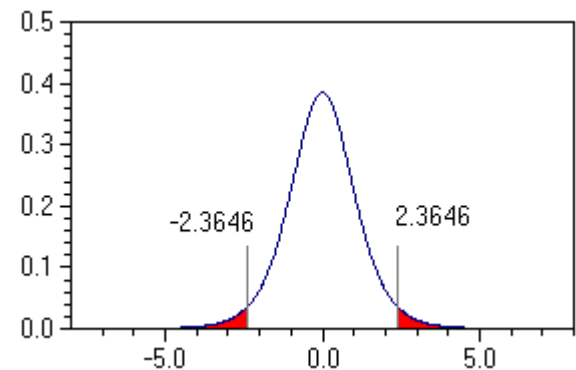
$$s_D^2 = \frac{\sum D_i^2 - (\sum D_i)^2 / n}{n-1} = \frac{2,08 - (1,64)^2 / 8}{7} = 0,25$$

$$s_{\bar{D}} = \sqrt{\frac{s_D^2}{n}} = \sqrt{\frac{0,25}{8}} = 0,176$$

Ο έλεγχος της σημαντικότητας της διαφοράς :

$$t = \frac{0,205}{0,176} = 1,16$$

Η κρίσιμη τιμή t για $\alpha = 0,05$ και βαθμούς ελευθερίας 7 ισούται με 2,36. Επειδή η τιμή που υπολογίστηκε (1,16) είναι μικρότερη από την κρίσιμη τιμή t , δεν απορρίπτεται η μηδενική υπόθεση H_0 , σε $\alpha = 0,05$.



Διάστημα εμπιστοσύνης

Ο τύπος του διαστήματος εμπιστοσύνης για τον έλεγχο της υπόθεσης $H_0: \mu_1 = \mu_2$ είναι:

$$\bar{D} \pm t_{\alpha/2, (n-1)} S_{\bar{D}}$$

$$0,205 \pm 2,3 * 0,176 = 0,205 \pm 0,4$$

Το κατώτερο όριο είναι: $l_1 = - 0,21$
και το ανώτερο όριο είναι: $l_2 = 0,62$

Επειδή το διάστημα εμπιστοσύνης περιέχει την τιμή 0, οι μέσοι όροι δεν διαφέρουν.

Πειραματικοί Σχεδιασμοί

```
> x=c(6.28, 7.05, 8.24, 5.2, 6.16, 6.67, 7.01, 5.46)
> y=c(6.85, 6.84, 7.07, 5.05, 6.18, 6.71, 6.49, 5.24)
> t.test(x, y, alternative=c("two.sided"), mu=0, paired=TRUE, var.equal=FALSE, conf.level=0.95)
```

Paired t-test

data: x and y

t = 1.1613, df = 7, p-value = 0.2836

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

-0.2124129 0.6224129

sample estimates:

mean of the differences

0.205

Βαθμός σφάλματος κατά σύγκριση ($\alpha_{\kappa\sigma}$)

Η πιθανότητα να υποθέσουμε σε σφάλμα τύπου I (πιθανότητα α) σε μία σύγκριση.

Βαθμός σφάλματος κατά πείραμα ($\alpha_{\kappa\pi}$)

Η πιθανότητα διάπραξης τουλάχιστον ενός σφάλματος τύπου I σε μία ομάδα συγκρίσεων και είναι:

$$\alpha_{\kappa\pi} = 1 - (1 - \alpha)^c$$

όπου c = ο αριθμός των συγκρίσεων ανά ζεύγη

για μικρό α και c

$$\alpha_{\kappa\pi} \approx c(\alpha)$$

Η πιθανότητα σφάλματος αυξάνεται με την αύξηση του αριθμού των μέσων που συγκρίνονται.

$$\alpha_{\kappa\eta} = 1 - (1 - \alpha)^c$$

όπου c = ο αριθμός των συγκρίσεων ανά ζεύγη

$$c = \frac{t(t-1)}{2}$$

Αν $t = 5$ μέσοι, όλες οι δυνατές συγκρίσεις είναι $c = 10$

$$\alpha_{\kappa\eta} = 1 - (1 - \alpha)^c = 1 - (1 - 0,05)^{10} = 0,40$$

Η πιθανότητα διάπραξης σφάλματος τύπου I είναι 40%

Σύγκριση δύο ή περισσότερων μέσων όρων με τη ανάλυση διακύμανσης (ANOVA)

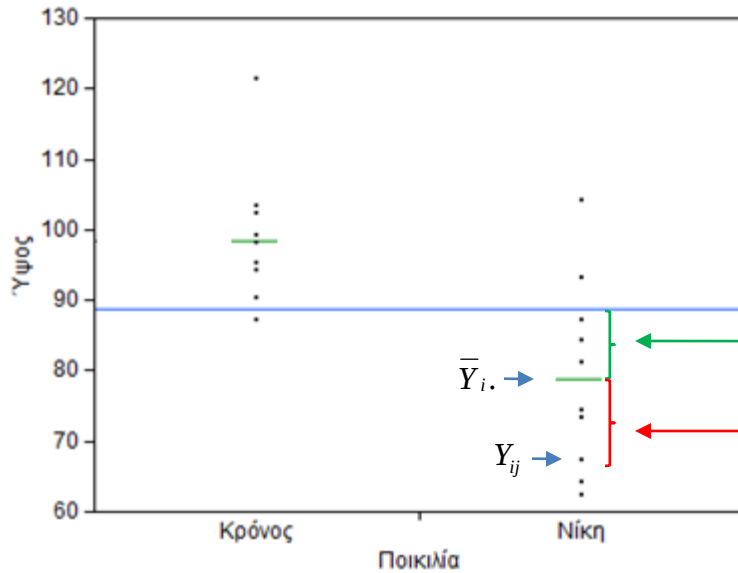
Η ανάλυση διακύμανσης επιτρέπει:

- α) τον υπολογισμό ενός κοινού σφάλματος για την σύγκριση των μέσων όρων
- β) παρέχει τη δυνατότητα ενός προκαταρκτικού ελέγχου της ύπαρξης διαφορών μεταξύ των μέσων όρων με τη χρήση του κριτηρίου F.

Προϋποθέσεις για την εφαρμογή της ανάλυσης διακύμανσης είναι οι πληθυσμοί από τους οποίους προέρχονται τα δείγματα να είναι κανονικοί με κοινή διακύμανση και ανεξάρτητες παρατηρήσεις.

Πειραματικοί Σχεδιασμοί

Κατάτμηση Αθροίσματος Τετραγώνων



Το γραμμικό πρότυπο:

$$Y_{ij} = \mu + \tau_i + \varepsilon_{ij}$$

$\bar{Y}_{..}$	εκτιμά	μ
$(\bar{Y}_{i.} - \bar{Y}_{..})$	»	τ_i
$(Y_{ij} - \bar{Y}_{i.})$	»	ε_{ij}

οπότε:

$$Y_{ij} = \bar{Y}_{..} + (\bar{Y}_{i.} - \bar{Y}_{..}) + (Y_{ij} - \bar{Y}_{i.})$$

$$Y_{ij} - \bar{Y}_{..} = (\bar{Y}_{i.} - \bar{Y}_{..}) + (Y_{ij} - \bar{Y}_{i.})$$

Αν υψώσουμε στο τετράγωνο την πιο πάνω ταυτότητα για κάθε παρατήρηση και αθροίσουμε όλες τις παρατηρήσεις:

$$\sum_{i=1}^a \sum_{j=1}^n (Y_{ij} - \bar{Y}_{..})^2 = \sum_{i=1}^a \sum_{j=1}^n (\bar{Y}_{i.} - \bar{Y}_{..})^2 + \sum_{i=1}^a \sum_{j=1}^n (Y_{ij} - \bar{Y}_{i.})^2 + 2 \sum_{i=1}^a \sum_{j=1}^n (Y_{ij} - \bar{Y}_{i.})(\bar{Y}_{i.} - \bar{Y}_{..})$$

$$\sum_{i=1}^a \sum_{j=1}^n (Y_{ij} - \bar{Y}_{..})^2 = n \sum_{i=1}^a (\bar{Y}_{i.} - \bar{Y}_{..})^2 + \sum_{i=1}^a \sum_{j=1}^n (Y_{ij} - \bar{Y}_{i.})^2$$

Πίνακας Ανάλυσης της Παραλλακτικότητας

Πηγή παρ/τητας	ΒΕ	ΑΤ	ΜΤ	F	ΘΣΜΤ
Επεμβάσεις	$a - 1$	$AT\varepsilon = n \sum_{i=1}^a (\bar{Y}_{i.} - \bar{Y}_{..})^2$	$MT\varepsilon = \frac{AT\alpha}{(\alpha - 1)}$	$\frac{MT\varepsilon}{MT\nu}$	$\sigma_e^2 + n\sigma_a^2$
Υπόλοιπο	$a(n - 1)$	$AT\nu = \sum_{i=1}^a \sum_{j=1}^n (Y_{ij} - \bar{Y}_{i.})^2$	$MT\nu = \frac{AT\nu}{\alpha(n - 1)}$		σ_e^2
Σύνολο	$an - 1$	$AT\sigma = \sum_{i=1}^a \sum_{j=1}^n (Y_{ij} - \bar{Y}_{..})^2$			

Η στήλη ΘΣΜΤ περιέχει τις μαθηματικά αναμενόμενες τιμές του μέσου τετραγώνου και εκτιμούνται από την στατιστική της στήλης ΜΤ. Το σ_e^2 εκτιμάται από την s_e^2 με το ΜΤυ. Όταν οι μέσοι όροι των επεμβάσεων είναι ίσοι, τότε το ΜΤε παρέχει επίσης ένα αμερόληπτο εκτιμητή της διακύμανσης σ_e^2 .

Παράδειγμα 2.

Τα δεδομένα αναφέρονται στο ύψος 10 φυτών από δύο ποικιλίες κριθαριού. Ελέγξτε αν οι δυο ποικιλίες διαφέρουν σημαντικά.

Κρόνος	Νίκη
95	104
99	74
103	84
90	81
121	62
98	67
102	64
94	93
87	73
95	87
$n_1 = 10$	$n_2 = 10$
$\bar{Y}_1 = 98,4$	$\bar{Y}_2 = 78,9$
$Y_{1.} = 984$	$Y_{2.} = 789$
$s^2_1 = 87,6$	$s^2_2 = 181,4$

Ο έλεγχος των υποθέσεων:

Μηδενική υπόθεση

$$H_0: \mu_1 = \mu_2$$

Εναλλακτική υπόθεση

$$H_1: \mu_1 \neq \mu_2$$

Δοκιμασία ή κριτήριο του F

Είναι μία γενική δοκιμασία, μονόπλευρη και ελέγχει την μηδενική υπόθεση H_0 .

$$F = \frac{\sigma_e^2 + n\sigma_a^2}{\sigma_e^2} = \frac{\text{ΜΤεπεμβάσεων}}{\text{ΜΤυπολοίπου}}$$

Πειραματικοί Σχεδιασμοί

Tr	Y	Y_{ij}	$e_{ij} = Y_{ij} - \bar{Y}_{i.}$	$\tau_i = \bar{Y}_{i.} - \bar{Y}_{..}$	$Y_{ij} - \bar{Y}_{..}$	$(Y_{ij} - \bar{Y}_{i.})^2$	$(\bar{Y}_{i.} - \bar{Y}_{..})^2$	$(Y_{ij} - \bar{Y}_{..})^2$
1	$Y_{1.1}$	95	-3,4	9,75	6,35	11,56	95,06	40,3
1	$Y_{1.2}$	99	0,6	9,75	10,35	0,36	95,06	107,1
1	$Y_{1.3}$	103	4,6	9,75	14,35	21,16	95,06	205,9
1	$Y_{1.4}$	90	-8,4	9,75	1,35	70,56	95,06	1,8
1	$Y_{1.5}$	121	22,6	9,75	32,35	510,8	95,06	1046,5
1	$Y_{1.6}$	98	-0,4	9,75	9,35	0,16	95,06	87,4
1	$Y_{1.7}$	102	3,6	9,75	13,35	12,96	95,06	178,2
1	$Y_{1.8}$	94	-4,4	9,75	5,35	19,36	95,06	28,6
1	$Y_{1.9}$	87	-11,4	9,75	-1,65	130	95,06	2,7
1	$Y_{1.10}$	95	-3,4	9,75	6,35	11,56	95,06	40,3
2	$Y_{2.1}$	104	25,1	-9,75	15,35	630	95,06	235,6
2	$Y_{2.2}$	74	-4,9	-9,75	-14,65	24,01	95,06	214,6
2	$Y_{2.3}$	84	5,1	-9,75	-4,65	26,01	95,06	21,6
2	$Y_{2.4}$	81	2,1	-9,75	-7,65	4,41	95,06	58,5
2	$Y_{2.5}$	62	-16,9	-9,75	-26,65	285,6	95,06	710,2
2	$Y_{2.6}$	67	-11,9	-9,75	-21,65	141,6	95,06	468,7
2	$Y_{2.7}$	64	-14,9	-9,75	-24,65	222	95,06	607,6
2	$Y_{2.8}$	93	14,1	-9,75	4,35	198,8	95,06	18,9
2	$Y_{2.9}$	73	-5,9	-9,75	-15,65	34,81	95,06	244,9
2	$Y_{2.10}$	87	8,1	-9,75	-1,65	65,61	95,06	2,7
Σύνολο		1773	0	0	0	2421,3	1901,25	4322,55

$$\bar{Y}_{1.} = 98,4$$

$$\bar{Y}_{2.} = 78,9$$

$$\bar{Y}_{..} = 88,65$$

Το συνολικό άθροισμα τετραγώνων υπολογίζεται ως εξής:

$$AT_{\text{συνόλου}} = \sum_{i=1}^a \sum_{j=1}^n (Y_{ij} - \bar{Y}_{..})^2 = \sum_{i=1}^a \sum_{j=1}^n Y_{ij}^2 - [\sum_{i=1}^a \sum_{j=1}^n Y_{ij}]^2 / an$$

Ο διορθωτικός όρος υπολογίζεται: $\Delta O = [\sum_{i=1}^a \sum_{j=1}^n Y_{ij}]^2 / an = \frac{Y_{..}^2}{an} = \frac{1773^2}{2*10} = 157176$

$$AT_{\text{συνόλου}} = 95^2 + 99^2 + \dots + 73^2 + 87^2 - 157176 = 4322,55$$

Το άθροισμα τετραγώνων επεμβάσεων δίνεται ως εξής:

$$AT_{\text{επεμβάσεων}} = n \sum_{i=1}^a (\bar{Y}_{i.} - \bar{Y}_{..})^2 = \sum_{i=1}^a \frac{Y_i^2}{n} - [\sum_{i=1}^a \sum_{j=1}^n Y_{ij}]^2 / an$$

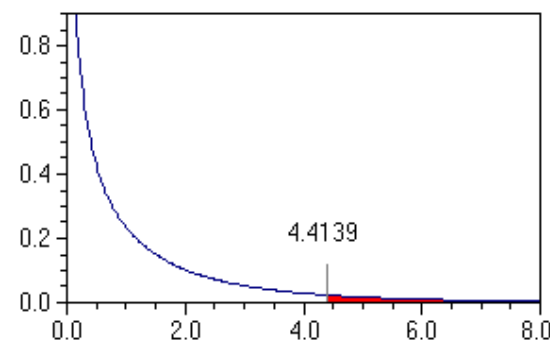
$$AT_{\text{επεμβάσεων}} = \frac{984^2}{10} + \frac{789^2}{10} - 157,176 = 96825 + 62252 - 157176 = 1901,25$$

Πίνακας Ανάλυσης της Παραλλακτικότητας

Πηγή παρ/τητας	ΒΕ	ΑΤ	ΜΤ	F	F πιν.	Prob > F
Επεμβάσεις	1	1901,25	1901,25	14,134	4,41	0,0014*
Υπόλοιπο	18	2421,30	134,52			
Σύνολο	19	4322,55				

$$t^2 = 3,7595^2 = 14,1336$$

Επειδή η τιμή F που υπολογίστηκε (14,13) είναι μεγαλύτερη από την κρίσιμη τιμή F (4,41), απορρίπτεται η μηδενική υπόθεση H_0 σε $\alpha = 0,05$. Άρα οι μέσοι όροι των δύο επεμβάσεων διαφέρουν στατιστικά σημαντικά.



Πειραματικοί Σχεδιασμοί

```
> Ftest=read.table("C:\\\\Actuar\\\\Ftest.txt", header=TRUE, dec=",")  
> attach(Ftest)  
> Ftest
```

	Tr	Y
1	1	95
2	1	99
3	1	103
4	1	90
5	1	121
6	1	98
7	1	102
8	1	94
9	1	87
10	1	95
11	2	104
12	2	74
13	2	84
14	2	81
15	2	62
16	2	67
17	2	64
18	2	93
19	2	73
20	2	87

Πειραματικοί Σχεδιασμοί

```
> Tr=factor(Tr)
> shapiro.test(Y)
```

Shapiro-Wilk normality test

```
data: Y
W = 0.96409, p-value = 0.6284
```

```
> bartlett.test(Y~Tr)
```

Bartlett test of homogeneity of variances

```
data: Y by Tr
Bartlett's K-squared = 1.1059, df = 1, p-value = 0.293
```

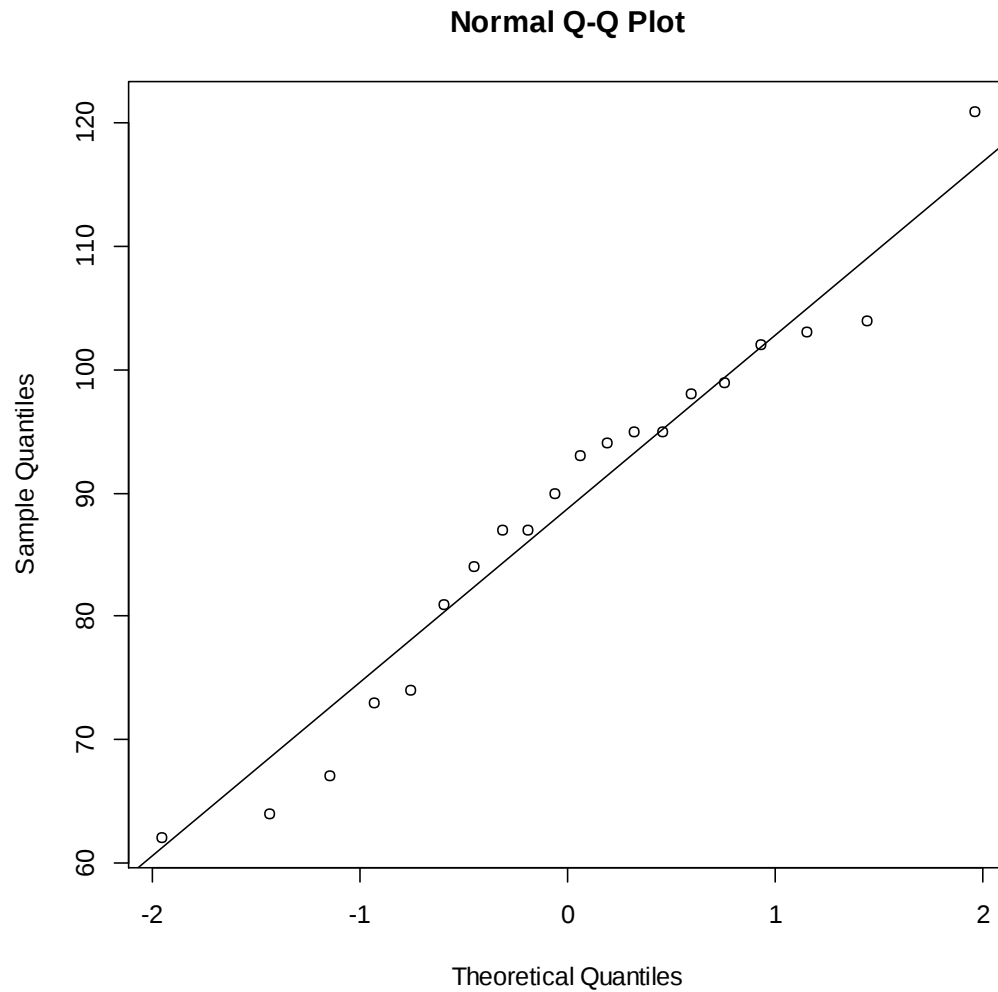
```
> fit=aov(Y~Tr)
> summary(fit)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Tr	1	1901	1901.3	14.13	0.00144 **
Residuals	18	2421	134.5		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

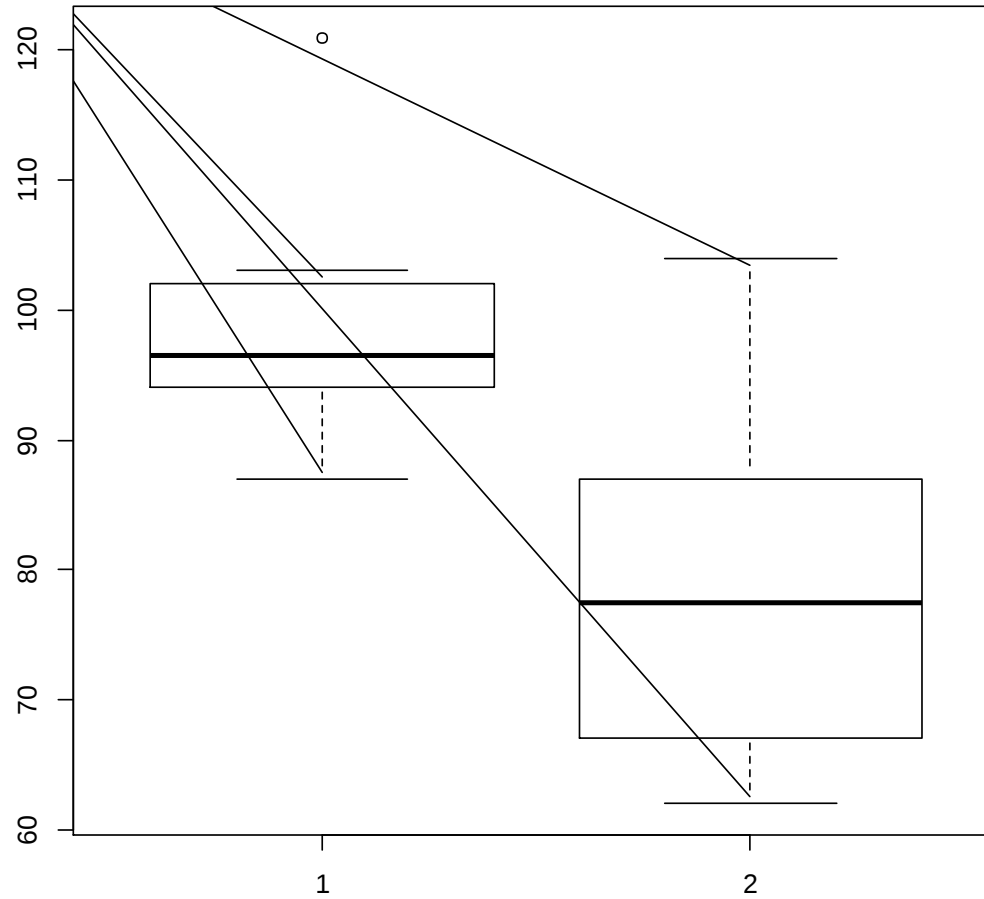
> qqnorm(Y)

> qqline(Y)



Πειραματικοί Σχεδιασμοί

> boxplot(Y~Tr)



Άσκηση 1. Μετρήθηκε η συγκέντρωση λυκοπενίου (mg/100g) σε φρέσκο χυμό και συμπύκνωμα τομάτας. Ελέγξτε αν υπάρχουν στατιστικά σημαντικές διαφορές.

Φρέσκος χυμός	9,2	8,5	9,4	8,9	9,2	9,3	9,0
Συμπύκνωμα	12,1	10,9	9,8	11,2	10,9	12,4	10,2

Άσκηση 2. Μετρήθηκε η συγκέντρωση σακχάρων με δυο μεθόδους (διαθλασίμετρο και όξινη υδρόλυση) σε έξι ποικιλίες γλυκού σόργου. Ελέγξτε αν υπάρχουν στατιστικά σημαντικές διαφορές.

Δείγμα	1	2	3	4	5	6
Διαθλασίμετρο	12,6	16,1	12,2	11,2	15,8	15,1
Όξινη υδρόλυση	10,7	14,6	12,3	10,3	14,9	12,4