

# Έλεγχοι υποθέσεων

Αναστάσιος Κατσιλέρος

Γεωπονικό Πανεπιστήμιο Αθηνών  
Εργαστήριο Βελτίωσης Φυτών και Γεωργικού Πειραματισμού

katsileros@aua.gr

Αθήνα 2020



Τα περιγραφικά μέτρα που υπολογίζονται από τον πληθυσμό ονομάζονται **παράμετροι του πληθυσμού** και είναι συνήθως άγνωστα, ενώ τα περιγραφικά μέτρα που υπολογίζονται από τμήμα του πληθυσμού ονομάζονται **στατιστικά** και αποτελούν εκτιμήσεις των αντίστοιχων, άγνωστων αλλά πραγματικών τιμών των παραμέτρων του πληθυσμού.

Τα στατιστικά  $\bar{Y}$  και  $s^2$  αποτελούν σημειακές εκτιμήσεις των αντίστοιχων παραμέτρων του πληθυσμού. Με βάση όμως τα δεδομένα ενός δείγματος είναι αδύνατον να προσδιορίσουμε ακριβώς τις πραγματικές τιμές των παραμέτρων. Η μέση τιμή του δείγματος παίρνει διαφορετικές τιμές αν πάρουμε άλλα δείγματα από τον ίδιο πληθυσμό. Επομένως για την εκτίμηση των παραμέτρων του πληθυσμού είναι χρήσιμο να ορίσουμε ένα **διάστημα εμπιστοσύνης** εντός του οποίου με βάση μια ορισμένη πιθανότητα, να περιέχει την πραγματική τιμή της παραμέτρου.

Έστω ένα δείγμα  $y_1, y_2, \dots, y_n$  από ένα πληθυσμό με κανονική κατανομή, με άγνωστη μέση τιμή  $\mu$  και με γνωστή διακύμανση  $\sigma^2$ . Επειδή η  $\bar{Y}$  ακολουθεί κανονική κατανομή  $N(\mu, \sigma^2/n)$ , η τυχαία μεταβλητή  $Z$  ακολουθεί την τυπική κανονική κατανομή.

$$Z = \frac{\bar{Y} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1) \quad (1)$$

Ορίζουμε ένα διάστημα  $[z_{\alpha/2}, z_{1-\alpha/2}]$  στο οποίο ανήκει η  $Z$ , με πιθανότητα  $1 - \alpha$ . Τα άκρα του διαστήματος  $z_{\alpha/2}$  και  $z_{1-\alpha/2}$  ονομάζονται **κρίσιμες τιμές**.

Ισχύει:

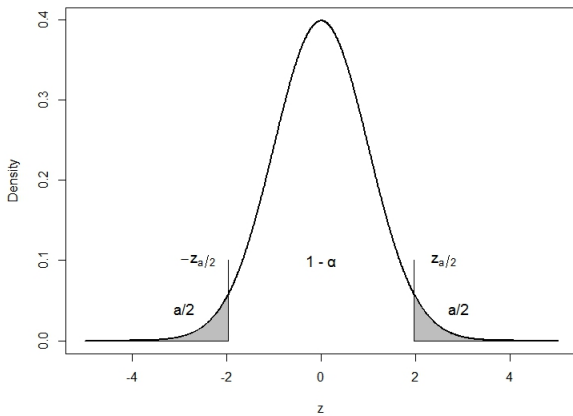
$$P(-z_{\alpha/2} \leq z \leq +z_{\alpha/2}) = 1 - \alpha$$

$$P(-z_{\alpha/2} \leq \frac{\bar{Y} - \mu}{\sigma/\sqrt{n}} \leq +z_{\alpha/2}) = 1 - \alpha$$

$$P(\bar{Y} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{Y} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}) = 1 - \alpha$$

Το διάστημα  $\bar{Y} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$  λέγεται **διάστημα εμπιστοσύνης** της μέσης τιμής  $\mu$  με πιθανότητα  $1 - \alpha$ .

Η πιθανότητα  $1 - \alpha$  ονομάζεται **επίπεδο εμπιστοσύνης** και οι ποσότητες  $\bar{Y} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$  και  $\bar{Y} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$  είναι τα **άνω και κάτω όρια εμπιστοσύνης**.

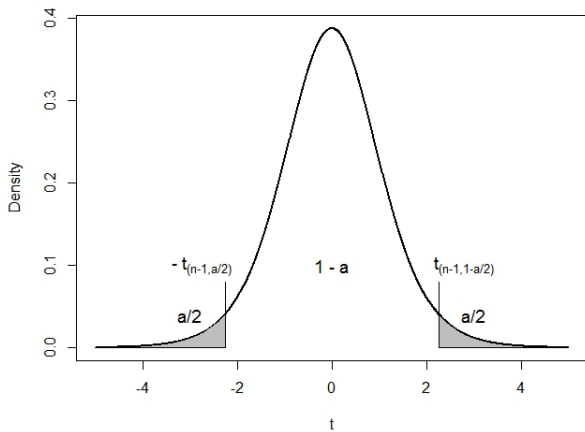


Έστω ένα δείγμα  $y_1, y_2, \dots, y_n$  από μια κανονική κατανομή με άγνωστη μέση τιμή  $\mu$  και άγνωστη διακύμανση  $\sigma^2$ .

$$t = \frac{\bar{Y} - \mu}{s/\sqrt{n}} \sim t_{n-1} \quad (2)$$

Ορίζουμε ένα διάστημα  $[t_{\alpha/2}, t_{1-\alpha/2}]$  στο οποίο ανήκει η  $t$ , με πιθανότητα  $1 - \alpha$ .

$$P(-t_{(n-1, \alpha/2)} \leq \frac{\bar{Y} - \mu}{s/\sqrt{n}} \leq t_{(n-1, \alpha/2)}) = 1 - \alpha$$
$$P(\bar{Y} - t_{(n-1, \alpha/2)} \frac{s}{\sqrt{n}} \leq \mu \leq \bar{Y} + t_{(n-1, \alpha/2)} \frac{s}{\sqrt{n}}) = 1 - \alpha$$





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** και η **εναλλακτική υπόθεση**. Η μηδενική υπόθεση ( $H_0$ ) είναι η υπόθεση με την οποία ξεκινάμε και υποθέτουμε πως η παράμετρος που μελετάμε (π.χ. η μέση τιμή  $\mu$ ) μπορεί να πάρει μια συγκεκριμένη τιμή ( $\mu_0$ ). Η εναλλακτική υπόθεση ( $H_1$ ) είναι η υπόθεση που δηλώνει ύπαρξης διαφοράς και ισχύει όταν απερρίφθη η μηδενική υπόθεση.

Τα στάδια που πραγματοποιούνται για τον έλεγχο υποθέσεων είναι τα εξής:

α) Αρχικά ο ερευνητής διαμορφώνει τη μηδενική  $H_0$  και εναλλακτική υπόθεση  $H_1$  και επιλέγει το επίπεδο σημαντικότητας  $\alpha$ .

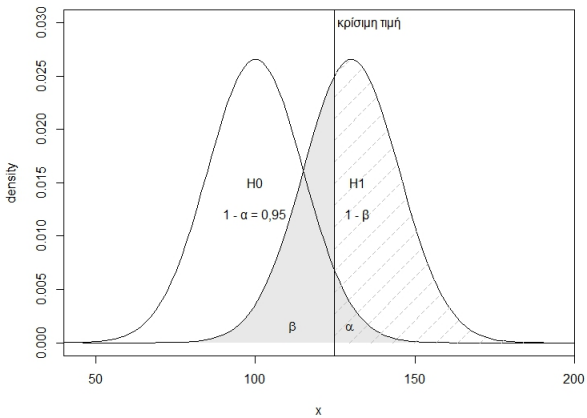
β) Επιλέγει τη στατιστική δοκιμασία ( $t$  – test,  $F$  – test) και με βάση την κατανομή της, του επιπέδου σημαντικότητας  $\alpha$  και το αν είναι αμφίπλευρη ή όχι η  $H_1$ , βρίσκει τις κρίσιμες τιμές που οριοθετούν την περιοχή απόρριψης ή όχι της μηδενικής υπόθεσης.

γ) Υπολογίζει την τιμή της στατιστική δοκιμασίας την οποία συγκρίνει με τις κρίσιμες τιμές και αποφασίζει για την απόρριψη ή όχι της μηδενικής υπόθεσης.

Πολλά στατιστικά προγράμματα δεν συγκρίνουν την τιμή της στατιστικής δοκιμασίας με τις κρίσιμες τιμές αλλά δίνουν απευθείας την τιμή **p-value** η οποία ορίζεται ως η ελάχιστη τιμή του επιπέδου σημαντικότητας για την οποία απορρίπτεται η  $H_0$ .

Η απόρριψη της μηδενικής υπόθεσης, όταν στην πραγματικότητα ισχύει λέγεται **Σφάλμα τύπου I**. Η πιθανότητα διάπραξης Σφάλματος τύπου I καθορίζεται από την επιλογή του επιπέδου σημαντικότητας  $\alpha$  από τον ερευνητή, που 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**.

Η μη απόρριψη μιας λανθασμένης μηδενικής υπόθεσης όταν πράγματι ισχύει λέγεται **Σφάλμα τύπου II**. Η πιθανότητα διάπραξης Σφάλματος τύπου II είναι  $\beta$  το οποίο δεν είναι γνωστό από πριν και δεν καθορίζεται. Για ένα ορισμένο αριθμό παρατηρήσεων η τιμή του  $\beta$  είναι αντίστροφα ανάλογη με αυτή του  $\alpha$ . 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.



## Έλεγχος υπόθεσης για τη μέση τιμή

Ο έλεγχος υπόθεσης για τη μέση τιμή ενός πληθυσμού που ακολουθεί κανονική κατανομή εξαρτάται από το αν η διακύμανση του πληθυσμού είναι γνωστή ή άγνωστη.

Στην περίπτωση που είναι γνωστή η διακύμανση χρησιμοποιείται η στατιστική δοκιμασία:

$$z = \frac{\bar{Y} - \mu_0}{\sigma/\sqrt{n}} \quad (3)$$

ενώ όταν είναι άγνωστη χρησιμοποιείται η στατιστική δοκιμασία:

$$t = \frac{\bar{Y} - \mu_0}{s/\sqrt{n}} \quad (4)$$

**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 εκατοστά.

Διατύπωση μηδενικής  $H_0 : \mu = 100$  και εναλλακτικής υπόθεσης  $H_1 : \mu \neq 100$ ,

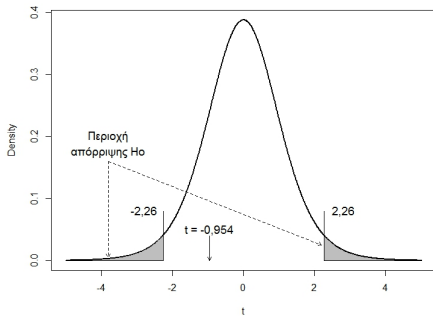
$$t = \frac{\bar{Y} - \mu_0}{s/\sqrt{n}} = \frac{98,5 - 100}{4,97/\sqrt{10}} = \frac{-1,5}{1,57} = -0,954$$

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





Επειδή ο έλεγχος είναι αμφίπλευρος η περιοχή απόρριψης της  $H_0$  περιλαμβάνει τιμές μικρότερες από  $-2,26$  και μεγαλύτερες από  $2,26$  ( $t=0,025,9$ ). Η τιμή από την δοκιμασία  $t$  ισούται με  $-0,954$  και δεν περιλαμβάνεται στην περιοχή απόρριψης της μηδενικής υπόθεσης, οπότε μπορούμε να υποθέσουμε ότι η μέση τιμή του πληθυσμού είναι 100 εκατοστά.

```
> x=c(101, 96, 93, 108, 102, 92, 98, 97, 95, 103)
> n=length(x)
> se = sqrt(var(x)/(n))
> t = (mean(x)-100)/se
> t
[1] -0.9539981
```

```
# Υπολογισμός κρίσιμης τιμής t (αμφίπλευρος έλεγχος  $\alpha/2$ )
> qt(p=0.975, df=9)
[1] 2.262157
# Επειδή η απόλυτη τιμή της t (0.9539) < της κρίσιμης τιμής
t, δεν απορρίπτεται η  $H_0$ 
# Υπολογισμός p-value
> 2*pt(-abs(-0.954), df=9)
[1] 0.3650026
# Επειδή p-value (0.365) > 0.05, δεν απορρίπτεται η  $H_0$ 
```

```
> 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

$t = -0.954$ ,  $df = 9$ ,  $p\text{-value} = 0.365$

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 υπόθεσης διαφοράς δύο μέσων τιμών

Αν οι διακυμάνσεις των δύο πληθυσμών  $\sigma_1^2$  και  $\sigma_2^2$  είναι γνωστές, η μεταβλητή  $Z$  ακολουθεί την τυπική κανονική κατανομή.

$$\frac{(\bar{Y}_1 - \bar{Y}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = Z \sim N(0, 1) \quad (5)$$

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 δοκιμασία:

$$Z = \frac{(\bar{Y}_1 - \bar{Y}_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \quad (6)$$

Όταν οι διακυμάνσεις 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 θα πρέπει να γίνει έλεγχος αν οι διακυμάνσεις των πληθυσμών από τους οποίους προήλθαν τα δείγματα είναι ίσες ή όχι.

Ο έλεγχος της υπόθεσης της ισότητας ή ομοιογένειας των διακυμάνσεων ( $H_0 : \sigma_1^2 = \sigma_2^2$ ), γίνεται με τη δοκιμασία:

$$F_{max} = \frac{s_{max}^2}{s_{min}^2} \quad (7)$$

Η τιμή από τη δοκιμασία συγκρίνεται 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 παρανομαστή.

Έστω δύο ανεξάρτητα τυχαία δείγματα από δύο κανονικούς πληθυσμούς με άγνωστες αλλά ίσες διακυμάνσεις, η στατιστική δοκιμασία για την σύγκριση των μέσων τιμών των δύο δειγμάτων είναι:

$$t = \frac{(\bar{Y}_1 - \bar{Y}_2)}{\sqrt{\frac{\sigma_p^2}{n_1} + \frac{\sigma_p^2}{n_2}}} \quad (8)$$

με  $n_1 + n_2 - 2$  βαθμούς ελευθερίας όπου  $\bar{Y}_1$  και  $\bar{Y}_2$  οι μέσες τιμές των δειγμάτων,  $n_1$  και  $n_2$  τα μεγέθη των δειγμάτων και  $s_p^2$  η εκτιμήτρια της κοινής διακύμανσης  $\sigma^2$  των δύο πληθυσμών.

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 - 1) + (n_2 - 1)} \quad (9)$$

Ο τύπος του διαστήματος εμπιστοσύνης δίνεται από τον τύπο:

$$(\bar{Y}_1 - \bar{Y}_2) \pm t_{\alpha/2, (n-1)} s_{(\bar{Y}_1 - \bar{Y}_2)} \quad (10)$$

όπου:

$$s_{\bar{Y}_1 - \bar{Y}_2} = \sqrt{\frac{\sigma_p^2}{n_1} + \frac{\sigma_p^2}{n_2}} \quad (11)$$

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                 |
| $n_1 = 10$         | $n_2 = 10$         |
| $\bar{Y}_1 = 98,4$ | $\bar{Y}_2 = 78,9$ |
| $s_1^2 = 87,6$     | $s_2^2 = 181,4$    |

Έλεγχος ομοιογένειας των δύο διακυμάνσεων

$$H_0 : \sigma_1^2 = \sigma_2^2, H_1 : \sigma_1^2 \neq \sigma_2^2,$$

$$F_{max} = \frac{s_{max}^2}{s_{min}^2} = \frac{181,4}{87,6} = 2,09$$

Η κρίσιμη τιμή  $F$  για  $\alpha/2 = 0,025$  και βαθμούς ελευθερίας για τον αριθμητή  $n_1 - 1 = 9$  και για τον παρανομαστή  $n_2 - 1 = 9$  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 συμπηφισμένη διακύμανση υπολογίζεται ως εξής:

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 - 1) + (n_2 - 1)} = \frac{s_1^2 + s_2^2}{2} = \frac{181,4 + 87,4}{2} = 134,05$$

Το τυπικό σφάλμα της διαφοράς:

$$\sigma_{\bar{Y}_1 - \bar{Y}_2} = \sqrt{\frac{2 * 134,05}{10}} = \sqrt{26,81} = 5,18$$

Εφαρμόζεται η δοκιμασία του t:

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

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.

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

$$\bar{Y} \pm t_{\alpha/2, (n-1)} s_{\bar{Y}} = (98,4 - 78,9) \pm 2,1 * 5,18 = 19,5 \pm 10,9$$

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

```
> 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

```
> x=c(95, 99, 103, 90, 121, 98, 102, 94, 87, 95)
> y=c(104, 74, 84, 81, 62, 67, 64, 93, 73, 87)
> n1=length(x)
> n2=length(y)
> se = sqrt( (1/n1 + 1/n2) * ((n1-1)*var(x) +
(n2-1)*var(y))/(n1+n2-2))
> se
[1] 5.186842
> df = n1+n2-2
> df
[1] 18
> t = (mean(x)-mean(y))/se
> t
[1] 3.759513
> 2*pt(-abs(t),df)
[1] 0.0014352
```

```
> 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

$t = 3.7595$ ,  $df = 18$ ,  $p\text{-value} = 0.001435$

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

Όταν οι διακυμάνσεις είναι άνισες δεν μπορούμε να υπολογίσουμε ένα κοινό  $s^2$  από τα δύο δείγματα. Ο έλεγχος της διαφοράς των μέσων τιμών των δειγμάτων γίνεται με τον τύπο:

$$t' = \frac{\bar{Y}_1 - \bar{Y}_2}{s_{\bar{Y}_1 - \bar{Y}_2}} \quad (12)$$

όπου:

$$s_{\bar{Y}_1 - \bar{Y}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \quad (13)$$

και οι διορθωμένοι βαθμοί ελευθερίας υπολογίζονται ως εξής:

$$BE' = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{(s_1^2/n_1)^2/(n_1 - 1) + (s_2^2/n_2)^2/(n_2 - 1)} \quad (14)$$



**Παράδειγμα 3.** Τα δεδομένα αναφέρονται στην ανάπτυξη εκφύτων σε θρεπτικό υπόστρωμα MS με και χωρίς προσθήκη φυτορρυθμιστικής ουσίας (IAA).

| 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               |
| $n_1 = 7$          | $n_2 = 8$          |
| $\bar{Y}_1 = 12,9$ | $\bar{Y}_2 = 27,7$ |
| $s_1^2 = 2,8$      | $s_2^2 = 28,3$     |

Έλεγχος ομοιογένειας των δύο διακυμάνσεων

$$H_0 : \sigma_1^2 = \sigma_2^2, H_1 : \sigma_1^2 \neq \sigma_2^2,$$

$$F_{max} = \frac{s_{max}^2}{s_{min}^2} = \frac{27,7}{12,9} = 10,13$$

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:

$$s_{\bar{Y}_1 - \bar{Y}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{12,9}{7} + \frac{27,7}{8}} = 1,98$$

Ο στατιστικός έλεγχος είναι:

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

$$BE' = \frac{(12,9/7 + 27,7/8)^2}{(12,9/7)^2/(7-1) + (27,7/8)^2/(8-1)} = 8,54 \approx 9$$

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 τα θρεπτικά υποστρώματα διαφέρουν στατιστικά σημαντικά ως προς την ανάπτυξη εκφύτων. Η πιθανότητα αυτό το συμπέρασμα να είναι λάθος, είναι το πολύ 0,05.

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

$$\bar{Y} \pm t_{\alpha/2, (n-1)} s_{\bar{Y}} = (12,9 - 27,7) \pm 2,26 * 1,98 = -14,7 \pm 4,5$$

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

```
> 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

```
> n1=length(x)
> n2=length(y)
> se = sqrt( (var(x)/n1) + (var(y)/n2) )
> df = ( (var(x)/n1 + var(y)/n2)^2 )/( (var(x)/n1)^2/(n1-1) +
(var(y)/n2)^2/(n2-1) )
> df
[1] 8.541184
> t = (mean(x)-mean(y))/se
> t
[1] -7.419805
> 2*pt(-abs(t),df)
[1] 5.309627e-05
```

```
> 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

$t = -7.4198$ ,  $df = 8.5412$ ,  $p\text{-value} = 5.31e-05$

alternative H: 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

**Παράδειγμα 4.** Τα δεδομένα αναφέρονται στην περιεκτικότητα σε mg βιταμίνης C ενός τροφίμου πριν και μετά από θερμική κατεργασία.

| <b>Ζεύγη</b> | <b>πριν</b> | <b>μετά</b> | <b>D</b> |
|--------------|-------------|-------------|----------|
| 1            | 6,85        | 6,28        | 0,57     |
| 2            | 7,05        | 6,84        | 0,21     |
| 3            | 8,24        | 7,07        | 1,17     |
| 4            | 5,2         | 5,05        | 0,15     |
| 5            | 6,18        | 6,16        | 0,02     |
| 6            | 6,71        | 6,67        | 0,04     |
| 7            | 7,01        | 6,49        | 0,52     |
| 8            | 5,46        | 5,24        | 0,22     |
| Μέση Τιμή    | 6,59        | 6,22        | 0,36     |
| Άθροισμα     | 52,7        | 49,8        | 2,9      |



$$s_{\bar{D}} = \frac{\sum D_i^2 - (\sum D_i)^2/n}{n-1} = \frac{2,081 - (2,9)^2/8}{7} = 0,147$$

$$s_{\bar{D}} = \sqrt{s_{\bar{D}}^2/n} = \sqrt{0,147/8} = 0,135$$

Ο έλεγχος της σημαντικότητας της διαφοράς:

$$t = \frac{\bar{D}}{s_{\bar{D}}} = \frac{0,36}{0,135} = 2,67$$

Η κρίσιμη τιμή  $t$  για  $\alpha = 0,05$  και βαθμούς ελευθερίας 7 ισούται με 2,36. Επειδή η τιμή που υπολογίστηκε (2,67) είναι μεγαλύτερη από την κρίσιμη τιμή  $t$ , απορρίπτεται η  $H_0$ .

Ο τύπος του διαστήματος εμπιστοσύνης είναι:

$$\bar{D} \pm t_{\alpha/2, (n-1)} s_{\bar{D}} = 0,36 \pm 2,36 * 0,135 = 0,36 \pm 0,32$$

```
> x=c(6.85, 7.05, 8.24, 5.2, 6.18, 6.71, 7.01, 5.46)
> y=c(6.28, 6.84, 7.07, 5.05, 6.16, 6.67, 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 = 2.673, df = 7, p-value = 0.03186

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

95 percent confidence interval:

0.04181681 0.68318319

sample estimates:

mean of the differences

0.3625

Η πιθανότητα να υποπέσουμε σε σφάλμα τύπου I σε μία σύγκριση των μέσων τιμών δύο δειγμάτων ονομάζεται **βαθμός σφάλματος κατά σύγκριση**  $\alpha_{K\Sigma}$  και ισούται με  $\alpha$ .

Η πιθανότητα διάπραξης τουλάχιστον ενός σφάλματος τύπου I σε μία ομάδα ανεξάρτητων συγκρίσεων ονομάζεται **βαθμός σφάλματος κατά πείραμα**  $\alpha_{K\Pi}$  και ισούται με  $\alpha_{K\Pi} = 1 - (1 - \alpha)^c$ , όπου  $c$  ο αριθμός των ανεξάρτητων συγκρίσεων

Η δοκιμασία του  $F$  ελέγχει αν οι παρατηρούμενες διαφορές ανάμεσα στις μέσες τιμές των δειγμάτων οφείλονται σε τυχαίους λόγους ή σε συστηματική διαφορά ανάμεσα στα δείγματα.

Προϋπόθεση για την εφαρμογή της δοκιμασίας του  $F$  είναι οι πληθυσμοί από τους οποίους προέρχονται τα δείγματα να ακολουθούν κανονική κατανομή, να έχουν ίσες διακυμάνσεις και οι παρατηρήσεις να είναι ανεξάρτητες.

Η δοκιμασία του  $F$  είναι ο λόγος δύο διακυμάνσεων και οι διακυμάνσεις που ελέγχει είναι οι δύο διαφορετικές εκτιμήσεις της πληθυσμιακής διακύμανσης  $\sigma^2$ . Η πρώτη προέρχεται από τον υπολογισμό απευθείας από την εντός των δειγμάτων διακύμανση και η δεύτερη εμμέσως από την μεταξύ των δειγμάτων διακύμανση.

Η εντός των δειγμάτων διακύμανση ή συμπηφισμένη διακύμανση  $\sigma_p^2$  αποτελεί μια καλή εκτίμηση της διακύμανσης  $\sigma^2$ , ανεξάρτητα αν οι μέσοι των δειγμάτων διαφέρουν και υπολογίζεται ως εξής:

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + \dots + (n_k - 1)s_k^2}{(n_1 - 1) + \dots + (n_k - 1)} \quad (15)$$

Η μεταξύ των δειγμάτων διακύμανση  $s_T^2$  αποτελεί μία άλλη εκτίμηση της διακύμανσης  $\sigma^2$ , αν υποθέσουμε όμως ότι δεν υπάρχουν διαφορές μεταξύ των μέσων τιμών των δειγμάτων. Αν θεωρήσουμε τις μέσες τιμές των  $k$  δειγμάτων ως ένα δείγμα, η διακύμανση των μέσων τιμών των δειγμάτων υπολογίζεται ως εξής:

$$s_{\bar{Y}}^2 = \frac{\sum_{k=1}^a (Y_i - \bar{Y})^2}{k - 1} \quad (16)$$

Επειδή όμως η διακύμανση του πληθυσμού  $\sigma^2$  είναι  $n$  φορές μεγαλύτερη της διακύμανσης των μέσων τιμών των δειγμάτων  $s_{\bar{Y}}^2 = \sigma^2/n$ , η μεταξύ των δειγμάτων διακύμανση  $s_T^2$  υπολογίζεται ως εξής:

$$s_T^2 = ns_{\bar{Y}}^2 = n \frac{\sum_{k=1}^a (Y_i - \bar{Y})^2}{k - 1} \quad (17)$$

Για να ελέγξουμε την ορθότητα ή μη της μηδενικής υπόθεσης ( $H_0 : \mu_1 = \mu_2 = \dots = \mu_k$ ), η γενική δοκιμασία  $F$  είναι η εξής:

$$F = \frac{\text{διακύμανση μεταξύ των δειγμάτων}}{\text{διακύμανση εντός των δειγμάτων}} = \frac{s_T^2}{s_P^2} \quad (18)$$

η οποία συγκρίνεται με την κρίσιμη τιμή του πίνακα  $F$  για  $k-1$  βαθμούς ελευθερίας του αριθμητή και  $N-k$  βαθμούς ελευθερίας του παρανομαστή.

Αν η μηδενική υπόθεση είναι αληθής, δηλαδή οι μέσες τιμές δεν διαφέρουν, αναμένουμε ο λόγος των δύο διακυμάνσεων να είναι κοντά στην μονάδα, ενώ αν δεν είναι αληθής, ο λόγος θα είναι πολύ μεγαλύτερος της μονάδας επειδή ο αριθμητής περιέχει και τη μεταβλητότητα που οφείλεται στη συστηματική διαφορά ανάμεσα στα δείγματα.

**Παράδειγμα 5.** Τα δεδομένα αφορούν πείραμα που πραγματοποίησε ένα μηχανικός, για να ελέγξει την επίδραση τριών διαφορικών συνθηκών αποθήκευσης, στην διάρκεια ζωής μπαταριών.

| 1η επέμβαση     | 2η επέμβαση        | 3η επέμβαση        |
|-----------------|--------------------|--------------------|
| 5               | 18                 | 6                  |
| 8               | 14                 | 12                 |
| 9               | 16                 | 14                 |
| 7               | 13                 | 17                 |
| 11              | 17                 | 13                 |
| $n_1 = 5$       | $n_2 = 5$          | $n_3 = 5$          |
| $\bar{Y}_1 = 8$ | $\bar{Y}_2 = 15,6$ | $\bar{Y}_3 = 12,4$ |
| $s_1^2 = 5$     | $s_2^2 = 4,3$      | $s_3^2 = 16,3$     |



Διατύπωση υποθέσεων:

Μηδενική υπόθεση  $H_0 : \mu_1 = \mu_2 = \mu_3$

Εναλλακτική υπόθεση  $H_1 : \mu_1 \neq \mu_2 \neq \mu_3$

Υπολογισμός εντός των δειγμάτων διακύμανσης  $s_p^2$

$$\begin{aligned} s_p^2 &= \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + (n_3 - 1)s_3^2}{(n_1 - 1) + (n_2 - 1) + (n_3 - 1)} = \frac{s_1^2 + s_2^2 + s_3^2}{3} = \\ &= \frac{5 + 4,3 + 16,3}{3} = 8,53 \end{aligned}$$

Υπολογισμός μεταξύ των δειγμάτων διακύμανσης  $s_\tau^2$

$$\begin{aligned} s_\tau^2 &= ns_{\bar{Y}}^2 = n \frac{\sum_{k=1}^a (Y_i - \bar{Y})^2}{k - 1} = \\ &= 5 \frac{[(8 - 12)^2 + (15,6 - 12)^2 + (12,4 - 12)^2]}{3 - 1} = 72,8 \end{aligned}$$

Έλεγχος των διακυμάνσεων με την δοκιμασία  $F$ :

$$F = \frac{s_T^2}{s_p^2} = \frac{72,8}{8,53} = 8,531$$

Η κρίσιμη τιμή από τον πίνακα του  $F$  για  $n - 1 = 2$  βαθμούς ελευθερίας του αριθμητή και  $N - k = 12$  βαθμούς ελευθερίας του παρανομαστή και σε επίπεδο σημαντικότητας  $\alpha = 0,05$ , είναι 3,24. Επειδή ο λόγος του  $F$  (8,531) είναι μεγαλύτερος από την κρίσιμη τιμή του  $F$  (3,885), απορρίπτεται η μηδενική υπόθεση και επομένως τουλάχιστον δύο επεμβάσεις διαφέρουν στατιστικά σημαντικά μεταξύ τους.

Για την ανάλυση σύνθετων πειραμάτων στα οποία υπάρχουν πολλές πηγές παραλλακτικότητας, εισήχθη από τον Sir Ronald A. Fisher το 1919, η στατιστική μέθοδος της **Ανάλυσης Διακύμανσης** ή **Παραλλακτικότητας** (Analysis of Variance, ANOVA), η οποία αποτελεί το βασικότερο στατιστικό εργαλείο στον πειραματικό σχεδιασμό.

Η ανάλυση διακύμανσης διαχωρίζει τη συνολική παραλλακτικότητα των δεδομένων του πειράματος στις διάφορες πηγές παραλλακτικότητας και στην συνέχεια τις ελέγχει με την δοκιμασία του  $F$ .

Το γραμμικό μοντέλο που περιγράφει τις επιδράσεις σε μία παρατήρηση, είναι το εξής:

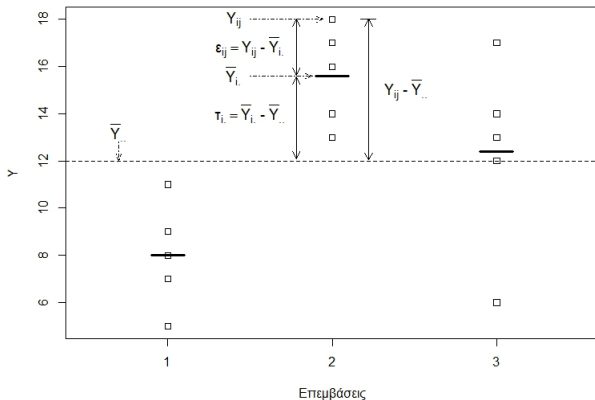
$$Y_{ij} = \mu + \tau_i + \epsilon_{ij} \quad (19)$$

όπου  $Y_{ij}$  η  $j$  παρατήρηση της  $i$  επέμβασης

$\mu$  ο γενικός μέσος του πληθυσμού

$\tau_i$  η επίδραση της  $i$  επέμβασης

$\epsilon_{ij}$  το πειραματικό σφάλμα,  $N(0, \sigma^2)$



Επειδή τα  $\mu$ ,  $\tau_i$  και  $\epsilon_{ij}$  δεν είναι γνωστά, εκτιμούνται από τα δεδομένα του πειράματος. Ο γενικός μέσος  $\mu$  εκτιμάται από το μέσο  $\bar{Y}_{..}$ , η επίδραση των επεμβάσεων  $\tau_i$  από την απόκλιση  $(\bar{Y}_i - \bar{Y}_{..})$  και το πειραματικό σφάλμα  $\epsilon_{ij}$  από την απόκλιση  $(Y_{ij} - \bar{Y}_i)$ .

Αν στο γραμμικό μοντέλο, αντικαταστήσουμε τις άγνωστες παραμέτρους με τους αντίστοιχους εκτιμητές, το μοντέλο γίνεται ως εξής:

$$Y_{ij} = \bar{Y}_{..} + (\bar{Y}_i - \bar{Y}_{..}) + (Y_{ij} - \bar{Y}_i) \quad (20)$$

Αν από την παραπάνω τύπο αφαιρεθεί από τις δύο πλευρές η μέση τμή  $\bar{Y}_{..}$  τότε

$$(Y_{ij} - \bar{Y}_{..}) = (\bar{Y}_i - \bar{Y}_{..}) + (Y_{ij} - \bar{Y}_i) \quad (21)$$

αν υψώσουμε και τα δύο μέλη της σχέσης στο τετράγωνο,

$$(Y_{ij} - \bar{Y}_{..})^2 = (\bar{Y}_i - \bar{Y}_{..})^2 + 2(Y_{ij} - \bar{Y}_i)(\bar{Y}_i - \bar{Y}_{..}) + (Y_{ij} - \bar{Y}_i)^2 \quad (22)$$

το επαναλάβουμε για όλες τις παρατηρήσεις και αθροίσουμε τα αποτελέσματα, η σχέση δίνεται ως εξής:

$$\begin{aligned} \sum_{i=1}^a \sum_{j=1}^n (Y_{ij} - \bar{Y}_{..})^2 &= \sum_{i=1}^a \sum_{j=1}^n (\bar{Y}_i - \bar{Y}_{..})^2 + \\ &2 \sum_{i=1}^a \sum_{j=1}^n (Y_{ij} - \bar{Y}_i)(\bar{Y}_i - \bar{Y}_{..}) + \sum_{i=1}^a \sum_{j=1}^n (Y_{ij} - \bar{Y}_i)^2 \end{aligned} \quad (23)$$

επειδή όμως  $\sum_{i=1}^a \sum_{j=1}^n (\bar{Y}_i - \bar{Y}_{..})^2 = n \sum_{i=1}^a (\bar{Y}_i - \bar{Y}_{..})^2$

και  $\sum_{j=1}^n (Y_{ij} - \bar{Y}_i) = Y_i - n\bar{Y}_i = Y_i - n(Y_i/n) = 0$

η σχέση γράφεται ως εξής:

$$\sum_{i=1}^a \sum_{j=1}^n (Y_{ij} - \bar{Y}_{..})^2 = n \sum_{i=1}^a (\bar{Y}_i - \bar{Y}_{..})^2 + \sum_{i=1}^a \sum_{j=1}^n (Y_{ij} - \bar{Y}_i)^2 \quad (24)$$

## Πίνακας Ανάλυσης Διακύμανσης

| Πηγή Παραλ/τητας                | Βαθμοί Ελευθερίας | Άθροισμα Τετραγώνων                                   | Μέσο Άθροισμα Τετραγώνων    | Δοκιμασία F                       | ΘΣΜΤ <sup>1</sup>                                | ΘΣΜΤ <sup>2</sup>           |
|---------------------------------|-------------------|---|-----------------------------|-----------------------------------|--|-----------------------------|
| Μεταξύ των επεμβάσεων           | a-1               | $n \sum_{i=1}^a (\bar{Y}_i - \bar{Y}_{..})^2$         | $AT_{\varepsilon\pi}/(a-1)$ | $MT_{\varepsilon\pi}/MT_{\nu\pi}$ | $\sigma^2 + \frac{n \sum_{i=1}^a \tau_i^2}{a-1}$ | $\sigma^2 + n\sigma_\tau^2$ |
| Εντός των επεμβάσεων ή Υπόλοιπο | a(n-1)            | $\sum_{i=1}^a \sum_{j=1}^n (Y_{ij} - \bar{Y}_i)^2$    | $AT_{\nu\pi}/a(n-1)$        |                                   | $\sigma^2$                                       | $\sigma^2$                  |
| Σύνολο                          | an-1              | $\sum_{i=1}^a \sum_{j=1}^n (Y_{ij} - \bar{Y}_{..})^2$ |                             |                                   |  |                             |

<sup>1</sup>Θεωρητική Σύσταση Μέσου Τετραγώνου με σταθερές επιδράσεις, <sup>2</sup>Θεωρητική Σύσταση Μέσου Τετραγώνου με τυχαίες επιδράσεις



```
> trt = factor(rep(c(1:3), each = 5))
> Y=c(5,8,9,7,11,18,14,16,13,17,6,12,14,17,13)
> data=data.frame(trt, Y)
> data
  trt Y
1 1 5
2 1 8
3 1 9
4 1 7
5 1 11
...
10 2 17
11 3 6
12 3 12
13 3 14
14 3 17
15 3 13
```

```
> grand_mean = mean(Y)
> trt_mean = tapply(Y, trt, mean)
> n = tapply(Y, trt, length)
> N=length(Y)
> a=length(trt_mean)
> SS_total = sum((Y - grand_mean)^2)
> SS_trt = sum(n*(trt_mean - grand_mean)^2)
> SS_error= SS_total - SS_trt
> df_total = N-1
> df_trt = a-1
> df_error= df_total-df_trt
> MS_trt = SS_trt/df_trt
> MS_error= SS_error/df_error
> F_value = MS_trt/MS_error
> p-value = pf(F_value, df_trt, df_error, lower.tail = FALSE)
```

```

> anova = matrix(c(df_trt, df_error, df_total, SS_trt,
SS_error, SS_total, MS_trt, MS_error, NA, F_value, NA, NA,
p_value, NA, NA), ncol = 5)
> dimnames(anova) = list(c("Treatment", "Error", "Total"),
c("Df", "Sum_Sq", "Mean_Sq", "F_value", "P-value"))
> printCoefmat(anova, signif.stars = TRUE, has.Pvalue =
TRUE, digits = 2, na.print = "")

```

|           | Df   | Sum_Sq | Mean_Sq | F_value | P_value  |
|-----------|------|--------|---------|---------|----------|
| Treatment | 2.0  | 145.6  | 72.8    | 8.5     | 0.005 ** |
| Error     | 12.0 | 102.4  | 8.5     |         |          |
| Total     | 14.0 | 248.0  |         |         |          |

---

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

## Πίνακας Ανάλυσης Διακύμανσης (ANOVA)

```
> fit=aov(Y~ trt, data)
```

```
> summary(fit)
```

|           | Df | Sum Sq | Mean Sq | F     | value Pr(>F) |
|-----------|----|--------|---------|-------|--------------|
| trt       | 2  | 145.6  | 72.80   | 8.531 | 0.00496 **   |
| Residuals | 12 | 102.4  | 8.53    |       |              |

---

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

Υπολογισμός κρίσιμης τιμής  $F$  για  $\alpha=0,05$ , ΒΕ αριθμητή (δειγμάτων) = 2 και ΒΕ παρανομαστή (υπολοίπου) = 12

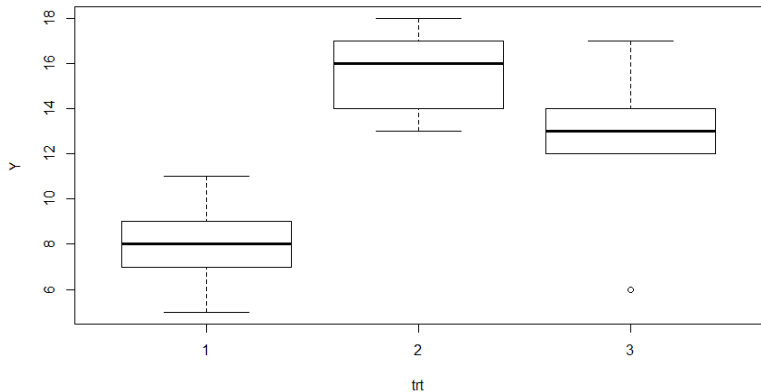
```
> qf(0.95,2,12)
[1] 3.885294
```

Υπολογισμός  $p$ -value για  $F = 8,531$

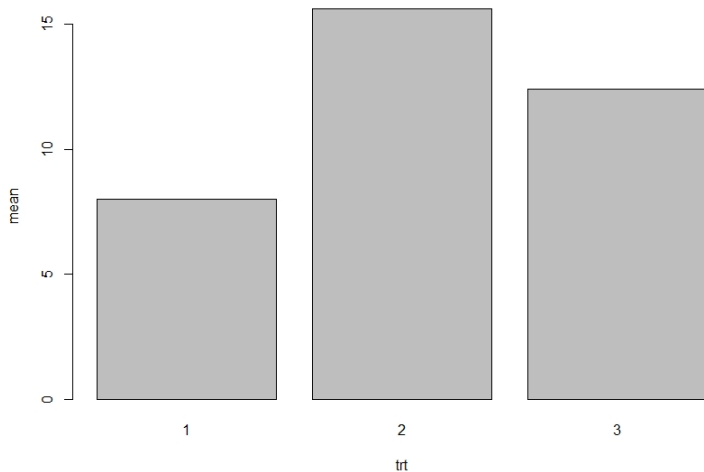
```
> pf(8.531, df1=2, df2=12, lower.tail = FALSE)
[1] 0.004956036
```

Επειδή  $F$  πειραματος ( $8,53$ )  $>$   $F$  κρίσιμης τιμής ( $3,88$ ) (ή  $p$ -value  $0,00496 < \alpha=0,05$ ), απορρίπτεται η  $H_0$ .

```
> plot(Y~trt)
```

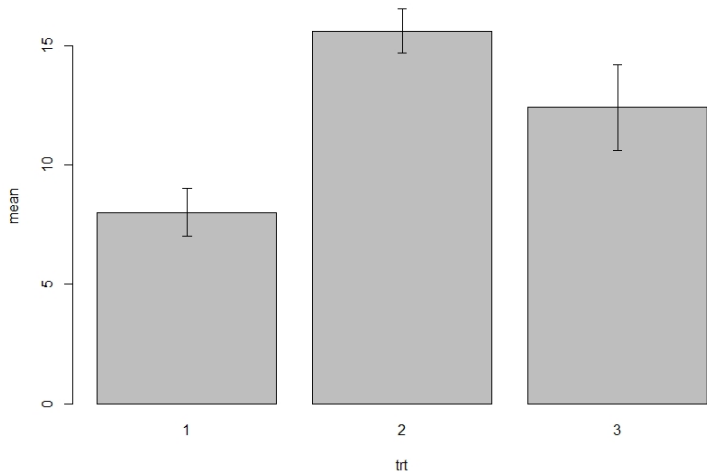


```
barplot(trt_mean, xlab = "trt", ylab = "mean")
```



```
trt_mean = tapply(Y, trt, mean)
trt_sd = tapply(Y, trt, sd)
trt_n=tapply(Y, trt, length)
trt_se = trt_sd/sqrt(trt_n)
plotmax = max(trt_mean+trt_se)
barcenters = barplot(trt_mean, ylim=c(0,plotmax), xlab =
"trt", ylab = "mean")
segments(barcenters, trt_mean-trt_se, barcenters,
trt_mean+trt_se, lwd=1)
arrows(barcenters, trt_mean - trt_se, barcenters, trt_mean+
trt_se, lwd = 1, angle = 90, code = 3, length = 0.05)
```





Εάν η στατιστική δοκιμασία του  $F$  είναι σημαντική, απορρίπτεται η μηδενική υπόθεση και ο ερευνητής αποδέχεται ότι υπάρχουν διαφορές μεταξύ των δειγμάτων - επεμβάσεων. Η ανάλυση διακύμανσης όμως δεν παρέχει πληροφορίες για το πως διαφέρουν οι επεμβάσεις και επομένως ο ερευνητής πρέπει να επιλέξει μία από τις μεθόδους πολλαπλών συγκρίσεων των επεμβάσεων, για να ελέγξει αυτές τις διαφορές.

## Μέθοδος Ελάχιστης Σημαντικής Διαφοράς

```
> library(agricolae)
```

```
> LSD.test(fit, "trt", console = TRUE)
```

```
Mean Square Error: 8.533333
```

```
trt, means and individual ( 95 %) CI
```

|   | Y    | std      | r | LCL       | UCL      | Min | Max |
|---|------|----------|---|-----------|----------|-----|-----|
| 1 | 8.0  | 2.236068 | 5 | 5.153611  | 10.84639 | 5   | 11  |
| 2 | 15.6 | 2.073644 | 5 | 12.753611 | 18.44639 | 13  | 18  |
| 3 | 12.4 | 4.037326 | 5 | 9.553611  | 15.24639 | 6   | 17  |

```
Alpha: 0.05 ; DF Error: 12
```

```
Critical Value of t: 2.178813
```

```
least Significant Difference: 4.025402
```

```
Treatments with the same letter are not significantly different.
```

|   | Y    | groups |
|---|------|--------|
| 2 | 15.6 | a      |
| 3 | 12.4 | a      |
| 1 | 8.0  | b      |

## Μέθοδος Duncan

```
> duncan.test(fit, "trt", console = TRUE)
```

Mean Square Error: 8.533333

trt, means

|   | Y    | std      | r | Min | Max |
|---|------|----------|---|-----|-----|
| 1 | 8.0  | 2.236068 | 5 | 5   | 11  |
| 2 | 15.6 | 2.073644 | 5 | 13  | 18  |
| 3 | 12.4 | 4.037326 | 5 | 6   | 17  |

Alpha: 0.05 ; DF Error: 12

Critical Range

| 2        | 3        |
|----------|----------|
| 4.025402 | 4.213441 |

Means with the same letter are not significantly different.

| Y      | groups |
|--------|--------|
| 2 15.6 | a      |
| 3 12.4 | a      |
| 1 8.0  | b      |

## Μέθοδος Student-Newman-Keul

```
> SNK.test(fit, "trt", console = TRUE)
```

Mean Square Error: 8.533333

trt, means

|   | Y    | std      | r | Min | Max |
|---|------|----------|---|-----|-----|
| 1 | 8.0  | 2.236068 | 5 | 5   | 11  |
| 2 | 15.6 | 2.073644 | 5 | 13  | 18  |
| 3 | 12.4 | 4.037326 | 5 | 6   | 17  |

Alpha: 0.05 ; DF Error: 12

Critical Range

| 2        | 3        |
|----------|----------|
| 4.025402 | 4.928934 |

Means with the same letter are not significantly different.

|   | Y    | groups |
|---|------|--------|
| 2 | 15.6 | a      |
| 3 | 12.4 | a      |
| 1 | 8.0  | b      |

## Μέθοδος Έντιμης Σημαντικής Διαφοράς

```
> HSD.test(fit, "trt", console = TRUE)
```

Mean Square Error: 8.533333

trt, means

|   | Y    | std      | r | Min | Max |
|---|------|----------|---|-----|-----|
| 1 | 8.0  | 2.236068 | 5 | 5   | 11  |
| 2 | 15.6 | 2.073644 | 5 | 13  | 18  |
| 3 | 12.4 | 4.037326 | 5 | 6   | 17  |

Alpha: 0.05 ; DF Error: 12

Critical Value of Studentized Range: 3.772929

Minimum Significant Difference: 4.928934

Treatments with the same letter are not significantly different.

|   | Y    | groups |
|---|------|--------|
| 2 | 15.6 | a      |
| 3 | 12.4 | ab     |
| 1 | 8.0  | b      |

## Μέθοδος Scheffe

```
> scheffe.test(fit, "trt", console = TRUE)
```

```
Mean Square Error : 8.533333
```

```
trt, means
```

|   | Y    | std      | r | Min | Max |
|---|------|----------|---|-----|-----|
| 1 | 8.0  | 2.236068 | 5 | 5   | 11  |
| 2 | 15.6 | 2.073644 | 5 | 13  | 18  |
| 3 | 12.4 | 4.037326 | 5 | 6   | 17  |

```
Alpha: 0.05 ; DF Error: 12
```

```
Critical Value of F: 3.885294
```

```
Minimum Significant Difference: 5.150107
```

```
Means with the same letter are not significantly different.
```

|   | Y    | groups |
|---|------|--------|
| 2 | 15.6 | a      |
| 3 | 12.4 | ab     |
| 1 | 8.0  | b      |

Οι βασικές προϋποθέσεις για την ανάλυση της διακύμανσης, είναι ότι οι πληθυσμοί από τους οποίους προήλθαν τα δείγματα - επεμβάσεις να ακολουθούν την κανονική κατανομή, οι διακυμάνσεις των πληθυσμών να είναι ίσες ή ομοιογενείς (ομοσκεδαστικότητα) και οι παρατηρήσεις να είναι ανεξάρτητες μεταξύ τους. Διαφορετικά αν για τον έλεγχο χρησιμοποιηθούν τα **υπόλοιπα**  $\varepsilon_{ij}$  (residuals) ή τα **τυποποιημένα υπόλοιπα**  $\varepsilon_{ijstd}$  (standardized residuals), αυτά πρέπει να είναι ανεξάρτητα, να ακολουθούν την κανονική κατανομή  $N(0, \sigma^2)$  και να έχουν κοινή διακύμανση.



| Επέμβαση | Y  | $\varepsilon_{ij} = Y_{ij} - \bar{Y}_i$ | $\varepsilon_{ijstd} = \varepsilon_{ij}/s_i$ |
|----------|----|---|--|
| 1        | 5  | -3                                      | -1,15  |
| 1        | 8  | 0                                       | 0  |
| 1        | 9  | 1                                       | 0,38   |
| 1        | 7  | -1                                      | -0,38  |
| 1        | 11 | 3                                       | 1,15   |
| 2        | 18 | 2,4                                     | 0,92   |
| 2        | 14 | -1,6                                    | -0,61  |
| 2        | 16 | 0,4                                     | 0,15   |
| 2        | 13 | -2,6                                    | -1   |
| 2        | 17 | 1,4                                     | 0,54   |
| 3        | 6  | -6,4                                    | -2,45  |
| 3        | 12 | -0,4                                    | -0,15  |
| 3        | 14 | 1,6                                     | 0,61   |
| 3        | 17 | 4,6                                     | 1,76   |
| 3        | 13 | 0,6                                     | 0,23   |

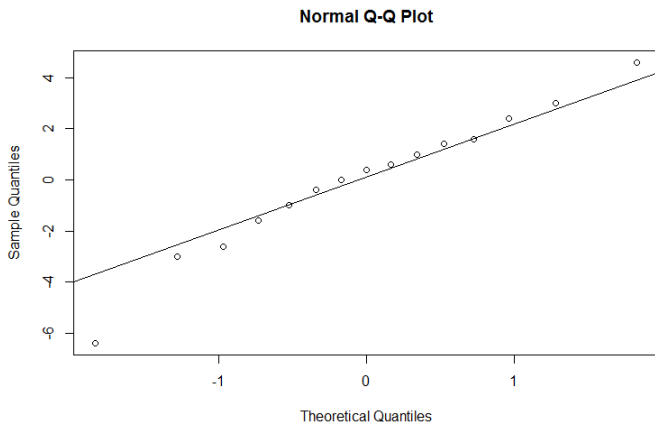
## Υπολογισμός υπολοίπων

```
> residuals=residuals(fit)
> round(residuals,2)
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
-3.0 0.0 1.0 -1.0 3.0 2.4 -1.6 0.4 -2.6 1.4 -6.4 -0.4 1.6 4.6 0.6
```

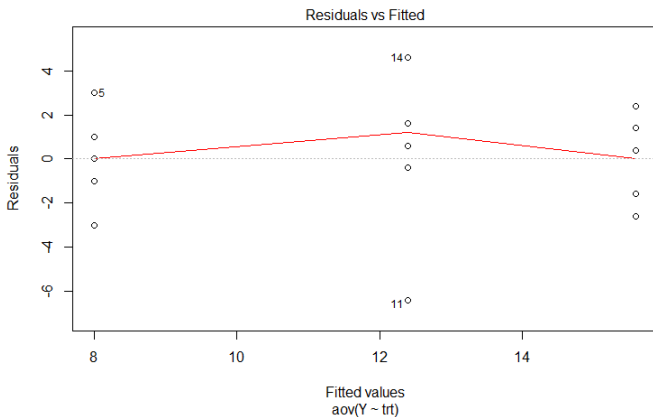
## Γραφικές Μέθοδοι Ελέγχου Κανονικότητας και Ομοσκεδαστικότητας

```
> qqnorm(residuals)
> qqline(residuals)
> plot(fit, which=1)
```

Ο έλεγχος κανονικότητας γίνεται με το **Q-Q διάγραμμα**, το οποίο συγκρίνει τα ποσοστιαία σημεία του δείγματος, με τα πληθυσμιακά ποσοστιαία σημεία της κανονικής κατανομής. Αν τα σημεία είναι κοντά στην ευθεία γραμμή δεν υπάρχει ένδειξη για απόκλιση από την κανονικότητα, ενώ παρεκκλίσεις δηλώνουν μη κανονικότητα.



Ο έλεγχος ότι τα υπόλοιπα των επεμβάσεων έχουν κοινή διακύμανση γίνεται με ένα **διάγραμμα διασποράς υπολοίπων** (residuals plot). Αν για όλα τα δείγματα – επεμβάσεις, τα υπόλοιπα είναι κατανεμημένα ομοιόμορφα και τυχαία γύρω από την οριζόντια ευθεία που διέρχεται από το 0, υπάρχει μια καλή ένδειξη ότι έχουν κοινή διακύμανση.



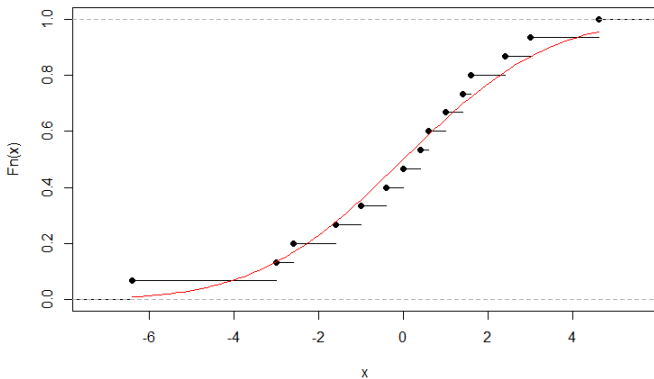
## Έλεγχος Kolmogorov-Smirnov (Lilliefors)

Ο έλεγχος Kolmogorov-Smirnov ελέγχει την απόσταση της εμπειρικής συνάρτησης κατανομής (Empirical Distribution Function) που προέρχεται από το δείγμα, με την αθροιστική συνάρτηση κατανομής. Βασική προϋπόθεση του έλεγχου είναι ότι η μηδενική κατανομή με την οποία ελέγχεται η κατανομή του δείγματος, να είναι πλήρως προσδιορισμένη και με γνωστές παραμέτρους.

Μία τροποποίηση του ελέγχου K-S είναι ο έλεγχος Lilliefors, που χρησιμοποιείται για να ελέγξει τα δεδομένα ενός δείγματος προέρχονται από έναν κανονικά κατανεμημένο πληθυσμό του οποίου οι παράμετροι είναι άγνωστοι και εκτιμούνται από το δείγμα. Η διαδικασία είναι η ίδια και στους δύο ελέγχους αλλά χρησιμοποιούν διαφορετικούς πίνακες κρίσιμων τιμών.

```
> plot(ecdf(residuals),main="Empirical and theoretical CDFs")  
> curve(pnorm(x,mean(residuals),sd(residuals)),-6.4, 4.6, +  
add=TRUE,col="red")
```

Empirical and theoretical CDFs



```
> ks.test(residuals, "pnorm", mean(residuals), sd(residuals))
```

One-sample Kolmogorov-Smirnov test

data: residuals

D = 0.10788, p-value = 0.987

alternative hypothesis: two-sided

```
> library(nortest)
```

```
> lillie.test(residuals)
```

Lilliefors (Kolmogorov-Smirnov) normality test

data: residuals

D = 0.10788, p-value = 0.9098

## Έλεγχος Anderson-Darling

Ο έλεγχος Anderson-Darling ανήκει επίσης στην οικογένεια των εμπειρικών συναρτήσεων κατανομών (EDF) και ελέγχει τη μηδενική υπόθεση ότι ένα δείγμα προήλθε από μία συγκεκριμένη κατανομή.

```
> ad.test(residuals)
Anderson-Darling normality test
data: residuals
A = 0.21292, p-value = 0.8195
```



## Έλεγχος Shapiro Wilk

Ο έλεγχος Shapiro-Wilk ελέγχει τη μηδενική υπόθεση ότι ένα δείγμα προήλθε από έναν κανονικά κατανομημένο πληθυσμό.

```
> shapiro.test(residuals)
Shapiro-Wilk normality test
data: residuals
W = 0.97088, p-value = 0.8709
```

## Έλεγχος Bartlett

Ο έλεγχος του Bartlett χρησιμοποιείται για να ελέγξει εάν τα  $k$  δείγματα προέρχονται από πληθυσμούς με ίσες διακυμάνσεις. Η δοκιμασία αυτή είναι ευαίσθητη στις αποκλίσεις από την κανονικότητα και την ανισότητα διασπορών.

```
> tapply(data$Y, data$trt, var)
1 2 3
5.0 4.3 16.3
> bartlett.test(Y~trt)
Bartlett test of homogeneity of variances
data: Y by trt
Bartlett's K-squared = 2.0618, df = 2, p-value = 0.3567
```

## Έλεγχος Levene

Ο έλεγχος Levene βασίζεται στην 'ανάλυση διακύμανσης' των παρατηρήσεων μετά την μετατροπή:  $Z_{ij} = |Y_{ij} - \bar{Y}_{i.}|$ . Αν αντί για τον μέσο  $\bar{Y}_i$  χρησιμοποιηθεί η διάμεσος  $\tilde{Y}$  ο έλεγχος ονομάζεται Brown-Forsythe.

```
> library(car)
```

```
> leveneTest(fit, center=mean)
```

```
Levene's Test for Homogeneity of Variance (center = mean)
```

|       | Df | F value | Pr(>F) |
|-------|----|---------|--------|
| group | 2  | 0.6085  | 0.5601 |
|       | 12 |         |        |

```
> leveneTest(fit, center=median) # Έλεγχος Brown-Forsythe
```

```
Levene's Test for Homogeneity of Variance (center = median)
```

|       | Df | F value | Pr(>F) |
|-------|----|---------|--------|
| group | 2  | 0.4386  | 0.6549 |
|       | 12 |         |        |