ΑΡΙΣΤΟΤΕΛΕΙΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΘΕΣΣΑΛΟΝΙΚΗΣ

ΓΕΩΠΟΝΙΚΗ ΣΧΟΛΗ
Α.Π.Θ.

Πλήρως Τυχαιοποιημένο Σχέδιο

Επιστημονική Επιμέλεια

Δρ. Γεώργιος Μενεξές

Τομέας Φυτών Μεγάλης Καλλιέργειας και Οικολογίας,
Εργαστήριο Γεωργίας

Πλήρως Τυχαιοποιημένο Σχέδιο (*Completely Randomized Design*-CRD)

Παράδειγμα 21 (Φασούλας, 2008, σ. 103).

Σε ένα πείραμα δοκιμάστηκαν 10 διαφορετικοί γενότυποι κριθαριού, για να φανεί κατά πόσο διαφέρουν ως προς την πρωϊμότητα ξεσταχιάσματος. Στο χωράφι σπάρθηκαν 10 φυτά από κάθε γενότυπο, οι δε αποστάσεις μεταξύ των φυτών ήταν αρκετές ώστε να αποκλείεται ανταγωνισμός μεταξύ των φυτών.

Να ελεγχθεί αν οι γενότυποι παρουσιάζουν στατιστικά σημαντικές διαφορές σε επίπεδο σημαντικότητας α=0,05.

Παραμετροποίηση -1

- **Πειραματικό Σχέδιο** (*Experimental Design*): Πλήρως Τυχαιοποιημένο Σχέδιο (…χωρίς Ομάδες)
- **Πλήθος Παραγόντων** (*Factors*): 1 (Γενότυπος)
- **Πλήθος Επιπέδων** (*Levels*) του Παράγοντα (π): 10
- **Αρ. Επαναλήψεων** (*Replications*) (n): 10
- **Συνολικό πλήθος μετρήσεων** (N): 100
- Σχέδιο: **Ισορροπημένο** (*Balanced*), δηλ. ίδιος αριθμός επαναλήψεων σε κάθε επέμβαση

Παραμετροποίηση -2

- **Εξαρτημένη** μεταβλητή (*Depended Variable*): Πρωϊμότητα ξεσταχιάσματος (ημέρες)
- **Ανεξάρτητη** μεταβλητή-Παράγοντας (*Independed Variable*) : Γενότυπος
- **Πρότυπο I** (*Model type I*) : Καθορισμένες Επιδράσεις (*Fixed Effects*)

Μεθοδολογία Εγκατάστασης Πειράματος

- Προηγούμενη εμπειρία και γνώση σχετικά με το πειραματικό υλικό
- Εμπειρία και γνώση σχετικά με προηγούμενα πειράματα στον ίδιο πειραματικό αγρό
- Έλεγχοι ομοιομορφίας και ομοιογένειας πειραματικού υλικού
- Διαστάσεις πειραματικών τεμαχίων
- Πλήθος φυτών
- Αποστάσεις
- Καλλιεργητική φροντίδα
- Περίοδος πειραματισμού
- Μέθοδος μέτρησης εξαρτημένης μεταβλητής
- Εγκυρότητα-Αξιοπιστία μετρήσεων
- Εδαφολογικά στοιχεία
- Κλιματολογικά στοιχεία
- Τήρηση Ημερολογίου Πειράματος

Πότε εφαρμόζεται το CRD

- Οι παράγοντες δεν επηρεάζονται από το περιβάλλον.
- Το περιβάλλον μπορεί να τεθεί υπό έλεγχο.
- Το πειραματικό υλικό είναι ομοιογενές.

Πίνακας Δεδομένων

Γενότυποι	Ε1	Ε2	Ε3	Ε4	Ε5	Ε6	Ε7	Ε8	Ε9	Ε10	Σύνολα
Α	5	6	6	5	2	3	2	2	5	5	**41**
Β	2	3	3	3	3	3	2	2	2	2	**25**
Γ	2	7	6	3	2	1	3	3	2	3	**32**
Δ	3	2	9	3	2	2	3	3	4	3	**34**
Ε	8	5	9	4	4	3	4	4	6	5	**52**
Ζ	1	8	8	4	3	2	4	2	3	1	**36**
Η	8	8	1	5	2	2	2	3	2	1	**34**
Θ	7	3	3	2	1	2	4	2	3	2	**29**
Ι	4	9	9	4	3	2	2	3	3	3	**42**
Κ	3	4	6	3	3	3	3	3	3	3	**34**
Σύνολα	**43**	**55**	**60**	**36**	**25**	**23**	**29**	**27**	**33**	**28**	**359**

Πίνακας Ανάλυσης Παραλλακτικότητας (ή Διακύμανσης)

Πηγή Παραλλακτικότητας	Βαθμοί Ελευθερίας	Άθροισμα Τετραγώνων	Μέσα Τετράγωνα	*F*
Γενότυποι (ή Παράγοντας)	π**-1**	**ΑΤΠ**	**ΜΤΠ=** $\frac{\text{ΑΤΠ}}{\pi-1}$	***F*=** $\frac{\text{ΜΤΠ}}{\text{ΜΤΣ}}$
Σφάλμα (ή Υπόλοιπο)	$\pi(n$**-1)**	**ΑΤΣ**	**ΜΤΣ=** $\frac{\text{ΑΤΣ}}{\pi(n-1)}$	
Ολική	πn**-1**	**ΣΑΤ**		

Η δειγματική τιμή *F* συγκρίνεται με την Κρίσιμη Τιμή (θεωρητική) της *F*-Κατανομής με (π-1) και [π(*n*-1)] β.ε., σε επίπεδο σημαντικότητας α.

One Way ANOVA Table (I)

Source of Variation	Degrees of Freedom (df)	Sum of Squares (SS)	Mean Squares (MS)	*F*
Treatments or Factor A (Genotype)	π-1	**Treatment SS (SSA)**	**MSA=** $\frac{Treatment\ SS}{\pi - 1}$	***F*=** $\frac{\text{MSA}}{\text{MSE}}$
Error	$\pi(n$-1)	**Error SS (SSE)**	**MSE=** $\frac{Error\ SS}{\pi(n-1)}$	
Total	πn-1	**Total SS (SST)**		

One Way ANOVA Table (II)

Source of Variation
Between Treatments
Within Treatments
Total

Πίνακας Δεδομένων

Γενότυποι	Ε1	Ε2	Ε3	Ε4	Ε5	Ε6	Ε7	Ε8	Ε9	Ε10	Σύνολα
Α	5	6	6	5	2	3	2	2	5	5	**41**
Β	2	3	3	3	3	3	2	2	2	2	**25**
Γ	2	7	6	3	2	1	3	3	2	3	**32**
Δ	3	2	9	3	2	2	3	3	4	3	**34**
Ε	8	5	9	4	4	3	4	4	6	5	**52**
Ζ	1	8	8	4	3	2	4	2	3	1	**36**
Η	8	8	1	5	2	2	2	3	2	1	**34**
Θ	7	3	3	2	1	2	4	2	3	2	**29**
Ι	4	9	9	4	3	2	2	3	3	3	**42**
Κ	3	4	6	3	3	3	3	3	3	3	**34**
Σύνολα	**43**	**55**	**60**	**36**	**25**	**23**	**29**	**27**	**33**	**28**	**359**

Υπολογισμοί (Ι)

Διορθωτικός Όρος (*Correction Term*), $\Delta O=\frac{359^2}{100}=1.288,81$

Συνολικό Άθροισμα Τετραγώνων, $\Sigma AT=(5^2+6^2+6^2+\ldots+3^2+3^2+3^2)-\Delta O=400,19$

Άθροισμα Τετραγώνων Παραγόντων, $ATΠ=(\frac{41^2+25^2+\ldots+34^2}{10})-\Delta O=51,49$

Άθροισμα Τετραγώνων Σφαλμάτων, $AT\Sigma=\Sigma AT-ATΠ=400,19-51,49=348,70$

Υπολογισμοί (II)

$$\text{ΑΤΠ} = SSA = SS_{Treatments} = \frac{1}{n}\sum_{i=1}^{\pi}\left(\sum_{j=1}^{n}\left(Y_{ij}\right)\right)^2 - \frac{\left(Y_{\bullet\bullet}\right)^2}{\pi \cdot n}$$

$$\Sigma\text{ΑΤ} = SST = SS_T = \sum_{i=1}^{\pi}\sum_{j=1}^{n}\left(Y_{ij}\right)^2 - \frac{\left(Y_{\bullet\bullet}\right)^2}{\pi \cdot n}$$

Πίνακας Ανάλυσης Παραλλακτικότητας (ή Διακύμανσης)

Πηγή Παραλλακτικότητας	Βαθμοί Ελευθερίας	Άθροισμα Τετραγώνων	Μέσα Τετράγωνα	F	$F_{0,05}$
Γενότυποι (Παράγοντας)	9	51,49	5,72	1,48	1,99
Σφάλμα	90	348,70	3,87		
Ολική	99	400,19			

Κρίσιμη Τιμή F(9, 90)=1,99, σε επίπεδο σημαντικότητας α=0,05

Επειδή 1,48<1,99 $\Rightarrow$ Οι γενότυποι δεν παρουσιάζουν στατιστικά σημαντικές διαφορές σε επίπεδο σημαντικότητας α=0,05.

Συντελεστής Παραλλακτικότητας (Coefficient of Variation), *CV*

$$CV = \frac{\sqrt{\text{ΜΤΣ}}}{\bar{Y}..} \times 100 = \frac{\sqrt{MSE}}{\bar{Y}..} \times 100$$

$$\text{Στο παράδειγμα, CV} = \frac{1,97}{3,59} \times 100 = 54,87\%$$

- **Εκφράζει το βαθμό ακρίβειας με την οποία συγκρίνονται οι επεμβάσεις.**
- **Εκφράζει την αξιοπιστία του πειράματος.**
- **Αποδεκτές τιμές: Ανάλογα με τη φύση του πειραματικού υλικού και το τι μετρούμε.**

<u>Παράδειγμα:</u>

Για το ρύζι όταν μετρούμε την απόδοση.
6%-8% για τη σύγκριση ποικιλιών
10%-12% για τη σύγκριση λιπασμάτων
Για την απόδοση 10%, για το ύψος 3%

Το Γενικό Γραμμικό Πρότυπο (*General Linear Model*)

$$Y_{ij} = \mu + t_i + e_{ij}$$

i: αναφέρεται στην κύρια επίδραση της επέμβασης i (i=1,…,π)
j: αναφέρεται στην μέτρηση ή παρατήρηση j (j=1,…,n)

Παράμετρος Πληθυσμού	Εκτιμητής
μ	$\bar{Y}_{..}$
t_i	$\bar{Y}_{i.} - \bar{Y}_{..}$
$\mu + t_i$	$\bar{Y}_{i.}$
σ_e^2	$s_e^2 = \text{ΜΤΣ (MSE)}$

Παρατηρούμενα (Δειγματικά) και Αναμενόμενα (Θεωρητικά) Μέσα Τετράγωνα (Expected Mean Squares)

Παρατηρούμενα ΜΤ	Αναμενόμενα
ΜΤΠ (MSA)	$\sigma_e^2 + n\sigma_a^2 = \sigma_e^2 + n\frac{\sum_{i=1}^{\pi} t_i^2}{\pi - 1}$
ΜΤΣ (MSE)	σ_e^2

Εκτιμητές:

$$\hat{\sigma}_a^2 = \frac{\sum_{i=1}^{\pi} (\bar{Y}_{i.} - \bar{Y}..)^2}{\pi - 1}$$

$$\hat{\sigma}_e^2 = \text{ΜΤΣ} = MSE$$

Παραδοχές και Προϋποθέσεις

Παραδοχές:

$$\sum_{i=1}^{\pi} t_i = 0$$

$$e_{ij} \sim N(0, \sigma_e^2)$$

Προϋποθέσεις:

Οι παρατηρήσεις προέρχονται από τυχαία δείγματα
Οι παρατηρήσεις είναι ανεξάρτητες η μία από την άλλη
Οι πληθυσμοί των παρατηρήσεων ακολουθούν Κανονική Κατανομή
Οι διασπορές των πληθυσμών είναι ίσες (Ομοσκεδαστικότητα)

Στατιστικοί Έλεγχοι

Μηδενικές Υποθέσεις

$$\mathrm{H}_0 : t_i = 0$$
$$\mathrm{H}_0 : \sigma_a^2 = 0$$
$$\mathrm{H}_0 : E(Y_{ij}) = 0$$
$$\mathrm{H}_0 : t_1 = t_2 = \cdots = t_\pi$$
$$\mathrm{H}_0 : \mu_1 = \mu_2 = \cdots = \mu_\pi$$

Εναλλακτικές Υποθέσεις

$$\mathrm{H}_1 \, \acute{\eta} \, \mathrm{H}_\alpha : \exists \, k \; (k = 1, \ldots \pi) : t_k \neq 0$$
$$\mathrm{H}_1 \, \acute{\eta} \, \mathrm{H}_\alpha : \sigma_a^2 > 0$$
$$\mathrm{H}_1 \, \acute{\eta} \, \mathrm{H}_\alpha : E(Y_{ij}) \neq 0$$
$$\mathrm{H}_1 \, \acute{\eta} \, \mathrm{H}_\alpha : \textit{μία τουλάχιστον κύρια επίδραση } t_k \neq 0, k = 1, \ldots, \pi$$
$$\mathrm{H}_1 \, \acute{\eta} \, \mathrm{H}_\alpha : \textit{τουλάχιστον 2 μέσοι όροι διαφέρουν}, \exists \, k, z, (k, z = 1, \ldots, \pi) : \mu_k \neq \mu_z$$

Άλλες Στατιστικές Αναλύσεις

- Αν η ANOVA ανιχνεύσει στατιστικά σημαντικές διαφορές ακολουθούν συγκρίσεις μέσων όρων

Σχέδιο: Μη Ισορροπημένο (Unbalanced Design): Πίνακας Ανάλυσης Παραλλακτικότητας (ή Διακύμανσης), Δείγματα Άνισα

Πηγή Παραλλακτικότητας	**Βαθμοί Ελευθερίας**	**Άθροισμα Τετραγώνων**	**Μέσα Τετράγωνα**	***F***
Γενότυποι (ή Παράγοντας)	**π-1**	**ΑΤΠ**	**ΜΤΠ=** $\frac{\text{ΑΤΠ}}{\pi - 1}$	***F*=** $\frac{\text{ΜΤΠ}}{\text{ΜΤΣ}}$
Σφάλμα (ή Υπόλοιπο)	***N*-π**	**ΑΤΣ**	**ΜΤΣ=** $\frac{\text{ΑΤΣ}}{N - \pi}$	
Ολική	***N*-1**	**ΣΑΤ**		

Υπολογισμοί

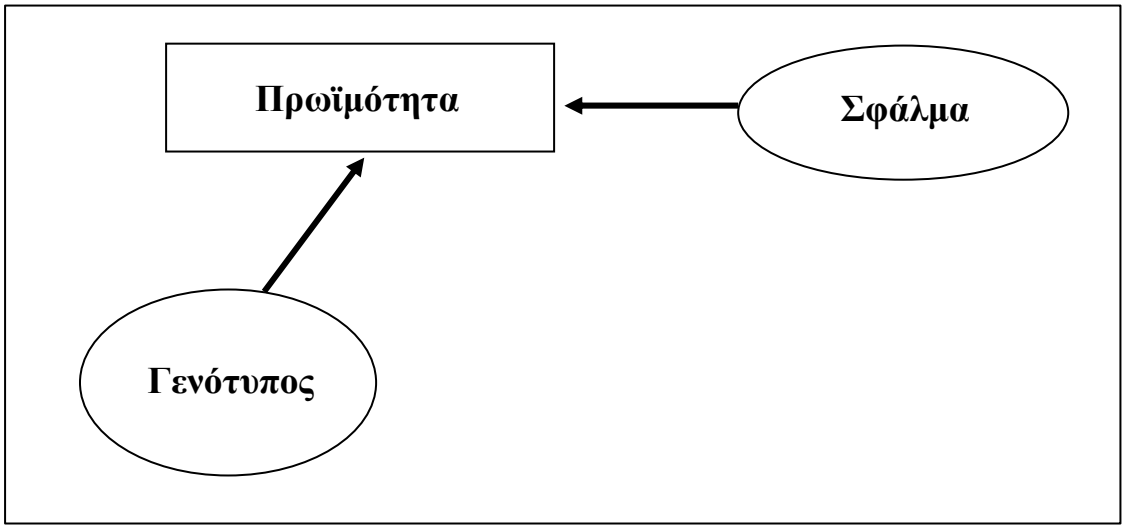
Οι υπολογισμοί γίνονται όπως και στο Ισορροπημένο Σχέδιο με τη διαφορά:

Άθροισμα Τετραγώνων Παραγόντων, ΑΤΠ=$\left(\frac{\Pi_1^2}{n_1} + \frac{\Pi_2^2}{n_2} + \cdots + \frac{\Pi_\pi^2}{n_\pi} \right) - \Delta O$,

Η δειγματική τιμή F συγκρίνεται με την Κρίσιμη Τιμή (θεωρητική) της F-Κατανομής με **(π-1) και (N-π) β.ε., σε επίπεδο σημαντικότητας α.**

όπου n_i είναι το πλήθος δειγμάτων (επαναλήψεων) στην i επέμβαση και Π_i είναι το άθροισμα των παρατηρήσεων της i επέμβασης (N= Συνολικό πλήθος μετρήσεων)

Διαγραμματική Αναπαράσταση του Υποδείγματος



Πρότυπο II (Model type II): Πρότυπο Τυχαίων Επιδράσεων – (*Random Effects Model*)

Τα στοιχεία του πίνακα της ANOVA υπολογίζονται κατά τον ίδιο τρόπο όπως και στο Πρότυπο I.

Διαφορές:

- $\mathrm{H}_0 : \sigma_a^2 = 0$

  $\mathrm{H}_1 \; \acute{\eta} \; \mathrm{H}_\alpha : \sigma_a^2 > 0$

- Δεν έχει νόημα η σύγκριση μέσων όρων
- Μελετώνται οι συνιστώσες της παραλλακτικότητας (*Variance Components*)

$$\hat{\sigma}^2 = \hat{\sigma}_e^2 + \hat{\sigma}_a^2 \qquad \hat{\sigma}_a^2 = \frac{\mathrm{ΜΤΠ} - \mathrm{ΜΤΣ}}{n} = \frac{MSA - MSE}{n}$$

Στο παράδειγμα

Εκφράζει την παραλλακτικότητα μεταξύ των γενοτύπων

$$\hat{\sigma}_a^2 = \frac{MSA - MSE}{n} \Rightarrow \hat{\sigma}_a^2 = \frac{5,721 - 3,874}{10} = \frac{1,847}{10} = 0,185$$

$$\hat{\sigma}^2 = \hat{\sigma}_e^2 + \hat{\sigma}_a^2 \Rightarrow \hat{\sigma}^2 = 3,874 + 0,185 = 4,06$$

Εκφράζει το 95,4% της Ολικής Παραλλακτικότητας

Εκφράζει το 4,6% της Ολικής Παραλλακτικότητας

Άλλες Εφαρμογές των Συνιστωσών Παραλλακτικότητας

$$Intraclass\ Correlation: \rho_I = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_e^2}$$

- Εκφράζει το βαθμό ομοιότητας των παρατηρήσεων μέσα στις ομάδες σε σχέση με την ομοιότητα των παρατηρήσεων μεταξύ των ομάδων.
- Στη Γενετική εκφράζει την κληρονομικότητα ποσοτικών χαρακτηριστικών
- Εκφράζει την αξιοπιστία, εσωτερική συνοχή ή συνέπεια, μεταξύ παρατηρήσεων που προέκυψαν από την ίδια διαδικασία
- Εκφράζει την επαναληψιμότητα διαδοχικών μετρήσεων στην ίδια πειραματική μονάδα

Συγκρίσεις Μέσων Όρων

Το Κριτήριο της Ελάχιστης (Στατιστικά) Σημαντικής Διαφοράς (ΕΣΔ-LSD)

Ισορροπημένο Σχέδιο:

$$\text{ΕΣΔ} = t_{\pi(n-1);a/2}\sqrt{\frac{2\times \text{ΜΤΣ}}{n}} = t_{\pi(n-1);a/2}\sqrt{\frac{2\times MSE}{n}}$$

Όπου $t_{\pi(n-1);a}$: Κρίσιμη τιμή της *t*-Κατανομής με $\pi(n-1)$ β.ε., σε επίπεδο σημαντικότητας $\alpha/2$

Μη Ισορροπημένο Σχέδιο:

$$\text{ΕΣΔ} = t_{(N-\pi);a/2}\sqrt{\frac{(n_i+n_j)\times \text{ΜΤΣ}}{n_i\times n_j}} = t_{(N-\pi);a/2}\sqrt{\frac{(n_i+n_j)\times MSE}{n_i\times n_j}}$$

Όπου $t_{N-\pi;a/2}$: Κρίσιμη τιμή της *t*-Κατανομής με ($N-\pi$) β.ε., σε επίπεδο σημαντικότητας $\alpha/2$

Στο Παράδειγμα

$$\text{ΕΣΔ}=1,99\times\sqrt{\frac{2\times 3,89}{10}}=1,99\times\sqrt{\frac{7,78}{10}}=1,99\times\sqrt{0,778}=1,99\times 0,882=1,75$$, σε επίπεδο σημαντικότητας α=0,05.

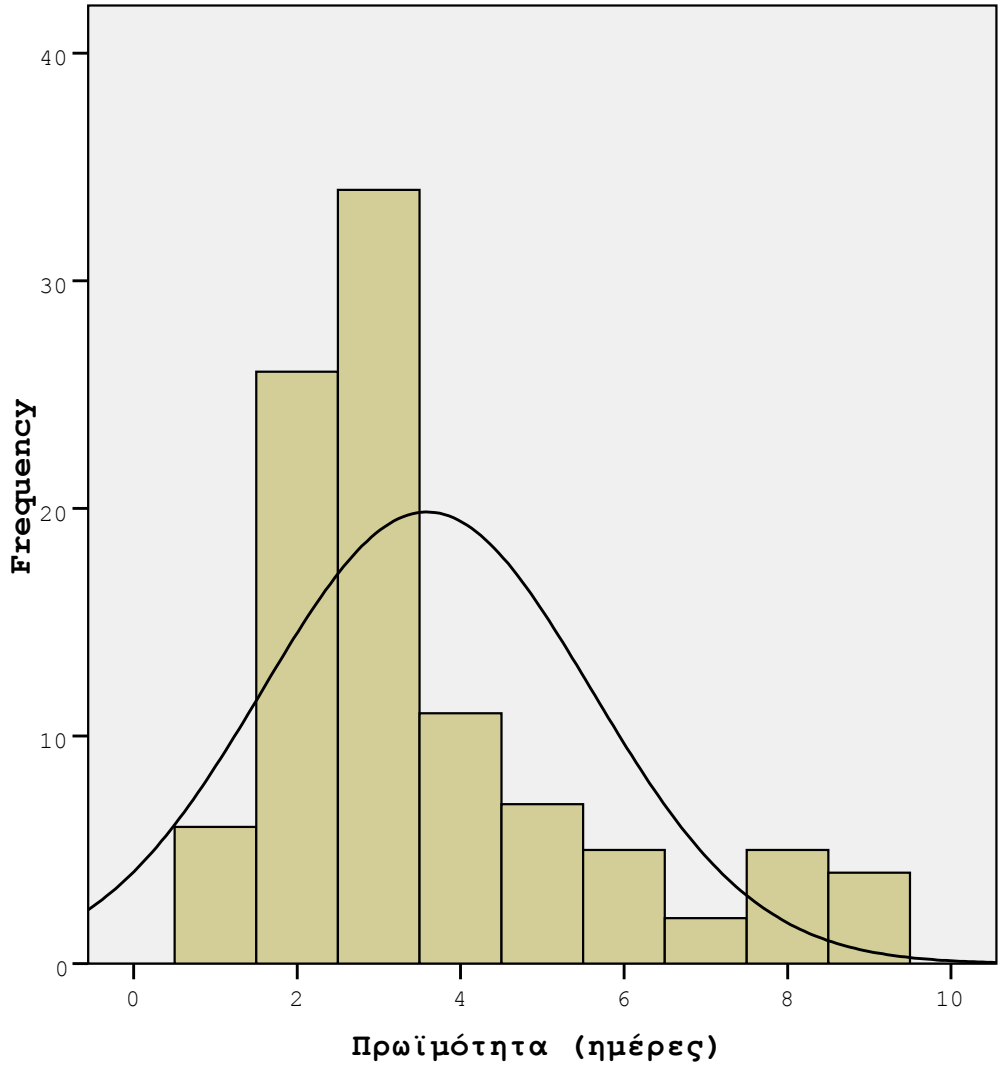
Στο Παράδειγμα

Γενότυποι	ΜΟ	ΤΑ	*N*
Α	**4,10**	**1,66**	**10**
Β	**2,50**	**0,53**	**10**
Γ	**3,20**	**1,87**	**10**
Δ	**3,40**	**2,07**	**10**
Ε	**5,20**	**1,93**	**10**
Ζ	**3,60**	**2,55**	**10**
Η	**3,40**	**2,67**	**10**
Θ	**2,90**	**1,66**	**10**
Ι	**4,20**	**2,62**	**10**
Κ	**3,40**	**0,97**	**10**
$\mathbf{E\Sigma\Delta_{0,10}}$	**1,46**		
$\mathbf{E\Sigma\Delta_{0,05}}$	**1,75**		
$\mathbf{E\Sigma\Delta_{0,01}}$	**2,32**		
$\mathbf{E\Sigma\Delta_{0,0011}}$	**2,97**		

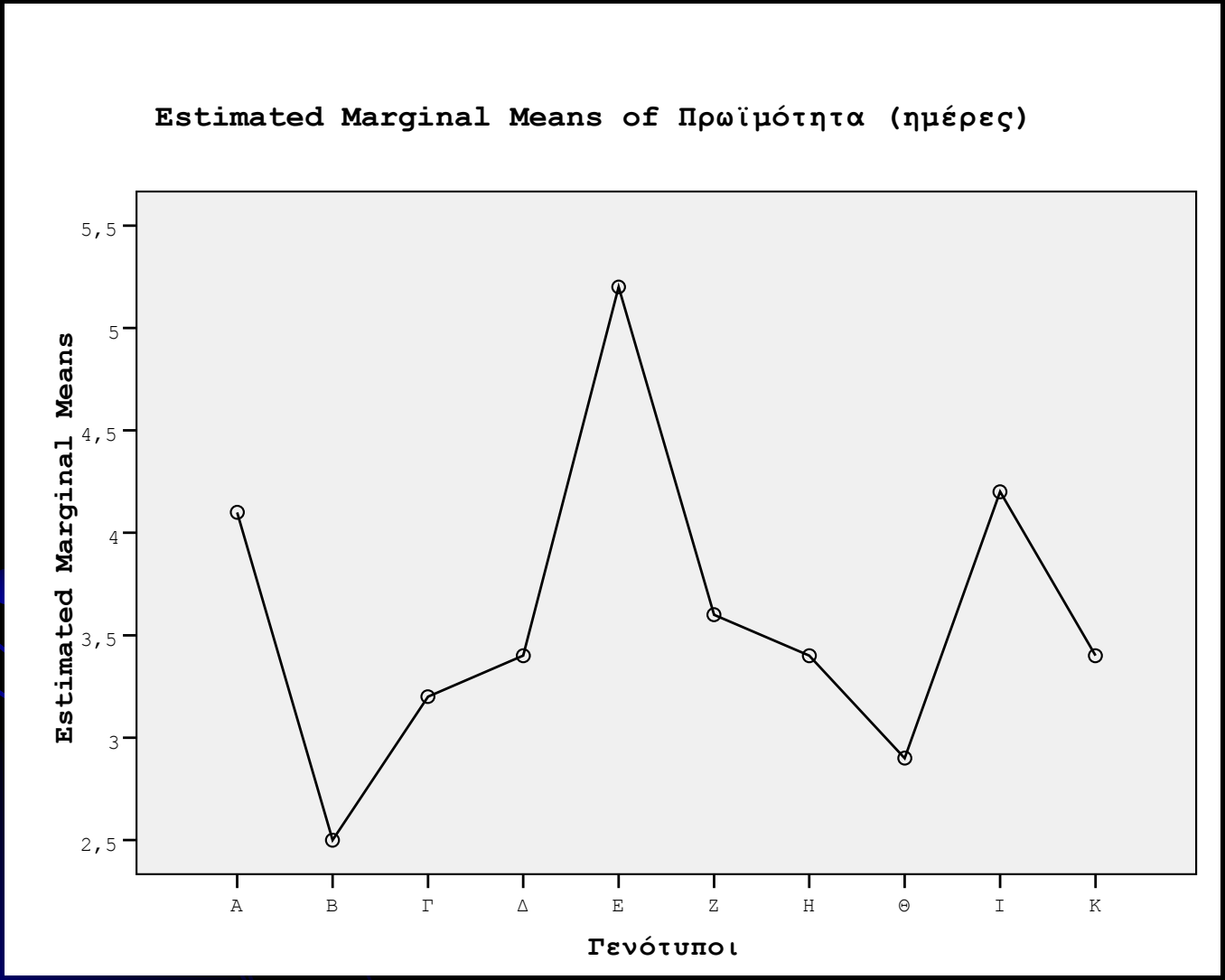
Αποτελέσματα με το SPSS (1)

		Πρωϊμότητα (ημέρες)
N	Statistic	100
Range	Statistic	8
Minimum	Statistic	1
Maximum	Statistic	9
Mean	Statistic	3.59
	Std. Error	.201
Std. Deviation	Statistic	2.011
Variance	Statistic	4.042
Skewness	Statistic	1.289
	Std. Error	.241
Kurtosis	Statistic	.995
	Std. Error	.478

Αποτελέσματα με το SPSS (2)



Αποτελέσματα με το SPSS (3)



Αποτελέσματα με το SPSS (4)

Tests of Between-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Observed Power(a)
Corrected Model	51.49(b)	9	5.72	1.48	.169	.667
Intercept	1288.81	1	1288.81	332.64	.000	1.000
Genotype	51.49	9	5.72	1.48	.169	.667
Error	348.70	90	3.87			
Total	1689.00	100				
Corrected Total	400.19	99				

R^2=0.129 (Συντελεστής Προσδιορισμού)

Αποτελέσματα με το SPSS (5)

- Έλεγχος Ομοιογένειας της Παραλλακτικότητας (π.χ. Έλεγχος του Levene)

F	df1	df2	Sig.
1.981	9	90	.051

Αποτελέσματα με το SPSS (5)

	Γενότυποι	Subset	
		2	1
Tukey HSD(a,b)	Β	2.50	
	Θ	2.90	
	Γ	3.20	
	Η	3.40	
	Δ	3.40	
	Κ	3.40	
	Ζ	3.60	
	Α	4.10	
	Ι	4.20	
	Ε	5.20	
	Sig.	.080	
Duncan(a,b)	Β	2.50	
	Θ	2.90	
	Γ	3.20	3.20
	Η	3.40	3.40
	Δ	3.40	3.40
	Κ	3.40	3.40
	Ζ	3.60	3.60
	Α	4.10	4.10
	Ι	4.20	4.20
	Ε		5.20
	Sig.	.107	.054

Παρουσίαση των Αποτελεσμάτων 1

Η ANOVA έδειξε ότι δεν υπάρχουν στατιστικά σημαντικές διαφορές, σε επίπεδο σημαντικότητας α=0,05, μεταξύ των 10 Γενοτύπων:

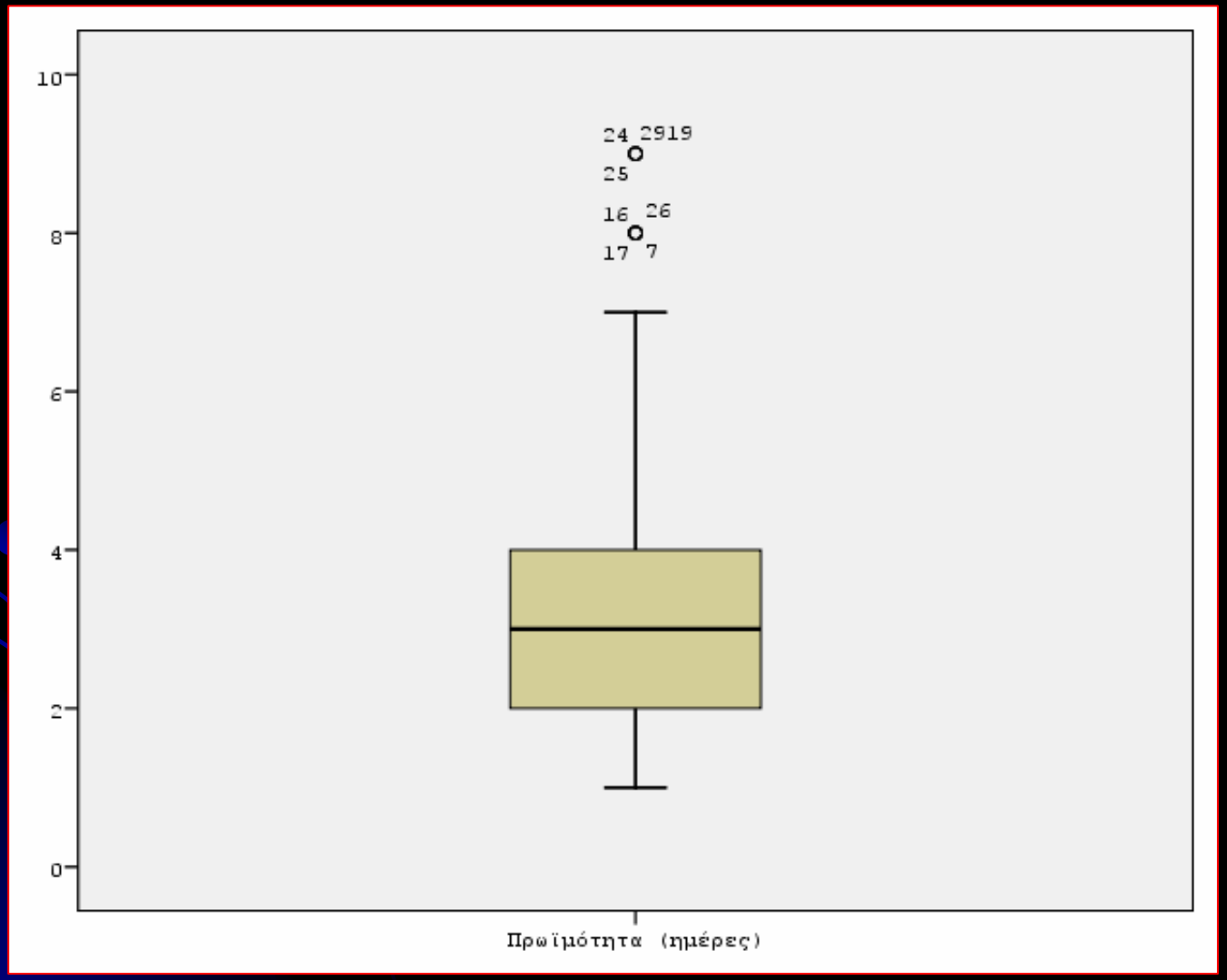
$(F(9,90)=1,48,\ p=0,169>0,05)$

Παρουσίαση των Αποτελεσμάτων 2

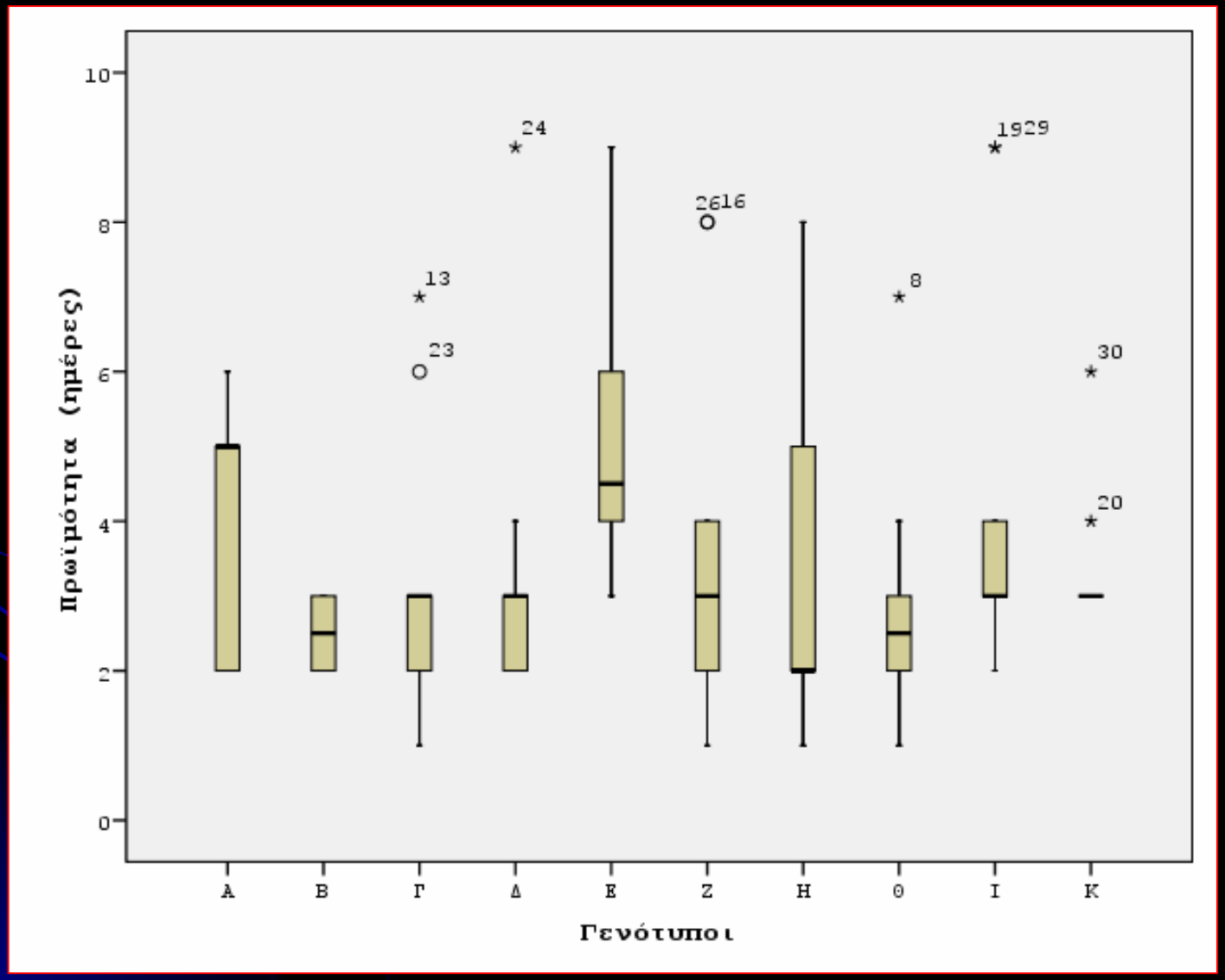
Γενότυποι	ΜΟ	ΤΑ	*N*
Α	4,10 ab	1,66	10
Β	2,50 b	0,53	10
Γ	3,20 ab	1,87	10
Δ	3,40 ab	2,07	10
Ε	5,20 a	1,93	10
Ζ	3,60 ab	2,55	10
Η	3,40 ab	2,67	10
Θ	2,90 b	1,66	10
Ι	4,20 ab	2,62	10
Κ	3,40 ab	0,97	10

Μέσοι όροι που ακολουθούνται από διαφορετικό γράμμα διαφέρουν στατιστικά σημαντικά, σε επίπεδο σημαντικότητας α=0,05, σύμφωνα με τα αποτελέσματα του ελέγχου Duncan

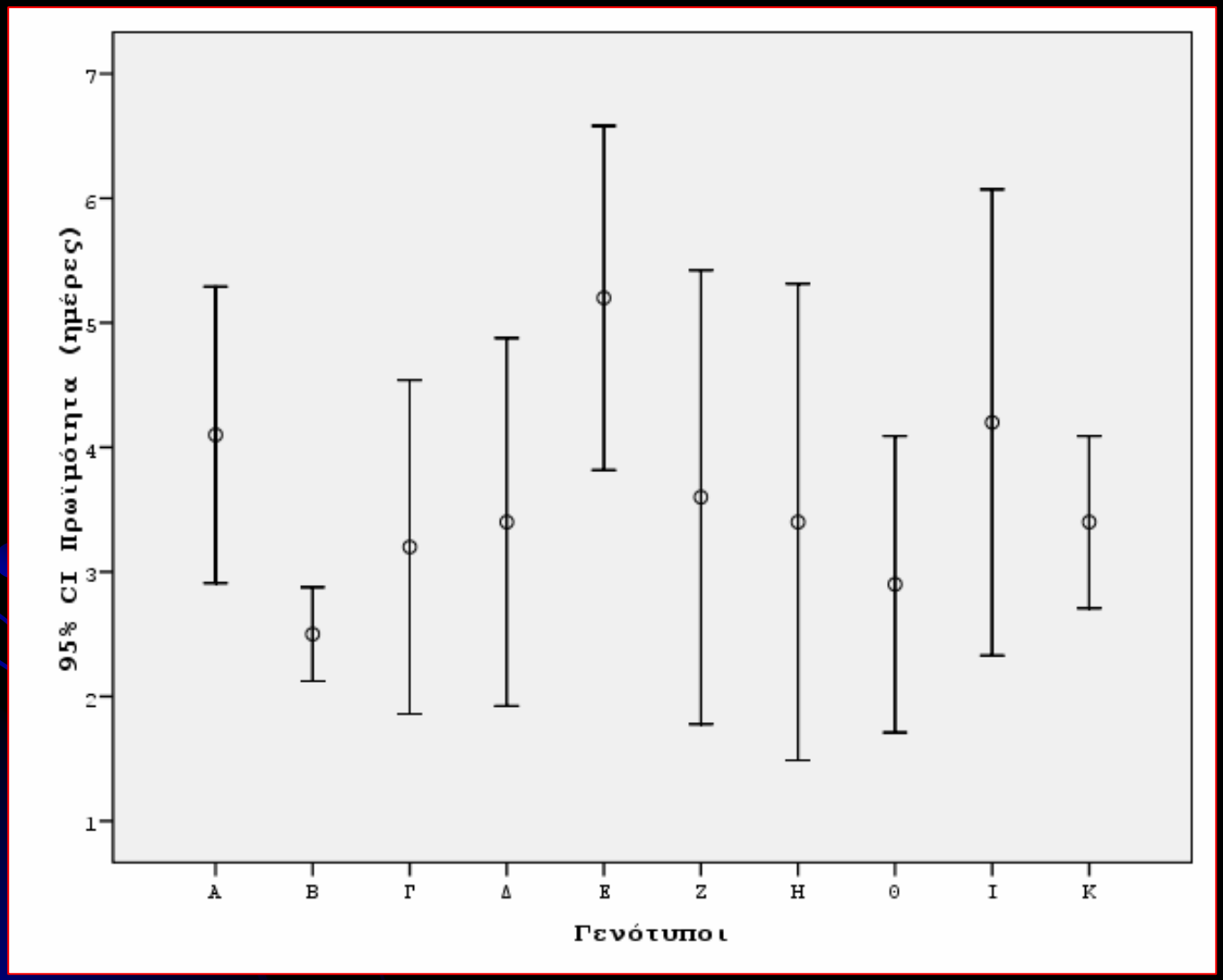
Άλλες Μορφές Διαγραμμάτων:
Box Plot



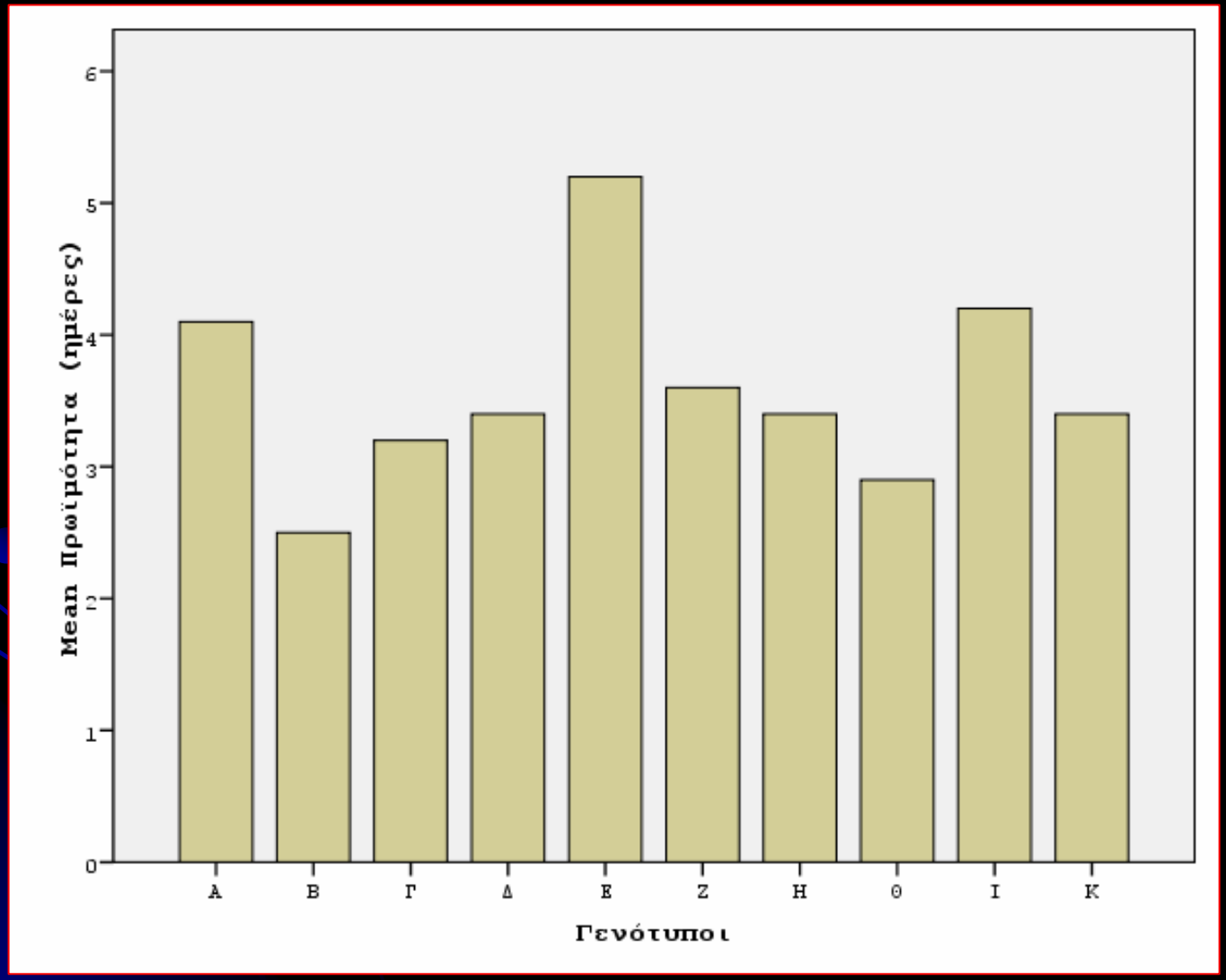
Συγκριτικό Διάγραμμα *Box Plot*



Διάγραμμα *Error bar*



Ραβδόγραμμα



Παραδείγματα (1)

7. [10] Η απόδοση σε *γάλα* (Kg/24h) μιας προβατίνας που έχει γεννήσει, υπολογίζεται ζυγίζοντας το νεογνό πριν και μετά τον θηλασμό. Πήραμε δείγματα από τρεις διαφορετικούς πληθυσμούς (φυλές) προβάτων και τα αποτελέσματα ήταν τα εξής:

Φυλές							
A_1	2.4	2.7	1.8	3.2	3.4	2.6	
A_2	3.2	3.4	4.1	2.8	2.9		
A_3	3.9	4.2	3.6	2.8	3.4	3.7	3.5

Υπάρχει διαφορά μεταξύ των τριών φυλών ως προς την απόδοση γάλακτος σε επίπεδο σημαντικότητας 5% ; Να διατυπώσετε με ακρίβεια τις υποθέσεις που ελέγχετε και να εξηγήσετε με σαφήνεια που στηρίζετε το συμπέρασμά σας.
(Δίνονται: SSA= 2.6, SST=6.5).

Παραδείγματα (2)

7. [10] Μετρήθηκε η ποσότητα πρωτεΐνης (gr/100ml) στο αίμα ατόμων που ζουν σε διαφορετικές συνθήκες στις γεωγραφικές περιοχές Α, Β, Γ και είχαμε τα παρακάτω αποτελέσματα:

Περιοχές									
A_1	7.64	7.04	7.43	7.57	7.74	7.63	8.06		
A_2	7.67	7.58	7.04	7.69	7.32	7.12	7.46	7.21	
A_3	7.98	7.91	7.11	7.65	8.17	8.28	7.21	7.41	6.37

Διατυπώστε κατάλληλο έλεγχο υποθέσεων και ελέγξτε σε επίπεδο σημαντικότητας 5% αν η ποσότητα πρωτεΐνης στο αίμα είναι η ίδια και στις τρεις περιοχές.
(Δίνονται: SSA=0.53, SST=4.77).

Παραδείγματα (3)

Παράδειγμα: Σε μία μελέτη, έχει καταγραφεί το βάρος του ήπατος (εκφρασμένο ως ποσοστό του βάρους του σώματος) ποντικιών που ανήκουν σε k=4 ομάδες που τράφηκαν με 4 δίαιτες. Θέλουμε να ερευνήσουμε αν υπάρχουν συστηματικές διαφορές μεταξύ των 4 ομάδων, δηλ., να θέλουμε να συγκρίνουμε τις μέσες τιμές των 4 ομάδων: $\mu_a, \mu_b, \mu_c, \mu_d$.

Παραδείγματα (4)

Παράδειγμα

Ας υποθέσουμε ότι ένας ερευνητής ενδιαφέρεται να μελετήσει την επίδραση που έχει το **γονεϊκό μοντέλο** διαπαιδαγώγησης στην **αυτοεκτίμηση του παιδιού**. Αφού κατατάξει τα άτομα που συμμετέχουν στην έρευνα σε τρεις διαφορετικές ομάδες ανάλογα με το μοντέλο με το οποίο διαπαιδαγωγήθηκαν (αυταρχικό, ενδιάμεσο, φιλελεύθερο), στη συνέχεια, τους ζητά να συμπληρώσουν ένα ερωτηματολόγιο αυτοεκτίμησης (Rosenberg Self-Esteem Scale)

- Τα αποτελέσματα από την έρευνα παρουσιάζονται στην επόμενη διαφάνεια.

Τα δεδομένα της Έρευνας

Αυταρχικό	Ενδιάμεσο	Φιλελεύθερο
17	10	3
14	11	9
19	8	2
16	12	5
18	9	1
20	11	7
23	8	6
21	12	
18	9	
	10	

Παραδείγματα (5)

Στον πίνακα 8.3 δίνεται μέρος από τα δεδομένα ενός πειράματος, που απέβλεπε στη μελέτη της σχέσεως μεταξύ της αποδόσεως των καπνοφύτων και του ρυθμού αναγωγής του αζώτου που προσλαμβάνουν τα φυτά από το έδαφος. Για λόγους απλότητας, τα ονόματα των ποικιλιών, που ήταν συνδυασμοί γραμμάτων και αριθμών, αντικαταστάθηκαν μ' ένα γράμμα το καθένα. Το πειραματικό τεμάχιο στο πείραμα αυτό αποτελούνταν από μία γλάστρα που περιείχε ένα φυτό. Οι γλάστρες ήταν γεμάτες με το ίδιο χώμα και είχαν τοποθετηθεί όλες μαζί στο κέντρο του θερμοκηπίου. Το πειραματικό σχέδιο ήταν το τελείως τυχαιοποιημένο, όπως έπρεπε στην περίπτωση αυτή.

ΠΙΝΑΚΑΣ 8.3

Παραγωγή ξηράς ουσίας (g/φυτό) πέντε ποικιλιών καπνού μέσα στο θερμοκήπιο σε γλάστρες

Ποικιλίες	Επαναλήψεις						Σύνολο	Μέσος όρος
	1	2	3	4	5	6		
Α	52,05	54,13	52,84	48,30	53,95	55,79	317,06	52,84
Β	49,19	41,14	52,52	37,16	52,63	53,25	285,89	47,65
Γ	54,82	46,50	42,09	37,39	55,01	42,83	278,64	46,44
Δ	31,15	32,63	33,53	34,31	41,16	36,86	209,64	34,94
Ε	32,96	32,77	32,91	30,42	27,49	32,42	188,97	31,50
							1280,20	42,67

Παραδείγματα (6)

Άσκηση 1. Σε μια έρευνα που έγινε στο πανεπιστήμιο Μελβούρνης επελέγη τυχαία ένα δείγμα αντρών και γυναικών διαφόρων ηλικιών το οποίο και υποβλήθηκε σε ένα τεστ αντοχής στον (σωματικό) πόνο. Στον ακόλουθο πίνακα δίνεται για κάθε ένα άτομο του δείγματος ο δείκτης αντοχής στον πόνο ο οποίος εξήχθη με βάση το τεστ αυτό (μεγαλύτερος δείκτης σημαίνει μεγαλύτερη αντοχή). Σε κάθε άτομο του δείγματος καταγράφεται επίσης και το (φυσικό) χρώμα των μαλλιών: ανοιχτό ξανθό (1), σκούρο ξανθό (2), ανοιχτό μελαχρινό (3), σκούρο μελαχρινό (4).

Έχει ενδιαφέρον να εξετάσουμε αν υπάρχουν διαφορές στην μέσο δείκτη αντοχής στον πόνο μεταξύ των ατόμων με διαφορετικό χρώμα μαλλιού.

Χρώμα μαλλιών	*Δείκτης αντοχής*
1	62
1	60
1	71
1	55
1	48

Χρώμα μαλλιών	*Δείκτης αντοχής*
2	63
2	57
2	52
2	41
2	43

Χρώμα μαλλιών	*Δείκτης αντοχής*
3	42
3	50
3	41
3	37

Χρώμα μαλλιών	*Δείκτης αντοχής*
4	32
4	39
4	51
4	30
4	35

McClave, J. T., and Dietrich II, F. H. (1991). *Statistics*. Dellen Publishing, San Francisco, Exercise 10.20.

Αποτελέσματα (1)

SCORE

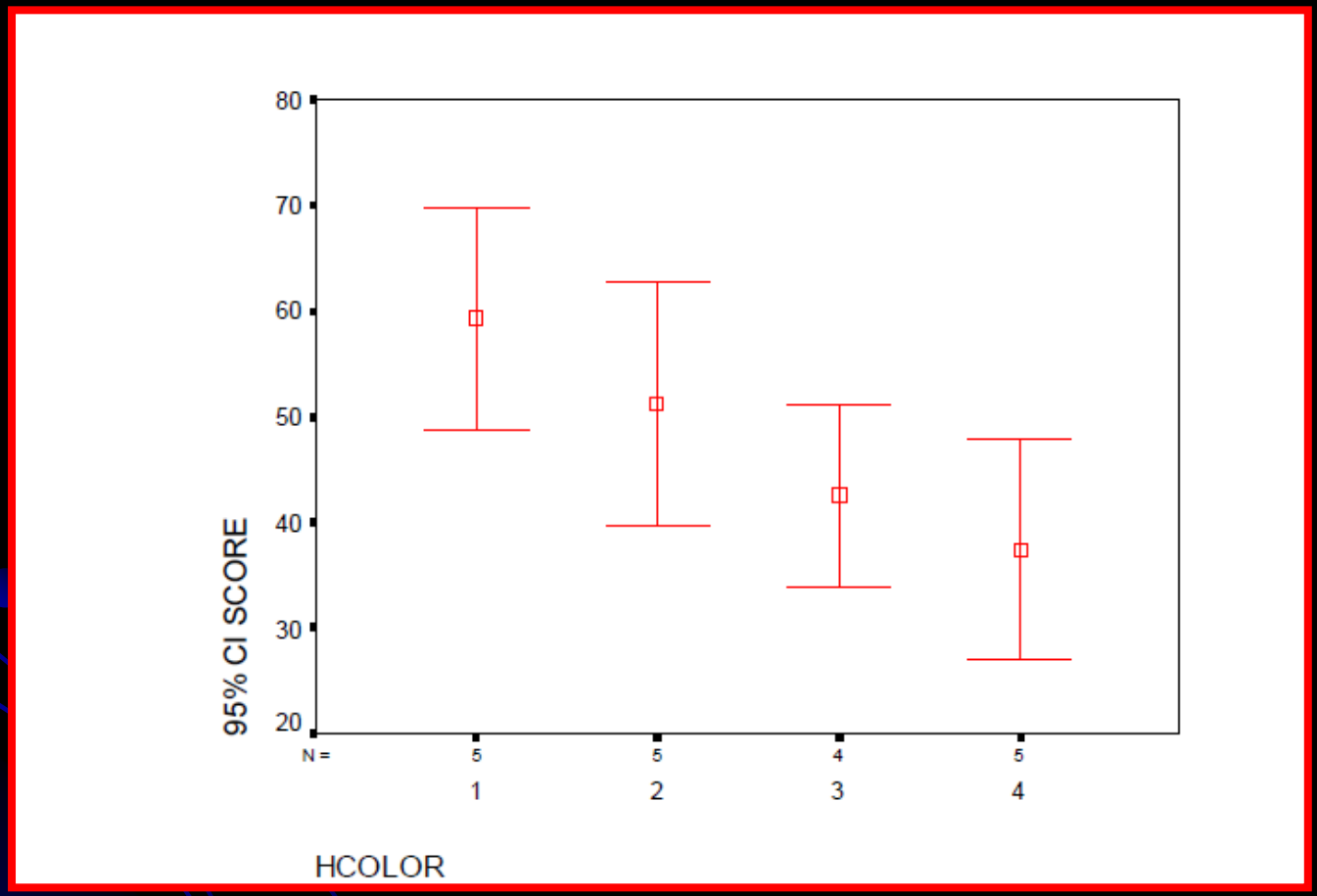
	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
1	5	59,20	8,53	3,81	48,61	69,79	48	71
2	5	51,20	9,28	4,15	39,67	62,73	41	63
3	4	42,50	5,45	2,72	33,83	51,17	37	50
4	5	37,40	8,32	3,72	27,06	47,74	30	51
Total	19	47,84	11,46	2,63	42,32	53,36	30	71

ANOVA

SCORE

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1360,726	3	453,575	6,791	,004
Within Groups	1001,800	15	66,787		
Total	2362,526	18			

Αποτελέσματα (2)



Αποτελέσματα (3)

Multiple Comparisons

Dependent Variable: SCORE

LSD

(I) HCOLOR	(J) HCOLOR	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1	2	8,00	5,17	,143	-3,02	19,02
	3	16,70*	5,48	,008	5,02	28,38
	4	21,80*	5,17	,001	10,78	32,82
2	1	-8,00	5,17	,143	-19,02	3,02
	3	8,70	5,48	,133	-2,98	20,38
	4	13,80*	5,17	,017	2,78	24,82
3	1	-16,70*	5,48	,008	-28,38	-5,02
	2	-8,70	5,48	,133	-20,38	2,98
	4	5,10	5,48	,367	-6,58	16,78
4	1	-21,80*	5,17	,001	-32,82	-10,78
	2	-13,80*	5,17	,017	-24,82	-2,78
	3	-5,10	5,48	,367	-16,78	6,58

*. The mean difference is significant at the .05 level.

Αποτελέσματα (4)

Multiple Comparisons

Dependent Variable: SCORE

	(I) HCOLOR	(J) HCOLOR	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Tukey HSD	1	2	8,00	5,17	,436	-6,90	22,90
		3	16,70*	5,48	,037	,90	32,50
		4	21,80*	5,17	,004	6,90	36,70
	2	1	-8,00	5,17	,436	-22,90	6,90
		3	8,70	5,48	,415	-7,10	24,50
		4	13,80	5,17	,074	-1,10	28,70
	3	1	-16,70*	5,48	,037	-32,50	-,90
		2	-8,70	5,48	,415	-24,50	7,10
		4	5,10	5,48	,789	-10,70	20,90
	4	1	-21,80*	5,17	,004	-36,70	-6,90
		2	-13,80	5,17	,074	-28,70	1,10
		3	-5,10	5,48	,789	-20,90	10,70
Bonferroni	1	2	8,00	5,17	,855	-7,69	23,69
		3	16,70*	5,48	,049	5,46E-02	33,35
		4	21,80*	5,17	,004	6,11	37,49
	2	1	-8,00	5,17	,855	-23,69	7,69
		3	8,70	5,48	,800	-7,95	25,35
		4	13,80	5,17	,105	-1,89	29,49
	3	1	-16,70*	5,48	,049	-33,35	-5,46E-02
		2	-8,70	5,48	,800	-25,35	7,95
		4	5,10	5,48	1,000	-11,55	21,75
	4	1	-21,80*	5,17	,004	-37,49	-6,11
		2	-13,80	5,17	,105	-29,49	1,89
		3	-5,10	5,48	1,000	-21,75	11,55

*. The mean difference is significant at the .05 level.

Αποτελέσματα (5)

Homogeneous Subsets

SCORE

	HCOLOR	N	Subset for alpha = .05	
			1	2
Tukey HSD[a,b]	4	5	37,40	
	3	4	42,50	
	2	5	51,20	51,20
	1	5		59,20
	Sig.		,086	,461

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 4,706.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

Αποτελέσματα (6)

Test of Homogeneity of Variances

SCORE

Levene Statistic	df1	df2	Sig.
,493	3	15	,692

Η “ανακάλυψη” της μεθόδου ANOVA αποδίδεται επίσημα στον…

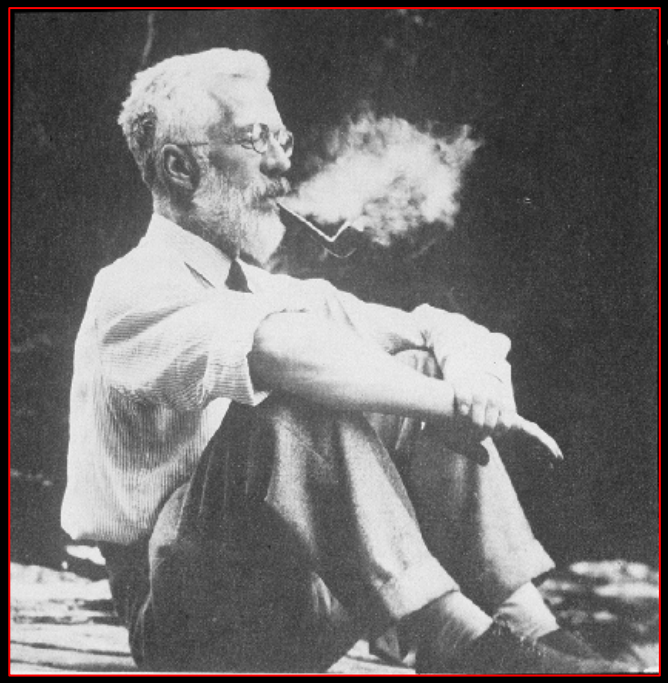
Sir Ronald Aylmer Fisher (1890-1962)

Born	17 February 1890 East Finchley, London , England
Died	29 July 1962 (aged 72) Adelaide, Australia
Residence	England Australia
Nationality	British
Fields	Statistician Evolutionary biologist Geneticist
Institutions	Rothamsted Experimental Station University College London Cambridge University CSIRO
Alma mater	Cambridge University
Academic advisors	Sir James Jeans F.J.M. Stratton
Doctoral students	C.R. Rao D. J. Finney

Known for	Fisher's fundamental theorem Maximum likelihood Fisher information Analysis of variance Fisher-Kolmogorov equation Coining the term 'null hypothesis' Fiducial inference Fisher's exact test Fisher's principle Fisherian runaway F-distribution
Influences	Leonard Darwin
Influenced	Joseph Oscar Irwin A. W. F. Edwards Georg Rasch
Notable awards	Royal Medal (1938) Guy Medal in Gold (1946) Copley Medal (1955) Linnean Society of London's Darwin-Wallace Medal (1958).
Religious stance	Church of England

Notes

He was the father-in-law of George E. P. Box.



- In 1919 Fisher started work at **Rothamsted Experimental Station** located at **Harpenden, Hertfordshire, England**. Here he started a major study of the extensive collections of data recorded over many years. This resulted in a series of reports under the general title *Studies in Crop Variation.* This began a period of great productivity. Over the next seven years, he pioneered the principles of the **design of experiments** and elaborated his studies of "**analysis of variance**". He furthered his studies of the statistics of small samples. Perhaps even more important, he began his systematic approach of the analysis of real data as the springboard for the development of new statistical methods. He began to pay particular attention to the labour involved in the necessary computations, and developed methods that were as practical as they were founded in rigour. In 1925, this work culminated in the publication of his first book, ***Statistical Methods for Research Workers***.[7] This went into many editions and translations in later years, and became a standard reference work for scientists in many disciplines. In 1935, this was followed by *The Design of Experiments,* which also became a standard.

In addition to "analysis of variance", Fisher invented the technique of maximum likelihood and originated the concepts of sufficiency, ancillarity, Fisher's linear discriminator and Fisher information. His 1924 article "On a distribution yielding the error functions of several well known statistics" presented Karl Pearson's chi-squared and Student's t in the same framework as the Gaussian distribution, and his own "analysis of variance" distribution z (more commonly used today in the form of the F distribution). These contributions easily made him a major figure in 20th century statistics.

- In defending the use of the z distribution when the data were not Gaussian, Fisher introduced the "randomization test". According to biographers Yates and Mather, "Fisher introduced the randomization test, comparing the value of t or z actually obtained with the distribution of the t or z values when all possible random arrangements were imposed on the experimental data."[8]
- However, Fisher wrote that randomization tests were "in no sense put forward to supersede the common and expeditious tests based on the Gaussian theory of errors." Fisher thus effectively began the field of non-parametric statistics, even though he didn't believe it was a necessary move.

Rothamsted is the largest agricultural research centre in the United Kingdom and almost certainly the oldest agricultural research station in the world

The **Rothamsted Experimental Station**, one of the oldest agricultural research institutions in the world, is located at Harpenden in Hertfordshire, England. It is now known as **Rothamsted Research**

Βιβλιογραφία

- Φασούλας, Α. Κ. (ανατ. 2008). *Στοιχεία Πειραματικής Στατιστικής*. Θεσσαλονίκη: Άγις-Σάββας Δ. Γαρταγάνης.
- Καλτσίκης, Π. Ι. (1997). *Απλά Πειραματικά Σχέδια*. Αθήνα: Εκδόσεις Α. Σταμούλη.
- Μιχαηλίδης, Ζ. (2005). *Βιομετρία-Γεωργικός Πειραματισμός*. ΑΤΕΙ Θεσσαλονίκης.
- Steel, R. & Torrie, J. (1986). *Principles and Procedures of Statistics: A Biometrical Approach*. Singapore: McGraw-Hill Book Company.
- Gomez, K. & Gomez, A. (1984). *Statistical Procedures for Agricultural Research*. Singapore: John Willey & Sons, Inc.
- Kuehl, R. (2000). *Designs of Experiments: Statistical Principles of Research Design and Analysis*. Pacific Grove: Duxbury Thomson Learning.

Σας ευχαριστώ για την προσοχή σας!!!

Rothamsted Park