Ανάλυση διασποράς (Ανάλυση διακύμανσης)

Σε προηγούμενο κεφάλαιο μελετήσαμε τους στατιστικούς ελέγχους υποθέσεων για τη σύγκριση των μέσων τιμών δύο πληθυσμών. Πολλές φορές όμως έχουμε πειραματικά δεδομένα από περισσότερους πληθυσμούς και μας ενδιαφέρει να συγκρίνουμε τις μέσες τιμές τους. Τα προβλήματα αυτά, καθώς και άλλα πολυπλοκότερα αντιμετωπίζονται με τη μέθοδο της ανάλυσης διασποράς ή ανάλυσης διακύμανσης.

Α) Ανάλυση διασποράς με ένα παράγοντα

Λαμβάνουμε κ δείγματα από κ πληθυσμούς για τους οποίους μας ενδιαφέρει να συγκρίνουμε τις μέσες τιμές τους. Έστω μ_i η μέση τιμή του i πληθυσμού, $i = 1, 2, \dots, \kappa$. Διατυπώνουμε τις στατιστικές υποθέσεις:

$H_0: \mu_1 = \mu_2 = \cdots = \mu_\kappa$ (οι μέσες τιμές *δεν διαφέρουν* ή ο παράγοντας που μελετούμε *δεν επιδρά*)

H_1: τουλάχιστον κάποιο $\mu_i \neq \mu_j$, $i, j = 1, 2, \dots, \kappa$ (τουλάχιστον δύο μέσες τιμές διαφέρουν ή ο παράγοντας που μελετούμε *επιδρά*)

Η απορριπτική περιοχή της H_0 είναι η εξής:

Εάν

$F = \frac{MSA}{MSE} > F_{\kappa-1, N-\kappa, \alpha}$ απορρίπτεται η H_0 (στατιστικό κριτήριο)

όπου η ποσότητα F υπολογίζεται στον παρακάτω πίνακα ανάλυσης διασποράς για ένα παράγοντα και $F_{\kappa-1, N-\kappa, \alpha}$ μια κριτική τιμή της F κατανομής με $\kappa - 1$ και $N - \kappa$ βαθμούς ελευθερίας για επίπεδο σημαντικότητας α (ο πίνακας της F κατανομής δίνεται στο τέλος του κειμένου).

Προϋποθέσεις για την εφαρμογή της μεθόδου: *Όλοι οι πληθυσμοί από τους οποίους προέρχονται τα κ δείγματα είναι κανονικοί με ίσες διακυμάνσεις.*

Πίνακας ανάλυσης διασποράς με ένα παράγοντα

Πηγή μεταβολής	Αθροίσματα τετραγώνων	Βαθμοί ελευθερίας	Μέσα τετράγωνα	Κριτήριο F
Μεταξύ των ομάδων (παράγοντας Α)	$SSA = \sum_{i=1}^{\kappa} \nu_i (\bar{y}_i - \bar{y})^2 =$	$\kappa - 1$	$MSA = \frac{SSA}{\kappa - 1}$	$F = \frac{MSA}{MSE}$
Εντός των ομάδων (σφάλματα ή υπόλοιπα)	$SSE = \sum_{i=1}^{\kappa} \sum_{j=1}^{\nu_i} (y_{ij} - \bar{y}_i)^2$	$N - \kappa$	$MSE = \frac{SSE}{N - \kappa}$	
Ολική	$SST = \sum_{i=1}^{\kappa} \sum_{j=1}^{\nu_i} (y_{ij} - \bar{y})^2$	$N - 1$		

όπου κ το πλήθος των πληθυσμών, των οποίων τις μέσες τιμές θέλουμε να συγκρίνουμε ή αλλιώς το πλήθος των σταθμών του παράγοντα που μελετούμε, ν_i το μέγεθος δείγματος που λαμβάνουμε από τον i πληθυσμό, $i = 1, 2, \dots, \kappa$, N το συνολικό μέγεθος δείγματος $N = \nu_1 + \nu_2 + \cdots + \nu_\kappa$ και $\bar{y}_i$ και $\bar{y}$ οι δειγματικές μέσες τιμές που υπολογίζονται ως εξής:

$$\bar{y}_i = \frac{1}{\nu_i} \sum_{j=1}^{\nu_i} y_{ij} \quad , \qquad \bar{y} = \frac{1}{N} \sum_{i=1}^{\kappa} \sum_{j=1}^{\nu_i} y_{ij}$$

Επιπλέον για το άθροισμα τετραγώνων της ολικής μεταβολής ισχύει: $SST = SSA + SSE$. Συνήθως το άθροισμα τετραγώνων των σφαλμάτων υπολογίζεται ως εξής: $SSE = SST - SSA$. Τα προβλήματα 1, 2, 3 που βρίσκονται λυμένα παρακάτω, αποτελούν εφαρμογές της μεθόδου ανάλυσης διασποράς με ένα παράγοντα (σελ. 5 - 9).

Β) Ανάλυση διασποράς με δύο παράγοντες χωρίς έλεγχο ύπαρξης αλληλεπίδρασης

Σε αυτήν την περίπτωση επιθυμούμε να διερευνήσουμε εάν το αποτέλεσμα ενός πειράματος επηρεάζεται από δύο παράγοντες, έστω Α και Β. Ο παράγοντας Α μπορεί να υπεισέρχεται στο πείραμα σε κ στάθμες, ενώ ο Β σε λ στάθμες. Για παράδειγμα έστω ότι ένας γεωπόνος επιθυμεί να διερευνήσει την επίδραση στην απόδοση τεσσάρων ποικιλιών αραβοσίτου ($\kappa = 4$) και τριών λιπασμάτων ($\lambda = 3$). Επομένως από τα πειραματικά δεδομένα ζητείται να εξαχθεί κάποιο συμπέρασμα για το αν υπάρχει στατιστικά σημαντική διαφορά στην απόδοση ανάλογα με την ποικιλία ή το λίπασμα που χρησιμοποιήθηκε. Για τη στατιστική ανάλυση των δεδομένων ενός πειράματος αυτής της μορφής χρησιμοποιείται η μέθοδος της ανάλυσης διασποράς για δύο παράγοντες. Παρακάτω δίνεται ο πίνακας της ανάλυσης διασποράς για δύο παράγοντες χωρίς έλεγχο ύπαρξης αλληλεπίδρασης και στις λύσεις των προβλημάτων 4 και 5 (σελ. 9 - 14) διατυπώνονται αναλυτικά οι στατιστικές υποθέσεις και τα στατιστικά κριτήρια για τη σημαντικότητα των δύο παραγόντων, που αποτελούν εφαρμογές της μεθόδου.

Πίνακας ανάλυσης διασποράς με δύο παράγοντες χωρίς έλεγχο ύπαρξης αλληλεπίδρασης

Πηγή μεταβολής	Αθροίσματα τετραγώνων	Βαθμοί ελευθερίας	Μέσα τετράγωνα	Κριτήριο F
Παράγοντας Α	$SSA = \lambda \sum_{i=1}^{\kappa} (\bar{y}_{i\bullet} - \bar{y}_{\bullet\bullet})^2$	$\kappa - 1$	$MSA = \frac{SSA}{\kappa - 1}$	$F_A = \frac{MSA}{MSE}$
Παράγοντας Β	$SSB = \kappa \sum_{i=1}^{\kappa} (\bar{y}_{\bullet j} - \bar{y}_{\bullet\bullet})^2$	$\lambda - 1$	$MSB = \frac{SSB}{\lambda - 1}$	$F_B = \frac{MSB}{MSE}$
Σφάλματα ή υπόλοιπα	$SSE = SST - SSA - SSB$	$(\kappa - 1)(\lambda - 1)$	$MSE = \frac{SSE}{(\kappa - 1)(\lambda - 1)}$	
Ολική	$SST = \sum_{i=1}^{\kappa} \sum_{j=1}^{\lambda} (y_{ij} - \bar{y}_{\bullet\bullet})^2$	$\kappa\lambda - 1$		

όπου κ και λ οι στάθμες των παραγόντων Α και Β και $\bar{y}_{i\bullet}$, $\bar{y}_{\bullet j}$ και $\bar{y}_{\bullet\bullet}$ οι δειγματικές μέσες τιμές που υπολογίζονται ως εξής:

$$\bar{y}_{i\bullet} = \frac{1}{\lambda} \sum_{j=1}^{\lambda} y_{ij} \quad , \quad \bar{y}_{\bullet j} = \frac{1}{\kappa} \sum_{i=1}^{\kappa} y_{ij} \quad , \quad \bar{y}_{\bullet\bullet} = \frac{1}{\kappa\lambda} \sum_{i=1}^{\kappa} \sum_{j=1}^{\lambda} y_{ij}$$

Γ) Ανάλυση διασποράς με δύο παράγοντες με έλεγχο ύπαρξης αλληλεπίδρασης

Σε αυτήν την περίπτωση επιθυμούμε να διερευνήσουμε εάν το αποτέλεσμα ενός πειράματος επηρεάζεται από δύο παράγοντες, έστω Α και Β, αλλά και να εξετάσουμε εάν υπάρχει αλληλεπίδραση Α×Β μεταξύ των παραγόντων του πειράματος. Παρακάτω δίνεται ο πίνακας της ανάλυσης διασποράς για δύο παράγοντες με έλεγχο ύπαρξης αλληλεπίδρασης και στις λύσεις των προβλημάτων 6 και 7 (σελ. 14 - 19) διατυπώνονται αναλυτικά οι στατιστικές υποθέσεις και τα στατιστικά κριτήρια για τη σημαντικότητα των δύο παραγόντων, καθώς και της μεταξύ τους αλληλεπίδρασης, που αποτελούν εφαρμογές της μεθόδου.

Πίνακας ανάλυσης διασποράς με δύο παράγοντες με έλεγχο ύπαρξης αλληλεπίδρασης

Πηγή μεταβολής	Αθροίσματα τετραγώνων	Βαθμοί ελευθερίας	Μέσα τετράγωνα	Κριτήριο F
Παράγοντας Α	$SSA = \lambda r \sum_{i=1}^{\kappa} (\bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet\bullet\bullet})^2$	$\kappa - 1$	$MSA = \frac{SSA}{\kappa - 1}$	$F_A = \frac{MSA}{MSE}$
Παράγοντας Β	$SSB = \kappa r \sum_{i=1}^{\kappa} (\bar{y}_{\bullet j\bullet} - \bar{y}_{\bullet\bullet\bullet})^2$	$\lambda - 1$	$MSB = \frac{SSB}{\lambda - 1}$	$F_B = \frac{MSB}{MSE}$
Αλληλεπί-δραση Α×Β	$SSAB = r \sum_{i=1}^{\kappa} \sum_{j=1}^{\lambda} (\bar{y}_{ij\bullet} - \bar{y}_{i\bullet\bullet} - \bar{y}_{\bullet j\bullet} - \bar{y}_{\bullet\bullet\bullet})^2$	$(\kappa - 1)(\lambda - 1)$	$MSAB = \frac{SSAB}{(\kappa - 1)(\lambda - 1)}$	$F_{AB} = \frac{MSAB}{MSE}$
Σφάλματα ή υπόλοιπα	$SSE = \sum_{i=1}^{\kappa} \sum_{j=1}^{\lambda} \sum_{\mu=1}^{r} (y_{ij\mu} - \bar{y}_{ij\bullet})^2$	$\kappa\lambda(r - 1)$	$MSE = \frac{SSE}{\kappa\lambda(r - 1)}$	
Ολική	$SST = \sum_{i=1}^{\kappa} \sum_{j=1}^{\lambda} \sum_{\mu=1}^{r} (y_{ij\mu} - \bar{y}_{\bullet\bullet\bullet})^2$	$\kappa\lambda - 1$		

όπου κ και λ οι στάθμες των παραγόντων Α και Β, r οι επαναλήψεις σε κάθε συνδυασμό των παραγόντων Α και Β και $\bar{y}_{i\bullet\bullet}$, $\bar{y}_{\bullet j\bullet}$, $\bar{y}_{ij\bullet}$, και $\bar{y}_{\bullet\bullet\bullet}$ οι δειγματικές μέσες τιμές που υπολογίζονται ως εξής:

$$\bar{y}_{i\bullet\bullet} = \frac{1}{\lambda r} \sum_{j=1}^{\lambda} \sum_{\mu=1}^{r} y_{ij\mu} \quad , \quad \bar{y}_{\bullet j\bullet} = \frac{1}{\kappa r} \sum_{i=1}^{\kappa} \sum_{\mu=1}^{r} y_{ij\mu} \quad , \quad \bar{y}_{ij\bullet} = \frac{1}{r} \sum_{\mu=1}^{r} y_{ij\mu}$$

$$\bar{y}_{\bullet\bullet\bullet} = \frac{1}{\kappa\lambda r} \sum_{i=1}^{\kappa} \sum_{j=1}^{\lambda} \sum_{\mu=1}^{r} y_{ij\mu}$$

Λύσεις των προβλημάτων από το **φυλλάδιο 8 – Προβλήματα ανάλυσης διασποράς**

1. Η απόδοση σε γάλα (Kg/24h) μιας προβατίνας που έχει γεννήσει υπολογίζεται ζυγίζοντας το νεογνό πριν και μετά το θηλασμό. Πήραμε δείγματα από τρεις 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

Σε επίπεδο σημαντικότητας 5% υπάρχει στατιστικά σημαντική διαφορά στη γαλακτοπαραγωγή ανάμεσα στις τρεις φυλές; Διατυπώστε με ακρίβεια τις στατιστικές υποθέσεις που ελέγχετε.
(Δίνονται: SSA=2.6, SST=6.5).

Θέλουμε να συγκρίνουμε τη μέση γαλακτοπαραγωγή των τριών φυλών, επομένως έχουμε ένα πρόβλημα ανάλυσης διασποράς με ένα παράγοντα (Α: φυλή). Έστω μ_i η μέση γαλακτοπαραγωγή της *i* φυλής $i = 1, 2, 3$. Διατυπώνουμε τις στατιστικές υποθέσεις:

H_0: $\mu_1 = \mu_2 = \mu_3$ (η μέση γαλακτοπαραγωγή είναι *ίδια* και για τις τρεις φυλές ή ο παράγοντας "φυλή" *δεν επιδρά* στη γαλακτοπαραγωγή)

H_1: τουλάχιστον κάποιο $\mu_i \neq \mu_j$, $i, j = 1, 2, 3$ (η μέση γαλακτοπαραγωγή *διαφοροποιείται* ανάλογα με τη φυλή ή ο παράγοντας "φυλή" *επιδρά* στη γαλακτοπαραγωγή)

Στατιστικό κριτήριο ελέγχου της σημαντικότητας του παράγοντα Α:
Εάν

$F = \frac{MSA}{MSE} > F_{\kappa-1, N-\kappa, \alpha}$ απορρίπτεται η H_0,

όπου η ποσότητα F υπολογίζεται στον παρακάτω πίνακα ανάλυσης διασποράς για ένα παράγοντα και $F_{\kappa-1, N-\kappa, \alpha}$ μια κριτική τιμή της F κατανομής με $\kappa - 1$ και $N - \kappa$ βαθμούς ελευθερίας για επίπεδο σημαντικότητας α (ο πίνακας της F κατανομής δίνεται στο τέλος του κειμένου).

Αντικαθιστώντας στον πίνακα ανάλυσης διασποράς με ένα παράγοντα έχουμε:

Πηγή μεταβολής	Αθροίσματα τετραγώνων	Βαθμοί ελευθερίας	Μέσα τετράγωνα	Κριτήριο F
Μεταξύ των ομάδων (παράγοντας A)	$SSA = \sum_{i=1}^{\kappa} \nu_i(\bar{y}_i - \bar{y})^2 =$ $= 2.6$	$\kappa - 1 =$ $3 - 1 =$ 2	$MSA = \frac{SSA}{\kappa - 1}$ $= \frac{2.6}{2} = 1.3$	$F = \frac{MSA}{MSE}$ $= \frac{1.3}{0.26} = 5$
Εντός των ομάδων (σφάλματα ή υπόλοιπα)	$SSE = \sum_{i=1}^{\kappa} \sum_{j=1}^{\nu_i} (y_{ij} - \bar{y}_i)^2$ $SSE = SST - SSA =$ $= 6.5 - 2.6 = 3.9$	$N - \kappa =$ $18 - 3 =$ 15	$MSE = \frac{SSE}{N - \kappa}$ $= \frac{3.9}{15} = 0.26$	
Ολική	$SST = \sum_{i=1}^{\kappa} \sum_{j=1}^{\nu_i} (y_{ij} - \bar{y})^2$ $= 6.5$	$N - 1 =$ $18 - 1 =$ 17		

Επομένως έχουμε:

$$F = \frac{MSA}{MSE} = 5 > F_{\kappa-1,N-\kappa,\alpha} = F_{3-1,18-3,0.05} = F_{2,15,0.05} = 3.68$$

Άρα σε επίπεδο σημαντικότητας 5%, ισχύει η ανισότητα του στατιστικού κριτηρίου και συνεπώς απορρίπτουμε την H_0. Επομένως υπάρχει στατιστικά σημαντική διαφορά ανάμεσα στις τρεις φυλές ως προς τη μέση γαλακτοπαραγωγή.

2. Μετρήθηκε η ποσότητα πρωτεΐνης (gr/100ml) στο αίμα ατόμων που ζουν σε διαφορετικές συνθήκες στις γεωγραφικές περιοχές A, B, 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 διαφέρει ανάλογα με τη γεωγραφική περιοχή.

(Δίδονται: SSA=0.53, SST =4.77)

Θέλουμε να συγκρίνουμε τη μέση ποσότητα πρωτεΐνης στο αίμα ατόμων που ζουν σε διαφορετικές γεωγραφικές περιοχές. Πρόκειται για ένα πρόβλημα ανάλυσης διασποράς με ένα παράγοντα (Α: γεωγραφική περιοχή). Έστω μ_i η μέση ποσότητα πρωτεΐνης στην i γεωγραφική περιοχή $i = 1, 2, 3$. Διατυπώνουμε τις στατιστικές υποθέσεις:

$H_0: \mu_1 = \mu_2 = \mu_3$	(η μέση ποσότητα πρωτεΐνης *δεν διαφοροποιείται* ανάλογα με τη γεωγραφική περιοχή ή ο παράγοντας "γεωγραφική περιοχή" *δεν επιδρά* στην ποσότητα πρωτεΐνης)
H_1: τουλάχιστον κάποιο $\mu_i \neq \mu_j$ $i, j = 1, 2, 3$	(η μέση ποσότητα πρωτεΐνης *διαφοροποιείται* ανάλογα με τη γεωγραφική περιοχή ή ο παράγοντας "γεωγραφική περιοχή" *επιδρά* στην ποσότητα πρωτεΐνης)

Αντικαθιστώντας στον πίνακα ανάλυσης διασποράς με ένα παράγοντα έχουμε:

Πηγή μεταβολής	Αθροίσματα τετραγώνων	Βαθμοί ελευθερίας	Μέσα τετράγωνα	Κριτήριο F
Μεταξύ των ομάδων (παράγοντας Α)	$SSA = \sum_{i=1}^{\kappa} \nu_i (\bar{y}_i - \bar{y})^2 =$ $= 0.53$	$\kappa - 1 =$ $3 - 1 =$ 2	$MSA = \frac{SSA}{\kappa - 1}$ $= \frac{0.53}{2} = 0.265$	$F = \frac{MSA}{MSE}$ $= \frac{0.265}{0.202} = 1.3$
Εντός των ομάδων (σφάλματα ή υπόλοιπα)	$SSE = \sum_{i=1}^{\kappa} \sum_{j=1}^{\nu_i} (y_{ij} - \bar{y}_i)^2$ $SSE = SST - SSA =$ $= 4.77 - 0.53 = 4.24$	$N - \kappa =$ $24 - 3 =$ 21	$MSE = \frac{SSE}{N - \kappa}$ $= \frac{4.24}{21} = 0.202$	
Ολική	$SST = \sum_{i=1}^{\kappa} \sum_{j=1}^{\nu_i} \left(y_{ij} - \bar{y}\right)^2$ $= 4.77$	$N - 1 =$ $24 - 1 =$ 23		

Στατιστικό κριτήριο ελέγχου της σημαντικότητας του παράγοντα Α:

Εάν

$F = \frac{MSA}{MSE} > F_{\kappa-1, N-\kappa, \alpha}$ απορρίπτεται η $H_{0.}$

Επομένως έχουμε:

$$F = \frac{MSA}{MSE} = 1.3 < F_{\kappa-1,N-\kappa,\alpha} = F_{3-1,24-3,0.05} = F_{2,21,0.05} = 3.47$$

Άρα σε επίπεδο σημαντικότητας 5%, δεν ισχύει η ανισότητα του στατιστικού κριτηρίου και συνεπώς δεν μπορούμε να απορρίψουμε την H_0. Επομένως δεν υπάρχει στατιστικά σημαντική διαφορά στη μέση ποσότητα πρωτεΐνης στο αίμα ανάμεσα στις τρεις γεωγραφικές περιοχές.

3. Παρακάτω δίνονται οι αποδόσεις σε κιλά ανά πειραματική μονάδα τεσσάρων ποικιλιών αραβοσίτου:

Ποικιλίες										
A_1	67	72	76	81	83	69	77	65	82	
A_2	56	73	65	71	55	69	72	63	69	67
A_3	77	82	88	79	67	83	91	78	95	
A_4	75	85	90	89	67	93	77	68	75	

Διατυπώστε κατάλληλο έλεγχο υποθέσεων και ελέγξτε σε επίπεδο σημαντικότητας 5%, εάν υπάρχει στατιστικά σημαντική διαφορά ως προς την απόδοση, μεταξύ των τεσσάρων ποικιλιών αραβοσίτου. (Δίδονται: SSA=1488.6, SST =3519.1)

Πρόκειται για ένα πρόβλημα ανάλυσης διασποράς με ένα παράγοντα (Α: ποικιλία αραβοσίτου). Έστω μ_i η μέση απόδοση της *i* ποικιλίας $i = 1, 2, 3, 4$. Διατυπώνουμε τις στατιστικές υποθέσεις:

$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$ (η μέση απόδοση *δεν διαφοροποιείται* ανάλογα με την ποικιλία ή ο παράγοντας "ποικιλία" *δεν επιδρά* στην απόδοση)

H_1: τουλάχιστον κάποιο $\mu_i \neq \mu_j$ $i, j = 1, 2, 3, 4$ (η μέση απόδοση *διαφοροποιείται* ανάλογα με την ποικιλία ή ο παράγοντας "ποικιλία" *επιδρά* στην απόδοση)

Στατιστικό κριτήριο ελέγχου της σημαντικότητας του παράγοντα Α:

Εάν

$F = \frac{MSA}{MSE} > F_{\kappa-1,N-\kappa,\alpha}$ απορρίπτεται η H_0.

όπου η ποσότητα *F* υπολογίζεται στον παρακάτω πίνακα ανάλυσης διασποράς για ένα παράγοντα:

Πηγή μεταβολής	Αθροίσματα τετραγώνων	Βαθμοί ελευθερίας	Μέσα τετράγωνα	Κριτήριο F
Μεταξύ των ομάδων (παράγοντας Α) (ποικιλία)	$SSA = \sum_{i=1}^{\kappa} \nu_i (\bar{y}_i - \bar{y})^2 =$ $= 1488.6$	$\kappa - 1 =$ $4 - 1 =$ 3	$MSA = \frac{SSA}{\kappa - 1} =$ $= \frac{1488.6}{3} = 496.2$	$F = \frac{MSA}{MSE} =$ $= \frac{496.2}{61.5} = 8.07$
Εντός των ομάδων (σφάλματα ή υπόλοιπα)	$SSE = \sum_{i=1}^{\kappa} \sum_{j=1}^{\nu_i} (y_{ij} - \bar{y}_i)^2$ $SSE = SST - SSA =$ $= 3519.1 - 1488.6 = 2030.5$	$N - \kappa =$ $37 - 4 =$ 33	$MSE = \frac{SSE}{N - \kappa}$ $= \frac{2030.5}{33} = 61.5$	
Ολική	$SST = \sum_{i=1}^{\kappa} \sum_{j=1}^{\nu_i} (y_{ij} - \bar{y})^2$ $= 3519.1$	$N - 1 =$ $37 - 1 =$ 36		

Επομένως έχουμε:

$$F = \frac{MSA}{MSE} = 8.07 > F_{\kappa-1,N-\kappa,\alpha} = F_{4-1,37-4,0.05} = F_{3,33,0.05} = 2.92$$

Άρα σε επίπεδο σημαντικότητας 5%, ισχύει η ανισότητα του στατιστικού κριτηρίου και συνεπώς απορρίπτουμε την H_0. Επομένως η μέση απόδοση διαφοροποιείται ανάλογα με την ποικιλία ή αλλιώς μπορούμε να πούμε ότι υπάρχει στατιστικά σημαντική διαφορά στη μέση απόδοση του αραβοσίτου που οφείλεται στον παράγοντα “ποικιλία”.

4. Τα παρακάτω δεδομένα δίνουν τις αποδόσεις σε gr μιας εργαστηριακής καλλιέργειας σιταριού, στην οποία χρησιμοποιήθηκαν τρία είδη λιπασμάτων και τέσσερα είδη ορμονών.

	ΟΡΜΟΝΗ			
ΛΙΠΑΣΜΑ	B_1	B_2	B_3	B_4
A_1	77	78	80	82
A_2	73	76	76	77
A_3	76	77	82	83

Αφού διατυπώσετε κατάλληλους ελέγχους υποθέσεων, ελέγξτε σε επίπεδο σημαντικότητας 5%, εάν η απόδοση της εργαστηριακής καλλιέργειας σιταριού διαφοροποιείται ανάλογα με το είδος του λιπάσματος και το είδος της ορμόνης που χρησιμοποιείται.
(Δίνονται : SSA = 40.2,   SSB = 50.9,   SST = 100.9)

Πρόκειται για ένα πρόβλημα ανάλυσης διασποράς με δύο παράγοντες (Α: λίπανση και Β: χρήση ορμόνης). Καθώς έχουμε μόνο μία μέτρηση σε κάθε συνδυασμό λιπάσματος και ορμόνης δεν είναι εφικτός ο έλεγχος για την ύπαρξη αλληλεπίδρασης μεταξύ των δύο παραγόντων του πειράματος. Διατυπώνουμε τις στατιστικές υποθέσεις:

$H_{0A}: \alpha_1 = \alpha_2 = \alpha_3$ (η μέση απόδοση *δεν διαφοροποιείται* ανάλογα με το είδος του λιπάσματος που χρησιμοποιήθηκε ή ο παράγοντας " λίπανση" *δεν επιδρά* στην απόδοση)

H_{1A}: τουλάχιστον κάποιο $\alpha_i \neq \alpha_j$, $i,j = 1,2,3$ (η μέση απόδοση *διαφοροποιείται* ανάλογα με το είδος του λιπάσματος που χρησιμοποιήθηκε ή ο παράγοντας " λίπανση" *επιδρά* στην απόδοση)

$H_{0B}: \beta_1 = \beta_2 = \beta_3 = \beta_4$, $i,j = 1,2,3$ (η μέση απόδοση *δεν διαφοροποιείται* ανάλογα με το είδος της ορμόνης που χρησιμοποιήθηκε ή ο παράγοντας "ορμόνη" *δεν επιδρά* στην απόδοση)

H_{1B}: τουλάχιστον κάποιο $\beta_i \neq \beta_j$, $i,j = 1,2,3,4$ (η μέση απόδοση *διαφοροποιείται* ανάλογα με το είδος της ορμόνης που χρησιμοποιήθηκε ή ο παράγοντας "ορμόνη" *επιδρά* στην απόδοση)

Στατιστικό κριτήριο ελέγχου της σημαντικότητας του παράγοντα Α:
Εάν

$F_A = \frac{MSA}{MSE} > F_{\kappa-1,(\kappa-1)(\lambda-1),\alpha}$ απορρίπτεται η $H_{0A,}$

όπου η ποσότητα F_A υπολογίζεται στον παρακάτω πίνακα ανάλυσης διασποράς με δύο παράγοντες χωρίς έλεγχο ύπαρξης αλληλεπίδρασης:

Πηγή μεταβολής	Αθροίσματα τετραγώνων	Βαθμοί ελευθερίας	Μέσα τετράγωνα	Κριτήρια F
Παράγοντας Α (Λίπανση)	$SSA = \lambda \sum_{i=1}^{\kappa} (\bar{y}_{i\blacksquare} - \bar{y}_{\blacksquare\blacksquare})^2$ $= 40.2$	$\kappa - 1 =$ $3 - 1 = 2$	$MSA = \frac{SSA}{\kappa - 1} =$ $\frac{40.2}{2} = 20.1$	$F_A = \frac{MSA}{MSE} =$ $\frac{20.1}{1.6} = 12.6$
Παράγοντας Β (Ορμόνη)	$SSB = \kappa \sum_{i=1}^{\kappa} (\bar{y}_{\blacksquare j} - \bar{y}_{\blacksquare\blacksquare})^2$ $= 50.9$	$\lambda - 1 =$ $4 - 1 = 3$	$MSB = \frac{SSB}{\lambda - 1} =$ $\frac{50.9}{3} = 17.0$	$F_B = \frac{MSB}{MSE} =$ $\frac{17.0}{1.6} = 10.6$
Σφάλματα ή υπόλοιπα	$SSE = SST - SSA - SSB$ =100.9 - 40.2 - 50.9 = 9.8	$(\kappa - 1)(\lambda - 1) =$ $(3 - 1)(4 - 1) =$ $2 \cdot 3 = 6$	$MSE = \frac{SSE}{(\kappa - 1)(\lambda - 1)}$ $= \frac{9.8}{6} = 1.6$	
Ολική	$SST = \sum_{i=1}^{\kappa} \sum_{j=1}^{\lambda} (y_{ij} - \bar{y}_{\blacksquare\blacksquare})^2$ $= 100.9$	$\kappa\lambda - 1$ $3 \cdot 4 - 1 = 11$		

Επομένως έχουμε:

$$F_A = \frac{MSA}{MSE} = 12.6 > F_{\kappa-1,(\kappa-1)(\lambda-1),\alpha} = F_{3-1,(3-1)(4-1),0.05} = F_{2,6,0.05} = 5.14$$

Άρα σε επίπεδο σημαντικότητας 5%, ισχύει η ανισότητα του στατιστικού κριτηρίου και συνεπώς απορρίπτουμε την H_{0A}. Επομένως η μέση απόδοση της εργαστηριακής καλλιέργειας σιταριού διαφοροποιείται ανάλογα με το είδος του λιπάσματος που χρησιμοποιείται ή αλλιώς μπορούμε να πούμε ότι ο παράγοντας λίπανση είναι στατιστικά σημαντικός σε επίπεδο σημαντικότητας 5%.

Στατιστικό κριτήριο ελέγχου της σημαντικότητας του παράγοντα Β:

Εάν

$F_B = \frac{MSB}{MSE} > F_{\lambda-1,(\kappa-1)(\lambda-1),\alpha}$ απορρίπτεται η H_{0B}.

Αντικαθιστώντας έχουμε:

$$F_B = \frac{MSB}{MSE} = 10.6 > F_{\lambda-1,(\kappa-1)(\lambda-1),\alpha} = F_{4-1,(3-1)(4-1),0.05} = F_{3,6,0.05} = 4.76$$

Άρα σε επίπεδο σημαντικότητας 5%, ισχύει η ανισότητα του στατιστικού κριτηρίου και συνεπώς απορρίπτουμε την H_{0B}. Επομένως η μέση απόδοση της εργαστηριακής καλλιέργειας σιταριού διαφοροποιείται ανάλογα με το είδος της ορμόνης που χρησιμοποιείται.

5. Κάποιο χημικό πείραμα έλαβε χώρα με 4 διαφορετικούς καταλύτες και σε 3 διαφορετικές θερμοκρασίες. Τα αποτελέσματα του πειράματος φαίνονται στον παρακάτω πίνακα.

	Καταλύτης			
Θερμοκρασία	A_1	A_2	A_3	A_4
B_1	53	59	58	50
B_2	57	65	62	60
B_3	52	62	54	52

Αφού διατυπώσετε κατάλληλους ελέγχους υποθέσεων, ελέγξτε σε επίπεδο σημαντικότητας 5%, εάν το αποτέλεσμα του πειράματος διαφοροποιείται ανάλογα με το είδος του καταλύτη και τη θερμοκρασία.
(Δίνονται : SSA = 132,   SSB = 96,   SST = 252)**.**

Πρόκειται για ένα πρόβλημα ανάλυσης διασποράς με δύο παράγοντες (Α: καταλύτης και Β: θερμοκρασία).
Διατυπώνουμε τις στατιστικές υποθέσεις:

$H_{0A}: \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$ (το αποτέλεσμα του πειράματος *δεν διαφοροποιείται* ανάλογα με το είδος του καταλύτη που χρησιμοποιείται ή ο παράγοντας “ καταλύτης” *δεν επιδρά* στο αποτέλεσμα του πειράματος)

H_{1A}: τουλάχιστον κάποιο $\alpha_i \neq \alpha_j$, $i, j = 1, 2, 3, 4$ (το αποτέλεσμα του πειράματος *διαφοροποιείται* ανάλογα με το είδος του καταλύτη που χρησιμοποιείται ή ο παράγοντας “ καταλύτης” *επιδρά* στο αποτέλεσμα του πειράματος)

$H_{0B}: \beta_1 = \beta_2 = \beta_3$ (το αποτέλεσμα του πειράματος *δεν διαφοροποιείται* ανάλογα με τη θερμοκρασία ή ο παράγοντας“ θερμοκρασία” *δεν επιδρά* στο αποτέλεσμα του πειράματος)

H_{1B}: τουλάχιστον κάποιο $\beta_i \neq \beta_j$, $i, j = 1, 2, 3$ (το αποτέλεσμα του πειράματος *διαφοροποιείται* ανάλογα με τη θερμοκρασία ή ο παράγοντας“ θερμοκρασία” *επιδρά* στο αποτέλεσμα του πειράματος)

Αντικαθιστώντας στον πίνακα ανάλυσης διασποράς για δύο παράγοντες χωρίς έλεγχο ύπαρξης αλληλεπίδρασης έχουμε:

Πηγή μεταβολής	Αθροίσματα τετραγώνων	Βαθμοί ελευθερίας	Μέσα τετράγωνα	Κριτήρια F
Παράγοντας Α (καταλύτης)	$SSA = \lambda \sum_{i=1}^{\kappa} (\bar{y}_{i\blacksquare} - \bar{y}_{\blacksquare\blacksquare})^2$ $= 132$	$\kappa - 1 =$ $4 - 1 = 3$	$MSA = \frac{SSA}{\kappa - 1} =$ $\frac{132}{3} = 44$	$F_A = \frac{MSA}{MSE} =$ $\frac{44}{4} = 11$
Παράγοντας Β (θερμοκρασία)	$SSB = \kappa \sum_{i=1}^{\kappa} (\bar{y}_{\blacksquare j} - \bar{y}_{\blacksquare\blacksquare})^2$ $= 96$	$\lambda - 1 =$ $3 - 1 = 2$	$MSB = \frac{SSB}{\lambda - 1} =$ $\frac{96}{2} = 48$	$F_B = \frac{MSB}{MSE} =$ $\frac{48}{4} = 12$
Σφάλματα ή υπόλοιπα	$SSE = SST - SSA - SSB$ $= 252 - 132 - 96 = 24$	$(\kappa - 1)(\lambda - 1) =$ $(4 - 1)(3 - 1) =$ $3 \cdot 2 = 6$	$MSE = \frac{SSE}{(\kappa - 1)(\lambda - 1)}$ $= \frac{24}{6} = 4$	
Ολική	$SST = \sum_{i=1}^{\kappa} \sum_{j=1}^{\lambda} (y_{ij} - \bar{y}_{\blacksquare\blacksquare})^2$ $= 252$	$\kappa\lambda - 1$ $4 \cdot 3 - 1 = 11$		

Στατιστικό κριτήριο ελέγχου της σημαντικότητας του παράγοντα Α:

Εάν

$F_A = \frac{MSA}{MSE} > F_{\kappa-1,(\kappa-1)(\lambda-1),\alpha}$ απορρίπτεται η H_{0A}.

Αντικαθιστώντας έχουμε:

$$F_A = \frac{MSA}{MSE} = 11 > F_{\kappa-1,(\kappa-1)(\lambda-1),\alpha} = F_{4-1,(4-1)(3-1),0.05} = F_{3,6,0.05} = 4.76$$

Άρα σε επίπεδο σημαντικότητας 5%, ισχύει η ανισότητα του στατιστικού κριτηρίου και συνεπώς απορρίπτουμε την H_{0A}. Επομένως το αποτέλεσμα του πειράματος *διαφοροποιείται* ανάλογα με το είδος του καταλύτη που χρησιμοποιείται ή αλλιώς μπορούμε να πούμε ότι ο παράγοντας “ καταλύτης” *επιδρά* στο αποτέλεσμα του πειράματος.

Στατιστικό κριτήριο ελέγχου της σημαντικότητας του παράγοντα Β:

Εάν

$F_B = \frac{MSB}{MSE} > F_{\lambda-1,(\kappa-1)(\lambda-1),\alpha}$ απορρίπτεται η H_{0B}.

Αντικαθιστώντας έχουμε:

$$F_B = \frac{MSB}{MSE} = 12 > F_{\lambda-1,(\kappa-1)(\lambda-1),\alpha} = F_{3-1,(3-1)(4-1),0.05} = F_{2,6,0.05} = 5.14$$

Άρα σε επίπεδο σημαντικότητας 5%, ισχύει η ανισότητα του στατιστικού κριτηρίου και συνεπώς απορρίπτουμε την H_{0B}. Επομένως το αποτέλεσμα του πειράματος *διαφοροποιείται* ανάλογα με τη θερμοκρασία ή αλλιώς μπορούμε να πούμε ότι ο παράγοντας "θερμοκρασία" *επιδρά* στο αποτέλεσμα του πειράματος ή ότι είναι στατιστικά σημαντικός σε επίπεδο σημαντικότητας 5%.

6. Τρεις ποικιλίες σίτου (A) δοκιμάστηκαν με τρία είδη λιπάσματος (B), φωσφορική αμμωνία, θειική αμμωνία και καθόλου λίπανση. Οι αποδόσεις σε κάθε συνδυασμό ποικιλία και λιπάσματος φαίνονται στον παρακάτω πίνακα:

	Ποικιλίες (Α)					
Λίπανση (Β)	Α1		Α2		Α3	
Φωσφορική	112	128	112	81	134	112
Αμμωνία	118	152	108	48	116	128
Θειική	168	116	61	98	125	106
Αμμωνία	144	80	58	98	110	110
Μάρτυρας	106	84	97	86	62	60
(καμιά λίπανση)	68	128	92	66	99	87

Ελέγξτε σε επίπεδο σημαντικότητας 5% εάν υπάρχει διαφορά στις αποδόσεις του σίτου που να οφείλεται στις διαφορετικές ποικιλίες, στη διαφορετική λίπανση, καθώς και αν υπάρχει αλληλεπίδραση μεταξύ ποικιλίας και λιπάσματος.

(Δίνονται: SSA = 6743.4, SSB = 4481.7, SSAB = 2789.4, SST = 27548.5).

Πρόκειται για ένα πρόβλημα *ανάλυσης διασποράς με δύο παράγοντες (A: ποικιλία και B: λίπανση) και έλεγχο σημαντικότητας της αλληλεπίδρασης μεταξύ των δύο παραγόντων.* Ο έλεγχος ύπαρξης αλληλεπίδρασης είναι εφικτός, καθώς έχουμε τέσσερις επαναλήψεις για κάθε συνδυασμό ποικιλίας και λίπανσης (r = 4 >1).

Διατυπώνουμε τις στατιστικές υποθέσεις:

$H_{0A}: \alpha_1 = \alpha_2 = \alpha_3$ (η μέση απόδοση *δεν διαφοροποιείται* ανάλογα με την ποικιλία ή ο παράγοντας “ποικιλία” *δεν επιδρά* στην απόδοση)

H_{1A}: τουλάχιστον κάποιο $\alpha_i \neq \alpha_j$, $i, j = 1, 2, 3$ (η μέση απόδοση *διαφοροποιείται* ανάλογα με την ποικιλία ή ο παράγοντας “ποικιλία” *επιδρά* στην απόδοση)

$H_{0B}: \beta_1 = \beta_2 = \beta_3$ (η μέση απόδοση *δεν διαφοροποιείται* ανάλογα με την λίπανση ή ο παράγοντας “λίπανση” *δεν επιδρά* στην απόδοση)

H_{1B}: τουλάχιστον κάποιο $\beta_i \neq \beta_j$, $i, j = 1, 2, 3$ (η μέση απόδοση *διαφοροποιείται* ανάλογα με την λίπανση ή ο παράγοντας “λίπανση” *επιδρά* στην απόδοση)

$H_{0AB}: (\alpha\beta)_{ij} = 0$ (Δεν υπάρχει αλληλεπίδραση)
$H_{1AB}: (\alpha\beta)_{ij} \neq 0$ (Υπάρχει αλληλεπίδραση)

Αντικαθιστώντας στον πίνακα ανάλυσης διασποράς με δύο παράγοντες με έλεγχο ύπαρξης αλληλεπίδρασης έχουμε:

Πηγή μεταβολής	Αθροίσματα τετραγώνων	Βαθμοί ελευθερίας	Μέσα τετράγωνα	Κριτήρια F
Παράγοντας A (ποικιλία)	$SSA = \lambda r \sum_{i=1}^{\kappa} (\bar{y}_{i\blacksquare\blacksquare} - \bar{y}_{\blacksquare\blacksquare\blacksquare})^2 = 6743.4$	$\kappa - 1 =$ $3 - 1 = 2$	$MSA = \frac{SSA}{\kappa - 1} =$ $\frac{6743.4}{2} = 3371.7$	$F_A = \frac{MSA}{MSE} =$ $\frac{3371.7}{502.6} = 6.7$
Παράγοντας B (λίπανση)	$SSB = \kappa r \sum_{i=1}^{\kappa} (\bar{y}_{\blacksquare j\blacksquare} - \bar{y}_{\blacksquare\blacksquare\blacksquare})^2 = 4481.7$	$\lambda - 1 =$ $3 - 1 = 2$	$MSB = \frac{SSB}{\lambda - 1} =$ $\frac{4481.7}{2} = 2240.9$	$F_B = \frac{MSB}{MSE} =$ $\frac{2240.9}{502.6} = 4.5$
Αλληλεπί-δραση A×B	$SSAB =$ $r \sum_{i=1}^{\kappa} \sum_{j=1}^{\lambda} (\bar{y}_{ij\blacksquare} - \bar{y}_{i\blacksquare\blacksquare} - \bar{y}_{\blacksquare j\blacksquare} - \bar{y}_{\blacksquare\blacksquare\blacksquare})^2$ $= 2789.4$	$(\kappa - 1)(\lambda - 1)$ $(3 - 1)(3 - 1)$ $= 2 \cdot 2 = 4$	$MSAB = \frac{SSAB}{(\kappa - 1)(\lambda - 1)}$ $= \frac{2789.4}{4} = 697.4$	$F_{AB} = \frac{MSAB}{MSE}$ $= \frac{2240.9}{502.6} = 1.4$
Σφάλματα ή υπόλοιπα	$SSE = \sum_{i=1}^{\kappa} \sum_{j=1}^{\lambda} \sum_{\mu=1}^{r} (y_{ij\mu} - \bar{y}_{ij\blacksquare})^2 =$ $SSE = SST - SSA - SSB - SSAB =$ 27548.5-6743.4-4481.7-2789.4=13570.0	$\kappa\lambda(r - 1) =$ $3 \cdot 3 \cdot (4 - 1) =$ 27	$MSE = \frac{SSE}{\kappa\lambda(r - 1)} =$ $\frac{13570}{27} = 502.6$	
Ολική	$SST = \sum_{i=1}^{\kappa} \sum_{j=1}^{\lambda} \sum_{\mu=1}^{r} (y_{ij\mu} - \bar{y}_{\blacksquare\blacksquare\blacksquare})^2$ $= 27548.5$	$\kappa\lambda r - 1 =$ $3 \cdot 3 \cdot 4 - 1 =$ 35		

Στατιστικό κριτήριο ελέγχου της σημαντικότητας του παράγοντα Α:

Εάν

$F_A = \frac{MSA}{MSE} > F_{\kappa-1,\kappa\lambda(r-1),\alpha}$ απορρίπτεται η $H_{0A.}$

Αντικαθιστώντας έχουμε:

$$F_A = \frac{MSA}{MSE} = 6.7 > F_{\kappa-1,\kappa\lambda(r-1),\alpha} = F_{3-1,3\cdot3(4-1),0.05} = F_{2,27,0.05} = 3.35$$

Άρα σε επίπεδο σημαντικότητας 5%, ισχύει η ανισότητα του στατιστικού κριτηρίου και συνεπώς απορρίπτουμε την H_{0A}. Επομένως η μέση απόδοση διαφοροποιείται ανάλογα με το είδος της ποικιλίας ή αλλιώς μπορούμε να πούμε ότι ο παράγοντας "ποικιλία" είναι στατιστικά σημαντικός σε επίπεδο σημαντικότητας 5%.

Στατιστικό κριτήριο ελέγχου της σημαντικότητας του παράγοντα Β:

Εάν

$F_B = \frac{MSB}{MSE} > F_{\lambda-1,\kappa\lambda(r-1),\alpha}$ απορρίπτεται η $H_{0B.}$

Αντικαθιστώντας έχουμε:

$$F_B = \frac{MSB}{MSE} = 4.5 > F_{\lambda-1,\kappa\lambda(r-1),\alpha} = F_{3-1,3\cdot3(4-1),0.05} = F_{2,27,0.05} = 3.35$$

Άρα σε επίπεδο σημαντικότητας 5%, ισχύει η ανισότητα του στατιστικού κριτηρίου και συνεπώς απορρίπτουμε την H_{0B}. Επομένως η μέση απόδοση διαφοροποιείται ανάλογα με το είδος του λιπάσματος ή αλλιώς μπορούμε να πούμε ότι ο παράγοντας "λίπανση" είναι στατιστικά σημαντικός σε επίπεδο σημαντικότητας 5%.

Στατιστικό κριτήριο ελέγχου της σημαντικότητας της αλληλεπίδρασης των δύο παραγόντων:

Εάν

$F_{AB} = \frac{MSAB}{MSE} > F_{(\kappa-1)(\lambda-1),\kappa\lambda(r-1),\alpha}$ απορρίπτεται η $H_{0AB.}$

Αντικαθιστώντας έχουμε:

$$F_{AB} = \frac{MSAB}{MSE} = 1.4 < F_{(\kappa-1)(\lambda-1),\kappa\lambda(r-1),\alpha} = F_{(3-1)(3-1),3\cdot3(4-1),0.05} = F_{4,27,0.05} = 2.73$$

Άρα σε επίπεδο σημαντικότητας 5%, δεν ισχύει η ανισότητα του στατιστικού κριτηρίου και συνεπώς δεν μπορούμε να απορρίψουμε την H_{0AB}. Επομένως δεν υπάρχει σημαντική αλληλεπίδραση μεταξύ των παραγόντων του πειράματος, δηλαδή μεταξύ ποικιλίας και λίπανσης.

7. Εξετάστηκε η επίδραση τριών σιτηρεσίων (Α) σε προβατίνες των φυλών Χίου και Καραγκούνικης (Β) κατά τη διάρκεια της κυοφορίας, στο ύψος της γαλακτοπαραγωγής μετά τον τοκετό. Παρακάτω δίνεται η γαλακτοπαραγωγή (Kg/24h):

	Σιτηρέσιο (A_1)	Σιτηρέσιο (A_2)	Σιτηρέσιο (A_3)
Χίου (B_1)	2.16 2.41 2.01 1.79 1.40 0.90 1.05	1.47 1.62 1.74 1.09 0.98 1.00 1.17	1.75 1.88 1.66 1.34 1.43 1.25 1.04
Καραγκούνικη (B_2)	1.05 0.61 0.55 1.73 0.96 0.95 1.06	0.85 0.49 0.46 1.07 0.73 0.77 0.52	0.63 0.78 0.75 0.55 0.77 0.60 0.88

Αφού διατυπώσετε κατάλληλους ελέγχους υποθέσεων, διερευνήστε σε επίπεδο σημαντικότητας 5% τις κύριες επιδράσεις των δύο παραγόντων του πειράματος, καθώς και την μεταξύ τους αλληλεπίδραση. (Δίνονται: SSA = 0.83, SSB = 4.92, SSAB = 0.03, SST = 10.16).

Πρόκειται για ένα πρόβλημα *ανάλυσης διασποράς με δύο παράγοντες (Α: σιτηρέσιο και Β: φυλή) και έλεγχο σημαντικότητας της αλληλεπίδρασης μεταξύ των δύο παραγόντων.*
Διατυπώνουμε τις στατιστικές υποθέσεις:

$H_{0A}: \alpha_1 = \alpha_2 = \alpha_3$ (η μέση γαλακτοπαραγωγή *δεν διαφοροποιείται* ανάλογα με το σιτηρέσιο ή ο παράγοντας "σιτηρέσιο" *δεν επιδρά* στη γαλακτοπαραγωγή)

H_{1A}: τουλάχιστον κάποιο $\alpha_i \neq \alpha_j$, $i, j = 1, 2, 3$ (η μέση γαλακτοπαραγωγή *διαφοροποιείται* ανάλογα με το σιτηρέσιο γαλακτοπαραγωγή ή ο παράγοντας "σιτηρέσιο" *επιδρά* στη γαλακτοπαραγωγή)

$H_{0B}: \beta_1 = \beta_2$ (η μέση γαλακτοπαραγωγή *δεν διαφοροποιείται* ανάλογα με τη φυλή ή ο παράγοντας "φυλή" *δεν επιδρά* στη γαλακτοπαραγωγή)

$H_{1B}: \beta_1 \neq \beta_2$ (η μέση γαλακτοπαραγωγή *διαφοροποιείται* ανάλογα με τη φυλή ή ο παράγοντας "φυλή" *επιδρά* στη γαλακτοπαραγωγή)

$H_{0AB}: (\alpha\beta)_{ij} = 0$ (Δεν υπάρχει αλληλεπίδραση)

$H_{1AB}: (\alpha\beta)_{ij} \neq 0$ (Υπάρχει αλληλεπίδραση)

Αντικαθιστώντας στον πίνακα ανάλυσης διασποράς για δύο παράγοντες με έλεγχο ύπαρξης αλληλεπίδρασης έχουμε:

Πηγή μεταβολής	Αθροίσματα τετραγώνων	Βαθμοί ελευθερίας	Μέσα τετράγωνα	Κριτήρια F
Παράγοντας Α (σιτηρέσιο)	$SSA = \lambda r \sum_{i=1}^{\kappa} (\bar{y}_{i\blacksquare\blacksquare} - \bar{y}_{\blacksquare\blacksquare\blacksquare})^2 = 0.83$	$\kappa - 1 =$ $3 - 1 = 2$	$MSA = \frac{SSA}{\kappa - 1} =$ $\frac{0.83}{2} = 0.415$	$F_A = \frac{MSA}{MSE} =$ $\frac{0.415}{0.12} = 3.46$
Παράγοντας Β (φυλή)	$SSB = \kappa r \sum_{i=1}^{\kappa} (\bar{y}_{\blacksquare j\blacksquare} - \bar{y}_{\blacksquare\blacksquare\blacksquare})^2 = 4.92$	$\lambda - 1 =$ $2 - 1 = 1$	$MSB = \frac{SSB}{\lambda - 1} =$ $\frac{4.92}{1} = 4.92$	$F_B = \frac{MSB}{MSE} =$ $\frac{4.92}{0.12} = 41.00$
Αλληλεπί-δραση Α×Β	$SSAB =$ $r \sum_{i=1}^{\kappa} \sum_{j=1}^{\lambda} (\bar{y}_{ij\blacksquare} - \bar{y}_{i\blacksquare\blacksquare} - \bar{y}_{\blacksquare j\blacksquare} - \bar{y}_{\blacksquare\blacksquare\blacksquare})^2$ $= 0.03$	$(\kappa - 1)(\lambda - 1)$ $(3 - 1)(2 - 1)$ $= 2 \cdot 1 = 2$	$MSAB = \frac{SSAB}{(\kappa - 1)(\lambda - 1)}$ $= \frac{0.03}{2} = 0.015$	$F_{AB} = \frac{MSAB}{MSE}$ $= \frac{0.015}{0.12} = 0.13$
Σφάλματα ή υπόλοιπα	$SSE = \sum_{i=1}^{\kappa} \sum_{j=1}^{\lambda} \sum_{\mu=1}^{r} (y_{ij\mu} - \bar{y}_{ij\blacksquare})^2 =$ $SSE = SST - SSA - SSB - SSAB =$ $= 10.16 - 0.83 - 4.92 - 0.03 = 4.38$	$\kappa\lambda(r - 1) =$ $3 \cdot 2 \cdot (7 - 1) =$ 36	$MSE = \frac{SSE}{\kappa\lambda(r - 1)} =$ $\frac{4.38}{36} = 0.12$	
Ολική	$SST = \sum_{i=1}^{\kappa} \sum_{j=1}^{\lambda} \sum_{\mu=1}^{r} (y_{ij\mu} - \bar{y}_{\blacksquare\blacksquare\blacksquare})^2$ $= 10.16$	$\kappa\lambda r - 1 =$ $3 \cdot 2 \cdot 7 - 1 =$ 41		

Στατιστικό κριτήριο ελέγχου της σημαντικότητας του παράγοντα Α:

Εάν

$F_A = \frac{MSA}{MSE} > F_{\kappa-1,\kappa\lambda(r-1),\alpha}$ απορρίπτεται η H_{0A}.

Αντικαθιστώντας έχουμε:

$$F_A = \frac{MSA}{MSE} = 3.46 > F_{\kappa-1,\kappa\lambda(r-1),\alpha} = F_{3-1,3\cdot 2(7-1),0.05} = F_{2,36,0.05} = 3.23$$

Άρα σε επίπεδο σημαντικότητας 5%, ισχύει η ανισότητα του στατιστικού κριτηρίου και συνεπώς απορρίπτουμε την H_{0A}. Επομένως η μέση γαλακτοπαραγωγή διαφοροποιείται ανάλογα με το σιτηρέσιο ή αλλιώς μπορούμε να πούμε ότι ο παράγοντας "σιτηρέσιο" επιδρά στη γαλακτοπαραγωγή ή ότι είναι στατιστικά σημαντικός σε επίπεδο σημαντικότητας 5%.

Στατιστικό κριτήριο ελέγχου της σημαντικότητας του παράγοντα Β:
Εάν

$F_B = \frac{MSB}{MSE} > F_{\lambda-1,\kappa\lambda(r-1),\alpha}$ απορρίπτεται η H_{0B}.

Αντικαθιστώντας έχουμε:

$$F_B = \frac{MSB}{MSE} = 41.00 > F_{\lambda-1,\kappa\lambda(r-1),\alpha} = F_{2-1,3\cdot2(7-1),0.05} = F_{1,36,0.05} = 4.08$$

Άρα σε επίπεδο σημαντικότητας 5%, ισχύει η ανισότητα του στατιστικού κριτηρίου και συνεπώς απορρίπτουμε την H_{0B}. Επομένως η μέση γαλακτοπαραγωγή διαφοροποιείται ανάλογα με τη φυλή ή αλλιώς μπορούμε να πούμε ότι ο παράγοντας “φυλή” επιδρά στη γαλακτοπαραγωγή ή ότι είναι στατιστικά σημαντικός σε επίπεδο σημαντικότητας 5%.

Στατιστικό κριτήριο ελέγχου της σημαντικότητας της αλληλεπίδρασης των δύο παραγόντων:
Εάν

$F_{AB} = \frac{MSAB}{MSE} > F_{(\kappa-1)(\lambda-1),\kappa\lambda(r-1),\alpha}$ απορρίπτεται η H_{0AB}.

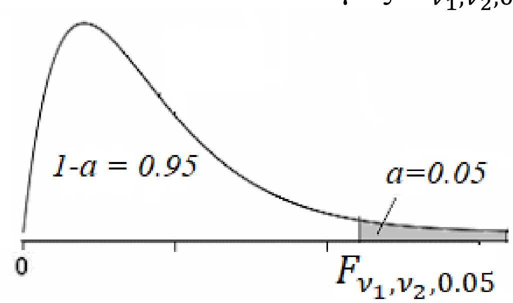
Αντικαθιστώντας έχουμε:

$$F_{AB} = \frac{MSAB}{MSE} = 0.13 < F_{(\kappa-1)(\lambda-1),\kappa\lambda(r-1),\alpha} = F_{(3-1)(2-1),3\cdot2(7-1),0.05} = F_{2,36,0.05} = 3.23$$

Άρα σε επίπεδο σημαντικότητας 5%, δεν ισχύει η ανισότητα του στατιστικού κριτηρίου και συνεπώς δεν μπορούμε να απορρίψουμε την H_{0AB}. Επομένως δεν υπάρχει σημαντική αλληλεπίδραση μεταξύ των παραγόντων του πειράματος, δηλαδή μεταξύ σιτηρεσίου και φυλής.

Κατανομή F

Τιμές $F_{\nu_1,\nu_2,0.05}$



Παράδειγμα: $F_{3,6,0.05} = 4.76$

$\nu_2 \backslash \nu_1$	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.3	254.3
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

Adapted from E. S. Pearson and H. O. Hartley, *Biometrika Tables for Statisticians*, Vol. 1, 1958, pp. 157–63, Table 18, by permission of the Biometrika Trustees.