"12º μάθημα" 20/12/21

$$Y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij} \quad \text{με } i=1,\dots,I \quad j=1,\dots,J$$

$$\underbrace{\sum_{i=1}^{I}\sum_{j=1}^{J}(Y_{ij}-\bar{Y}_{\cdot\cdot})^2}_{SS_{tot}} = \underbrace{J\cdot\sum_{i=1}^{I}(\bar{Y}_{i\cdot}-\bar{Y}_{\cdot\cdot})^2}_{SS_A} + \underbrace{I\cdot\sum_{j=1}^{J}(\bar{Y}_{\cdot j}-\bar{Y}_{\cdot\cdot})^2}_{SS_B} + \underbrace{\sum_{i=1}^{I}\sum_{j=1}^{J}(\bar{Y}_{i\cdot}-\bar{Y}_{\cdot j}+\bar{Y}_{\cdot\cdot})^2}_{SS_{res}}$$

$$e_{ij} = Y_{ij} - \hat{Y}_{ij} = Y_{ij} - (\hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j) =$$

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

$$= Y_{ij} - \bar{Y}_{i\cdot} - \bar{Y}_{\cdot j} + \bar{Y}_{\cdot\cdot}$$

Πίνακας ΑΝΑΔΙΑ του μοντέλου Αναλ. Διακύμανσης 2 παράγ.

Πηγή Μετ.	SS	Β.Ε.	MS	F-πηλίκα
Παράγ. Α.	SS_A	$I-1$	$MS_A = \frac{SS_A}{I-1}$	$F_A = \frac{MS_A}{MS_{res}}$
Παράγ. Β	SS_B	$J-1$	$MS_B = \frac{SS_B}{J-1}$	$F_B = \frac{MS_B}{MS_{res}}$
Υπόλοιπα	SS_{res}	$(I-1)(J-1)$	$MS_{res} = \frac{SS_{res}}{(I-1)(J-1)}$	
Ολική	SS_{tot}	$IJ-1$		

Πρόταση: Υπό τις υποθέσεις για τα σφάλματα και τις πλευρικές συνθήκες (α) $E(MS_A) = \sigma^2 + \frac{J}{I-1}\sum_{i=1}^{I} a_i^2$

(β) $E(MS_B) = \sigma^2 + \frac{I}{J-1}\sum_{j=1}^{J} \beta_j^2$

(γ) $E(MS_{res}) = \sigma^2$

Απόδειξη (αλγεβρική):

(α) $E(SS_A) = E\left\{ J\sum_{i=1}^{I}(\overline{Y}_{i\cdot}-\overline{Y}_{\cdot\cdot})^2 \right\} = J\cdot\sum_{i=1}^{I} E(\overline{Y}_{i\cdot}-\overline{Y}_{\cdot\cdot})^2 =$

$= J\cdot\sum_{i=1}^{I}\left\{ Var(\overline{Y}_{i\cdot}-\overline{Y}_{\cdot\cdot}) + (E(\overline{Y}_{i\cdot}-\overline{Y}_{\cdot\cdot}))^2 \right\} =$

$= J\cdot\sum_{i=1}^{I} Var(\overline{Y}_{i\cdot}-\overline{Y}_{\cdot\cdot}) + J\cdot\sum_{i=1}^{I}\left\{ E(\overline{Y}_{i\cdot}-\overline{Y}_{\cdot\cdot})\right\}^2$

$= J\cdot\sum_{i=1}^{I} Var(\overline{Y}_{i\cdot}-\overline{Y}_{\cdot\cdot}) + J\cdot\sum_{i=1}^{I} a_i^2 \quad (1)$

$\overline{Y}_{i\cdot}-\overline{Y}_{\cdot\cdot} = \overline{Y}_{i\cdot} + \frac{1}{I}\sum_{i=1}^{I}\overline{Y}_{i\cdot} = \frac{-1}{I}\overline{Y}_{i\cdot} - \dots - \frac{1}{I}\overline{Y}_{(i-1)\cdot} +$

$+ \left(1-\frac{1}{I}\right)\overline{Y}_{i\cdot} - \frac{1}{I}\overline{Y}_{(i+1)\cdot} - \dots - \frac{1}{I}\overline{Y}_{I\cdot}$

Άρα $Var(\overline{Y}_{i\cdot}-\overline{Y}_{\cdot\cdot}) \underset{Y_{ij}}{\overset{\alpha\nu\varepsilon\xi}{=}} \left(\frac{-1}{I}\right)^2 \cdot Var\,\overline{Y}_{\cdot\cdot} + \dots + \left(\frac{-1}{I}\right)^2 \cdot Var(\overline{Y}_{(i-1)}$

$+ \left(\frac{I-1}{I}\right)^2 \cdot Var(\overline{Y}_{i\cdot}) + \dots + \left(\frac{-1}{I}\right)^2 \cdot Var(Y_{I\cdot}) \quad (2)$

$Var(\overline{Y}_{i\cdot}) = Var\left(\frac{1}{J}\cdot\sum_{j=1}^{J} Y_{ij}\right) \underset{\alpha\nu\varepsilon\xi}{\overset{Y_{ij}}{=}} \frac{1}{J^2}\cdot\sum_{j=1}^{J} Var\,Y_{ij} =$

$= \frac{1}{J^2}\cdot\sum_{j=1}^{J}\sigma^2 = \frac{\sigma^2}{J} \quad (3)$

Από (2),(3):

$$Var(\overline{Y}_{i\cdot}-\overline{Y}_{\cdot\cdot})^2=(I-1)\cdot\frac{1}{I^2}\cdot\frac{\sigma^2}{J}+\frac{(I-1)^2}{I^2}\cdot\frac{\sigma^2}{J}+\frac{(I-1)^2}{I^2}\cdot\frac{\sigma^2}{J}=$$

$$=\frac{I-1}{J\cdot I}\quad(4)$$

Από (1),(4):

$$E(SS_A)=\cancel{J}\cdot\sum_{i=1}^{I}\frac{I-1}{I\cancel{J}}\sigma^2+J\cdot\sum_{i=1}^{I}a_i^2=(I-1)\sigma^2+J\sum a_i^2$$

Άρα $E(MS_A)=\dfrac{E(SS_A)}{I-1}=\sigma^2+\dfrac{J}{I-1}\displaystyle\sum_{i=1}^{I}a_i^2$

(Μπορεί να μπει κ' ως θέμα μια τέτοια απόδειξη)

$$E(SS_{tot})=E\left(\sum_{i=1}^{I}\sum_{j=1}^{J}(Y_{ij}-\overline{Y}_{\cdot\cdot})^2\right)=\sum_i\sum_j E(Y_{ij}-\overline{Y}_{\cdot\cdot})^2=$$

$$=\sum_i\sum_j\left\{Var(Y_{ij}-\overline{Y}_{\cdot\cdot})+\left[E(Y_{ij}-\overline{Y}_{\cdot\cdot})\right]^2\right\}=$$

$$=J\cdot\sum_i a_i^2+I\cdot\sum_j b_j^2+(IJ-1)\sigma^2$$

$$E(SS_{res})=E(SS_{tot})-E(SS_A)-E(SS_B)=$$
$$=(I-1)(J-1)\sigma^2$$

$$E(MS_{res})=\frac{E(SS_{res})}{(I-1)(J-1)}=\sigma^2$$

Έλεγχος Υποθέσεων

$H_0^A: \alpha_1 = \alpha_2 = \ldots = \alpha_I$ και $H_0^B: \beta_1 = \beta_2 = \ldots = \beta_J$

Τα επίπεδα του A είναι το ίδιο σημαντικά, ισοεπιδραίν στην Y.
Τα επίπεδα του B είναι το ίδιο σημαντικά, ισοεπιδραίν στην Y.

$H_0^A: \alpha_1 = \alpha_2 = \ldots = \alpha_I = 0$

$H_0^B: \beta_1 = \beta_2 = \ldots = \beta_J = 0$

Υπό την H_0^A ισχύει: $E(MSA) = E(MS_{res})$	Υπό την H_0^B ισχύει: $E(MS_B) = E(MS_{res})$
Άρα ένα τεστ για έλεγχο της H_0^A θα στηριχτεί στο MSA και MS_{res}	— " — της H_0^B — " — στο MS_B και MS_{res}

<u>Πρόταση</u>: Υπό τις υποθέσεις για τα σφάλματα και τις πλευρικές συνθήκες

(a) $\dfrac{SS_{res}}{\sigma^2} \sim X^2_{(I-1)(J-1)}$

(β) Όταν η $H_0^A: a_1 = a_2 = \ldots = a_I = 0$ είναι αληθής τότε $\dfrac{SSA}{\sigma^2} \sim X^2_{(I-1)}$

(γ) Όταν η $H_0^B: \beta_1 = \beta_2 = \ldots = \beta_J = 0$ είναι αληθής τότε $\dfrac{SSB}{\sigma^2} \sim X^2_{(J-1)}$

Απόδειξη: $SSres = \sum_i \sum_j (Y_{ij} - \bar{Y}_{i\cdot} - \bar{Y}_{\cdot j} + \bar{Y}_{\cdot\cdot})^2$

$$\bar{Y}_{\cdot\cdot} = \frac{1}{I} \cdot \sum_{i=1}^{I} \bar{Y}_{i\cdot} = \frac{1}{J} \cdot \sum_{j=1}^{J} \bar{Y}_{\cdot j}$$

$$Y_{ij} = \mu + a_i + b_j + \varepsilon_{ij} \underset{\text{Normal}}{\overset{\varepsilon_{ij}}{\sim}} N(\mu + a_i + b_j + \sigma^2)$$

$$\frac{Y_{ij} - \mu - a_i - b_j}{\sigma} \sim N(0,1) \Rightarrow \frac{(Y_{ij} - \mu - a_i - b_j)^2}{\sigma^2} \sim X^2_1$$

$$\underset{\text{ανεξ}}{\overset{Y_{ij}}{\Rightarrow}} \sum_{i=1}^{I} \frac{(Y_{ij} - \mu - a_i - b_j)^2}{\sigma^2} \sim X^2_I \Rightarrow$$

$$\frac{\sum_{i=1}^{I} (Y_{ij} - \mu - a_i - b_j)^2}{\sigma^2} \sim X^2_I$$

$$\hat{\mu} = \bar{Y}_{\cdot\cdot}, \quad \hat{a}_i = \bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot}, \quad \hat{b}_j = \bar{Y}_{\cdot j} - \bar{Y}_{\cdot\cdot}$$

$$\boxed{\frac{1}{\sigma^2} \cdot \sum_{i=1}^{I} (Y_{ij} - \bar{Y}_{i\cdot} - \bar{Y}_{\cdot j} + \bar{Y}_{\cdot\cdot})^2 \sim X^2_{I-1}}$$

Αναλόγως: $\boxed{\frac{1}{\sigma^2} \cdot \sum_{j=1}^{J} (Y_{ij} - \bar{Y}_{i\cdot} - \bar{Y}_{\cdot j} + \bar{Y}_{\cdot\cdot})^2 \sim X^2_{J-1}}$

Λόγω ανεξαρτησίας Y_{ij}:

$$\boxed{\begin{gathered} \frac{1}{\sigma^2} \sum_{i=1}^{I} \sum_{j=1}^{J} (Y_{ij} - \bar{Y}_{i\cdot} - \bar{Y}_{\cdot j} + \bar{Y}_{\cdot\cdot})^2 \sim X^2_{(I-1)(J-1)} \\ \Downarrow \\ \frac{SSres}{\sigma^2} \sim X^2_{(I-1)(J-1)} \end{gathered}}$$

β) $SS_A = J \cdot \sum_{i=1}^{I} (\overline{Y}_{i\circ} - \overline{Y}_{\circ\circ})^2$

Αφού $Y_{ij} \sim N(\mu + \alpha_i + \beta_j, \sigma^2)$ με $j = 1, \dots, J$

$$\Downarrow$$

$$\overline{Y}_{i\circ} \sim N\left(\mu + \alpha_i + \cancel{\sum_{j=1}^{J} \beta_j}^{\,0}, \frac{\sigma^2}{J}\right)$$

Άρα, υπό την H_0^A: $\alpha_1 = \dots = \alpha_I = 0$

Επομένως: $\overline{Y}_{i\circ} \sim N\left(\mu, \frac{\sigma^2}{J}\right) \xRightarrow[\text{μετχ}]{\text{τυπ}} \frac{\overline{Y}_{i\circ} - \mu}{\sqrt{\frac{\sigma^2}{J}}} \sim N(0,1)$

$\Rightarrow \frac{J(\overline{Y}_{i\circ} - \mu)^2}{\sigma^2} \sim X_1^2 \xRightarrow[\text{ανεξ.}]{Y_{ij}} \frac{J \cdot \sum_{i=1}^{I} (\overline{Y}_{i\circ} - \mu)}{\sigma^2} \sim X_I^2$

$$\Rightarrow \boxed{\frac{J \cdot \sum_{i=1}^{I} (\overline{Y}_{i\circ} - \overline{Y}_{\circ\circ})^2}{\sigma^2} \sim X_{I-1}^2}$$

γ) Εντελώς ανάλογη με αυτή της β).

Για τον έλεγχο της H_0^A: $\alpha_1 = \dots = \alpha_I$, η ΣΣΤ
$F_A = \frac{MSA}{MSres} \sim F_{I-1, (I-1)(J-1)}$ υπό την H_0^A και

κ.π. : $F_A \geq F_{I-1, (I-1)(J-1) a}$

Για τον έλεγχο της H_0^B: $\beta_1 = \dots = \beta_J$, η ΣΣΤ
$F_B = \frac{MSB}{MSres} \sim F_{J-1, (I-1)(J-1)}$ υπο την H_0^B και

κ.π : $F_B \geq F_{J-1, (I-1)(J-1), a}$

Αν οι H_0^A ή/και η H_0^B απορριφθεί, τότε
→ Πολλαπλές Συγκρίσεις → Ελαχ. Σημαντική Διαφορά
→ Γραμμικοί Αντιθέσεων

<u>ΕΣΔ</u>: Αν η H_0^A απορριφθεί, τότε για τον έλεγχο της $H_0: \alpha_i = \alpha_{i'}$ έναντι $H_A: \alpha_i \neq \alpha_{i'}$ με $i, i' = 1, \ldots, I$, με $i \neq i'$, η H_0 απορρίπτεται αν: $|\bar{Y}_{i\cdot} - \bar{Y}_{i'\cdot}| \geq ΕΣΔ \overset{op}{=} t_{(I-1)(J-1), \frac{a}{2}} \sqrt{}$

Αν η $H_0: \alpha_i = \alpha_{i'}$ απορριφθεί, τότε:

- Αν $\bar{Y}_{i\cdot} - \bar{Y}_{i'\cdot} > 0$, τότε το i-επίπεδο είναι σημαντικότερο από το i'-επίπεδο.

- Αν $\bar{Y}_{i\cdot} - \bar{Y}_{i'\cdot} < 0$, τότε το i'-επίπεδο είναι σημαντικότερο από το i-επίπεδο.

Αν η H_0^B απορριφθεί, τότε για τον έλεγχο της $H_0: \beta_j = \beta_{j'}$ έναντι της $H_a: \beta_j \neq \beta_{j'}$ με $j, j' = 1, \ldots, J$ και $j \neq j'$ η H_0 απορρίπτεται αν:

$$|\bar{Y}_{\cdot j} - \bar{Y}_{\cdot j'}| \geq ΕΣΔ \overset{op}{=} t_{(I-1)(J-1), \frac{a}{2}} \sqrt{\frac{MS_{res}}{I}}$$

Αν η H_0 απορριφθεί, τότε:

- Αν $\bar{Y}_{\cdot j} - \bar{Y}_{\cdot j'} > 0$, το j-επίπεδο σημαντικότερο του j'-επιπέδου.

- Αν $\bar{Y}_{\cdot j} - \bar{Y}_{\cdot j'} < 0$, το j'-επίπεδο σημαντικότερο του j-επιπέδου.

Μέθοδος Γραμμικών Αντιθέσεων ή Scheffe

Έστω γραμμική αντίθεση $L_A = \sum_{i=1}^{I} c_i a_i$ με $\sum_{i=1}^{I} c_i = 0$

Για τον έλεγχο της $H_0: L_A = 0$, η ΣΣΤ:

$$F_{L_A} = \frac{MS_{L_A}}{MS_{res}}, \quad MS_{L_A} \overset{\text{ορ}}{=} \frac{\hat{L}_A^2}{\frac{1}{J}\sum_{i=1}^{I} c_i^2}, \quad \hat{L}_A = \sum_{i=1}^{I} c_i \bar{Y}_{i\cdot}$$

με κατανομή $F_{1,(I-1)(J-1)}$ υπό την $H_0: L_A = 0$
και Κ.Π. $f_{L_A} \geq F_{1,(I-1)(J-1),\alpha}$

Έστω γραμμική αντίθεση $L_B = \sum_{j=1}^{J} c_j b_j$ με $\sum_{j=1}^{J} c_j = 0$

Για τον έλεγχο της $H_0: L_B = 0$ η ΣΣΤ είναι

$$F_{L_B} = \frac{MS_{L_B}}{MS_{res}}, \quad MS_{L_B} \overset{\text{ορ}}{=} \frac{\hat{L}_B^2}{\frac{1}{I}\sum_{j=1}^{J} c_j^2}, \quad \hat{L}_B = \sum_{j=1}^{J} c_j \bar{Y}_{\cdot j}$$

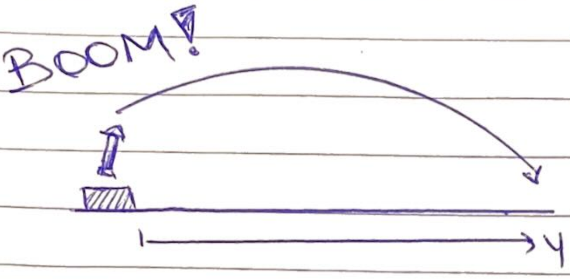
με κατανομή $F_{1,(I-1)(J-1)}$ υπό την $H_0: L_B = 0$
και Κ.Π. $f_{L_B} \geq F_{1,(I-1)(J-1),a}$

Άσκηση: (Παρόμοια και σε εξετάσεις)

4 διαφορετικά καύσιμα
3 διαφορετικοί τρόποι πάνω στο βεληνικές μιας ρουκέτας. Τα δεδομένα συγκεντρωτικά, είναι:

τρόπος εκτόξευσης \ καύσιμα	1	2	3	4
1	45.9	57.6	52.2	41.7
2	46	51	50.1	38.8
3	45.7	56.9	55.3	48.1

Δεδομένα → Βελτιωτικές σε Km



Y → Βελτιωτικές → Εξαρτ. Μεταβλητή

Τρόπος Εκτόξευσης ← 3 επίπεδα ← Παράγοντας A

Καύσιμο ← 4 επίπεδα ← Παράγοντας B

Μοντέλο Ανάλ/σια κατα 2 παράγοντες

$$Y_{ij} = \mu + a_i + b_j + \varepsilon_{ij} \quad \text{με} \quad \underbrace{i = 1,2,3}_{\text{παραγ. A}} \quad \underbrace{j = 1,2,3,4}_{\text{παραγ. B}}$$

- Υπο τις συνθήκες για τα σφάλματα
- Πλευρικές συνθήκες

Πηγή Μεταβλητότητας	SS	β.ε	MS	F-πηλίκα
Πηγή A	50,852	2	25,426	$F_A = 4,428$
Πηγή B	293,702	3	97,9	$F_B = 17,048$
Υπόλοιπα	34,455	6	5,743	
Ολική	379,009	11		

Έχει ουσία (δηλ. ξεφεύγει)

Πίνακας Αναλ/σια

Έλεγχος : $H_0^A: a_1 = a_2 = a_3$

Επειδή $F_A = 4{,}428 < F_{2,6,0.05} = 5{,}14$
$F_{2,6,0.01} = 10{,}9$

Άρα η H_0^A δεν μπορεί να απορριφθεί. Επομένως ο παράγοντας A δεν είναι σημαντικός.

Έλεγχος : $H_0^B: b_1 = b_2 = b_3 = b_4$

Επειδή $F_B = 17{,}048 > F_{3,6,0.05} = 4{,}76$
$> F_{3,6,0.01} = 9{,}78$ } η H_0^B απορρίπτεται
↓
ΕΣΔ

(Προχωράμε δηλαδή σε πολλαπλές συγκρίσεις).

ΕΣΔ για τα επίπεδα b_j με $j = 1, \ldots, 4$ του παραγ. B

$$ΕΣΔ = t_{(I-1)(J-1), \frac{\alpha}{2}} \sqrt{\frac{2MSres}{I}} = 4{,}793$$

$|\bar{Y}_{\cdot 1} - \bar{Y}_{\cdot 2}| = |-9{,}3| = 9{,}3 > 4{,}793$, η $b_1 = b_2$ απορ/ται

Επειδή $\bar{Y}_{\cdot 1} - \bar{Y}_{\cdot 2} < 0 \Rightarrow b_2$ σημαντικότερο του b_1

$|\bar{Y}_{\cdot 1} - \bar{Y}_{\cdot 3}| = |-6{,}66| > ΕΣΔ$, η $b_1 = b_3$ απορ/ται

Επειδή $\bar{Y}_{\cdot 1} - \bar{Y}_{\cdot 3} < 0 \Rightarrow b_3$ σημαντικότερο του b_1

Κ.Ο.Κ : $b_1 = b_4$
$b_2 = b_3$
$b_2, b_3 > b_4$
$b_2, b_3 > b_1$

– Καλές Γιορτές – (Τέλος Ύλης)