Δευτέρα 21/12/2020

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

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

$\sum a_i = \sum \beta_j = 0$ | Αμεροληψία ΕΕΤ.
Υποθέσεις για ε_{ij}. | ΕΕΤ = ΕΜΠ.

Θεώρημα:

$\pm \bar{Y}_{i.}, \bar{Y}_{.j}$

$$\underbrace{\sum_{i=1}^{I}\sum_{j=1}^{J}(Y_{ij} - \bar{Y}_{..})^2}_{SStot} = \underbrace{J\sum_{i=1}^{I}(\bar{Y}_{i.} - \bar{Y}_{..})^2}_{SS_A} + \underbrace{I\sum_{j=1}^{J}(\bar{Y}_{.j} - \bar{Y}_{..})^2}_{SS_B} + \underbrace{\sum\sum(Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..})^2}_{SSres}$$

Παρατήρηση: (Ελέγχουμε αν το SSres ορίζεται όπως στο υπόδειγμα.)

$$e_{ij} = Y_{ij} - \hat{Y}_{ij} = Y_{ij} - (\hat{\mu} + \hat{a}_i + \hat{\beta}_j) = Y_{ij} - (\cancel{\bar{Y}_{..}} + \bar{Y}_{i.} - \cancel{\bar{Y}_{..}} + \bar{Y}_{.j} - \bar{Y}_{..})$$

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

ΠΙΝΑΚΑΣ ΑΝΑΔΙΑ / ΚΑΤΑ ΔΥΟ ΠΑΡΑΓΟΝΤΕΣ

ελέγχει αν έχει νόημα το μοντέλο αυτό

Πηγή μεταβλητότητας	SS.	Β.Ε	MS	F-πηλίκο
Παράγοντας Α	$SSA=J\sum_{i=1}^{I}(\bar{Y}_{i\cdot}-\bar{Y}_{\cdot\cdot})^2$	$I-1$	$SSA/I-1$	$F_A=\frac{MSA}{MSres}$
Παράγοντας Β	$SS_B=I\sum_{j=1}^{J}(\bar{Y}_{\cdot j}-\bar{Y}_{\cdot\cdot})^2$	$J-1$	$SSB/J-1$	
Υπόλοιπα	$SSres=\sum_i\sum_j(Y_{ij}-\bar{Y}_{i\cdot}-\bar{Y}_{\cdot j}+\bar{Y}_{\cdot\cdot})^2$	$(I-1)(J-1)$	$SSres/(I-1)(J-1)$	$F_B=\frac{MSB}{MSres}$
Ολική μεταβλητότητα.	SStot.	$IJ-1$	//////	

αδιάλ αν ισοσταθμίζουν οι δυο παράγοντες

Πρόταση:

Υπό τις υποθέσεις για τα σφάλματα ($\varepsilon_{ij}\sim N(0,\sigma^2)$ και ασυσχέτιστα) και τις πλευρικές συνθήκες.

i) • $E(MSA)=\sigma^2+\frac{J}{I-1}\sum_{i=1}^{I}a_i^2$

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

iii) $E(MSres)=\sigma^2$.

Απόδειξη (i)

$MS_A=\frac{SSA}{I-1}$, $SSA=J\sum_{i=1}^{I}(\bar{Y}_{i\cdot}-\bar{Y}_{\cdot\cdot})^2.$

$E(SSA)=J\sum_{i=1}^{I}E(\bar{Y}_{i\cdot}-\bar{Y}_{\cdot\cdot})^2=J\sum_{i=1}^{I}[Var(\bar{Y}_{i\cdot}-\bar{Y}_{\cdot\cdot})+(E(\bar{Y}_{i\cdot}-\bar{Y}_{\cdot\cdot}))^2]$

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

Αλλά $\sum_{i=1}^{I}(E(\bar{Y}_{i\cdot}-\bar{Y}_{\cdot\cdot}))^2 = \sum(E(\hat{a}_i))^2 = \sum_{i=1}^{I} a_i^2$ (2).

$\bar{Y}_{i\cdot}-\bar{Y}_{\cdot\cdot} = \bar{Y}_{i\cdot} - \frac{1}{I}\sum_{i=1}^{I}\bar{Y}_{i\cdot} = \bar{Y}_{i\cdot} - \frac{1}{I}\bar{Y}_{1\cdot} - \dots - \frac{1}{I}\bar{Y}_{i\cdot} - \dots - \frac{1}{I}\bar{Y}_{I\cdot} =$

$= -\frac{1}{I}\bar{Y}_{1\cdot} - \dots - \left(1-\frac{1}{I}\right)\bar{Y}_{i\cdot} - - \frac{1}{I}\bar{Y}_{I\cdot}\Big)$

Άρα $Var(\bar{Y}_{i\cdot}) = Var\left(\frac{1}{J}\sum_{j=1}^{J} Y_{ij}\right) = \frac{1}{J^2}\sum_{j=1}^{J} Var(Y_{ij}) = \frac{1}{J^2}\sum_{j=1}^{J}\sigma^2 = \frac{\sigma^2}{J^2}$.

Επομένως $Var(\bar{Y}_{i\cdot}-\bar{Y}_{\cdot\cdot}) \underset{\text{Yi ανεξ.}}{=}$

$= \frac{1}{I^2}Var(\bar{Y}_{1\cdot}) + + \frac{(I-1)^2}{I^2}Var(\bar{Y}_{i\cdot}) + \quad + \frac{1}{I^2}Var(\bar{Y}_{I\cdot})$.

$\Rightarrow Var(\bar{Y}_{i\cdot}-\bar{Y}_{\cdot\cdot}) = \frac{\sigma^2}{J}\left\{\frac{I-1}{I^2}+\frac{(I-1)^2}{I}\right\} = \frac{I-1}{I}\cdot\frac{\sigma^2}{J}$ (3)

Από (1), (2), (3) τότε

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

$E(MSA) = \sigma^2 + \frac{J}{I-1}\sum a_i^2$.

γ) Κατ' αναλογία με 6)

$E(SStot) = E(\sum\sum(Y_{ij}-\bar{Y})^2) = \sum\sum E(Y_{ij}-\bar{Y}_{\cdot\cdot})^2 =$

$= \sum\sum[Var(Y_{ij}-\bar{Y}_{\cdot\cdot}) + (E(Y_{ij}-\bar{Y}_{\cdot\cdot}))^2] =$

$= J\sum a_i^2 + I\sum b_j^2 + (IJ-1)\sigma^2$ (4)

Από (α), (β), (γ) $E(SSres) = (I-1)(J-1)\sigma^2$.

Άρα $E(MSres) = \sigma^2$.

Έλεγχοι υποθέσεων σχετικά με την αλληλεπίδραση των επιπέδων των παραγόντων A και B στην εξαρτ. μεταβλητή Y

$H_0^A: a_1 = a_2 = \dots = a_I = 0$ και $H_0^B: b_1 = b_2 = \dots = b_J = 0$ {θέλω να απορ. τουλ. μια από αυτές τις υποθέσεις}

Ο έλεγχος της H_0^A στηρίζεται στην σύγκριση MSA με το MSres. (F_A)
» H_0^B » » MSB » (F_B)

Πρόταση:
Αν οι υποθέσεις για τα σφάλματα ικανοποιούνται και οι πλευρικές συνθήκες ισχύουν τότε α) $\frac{SSres}{\sigma^2} \sim \chi^2_{(I-1)(J-1)}$

β) Υπό την H_0^A: $\frac{SS_A}{\sigma^2} \sim \chi^2_{I-1}$

γ) Υπό την H_0^B: $\frac{SS_B}{\sigma^2} \sim \chi^2_{J-1}$

Απόδειξη:
$Y_{ij} \sim N(\mu + a_i + b_j, \sigma^2)$ και είναι ανεξ. ασυσχ. $i = 1 \dots I$, $j = 1 \dots J$.

$\frac{(Y_{ij} - \mu - a_i - b_j)^2}{\sigma^2} \sim N(0,1)^2 \equiv \chi_1^2 \Rightarrow \sum_{i=1}^{I} \frac{(Y_{ij} - \mu - a_i - b_j)^2}{\sigma^2} \sim \chi_I^2$

Αν αντικαταστήσω τα μ, a_i, b_j με τους ΕΕΤ $\hat{\mu}, \hat{a}_i, \hat{b}_j$ έχουμε $\sum_{i=1}^{I} \frac{(Y_{ij} - \hat{\mu} - \hat{a}_i - \hat{b}_j)^2}{\sigma^2} = \sum_{i=1}^{I} \frac{(Y_{ij} - \bar{Y}_{i\cdot} - \bar{Y}_{\cdot j} + \bar{Y}_{\cdot\cdot})^2}{\sigma^2} \sim \chi^2_{I-1}$

αφού $\bar{Y}_{\cdot\cdot} = \frac{1}{J} \bar{Y}_{i\cdot} = \frac{1}{I} \bar{Y}_{\cdot j}$ (υπάρχει μόνο μια σχέση ή ένας δεσμός μεταξύ τους)

Ανάλογα $\frac{\sum_{j=1}^{J}(Y_{ij}-\bar{Y}_{i\cdot}-\bar{Y}_{\cdot j}+\bar{Y}_{\cdot\cdot})^2}{\sigma^2} \sim \chi^2_{(J-1)}$.

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

β) $SSA = J\sum_{i=1}^{I}(\bar{Y}_{i\cdot}-\bar{Y}_{\cdot\cdot})^2$.

$\bar{Y}_{i\cdot} = \frac{1}{J}\sum_{j=1}^{J}Y_{ij} \sim \text{Normal}\left(\mu+a_i+\Sigma\beta_j, \frac{\sigma^2}{J}\right)$ $i=1,..I$

Λόγω της πλευρικής συνθήκης $\Sigma\beta_j=0$ και $H_0^A: a_1 = a_I = 0$
έχουμε $\bar{Y}_{i\cdot} \sim N(\mu, \frac{\sigma^2}{J})$.

$\Rightarrow \frac{\bar{Y}_{i\cdot}-\mu}{\sigma/\sqrt{J}} \sim N(0,1) \Rightarrow \frac{(\bar{Y}_{i\cdot}-\mu)^2}{\sigma^2/J} \sim N(0,1)^2 \Rightarrow$

$\Rightarrow \frac{J(\bar{Y}_{i\cdot}-\mu)^2}{\sigma^2} \sim \chi^2_1 \xRightarrow{\text{Y}_i \text{ ανεξ.}} \frac{J\sum_{i=1}^{I}(\bar{Y}_{i\cdot}-\mu)^2}{\sigma^2} \sim \chi^2_I$

$\Rightarrow \frac{J\Sigma(\bar{Y}_{i\cdot}-\bar{Y}_{\cdot\cdot})^2}{\sigma^2} \sim \chi^2_{I-1}$

F-τεστ για τον έλεγχο της H_0^A και H_0^B

Για τον έλεγχο της $H_0^A = a_1 = \ldots = a_I = 0$ χρησιμοποιείται η ΣΣΤ.
$F_A = \frac{MSA}{MSres} \sim F_{I-1,(I-1)(J-1)}$ υπό την H_0^A και κρίσιμη περιοχή
$F_A \geq F_{(I-1),(I-1)(J-1),a}$.

Για τον έλεγχο της H_0^B: $\beta_1 = \ldots = \beta_J = 0$ η ΣΣΤ είναι
$\frac{MS_B}{MSres} \sim F_{J-1,(I-1)(J-1)}$, με κ.π. $F_B \geq F_{(J-1),(I-1)(J-1),a}$

Στην πράξη αναμένουμε οι υποθέσεις αυτές να απορρίπτονται.
Αν απορρίπτονται σημαίνει ότι υπάρχουν επίπεδα των A και B που ασκούν σημαντικότερη επίδραση στην Y.
Ποια επίπεδα των A και B ασκούν πιο σημαντική επίδραση στην Y;
μπορούμε να το δούμε με τις πολλαπλές συγκρίσεις.

Αν τουλάχιστον μια από τις H_0^A και H_0^B απορριφθεί ⟶ Πολλαπλές συγκρίσεις.

Μέθοδος Ελάχιστης Σημαντικής διαφοράς (ΕΣΔ)

I) Για τον A. Αν η $H_0^A: a_1 = \dots = a_I = 0$ απορ. τότε ελέγχουμε ανά δύο
$H_0^A: a_i = a_{i'}$ έναντι $H_a^A: a_i \neq a_{i'}$, $i, i' = 1, \dots, I$, $i \neq i'$

Για τον έλεγχο χρησιμοποιείται η: $\dfrac{\bar{Y}_{i\cdot} - \bar{Y}_{i'\cdot}}{\sqrt{\dfrac{2MS_{res}}{J}}} \sim t_{(I-1)(J-1)}$

και κ.π. $|\bar{Y}_{i\cdot} - \bar{Y}_{i'\cdot}| > t_{(I-1)(J-1), \frac{a}{2}} \sqrt{\dfrac{2MS_{res}}{J}}$

Αν η H_0^A απορριφθεί ελέγχω το πρόσημο της διαφοράς $\bar{Y}_{i\cdot} - \bar{Y}_{i'\cdot}$
Αν $\bar{Y}_{i\cdot} - \bar{Y}_{i'\cdot} > 0$ τότε το a_i σημαντικότερο του $a_{i'}$
Αν $\bar{Y}_{i\cdot} - \bar{Y}_{i'\cdot} < 0$ τότε το $a_{i'}$ » » a_i

II) Για τον παράγοντα B
$H_0^B: \beta_j = \beta_{j'}$, $H_a^B: \beta_j \neq \beta_{j'}$, $j = 1 \dots J$, $j \neq j'$

Για τον έλεγχο χρησιμοποιείται η $\dfrac{\bar{Y}_{\cdot j} - \bar{Y}_{\cdot j'}}{\sqrt{\dfrac{2MS_{res}}{I}}} \sim t_{(I-1)(J-1)}$

κ.π. $|\bar{Y}_{\cdot j} - \bar{Y}_{\cdot j'}| > t_{(I-1)(J-1), \frac{a}{2}} \sqrt{\dfrac{2MS_{res}}{I}}$

$\bar{Y}_{\cdot j} - \bar{Y}_{\cdot j'} > 0$ πιο σημαντική επίδραση ασκεί το j,
$\bar{Y}_{\cdot j} - \bar{Y}_{\cdot j'} < 0$ " " " j'

Μέθοδος Γραμμικών Παινγραμμών / Scheffe.

Έστω $L_A: \sum_{i=1}^{I} c_i a_i$ με $\sum c_i = 0$ μια γραμμική αντίθεση στα επίπεδα του παράγοντα Α. Για τον έλεγχο της H_0^{LA} ότι δηλ. $L_A = 0$ χρησιμοποιείται η σ.σ.

$L_A = \frac{MSL_A}{MSres}$, όπου $MSL_A = \frac{\hat{L}_A^2}{\frac{1}{J}\sum c_i^2}$ και $L_A = \sum_{i=1}^{J} c_i \bar{Y}_{i\cdot}$ με

κατανομή την $F_{LA} \sim F_{1,(I-1)(J-1)}$ υπό την H_0^{LA} και κρ. περιοχή $F_{LA} \geq F_{1,(I-1)(J-1),a}$

Κατα αντιστοιχία για τον παράγοντα Β.

$L_B = \sum_{j=1} c_j b_j$ με $\sum c_j = 0$

$H_0^{LB}: L_B = 0$ παίρνουμε τη σσ

$L_B = \frac{MSL_B}{MSres}$ με $MSL_B = \frac{\hat{L}_B^2}{\frac{1}{I}\sum c_j}$ και $L_B = \sum_{j=1}^{J} c_j \bar{Y}_{\cdot j}$ με

κατανομή $F_{LB} \sim F_{1,(I-1)(J-1)}$ υπό την H_0^{LB} και κρ. περιοχή μεγέθους a: $F_{LB} \geq F_{1,(I-1)(J-1)a}$.

Ανάλυση υπολοίπων για να διερευνηθεί η ισχύς των υποθέσεων για τα σφάλματα όμοια με τα προηγούμενα.

Άσκηση: Σε ένα πείραμα θέλουμε να εξετάσουμε την επίδραση 4 διαφορετικών καυσίμων και 3 διαφορετικών τρόπων εκτόξευσης πάνω στο βεληνεκές κάποιου τύπου ρουκέτας. Τα δεδομένα που έχουμε στη διάθεση μας είναι τα ακόλουθα, όπου το βεληνεκές μετριέται σε km.

		Καύσιμα			
		1	2	3	4
Τρόπος Εκτόξευσης.	1	45.9	57,6	52.2	41,7
	2	46	51	50,1	38,6
	3	45,7	56,9	55,3	48,1