Υποθέσεις για τα σφάλματα

1) $E(\varepsilon_{ij})=0$
2) $Var(\varepsilon_{ij})=\sigma^2$
3) $Cov(\varepsilon_{ij},\varepsilon_{kl})=0$
4) $\varepsilon_{ij}\sim N(0,\sigma^2)$

Συνέπειες σφαλμάτων στο μοντέλο

1) $E(Y_{ij})=\mu+\alpha_i+\beta_j$
2) $Var(Y_{ij})=\sigma^2$
3) $Cov(Y_{ij}\ Y_{kl})=0$
4) $Y_{ij}\sim N(\mu+\alpha_i+\beta_j,\sigma^2)$

ΙΔΙΟΤΗΤΕΣ ΤΩΝ ΕΕΤ:

1) Οι ΕΕΤ εκτιμούν αμερόληπτα τις αντίστοιχες παραμέτρους
$E(\hat{\mu})=\mu$, $E(\hat{\alpha_i})=\alpha_i$, $E(\hat{b_j})=b_j$, $i=1,2,\dots,I$, $j=1,2,\dots,J$

Απόδειξη

$E(\hat{\alpha_i}) = E(\bar{Y}_{i.}-\bar{Y}_{..}) = E(\bar{Y}_{i.})-E(\bar{Y}_{..}) = E\left(\frac{1}{J}\sum Y_{ij}\right)-E\left(\frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J} Y_{ij}\right) =$

$= E\left(\frac{1}{J}\sum_j(\mu+\alpha_i+b_j+\varepsilon_{ij})\right) - E\left(\frac{1}{IJ}\sum_i\sum_j(\mu+\alpha_i+b_j+\varepsilon_{ij})\right) =$

$= \frac{1}{J}\sum_j(\mu+\alpha_i+b_j+\cancelto{0}{E(\varepsilon_{ij})}) - \frac{1}{IJ}\sum_i\sum_j(\mu+\alpha_i+b_j+\cancelto{0}{E(\varepsilon_{ij})}) =$

$= \frac{1}{J}\left(J\mu+J\alpha_i+\sum_j b_j\right) - \frac{1}{IJ}\left(IJ\mu+J\sum_i\alpha_i+I\sum_j b_j\right) =$

$= \mu+\alpha_i-\mu = \alpha_i$

2) Υπό τις υποθέσεις για τα σφάλματα , ισχύει ότι:
$EET(\mu,\alpha_i,b_j) = EM\Pi(\mu,\alpha_i,b_j)$

Απόδειξη

Η πιθανοφάνεια είναι:

$L = \prod_{i=1}^{I}\prod_{j=1}^{J} f_{Y_{ij}}(y_{ij}) = \prod_{i=1}^{I}\prod_{j=1}^{J}\frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2\sigma^2}(y_{ij}-\mu-\alpha_i-b_j)^2} =$

$= \frac{1}{\sigma^{IJ}(2\pi)^{-IJ/2}}\cdot e^{-\frac{1}{2\sigma^2}\sum\sum(y_{ij}-\mu-\alpha_i-b_j)^2}$

Έπειτα παίρνουμε το λογάριθμο της L και εκτελούμε τα ίδια που κάναμε σε προηγούμενα μαθήματα.

3) Ολική Μεταβλητότητα στο Μοντέλο Αναλ. διακ. κατά 2 παρ

Ολική Μεταβλητότητα $\equiv$ Δειγματική διακύμανση εκτός του όρου που σχετίζεται με το μέγεθος του δείγματος

Επομένως,

$SS_{tot} := \sum_{i=1}^{I}\sum_{j=1}^{J}(Y_{ij}-\bar{Y}_{..})^2 = \sum_{i=1}^{I}\sum_{j=1}^{J}(Y_{ij}\pm\bar{Y}_{i.}-\bar{Y}_{..}\pm\bar{Y}_{.j})^2 = \dots =$

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

Διαπιστωση του SS_{res}:

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

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

ΠΙΝΑΚΑΣ ΑΝΑΔΙΑ

Πηγή Μεταβλητότητας	SS	β.ε	MS	Επιπλέον
Παράγοντας A	SS_A	$I-1$	$SS_A/(I-1)$	$F_A = \frac{MS_A}{MS_{res}}$
Παράγοντας B	SS_B	$J-1$	$SS_B/(J-1)$	
Υπόλοιπα	SS_{res}	$(I-1)(J-1)$	$SS_{res}/(I-1)(J-1)$	$F_B = \frac{MS_B}{MS_{res}}$
Ολική Μεταβλητότητα	SS_{tot}	$IJ-1$		

Θεώρημα

Υπό τις υποθέσεις για τα σφάλματα και τις πλευρικές συνθήκες έχουμε τα εξής:

i) $E(MS_A) = \sigma^2 + \frac{J}{I-1} \sum_{i=1}^{I} \alpha_i^2$

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

iii) $E(MS_{res}) = \sigma^2$

Απόδειξη

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

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

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

$$= J \cdot \sum_{i=1}^{I} Var(\overline{Y_{i.}} - \overline{Y_{..}}) + J \sum_{i=1}^{I} (E(\hat{\alpha}_i))^2 =$$

$$= J \sum_{i=1}^{I} Var(\overline{Y_{i.}} - \overline{Y_{..}}) + J \sum_{i=1}^{I} \alpha_i^2 \quad (1)$$

$$\overline{Y_{i.}} - \overline{Y_{..}} = \overline{Y_{i.}} - \frac{1}{I} \sum_{i} \overline{Y_{i.}} = \overline{Y_{i.}} - \frac{1}{I} \overline{Y_{1.}} - \dots - \frac{1}{I} \overline{Y_{i.}} - \dots - \frac{1}{I} \overline{Y_{I.}} =$$

$$= -\frac{1}{I} \overline{Y_{1.}} - \dots - \left(1 - \frac{1}{I}\right) \overline{Y_{i.}} - \dots - \frac{1}{I} \overline{Y_{I.}} =$$

$$= -\frac{1}{I} \overline{Y_{1.}} - \dots - \frac{I-1}{I} \overline{Y_{i.}} - \dots - \frac{1}{I} \overline{Y_{I.}}$$

Αρα,

$$Var(\overline{Y_{i.}} - \overline{Y_{..}}) = Var\left(-\frac{1}{I} \overline{Y_{1.}} - \dots - \frac{I-1}{I} \overline{Y_{i.}} - \dots - \frac{1}{I} \overline{Y_{I.}}\right) =$$

$$\overset{Y_{ij} \text{ ανεξ}}{\underset{\overline{Y_{i.}} \text{ ανεξ.}}{=\!=\!=}} \frac{1}{I^2} Var(\overline{Y_{1.}}) + \dots + \frac{(I-1)^2}{I^2} Var(\overline{Y_{i.}}) + \dots + \frac{1}{I^2} Var(\overline{Y_{I.}}) \quad (2)$$

Τα δεδομένα της i γραμμής είναι:
$Y_{i1}, Y_{i2}, \dots, Y_{iJ}$ και λόγω των υποθέσεων των σφαλμάτων.
έστω $Y_{i1}, Y_{i2}, \dots, Y_{iJ} \sim N(\mu + \alpha_i + \beta_j, \sigma^2)$ ανεξάρτητες
Τότε, $\overline{Y_{i.}} \sim N\left(\mu + \alpha_i, \frac{\sigma^2}{J}\right)$, $\forall i = 1, \dots, I$
Αρα, $Var(\overline{Y_{i.}}) = \frac{\sigma^2}{J}$, $\forall i = 1, \dots, I$ (3)
Απο (2) + (3) $\Rightarrow$

$$\Rightarrow Var(\overline{Y_{i.}} - \overline{Y_{..}}) = \frac{1}{I^2} \frac{\sigma^2}{J} + \dots + \frac{(I-1)^2}{I^2} \cdot \frac{\sigma^2}{J} + \dots + \frac{1}{I^2} \cdot \frac{\sigma^2}{J} =$$

$$= \frac{(I-1)}{I^2 J} \sigma^2 + \frac{(I-1)^2}{I^2 J} \cdot \sigma^2 \Rightarrow Var(\overline{Y_{i.}} - \overline{Y_{..}}) = \frac{I-1}{I} \cdot \frac{\sigma^2}{J} \quad (4)$$

Απο την (1) + (4) $\Rightarrow$

$$\Rightarrow E(SSA) = J \cdot \sum_{i=1}^{I} \frac{I-1}{I} \frac{\sigma^2}{J} + J \cdot \sum_{i=1}^{I} \alpha_i^2 = JI \frac{I-1}{I} \frac{\sigma^2}{J} + J \sum_{i=1}^{I} \alpha_i^2 =$$

$$= (I-1)\sigma^2 + J \cdot \sum_{i=1}^{I} \alpha_i^2$$

Επομένως,

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

Στατιστικά Τεστ για έλεγχο τοοεπίδρασης των επιπέδων των παραγόντων A και B στην Y.

$H_0^A: \alpha_1 = \alpha_2 = \dots = \alpha_I \ (=0)$

$H_0^B: b_1 = b_2 = \dots = b_J \ (=0)$

Αν H_0^A είναι αληθής τότε $E(MS_A) = E(MS_{res})$
Άρα, ένα τεστ θα στηριχτεί στο $\frac{MS_A}{MS_{res}} = F_A$
Επίσης, αν H_0^B είναι αληθής τότε $E(MS_B) = E(MS_{res})$
Άρα, ένα τεστ θα στηριχτεί στο $\frac{MS_B}{MS_{res}} = F_B$

Θεώρημα:
Υπό τις υποθέσεις για τα σφάλματα και τις πλευρικές συνθήκες
α) $\frac{SS_{res}}{\sigma^2} \sim \chi^2_{(I-1)(J-1)}$
β) υπό την $H_0^A: \alpha_1 = \alpha_2 = \dots = \alpha_I = 0$, ισχύει: $\frac{SS_A}{\sigma^2} \sim \chi^2_{I-1}$
γ) " " $H_0^B: b_1 = b_2 = \dots = b_J = 0$, ισχύει: $\frac{SS_B}{\sigma^2} \sim \chi^2_{J-1}$

Απόδειξη
α) $Y_{ij} \sim N(\mu + a_i + b_j, \sigma^2)$, $\forall i = 1, \dots, I$, $j = 1, \dots, J$ ανεξ/τ
Άρα,
$$\frac{Y_{ij} - \mu - a_i - b_j}{\sigma} \sim N(0,1)$$
$$\frac{(Y_{ij} - \mu - a_i - b_j)^2}{\sigma^2} \sim \chi^2_1$$
$$\frac{\sum_{i=1}^{I} (Y_{ij} - \mu - a_i - b_j)^2}{\sigma^2} \sim \chi^2_I$$
Τότε, $\frac{\sum_{i=1}^{I} (Y_{ij} - \hat{\mu} - \hat{a}_i - \hat{b}_j)^2}{\sigma^2} \sim \chi^2_{I-1}$

Αλλά, $\overline{Y}_{..} = \frac{1}{J} \cdot \overline{Y}_{i.} = \frac{1}{I} \cdot \overline{Y}_{.j}$

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

$$\left.\begin{array}{l} \dfrac{\sum_{i=1}^{I}(Y_{ij}-\bar{Y}_{i.}-\bar{Y}_{.j}+\bar{Y}_{..})^2}{\sigma^2} \sim \chi^2_{I-1} \\ \text{ομοίως, από την άλλη} \\ \dfrac{\sum_{j=1}^{J}(Y_{ij}-\bar{Y}_{i.}-\bar{Y}_{.j}+\bar{Y}_{..})^2}{\sigma^2} \sim \chi^2_{J-1} \end{array}\right\} \dfrac{\sum_{i=1}^{I}\sum_{j=1}^{J}(Y_{ij}-\bar{Y}_{i.}-\bar{Y}_{.j}+\bar{Y}_{..})^2}{\sigma^2} \sim \chi^2_{(I-1)(J-1)}$$

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

$\bar{Y}_{i.} = \frac{1}{J}\cdot\sum_{j=1}^{J} Y_{ij} \sim$ Normal

$E(\bar{Y}_{i.}) = E(\frac{1}{J}\cdot\sum_j Y_{ij}) = \frac{1}{J}\cdot\sum_j E(Y_{ij}) = \frac{1}{J}\cdot\sum_j(\mu+a_i+b_j) =$

$= \mu+a_i+\frac{1}{J}\sum_j b_j$ ①

Υπό την H_0^A και την παραδοχή ότι $\sum_j b_j = 0$ ②

Από τις ① και ② έπεται: $E(\bar{Y}_{i.}) = \mu$.

$Var(\bar{Y}_{i.}) = Var(\frac{1}{J}\sum_j Y_{ij}) \overset{Y_{ij}\ \text{ανεξ.}}{=} \frac{1}{J^2}\cdot\sum_j Var(Y_{ij}) =$

$= \frac{1}{J^2}\cdot\sum_{j=1}^{J}\sigma^2 = \frac{\sigma^2}{J}$

Άρα,

$\bar{Y}_{i.} \sim N(\mu, \frac{\sigma^2}{J})$, $\forall i=1,2,\dots,I$ υπό την H_0^A και $\sum_j b_j = 0$

Έτσι,

$\frac{\bar{Y}_{i.}-\mu}{\sigma/\sqrt{J}} \sim N(0,1) \Rightarrow \frac{\sqrt{J}(\bar{Y}_{i.}-\mu)}{\sigma} \sim N(0,1) \Rightarrow \frac{J\cdot(\bar{Y}_{i.}-\mu)^2}{\sigma^2} \sim \chi^2_1$

$\Rightarrow \frac{J\sum_{i=1}^{I}(\bar{Y}_{i.}-\mu)^2}{\sigma^2} \sim \chi^2_I$, τότε $\frac{J\cdot\sum_{i=1}^{I}(\bar{Y}_{i.}-\hat{\mu})^2}{\sigma^2} \sim \chi^2_{I-1}$

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

γ) ομοίως.

Έλεγχοι ισοεπίδρασης των επιπέδων των παραγόντων A & B

Για τον έλεγχο της $H_0^A: \alpha_1=\alpha_2=\dots=\alpha_I(=0)$ η σστ είναι
$\frac{MS_A}{MS_{res}} \sim F_{I-1,(I-1)(J-1)}$, υπο την H_0^A και κρίσιμη περιοχή

μεγέθους α την $F_A \geq F_{(I-1),(I-1)\cdot(J-1),\alpha}$

Για τον έλεγχο της $H_0^B: b_1=b_2=\dots=b_J(=0)$ η σστ είναι
$\frac{MS_B}{MS_{res}} \sim F_{J-1,(I-1)(J-1)}$, υπο την H_0^B και κρίσιμη περιοχή

μεγέθους α την $F_B \geq F_{(J-1),(I-1)\cdot(J-1),a}$

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 την πιο σημαντική
επίδραση στην Y;

Απάντηση

Αυτό μπορούμε να το δούμε με τις πολλαπλές συγκρίσεις

Μέθοδος Ελάχιστης Σημαντικής Διαφοράς (ΕΣΔ)

ΕΣΔ για τον παράγοντα A $\longrightarrow$ έλεγχος της υπόθεσης
$H_0^A: \alpha_i=\alpha_{i'}$ vs $H_a^A: \alpha_i\neq\alpha_{i'}$
Η H_0^A απορρίπτεται αν $|\bar{Y}_{i.}-\bar{Y}_{i'.}| > t_{(I-1)(J-1),\frac{a}{2}}\sqrt{\frac{2MS_{res}}{J}} = ΕΣΔ(A)$

Αν τώρα $H_0^A: \alpha_i=\alpha_{i'}$ απορρίπτεται τότε
αν $\bar{Y}_{i.}-\bar{Y}_{i'.}>0$ το i-επίπεδο ασκεί πιο σημαντική
επίδραση από το i'-επίπεδο ενώ
αν $\bar{Y}_{i.}-\bar{Y}_{i'.}<0$ το i'-επίπεδο ασκεί πιο σημαντική
επίδραση από το i-επίπεδο
και αυτό για κάθε $i=1,2,\dots,I$
Ακριβώς ίδια περίπτωση για τον έλεγχο της υπόθεσης:
$H_0^B: b_j=b_{j'}$ vs $H_a^B: b_j\neq b_{j'}$

Η H_0^B απορρίπτεται αν : $|\bar{Y}_{.j} - \bar{Y}_{.j'}| > t_{(I-1)(J-1),\frac{a}{2}} \sqrt{\frac{2MS_{res}}{I}} = ΕΣΔ_{(B)}$

Αν $H_0^B: b_j = b_{j'}$ απορρίπτεται τότε

αν $\bar{Y}_{.j} - \bar{Y}_{.j'} > 0$ πιο σημαντική επίδραση ασκεί το j

αν $\bar{Y}_{.j} - \bar{Y}_{.j'} < 0$ " " " " " j'

και αυτό $\forall j = 1,2,\dots,J$.

ΜΕΘΟΔΟΣ ΓΡΑΜΜΙΚΩΝ ΑΝΤΙΘΕΣΕΩΝ:

Έστω $d_A = \sum_{i=1}^{I} c_i a_i$ (με $\sum_{i=1}^{I} c_i = 0$) μια γραμμική αντίθεση στα επίπεδα του παράγοντα Α. Για τον έλεγχο της $H_0^{d_A}$ ότι δηλαδή $d_A = 0$ χρησιμοποιείται η σστ

$F_{d_A} = \frac{MS_{d_A}}{MS_{res}}$ όπου $MS_{d_A} = \frac{\hat{d}_A^2}{\frac{1}{J}\sum_{i=1}^{I} c_i^2}$ και $\hat{d}_A = \sum_{i=1}^{I} c_i \bar{Y}_{i.}$ με

κατανομή της $F_{d_A} \sim F_{1,(I-1)(J-1)}$ υπό την $H_0^{d_A}$ και κριτ. περ. μεγέθους α : $F_{d_A} \geq F_{1,(I-1)(J-1),a}$

Κατά αντιστοιχία και για τον Παράγοντα Β

$d_B = \sum_{j=1}^{J} c_j \beta_j$ (με $\sum_{j=1}^{J} c_j = 0$)

$H_0^{d_B}: d_B = 0$ παίρνουμε τη σστ

$F_{d_B} = \frac{MS_{d_B}}{MS_{res}}$ με $MS_{d_B} = \frac{\hat{d}_B^2}{\frac{1}{I}\sum_{j=1}^{J} c_j}$ και $\hat{d}_B = \sum_{j=1}^{J} c_j \bar{Y}_{.j}$ με

κατανομή της $F_{d_B} \sim F_{1,(I-1)(J-1)}$ υπό την $H_0^{d_B}$ και κρ. περ. μεγέθους α : $F_{d_B} \geq F_{1,(I-1)\cdot(J-1),a}$

Το μοντέλο της ανάλυσης διακύμανσης κατά 2 παράγοντες ως ειδική περίπτωση του γενικού γραμμικού μοντέλου $\underset{\sim}{Y} = X\underset{\sim}{b} + \underset{\sim}{\varepsilon}$

$Y_{ij} = \mu + a_i + b_j + \varepsilon_{ij}$, $i = 1,2,\dots,I$, $j = 1,2,\dots,J$.

$\underset{\sim}{Y} = (Y_{11}, Y_{12},\dots, Y_{1J}, Y_{21}, Y_{22},\dots, Y_{2J},\dots, Y_{I1}, Y_{I2},\dots, Y_{IJ})^T_{IJ\times 1}$

και ο κατάλληλος πίνακας σχεδιασμού είναι ο εξής:

$$
X = \begin{array}{c}
\begin{array}{cccccccccc} & \longleftarrow & I & \longrightarrow & & \longleftarrow & J & \longrightarrow & \end{array} \\
\begin{bmatrix}
1 & 1 & 0 & \cdots & 0 & 1 & 0 & \cdots & \cdots & 0 \\
1 & 1 & 0 & \cdots & 0 & 0 & 1 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & & \vdots & 0 & 0 & & \ddots & \vdots \\
1 & 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & \cdots & 1 \\
1 & 0 & 1 & \cdots & 0 & 1 & 0 & \cdots & \cdots & 0 \\
1 & 0 & 1 & \cdots & 0 & 0 & 1 & & & \vdots \\
\vdots & \vdots & \vdots & & \vdots & \vdots & & & \ddots & 0 \\
1 & 0 & 1 & \cdots & 0 & 0 & \cdots & \cdots & 0 & 1 \\
1 & 0 & 0 & \cdots & 1 & 1 & 0 & \cdots & \cdots & 0 \\
1 & 0 & 0 & \cdots & 1 & 0 & 1 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & & \vdots & \vdots & & & \ddots & 0 \\
1 & 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 0 & 1
\end{bmatrix}_{(IJ)\times(I+J+1)}
\end{array}
$$

$$\underset{\sim}{b} = \left(\mu, a_1, a_2, \ldots, a_i, b_1, b_2, \ldots, b_j \right)_{(I+J+1)\times 1}$$

$$\underset{\sim}{\varepsilon} = \left(\varepsilon_{11}, \varepsilon_{12}, \ldots, \varepsilon_{1J}, \varepsilon_{21}, \varepsilon_{22}, \ldots, \varepsilon_{2J}, \ldots, \varepsilon_{I1}, \varepsilon_{I2}, \ldots, \varepsilon_{IJ} \right)_{IJ\times 1}$$