Μάθημα 9ο — Παρασκευή 08/12 — 04/12/17
Αναπλήρωση της 11/12
12:00– 14:30

Μέθοδος γραμμικών αντιθέσεων ή μέθοδος Scheffe

ΟΡΙΣΜΟΣ: Κάθε γραμμικός συνδυασμός των κύριων επιδράσεων a_i, $i=1,\dots,I$ των επιπέδων του παράγοντα στην εξαρτημένη μεταβλητή, της μορφής:
$L=\sum_{i=1}^{I} c_i a_i$ με $\sum_{i=1}^{I} c_i=0$ λέγεται **γραμμική αντίθεση**.

Η γραμμική αντίθεση επιτρέπει τη διατύπωση συγκρίσεων περισσότερων των δύο a_i, $i=1,\dots,I$

α) π.χ. Έστω ενδιαφέρει $a_i=a_{i'}$ (ΕΣΔ)

$$\left.\begin{array}{l} \text{Αν } c_k=0,\ k\neq i,i' \\ \text{Αν } c_i=1,\ c_{i'}=-1 \end{array}\right\}^{k=1,\dots,I} \Rightarrow L=\sum_{i=1}^{I} c_i a_i=0 \Leftrightarrow a_i=a_{i'} \Leftrightarrow a_i-a_{i'}=0$$

Άρα $H_0: a_i=a_{i'} \Leftrightarrow L=0$

β) $c_k=0$, $\forall k\neq i,m,l$ ← Είναι η επίδραση του κ-επιπέδου ισοδύναμη με το μέσο όρο των επιδράσεων των l και m επιπέδων;
$c_i=1$, $c_m=\frac{1}{2}=c_l$

$$L=\sum_{i=1}^{I} c_i a_i=a_i-\frac{1}{2}a_m-\frac{1}{2}a_l$$

$$L=0 \Leftrightarrow a_i=\frac{1}{2}a_m+\frac{1}{2}a_l$$

Η επίδραση του i-επιπέδου είναι ίση με το ημιάθροισμα των επιδράσεων των επιπέδων m, l.

- Έχει ενδιαφέρον η **κατασκευή** στατιστικού τεστ για έλεγχο $H_0: L=0$ έναντι $H_a: L\neq 0$.

Θεωρώ ότι οι υποθέσεις για τα σφάλματα ικανοποιούνται: δηλ. $\varepsilon_{ij}\sim N(0,\sigma^2)$; $i=1,\dots,I$, $j=1,\dots,J_i$

και έστω ανεξάρτητα. Λόγω των υποθέσεων για τα σφάλματα $Y_{ij}\sim N(\mu+a_i,\sigma^2)$, άρα $\overline{Y_{i\cdot}}\sim N\left(\mu+a_i,\frac{\sigma^2}{J_i}\right)$, $i=1,\dots,I$ ανεξάρτητα, λόγω της ανεξαρτησίας των Y_{ij}

λόγω ασυσχέτιστου και ανεξαρτησίας των σφαλμάτων.

Με την κατά Wald προσέγγιση ένα τεστ για έλεγχο $H_0: L=0$.
θα στηρίζεται σε ένα εκτιμητή της L.
Θεωρώ τον εξής εκτιμητή της L: $\hat{L}=\sum_{i=1}^{I} c_i\hat{a}_i$, $\hat{a}_i=\bar{Y}_{i.}-\bar{Y}_{..}$, $i=1,\dots,I$.

$$\hat{L}=\sum_{i=1}^{I} c_i\hat{a}_i=\sum_{i=1}^{I} c_i(\bar{Y}_{i.}-\bar{Y}_{..})=\sum_{i=1}^{I} c_i\bar{Y}_{i.}-\bar{Y}_{..}\cancel{\sum_{i=1}^{I} c_i}^{\,0}\Rightarrow \hat{L}=\sum_{i=1}^{I} c_i\bar{Y}_{i.}$$

Το $\hat{L}$ είναι ένας γραμμικός συνδυασμός των ανεξάρτητων $\bar{Y}_{i.}\sim N(\mu+a_i,\frac{\sigma^2}{J_i})$, $i=1,\dots$
Άρα, $\hat{L}\sim$ Normal.

- $E(\hat{L})=E\left(\sum_{i=1}^{I} c_i\bar{Y}_{i.}\right)=\sum_{i=1}^{I} c_i E(\bar{Y}_{i.})=\sum_{i=1}^{I} c_i(\mu+a_i)=\mu\cancel{\sum_{i=1}^{I} c_i}^{\,0}+\sum_{i=1}^{I} c_i a_i\Rightarrow$

$$E(\hat{L})=\sum_{i=1}^{I} c_i a_i=L$$

- $Var(\hat{L})=Var\left(\sum_{i=1}^{I} c_i\bar{Y}_{i.}\right)\xrightarrow[\substack{\text{λόγω κανονικότητας}\\ \text{(συνδιακυμάνσεις=0)}}]{\bar{Y}_{i.}\text{ ανεξ.}}\sum_{i=1}^{I} c_i^2 Var\bar{Y}_{i.}=\sum_{i=1}^{I} c_i^2\frac{\sigma^2}{J_i}\Rightarrow$

$$Var(\hat{L})=\sigma^2\sum_{i=1}^{I}\left(\frac{c_i^2}{J_i}\right),\ i=1,\dots,I$$

Άρα, $\hat{L}\sim N\left(L,\sigma^2\sum_{i=1}^{I}\frac{c_i^2}{J_i}\right)\Rightarrow\frac{\hat{L}-L}{\sqrt{\sum_{i=1}^{I} c_i^2\frac{\sigma^2}{J_i}}}\sim N(0,1)$

Υπό την $H_0: L=0$, το $\hat{L}\sim N\left(0,\sigma^2\sum_{i=1}^{I}\frac{c_i^2}{J_i}\right)\Rightarrow\frac{\hat{L}}{\sqrt{\sum_{i=1}^{I}\sigma^2\frac{c_i^2}{J_i}}}\sim N(0,1)$

Υπό την $H_0: L=0$, το $\frac{\hat{L}}{\sigma\cdot\left(\sum_{i=1}^{I}\frac{c_i^2}{J_i}\right)^{1/2}}\sim N(0,1)\Rightarrow$ θέλω να το διώξω γιατί δεν το γνωρίζω.

$$\frac{\hat{L}^2}{\sigma^2\sum_{i=1}^{I}\frac{c_i^2}{J_i}}\sim\chi_1^2 \text{ υπό την } H_0: L=0.$$

Θυμόμαστε $\frac{SSres}{\sigma^2}\sim\chi^2_{N-I}$ (MS_L: μέσο τετράγωνο που οφείλεται στη γραμμική αντίθεση L)

Συμβολίζω $MS_L=\frac{\hat{L}^2}{\sum\frac{c_i}{J_i}}$ και θεωρώ το $F_L\overset{\text{ορ.}}{=}\frac{MS_L}{MSres}=\frac{\hat{L}^2/\sum_{i=1}^{I} c_i^2/J_i}{SSres/(N-I)}=$

$$= \frac{\dfrac{\hat{L}^2}{\sigma^2 \sum_{i=1}^{I} c_i^2/J_i}}{\dfrac{SSres}{\sigma^2(N-I)}} \sim \frac{\chi_1^2}{\chi_{N-I}^2/(N-I)} \overset{\hat{L},\, SSres\ \text{ανεξ.}}{\underset{\text{ανεξ.}}{=\!=\!=}} \frac{MS_L}{MSres} \quad F_{1,N-I} \text{ υπό } H_0: L=0.$$

<u>Μορφή της κρίσιμης περιοχής</u>: Μεγάλες τιμές του F_L οδηγούν σε μεγάλες τιμές του $\hat{L}$ και επειδή $\hat{L}$ εκτιμητής του L, μεγάλες τιμές του F_L οδηγούν σε απόρριψη της $H_0: L=0$.

Κ.Π. $F_L \geq C$.

Προσδιορισμός C: $\alpha = P(\text{απορ. } H_0 \mid H_0 \text{ αληθ.}) = P(F_L \geq C \mid F_L \sim F_{1,N-I}) \Rightarrow$

$\Rightarrow \alpha = P(F_{1,N-I} \geq C) \Rightarrow$

$\Rightarrow C = F_{1,N-I,\alpha}$.

Συγκεντρωτικά: Για τον έλεγχο της $H_0: L=0$, χρησιμοποιείται η στατιστική συνάρτηση $F_L = \dfrac{MS_L}{MSres}$, $MS_L = \dfrac{\hat{L}^2}{\sum_{i=1}^{I} \frac{c_i^2}{J_i}}$ με κατανομή την $F_{1,N-I}$ υπό $H_0: L=0$ και κ.π. μεγέθους α την $F_L \geq F_{1,N-I,\alpha}$.

Ανάλυση υπολοίπων

Υπόλοιπα: $e_{ij} = Y_{ij} - \hat{Y}_{ij} = Y_{ij} - (\hat{\mu} + \hat{\alpha}_i) = Y_{ij} - (\bar{Y}_{..} + \bar{Y}_{i.} - \bar{Y}_{..}) \Rightarrow$

$e_{ij} = Y_{ij} - \bar{Y}_{i.}$, $i = 1, \dots, I$, $j = 1, \dots, J_i$.

Τα e_{ij} χρησιμοποιούνται όπως και στα μοντέλα παλινδρόμησης για τον έλεγχο ικανοποίησης των υποθέσεων για τα σφάλματα.

Η ανάλυση διακύμανσης κατά ένα παράγοντα ως ειδική περίπτωση του γενικού γραμμικού μοντέλου $\underset{\sim}{Y} = X\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$

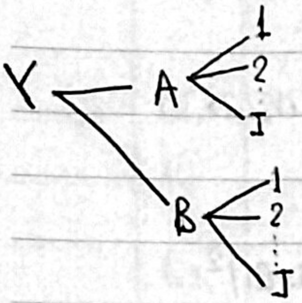
Μοντέλο ανάλυσης διακύμανσης κατά ένα παράγοντα.

$$\underset{\uparrow \atop \beta_i}{Y_{ij} = \mu_i} + \varepsilon_{ij} \Leftrightarrow Y_{ij} = \underset{\uparrow \atop \beta_i}{\underbrace{\mu + \alpha_i}} + \varepsilon_{ij}, \quad i = 1, \dots, I, \quad j = 1, \dots, J_i.$$

Μοντέλο ανάλυσης Διακύμανσης κατά δυο Παραγοντες

Η ανάλυση διακύμανσης κατά δύο παράγοντες διερευνά τη σχέση ανάμεσα σε μια ποσοτική μεταβλητή Y και σε δύο παράγοντες (ποιοτικές μεταβλητές) A και B, που ο A εμφανίζεται σε I επίπεδα και ο B εμφανίζεται σε J επίπεδα.

Ενδιαφέρον εντοπίζεται στην ανίχνευση των επιπέδων των δύο παραγόντων που ασκούν σημαντικότερη επίδραση στην εξαρτημένη μεταβλητή Y.



Y : εξαρτημένη ποσοτική
A, B : παράγοντες (ποιοτική ανεξάρτητη μεταβλητή)
$1,2,\ldots,I$, $1,2,\ldots,J$ επίπεδα

Μορφή δεδομένων

	παράγοντας B				
παράγοντας A	1	2	⋯	J	Σύνολα
1	Y_{11}	Y_{12}	⋯	Y_{1J}	$Y_{1.}$ $\bar{Y}_{1.}$
2	Y_{21}	Y_{22}	⋯	Y_{2J}	$Y_{2.}$ $\bar{Y}_{2.}$
⋮		Y_{ij}			
I	Y_{I1}	Y_{I2}	⋯	Y_{IJ}	$Y_{I.}$
Σύνολα	$Y_{.1}$ $\bar{Y}_{.1}$	$Y_{.2}$ $\bar{Y}_{.2}$	⋯	$Y_{.J}$ $\bar{Y}_{.J}$	

$Y_{ij} \rightarrow$ παριστά παρατήρηση της Y υπό το επίπεδο i του A και το επίπεδο j του B, $i=1,\ldots,I$, $j=1,\ldots,J$

Θεωρώ μια παρατήρηση ανά κυψελίδα, μια παρατήρηση ανά δοκιμασία.
Περισσότερες παρατηρήσεις ανά δοκιμασία $\longrightarrow$ μοντέλα με αλληλεπίδραση

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

$$Y_{.j} = \sum_{i=1}^{I} Y_{ij}, \quad \bar{Y}_{.j} = \frac{1}{I} \sum_{i=1}^{I} Y_{ij}$$

$$Y_{..} = \sum_{i=1}^{I} \sum_{j=1}^{J} Y_{ij} = \sum_{i=1}^{I} Y_{i.} = \sum_{j=1}^{J} Y_{.j}$$

$$\bar{Y}_{..} = \frac{1}{IJ} \sum_{i=1}^{I} \sum_{j=1}^{J} Y_{ij} = \frac{1}{I} \sum_{i=1}^{I} \bar{Y}_{i.} = \frac{1}{J} \sum_{j=1}^{J} \bar{Y}_{.j} = \frac{1}{IJ} Y_{..}$$ γενικός δειγματικός μέσος

Μοντέλο

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

- Y_{ij}: παρατήρηση της εξαρτημένης μεταβλητής
- μ: μια κοινή επίδραση στην Y όλων των επιπέδων των δύο παραγόντων
- a_i: ατομική επίδραση στην Y του i-επιπέδου του A
- b_j: ατομική επίδραση στην Y του j-επιπέδου του B.
- ε_{ij}: σφάλματα μοντέλου.

Εκτιμητές Ελαχίστων Τετραγώνων του μοντέλου

Ελαχιστοποίηση $S = \sum_{i=1}^{I}\sum_{j=1}^{J}\varepsilon_{ij}^2 = \sum_{i=1}^{I}\sum_{j=1}^{J}(Y_{ij}-\mu-a_i-b_j)^2$

$\frac{\partial S}{\partial \mu} = 0$	$\frac{\partial S}{\partial a_i} = 0$, $i=1,\dots,I$	$\frac{\partial S}{\partial b_j} = 0$, $j=1,\dots,J$

κανονικές εξισώσεις
$$\begin{cases} IJ\mu + J\sum_{i=1}^{I}a_i + I\sum_{j=1}^{J}b_j = Y_{..} \\ J\mu + Ja_i + \sum_{j=1}^{J}b_j = Y_{i.}, \quad i=1,\dots,I \\ I\mu + \sum_{i=1}^{I}a_i + Jb_j = Y_{.j}, \quad j=1,\dots,J \end{cases}$$

(γιατί η 1η εξίσωση μπορεί να παραχθεί από τις υπόλοιπες)

$\nexists$ μοναδική λύση. Πετυχω μοναδική λύση καταφεύγω στην υιοθέτηση πλευρικών συνθηκών. Διαισθητικά ένας εκτιμητής του μ πρέπει να είναι $\overline{Y}_{..}$

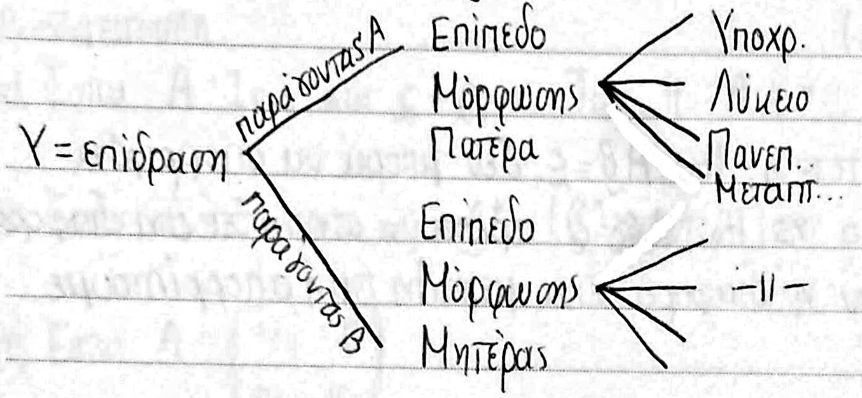
Πρέπει $\sum_{i=1}^{I}a_i = \sum_{j=1}^{J}b_j = 0$. (για να μπορούμε να εκτιμήσουμε το μ.)

Αν υιοθετήσω ως πλευρικές συνθήκες τις $\sum_{i=1}^{I}a_i = \sum_{j=1}^{J}b_j = 0$, τότε οι μοναδικές λύσεις του συστήματος των κανονικών εξισώσεων είναι: $\hat{\mu} = \overline{Y}_{..}$

$\hat{a}_i = \overline{Y}_{i.} - \overline{Y}_{..}$, $i=1,\dots,I$

$\hat{b}_j = \overline{Y}_{.j} - \overline{Y}_{..}$ $j=1,\dots,J$

ΠΑΡΑΔΕΙΓΜΑ: $Y = \mu + a_i + \varepsilon_{ij}$



Ανάλυση διακύμανσης κατα δύο παράγοντες και γενικό γραμμικό μοντέλο

$$\underset{\sim}{Y} = \begin{bmatrix} Y_{11} \\ Y_{12} \\ \vdots \\ Y_{1J} \\ Y_{21} \\ Y_{22} \\ \vdots \\ Y_{2J} \\ \vdots \\ Y_{I1} \\ Y_{I2} \\ \vdots \\ Y_{IJ} \end{bmatrix}, \quad \underset{\sim}{X} = \begin{array}{c} \\ J \\ \\ \\ J \\ \\ \\ \\ J \\ \\ \end{array} \begin{array}{c} \begin{array}{ccccccc} & \longleftarrow & I & \longrightarrow & \longleftarrow & J & \longrightarrow \end{array} \\ \begin{bmatrix} 1 & 1 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & & & & & & \\ 1 & 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 1 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 1 & \cdots & 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & & & & & & \\ 1 & 0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 1 \\ \vdots & & & & & & & & \\ 1 & 0 & 0 & \cdots & 1 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 1 & 0 & 1 & \cdots & 0 \\ \vdots & & & & & & & & \\ 1 & 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 1 \end{bmatrix}_{IJ \times (I+J+1)} \end{array}, \quad \underset{\sim}{\beta} = \begin{bmatrix} \mu \\ a_1 \\ \vdots \\ a_I \\ \beta_1 \\ \vdots \\ \beta_J \end{bmatrix}_{(I+J+1)\times 1}$$

$\underset{\sim}{\varepsilon}$

$$\hat{\beta} = (X'X)^{-1} X' \underset{\sim}{Y}$$

Ανάλυση διακύμανσης κατά δύο παράγοντες

$$Y \begin{cases} \text{παράγοντας } A \begin{cases} 1 \\ 2 \\ \vdots \\ I \end{cases} \\ \text{παράγοντας } B \begin{cases} 1 \\ 2 \\ \vdots \\ J \end{cases} \end{cases}$$

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

ΕΕΓ:
$$\left.\begin{aligned} \hat{\mu} &= \overline{Y}_{\cdot\cdot} \\ \hat{a}_i &= \overline{Y}_{i\cdot} - \overline{Y}_{\cdot\cdot} \quad , i = 1,\dots,I \\ \hat{b}_j &= \overline{Y}_{\cdot j} - \overline{Y}_{\cdot\cdot} \quad , j = 1,\dots,J \end{aligned}\right\} \text{ΠΣ: } \sum_{i=1}^{I} a_i = \sum_{j=1}^{J} b_j = 0$$

Υποθέσεις σφάλματα

$$\left.\begin{aligned} &1)\ E(\varepsilon_{ij}) = 0 \\ &2)\ Var(\varepsilon_{ij}) = \sigma^2 \\ &3)\ Cov(\varepsilon_{ij}, \varepsilon_{kl}) = 0 \\ &4)\ \varepsilon_{ij} \sim N(0, \sigma^2) \end{aligned}\right\} \begin{aligned} &\varepsilon_{ij} \sim N(0,\sigma^2) \\ &\text{ασυσχέτιστα} \\ &(\text{ανεξάρτητα}) \end{aligned} \qquad \left.\begin{aligned} &1)\ E(Y_{ij}) = \mu + a_i + b_j \\ &2)\ Var(Y_{ij}) = \sigma^2 \\ &3)\ Cov(Y_{ij}, Y_{kl}) = 0 \\ &4)\ Y_{ij} \sim N(\mu + a_i + b_j, \sigma^2) \end{aligned}\right\} \begin{aligned} &Y_{ij} \sim N(\mu + a_i + b_j, \sigma^2) \\ &\text{ασυσχέτιστα} \\ &(\text{ανεξάρτητα}) \end{aligned}$$

<u>ΙΔΙΟΤΗΤΕΣ των ΕΕΓ</u>

1) Οι ΕΕΓ εκτιμούν αμερόληπτα τις αντίστοιχες παραμέτρους

$E(\hat{\mu}) = \mu$, $E(\hat{a}_i) = a_i$, $E(\hat{b}_j) = b_j$, $i = 1,\dots,I$, $j = 1,\dots,J$.

ΑΠΟΔΕΙΞΗ

$$E(\hat{a}_i) = E(\overline{Y}_{i\cdot} - \overline{Y}_{\cdot\cdot}) = E(\overline{Y}_{i\cdot}) - E(\overline{Y}_{\cdot\cdot}) = E\left(\frac{1}{J}\sum_{j=1}^{J} Y_{ij}\right) - E\left(\frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J} Y_{ij}\right) =$$

$$= E\left(\frac{1}{J}\sum_{j=1}^{J}(\mu + a_i + b_j + \varepsilon_{ij})\right) - E\left(\frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J}(\mu + a_i + b_j + \varepsilon_{ij})\right) =$$

$$= \frac{1}{J}\sum_{j=1}^{J}(\mu + a_i + b_j + \cancelto{0}{E(\varepsilon_{ij})}) - \frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J}(\mu + a_i + b_j + \cancelto{0}{E(\varepsilon_{ij})}) =$$

$$= \frac{1}{J}(J\mu + Ja_i + \cancelto{0}{\sum_{j=1}^{J} b_j}) - \frac{1}{IJ}(IJ\mu + \cancelto{0}{J\sum_{i=1}^{I} a_i} + \cancelto{0}{I\sum b_j}) = \mu + a_i - \mu = a_i$$

2) Υπό τις υποθέσεις για τα σφάλματα οι ΕΕΤ(μ, a_i, b_j) = ΕΜΠ(μ, a_i, b_j)

ΑΠΟΔΕΙΞΗ

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

$\log L = \ldots$ συνεχίζω όπως στα άλλα μοντέλα

3) Ολική μεταβλητότητα στο μοντέλο Ανάλυσης διακύμανσης κατά 2 παράγοντες — Πίνακας ΑΝΑΔΙΑ ανάλυσης διακύμανσης κατά 2 παράγοντες

Ολική μεταβλητότητα ≡ Δειγματική διακύμανση εκτός του όρου που σχετίζεται με μέγεθος δείγματος.

$$SS_{tot} = \sum_{i=1}^{I}\sum_{j=1}^{J}(\overline{Y}_{ij} \overbrace{\;}^{\pm\overline{Y}_{i\cdot}\,\pm\overline{Y}_{\cdot j}} - \overline{Y}_{\cdot\cdot})^2 \Rightarrow SS_{tot} = J\sum_{i=1}^{I}(\overline{Y}_{i\cdot}-\overline{Y}_{\cdot\cdot})^2 + I\sum_{j=1}^{J}(\overline{Y}_{\cdot j}-\overline{Y}_{\cdot\cdot})^2 + \sum_{i=1}^{I}\sum_{j=1}^{J}(Y_{ij}-Y_{i\cdot}-Y_{\cdot j}+Y_{\cdot\cdot})^2$$

$$SS_{tot} = SS_A + SS_B + SS_{res}$$

Υπόλοιπα μοντέλο ανάλυσης διακύμανσης κατά δύο παράγοντες

$$e_{ij} = Y_{ij} - \hat{Y}_{ij} = Y_{ij} - (\mu - \hat{a}_i - \hat{b}_j) = Y_{ij} - (\cancel{\overline{Y}_{\cdot\cdot}} + \overline{Y}_{i\cdot} - \cancel{\overline{Y}_{\cdot\cdot}} + \overline{Y}_{\cdot j} - \overline{Y}_{\cdot\cdot}) \Rightarrow$$

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