"11ο μάθημα" 13/12/21 ①

▶ Ανάλυση Διακύμανσης κατά δύο παράγοντες.

Ποσοτική

↓

Y ← Ⓐ Ποιοτική - Παράγοντας, I επίπεδα: $\alpha_1, \alpha_2, \dots, \alpha_I$

Y ← Πως επηρεάζεται από μια άλλη Ⓑ Ποιοτική - Παράγοντας, J επίπεδα: $\beta_1, \beta_2, \dots, \beta_J$

Μορφή Δεδομένων (να έχουν τουλ. μια παρατήρηση σε κάθε κυψελίδα).

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

Y_{ij}: Παρατήρηση της Y που έχει αλλάξει στο i-επίπεδο του A και στο j-επίπεδο του B

Αλγεβρικά: $\boxed{Y_{i\cdot} = \sum_{j=1}^{J} Y_{ij} \qquad \overline{Y}_{i\cdot} = \frac{1}{J}\sum_{j=1}^{J} Y_{ij}}$

$$\boxed{Y_{\cdot j} = \sum_{i=1}^{I} Y_{ij} \qquad \overline{Y}_{\cdot j} = \frac{1}{I}\sum_{i=1}^{I} Y_{ij}}$$

$$Y_{\cdot\cdot} = \sum_i \sum_j Y_{ij} \qquad \overline{Y}_{\cdot\cdot} = \frac{1}{IJ}\sum_i \sum_j Y_{ij} =$$

$$= \frac{1}{I}\sum_i \overline{Y}_{i\cdot} = \frac{1}{J}\sum_j \overline{Y}_{\cdot j} =$$

$$= \frac{1}{IJ} Y_{\cdot\cdot}$$

▶ Μοντέλο Ανάλυσης Διακύμανσης κατά δύο Παράγοντες

κοινή επίδραση όλων των επιπέδων των παραγόντων Α και Β στην Υ

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

Παρατηρήσεις στην εξαρτ. μεταβλητή Υ

Ατομική επίδραση των επιπέδων του Β στην Υ

Ατομική επίδραση των επιπέδων του Α στην Υ

Σφάλματα που οφείλονται στην πιθανή ύπαρξη άλλων παραγόντων που επηρεάζουν την Υ και τους οποίους λαμβάνουμε υπόψιν

Παράμετροι Μοντέλου: μ, α_i, β_j

Εκτιμητές Ελαχ. Τετραγώνων των μ, α_i, β_j με $i=1,\dots,I$, $j=1,\dots,J$

$$S^2 = \sum_i \sum_j \varepsilon_{ij}^2 = \sum_i \sum_j (Y_{ij} - \mu - \alpha_i - \beta_j)^2$$

$$\frac{\partial S^2}{\partial \mu} = 0, \quad \frac{\partial S^2}{\partial \alpha_i} = 0, \quad \frac{\partial S^2}{\partial \beta_j} = 0 \qquad i=1,\dots,I \quad j=1,\dots,J$$

↓

Κανονικών Εξισώσεων

$$IJ\mu + J\cdot\sum_{i=1}^{I} \alpha_i + I\cdot\sum_{j=1}^{J} b_j = Y_{..}$$

$$J\mu + J\alpha_i + \sum_{j=1}^{J} b_j = Y_{i.}\ ,\ i=1,\dots,I$$

$$I\mu + Ib_j + \sum_{i=1}^{I} \alpha_i = Y_{.j}\ ,\ j=1,\dots,J$$

(όχι μοναδική λύση καθώς δεν είναι γραμμικές. Ίσως προς μαθ
Άρα δεν έχω μοναδικούς εκτιμητές

<u>Υιοθέτηση Πλευρικών Συνθηκών</u>

$$\sum_{i=1}^{I} \alpha_i = 0 \quad \text{και} \quad \sum_{j=1}^{J} b_j = 0$$

επόμενα προκύπτει ότι $\hat{\mu} = \bar{Y}_{..}$, $\hat{\alpha}_i = \bar{Y}_{i.} - \bar{Y}_{..}$ και $\hat{b}_j = \bar{Y}_{.j} - \bar{Y}_{..}$ με $i=1,\dots,I$ και $j=1,\dots,J$.

<u>Παρατήρηση</u>: Το μοντέλο ανάλυσης διακύμανσης κατά δύο παράγοντες, ειδική περίπτωση του: $\underset{\sim}{Y} = X\underset{\sim}{b} + \underset{\sim}{\varepsilon}$.

$$\underset{\sim}{Y} = \begin{pmatrix} Y_{11}\\ Y_{12}\\ \vdots \\ Y_{1J}\\ Y_{21}\\ Y_{22}\\ \vdots \\ Y_{2J}\\ \vdots \\ Y_{I1}\\ Y_{I2}\\ \vdots \\ Y_{IJ} \end{pmatrix}$$

IJx1 (διάστατο) διάνυσμα

Ορίζω X = (πίνακας σχεδιασμού)

$$X = \begin{array}{c} \\ J\left\{\begin{matrix}\\ \\ \\ \\ \end{matrix}\right. \\ J\left\{\begin{matrix}\\ \\ \\ \\ \end{matrix}\right. \\ \\ J\left\{\begin{matrix}\\ \\ \\ \\ \end{matrix}\right. \end{array} \begin{array}{c} \qquad \xleftarrow{\quad I \quad} \qquad \xleftarrow{\quad J \quad} \\ \left[\begin{array}{ccccc|cccc} 1 & 1 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 1 & 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \\ \hline 1 & 0 & 1 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 1 & \cdots & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 1 & 0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 1 \\ \hline & & & & \vdots & & & & \\ \hline 1 & 0 & 0 & \cdots & 1 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 1 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 1 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \end{array}\right] \end{array}$$

$(IJ)\times(I+J+1)$

$$\underset{\sim}{b} = \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_I \\ b_1 \\ b_2 \\ \vdots \\ b_J \end{pmatrix}$$

$(I+J+1)\times 1$

$$\underset{\sim}{Y} = X\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$$

$$\Updownarrow$$

$$\underset{\sim}{\varepsilon} = \begin{pmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \vdots \\ \varepsilon_{1J} \\ \hline \varepsilon_{21} \\ \varepsilon_{22} \\ \vdots \\ \varepsilon_{2J} \\ \hline \vdots \\ \hline \varepsilon_{I1} \\ \varepsilon_{I2} \\ \vdots \\ \varepsilon_{IJ} \end{pmatrix} \qquad Y_{ij} = \mu + a_i + b_j + \varepsilon_{ij}$$

$i = 1, \dots, I$ και $j = 1, \dots, J$

Υποθέσεις για τα Σφάλματα

① $E(\varepsilon_{ij}) = 0$

② $Var(\varepsilon_{ij}) = \sigma^2$

③ Τα σφάλματα ανά δύο ασυσχέτιστα, δηλ. $Cov(\varepsilon_{ij}, \varepsilon_{k\ell}) = 0$
$\forall\ i = 1, \dots, I,\ j = 1, \dots, J$ με $i \neq k$ και $j \neq \ell$.
$k = 1, \dots, I \quad \ell = 1, \dots, J$

④ $\varepsilon_{ij} \sim N(0, \sigma^2)$

Συνέπειες στα Y_{ij} $i = 1, \dots, I$ $j = 1, \dots, J$

① $E(Y_{ij}) = \mu + a_i + b_j$

② $Var(Y_{ij}) = \sigma^2$

③ Y_{ij} ασυσχέτιστα ανά δύο $i, k = 1, \dots, I$ και $j, \ell = 1, \dots, J$
$Cov(Y_{ij}, Y_{k\ell}) = 0$ με $i \neq k$ και $j \neq \ell$

④ $Y_{ij} \sim N(\mu + a_i + b_j, \sigma^2)$

Ιδιότητες των ΕΕΤ

① Υπό τις υποθέσεις για τα σφάλματα (και τις παραμετρικές συνθήκες), οι ΕΕΤ των μ, a_i, b_j είναι αμερόληπτοι
$E(\hat{\mu}) = \mu$, $E(\hat{a}_i) = a_i$, $E(\hat{b}_j) = b_j$ $i = 1, \dots, I$ $j = 1, \dots, J$

$$E(\hat{\alpha}_i) = E(\overline{Y}_{i.} - \overline{Y}_{..}) = E(\overline{Y}_{i.}) - E(\overline{Y}_{..}) =$$

$$= E\left(\frac{1}{J}\sum_j Y_{ij}\right) - E\left(\frac{1}{IJ}\cdot\sum_i\sum_j Y_{ij}\right) =$$

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

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

$$= \frac{1}{\cancel{J}}\left(\cancel{J}\mu + \cancel{J}\alpha_i + \underbrace{\cancel{\sum_j \beta_j}}_{0}\right) - \frac{1}{\cancel{IJ}}\left(\cancel{IJ}\mu + J\cdot\underbrace{\cancel{\sum_i \alpha_i}}_{0} + I\underbrace{\cancel{\sum_j \beta_j}}_{0}\right)$$

(από πλευρικές συνθήκες)

$$= \cancel{\mu} + \alpha_i - \cancel{\mu} = \alpha_i$$

Ομοίως δείχνω ότι $E(\hat{\beta}_i) = \beta_i$

② Υπό τις υποθέσεις για τα σφάλματα και τις πλευρικές συνθήκες, οι ΕΕΤ των μ, α_i, β_j ταυτίζονται με τους ΕΜΠ των αντίστοιχων παραμέτρων.

Οι ΕΕΤ προκύπτουν από ελαχιστοποίηση της:

$$S^2 = \sum_i\sum_j (Y_{ij} - \mu - \alpha_i - \beta_j)^2 \quad (\alpha)$$

L = Πιθανοφάνεια = Η από κοινού κατανομή των παρατηρήσεων στην $Y_{ij} \overset{\text{ανεξ.}}{=} \prod_{i=1}^{I}\prod_{j=1}^{J} f_{Y_{ij}}(y_{ij}) =$

$$= \prod_i\cdot\prod_j \frac{1}{\sigma\sqrt{2\pi}}\cdot e^{-\frac{1}{2\sigma^2}(y_{ij}-\mu-\alpha_i-\beta_j)^2} =$$

$$= \frac{1}{(\sigma\sqrt{2\pi})^{IJ}}\cdot e^{-\frac{1}{2\sigma^2}\cdot\sum_i\sum_j(y_{ij}-\mu-\alpha_i-\beta_j)^2}$$

Οι ΕΜΠ προκύπτουν από τη μεγιστοποίηση του L ή του log L ή από μεγιστοποίηση της:

$$-IJ\log(\sigma\sqrt{2\pi}) - \frac{1}{2\sigma^2}\sum_i\sum_j (y_{ij}-\mu-a_i-b_j)^2$$

ή από μεγιστοποίηση του:

$$-\sum_i\sum_j (y_{ij}-\mu-a_i-b_j)^2$$

ή από ελαχιστοποίηση του:

$$\sum_i\sum_j (y_{ij}-\mu-a_i-b_j)^2 \quad (**)$$

Από (*), (**) οι ΕΕΤ $(\mu, a_i, b_j) \equiv$ ΕΜΠ (μ, a_i, b_j)

Σπίτι

Άσκηση: Να βρεθεί ο ΕΜΠ σ^2 της κοινής διασποράς σ^2.
Υπόδειξη: Αρκεί να μεγιστοποιηθεί ως προς σ^2 ο $\log L$ (λογάριθμος πιθανοφάνειας), δηλ. $\log L(\hat{\mu}, \hat{a}_i, \hat{b}_j)$.

Άσκηση: Αεροπλάνο κ όργανα

$$\left.\begin{array}{ll} \text{Μεθ.1}: & 52.9,\ 62.1,\ \ldots,\ 53.1 \\ \text{Μεθ.2}: & 58.4,\ 55,\ \ldots,\ 58.4 \\ \text{Μεθ.3}: & \underbrace{71.3,\ 66.6,\ \ldots,\ 67.3}_{\text{χρόνοι φωσφορισμού του οργάνου}} \end{array}\right\}$$

Έχω $\underset{\text{ποσοτ. μεταβλητή}}{Y}$ = χρόνος φωσφορισμού του οργάνου ← εξαρτάται από τη μέθοδο παραγωγής του

Παράγοντας → Μέθοδο Παραγωγής → 3 επίπεδα ← Μεθ.1, Μεθ.2, Μεθ.3

Πρόκειται για πρόβλημα αναδια κατά 1 παράγοντα:
$Y_{ij} = \mu + a_i + \varepsilon_{ij}$ με $i = 1,2,3$ (επίπεδα) $j = 1, \dots, 8$ (μέγεθος)

Υιοθετώ: τις υποθέσεις για τα σφάλματα και την πλευρική συνθήκη για το παραπάνω μοντέλο.

ΠΙΝΑΚΑΣ ΑΝΑΔΙΑ

Πηγή Μεταβλητότητας	SS	β.ε	MS	F-πηλίκο
Παράγοντας (Μέθοδοι Παραγωγής)	$SStr = 581,41$	$I-1=2$	$MStr = 292,005$	$F = \frac{MStr}{MSres} = 17,044$
Υπόλοιπα	$SSres = 360,015$	$N-I=21$	$MSres = 17,144$	
Ολική	$SStot = 944,425$	23		

→ Ελέγχω τη μηδενική υπόθεση ότι τα 3 επίπεδα είναι ταυτόσημα. Δηλ. $H_0: \alpha_1 = \alpha_2 = \alpha_3$.

→ Πως ελέγχεται; - Με το F-πηλίκο

Παρατηρώ: $F = 17,044 >> F_{a,2,21} =$
- $F_{0.05,2,21} = 5,78 << 17,044$
- $F_{0.01,2,21} = 4,32 << 17,044$

Και στις δύο περιπτώσεις F-πηλίκο = 17,044 πολύ μεγαλύτερο, άρα απορρίπτεται η μηδενική (κατά πολύ)

Άρα υπάρχει κάποια μέθοδος που είναι καλύτερη.
→ Πως θα την βρω;

Πολλαπλές Συγκρίσεις: $\bar{Y}_{1\cdot} = 57,1$ $\bar{Y}_{2\cdot} = 59.175$ $\bar{Y}_{3\cdot} = 68.45$

φαίνεται να είναι καλύτερο.

$$ΕΣΔ = t_{\frac{\alpha}{2}, N-I}\sqrt{MSres\left(\frac{1}{J_i}+\frac{1}{J_{i'}}\right)} = t_{\frac{\alpha}{2}, N-I}\sqrt{MSres\frac{2}{J}} \overset{a=5\%}{=\!=} 4,306$$

$J_i = J_{i'} = 8$

Επειδή $|\bar{Y}_{1\cdot} - \bar{Y}_{2\cdot}| = |-2,075| < ΕΣΔ = 4,306$

Άρα η υπόθεση $\alpha_1 = \alpha_2$ δεν μπορεί να απορριφθεί. Άρα είναι σχεδόν ταυτόσημες Μεθ1 ≈ Μεθ.2
(δεν μπορούσα να το ισχυριστώ με σιγουριά κάτι)

Όμοιως, επειδή $|\bar{Y}_{1\cdot} - \bar{Y}_{3\cdot}| = |-11,35| > ΕΣΔ = 4,306$, η υπόθεση $\alpha_1 = \alpha_3$ απορρίπτεται

Επειδή $\bar{Y}_{1\cdot} - \bar{Y}_{3\cdot} < 0 \Rightarrow$ α_3 πιο σημαντικό από το α_1, επομένως Μεθ.1 < Μεθ.3

Επειδή $|\bar{Y}_{2\cdot} - \bar{Y}_{3\cdot}| = |-9,275| > ΕΣΔ = 4,306$, η υπόθεση $\alpha_2 = \alpha_3$ πρέπει να απορριφθεί

Επειδή $\bar{Y}_{2\cdot} - \bar{Y}_{3\cdot} < 0 \Rightarrow$ α_3 πιο σημαντικό από το α_2 επομένως Μεθ.2 < Μεθ.3

Συμπέρασμα: Μεθ.1 ≡ Μεθ.2 < Μεθ.3