# ΑΝΑΛΥΣΗ ΔΙΑΚΥΜΑΝΣΗΣ

Επικ. Καθ. Στέλιος Ζήμερας

Τμήμα Μαθηματικών
Κατεύθυνση Στατιστικής και Αναλογιστικά – Χρηματοοικονομικά Μαθηματικά

2017

# Μοντελοποίηση περιορισμών

Μοντέλο 1:

$$E\left(Y_{jk}\right)=\mu_j \quad \Rightarrow \quad E\left(Y_i\right)=\sum_{j=1}^{J} x_{ij}\mu_j$$

όπου $x_{ij}$ = 1 όταν η απόκριση $Y_i$ αντιστοιχεί στο επίπεδο j ενώ $x_{ij}$=0 στις υπόλοιπες περιπτώσεις.

$$E\left(y\right)=X\beta \quad \text{με} \quad \beta=\begin{pmatrix}\mu_1\\ \mu_2\\ .\\ .\\ .\\ \mu_j\end{pmatrix} \quad \text{και} \quad X=\begin{pmatrix}1 & 0 & \dots & 0\\ 0 & 1 & \dots & 0\\ . & . & O & .\\ . & O & . & 0\\ 0 & . & . & 1\end{pmatrix} \quad \text{όπου} \quad 0 \ \text{και} \ 1 \quad \text{είναι}$$

διανύσματα μήκους K από μηδενικά και μονάδες αντίστοιχα.

# Μοντελοποίηση περιορισμών

Ο $X^T X$ είναι ο $JxJ$ διαγώνιος πίνακας.

$$X^T X = \begin{pmatrix} K & . & . & . & . \\ . & . & . & O & . \\ . & . & K & . & . \\ . & O & . & . & . \\ . & . & . & . & K \end{pmatrix} \quad \text{και} \quad X^T y = \begin{pmatrix} Y_{1.} \\ Y_{2.} \\ . \\ . \\ . \\ Y_{j.} \end{pmatrix}$$

$$b = \frac{1}{K} \begin{pmatrix} Y_{1.} \\ Y_{2.} \\ . \\ . \\ . \\ Y_{j.} \end{pmatrix} = \begin{pmatrix} \bar{y}_1 \\ \bar{y}_2 \\ . \\ . \\ . \\ \bar{y}_j \end{pmatrix} \quad \text{και} \quad b^T X^T y = \frac{1}{K} \sum_{j=1}^{J} Y_{j.}^2$$

Οι προσαρμοσμένες τιμές είναι $\hat{y} = \left[\bar{y}_1, \bar{y}_1, \ldots, \bar{y}_1, \bar{y}_2, \ldots \bar{y}_j\right]^T$

# Μοντελοποίηση περιορισμών

Μοντέλο 2:

$$E\left(Y_{jk}\right) = \mu + \alpha_j$$

Το μ παριστάνει τη μέση επίδραση (αποτέλεσμα) για όλα τα επίπεδα . Το $\alpha_j$ είναι μία επιπρόσθετη επίδραση λόγω του επιπέδου j. Υπάρχουν J +1 παράμετροι.

$$\beta = \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ . \\ . \\ . \\ \alpha_J \end{pmatrix} \quad , \quad X = \begin{pmatrix} 1 & 1 & 0 & \dots & 0 \\ 1 & 0 & 1 & & \\ \vdots & & & O & \\ \vdots & O & & & \\ 1 & & & & 1 \end{pmatrix} \quad \text{όπου τα 1 και 0 είναι διανύσματα}$$

μήκους K.

# Μοντελοποίηση περιορισμών

Συνεπώς $X^T y = \begin{pmatrix} Y_{..} \\ Y_{1.} \\ \vdots \\ Y_{J.} \end{pmatrix}$ και $X^T X = \begin{pmatrix} N & & K & \dots & K \\ K & & K & & \\ \vdots & & & O & \\ \vdots & O & & & \\ K & & & & K \end{pmatrix}$

Η πρώτη γραμμή του (J+1) x (J+1) πίνακα $X^T X$ είναι το άθροισμα των υπόλοιπων γραμμών, οπότε ο πίνακας $X^T X$ είναι μοναδιαίος και δεν υπάρχει μοναδική λύση για τις κανονικές εξισώσεις $X^T X b = X^T y$ . Η γενική λύση μπορεί να γραφτεί ως :

$$b = \begin{pmatrix} \hat{\mu} \\ \hat{\alpha}_1 \\ \vdots \\ \hat{\alpha}_J \end{pmatrix} = \frac{1}{K} \begin{pmatrix} 0 \\ Y_{1.} \\ \vdots \\ Y_{J.} \end{pmatrix} - \lambda \begin{pmatrix} -1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

Όπου το λ είναι μία αυθαίρετη σταθερά.

# Μοντελοποίηση περιορισμών

Παραδοσιακά απαιτούμε να ισχύει ο περιορισμός του αθροίσματος στο μηδέν

$\sum_{j=1}^{J} \alpha_j = 0$ , οπότε $\frac{1}{K}\sum_{j=1}^{J} Y_{j.} - J\lambda = 0$ και συνεπώς θα είναι $\lambda = \frac{1}{JK}\sum_{j=1}^{J} Y_{j.} = \frac{Y_{..}}{N}$

Αυτά δίνουν τη λύση $\hat{\mu} = \frac{Y_{..}}{N}$ και $\hat{\alpha}_j = \frac{Y_{j.}}{K} - \frac{Y_{..}}{N}$ για $j = 1, \ldots, J$.

Οπότε έχουμε $b^T X^T y = \frac{Y_{..}^2}{N} + \sum_{j=1}^{J} Y_{j.}\left(\frac{Y_{j.}}{K} - \frac{Y_{..}}{N}\right) = \frac{1}{K}\sum_{j=1}^{J} Y_{j.}^2$

# Μοντελοποίηση περιορισμών

## Μοντέλο 3:

$$E(Y_{jk}) = \mu + \alpha_j \Rightarrow$$

$\alpha_1 = 0$ λέγεται γωνιακή παραμετροποίηση.

Για αυτή την έκδοση έχουμε $J$ παραμέτρους.

$$\beta = \begin{pmatrix} \mu \\ \alpha_2 \\ \vdots \\ \alpha_j \end{pmatrix} \quad \text{και} \quad X = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 1 & 1 & & \\ \vdots & . & . & O \\ \vdots & O & . & . \\ 1 & & & 1 \end{pmatrix} \Rightarrow X^T y = \begin{pmatrix} Y_{..} \\ Y_{2.} \\ \vdots \\ Y_{j.} \end{pmatrix} \quad \text{και} \quad X^T X = \begin{pmatrix} N & K & \dots & K \\ K & K & \dots & O \\ \vdots & \dots & \dots & \vdots \\ K & O & \dots & K \end{pmatrix}$$

Ο JxJ πίνακας $X^T X$ δεν είναι μοναδιαίος, συνεπώς υπάρχει μία μοναδική

$$\text{λύση} \; : \; b = \frac{1}{K}\begin{pmatrix} Y_{1.} \\ Y_{2.} - Y_{1.} \\ \vdots \\ Y_{j.} - Y_{1.} \end{pmatrix} \Rightarrow b^T X^T y = \frac{1}{K}\left[ Y_{..} Y_{1.} + \sum_{j=2}^{J} Y_{j.} \left( Y_{j.} - Y_{1.} \right) \right] = \frac{1}{K}\sum_{j=1}^{J} Y_{j.}^2$$

# Παράδειγμα

**Πίνακας 15 : Βάρη αποξηραμένων φυτών, σε συνθήκες ελέγχου και δύο θεραπειών**

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Έλεγχος | 4.17 | 5.58 | 5.18 | 6.11 | 4.50 | 4.61 | 5.17 | 4.53 | 5.33 | 5.14 |
| Θεραπεία Α | 4.81 | 4.17 | 4.41 | 3.59 | 5.87 | 3.83 | 6.03 | 4.89 | 4.32 | 4.69 |
| Θεραπεία Β | 6.31 | 5.12 | 5.54 | 5.50 | 5.37 | 5.29 | 4.92 | 6.15 | 5.80 | 5.26 |

Εδώ έχουμε N=30 , J=3 , K=10

$\frac{Y_{..}^2}{N} = 772.0599$ (εδώ N = 30 ) και $\frac{1}{K}\sum_{j=1}^{J} Y_{j.}^2 = 775.8262$ (εδώ $K$=10, $J$=3)

$$\sum_{j=1}^{J}\sum_{k=1}^{K} Y_{jk}^2 = 786.3138$$

| Πηγή Μεταβλητότητας | Βαθμοί Ελευθερίας | Άθροισμα Τετραγώνων | Μέσο Άθροισμα | f |
|---|---|---|---|---|
| Μέσος | 1 | 772.0599 | | |
| Μεταξύ Θεραπειών | 2 | 3.7663 | 1.883 | 4.85 |
| Κατάλοιπα | 27 | 10.4921 | 0.389 | |
| Σύνολο | 30 | 786.3138 | | |

# Παράδειγμα

Εφ' όσον η τιμή f=4.85 είναι σημαντική στο επίπεδο 5% συγκρινόμενη με την τιμή $F_{2,\ 27} = 3,\ 3541$ της κατανομής, βγάζουμε το συμπέρασμα ότι οι μέσοι των ομάδων διαφέρουν.

Για να ερευνήσουμε περαιτέρω το αποτέλεσμα αυτό, είναι βολικό να χρησιμοποιήσουμε τη πρώτη μορφή του μοντέλου $E\left(Y_{jk}\right) = \mu_j$. Οι εκτιμηθέντες μέσοι είναι : $b = \begin{pmatrix} \hat{\mu}_1 \\ \hat{\mu}_2 \\ \hat{\mu}_3 \end{pmatrix} = \begin{pmatrix} 5.032 \\ 4.661 \\ 5.526 \end{pmatrix}$