Πολλαπλό Γραμμικό Υπόδειγμα

Η Υ εξαρτάται από κ-1 ερμηνευτικές μεταβλητές $X_2,\ldots,X_k$

$$Y_i=\beta_1+\beta_2 X_{i2}+\beta_3 X_{i3}+\ldots+\beta_k X_{i,k}+\varepsilon_i\ ,\quad \forall i=1,2,\ldots,\mathrm{N}$$

Περιγραφή με μήτρες:

$Y=X\beta+\varepsilon$ όπου

$$\begin{bmatrix}Y_1\\Y_2\\\vdots\\\vdots\\Y_n\end{bmatrix}_{n\times1}=\begin{bmatrix}1&X_{12}&X_{13}&\ldots&X_{1k}\\1&X_{22}&X_{23}&\ldots&X_{2k}\\\vdots&&&&\\\vdots&&&&\\1&X_{n2}&X_{n3}&\ldots&X_{nk}\end{bmatrix}_{n\times k}\cdot\begin{bmatrix}\beta_1\\\beta_2\\\vdots\\\beta_\kappa\end{bmatrix}_{k\times1}+\begin{bmatrix}\varepsilon_1\\\varepsilon_2\\\vdots\\\vdots\\\varepsilon_n\end{bmatrix}_{n\times1}$$

Παρατήρηση: Ο $X'X$ και ο $(X'X)^{-1}$ είναι συμμετρικός πίνακας. Ισχύει

$$X'X=\begin{bmatrix}1&1&1&\ldots&1\\X_{12}&X_{22}&\ldots&\ldots&X_{n2}\\X_{13}&X_{23}&\ldots&\ldots&X_{n3}\\\vdots&&&&\\X_{1k}&X_{2k}&\ldots&\ldots&X_{nk}\end{bmatrix}_{kxn}\begin{bmatrix}1&X_{12}&X_{13}&\ldots&X_{1k}\\1&X_{22}&X_{23}&\ldots&X_{2k}\\\vdots&&&&\\\vdots&&&&\\1&X_{n2}&X_{n3}&\ldots&X_{nk}\end{bmatrix}_{n\times k}\Rightarrow$$

$$X'X=\begin{bmatrix}N&\sum X_{i2}&\sum X_{i3}&\ldots&\sum X_{ik}\\\sum X_{i2}&\sum X_{i2}^2&\sum X_{i2}X_{i3}&\ldots&\sum X_{i2}X_{ik}\\\sum X_{i3}&\sum X_{i2}X_{i3}&&&\\\vdots&&&&\\\sum X_{ik}&\sum X_{i2}X_{ik}&\ldots&\ldots&\sum X_{ik}^2\end{bmatrix}_{k\times k}$$

και $$X'Y=\begin{bmatrix}1&1&1&\ldots&1\\X_{12}&X_{22}&\ldots&\ldots&X_{n2}\\X_{13}&X_{23}&\ldots&\ldots&X_{n3}\\\vdots&&&&\\X_{1k}&X_{2k}&\ldots&\ldots&X_{nk}\end{bmatrix}_{kxn}\begin{bmatrix}Y_1\\Y_2\\\vdots\\Y_n\end{bmatrix}_{nx1}=\begin{bmatrix}\sum \mathrm{Y_i}\\\sum X_{i2}Y_i\\\vdots\\\sum X_{ik}Y_i\end{bmatrix}_{kx1}$$

Το πρώτο στοιχείο του πίνακα $X'Y$ ισούται με το άθροισμα των Υ.
Άρα μπορώ να βρώ τον

$$\bar{Y}=\frac{\sum \mathrm{Y}}{N}$$

Το Ν από τον $X'X$

<u>**Ελάχιστα Τετράγωνα στη Πολυμεταβλητή Παλινδρόμηση:**</u>

$\hat{Y}=X\hat{\beta}$ άρα $\hat{\varepsilon}=Y-\hat{Y}=Y-X\hat{\beta}$ το διάνυσμα των καταλοίπων

Ελαχιστοποίηση: $RSS=\sum_{i=1}^{n}\hat{\varepsilon}_i^2=\hat{\varepsilon}'\hat{\varepsilon}=\left(Y-X\hat{\beta}\right)'\left(Y-X\hat{\beta}\right)$

Εκτιμητής Ε.Τ. : $\underline{\hat{\beta}=(X'X)^{-1}X'Y}$

Ερμηνεία των συντελεστών κλίσης $\hat{\beta}_j$, j=2,…,k **:** μεταβολή στην αναμενόμενη τιμή της Υ όταν η X_j μεταβάλλεται κατά μία μονάδα, <u>ενώ οι υπόλοιπες ανεξάρτητες (ερμηνευτικές) μεταβλητές παραμένουν σταθερές</u> (*ceteris paribus*)

Στατιστικός Έλεγχος Υποδείγματος

Η διακύμανση των σφαλμάτων, σ^2 εκτιμάται από τη ποσότητα $\boxed{\hat{\sigma}^2 = S^2 = \frac{\sum \hat{\varepsilon}_i^2}{N-k} = \frac{RSS}{N-k}}$

Ο πίνακας διακυμάνσεων-συνδιακυμάνσεων (covariance matrix) εκτιμάται από

$$Var(\hat{\beta}) = S^2(X'X)^{-1} = \begin{bmatrix} Var(\hat{\beta}_1) & Cov(\hat{\beta}_1,\hat{\beta}_2) & \dots & Cov(\hat{\beta}_1,\hat{\beta}_k) \\ Cov(\hat{\beta}_2,\hat{\beta}_1) & Var(\hat{\beta}_2) & \dots & Cov(\hat{\beta}_2,\hat{\beta}_k) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(\hat{\beta}_k,\hat{\beta}_1) & Cov(\hat{\beta}_k,\hat{\beta}_2) & \dots & Var(\hat{\beta}_k) \end{bmatrix}$$

δηλαδή οι ρίζες των διαγωνίων στοιχείων του πίνακα, δίνουν τα τυπικά σφάλματα των συντελεστών $se\left(\hat{\beta}_j\right)$.

Έλεγχος υποθέσεων – t Έλεγχος:

$\boxed{H_0 : \beta_j = c, \;\; j = 1,2..,k}$, Στατιστική ελέγχου: $t = \frac{\hat{\beta}_j - c}{\mathrm{SE}(\hat{\beta}_j)} \sim t_{N-k}$

Με εναλλακτική υπόθεση,

- $H_a : \beta_j \neq c$ (δικατάληκτος έλεγχος) η H_0 απορρίπτεται σε επ.σημ. α αν $|t| > t_{(a/2,N-k)}$
- $H_a : \beta_j > c$ (μονοκατάληκτος έλεγχος) η H_0 απορρίπτεται σε επ.σημ. α αν $t > t_{(a,N-k)}$
- $H_a : \beta_j < c$ (μονοκατάληκτος έλεγχος) η H_0 απορρίπτεται σε επ.σημ. α αν $t < -t_{(a,N-k)}$

Ειδικότερα αν c=0 τότε,

- Αν η H_0 απορριφθεί τότε ο β_j είναι στατιστικά σημαντικός, δηλαδή η X_j επηρεάζει σημαντικά τη Υ.
- Αν η H_0 δεν απορριφθεί τότε ο β_j δεν είναι στατιστικά σημαντικός και η μεταβλητή X_j δε φαίνεται να επηρεάζει σημαντικά τη Υ

Διαστήματα Εμπιστοσύνης: (1-α)% Δ.Ε. του β_j: $\hat{\beta}_j \pm t_{N-k,a/2} \cdot Se\left(\hat{\beta}_j\right)$

Διάστημα Πρόβλεψης για την Y_0:

$\hat{Y}_0 = \hat{\beta}_1 + \hat{\beta}_2 X_{02} + \dots + \hat{\beta}_k X_{0k}$ η πρόβλεψη της Υ για ένα δεδομένο διάνυσμα των ανεξάρτητων μεταβλητών $x_0 = \begin{bmatrix} 1 \\ X_{02} \\ .. \\ X_{0k} \end{bmatrix}$ (ισοδύναμα με μήτρες $\hat{Y}_0 = x_0{}'\hat{\beta}$).

100(1-α)% Δ.Ε. για την Y_0

$$\hat{Y}_0 - t_{N-k,a/2}\sqrt{Var(\hat{Y}_0 - Y_0)} \le Y_0 \le \hat{Y}_0 + t_{N-k,a/2}\sqrt{Var(\hat{Y}_0 - Y_0)}$$

This page intentionally left blank

ΠΑΡΑΔΕΙΓΜΑ ΕΞΕΤΑΣΗΣ ΣΤΗΝ ΟΙΚΟΝΟΜΕΤΡΙΑ Ι

ΘΕΜΑ 1 (ΥΠΟΧΡΕΩΤΙΚΟ, ΜΟΝΑΔΕΣ 6/10)

Θεωρείστε την συνάρτηση ζήτησης: $Y_t = \beta_1 + \beta_2 X_{t2} + \beta_3 X_{t3} + u_t, t = 1,2,\ldots,T = 25$ *παρατηρήσεις.*

ή σε μορφή γραμμικής άλγεβρας $Y = X\beta + u$

Αν

$$(X'X)^{-1} = \begin{bmatrix} 1.70 & -0.06 & -0.08 \\ -0.06 & 0.92 & 0.55 \\ -0.08 & 0.55 & 0.54 \end{bmatrix}, X'Y = \begin{bmatrix} -1.60 \\ -30.0 \\ 47.0 \end{bmatrix}, Y'Y = \sum Y_t^2 = 1.800 \text{ και}$$

$u'u = \sum u_t^2 = 6.0$, τότε απαντήστε στα παρακάτω ερωτήματα:

Α) Να βρεθούν με την μέθοδο των ελαχίστων τετραγώνων οι εκτιμήσεις των συντελεστών της παραπάνω συνάρτησης ζήτησης, τα τυπικά τους σφάλματα και να διεξαχθούν οι έλεγχοι αν οι συντελεστές αυτοί είναι στατιστικά σημαντικοί. Δίνεται επίπεδο σημαντικότητας α=5%.

Β)Να βρεθεί και να ερμηνευτεί ο συντελεστής προσδιορισμού R^2. Επίσης να βρεθεί ο διορθωμένος συντελεστής προσδιορισμού $\bar{R}^2$.

Β2) Να ελεγχθεί με F έλεγχο αν ισχύει $\beta_2 = \beta_3 = 0$ (σημαντικότητα υποδείγματος)

Γ) Ελέγξατε αν ισχύουν από κοινού οι παρακάτω περιορισμοί

$$\beta_2 = -\beta_3 \quad , \quad \beta_1 = 0$$

Γ2)Ελέγξτε αν ισχύει ξεχωριστά ο περιορισμός $\beta_3 = -2\beta_1$

Δ) Αν η τιμή του αγαθού είναι $X_{02} = 3.0$ και το εισόδημα του καταναλωτή είναι $X_{03} = 10.0$, τότε βρείτε ποια θα είναι η προβλεπόμενη ζητούμενη ποσότητα Yο και το διάστημα εμπιστοσύνης της.

Ε) Για να ελέγξει την σωστή εξειδίκευση του διαταρακτικού όρου ένας ερευνητής εκτίμησε τις ακόλουθες βοηθητικές παλινδρομήσεις:

$u_t^2 = 2.5 + 1.2X_{t2}, \quad R^2 = 0.18$ (1)

$u_t^2 = 2.5 + 1.2X_{t2} - 1.4X_{t1} \cdot X_{t2}, \quad R^2 = 0.18$ (2)

$u_t = -543.1 + 0.002X_{1t} - 1.23X_{2t} + 0.781u_{t-1}, \quad R^2 = 0.08$ (3)

$u_t = -544.6 + 0.002X_{1t} - 1.3X_{2t} + 0.780u_{t-1} + 0.002u_{t-2}, \quad R^2 = 0.18$ (4)

Με βάση τα αποτελέσματα των βοηθητικών εκτιμήσεων αυτών εξηγήστε τι προβλήματα παρουσιάζονται στα κατάλοιπα του υποδείγματος u_t και <u>προτείνετε</u> έναν εκτιμητή ή κάποια μέθοδο που να τα διορθώνει.

(όλοι οι έλεγχοι της άσκησης να γίνουν σε επίπεδο σημαντικότητας α=5%)