Εισαγωγή στην οικονομετρία (1404)

Διάλεξη 10

Άγγελος Αλεξόπουλος
Οικονομικό Πανεπιστήμιο Αθηνών -
Τμήμα Οικονομικής Επιστήμης
angelos@aueb.gr

Περίγραμμα Διάλεξης

✓ Εκτιμητής Ελαχίστων Τετραγώνων (ΕΤ)
- Σταθερός όρος & Συντελεστής κλίσης
- Αποκλίσεις από τον μέσο

✓ Ιδιότητες καταλοίπων και εξαρτημένης μεταβλητής

✓ Παράδειγμα εκτίμησης ΕΤ

Κλασικές υποθέσεις

Αν X_i είναι σταθερή σε επαναλαμβανόμενα δείγματα

1. Γραμμικότητα ως προς τις παραμέτρους
2. $E(u_i) = 0 \; or \; E(u_i|x_i) = 0 \; \forall i$
3. $var(u_i) = E(u_i^2) = \sigma^2$ or $var(u_i|x_i) = E(u_i^2|x_i) = \sigma^2, \forall i$
4. $cov(u_i, u_j) = 0, \forall i \neq j$
5. $u_i \sim N(0, \sigma^2), \forall i$

Εκτίμηση συντελεστών απλού γραμμικού υποδείγματος

- Συνάρτηση παλινδρόμησης πληθυσμού (population regression function, PRF)

$$Y_i = \beta_0 + \beta_1 X_i + u_i, i = 1, \dots, n \quad (1)$$

- Συνάρτηση παλινδρόμησης δείγματος (sample regression function, SRF): εκτιμημένο υπόδειγμα

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i, \forall i$$

$\hat{Y}_i$ εκτιμητής $E(Y|X_i)$

$\hat{\beta}_0$ και $\hat{\beta}_1$ εκτιμητές των β_0 και β_1

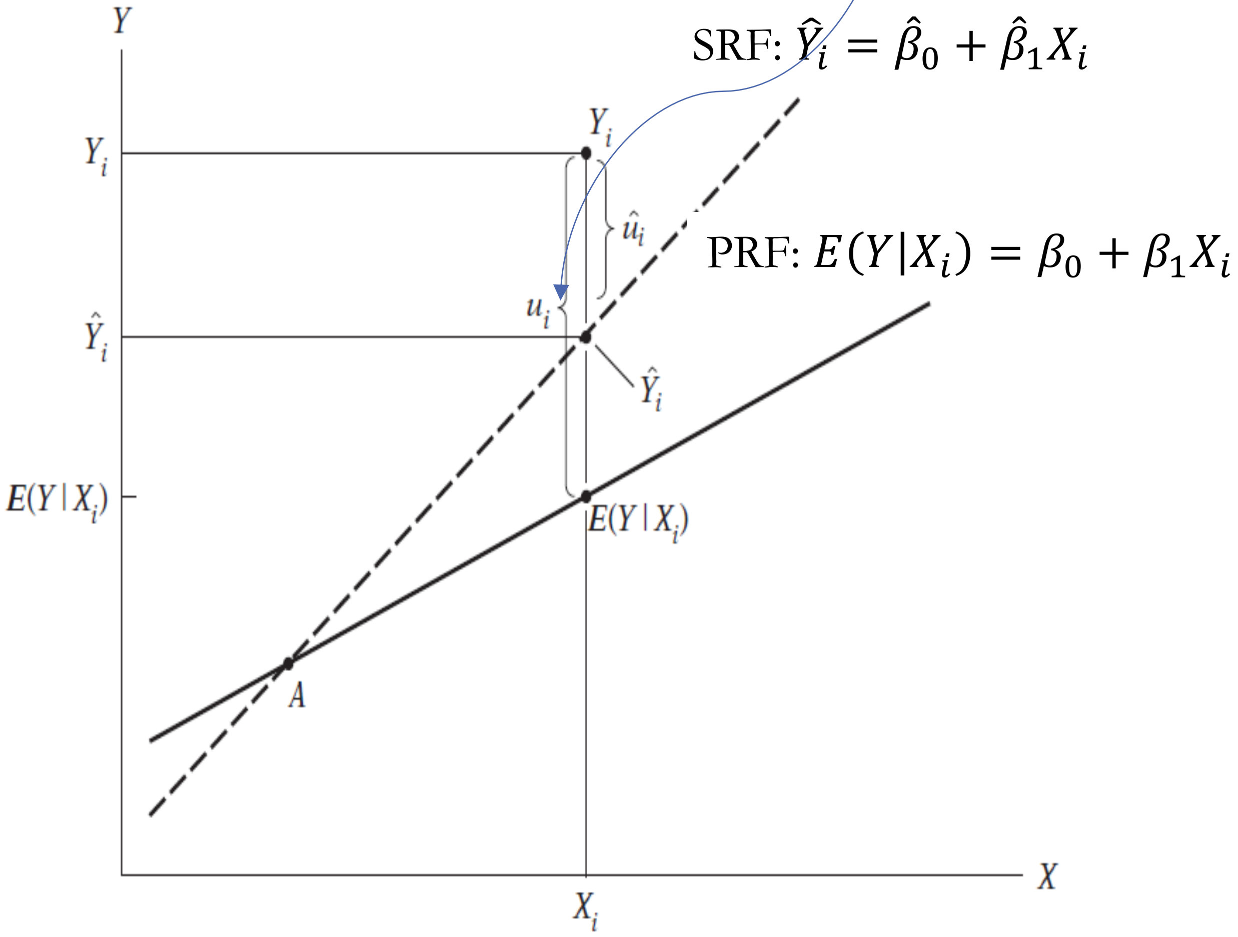
Κατάλοιπα

$$\boxed{u_i = Y_i - E(Y|X_i)}$$

✓ Κατάλοιπα (residuals) της εκτίμησης του γραμμικού υποδείγματος: απόσταση μεταξύ παρατηρούμενων και εκτιμημένων τιμών της εξαρτημένης μεταβλητής

$$\hat{u}_i = Y_i - \hat{Y}_i$$

- Οι τιμές των καταλοίπων μπορεί να είναι θετικές ή αρνητικές
- Ανάλογα αν οι παρατηρούμενες τιμές της Y_i βρίσκονται πάνω ή κάτω από τις εκτιμήσεις $\hat{Y}_i$



Μέθοδος Ελαχίστων Τετραγώνων

✓ Μέθοδος ελαχίστων τετραγώνων (ΕΤ, ordinary least squares, OLS ή LS): εκτιμητές (εκτιμήσεις) για τους συντελεστές β_0 και β_1 που ελαχιστοποιούν την συνάρτηση του αθροίσματος των τετραγώνων των καταλοίπων (residual sum of squares, RSS):

$$\mathrm{RSS}(\hat{\beta}_0, \hat{\beta}_1) \equiv \sum_{i=1}^{n} \hat{u}_i^2$$

✓Το **προβλημα της μεθόδου ΕΤ** λύνεται ως εξής:

$$\underset{\hat{\beta}_0,\hat{\beta}_1}{\text{ελαχ}} \ \mathrm{RSS}(\hat{\beta}_0,\hat{\beta}_1) \equiv \underset{\hat{\beta}_0,\hat{\beta}_1}{\text{ελαχ}} \ \sum_{i=1}^{n} \hat{u}_i^2 = \underset{\hat{\beta}_0,\hat{\beta}_1}{\text{ελαχ}} \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2$$

$$= \underset{\hat{\beta}_0,\hat{\beta}_1}{\text{ελαχ}} \ \sum_{i=1}^{n} \left(Y_i - (\hat{\beta}_0 + \hat{\beta}_1 X_i)\right)^2 = \underset{\hat{\beta}_0,\hat{\beta}_1}{\text{ελαχ}} \sum_{i=1}^{n} (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)^2 \quad (2)$$

- Οι λύσεις που προκύπτουν από το πρόβλημα (2) είναι οι εκτιμητές ελαχίστων τετραγώνων

Μέθοδος Ελαχίστων Τετραγώνων

Για τη λύση της (2) απαιτείται ικανοποίηση των συνθηκών πρώτης και δεύτερης τάξης:

Συνθήκες πρώτης τάξης (first order conditions - FOC):

$$\frac{\partial}{\partial \hat{\beta}_0} \sum_{i=1}^{n} \hat{u}_i^2 = 0 \text{ και } \frac{\partial}{\partial \hat{\beta}_1} \sum_{i=1}^{n} \hat{u}_i^2 = 0 \qquad (3α)$$

Συνθήκες δεύτερης τάξης (second order conditions-SOC):

$$\frac{\partial^2}{\partial \hat{\beta}_0^2} \sum_{i=1}^{n} \hat{u}_i^2 > 0 \text{ και } \frac{\partial^2}{\partial \hat{\beta}_1^2} \sum_{i=1}^{n} \hat{u}_i^2 > 0 \qquad (3β)$$

- από τις συνθήκες πρώτης τάξης προκύπτουν οι εκτιμητές ΕΤ
- οι συνθήκες δεύτερης τάξης είναι οι ικανές συνθήκες για την επίτευξη του ελάχιστου στην συνάρτηση (2)

Μέθοδος Ελαχίστων Τετραγώνων

Από τις συνθήκες πρώτης τάξης προκύπτουν οι εκτιμητές ελαχίστων τετραγώνων $\hat{\beta}_0$ και $\hat{\beta}_1$ ως

$$\frac{\partial}{\partial\hat{\beta}_0}\sum_{i=1}^{n}\hat{u}_i^2 = \frac{\partial}{\partial\hat{\beta}_0}\sum_{i=1}^{n}(Y_i-\hat{\beta}_0-\hat{\beta}_1X_i)^2=0$$

$$=\sum_{i=1}^{n}\frac{\partial}{\partial\hat{\beta}_0}(Y_i-\hat{\beta}_0-\hat{\beta}_1X_i)^2=0$$

$$=\sum_{i=1}^{n}2\left(Y_i-\hat{\beta}_0-\hat{\beta}_1X_i\right)\frac{\partial}{\partial\hat{\beta}_0}(Y_i-\hat{\beta}_0-\hat{\beta}_1X_i)=0$$

$$=\sum_{i=1}^{n}2\left(Y_i-\hat{\beta}_0-\hat{\beta}_1X_i\right)(-1)=0$$

$$\Rightarrow \sum_{i=1}^{n}\left(Y_i-\hat{\beta}_0-\hat{\beta}_1X_i\right)=0 \text{ ή } \sum_{i=1}^{n}Y_i=n\hat{\beta}_0+\hat{\beta}_1\sum_{i=1}^{n}X_i \qquad (4\alpha)$$

Μέθοδος Ελαχίστων Τετραγώνων

$\kappa\alpha\iota$

$$\frac{\partial}{\partial \hat{\beta}_1} \sum_{i=1}^{n} \hat{u}_i^2 = \frac{\partial}{\partial \hat{\beta}_1} \sum_{i=1}^{n} (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)^2 = 0$$

$$= \sum_{i=1}^{n} \frac{\partial}{\partial \hat{\beta}_1} (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)^2 = 0$$

$$= \sum_{i=1}^{n} 2\left(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i\right) \frac{\partial}{\partial \hat{\beta}_1} (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i) = 0$$

$$= \sum_{i=1}^{n} 2\left(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i\right)(-X_i) = 0$$

$$\Rightarrow \sum_{i=1}^{n}\left(Y_i X_i - \hat{\beta}_0 X_i - \hat{\beta}_1 X_i^2\right) = 0 \text{ ή } \sum_{i=1}^{n} Y_i X_i = \hat{\beta}_0 \sum_{i=1}^{n} X_i + \hat{\beta}_1 \sum_{i=1}^{n} X_i^2 \quad (4\beta)$$

Μέθοδος Ελαχίστων Τετραγώνων

Το σύστημα αυτό των δύο εξισώσεων (4α)-(4β), που συνοψίζεται ως

$$\left.\begin{aligned} &(4\alpha): && \sum_{i=1}^{n} Y_i = n\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^{n} X_i \\ &(4\beta): && \sum_{i=1}^{n} Y_i X_i = \hat{\beta}_0 \sum_{i=1}^{n} X_i + \hat{\beta}_1 \sum_{i=1}^{n} X_i^2 \end{aligned}\right\}$$

αναφέρεται στη βιβλιογραφία ως **σύστημα κανονικών εξισώσεων**

Μέθοδος Ελαχίστων Τετραγώνων

Λύνοντάς το σύστημα εξισώσεων με τη μέθοδο των οριζουσών (ή μέθοδος Cramer), έχουμε τους εξής γενικούς τύπους των εκτιμητών ελαχίστων τετραγώνων:

$$\hat{\beta}_0 = \frac{\begin{vmatrix} \sum_{i=1}^{n} Y_i & \sum_{i=1}^{n} X_i \\ \sum_{i=1}^{n} Y_i X_i & \sum_{i=1}^{n} X_i^2 \end{vmatrix}}{\begin{vmatrix} n & \sum_{i=1}^{n} X_i \\ \sum_{i=1}^{n} X_i & \sum_{i=1}^{n} X_i^2 \end{vmatrix}} = \frac{\sum_{i=1}^{n} Y_i \sum_{i=1}^{n} X_i^2 - \sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i X_i}{n \sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i\right)^2} \tag{5α}$$

Μέθοδος Ελαχίστων Τετραγώνων

$\kappa\alpha\iota$

$$\hat{\beta}_1 = \frac{\begin{vmatrix} n & \sum_{i=1}^{n} Y_i \\ \sum_{i=1}^{n} X_i & \sum_{i=1}^{n} Y_i X_i \end{vmatrix}}{\begin{vmatrix} n & \sum_{i=1}^{n} X_i \\ \sum_{i=1}^{n} X_i & \sum_{i=1}^{n} X_i^2 \end{vmatrix}} = \frac{n \sum_{i=1}^{n} Y_i X_i - \sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i}{n \sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i\right)^2}$$

(5β)

Μέθοδος Ελαχίστων Τετραγώνων

Από τις δεύτερες παραγώγους της συνάρτησης (2) έχουμε:

$$\frac{\partial^2}{\partial \hat{\beta}_0^2} \sum_{i=1}^{n} \hat{u}_i^2 = \frac{\partial}{\partial \hat{\beta}_0} \sum_{i=1}^{n} 2(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)\ (-1) = 2 \sum_{i=1}^{n} 1 = 2n > 0$$

και

$$\frac{\partial^2}{\partial \hat{\beta}_1^2} \sum_{i=1}^{n} \hat{u}_i^2 = \frac{\partial}{\partial \hat{\beta}_1} \sum_{i=1}^{n} 2\left(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i\right)\ (-X_i) = 2 \sum_{i=1}^{n} X_i^2 > 0$$

οι θετικές τιμές επιβεβαιώνουν ότι οι λύσεις των LS εκτιμητών αποτελούν λύσεις στο ελάχιστο σημείο της συνάρτησης RSS($\hat{\beta}_0$,$\hat{\beta}_1$).

Ο εκτιμητής ΕΤ σε αποκλίσεις από τους μέσους

✓Εναλλακτική παρουσίαση εκτιμητών ΕΤ: αποκλίσεις των τιμών των μεταβλητών X_i και Y_i από τους μέσους τους στο δείγμα, $X_i - \bar{X}$ και $Y_i - \bar{Y}$

- καλύτερη ερμηνεία του εκτιμητή ΕΤ ιδιαίτερα του συντελεστή κλίσης $\hat{\beta}_1$
- απλοποιεί σημαντικά κάποια από τα θεωρητικά αποτελέσματα για τον εκτιμητή (στατιστικές ιδιότητες εκτιμητή)

Ο εκτιμητής ΕΤ σε αποκλίσεις από τους μέσους

- Από την κανονική εξίσωση (4α) συνεπάγεται ότι ο εκτιμητής $\hat{\beta}_0$ γράφεται ως:

$$\sum_{i=1}^{n} Y_i = n\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^{n} X_i \Rightarrow$$

$$\hat{\beta}_0 = \frac{1}{n}\sum_{i=1}^{n} Y_i - \hat{\beta}_1 \frac{1}{n}\sum_{i=1}^{n} X_i \quad \Rightarrow$$

$$\hat{\beta}_0 = \overline{\mathrm{Y}} - \hat{\beta}_1 \overline{\mathrm{X}}$$

όπου $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ και $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$ οι μέσοι των τιμών των X_i και Y_i στο δείγμα (δειγματικοί μέσοι)

Ο εκτιμητής ΕΤ σε αποκλίσεις από τους μέσους

- Αντικαθιστώντας τη λύση του εκτιμητή $\hat{\beta}_0$

$$\hat{\beta}_0 = \overline{\mathrm{Y}} - \hat{\beta}_1 \overline{\mathrm{X}}$$

στην κανονική εξίσωση (4β)

$$\sum_{i=1}^{n} Y_i X_i = \hat{\beta}_0 \sum_{i=1}^{n} X_i + \hat{\beta}_1 \sum_{i=1}^{n} X_i^2$$

έχουμε:

$$\sum_{i=1}^{n} Y_i X_i = (\bar{Y} - \hat{\beta}_1 \bar{X}\ ) \sum_{i=1}^{n} X_i + \hat{\beta}_1 \sum_{i=1}^{n} X_i^2$$

Διαιρώντας με n και τα δύο μέλη της τελευταίας σχέσης και κάνοντας πράξεις θα πάρουμε τον εκτιμητή $\hat{\beta}_1$ σε αποκλίσεις από τους μέσους

Ο εκτιμητής ΕΤ σε αποκλίσεις από τους μέσους

$$\sum_{i=1}^{n} Y_i X_i = (\bar{Y} - \hat{\beta}_1 \bar{X}) \sum_{i=1}^{n} X_i + \hat{\beta}_1 \sum_{i=1}^{n} X_i^2$$

$$\frac{1}{n}\sum_{i=1}^{n} Y_i X_i = (\bar{Y} - \hat{\beta}_1 \bar{X}) \frac{1}{n}\sum_{i=1}^{n} X_i + \hat{\beta}_1 \frac{1}{n}\sum_{i=1}^{n} X_i^2 \Rightarrow$$

$$\frac{1}{n}\sum_{i=1}^{n} Y_i X_i - \bar{Y}\frac{1}{n}\sum_{i=1}^{n} X_i = \hat{\beta}_1 \left(\frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}\frac{1}{n}\sum_{i=1}^{n} X_i\right) \Rightarrow$$

$$\frac{1}{n}\sum_{i=1}^{n} Y_i X_i - \bar{Y}\bar{X} = \hat{\beta}_1 \left(\frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}^2\right) \Rightarrow$$

$$\frac{1}{n}\sum_{i=1}^{n} (Y_i X_i - \bar{Y}\bar{X}) = \hat{\beta}_1 \frac{1}{n}\sum_{i=1}^{n} \left(X_i^2 - \bar{X}^2\right) \Rightarrow$$

$$\hat{\beta}_1 = \frac{\frac{1}{n}\sum_{i=1}^{n}(Y_i X_i - \bar{Y}\bar{X})}{\frac{1}{n}\sum_{i=1}^{n}\left(X_i^2 - \bar{X}^2\right)} = \frac{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \bar{Y})(X_i - \bar{X})}{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2} = \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})(X_i - \bar{X})}{\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

Ο εκτιμητής ΕΤ σε αποκλίσεις από τους μέσους

$$\hat{\beta}_1 = \frac{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \bar{Y})(X_i - \bar{X})}{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2} = \frac{\hat{\sigma}_{YX}}{\hat{\sigma}_X^2} \text{ ή } \frac{s_{YX}}{s_X^2}$$

όπου $\hat{\sigma}_{YX} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \bar{Y})(X_i - \bar{X})$ και $\hat{\sigma}_X^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$ οι εκτιμητές της συνδιακύμανσης μεταξύ των τιμών των μεταβλητών X_i και Y_i και της διακύμανσης των τιμών της X_i με βάση τις παρατηρήσεις του δείγματος

Ο εκτιμητής ΕΤ σε αποκλίσεις από τους μέσους

$$\hat{\beta}_1 = \frac{\hat{\sigma}_{YX}}{\hat{\sigma}_X^2} = \frac{\hat{r}_{YX}\hat{\sigma}_Y\hat{\sigma}_X}{\hat{\sigma}_X^2} = \hat{r}_{YX}\frac{\hat{\sigma}_Y}{\hat{\sigma}_X}$$

- όπου $\hat{r}_{YX} = \frac{\hat{\sigma}_{YX}}{\hat{\sigma}_Y\hat{\sigma}_X}$ (εκτιμητής του συντελεστή συσχέτισης των τιμών των μεταβλητών Y_i και X_i)
- και $\hat{\sigma}_Y = \sqrt{\hat{\sigma}_Y^2}$, $\hat{\sigma}_X = \sqrt{\hat{\sigma}_X^2}$ (εκτιμητές των τυπικών αποκλίσεων) και $\hat{\sigma}_{YX}$ εκτιμητής συνδιακύμανσης

➢ Ο εκτιμητής ΕΤ $\hat{\beta}_1$ εξαρτάται από κοινού από τη σχετική μεταβλητότητα (διακύμανση) των μεταβλητών X_i και Y_i στο δείγμα και το βαθμό συσχέτισής τους

➢ Αν η $\hat{\sigma}_X$ είναι πολύ μεγαλύτερη σε σχέση με την $\hat{\sigma}_Y$ τότε $\frac{\hat{\sigma}_Y}{\hat{\sigma}_X}$ κοντά στο μηδέν, οι τιμές του $\hat{\beta}_1$ μπορεί να είναι πολύ κοντά στο μηδέν, παρόλο που οι μεταβλητές X_i και Y_i μπορεί να έχουν μεγάλη συσχέτιση

Το γραμμικό υπόδειγμα σε αποκλίσεις από τους μέσους

Από τον ορισμό τον καταλοίπων έχουμε

$$\hat{u}_i = Y_i - \hat{Y}_i \Rightarrow Y_i = \hat{Y}_i + \hat{u}_i \Rightarrow Y_i = \hat{\beta}_0 + \hat{\beta}_1 X_i + \hat{u}_i$$

αντικαθιστώντας τη λύση του εκτιμητή του σταθερού όρου του υποδείγματος $\hat{\beta}_0 = \overline{\mathrm{Y}} - \hat{\beta}_1 \overline{\mathrm{X}}$ έχουμε

$$Y_i = \overline{\mathrm{Y}} - \hat{\beta}_1 \overline{\mathrm{X}} + \hat{\beta}_1 X_i + \hat{u}_i \Rightarrow$$
$$(Y_i - \overline{\mathrm{Y}}) = \hat{\beta}_1 (X_i - \overline{\mathrm{X}}) + \hat{u}_i$$

Από την τελευταία σχέση μπορεί να εκτιμηθεί ο συντελεστής κλίσης του γραμμικού υποδείγματος $\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})(X_i - \bar{X})}{\sum_{i=1}^{n}(X_i - \bar{X})^2}$ (ίδιος εκτιμητής με αυτόν που προέκυψε από την (5β))

Ιδιότητες καταλοίπων

i. Το **άθροισμα των καταλοίπων είναι μηδέν**, δηλαδή ισχύει

$$\sum_{i=1}^{n} \hat{u}_i = 0$$

- Από την κανονική εξίσωση (4α) έχουμε $\sum_{i=1}^{n}(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i) = 0$ και εξ ορισμού ισχύει ότι $\hat{u}_i = Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i$ επομένως $\sum_{i=1}^{n} \hat{u}_i = 0$

- Εφόσον $\sum_{i=1}^{n} \hat{u}_i = 0$ άρα και ο μέσος όρος των καταλοίπων στο δείγμα θα είναι μηδέν $\bar{\hat{u}} = \frac{1}{n}\sum_{i=1}^{n} \hat{u}_i = 0$

Ιδιότητες καταλοίπων

ii. Οι **τιμές της ανεξάρτητης μεταβλητής και των καταλοίπων είναι ορθογώνιες** (μη συσχετιζόμενες) μεταξύ τους, δηλ.

$$\sum_{i=1}^{n} X_i \hat{u}_i = 0$$

(απόδειξη στο φροντιστήριο)

Ιδιότητες εξαρτημένης μεταβλητής

i. Η εκτιμημένη γραμμική σχέση $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$ περνάει από **το σημείο των μέσων όρων** των τιμών της εξαρτημένης και ανεξάρτητης μεταβλητής στο δείγμα. Το σημείο αυτό ορίζεται ως εξής:

$$\left(\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \,,\; \bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i\right)$$

και ισχύει ότι

$$\bar{Y} = \bar{\hat{Y}}$$

Ο μέσος όρος των εκτιμημένων τιμών της εξαρτημένης μεταβλητής $\bar{\hat{Y}} = \frac{1}{n}\sum_{i=1}^{n} \hat{Y}_i$ θα πρέπει να ισούται με το μέσο όρο που βασίζεται στις παρατηρούμενές τιμές της μεταβλητής $\bar{Y}$

Ιδιότητες εξαρτημένης μεταβλητής

Απόδειξη: Αν γράψουμε το γραμμικό υπόδειγμα ως εξής:

$$Y_i = \widehat{Y}_i + \hat{u}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i + \hat{u}_i,$$

και πάρουμε τους μέσους όρους και των δύο μελών της παραπάνω εξίσωσης έχουμε

$$\frac{1}{n}\sum_{i=1}^{n} Y_i = \frac{1}{n}\sum_{i=1}^{n} \widehat{Y}_i + \frac{1}{n}\sum_{i=1}^{n} \hat{u}_i = \frac{1}{n} n\hat{\beta}_0 + \hat{\beta}_1 \frac{1}{n}\sum_{i=1}^{n} X_i + \frac{1}{n}\sum_{i=1}^{n} \hat{u}_i \Rightarrow$$

$$\bar{Y} = \bar{\widehat{Y}} = \hat{\beta}_0 + \hat{\beta}_1 \bar{X},$$

λόγω της ιδιότητας των καταλοίπων: (i) $\sum_{i=1}^{n} \hat{u}_i = 0$.

Ιδιότητες εξαρτημένης μεταβλητής

ii. Οι **τιμές $\widehat{\boldsymbol{Y}}_{\boldsymbol{i}}$ είναι ορθογώνιες με τα κατάλοιπα $\widehat{\boldsymbol{u}}_{\boldsymbol{i}}$**, δηλαδή ισχύει

$$\sum_{i=1}^{n} \hat{Y}_i \hat{u}_i = 0$$

Απόδειξη:

Πολλαπλασιάζουμε και τα δύο μέλη της σχέσης $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$ με $\hat{u}_i$ και αθροίζουμε για όλες τις παρατηρήσεις i. Τότε, θα έχουμε

$$\sum_{i=1}^{n} \hat{Y}_i \hat{u}_i = \sum_{i=1}^{n} (\hat{\beta}_0 + \hat{\beta}_1 X_i) \hat{u}_i = \hat{\beta}_0 \sum_{i=1}^{n} \hat{u}_i + \hat{\beta}_1 \sum_{i=1}^{n} X_i \hat{u}_i = 0,$$

λόγω των ιδιοτήτων των καταλοίπων: (i) $\sum_{i=1}^{n} \hat{u}_i = 0$ και (ii) $\sum_{i=1}^{n} X_i \hat{u}_i = 0$.

Άσκηση

Σας δίνεται ένα δείγμα 13 παρατηρήσεων για τις μεταβλητές Y_i (μέσο ωρομίσθιο) X_i (έτη εκπαίδευσης). Η γραμμική σχέση που τις συνδέει με βάση την οικονομική θεωρία είναι η εξής $Y_i = \beta_0 + \beta_1 X_i + u_i$

a. Υπολογίστε τους εκτιμητές ΕΤ $\hat{\beta}_0$ και $\hat{\beta}_1$

b. Επιβεβαιώστε υπολογιστικά ότι $\sum_{i=1}^{n} \hat{u}_i = 0$ και $\sum_{i=1}^{n} X_i \hat{u}_i = 0$

Υπόδειξη: χρησιμοποιήστε αποκλίσεις από τους μέσους

n	Y_i	X_i
1	4.4567	6
2	**5.77**	7
3	5.9787	8
4	7.3317	9
5	7.3182	10
6	6.5844	11
7	7.8182	12
8	7.8351	13
9	11.0223	14
10	10.6738	15
11	10.8361	16
12	13.615	17
13	13.531	18

Άσκηση

Λύση:

a. αν $y_i = Y_i - \bar{Y}$, $x_i = X_i - \bar{X}$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})(X_i - \bar{X})}{\sum_{i=1}^{n}(X_i - \bar{X})^2} = \frac{131.7856}{182}$$

$$= 0.724097$$

$$\hat{\beta}_0 = \bar{\mathrm{Y}} - \hat{\beta}_1 \bar{\mathrm{X}}$$

$$= 8.674708 - 0.724097 * 12$$

$$= -0.01445$$

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{1}{13} 156 = 12$$

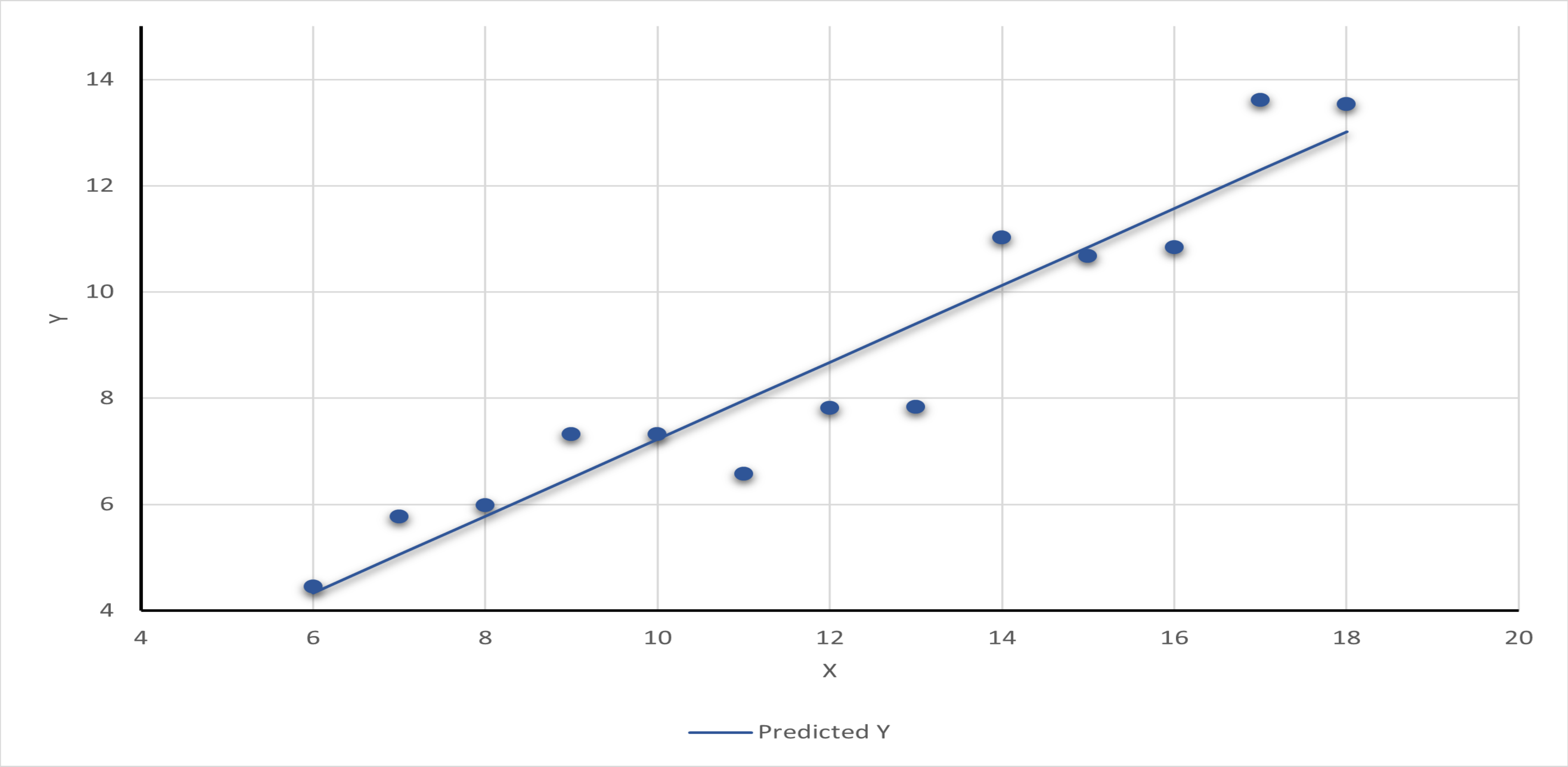
$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i = \frac{1}{13} 112.7712$$

$$= 8.674708$$

n	Y_i	X_i	y_i	x_i	$y_i\, x_i$	x_i^2	$\hat{Y}_i$
1	4.4567	6	-4.2180	-6	25.3080	36	4.3301
2	**5.7700**	7	-2.9047	-5	14.5235	25	5.0542
3	5.9787	8	-2.6960	-4	10.7840	16	5.7783
4	7.3317	9	-1.3430	-3	4.0290	9	6.5024
5	7.3182	10	-1.3565	-2	2.7130	4	7.2265
6	6.5844	11	-2.0903	-1	2.0903	1	7.9506
7	7.8182	12	-0.8565	0	0.0000	0	8.6747
8	7.8351	13	-0.8396	1	-0.8396	1	9.3988
9	11.0223	14	2.3476	2	4.6952	4	10.1229
10	10.6738	15	1.9991	3	5.9973	9	10.8470
11	10.8361	16	2.1614	4	8.6456	16	11.5711
12	13.615	17	4.9403	5	24.7015	25	12.2952
13	13.531	18	4.8563	6	29.1378	36	13.0193
sum	112.7712	156	0.00	0.00	131.7856	182.00	112.7712

Άσκηση

$$\hat{Y}_i = -0.01445 + 0.724097X_i$$



Άσκηση

b. $\sum_{i=1}^{n} \hat{u}_i = 0$

$\hat{u}_i = Y_i - \hat{Y}_i$

$\sum_{i=1}^{n} X_i \hat{u}_i = 0$

n	Y_i	X_i	y_i	x_i	$y_i\ x_i$	x_i^2	$\hat{Y}_i$	$\hat{u}_i$	$X_i\ \hat{u}_i$
1	4.4567	6	-4.2180	-6	25.3080	36	4.3301	0.1266	0.7594
2	**5.7700**	7	-2.9047	-5	14.5235	25	5.0542	0.7158	5.0104
3	5.9787	8	-2.6960	-4	10.7840	16	5.7783	0.2004	1.6030
4	7.3317	9	-1.3430	-3	4.0290	9	6.5024	0.8293	7.4635
5	7.3182	10	-1.3565	-2	2.7130	4	7.2265	0.0917	0.9169
6	6.5844	11	-2.0903	-1	2.0903	1	7.9506	-1.3662	-15.0283
7	7.8182	12	-0.8565	0	0.0000	0	8.6747	-0.8565	-10.2781
8	7.8351	13	-0.8396	1	-0.8396	1	9.3988	-1.5637	-20.3282
9	11.0223	14	2.3476	2	4.6952	4	10.1229	0.8994	12.5916
10	10.6738	15	1.9991	3	5.9973	9	10.8470	-0.1732	-2.5980
11	10.8361	16	2.1614	4	8.6456	16	11.5711	-0.7350	-11.7599
12	13.615	17	4.9403	5	24.7015	25	12.2952	1.3198	22.4367
13	13.531	18	4.8563	6	29.1378	36	13.0193	0.5117	9.2108
sum	112.7712	156	0.00	0.00	131.7856	182.00	112.7712	0.00	0.00

Περίληψη

✓Εκτιμητής Ελαχίστων Τετραγώνων

- Σταθερά & Συντελεστής κλίσης
- Αποκλίσεις από τον μέσο

✓Ιδιότητες καταλοίπων και εξαρτημένης μεταβλητής

- Το άθροισμα των καταλοίπων είναι μηδέν
- Οι τιμές της ανεξάρτητης μεταβλητής και των καταλοίπων είναι ορθογώνιες
- Η εκτιμημένη γραμμική σχέση $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$ περνάει από το σημείο των μέσων όρων των τιμών της εξαρτημένης και ανεξάρτητης μεταβλητής στο δείγμα
- Οι τιμές $\hat{Y}_i$ είναι ορθογώνιες με τα κατάλοιπα $\hat{u}_i$

Βιβλιογραφία

- ✓ Stock & Watson, κεφ. 4 (4.2 & Παράρτημα 4.2)
- ✓ Gujarati & Porter, κεφ. 3
- ✓ Τζαβαλής, κεφ. 2

Ευχαριστώ!

Ερωτήσεις?