ΟΙΚΟΝΟΜΕΤΡΙΑ

Ασκήσεις: 05

Βιολέττα Δάλλα

Τμήμα Οικονομικών Επιστημών

Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών

Θέμα 1, Ιούνιος 2019

α) OLS εκτιμητής του β:

$$\widehat{\beta} = (X^{'}X)^{-1}X^{'}Y = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 10 \\ 6 \\ 0,5 \end{pmatrix} = \begin{pmatrix} 0,5 & 0 & 0 \\ 0 & 0,25 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 10 \\ 6 \\ 0,5 \end{pmatrix} = \begin{pmatrix} 5 \\ 1,5 \\ 0,5 \end{pmatrix}$$

Εκτιμώμενη γραμμή/εξίσωση παλινδρόμησης:

$$\widehat{Y}_t = \widehat{\beta}_0 + \widehat{\beta}_1 S_t + \widehat{\beta}_2 S_t^2 = 5 + 1,5S_t + 0,5S_t^2$$

Συντελεστής προσδιορισμού:

$$R^2 = \frac{SSR}{SST} = \frac{3}{5} = 0,6$$

Ερμηνεία συντελεστή προσδιορισμού:

Το 60% της συνολικής μεταβλητότητας του ωρομισθίου ερμηνεύεται από τη γραμμική επίδραση της εκπαίδευσης και του τετραγώνου της εκπαίδευσης.

β) OLS εκτιμητής του $V(\widehat{\beta})$:

$$\widehat{V}(\widehat{\beta}) = s^2(X^{'}X)^{-1} = s^2 \begin{pmatrix} 0,5 & 0 & 0 \\ 0 & 0,25 & 0 \\ 0 & 0 & 1 \end{pmatrix} = 0,1 \cdot \begin{pmatrix} 0,5 & 0 & 0 \\ 0 & 0,25 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0,05 & 0 & 0 \\ 0 & 0,025 & 0 \\ 0 & 0 & 0,1 \end{pmatrix}$$

όπου

$$s^2 = \frac{1}{T-K-1} SSE = \frac{1}{23-2-1} 2 = 0,1$$

$$SST = SSR + SSE \Rightarrow SSE = SST - SSR = 5 - 3 = 2$$

γ) 95% διάστημα πρόβλεψης για το ωρομίσθιο (Y_f) ατόμου με 5 έτη εκπαίδευσης ($S_f = 5$):

$$\Delta\Pi_{Y_f}(95\%) = \left[\widehat{Y}_f \pm t_{T-K-1,\frac{0,05}{2}} s_{\widehat{Y}_f}\right] = [25 \pm 2,086 \cdot 7,955] = [8,406 \ , 41,594]$$

όπου

$$\widehat{Y}_f = \widehat{\beta}_0 + \widehat{\beta}_1 S_f + \widehat{\beta}_2 S_f^2 = 5 + 1,5 \cdot 5 + 0,5 \cdot 5^2 = 25$$

$$t_{T-K-1,\frac{0,05}{2}} = t_{23-2-1,\,0,025} = t_{20,\,0,025} = 2,086$$

$$s_{\widehat{Y}_f} = \sqrt{s^2 + X_f'\widehat{V}(\widehat{\beta})X_f} = \sqrt{0,1 + \begin{pmatrix}1 & 5 & 25\end{pmatrix}\begin{pmatrix}0,05 & 0 & 0\\ 0 & 0,025 & 0\\ 0 & 0 & 0,1\end{pmatrix}\begin{pmatrix}1\\5\\25\end{pmatrix}}$$

$$= \sqrt{63,275} = 7,955$$

$$X_f = \begin{pmatrix}1\\ S_f\\ S_f^2\end{pmatrix} = \begin{pmatrix}1\\5\\5^2\end{pmatrix} = \begin{pmatrix}1\\5\\25\end{pmatrix}$$

δ) Μέσο ωρομίσθιο: $E(Y_t) = \beta_0 + \beta_1 S_t + \beta_2 S_t^2$

Μέσο ωρομίσθιο είναι 5€ όταν η εκπαίδευση είναι 2 έτη $\Leftrightarrow E(Y_t) = 5$ όταν $S_t = 2 \Leftrightarrow$

$\beta_0 + \beta_1 S_t + \beta_2 S_t^2 = 5$ όταν $S_t = 2 \Leftrightarrow$

$\beta_0 + \beta_1 \cdot 2 + \beta_2 \cdot 2^2 = 5 \Leftrightarrow$

$\beta_0 + 2\beta_1 + 4\beta_2 = 5$

Στατιστικός έλεγχος για γραμμικό περιορισμό

Υποθέσεις: $H_0 : \beta_0 + 2\beta_1 + 4\beta_2 = 5$ έναντι $H_1 : \beta_0 + 2\beta_1 + 4\beta_2 \neq 5$

$\Leftrightarrow H_0 : R\beta = c$ έναντι $H_1 : R\beta \neq c$

όπου

$$R = \begin{pmatrix} 1 & 2 & 4 \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix}, \quad c = 5$$

Στατιστική ελέγχου: $F = (R\widehat{\beta} - c)'(R\widehat{V}(\widehat{\beta})R')^{-1}(R\widehat{\beta} - c)/q$

$$= 5' \cdot (1,75)^{-1} \cdot 5/1 = 5 \cdot (1,75)^{-1} \cdot 5/1 = 14,286$$

όπου

$$R\widehat{\beta} - c = \begin{pmatrix} 1 & 2 & 4 \end{pmatrix} \begin{pmatrix} 5 \\ 1,5 \\ 0,5 \end{pmatrix} - 5 = 5$$

$$R\widehat{V}(\widehat{\beta})R' = \begin{pmatrix} 1 & 2 & 4 \end{pmatrix} \begin{pmatrix} 0,05 & 0 & 0 \\ 0 & 0,025 & 0 \\ 0 & 0 & 0,1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix} = 1,75$$

Κρίσιμη περιοχή: $F > F_{q,\,T-K-1,\,\alpha} = F_{1,\,23-2-1,\,0{,}05} = F_{1,\,20,\,0{,}05} = 4,351$

Απόφαση: Απορρίπτουμε H_0 σε επίπεδο σημαντικότητας $\alpha = 0,05$.

Σχόλιο: Μέσο ωρομίσθιο δεν είναι 5€ όταν η εκπαίδευση είναι 2 έτη.

ε) Ορίζουμε τη ψευδομεταβλητή D:

$$D_t = \begin{cases} 1, & \text{αν άτομο } t = \text{άνδρας} \\ 0, & \text{αλλού} \end{cases}$$

Θέλουμε να ελέγξουμε αν όλοι οι συντελεστές παλινδρόμησης δεν διαφοροποιούνται για άνδρες και γυναίκες. Άρα, εισάγουμε στο υπόδειγμα παλινδρόμησης (1) την ψευδομεταβλητή D και τις πολλαπλασιαστικές ψευδομεταβλητές $S \cdot D$ και $S^2 \cdot D$. Έχουμε το υπόδειγμα παλινδρόμησης:

$$(2)\ Y_t = \alpha_0 + \delta_0 D_t + \alpha_1 S_t + \delta_1 (S_t \cdot D_t) + \alpha_2 S_t^2 + \delta_2 (S_t^2 \cdot D_t) + \varepsilon_t$$

Για άνδρες ($D_t = 1$):

$$(2) \Rightarrow Y_t = \alpha_0 + \delta_0 \cdot 1 + \alpha_1 S_t + \delta_1 (S_t \cdot 1) + \alpha_2 S_t^2 + \delta_2 (S_t^2 \cdot 1) + \varepsilon_t \Rightarrow$$

$$Y_t = (\alpha_0 + \delta_0) + (\alpha_1 + \delta_1)S_t + (\alpha_2 + \delta_2)S_t^2 + \varepsilon_t$$

Για γυναίκες ($D_t = 0$):

$$(2) \Rightarrow Y_t = \alpha_0 + \delta_0 \cdot 0 + \alpha_1 S_t + \delta_1(S_t \cdot 0) + \alpha_2 S_t^2 + \delta_2(S_t^2 \cdot 0) + \varepsilon_t \Rightarrow$$
$$Y_t = \alpha_0 + \alpha_1 S_t + \alpha_2 S_t^2 + \varepsilon_t$$

Οι συντελεστές παλινδρόμησης δεν διαφοροποιούνται για άνδρες και γυναίκες $\Leftrightarrow$
$\alpha_0 + \delta_0 = \alpha_0$ και $\alpha_1 + \delta_1 = \alpha_1$ και $\alpha_2 + \delta_2 = \alpha_2 \Leftrightarrow$
$\delta_0 = 0$ και $\delta_1 = 0$ και $\delta_2 = 0$

Έχουμε τα υποδείγματα:

$$(2)\ Y_t = \alpha_0 + \delta_0 D_t + \alpha_1 S_t + \delta_1(S_t \cdot D_t) + \alpha_2 S_t^2 + \delta_2(S_t^2 \cdot D_t) + \varepsilon_t \ \ (U)$$

$$(1)\ Y_t = \beta_0 + \beta_1 S_t + \beta_2 S_t^2 + u_t \ \ (R)$$

Το υπόδειγμα (1) είναι ειδική περίπτωση του (2) με γραμμικούς περιορισμούς $\delta_0 = 0$ και $\delta_1 = 0$ και $\delta_2 = 0$.
Βάσει των (1) και (2) ελέγχουμε την υπόθεση ότι οι γραμμικοί περιορισμοί $\delta_0 = 0$ και $\delta_1 = 0$ και $\delta_2 = 0$ ισχύουν.

Στατιστικός έλεγχος για γραμμικούς περιορισμούς

Υποθέσεις: $H_0 : \delta_0 = \delta_1 = \delta_2 = 0$ έναντι $H_1 : \delta_0 \neq 0$ ή/και $\delta_1 \neq 0$ ή/και $\delta_2 \neq 0$

Στατιστική ελέγχου: $F = \frac{(SSE_R - SSE_U)/q}{SSE_U/(T-K-1)}$

όπου

$SSE_R = 2$, $q = 3$, $T = 23$ και $K = 5$

Κρίσιμη περιοχή: $F > F_{q,T-K-1,\alpha} = F_{3,23-5-1,\alpha} = F_{3,17,0{,}05} = 3{,}197$

- Αν απορρίπταμε H_0 σε επίπεδο σημαντικότητας $\alpha = 0{,}05$.

  Σχόλιο: Οι συντελεστές παλινδρόμησης διαφοροποιούνται για άνδρες και γυναίκες.

- Αν δεν απορρίπταμε H_0 σε επίπεδο σημαντικότητας $\alpha = 0{,}05$.

  Σχόλιο: Οι συντελεστές παλινδρόμησης δεν διαφοροποιούνται για άνδρες και γυναίκες.

Θέμα 1, Φεβρουάριος 2010

α) Πρόβλεψη για εισόδημα (Y_f^A) άνδρα ($D_f = 0$) με 6 χρόνια μόρφωσης ($M_f = 6$):

$$\widehat{Y}_f^A = \widehat{\beta}_0 + \widehat{\beta}_1 M_f + \widehat{\beta}_2 D_f = 8,4 + 0,3 \cdot 6 - 0,96 \cdot 0 = 10,2$$

Πρόβλεψη για εισόδημα (Y_f^Γ) γυναίκας ($D_f = 1$) με 9 χρόνια μόρφωσης ($M_f = 9$):

$$\widehat{Y}_f^\Gamma = \widehat{\beta}_0 + \widehat{\beta}_1 M_f + \widehat{\beta}_2 D_f = 8,4 + 0,3 \cdot 9 - 0,96 \cdot 1 = 10,14$$

β) Υπόδειγμα παλινδρόμησης:

$$(1)\ Y_i = \beta_0 + \beta_1 M_i + \beta_2 D_i + u_i$$

Στατιστικός έλεγχος για σημαντικότητα υποδείγματος

Υποθέσεις: $H_0 : \beta_1 = \beta_2 = 0$ έναντι $H_1 : \beta_1 \neq 0$ ή/και $\beta_2 \neq 0$

Στατιστική ελέγχου: $F = \frac{R^2/K}{(1-R^2)/(T-K-1)} = \frac{0{,}4/2}{(1-0{,}4)/(24-2-1)} = 7$

Κρίσιμη περιοχή: $F > F_{K,\,T-K-1,\,\alpha} = F_{2,\,21,\,0{,}05} = 3,47$

Απόφαση: Απορρίπτουμε H_0 σε επίπεδο σημαντικότητας $\alpha = 0,05$.

Σχόλιο: Το υπόδειγμα είναι σημαντικό.

γ) Μέσο εισόδημα: $E(Y_i) = \beta_0 + \beta_1 M_i + \beta_2 D_i$

Μέσο εισόδημα άνδρα ($D_i = 0$) με δεδομένη μόρφωση ($M_i = M^*$):
$E(Y_i^A) = \beta_0 + \beta_1 M_i + \beta_2 D_i = \beta_0 + \beta_1 M^* + \beta_2 \cdot 0 = \beta_0 + \beta_1 M^*$

Μέσο εισόδημα γυναίκας ($D_i = 1$) με δεδομένη μόρφωση ($M_i = M^*$):
$E(Y_i^\Gamma) = \beta_0 + \beta_1 M_i + \beta_2 D_i = \beta_0 + \beta_1 M^* + \beta_2 \cdot 1 = (\beta_0 + \beta_2) + \beta_1 M^*$

Για δεδομένο επίπεδο μόρφωσης, το μέσο εισόδημα άνδρα υπερβαίνει το μέσο εισόδημα γυναίκας $\Leftrightarrow$
$E(Y_i^A) > E(Y_i^\Gamma) \Leftrightarrow$
$\beta_0 + \beta_1 M^* > (\beta_0 + \beta_2) + \beta_1 M^* \Leftrightarrow$
$\beta_2 < 0$

Στατιστικός έλεγχος για συντελεστή παλινδρόμησης

Υποθέσεις: $H_0 : \beta_2 = 0$ έναντι $H_1 : \beta_2 < 0$

Στατιστική ελέγχου: $t = \frac{\widehat{\beta}_2 - \beta_2^*}{s_{\widehat{\beta}_2}} = \frac{-0{,}96-0}{0{,}04} = -24$

Κρίσιμη περιοχή: $t < -t_{T-K-1,\,\alpha} = -t_{24-2-1,\,0{,}05} = -t_{21,\,0{,}05} = -1,721$

Απόφαση: Απορρίπτουμε H_0 σε επίπεδο σημαντικότητας $\alpha = 0,05$.

Σχόλιο: Για δεδομένο επίπεδο μόρφωσης, το μέσο εισόδημα άνδρα υπερβαίνει το μέσο εισόδημα γυναίκας.

δ) Εισάγουμε στο υπόδειγμα παλινδρόμησης (1) την ερμηνευτική μεταβλητή Π. Έχουμε το υπόδειγμα παλινδρόμησης:

$$Y_i = \beta_0 + \beta_1 M_i + \beta_2 D_i + \beta_3 \Pi_i + \varepsilon_i$$

Ισχύει ότι: $\Pi_i = H_i - M_i - 5 \Leftrightarrow M_i + \Pi_i = H_i - 5$

- Αν η ηλικία ήταν σταθερή στο δείγμα, δηλαδή $H_i = H^*$.
  Ισχύει ότι: $M_i + \Pi_i = H^* - 5 \Rightarrow$ τέλεια γραμμική σχέση ανάμεσα στις ερμηνευτικές

μεταβλητές M και Π $\Rightarrow$ υπάρχει τέλεια πολυσυγγραμμικότητα $\Rightarrow$ ο πίνακας X δεν είναι πλήρους βαθμού $\Rightarrow$ ο πίνακας $X^{'}X$ δεν είναι αντιστρέψιμος $\Rightarrow$ οι OLS εκτιμητές δεν υπολογίζονται.

- Αν η ηλικία δεν ήταν σταθερή στο δείγμα, δηλαδή $H_i \neq H^*$.
  Ισχύει ότι: $M_i + \Pi_i = H_i - 5$ $\Rightarrow$ σχεδόν τέλεια γραμμική σχέση ανάμεσα στις ερμηνευτικές μεταβλητές M και Π $\Rightarrow$ υπάρχει πολυσυγγραμμικότητα $\Rightarrow$ ο πίνακας X είναι πλήρους βαθμού $\Rightarrow$ ο πίνακας $X^{'}X$ είναι αντιστρέψιμος $\Rightarrow$ οι OLS εκτιμητές υπολογίζονται και οι στατιστικές ιδιότητες τους δεν επηρεάζονται από την ύπαρξη της πολυσυγγραμμικότητας (εφόσον οι υπόλοιπες βασικές υποθέσεις ισχύουν).
  Λόγω της πολυσυγγραμμικότητας, οι διακυμάνσεις και άρα και τα τυπικά σφάλματα των OLS εκτιμητών των συντελεστών παλινδρόμησης μπορεί να είναι μεγάλα, ειδικά σε μικρά δείγματα (εδώ $T = 24$). Στην περίπτωση αυτή, οι t στατιστικές για τη σημαντικότητα μίας ερμηνευτικής μεταβλητής θα είναι μικρές (ειδικά για τις ερμηνευτικές μεταβλητές M και Π) και θα συμπεραίναμε πιθανώς λανθασμένα ότι κάθε μία από αυτές δεν είναι σημαντική. Αν λανθασμένα αφαιρούσαμε αυτές τις ερμηνευτικές μεταβλητές θα κάναμε σφάλμα στην εξειδίκευση του υποδείγματος παλινδρόμησης.

Συνιστάται να ελεγχθεί η από κοινού σημαντικότητα των ερμηνευτικών μεταβλητών αυτών με F στατιστική για γραμμικούς περιορισμούς προτού αυτές αφαιρεθούν για αν αποφευχθεί πιθανό σφάλμα εξειδίκευσης του υποδείγματος παλινδρόμησης.