

ΚΕΦΑΛΑΙΟ 3

3. ΕΙΣΑΓΩΓΗ ΣΤΗΝ ΟΙΚΟΝΟΜΕΤΡΙΑ

Ας υποθέσουμε ότι έχουμε δύο μεταβλητές, την Y_t και την X_t και ότι για κάθε μια απ' αυτές διαθέτουμε η τιμές. Υποθέτουμε επίσης ότι η Y_t είναι τυχαία μεταβλητή και οι τιμές της εξαρτώνται από τη μεταβλητή X_t . Για το λόγο αυτό η Y_t καλείται εξαρτημένη και η X_t ανεξάρτητη ή ερμηνευτική μεταβλητή. Η σχέση που συνδέει τις δύο μεταβλητές είναι της μορφής

$$Y_t = f(X_t)$$

Η απλούστερη μορφή της παραπάνω συνάρτησης δίνεται από τη σχέση

$$Y_t = \alpha + \beta X_t$$

όπου α σταθερά και β ο συντελεστής της απλής παλινδρόμησης που εκφράζει το ποσό της μεταβολής της εξαρτημένης μεταβλητής Y_t μετά τη μεταβολή κατά μια μονάδα της ανεξάρτητης μεταβλητής X_t .

Λόγω της χρησιμοποίησης παρατηρήσεων από ένα δείγμα του πληθυσμού η σχέση μεταξύ των Y_t και X_t δεν είναι ακριβής και για το λόγο αυτό χρησιμοποιούμε και ένα όρο σφάλματος ε_t .

Έτσι η προηγούμενη σχέση γίνεται

$$Y_t = \alpha + \beta X_t + \varepsilon_t$$

Οι βασικές υποθέσεις που πρέπει να πληρούνται στο παραπάνω υπόδειγμα για την εφαρμογή της μεθόδου των ελάχιστων τετραγώνων, που θα αναπτύξουμε στη συνέχεια, στην προσπάθειά μας να εκτιμήσουμε τους συντελεστές της συνάρτησης είναι:

- α) ο όρος σφάλματος να είναι τυχαία μεταβλητή, δηλαδή να μην περιλαμβάνει συστηματικούς παράγοντες.
 β) ο αριθμητικός μέσος του όρου σφάλματος να είναι ίσος με μηδέν.
 γ) ο όρος σφάλματος θα πρέπει να κατανέμεται κανονικά γύρω από τον αριθμητικό μέσο με διακύμανση σταθερή.
 δ) η συνδιακύμανση να είναι ίση με μηδέν.
 Η συνάρτηση σε μορφή μητρών γράφεται $y = x\beta + \varepsilon$

$$\text{όπου } y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_T \end{bmatrix} \quad x = \begin{bmatrix} 1 & x_{11} & x_{12} \dots x_{1N} \\ 1 & x_{21} & x_{22} \dots x_{2N} \\ \vdots & \vdots & \vdots \\ 1 & x_{T1} & x_{T2} \dots x_{TN} \end{bmatrix} \quad \beta = \begin{bmatrix} \alpha \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_N \end{bmatrix} \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_T \end{bmatrix}$$

Υποθέσεις:

- 1) $y = x\beta + \varepsilon$ δείχνει τη γραμμικότητα του υποδείγματος
- 2) $E\varepsilon = 0$, όπου ε στοχαστική μεταβλητή με μέσο μηδέν
- 3) $V(\varepsilon) = E\varepsilon\varepsilon' = \sigma^2 I$

$$E\varepsilon\varepsilon' = E \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_T \end{bmatrix} \begin{bmatrix} \varepsilon_1 & \varepsilon_2 & \dots & \varepsilon_T \end{bmatrix} = \begin{bmatrix} E\varepsilon_1^2 & E\varepsilon_1\varepsilon_2 \dots E\varepsilon_1\varepsilon_T \\ E\varepsilon_2\varepsilon_1 & E\varepsilon_2^2 \dots E\varepsilon_2\varepsilon_T \\ \vdots & \vdots & \vdots \\ E\varepsilon_T\varepsilon_1 & E\varepsilon_T\varepsilon_2 & E\varepsilon_T^2 \end{bmatrix} =$$

$$= \begin{bmatrix} \sigma^2 & 0 & \dots & 0 \\ 0 & \sigma^2 & \dots & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & \sigma^2 \end{bmatrix} = \sigma^2 \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} = \sigma^2 I$$

γιατί

$$E\epsilon_1^2 = E(\epsilon_1 - E\epsilon_1)^2 = V(\epsilon_1) = \sigma^2$$

$$E(\epsilon_1 \epsilon_2) = E(\epsilon_1 - E\epsilon_1)(\epsilon_2 - E\epsilon_2) = \text{cov}(\epsilon_1, \epsilon_2) = 0$$

$\text{cov}(\epsilon_t, \epsilon_s) = 0, \quad t \neq s$ άρα τα ϵ ασυσχέτιστα.

Όπως είναι φανερό από τα παραπάνω οι διακυμάνσεις των ϵ εξ' υποθέσεως είναι ίσες μεταξύ τους και ίσες με σ^2 , δηλ. υπάρχει ομοσκεδαστικότητα. Εξ' άλλου οι συνδιακυμάνσεις των ϵ είναι μηδέν εξ' υποθέσεως δηλ. τα ϵ είναι ασυσχέτιστα.

4) Η μήτρα των x είναι σταθερή σε επαναλαμβανόμενα δείγματα, δηλ. μη στοχαστική.

Σαν επακόλουθο της υπόθεσης αυτής τα ϵ και x είναι ανεξάρτητα και επομένως ισχύουν οι σχέσεις $E(\epsilon/x) = E\epsilon = 0$ και

$$V(\epsilon/x) = E(\epsilon\epsilon'/x) = E\epsilon\epsilon' = \sigma^2 \cdot I$$

5) Ο βαθμός της μήτρας x είναι $N + 1 \leq T$, όπου T το μέγεθος του δείγματος. Η υπόθεση γίνεται για να υπάρχει η αντίστροφη μήτρα, της $(x'x)$, που είναι απαραίτητη για την εκτίμηση του υποδείγματος.

3.1. ΕΚΤΙΜΗΣΗ ΤΟΥ ΔΙΜΕΤΑΒΛΗΤΟΥ ΚΛΑΣΣΙΚΟΥ ΓΡΑΜΜΙΚΟΥ ΥΠΟΔΕΙΓΜΑΤΟΣ

Εφαρμόζοντας τη μέθοδο των ελαχίστων τετραγώνων καταλήγουμε στο σύστημα των δύο κανονικών εξισώσεων, η επίλυση του οποίου μας δίνει τις εκτιμήσεις των παραμέτρων.

Αν το υπόδειγμα δεν είναι σε μορφή μητρών, τότε οι εκτιμήσεις δίνονται από τις παρακάτω σχέσεις: (εξίσωση $\hat{Y} = a + bx$)

$$b = \frac{\text{cov}(X,Y)}{V(X)} = \frac{\sum (X-\bar{X})(Y-\bar{Y}) / (T-1)}{\sum (X-X)^2 / (T-1)} = \frac{\sum (X-\bar{X})(Y-\bar{Y})}{\sum (X-X)^2} = \frac{T\sum XY - \sum X\sum Y}{T\sum X^2 - (\sum X)^2}$$

$$a = \bar{Y} - b\bar{X}$$

Αν η εξίσωση γράφεται σε αποκλίσεις από τους μέσους παίρνει την εξής μορφή:

$$\left. \begin{aligned} \hat{Y} &= a + bX \\ \bar{Y} &= a + b\bar{X} \end{aligned} \right\} \text{ αφαιρώ κατά μέλη}$$

$$\begin{array}{ccc} Y - \bar{Y} & = & b (X - \bar{X}) \\ \uparrow & & \uparrow \\ y & = & b x \end{array}$$

$$\text{και} \quad b = \frac{\sum xy}{\sum x^2}$$

Αν οι μεταβλητές έχουν εκφραστεί σε μήτρες τότε η εξίσωση είναι η $y = xb + \epsilon$. Η παλινδρόμηση είναι $\hat{y} = xb$ και η εκτίμηση των b δίνεται από το τύπο:

$$b = (x'x)^{-1} \cdot x'y$$

Ο τύπος αυτός ισχύει για όσες μεταβλητές έχουμε. Οι διακυμάνσεις των b δίνονται από τον τύπο $s_b^2 = s^2 (x'x)^{-1}$

ΑΣΚΗΣΗ

Δίνεται το διμεταβλητό υπόδειγμα $Y = \alpha + \beta X + \epsilon$. Να εκτιμηθούν τα α και β με βάση τα παρακάτω δεδομένα.

X	Y
2	8
4	6
5	4
7	4

ΛΥΣΗ

$$b = \frac{T\sum XY - \sum X \cdot \sum Y}{T\sum X^2 - (\sum X)^2} = \frac{4 \cdot 88 - 18 \cdot 22}{4 \cdot 94 - 18^2} = \frac{-44}{52} = -0,8461538$$

X	Y	XY	X ²
2	8	16	4
4	6	24	16
5	4	20	25
7	4	28	49
18	22	88	94

$$a = \bar{Y} - b\bar{X} = \frac{22}{4} - (-0,8461538) \cdot \frac{18}{4} = 9,307623$$

Αν κάνουμε τους υπολογισμούς με μήτρες θα έχουμε:

$$y = xb \quad y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_T \end{bmatrix} \quad x = \begin{bmatrix} 1 & x_{11} \\ 1 & x_{12} \\ \vdots & \vdots \\ 1 & x_{1T} \end{bmatrix} \quad \text{και} \quad b = \begin{bmatrix} a \\ b \end{bmatrix}$$

$$b = (x'x)^{-1} \cdot x'y$$

$$y = \begin{bmatrix} 8 \\ 6 \\ 4 \\ 4 \end{bmatrix} \quad x = \begin{bmatrix} 1 & 2 \\ 1 & 4 \\ 1 & 5 \\ 1 & 7 \end{bmatrix} \quad x' = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 4 & 5 & 7 \end{bmatrix}$$

$$x'x = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 4 & 5 & 7 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ 1 & 4 \\ 1 & 5 \\ 1 & 7 \end{bmatrix} = \begin{bmatrix} 4 & 18 \\ 18 & 94 \end{bmatrix}$$

$$\text{adj } (x'x) = \begin{bmatrix} +94 & -18 \\ -18 & +4 \end{bmatrix} \quad |x'x| = 4 \cdot 94 - 18 \cdot 18 = 52$$

$$(x'x)^{-1} = \frac{\text{adj } (x'x)}{|x'x|} = \frac{1}{52} \begin{bmatrix} 94 & -18 \\ -18 & 4 \end{bmatrix} = \begin{bmatrix} 1,8076923 & -0,3461538 \\ -0,3461538 & 0,0769231 \end{bmatrix}$$

$$x'y = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 4 & 5 & 7 \end{bmatrix}_{2 \times 4} \cdot \begin{bmatrix} 8 \\ 6 \\ 4 \\ 4 \end{bmatrix}_{4 \times 1} = \begin{bmatrix} 22 \\ 88 \end{bmatrix}$$

$$b = (x'x)^{-1} \cdot x'y = \begin{bmatrix} 1,8076923 & -0,3461538 \\ -0,3461538 & 0,0769231 \end{bmatrix}_{2 \times 2} \cdot \begin{bmatrix} 22 \\ 88 \end{bmatrix}_{2 \times 1} = \begin{bmatrix} 9,3076962 \\ -0,8461508 \end{bmatrix}$$

οπότε $\alpha = 9,307692$ και $b = -0,8461508$

3.2. ΔΙΑΚΥΜΑΝΣΗ ΠΕΡΙ ΤΗΝ ΚΑΜΠΥΛΗ ΠΑΛΙΝΔΡΟΜΗΣΗΣ

Η διακύμανση περί την καμπύλη παλινδρόμησης δίνεται από τη σχέση:

$$s^2 = \frac{\sum e_t^2}{T - N - 1} = \frac{\sum (Y - \alpha - bX)^2}{T - 2} = \frac{\sum Y^2 - \alpha \sum Y - b \sum XY}{T - 2}$$

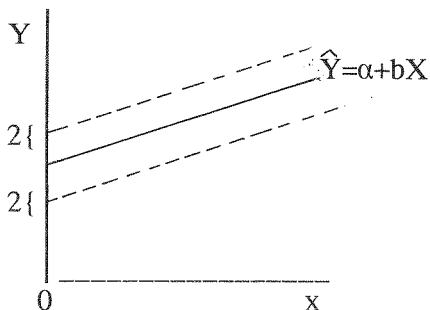
Σε μορφή μητρώων η διακύμανση δίνεται από τη σχέση:

$$s^2 = \frac{e'e}{T - N - 1} = \frac{y'y - b'x'y}{T - N - 1} \quad [\text{οι αριθμητές είναι μήτρες}]$$

$N = \text{αριθμός ανεξάρτητων μεταβλητών}$

Η τυπική απόκλιση περί την καμπύλη παλινδρόμησης μετράει τη διασπορά των τιμών της Y γύρω από την καμπύλη παλινδρόμησης.

$$\text{π.χ. } s^2 = 4 \\ S = 2$$



3.3. ΣΥΝΤΕΛΕΣΤΗΣ ή ΔΕΙΚΤΗΣ ΠΡΟΣΔΙΟΡΙΣΜΟΥ

$$0 \leq R^2 \leq 1$$

Δίνεται από τις σχέσεις:

$$R^2 = \frac{\text{cov}^2(x, y)}{s_x^2 \cdot s_y^2} = \frac{[\sum (X - \bar{X})(Y - \bar{Y})]^2}{\sum (X - \bar{X})^2 \cdot \sum (Y - \bar{Y})^2} = \frac{[\sum XY - (\sum X)(\sum Y)]^2}{[\sum X^2 - (\sum X)^2] \cdot [\sum Y^2 - (\sum Y)^2]}$$

Γενικός τύπος
$$R^2 = 1 - \frac{(T - N - 1) \cdot s^2}{(T - 1) \cdot s_y^2}$$

$$R^2 = 1 - \frac{\sum e_t^2}{s_y^2} \left\{ \begin{array}{l} \text{σε αποκλίσεις από τους μέσους} \end{array} \right.$$

$$R^2 = \frac{b \cdot \sum xy}{\sum y^2} \left\{ \begin{array}{l} \text{σε αποκλίσεις από τους μέσους} \end{array} \right. \quad R^2 = \frac{(\sum xy)^2}{\sum x^2 \cdot \sum y^2}$$

Ερμηνεία: Το R^2 μετράει το ποσοστό της μεταβλητικότητας της Y που οφείλεται στην μεταβλητικότητα της ανεξάρτητης μεταβλητής X . Είναι καθαρός αριθμός, ανεξάρτητος των μονάδων μέτρησης των μεταβλητών. Αποτελεί μέτρο προσαρμογής των δεδομέ-

νων στην εξίσωση δηλ. όσο μεγαλύτερη είναι η τιμή του, τόσο καλύτερα τα δεδομένα αναπαρίστανται από την εξίσωση. Παίρνει τιμές από μηδέν έως ένα.

Αν πάρει την τιμή ένα, η σχέση της εξίσωσης είναι μαθηματική, δηλ. οι μεταβολές της Y εξαρτώνται εξ' ολοκλήρου από τις μεταβολές της X κι αν $R^2 = 0$, οι μεταβλητές είναι ασυσχέτιστες, που σημαίνει ότι μεταβαλλομένης της μιας, η άλλη παραμένει αμετάβλητη.

Σε μορφή μητρώων το R^2 δίνεται από:

$$R^2 = \frac{b' x' y}{y' y}$$

3.4. ΣΥΝΤΕΛΕΣΤΗΣ ΣΥΣΧΕΤΙΣΗΣ

Ο συμβολισμός του είναι R (ή r) και ισχύει $-1 \leq R \leq 1$.

Μετράει την ένταση ή βαθμό εξάρτησης των μεταβλητών X και Y στο δείγμα. Αν $R = 0$ τα X και Y είναι ασυσχέιστα. Αν $R = \pm 1$ τα X και Y σχετίζονται με τέλεια συσχέτιση (δηλ. έχουμε μαθηματική σχέση)

$$r = b \cdot \frac{S_X}{S_Y} \quad r = \frac{\sum xy}{\sqrt{\sum x^2} \cdot \sqrt{\sum y^2}}$$

3.5. ΣΤΑΤΙΣΤΙΚΟΣ ΕΛΕΓΧΟΣ ΤΟΥ ΥΠΟΔΕΙΓΜΑΤΟΣ

Αφορά το στατιστικό έλεγχο του συντελεστή παλινδρόμησης b , του συντελεστή συσχέτισης r , και του συντελεστή προσδιορισμού R^2 . Και οι τρεις έλεγχοι είναι ισοδύναμοι στο διμεταβλητό υπόδειγμα και επομένως αρκεί ένας από τους τρεις για να γίνει ο στατιστικός έλεγχος του υποδείγματος. Οι έλεγχοι αυτοί είναι:

1ος έλεγχος στατιστικής σημαντικότητας του συντελεστή παλινδρόμησης b.

Υποθέσεις

$H_0 : \beta = 0$ β πληθυσμού

$H_1 : \beta \neq 0$ b δείγματος

Ο έλεγχος γίνεται από τον τύπο: $t = \frac{|b - \beta|}{s_b}$ και με βάση την

υπόθεση H_0 ο τύπος γίνεται: $t = \frac{|b|}{s_b}$.

Το s_b (τυπική απόκλιση) δίνεται από τον τύπο:

$$s_b = \frac{s}{\sqrt{\sum (X_i - \bar{X})^2}} = \frac{s}{s_x \sqrt{T-1}} \text{ γιατί } s_x^2 = \frac{\sum (X - \bar{X})^2}{T-1} \Rightarrow \sum (X - \bar{X})^2 = (T-1) s_x^2$$

Αν $t(\text{δείγματος}) \geq t_{v,\alpha}$ απορρίπτεται η υπόθεση H_0 και δεχόμαστε την H_1 . Συμπεραίνουμε δε ότι υπάρχει γραμμική παλινδρόμηση στον πληθυσμό, ή λέμε ότι η τιμή του b είναι στατιστικά σημαντική. Το ν είναι βαθμοί ελευθερίας $v = T - N - 1 = T - 2$ (για το διμεταβλητό είναι $N = 1$).

N είναι ο αριθμός των ανεξαρτήτων μεταβλητών.

Το α είναι το επίπεδο της στατιστικής σημαντικότητας.

Αν $t < t_{v,\alpha}$ δεκτή η H_0 , αντίθετα συμπεράσματα.

2ος έλεγχος στατιστικής σημαντικότητας του συντελεστή συσχέτισης r

Υποθέσεις

$H_0 : \rho = 0$ ρ συντελεστής συσχέτισης πληθυσμού

$H_1 : \rho \neq 0$ r δείγματος

Ο τύπος ελέγχου είναι: $t = \frac{r \sqrt{T - 2}}{\sqrt{1 - r^2}}$

Αν $t(\text{δείγματος}) \geq t_{\alpha}$ απορρίπτεται η υπόθεση H_0 και δεχόμαστε την H_1 . Σ' αυτή την περίπτωση υπάρχει γραμμική συσχέτιση στον πληθυσμό ή αλλιώς υπάρχει γραμμική παλινδρόμηση στον πληθυσμό. Αν $t < t_{\alpha}$ δεχόμαστε την H_0 , αντίθετα συμπεράσματα.

3ος έλεγχος του συντελεστή προσδιορισμού R^2

Υποθέσεις

$$\left. \begin{array}{l} H_0 : R^2 = 0 \\ H_1 : R^2 \neq 0 \end{array} \right\} \text{πληθυσμού}$$

Ο έλεγχος γίνεται με την κατανομή F του Snedecor. Ο τύπος

ελέγχου είναι:
$$F = \frac{\frac{R^2}{N}}{\frac{(1 - R^2)}{T - N - 1}} \quad N = \text{πλήθος ανεξαρτήτων μεταβλητών}$$

Ισχύει για όλες τις παλινδρομήσεις.

Αν $F \geq F_{\alpha, v_1, v_2}$, α απορρίπτω την H_0 , και δέχομαι την H_1 . Σ' αυτή την περίπτωση υπάρχει γραμμική παλινδρόμηση στον πληθυσμό.

$$(V_1 = N \quad \text{και} \quad V_2 = T - N - 1)$$

3.6. ΟΙΚΟΝΟΜΙΚΟΣ ΕΛΕΓΧΟΣ ΤΟΥ ΥΠΟΔΕΙΓΜΑΤΟΣ

Αναφέρεται κυρίως στον έλεγχο των προσήμων των συντελεστών παλινδρόμησης δηλ. εξετάζουμε αν τα πρόσημα τους, συμφωνούν με τα πρόσημα των μεταβλητών που θα αναμέναμε να βγουν σύμφωνα με την οικονομική θεωρία.

ΑΣΚΗΣΗ

Δίνεται η ζήτηση και η τιμή ενός αγαθού για μια σειρά 5 ετών.

Τιμή (σε εκατομ.)	Ζήτηση (σε χιλ. τον.)
2	20
4	15
5	14
7	10
10	7

Να υπολογιστούν: 1) η εξίσωση ζήτησης γραμμικής μορφής, 2) η διακύμανση περί την καμπύλη παλινδρόμησης και η τυπική απόκλιση. Να αξιολογηθεί η τιμή της διακύμανσης 3) ο δείκτης προσδιορισμού και ο συντελεστής συσχέτισης (να ερμηνευτούν) 4) να γίνει οικονομικός έλεγχος του υποδείγματος 5) να γίνει στατιστικός έλεγχος του υποδείγματος 6) να υπολογιστεί η ελαστικότητα της συνάρτησης για την τιμή του $x = 11$ μονάδες και η μέση ελαστικότητα (να ερμηνευτούν).

ΛΥΣΗ

X	Y	XY	x^2	y^2
2	20	40	4	400
4	15	60	16	225
5	14	70	25	196
7	10	70	49	100
10	7	70	100	49
28	66	310	194	970

$$1) \quad \hat{Y} = \alpha + bX$$

$$b = \frac{T\Sigma XY - \Sigma X \Sigma Y}{T\Sigma X^2 - (\Sigma X)^2} = \frac{5 \cdot 310 - 28 \cdot 66}{5 \cdot 194 - 28^2} = -1,602$$

$$\alpha = \bar{Y} - b\bar{X} = \frac{66}{5} - (-1,602) \cdot \frac{28}{5} = 22,1712$$

$$\text{Αρα } \hat{Y} = 22,1712 - 1,602X$$

$$2) \quad S^2 = \frac{\Sigma Y^2 - \alpha \Sigma Y - b \Sigma XY}{T - N - 1} = \frac{970 - 22,1712 \cdot 66 + 1,602 \cdot 310}{5 - 1 - 1} = 1,10693$$

$$\text{και } S = 1,0521$$

Η διακύμανση και η τυπική απόκλιση περί την καμπύλη παλινδρόμησης δεν μπορεί να αξιολογηθεί εύκολα, γιατί δεν υπάρχει μέτρο σύγκρισης της, για να πούμε ότι είναι μικρή ή μεγάλη η διασπορά. Εξ' άλλου οι τιμές της μεταβάλλονται όταν αλλάξουν οι μονάδες μέτρησης των μεταβλητών και η αξιολόγησή της είναι ακόμη δυσκολότερη. Αν λάβουμε υπ' όψη μας τα εμπειρικά δεδομένα της άσκησης μπορούμε να πούμε με επιφύλαξη ότι η διασπορά τιμών της y είναι σχετικά μικρή.

$$3) \quad R^2 = \frac{[T\Sigma XY - (\Sigma X)(\Sigma Y)]^2}{[T\Sigma X^2 - (\Sigma X)^2][T\Sigma Y^2 - (\Sigma Y)^2]} = \frac{(5 \cdot 310 - 28 \cdot 66)^2}{(5 \cdot 194 - 28^2)(5 \cdot 970 - 66^2)} = 0,9664$$

$$\text{και } |r| = R = 0,9830 \quad (\text{η σχέση είναι αρνητική λόγω της } b)$$

Ερμηνεία: Το $R^2 = 96,64\%$ σημαίνει ότι οι μεταβολές της Y ερμηνεύονται κατά 96,64% από τις μεταβολές της ερμηνευτικής μεταβλητής X . Το $R = 0,9830$ δείχνει μεγάλη ένταση εξάρτησης μεταξύ των μεταβλητών X και Y , δηλαδή εντονότατη συσχέτιση.

(4) Το πρόσημο του συντελεστή της μεταβλητής X συμφωνεί με την οικονομική θεωρία γιατί όταν αυξάνει η τιμή, μειώνεται η ζήτηση και αντιστρόφως.

5) 1ος έλεγχος

Υποθέσεις

$$H_0 : \beta = 0$$

$$H_1 : \beta \neq 0 \quad t = \frac{|b|}{s_b}$$

$$s_b = \frac{s}{\sqrt{\sum (X_i - \bar{X})^2}} = \frac{s}{\sqrt{\sum X_i^2 - \frac{(\sum X_i)^2}{T}}} = \frac{1,0521}{\sqrt{194 - \frac{28^2}{5}}} = \frac{1,0521}{6,099} = 0,172$$

$$t = \frac{1,602}{0,172} = 9,313$$

Υποθετική τιμή πινάκων

$$t_{3,0,05} = 2,71$$

Επειδή $t = 9,313 > t_{3,0,05} = 2,71$ απορρίπτεται η H_0 και δεχόμαστε την H_1 πράγμα που σημαίνει ότι η τιμή του β είναι στατιστικά σημαντική (δηλ. $\neq 0$), ή αλλιώς υπάρχει γραμμική παλινδρόμηση στον πληθυσμό.

2ος έλεγχος

Υποθέσεις

$$H_0 : \rho = 0$$

$$H_1 : \rho \neq 0 \quad t = \frac{r \sqrt{T - 2}}{\sqrt{1 - r^2}} = \frac{0,9830 \sqrt{5 - 2}}{\sqrt{1 - 0,9830^2}} = 9,263$$

Επειδή $t = 9,263 > t_{3,0,05} = 2,71$ απορρίπτεται η H_0 και γίνεται δεκτή η H_1

3ος έλεγχος

Υποθέσεις

$$H_0 : R^2 = 0$$

$$H_1 : R^2 \neq 0 \quad F = \frac{\frac{R^2}{N}}{\frac{(1 - R^2)}{T - N - 1}} = \frac{\frac{0,9664}{1}}{\frac{1 - 0,9664}{5 - 1 - 1}} = \frac{0,9664 \cdot 3}{1 - 0,9664} = 87,85$$

Υποθετική τιμή πινάκων

$$F_{1,3,0,05} = 4,24$$

Επειδή $F = 87,85 > F_{1,3,0,05} = 4,24$ απορρίπτεται η H_0 και γίνεται δεκτή η H_1

6) $\hat{Y} = 22,17 - 1,6X$

$$X = 11$$

$$\hat{Y} = 22,17 - 1,6 \cdot 11 = 4,57$$

για $X = 11$ το Y είναι $\hat{Y} = 4,57$

Η ελαστικότητα ζήτησης θα είναι:

$$\varepsilon_{XY} = b \cdot \frac{X}{\hat{Y}} = -1,6 \cdot \frac{11}{4,57} = -3,85$$

Η ζήτηση του αγαθού είναι ελαστική για $X = 11$ γιατί $|-3,85| > 1$, είναι ελαστική η συνάρτηση ζήτησης. Δηλ. αν αυξηθεί η τιμή του αγαθού κατά 10% τότε η ζήτηση θα μειωθεί κατά $0,1 \cdot 3,85 = 0,385$ ή 38,5%. Αρα δε συμφέρει να αυξήσουμε την τιμή.

Η μέση ελαστικότητα θα είναι:

$$\varepsilon_{\bar{X}\bar{Y}} = b \cdot \frac{\bar{X}}{\bar{Y}} = -1,6 \cdot \frac{5,6}{13,2} = -0,678$$

Η μέση ελαστικότητα είναι ανελαστική (< 1)

3.7. ΤΡΙΜΕΤΑΒΛΗΤΟ ΥΠΟΔΕΙΓΜΑ (ΠΟΛΥΜΕΤΑΒΛΗΤΟ)

Το υπόδειγμα είναι της μορφής: $Y_t = \alpha + \beta_1 X_{1t} + \beta_2 X_{2t} + \varepsilon_t$

Η εξίσωση παλινδρόμησης στο δείγμα, είναι της μορφής:

$$\hat{Y}_t = a + b_1 X_{1t} + b_2 X_{2t}$$

Αν η εξίσωση παλινδρόμησης γραφεί σε αποκλίσεις από τους μέσους (λείπει ο σταθερός όρος) θα έχουμε: $\hat{y}_t = b_1 x_{1t} + b_2 x_{2t}$

όπου $y_t = Y_t - \bar{Y}$, $x_1 = X_1 - \bar{X}_1$, $x_2 = X_2 - \bar{X}_2$

Εφαρμόζοντας τη μέθοδο των ελαχίστων τετραγώνων στο υπό-δείγμα προκύπτουν οι παρακάτω εκτιμήσεις των παραμέτρων:

$$b_1 = \frac{\sum x_1 y \cdot \sum x_2^2 - \sum x_2 y \cdot \sum x_1 x_2}{\sum x_1^2 \cdot \sum x_2^2 - (\sum x_1 x_2)^2} \quad \text{σε αποκλίσεις από τους μέσους}$$

$$b_2 = \frac{\sum x_2 y \cdot \sum x_1^2 - \sum x_1 y \cdot \sum x_1 x_2}{\sum x_1^2 \cdot \sum x_2^2 - (\sum x_1 x_2)^2} \quad \text{σε αποκλίσεις από τους μέσους}$$

$$a = Y - b_1 x_1 - b_2 x_2 \quad \text{σε αρχικές τιμές}$$

Η διακύμανση περί την καμπύλη παλινδρόμησης είναι:

$$s^2 = \frac{\sum (Y - \hat{Y})^2}{T - N - 1} = \frac{\sum Y^2 - a \sum Y - b_1 \sum x_1 Y - b_2 \sum x_2 Y}{T - 3} \quad \text{σε αρχικές τιμές}$$

$$N = 2 \quad \swarrow$$

Διακυμάνσεις των συντελεστών παλινδρόμησης:

$$s_{b_1}^2 = s^2 \cdot \frac{\sum x_2^2}{\Delta} \quad \text{και} \quad s_{b_1}^2 = s^2 \cdot \frac{1}{\sum x_1^2 (1 - R^2)}$$

$$s_{b_2}^2 = s^2 \cdot \frac{\Sigma x_1^2}{\Delta} \quad (\Delta = \Sigma x_1^2 \cdot \Sigma x_2^2 - (\Sigma x_1 x_2)^2)$$

$$\text{cov} (b_1, b_2) = \frac{s^2 (-\Sigma x_1 x_2)}{\Delta} \quad (\text{συνδιακύμανση των } b_1, b_2)$$

Συντελεστής προσδιορισμού: $R^2 = 1 - \frac{(T - N - 1) s^2}{(T - 1) s_Y^2}$

$$R^2 = \frac{b_1 \Sigma x_1 y + b_2 \Sigma x_2 y}{\Sigma y^2} \quad \text{σε αποκλίσεις από τους μέσους}$$

ΑΣΚΗΣΗ

Δίνεται το τριμεταβλητό υπόδειγμα $Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \varepsilon_t$.

Να εκτιμηθεί το υπόδειγμα και να γίνει ο στατιστικός έλεγχος με βάση τα παρακάτω δεδομένα:

Y	X ₁	X ₂
5	1	4
6	2	6
10	4	10
11	5	12

ΛΥΣΗ

Y	X ₁	X ₂	y=Y- $\bar{Y}$	x ₁ =X- $\bar{X}_1$	x ₂ =X ₂ - $\bar{X}_2$	x ₁ y	x ₂ ²	x ₂ y	x ₁ x ₂	x ₁ ²
5	1	4	-3	-2	-4	6	16	12	8	4
6	2	6	-2	-1	-2	2	4	4	2	1
10	4	10	2	1	2	2	4	4	2	1
11	5	12	3	2	4	6	16	12	8	4
32	12	32				16	40	32	20	10

$$b_1 = \frac{\Sigma x_1 y \cdot \Sigma x_2^2 - \Sigma x_2 y \cdot \Sigma x_1 x_2}{\Sigma x_1^2 \cdot \Sigma x_2^2 - (\Sigma x_1 x_2)^2} = \frac{16 \cdot 40 - 32 \cdot 20}{10 \cdot 40 - 20^2} = \frac{0}{0} \quad \text{απροσδιόριστη μορφή.}$$

όμοια για b_2

$$s_{b_1}^2 = \frac{s^2 \cdot \Sigma x_2^2}{\Sigma x_1^2 \cdot \Sigma x_2^2 - (\Sigma x_1 x_2)^2}, \quad s^2 = \frac{\Sigma Y^2 - a \Sigma Y - b_1 \Sigma x_1 Y - b_2 \Sigma x_2 Y}{T - 3} \quad (\text{σε αριθμικές τιμές})$$

$$a = \bar{Y} - b_1 \bar{X}_1 - b_2 \bar{X}_2$$

1ος έλεγχος

α) Ο στατιστικός έλεγχος αφορά την υπόθεση $H_0 : \beta_1 = 0$

$$H_1 : \beta_1 \neq 0$$

$$t = \frac{|b_1|}{s_{b_1}} \quad \text{Αν} \quad t \geq t_{\alpha, \nu} \quad \text{απορρίπτουμε την } H_0$$

$$V = T - N - 1 = T - 2 - 1 = T - 3$$

Η τιμή του συντελεστή παλινδρόμησης b_1 είναι στατιστικά σημαντική (δηλ. $\neq 0$)

β) $H_0 : \beta_2 = 0$

$$H_1 : \beta_2 \neq 0 \quad t = \frac{|b_2|}{s_{b_2}} \quad \text{όπου} \quad s_{b_2}^2 = \frac{s^2 \cdot \Sigma x_1^2}{\Delta}$$

αν $t \geq t_{\alpha, \nu}$ απορρίπτεται η H_0 . Η τιμή του συντελεστή παλινδρόμησης b_2 είναι στατιστικά σημαντική (δηλ. $\neq 0$)

2ος έλεγχος

Ελέγχω την υπόθεση ότι όλα τα β είναι μηδέν.

$$H_0 : \beta_1 = \beta_2 = 0$$

H_1 : ένα από τα βήτα τουλάχιστον, είναι διάφορο του μηδενός

$$F = \frac{\frac{R^2}{N}}{\frac{1 - R^2}{T - N - 1}} \quad V_1 = N, \quad V_2 = T - N - 1$$

Αν $F \geq F_{V_1, V_2, \alpha}$ απορρίπτεται η H_0 .

3.8 ΤΟ ΚΛΑΣΣΙΚΟ ΓΡΑΜΜΙΚΟ ΥΠΟΔΕΙΓΜΑ ΣΤΗΝ ΓΕΝΙΚΗ ΤΟΥ ΜΟΡΦΗ

Το υπόδειγμα είναι της μορφής :

$$Y_t = \alpha + \beta_1 X_{1t} + \beta_2 X_{2t} + \beta_3 X_{3t} + \dots + \beta_N X_{Nt} + \varepsilon_t$$

Και στην περίπτωση αυτή ισχύουν οι υποθέσεις που δόθηκαν προηγουμένως στο διμεταβλητό υπόδειγμα.

ΑΣΚΗΣΗ

Δίνεται το υπόδειγμα $y = x\beta + u$ όπου $x = \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix}$ και $y = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$

ΛΥΣΗ

Ο b θα είναι μήτρα της μορφής $b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$.

Επειδή στη μήτρα των x λείπει η μοναδιαία στήλη δεν υπάρχει στην b , ο όρος α , δηλ. ο σταθερός όρος.

Επειδή η μήτρα x είναι 2×2 , δηλ. έχω δύο στήλες, έχω τριμεταβλητό υπόδειγμα με δύο ανεξάρτητες μεταβλητές x_1, x_2 και μια εξηρημένη την y . Θα είναι δηλ. της μορφής:

$$y_t = \beta_1 x_{1t} + \beta_2 x_{2t} + \varepsilon_t$$

$$b = (x'x)^{-1} \cdot x'y$$

$$x'x = \begin{bmatrix} 4 & 2 \\ 3 & 1 \end{bmatrix} \cdot \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 20 & 14 \\ 14 & 10 \end{bmatrix}$$

$$(x'x)^{-1} = \frac{\text{adj } x'x}{|x'x|} = \frac{\begin{bmatrix} 10 & -14 \\ -14 & 20 \end{bmatrix}}{4} = \frac{1}{4} \begin{bmatrix} 10 & -14 \\ -14 & 20 \end{bmatrix} = \begin{bmatrix} 10/4 & -14/4 \\ -14/4 & 20/4 \end{bmatrix} = \begin{bmatrix} 2,5 & -3,5 \\ -3,5 & 5 \end{bmatrix}$$

$$\text{adj } x'x = \begin{bmatrix} 10 & -14 \\ -14 & 20 \end{bmatrix}, \quad |x'x| = \begin{vmatrix} 20 & 14 \\ 14 & 10 \end{vmatrix} = 20 \cdot 10 - 14 \cdot 14 = 200 - 196 = 4$$

$$x'y = \begin{bmatrix} 4 & 2 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \end{bmatrix} = \begin{bmatrix} 26 \\ 17 \end{bmatrix}$$

$$b = (x'x)^{-1} \cdot x'y = \begin{bmatrix} 2,5 & -3,5 \\ -3,5 & 5 \end{bmatrix} \cdot \begin{bmatrix} 26 \\ 17 \end{bmatrix} = \begin{bmatrix} 5,5 \\ -6 \end{bmatrix}$$

$$\text{άρα } b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} 5,5 \\ -6 \end{bmatrix} \text{ και } b_1 = 5,5 \text{ και } b_2 = -6$$

3.9. ΕΛΕΓΧΟΣ ΤΟΥ ΚΛΑΣΣΙΚΟΥ ΓΡΑΜΜΙΚΟΥ ΥΠΟΔΕΙΓΜΑΤΟΣ

Αν στο κλασσικό γραμμικό υπόδειγμα προστεθεί η υπόθεση ότι τα ε_t ακολουθούν την κανονική κατανομή με μέσο μηδέν και διακύμανση σ^2 , δηλ. $\varepsilon_t \approx N(0, \sigma^2)$, τότε έχουμε το κανονικό κλασσικό γραμμικό υπόδειγμα. Η υπόθεση που κάναμε μπορεί να στηριχθεί στο κεντρικό οριακό θεώρημα και είναι απαραίτητη για τον στατιστικό έλεγχο του υποδείγματος ή για την εφαρμογή της μεθόδου μέγιστης πιθανοφάνειας.

Οι έλεγχοι του υποδείγματος είναι οι εξής:

1) Έλεγχος της υπόθεσης ότι κάθε β χωριστά είναι ίσο με το μηδέν

$$H_0 : \beta_v = 0$$

$$H_1 : \beta_v \neq 0 \quad v = 1, 2, 3, \dots$$

Ο τύπος ελέγχου είναι: $t = \frac{|b_v|}{s_{b_v}}$

Αν $t \geq t_{\alpha, \nu}$ απορρίπτεται η H_0 . Βαθμοί ελευθερίας $\nu = T - N - 1$.

2) Έλεγχος της υπόθεσης ότι όλα τα β συγχρόνως είναι ίσα με το μηδέν.

$H_0 : \beta_1 = \beta_2 = \beta_3 = \dots = \beta_v = 0$

H_1 : Υπάρχει ένα τουλάχιστον β διάφορο του μηδενός.

Τύπος ελέγχου:
$$F = \frac{\frac{R^2}{N}}{\frac{1 - R^2}{T - N - 1}}$$

Αν $F \geq F_{\alpha, v_1, v_2}$, α απορρίπτεται η H_0 .

3) Έλεγχος της υπόθεσης ότι ένα υποσύνολο των β είναι ίσο με το μηδέν.

$H_0 : \beta_\theta = \beta_{\theta+1} = \beta_{\theta+2} = \dots = \beta_N = 0$

H_1 : ένα τουλάχιστον από αυτά, διάφορο του μηδενός

Ο έλεγχος γίνεται με την κατανομή F .

3.10. ΕΞΕΙΔΙΚΕΥΣΗ ΤΟΥ ΥΠΟΔΕΙΓΜΑΤΟΣ

Η εξειδίκευση του υποδείγματος αναφέρεται στον καθορισμό των μεταβλητών που θα συμπεριληφθούν στο υπόδειγμα και στον προσδιορισμό της μαθηματικής σχέσης που τα συνδέει.

3.11. ΣΤΑΔΙΑ ΤΗΣ ΟΙΚΟΝΟΜΕΤΡΙΚΗΣ ΑΝΑΛΥΣΗΣ ΕΝΟΣ ΥΠΟΔΕΙΓΜΑΤΟΣ

- 1) Εξειδικεύουμε το υπόδειγμα
- 2) Εκτιμούμε τους συντελεστές του υποδείγματος (μέθοδος ελαχίστων τετραγώνων ή μέγιστης πιθανοφάνειας)
- 3) Κάνουμε οικονομικό έλεγχο του υποδείγματος, ελέγχοντας τα πρόσημα των συντελεστών που προκύπτουν από τα εμπειρικά

δεδομένα, αν συμφωνούν με τα πρόσχημα που θα περιμέναμε να έχουμε σύμφωνα με την οικονομική θεωρία.

- 4) Κάνουμε στατιστικό έλεγχο του υποδείγματος
- 5) Κάνουμε οικονομετρικό έλεγχο του υποδείγματος (αυτοσχέτιση, ετεροσκεδαστικότητα κλπ)
- 6) Αξιολογούμε τα αποτελέσματα και βγάζουμε συμπεράσματα και προτάσεις για την ακολουθητέα οικονομική πολιτική.

π.χ. Έστω ότι έχουμε να μελετήσουμε τη ζήτηση βοδινού κρέατος

1)	Ζήτηση βοδινού	τιμή βοδινού	εισόδημα	τιμή αμνοεργιών	τιμή κοτόπουλων	τιμή χοιρινού
	Y	X ₁	X ₂	X ₃	X ₄	X ₅

Έστω ότι η σχέση που συνδέει τις μεταβλητές είναι γραμμικής μορφής

$$Y = a + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \varepsilon_t$$

- 2) Εκτιμώ το υπόδειγμα (υπολογίζω τα $a, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5$) με την μέθοδο των ελαχίστων τετραγώνων και υπολογίζω την εξίσωση παλινδρόμησης $\hat{Y} = a + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5$

Έστω ότι εκτιμήσαμε την εξίσωση παλινδρόμησης και είναι η:

$$Y_t = 100 - 0,1X_1 + 0,15X_2 + 1,2X_3 + 0,2X_4 + 1,1X_5$$

- 3) Το πρόσχημο της τιμής του συντελεστή της μεταβλητής X_1 , συμφωνεί με την οικονομική θεωρία, γιατί όταν αυξάνει η τιμή μειώνεται η ζήτηση.

Το πρόσχημο της τιμής του συντελεστή της μεταβλητής X_2 , συμφωνεί με την οικονομική θεωρία, γιατί όταν αυξάνει το εισόδημα αυξάνει η ζήτηση. Εξ' άλλου ο συντελεστής του X_2 , είναι μέσα στα όρια μηδέν και ένα που καθορίζονται από την οικονομική θεωρία.

Με τον ίδιο τρόπο εξετάζουμε και τα πρόσχημα για τις X_3, X_4, X_5

- 4) Ο στατιστικός έλεγχος είναι απαραίτητος γιατί μπορεί π.χ. να ισχύει ότι $H_0 : \beta_4 = 0$, οπότε το X_4 παραλείπεται.

Κάνουμε έλεγχο για κάθε β χωριστά.

Κάνουμε τον 1ο και 2ο έλεγχο.

5) Συνεχίζουμε με τον οικονομετρικό έλεγχο του υποδείγματος, έτσι ώστε, αν διαπιστωθούν οικονομετρικά προβλήματα, δηλ. πα-
ραβίαση μιας ή περισσότερων υποθέσεων του κλασσικού γραμμικού
υποδείγματος, να τα λάβουμε υπ' όψην μας κατά την αξιολόγηση
του υποδείγματος, εφ' όσον δε βρούμε τρόπους αντιμετώπισής
τους.

3.12. ΣΧΕΣΕΙΣ ΜΕΤΑΞΥ ΤΩΝ ΣΥΝΤΕΛΕΣΤΩΝ ΠΑΛΙΝΔΡΟΜΗΣΗΣ

Θα εξετάσουμε τι συμβαίνει σε ένα υπόδειγμα αν προσθέσουμε
ή αφαιρέσουμε ανεξάρτητες μεταβλητές ως προς τους συντελεστές
παλινδρόμησης, δηλ. το ενδιαφέρον μας συγκεντρώνεται στο να
εξετάσουμε την μεταβολή που υφίστανται οι συντελεστές παλινδρό-
μησης από την προσθήκη ή αφαίρεση ανεξαρτήτων μεταβλητών σε
ένα υπόδειγμα.

Εστω το διμεταβλητό υπόδειγμα σε ένα δείγμα και σε αποκλί-
σεις από τους μέσους: $y_t = b_{e1} \cdot x_1 + e_1$ (1)

Αν στο υπόδειγμα (1) προσθέσουμε τη μεταβλητή x_2 , τότε θα
πάρουμε το τριμεταβλητό υπόδειγμα: $y_t = b_1 x_1 + b_2 x_2 + e_{12}$ (2)

Συνήθως όμως στα οικονομικά φαινόμενα υπάρχει συσχέτιση
μεταξύ των ανεξαρτήτων μεταβλητών. Αν υποθέσουμε ότι μεταξύ
των x_1 και x_2 υπάρχει κάποια γραμμική σχέση εξάρτησης, θα
έχουμε την εξίσωση σε αποκλίσεις από τους μέσους:

$$x_2 = b_{u2} x_1 + u_2 \quad (3)$$

Μεταξύ των συντελεστών παλινδρόμησης των 3 παραπάνω εξι-
σώσεων αποδεικνύεται ότι ισχύει η σχέση: $b_{e1} = b_1 + b_{u2} \cdot b_2$ (4)

Η προσθήκη (ή αφαίρεση) μιας ή περισσότερων ανεξαρτήτων με-
ταβλητών αλλάζει τους συντελεστές παλινδρόμησης εκτός αν η προ-
στιθεμένη μεταβλητή και η συμπεριλαμβανόμενη στο υπόδειγμα εί-
ναι ασυσχέτιστες, οπότε $b_{u2} = 0$ και $b_{e1} = b_1$

Απόδειξη της (4)

$$b_{e1} = \frac{\Sigma y x_1}{\Sigma x_1^2}, \quad b_{u2} = \frac{\Sigma x_1 x_2}{\Sigma x_1^2}, \quad b_1 = \frac{\Sigma x_1 y \cdot \Sigma x_2^2 - \Sigma x_2 y \cdot \Sigma x_1 x_2}{\Sigma x_1^2 \cdot \Sigma x_2^2 - (\Sigma x_1 x_2)^2}$$

$$b_2 = \frac{\Sigma x_2 y \cdot \Sigma x_1^2 - \Sigma x_1 y \cdot \Sigma x_1 x_2}{\Sigma x_1^2 \cdot \Sigma x_2^2 - (\Sigma x_1 x_2)^2}$$

Κάνω τις πράξεις

$$\begin{aligned} b_1 + b_{u2} \cdot b_2 &= \frac{\Sigma x_1 y \cdot \Sigma x_2^2 - \Sigma x_2 y \cdot \Sigma x_1 x_2}{\Sigma x_1^2 \cdot \Sigma x_2^2 - (\Sigma x_1 x_2)^2} + \frac{\Sigma x_1 x_2}{\Sigma x_1^2} \cdot \frac{\Sigma x_2 y \cdot \Sigma x_1^2 - \Sigma x_1 y \cdot \Sigma x_1 x_2}{\Sigma x_1^2 \cdot \Sigma x_2^2 - (\Sigma x_1 x_2)^2} = \\ &= \frac{\Sigma x_1 y \cdot \Sigma x_2^2 - \Sigma x_2 y \cdot \Sigma x_1 x_2}{\Sigma x_1^2 \cdot \Sigma x_2^2 - (\Sigma x_1 x_2)^2} + \frac{\Sigma x_1 x_2 \cdot \Sigma x_2 y \cdot \Sigma x_1^2 - \Sigma x_1 y (\Sigma x_1 x_2)^2}{\Sigma x_1^2 [\Sigma x_1^2 \cdot \Sigma x_2^2 - (\Sigma x_1 x_2)^2]} = \\ &= \frac{\Sigma x_1^2 \cdot \Sigma x_1 y \cdot \Sigma x_2^2 - \Sigma x_1^2 \cdot \Sigma x_2 y \cdot \Sigma x_1 x_2 + \Sigma x_1 x_2 \cdot \Sigma x_2 y \cdot \Sigma x_1^2 - \Sigma x_1 y (\Sigma x_1 x_2)^2}{\Sigma x_1^2 [\Sigma x_1^2 \cdot \Sigma x_2^2 - (\Sigma x_1 x_2)^2]} = \\ &= \frac{\Sigma x_1^2 \cdot \Sigma x_1 y \cdot \Sigma x_2^2 - \Sigma x_1 y \cdot (\Sigma x_1 x_2)^2}{\Sigma x_1^2 [\Sigma x_1^2 \cdot \Sigma x_2^2 - (\Sigma x_1 x_2)^2]} = \frac{\Sigma x_1 y [\Sigma x_1^2 \cdot \Sigma x_2^2 - (\Sigma x_1 x_2)^2]}{\Sigma x_1^2 [\Sigma x_1^2 \cdot \Sigma x_2^2 - (\Sigma x_1 x_2)^2]} = \\ &= \frac{\Sigma x_1 y}{\Sigma x_1^2} = b_{e1} \end{aligned}$$

3.1. ΣΧΕΣΕΙΣ ΜΕΤΑΞΥ ΤΩΝ ΣΥΝΤΕΛΕΣΤΩΝ ΠΡΟΣΔΙΟΡΙΣΜΟΥ

Η προσθήκη μιας ή περισσότερων ανεξαρτήτων μεταβλητών σε ένα υπόδειγμα αυξάνει την τιμή του συντελεστή προσδιορισμού, είτε οι προστιθέμενες μεταβλητές προσφέρουν επιπλέον ερμηνεία στη μεταβλητικότητα της y είτε όχι.

Έστω τα υποδείγματα της προηγούμενης παραγράφου με τους αντίστοιχους δείκτες προσδιορισμού: για την (1) R_{y1}^2 , για την (2)

R_{y12}^2 , για την (3) R_{12}^2

Αποδεικνύεται ότι μεταξύ των συντελεστών προσδιορισμού ισχύει η σχέση:

$$R_{y.12}^2 = R_{y1}^2 + (1 - R_{y1}^2) \cdot R_{e_1u_2}^2 \quad (5)$$

Από την (1) έχουμε : $e_1 = y_1 - b_{e_1}x_1$

Από την (3) έχουμε : $u_2 = x_2 - b_{u_2}x_1$

Όπου $R_{e_1u_2}^2$ είναι ο συντελεστής μερικού προσδιορισμού στο υπόδειγμα (2) και μετράει την <<καθαρή επίδραση>> της μεταβλητής x_2 πάνω στην y αφού προηγουμένως και από τη y και από τη x_2 , έχει αφαιρεθεί η επίδραση της x_1

Το $R_{e_1u_2}^2$ συνήθως συμβολίζεται με $R_{y2.1}^2$

ΑΣΚΗΣΗ

Να εκτιμηθεί το παρακάτω υπόδειγμα και να γίνει ο στατιστικός έλεγχος. Κρίνεται το υπόδειγμα από οικονομετρικής πλευράς. Ποιά προβλήματα παρουσιάζει η εκτίμηση του με την OLS που χρησιμοποιήσατε;

$$C_i = \beta_0 + \beta_1 Y_1 + \beta_2 Y_2 + u_i$$

Δίνονται:

$$ee' = \begin{bmatrix} 0,004 & 0 & \dots\dots 0 \\ 0 & 0,004 & \dots\dots 0 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ 0\dots\dots 0 & \dots\dots 0,004 \end{bmatrix} \quad x'x = \begin{bmatrix} 5 & 17 & 15 \\ 17 & 61 & 52 \\ 15 & 52 & 49 \end{bmatrix} \quad x'y = \begin{bmatrix} 4,33 \\ 15,32 \\ 13,12 \end{bmatrix}$$

$$\bar{x} = \begin{bmatrix} 1 \\ 3,4 \\ 3 \end{bmatrix}, \quad \bar{y} = 0,866, \quad s^2 = 0,004$$

και για $\alpha = 0,005$ $t_{\alpha/2, 2} = 4,305$

ΛΥΣΗ

$$b = \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix}, \quad b = (x'x)^{-1} \cdot x'y, \quad (x'x)' = \begin{bmatrix} 5 & 17 & 15 \\ 17 & 61 & 52 \\ 15 & 52 & 49 \end{bmatrix}$$

$$\text{adj } (x'x) = \begin{bmatrix} + (61 \cdot 49 - 52^2) & - (17 \cdot 49 - 15 \cdot 52) & + (17 \cdot 52 - 51 \cdot 61) \\ - (17 \cdot 49 - 52 \cdot 15) & + (5 \cdot 49 - 15^2) & - (5 \cdot 52 - 15 \cdot 17) \\ + (17 \cdot 52 - 15 \cdot 61) & - (5 \cdot 52 - 17 \cdot 15) & + (5 \cdot 61 - 17^2) \end{bmatrix} = \begin{bmatrix} 285 & -53 & -31 \\ -53 & 20 & -5 \\ -31 & -5 & 16 \end{bmatrix}$$

$$(x'x)^{-1} = \frac{\text{adj } (x'x)}{|x'x|} = \frac{1}{59} \begin{bmatrix} 285 & -53 & -31 \\ -53 & 20 & -5 \\ -31 & -5 & 16 \end{bmatrix} = \begin{bmatrix} 4,8 & -0,9 & -0,5 \\ -0,9 & 0,3 & -0,08 \\ -0,5 & -0,08 & 0,27 \end{bmatrix}$$

$$|x'x| = \begin{vmatrix} 5 & 17 & 15 \\ 17 & 61 & 52 \\ 15 & 52 & 49 \end{vmatrix} = \begin{vmatrix} 5 & 17 \\ 17 & 61 \\ 15 & 52 \end{vmatrix} = 5 \cdot 61 \cdot 49 + 17 \cdot 52 \cdot 15 + 15 \cdot 17 \cdot 52 - 15 \cdot 61 \cdot 15 - 52 \cdot 52 \cdot 5 - 49 \cdot 17 \cdot 17 = 59$$

$$b = (x'x)^{-1} \cdot x'y = \begin{bmatrix} 4,8 & -0,9 & -0,5 \\ -0,9 & 0,3 & -0,08 \\ -0,5 & -0,08 & 0,27 \end{bmatrix} \cdot \begin{bmatrix} 4,33 \\ 15,32 \\ 13,12 \end{bmatrix} = \begin{bmatrix} 4,8 \cdot 4,33 - 0,9 \cdot 15,32 - 0,5 \cdot 13,12 \\ -0,9 \cdot 4,33 + 0,3 \cdot 15,32 - 0,08 \cdot 13,12 \\ -0,5 \cdot 4,33 - 0,08 \cdot 15,32 - 0,27 \cdot 13,12 \end{bmatrix} =$$

3x3 3x1

$$= \begin{bmatrix} 0,436 \\ -0,35 \\ 0,15 \end{bmatrix} \quad \text{Άρα } b_0 = 0,436, b_1 = -0,35 \text{ και } b_2 = 0,15$$

Στατιστικός έλεγχος

Για να γίνει χρειάζονται οι διακυμάνσεις των b που υπολογίζονται από τη μήτρα διακυμάνσεων-συνδιακυμάνσεων και από τον τύπο:

$$s_b^2 = s^2 (x'x)^{-1}$$

$$s_b^2 = 0,004 \cdot \begin{bmatrix} 4,8 & -0,9 & -0,5 \\ -0,9 & 0,3 & -0,08 \\ -0,5 & -0,08 & 0,27 \end{bmatrix} = \begin{bmatrix} 0,019 & -0,0036 & -0,02 \\ -0,0036 & 0,0012 & -0,00032 \\ -0,02 & -0,00032 & 0,0011 \end{bmatrix}$$

Τα διαγώνια στοιχεία της s_b^2 είναι οι διακυμάνσεις των b_0, b_1, b_2 αντίστοιχα και τα μη διαγώνια οι συνδιακυμάνσεις των b .

$$\begin{bmatrix} s_{b_0}^2 & \text{cov}(b_0, b_1) & \text{cov}(b_0, b_2) \\ \text{cov}(b_0, b_1) & s_{b_1}^2 & \text{cov}(b_1, b_2) \\ \text{cov}(b_0, b_2) & \text{cov}(b_1, b_2) & s_{b_2}^2 \end{bmatrix}$$

$$\text{Άρα } s_{b_0}^2 = 0,019, s_{b_1}^2 = 0,0012, s_{b_2}^2 = 0,0011$$

α) Υποθέσεις

$$H_0 : \beta_0 = 0$$

$$H_1 : \beta_0 \neq 0$$

$$t = \frac{|b_0|}{s_{b_0}} = \frac{|0,436|}{\sqrt{0,019}} = 3,16$$

Έχουμε $t_{\alpha/2, 2} = 4,305$. Επειδή $t < t_{\alpha/2, 2}$ γίνεται δεκτή η H_0 . Δηλαδή δεν υπάρχει σταθερός όρος.

β) Υποθέσεις

$$H_0 : \beta_1 = 0$$

$$H_1 : \beta_1 \neq 0 \quad t = \frac{|b_1|}{s_{b_1}} = \frac{|-0,35|}{\sqrt{0,0012}} = 10,1$$

Επειδή $t > t_{\alpha/2, 2}$ απορρίπτεται η H_0 γίνεται δεκτή η H_1 . Η τιμή του b_1 είναι στατιστικά σημαντική (δηλ. στον πληθυσμό υπάρχει b_1).

γ) Υποθέσεις

$$H_0 : \beta_2 = 0$$

$$H_1 : \beta_2 \neq 0 \quad t = \frac{|b_2|}{s_{b_2}} = \frac{|0,15|}{\sqrt{0,0011}} = 4,52$$

Επειδή $t > t_{\alpha/2, 2}$ απορρίπτεται η H_0 και γίνεται δεκτή η H_1 . Η τιμή του b_2 είναι στατιστικά σημαντική.

Λόγω έλλειψης στοιχείων δεν μπορούμε να κάνουμε έλεγχο της υπόθεσης ότι όλα τα β συγχρόνως είναι ίσα με το μηδέν δηλ. να εξετάσουμε την υπόθεση $H_0 : \beta_1 = \beta_2 = 0$

H_1 : ένα τουλάχιστο β είναι διάφορο του μηδενός

ΑΣΚΗΣΗ

Από ένα δείγμα 15 ετήσιων παρατηρήσεων για τη ζητούμενη ποσότητα (Y) ενός αγαθού, την τιμή του (X_1), το εισόδημα των καταναλωτών (X_2) και την τιμή του υποκατάστατου αγαθού (X_3), εκτιμήθηκε η ακόλουθη συνάρτηση:

$$Y = 79,1 - 4,93X_1 + 0,016X_2 + 0,175X_3 \quad R^2 = 0,95$$

(1,61) (0,074) (0,64)

Οι αριθμοί στις παρενθέσεις είναι τα τυπικά σφάλματα των συντελεστών.

α) Είναι σωστά εξειδικευμένη η συνάρτηση; β) Είναι τα αποτελέσματα της εκτίμησης σύμφωνα με την οικονομική θεωρία; γ) Να γίνει ο στατιστικός έλεγχος του υποδείγματος, αν για $\alpha=0,05$, $t_{11} = 1,8$ και $F_{3,11} = 3,6$

δ) Σε τι συνίσταται ο οικονομετρικός έλεγχος ενός υποδείγματος;

ΛΥΣΗ

Έχουμε $S_{b_1} = 1,61$, $S_{b_2} = 0,074$, $S_{b_3} = 0,64$

α) Είναι σωστές οι μεταβλητές που συμπεριλήφθηκαν στο υπόδειγμα γιατί η ζήτηση ενός αγαθού εξαρτάται από την τιμή του, το εισόδημα του καταναλωτή και την τιμή υποκατάστατου αγαθού.

Ως προς τη μορφή της μαθηματικής σχέσης μεταξύ των μεταβλητών πράγματι μπορεί να είναι γραμμική.

β) Τα πρόσημα των συντελεστών είναι σωστά γιατί όταν αυξάνει η τιμή μειώνεται η ζήτηση, όταν αυξάνει το εισόδημα αυξάνει η ζήτηση, όταν αυξάνει η τιμή υποκατάστατου αυξάνεται η ζήτηση. Επίσης ο συντελεστής εισοδήματος βρίσκεται μεταξύ 0 και 1 και συμφωνεί με την οικονομική θεωρία.

γ) 1ος έλεγχος

Υποθέσεις

$$H_0 : \beta_1 = 0$$

$$H_1 : \beta_1 \neq 0$$

$$t = \frac{|b_1|}{S_{b_1}} = \frac{4,93}{1,61} = 3,062$$

Επειδή $t > t_{11} = 1,8$ απορρίπτεται, η H_0 και γίνεται δεκτή η H_1

Υποθέσεις

$$H_0 : \beta_2 = 0$$

$$H_1 : \beta_2 \neq 0$$

$$t = \frac{|b_2|}{S_{b_2}} = \frac{0,016}{0,074} = 0,21$$

Επειδή $t < t_{11} = 1,8$ γίνεται δεκτή η H_0 , δηλαδή η b_2 είναι στατιστικά ασήμαντη. Φαίνεται ότι η X_2 δεν επηρεάζει τη ζήτηση του προϊόντος

Υποθέσεις

$$H_0 : \beta_3 = 0$$

$$H_1 : \beta_3 \neq 0 \quad t = \frac{|b_3|}{s_{b_3}} = \frac{0,175}{0,64} = 0,27$$

Επειδή $t < t_{11} = 1,8$ γίνεται δεκτή η H_0 . Κι εδώ η τιμή του b_3 είναι στατιστικά ασήμαντη.

2ος έλεγχος

Υποθέσεις

$$H_0 : \beta_1 = \beta_2 = \beta_3 = 0$$

H_1 : ένα τουλάχιστον β είναι διάφορο του μηδενός

$$F = \frac{\frac{R^2}{N}}{\frac{1 - R^2}{T - N - 1}} = \frac{\frac{0,95}{3}}{\frac{1 - 0,95}{15 - 3 - 1}} = 69,66$$

Επειδή $F > F_{3,11} = 3,6$ απορρίπτεται η H_0 και γίνεται δεκτή η H_1 .

3.14. ΔΙΟΡΘΩΜΕΝΟΣ ΣΥΝΤΕΛΕΣΤΗΣ ΠΡΟΣΔΙΟΡΙΣΜΟΥ

Όπως αναφέραμε στη σχέση μεταξύ των συντελεστών προσδιορισμού, η προσθήκη μιας ή περισσότερων ανεξάρτητων μεταβλητών σε ένα υπόδειγμα αλλάζει την τιμή του συντελεστή προσδιορισμού και συγκεκριμένα την αυξάνει, είτε η προστιθέμενη μεταβλητή εί-

ναι σημαντική, είτε όχι. Επομένως δεν είναι δυνατό να συγκρίνουμε υποδείγματα με διάφορο αριθμό ανεξαρτήτων μεταβλητών (για τους παραπάνω λόγους) με κριτήριο το R^2 .

Για να συγκρίνουμε εξισώσεις με διάφορο αριθμό ανεξαρτήτων μεταβλητών ή με τον ίδιο αριθμό ανεξαρτήτων μεταβλητών, αλλά διαφορετικό μέγεθος δείγματος χρησιμοποιούμε το διορθωμένο συντελεστή προσδιορισμού $\bar{R}^2$, ο οποίος αντισταθμίζει την αύξηση που παίρνει ο απλός συντελεστής προσδιορισμού από την προσθήκη ανεξάρτητης μεταβλητής και πλησιάζει περισσότερο στην πραγματική τιμή:

$$\bar{R}^2 = R^2 - (1 - R^2) \frac{N}{T - N - 1}$$

Ο $\bar{R}^2$ μπορεί να πάρει και αρνητικές τιμές, όταν οι τιμές του R^2 είναι χαμηλές, ή όταν αυξάνεται ο αριθμός των ανεξαρτήτων μεταβλητών ενώ το μέγεθος του δείγματος παραμένει σταθερό.