Επαναληπτικές Ασκήσεις στην Οικονομετρία

1). Έστω ότι θέλουμε να εξετάσουμε αν οι καταναλωτικές δαπάνες εμφανίζουν γραμμική σχέση με το εισόδημα και τον πλούτο. Από τη μελέτη αυτή προκύπτει το ακόλουθο μοντέλο

$$\hat{y}_i = 24,7747 + 0,9415X_{2i} - 0,0424X_{3i}$$

τυπικά σφάλματα	(6,7525)	(0,8229)	(0,0807)
t	(3,6690)	(1,1442)	(−0,5221)

$R^2 = 0,9635$, $R^2_{adj} = 0,9531$, $F = 92,4019$

όπου y_i ~> καταναλωτικές δαπάνες

X_2 ~> Εισόδημα

X_3 ~> Πλούτος

$n = 10$ (αριθμός παρατηρήσεων)

Να σχολιαστούν τα αποτελέσματα

Σχολιασμός

- $R^2 = 0,9635$ ~> Το 96,35% της μεταβολής των καταναλωτικών δαπανών ερμηνεύεται από το εισόδημα και τον πλούτο

- Όταν το εισόδημα και ο πλούτος είναι μηδέν (δηλ. μηδενικό εισόδημα και μηδενικός πλούτος τότε οι καταναλωτικές δαπάνες είναι ίσες με 24,7747.

Αν το εισόδημα αυξηθεί κατά 1 μονάδα και ο πλούτος παραμείνει σταθερός, τότε οι καταναλωτικές δαπάνες θα αυξηθούν κατά 0,9415

Αν ο πλούτος αυξηθεί κατά 1 μονάδα και το εισόδημα παραμείνει σταθερό τότε οι καταναλωτικές δαπάνες θα μειωθούν κατά 0,0424. Η [illegible]

Το πρόσημο της εκτίμησης αυτής δεν είναι το αναμενόμενο σύμφωνα με την οικονομική θεωρία. Αφού, κάποιος θα περίμενε μια θετική σχέση μεταξύ της κατανάλωσης και του πλούτου

Θα ελέγξουμε τώρα αν η πάνω εκτίμηση είναι στατιστικά σημαντική

$H_0: b_2 = 0$ vs $H_1: b_2 \neq 0$

$t = \frac{\hat{b}_2 - b_{20}}{S_{\hat{b}_2}}$ το οποίο δίνεται και δεν

χρειάζεται να υπολογίσει t για την $\hat{b}_2$ $= 1,1442$ θα το συγκρίνω με $t_{n-3,\alpha/2}$

$= t_{7,\,0.025} = 2,365$ (για $\alpha = 0,05$)

$1,1442 < 2,365 \Rightarrow$ δεν απορρ. την $H_0 \Rightarrow$ δεν είναι στατιστικά σημαντική

$H_0: b_3 = 0$ vs $H_3: b_3 \neq 0$

$t = -0,5261$ θα το συγκρίνω με το $-t_{n-3,\alpha/2} = -2,365$

Το $-0,5261 > -2,365 \Rightarrow$ δεν απορρίπτουμε την $H_0 \Rightarrow$ δεν είναι στατιστικά σημαντική η εκτίμηση

Τώρα θα ελέγξουμε αν η παλινδρόμηση είναι στατιστικά σημαντική

$H_0: b_2 = b_3 = 0$ vs. $H_1:$ κάποιο $b_j \neq 0$
$j = 2, 3$

$F = 92,4019 \sim F_{(k-1),(n-k)} = F_{2,7}$

όπου $k=3$, $n=10$

Θα χρησιμοποιήσω $\alpha = 0,05$

Άρα κοιτάω στους πίνακες $F_{2,7,0.05} = 4,74$

$92,4019 > 4,74 \Rightarrow$ απορρ. $H_0 \Rightarrow$ η παλινδρόμηση είναι στατιστικά σημαντική

Καταλήγουμε στο συμπέρασμα ότι παρόλο που οι b_2 και b_3 μεμονωμένα είναι στατιστικά μη σημαντικοί η παλινδρόμηση είναι στατιστικά σημαντική. $\Rightarrow$ καταλήγουμε στο συμπέρασμα ότι υπάρχει πολυσυγγραμμικότητα

2) Έστω ότι μας ενδιαφέρει να εξετάσουμε την κατανάλωση (Consum) με το εισόδημα (GDP) και τον πλούτο (Wealth). Υποθέτουμε ότι η μεταβλητή πλούτος αντιπροσωπεύει την άθροιση των μισθών και των κερδών. Αρχικά παλινδρομούμε την κατανάλωση πάνω στο εισόδημα και στον πλούτο και προκύπτουν τα αποτελέσματα του Πίνακα 1. Εν συνεχεία παλινδρομούμε την εξαρτημένη μεταβλητή με κάθε μια ανεξάρτητη μεταβλητή και προκύπτουν τα αποτελέσματα των Πινάκων 2 και 3 αντίστοιχα. Να σχολιασθούν τα αποτελέσματα.

Πίνακας 1

The regression equation is

Consum = -753 + 3,8 GDP - 0,89 Wealth

Predictor	Coef	SE Coef	T	P	VIF
Constant	-752,9	220,4	-3.42	0.002	
GDP	3.80	26.50	0.14	0.887	867560.2
Wealth	-0.889	7.793	-0.11	0.910	867560,2

S = 432,189 , R-Sq = 95,5% , R-Sq (adj) = 95,3%

Analysis of Variance

Source	Df	SS	MS	F	P
Regression	2	140097329	70048665	375,02	0.000
Residual Error	35	6537551	186787		
Total	37	146634880			

Source	Df	Seq SS
GDP	1	140094900
Wealth	1	2429

Πίνακας 2

The regression equation is
Consum = - 753 + 0,779 GDP

Predictor	Coef	SE Coef	T	P
Constant	-752,7	217,3	-3.46	0,001
GDP	0.77903	0.02805	27.77	0.000

S = 426,223   R-Sq = 95,5 %   R-Sq(adj) = 95,4%

Analysis of Variance

Source	Df	SS	MS	F	p
Regression	1	140094900	140094900	771,17	0.000
Residual Error	36	6539980	181666		
Total	37	146634880			

Πίνακας 3

The regression equation is
Consum = - 753 + 0,229 Wealth

Predictor	Coef	SE Coef	T	P
Constant	-752,7	217.4	-3.46	0.001
Wealth	0.229137	0.008252	27,77	0.000

S = 426,269   R-Sq = 95,5%   R-sq(adj) = 95,4%

Analysis of Variance

Source	Df	SS	MS	F	p
Regression	1	140094227	140094227	771,08	0.0000
Residual Error	36	6540653	181685		για υπάρχ.
Total	37	146634880			

Σχολιασμός

Πίνακας 1 ~> Τα αποτελέσματα είναι παράδοξα.
- Η οριακή ροπή προς κατανάλωση είναι ίση με 3,8 > 1. Ενώ ο πλούτος έχει αρνητικό πρόσημο σε αντίθεση με το αναμενόμενο της οικονομικής θεωρίας.
R^2 = 95,5% αρκετά μεγάλο ποσοστό.
Ερμηνεία του R^2: Το 95,5% της κατανάλωσης ερμηνεύεται από το εισόδημα και τον πλούτο και το υπόλοιπο 4,5% από άλλους τυχαίους παράγοντες.

• $H_0: b_1 = 0$ vs $H_1: b_1 \neq 0$ ~> έλεγχος για την εκτίμηση του συντελεστή της ανεξάρτητης μεταβλητής «εισόδημα»

p-value = 0,887 > α = 0,05

⇒ δεν απορρ. η H_0 δηλ. η εκτίμηση δεν είναι στατιστικά σημαντική.

• $H_0: b_2 = 0$ vs $H_1: b_2 \neq 0$ ~> έλεγχοι για την εκτίμηση του συντελεστή της ανεξάρτητης μεταβλητής «πλούτος»

p-value = 0,910 > 0,05 ⇒

⇒ δεν απορρίπτεται η H_0 δηλ. η εκτίμηση δεν είναι στατιστικά σημαντική.

- VIF = 867500,2 > 10. Ο συντελεστής αυτός επιβεβαιώνει ήδη το συμπέρασμα που έχω προκύψει πιο πάνω. Άρα υπάρχει έντονη πολυσυγγραμμικότητα

Κοιτώντας τώρα τους άλλους δύο Πίνακες διαπιστώνουμε ότι τα αποτελέσματα τώρα διαφοροποιούνται, και υπάρχει πλέον στατιστική σημαντικότητα των ανεξάρτητων μεταβλητών. Επίσης διαπιστώνεται ότι τα προβλήματα που είχαν δημιουργηθεί έχουν ήδη επιλυθεί. Η οριακή ροπή προκατανάλωση είναι ίση με 0,7779 (θετική και <1) ενώ ο πλούτος έχει θετικό πρόσημο.

3). Έστω ότι Y η ~~...~~ εξαρτημένη μεταβλητή μας και X_1, X_2, X_3 οι ανεξάρτητες μεταβλητές. Τα αποτελέσματα της παλινδρόμησης της Y πάνω στις X_1, X_2, X_3 παρουσιάζονται παρακάτω αρχικά για όλες τις παρατηρήσεις $(n=20)$ και μετά για τις 8 μικρότερες και για τις 8 μεγαλύτερες.

Εξίσωση πολλαπλής παλ/σης (όλες οι παρατηρήσεις)

$$\hat{Y} = -6.39 + 0.371X_1 - 0.058X_2 + 0.0375X_3$$
$$(2.149) \quad (0.2987) \quad (0.1015) \quad (0.01026)$$

$R^2 = 91,3\%$, $SSE = 1,8759$

Εξίσωση (οι 8 μικρότερες)

$$\hat{Y} = -15.7 + 0.722X_1 + 0.077X_2 + 0.0175X_3$$
$$(3.756) \quad (0.4355) \quad (0.1847) \quad (0.007485)$$

$R^2 = 97\%$, $SSE_1 = 0.0940$

Εξίσωση (οι 8 μεγαλύτερες)

$$\hat{Y} = 7.9 - 0.132X_1 + 0.534X_2 - 0.0347X_3$$
$$(11.31) \quad (0.7434) \quad (0.6954) \quad (0.07580)$$

$R^2 = 42,2\%$, $SSE_2 = 0.8429$

a. Ελέγξτε την ύπαρξη ετεροσκεδαστικότητας.

b. Υποθέτοντας ότι η διασπορά του σφάλματος είναι ανάλογη προς την X_3^2 διορθώσαμε το πρόβλημα της ετεροσκεδαστικότητας και πήραμε τα ακόλουθα αποτελέσματα. Σχολιάστε

$$\hat{Y}/X_3 = 0.0378 - 6.721/X_3 + 0.410X_1/X_3 - 0.0648X_2/X_3$$
$$(0.009319) \quad (1.994) \quad (0.2946) \quad (0.09708)$$

Λύση

α. Υπολογίζουμε τον λόγο

$$\lambda = \frac{SSE_2}{SSE_1} = \frac{0,8429}{0,094} = 8,97$$

Θα πρέπει να συγκριθεί με την κρίσιμη τιμή της F κατανομής για β.ε. $(n-c-2k)/2$
$= (20-4-2\times4)/2 = 4$

δηλ. $F_{4,4,0.05} = 6,39$

$8,97 > 6,39 \Rightarrow$ απορρίπτουμε την Ho
$\Rightarrow$ έχουμε πρόβλημα ετεροσκεδαστικότητας.

H_0: υπάρχει ~~ομοσκεδαστικότητα~~ κατά
H_1: υπάρχει ετεροσκεδαστικότητα
$\Rightarrow$ ο έλεγχος

β. Θα κάνουμε τον ακόλουθο έλεγχο για να δούμε αν διορθώθηκε η ετεροσκεδαστικότητα.
Ο έλεγχος είναι ο παρακάτω.
Θα χρησιμοποιήσω το τεστ κατά White

$$\chi^2 = nR^2 = 20 \cdot 0,91 = 18,2$$

$\chi^2_k = \chi^2_3$ k ⇒ αριθμός ανεξάρτητων μεταβλητών.

Για $\alpha = 0,05$

$$\chi^2_{3,0.05} = 7,815$$

$18,2 > 7,815 \Rightarrow$ απορρ. $\Rightarrow$ Υπάρχει ετεροσκεδαστικότητα

4) Έστω ότι από την παλινδρόμηση ανάμεσα στη Y και X από ένα δείγμα 40 παρατηρήσεων προέκυψε ότι $\sum \hat{u}^2 = 0,82$ και $d = 0,55$. Επιπλέον, το άθροισμα των τετραγώνων των υπολοίπων από την παλινδρόμηση ανάμεσα στις πρώτες διαφορές των δύο μεταβλητών είναι $\sum \hat{e}^2 = 0,94$. Μπορεί να γίνει δεκτή η υπόθεση $\rho = 1$; ($a = 0,01$)

Λύση

H_0: δεν υπάρχει αυτοσυσχέτιση

vs

H_1: ~~δεν~~ υπάρχει αυτοσυσχέτιση

$d = 0.55 < d_L = 1,25$ (από πίνακα) για $a = 0.01$

Διαπιστώνεται ότι υπάρχει αυτοσυσχέτιση.

Τώρα, για τον έλεγχο $H_0: \rho = 1$ vs $H_1: \rho \neq 1$

θα εφαρμοστεί ο έλεγχος Berenblutt-Webb

$$g = \frac{\sum \hat{e}^2}{\sum \hat{u}^2} = \frac{0,94}{0,82} = 1,146$$

Επειδή $g < 1,25$ για $a = 0.01$ μπορεί να γίνει δεκτό ότι $\rho = 1$.

5) Έστω ότι η καθαρή επένδυση (Y) και η μεταβολή στις πωλήσεις (X) συνδέονται με το απλό γραμμικό μοντέλο $Y = \beta_0 + \beta_1 X + \varepsilon$.
Ένα τυχαίο δείγμα από 5 στοιχεία έδωσε τα παρακάτω αποτελέσματα (μονάδα μέτρησης και των δύο μεταβλητών εκατομμύρια ευρώ).

Y	10	24	41	14	6
X	30	20	70	50	-20

α. Να εκτιμηθούν τα β_0, β_1 με τη μέθοδο των ελαχίστων τετραγώνων.
β. Είναι η σχέση ανάμεσα στις επενδύσεις και στη μεταβολή των πωλήσεων στατιστικά σημαντική;
γ. Πως είναι γνωστό το παραπάνω μοντέλο από την οικονομική θεωρία
δ. Πόσο θα γίνουν οι επενδύσεις αν οι πωλήσεις αυξηθούν κατά 10 εκατομμύρια.

Λύση

Πρέπει να υπολογιστούν αναλυτικά τα ανάλογα

$\Sigma xy = 4.230$

$\Sigma x^2 = 9.100$

$\Sigma y^2 = 2589$

$\bar{y} = 30$

$\bar{x} = 19$

(Τα παραπάνω υπολογίζονται με τη μεθοδολογία που έχει εξηγηθεί αναλυτικά στην Στατιστική)

$$\hat{\beta}_1 = \frac{n\Sigma xy - (\Sigma x)(\Sigma y)}{n\Sigma x^2 - (\Sigma x)^2} = 0.3$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = 10$$

Άρα $\hat{y} = 10 + 0,3x$

Ερμηνείες κατά τα γνωστά. Δηλ.

Όταν οι πωλήσεις είναι μηδενικές οι επενδύσεις είναι ίσες με 10 εκατ. ευρώ. Όταν οι πωλήσεις αυξηθούν κατά 1 εκ. ευρώ τότε θα αυξηθούν και οι επενδύσεις κατά 300.000 ευρώ.

β. Θα γίνει ο έλεγχος $H_0: b_1=0$ κατά $H_1: b_1 \neq 0$

Για τον έλεγχο αυτόν μπορούμε να χρησιμοποιήσουμε τη στατιστική συνάρτηση $t = \frac{\hat{b}}{S_{\hat{b}}}$ ή την F που ελέγχει την κλίση

$$F = \frac{MSR}{MSE} = \frac{SSR/1}{SSE/n-2} = \frac{\sum(\hat{y}_i-\bar{y})^2}{\sum(y_i-\hat{y}_i)^2/n-2} =$$

$$= \frac{414}{123,33} = 3,357$$

Από την εξίσωση $\hat{y} = 10 + 0,3x$ για κάθε τιμή x_i υπολογίζουμε την αντίστοιχη $\hat{y}_i$. Από κάθε $\hat{y}_i$ αφαιρούμε $\bar{y}$ μετά υψώνουμε στο τετράγωνο και μετά αθροίζουμε για όλες τις τιμές και έτσι υπολογίζουμε το $\sum(\hat{y}_i-\bar{y})^2$. Για το $\sum(y_i-\hat{y}_i)^2$ αφαιρούμε από την κάθε παρατηρούμενη τιμή y την εκτιμηθείσα $\hat{y}$ υψώνουμε στο τετράγωνο και μετά παίρνουμε το άθροισμα για όλες τις τιμές.

Την τιμή $F = 3,357$ θα την συγκρίνουμε με το κρίσιμο σημείο $F_{1,3,0.05} = 10.1$
Άρα δεν απορρ. η H_0 άρα η εκτίμηση δεν είναι στατιστικά σημαντική. Αυτό ίσως να οφείλεται στο ότι το δείγμα είναι μικρό. οπότε η δύναμη του κριτηρίου είναι μικρή.

γ. Το μοντέλο αυτό μπορεί να θεωρηθεί ότι περιγράφει την αρχή του επιταχυντή σύμφωνα με την οποία οι επενδύσεις είναι συνάρτηση της μεταβολής του εισοδήματος (προϊόντος).

δ. Για να κάνουμε πρόβλεψη θα χρησιμοποιήσουμε την εκτιμηθείσα εξίσωση

$\hat{y} = 10 + 0,3x$ και όπου x θα βάλουμε το 10

Άρα $\hat{y} = 10 + 0,3 \cdot 10 = 13$

Άρα για 10 εκατ. ευρώ αύξηση οι επενδύσεις θα είναι 13 εκατ. ευρώ.