

Ψευδομεταβλητες (Απαντησεις- Ασκηση 1)

Αναλυση δεδομενων: wage.wf1

Ls wage c female

Dependent Variable: WAGE
Method: Least Squares
Date: 11/06/22 Time: 15:37
Sample: 1 1000
Included observations: 1000

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	11.52545	0.271476	42.45481	0.0000
FEMALE	-2.656750	0.386249	-6.878330	0.0000
R-squared	0.045261	Mean dependent var		10.21302
Adjusted R-squared	0.044304	S.D. dependent var		6.246641
S.E. of regression	6.106698	Akaike info criterion		6.458647
Sum squared resid	37217.17	Schwarz criterion		6.468463
Log likelihood	-3227.324	Hannan-Quinn criter.		6.462378
F-statistic	47.31143	Durbin-Watson stat		0.097275
Prob(F-statistic)	0.000000			

series male=1-female

Ls wage c male

Dependent Variable: WAGE
Method: Least Squares
Date: 11/06/22 Time: 15:40
Sample: 1 1000
Included observations: 1000

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	8.868704	0.274753	32.27879	0.0000
MALE	2.656750	0.386249	6.878330	0.0000
R-squared	0.045261	Mean dependent var		10.21302
Adjusted R-squared	0.044304	S.D. dependent var		6.246641
S.E. of regression	6.106698	Akaike info criterion		6.458647
Sum squared resid	37217.17	Schwarz criterion		6.468463
Log likelihood	-3227.324	Hannan-Quinn criter.		6.462378
F-statistic	47.31143	Durbin-Watson stat		0.097275
Prob(F-statistic)	0.000000			

ls wage male female

Dependent Variable: WAGE
Method: Least Squares
Date: 11/06/22 Time: 15:41
Sample: 1 1000
Included observations: 1000

Variable	Coefficient	Std. Error	t-Statistic	Prob.
MALE	11.52545	0.271476	42.45481	0.0000
FEMALE	8.868704	0.274753	32.27879	0.0000
R-squared	0.045261	Mean dependent var		10.21302
Adjusted R-squared	0.044304	S.D. dependent var		6.246641
S.E. of regression	6.106698	Akaike info criterion		6.458647
Sum squared resid	37217.17	Schwarz criterion		6.468463
Log likelihood	-3227.324	Hannan-Quinn criter.		6.462378
Durbin-Watson stat	0.097275			

View → Representations

Estimation Command:

```
=====
LS WAGE MALE FEMALE
```

Estimation Equation:

```
=====
WAGE = C(1)*MALE + C(2)*FEMALE
```

Substituted Coefficients:

```
=====
WAGE = 11.5254545455*MALE + 8.86870445344*FEMALE
```

Θελουμε να τεσταρουμε αν μαζι τα C(1) και C(2) ειναι στατιστικα σημαντικα η οχι (joint test). Οποτε

View → Coefficient tests → Wald και μετα μεσα στο παραθυρο γραφουμε C(1)=0, C(2)=0 → OK

Wald Test:

Equation: Untitled

Test Statistic	Value	df	Probability
F-statistic	1422.166	(2, 998)	0.0000
Chi-square	2844.331	2	0.0000

Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
C(1)	8.868704	0.274753
C(2)	11.52545	0.271476

Restrictions are linear in coefficients.

Is wage c male female

Wrong!

Is wage c female educ

Dependent Variable: WAGE
Method: Least Squares
Date: 11/07/22 Time: 16:09
Sample: 1 1000
Included observations: 1000

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-3.505029	0.961422	-3.645671	0.0003
FEMALE	-2.526996	0.344133	-7.343090	0.0000
EDUC	1.126563	0.069744	16.15284	0.0000

R-squared	0.243291	Mean dependent var	10.21302
Adjusted R-squared	0.241773	S.D. dependent var	6.246641
S.E. of regression	5.439340	Akaike info criterion	6.228188
Sum squared resid	29497.66	Schwarz criterion	6.242911
Log likelihood	-3111.094	Hannan-Quinn criter.	6.233784
F-statistic	160.2735	Durbin-Watson stat	0.470282
Prob(F-statistic)	0.000000		

ls wage c female educ female*educ

Dependent Variable: WAGE
 Method: Least Squares
 Date: 11/07/22 Time: 17:41
 Sample: 1 1000
 Included observations: 1000

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-3.562477	1.280930	-2.781165	0.0055
FEMALE	-2.400561	1.893249	-1.267959	0.2051
EDUC	1.130869	0.094280	11.99477	0.0000
FEMALE*EDUC	-0.009522	0.140199	-0.067915	0.9459
R-squared	0.243294	Mean dependent var		10.21302
Adjusted R-squared	0.241015	S.D. dependent var		6.246641
S.E. of regression	5.442057	Akaike info criterion		6.230183
Sum squared resid	29497.52	Schwarz criterion		6.249814
Log likelihood	-3111.092	Hannan-Quinn criter.		6.237644
F-statistic	106.7439	Durbin-Watson stat		0.470262
Prob(F-statistic)	0.000000			

View → Representations

Estimation Command:

=====

LS WAGE C FEMALE EDUC FEMALE*EDUC

Estimation Equation:

=====

WAGE = C(1) + C(2)*FEMALE + C(3)*EDUC + C(4)*FEMALE*EDUC

Substituted Coefficients:

=====

WAGE = -3.56247732836 - 2.40056111733*FEMALE + 1.13086854216*EDUC - 0.00952158897031*FEMALE*EDUC

Θελουμε να τεσταρουμε αν μαζι τα C(2) και C(4) ειναι στατιστικα σημαντικα η οχι (joint test). Οποτε

View → Coefficient tests → Wald και μετα μεσα στο παραθυρο γραφουμε C(2)=0, C(4)=0 → OK

Wald Test:
Equation: Untitled

Test Statistic	Value	df	Probability
F-statistic	26.93587	(2, 996)	0.0000
Chi-square	53.87174	2	0.0000

Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
C(2)	-2.400561	1.893249
C(4)	-0.009522	0.140199

Restrictions are linear in coefficients.

ls wage c educ female black

Dependent Variable: WAGE
Method: Least Squares
Date: 11/09/22 Time: 20:09
Sample: 1 1000
Included observations: 1000

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-3.270737	0.963506	-3.394622	0.0007
EDUC	1.117967	0.069647	16.05180	0.0000
FEMALE	-2.500580	0.343399	-7.281858	0.0000
BLACK	-1.512969	0.606636	-2.494033	0.0128
R-squared	0.247987	Mean dependent var		10.21302
Adjusted R-squared	0.245722	S.D. dependent var		6.246641
S.E. of regression	5.425155	Akaike info criterion		6.223962
Sum squared resid	29314.58	Schwarz criterion		6.243593
Log likelihood	-3107.981	Hannan-Quinn criter.		6.231423
F-statistic	109.4819	Durbin-Watson stat		0.479666
Prob(F-statistic)	0.000000			

ls wage c educ female black female*black

Dependent Variable: WAGE
 Method: Least Squares
 Date: 11/09/22 Time: 21:10
 Sample: 1 1000
 Included observations: 1000

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-3.230327	0.967499	-3.338841	0.0009
EDUC	1.116823	0.069714	16.01998	0.0000
FEMALE	-2.552070	0.359686	-7.095280	0.0000
BLACK	-1.831240	0.895726	-2.044418	0.0412
FEMALE*BLACK	0.587905	1.216954	0.483096	0.6291

R-squared	0.248164	Mean dependent var	10.21302
Adjusted R-squared	0.245141	S.D. dependent var	6.246641
S.E. of regression	5.427245	Akaike info criterion	6.225728
Sum squared resid	29307.71	Schwarz criterion	6.250266
Log likelihood	-3107.864	Hannan-Quinn criter.	6.235054
F-statistic	82.10655	Durbin-Watson stat	0.480319
Prob(F-statistic)	0.000000		

View → Representations

Estimation Command:

```
=====
LS WAGE C EDUC FEMALE BLACK FEMALE*BLACK
```

Estimation Equation:

```
=====
WAGE = C(1) + C(2)*EDUC + C(3)*FEMALE + C(4)*BLACK + C(5)*FEMALE*BLACK
```

Substituted Coefficients:

```
=====
WAGE = -3.23032714191 + 1.11682345046*EDUC - 2.55207038819*FEMALE - 1.83123950642*BLACK +
0.587905226322*FEMALE*BLACK
```

Θέλουμε να τεσταρουμε αν απο κοινου τα C(3), C(4) και C(5) είναι στατιστικά σημαντικά η όχι. Οποτε

View → Coefficient tests → Wald και μετα μεσα στο παραθυρο γραφουμε C(3)=0, C(4)=0, C(5)=0 → OK

Wald Test:
 Equation: Untitled

Test Statistic	Value	df	Probability
F-statistic	20.20346	(3, 995)	0.0000
Chi-square	60.61037	3	0.0000

Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
C(3)	-2.552070	0.359686
C(4)	-1.831240	0.895726
C(5)	0.587905	1.216954

Restrictions are linear in coefficients.

ls wage c educ south midwest west

Dependent Variable: WAGE
 Method: Least Squares
 Date: 11/10/22 Time: 20:02
 Sample: 1 1000
 Included observations: 1000

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-3.948547	1.054772	-3.743507	0.0002
EDUC	1.119975	0.071672	15.62639	0.0000
SOUTH	-1.558710	0.486811	-3.201881	0.0014
MIDWEST	-0.429687	0.518445	-0.828798	0.4074
WEST	-0.560673	0.527315	-1.063259	0.2879
R-squared	0.211658	Mean dependent var		10.21302
Adjusted R-squared	0.208489	S.D. dependent var		6.246641
S.E. of regression	5.557442	Akaike info criterion		6.273140
Sum squared resid	30730.73	Schwarz criterion		6.297679
Log likelihood	-3131.570	Hannan-Quinn criter.		6.282467
F-statistic	66.78581	Durbin-Watson stat		0.414173
Prob(F-statistic)	0.000000			

Chow ελεγχος

Χρησιμοποιώντας το Chow test, θέλουμε να δούμε αν στην παλινδρόμηση $wage = c + educ + female + black + female*black$ οι συντελεστές διαφοροποιούνται αναλόγα με το αν κάποιος μένει στο South ή όχι.

Το Chow test εφαρμόζεται ως εξής:

- 1) Στο παλινδρόμηση $wage = c + educ + female + black + female*black$ πολλαπλασιάζουμε όλες τις ανεξαρτητες μεταβλητες με την ψευδομεταβλητη South και αρα τρεχουμε το υποδειγμα

Is $wage = c + educ + female + black + female*black + south + educ*south + female*south + black*south + female*black*south$

Dependent Variable: WAGE
Method: Least Squares
Date: 11/12/22 Time: 22:51
Sample: 1 1000
Included observations: 1000

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-3.577536	1.151332	-3.107301	0.0019
EDUC	1.165847	0.082408	14.14719	0.0000
FEMALE	-2.754044	0.425705	-6.469368	0.0000
BLACK	-0.431165	1.348249	-0.319796	0.7492
FEMALE*BLACK	0.067320	1.906318	0.035314	0.9718
SOUTH	1.302260	2.114735	0.615803	0.5382
EDUC*SOUTH	-0.191725	0.154240	-1.243036	0.2141
FEMALE*SOUTH	0.911939	0.795976	1.145686	0.2522
BLACK*SOUTH	-1.744432	1.826695	-0.954966	0.3398
FEMALE*BLACK*SOUTH	0.542833	2.511154	0.216169	0.8289
R-squared	0.255731	Mean dependent var		10.21302
Adjusted R-squared	0.248965	S.D. dependent var		6.246641
S.E. of regression	5.413481	Akaike info criterion		6.225611
Sum squared resid	29012.71	Schwarz criterion		6.274689
Log likelihood	-3102.806	Hannan-Quinn criter.		6.244264
F-statistic	37.79605	Durbin-Watson stat		0.499707
Prob(F-statistic)	0.000000			

- 2) Τεσταρουμε την απο κοινου στατιστικη σημαντικοτητα ολων των νεων συντελεστων που προσθεσαμε στην παλινδρομηση

View → Representations

Estimation Command:

```
=====
LS WAGE C EDUC FEMALE BLACK FEMALE*BLACK SOUTH EDUC*SOUTH FEMALE*SOUTH
BLACK*SOUTH FEMALE*BLACK*SOUTH
```

Estimation Equation:

```
=====
WAGE = C(1) + C(2)*EDUC + C(3)*FEMALE + C(4)*BLACK + C(5)*FEMALE*BLACK + C(6)*SOUTH +
C(7)*EDUC*SOUTH + C(8)*FEMALE*SOUTH + C(9)*BLACK*SOUTH + C(10)*FEMALE*BLACK*SOUTH
```

Substituted Coefficients:

```
=====
WAGE = -3.57753582529 + 1.1658472088*EDUC - 2.75404437009*FEMALE - 0.431164989279*BLACK +
0.0673197055694*FEMALE*BLACK + 1.30226006876*SOUTH - 0.191725339676*EDUC*SOUTH +
0.911938577422*FEMALE*SOUTH - 1.74443200463*BLACK*SOUTH +
0.542833007449*FEMALE*BLACK*SOUTH
```

Δηλαδή εξετάζουμε αν τα C(6), C(7), C(8), C(9) και C(10) είναι από κοινού στατιστικά σημαντικά

View → Coefficient tests → Wald και γράφουμε C(6)=0, C(7)=0, C(8)=0, C(9)=0, C(10)=0 → OK

Wald Test:

Equation: Untitled

Test Statistic	Value	df	Probability
F-statistic	2.013212	(5, 990)	0.0744
Chi-square	10.06606	5	0.0734

Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
C(6)	1.302260	2.114735
C(7)	-0.191725	0.154240
C(8)	0.911939	0.795976
C(9)	-1.744432	1.826695
C(10)	0.542833	2.511154

Restrictions are linear in coefficients.

Για το πως προκύπτει η τιμή της F στατιστικής (2.013212) δείτε της σημειώσεις του φροντιστηρίου.

$$F = \frac{[SSR_R - SSR_U] / J_1}{SSR_U / (N - M)} = \frac{[29307.71 - 29012.71] / 5}{29012.71 / (1000 - 10)} = 2.013$$

Ισοδυναμως, το Chow test εφαρμοζεται και ως εξης (δείτε της σημειώσεις του φροντιστηριου.)

- 1) Εκτιμούμε την παλινδρομηση

wage c educ female black female*black

χρησιμοποιώντας **ολες** τις παρατηρησεις. Αυτό το κάναμε πιο πάνω και το άθροισμα των τετραγωνων των καταλοιπων (sum of squared residuals) είναι 29307.71

- 2) Εκτιμούμε την ίδια παλινδρομηση χρησιμοποιώντας μόνο το δείγμα που αναφέρεται στα άτομα που διαμένουν στο South

Quick → Estimate equation → wage c educ female black female*black και γράφουμε στο Sample: if south=1

Dependent Variable: WAGE
Method: Least Squares
Date: 11/12/22 Time: 23:08
Sample: 1 1000 **IF SOUTH=1**
Included observations: 315

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-2.275276	1.554999	-1.463201	0.1444
EDUC	0.974122	0.114294	8.522983	0.0000
FEMALE	-1.842106	0.589593	-3.124369	0.0020
BLACK	-2.175597	1.080434	-2.013633	0.0449
FEMALE*BLACK	0.610153	1.432911	0.425813	0.6705
R-squared	0.224292	Mean dependent var		9.183968
Adjusted R-squared	0.214283	S.D. dependent var		5.353738
S.E. of regression	4.745589	Akaike info criterion		5.968054
Sum squared resid	6981.392	Schwarz criterion		6.027619
Log likelihood	-934.9685	Hannan-Quinn criter.		5.991852
F-statistic	22.40874	Durbin-Watson stat		0.417787
Prob(F-statistic)	0.000000			

Εδώ το άθροισμα των τετραγωνων των καταλοιπων (sum of squared residuals) είναι 6981.39.

- 3) Εκτιμούμε την ίδια παλινδρομηση χρησιμοποιώντας μόνο το δείγμα που αναφέρεται στα άτομα που ΔΕΝ διαμένουν στο South

Quick → Estimate equation → wage c educ female black female*black και γραφουμε στο
 Sample: if south=0

Dependent Variable: WAGE
 Method: Least Squares
 Date: 11/12/22 Time: 23:10
 Sample: 1 1000 IF SOUTH=0
 Included observations: 685

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-3.577536	1.210570	-2.955249	0.0032
EDUC	1.165847	0.086648	13.45492	0.0000
FEMALE	-2.754044	0.447608	-6.152797	0.0000
BLACK	-0.431165	1.417618	-0.304147	0.7611
FEMALE*BLACK	0.067320	2.004401	0.033586	0.9732
R-squared	0.253036	Mean dependent var		10.68623
Adjusted R-squared	0.248643	S.D. dependent var		6.566630
S.E. of regression	5.692012	Akaike info criterion		6.323277
Sum squared resid	22031.32	Schwarz criterion		6.356339
Log likelihood	-2160.722	Hannan-Quinn criter.		6.336070
F-statistic	57.58808	Durbin-Watson stat		0.490233
Prob(F-statistic)	0.000000			

Εδω το αθροισμα των τετραγωνων των καταλοιπων (sum of squared residuals) είναι 22031.32

4)

Κατασκευαζουμε την F στατιστικη

$$F = \frac{[SSR_R - (SSR_{south} + SSR_{nonsouth})]/J_2}{(SSR_{south} + SSR_{nonsouth})/(N-2K)} = \frac{[29307.71 - (6981.392 + 22031.32)]/5}{(6981.392 + 22031.32)/(1000-10)} = 2.013 \text{ (οπως και πριν!)}$$

APA:

$$F = \frac{[SSR_R - SSR_u]/J_1}{SSR_u/(N-M)} = \frac{[SSR_R - (SSR_{south} + SSR_{nonsouth})]/J_2}{(SSR_{south} + SSR_{nonsouth})/(N-2K)}$$

αφου ουσιαστικα $SSR_u = SSR_{south} + SSR_{nonsouth}$,

$$J_1 = J_2, \text{ και } M = 2K$$

