

## Семінар 4. Множинна регресія

В цьому розділі ми розглянемо побудову лінійної множинної регресії, а також моделі, що можуть бути зведені до стандартної лінійної регресії:

1. Аналіз за допомогою лінійної регресії
2. Поліноміальна регресія

### Аналіз за допомогою лінійної регресії

Побудуємо модель залежності витрат на харчування (*FOOD*) від доходу у розпорядженні (*DPI*) та індексу цін на продукти харчування (*PFOOD*)<sup>1</sup>

$$FOOD = \beta_0 + \beta_1 * DPI + \beta_2 * PFOOD + e$$

Після цього необхідно оцінити параметри моделі множинної лінійної регресії, виписати вибірку регресійну функцію, зробити висновки про значущість коефіцієнтів регресії та моделі в цілому, знайти чому дорівнює коефіцієнт детермінації, загальна сума квадратів, пояснена сума квадратів та сума квадратів залишків.

Отже, результати оцінювання даної моделі матимуть наступний вигляд:

Equation: UNTITLED    Workfile: EXPEND::Expend\				
View   Proc   Object   Print   Name   Freeze   Estimate   Forecast   Stats   Resids				
Dependent Variable: FOOD				
Method: Least Squares				
Date: 08/27/13    Time: 12:46				
Sample: 1959 1983				
Included observations: 25				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	55.97158	3.421351	16.35950	0.0000
DPI	0.091004	0.008594	10.58884	0.0000
PFOOD	0.009373	0.032146	0.291577	0.7733
<hr/>				
R-squared	0.977600	Mean dependent var		128.0840
Adjusted R-squared	0.975564	S.D. dependent var		18.79700
S.E. of regression	2.938366	Akaike info criterion		5.105751
Sum squared resid	189.9478	Schwarz criterion		5.252016
Log likelihood	-60.82188	Hannan-Quinn criter.		5.146319
F-statistic	480.0737	Durbin-Watson stat		0.813574
Prob(F-statistic)	0.000000			

<sup>1</sup> Файл expend.wf1

Вибіркова регресійна функція дорівнюватиме:

$$FOOD = 55.97158 + 0.091004 * DPI + 0.009373 * PFOOD$$

### **Перевірка значущості регресії**

Оскільки  $\text{Prob}(F\text{-statistic}) = 0.000000$ , що менше ніж 0,05 та 0,01 (рівень значущості), тоді можна зробити висновок про значущість регресії при будь-якому рівні значущості.

#### **Перевірка значущості коефіцієнтів моделі**

З таблиці результатів оцінювання моделі використовуємо наступні дані:

- $\text{Prob.}(DPI) = 0.0000$ , що менше ніж 0,01 та 0,05 (рівень значущості), тоді можна зробити висновок про значущість змінної  $DPI$  при будь-якому рівні значущості.
- $\text{Prob.}(PFOOD) = 0.7733$ , що більше ніж 0,05 (рівень значущості), тоді можна зробити висновок про незначущість змінної  $PFOOD$  при рівні значущості 0,05.

**Коефіцієнт детермінації** дорівнює 0.977600, що є досить високим рівнем, отже, модель можна вважати вдалою.

**Сума квадратів залишків** дорівнює 189.9478.

### **Поліноміальна регресія**

Розглянемо випадок, коли необхідно побудувати регресійну модель невизначеного вигляду, якщо є припущення, що одна зі змінних  $FOOD$  може залежати від двох змінних  $DPI$  та  $PFOOD$ :

Побудуємо наступні моделі:

$$FOOD = \beta_0 + \beta_1 * DPI^2 + \beta_2 * PFOOD + e$$

$$FOOD = \beta_0 + \beta_1 * DPI + \beta_2 * PFOOD + \beta_3 * PFOOD^2 + e$$

У вікні специфікації рівняння для першої моделі запишемо:

Equation Estimation

Specification Options

Equation specification

Dependent variable followed by list of regressors including ARMA and PDL terms, OR an explicit equation like  $Y=c(1)+c(2)*X$ .

food c dpi^2 pfood

Estimation settings

Method: LS - Least Squares (NLS and ARMA)

Sample: 1959 1983

OK Скасувати

Для другої моделі специфікація матиме наступний вигляд:

Equation Estimation

Specification Options

Equation specification

Dependent variable followed by list of regressors including ARMA and PDL terms, OR an explicit equation like  $Y=c(1)+c(2)*X$ .

food c dpi pfood^2

Estimation settings

Method: LS - Least Squares (NLS and ARMA)

Sample: 1959 1983

OK Скасувати

Отримаємо наступні результати оцінювання першої моделі та другої моделі:

Equation: UNTITLED Workfile: EXPEND::Expend\

ViewProcObjectPrintNameFreezeEstimateForecastStatsResids

Dependent Variable: FOOD  
Method: Least Squares  
Date: 08/27/13 Time: 12:55  
Sample: 1959 1983  
Included observations: 25

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	91.32095	1.488093	61.36775	0.0000
DPI^2	7.91E-05	7.42E-06	10.66702	0.0000
PFOOD	-0.119817	0.043464	-2.756683	0.0115

R-squared0.977874Mean dependent var128.0840

Adjusted R-squared0.975863S.D. dependent var18.79700

S.E. of regression2.920331Akaike info criterion5.093437

Sum squared resid187.6233Schwarz criterion5.239703

Log likelihood-60.66797Hannan-Quinn criter.5.134005

F-statistic486.1578Durbin-Watson stat0.883084

Prob(F-statistic)0.000000

Equation: UNTITLED Workfile: EXPEND::Expend\

ViewProcObjectPrintNameFreezeEstimateForecastStatsResids

Dependent Variable: FOOD  
Method: Least Squares  
Date: 08/27/13 Time: 12:54  
Sample: 1959 1983  
Included observations: 25

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	57.74610	3.858171	14.96722	0.0000
DPI	0.088642	0.006475	13.68999	0.0000
PFOOD^2	6.96E-05	8.50E-05	0.818693	0.4217

R-squared	0.978178	Mean dependent var	128.0840
Adjusted R-squared	0.976195	S.D. dependent var	18.79700
S.E. of regression	2.900190	Akaike info criterion	5.079596
Sum squared resid	185.0443	Schwarz criterion	5.225861
Log likelihood	-60.49495	Hannan-Quinn criter.	5.120164
F-statistic	493.0869	Durbin-Watson stat	0.836724
Prob(F-statistic)	0.000000		

Таким чином, найкращі результати в цьому випадку продемонструвала нелінійна модель.

### Самостійна робота

1. Зробіть прогноз за наведеною моделлю, якщо в наступному році очікується:  $DPI=1125$ ,  $PFOOD=231$ . Визначте зміну споживання продуктів харчування при зміні кожного з факторів на 1.

2. Побудуйте регресію видатків на харчування по часу (змінна *time*). Наведіть інтерпретацію коефіцієнтам регресії. Розрахуйте видатки на харчування в 1984 та 1985 роках, якщо  $DPI_{1984}=1132$ ,  $DPI_{1985}=1151$ ,  $PFOOD_{1984}=162$ ,  $PFOOD_{1985}=166$ .

3. За допомогою даних файлу M2.xls побудуйте найкращу нелінійну регресію.

4. На основі даних файлу *phillips\_curve.wfl* розрахуйте темп інфляції (*inf*), побудуйте нелінійну регресію темпу інфляції по безробіттю (*unr*).

Розрахуйте величину, обернену до рівня безробіття (*runr*), побудуйте нелінійну регресію темпу інфляції по величині, оберненій до рівня безробіття.

Побудуйте графіки для обох регресій.