

Logit та Probit моделі у EViews

Доц. Андрій Ставицький

Файл:Binary.WF1

- ▶ GRADE показує наявність покращення навчання
- ▶ PSI наявність нової технології викладання
- ▶ GPA та TUCE – дві підсумкові оцінки знань студентів

EViews - [Group: UNTITLED Workfile: BINARY::undated\]					
File Edit Object View Proc Quick Options Add-ins Win					
View	Proc	Object	Print	Name	Freeze
Default Sort Transpo					
obs	GRADE	GPA	TUCE	PSI	
1	0.0000	2.6600	20.000	0.0000	
2	0.0000	2.8900	22.000	0.0000	
3	0.0000	3.2800	24.000	0.0000	
4	0.0000	2.9200	12.000	0.0000	
5	1.0000	4.0000	21.000	0.0000	
6	0.0000	2.8600	17.000	0.0000	
7	0.0000	2.7600	17.000	0.0000	
8	0.0000	2.8700	21.000	0.0000	
9	0.0000	3.0300	25.000	0.0000	
10	1.0000	3.9200	29.000	0.0000	
11	0.0000	2.6300	20.000	0.0000	
12	0.0000	3.3200	23.000	0.0000	
13	0.0000	3.5700	23.000	0.0000	
14	1.0000	3.2600	25.000	0.0000	
15	0.0000	3.5300	26.000	0.0000	
16	0.0000	2.7400	19.000	0.0000	
17	0.0000	2.7500	25.000	0.0000	
18	0.0000	2.8300	19.000	0.0000	
19	0.0000	3.1200	23.000	1.0000	
20	1.0000	3.1600	25.000	1.0000	
21	0.0000	2.0600	22.000	1.0000	
22	1.0000	3.6200	28.000	1.0000	
23	0.0000	2.8900	14.000	1.0000	
24	0.0000	3.5100	26.000	1.0000	
25	1.0000	3.5400	24.000	1.0000	
26	1.0000	2.8300	27.000	1.0000	
27	1.0000	3.3900	17.000	1.0000	
28	0.0000	2.6700	24.000	1.0000	
29	1.0000	3.6500	21.000	1.0000	
30	1.0000	4.0000	23.000	1.0000	
31	0.0000	3.1000	21.000	1.0000	
32	1.0000	2.3900	19.000	1.0000	

Специфікація моделі

Equation Estimation

Specification Options

Equation specification

Binary dependent variable followed by list of regressors, OR
an explicit equation like $Y=c(1)+c(2)*X$.

grade c gpa tuce psi

Binary estimation method: ☒ Probit ☐ Logit ☐ Extreme value

Estimation settings

Method: BINARY - Binary Choice (Logit, Probit, Extreme Value)

Sample: 1 32

OK Скасувати

Опції

The image shows a software window titled "Equation Estimation" with a close button (X) in the top right corner. It has two tabs: "Specification" and "Options", with "Options" currently selected. The window is divided into four main sections:

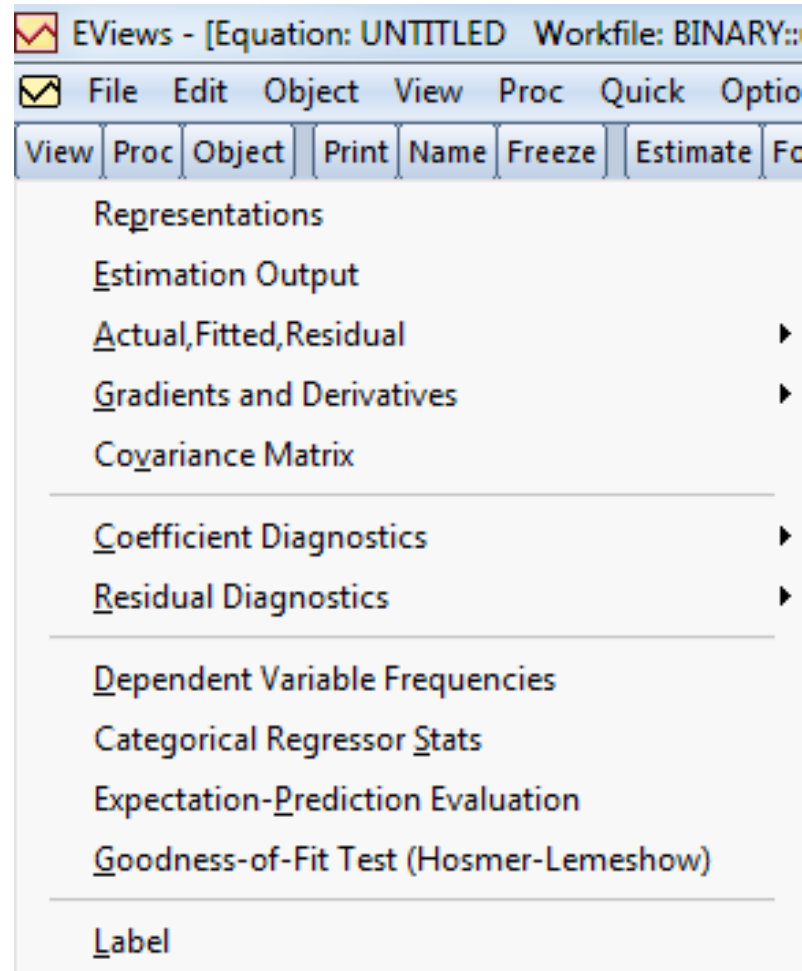
- Covariance:** Contains three radio button options: "Robust Covariances" (unchecked), "Huber/White" (selected), and "GLM" (unchecked).
- Optimization algorithm:** Contains three radio button options: "Quadratic Hill Climbing" (selected), "Newton-Raphson" (unchecked), and "Berndt-Hall-Hall-Hausman" (unchecked).
- Iteration control:** Includes two input fields: "Max Iterations:" with the value "500" and "Convergence:" with the value "0.0001". Below these is a label "Starting coefficient values:" followed by a dropdown menu showing "EViews Supplied". At the bottom of this section is a checkbox for "Display settings" which is unchecked.
- Derivatives (for index):** Includes a label "Select method to favor:" followed by two radio button options: "Accuracy" (selected) and "Speed" (unchecked). Below this is a checkbox for "Use numeric only" which is unchecked.

At the bottom right of the window are two buttons: "OK" and "Скасувати" (Cancel).

Оцінка моделі

EViews - [Equation: UNTITLED Workfile: BINARY::undated\]					
File	Edit	Object	View	Proc	Quick Options Add-ins Window Help
View	Proc	Object	Print	Name	Freeze Estimate Forecast Stats Resids
Dependent Variable: GRADE					
Method: ML - Binary Probit (Quadratic hill climbing)					
Date: 03/24/13 Time: 17:42					
Sample: 1 32					
Included observations: 32					
Convergence achieved after 5 iterations					
Covariance matrix computed using second derivatives					
Variable	Coefficient	Std. Error	z-Statistic	Prob.	
C	-7.452320	2.542472	-2.931131	0.0034	
GPA	1.625810	0.693882	2.343063	0.0191	
TUCE	0.051729	0.083890	0.616626	0.5375	
PSI	1.426332	0.595038	2.397045	0.0165	
McFadden R-squared	0.377478	Mean dependent var		0.343750	
S.D. dependent var	0.482559	S.E. of regression		0.386128	
Akaike info criterion	1.051175	Sum squared resid		4.174660	
Schwarz criterion	1.234392	Log likelihood		-12.81880	
Hannan-Quinn criter.	1.111907	Deviance		25.63761	
Restr. deviance	41.18346	Restr. log likelihood		-20.59173	
LR statistic	15.54585	Avg. log likelihood		-0.400588	
Prob(LR statistic)	0.001405				
Obs with Dep=0	21	Total obs		32	
Obs with Dep=1	11				

Меню для аналізу бінарних моделей



Dependent Variable Frequencies

- ▶ Аналіз частот та кумулятивних частот залежної змінної.

EViews - [Equation: UNTITLED Workfile: BINARY::undated\]

File Edit Object View Proc Quick Options Add-ins Window

View Proc Object Print Name Freeze Estimate Forecast Stats Resids

Dependent Variable Frequencies
Equation: UNTITLED
Date: 03/24/13 Time: 18:10

Dep. Value	Count	Percent	Cumulative Count	Cumulative Percent
0	21	65.63	21	65.63
1	11	34.38	32	100.00

Categorical Regressor Stats

- Описова статистика для кожного з регресорів.

EViews - [Equation: UNTITLED Workfile: BINARY::undated\]

File Edit Object View Proc Quick Options Add-ins Window

View Proc Object Print Name Freeze Estimate Forecast Stats Resids

Categorical Descriptive Statistics for Explanatory Variables
Equation: UNTITLED
Date: 03/24/13 Time: 18:08

Variable	Dep=0	Mean Dep=1	All
C	1.000000	1.000000	1.000000
GPA	2.951905	3.432727	3.117188
TUCE	21.09524	23.54545	21.93750
PSI	0.285714	0.727273	0.437500

Variable	Dep=0	Standard Deviation Dep=1	All
C	0.000000	0.000000	0.000000
GPA	0.357220	0.503132	0.466713
TUCE	3.780275	3.777926	3.901509
PSI	0.462910	0.467099	0.504016

Variable	Dep=0	Dep=1	All
Observations	21	11	32

Expectation-Prediction (Classification) Table

- ▶ Аналіз вірних та помилкових класифікацій залежної змінної.

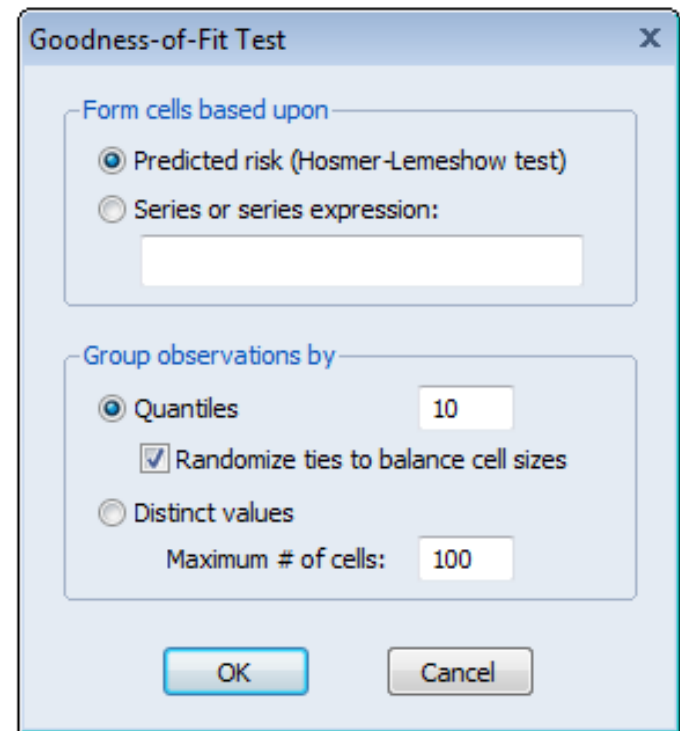
EViews - [Equation: UNTITLED Workfile: BINARY::undated\]						
<input checked="" type="checkbox"/> File <input type="checkbox"/> Edit <input type="checkbox"/> Object <input type="checkbox"/> View <input type="checkbox"/> Proc <input type="checkbox"/> Quick <input type="checkbox"/> Options <input type="checkbox"/> Add-ins <input type="checkbox"/> Window <input type="checkbox"/> Help						
<input type="checkbox"/> View	<input type="checkbox"/> Proc	<input type="checkbox"/> Object	<input type="checkbox"/> Print	<input type="checkbox"/> Name	<input type="checkbox"/> Freeze	<input type="checkbox"/> Estimate <input type="checkbox"/> Forecast <input type="checkbox"/> Stats <input type="checkbox"/> Resids
Expectation-Prediction Evaluation for Binary Specification Equation: UNTITLED Date: 03/24/13 Time: 18:14 Success cutoff: C = 0.5						
	Estimated Equation			Constant Probability		
	Dep=0	Dep=1	Total	Dep=0	Dep=1	Total
P(Dep=1)≤C	18	3	21	21	11	32
P(Dep=1)>C	3	8	11	0	0	0
Total	21	11	32	21	11	32
Correct	18	8	26	21	0	21
% Correct	85.71	72.73	81.25	100.00	0.00	65.63
% Incorrect	14.29	27.27	18.75	0.00	100.00	34.38
Total Gain*	-14.29	72.73	15.63			
Percent Gain**	NA	72.73	45.45			
	Estimated Equation			Constant Probability		
	Dep=0	Dep=1	Total	Dep=0	Dep=1	Total
E(# of Dep=0)	16.89	4.14	21.03	13.78	7.22	21.00
E(# of Dep=1)	4.11	6.86	10.97	7.22	3.78	11.00
Total	21.00	11.00	32.00	21.00	11.00	32.00
Correct	16.89	6.86	23.74	13.78	3.78	17.56
% Correct	80.42	62.32	74.20	65.63	34.38	54.88
% Incorrect	19.58	37.68	25.80	34.38	65.63	45.12
Total Gain*	14.80	27.95	19.32			
Percent Gain**	43.05	42.59	42.82			

*Change in "% Correct" from default (constant probability) specification

**Percent of incorrect (default) prediction corrected by equation

Goodness-of-Fit Tests – 1

- ▶ Вікно дозволяє провести тест Пірсона χ^2 якості моделі. Ідея тесту – порівняння прогнозних та реальних значень за групами. Якщо різниці великі, то модель відкидається.



Goodness-of-Fit Tests – 2

EViews - [Equation: UNTITLED Workfile: BINARY::undated\]											
File Edit Object View Proc Quick Options Add-ins Window Help											
View	Proc	Object	Print	Name	Freeze	Estimate	Forecast	Stats	Resids		
Goodness-of-Fit Evaluation for Binary Specification											
Andrews and Hosmer-Lemeshow Tests											
Equation: UNTITLED											
Date: 03/24/13 Time: 18:38											
Grouping based upon predicted risk (randomize ties)											
	Quantile of Risk		Dep=0		Dep=1		Total	H-L			
	Low	High	Actual	Expect	Actual	Expect	Obs	Value			
1	0.0161	0.0185	3	2.94722	0	0.05278	3	0.05372			
2	0.0186	0.0272	3	2.93223	0	0.06777	3	0.06934			
3	0.0309	0.0457	3	2.87888	0	0.12112	3	0.12621			
4	0.0531	0.1088	3	2.77618	0	0.22382	3	0.24186			
5	0.1235	0.1952	2	3.29779	2	0.70221	4	2.90924			
6	0.2732	0.3287	3	2.07481	0	0.92519	3	1.33775			
7	0.3563	0.5400	2	1.61497	1	1.38503	3	0.19883			
8	0.5546	0.6424	1	1.20962	2	1.79038	3	0.06087			
9	0.6572	0.8342	0	0.84550	3	2.15450	3	1.17730			
10	0.8400	0.9522	1	0.45575	3	3.54425	4	0.73351			
Total			21	21.0330	11	10.9670	32	6.90863			
H-L Statistic			6.9086		Prob. Chi-Sq(8)		0.5465				
Andrews Statistic			20.6045		Prob. Chi-Sq(10)		0.0240				

p -значення для HL тест велике, а для тесту Ендрюса – мале, що дає різні результати.

Прогноз

- Proc/Forecast (Fitted Probability/Index)...., розраховує ймовірність або індекс для залежної змінної.

The screenshot shows the 'Forecast' dialog box with the following settings:

- Forecast equation:** UNTITLED
- Series to forecast:** ☒ Probability, ☐ Index - where Prob = 1-F(-Index)
- Series names:**
 - Forecast name: gradef
 - S.E. (optional):
 - GARCH(optional):
- Method:**
 - Static forecast (no dynamics in equation)
 - ☐ Structural (ignore ARMA)
 - ☐ Coef uncertainty in S.E. calc
- Forecast sample:** 1 32
- Output:**
 - ☒ Forecast graph
 - ☒ Forecast evaluation
- ☒ Insert actuals for out-of-sample observations
- Buttons:** OK, Cancel



EViews - [Equation: UNTITLED Workfile: BINARY::undated\]

File

Edit

Object

View

Proc

Quick

Options

Add-ins

Window

Help

View

Proc

Object

Print

Name

Freeze

Estimate

Forecast

Stats

Resids

Dependent Variable: GRADE

Method: ML - Binary Probit (Quadratic hill climbing)

Date: 03/24/13 Time: 17:42

Sample: 1 32

Included observations: 32

Convergence achieved after 5 iterations

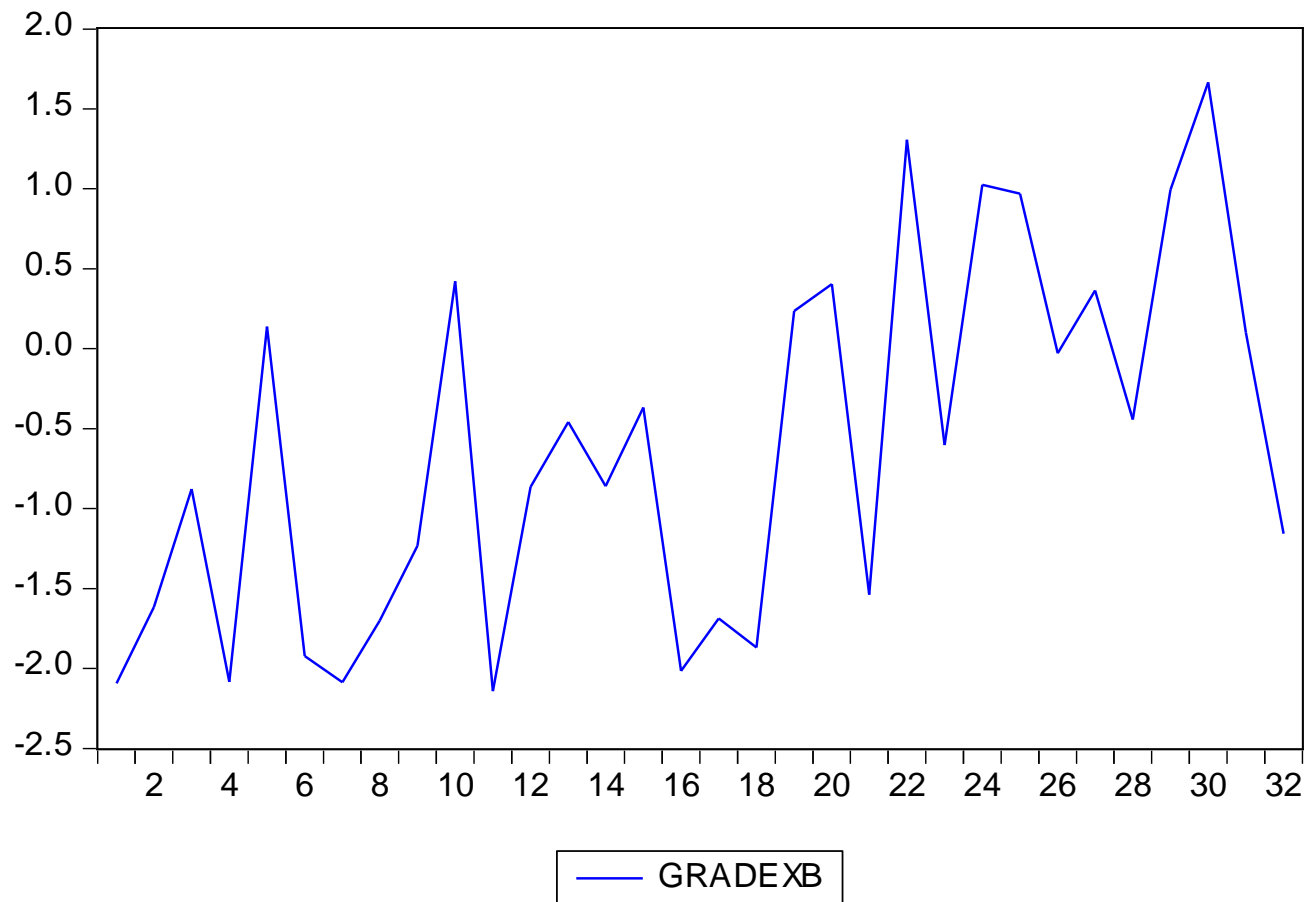
Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-7.452320	2.542472	-2.931131	0.0034
GPA	1.625810	0.693882	2.343063	0.0191
TUCE	0.051729	0.083890	0.616626	0.5375
PSI	1.426332	0.595038	2.397045	0.0165

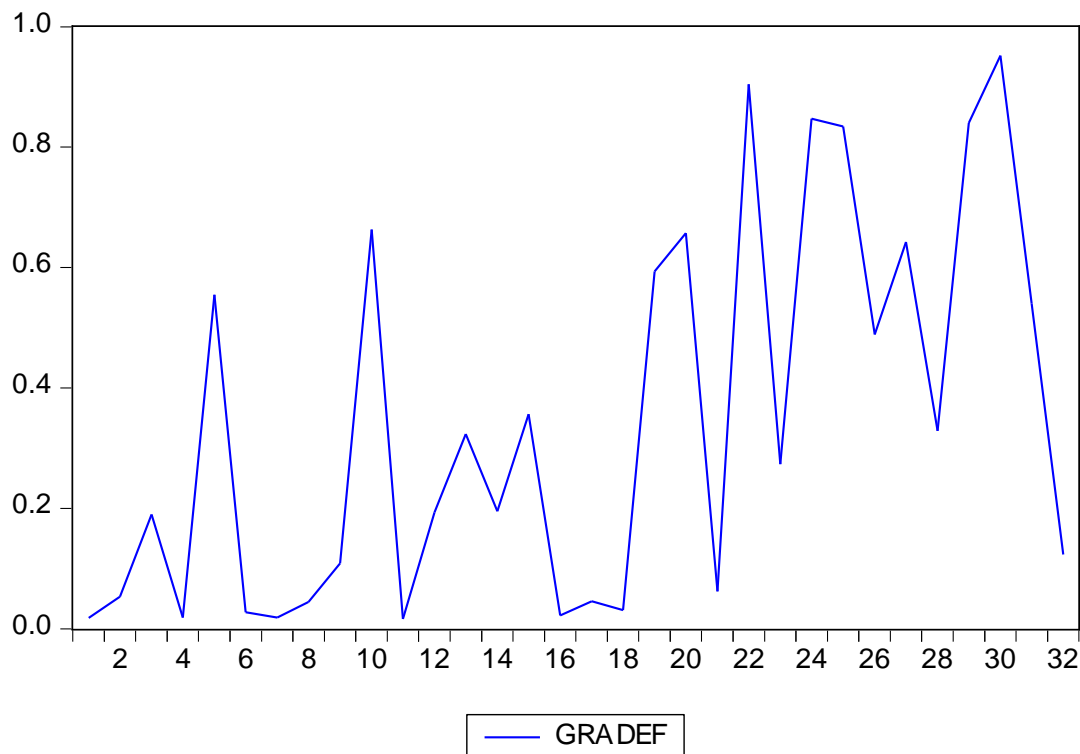
McFadden R-squared	0.377478	Mean dependent var	0.343750
S.D. dependent var	0.482559	S.E. of regression	0.386128
Akaike info criterion	1.051175	Sum squared resid	4.174660
Schwarz criterion	1.234392	Log likelihood	-12.81880
Hannan-Quinn criter.	1.111907	Deviance	25.63761
Restr. deviance	41.18346	Restr. log likelihood	-20.59173
LR statistic	15.54585	Avg. log likelihood	-0.400588
Prob(LR statistic)	0.001405		

Obs with Dep=0	21	Total obs	32
Obs with Dep=1	11		

Прогноз: індекс



Прогноз: ймовірність



Forecast: GRADEF
Actual: GRADE
Forecast sample: 1 32
Included observations: 32
Root Mean Squared Error 0.361190
Mean Absolute Error 0.257987
Mean Abs. Percent Error 12.95082
Theil Inequality Coefficient 0.343928
Bias Proportion 0.000008
Variance Proportion 0.201997
Covariance Proportion 0.797995

Залежність між індексом та ймовірністю

- ▶ Для першого спостереження ($t=1$):

$$\log[P_1] = -2.093086,$$

$$P_1 = \exp(-2.093086) = 0.123306.$$

Logit

EViews - [Equation: EQ01 Workfile: BINARY::undated\]				
	File	Edit	Object	View
	Proc	Quick	Options	Add-ins
	Window	H		
View	Proc	Object	Print	Name
	Freeze	Estimate	Forecast	Stats
	Resids			

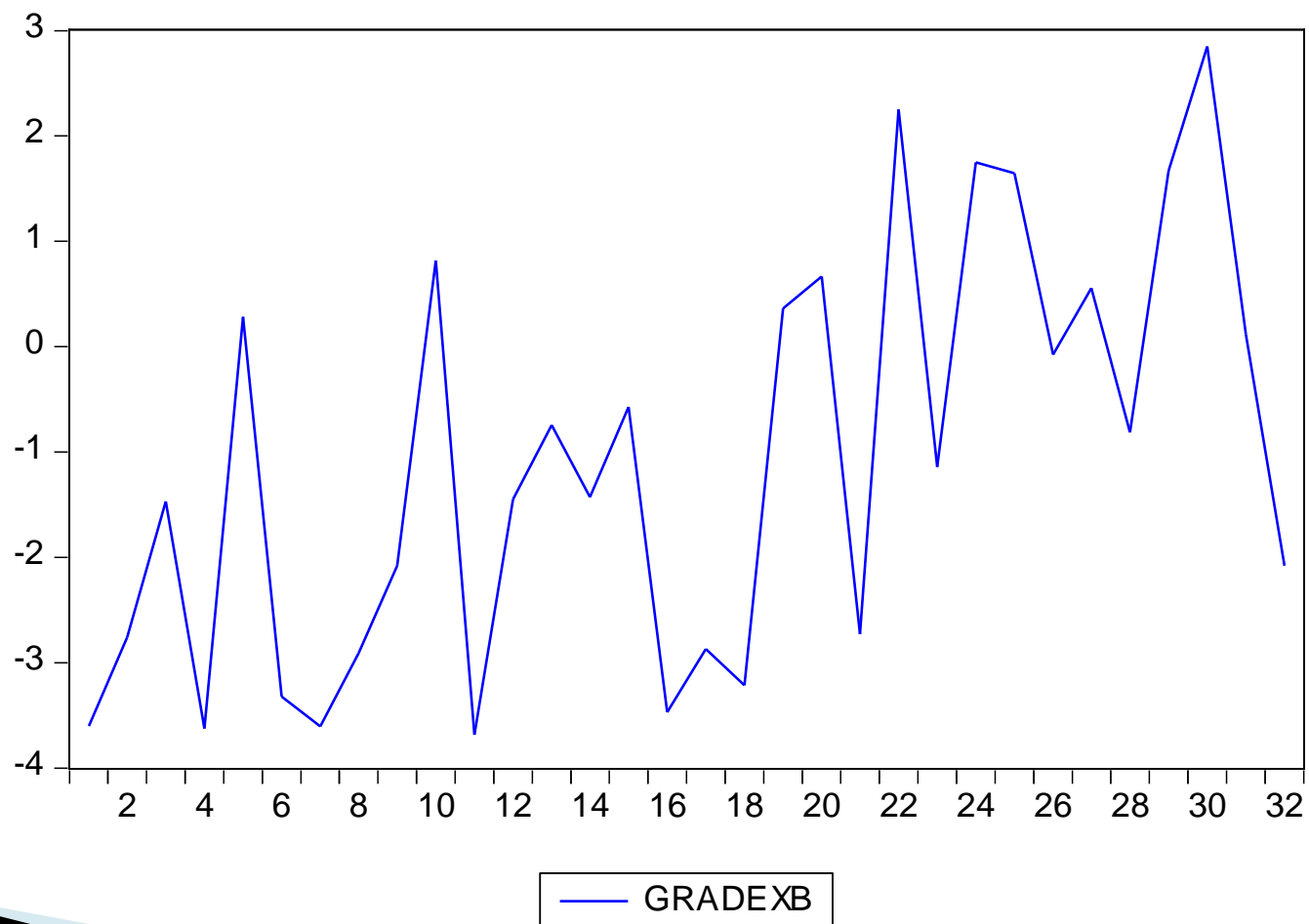
Dependent Variable: GRADE
Method: ML - Binary Logit (Quadratic hill climbing)
Date: 03/24/13 Time: 19:14
Sample: 1 32
Included observations: 32
Convergence achieved after 5 iterations
Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-13.02135	4.931324	-2.640537	0.0083
GPA	2.826113	1.262941	2.237723	0.0252
TUCE	0.095158	0.141554	0.672235	0.5014
PSI	2.378688	1.064564	2.234424	0.0255

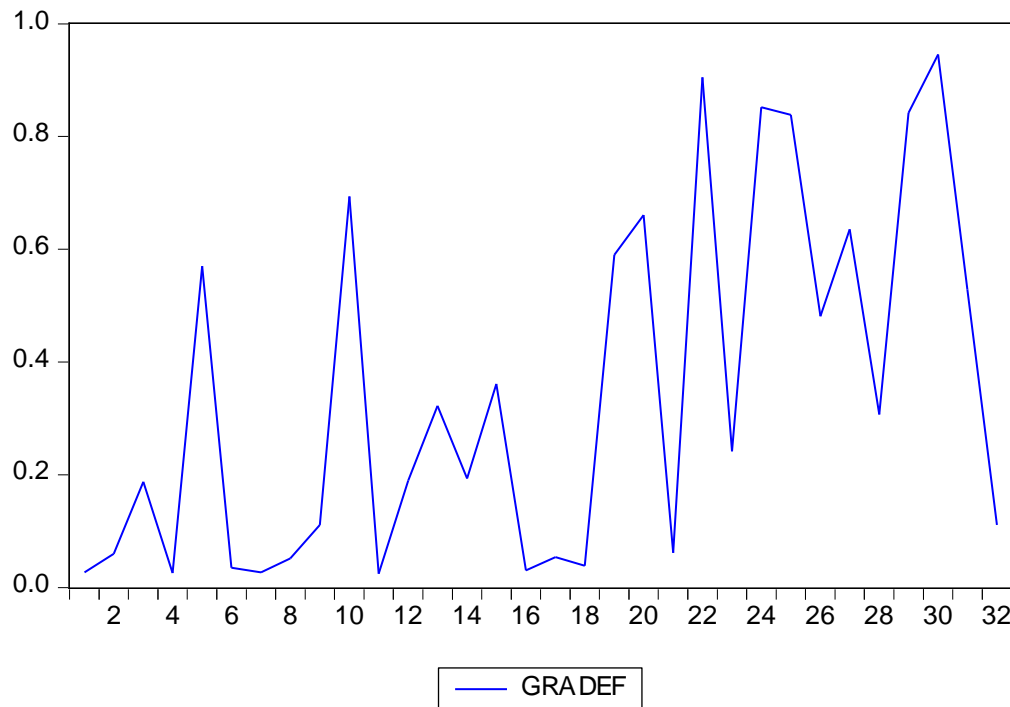
McFadden R-squared	0.374038	Mean dependent var	0.343750
S.D. dependent var	0.482559	S.E. of regression	0.384716
Akaike info criterion	1.055602	Sum squared resid	4.144171
Schwarz criterion	1.238819	Log likelihood	-12.88963
Hannan-Quinn criter.	1.116333	Deviance	25.77927
Restr. deviance	41.18346	Restr. log likelihood	-20.59173
LR statistic	15.40419	Avg. log likelihood	-0.402801
Prob(LR statistic)	0.001502		

Obs with Dep=0	21	Total obs	32
Obs with Dep=1	11		

Прогноз: індекс



Прогноз: ймовірність



Forecast: GRADEF
Actual: GRADE
Forecast sample: 1 32
Included observations: 32
Root Mean Squared Error 0.359869
Mean Absolute Error 0.257802
Mean Abs. Percent Error 12.89010
Theil Inequality Coefficient 0.342578
Bias Proportion 0.000000
Variance Proportion 0.205275
Covariance Proportion 0.794725

Залежність між індексом та ймовірністю

- ▶ Для першого спостереження ($t=1$)

$$\log[P_i/(1-P_i)] = -3.600734,$$

$$P_i/(1-P_i) = \exp(-3.600734) = 0.027304,$$

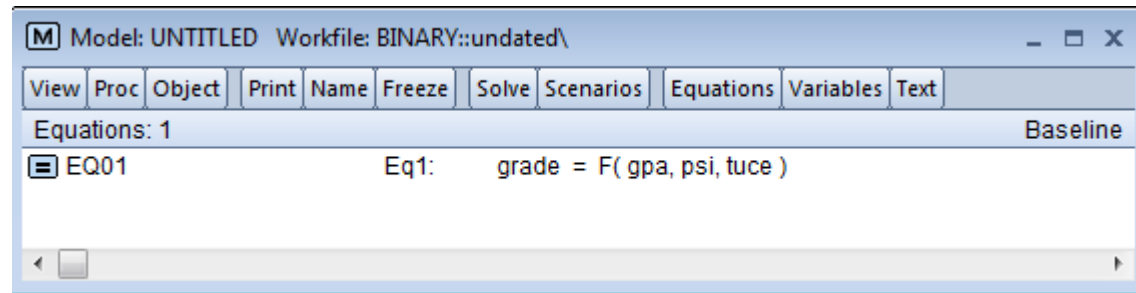
$$P_i = 0.027304/(1 + 0.027304) = 0.026578.$$

Застосування: чи впливає PSI?

- ▶ Потрібно побудувати графік ймовірностей GRADE як функції від GPA, де PSI приймає два значення (1 – нова методика використовується, 0 – ні).

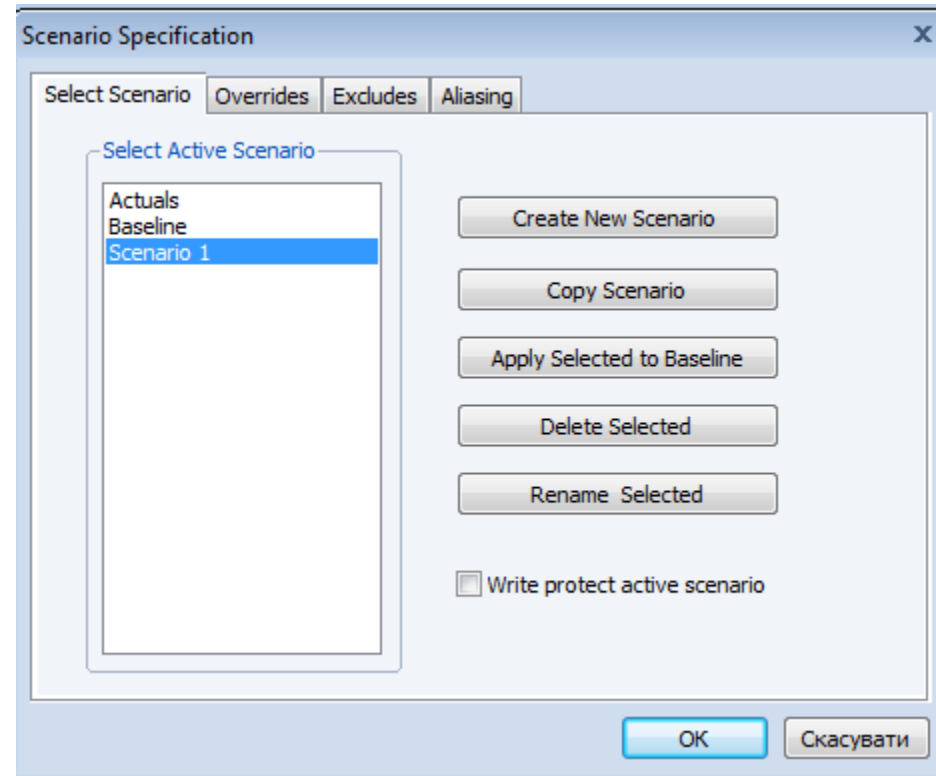
Розв'язок – 1

- ▶ Створюємо модель на основі отриманого рівняння **Proc/Make Model** з меню моделі.



Розв'язок – 2

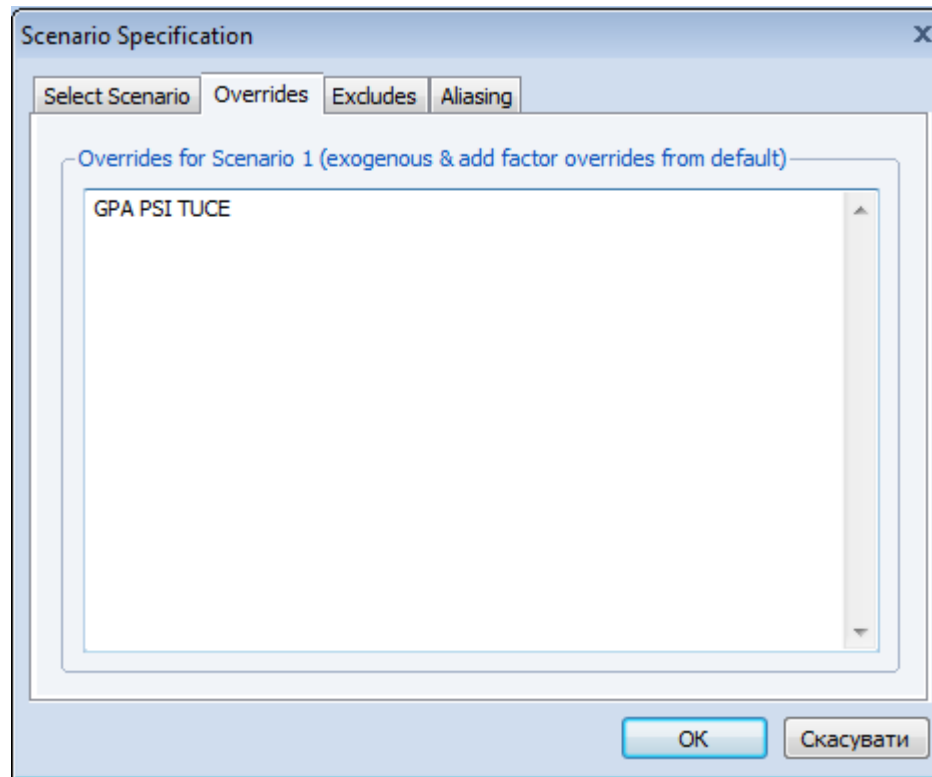
- Визначаємо сценарії моделі (Scenarios–Scenario Specification – Scenario 1).



Розв'язок – 3

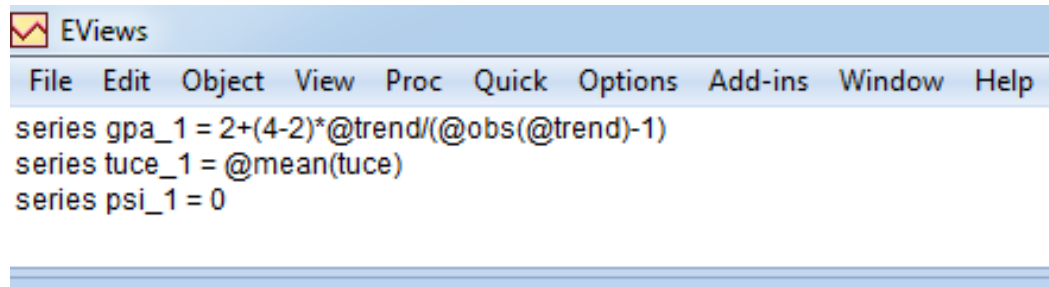
- ▶ Діалогове вікно **Scenario Specification** дозволяє визначити набір припущень для моделі.
- ▶ Клацніть по закладці **Overrides** та введіть "GPA PSI TUCE". Це дозволить EViews використовувати такі зміні GPA_1, PSI_1, and TUCE_1 замість оригінальних GPA, PSI та TUCE при розв'язку моделі відносно GRADE у сценарії 1.

Розв'язок – 4



Розв'язок – 5

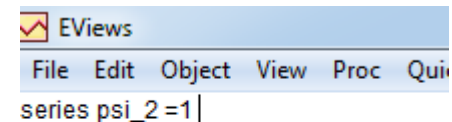
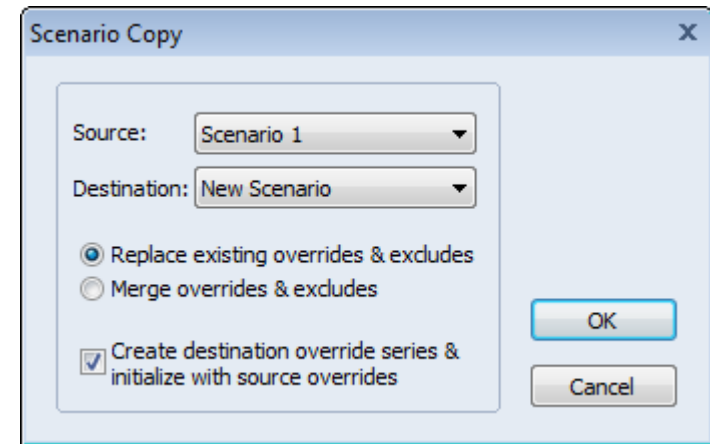
- ▶ Тепер треба створити зміні GPA_1, PSI_1 та TUCE_1 у робочому файлі. Потрібно використовувати ці часові ряди для обчислення ймовірностей GRADE для різних значень GPA від 2 до 4 з однаковим кроком, причому TUCE буде дорівнювати середньому рівню для всіх спостережень. Зміна PSI буде дорівнювати 0:



```
EViews
File Edit Object View Proc Quick Options Add-ins Window Help
series gpa_1 = 2+(4-2)*@trend/(@obs(@trend)-1)
series tuce_1 = @mean(tuce)
series psi_1 = 0
```

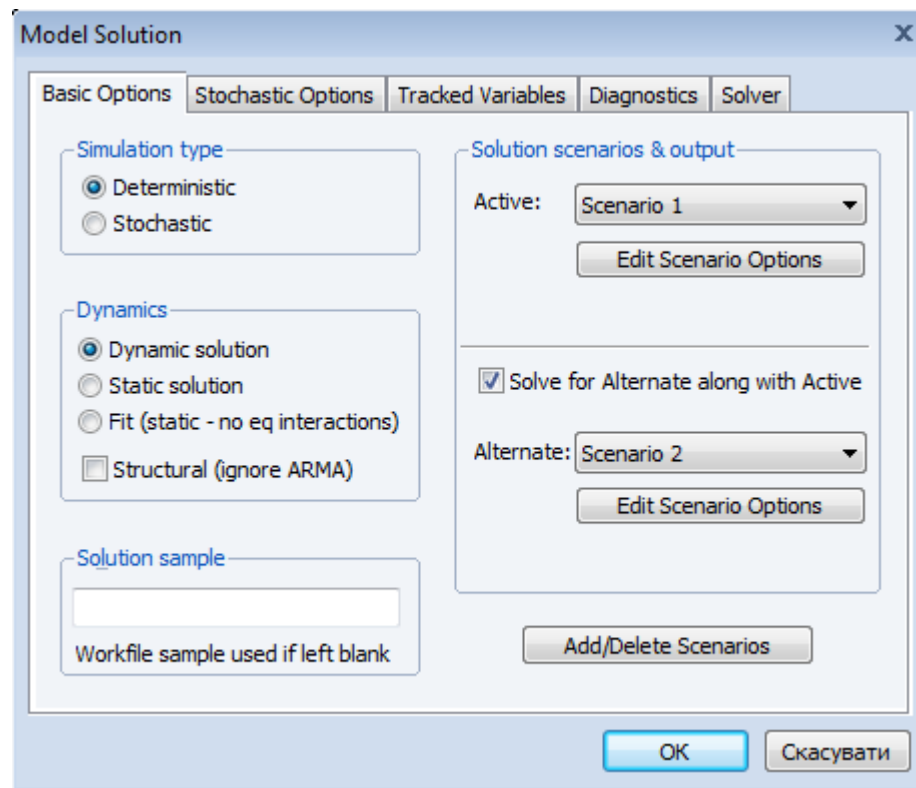
Розв'язок – 6

- ▶ Після створення даних необхідно скопіювати їх до нового сценарію, де зміна $PSI=1$.
- ▶ Вибираємо закладинку **Select Scenario, Copy Scenario**, потім – **Scenario 1** як Джерело, та **New Scenario** як Призначення.
- ▶ Тепер створюємо $PSI_2=1$



Розв'язок – 7

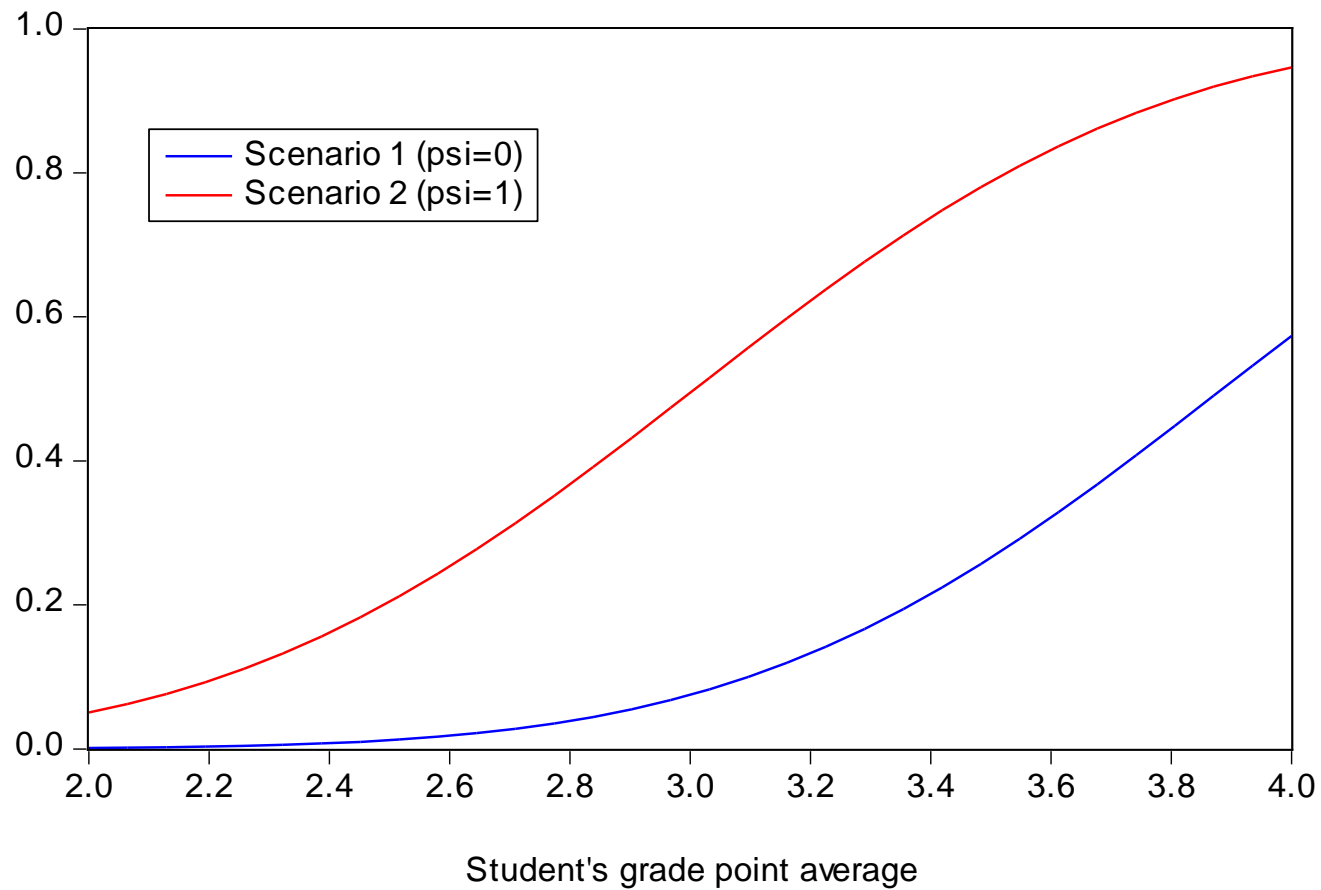
- Розв'язуємо модель для двох сценаріїв



Розв'язок – 8

- ▶ Показуємо результати у одній групі
(Object/New Object.../Group) for series:
gpa_1 grade_1 grade_2

Порівняння сценаріїв



Питання?



Самостійна робота