

ЕКОНОМІКА

DOI 10.32782/2786-8559/2023-2-1
УДК 351.851

Волошин Володимир Степанович

кандидат економічних наук, доцент,
доцент кафедри комп'ютерних технологій та економічної кібернетики,
Національний університет водного господарства та природокористування
ORCID ID: <https://orcid.org/0000-0001-8108-0126>

МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ ЧИСЕЛЬНОСТІ ЗАКЛАДІВ ВИЩОЇ ОСВІТИ УКРАЇНИ

Встановлено, що вітчизняні заклади вищої освіти поступово інтегруються у світовий науковий простір, працюють в умовах зростаючої конкуренції. Суттєвий вплив на їх кількість справляє потенційна кількість майбутніх здобувачів вищої освіти – бакалаврів, магістрів, аспірантів. З метою дослідження закладів вищої освіти в якості залежної змінної моделі у було розглянуто показник їх кількості в Україні. У якості пояснюючих змінних, виходячи з економічних міркувань та міркувань доступності інформації, було відібрано показники: абітурієнти, уперше прийняті до закладів вищої освіти; випущено бакалаврів та магістрів; кількість аспірантів. Використано шкалу Чеддока для визначення найбільш взаємозалежних показників. Під час тестування масиву пояснюючих змінних на мультиколінеарність та процедури покрокового відбору пояснюючих змінних було виявлено неповну мультиколінеарність. Розраховано коефіцієнт кореляції, що має значення 0,719. Тобто можна стверджувати що на 71,9% можливо пояснити залежну змінну «кількість ЗВО» незалежною змінною як x -особи, уперше прийняті до ЗВО, а на 28,1% іншими випадковими факторами. Під час виконання специфікації економетричної було розглянуто модель багатофакторної регресії. Було проведено теоретичне та статистичне обґрунтування економетричної моделі, проведено тестування моделі на мультиколінеарність, функціональної форми економетричної моделі, відбір суттєвих пояснюючих змінних до економетричної моделі та обрано за основу модель множинної регресії. Проведено графічний аналіз залишків. З ймовірністю 90% можна стверджувати, що при збільшенні кількості абітурієнтів уперше прийнятих до ЗВО на 1000 осіб мінімальне збільшення кількості ЗВО буде 0,401 одиниць, а максимальне 0,61 одиниць. При $R=95\%$ можна стверджувати, що при збільшенні кількості абітурієнтів уперше прийнятих до ЗВО на 1000 осіб мінімальне збільшення кількості ЗВО буде 0,39 одиниць, а максимальне 0,63 одиниць. З ймовірністю 99% можна стверджувати, що при збільшенні кількості абітурієнтів уперше прийнятих до ЗВО на 1000 осіб мінімальне збільшення кількості ЗВО буде 0,34 одиниць, а максимальне 0,68 одиниць.

Ключові слова: заклад вищої освіти, моделювання, статистика, інформаційні технології, кореляція, автокореляція, прогнозування.

Volodymyr Voloshyn

The National University of Water and Environmental Engineering

MATHEMATICAL MODELING OF THE NUMBER OF HIGHER EDUCATION INSTITUTIONS IN UKRAINE

In order to study higher education institutions, the number of higher education institutions in Ukraine was considered as the dependent variable of the model. The following indicators were selected as explanatory variables, based on economic and information accessibility considerations: applicants admitted to higher education institutions for the first time; graduates; and the number of postgraduate students. The Chaddock scale was used to identify the most interdependent indicators. When testing the array of explanatory variables for multicollinearity and the procedure of stepwise selection of explanatory variables, incomplete multicollinearity was found. The correlation coefficient was calculated and was 0,719. That is, it can be argued that 71,9% of the dependent variable «number of higher education institutions» can be explained by the independent variable as x -persons admitted to higher education institutions for the first time, and 28,1% by other random

factors. When performing the econometric specification, a multivariate regression model was considered. The theoretical and statistical justification of the econometric model was carried out, the model was tested for multicollinearity, the functional form of the econometric model, the selection of significant explanatory variables for the econometric model, and the multiple regression model was chosen as the basis. A graphical analysis of the residuals is carried out. With a 90% probability, it can be stated that with an increase in the number of applicants admitted to HEIs for the first time by 1000 people, the minimum increase in the number of HEIs will be 0,401 units, and the maximum 0,61 units. At $P=95\%$, it can be stated that with an increase in the number of applicants admitted to the HEI for the first time by 1000 people, the minimum increase in the number of HEIs will be 0,39 units, and the maximum will be 0,63 units. With a probability of 99%, it can be stated that with an increase in the number of applicants for the first time admitted to the HEI by 1000 people, the minimum increase in the number of HEIs will be 0,34 units, and the maximum 0,68 units.

Keywords: higher education institution, modeling, statistics, information technology, correlation, autocorrelation, forecasting.

Вступ. Конкуренція серед освітніх закладів є нормальним станом ринкових механізмів, який є основою діяльності сучасних університетів. Якщо на Заході конкурентні відносини у сфері освіти досить добре розвинені, то їх формування в Україні лише відбувається. При цьому конкуренція виконує фундаментальні функції, які сприяють стимулюванню якості освіти в цілому.

Заклади вищої освіти (ЗВО) України поступово підвищують свій авторитет у світовому науково-освітньому просторі. Основними показниками успішності ЗВО є глобальні рейтинги. Досить важливо, при цьому, математичне моделювання кількості ЗВО України.

Матеріали та методи. У процесі написання статті застосовано методи, які використовуються як на емпіричному, так і на теоретичному рівнях дослідження. Також використано спостереження, що дає змогу дослідникові не вдаватися до суцільного дослідження, отримати узагальнюючі дані, які дозволяють правомірно відобразити характеристики всієї сукупності предметів або явищ.

Результати. З метою дослідження закладів вищої освіти в якості залежної змінної моделі у було розглянуто показник:

y – кількість ЗВО в Україні.

За результатами провізорного кореляційного аналізу та використовуючи методи оцінювання моделей на основі кореляційно регресійного

аналізу було прийнято у якості залежної змінної обрати змінну «кількість ЗВО» за період з 1990 по 2019 рік, а у якості пояснюючих змінних, виходячи з економічних міркувань та міркувань доступності інформації, було відібрано показники:

x_1 – абітурієнти, уперше прийняті до ЗВО;

x_2 – випущено бакалаврів та магістрів;

x_3 – кількість аспірантів.

На етапі специфікації за результатами кореляційного аналізу статистичних показників по Україні (вибірка розміром $n=30$) було отримано результати, наведені на рис. 1. Використовуючи шкалу Чеддока було визначено найбільш взаємозалежні показники.

Під час тестування масиву пояснюючих змінних на мультиколінеарність та процедури покрокового відбору пояснюючих змінних було виявлено не повну мультиколінеарність, оскільки виконується $x_2 > x_{2кр}$ також було отримано підтвердження за критерієм Фішера $F_{1-3} > F_{кр}$ кожна пояснююча змінна є мультиколінеарною (рис. 2).

$$y = 34,1894 + 0,0002x_1 - 0,003x_2 + 0,011x_3 + \varepsilon_t, \quad (1)$$

Після послідовного відбору до тих пір поки показники якості моделі почали відповідати загальноприйнятим вимогам, опираючись на результати тесту на гетероскедастичність, перевірки на присутність мультиколінеарності, результатів тесту на основі VIF-індексу та помилки специфікації за тестом Ремзі було отримано наступний вигляд економетричної моделі рис. 3.

Table: CORREL Workfile: VUZ_U::Vuz_u\									
View	Proc	Object	Print	Name	Edit+/-	CellFmt	Grid+/-	Title	Comments+/-
Correlation									
		A	B	C	D	E			
1			Y	X1	X2	X3			
2									
3		Y	1.000000	0.848108	0.787012	0.885884			
4		X1	0.848108	1.000000	0.677359	0.767022			
5		X2	0.787012	0.677359	1.000000	0.962164			
6		X3	0.885884	0.767022	0.962164	1.000000			

Рисунок 1 – Результати кореляційного аналізу статистичних показників ЗВО

Джерело: власна розробка виконана на основі [1]

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1	0.000190	6.76E-05	2.809153	0.0093
X2	-0.000287	0.000121	-2.370091	0.0255
X3	0.011625	0.002824	4.116715	0.0003
C	34.18942	25.17522	1.358058	0.1861

R-squared	0.880590	Mean dependent var	290.9667
Adjusted R-squared	0.866812	S.D. dependent var	62.83557
S.E. of regression	22.93176	Akaike info criterion	9.226489
Sum squared resid	13672.51	Schwarz criterion	9.413315
Log likelihood	-134.3973	Hannan-Quinn criter.	9.286256
F-statistic	63.91267	Durbin-Watson stat	0.660202
Prob(F-statistic)	0.000000		

Рисунок 2 – Початковий варіант економетричної моделі

Джерело: власна розробка виконана на основі [1]

	A	B	C	D	E
1	Dependent Variable: Y				
2	Method: Stepwise Regression				
3	Date: 07/09/22 Time: 15:30				
4	Sample: 1990 2019				
5	Included observations: 30				
6	Number of always included regressors: 1				
7	Number of search regressors: 1				
8	Selection method: Stepwise forwards				
9	Stopping criterion: p-value forwards/backwards = 0.05/0.05				
10					
11	Variable	Coefficient	Std. Error	t-Statistic	Prob.*
12					
13	C	129.8729	19.99935	6.493859	0.0000
14	X1	0.000509	6.01E-05	8.470296	0.0000
15					
16	R-squared	0.719287	Mean dependent var	290.9667	
17	Adjusted R-squared	0.709261	S.D. dependent var	62.83557	
18	S.E. of regression	33.88106	Akaike info criterion	9.947930	
19	Sum squared resid	32141.94	Schwarz criterion	10.04134	
20	Log likelihood	-147.2190	Hannan-Quinn criter.	9.977814	
21	F-statistic	71.74591	Durbin-Watson stat	0.338957	
22	Prob(F-statistic)	0.000000			
23					
24	Selection Summary				
25					
26	Added X1				
27					
28	*Note: p-values and subsequent tests do not account for stepwise selection.				
29					

Рисунок 3 – Покроковий відбір пояснюючих змінних

Отримані показники можна охарактеризувати наступним чином: R-квадрат має значення 0,719. Тобто можна стверджувати що на 71,9% можливо пояснити залежну змінну «кількість ЗВО» незалежною змінною як x_1 -особи, уперше прийняті до ЗВО, а на 28,1% іншими випадковими факторами.

$$y = 129,8729 + 0,0005 * x_1 + \varepsilon_i, \quad (2)$$

Під час виконання специфікації економетричної було розглянуто модель багатофакторної регресії. Було проведено теоретичне та статистичне обґрунтування економетричної моделі, проведено тестування моделі на мультиколінеарність, функціональної форми економетричної моделі,

відбір суттєвих пояснюючих змінних до економетричної моделі та обрано за основу модель множинної регресії. Проведено графічний аналіз залишків (рис. 4).

Результатом чого було отримано криву нормального розподілу та гістограму розподілу залишків, що дає можливість візуально оцінити близькість розподілу залишків до нормального. Оскільки $p > \alpha$ залишки відповідають нормальному закону розподілу.

Протестовано залишки на гетероскедестичність за тестами: White, Glejser та Breush-Pagan-Godfrey (рис. 5).

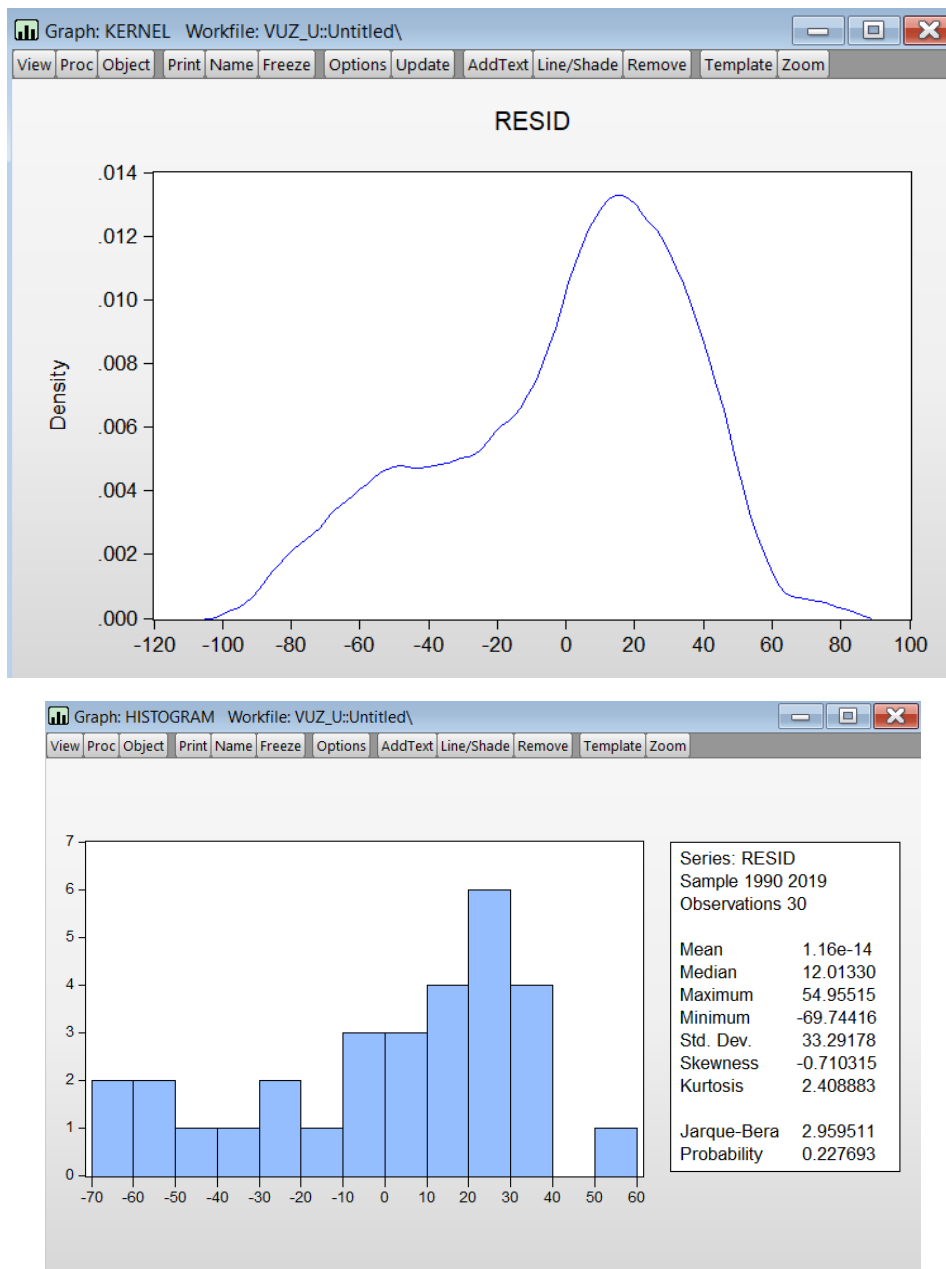


Рисунок 4 – Результати тесту Жарке-Бера для усіх варіантів моделі

Джерело: власна розробка

View	Proc	Object	Print	Name	Edit+/-	CellFmt	Grid+/-	Title	Comments+/-
		A	B	C	D	E			
1	Heteroskedasticity Test: Breusch-Pagan-Godfrey								
2									
3	F-statistic		5.575776	Prob. F(1,28)					0.0254
4	Obs*R-squared		4.981963	Prob. Chi-Square(1)					0.0256
5	Scaled explained SS		3.057167	Prob. Chi-Square(1)					0.0804
6									
7									
8	Test Equation:								
9	Dependent Variable: RESID^2								
10	Method: Least Squares								
11	Date: 07/09/22 Time: 15:04								
12	Sample: 1990 2019								
13	Included observations: 30								
14									
15		Variable	Coefficient	Std. Error	t-Statistic	Prob.			
16									
17		C	2664.750	709.5692	3.755447	0.0008			
18		X1	-0.005037	0.002133	-2.361308	0.0254			
19									
20	R-squared		0.166065	Mean dependent var		1071.398			
21	Adjusted R-squared		0.136282	S.D. dependent var		1293.451			
22	S.E. of regression		1202.087	Akaike info criterion		17.08585			
23	Sum squared resid		40460378	Schwarz criterion		17.17926			
24	Log likelihood		-254.2877	Hannan-Quinn criter.		17.11573			
25	F-statistic		5.575776	Durbin-Watson stat		0.675607			
26	Prob(F-statistic)		0.025403						
27									

Рисунок 5 – Тест Бреуша-Пагана-Голдфрі для усіх варіантів моделі

Джерело: власна розробка

Результати тестів Бреуша-Пагана-Голдфрі та Вайта для усіх варіантів моделі гетероскедастичність присутня за тестом Глейзера відсутня. Отже, за двома тестами з трьох випадкові величини мають різну дисперсію – можна стверджувати про те, що модель є гетероскедастична. Побудована оцінена модель з фіксованим ефектом та ϵ з фіктивною зміною.

Тестування залишків на автокореляцію:

$$DW = 0.339$$

$$dL = 1,352 \quad dU = 1,489$$

$$4 - dL = 2,65 \quad 4 - dU = 2,51$$

Оскільки результат тесту Дарбіна – Уотсона на автокореляцію не попадає в проміжок $2,51 < DW < 2,65$, отже автокореляція відсутня.

Тест *Бреуша-Голдфрі* також підтвердив відсутність автокореляції. Інтервали довіри для параметрів моделі зображено на рис. 6.

Часткові коефіцієнти еластичності та стандартизовані коефіцієнти регресії зображено на рис. 7.

При збільшенні абітурієнтів уперше прийнятих до ЗВО на 1% кількість ЗВО в середньому збільшиться на 0,553% (рис. 8).

Equation: EQX1 Workfile: VUZ_U:Untitled\									
View	Proc	Object	Print	Name	Freeze	Estimate	Forecast	Stats	Resids
Coefficient Confidence Intervals									
Date: 07/09/22 Time: 21:03									
Sample: 1990 2019									
Included observations: 30									
Variable	Coefficient	90% CI		95% CI		99% CI			
		Low	High	Low	High	Low	High		
X1	0.000509	0.000407	0.000612	0.000386	0.000632	0.000343	0.000675		
C	129.8729	95.85143	163.8944	88.90614	170.8397	74.60950	185.1364		

Рисунок 6 – Інтервали довіри для параметрів моделі

Джерело: власна розробка

Equation: EQX1 Workfile: VUZ_U:Untitled\			
View	Proc	Object	Print Name Freeze Estimate Forecast Stats Resids
Scaled Coefficients			
Date: 07/09/22 Time: 21:12			
Sample: 1990 2019			
Included observations: 30			
Variable	Coefficient	Standardized Coefficient	Elasticity at Means
X1	0.000509	0.848108	0.553650
C	129.8729	NA	0.446350

Рисунок 7 – Часткові коефіцієнти еластичності та стандартизовані коефіцієнти регресії

Джерело: власна розробка

Значення прогнозу на 2020 рік дорівнює 258 ЗВО. Нижня межа допустимого відхилення від прогнозу становить 188, а верхня дорівнює 327. Передбачувальна сила знаходиться на рівні 94,6%.

Висновки. З ймовірністю 90% можна стверджувати, що при збільшенні кількості абітурієнтів уперше прийнятих до ЗВО на 1000 осіб мінімальне збільшення кількості ЗВО буде 0,401 од, а максимальне 0,61 од. При P=95% можна стверджувати, що при збільшенні кількості абітурієнтів уперше прийнятих до ЗВО на 1000 осіб мінімальне збільшення кількості ЗВО буде 0,39 од, а максимальне 0,63 од. З ймовірністю 99% можна стверджувати, що при збільшенні кількості абітурієнтів уперше прийнятих до ЗВО на 1000 осіб мінімальне збільшення кількості ЗВО буде 0,34 од, а максимальне 0,68 од.

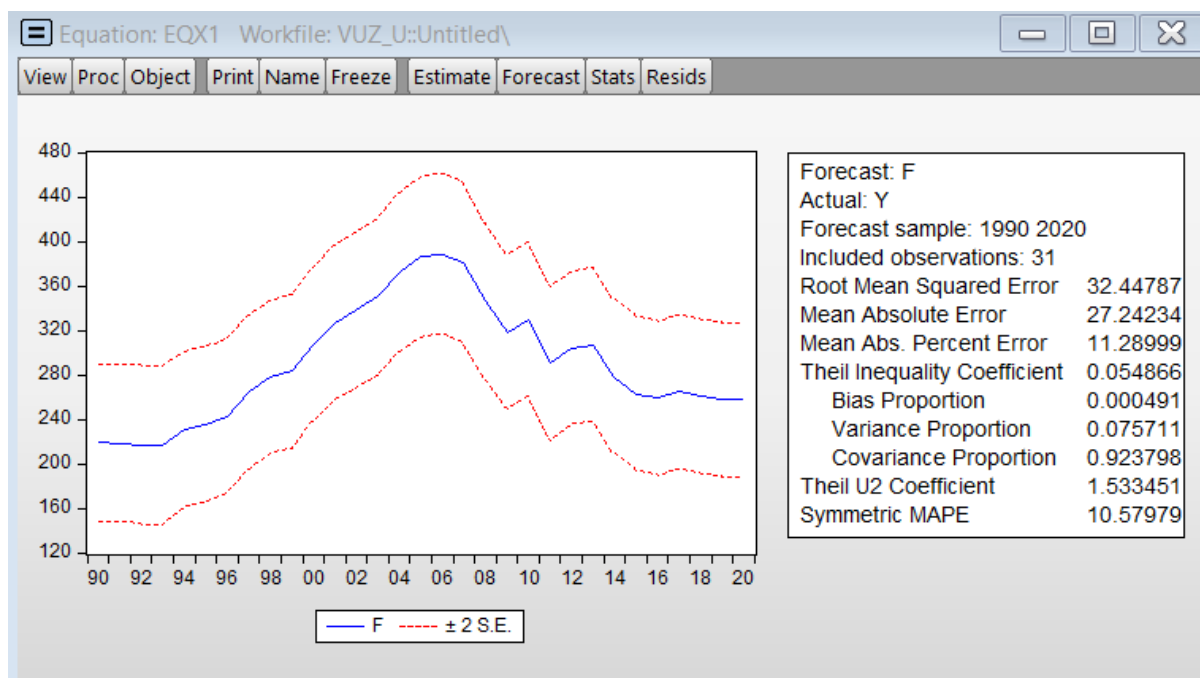


Рисунок 8 – Графік прогнозних значень залежної змінної моделі

Джерело: власна розробка

Література:

1. Державна служба статистики України : веб-сайт. URL: <https://www.ukrstat.gov.ua/> (дата звернення: 09.02.2023).
2. Кількість отриманих патентів вченими університету : веб-сайт. URL: <http://www.euroosvita.net> (дата звернення: 09.02.2023).
3. Міністерство освіти і науки України : веб-сайт. URL: <https://mon.gov.ua/ua/osvita/cifrova-osvita/elektronna-naukovo-informacijna-sistema> (дата звернення: 09.02.2023).

References:

1. Derzhavna sluzhba statystyky Ukrayiny: veb-sayt. Available at: <https://www.ukrstat.gov.ua/> (accessed February 9, 2023). (in Ukrainian)
2. Kilkist otrymanykh patentiv vchenymy universytetu: veb-sayt. Available at: <http://www.euroosvita.net> (accessed February 9, 2023). (in Ukrainian)
3. Ministerstvo osvity i nauky Ukrayiny: veb-sayt. Available at: <https://mon.gov.ua/ua/osvita/cifrova-osvita/elektronna-naukovo-informacijna-sistema> (accessed February 9, 2023). (in Ukrainian)

Стаття надійшла до редакції 15.06.2023 р.