Лабораторна робота № 4

Мета роботи: навчитися будувати нелінійну багатофакторну економетричну модель та виконувати економічний аналіз характеристик взаємозв'язку.

Приклад виконання лабораторної роботи

Завдання. За допомогою табличного редактора Microsoft Excel на основі статистичних даних (табл. 4.1) побудувати та дослідити економетричну модель парної нелінійної регресії.

Необхідно:

1. Виконати ідентифікацію змінних та привести кожну нелінійну модель (табл. 4.2) до лінійного виду.
2. Оцінити параметри моделей.
3. Визначити коефіцієнт детермінації, індекс кореляції та коефіцієнт еластичності.
4. Перевірити адекватність та точність економетричних моделей, обрати найкращу та визначити точковий прогноз для заданого значення незалежної змінної.

Таблиця 4.1 – Вихідні дані

№	Працезатрати, тис. люд.-дні	Основні засоби, тис. грн.	Обсяг виробництва, млн. грн.
1	33,3	44,4	70,1
2	38,5	47,3	78,6
3	40,4	48,2	81,7
4	45,8	49,8	88,7
5	46,2	52,2	90,8
6	50,0	54,3	96,4
7	52,1	54,4	99,7
8	56,7	58,7	106,1
9	57,9	61,5	109,2
10	60,5	63,8	112,8
11	63,0	64,2	117,1
12	65,1	65,0	119,1
13	66,6	65,4	123,2
14	67,2	66,0	124
прогнозні значення	*33,3*	*44,4*	

Розв'язання.

1. Ідентифікація змінних та приведення нелінійної моделі до лінійного виду.

Ідентифікуємо змінні економетричної моделі:

X_1, X_2 – незалежні змінні;

Y − результативна (залежна, ендогенна) змінна;

u − вектор залишків (стохастична складова).

Загальний вигляд моделі:

$$Y = f\left(X_1,\, X_2,\, u\right).$$

Розглянемо специфікації економетричної моделі $Y = f\left(X_1,\, X_2,\, u\right)$, використавши таблицю 4.2.

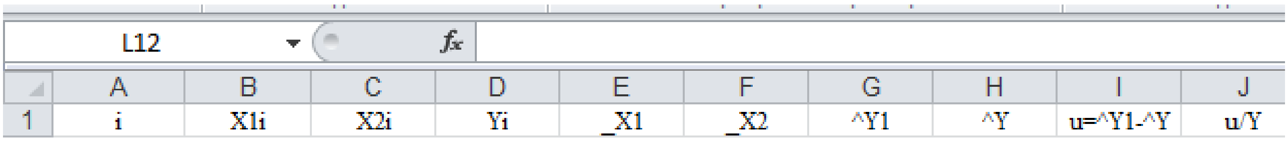
Таблиця 4.2 – Специфікація економетричної моделі

№	Вид моделі	Операція (якщо необхідно)	Заміна
1	$y = a_0 + \frac{a_1}{x_1} + \frac{a_2}{x_2}$	–	$X_1 = \frac{1}{x_1}$, $X_2 = \frac{1}{x_2}$
2	$y = a_0 \cdot a_1^{x_1} \cdot a_2^{x_2}$	логарифмування обох частин залежності за натуральною основою $\ln y = \ln a_0 + x_1 \ln a_1 + x_2 \ln a_2$	$Y = \ln y$, $\bar{a}_0 = \ln a_0$, $\bar{a}_1 = \ln a_1$, $\bar{a}_2 = \ln a_2$
3	$y = a_0 \cdot x_1^{a_1} \cdot x_2^{a_2}$	логарифмування обох частин залежності за натуральною основою $\ln y = \ln a_0 + a_1 \ln x_1 + a_2 \ln x_2$	$Y = \ln y$, $\bar{a}_0 = \ln a_0$, $X_1 = \ln x_1$, $X_2 = \ln x_2$
4	$y = a_0 + a_1 x_1 + a_2 x_2^2$	–	$X_1 = x_1$, $X_2 = x_2^2$
5	$y = a_0 + a_1 \lg x_1 + a_2 \lg x_2$	–	$X_1 = \lg x_1$, $X_2 = \lg x_2$
6	$y = a_0 \cdot e^{a_1 x_1} \cdot e^{a_2 x_2}$	логарифмування обох частин залежності за натуральною основою $\ln y = \ln a_0 + a_1 x_1 + a_2 x_2$	$Y = \ln y$, $\bar{a}_0 = \ln a_0$

Відповідно, будемо мати $Y_1 = a_0 + a_1 X_1 + a_2 X_2 + u_1$, $\widehat{Y_1} = \widehat{a}_0 + \widehat{a}_1 X_1 + \widehat{a}_2 X_2$.

2. Оцінка параметрів моделей.

Для оцінки параметрів моделей застосуємо функцію ЛИНЕЙН(Y;X;1;1) для кожного з випадків. Для цього в розрахункову таблицю треба додати відповідні стовпчики X_1, Y_1, $\widehat{Y_1}$, $u = Y - \widehat{Y}$, $\frac{|u|}{Y} \cdot 100\%$, а також значення залежних змінних моделі у вигляді оберненої заміни та за допомогою знайдених параметрів (рис. 4.1). Для зручності кожну модель треба розмістити в окремих книгах, за назви яких можна взяти їх номери з таблиці 4.1.



	A	B	C	D	E	F	G	H	I	J
1	i	X1i	X2i	Yi	_X1	_X2	^Y1	^Y	u=^Y1-^Y	u/Y

Рисунок 4.1 – Розрахункова таблиця зі значеннями замінених змінних на прикладі функції 1 з таблиці 4.2

Виділивши область розміром 5 рядків і 3 стовпці відповідно під кожною з шести функцій (табл. 4.2), застосуємо до кожної з них функцію ЛИНЕЙН(Y;X;1;1) і заповнимо таблицю 4.3.

Таблиця 4.3 – Результати застосування функції ЛИНЕЙН

для функцій 1					
$\hat{a}_2$	$\hat{a}_1$	$\hat{a}_0$	-5616,43919	-709,62187	215,86198
$S_{\hat{a}_2}$	$S_{\hat{a}_1}$	$S_{\hat{a}_0}$	1358,22244	732,96689	10,60140
R^2	σ_u		0,98321	2,45448	#Н/Д
F	$k=n-m$		322,12306	11	#Н/Д
$SSR=\sum_{i=1}^{n}\left(\hat{Y}_i-\bar{Y}\right)^2$	$SSE=\sum_{i=1}^{n}\left(\hat{Y}_i-Y_i\right)^2$		3881,24576	66,26924	#Н/Д
для функцій 2					
$\hat{a}_2$	$\hat{a}_1$	$\hat{a}_0$	-0,00239	0,01785	3,79142
$S_{\hat{a}_2}$	$S_{\hat{a}_1}$	$S_{\hat{a}_0}$	0,00375	0,00259	0,07996
R^2	σ_u		0,99420	0,01490	#Н/Д
F	$k=n-m$		943,12391	11	#Н/Д
$SSR=\sum_{i=1}^{n}\left(\hat{Y}_i-\bar{Y}\right)^2$	$SSE=\sum_{i=1}^{n}\left(\hat{Y}_i-Y_i\right)^2$		0,41861	0,00244	#Н/Д
для функцій 3					
$\hat{a}_2$	$\hat{a}_1$	$\hat{a}_0$	0,19405	0,69176	1,08825
$S_{\hat{a}_2}$	$S_{\hat{a}_1}$	$S_{\hat{a}_0}$	0,07478	0,04624	0,12501
R^2	σ_u		0,99902	0,00612	#Н/Д
F	$k=n-m$		5619,19454	11	#Н/Д
$SSR=\sum_{i=1}^{n}\left(\hat{Y}_i-\bar{Y}\right)^2$	$SSE=\sum_{i=1}^{n}\left(\hat{Y}_i-Y_i\right)^2$		0,42064	0,00041	#Н/Д
для функцій 4					
$\hat{a}_2$	$\hat{a}_1$	$\hat{a}_0$	0,00135	1,47062	18,73154
$S_{\hat{a}_2}$	$S_{\hat{a}_1}$	$S_{\hat{a}_0}$	0,00115	0,08934	1,20159
R^2	σ_u		0,99901	0,59469	#Н/Д
F	$k=n-m$		5575,5233	11	#Н/Д

$SSR=\sum_{i=1}^{n}\left(\widehat{Y}_i-\bar{Y}\right)^2$	$SSE=\sum_{i=1}^{n}\left(\widehat{Y}_i-Y_i\right)^2$		3943,6248	3,89021	#Н/Д
для функції 5					
$\widehat{a}_2$	$\widehat{a}_1$	$\widehat{a}_0$	123,1717	104,5525	-293,7379
$S_{\widehat{a}_2}$	$S_{\widehat{a}_1}$	$S_{\widehat{a}_0}$	45,20221	27,94895	32,81835
R^2	σ_u		0,99281	1,60605	#Н/Д
F	$k=n-m$		759,7026	11	#Н/Д
$SSR=\sum_{i=1}^{n}\left(\widehat{Y}_i-\bar{Y}\right)^2$	$SSE=\sum_{i=1}^{n}\left(\widehat{Y}_i-Y_i\right)^2$		3919,1417	28,37331	#Н/Д
для функції 6					
$\widehat{a}_2$	$\widehat{a}_1$	$\widehat{a}_0$	-0,00239	0,01785	3,79142
$S_{\widehat{a}_2}$	$S_{\widehat{a}_1}$	$S_{\widehat{a}_0}$	0,00375	0,00259	0,07996
R^2	σ_u		0,99420	0,01490	#Н/Д
F	$k=n-m$		943,12391	11	#Н/Д
$SSR=\sum_{i=1}^{n}\left(\widehat{Y}_i-\bar{Y}\right)^2$	$SSE=\sum_{i=1}^{n}\left(\widehat{Y}_i-Y_i\right)^2$		0,41861	0,00244	#Н/Д

3. Визначення коефіцієнту детермінації, індексу кореляції та коефіцієнту еластичності.

Для визначення необхідних коефіцієнтів зробимо попередні розрахунки (табл. 4.4, рис. 4.2).

Таблиця 4.4 – Основні елементи базової таблиці дисперсійного аналізу функції 1

позначення	*статистична функція*	*значення*
джерело варіації регресії	$k_1=m-1$	2
джерело варіації залишків	$k_2=n-m$	11
джерело варіації загальної змінної	$n-1$	13
SSR	=КВАДРОТКЛ($\widehat{Y}$)	3881,24576
SSE	=СУММКВРАЗН($Y;\widehat{Y}$)	66,2692449
SST	=КВАДРОТКЛ(Y)	3947,515
MSR	$\frac{SSR}{m-1}$	1940,62288
MSE	$\frac{SSE}{n-m}$	6,02447681

MST	$\frac{SST}{n-1}$	303,655

N19

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	i	X1i	X2i	Yi	_X1	_X2	^Y1	^Y	u=^Y1-^Y	мод(u/Y)		-5616,43919	-709,62187	215,86198	
2	1	33,3	44,4	70,1	0,030030	0,022523	68,055640	68,055640	2,044360	0,029163		1358,22244	732,96689	10,60140	
3	2	38,5	47,3	78,6	0,025974	0,021142	78,689461	78,689461	-0,089461	0,001138		0,98321	2,45448	#Н/Д	
4	3	40,4	48,2	81,7	0,024752	0,020747	81,773452	81,773452	-0,073452	0,000899		322,12306	11	#Н/Д	
5	4	45,8	49,8	88,7	0,021834	0,020080	87,588154	87,588154	1,111846	0,012535		3881,24576	66,26924	#Н/Д	
6	5	46,2	52,2	90,8	0,021645	0,019157	92,907583	92,907583	-2,107583	0,023211					
7	6	50	54,3	96,4	0,02	0,018416	98,236044	98,236044	-1,836044	0,019046	n	14	k_1	2	
8	7	52,1	54,4	99,7	0,019194	0,018382	98,998235	98,998235	0,701765	0,007039	m	3	k_2	11	
9	8	56,7	58,7	106,1	0,017637	0,017036	107,666212	107,666212	-1,566212	0,014762					
10	9	57,9	61,5	109,2	0,017271	0,016260	112,281779	112,281779	-3,081779	0,028221	SSR	3881,24576	MSR	1940,62288	
11	10	60,5	63,8	112,8	0,016529	0,015674	116,100735	116,100735	-3,300735	0,029262	SSE	66,26924	MSE	6,02447681	2,454481
12	11	63	64,2	117,1	0,015873	0,015576	117,114669	117,114669	-0,014669	0,000125	SST	3947,51500	MST	303,655	
13	12	65,1	65	119,1	0,015361	0,015385	118,554738	118,554738	0,545262	0,004578					
14	13	66,6	65,4	123,2	0,015015	0,015291	119,328726	119,328726	3,871274	0,031423	проверка	3947,515			
15	14	67,2	66	124	0,014881	0,015152	120,204572	120,204572	3,795428	0,030608					

Рисунок 4.2 – Розрахунок ANOVA – таблиці для функції 1

Перевіримо розрахунки, відповідно до теореми додавання дисперсій:

$$SSR + SSE = SST .$$

Наприклад, для функції 1:

$$3881,24576 + 66,26924 = 3947,515 .$$

Чим менша стандартна помилка залишків $\hat{\sigma}_u = \sqrt{MSE}$, тим краще підібрана функція. Порівняємо стандартні помилки для всіх функції (табл. 4.5).

Таблиця 4.5 – Стандартні помилки залишків $\hat{\sigma}_u$

номер функції	значення стандартної помилки залишків	вибір кращої функції
1	2,454481	
2	0,014897	
3	0,006118	краща функція
4	0,594689	
5	1,606049	
6	0,014897	

Значення обчислених коефіцієнтів детермінації (без урахування числа ступенів свободи та скорегований), індексу кореляції, а також характеристику зв'язку між незалежними змінними (додаток А) відповідно, внесемо до табл. 4.6.

Таблиця 4.6 – Коефіцієнти детермінації та індекс кореляції

номер функції	$R^2 = \frac{SSR}{SST}$	$\bar{R}^2 = 1 - \frac{MSE}{MST}$	$\eta = \sqrt{R^2}$	характеристика зв'язку
1	0,983212	0,980160	0,991571	достатньо сильний
2	0,994202	0,993148	0,997097	достатньо сильний

3	0,999022	0,998844	0,999511	дуже сильний
4	0,999015	0,998835	0,999507	дуже сильний
5	0,992812	0,991506	0,996400	достатньо сильний
6	0,994202	0,993148	0,997097	достатньо сильний

Значення обчислених частинних коефіцієнтів еластичності та загальної еластичності внесемо до табл. 4.7. В останній стовпчик таблиці внесено як змінюються обсяги виробництва, коли всі враховані фактори збільшуються одночасно на 1%.

Таблиця 4.7 – Частинні коефіцієнти еластичності та загальна еластичність

номер функції	$E_{Y/X_1} = \hat{a}_1 \cdot \frac{\overline{X}_1}{\overline{Y}}$	$E_{Y/X_2} = \hat{a}_2 \cdot \frac{\overline{X}_2}{\overline{Y}}$	$A = \sum_{i=1}^{n} E_{Y/X_i}$	збільшення або зменшення (%) обсягів виробництва при одночасному збільшенні всіх врахованих факторів на 1%
1	-372,10718	-3150,753	-3522,8602	зменшуються на 352286,02%
2	0,533815	0,559647	1,093463	збільшуються на 109,34%
3	0,3627423	0,1088589	0,4716013	збільшуються на 47,16%
4	0,7711538	0,0007592	0,771913	збільшуються на 77,19%
5	54,824589	69,097815	123,9224	збільшуються на 12392,34%
6	0,009358	-0,001342	0,008016	збільшуються на 0,8%

4. Перевірка адекватність та точність економетричних моделей, обрання найкращої та визначення точкового прогнозу для заданого значення незалежної змінної.

Перевірку точності та адекватності економетричної моделі визначимо відповідно за допомогою середньої відносної похибки апроксимації та критерію Фішера.

За допомогою знайдених значень параметрів заповнюємо розрахункову таблицю і обчислюємо середню відносну похибку апроксимації $\overline{\varepsilon} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \cdot 100\% = \frac{1}{n} \cdot \left|\frac{u}{Y}\right| \cdot 100\%$ (рис. 4.3).

	A	B	C	D	E	F	G	H	I	J
1	i	X1i	X2i	Yi	_X1	_X2	^Y1	^Y	u=^Y1-^Y	мод(u/Y)
2	1	33,3	44,4	70,1	0,030030	0,022523	68,055640	68,055640	2,044360	0,029163
3	2	38,5	47,3	78,6	0,025974	0,021142	78,689461	78,689461	-0,089461	0,001138
4	3	40,4	48,2	81,7	0,024752	0,020747	81,773452	81,773452	-0,073452	0,000899
5	4	45,8	49,8	88,7	0,021834	0,020080	87,588154	87,588154	1,111846	0,012535
6	5	46,2	52,2	90,8	0,021645	0,019157	92,907583	92,907583	-2,107583	0,023211
7	6	50	54,3	96,4	0,02	0,018416	98,236044	98,236044	-1,836044	0,019046
8	7	52,1	54,4	99,7	0,019194	0,018382	98,998235	98,998235	0,701765	0,007039
9	8	56,7	58,7	106,1	0,017637	0,017036	107,666212	107,666212	-1,566212	0,014762
10	9	57,9	61,5	109,2	0,017271	0,016260	112,281779	112,281779	-3,081779	0,028221
11	10	60,5	63,8	112,8	0,016529	0,015674	116,100735	116,100735	-3,300735	0,029262
12	11	63	64,2	117,1	0,015873	0,015576	117,114669	117,114669	-0,014669	0,000125
13	12	65,1	65	119,1	0,015361	0,015385	118,554738	118,554738	0,545262	0,004578
14	13	66,6	65,4	123,2	0,015015	0,015291	119,328726	119,328726	3,871274	0,031423
15	14	67,2	66	124	0,014881	0,015152	120,204572	120,204572	3,795428	0,030608
16										
17	*прогн*	*33,3*	*44,4*			_a0	215,86198			
18	*сума*					_a1	-709,62187			0,232011
19	*n*			14		_a2	-5616,43919		**відн пох**	1,66%
20										

Рисунок 4.3 – Розрахункова таблиця зі значеннями замінених змінних на прикладі функції 1 з таблиці 4.2

Для обчислення середньої відносної похибки апроксимації для кожної моделі треба знайти суму значень $\frac{|u|}{Y} \cdot 100\%$ і помножити її на $\frac{1}{n}$ (рис. 4.3).

Використовуючи таблицю 4.3, а саме число, записане в четвертому рядку і першому стовпці, визначимо фактичне значення F–критерію. Табличне значення для заданого рівня значущості $\alpha = 0,05$ та числа ступенів свободи $k_1 = m - 1 = 2$ і $k_2 = n - 3 = 11$ знайдемо використавши функцію редактора Excel:

$$F_{\text{табл}} = \text{F.ОБР.ПХ}\left(\alpha; k_1; k_2\right) = 3{,}982298.$$

Порівнюємо фактичне і табличне значення критерію і робимо висновок про адекватність моделі (рис. 4.4).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	i	X1i	X2i	Yi	_X1	_X2	^Y1	^Y	u=^Y1-^Y	мод(u/Y)		-5616,43919	-709,62187	215,86198	
2	1	33,3	44,4	70,1	0,030030	0,022523	68,055640	68,055640	2,044360	0,029163		1358,22244	732,96689	10,60140	
3	2	38,5	47,3	78,6	0,025974	0,021142	78,689461	78,689461	-0,089461	0,001138		0,98321	2,45448	#Н/Д	
4	3	40,4	48,2	81,7	0,024752	0,020747	81,773452	81,773452	-0,073452	0,000899		322,12306	11	#Н/Д	
5	4	45,8	49,8	88,7	0,021834	0,020080	87,588154	87,588154	1,111846	0,012535		3881,24576	66,26924	#Н/Д	
6	5	46,2	52,2	90,8	0,021645	0,019157	92,907583	92,907583	-2,107583	0,023211					
7	6	50	54,3	96,4	0,02	0,018416	98,236044	98,236044	-1,836044	0,019046	*n*	14	*k_1*	2	
8	7	52,1	54,4	99,7	0,019194	0,018382	98,998235	98,998235	0,701765	0,007039	*m*	3	*k_2*	11	
9	8	56,7	58,7	106,1	0,017637	0,017036	107,666212	107,666212	-1,566212	0,014762					
10	9	57,9	61,5	109,2	0,017271	0,016260	112,281779	112,281779	-3,081779	0,028221	*SSR*	3881,24576	*MSR*	1940,62288	
11	10	60,5	63,8	112,8	0,016529	0,015674	116,100735	116,100735	-3,300735	0,029262	*SSE*	66,26924	*MSE*	6,02447681	2,454481
12	11	63	64,2	117,1	0,015873	0,015576	117,114669	117,114669	-0,014669	0,000125	*SST*	3947,51500	*MST*	303,655	
13	12	65,1	65	119,1	0,015361	0,015385	118,554738	118,554738	0,545262	0,004578					
14	13	66,6	65,4	123,2	0,015015	0,015291	119,328726	119,328726	3,871274	0,031423	**проверка**	3947,515			
15	14	67,2	66	124	0,014881	0,015152	120,204572	120,204572	3,795428	0,030608					
16											R^2	0,983212	0,991571	F_факт	322,12306
17	*прогн*	*33,3*	*44,4*			_a0	215,86198				_R^2	0,980160		F_табл	3,982298
18	*сума*					_a1	-709,62187			0,232011					
19	*n*			14		_a2	-5616,43919		**відн пох**	1,66%					**адекватна**
20															

Рисунок 4.4 – Розрахункова таблиця зі значеннями замінених змінних на прикладі функції 1 з таблиці 4.2

Для вибору найкращої моделі занесемо в таблицю значення відносних похибок апроксимації всіх моделей (табл. 4.8) і порівняємо їх за допомогою додатку Г.

Таблиця 4.8 – Значення відносних похибок апроксимації всіх моделей

№ моделі	Значення відносної похибки апроксимації	рівень прогнозу	вибір
1	1,657%	менше як 10%, висока якість	
2	1,069%	менше як 10%, висока якість	
3	0,449%	менше як 10%, висока якість	
4	0,447%	менше як 10%, висока якість	найкраща
5	1,066%	менше як 10%, висока якість	
6	1,069%	менше як 10%, висока якість	

Оскільки, згідно табл. 4.8, найкращою моделлю є модель №4, то обчислимо прогнозне значення:

$$Y_{\text{прогн}} = 18,73154 + 1,47062 \cdot 33,3 + 0,00135 \cdot 44,4^2 = 70,37099.$$