

ЛАБОРАТОРНА РОБОТА

ПОБУДОВА РЕГРЕСИВНОЇ МОДЕЛІ З ДВОМА ЗМІННИМИ

При моделюванні та аналізі багатьох соціально-економічних явищ та процесів виникає задача виявлення та оцінки зв'язку між ними, одне з яких є незалежною змінною (x), чи фактором, а інше (y) – залежною, або результативною ознакою.

Форма зв'язку між змінними x та y встановлюється шляхом логічного аналізу їх природи та зовнішнього вигляду кореляційного поля та емпіричної лінії регресії, а тіснота зв'язку – величиною коефіцієнта кореляції.

В економічних дослідженнях досить часто використовуються лінійні економетричні моделі з двома змінними.

У випадку лінійного зв'язку між змінними x та y тіснота (щільність) зв'язку оцінюється коефіцієнтом парної кореляції, який можна розрахувати за наступною формулою:

$$t_{xy} = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{s_x \cdot s_y}, \quad (1.1)$$

де $\overline{xy}$ – середнє значення добутку змінної x на змінну y ; $\bar{x}$, $\bar{y}$ – середнє значення змінних x та y .

Параметри лінійної економетричної моделі з двома змінними a_0 і a_1 оцінюються на основі методу найменших квадратів шляхом розв'язання системи нормальних рівнянь:

$$\begin{cases} n \cdot a_0 + a_1 \sum_{i=1}^n x_i \cdot y_i \\ a_0 \cdot \sum_{i=1}^n x_i + a_1 \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i \cdot y_i \end{cases}, \quad (1.2)$$

де n – кількість спостережень або довжина вибірки.

Лінійна залежність між змінними x та y є частинним випадком більш загальної форми зв'язку – нелінійної. В економетричних дослідженнях досить часто використовуються такі нелінійні моделі:

- 1) гіпербола: $y = a_0 + \frac{a_1}{x} + u$;
- 2) показникова функція: $y = a_0 \cdot a_1^x + u$;
- 3) степенева функція: $y = a_0 \cdot x^{a_1} + u$;
- 4) напівлогарифмічна функція: $y = a_0 + a_1 \ln x + u$;
- 5) логарифмічна функція: $\ln y = a_0 + a_1 \cdot x + u$;
- 6) парабола: $y = a_0 + a_1 \cdot \sqrt{x} + u$, $y = a_0 + a_1 \cdot x + u$, $y = a_0 + a_1 \cdot x + a_2 \cdot x^2 + u$.

У випадках, коли функція є нелінійною (але не поліномом), її зводять до лінійної шляхом логарифмування та наступною заміною нелінійних змінних на лінійні (показникова та степенева функції) або просто однією заміною в залежності від виду функції. Цей метод приведення нелінійної функції до лінійного виду називають лінеаризацією рівняння регресії.

Розглянемо методику лінеаризації наведених вище функцій:

1. Якщо $y = a_0 + \frac{a_1}{x}$, то зміною аргументу $x' = \frac{1}{x}$ зводимо рівняння до виду $y = a_0 + a_1 \cdot x'$.
2. Якщо $y = a_0 \cdot a_1^x$, то логарифмуючи це рівняння, отримуємо $\ln y = \ln a_0 + x \cdot \ln a_1$. Заміною змінних $y' = \ln y$, $a'_0 = \ln a_0$, $a'_1 = \ln a_1$ отримуємо $y' = a'_0 + a'_1 \cdot x$.
3. Аналогічно, якщо $y = a_0 \cdot x^{a_1}$, то логарифмуючи рівняння, отримуємо $\ln y = \ln a_0 + a_1 \cdot \ln x$. Заміною змінних $y' = \ln y$, $x' = \ln x$, $a'_0 = \ln a_0$ отримуємо рівняння $y' = a'_0 + a'_1 \cdot x$.
4. У випадку $y = a_0 + a_1 \cdot \ln x$, приймаючи $x' = \ln x$, отримуємо лінеаризовану функцію $y = a_0 + a_1 \cdot x'$.
5. Якщо $y = a_0 + a_1 \cdot x^2$ чи $y = a_0 + a_1 \cdot \sqrt{x}$, то заміною $x' = \sqrt{x}$ або $x' = x^2$ зведемо їх до лінійної функції.
6. Для функції $\ln y = a_0 + a_1 \cdot x$ робимо заміну $y' = \ln y$ і отримуємо лінійну економетричну модель $y' = a_0 + a_1 \cdot x$.

Такими ж методами можна лінеаризувати інші нелінійні функції.

Приклад. На основі статистичних даних показника Y і фактора X знайти оцінки параметрів економетричної моделі, якщо припустити, що стохастична залежність між фактором X і показником Y має вигляд $y = a_0 + a_1 \ln x$.

X	1	1,5	2	2,5	3	3,5	4	4,5	5	5,5	6	6,5	7	7,5	8,5
Y	7,2	5,5	5	9,2	8,8	12,2	11,6	13,9	15,6	13,3	18,8	20,8	18,8	20,7	?

Розв’язок. Розраховуємо параметри моделі a_0 і a_1 . Записуємо систему нормальних рівнянь:

$$\begin{cases} n \cdot a_0 + a_1 \cdot \sum_{i=1}^n \ln x_i = \sum_{i=1}^n y_i, \\ a_0 \cdot \sum_{i=1}^n \ln x_i + a_1 \cdot \sum_{i=1}^n \ln^2 x_i = \sum_{i=1}^n y_i \ln x_i. \end{cases}$$

$$\begin{cases} 14a_0 + 18,2a_1 = 181,4, \\ 18,2a_0 + 28,5a_1 = 274,76. \end{cases}$$

Звідки $a_1 = 8,03$; $a_0 = 2,52$. Запишемо економетричну модель $y = 2,52 + 8,03 \ln x + u$.

	x_i	y_i	$\ln x_i$	$\ln^2 x_i$	$y_i \ln x_i$	$\hat{y}_i$	$(y_i - \hat{y}_i)^2$	$(y_i - \bar{y})^2$	$e^{(\ln x_i - \bar{x})^2}$
	1	7,2	0	0	0	2,52	21,94	33,14	5,41
	1,5	5,5	0,405	0,164	2,230	5,77	0,07	55,61	2,22
	2	5	0,693	0,480	3,466	8,08	9,51	63,32	1,44
	2,5	9,2	0,916	0,840	8,430	9,88	0,46	14,12	1,16
	3	8,8	1,099	1,207	9,668	11,34	6,46	17,28	1,04
	3,5	12,2	1,253	1,569	15,284	12,58	0,14	0,57	1,00
	4	11,6	1,386	1,922	16,081	13,65	4,22	1,84	1,01
	4,5	13,9	1,504	2,262	20,907	14,60	0,49	0,89	1,04
	5	15,6	1,609	2,590	25,107	15,45	0,02	6,98	1,10
	5,5	13,3	1,705	2,906	22,673	16,21	8,48	0,12	1,18
	6	18,8	1,792	3,210	33,685	16,91	3,57	34,14	1,27
	6,5	20,8	1,872	3,504	38,933	17,55	10,54	61,51	1,39
	7	18,8	1,946	3,787	36,583	18,15	0,42	34,14	1,52
	7,5	20,7	2,015	4,060	41,708	18,70	3,99	59,95	1,67
Σ	59,5	181,4	18,20	28,50	274,76	-	70,32	383,61	22,46
середні	4,25	12,96	1,30						

Завдання до лабораторної роботи. На основі статистичних даних показника Y і фактора X (табл. 1) знайти оцінки параметрів економетричної моделі, якщо припустити, що стохастична залежність між фактором X і показником Y має вигляд, наведений в табл. 1.

Таблиця 1

Варіант	1	2	3	4	5	6	7
X	Y	Y	Y	Y	Y	Y	Y
1	5,2	2,5	7,2	26,7	3,3	3,7	3,2
1,5	4,5	2,8	5,5	23,9	3,6	5,1	4
2	4,1	2,2	5	22,8	3,8	4,5	4,7
2,5	4,9	3,2	9,2	24,5	4,1	5,6	3,8
3	4	3,6	8,8	20,5	3,8	5,2	5,5
3,5	3,6	4	12,2	19,3	4,1	6,5	4,2
4	3,7	3,1	11,6	18,9	4,4	5,8	6,1
4,5	4,9	4,8	13,9	22,2	4,6	7,6	6,1
5	3,7	4,2	15,6	18,6	4,6	7,7	5,5
5,5	3,5	5,5	13,3	19,8	4,7	8	6,4
6	4,4	6,6	18,8	17,5	4,8	7,1	5,8
6,5	3,5	6	20,8	17,2	4,4	9,1	6,9
7	3,8	7,4	18,8	17,8	4,9	8,5	6,9
7,5	3,3	8	20,7	15,8	5	10	7,2
8,5	?	?	?	?	?	?	?
Залежність	$y = a_0 + \frac{a_1}{x}$	$y = a_0 + a_1 \ln x$	$y = a_0 \cdot x^{a_1}$	$y = \frac{1}{a_0 + a_1 x}$	$y = a_0 + \frac{a_1}{x}$	$y = \frac{x}{a_0 + a_1 x}$	$y = a_0 + a_1 \ln x$