

Тема 2. Методи побудови загальної лінійної моделі

Лекція 2. Особливості економетричних моделей

Зміст

- [1. Роль і місце економетричних моделей в управлінні економічними системами](#)
- [2. Формування сукупності спостережень](#)
- [3. Поняття однорідності спостережень](#)
- [4. Точність економічних даних](#)
- [5. Вибір змінних і структура зв'язків](#)
- [6. Випадкова складова економетричної моделі](#)
- [7. Оцінювання параметрів моделі методом найменших квадратів](#)
- [8. Коефіцієнт кореляції, перевірка гіпотези про значущість коефіцієнта кореляції](#)
- [9. Індекс кореляції](#)
- [10. Перевірка моделі на адекватність за критерієм Фішера](#)

1. Роль і місце економетричних моделей в управлінні економічними системами.

Сучасні методи управління економічними системами та процесами базуються на широкому використанні математичних методів та ЕОМ.

На сучасному етапі економічна наука сягнула високих рубежів у своєму розвитку і розглядає задачі, для розв'язування яких недостатньо використовувати традиційні економічні методи, тому математика займає в економіці одне з провідних місць. Сформувався напрям – економіко-математичне моделювання. Математичне

моделювання є вираженням процесу математизації наукового економічного знання. Основним поняттям є математична модель.

Математична модель об'єкта містить у собі 3 групи елементів:

- 1) характеристику об'єкта, яку потрібно визначити (невідомі величини) – вектор $Y=(y_j)$;
- 2) характеристики зовнішніх щодо модельованого об'єкта умов, які змінюються – вектор $X=(x_j)$;
- 3) сукупність внутрішніх параметрів об'єкта – A .

Множини умов X і параметрів A розглядаються як екзогенні величини (такі, що визначаються поза рамками моделі), а величини, що належать вектору Y – ендогенні (такі, що визначаються за допомогою моделі).

Математичну модель можна розуміти як перетворювач зовнішніх умов об'єкта Y (виходу), які мають бути знайдені.

Математичні моделі поділяють на дві групи:

- структурні;
- функціональні.

Вивчення структурних моделей дає одночасно цінну інформацію про поведінку об'єкта. Структурні моделі розкривають внутрішню організацію об'єкта : його складові частини, внутрішні параметри, їх зв'язок з входом і виходом. Є три види структурних моделей :

- 1) $y_j = f_j(A,X)$ – явні функції від зовнішніх умов X і внутрішніх параметрів A ;
- 2) $\psi_i (A,X,Y)=0$ – неявні функції, всі невідомі визначаються одночасно із системи рівнянь;
- 3) імітаційні моделі – конкретний вигляд співвідношень невідомий.

Моделі типу (1) і (2) можна розв'язати за допомогою чисельних алгоритмів.

Імітаційні моделі не зводяться до чітко визначених математичних задач і тому їм належить проміжне місце між структурними і функціональними моделями.

Основна ідея функціональних моделей – пізнання суті об'єкта через його поведінку, діяльність, функціонування.

Функціональна модель описує поводження об'єкта так, що задаючи значення входу X можна отримати значення виходу Y без участі інформації про параметри :

$$Y = A(X),$$

де A – оператор, який пов'язує X і Y .

При вивченні функціональних моделей необхідно сформулювати гіпотезу про внутрішню структуру об'єкта.

Економетричні моделі належать до функціональних моделей. Вони кількісно описують зв'язок між вхідними показниками економічної системи (X) та результативним показником (Y).

У загальному вигляді економетричну модель можна записати так :

$$y=f(x;u),$$

де x – вхідні економічні показники; u – випадкова або стохастична складова.

Показники X найчастіше бувають детермінованими, тобто не залежать від впливу випадкових факторів.

Складова u – випадкова змінна, але оскільки залежна змінна Y залежить від u , то вона є теж *стохастичною*, описує кореляційно-регресійний зв'язок між економічними величинами.

Отже, для побудови економетричної моделі необхідно :

- 1) мати досить велику сукупність спостережень даних;
- 2) забезпечити однорідність сукупності спостережень;
- 3) забезпечити точність вхідних даних.

[До змісту](#)

2. Формування сукупності спостережень.

Сукупність спостереження розглянута в теорії статистики. Слід розрізняти джерело даних – одиницю спостереження та одиницю сукупності. Під час перепису : одиниця спостереження – сім'я, а одиниця сукупності – окрема людина. Ці поняття часто збігаються. Тому в економетрії здебільшого йдеться про одиницю сукупності.

Нехай X – певна ознака спостереження, то X_{ij}^t або X_{ijt} означає j ознаку i – го спостереження в період t , ($i=1;m$), ($j=1;n$), ($t=1;T$) :

n – число одиниць сукупності;

m – число ознак, які описують кожну одиницю;

T – проміжок часу, за який вивчається ознака певного спостереження.

За одиницю сукупності часто беруть певний економічний об'єкт, що функціонує – технологічний агрегат, цех, підприємство, галузь, людину, тварину.

Існує три способи формування вибірки :

- часова, коли розглядається функціонування окремого об'єкта в динаміці;
- просторова, коли сукупність спостережень вивчається у статиці;
- просторово-часова – комбінація просторової і часової вибірки.

[До змісту](#)

3. Поняття однорідності спостережень.

Економічні сукупності як правило неоднорідні. Поняття однорідності сукупності спостережень охоплює якісну і кількісну однорідність. Під

якісною розуміють *однотипність* економічних об'єктів, їх *однакову якість та призначення*, а під кількісною – *однорідність* групи одиниць спостережень, що визначається на основі *кількісних оцінок*.

Кількісна однорідність можлива лише за наявності *однотипності* явищ та процесів, що утворюють сукупність спостережень, тобто ці *поняття однорідності діалектично взаємопов'язані*.

Дані вхідної сукупності спостережень повинні мати :

- однаковий ступінь агрегування;
- однорідну структуру одиниць сукупності;
- однакові методи розрахунку показників у часі;
- однакову періодичність обліку окремих змінних;
- порівнянні ціни та однакові інші зовнішні економічні умови.

[До змісту](#)

4. Точність економічних даних.

При економічних розрахунках постає питання про точність (помилки) економічних показників. Помилки показників виникають при побудові алгоритму розрахунку, формуванні даних, у процесі обчислення. Усі помилки поділяють на *систематичні* і *випадкові*.

Систематичні помилки мають або *постійну* величину або змінюються за певною *функціональною* залежністю. Вони є *однонаправленими* і можуть бути істотними за величиною.

Випадкові помилки зумовлюються впливом *випадкових чинників* при формуванні показників. При формуванні сукупності спостережень для економетричних моделей слід звертати увагу на можливість існування помилок саме у вхідних даних.

[До змісту](#)

5. Вибір змінних і структура зв'язків.

При економетричному моделюванні треба добре знати об'єкт дослідження, а саме :

- визначити набір змінних, що описують процес функціонування досліджуваних об'єктів;
- проаналізувати взаємозв'язки між окремими змінними;
- вибрати раціональний тип економетричної моделі.

Вибір результативних ознак вирішується, виходячи з *мети дослідження*. Вибір незалежних змінних є процесом уточнення початкової гіпотези. У цьому процесі можна вирізнити такі етапи :

- формування початкової гіпотези про набір незалежних змінних;
- експертна оцінка цього набору;
- аналіз взаємозв'язків;
- добір і звуження кола істотних для моделювання змінних.

[До змісту](#)

6. Випадкова складова економетричної моделі.

Розглянемо економетричну модель з двома змінними у загальному вигляді:

$$y = f(x) + u,$$

де y – залежна змінна; x – незалежна змінна;

u – випадкова складова. Це проста економетрична модель.

Одні і ті самі економічні процеси можна описати різними функціями, але за допомогою статистичного аналізу потрібно вибрати одну.

Найпростіша форма зв'язку між двома змінними лінійна:

$$y = a_0 + a_1x,$$

де a_0, a_1 – невідомі параметри.

Можливі й інші форми залежностей між двома змінними. Як правило нелінійні функції зводять до лінійних шляхом логарифмування або заміни :

$$y = a_0 e^{a_1 x}, \ln y = \ln a_0 + a_1 x;$$

$$y = a_0 x^{a_1}, \ln y = \ln a_0 + a_1 \ln x, \ln x = z;$$

$$y = a_0 + a_1/x, 1/x = z, y = a_0 + a_1 z;$$

$$y = \alpha(1-r)^x, \ln y = \ln \alpha + x \ln(1-r);$$

$$y = e^{a_0 + a_1 x}, \ln y = a_0 + a_1 x;$$

$$y = 10^{a_0 + a_1 x}, \lg y = a_0 + a_1 x.$$

Нехай ми вибрали якусь із цих залежностей, визначили невідомі параметри і одержали рівняння регресії.

Як відомо певна частина фактичних спостережень над змінною лежатиме вище або нижче від значень, обчислених згідно з вибраною функцією. А це свідчить про те, що вибрана функція не адекватна реальному процесу взаємозв'язків у економіці. Щоб розв'язати цю задачу, до економетричної моделі вводять стохастичну (випадкову) складову, яка акумулює в собі всі відхилення фактичних спостережень змінної у від обчислених згідно з моделлю:

$$y = a_0 + a_1 x + u.$$

Стохастичну складову u економетричної моделі називають похибкою (залишком, збуренням, відхиленням).

У класичній лінійній економетричній моделі змінна u інтерпретується як випадкова змінна, що має розподіл з середньою величиною, яка дорівнює нулю і сталою дисперсією σ_u^2 . Оскільки u охоплює вплив багатьох факторів, які є незалежними, то згідно з центральною граничною теоремою теорії ймовірності можна зробити висновок, що стохастична (випадкова) складова економетричної моделі розподілена за нормальним законом.

[До змісту](#)

7. Оцінювання параметрів моделі методом найменших квадратів.

Щоб оцінити параметри моделі, потрібно сформувати сукупність спостережень, кожна одиниця якої характеризується відповідними значеннями $(x; y)$.

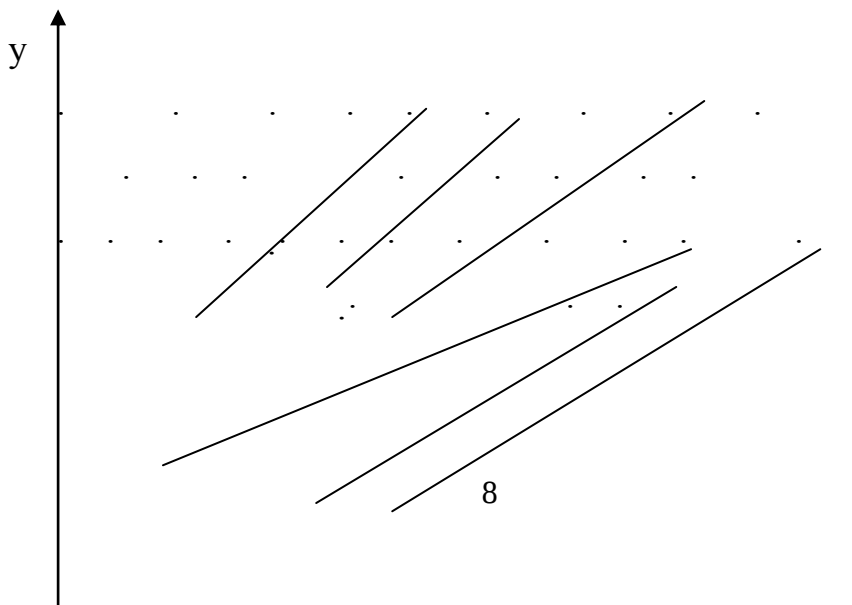
Нехай побудували кореляційне поле точок $(x; y)$, через які можна провести безліч прямих ліній $y = a_0 + a_1 x$ (рис. 1).

Вони різняться між собою параметрами a_0 і a_1 . Треба вибрати ту лінію, для якої б відхилення фактичних y від розрахункових $\hat{y}$ було б найменшим.

Принцип найменших квадратів відхилень полягає в знаходженні таких $\hat{a}_0$ і $\hat{a}_1$, для якої $\sum u_i^2$ була б найменшою ($\hat{y} = \hat{a}_0 + \hat{a}_1 x$ — для розрахункових).

$S = \sum u_i^2 = \sum (y_i - \hat{y}_i)^2 = \sum (y_i - \hat{a}_0 - \hat{a}_1 x)^2 \rightarrow \min$. Для виконання цієї умови візьмемо перші похідні за невідомими параметрами.

$$\begin{cases} \frac{\partial (\sum_{i=1}^n u_i^2)}{\partial a_0} = -2 \sum_{i=1}^n (y_i - \hat{a}_0 - \hat{a}_1 x_i) = 0; \\ \frac{\partial (\sum_{i=1}^n u_i^2)}{\partial a_1} = -2 \sum_{i=1}^n x_i (y_i - \hat{a}_0 - \hat{a}_1 x_i) = 0. \end{cases}$$



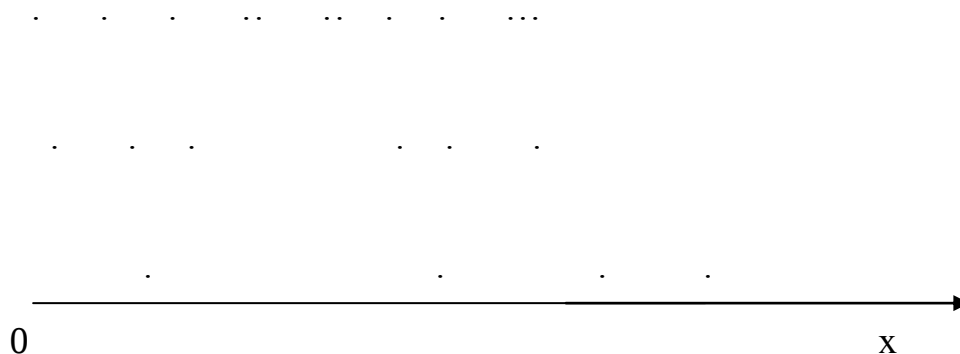


Рис. 1. Кореляційне поле точок

Виконавши елементарні перетворення, дістанемо систему нормальних рівнянь

$$\begin{cases} n\hat{a}_0 + \hat{a}_1 \sum_{i=1}^n x_i = \sum_{i=1}^n y_i; \\ \hat{a}_0 \sum_{i=1}^n x_i + \hat{a}_1 \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i y_i. \end{cases}$$

Підставимо в систему рівнянь значення $\sum x_i$, $\sum y_i$, $\sum x_i^2$, $\sum x_i y_i$, які обчислюються на основі сукупності спостережень, і розв'яжемо її відносно невідомих оцінок параметрів $\hat{a}_0$ і $\hat{a}_1$:

$$\hat{a}_0 = \frac{\sum x \sum xy - \sum x^2 \sum y}{(\sum x)^2 - n \sum x^2} \quad \hat{a}_1 = \frac{\sum x \sum y - n \sum xy}{(\sum x)^2 - n \sum x^2}$$

$$\hat{a}_0 = \frac{\sum x_i^2 \sum y_i - \sum x_i \sum x_i y_i}{n \sum x_i^2 - (\sum x_i)^2}; \quad \hat{a}_1 = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - (\sum x_i)^2}.$$

Оцінки параметрів $\hat{a}_0$ і $\hat{a}_1$ за методом найменших квадратів такі, що лінія регресії обов'язково проходить через точку $(\bar{X}, \bar{Y})$, тобто $\bar{Y} = \hat{a}_0 + \hat{a}_1 \bar{X}$.

Є ще інший альтернативний спосіб обчислення параметрів $\hat{a}_0$ і $\hat{a}_1$:

$$\hat{a}_1 = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$

де $\hat{a}_0$ знаходиться з рівняння:

$$\hat{a}_0 = \bar{y} - \hat{a}_1 \bar{x}.$$

За нашим прикладом $\bar{x} = 1008/10=100,8$, $\bar{y} = 542/10=54,2$.

Приклад. Побудувати економетричну модель залежності витрат ресурсів на одиницю продукції (у) від рівня фондомісткості продукції (х) (дані в умовних одиницях). Розрахунки проводимо в табл.1.

Запишемо систему нормальних рівнянь:

$$\begin{cases} n\hat{a}_0 + \hat{a}_1 \sum x_i = \sum y_i & ; \\ \hat{a}_0 \sum x_i + \hat{a}_1 \sum x_i^2 = \sum y_i x_i & ; \end{cases} \quad \begin{cases} 10\hat{a}_0 + \hat{a}_1 1008 = 542 & ; \\ 1008\hat{a}_0 + \hat{a}_1 104784 = 56221 & ; \end{cases}$$

$$\hat{a}_0 = \frac{542 \times 104784 - 56221 \times 1008}{10 \times 104784 - 1008^2} = \frac{56792928 - 56670768}{1047840 - 1016064} = \frac{122160}{31776} = 3,84;$$

$$\hat{a}_1 = \frac{10 \times 56221 - 1008 \times 542}{31776} = \frac{562210 - 546336}{31776} = \frac{15874}{31776} = 0,4993 \approx 0,5.$$

Економетрична модель має вигляд (для фактичних значень незалежної змінної)

$$Y = 3,8 + 0,5 x + u.$$

Скориставшись альтернативним способом обчислення за допомогою відхилень середніх арифметичних, одержимо:

$$\hat{a}_1 = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} = \frac{1587,5}{3376} = 0,4702 \approx 0,5,$$

де $\hat{a}_0$ знаходиться з рівняння :

$$\hat{a}_0 = \bar{Y} - \hat{a}_1 \bar{x} = 54,2 - 50,4 = 3,8.$$

Оскільки вільний член моделі $\hat{a}_0 \neq 0$ то рівень витрат на одиницю продукції не є строго пропорційним до рівня фондомісткості, $\hat{a}_1 = 0,5$ показує, що граничне збільшення витрат при зростанні фондомісткості продукції на 1 ум. од. становить 0,5 ум. од.

Еластичність витрат щодо фондомісткості продукції визначається коефіцієнтом *еластичності* :

$$E = \frac{\partial Y}{\partial X} / \frac{Y}{X};$$

$$E = 0,5 \cdot 1,8519 = 0,93.$$

Таблиця 1.

Вихідні і розрахункові дані для побудови економетричної моделі

№ п/п	x_i	y_i	x_i^2	y_i^2	$x_i y_i$	$x_i - \bar{x}$	$y_i - \bar{y}$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})(y_i - \bar{y})$	$\hat{y}_i$	u_i	u_i^2	$(y_i - \bar{y})^2$	$(\hat{y}_i - \bar{y})^2$
1	90	50	8100	2500	4500	-10,8	-4,2	116,64	45,36	48,8	1,2	1,44	17,64	29,16
2	75	40	5625	1600	3000	-25,8	-14,2	665,64	336,36	41,3	-1,3	1,69	201,64	166,41
3	120	65	14400	4225	7800	19,2	10,8	368,64	207,36	63,8	1,2	1,44	116,64	92,16
4	100	55	10000	3025	5500	-0,8	0,8	0,64	-0,64	53,8	1,2	1,44	0,64	0,16
5	80	45	6400	2025	3600	-20,8	-9,2	432,64	181,36	43,8	1,2	1,44	84,64	108,16
6	78	42	6084	1764	3276	-22,8	-12,2	519,64	278,16	42,8	-0,8	0,64	148,84	129,96
7	110	56	12100	3136	6160	9,2	1,8	84,64	16,56	58,8	-2,8	7,84	3,24	21,16
8	115	60	13225	3600	6900	14,2	5,8	201,64	82,36	61,3	-1,3	1,69	33,64	50,41
9	115	64	13225	4096	7350	14,2	9,8	201,64	139,16	61,3	2,7	7,29	96,04	50,41
10	125	65	15625	4225	8125	24,2	10,8	585,64	261,26	66,3	-1,3	1,69	116,64	146,41
Σ	1008	542	104784	30196	56221	0	0	3376	1587,5	542	0	26,6	819,6	794,4

Значення цього коефіцієнта слід тлумачити так : при збільшенні фондомісткості продукції на 1 % витрати на одиницю продукції зростуть на 0,93 %.

Залишки обчислюються згідно з рівністю :

$$u_i = y_i - \hat{y}_i.$$

Оцінка дисперсії залишків подається так :

$$\sigma^2 = \frac{\sum U_i^2}{n-2} = \frac{26,6}{8} = 3,325.$$

Для залишків u_i можна задати певну функцію закону розподілу, наприклад функцію нормального розподілу.

[До змісту](#)

8. Коефіцієнт кореляції, перевірка гіпотези про значущість коефіцієнта кореляції.

Оцінкою коефіцієнта кореляції є вибірковий коефіцієнт кореляції, який можна обчислити за формулою :

$$r = \frac{\overline{xy} - \bar{x}\bar{y}}{\sigma_x \times \sigma_y} ;$$

$$\sigma_x^2 = \overline{x^2} - \bar{x}^2 = \frac{104784}{10} - 100,8^2 = 317,76 ; \quad \sigma_x = \sqrt{317,76} = 17,8258 ;$$

$$\sigma_y^2 = \overline{y^2} - \bar{y}^2 = \frac{30198}{10} - 54,2^2 = 81,96 ; \quad \sigma_y = \sqrt{81,96} = 9,0532 ;$$

$$r = \frac{\overline{xy} - \bar{x}\bar{y}}{\sigma_x \times \sigma_y} = \frac{5622,1 - 100,8 \times 54,2}{17,8 \times 9,05} = 0,98.$$

Оскільки коефіцієнт кореляції додатний зв'язок між досліджуваними ознаками прямий. Він прямує до 1, а це свідчить про тісний зв'язок між ознаками.

Вибірковий коефіцієнт кореляції є точковою оцінкою коефіцієнта кореляції і тому потребує перевірки. Вона базується на критерії Стюдента за формулою :

$$t = \frac{|r|\sqrt{n-m}}{\sqrt{1-r^2}},$$

де r – вибірковий коефіцієнт кореляції; $n-m$ – число ступенів вільності.

Якщо $t > t_{\text{табл. } \alpha}$, де $t_{\text{табл. } \alpha}$ – відповідне табличне значення t розподілу з $(n-m)$ ступенями вільності, то можна зробити висновок про значущість коефіцієнта кореляції між залежною і незалежною змінними моделі. Для нашого прикладу :

$$t = \frac{|r|\sqrt{n-m}}{\sqrt{1-r^2}} = \frac{0,98\sqrt{10-2}}{\sqrt{1-0,98^2}} = \frac{2,7719}{0,1990} = 13,9291.$$

Табличне значення t – критерія для рівня значущості $\alpha = 0,05$ і $n-m=8$ ступенів вільності дорівнює 2,306. Оскільки $t > t_{\text{табл. } \alpha}$ робимо висновок, що коефіцієнт кореляції є значущим і зв'язок між x та y існує.

[До змісту](#)

9. Індекс кореляції.

Для аналізу якості опису існуючої залежності між двома ознаками часто використовують індекс кореляції. Він розраховується за формулою :

$$\eta = \sqrt{\frac{\sum(\tilde{y} - \bar{y})^2}{\sum(y - \bar{y})^2}} = \sqrt{\frac{794,4}{819,6}} = 0,9845 \approx 0,98.$$

Оскільки коефіцієнт кореляції прямує до 1 робимо висновок, що між досліджуваними ознаками існує тісний зв'язок. Виконується рівність між коефіцієнтом кореляції r і індексом кореляції η , що підтверджує наявність прямого лінійного, тісного зв'язку між досліджуваними ознаками.

[До змісту](#)

10. Перевірка моделі на адекватність за критерієм Фішера.

Перевірити модель на адекватність можна за допомогою також F-критерію :

$$F_{k-1, n-k} = \frac{\sum (\tilde{Y} - \bar{Y})^2}{\sigma_u^2}.$$

За даними табл. 1 фактичне значення критерію Фішера буде дорівнювати:

$$F_{1,8} = \frac{\sum (\tilde{Y} - \bar{Y})^2}{\sigma_u^2} = \frac{794,4}{3,325} = 238,9.$$

Оскільки $F_{\text{факт}} > F_{\text{табл}}$, можна зробити висновок, що побудована модель адекватна реальній дійсності і може бути використана у подальших економічних дослідженнях, а саме у прогнозуванні економічних показників для прийняття обґрунтованих, ефективних економічних рішень.

[До змісту](#)