

УДК 339.175

Г.В. Іванець

Харківський торгово-економічний інститут  
Київського національного торгово-економічного університету, Харків

## МЕТОДИКА ВИБОРУ МАТЕМАТИЧНОЇ МОДЕЛІ ЗАЛЕЖНОСТІ ЕКОНОМІЧНИХ ПОКАЗНИКІВ ВІД ІНШИХ ЧИННИКІВ НА ОСНОВІ СТАТИСТИЧНИХ ДАНИХ В УМОВАХ ЧАСТКОВОЇ НЕВИЗНАЧЕНОСТІ

*Розглянута методика вибору математичної моделі залежності економічних показників від інших чинників на основі статистичних даних в умовах часткової невизначеності. При виборі моделі необхідно, щоб вона була адекватною явищу, що вивчається, і досить добре відтворювала реальну ситуацію, а випадкова складова моделі була найменшою. Якщо заздалегідь невідома форма моделі, то при її виборі в якості критерію беруть найменше значення суми квадратів випадкових складових за статистичними даними, якому відповідає найбільший коефіцієнт детермінації серед можливих форм моделей. Оскільки моделі можуть описуватись як лінійними, так і нелінійними функціями, то перед дослідженням всі нелінійні функції необхідно привести до лінійного вигляду.*

**Ключові слова:** математична модель, адекватність моделі, коефіцієнт детермінації, сума квадратів випадкових складових.

### Вступ

**Загальна постановка проблеми.** В сучасному конкурентному середовищі для здобуття успіху економісти намагаються передбачити майбутні значення тих чи інших економічних показників, запропонувати інструменти і способи зміни ситуації в бажаному напрямку. Одним із основних завдань економічного аналізу є моделювання розвитку економічних явищ і процесів при створенні тих чи інших умов. В умовах часткової невизначеності зв'язків між досліджуваними параметрами і факторами, які впливають на них, виникає проблема вибору структури моделі, яка б якомога точніше описувала і пояснювала процеси на основі статистичних даних.

**Мета статті.** Розробка методики вибору математичної моделі залежності економічних показників від інших чинників на основі статистичних даних в умовах часткової невизначеності.

### Викладення матеріалів досліджень

Для економіки неможливе будь-яке експериментування, тому особливого значення набуває математичне моделювання. Моделювання – це конкретне відтворення певних характеристик досліджуваного явища, що дає змогу вивчати можливу поведінку явища без проведення експериментів над ним. Математична модель повинна відповідати таким вимогам [2, 3]:

- будуватися на основі економічної теорії й відображати об'єктивні закономірності процесів;
- правильно відтворювати функцію та структуру реальної економічної системи;
- мати рішення, узгоджені розмірності тощо.

Будь-яке дослідження завжди поєднує математичні моделі і статистичні дані. Економічні дані є

кількісними характеристиками економічних об'єктів. Вони формуються під дією багатьох факторів, які не завжди можна проконтролювати ззовні. Неконтрольовані фактори можуть набувати випадкових значень і тим самим зумовлювати випадковість даних. У багатьох дослідженнях виявляється, що деяка результативна ознака змінюється під впливом кількох факторів. Розглянемо можливі моделі опису зв'язку між результативною змінною і чинниками (пояснюючими змінними), які впливають на неї.

Вплив багатьох чинників на результативну змінну може описуватися лінійною моделлю:

$$y = a_0 + a_1x_1 + a_2x_2 + \dots + a_mx_m + e, \quad (1)$$

де  $y$  – досліджувана (залежна змінна);

$x_1, x_2, \dots, x_m$  – незалежні, пояснюючі змінні;

$a_0, a_1, \dots, a_m$  – параметри моделі;

$e$  – випадкова складова моделі.

На основі статистичних даних можна знайти параметри цієї моделі, застосовуючи відомий метод найменших квадратів (МНК) [1, 2, 5].

Якщо модель описується нелінійними функціями, то спочатку їх зводять до лінійних, оскільки лінійні функції мають багато переваг перед іншими і добре вивчені [3, 5]. Розглянемо методи зведення деяких нелінійних функцій до лінійних.

Степенева функція:

$$y = a \cdot x_1^{a_1} \cdot x_2^{a_2} \cdot \dots \cdot x_m^{a_m}, \quad (2)$$

після логарифмування набуває вигляду:

$$\ln y = \ln a + a_1 \ln x_1 + a_2 \ln x_2 + \dots + a_m \ln x_m, \quad (3)$$

і після заміни  $\ln a = a_0$  є лінійною відносно параметрів  $a_0, a_1, \dots, a_m$ .

Показникова функція:

$$y = b_0 \cdot b_1^{x_1} \cdot b_2^{x_2} \cdot \dots \cdot b_m^{x_m}, \quad (4)$$

після логарифмування набирає вигляду:

$$\ln y = \ln b_0 + x_1 \cdot \ln b_1 + x_2 \cdot \ln b_2 + \dots + x_m \cdot \ln b_m, \quad (5)$$

і після заміни  $\ln b_i = a_i$ ,  $i = 0, 1, 2, \dots, m$  є лінійною відносно параметрів  $a_0, a_1, \dots, a_m$ . Зокрема, відома виробнича функція Кобба-Дугласа, яка описує залежність між обсягом виробленої продукції  $Y$  і витратами праці  $L$  та капіталу  $K$ :

$$Y = A \cdot K^{\beta_1} \cdot L^{\beta_2}, \quad (6)$$

після логарифмування набирає вигляду:

$$\ln Y = \ln A + \beta_1 \cdot \ln K + \beta_2 \cdot \ln L, \quad (7)$$

і після заміни  $\ln A = \beta_0$  є лінійною відносно параметрів  $\beta_0, \beta_1, \beta_2$ .

Гіперболічна функція:

$$y = a_0 + \frac{a_1}{x_1} + \frac{a_2}{x_2} + \dots + \frac{a_m}{x_m} \quad (8)$$

шляхом заміни змінних  $z_i = \frac{1}{x_i}$ ,  $i = 1, 2, \dots, m$  зводиться до лінійного вигляду:

$$y = a_0 + a_1 z_1 + a_2 z_2 + \dots + a_m z_m. \quad (9)$$

Квадратична функція:

$$y = a_0 + a_1 x_1^2 + a_2 x_2^2 + \dots + a_m x_m^2, \quad (10)$$

шляхом заміни змінних  $z_i = x_i^2$ ,  $i = 1, 2, \dots, m$  зводиться до лінійного вигляду:

$$y = a_0 + a_1 z_1 + a_2 z_2 + \dots + a_m z_m. \quad (11)$$

Поліноміальна функція:

$$y = a_0 + a_1 x_1 + a_2 x_2^2 + \dots + a_m x_m^m, \quad (12)$$

шляхом заміни змінних  $z_i = x_i^i$ ,  $i = 1, 2, \dots, m$  зводиться до лінійного вигляду:

$$y = a_0 + a_1 z_1 + a_2 z_2 + \dots + a_m z_m. \quad (13)$$

При виборі моделі необхідно, щоб вона була адекватною явищу, що вивчається, і досить добре відтворювала реальну ситуацію, а випадкова складова моделі  $e$  була найменшою. Тому, очевидно, що із усіх моделей необхідно вибирати ту, в якій сума квадратів випадкових складових  $e_i$  за статистичними даними є найменшою:

$$\sum_{i=1}^n e_i^2 \Rightarrow \text{найменша}. \quad (14)$$

Звичайно, що при дослідженні конкретного економічного явища першочерговим завданням є пошук форми опису статистичного зв'язку між його показниками. Форма залежності повинна мати певне економічне обґрунтування. Якщо заздалегідь невідома форма залежності, то при її виборі в якості критерію беруть найменше значення суми квадратів випадкових складових  $e_i$  за статистичними даними.

Оскільки серед функцій для вибору форми залежності можуть бути як лінійні, так і нелінійні, то перед дослідженням необхідно всі нелінійні функції привести до лінійного вигляду. Після цього для подальшого

вибору залишають ті функції, для яких коефіцієнт детермінації найбільший. Коефіцієнт детермінації:

$$R^2 = \frac{\sum (y_i^* - \bar{y})^2}{\sum (y_i - \bar{y})^2}, \quad (15)$$

де  $y_i^*$  – розрахункове значення залежної змінної;

$y_i$  – статистичне значення залежної змінної,

$$\bar{y} = \frac{1}{n} \sum_{s=1}^n y_i \text{ – середнє значення залежної змінної}$$

за статистичними даними є кількісною мірою тісноти зв'язку між залежною і незалежними змінними.

Чим сильніший зв'язок (чим більше зміна залежної змінної визначається зміною незалежних змінних), тим ближче коефіцієнт детермінації до одиниці. І, навпаки, при відсутності зв'язку коефіцієнт детермінації прямує до нуля.

Для дослідження було взято статичні дані залежності деякого економічного параметру  $y$  (залежної змінної) від інших економічних чинників  $x_1, x_2, x_3$  (незалежних змінних), зв'язок між якими описувався лінійною функцією  $y = a_0 + a_1 x_1 + a_2 x_2 + x_3 x_3 + e$ . Вважалося, що досліднику невідома форма опису статистичного зв'язку між змінними. Статистичні дані мали вигляд:

Таблиця 1

Статистичні дані

$X_1$	$X_2$	$X_3$	$Y$
2	10	100	7,653
8	5	400	14,34
4	32	200	9,894
8	20	600	15,72
9	5	600	19,49
12	20	800	19,33
10	12	400	12,63
5	28	200	10,52
7	15	250	10,07
7	30	250	9,78
3	50	100	6,61
6	70	150	5,89
1	30	220	10,25
15	25	900	21,02
14	46	850	18,73
11	12	700	21,35
7	17	520	15,36
4	23	300	11,95
2	75	80	4,55
1	55	20	6,1

Завдання полягало у виборі форми опису невідомої математичної моделі, яка б найкраще відображала статистичний зв'язок між змінними. У виборі приймали участь лінійна функція, степенева функція, показникові функція, гіперболічна функція, поліноміальна і квадратична функція. При цьому нелінійні функції перед дослідженням були приведені до лінійного вигляду. В результаті дослідження були отримані дані, наведені в табл. 2.:

Результати досліджень

Форма моделі	Значення коефіцієнту детермінації $R^2$	Адекватність моделі	Сума квадратів випадкових складових ( $e_i$ )
Лінійна	0,967054	адекватна	17,87417265
Степенева	0,856262	адекватна	54,06985
Показникова	0,963372	адекватна	19,34077
Гіперболічна	0,404877464	не адекватна	322,8711
Поліноміальна	0,889414454	адекватна	59,99585
Квадратична	0,927029605	адекватна	39,58854

Аналізуючи отримані дані, можна зробити наступні висновки:

- з подальшого дослідження необхідно виключити гіперболічну модель, як завідомо неадекватну;
- для адекватних моделей простежується залежність між значенням коефіцієнту детермінації  $R^2$  і величиною суми квадратів випадкових складових  $\sum_{i=1}^n e_i^2$ .

Чим більше значення коефіцієнту детермінації, тим менша величина суми квадратів випадкових складових;

- найкраще відображає статистичний зв'язок між змінними лінійна форма моделі, яка має найвищий коефіцієнт детермінації  $R^2 = 0,967054$  і найменшу величину суми квадратів випадкових складових  $\sum_{i=1}^n e_i^2 = 17,87417265$ ;

- результати дослідження повністю узгоджуються із заздалегідь відомою формою статистичного зв'язку між змінними.

статистичний зв'язок між змінними, якщо заздалегідь вона невідома. При цьому в якості критерію беруть найменше значення суми квадратів випадкових складових за статистичними даними. Дослідження підтвердили, що найбільшому значенню коефіцієнту детермінації відповідає найменше значення суми квадратів випадкових складових за статистичними даними. Це дозволяє при виборі моделі взяти ту, для якої коефіцієнт детермінації найвищий.

### Список літератури

1. Грубер Й. Економетрія: Вступ до множинної регресії та економетрії: Підручник Т. 2 / Й. Грубер. – К.: Нічлава, 1988. – 408 с.
2. Джонсон Дж. Економетрические методы: Учебник / Дж. Джонсон. – М.: Статистика, 1980. – 444 с.
3. Ліцинський О.Л. Економетрія: Навчальний посібник / О.Л. Ліцинський, В.В. Рязанцева, О.О. Юнькова. – К.: МАУП, 2003. – 208 с.
4. Иберла К. Факторный анализ / К. Иберла. – М.: Статистика, 1980. – 398 с.

### Висновки

Надійшла до редколегії 15.02.2013

Розглянута методика вибору форми моделі дозволяє вибрати таку, яка найкращим чином описує

**Рецензент:** канд. екон. наук проф. В.І. Ганін, Харківський торгово-економічний інститут КНТЕУ, Харків.

### МЕТОДИКА ВИБОРА МАТЕМАТИЧЕСКОЙ МОДЕЛИ ЗАВИСИМОСТИ ЭКОНОМИЧЕСКИХ ПОКАЗАТЕЛЕЙ ОТ ДРУГИХ ФАКТОРОВ НА ОСНОВЕ СТАТИСТИЧЕСКИХ ДАННЫХ В УСЛОВИЯХ ЧАСТИЧНОЙ НЕОПРЕДЕЛЕННОСТИ

Г. В. Иванец

*Рассмотрена методика выбора математической модели зависимости экономических показателей от других факторов на основе статистических данных в условиях частичной неопределенности. При выборе модели необходимо, чтобы она была адекватной явлению, которое изучается, и довольно хорошо отображала реальную ситуацию, а случайная составляющая модели была наименьшей. Если заведомо неизвестна форма модели, то при ее выборе в качестве критерия берут наименьшее значение суммы квадратов случайных составляющих по статистическим данным, которому соответствует наибольший коэффициент детерминации среди возможных форм моделей. Поскольку модели могут описываться как линейными, так и нелинейными функциями, то перед исследованием все линейные функции необходимо привести к линейному виду.*

**Ключевые слова:** математическая модель, адекватность модели, коэффициент детерминации, сумма квадратов случайных составляющих.

### METHOD CHOICE OF MATHEMATICAL MODEL DEPENDENCE ECONOMIC INDICATORS FROM OTHER FACTORS ON THE BASIS STATISTICAL INFORMATION IN THE CONDITIONS PARTIAL VAGUENESS

G. V. Ivanets

*The method choice of mathematical model dependence economic indicators is considered from other factors on the basis statistical information in the conditions partial vagueness. It is necessary at the choice model, that it was adequate the phenomenon which is studied, and enough well represented the real situation, and a casual constituent model was the least. If a model form is scianter unknown, at its choice as a criterion take the least value sum of squares casual constituents from statistical data, which the most coefficient determinacy among the possible forms models corresponds. As models can be described both linear and nonlinear functions, before research all of linear functions must be led to the linear kind.*

**Keywords:** mathematical model, model adequacy, coefficient of determinacy, sum squares casual constituents.