

ЗАСТОСУВАННЯ БАГАТОФАКТОРНОГО РЕГРЕСІЙНОГО АНАЛІЗУ У МОДЕЛЮВАННІ ЕКОНОМІЧНИХ ПРОЦЕСІВ

Надь Н.М.

У статті розглянуто теоретичні підходи до економетричного моделювання економічних процесів та явищ. Розкрито основні етапи побудови багатофакторних регресій в економіці.

Кількість бібліографічних посилань – 6, мова – українська.

Ключові слова: багатофакторна регресія, кореляційний зв'язок, коефіцієнт кореляції, F-критерій Фішера, t-статистика Ст'юдента.

ВСТУП

На діяльність підприємства впливає ряд чинників, як позитивних, так і негативних. Оцінити результат їхньої дії можливо за допомогою кількісних методів дослідження, основу яких становлять побудова і аналіз економіко-математичних моделей. В економічних дослідженнях велике значення має теорія і практика економетричного моделювання, що дає змогу вивчити та кількісно визначити внутрішні й зовнішньо-наслідкові зв'язки між показниками економічних систем, установлювати закономірності їхнього формування і тенденції розвитку.

Питаннями математичного моделювання економічних процесів займалися такі дослідники, як Лондар С. (моделювання податково-бюджетних процесів) [2], Сирський М. (моделювання факторів прямого іноземного інвестування) [4], Яркіна Н.М. (моделювання підприємницького ризику) [6], Піскунова О.В., Рядно О.А., Горбачук В., Домрачев В., Литвин О. (дослідження економічного зростання). Основи побудови регресійних рівнянь висвітлено також у працях Бараза В.Р. [1], Лук'яненко І.Г. [3], Толбатова Ю.А. [5].

При написанні статті були використані методи економіко-математичного моделювання. Метою статті є визначення та аналіз основних етапів побудови багатофакторних регресій в економіці.

РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

Будь-який економічний процес, особливо якщо його всебічно можна виразити чи визначити через кількісні характеристики, часто піддається математичному моделюванню, тобто його вираженню через рівняння регресії. При цьому, під регресією розуміється залежність однієї випадкової величини від іншої випадкової величини. Такий процес побудови математичних залежностей між факторною (факторними) ознакою та залежною змінною дозволяє не тільки визначити наявний тісний зв'язок між даними показниками (що вирішується за допомогою кореляційного аналізу), але й прогнозувати одну (залежну) змінну (y) на основі іншої (інших) змінної (x). Таким чином, регресійний аналіз дозволяє робити аргументовані висновки щодо розвитку даного економічного процесу, що базуються та підкріплюються конкретними математичними розрахунками.

Звичайно, більш актуальними для економічного моделювання є багатофакторні регресії, оскільки рідко економічне явище можна пояснити переважно одним фактором впливу.

У процесі побудови багатофакторних регресійних моделей можна виділити такі етапи:

1. Вибір та аналіз всіх можливих факторів, які впливають на процес (або показник), що вивчається.
 2. Вимір та аналіз знайдених факторів – якщо деякі фактори неможливо кількісно чи якісно визначити або для них недоступна статистика, то вони вилучаються з подальшого розгляду.
 3. Математико-статистичний аналіз факторів – на цьому етапі при наявності у динамічних рядах недостатньої інформації за допомогою спеціальних методів проводиться її відтворення, а також здійснюється перевірка основних припущень класичного регресійного аналізу.
 4. Вибір вигляду регресійної багатофакторної моделі.
 5. Оцінка невідомих параметрів регресійної моделі.
 6. Перевірка значимості знайдених параметрів моделі та її оцінка на адекватність реальній дійсності – здійснюється за допомогою F-критерію Фішера та t-критерію Ст'юдента.
- F-статистика Фішера розраховується з m та (n-m-1) ступенями вільності:

$$F = \frac{\frac{\sum_{i=1}^n (y_{ip} - y_c)^2}{m}}{\frac{\sum_{i=1}^n (y_i - y_{ip})^2}{n - m - 1}} \quad (1)$$

де m – кількість факторів, які увійшли в модель;

n – загальна кількість спостережень;

y_{ip} – розрахункове значення залежної змінної при i -му спостереженні;

y_c – середнє значення залежної змінної;

y_i – значення залежної змінної при i -му спостереженні.

За F - таблицями Фішера знаходиться критичне значення $F_{кр}$ з m та $(n-m-1)$ ступенями вільності, задавши попередньо рівень довіри $(1 - \alpha) * 100\%$. Якщо $F > F_{кр}$, то це свідчить про адекватність побудованої моделі.

Якщо модель неадекватна, то необхідно повернутися до етапу побудови моделі і, можливо, ввести додаткові фактори, або перейти до нелінійної моделі.

t – статистика для параметрів багатofакторної регресії має вигляд:

$$t = \frac{a_i}{\sigma_{a_i}^2} \quad (2)$$

де a_i – оцінка i -го параметра;

$\sigma_{a_i}^2$ – середньоквадратичне відхилення оцінки i -го параметра.

Якщо t -значення перевищує критичне значення, яке знаходиться за таблицею t – критерію, то вважається, що відповідний параметр статистично значимий і має суттєвий вплив на узагальнюючий показник.

7. Розрахунок основної характеристики (коефіцієнт множинної кореляції R), аналіз отриманих результатів, висновки.

Коефіцієнт множинної кореляції є основним показником щільності кореляційного зв'язку узагальнюючого показника з факторами. Якщо його значення, розраховане за формулою (3), прямує до 1, то взаємозв'язок між показником і факторами вважається щільним.

$$R = \sqrt{1 - \frac{\sum_{i=1}^n (y_i - y_{ip})^2}{\sum_{i=1}^n (y_i - y_c)^2}} \quad (3)$$

Процедура пошуку рівняння регресії між різними числовими сукупностями, як правило, включає наступне:

1. встановлення значимості зв'язку між ними;
2. можливість представлення даної залежності у формі математичного виразу (рівняння регресії).

Перший етап передбачає визначення коефіцієнтів парної (між двома числовими масивами – x_i та y_i) кореляції. Взаємозв'язок між значенням даних коефіцієнтів та характеристикою сили наявного зв'язку між ними подано у таблиці 1.

Таблиця 1 Якісна оцінка тісноти зв'язку

Величина коефіцієнту парної кореляції	Характеристика сили зв'язку
До 0,3	Практично відсутній
0,3-0,5	Слабий
0,5-0,7	Помітний
0,7-0,9	Сильний

0,9-0,99	Дуже сильний
----------	--------------

Джерело [1]

Враховуючи дані таблиці, прийнято, що для врахування показників у якості факторів, придатних для моделювання, їхня тіснота зв'язку з результативною змінною моделі має бути сильна або дуже сильна, тобто дорівнювати більше 0,7. Якщо ж значення коефіцієнтів кореляції складає менше 0,7, то вважається, що зв'язок між даною групою факторів незначний і його (фактор) не варто розглядати як такий, що має суттєвий вплив на дане економічне явище. Таким чином, процедура розрахунку парних кореляційних коефіцієнтів дає змогу з усієї сукупності факторів вибрати ті, які дійсно виявляють тісний зв'язок із показником, що обраний у якості залежної змінної визначеного економічного процесу, та відкинути фактори, що практично не впливають на цей процес.

Є декілька методів визначення рівня коефіцієнта кореляції. Найбільш відомим є метод найменших квадратів. Проте цей доволі трудомісткий розрахунок можна успішно замінити використанням функціонально-статистичних залежностей табличного редактора Excel, що дозволяє пошук кореляційних коефіцієнтів звести до справи кількох хвилин.

Слід зазначити, що у рівнянні багатфакторної регресії часто виникають проблеми, так званої, мультиколінеарності: коли факторні ознаки мають тісний зв'язок не лише з функціональною (залежною) однакою, але й між собою. У таких випадках функціональний зв'язок може бути викривлений і не адекватно описувати змодельований процес. Тому часто крім розрахунку парних коефіцієнтів кореляції між незалежними змінними та залежною змінною доцільно визначати коефіцієнти парної кореляції між самими факторними ознаками та вилучати ті, які проявляють мультиколінеарність і мають при цьому менш тісний зв'язок з функціональною змінною.

Другим етапом при побудові регресійних моделей є вибір формули самого рівняння регресії. Найбільш простий вигляд має лінійна багатфакторна регресія, яка описує лінійний зв'язок між досліджуваними даними:

$$y = a_0 + a_1 x_1 + \dots + a_n x_n \quad (4)$$

де y – залежна змінна, функція;
 a_0, \dots, a_n – коефіцієнти регресії;
 x_1, \dots, x_n – залежні змінні.

У випадку криволінійних залежностей використовуються формули наступного виду:

$$y = a_0 + a_1 / x_1 + \dots + a_n / x_n \quad (\text{гіперболічна}) \quad (5)$$

$$y = a_0 + a_1 x_1^{X1} + \dots + a_n x_n^{Xn} \quad (\text{показникова}) \quad (6)$$

$$y = a_0 x_1^{A1} \dots x_n^{An} \quad (\text{степенева}) \quad (7)$$

$$y = a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n \quad (\text{параболічна}) \quad (8)$$

$$y = a_0 + a_1 \lg x_1 + \dots + a_n \lg x_n \quad (\text{логарифмічна}) \quad (9)$$

$$y = a_0 \exp(a_1 x_1) \dots \exp(a_n x_n) \quad (\text{експоненціальна}) \quad (10)$$

Емпіричні формули можуть бути різноманітними, оскільки при виборі аналітичної залежності керуються не якимись строгими теоріями (фізичними чи економічними), а ставиться тільки одна умова – можлива близька відповідність значень, розрахованих за формулою дослідним даним. Відповідні регресії можна побудувати «вручну», проте такі розрахунки є доволі громіздкими та трудомісткими. Найбільш простим, швидким та зручним методом є застосування можливостей Excel за допомогою наступних серій команд: сервіс / аналіз даних / регресія.

Висновок про придатність (значимість) рівняння регресії проводиться за допомогою таких кроків:

1. спочатку проводиться загальна перевірка за допомогою F-тесту (F-статистики Фішера). Метою даного методу є з'ясування того, чи пояснюють x - змінні значну частину варіації y -змінної. Якщо відповідь негативна, тобто побудована регресія є незначимою, неадекватною експериментальним даним, то слід шукати, будувати інший вид регресії і знов перевіряти її таким чином на адекватність.

2. Якщо ж регресія виявляється адекватною (значимою), то продовжується аналіз, використовуючи t -тести для окремих коефіцієнтів регресії (t -статистика Ст'юдента). Даний аналіз дає змогу визначити значимість впливу x_i -го фактору на залежну змінну рівняння y (при умові, що всі інші фактори залишаються незмінними). Цілом прийнятною вважається значимість коефіцієнта регресії, як і адекватність регресійної моделі, на рівні 95%. Звісно, чим вищою є ця значимість, тим побудована модель буде достовірніше відображати математичні взаємозв'язки між

факторами та залежною змінною економічних явищ та процесів.

ВИСНОВКИ

За допомогою побудованих багатofакторних регресійних моделей, які є адекватними наявним статистичним та розрахунковим даним та мають високі ступені значимості оцінених параметрів, можна здійснювати прогнозування зміни змодельованого економічного явища в результаті зміни одного чи більше його факторів. Слід констатувати, що такі моделі на сьогодні не є сталими в довгостроковому періоді, оскільки в українській економіці зовнішнє середовище та його умови змінюються досить часто. Тому в майбутньому доцільно розрахувати нові параметри регресійних моделей на основі представлених методів та здійснювати прогнозування згідно отриманих даних.

ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Бараз В.Р. Корреляционно-регрессионный анализ связи показателей коммерческой деятельности с использованием программы Excel / Бараз В.Р. – Екатеринбург, 2005. – 103с
2. Лондар С. Тенденції податково-бюджетних процесів в Україні / С. Лондар // Factors of Economic Growth in Ukraine and Neighboring Countries: міжнар. наук.-практ. конф., 21-22 листопада 2003р. : тези доп. – Київ, 2003. – С. 47-52.
3. Лук'яненко І.Г. Економетрика: Підручник / І.Г. Лук'яненко, Л.І. Краснікова. – К.: Товариство “Знання”, КОО, 1998. – 494 с.
4. Сирський М. Статистичний аналіз факторів прямого іноземного інвестування в економіку України / М. Сирський // Factors of Economic Growth in Ukraine and Neighboring Countries: міжнар. наук.-практ. конф., 21-22 листопада 2003р. : тези доп. – Київ, 2003.– С. 65-69.
5. Толбатов Ю.А. Економетрика: Підручник для студентів екон. спеціальн. вищ. навч. закл. / Толбатов Ю.А. – К.: Четверта хвиля, 1997. – 320 с.
6. Яркіна Н.М. Економетричне моделювання в управлінні підприємницьким ризиком / Н.М. Яркіна // Фінанси України. – 2003. - №11. – С. 77-80.