УДК: 519.252 DOI 10.33111/mise.100.8

Колєчкіна Л.М., д.фіз.-мат.н., професор
кафедри комп'ютерної математики та інформаційної безпеки,
Київський національний економічний університет
імені Вадима Гетьмана

Колєчкін В.О., старший викладач
кафедри комп'ютерної математики та інформаційної безпеки,
Київський національний економічний університет
імені Вадима Гетьмана

Koliechkina L.M., doctor of physical and mathematical sciences
professor of the Department of Computer Mathematics and Information Security,
KNEU named after V. Hetman

Koliechkin V.O., senior lecturer
of the Department of Computer Mathematics and Information Security,
KNEU named after V. Hetman

БАГАТОФАКТОРНА МОДЕЛЬ ЕКОНОМІЧНОЇ ЗАДАЧІ ОПТИМІЗАЦІЇ РОБОТИ ПІДПРИЄМСТВА ТА АНАЛІЗ МЕТОДІВ ЇЇ РОЗВ'ЯЗАННЯ

MULTIFACTOR MODEL OF THE ECONOMIC PROBLEM OF OPTIMIZING THE WORK OF THE ENTERPRISE AND ANALYSIS OF THE METHODS OF ITS SOLUTION

Анотація. *Різні економічні процеси часто можна описати за допомогою математичного моделюванню, тобто виразити певні залежності між економічними показниками через рівняння регресії. У процесі побудови багатофакторних регресійних моделей можуть виникнути проблема визначення закону розподілення спостережень і знаходження параметрів цього розподілу. В загальному така проблема може бути вирішена шляхом перевірки статистичних гіпотез, на основі числових статистичних тестів і критеріїв з урахуванням помилок першого і другого роду. Розглянуто економіку-математичну модель, яка може бути прикладом моделі прикладних задач. Для реалізації моделі і її застосування описано різні методи статистичного аналізу та їх порівняльні характеристики. Проведено числовий експеримент. Показано, що модель задачі носить загальний характер, але може бути використана для практичної проблеми, що пов'язана з аналізом та оптимізацією економічних показників підприємства. Під час застосування відповідного програмного забезпечення (MS Excel, GPSS World, AnyLogic і та ін.) можна швидко отримати точні показники діяльності підприємства та зробити відповідні висновки про його функціонування на даний момент, розрахувати прогнозовані значення на майбутнє.*

Ключові слова: *багатофакторна модель, оптимізація, статистичний аналіз даних, кореляційно-регресійний аналіз*

© Л.М.Колєчкіна, В.О.Колєчкін, 2022

Abstract. *Various economic processes can often be described with the help of mathematical modeling, that is, express certain dependencies between economic indicators through regression equations. In the process of building multivariate regression models, the problem of determining the law of the distribution of observations and finding the parameters of this distribution may arise. In general, such a problem can be solved by testing statistical hypotheses, based on numerical statistical tests and criteria taking into account errors of the first and second kind.*

The paper considers an economic-mathematical model, which can be an example of a model of applied problems. To implement the model and its application, various methods of statistical analysis and their comparative characteristics are described. A numerical experiment was conducted. The problem model has a general nature, but can be used for a practical problem related to the analysis and optimization of the economic indicators of the enterprise. When using the appropriate software (MS Excel, GPSS World, AnyLogic, etc.), you can quickly obtain accurate indicators of the company's activity and make appropriate conclusions regarding its current functioning, calculate forecast values for the future.

Keywords: *multifactor model, optimization, statistical data analysis, correlation-regression analysis*

Постановка проблеми. Різні економічні процеси часто можна описати за допомогою математичного моделюванню, тобто виразити певні залежності між економічними показниками через рівняння регресії. При цьому під регресією розуміється залежність між величинами, що є випадковими [1–3]. Процес побудови математичних залежностей між факторами дозволяє визначити наявний тісний зв'язок між даними показниками і спрогнозувати одну залежну змінну через інші незалежні змінні. Маємо статистичний аналіз, на основі якого можемо робити висновки про розвиток економічних показників та їх зміни. Найбільш актуальними для моделювання економічних процесів є багатофакторні регресії, так як більшість економічних показників описуються як залежність від двох і більше факторів [4–7]. У процесі побудови багатофакторних регресійних моделей можуть виникнути проблема визначення закону розподілення спостережень і знаходження параметрів цього розподілу. У загальному така проблема може бути вирішена шляхом перевірки статистичних гіпотез на основі числових статистичних тестів і критеріїв з урахуванням помилок першого і другого роду.

Аналіз останніх досліджень і публікацій. Ця теорія детально розроблена, описана в літературі і є загальновідомою [6, 7]. Але слід зазначити, що з різних причин на практиці її досить рідко використовують. Досить часто для проведення статистичного аналізу користуються можливостями табличного процесору MS Excel або іншими статистичними пакетами обробки даних. При цьому важливим аспектом є урахування специфіки вихідної зада-

чі та її властивостей і особливостей. В цьому плані є доцільним використання прикладної статистики.

Цікавим є підхід, який описаний у [8, 9] на базі застосування теорії інформації для розв'язування різного класу статистичних задач.

У статтях описано [3, 5] метод найменших квадратів (МНК), який є широко відомим і користується достатньою популярністю. Разом з тим не припиняються спроби вдосконалення цього методу. Як результат, з'являються нові модифікації і версії МНК, однієї з яких є зважений метод найменших квадратів (ЗМНК). Суть ЗМНК полягає в тому, щоб надати спостереженням вагу, що обернено-пропорційна похибкам їх апроксимації. Отже, спостереження ігноруються тим більшою мірою, чим складніше їх апроксимувати [3, 5]. В результаті такого підходу формально похибка апроксимації знижується, проте фактично це відбувається шляхом часткової відмови від перегляду проблемних спостережень, що вносять велику помилку.

Під прикладної статистикою розуміють частину математичної статистики, присвячену методам обробки реальних статистичних даних, а також відповідне математичне і програмне забезпечення. Звідси суто математичні завдання і не включають у прикладну статистику.

Під статистичними даними розуміють числові або нечислові значення контрольованих параметрів (ознак) досліджуваних об'єктів, які отримані в результаті спостережень (вимірювань, аналізів, випробувань, дослідів іта ін.) певного числа ознак, у кожної одиниці, що увійшла в дослідження. Способи отримання статистичних даних та обсяги вибірок встановлюють виходячи з постановок конкретної прикладної задачі на основі методів математичної теорії планування експерименту.

Результат спостереження x_i досліджуваної ознаки X чи їх сукупності, y_i-ї одиниці вибірки відображає кількісні та якісні властивості обстеженої ознаки з номером i (де $i = 1, 2,..., n$, а n — обсяг вибірки). На основі вибраних даних слід проаналізувати вибірку та зробити висновок про подальші дослідження ознак і характеристик математичної моделі.

Результати спостережень $x_1, x_2,..., x_n$, де x_i — результат спостереження i-ї одиниці вибірки, або результати спостережень для декількох вибірок, обробляють за допомогою методів прикладної статистики, відповідно до поставленої задачі. Використовують, як правило, аналітичні методи, тобто методи, засновані на числових розрахунках. В окремих випадках допустимо застосування графічних методів візуального аналізу [5, 6].

Кількість розроблених дотепер методів обробки даних дуже велике. Вони описані в статтях і підручниках, а також в стандартах та інших нормативно-технічних та інструктивно-методичних документах. Це пов'язано з необхідністю обробки великих масивів даних і їх аналізу [1–7].

Багато методів прикладної статистики вимагають проведення трудомістких розрахунків, і тому для їх реалізації необхідно використовувати комп'ютери. Програми розрахунків на ЕОМ повинні відповідати сучасному науковому рівню. Однак для одиничних розрахунків при відсутності відповідного програмного забезпечення успішно використовують мікрокалькулятори.

Аналіз методів математичної статистики та їх порівняння. На сьогодні в математичній статистиці розроблений ряд загальних методів визначення оцінок і похибок, такі як метод моментів, метод максимальної правдоподібності, метод однокрокових оцінок, метод стійких робасних оцінок, метод незміщене оцінок і та ін. Теоретичні основи різних методів оцінювання та отримані з їх допомогою конкретні правила визначення оцінок і довірчих кордонів для тих чи інших параметрів розподілів розглянуто в [2–5].

Більшість статистичних методів належать до методів параметричної статистики, в основу яких покладено припущення, що випадковий вектор змінних утворює деякий багатовимірний розподіл. Якщо це припущення не підтверджується, слід скористатися непараметричними методами математичної статистики.

Наприклад, відомий метод моментів, що заснований на використанні виразів для моментів розглянутих випадкових величин через параметри їх функцій розподілу дає можливість отримати оцінки, підставляючи вибіркові моменти замість теоретичних в функції, що виражають параметри.

Що стосується оцінок максимальної правдоподібності, то, як правило, вони є ефективними і мають меншу дисперсію, ніж оцінки методу моментів. В окремих випадках доцільно використовувати не оцінки максимальної правдоподібності, а інші види оцінок, наприклад, однокрокові оцінки, що називаються в літературі наближені оцінки максимальної правдоподібності. При досить великих обсягах вибірок вони мають досить гарні властивості, як і оцінки максимальної правдоподібності. Тому їх слід розглядати не як «наближені», а як оцінки, отримані по іншому методу, ніж метод максимальної правдоподібності.

Під час вивчення причинно-наслідкових зв'язків показники моделі поділяють на результатні та факторні. Це групування не постійне, воно залежить від конкретних ситуацій, мети аналізу. Напри-

клад, в оцінці змін прибутку від реалізації продукції показник собівартості продукції розглядають як факторний. Водночас у ході вивчення затрат на виробництво собівартість розглядають як результатний показник, що залежить від багатьох факторів виробництва.

Слід зазначити, що у використанні статистичного показника необхідно розглядати такі задачі: опис структури економіки підприємства, опис тенденцій розвитку економіки в майбутньому, аналіз і прогнозування різних економічних явищ, вияв факторів розвитку економіки підприємства для прийняття управлінських рішень [3–5].

У процесі моделювання економічних показників підприємства засобами статистичного аналізу можуть бути застосовані такі підходи:

1. *Кореляційний аналіз* дозволяє встановити силу і напрям стохастичного взаємозв'язку між змінними (випадковими величинами).

2. *Регресійний аналіз* забезпечує моделювання взаємозв'язку однієї випадкової змінної від однієї або декількох інших випадкових змінних. При цьому перша змінна називається залежною, а решта — незалежними. Вибір або призначення залежною і незалежних змінних є довільним / умовним і здійснюється дослідником залежно від розв'язуваної ним задачі.

3. *Канонічний аналіз* призначений для аналізу залежностей між двома списками ознак (незалежних змінних), що характеризують об'єкти.

4. *Дисперсійний аналіз* можна визначити як параметричний, статистичний метод, призначений для оцінки впливу різних чинників на результат експерименту, а також для подальшого планування експериментів.

5. *Кластерний аналіз* — це метод класифікаційного аналізу; його основне призначення — розбиття множини досліджуваних об'єктів і ознак на однорідні в деякому сенсі групи, або кластери [1–6].

Одним із важливих напрямів у статистичному аналізі є багатовимірний статистичний аналіз, який застосовують під час вирішення таких завдань:

- дослідження залежності між ознаками;
- класифікація об'єктів або ознак, заданих векторами;
- зниження розмірності простору ознак.

При цьому результатом спостереження є вектор значень фіксованого числа кількісних і іноді якісних ознак, виміряних у об'єкта. Нагадаємо, що кількісна ознака — це ознака, яку можна безпосередньо виразити числом і одиницею виміру. Кількісна ознака протиставляється якісній, що визначається віднесенням до

однієї з двох або більше умовних категорій (якщо є рівно дві категорії, то ознака називається альтернативною.

Статистичний аналіз якісних ознак — це частина статистики об'єктів нечислової природи. Кількісні ознаки діляться на ознаки, виміряні в шкалах інтервалів, відносин, різниць, абсолютних похибок, а якісні — на ознаки, виміряні в шкалі найменувань і порядкової шкали. Методи обробки даних повинні бути узгоджені зі шкалами, в яких виміряні розглянуті ознаки.

Метою дослідження залежності між ознаками є доказ наявності зв'язку між ознаками і вивчення зв'язку між ними. Для доказу наявності зв'язку між двома випадковими величинами X і Y застосовують кореляційний аналіз [2, 4, 5].

Якщо спільний розподіл X і Y є нормальним, то статистичні висновки засновують на вибірковому коефіцієнті лінійної кореляції, в інших випадках використовують коефіцієнти рангової кореляції Кендалла і Спірмена, а для якісних ознак — критерій x_i-квадрат.

Регресійний аналіз застосовують для вивчення функціональної залежності кількісної ознаки. Основне завдання регресійного аналізу полягає в оцінці невідомих параметрів a і b, які задають лінійну залежність y від x. Для вирішення цього завдання застосовують розроблений К. Гаусом метод найменших квадратів, тобто знаходять оцінки невідомих параметрів моделі a і b з умови мінімізації суми квадратів по змінних a і b [2, 4–5].

Теорія регресійного аналізу описана і розрахункові формули дані в спеціальній літературі [2, 4–7]. У цій теорії розроблені методи точкового та інтервального оцінювання параметрів, які задають функціональну залежність, а також непараметричні методи оцінювання цієї залежності, методи перевірки різних гіпотез, пов'язаних з регресійної залежності. Вибір планів експерименту, тобто точок x_i, у яких будуть проводитися експерименти зі спостереження y_i — предмет теорії планування експерименту [18].

Дисперсійний аналіз застосовують для вивчення впливу якісних ознак на кількісну змінну [4–7].

Постановка та формалізація задачі

Багатофакторна лінійна регресійна модель може бути записана у такому вигляді:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_p + \varepsilon,$$

де y — залежна змінна; $x_1, x_2, \ldots, x_p$ — незалежні змінні (або фактори); $\beta_0, \beta_1, \ldots, \beta_p$ — невідомі параметри, які потрібно оцінити; ε — випадкова величина.

Приклад. На основі статистичних даних показника y та факторів x_1, x_2 та x_3 провести аналіз та побудувати математичну модель залежності y від x_1, x_2 та x_3, попередньо дослідивши її на наявність мультиколінеарності. Вихідні дані приведені у табл. 1.

Таблиця 1

ВИХІДНІ ДАНІ ЗАДАЧІ

Фондовіддача, x_1, тис. грн	Продуктивність праці, x_2, тис. грн	Питомі інвестиції, x_3, тис. грн	Прибуток за місяць, y, тис. грн
12	5	15	40
17	7	18	45
13	6	16	40
14	7	17	43
16	6	20	48
15	5	15	39
14	6	16	42
17	9	18	45
12	5	19	38
18	10	20	48
20	11	22	50
17	10	21	48
18	12	21	49
19	8	20	45
20	9	22	49
22	14	23	52
24	15	24	54
21	13	20	51
25	16	24	55
27	18	25	56

Виклад основного матеріалу. На першому кроці будуємо економетричну модель за замовчуванням у пакеті SPSS, тобто модель, до якої у примусовому порядку включені усі фактори — x_1, x_2 та x_3.

При цьому отримаємо такі результати:

ВВЕДЕННЯ І ДОДАВАННЯ ЗМІННИХ

Модель	Включені змінні	Виключені змінні	Метод
1	X_3, X_2, X_1	–	Примусове включення

Таблиця 2

ЗВЕДЕННЯ ДЛЯ МОДЕЛІ 1

Модель	H	R-квадрат	Скоректований R-квадрат	Стд. похибка	Зміна статистик					Дурбін — Уотсон
					Зміна R квадрат	Зміна F	ст.св.1	ст.св.2	Знч. Зміна F	
1	,963[a]	,927	,913	1,58572	,927	67,649	3	16	,000	2,404

Коефіцієнти										
Модель		Нестандартизовані коефіцієнти		Стандартизовані коефіцієнти	t	Знч.	95,0 %% довірчий інтервал для B		Статистики колінеарності	
		B	Стд. похибка	Бета			Нижня границя	Верхня границя	Толерантність	КРД
1	(Константа)	23,309	3,548		6,569	,000	15,787	30,831		
	X1	,483	,279	,382	1,730	,103	-,109	1,074	,094	10,671
	X2	,410	,281	,302	1,460	,164	-,186	1,007	,106	9,393
	X3	,550	,272	,310	2,019	,061	-,027	1,127	,193	5,169

Отже, для отриманої моделі:

- коефіцієнт детермінації $R^2 = 0,927$; це значить, що 92,7 % вихідних даних підпорядковуються лінійній регресії;
- стандартна похибка регресії E = 1,58572. У відсотковому відношенні це складає 3,38 %; оскільки E% < 15 %, то отримана модель добре пояснюється отриманою регресією;
- значення критерію Фішера F = 67,649; порівнюючи його з критичним значенням даного критерію на рівні значимості $\alpha = 0,05$ та степенями свободи $m_1 = 3$ і $m_2 = 16$: $F_{кр} = 3,24$, можна бачити, що отримана модель буде достовірною з ймовірністю 0,95.

У другій звітній таблиці, ми бачимо, що отримана модель має вигляд:

$$y = 23{,}31 + 0{,}48x_1 + 0{,}41x_2 + 0{,}55x_3.$$

В останній колонці звіту коефіцієнтів (Статистика колінеарності), можемо бачити, що значення коефіцієнта близькі до критичного значення 10, а отже, ведуть до мультиколінеарності [2, 5, 10].

Щоб позбутися ефекту мультиколінеарності, на другому кроці будуємо модель покрокової регресії, яка дасть змогу частково усунути мультиколінеарність. Для цього у вікні регресії обираємо модель покрокової регресії і зробимо перерахунок моделі. Отримаємо:

Таблиця 3

ЗВЕДЕННЯ ДЛЯ МОДЕЛІ 2

Модель	H	R-квадрат	Скоректований R-квадрат	Стд. Помилка похибки	зміна статистик					Дурбін-Уотсон
					зміна R квадрат	зміна F	ст.св.1	ст.св.2	Знч. зміни F	
1	,943[a]	,890	,884	1,83326	,890	145,812	1	18	,000	
2	,958[b]	,917	,907	1,63766	,027	5,557	1	17	,031	2,369

Коефіцієнти										
Модель		Нестандартизовані коефіцієнти		Стандартизовані коефіцієнти	t	Знч.	95,0%% довірчий інтервал для B		Статистики колінеарності	
		B	Стд. похибка	Бета			Нижня границя	Верхня границя	Толерантність	КРД
1	(Константа)	25,336	1,828		13,858	,000	21,495	29,177		
	X1	1,192	,099	,943	12,075	,000	,985	1,399	1,000	1,000
2	(Константа)	19,970	2,802		7,128	,000	14,058	25,881		
	X1	,783	,195	,619	4,018	,001	,372	1,193	,205	4,879
	X3	,644	,273	,363	2,357	,031	,068	1,221	,205	4,879

На даному етапі ми отримали дві найкращі моделі:

- $y = 25{,}34 + 1{,}19x_1$
- $y = 19{,}97 + 0{,}783x_1 + 0{,}64x_3$

При цьому для подальшого використання підходить як перша, так і друга моделі.

Модель задачі носить загальний характер, але може бути використана для практичної проблеми, що пов'язана з аналізом та оптимізацією економічних показників підприємства. Подібні задачі, але інший підхід до розв'язання, були розглянуті в різних літературних джерелах [4, 5, 10].

Висновки. Отже, використання статистичного аналізу для моделювання економічних показників підприємства є ефективним інструментом. У разі застосування відповідного програмного забезпечення (MS Excel, GPSS World, AnyLogic і та ін.) можна швидко отримати точні показники діяльності підприємства та зробити відповідні висновки про його функціонування на даний момент, розрахувати прогнозовані значення на майбутнє. Проте при цьому слід враховувати специфіку задачі та особливості методів застосування.

Бібліографічні посилання

1. Strutz, T. (2016). Data Fitting and Uncertainty (A practical introduction to weighted least squares and beyond). Springer Vieweg.
2. Карташов М. В. Імовірність, процеси, статистика. Київ: ВПЦ Київський університет, 2007. 504 с.
3. Jaynes, E.T. (1957). "Information Theory and Statistical Mechanics". Phys. Rev. 106 (4): 620. Bibcode:1957PhRv.106.620J. doi:10.1103/ physrev.106.620.
4. Єріна А.М. Статистичне моделювання та прогнозування: навч. посібник. Київ: КНЕУ, 2001. 170 с.
5. Greene, William (2012). Econometric Analysis (7th ed.). Pearson Education. pp. 34, 41–42. ISBN 9780273753568.
6. Опря А.Т. Статистика (модульний варіант з програмованою формою контролю знань): навч. посібник. Київ: Центр учбової літератури, 2012. 448 с.
7. Wooldridge, Jeffrey (2013). Introductory Econometrics, A modern approach. South-Western, Cengage learning. ISBN 978-1-111-53104-1.
8. Burnham, K. P. and Anderson D. R. (2002) Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach, Second Edition (Springer Science, New York) ISBN 978-0-387-95364-9.
9. Wolberg, J. (2005). Data Analysis Using the Method of Least Squares: Extracting the Most Information from Experiments. Berlin: Springer. ISBN 978-3-540-25674-8.
10. Колєчкіна Л.М., Литвиненко Ю.О. Економетрія: навч.-метод. посіб. для самост. вивч. дисц. Полтава: ПУЕТ, 2015. 157 с.

Статтю подано до редакції 28.11.2022