УКРАЇНСЬКИЙ ДЕРЖАВНИЙ УНІВЕРСИТЕТ
ЗАЛІЗНИЧНОГО ТРАНСПОРТУ

ЕКОНОМІЧНИЙ ФАКУЛЬТЕТ

Кафедра фінансів, обліку і аудиту

О. М. Тройнікова

ЕКОНОМЕТРИКА

Конспект лекцій

Харків 2026

Тройнікова О. М. Економетрика: Конспект лекцій. – 2-ге вид. – Харків: УкрДУЗТ, 2026. – 60 с.

Конспект лекцій з освітньої компоненти «Економетрика» для здобувачів вищої освіти першого (бакалаврського) рівня спеціальності D2 «Фінанси, банківська справа, страхування та фондовий ринок» усіх форм здобуття освіти.

У конспекті лекцій наведено основні етапи проведення економічного аналізу. Надано методологію економіко-математичного моделювання з використанням економетричних методів.

Конспект лекцій має на меті сформувати у здобувачів компетентності згідно з ОПП спеціальностей здобувачів вищої освіти.

Конспект лекцій можна використовувати як теоретичне джерело для написання розділу кваліфікаційної роботи про фінансовий аналіз і аналіз діяльності підприємства, використання моделей для аналізу та прогнозування економічних явищ.

Іл. 1, бібліогр.: 11 назв.

Конспект лекцій розглянуто і рекомендовано до друку на засіданні кафедри фінансів, обліку і аудиту 9 лютого 2026 р., протокол № 6.

Рецензент

доц. О. Д. Стешенко

ЗМІСТ

Вступ 4
Тематичний план 5
Тема1. Основні положення курсу «Економетрика». Роль і етапи економетричного дослідження. Метод експертних оцінок 6
1.1 Еволюція економетрики як науки 6
1.2 Економетричні моделі. Етапи проведення економетричного аналізу 9
1.3 Метод експертних оцінок 13
Тема 2. Специфікація економетричної моделі 21
2.1 Вибір виду функції 21
2.2 Основні математичні функції, що найчастіше використовують для економетричного аналізу 23
Тема 3. Проста вибіркова лінійна регресія 27
3.1 Загальне поняття про вибіркову лінійну регресію 27
3.2 Визначення параметрів лінійної регресії 28
3.3 Визначення якості моделі 29
3.4 Перевірка простої регресійної моделі на адекватність 29
Тема 4. Загальна проста лінійна регресія 32
4.1 Загальна проста лінійна модель 33
4.2 Перевірка на значущість параметрів загальної простої лінійної регресії. Інтервали довіри 33
4.3 Прогнозування значення результативного показника загальної простої лінійної регресії 35
Тема 5. Загальна лінійна багатофакторна регресія 37
5.1 Класична лінійна багатофакторна модель 38
5.2 Етапи побудови багатофакторної регресійної моделі 42
5.3 Визначення невідомих параметрів багатофакторної регресії за методом найменших квадратів (МНК) 44
5.4 Аналіз багатофакторної економетричної моделі 45
Тема 6. Мультиколінеарність 50
6.1 Поняття про мультиколінеарність і її вплив на оцінювання параметрів моделі 50
6.2 Тестування наявності мультиколінеарності 52
Тема 7. Нелінійна регресія 55
7.1 Рівняння нелінійної парної регресії 55
7.2 Найпростіші перетворення нелінійних моделей у лінійні 57
Список літератури 60

Вступ

Основним завданням конспекту лекцій з освітньої компоненти «Економетрика» є вивчення методології економіко-математичного моделювання з використанням економетричних методів і ознайомлення з прикладами практичного використання прикладних економетричних моделей у фінансовій діяльності для проведення фінансового аналізу та аналізу господарчої діяльності підприємства, використання моделей для прогнозування та аналізу економічних явищ. Конспект лекцій має на меті сформувати у здобувачів компетентності згідно з ОПП спеціальностей здобувачів вищої освіти.

У конспекті лекцій наведено основні етапи проведення економічного аналізу. Надано методологію економіко-математичного моделювання з використанням економетричних методів.

Конспект лекцій можна використовувати як теоретичне джерело для написання розділу кваліфікаційної роботи про фінансовий аналіз і аналіз діяльності підприємства, використання моделей для аналізу та прогнозування економічних явищ.

Курс починається із загальної характеристики науки «Економетрика», історії її виникнення, сутності та економічної складової в цій дисципліні, методів економетричного аналізу та моделювання.

Конспект лекцій складається з семі тем, кожна тема містить глосарій основних понять і контрольні запитання.

Тематичний план

Тема	Годин
1 Основні положення курсу «Економетрика». Роль і етапи економетричного дослідження. Метод експертних оцінок	2
2 Специфікація економетричної моделі	2
3 Проста вибіркова лінійна регресія	
4 Загальна проста лінійна регресія	2
5 Загальна лінійна багатофакторна регресія	3
6 Мультиколінеарність	2
7 Нелінійна регресія	2

Тема 1. Основні положення курсу «Економетрика». Роль і етапи економетричного дослідження. Метод експертних оцінок

План

1.1 Еволюція економетрики як науки.

1.2 Економетричні моделі. Етапи проведення економетричного аналізу.

1.3 Метод експертних оцінок.

Ключові слова: економетрика, економетрична модель, економетричний аналіз, метод експертних оцінок, коефіцієнт конкордації, критерій Пірсона.

1.1 Еволюція економетрики як науки

Економетрика є економічною дисципліною з широким використанням методів прикладної математики, теорії ймовірностей, математичної статистики та економічної теорії.

В економетричних дослідженнях використовують економічну теорію, математичні методи і статистику. Тому економетричні моделі часто ототожнюють з економіко-статистичними.

Формування її в наукову теорію можна датувати груднем 1930 року, коли з ініціативи Ірвін Фішера, Равнер Фріша, Яна Тінберґена, Й. Шумпетера, О. Андерсона та інших учених було засноване Міжнародне економетричне товариство, яке визначило себе як «Міжнародне товариство для розвитку економічної теорії і її зв'язку зі статистикою та математикою». В установчих документах товариства міститься і перше визначення сутності економетрики – «математична статистика для економістів» [2, 10].

Пізніше відомий економетрист Цві Гріллiхес визначив сутність економетрики більш емоційно: «Економетрика є одночасно нашим

телескопом і нашим мікроскопом для вивчення оточуючого нас економічного світу».

Термін «економетрика» уперше запропонував львівський учений Петро Чомпа у книзі «Нариси економетрики і природної теорії бухгалтерії, яка ґрунтується на політичній економії», що вийшла друком у Львові в 1910 році. Цим терміном користувалися у своїх працях з економетрики вчені у 10-20-х роках минулого століття. Уже тоді з'явилися спроби прогнозувати процеси економіки, поводження товарного і грошового ринків, циклічного розвитку виробництва товарів і послуг на основі так званих «барометрів розвитку». Потужний поштовх отримали економетричні дослідження за часів великої депресії в економіці США на початку 1930-х років. Саме тоді була розроблена класична економетрична модель – виробнича функція Кобба-Дугласа [4].

Значний внесок у становлення економетрики зробили дослідження циклічності економіки - форма руху національної економіки та світового господарства в цілому, що припускає зміну революційних і еволюційних стадій розвитку економіки, економічного прогресу. Економічні процеси розвиваються нерівномірно, циклічно, послідовно перебуваючи у фазах стабільного розвитку, кризи, депресії, пожвавлення, піднесення.

Економіка періодично переходила і переходить від одного відносно рівноважного, сталого стану до іншого рівноважного стану через перехідний кризовий період, для якого характерні посилення нестійкості, спад виробництва, хаотичні зміни, перебудова структури економіки. При цьому економічна система очищується від застарілих елементів і збагачується додатковим вмістом за рахунок притаманній їй спадковій мінливості, яка сприяє адаптації до нових умов розвитку суспільства.

Важливим етапом формування економетрики стало розроблення економічних барометрів. На початку XX століття в деяких країнах були спроби скласти так звані «барометри розвитку». Найвідоміший із них

«гарвардський барометр», за допомогою якого в 1920-ті роки намагалися передбачити поведінку товарного і грошового ринків.

Успіх використання гарвардського барометра викликав появу багатьох аналогічних барометрів в інших країнах. Проте приблизно з 1925 році він втратив свою чутливість. Його крах пояснюють появою потужного регулюючого фактора в економіці США.

Починаючи з 1930-х років відомі економісти Я. Тінберґен, Л. Клейн, Р. Стоун та інші розробили моделі економіки, які описували статистичні зв'язки виробництва, кінцевого індивідуального і державного попиту, цін, податків, зовнішньої торгівлі, пропозиції робочої сили, накопичення та зношування капіталу. Такі моделі складалися вже з багатьох рівнянь, у зв'язку з чим стало значно складніше оцінювати невідомі параметри. А це призвело до необхідності використання нового математичного апарату і розширило можливості практичного використання економетрики.

У 1933 році Р. Фріш заснував журнал «Економетрика», який і зараз має велике значення для розвитку економетрики. А вже в 1941 році з'являється перший підручник із нової наукової дисципліни, написаний Я. Тінберґеном.

У 1969 році Фріш і Тінберґен отримали Нобелівську премію з економіки, як вказано в офіційному повідомленні нобелівського комітету «за створення і застосування динамічних моделей для аналізу економічних процесів».

До 1970-х років економетрику сприймали як емпіричну оцінку моделей, створених у рамках економічної теорії.

Важливою подією для розвитку економетрики стала поява комп'ютерів. Завдяки їм потужно став розвиватися статистичний аналіз часових рядів. Автори Г. Бокс і Г. Дженкінс створили ARIMA-модель (інтегрована модель авторегресії (ковзної середньої) — модель і методологія аналізу часових рядів) [11].

Стимулював економетричні дослідження і бурхливий розвиток фінансових ринків і похідних інструментів.

Як галузь економічної науки економетрика пройшла тривалий шлях формування найголовніших ознак самостійної дисципліни – предмета і методології. Внесок у формування та розвиток економетричних методів моделювання зробили лауреати Нобелівської премії з економіки Р. Фріш і Я. Тінберґен (1969), Л. Клейн (1980), Т. Хаавельмо (1989), Дж. Хекман (2000) [4].

1.2 Економетричні моделі. Етапи проведення економетричного аналізу

Економетрика складається з двох частин:

1) економетричні методи;
2) економетричні моделі економічних процесів і явищ.

Отже, ***економетрика*** **–** це наука, що вивчає кількісні закономірності та взаємозв'язки економічних об'єктів і процесів за допомогою математико-статистичних методів і моделей.

Економетричні методи можна умовно поділити на чотири групи. До першої групи входять методи оцінювання параметрів класичної економетричної моделі за методом найменших квадратів, їх верифікація. До другої групи належать методи оцінювання параметрів узагальненої моделі, коли порушені деякі передумови використання методу найменших квадратів. До третьої групи входять методи оцінювання параметрів динамічних економетричних моделей, їх верифікація. Четверта група охоплює методи оцінювання параметрів економетричних моделей, які побудовані на основі системи одночасних структурних рівнянь [2, 4].

Економетричні моделі кількісно описують зв'язок між вхідними показниками економічної системи, тобто факторами X і результативним

показником Y. У загальному вигляді економетричну модель можна записати так:

$$Y = f(X,u), \tag{1.1}$$

де X — вхідні економічні показники;

u — випадкова складова.

Показники **X** найчастіше бувають детермінованими. Адитивна складова **u** є випадковою змінною, яка включає помилки розрахунків, вимірювань і ті фактори, які невраховані в моделі, але вплив яких неможливо викреслити.

Існує багато відомих нам моделей. Наприклад, модель споживання, яка має вигляд

$$C_j = f(r_j), \tag{1.2}$$

де C – споживання;

r – дохід.

До типових економіко-математичних моделей, які на сьогодні розроблено і вивчено в економетриці, належать виробничі функції, функції попиту різних груп споживачів і цільові функції переваги споживачів, статистичні та динамічні міжгалузеві моделі виробництва, розподілу і споживання продукції, моделі загальної економічної рівноваги.

Крім того, сучасні економетричні методи широко використовують для порівняння ефективності різних економічних гіпотез і послідовного уточнення їх.

Етапи проведення економетричного аналізу [1-4]:

1 Формування теорії чи гіпотези.

2 Розроблення економетричної моделі для перевірки цієї моделі.

3 Формування вихідної інформації.

4 Оцінювання параметрів вибраної моделі.

5 Оцінювання якості моделі.

6 Оцінювання адекватності моделі. Перевірка моделі, статистичні висновки.

7 Верифікація моделі.

Але доводиться аналізувати не тільки результативні показники, на які впливають фактори, що мають кількісне вимірювання. У таких випадках для вилучення факторів, які мають найбільший вплив на аналізований результативний показник, використовують метод експертних оцінок [1, 3].

На першому етапі необхідно чітко сформулювати завдання дослідження, ознайомитися з економічною теорією, побудувати гіпотези взаємозв'язку (припущення).

Задача встановлення взаємозв'язків між економічними величинами є першою принциповою задачею кожного, хто вивчає економіку.

Розрізняють два види взаємозв'язків:

1) функціональний – кожному значенню фактора Хі відповідає одне конкретне значення результуючого показника У;

2) стохастичний (кореляційний) – кожному значенню фактора Хі може відповідати кілька різних значень показника Y, тобто існує статистичний розподіл.

На другому етапі необхідно сформулювати теоретичні уявлення і прийняти гіпотези у вигляді математичних рівнянь. Для цього використовують усі ті форми функцій, які можуть бути застосовані для вивчення взаємозв'язків. Ці рівняння встановлюють зв'язки між основними змінними, припускаючи, що всі інші змінні є випадковими. Це етап специфікації моделі. Зазвичай вибір моделі починається з лінійної. Крім того, існують нелінійні моделі опису змінних.

Третій етап пов'язаний зі збиранням і підготовкою економічної інформації, тобто формуванням масивів вихідної інформації відповідно до мети і завдань дослідження.

Усі спостереження мають бути випадковими та незалежними. Масив вихідних даних повинен мати достатній обсяг спостережень – кількість спостережень має щонайменше у шість-вісім разів перевищувати кількість факторів, які включені в цю модель.

На четвертому етапі оцінюють параметри економетричної моделі за методом найменших квадратів і розраховують інтервали довіри, тобто інтервали, до яких потраплятиме обчислювана величина із заданим рівнем імовірності.

П'ятий етап передбачає оцінювання якості побудованої моделі (за допомогою певних тестів і критеріїв) і аналіз залишків, щоб відповісти на питання, чи не вступає специфікація моделі в протиріччя з передумовами «класичної» моделі лінійної регресії. Якщо деякі передумови моделі не виконувані, то для продовження аналізу треба змінити специфікацію моделі (тобто повернутися до другого етапу роботи) або застосовувати інші методи оцінювання параметрів (тобто повернутися до четвертого етапу роботи).

На шостому етапі оцінюють адекватність моделі (перевірка статистичної значущості рівняння залежності в цілому і його окремих параметрів; перевірка відповідності формальних властивостей оцінок завданням дослідження).

На сьомому етапі аналізують вірогідність (достовірність) моделі та розглядають методи прогнозу для розрахованої моделі. Дуже важливими на цій стадії є оцінки дисперсії помилок в кожній моделі. Ці оцінки відіграють вирішальну роль для з'ясування якості економетричних зв'язків, вони необхідні для визначення надійності розрахованих параметрів і застосування моделей із метою прогнозування.

1.3 Метод експертних оцінок

Методи експертних оцінок – це спосіб прогнозування та оцінювання майбутніх результатів дій на основі досліджень, аналізу та прогнозів фахівців. Застосовуючи метод експертних оцінок, опитують спеціальну групу експертів із метою визначення певних змінних величин, необхідних для оцінювання досліджуваного питання. До складу експертів слід включати людей із різними типами мислення – образне і словесно логічне, що сприяє успішному розв'язанню проблеми. Залучені експерти можуть висловити свою думку щодо найкращих способів мобілізації резервів, залучення інвестицій, строків досягнення поставлених завдань, критеріїв відбору оптимальних варіантів рішення тощо.

Необхідною умовою ефективного застосування методів експертної оцінки є достатня обізнаність експерта з досліджуваної проблеми, високий рівень ерудиції, здатність його давати чіткі вичерпні відповіді, до того ж експромтом. Крім того, експерт не повинен бути зацікавленим у тому чи іншому варіанті розв’язання поставленої перед ним проблеми. Експертів вибирають за ознакою їхнього формального професійного статусу – посади, наукового ступеня, стажу роботи та ін. Такий підбір сприяє тому, що серед експертів є високопрофесійні, із великим практичним досвідом у певній галузі спеціалісти [1-4, 8].

Отже, методи експертної оцінки потребують ретельної підготовки експертів, робота яких полягає:

а) у чіткому визначенні мети і завдань, а в деяких випадках об'єднанні та систематизації висновків;

б) виборі достатньо компетентних незалежних експертів у сфері відповідних об'єктів;

в) обговоренні питання в групі експертів чи виключенні безпосереднього спілкування між ними;

г) наданні учасникам експертизи на кожному наступному етапі результатів і висновків попереднього етапу.

У результаті можна зробити певні висновки, які поділяють більшість експертів; вибрати оптимально прийнятні методи обробки висновків експертів; точно формулювати підсумкові висновки в експертній роботі.

Як уже було зазначено, фактично метод експертних оцінок є методом прогнозування, основним критерієм якого є досягнення згоди серед усіх членів експертної групи.

Експерти, обізнані у взаємопов'язаних сферах діяльності, детально відповідають на запитання, наприклад, анкети, пов'язаної з досліджуваною проблемою. Кожен із них фіксує свою думку про проблему, а потім повідомляє свою відповідь колегам. У випадку розбіжності його прогнозу з думкою інших експерт зобов'язаний пояснити причину такої невідповідності.

Далі процедуру повторюють до тих пір, поки думки експертів не збіжаться. При цьому потрібно дотримуватися анонімності, що допомагає уникнути можливості групових роздумів над проблемною ситуацією.

Завдяки застосуванню експертних оцінок отримують два види інформації, на підставі якої вирішують два види завдань різної значущості і на різних рівнях управління:

1 Інформація про поодинокі причинно-наслідкові зв'язки в конкретних умовах місця і часу. Здебільшого цю інформацію одержують у результаті опитування керівників виробничих підрозділів підприємства (бригадири, керівник відділення, начальник цеху) і працівників. Вона призначена для пошуку напрямів підвищення ефективності виробництва і реалізації продукції після визначення причин непродуктивного використання ресурсів і формування дієвих заходів для їх усунення.

2 Інформація про типові взаємозв'язки досліджуваних економічних явищ і процесів. Таку інформацію здатні надати тільки експерти високого

класу, професіонали, глибоко обізнані з сутністю і закономірностями прояву вказаних

явищ за різних умов господарювання.

Основними завданнями, які найчастіше вирішують на практиці на основі отриманої від експертів інформації, є:

а) ранжування (впорядкування, розміщення за зростанням чи спаданням) факторів і відповідних показників, що їх характеризують, за їхньою значущістю в розвитку досліджуваного явища, процесу;

б) ранжування підприємств чи їхніх структурних виробничих підрозділів (бригад, підрозділів, ділянок) за рейтингом, в основу якого покладено сукупність різних показників, що характеризують результати фінансово-господарської діяльності чи окремих її видів (фінансовий стан, рентабельність, платоспроможність тощо);

в) переднє оцінювання виконання плану за певним показником.

Цільовий аналіз, що ґрунтований на результатах експертних оцінок, здійснюють у декілька етапів:

1) визначення мети дослідження;

2) визначення необхідного кількісного та якісного складу групи експертів;

3) створення групи експертів;

4) визначення способу опитування;

5) складання програми обстеження і анкети (листка) опитування;

6) опитування;

7) зведення, групування та аналіз отриманої від експертів інформації;

8) узагальнення результатів експертизи і розроблення варіантів рішень.

Загалом усі експертні методи поділяють на дві групи:

– індивідуальні;

– колективні.

Індивідуальні експертні методи – це використання думок експертів, які сформульовані особисто кожним із них самостійно без урахування думок інших експертів. До індивідуальних експертних методів належать інтерв'ю, анкетування.

Сутність методу інтерв'ю полягає в організації співбесіди аналітика з експертом, під час якої експерт дає відповіді на запитання аналітика про фактори впливу на досліджуваний об'єкт, очікувані результати господарювання, невикористані резерви, шляхи виходу з кризи, напрями підвищення ефективності виробництва тощо.

Метод анкетування (аналітичного експертного оцінювання) полягає в наданні експертом письмових відповідей на запитання анкети. Проте цей метод має певні недоліки, зокрема експерт може не зрозуміти запитання анкети, проявити суб'єктивізм, небажання критикувати керівництво і залишати свою письмову відповідь тощо.

Основними перевагами індивідуальних методів експертних оцінок є простота організації обстеження, зрозумілість, урахування і використання набутих знань і досвіду кожного експерта. Обмеженням застосування цих методів є недостатні знання, інформація експертів із суміжних сфер діяльності. Виходячи з цього, більшого поширення на практиці набули колективні експертні методи.

Колективні експертні методи – це методи, які забезпечують формування єдиної спільної думки в результаті взаємодії залучених фахівців-експертів.

До розповсюджених колективних методів експертної оцінки відносять:

– метод комісії;
– метод відстороненого оцінювання;
– метод Дельфі;
– метод конференції ідей та ін.

Метод комісії полягає у виробленні експертами кращого варіанта досягнення поставленої мети з урахуванням усіх висловлених на нараді пропозицій, ідей.

Позитивною ознакою цього методу є можливість залучення для експертизи фахівців із широким діапазоном знань із суміжних галузей науки та практики. Негативним є можливий суб'єктивізм, наявні стереотипи мислення, що склалися в експертів, їхня схильність до компромісу.

Метод відстороненого оцінювання полягає у виборі оптимального незалежного рішення з тих, що висловили експерти на нараді. Робота наради поділена на дві частини: висунення ідей і їх критичний аналіз.

Метод Дельфі – один із методів колективної експертної оцінки, який передбачає проведення експертного опитування серед групи спеціалістів у кілька турів (найчастіше три-чотири тури) для вибору найкращого із рішень. Метод Дельфі отримав свою назву від назви містечка Дельфі у Стародавній Греції, у якому жили оракули-провидці при храмі бога Аполлона. Слово головного оракула не підлягало сумніву, його приймали за істину.

Метод конференції ідей подібний до мозкового штурму, але відрізняється від нього темпом проведення нарад і дозволеною короткою доброзичливою критикою ідей у формі реплік і коментарів. При цьому стимулюють поєднання кількох пропозицій, фантазування, що сприяє підвищенню якості ідей. Усі висунуті ідеї занотовують у протоколі без зазначення їхніх авторів. До складу учасників конференції ідей включають не лише висококваліфікованих фахівців, а й новачків, неспеціалістів – незаангажованих і здатних висувати свіжі, нові, неординарні підходи.

Отже, методи експертної оцінки потребують ретельної підготовки експертів, робота яких полягає в чіткому визначенні мети і завдань, а в деяких випадках об'єднанні та систематизації висновків; виборі достатньо компетентних незалежних експертів у сфері відповідних об'єктів; обговоренні питання в групі експертів чи виключенні безпосереднього

спілкування між ними; наданні учасникам експертизи на кожному наступному етапі результатів і висновків попереднього етапу. Це дає змогу зробити певні висновки, які поділяють більшість експертів.

Отже, **суть експертної процедури** полягає в тому, що якомога більшій кількості експертів пропонують оцінити вплив кожного з установлених факторів на досліджуваний економічний показник, розташовуючи фактори в порядку зменшення ступеня впливу. При цьому мінімальна кількість експертів дорівнює кількість ранжованих факторів +2. Зазвичай найбільш впливовому фактору присвоюють ранг 1 і далі, зі зменшенням впливу, – ранги 2, 3 і т. ін.

Порядок проведення експертизи:

1 Постановка завдання або висунення гіпотези.

2 Вибір експертів із осіб, компетентних у досліджуваній галузі.

3 Ранжирування факторів експертами.

4 Обробка інформації, отриманої від експертів.

5 Перевірка узгодженості і вірогідності експертних оцінок.

6 Використання результатів дослідження.

Із ранжируванням деякі фактори можуть одержати однаковий ранг, тому додатково для таких факторів експерту пропонують за результатами повторної експертизи уточнити місця, які одержать ці фактори в разу загального упорядкування всіх факторів зі зменшенням рангів, тобто зменшенням ступеня їхнього впливу на залежну змінну. Залежно від рангу і місця, які вони займають, фактори будуть мати ту або іншу вагу. У розрахунках приймають, що вага місця дорівнює його номеру, але для факторів, які мають однаковий ранг, вага дорівнює середній арифметичній вазі відповідних місць.

Для перевірки узгодженості думок експертів визначають **коефіцієнт конкордації:**

$$W = \frac{S}{\frac{1}{12}m^2(n^3 - n) - \frac{m}{12}\sum_{j=1}^{m} T_j}, \qquad (1.1)$$

де **S** – сума квадратів відхилень;

m – кількість експертів;

n – кількість факторів;

T_j – показник, що враховує збіг рангів,

$$T_j = \sum_{k=1}^{n}(t_k^3 - t_k),$$

де t_k – кількість повторень рангу **k** із ранжируванням факторів **j**-м експертом.

Якщо збігу рангів не було, то $t_k = 1$ і $T_j = 0$.

Коефіцієнт конкордації змінюється в межах $0 \le W \le 1$.

Узгодженість вважають задовільною, якщо $W \ge 0{,}5$, якщо ж $W \ge 0{,}7$, то узгодженість вважають доброю. Із повним збігом думок експертів $W = 1$.

Вірогідність коефіцієнта конкордації перевіряють за критерієм Пірсона, розрахункове значення якого визначають за формулою

$$\chi^2_{розр} = \frac{S}{\frac{1}{12} \cdot m \cdot n \cdot (n+1) - \frac{1}{(n-1)} \cdot \frac{1}{12} \cdot \sum_{j=1}^{m} T_j}.$$

Розрахункове значення $\chi^2_{розр}$ порівнюють із табличним $\chi^2_{табл}$.

За заданих рівні значущості і кількості ступенів свободи $\gamma = n - 1$, де n – кількість факторів, вибираємо табличне значення критерію Пірсона.

Якщо розрахункове значення $\chi^2_{розр}$ перевищує табличне $\chi^2_{табл}$, то гіпотезу про наявність згоди думок семи експертів із ранжируванням п’яти факторів приймають, тобто можна стверджувати, що має місце невипадкова узгодженість думок експертів, тому, на думку вибраних експертів, можна

вилучати фактори, що найбільше впливають на результативний показник. Для вилучення найбільш впливових факторів будують діаграму зменшення впливу факторів.

Глосарій

Екзогенні змінні – це незалежні змінні (фактори) від внутрішньої структури економічного явища, їхні величини задані поза моделлю.

Економетрична модель – це функція чи система функцій, що описує кореляційно-регресійний зв'язок між економічними показниками, причому залежно від причинних зв'язків між ними один чи кілька із цих показників розглядають як залежні змінні, а інші – як незалежні.

Ендогенні змінні – залежні змінні (результативні показники), які визначені внутрішньою структурою досліджуваного економічногоявища, тобто їхні величини обчислюють на основі економетричної моделі.

Коефіцієнт конкордації – це показник, який використовують для перевірки узгодженості думок експертів.

Критерій Пірсона – це показник, за яким перевіряють вірогідність коефіцієнта конкордації.

Методи експертних оцінок – це спосіб прогнозування та оцінювання майбутніх результатів дій на основі досліджень, аналізу та прогнозів фахівців. Застосовуючи метод експертних оцінок, опитують спеціальну групу експертів із метою визначення певних змінних величин, необхідних для оцінювання досліджуваного питання.

Рівняння регресії – характеризує кореляційну залежність ендогенної змінної від екзогенних змінних.

Контрольні запитання

1 Охарактеризуйте етапи проведення економетричного аналізу.

2 Дайте загальне визначення методам експертної оцінки.

3 Які методи належать до індивідуальних методів?

4 Які методи належать до колективних методів?

5 Які недоліки мають індивідуальні методи експертної оцінки?

6 Назвіть етапи перевірки узгодженості думок експертів для аналізу результатів досліджень за методом експертних оцінок.

7 Додатковий етап перевірки за значенням критерію Пірсона. У яких випадках необхідно його застосовувати?

Тема 2. Специфікація економетричної моделі

План

2.1 Вибір виду функції.

2.2 Основні математичні функції, що найчастіше використовують для економетричного аналізу.

Ключові слова: специфікація моделі, види математичних функцій, проста регресія, багатофакторна регресія, параметри регресії, метод найменших квадратів.

2.1 Вибір виду функції

Використовуючи всі ті форми функцій, які можуть бути застосовані для вивчення взаємозв'язків, необхідно сформулювати теоретичні уявлення і прийняті гіпотези у вигляді математичних рівнянь. Ці рівняння

встановлюють зв'язки між основними визначальними змінними за припущення, що всі інші змінні є випадковими.

Економетрична модель базована на єдності двох аспектів — теоретичного, якісного аналізу взаємозв'язків та емпіричної інформації. Теоретична інформація знаходить своє відображення в специфікації моделі.

Специфікація моделі — це аналітична форма економетричної моделі. Вона складається з певного виду функції чи функцій, що використовують для побудови моделей, має ймовірнісні характеристики, які притаманні стохастичним залишкам моделі, тобто обґрунтування лінійної або певного типу криволінійної форми регресії. Перевагу слід надавати теоретичному аналізу форми залежностей, використовуючи як допоміжні графічні й аналітичні засоби обґрунтування.

Із досвіду економетричних досліджень, а також на підставі якісного теоретичного аналізу взаємозв'язків між економічними показниками можна навести клас функцій, які можуть описувати ці взаємозв'язки:

1) лінійна функція

$$y = a_0 + a_1 x_1 + a_2 x_2 + \ldots + a_m x_m; \tag{2.1}$$

2) степенева функція

$$y = a_0 x_1^{a_1} x_2^{a_2} \ldots x_m^{a_m} \rightarrow \ln y = \ln a_0 + \\ + a_1 \ln x_1 + a_2 \ln x_2 + \ldots + a_m \ln x_m; \tag{2.2}$$

3) гіпербола

$$y = a_0 + \frac{a_1}{x_1} + \frac{a_2}{x_2} + \ldots + \frac{a_m}{x_m} \rightarrow y = a_0 + \\ + a_1 z_1 + a_2 z_2 + \ldots + a_m z_m, \tag{2.3}$$

де

$$z_j = \frac{1}{x_j}; \tag{2.4}$$

4) квадратична функція

$$y = a_0 + a_1 x_1^2 + a_2 x_2^2 + ... + a_m x_m^2 \rightarrow y = a_0 + \\ + a\ t_1 + a_2 t_2 + ... + a_m t_m, \tag{2.5}$$

де

$$t_j = x_j^2 . \tag{2.6}$$

У цих функціях **y** - залежна (пояснювана) змінна; $\mathbf{x_j}$, **j = l, m** – незалежні змінні (фактори); $\mathbf{a_j}$ **–** параметри функцій, якщо j = 0, m.

Серед наведених щойно видів функцій три останні є нелінійними. Але за допомогою перетворення залежної і незалежних змінних ці функції можна звести до лінійного вигляду.

2.2 Основні математичні функції, що найчастіше використовують для економетричного аналізу

Лінійні функції найпоширеніші в економетричному моделюванні, тому обґрунтування економетричних методів розглянемо на базі лінійних моделей.

Розуміючи, що вибір аналітичної форми економетричної моделі не можна розглядати без конкретного переліку незалежних змінних, специфікація моделі передбачає відбір факторів для економетричного дослідження.

При цьому в процесі такого дослідження можна кілька разів повертатися до етапу специфікації моделі, уточнюючи перелік незалежних змінних і вид функції, що застосована. Адже коли вид функції та її складові не відповідають реальним процесам, то йдеться про помилки специфікації.

Помилки специфікації моделі можуть бути трьох видів:

- ігнорування з побудовою економетричної моделі істотної пояснюючої змінної;

- введення в модель незалежної змінної, яка не є істотною для вимірюваного зв'язку, тобто зайва;
- використання невідповідних математичних форм залежності;
- використання неточних або агрегованих статистичних даних.

В економетричну модель включена обмежена кількість найбільш важливих факторів і вона набуває вигляду

$$y^{\wedge} = f\,(x1, x2\ , ..., xm)\ , \qquad (2.7)$$

де $\hat{y}$ – математичне сподівання, або найбільш імовірне, середнє з можливих значень y.

Таке спрощення дійсності зумовлює виникнення помилки апроксимації, або прогнозу, бо в такому разі

$$y = f\,(x1, x2\ , ..., xm\ ) + e\ , \qquad (2.8)$$

де e – випадкова складова y, яку пояснюють впливом інших (неврахованих іневідомих) факторів.

Це означає, що економетричні моделі є стохастичними, імовірнісними.

Стохастичні залежності в економетрії називають також кореляційними.

Термін «***кореляція***» (лат. correlatio) ввів до наукової термінології англійський учений Ф. Гальтон у 1877 році. Цей термін буквально означає співвідношення, пропорцію, зв’язок, залежність.

Для визначення залежностей в економетрії використовують такі позначення:

y – залежна (ендогенна) змінна;

x_i – незалежні (екзогенні), пояснюючі змінні, або фактори.

Назва «рівняння регресії» введена англійським антропологом К. Пірсоном на початку XX сторіччя. Отже, предметом економетрії, по-перше, є методологія побудови – ***специфікація*** – рівнянь регресії в економіці.

Рівняння регресії можуть мати лінійну

$$y = a_0 + a_1 x_1 + a_2 x_2 + \ldots + a_m x_m + e , \qquad (2.9)$$

і нелінійну форму, наприклад, степеневу

$$y = a_0 x_1^{a_1} x_2^{a_2} \ldots x_m^{a_m} + \varepsilon , \qquad (2.10)$$

де a_i ($i = 0, m$) – параметри рівняння регресії, або коефіцієнти регресії. Вони визначають вплив змінення факторів x_i на рівень y.

Якщо модель лінійна, то a_i показує, на скільки одиниць змінюється y зі збільшенням x_i на одиницю. У степеневій моделі коефіцієнт регресії a_i є відносною мірою зміни, він визначає, на скільки відсотків зміниться y, якщо x_i збільшиться на один відсоток.

По-друге, предметом економетрії є методи оцінювання параметрів рівнянь регресії.

Оскільки кількість незалежних змінних у дійсності невизначено велика, і істинна аналітична форма залежностей y від x_i невідома, коефіцієнти регресії a_i і величину e визначити абсолютно точно об’єктивно неможливо. Оцінювання цих параметрів моделі зводять до визначення їхніх емпіричних, імовірнісних оцінок, відповідно a_i і e. Тоді, наприклад, рівняння лінійної регресії набуває такого вигляду:

$$y = a_0 + a_1 x_1 + \ldots + a_m x_m + e . \qquad (2.11)$$

Отже, підсумовуючи сказане, предметом економетрії є специфікація (визначення змінних і аналітичної форми) і оцінювання параметрів рівнянь регресії.

Глосарій

Багатофакторна регресія – це залежність між двома та більше змінними.

Випадкова складова – це складова економетричної моделі, яку пояснюють впливом інших неврахованих факторів, помилкою в розрахунках і похибкою у вимірюванні.

Кореляція – це залежність, зв'язок.

Параметри рівняння регресії або коефіцієнти регресії – визначають вплив зміни фактора x_i на рівень y.

Проста регресія – це залежність між двома змінними.

Специфікація моделі – це аналітична форма економетричної моделі, яка складається з певного виду функції чи функцій.

Контрольні запитання

1 Параметри регресійної моделі як показники змін результативного показника зі зміною факторів.

2 Дайте визначення специфікації моделі.

3 Визначте етапи специфікації моделі.

4 Поясніть методи вибору виду економетричної моделі.

5 Поясніть наявність випадкової величини в регресійних моделях.

Тема 3. Проста вибіркова лінійна регресія

План

3.1 Загальне поняття про вибіркову лінійну регресію.

3.2 Визначення параметрів лінійної регресії.

3.3 Визначення якості моделі.

3.4 Перевірка простої лінійної регресійної моделі на адекватність.

Ключові слова: проста вибіркова лінійна регресія, кореляція, коефіцієнт кореляції, адекватність моделі, коефіцієнт детермінації, критерій Фішера.

3.1 Загальне поняття про вибіркову лінійну регресію

Зазвичай перед початком економетричного аналізу, але після формування теорії про залежність, яка нас цікавить, висувають гіпотезу про існування лінійного зв'язку між фактором і результативним показником, тому що саме лінійна функція є базовою, і через математичні перетворення майже кожній функції можна надати лінійного вигляду.

Прості парні лінійні регресійні моделі встановлюють лінійну залежність між двома змінними. При цьому одну зі змінних вважають залежною ***y***, тобто ендогенною змінною, і розглядають як функцію незалежної змінної ***x***, тобто екзогенної.

У загальному вигляді проста вибіркова регресійна модель

$$\mathbf{y = a_0 + a_1 x + u}, \qquad (3.1)$$

де **y** – результативний показник (залежна змінна);

x – фактор (незалежна змінна);

$\mathbf{a_0, a_1}$ – невідомі параметри регресійної моделі;

$\mathbf{u}$ – випадкова величина (помилки або відхилення).

Регресійну модель називають лінійною, якщо вона лінійна за своїми параметрами. Її можна трактувати як пряму на площині, де a_0 – перетин з віссю ординат, тобто параметр a_0 показує, чому дорівнює результативний показник, якщо значення фактора *0*, а a_1 – це параметр, за значенням якого можна оцінити, наскільки зміниться результативний показник зі збільшенням фактора на 1. При цьому за знаком параметра можна визначити тип функції – пряма чи обернена. Так, за додатного значення a_1 можна дійти висновку, що зі збільшенням фактора результативний показник також збільшується, а за від'ємного значення цього параметра за умов збільшення незалежної змінної ендогенна змінна зменшується.

3.2 Визначення параметрів лінійної регресії

Для подальшого аналізу вибраної залежності необхідно оцінити (знайти) невідомі параметри $\mathbf{a_0, a_1}$ цієї моделі. Розглянемо пряму $\tilde{\mathbf{y}} = \mathbf{a_0} + \mathbf{a_1 x}$.

Відхилення, або помилки, іноді називають залишками. Логічно, що пряму треба проводити так, щоб сума квадратів помилок була мінімальною. У цьому і полягає критерій найменших квадратів.

Невідомі параметри $\mathbf{a_0, a_1}$ розраховують за формулами

$$\mathbf{a_1} = \frac{\frac{1}{n}\left(\sum_{i=1}^{n} x_i y_i\right) - \bar{x} \cdot \bar{y}}{\frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{x}^2}, \text{ де } \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i ; \ \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i . \qquad (3.2)$$

При цьому параметр $\mathbf{a_0}$ (перетин)

$$\mathbf{a_0 = \bar{y} - a_1 \bar{x}}. \tag{3.3}$$

3.3 Визначення якості моделі

Критерієм, що характеризує щільність зв'язку між залежною змінною ***y*** і незалежною ***x***, тобто наскільки значним є вплив змінної ***x*** на ***y***, є ***коефіцієнт кореляції***:

$$r_{yx} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 * \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2}}. \tag{3.4}$$

Коефіцієнт кореляції є відносною мірою зв'язку між двома змінними. Значення коефіцієнта кореляції завжди знаходиться в межах -1 і +1 ($-1 \le r_{yx} \le +1$). Додатне значення коефіцієнта кореляції свідчить про прямий, а від`ємне - про зворотний зв'язок між змінними. Коли коефіцієнт кореляції прямує за абсолютною величиною до 1 ($r_{yx} \to \pm 1$), то це свідчить про наявність міцного зв'язку, тобто щільність зв'язку велика. У протилежному випадку, коли коефіцієнт кореляції прямує до 0 ($r_{yx} \to 0$) – зв'язок слабкий або зовсім відсутній.

Знаки параметра a_1 і коефіцієнта кореляції мають збігатися [4-7].

3.4 Перевірка простої лінійної регресійної моделі на адекватність

Наступним етапом проведення економетричного аналізу є підтвердження або спростування первинної гіпотези про лінійність зв'язку

між вибраними змінними. Для цього використовують коефіцієнт детермінації, який позначають як R^2. Коефіцієнт детермінації використовують як критерій адекватності моделі.

Коефіцієнт детермінації завжди додатний і знаходиться в межах від нуля до одиниці ($0 \le R^2 \le 1$). коефіцієнт детермінації дорівнює квадрату коефіцієнта кореляції:

$$R^2 = r_{yx}^2 . \qquad (3.5)$$

За допомогою коефіцієнта детермінації можна перевірити адекватність простої регресивної моделі. Якщо його значення близьке до одиниці, то можна вважати, що модель адекватна. Якщо його значення близьке до нуля, то модель неадекватна, тобто немає лінійного зв'язку між залежною та незалежною змінними.

$0 \le R^2 < 0{,}45$ – модель неадекватна, тобто не існує лінійного зв'язку між змінними;

$0{,}55 < R^2 \le 1$– модель адекватна, тобто існує лінійний зв'язку між змінними;

$0{,}45 \le R^2 \le 0{,}55$ – невизначений коефіцієнт детермінації, тобто за його значенням неможливо визначити адекватність моделі.

У таких випадках необхідно використовувати інші показники. Отже, потрібен інший критерій, який би однозначно давав би відповідь на запитання про адекватність побудованої моделі. Найбільш поширеним із таких критеріїв є *критерій Фішера*.

Перевірка моделі за F-критерієм Фішера складається з певних етапів.

На першому етапі розраховуємо величину так званого F-відношення (розрахункове значення):

$$\mathbf{F_{(k-1,n-k)} = \frac{R^2/(k-1)}{(1-R^2)/(n-k)}}, \qquad (3.6)$$

де **n** – кількість спостережень;

k – кількість параметрів моделі (для простої регресії k = 2).

На другому етапі задаємо рівень значущості β.

На третьому етапі за статистичними таблицями F-розподілу Фішера зі (**k-1**,**n-k**) ступенями свободи та рівнем значущості β знаходимо критичне значення $F_{кр}$ (критичне значення критерію Фішера визначають за довідником критичних значень математичних показників).

Якщо розраховане значення $\mathbf{F} \rangle \mathbf{F_{кр}}$, то можна зробити висновок, що побудована регресійна модель адекватна, тобто дійсно існує лінійний зв'язок між **x** і **y**. Якщо $F \langle F_{кр}$, то модель неадекватна, тобто зв'язок між досліджуваними змінним не лінійного вигляду.

Глосарій

Коефіцієнт детермінації — це критерій, який використовують для визначення адекватності моделі, тобто підтвердження або спростування первинної гіпотези про лінійність зв'язку між вибраними змінними.

Коефіцієнт кореляції — це критерій, що характеризує щільність зв'язку між залежною змінною ***y*** і незалежною ***x***, тобто наскільки значним є вплив змінної ***x*** на ***y,*** і є відносною мірою зв'язку між двома змінними.

Кореляція (лат. correlatio) — це співвідношення, зв'язок або залежність між змінними регресії.

Критерій Фішера — це показник, за яким визначають адекватність моделей у випадках, коли коефіцієнт детермінації потрапляє в невизначений інтервал.

Проста лінійна регресія — це модель, яку вважають лінійною, якщо вона лінійна за своїми параметрами. Її можна трактувати як пряму на площині.

Контрольні запитання

1 Поясніть зв'язок між результативним показником y і параметром a_1, тобто коефіцієнтом при змінній.

2 Поясніть зв'язок між результативним показником y і параметром a_0, тобто вільним членом економетричної моделі.

3 Як оцінити щільність зв’язку між змінними та результативним показником у регресійній моделі?

4 Поясніть значення адекватності моделі. Як її перевірити?

5 У яких випадках необхідно використовувати критерій Фішера?

Тема 4. Загальна проста лінійна регресія

План

4.1 Загальна проста лінійна модель.

4.2 Перевірка на значущість параметрів загальної простої лінійної регресії. Інтервали довіри.

4.3 Прогнозування значення результативного показника загальної простої лінійної регресії.

Ключові слова: загальна регресійна модель, t-тест Стьюдента, статистична значущість параметрів загальної регресійної моделі, прогнозування ендогенної змінної, точковий метод прогнозування значення результативного показника, інтервальний метод прогнозування значення результативного показника.

4.1 Загальна проста лінійна модель

Вибіркові лінійні регресійні моделі правильні тільки для однієї вибірки. Якщо розглянути всю генеральну сукупність, із якої вибирають вибірку, то правильною для неї буде модель

$$y = \alpha_0 + \alpha_1 x + \varepsilon, \tag{4.1}$$

де α_0, α_1 - дійсні параметри всієї генеральної сукупності;

ε - неспостережувана випадкова величина.

Цю модель називають ***узагальненою регресивною моделлю*** [1, 3, 7].

4.2 Перевірка на значущість параметрів загальної простої лінійної регресії. Інтервали довіри

Якщо параметри вибіркової лінійної моделі розраховані за методом найменших квадратів, то з виконанням класичних припущень загальної лінійної моделі математичні сподівання параметрів a_0, a_1 дорівнюють значенням параметрів узагальненої моделі α_0, α_1 (яка є дійсною для всієї генеральної сукупності).

Для перевірки статистичної значущості параметрів узагальненої простої лінійної регресії a_0 і a_1 використовують ***t-тест Стьюдента***, розрахункове значення t-критерію можна отримати за формулою

$$t^* = \frac{a_j}{\sqrt{\sigma^2_{a_j}}}, \tag{4.2}$$

де $\sigma^2_{a_0} = \sigma^2_u \frac{\sum_{i=1}^{n} x_i^2}{n\sum_{i=1}^{n}(x_i - \bar{x})^2}$; $\sigma^2_{a_1} = \sigma^2_u \frac{1}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$; $\sigma^2_u = \frac{\sum_{i=1}^{n} u_i^2}{n-k}$; (4.3)

aj – відповідний параметр вибіркової простої регресії;

σ^2_u – оцінка дисперсії залишків;

$\sigma^2_{a_0}, \sigma^2_{a_1}$ – дисперсія оцінок параметрів моделі;

k – кількість оцінених параметрів (для простої регресії k = 2, бо оцінюють параметри a_0 і a_1).

Значення $t^*_{розр}$ порівнюють із табличним значенням, що дає змогу знайти критичну область із (n-k) ступенями свободи. Для знаходження критичного значення потрібно задати рівень значущості β. Потім за таблицями t-розподілу Стьюдента (довідник критичних значень математичних показників) за заданого рівня значущості β і ступенях свободи (n-k) знайти відповідне критичне значення ($t_{\beta/2}$) розрахункового значення t-статистики (рисунок 4.1) [2, 4, 5].

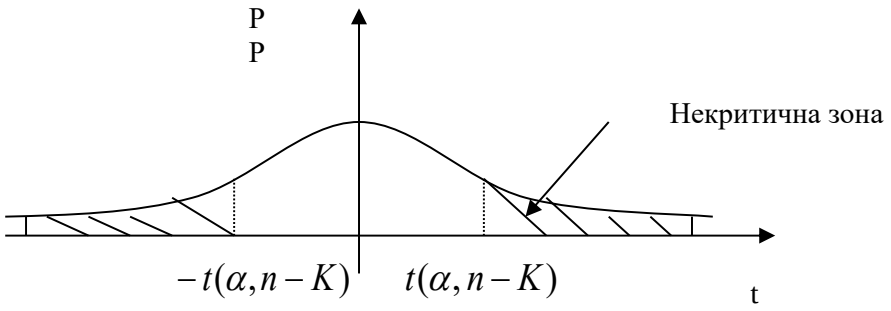

Рисунок 4.1 – Графічне зображення критичної зони для розрахункового значення t-статистики

Якщо значення $\mathbf{t^*}$ потрапляє в критичну зону, тобто $-\mathbf{t}_{\beta/2} \langle \mathbf{t^*} \langle \mathbf{t}_{\beta/2}$, то можна зробити висновок, що з імовірністю $(\mathbf{1-\beta})$ оцінка $\mathbf{a_j}$ є статистично незначущою (тобто приймають гіпотезу, що $\boldsymbol{\alpha_j = 0}$). У протилежному випадку оцінка $\mathbf{a_j}$ є статистично значущою.

Для того щоб визначити, як параметри a_0 і a_1 пов'язані з параметрами α_0 і α_1, потрібно побудувати ***інтервали довіри*** для параметрів узагальненої регресійної моделі, тобто такі інтервали, у які із заданою ймовірністю потрапляють їхні значення. Для цього розраховуємо межі інтервалів:

$$\alpha_j = a_j \pm t_{\beta/2} \cdot \sigma_{a_j}, \tag{4.4}$$

де $\sigma_{a_j} = \sqrt{\sigma^2_{a_j}}$ - стандартна помилка оцінювання параметра a_j.

4.3 Прогнозування значення результативного показника загальної простої лінійної регресії

Якщо модель адекватна, то можна знаходити прогнозні значення залежної змінної ***y*** виходячи з побудованої моделі. При цьому можна отримувати два типи прогнозів: точковий та інтервальний. ***Точковий прогноз*** дає теоретичне значення залежної змінної для відповідного значення x_{n+1} із побудованої вибіркової моделі без урахування усіх інших можливих факторів, які не враховані в моделі, помилок і відхилень:

$$\tilde{y}_{n+1} = a_0 + a_1 \cdot x_{n+1}. \tag{4.5}$$

При цьому, виходячи з узагальненої моделі, дійсне значення y для відповідного значення x_{n+1} буде дорівнювати

$$y_{n+1} = \alpha_0 + \alpha_1 \cdot x_{n+1} + \varepsilon_{n+1}, \quad (4.6)$$

де ε_{n+1} – значення випадкової величини, не спостережуваної для значення x_{n+1}. Дійсне значення y_{n+1} знайти неможливо, можна лише оцінити його за допомогою прогнозу.

Отже, застосування ***інтервального методу прогнозу*** результативного показника дає змогу визначити інтервал, у який можливо очікувати потрапляння значення y_{n+1}, але з урахуванням впливу всіх інших можливих факторів помилок і відхилень.

Такий інтервал довіри за заданого рівня значущості β для y_{n+1}

$$(a_0 + a_1 \cdot x_{n+1}) \pm t_{\beta/2} \cdot \sigma_u \sqrt{\left(\frac{1}{n} + \frac{(x_{n+1} - \bar{x})^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \right)}. \quad (4.7)$$

Глосарій

Загальна регресійна модель — це економетрична модель, яка дійсна для всієї сукупності екзогенних змінних та ендогенних змінних, що їй відповідають.

Інтервальний метод прогнозування значення результативного показника — дає змогу визначити лише інтервал, у який можливо очікувати потрапляння значення y_{n+1}, але з урахуванням впливу всіх інших можливих факторів помилок і відхилень [1, 3-7].

Статистична значущість параметрів загальної регресійної моделі — параметри є статистично значущими, коли відповідне значення тесту Стьюдента не дорівнює 0, тобто не потрапляють у критичну зону.

Точковий метод прогнозування значення результативного показника — дає теоретичне значення залежної змінної для відповідного значення $\mathbf{x_{n+1}}$ із побудованої вибіркової моделі без урахування всіх інших можливих факторів, які не враховані в моделі, помилок і відхилень.

t-тест Стьюдента — це показник, за яким визначають статистичну значущість параметрів економетричної моделі.

Контрольні запитання

1 Дайте визначення загальної регресійної моделі.

2 Як визначити статистичну значущість параметрів загальної регресійної моделі?

3 Із застосуванням яких методів можна прогнозувати результативний показник загальної економетричної моделі?

Тема 5. Загальна лінійна багатофакторна регресія

План

5.1 Класична лінійна багатофакторна модель.

5.2 Етапи побудови багатофакторної регресійної моделі.

5.3 Визначення невідомих параметрів багатофакторної регресії за методом найменших квадратів (МНК).

5.3 Аналіз багатофакторної економетричної моделі.

Ключові слова: лінійна багатофакторна модель, параметри багатофакторної регресії, множинна кореляція, множинна детермінація,

коефіцієнт множинної кореляції, коефіцієнт множинної детермінації, оцінений коефіцієнт множинної детермінації.

5.1 Класична лінійна багатофакторна модель

У загальному випадку *модель* – це штучно створений аналог реального об'єкта (оригіналу), який певною мірою відтворює структуру, властивості та характеристики оригіналу.

Існують різні класифікації моделі залежно від ознаки, покладеної в основу класифікації.

Розрізняють предметні, фізичні, знакові та інші моделі. Однак серед множини моделей пріоритетне місце міцно займають *математичні моделі*, що можна пояснити рядом їхніх визначальних особливостей: компактною формою подання інформації, практичною відсутністю витрат на реалізацію, можливістю застосування математико-аналітичних методів аналізу та оптимізації, практично необмеженим ресурсом використання засобів обчислювальної техніки для потреб моделювання.

Типова *постановка задачі* побудови математичної моделі є об'єктом дослідження, функціонування якого може бути описане двома групами змінних: першу з них становлять так звані вхідні, незалежні змінні або фактори, упорядкована сукупність яких утворює вектор $X = [x_1, x_2, ..., x_l]^T$, другу – вихідні, залежні змінні, вектор $Y = [y_1, y_2, ..., y_m]^T$.

Під математичною моделлю об'єкта досліджень розуміють формальну систему, що складається з елементів $y, x_1, x_2, ..., x_l$.

Необов'язково всі вхідні змінні будуть використані, деякі з них можуть зовсім не впливати на залежну змінну y і під час побудови математичної моделі мають бути вилучені з подальшого розгляду.

Для побудови математичної моделі зазвичай використовують інформацію, отриману через спостереження за об’єктом досліджень у

режимі його природного функціонування (пасивний експеримент) чи під час спеціально організованих досліджень (активний експеримент). При цьому процедуру побудови математичної моделі за експериментальними даними, отриманими зі спостереженням вхідних і вихідних змінних об'єкта дослідження, називають *ідентифікацією об'єкта*.

У загальному випадку маємо такі *етапи ідентифікації*:

1) визначення сукупності «суттєвих» факторів, із яких формується вектор X';

2) встановлення характеру зв'язків (відношень) між елементами математичної моделі;

3) оцінювання значень вектора $A=[a_1, a_2, ..., a_p]$ параметрів (коефіцієнтів) математичної моделі;

4) верифікація математичної моделі, тобто перевірка її придатності для прикладного застосування.

Отже, перший і другий етапи, результатом виконання яких є визначення структури математичної моделі, часто поєднують під назвою *структурна ідентифікація*. Безумовно, структурна ідентифікація є визначальним етапом ідентифікації в цілому. Отримані на цьому етапі негативні результати унеможливлюють позитивне розв'язання всієї задачі. Однак інколи структура математичної моделі може бути визначена ще до проведення експерименту з об'єктом дослідження через безпосередній аналіз причинно-наслідкових зв'язків, що дають змогу визначити механізми (алгоритми) дії об'єкта дослідження, або застосування сучасної системи знань про технічні, фізичні, біологічні, соціальні та інші принципи функціонування механізмів об'єкта дослідження, що допомагає виключно теоретично визначити структуру моделі. У цьому випадку немає потреби в структурній ідентифікації, і процедуру ідентифікації слід починати з третього етапу – *параметричної ідентифікації* або, як її називають, ідентифікацією у вузькому сенсі. Зміст задачі параметричної ідентифікації

– якнайточніше обчислити кількісні значення параметрів $a_0, a_1, a_2, \ldots$ математичної моделі за експериментальними даними, які разом із корисною інформацією містять випадкову складову, що й обумовлює появу помилок у розрахованих значеннях (оцінках) параметрів.

Останній етап ідентифікації має на меті перевірку «споживчих» якостей математичної моделі, у першу чергу можливості її ефективного і надійного використання відповідно до мети ідентифікації. Здебільшого механізм підтвердження моделі полягає в співставленні вихідних реакцій об'єкта дослідження та математичної моделі за однією і тією самою вибіркою вхідних сигналів. Практичну однаковість цих реакцій вважають достатнім підтвердженням якості моделі.

Однак об'єктивнішими слід вважати висновки з аналізу якості математичної моделі, отримані з урахуванням ряду аспектів, пов'язаних із цільовим призначенням математичної моделі, яке має суттєвий вплив на особливості ідентифікації та використання моделей. Зокрема, одним із найпоширеніших видів математичної моделі є так звані *пізнавальні моделі,* що являють собою системи для накопичення, концентрації, зберігання знань та інформації навчально-пізнавального і наукового характеру. Пізнавальні моделі, не орієнтовані на застосування в конкретних прикладних задачах, тяжіють до масштабного, глобального відображення суті проблем, явищ, процесів. Приклад таких моделей – деякі фундаментальні концепції та природні закони: модель Сонячної системи, закон всесвітнього тяжіння, моделі земної кулі та ін. Різновид пізнавальних моделей - так звані *феноменологічні, або концептуальні моделі*, які дають спрощені схематичні описи явищ, процесів, об'єктів, уникаючи деталізації чи локальних уточнень і зосереджуючи увагу на найсуттєвіших елементах механізму причинно-наслідкових зв'язків у досліджуваних явищах, процесах, не враховуючи впливу менш значущих, другорядних факторів.

Інший аспект використання моделей – *прогностичні моделі*, призначені насамперед для застосування в системах управління. Цей прогноз можна отримати як за допомогою засобів математичного моделювання, так і застосуванням моделей інших типів, наприклад фізичних.

Серед математичних моделей поширені так звані *апроксимативні моделі*, характерною особливістю яких є те, що структура цих моделей не відображає внутрішнього механізму тих явищ і процесів, що визначають зміст об'єкта ідентифікації. Вибирають структуру апроксимативної моделі (від англ. approximation – наближення) тільки формально за критерієм близькості виходу моделі до значень експериментально отриманих вихідних даних за умов подання на вхід моделі реального вхідного впливу.

Сфера практичного використання апроксимативних моделей надзвичайно широка. Обставинами, що сприяють цьому, є як певні особливості об'єкта дослідження, так і умови проведення самого дослідження: погані структурованість і вимірювання даних про об'єкт дослідження, низький рівень формалізації відомостей та інформованості про спосіб його функціонування. У подібній ситуації єдиним джерелом інформації про об'єкт дослідження часто є лише вибірка експериментально отриманих даних.

Узагальнена багатофакторна лінійна регресійна модель може бути записана в такому вигляді:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n + e , \qquad (5.1)$$

де y – залежна змінна;

$x_1, x_2, \ldots, x_n$ – незалежні змінні (або фактори);

$\beta_0, \beta_1, \ldots, \beta_n$ – параметри моделі (константи), які потрібно оцінити;

e – неспостережувана випадкова величина.

Узагальнена регресійна модель – це модель, яка дійсна для всієї генеральної сукупності. Невідомі параметри узагальненої моделі є константами, а випадкова величина – неспостережувана, і ми можемо зробити тільки припущення відповідно до закону її розподілу. На відміну від узагальненої регресійної моделі, вибіркова модель побудована для певної вибірки; невідомі параметри вибіркової моделі є випадковими величинами, математичне сподівання яких дорівнює параметрам узагальненої моделі (випадок класичної лінійної регресії), випадкові величини (помилки) можна оцінити, виходячи з вибіркових даних [1-4].

Вибіркова лінійна багатофакторна модель має такий вигляд:

$$y = b_0 + b_1 x_1 + b_2 x_2 + \ldots + b_n x_n + e, \qquad (5.2)$$

де y – залежна змінна;

$x_1, x_2, \ldots, x_n$ – незалежні змінні (або фактори);

$b_0, b_1, \ldots, b_n$ – оцінки невідомих параметрів узагальненої моделі;

e – випадкова величина (помилка).

Лінійною регресійною моделлю називають модель, лінійну за своїми параметрами.

Багатофакторна лінійна регресійна модель має p незалежних змінних, або факторів, які впливають на залежну змінну y, і $(p+1)$ невідомих параметрів, які потрібно оцінити.

5.2 Етапи побудови багатофакторної регресійної моделі

Процес побудови багатофакторної регресійної моделі більш складний, ніж процес побудови простої лінійної регресії. Він складається з багатьох етапів. Серед них можна виділити такі:

1 Вибір і аналіз усіх можливих факторів, які впливають на процес (або показник), що вивчають.

2 Вибір і аналіз знайдених факторів.

3 Математико-статистичний аналіз факторів.

4 Вибір методу і побудова регресійної багатофакторної моделі.

5 Оцінювання невідомих параметрів регресійної моделі.

6 Перевірка моделі на адекватність.

7 Розрахунок основних характеристик і побудова інтервалів довіри.

8 Аналіз отриманих результатів, висновки.

Метод побудови регресійної багатофакторної моделі неможливо відокремити від самої моделі, вони дуже тісно пов'язані між собою. Іншими словами, саме вибраний метод впливає на остаточний вигляд регресійної моделі.

Оцінюють невідомі параметри $\beta_0, \beta_1, \ldots, \beta_n$ у лінійних регресійних моделях за *методом найменших квадратів*.

Після того як параметри знайдено, перевіряють моделі на адекватність за допомогою *F-критерію Фішера*, а також значущість знайдених параметрів за *t-критерієм Стьюдента*. Якщо модель неадекватна, то необхідно повернутися до етапу побудови моделі і, можливо, від лінійної моделі перейти до нелінійної, або ввести додаткові фактори.

Якщо модель адекватна, то можемо працювати далі: прогнозувати, вивчати вплив окремих факторів на залежний показник, будувати інтервали довіри, аналізувати та інтерпретувати отримані результати. Для того щоб розглянути, як можна проінтерпретувати параметри регресійної моделі, повернемося до загальної моделі багатофакторного регресійного аналізу та знайдемо математичне очікування обох частин. Виходячи з основних припущень, отримаємо

$$E(y_i \mid x_{1i}, x_{2i}, \ldots, x_{ni}) = \beta_0, + \beta_1, x_{1i} + \ldots + \beta_n x_{ni} . \qquad (5.3)$$

Рівняння дає умовне математичне сподівання y за фіксованих значень x.

Параметри $\beta_0, \beta_1, \ldots, \beta_n$ ще називають коефіцієнтами регресії. Кожен із них вимірює вплив відповідної змінної за умови, що всі інші залишаються, тобто є константами.

5.3 Визначення невідомих параметрів багатофакторної регресії за методом найменших квадратів (МНК)

Маємо ряд спостережень за залежною змінною $y = \{y_1, y_2,\ldots,y_n\}$ і незалежними змінними, або факторами:

$$x_1 = \{x_{11}, x_{21},\ldots,x_{1n}\}; \; x_2 = \{x_{21}, x_{22},\ldots,x_{2n}\};$$
$$x_n = \{x_{n1}, x_{n2},\ldots,x_{pn}\}. \qquad (5.4)$$

На підставі цих спостережень побудуємо лінійну вибіркову багатофакторну модель:

$$y = b_0 + b_1 x_1 + b_2 x_2 +\ldots+ b_p x_p, \qquad (5.5)$$

де y – залежна змінна;

$x_1, x_2, \ldots, x_p$ - незалежні змінні, або фактори;

$b_0, b_1, \ldots, b_p$ - невідомі параметри;

e – випадкова величина, або помилка.

Як і у випадку простої лінійної регресії, знайдемо невідомі параметри за методом найменших квадратів, тобто мінімізуючи суму квадратів відхилень фактичних даних від теоретичних (даних, які ми отримуємо з регресійної моделі):

$$F = (b_0, b_1,\ldots,b_n) = \min \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - b_0 - b_1 x_1 - \ldots - b_p x_p)^2 \quad . \qquad (5.6)$$

Для того щоб знайти мінімум виразу, необхідно прирівняти до нуля часткові похідні функції F за аргументами $b_0, b_1, \ldots, b_n$. Отримаємо систему нормальних рівнянь. Зважаючи на досить громіздкий вигляд системи нормальних рівнянь у загальному випадку, ми не будемо її наводити. Загальний вираз для розрахунку невідомих параметрів моделі розглянемо пізніше, коли повернемося до матричного підходу в багатофакторному аналізі. Зазначимо тільки, що перетин (параметр b_0) розраховують аналогічно простій регресії за допомогою середніх значень:

$$b_0 = \bar{y} - b_1\bar{x}_1 - \ldots - b_p\bar{x}_p \quad . \tag{5.7}$$

5.4 Аналіз багатофакторної економетричної моделі

Аналізують багатофакторну економетричну модель за допомогою коефіцієнта множинної кореляції та детермінації, який є корисною мірою ступеня відповідності даних, $\{y_i, i = \overline{1,n}\}$, отриманих із регресійної моделі, фактичним даним $\{x_i, i = \overline{1,n}\}$. Коефіцієнт множинної кореляції визначають як коефіцієнт кореляції між y та x:

$$R = \frac{\sum_{i=1}^{n}(y_i - \bar{y})(x_i - \bar{x})}{\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2 (x_i - \bar{x})^2}} \quad . \tag{5.8}$$

Квадрат коефіцієнта множинної кореляції, як і у випадку простої регресії, називають *коефіцієнтом детермінації* і позначають через R^2. Можна показати, що вигляд коефіцієнта детермінації у випадку багатофакторної регресії ідентичний коефіцієнту детермінації простої лінійної регресії:

$$R^2 = \frac{\sum_{i=1}^{n} x_i - \bar{y})^2}{\sum_{i=1}^{n} y_i - \bar{y})^2} \quad . \qquad (5.9)$$

Коефіцієнт детермінації r^2 та оцінений коефіцієнт детермінації $\bar{R}^2$

Важливою властивістю коефіцієнта детермінації R^2 є те, що він – неспадна функція від кількості факторів, які входять до моделі. Якщо кількість факторів зростає, то R^2 також зростає і ніколи не зменшується. Тобто, якщо ми додаємо новий фактор у регресійну модель, це тільки збільшує значення коефіцієнта детермінації R^2, що легко побачити з його визначення:

$$R^2 = 1 - \frac{\sum_{i-1}^{n} e_i^2}{\sum_{i-1}^{n} \left(y_i - -\bar{y}\right)^2} \quad . \qquad (5.10)$$

У цьому виразі знаменник не залежить від кількості факторів *x,* тоді як чисельник, навпаки, залежить. Інтуїтивно можна зрозуміти, що, якщо кількість факторів *x* зростає, величина $\sum_{i=1}^{n} e_i^2$ спадає (або хоча б не зростає). Якщо ми порівнюватимемо дві регресійні моделі з однаковою залежною змінною, але різною кількістю факторів x, то, звичайно, переважною буде та, яка має більше значення R^2.

Зразу ж виникає запитання, що робити, якщо ми хочемо порівняти значення коефіцієнтів детермінації в різних моделях. У таких випадках потрібно корегувати коефіцієнт кореляції з урахуванням кількості факторів *x,* які входять у різні моделі, тобто зменшити вплив залежності значення коефіцієнта детермінації від кількості факторів. Для цього вводять спеціальний коефіцієнт детермінації:

$$\overline{R}^2 = 1 - \frac{\frac{1}{(n-k)}\sum_{i=1}^{n} e_i^2}{\frac{1}{(n-1)}\sum_{i=1}^{n}(y_i - \overline{y})^2} \quad , \tag{5.11}$$

де *k* – кількість параметрів регресійної моделі, включаючи перетин.

На відміну від простого *коефіцієнта детермінації*, *оцінений коефіцієнт детермінації* корегують з урахуванням ступенів вільності суми квадратів остачі і загальної суми квадратів [4].

Оцінений коефіцієнт детермінації $\overline{R}^2$ і коефіцієнт детермінації R^2 пов'язані між собою такою залежністю:

$$\overline{R}^2 = 1 - (1 - R^2)\frac{n-1}{n-k} \, . \tag{5.12}$$

Із виразу видно, якщо $k > 1$, то $\overline{R}^2 < R^2$. Крім того, якщо кількість факторів *x* зростає, оцінений коефіцієнт детермінації зростає повільніше, ніж просто коефіцієнт детермінації. Отже, зменшується вплив кількості факторів на величину коефіцієнта детермінації, тому на практиці більше використовують оцінений коефіцієнт детермінації, особливо для порівняння різних регресійних моделей. Слід зазначити, що оцінений коефіцієнт детермінації може бути і від'ємним, на відміну від R^2, який має додатне значення. Крім того, коли $R^2 = 1$, оцінений коефіцієнт кореляції також дорівнює одиниці. Коли $\overline{R}^2$ прямує до від'ємної величини, R^2 прямує до нуля.

Перевірка моделі на адекватність за F-критерієм Фішера

Для перевірки адекватності багатофакторної регресійної моделі, як і у випадку простої лінійної моделі, використовують **F-критерій Фішера**.

При цьому нуль-гіпотеза узагальнена і має вигляд

$$H_0: \beta_1 = \beta_2 = \ldots = \beta_p = 0 \ . \qquad (5.13)$$

Проти альтернативної гіпотези H_1 хоча б одне значення β_1 відмінне від нуля.

Якщо нуль-гіпотеза H_0 неправильна, тоді правильна гіпотеза H_1, тобто не всі параметри незначною мірою відрізняються від нуля, що дає підставу вважати, що відібрані фактори пояснюють зміну залежної величини *y*. Для перевірки H_0-гіпотези розраховують *F*-статистику Фішера з *p* і (n-p-1) ступенями вільності:

$$F_{p,n-p-1} = \frac{\dfrac{\sum_{i=1}^{n}(x_i - \bar{y})^2}{p}}{\dfrac{\sum_{i=1}^{n}(y_i - x_i)^2}{n-p-1}}., \qquad (5.14)$$

де *p* – кількість факторів, які увійшли в модель;

n – загальна кількість спостережень.

За *F*-таблицями Фішера, як і у випадку простої регресії, знаходимо критичне значення $F_{кр}$ із *p* і *(n-p-1)* ступенями вільності, задавши попередньо рівень помилки $\alpha*100$ % (або рівень довіри *(1-*α*)*100 %).*

Якщо $F>F_{кр}$, тоді нуль-гіпотеза відкидають, що свідчить про адекватність побудованої моделі. У протилежному випадку її приймають і модель вважають неадекватною.

Глосарій

Коефіцієнт множинної детермінації — це показник, який оцінює якість побудованої моделі. Індекс множинної детермінації.

Коефіцієнт множинної кореляції — це показник, який є корисною мірою ступеня відповідності побудованої регресії фактичним даним, визначають як коефіцієнт кореляції між фактичним і теоретичним значеннями результативного показника.

Лінійна багатофакторна модель — це лінійна регресійна модель, що має p незалежних змінних, або факторів, які впливають на залежну змінну y, і (p+1) невідомих параметрів, при цьому $p = 2$ або більше.

Множинна кореляція — це відповідність побудованої регресії фактичним даним.

Оцінений коефіцієнт множинної детермінації — це порівняний показник коефіцієнтів детермінації в різних моделях із різними функціями залежності з урахуванням кількості факторів у моделях.

Параметри багатофакторної регресії, або незмінні коефіцієнти регресії, визначають вплив зміни факторів x_i на рівень y.

Контрольні запитання

1 Дайте визначення багатофакторній лінійній моделі та етапам її побудови.

2 Як визначити вплив зміни факторів x_i на рівень y?

3 Як оцінити відповідність побудованої регресії фактичним даним?

4 Як оцінити якість побудованої моделі та ступеня відповідності побудованої регресії фактичним даним?

Тема 6. Мультиколінеарність

План

6.1 Поняття про мультиколінеарність і її вплив на оцінювання параметрів моделі.

6.2. Тестування наявності мультиколінеарності

Ключові слова: мультиколінеарність, суть мультиколінеарності, тестування наявності мультиколінеарності.

6.1 Поняття про мультиколінеарність і її вплив на оцінювання параметрів моделі

Умовою застосування для оцінювання параметрів лінійних багатофакторних моделей методу найменших квадратів є відсутність лінійних зв'язків між незалежними змінними моделі. Якщо такі зв'язки існують, то це явище називають ***мультиколінеарністю*** [2].

Термін «мультиколінеарність» вперше запровадив Р. Фріш. Суть мультиколінеарності полягає в тому, що в багатофакторній регресійній моделі дві або більше незалежних змінних пов'язані між собою лінійною залежністю або, іншими словами, мають високий ступінь кореляції

Наявність мультиколінеарності створює певні проблеми для розроблення моделей. Насамперед визначник матриці спостережень наближається до нуля, і оператор оцінювання за звичайним МНК стає надзвичайно чутливий до похибок вимірювань і похибок обчислень. При цьому МНК-оцінки можуть мати значне зміщення відносно дійсних оцінок узагальненої моделі, а в деяких випадках навіть стати взагалі беззмістовними.

Регресійне рівняння задовільно описує рух залежної змінної, якщо коефіцієнт множинної кореляції досить високий, а кореляція між

незалежними факторами незначна. Мультиколінеарність незалежних змінних веде до зміщення оцінок параметрів, а отже, до неможливості коректної інтерпретації результатів.

Наприклад, коли вивчають залежність між ціною акції, дивідендами на акцію та отриманим прибутком на акцію, то дивіденди і отриманий прибуток на одну акцію мають високий ступінь кореляції. Іншими словами, виникає ситуація, коли два колінеарних фактори змінюються в одному напрямку. У такому разі майже неможливо оцінити вплив кожного з них на досліджуваний показник.

З'ясуємо, до яких наслідків може призвести мультиколінеарність. Це одне з найважливіших питань, яке потрібно зрозуміти для розроблення економетричних моделей.

Практичні наслідки мультиколінеарності:

- мультиколінеарність незалежних змінних (факторів) призводить до зміщення оцінок параметрів моделі, які розраховують за методом найменших квадратів. На основі цих оцінок неможливо зробити конкретні висновки про результати взаємозв'язку між показником і факторами;
- збільшення дисперсії та коваріації оцінок параметрів, обчислених за методом найменших квадратів;
- збільшення довірчого інтервалу (оскільки збільшується середній квадрат відхилення параметрів);
- незначущість t-статистик.

Зауваження. Мультиколінеарність не є проблемою, якщо єдиною метою регресійного аналізу є прогноз (оскільки чим більше значення R^2, тим точніший прогноз). Якщо метою аналізу є не прогноз, а дійсне значення параметрів, то мультиколінеарність перетворюється на проблему, оскільки її наявність призводить до значних стандартних похибок оцінок параметрів.

6.2 Тестування наявності мультиколінеарності

Єдиного способу визначення мультиколінеарності, на жаль, немає. Зовнішні ознаки наявності мультиколінеарності такі:

- велике значення R^2 і незначущість t-статистики. Наявність цих двох факторів одночасно є «класичною» ознакою мультиколінеарності.

З одного боку, незначущість t-статистики Стьюдента означає, що один або більше оцінених параметрів статистично незначуще відрізняються від нуля. З іншого боку, якщо значення R^2 велике, ми приймаємо з великою ймовірністю F-критерій Фішера, який відкидає нульову гіпотезу. Суперечність свідчить про наявність мультиколінеарності;

- велике значення парних коефіцієнтів кореляції.

Якщо значення хоча б одного коефіцієнта кореляції $r_{xy} > 0,8$, то мультиколінеарність є серйозною проблемою.

Зазначимо, що велике значення парних коефіцієнтів кореляції — достатня, але не необхідна умова наявності мультиколінеарності. Мультиколінеарність може мати місце навіть за відносно невеликих значень парних коефіцієнтів кореляції у більш ніж двофакторній регресійній моделі.

Для визначення мультиколінеарності здебільшого застосовують такі тести:

- F-mecm, запропонований Глобером і Фарраром (він має й іншу назву:побудова допоміжної регресії);

- характеристичні значення та умовний індекс. Розглянемо їх більш детально.

Перший із них базований на тому, що за наявності мультиколінеарності один чи більше факторів пов'язані між собою лінійною залежністю. Одним із способів визначення щільності регресійного зв'язку є побудова регресійної залежності кожного фактора x_i з усіма іншими факторами. Тому *F-тест* має іншу назву: побудова допоміжної регресії.

Обчислення відповідного коефіцієнта детермінації для цього допоміжного регресійного рівняння та його перевірка за допомогою F-критерію дають змогу виявити лінійні зв'язки між незалежними змінними.

Нехай $R^2x_1x_2.....x_m$ — коефіцієнт детермінації в регресії, яка пов'язує фактор x_i з іншими факторами. Тоді *F-тест* виконують так, як зазначено в темі 3 цього конспекту лекцій, і за його значенням зробимо висновок про наявність лінійного зв’язку між факторами, що аналізують і, як наслідок, корелюють між собою.

Але цей метод не є універсальним і має недолік: він не проводить чіткої межі між тим, що треба вважати «суттєвою» мультиколінеарністю, яку необхідно враховувати, і тим, коли нею можна знехтувати.

1 Між незалежними змінними може існувати лінійна залежність, однак вона може і не бути явищем мультиколінеарності змінних, а тому не впливатиме на кількісні оцінки параметрів моделі, розраховані за допомогою звичайного МНК.

2 Якщо $F_{роз} > Ft_{абл}$, то X_k залежить від усіх інших незалежних змінних, і потрібно вирішити питання про її виключення з переліку змінних.

3 Якщо tроз > tкр, то X_k і X_j щільно пов'язані між собою.

4 Аналізуючи F- і t-критерії, робимо висновок, яку зі змінних треба виключити з моделі (зрозуміло, якщо це можливо з економіко-логіко-теоретичних міркувань).

5 Якщо виконавши пункти 2-4, ми не досягли мети, тобто не усунули мультиколінеарність, оцінювати параметри моделі потрібно за допомогою іншого методу, наприклад методу головних компонентів (абооднієї з його модифікацій).

Засоби усунення мультиколінеарності. Метод головних компонентів

Виявлення мультиколінеарності є лише частиною справи. Інша частина

— як її усунути. Безпомилкових і абсолютно правильних порад немає, оскільки мультиколінеарність є прикладною проблемою.

Звичайно, усе залежить від ступеня мультиколінеарності, однак у будь-якому разі можна запропонувати кілька простих методів усунення мультиколінеарності:

1) використання додаткової або первинної інформації;
2) об'єднання інформації;
3) відкидання змінної з високою кореляцією;
4) перетворення даних (використання перших різниць);
5) збільшення кількості спостережень.

Які поради спрацюють на практиці, залежить від істотності проблемита її характеру.

Якщо переліченими методами не вдається усунути мультиколінеарність, то для оцінювання параметрів багатовимірної моделі доцільно застосувати метод головних компонентів. Цей метод доцільно застосовувати, по-перше, для оцінювання параметрів моделей із великою кількістю факторів,по-друге, для моделей, у яких незалежні змінні (стовпці матриці спостережень *X*) мають однакові одиниці вимірювання [2, 4].

Глосарій

Одна з передумов застосування методу найменших квадратів для оцінювання параметрів лінійних багатофакторних моделей — відсутність лінійних зв'язків між незалежними змінними моделі. Якщо такі зв'язки існують, то це явище називають ***мультиколінеарністю***.

Мультиколінеарність полягає в тому, що в багатофакторній регресійній моделі дві або більше незалежних змінних пов'язані між собою лінійною залежністю або, іншими словами, мають високий ступінь кореляції.

Контрольні запитання

1 У чому проявляється мультиколінеарність незалежних змінних?

2 Які наслідки наявності мультиколінеарності для проведення економетричного аналізу?

3 Як впливає наявність мультиколінеарності на оцінювання параметрів моделі?

4 Які існують методи усунення мультиколінеарності?

Тема 7. Нелінійна регресія

План

7.1 Рівняння нелінійної парної регресії.

7.2 Найпростіші перетворення нелінійних моделей у лінійні.

Ключові слова: нелінійна парна регресія, криві зростання, лінеаризація.

7.1 Рівняння нелінійної парної регресії

За практичного використання лінійне моделювання економічних процесів дає позитивний результат і допомагає ефективно аналізувати і прогнозувати явища. Але внаслідок складності економетричних процесів доволі часто неможливо проводити дослідження, застосовуючи лише лінійні моделі. У таких випадках доцільно використовувати нелінійні економетричні моделі.

Нелінійна економетрична модель – це регресійна модель, яка встановлює нелінійну залежність між економічними показниками, один із

яких є залежною (пояснюваною) змінною, а інші – незалежними (пояснювальними) змінними.

Виділяють *два класи нелінійних регресійних моделей*:

1 Нелінійні відносно включених в аналіз пояснювальних змінних, але лінійні за оцінюваними параметрами (поліноміальна функція, гіперболічна).

2 Нелінійні як відносно включених в аналіз пояснювальних змінних, так і за оцінюваними параметрами – ***квазілінійні*** (степенева функція, показникові). Прикладом квазілінійної моделі може бути модель попиту, для якої функція попиту має такий вигляд:

$$Q = \beta_0 + \beta_1 P + \beta_2 P^2, \qquad (7.1)$$

де Q – попит;

P – ціна;

β_0, β_1 і β_2 – параметри моделі.

Головна відмінність між наведеними класами полягає в можливості застосування звичайного методу найменших квадратів (ЗМНК).

Для першого класу (нелінійні за змінними) достатньо зробити заміну змінних, щоб модель стала лінійною. Для другого класу (квазілінійних моделей) зазвичай використовують процедуру лінеаризації через логарифмування.

Порівнюючи лінійні і нелінійні економетричні моделі, можна зробити такі висновки:

1) лінійна класична регресійна модель і відповідні методи оцінювання її параметрів, тестування і прогнозування в цілому теоретично краще обґрунтовані, ніж нелінійні;

2) щодо простоти обчислювального апарату, то лінійна модель регресії також не перевершена ніякими іншими типами моделей [2, 4].

7.2 Найпростіші перетворення нелінійних моделей у лінійні

На практиці економетричного моделювання за необхідності застосування нелінійних моделей практикують такі два підходи:

1) замість складної нелінійної функціональної залежності навіть із невеликою кількістю факторів намагаються застосувати лінійну регресійну функцію з великою кількістю факторів;

2) за допомогою математичних перетворень намагаються звести (перетворити) нелінійну функцію на лінійну.

Для оцінювання параметрів нелінійних моделей використовують два основні підходи:

– перший основний прийом, за допомогою якого спрощено процес оцінювання параметрів нелінійної моделі, – це лінеаризація.

Лінеаризація – перехід від нелінійних зв'язків (показникової, степеневої функції тощо) до лінійного зв'язку за допомогою різних перетворень, що дає змогу надалі використовувати звичайний метод найменших квадратів (МНК). Отже, для оцінювання параметрів заміняють змінні, логарифмують відповідні частини рівняння або застосовують комбінований підхід для моделей нелінійних за змінними та параметрами;

– другий підхід використовують, коли підібрати відповідні перетворення не вдається, у такому випадку використовують методи нелінійної оптимізації на основі вихідних змінних.

Загальний вигляд регресії, лінійної відносно параметрів:

$$y_i = a_0 + a_1 j(x_i) + l_i. \qquad (7.2)$$

Заміною змінної $z_i = j(x_i)$, $i = 1...n$, нелінійна парна модель зведена до лінійної $y_i = a_0 + a_1 z_i + l_i$.

В економетриці використовують такі лінії, які носять назву **кривих зростання**. Криві зростання описують різні тенденції економічних процесів, наприклад життєвий цикл товару, процес нагромадження капіталу, маркетингову діяльність фірм тощо. Щоб правильно підібрати найкращу функцію для моделювання і прогнозування економічного явища, необхідно знати особливості кожного виду кривих (про це детально в темі 2).

1 Показникова (експоненційна) модель: основна $y = \alpha\beta^x$, еквівалентна $y = \alpha e^{\beta x}$, де α, β – параметри моделі.

2 Степенева (мультиплікативна) $y = \alpha x^{\beta}$, де α, β – параметри моделі.

3 Зворотна функція регресії $y = \beta_0 + \beta_1 \frac{1}{x}$.

4 Поліном m–го ступеня $y = a_0 + a_1 x + \ldots + a_m x^m$.

5 Проста експонента $y = \alpha\beta^x$, де α і β – додатні числа, при цьому якщо $\beta > 1$, то функція зростає зі зростанням часу t, якщо $\beta < 1$ – функція спадає.

6 Модифікована експонента $y = \alpha\beta^x + \gamma$, де постійні величини $\alpha < 0$, $0 < \beta > 1$, а константа γ має назву асимптоти цієї функції, тобто значення функції необмежено наближуються (знизу) до величини γ.

7 Крива Гомперца $y = e^{\alpha\beta^x - \gamma}$, де α, β – додатні параметри, причому $\beta < 1$; параметр γ – асимптота функції.

8 Логістична крива $y = \frac{1}{\alpha\beta^x + \gamma}$.

Найчастіше в економіці використовують поліноміальні функції другого й третього ступенів. Такі функції легко лінеаризовані за допомогою переходу до нових змінних. У практиці криволінійного вирівнювання найчастіше використовують два види перетворень: логарифмування (ln$y$) і зворотне перетворення $\frac{1}{x}$. При цьому можливі перетворення як залежної змінної y , так і незалежної x або обох одночасно [2, 4].

Глосарій

Криві зростання - графіки нелінійних функцій, що описують різні тенденції економічних процесів.

Лінеаризація – перехід від нелінійних зв'язків (показникової, степеневої функції тощо) до лінійного зв'язку за допомогою різних перетворень, що дає змогу надалі використовувати звичайний метод найменших квадратів (МНК).

Нелінійна економетрична модель – це регресійна модель, яка встановлює нелінійну залежність між економічними показниками, один із яких є залежною (пояснюваною) змінною, а інші – незалежними (пояснювальними) змінними.

Контрольні запитання

1 Для чого застосовують перетворення нелінійних моделей?

2 Які є класи нелінійних регресійних моделей?

3 Які підходи використовують для оцінювання параметрів нелінійних моделей?

4 У яких випадках застосовують криві зростання?

5 Для чого застосовують лінеаризацію в економетричному аналізі?

6 У чому полягає відмінність двох розглянутих класів нелінійних регресійних моделей?

Список літератури

1 Волошин О., Галайко Н. Економетрія: навч. посіб. Львів: Львівський державний університет внутрішніх справ, 2012. Ч. 1. 192 с.

2 Диха М., Мороз В. Економетрія: навч. посіб. Київ: Центр учбової літератури, 2021. 206 с.

3 Єрьоменко В., Алілуйко А., Березька К. Економетрика: навч. посіб. Тернопіль: Підручники і посібники, 2023. 168 с.

4 Єрьоміна М. О. Економетрика: конспект лекцій. Харків: УкрДАЗТ, 2015. 43 с.

5 Єрьоміна М. О. Методичні вказівки до курсової роботи з дисципліни «Економетрика» для студентів економічних спеціальностей усіх форм навчання. Харків : УкрДУЗТ, 2016. 27 с.

6 Васильєва Н. К., Мороз С. І. Моделі оптимізації та економетрики – застосування в аграрній економіці: навч. посіб. Дніпро: Видавець Біла К. О., 2023. 190 с.

7 Наконечний С. І., Терещенко Т. О., Романюк Т. П. Економетрія: підручник. Вид. 3-тє, доп. і перероб. Київ: КНЕУ, 2004. 520 с.

8 Новосад Р. Г., Селіверстов І. І. Кількісні методи експертного оцінювання: наук.-метод. розробка / уклад. В. П. Артим. Київ: НАДУ, 2009. 36 с.

9 Економетрика в електронних таблицях: навч. посіб. / Н. К. Васильєва, О. А. Мироненко, Н. М. Самарець, Н. О. Чорна; за заг. ред. Н. К. Васильєвої. Дніпро: Біла К. О., 2017. 149 с.

10 On the founding of the econometric society. URL: https://www.cambridge.org/core/journals/journal-of-the-history-of-economic-thought/article/on-the-founding-of-the-econometric-society/BCF0EF9A40C8DDE863DADD4174140A91.

11 Box G. E. P. & Jenkins G. M. (1970). Time Series Analysis: Forecasting and Control. San Francisco: Holden-Day. 240 p.

О. М. Тройнікова

ЕКОНОМЕТРИКА

Конспект лекцій

Відповідальний за випуск Тройнікова О. М.

Редактор Ібрагімова Н. В.

Підписано до друку 11.03.2026 р.
Умовн. друк. арк. 3,5. Тираж . Замовлення № .

Видавець та виготовлювач Український державний університет
залізничного транспорту,
61050, Харків-50, майдан Фейєрбаха,7.
Свідоцтво суб'єкта видавничої справи ДК № 6100 від 21.03.2018 р.