МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ ЕКОНОМІЧНИЙ УНІВЕРСИТЕТ
ІМЕНІ СЕМЕНА КУЗНЕЦЯ

ЕКОНОМЕТРИКА

Навчальний посібник

Харків. ХНЕУ ім. Кузнеця, 2015

УДК 330.43(075)
ББК 65в611я7
Е 45

Рецензенти: докт. екон. наук, професор, завідувач кафедри економічної кібернетики та прикладної економіки Харківського національного університету ім. В. Н. Каразіна *Меркулова Т. В.*; докт. екон. наук, доцент кафедри вищої математики Буковинського державного фінансово-економічного університету *Якимова Л. П.*

Рекомендовано до видання рішенням вченої ради Харківського національного економічного університету імені Семена Кузнеця.
Протокол № 11 від 18.05.2015 р.

Авторський колектив: Гур'янова Л. С.
Клебанова Т. С.
Сергієнко О. А.
Прокопович С. В.

Е 45 Економетрика : навчальний посібник для студентів напряму підготовки "Економічна кібернетика" всіх форм навчання / Л. С. Гур'янова, Т. С. Клебанова, О. А. Сергієнко, С. В. Прокопович. – Х. : ХНЕУ ім. С. Кузнеця, 2015. – 384 с. (Укр. мов.)
ISBN 978-966-676-615-4

Розглянуто основні положення економетричного моделювання як методу наукового пізнання. Досліджено методи побудови загальної лінійної моделі, методи оцінювання ступеня мультиколінеарності та її усунення, методи побудови моделей в умовах автокореляції, гетероскедастичності залишків, нелінійні економетричні моделі, виробничі функції, економетричні моделі динаміки, моделі розподіленого лагу, системи одночасних рівнянь. Подано теоретичний матеріал і демонстраційні приклади, що дозволяють засвоїти зміст та методику застосування економетричних методів і моделей для дослідження економічних процесів. Наведено запитання для самодіагностики; тести; задачі для самостійного розв'язання за темами, глосарій.

Рекомендовано для студентів економічних спеціальностей, аспірантів, які проводять дослідження, пов'язані із завданнями економетричного моделювання.

УДК 330.43(075)
ББК 65в611я7

ISBN 978-966-676-615-4

Вступ

Сучасні умови функціонування економічних систем характеризуються невизначеністю, динамічністю, появою нових зв'язків, що призводить до істотних ризиків прийняття управлінських рішень. У цих умовах формування обґрунтованих управлінських рішень неможливе без застосування кількісних методів аналізу, основу яких становить економетрика. Економетричні методи й моделі дозволяють визначити ключові змінні, досліджувати природу їх впливу, провести аналіз механізмів економічних процесів, розробити науково обґрунтований прогноз, визначити найбільш імовірні сценарії розвитку досліджуваної економічної системи.

У зв'язку з цим навчальна дисципліна *"Економетрика"* є однією з базових дисциплін економіко-математичного циклу, має одночасно теоретичне, методологічне й прикладне значення. Дана навчальна дисципліна містить знання про якісні властивості економічних систем, про оцінювання взаємозв'язків показників розвитку економіки й економетричні моделі економічних систем і процесів.

Метою вивчення навчальної дисципліни є побудова економетричних моделей, що кількісно описують взаємозв'язки між економічними змінними.

Завданням навчальної дисципліни є засвоєння студентами основних принципів, методів та інструментарію щодо постановки завдань економетричного моделювання, методів їх розв'язання та аналізу з метою широкого використання в економіці та підприємництві.

Пререквізитами дисципліни "Економетрика" є загальна економічна теорія, макро- і мікроекономіка, теорія ймовірностей і математична статистика, теорія випадкових процесів. Навчальна дисципліна забезпечує вивчення таких дисциплін: "Прогнозування соціально-економічних процесів", "Теорія економічного ризику", "Методи економіко-статистичних досліджень", "Моделі економічної динаміки", "Нелінійні моделі економічної динаміки".

Вивчення навчальної дисципліни *"Економетрика"* передбачає формування у студентів таких *компетентностей:*

здатність проводити аналіз причинно-наслідкових зв'язків об'єкта дослідження, здійснювати постановку мети аналізу, визначати факторні та результативні ознаки, здійснювати оптимізацію складу інформаційних джерел, вибір методів аналізу;

здатність застосовувати економетричні методи для побудови моделей аналізу та прогнозування соціально-економічних систем, зокрема, методи оцінювання параметрів лінійних та нелінійних моделей, методи оцінювання параметрів в умовах мультиколінеарності, автокореляції, гетероскедастичності, наявності лагових змінних, двосторонніх причинно-наслідкових зв'язків змінних, методи дослідження часових рядів;

здатність визначати можливості застосування економетричних моделей для імітації та прогнозування, формувати сценарії розвитку соціально-економічних систем за різних умов, здійснювати аналіз тенденцій розвитку соціально-економічних систем, прихованих механізмів економічних процесів для прийняття ефективних управлінських рішень.

Навчальний посібник складено відповідно до програми дисципліни та містить інформацію за такими темами: "Економетричне моделювання як метод наукового пізнання", "Методи побудови загальної лінійної моделі", "Мультиколінеарність та її вплив на оцінки параметрів моделі", "Узагальнений метод найменших квадратів", "Побудова моделі з автокорельованими залишками", "Емпіричні методи кількісного аналізу на основі статистичних рівнянь", "Економетричні моделі динаміки", "Моделі розподіленого лага", "Економетричні моделі на основі системи структурних рівнянь". У рамках кожної теми поданий теоретичний матеріал; демонстраційні приклади, що дозволяють засвоїти зміст та методику застосування економетричних методів і моделей для дослідження економічних процесів; запитання для самодіагностики; тести; задачі для самостійного розв'язання; ключові слова. Особливу увагу приділено програмним засобам економетричного моделювання. Навчальний посібник містить лабораторний практикум з попереднього аналізу даних, побудови множинних лінійних і нелінійних моделей, оцінювання ступеня мультиколінеарноісті та її усунення, побудови економетричних моделей динаміки, моделей з лаговими змінними, систем одночасних рівнянь за допомогою пакета прикладних програм *Statistica*.

Навчальний посібник призначений для студентів напряму підготовки 6.030502 "Економічна кібернетика", а також для студентів інших економічних спеціальностей, які вивчають теми "Парна лінійна регресія", "Множинна лінійна регресія", "Проблема мультиколінеарності у регресійних моделях", "Узагальнені економетричні моделі", а також для аспірантів, які виконують дослідження, пов'язані із завданнями економетричного моделювання.

Розділ 1. Економетричне моделювання як метод наукового пізнання

1.1. Історія розвитку економетрики.

1.2. Визначення економетрики. Приклади економетричних досліджень.

1.3. Поняття і класифікація економетричних моделей. Етапи побудови економетричної моделі.

1.1. Історія розвитку економетрики

Економетрика як самостійний науковий напрям виникла в 20-х – 30-х роках XX століття. Термін "економетрія" вперше був використаний у 1910 р. польським бухгалтером Павлом Сіомпой, який позначив мету економетрії як математичний опис рядів економічних даних та їх відображення в графічній формі. Буквально слово "економетрика" означає вимір економіки. Норвезький економіст-лауреат Нобелівської премії з економіки Рагнар Фріш у своїх роботах підкреслював міждисциплінарний характер економетрики. Як редактор-засновник журналу "Економетрика", в 1933 р. у першому номері цього журналу Фріш визначає економетрику більш широко: "Економетрика – це ні в якому разі не те ж саме, що й економічна статистика. Вона зовсім не ідентична тому, що називається загальною економічною теорією, хоча значна частка цієї теорії має виражено кількісний характер. Також економетрика не повинна сприйматися як синонім застосування математики в економіці. Досвід показує, що і статистика, і економічна теорія, і математика, взяті окремо, є необхідними, але недостатніми для дійсного розуміння кількісних відносин у сучасному економічному просторі. Саме об'єднання всіх трьох частин дає потужний ефект. Саме це об'єднання і становить економетрику" [2].

Велика депресія 1929 – 1933 рр., що зачепила економічно розвинені країни світу, показала необхідність розгляду економіки як системи, що включає певні елементи і взаємозв'язки між ними. Складність економіки як об'єкта управління, підвищення невизначеності зовнішнього середовища внаслідок появи нових форм зв'язків економік різних країн, послужило імпульсом для розвитку економетричних досліджень.

У 1930 р. утворено Економетричне товариство, засновниками якого виступили І. Фішер, Р. Фріш, О. Андерсон, О. Моргенштерн та ін. У 1928 р. Ч. Коббом і П. Дугласом була побудована виробнича функція, яка стала однією з найпопулярніших економетричних моделей. У 1930 – 1939 рр. Я. Тінбергеном розпочато макроекономічне моделювання, яке спиралося на теорію Дж. Кейнса і розробку системи національних рахунків у США та інших країнах. У 1955 р. Л. Клейном і А. Голдбергером була побудована одна з перших комплексних економетричних моделей. У 1969 р. Р. Фріш і Я. Тінберген стали першими дослідниками, які отримали Нобелівську премію з економіки за створення і застосування динамічних моделей до аналізу економічних процесів. У 1980 р. другу економетричну Нобелівську премію з економіки отримав Л. Клейн за створення економічних моделей і їх застосування до аналізу коливань економіки і економічної політики.

Розробка методів оцінювання параметрів системи одночасних рівнянь Хаавельмо Т. М., Купмансом Т. Ч., Андерсоном Т. В., Клейном Л. Р. та іншими ознаменувала остаточне створення спеціальних економетричних методів.

У цей же час активно розвивалася не тільки макро-, але і мікроеконометрика. Основоположниками цього напряму є Д. Хекман і Д. Макфадден, які розробили теорію і методи, які широко використовуються в статистичному аналізі поведінки індивідуумів і домогосподарств. Так, Дж. Хекман вирішив проблему зсуву вибірки через селективність даних. Основний внесок Д. Макфаддена полягає у розвитку методів для аналізу дискретного вибору. У 1974 р. він розробив логіт-аналіз, який був визнаний фундаментальним досягненням економічної науки; економетричні методи для оцінювання виробничих технологій і дослідження факторів, що формують попит фірм на капітал і робочу силу.

Вирішальною подією для розвитку економетрики стала поява комп'ютерів. Завдяки їм потужний розвиток отримав статистичний аналіз часових рядів. Дж. Бокс і Г. Дженкінс створили ARIMA-модель у 1970 р., а К. Сімс та інші – VAR-моделі на початку 1980-х рр.

Розвиток фінансових ринків та похідних фондових інструментів призвело до виникнення спеціальних методів аналізу часових рядів. У 1970-х рр. були створені моделі лінійної фільтрації. Наступним кроком стала розробка методів, спрямованих на виявлення "шоків", тестування рядів на стаціонарність, дослідження коінтегрованості часових рядів, розроблення моделей з мінливою в часі волатильністю і т. д. У цій області слід зазначити заслуги Р. Інгла і К. Гренджера, які стали лауреатами Нобелівської премії в 2003 р.

Таким чином, можна виділити такі основні частини економетрики як наукової дисципліни (рис. 1.1) [27; 31].

Рис. 1.1. **Основні частини економетрики**

Сьогодні економетричні методи застосовуються в різних галузях прикладної економіки, починаючи з дослідження витрат домашніх господарств і підприємницьких інвестицій і закінчуючи організацією виробництв, ринків праці, ефектів державної політики. Економетричні дослідження ведуться у всіх значущих наукових центрах, що займаються проблемами економіки. Різноманітність спектра завдань, які вирішуються за допомогою економетрики, можна умовно розподілити на три напрями:

за кінцевими прикладними цілями – прогноз економічних і соціально-економічних показників, що характеризують стан і розвиток аналізованої системи; розробка різних можливих сценаріїв розвитку аналізованої системи;

за рівнем ієрархії аналізованої економічної системи – дослідження може бути спрямоване на аналіз економіки в цілому (макрорівень), регіону, галузі (мезорівень) і підприємства, фірми, сім'ї (мікрорівень);

за профілем аналізованої економічної системи – дослідження може бути сконцентровано на аналізі кон'юнктури ринку, інвестиційної, фінансової або соціальної політики, ціноутворення, попиту та пропозиції і т. д.

1.2. Визначення економетрики. Приклади економетричних досліджень

Із сучасних позицій економетрику можна розглядати як науку про моделювання економічних явищ, що дозволяє пояснювати і прогнозувати їх розвиток, виявляти і вимірювати визначальні чинники.

На основі економічної теорії розробляються концепції розвитку досліджуваних процесів; за допомогою статистики ці процеси виражаються в статистичних показниках; математико-статистичні методи дозволяють будувати моделі досліджуваних процесів, оцінювати їх параметри, ступінь відповідності реальним даним і, нарешті, прогнозувати розвиток досліджуваного явища.

Економетрика – один із напрямів економіко-математичних методів аналізу, який полягає в статистичному вимірюванні (оцінюванні) параметрів математичних виразів, що характеризують деяку економічну концепцію про взаємозв'язок і розвиток об'єкта, явища, що необхідно для отримання конкретних економічних висновків на основі економетричних моделей.

Об'єктом навчальної дисципліни "Економетрика" є сукупність різних соціально-економічних процесів, що відбуваються в економічній системі.

Предметом навчальної дисципліни "Економетрика" є економетричні методи і моделі, що дозволяють визначити і вивчити кількісні взаємозв'язки між соціально-економічними явищами.

Слід розглянути приклади економетричних досліджень.

Приклад у галузі вимірювання залежності "ціна – якість" на часовому проміжку [2].

У кінці 1930-х рр. у зв'язку з великими масштабами виробництва *General Motors (GM)* і наявністю істотного циклічного безробіття в країні, в Конгресі США та інших громадських і політичних структурах велися гострі дебати з приводу того, чи слід вимагати від *General Motors* встановлення меж варіювання цін, щоб стабілізувати рівень виробництва і зайнятості. Деякі критично налаштовані спостерігачі стверджували, що *GM* безсоромно користується монопольним становищем і що питання зайнятості повинні вирішуватися більш конструктивно. На підтвердження цього зазначалося, що в період 1925 – 1935 рр. індекс цін на автомобілі за даними Американського бюро статистики праці (БСП), зріс для машин *GM* на 45 %; більше того індекс обсягу продажів нових автомобілів значно коливався з року в рік. Ці зауваження стурбували перспективою урядового втручання компанію *General Motors*, в 1938 р. вона профінансувала дослідження Е. Корта з Асоціації виробників автомобілів щодо оцінювання впливу зміни цін на автомобілі на загальний обсяг їх продажів.

Е. Корт запропонував оцінити параметри такого рівняння регресії:

$$lnP_i = a_0 + a_1 \cdot D_{1926,i} + a_2 \cdot D_{1927,i} + \beta_1 \cdot WT_i + \beta_2 \cdot LH_i + \beta_3 \cdot HP_i + u_i,$$

де P_i – ціна i-ї моделі машини;

α, β – невідомі параметри, що підлягають оцінюванню;

$D_{1926,i}$ – фіктивна змінна, що набуває значення 1 для 1926 р. і 0 – в усі інші роки;

$D_{1927,i}$ – фіктивна змінна, що набуває значення 1 для 1927 р. і 0 – в усі інші роки;

WT_i – вага i-ї моделі машини у фунтах;

LH_i – довжина осі колеса в дюймах;

HP_i – заявлена потужність;

u_i – однаково розподілені незалежні нормальні випадкові величини з нульовим середнім і дисперсією σ^2.

Дані, які були необхідні для оцінювання параметрів рівняння, складалися з вибірки спостережень за різними моделями машин, проданих протягом трьох років. У ході цього такі змінні, як вага машини, довжина осі колеса, заявлена потужність розглядалися як характеристики якості автомобіля. Оцінки коефіцієнтів при фіктивних змінних відображають зміну в часі натурального логарифма індексу цін за умови фіксованої якості.

Рівняння ціни, виведене Е. Кортом, дозволяє перевірити ряд статистичних гіпотез. Одна з гіпотез формулюється так: якість не відіграє ніякої ролі, тобто розмах ціни не пов'язаний зі зміною якості. Цю гіпотезу можна перевірити емпірично, розглядаючи спільну нульову гіпотезу $\beta_1 = \beta_2 = \beta_3 = 0$ проти альтернативної гіпотези $\beta_1 \neq \beta_2 \neq \beta_3 \neq 0$.

Інша гіпотеза полягає в тому, що всі відхилення ціни пов'язані з відхиленнями якості і, отже, протягом цього періоду індекс цін із поправкою на якість залишався незмінним, і була відсутня інфляція. Це гіпотеза $\alpha_1 = \alpha_2 = 0$, у той час як альтернатива їй $\alpha_1 \neq \alpha_2 \neq 0$.

Таким чином, наведена модель дозволяє перевести "проблему якості" в кількісну категорію.

Зрештою, результати досліджень показали, що в той час, як офіційний індекс цін БСП виріс на 45 %, запропонований Е. Кортом індекс цін із поправкою на якість скоротився приблизно на 55 % (тобто майже до 20 %). Керівництво *General Motors* використовувало ці результати нарівні з іншими даними під час підбору аргументів на користь того, що виробники автомобілів уже скоротили ціни з поправкою на якість і що подальше їх зниження, спрямоване на зниження безробіття, може призвести до банкрутства, оскільки рівень беззбитковості збільшиться набагато більше, ніж попит, викликаний зниженням ціни.

Приклад у галузі вимірювання залежності "рівень інфляції – рівень безробіття"[40].

У 1958 р. О. Філліпс, спираючись на емпіричні дані по Англії за 1861 – 1957 рр., вивів кореляційну залежність між рівнем безробіття і зміною приросту грошової заробітної плати. Ця залежність була подана у вигляді кривої, що характеризує функціональний зв'язок цих двох величин: чим вище безробіття, тим менше приріст грошової заробітної плати, тим нижче зростання цін; і навпаки, чим нижче безробіття і вище зайнятість, тим більше приріст грошової заробітної плати, тим вище темп зростання цін. "Криві Філліпса" були покладені в основу кейнсіанської теорії інфляції, в якій зв'язок між рівнем зайнятості та інфляцією встановлювалася через динаміку грошової заробітної плати. У загальному вигляді залежність, відображена кривої Філліпса, має вигляд:

$$\pi_t - \pi_{t-1} = -\beta \cdot (u_t - u_t^e) + \varepsilon_t,$$

де π_t – рівень інфляції в момент часу t;

u_t – рівень безробіття в момент часу t;

u_t^e – рівень природного безробіття;

ε_t – вплив шокових змін;

β – коефіцієнт або параметр моделі.

Графік залежності, виражений кривою Філліпса, поданий на рис. 1.2.

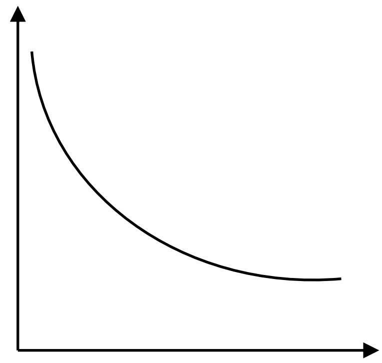


Рис. 1.2. **Крива Філліпса**

Можна зауважити, що не завжди збільшення рівня безробіття призводить до зниження інфляції, для ряду країн спостерігаються такі періоди, коли обидва ці показники зростають одночасно. Тому крива Філліпса вивчалася для різних економічних систем і в різні періоди часу. Так, історія інфляції і безробіття в США 60-х років XX століття досить точно відображалася цією кривою. Однак аналіз статистичних даних першої половини 70-х років XX століття свідчить, що крива Філліпса не завжди є справедливою. У 1973 – 1975 рр. у США спостерігалося одночасне збільшення інфляції та безробіття. У 1976 – 1979 рр. розвиток економіки відбувався згідно з кривою Філліпса.

Приклад у галузі вимірювання залежності "бюджетні надходження – ставка податку" [30].

Залежність, що відображає взаємозв'язок між бюджетними надходженнями і ставками податку на прибуток і заробітну плату, була досліджена американським економістом Артуром Лаффером у 80-х роках XX століття. Теорія А. Лаффера припускає розширення податкової бази за умови зниження ставок податків. При цьому передбачається дотримання принципу пропорційності – економічної збалансованості інтересів платників податків і доходів бюджету, що виявляється через співвідношення його наповнюваності та наслідків оподаткування для платника.

Абстрактна модель А. Лаффера, що моделює залежність зростання ставок і податкових надходжень, показує, що лібералізація податкового законодавства до певної межі здатна привести до збільшення податкової бази, а розширення бази стимулює зростання податкових відрахувань. Найбільш спірним питанням у цьому контексті є пошук точки оптимуму, понад яку зниження податкових ставок веде до зворотного зниження податкових надходжень. Для економік різних країн положення даної точки знаходиться експериментально й істотно розрізняється. Графічно форма кривої Лаффера для стандартного варіанта залежності наведена на рис. 1.3.

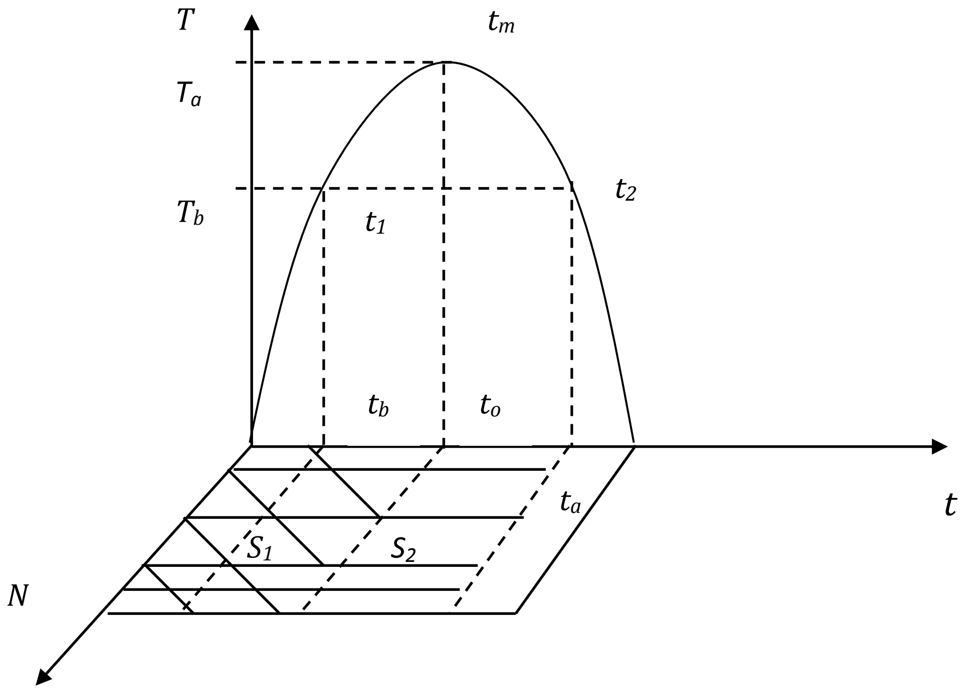


Рис. 1.3. **Графічне подання кривої Лаффера**

У даний час методологія моделювання виробничо-фіскальних ефектів знайшла найбільш повне відображення в уявленні про "поділ" впливу податків на дві складові. Перша з них пов'язана з вивченням виробничої кривої $Y = Y(q)$, в полі координат "податкове навантаження (q) – обсяг випуску *(Y)".* У цю функцію входять два фактори – праця і капітал, а значення податкового навантаження присутні у вигляді ступеня в нелінійній залежності. Загальний вигляд моделі можна подати таким чином:

$$Y = \gamma \cdot D \cdot K^{q\cdot(a+b\cdot q)} L^{q\cdot(n+m\cdot q)},$$

де Y – випуск (обсяг ВВП);

D – трендовий оператор, який набуває вигляду $D = e^{\beta t}$;

t – період часу;

K – капітал;

L – чисельність зайнятих;

$\beta,\ \gamma,\ \alpha,\ b,\ m,\ n$ – параметри моделі;

q – податкове навантаження.

У цілому, оцінювання параметрів моделі дає можливість додаткового дослідження точок екстремуму функції. Дана крива досягає локального максимуму в точці q^*, яка називається точкою Лаффера 1-го роду і для якої виконуються такі умови:

$$\frac{\partial Y(q^*)}{\partial q} = 0, \frac{\partial^2 Y(q^*)}{\partial q} < 0.$$

Тобто це таке значення податкового навантаження, за якого досягається максимальний випуск (ВВП). Економічно точка Лаффера 1-го роду означає межу податкового тягаря, відповідно до якого виробнича система ще не переходить у режим рецесії.

Друга складова пов'язана з побудовою фіскальної кривої $T = T(q)$ у координатному полі "податкове навантаження – обсяг податкових надходжень". Фіскальна крива, тобто залежність між доходами держави у вигляді податків і фіскальної навантаженням, описується моделлю, що має специфікацію такого вигляду:

$$T = \gamma \cdot D \cdot K^{q\cdot(a+b\cdot q)} L^{q\cdot(n+m\cdot q)},$$

де T – податкові надходження.

Таке значення податкового навантаження, за якого у фіскальній функції максимальні прибутки держави (податкові надходження) називають точкою Лаффера 2-го роду. В цілому значення фіскальної функції можливо отримати шляхом множення виробничої функції і величини податкового навантаження. Для точки Лаффера 2-го роду виконуються такі умови:

$$\frac{\partial T(q^{**})}{\partial q} = 0, \frac{\partial^2 T(q^{**})}{\partial q} < 0.$$

Точка Лаффера 2-го роду вказує величину податкового тягаря, за межами якого зростання маси податкових надходжень стає неможливим. Ідентифікація точок Лаффера 1-го і 2-го родів та їх зіставлення з фактичним і номінальним податковими тягарями дозволяє оцінити ефективність кількісного налаштування податкової системи та визначити напрями його оптимізації.

Сенс розгляду двох точок Лаффера полягає в тому, що між регулюючою та фіскальною функціями податкової системи завжди існує певний антагонізм: сприяючи поповненню бюджету країни, збільшення податкового тягаря зменшує ділову активність економічних агентів, погашає стимули до розширення виробництва. У зв'язку з цим основна проблема фіскальної політики полягає в тому, щоб знайти компроміс між інтересами виробника і бюджету.

Наведені приклади економетричних досліджень свідчать про широке застосування економетричних методів для обґрунтування економічних законів і положень. Саме *синтез математики, економіки та статистики* зробив економетрію потужним інструментом для теоретичного і прикладного аналізу різних соціально-економічних систем.

Соціально-економічна система може бути подана нескінченним числом структурних та функціональних інваріантів, що відображають взаємозв'язки між різними процесами, які відбуваються в цій системі (економічними, соціальними, екологічними, демографічними і т. д.). Опис системи проводиться за допомогою її якісних і кількісних характеристик, іменованих параметрами. Параметри становлять основу мов опису систем, а під час формалізації ототожнюються з незалежними змінними математичного опису процесу функціонування систем.

У ході побудови економетричної моделі реалізується *метод моделювання за принципом "чорного ящика"*, *коли досліднику не відомий механізм процесів, що відбуваються у системі, вивчити який можна за вхідними і вихідними характеристиками системи. Вхідні* і *вихідні* характеристики системи часто ототожнюють з *екзогенними* та *ендогенними* змінними, або в кореляційно-регресійному аналізі вживають терміни незалежні (факторні) змінні, або ознаки, і залежні (результативні) змінні, або ознаки. Графічно принцип "чорного ящика" зображений на рис. 1.4.

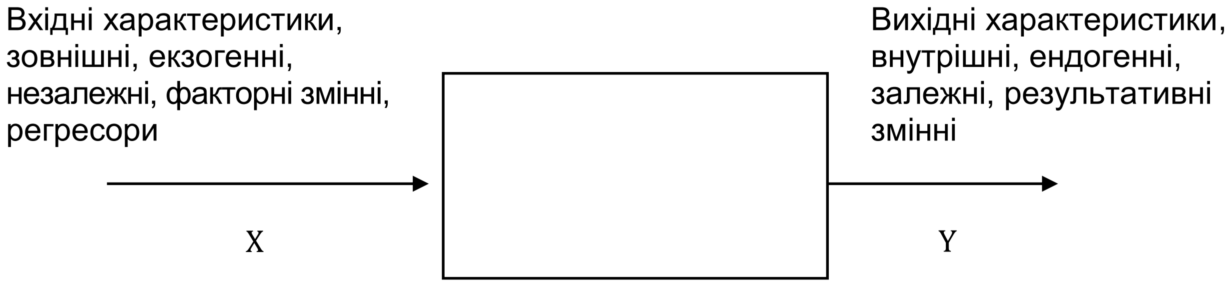


Рис. 1.4. **Дослідження системи за принципом "чорного ящика"**

Досліднику необхідно виділити вхідні і вихідні характеристики і на підставі економетричних методів встановити характер причинно-наслідкових зв'язків, що лежать в основі механізму функціонування соціально-економічної системи. Заключним етапом у процедурі послідовної формалізації опису процесів функціонування соціально-економічних систем є вираження залежності у вигляді деяких математичних співвідношень (функцій, систем рівнянь і т. д.), що відображають зв'язок між певними явищами.

1.3. Поняття і класифікація економетричних моделей. Етапи побудови економетричної моделі

Економетрична модель – це особливий клас економіко-математичних моделей, в яких дослідник вирішує цілий ряд завдань:

1. Вибір форми математичної залежності, яка описує поведінку економічного об'єкта, на основі системи спостережень.
2. Оцінювання параметрів даної моделі різними методами (метод найменших квадратів, метод максимальної правдоподібності тощо).
3. Перевірка статистичної значущості моделі.

Часто економетрична модель у загальному вигляді подається як система лінійних рівнянь:

$$BY = AX + E,$$

де B – матриця коефіцієнтів при ендогенних змінних;
Y – вектор ендогенних змінних;
A – матриця коефіцієнтів при екзогенних змінних;
X – вектор екзогенних змінних;
E – вектор випадкових збурень (помилок, відхилень).

Економетричні моделі включають різні класи моделей. Наведемо одну з класифікацій економетричних моделей (рис. 1.5). *За способом математичного уявлення* економетричні моделі можна умовно розподілити на *прості* і *складні*. *Прості економетричні моделі* подані одним рівнянням, однією залежністю, *складні* – кількома рівняннями, кількома залежностями.

За кількістю факторних ознак, що включаються в модель, прості економетричні моделі можна розподілити на *однофакторні* і *багатофакторні.*

Однофакторні моделі містять одну незалежну ознаку, багатофакторні моделі – дві та більше незалежних ознак. *Однофакторні і багатофакторні моделі* можуть бути подані *лінійними* і *нелінійними* функціями. *Складні економетричні моделі* можуть бути подані трьома видами систем одночасних рівнянь, залежно від форми включення в праву частину ендогенних змінних. Зазвичай виділяють три типи систем: *системи, вирішені щодо ендогенних змінних; рекурсивні системи; системи, невирішені щодо ендогенних змінних.* Залежно від наявності (відсутності) в моделі фактора часу розрізняють динамічні і статичні моделі. Прикладами динамічних моделей є такі моделі часових рядів, як трендові моделі; моделі згладжування; моделі декомпозиції часового ряду; авторегресії – ковзного середнього.

Побудова економетричної моделі проводиться в кілька основних етапів.

Етап 1. Якісний аналіз (постановка мети аналізу, визначення сукупності, визначення результативних і факторних ознак, вибір періоду, за який проводиться аналіз, вибір методу аналізу).

Етап 2. Попередній аналіз модельованої сукупності (перевірка однорідності сукупності, виключення аномальних спостережень, уточнення необхідного обсягу ознак, установлення законів розподілу ознак).

Етап 3. Побудова економетричної моделі (встановлення переліку чинників, розрахунок оцінок параметрів рівнянь регресії, перебір конкуруючих варіантів моделі).

Етап 4. Оцінювання адекватності моделі (перевірка статистичної значущості рівняння залежності в цілому і його окремих параметрів; перевірка відповідності формальних властивостей оцінок завданням дослідження).

Етап 5. Економічна інтерпретація і практичне використання моделі.

Під час побудови економетричної моделі у випадку, якщо *верифікація (оцінювання точності)* показує, що модель не є адекватною, то здійснюється повернення до вибору специфікації моделі. Виділяють такі типи помилок специфікації економетричної моделі: в моделі помилково враховано чинник, який не робить істотного впливу на результуючу змінну; в моделі не врахований фактор, що робить істотний вплив на результуючу змінну; обрана неправильна форма математичної залежності, яка описує взаємозв'язок між досліджуваними змінними.

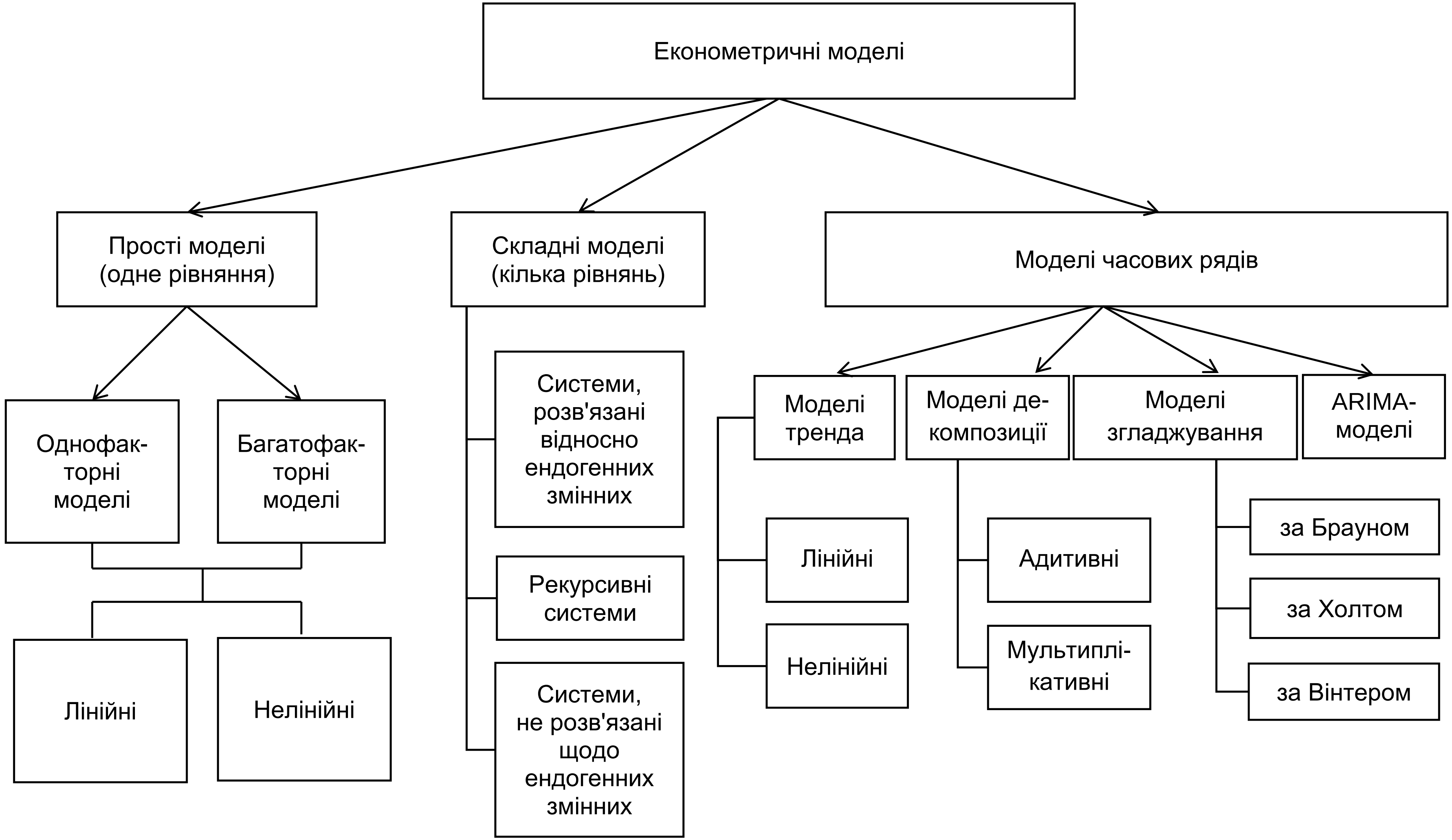


Рис. 1.5. **Класифікація економетричних моделей**

Якщо модель задовольняє висунуті вимоги і статистично значуща, то вона може бути використана або для прогнозування, або для пояснення (аналізу) прихованих механізмів досліджуваних процесів.

Контрольні запитання для самодіагностики

1. Дайте визначення поняттю "економетрика". Що є передумовою для розвитку економетрики?

2. Наведіть приклади використання економетричних моделей для вивчення економічних процесів.

3. У чому особливості економетричного моделювання?

4. Сформулюйте визначення "економетрична модель".

5. Які економетричні методи використовуються під час побудови й аналізу економетричних моделей?

6. Наведіть класифікацію економетричних моделей.

7. Які моделі належать до моделей часових рядів? У чому їх особливості?

8. Назвіть основні типи систем одночасних рівнянь. Наведіть приклади застосування комплексних моделей в економіці.

9. Назвіть етапи побудови економетричних моделей.

10. У чому сутність якісного аналізу у ході побудови економетричної моделі?

Тести

1. Економетрика – це:

а) наукова дисципліна, що вивчає можливості математичного опису часових рядів економічних індикаторів;

б) один із напрямів економіко-математичних методів аналізу, який полягає в статистичному вимірюванні (оцінюванні) параметрів математичних виразів, що характеризують деяку економічну концепцію про взаємозв'язок і розвиток об'єкта, явища;

в) наукова дисципліна, що застосовує математичні методи в економіці.

2. Предметом економетрики як наукової дисципліни є:

а) методи і моделі, що дозволяють визначити і вивчати кількісні взаємозв'язки між соціально-економічними явищами;

б) моделі, які дозволяють вивчати загальні властивості економіки;

в) моделі, які дозволяють досліджувати жорсткі функціональні зв'язки між економічними змінними.

3. Економетрична модель може бути використана для:

а) аналізу взаємозв'язку економічних змінних;

б) визначення чинників, що роблять найбільш значущий вплив на досліджуваний процес;

в) прогнозування економічних процесів;

г) усі відповіді правильні.

4. Економетрична модель є:

а) стохастичною;

б) детермінованою;

в) концептуальною;

г) структурною.

5. У загальному вигляді економетрична модель може бути подана таким чином:

а) $BY = AX + E$;

б) $BX = AY + E$;

в) $BX = AX + E$.

6. Ендогенні змінні – це:

а) результуючі змінні;

б) факторні (незалежні) змінні;

в) специфічні змінні, які вводяться в модель для урахування якісних характеристик економічних процесів.

7. Виберіть правильну послідовність етапів побудови економетричної моделі:

а) якісний аналіз; попередній аналіз; оцінювання параметрів моделі; оцінювання адекватності; практичне використання моделі;

б) практичне використання моделі; якісний аналіз; оцінювання адекватності;

в) попередній аналіз; якісний аналіз; оцінювання параметрів моделі; оцінювання адекватності; практичне використання моделі.

8. Серед наведених далі моделей виберіть ту, яка відноситься до моделей часових рядів:

а) $y = a_0 + a_1 \cdot t + a_2 \cdot t^2 + \varepsilon$;

б) $y = a_0 + a_1 \cdot x + a_2 \cdot x^2 + \varepsilon$;

в) $y = a_0 + a_1 \cdot x_{t-1} + a_2 \cdot x_{t-2} + \varepsilon$.

9. Перевірка однорідності сукупності здійснюється на етапі:

а) якісного аналізу;

б) попереднього аналізу;

в) оцінювання параметрів моделі;

г) оцінювання адекватності моделі.

10. Основними складовими економетрики як наукової дисципліни є:

а) економетричні методи;

б) економетричні моделі;

в) усі відповіді правильні.

11. Серед наведених далі моделей виберіть ту, яка належить до складних економетричних моделей:

а) $y = a_0 + a_1 \cdot x_1^{b1} + a_2 \cdot x_2^{b2} + \varepsilon$;

б) $\begin{aligned} y_1 &= a_1 \cdot x_1 + a_2 \cdot y_3 + \varepsilon_1 \\ y_2 &= a_3 \cdot x_1 + a_4 \cdot y_1 + \varepsilon_2 \\ y_3 &= a_5 \cdot x_2 + a_6 \cdot y_2 + \varepsilon_3 \end{aligned}$;

в) $y = a_0 + a_1 \cdot x_1 + a_2 \cdot x_2 + a_3 \cdot x_3 + \varepsilon$.

12. Формування гіпотези про склад факторних ознак економетричної моделі здійснюється на етапі:

а) попереднього аналізу;

б) якісного аналізу;

в) оцінювання адекватності моделі.

13. Серед наведених далі моделей виберіть ту, яку відносять до нелінійних багатофакторних моделей:

а) $y = a_0 + a_1 \cdot x + a_2 \cdot x^2 + \varepsilon$;

б) $y = a_0 \cdot x_1^{a_1} \cdot x_2^{a_2} + \varepsilon$;

в) $y = k + a_0 \cdot a_1^x + \varepsilon$.

14. Перевірка статистичної значущості рівняння залежності в цілому і його окремих параметрів здійснюється на етапі:

а) якісного аналізу;

б) інтерпретації рівняння регресії;

в) оцінювання адекватності моделі.

15. Серед наведених далі складних економетричних моделей виберіть ту, яка належить до класу рекурсивних систем:

а) $\begin{aligned} y_1 &= a_1 \cdot x_1 + a_2 \cdot x_3 + \varepsilon_1 \\ y_2 &= a_3 \cdot x_4 + a_4 \cdot y_1 + \varepsilon_2 \\ y_3 &= a_5 \cdot x_2 + a_6 \cdot y_2 + \varepsilon_3 \end{aligned}$; б) $\begin{aligned} y_1 &= a_1 \cdot x_1 + a_2 \cdot y_3 + \varepsilon_1 \\ y_2 &= a_3 \cdot x_4 + a_4 \cdot y_1 + \varepsilon_2 \\ y_3 &= a_5 \cdot x_2 + a_6 \cdot y_2 + \varepsilon_3 \end{aligned}$.

Практичні завдання

1. Розглядається завдання прогнозування та виявлення факторів, які мають найбільш сильний вплив на вартість житлової, офісної, складської

нерухомості. Сформувати перелік факторних ознак (x_1, x_2, …, x_m), які, на вашу думку, впливають на вартість нерухомості різного функціонального призначення (y_1 – вартість 1 м^2 житлової нерухомості, y_2 – вартість 1 м^2 офісної нерухомості, y_3 – вартість 1 м^2 складської нерухомості). Обґрунтувати тип зв'язку між змінними (прямий зв'язок – зі збільшенням факторної змінної значення результуючої ознаки збільшується, зворотний – зі збільшенням факторної змінної значення результуючої ознаки зменшується).

2. Необхідно побудувати модель залежності рівня фінансової безпеки підприємства (y) від величини бюджету страхування ризиків (x). Серед наведених у табл. 1.1 залежностей необхідно вибрати найбільш ймовірні для опису взаємозв'язку досліджуваних змінних. Обґрунтувати відповідь.

Таблиця 1.1

Види залежностей

Назва функції	Модель
Лінійна	$y = a_0 + a_1 \cdot x$
Показова	$y = a_0 \cdot a_1^x$
Модифікована показова	$y = k + a_0 \cdot a_1^x$

3. Треба сформулювати найбільш повний список змінних, що характеризують функціонування економічної системи (банку, туристичної фірми, інвестиційного фонду, страхової компанії, регіону тощо). Виділити серед них ендогенні та екзогенні змінні системи. Уявити причинно-наслідкові зв'язки між основними змінними у вигляді рекурсивної системи одночасних рівнянь.

Ключові слова

Економетрика. Предмет економетрики. Об'єкт економетрики. Складові економетрики. Економетрична модель. Екзогенні змінні. Ендогенні змінні. Класифікація економетричних моделей. Етапи побудови економетричної моделі.

Розділ 2. Методи побудови загальної лінійної моделі

2.1. Оцінювання параметрів парної лінійної регресії методом найменших квадратів.

2.2. Оцінювання значущості лінійної парної регресії та її параметрів.

2.3. Оцінювання параметрів множинної лінійної регресії.

2.4. Оцінювання тісноти та значущості зв'язку між змінними в рівнянні множинної регресії.

2.5. Стандартизована форма моделі множинної регресії.

2.1. Оцінювання параметрів парної лінійної регресії методом найменших квадратів

Одним з основних завдань під час побудови економетричної моделі є вибір форми математичної залежності, яка описує поведінку економічного об'єкта на основі системи спостережень. Спостерігаючи статистичний зв'язок між ознаками, можна наближено подати значення результативної ознаки у вигляді деякої функції від величини одного або декількох факторних ознак, прагнучи у ході цього, щоб спостережувані дані якомога чіткіше відтворювалися взятою функцією. *Функція, що відображає статистичний зв'язок між ознаками, називається рівнянням регресії.* Модель, що описує кореляційно-регресійний зв'язок між економічними показниками називається загальною. Якщо рівняння регресії пов'язує лише дві ознаки, це *рівняння парної регресії:*

$$Y = f(x).$$

Функція може бути як лінійною $y = a_0 + a_1 \cdot x$, так і нелінійною, наприклад, $y = a_0 + a_1 \cdot x + a_2 \cdot x^2$, $y = a_0 \cdot a_1^x$, $y = \frac{k}{1+a_1 a_2^x}$ і т. д.

Модель множинної регресії – це рівняння, що відображає кореляційний зв'язок між результатом і декількома факторами:

$$Y = f\ (x_1, x_2, \dots, x_m).$$

Існують особливості формування сукупності спостережень (вибірки) під час побудови регресійних рівнянь. Дані вибірки можуть бути отримані за об'єктами: страховими компаніями, інвестиційними компаніями, банками, фірмами, регіонами, країнами і так далі на певний момент часу. Це так звані *просторові вибірки* (*cross-sectional data*). А можуть бути сформовані як *часові (динамічні) ряди* (*time-series data*) – послідовності спостережень, впорядкованих у часі. *Просторово-часова вибірка* є комбінацією просторової та часової вибірок

Під час проведення кореляційно-регресійного аналізу передбачається, що спостереження були отримані з однієї сукупності одиниць. Тобто механізм впливу факторів приблизно однаковий. Для забезпечення статистичної достовірності моделі кількість спостережень має бути в 8 – 10 разів більше числа факторів, включених у модель. У процесі побудови економетричної моделі передбачається, що фактори впливають на результат, причому вплив одного фактора не залежить від впливу інших факторів. В іншому випадку зміна одного з факторів призведе як до прямого, так і опосередкованого (через інші фактори) впливу, що спричинить за собою помилки в інтерпретації результатів регресійного аналізу.

Найбільш простим і поширеним випадком є подання економетричної моделі у вигляді рівняння парної (або однофакторної) лінійної регресії. Лінійна економетрична однофакторна модель може бути наведена у вигляді

$$y = a_0 + a_1 \cdot x.$$

Оскільки лінійний зв'язок між ознаками X і Y– теоретичний, то над $Y = f(x)$ ставиться знак ^. Завдання полягає в тому, щоб, вибравши зазначену (лінійну) форму залежності, визначити параметри рівняння a_0 і a_1 так, щоб відхилення спостережуваних (реальних) значень ознаки Y_i від теоретичних значень $\hat{Y}_i$ були мінімальними. Ці відхилення також називають *збуреннями, похибками, залишками*.

Існує багато методів визначення оцінок параметрів моделі. Деякі методи обчислення оцінок на практиці використовуються, в основному, через свою простоту і наочність. Однак ці методи не завжди приводять до отримання "хороших" оцінок. Найбільш результативним підходом отримання оцінок параметрів є запропонований І. Фішером метод максимальної правдоподібності (ММП), заснований на функції правдоподібності. Суть ММП полягає в такому [25]:

Нехай відомі істинні значення параметрів a_j $j = \overline{0, m}$ і модель $y = f(X, a_0, \dots, a_m)$. Тоді можна отримати теоретичні значення $\hat{y} = f(X, a_0, \dots, a_m)$, обчислені за моделлю. Маючи модель, можна отримати *відхилення* $e_i = y_i - \hat{y}_i$. Ці відхилення називають *залишками, помилками, збурюваннями*.

Оскільки $\hat{y}_i$ залежать від параметрів a_j, то й відхилення будуть залежати від параметрів. Нехай відома функція щільності ймовірності для ε_i. Виникає завдання: якими мають бути значення параметрів, щоб отримані відхилення мали задані характеристики? Згідно з методом І. Фішера, на безлічі відхилень будується функція правдоподібності і параметри підбираються з умови максимізації цієї функції.

Окремим випадком методу правдоподібності є метод найменших квадратів (*МНК, Least Square Method*). Історично склалося так, що цей метод був розроблений і отримав теоретичне обґрунтування раніше, ніж метод максимальної правдоподібності. Тому в літературі і практичних застосуваннях МНК описується як самостійний метод визначення оцінок параметрів.

Слід розглянути суть *методу найменших квадратів (МНК)*. У методі найменших квадратів оцінки знаходяться з умови мінімізації суми квадратів відхилень фактичних значень результівної змінної від теоретичних, обчислених за моделлю. Обґрунтуванням вибору такої умови для визначення оцінок параметрів є теорема Гаусса, яка формулюється таким чином:

Теорема Гаусса. Якщо помилки ε_t некорельовані, тобто $cov(\varepsilon_j, \varepsilon_k) = 0$ у разі $j \neq k$, мають нульове математичне очікування $M(\varepsilon_t)$ і однакову постійну дисперсію σ^2, то оптимальними вибірковими оцінками параметрів є значення a_i, що мінімізують суму квадратів відхилень між спостережуваними значеннями y_i і значеннями, отриманими за моделлю $y_i = f(X, a_0, \dots, a_m)$, тобто суму квадратів:

$$F(a_0, \dots, a_m) = \sum_{i=1}^{n}(y_i - f(X, a_0, \dots, a_m))^2.$$

Для знаходження невідомих коефіцієнтів рівняння моделі використовуються методи регресійного аналізу, які базуються на таких припущеннях, що мають назву обмеження Гаусса – Маркова.

У класичній моделі регресійного аналізу передбачаються виконаними такі припущення (передумови регресійного аналізу):

Умова 1. Величини ε_j є випадковими.

Умова 2. Математичне сподівання залишків дорівнює нулю: $M(\varepsilon_j) = 0$.

Умова 3. Залишки ε_j та ε_k некорельовані: $M(\varepsilon_j, \varepsilon_k) = 0, j \neq k$.

Умова 4. Дисперсія залишків постійна для кожного спостереження j:$D(\varepsilon_j) = \sigma^2$. Ця умова називається умовою гомоскедастичності. Порушення цієї умови називається гетероскедастичністю.

Умова 5. Величини ε_j взаємно незалежні зі значеннями пояснювальних змінних.

Умова 6. Спільний розподіл випадкових величин $\varepsilon_1, \ldots, \varepsilon_n$ є нормальним.

Отримавши оцінки за допомогою МНК, треба перевірити здійсниміcть зазначених умов для відхилень:

$$e_i = y_i - (a_0 + a_1 \cdot x).$$

Математичне сподівання відхилень e_i буде завжди рівним нулю через специфіку методу найменших квадратів. Перевірку сталості дисперсії можна провести за допомогою критерію І. Фішера. Нормальність розподілу помилок за умови великої кількості спостережень перевіряється за допомогою відомих із математичної статистики методів, використовуючи критерій χ^2 або критерій Колмогорова. У разі малої кількості спостережень надійну перевірку нормальності розподілу помилок зробити не можна і припущення про нормальність розподілу помилок доводиться брати на віру.

Треба визначити значення a_0 та a_1, які мінімізують вираз:

$$Q = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - a_0 - a_1 \cdot x_i)^2 \to min \qquad (2.1)$$

Мінімум функції (2.1) досягається за умови, коли перші похідні дорівнюють нулю, тобто:

$$\frac{\partial(\sum_{i=1}^{n} e_i^2)}{\partial a_0} = -2 \sum_{i=1}^{n} (y_i - a_0 - a_1 \cdot x_i) = 0;$$

$$\frac{\partial\left(\sum_{i=1}^{n} e_i^2\right)}{\partial a_1} = -2\sum_{i=1}^{n}(y_i - a_0 - a_1 \cdot x_i)x_i = 0.$$

Звідси система нормальних рівнянь для визначення оцінок параметрів моделі a_i буде мати вигляд:

$$\begin{cases} n \cdot a_0 + a_1 \cdot \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i \\ a_0 \cdot \sum_{i=1}^{n} x_i + a_1 \cdot \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i \end{cases}.$$

Тоді можна отримати формули для знаходження параметрів a_0 та a_1із системи нормальних рівнянь:

$$a_1 = \frac{\sum_{i=1}^{n}(x_i - x) \cdot (y_i - y)}{\sum_{i=1}^{n}(x_i - x)^2},$$

$$a_0 = y - a_1 \cdot x,\ x = \frac{\sum_{i=1}^{n} x_i}{n},\ y = \frac{\sum_{i=1}^{n} y_i}{n}.$$

Оцінки параметрів (коефіцієнти регресії) a є вибірковими характеристиками сили зв'язку, що мають такі властивості (рис. 2.1) [27].

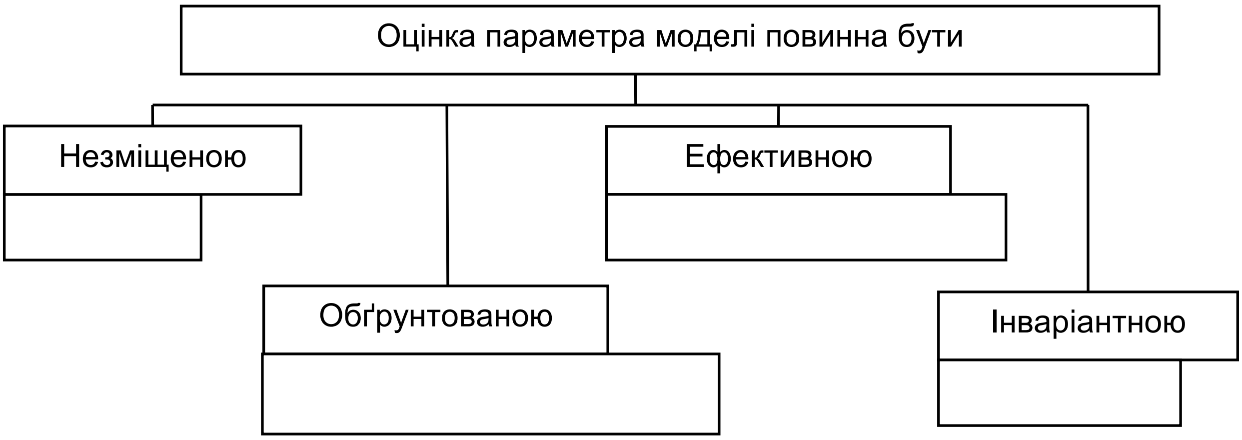


Рис. 2.1. **Властивості оцінок параметрів моделі**

Незміщеністю вибіркової оцінки a параметрів a називається така властивість, яка відповідає умові:

$$M\ a = a.$$

Тобто знайдена оцінка параметра моделі a_i може розглядатися як середнє значення з можливої великої кількості незміщених оцінок. Якщо оцінки мають властивість незміщеності, то їх можна порівнювати за різними дослідженнями.

Друга властивість оцінок a – *ефективність* – пов'язана зі значенням дисперсії оцінок і відповідає умові:

$$Var\ a\ = minVar(a^*),$$

де $Var\ a$ – дисперсія оцінок a згідно з МНК;

$Var(a^*)$ – дисперсія оцінок a, визначених іншими методами.

Оцінки вважаються ***ефективними,*** якщо вони характеризуються найменшою дисперсією. У практичних дослідженнях це означає можливість переходу від точкового прогнозу до інтервального.

Оцінки a є *обґрунтованими*, якщо під час збільшення числа спостережень ($n \to \infty$) вони будуть прагнути до істинних значень параметрів, тобто іншими словами – якщо для як завгодно малої величини $\varepsilon\ \ \varepsilon > 0$ виконується співвідношення $\lim_{n\to\infty} P\ \ a - a\ < \varepsilon\ = 1$.

Обґрунтованість оцінок характеризує збільшення їхньої точності зі збільшенням обсягу вибірки.

Інваріантність оцінки базується на тому, що в разі перетворення параметрів a за допомогою деякої функції g таке саме перетворення, виконане щодо a, дає оцінку $g(a)$ нового параметра.

Приклад 2.1. Необхідно дослідити залежність доходу банків (ум. грош. од.) від обсягу залучених коштів (ум. грош. од.) (табл. 2.1). Дохід банку, що є залежною змінною, слід позначити через Y, а обсяг залучених коштів як незалежну змінну (фактор) – через X. На величину доходу також впливають інші фактори (обсяг кредитів, структура кредитного та інвестиційного портфелів тощо). Слід припустити, що в досліджуваній групі банків значення цих останніх факторів приблизно однакові. Тому досліджуємо вплив у явному виді тільки одного фактора (обсягу залучених коштів) на величину доходу.

З аналізу реальних даних видно, що зі збільшенням величини залучених коштів величина доходу збільшується. Слід припустити, що зв'язок між цими факторами лінійний. Допоміжні розрахунки для параметрів моделі наведені в табл. 2.1. Середні значення факторів були розраховані як:

$$x = \frac{{}_{i=1}^{n} x_i}{n} = \frac{1356}{8} = 169{,}8; \; y = \frac{{}_{i=1}^{n} y_i}{n} = \frac{17}{8} = 2{,}1.$$

Таблиця 2.1

Вихідні дані завдання й проміжні розрахунки

i	X_i	Y_i	$\Delta X_i = X_i - X$	$\Delta Y_i = Y_i - Y$	$(\Delta X_i)^2$	$\Delta X_i \Delta Y_i$
1	110	0,9	-59,8	-1,2	3 570,1	72,4
2	128	1,2	-41,8	-0,9	1 743,1	38,1
3	153	1,6	-16,8	-0,5	280,6	8,6
4	173	1,9	3,3	-0,2	10,6	-0,7
5	183	2,5	13,3	0,4	175,6	5,1
6	189	2,5	19,3	0,4	370,6	7,5
7	201	3,0	31,3	0,9	976,6	27,7
8	221	3,3	51,3	1,2	2 626,6	60,9
Σ	1 356	17	0	0	9 753,5	219,6

Треба визначити на основі методу найменших квадратів (МНК) емпіричні коефіцієнти регресії, тобто оцінки параметрів лінійної моделі a_0 та a_1, отримані за результатами спостережень.

$$a_1 = \frac{{}_{i=1}^{n}(x_i - x) \cdot (y_i - y)}{{}_{i=1}^{n}(x_i - x)^2} = \frac{219{,}6}{9753{,}5} = 0{,}022518,$$

$$a_0 = y - a_1 \cdot x = 2{,}1 - 0{,}022518 \cdot 169{,}8 = -1{,}709855.$$

Варто зауважити, що оцінки параметрів моделі за методом МНК є досить чутливими до точності розрахунків та адекватності аналітичної форми моделі. Оскільки вільний член моделі a_0 = -1,709855 ≠ 0, то величина доходу не є строго пропорційною до величини залучених коштів. Кількісна оцінка параметра a_1 = 0,022518 показує, що граничне збільшення доходу зі зростанням величини залучених коштів на 1 ум. грош. од. становить 0,0225184.

Еластичність доходу відносно величини залучених коштів визначається коефіцієнтом еластичності:

$$E_{\frac{y}{x}} = \frac{\partial Y}{\partial X} \cdot \frac{X}{Y} = 0{,}022518 \cdot \frac{169{,}8}{2{,}1} = 1{,}8094.$$

Таким чином, у разі збільшення величини залучених коштів на 1 % величина доходу гранично зросте на 1,8094 %.

2.2. Оцінювання значущості лінійної парної регресії та її параметрів

Після визначення оцінок параметрів a_0 і a_1 слід перевірити їх статистичну надійність й істотність. Для цього використовують критерій Стьюдента. Знаходять для кожного параметра розрахункове значення за критерієм Стьюдента:

$$t_{a_i} = \frac{a_i}{\sigma_{a_i}},$$

де σ_{a_i}– середньоквадратична похибка оцінювання параметрів моделі.

Для парної лінійної регресії:

$$\sigma_{a_0} = \sqrt{\frac{\sum_{i=1}^{n} x_i^2 \cdot \sigma_e^2}{n \cdot \sum_{i=1}^{n}(x_i - x)^2}},$$

$$\sigma_{a_1} = \sqrt{\frac{\sigma_e^2}{\sum_{i=1}^{n}(x_i - x)^2}},$$

де σ_e^2 – оцінювання дисперсії випадкової величини, яка розраховується за формулою:

$$\sigma_e^2 = \frac{\sum_{i=1}^{n} e_i^2}{n - m - 1}.$$

Якщо $t_{a_i} \leq t_{\text{табл}}$, то з імовірністю помилки α приймається гіпотеза H_0: $a_i = 0$ (відповідний фактор незначно впливає на результативну ознаку).

В іншому випадку, тобто у разі $t_{a_i} > t_{\text{табл}}$, приймається гіпотеза H_1: $a_i \neq 0$ (відповідний фактор значно впливає на результативну ознаку). У даному критерії $t_{\text{табл}}$ – значення порогове (критичне) для критерію Стьюдента, яке береться з таблиць для відповідного рівня значущості α і кількості ступенів свободи $k = n - m - 1$ (де n – кількість спостережень, m – кількість незалежних факторів у моделі).

Мірою того, наскільки добре модель (рівняння регресії) описує дану систему спостережень, слугує коефіцієнт детермінації. Коефіцієнт детермінації R^2є відношенням поясненої (через регресійну модель) варіації результативної ознаки до всієї варіації в цілому, тобто:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} \left(Y_i - \hat{Y}_i\right)^2}{\sum_{i=1}^{n} \left(Y_i - \bar{Y}\right)^2}.$$

Таким чином, чим ближче цей коефіцієнт до 1, тим краще підібрана модель для опису конкретного економічного явища.

Тісноту лінійного зв'язку між залежною змінною y та незалежною змінною x оцінюють за допомогою лінійного парного коефіцієнта кореляції:

$$r_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}.$$

Значення r лежить у діапазоні від -1 до +1. У разі $r = 0$ змінні не можуть мати лінійного кореляційного зв'язку. Ступінь тісноти їх лінійної залежності зростає під час наближення r до ± 1. Цей коефіцієнт характеризує не тільки тісноту зв'язку, а і його напрям: коли $r > 0$, то зв'язок між показниками прямий, якщо $r < 0$ – обернений. Для парної лінійної регресії виконується таке співвідношення: $r_{xy} = \pm \sqrt{R^2}$.

Оскільки коефіцієнт кореляції є також вибірковою характеристикою, то його значущість перевіряється за допомогою критерію Стьюдента:

$$t_R = \frac{r_{xy} \cdot \sqrt{n - 2}}{\sqrt{1 - r_{xy}^2}}.$$

Статистичну значущість моделі в цілому перевіряють за допомогою критерію Фішера:

$$F = \frac{\sum_{i=1}^{n}\left(\hat{Y}_i - \bar{Y}\right)^2}{m} \quad \frac{\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2}{n-m-1} = \frac{R^2}{1-R^2} \cdot \frac{n-m-1}{m}.$$

Розраховане значення статистики Фішера необхідно порівняти з табличним $F_{\text{табл}}$ для числа ступенів свободи $k_1 = m$, $k_2 = m - n - 1$, рівня значущості α. Можлива помилка (рівень значущості) α може прийматися 0,05 або 0,01. Це означає, що у 5 % або 1 % випадків можна помилитися, а у 95 % або 99 % випадків (рівень довіри) висновки будуть правильними. За умови $F > F_{\text{табл}}$ побудова регресійної моделі відповідає реальній дійсності.

Якщо значення лінійного коефіцієнта кореляції статистично істотно, модель і її параметри статистично значущі, то підібране рівняння регресії можна використовувати для прогнозу досліджуваних економічних явищ або об'єктів. У ході цього не можна обійтися тільки точковим прогнозом значення результативної ознаки $Y_{\text{пр}} = a_0 + a_1 \cdot X_{\text{пр}}$, отриманим за моделлю, необхідно знайти ще й довірчі інтервали або довірчі межі для знайденого прогнозного значення $Y_{\text{пр}}$, тобто інтервальний прогноз:

$$Y\text{пр} - \Delta Y\text{пр} \leq Y\text{пр} \leq Y\text{пр} + \Delta Y\text{пр},$$

де $Y\text{пр}$ – набір можливих прогнозних значень, які лежать усередині кордонів довірчих інтервалів точкового прогнозу $Y\text{пр}$;

$\Delta Y\text{пр}$ – помилка прогнозу, яка розраховується за формулою:

$$\Delta Y\text{пр} = t_{\text{табл}} \cdot \sigma_e \cdot \sqrt{\frac{1}{n} + \frac{\left(X\text{пр} - \bar{X}\right)^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2}}.$$

Приклад 2.2. За даними прикладу 2.1 оцінити статистичну значущість отриманого рівняння регресії. Знайти прогнозне значення величини доходу банку, якщо величина залучених коштів складе $X\text{пр} = 250$ ум. грош. од.

Рівняння залежності між факторами X й Y буде мати такий вигляд:

$$Y = a_0 + a_1 \cdot X = -1{,}709855 + 0{,}022518 \cdot X.$$

Для перевірки гіпотези про те, чи значно відрізняється від нуля вибірковий коефіцієнт a_1, слід перевірити статистичну значущість цього коефіцієнта, використовуючи критерій Стьюдента, тобто перевірити гіпотезу H$_0$: $a_1 = 0$ проти альтернативної гіпотези H$_1$: $a_1 \neq 0$. Проміжні розрахунки наведено в табл. 2.2.

Таблиця 2.2

Вихідні дані й проміжні розрахунки

i	X_i	Y_i	Y_i	$e_i = Y_i - Y_i$	${e_i}^2$	$(\Delta X_i)^2$	$\Delta Y_i = Y_i - Y$	$(\Delta Y_i)^2$
1	110	0,9	0,77	0,13	0,018	3570,1	-1,2	1,5
2	128	1,2	1,17	0,03	0,001	1743,1	-0,9	0,8
3	153	1,6	1,74	-0,14	0,018	280,6	-0,5	0,3
4	173	1,9	2,19	-0,29	0,082	10,6	-0,2	0,0
5	183	2,5	2,41	0,09	0,008	175,6	0,4	0,2
6	189	2,5	2,55	-0,05	0,002	370,6	0,4	0,2
7	201	3,0	2,82	0,18	0,034	976,6	0,9	0,8
8	221	3,3	3,27	0,03	0,001	2626,6	1,2	1,4
Σ	1356	17	16,90	0,00	0,163	9753,5	0	5,1

Слід розрахувати середньоквадратичне відхилення помилок:

$$\sigma_e = \sqrt{\frac{\sum_{i=1}^{n}(y_i - y_i)^2}{n-2}} = \sqrt{\frac{0{,}163}{6}} = 0{,}164991;$$

середньоквадратичну похибку оцінки параметра a_1:

$$\sigma_{a_1} = \frac{\sigma_e}{\sqrt{\sum_{i=1}^{n}(x_i - x)^2}} = \frac{0{,}164991}{\sqrt{9\ 753{,}5}} = 0{,}001671;$$

критерій Стьюдента:

$$t_{a_1} = \frac{0{,}022518}{0{,}001671} = 13{,}48.$$

Знайти за таблицями розподілу Стьюдента значення $t_{\text{табл}}$ для числа ступенів свободи k = 8 – 2 = 6, рівня значущості α = 0,05 ($t_{\text{табл}}$ = 1,943). Порівнюючи значення $|t_{a_1}|$ й $t_{\text{табл}}$, можна дійти висновку, що отримане значення параметра a_1 статистично значуще (|13,48| > 1,943), у такий спосіб приймається гіпотеза H_1, тобто зміна обсягу залучених коштів істотно впливає на величину доходу.

Слід знайти інтервальні оцінки параметра a_1:

$$a_1 - \Delta a_1 \le a_1 \le a_1 + \Delta a_1, \qquad \Delta a_1 = t_{\text{табл}} \cdot \sigma_{a_1}.$$

З імовірністю 0,95 істинні значення параметра a_1 лежать у таких межах:

$$0{,}022518 - 0{,}003246 \le a_1 \le 0{,}022518 + 0{,}003246,$$
$$0{,}019272 \le a_1 \le 0{,}025764.$$

Розрахуємо коефіцієнт детермінації для моделі:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2}{\sum_{i=1}^{n}\left(Y_i - \bar{Y}\right)^2} = 1 - \frac{0{,}163}{5{,}1} = 0{,}968.$$

Коефіцієнт детермінації показує, що 96,8 % (R^2 · 100 %) загальної зміни доходу банку пояснюється зміною величини залучених коштів, у той час, як на інші фактори доводиться лише 3,2 % зміни.

Розрахувати коефіцієнт кореляції та значення критерію Стьюдента *(t-критерію)* для оцінювання його статистичної значущості:

$$R = \sqrt{R^2} = \sqrt{0{,}968} = 0{,}9839, \qquad t_R = \frac{R \cdot \sqrt{n-2}}{\sqrt{1-R^2}} = \frac{0{,}9839 \cdot \sqrt{6}}{\sqrt{1-0{,}9839}} = 75{,}38.$$

Оскільки $|t_R|$ > $t_{\text{табл}}$ (75,38 > 1,943), то можна зробити висновок про значущість коефіцієнта кореляції між залежною і пояснювальною змінними.

Треба перевірити статистичну значущість моделі в цілому за допомогою критерію Фішера. Розрахувати статистику Фішера за формулою:

$$F_R = \frac{R^2}{1-R^2}\cdot\frac{n-2}{1} = \frac{0{,}968}{1-0{,}968}\cdot\frac{8-2}{1} = 181{,}67.$$

Розраховане значення статистики Фішера порівнюється з табличним для числа ступенів свободи k_1 = 1, k_2 = 8 – 2 = 6, рівня значущості α = 0,01 ($F_{0,01}$ (1;6) = 13,74). Оскільки 181,67 > 13,74, то приймається гіпотеза, що побудована модель є статистично значущою, тобто зв'язок між залежною та пояснювальною змінними істотний.

Оскільки побудована лінійна економетрична модель є адекватною і статистично значущою, то отримане рівняння залежності доходу від величини залучених коштів може бути використане для прогнозу. Прогнозована величина доходу за рівнянням регресії складе:

$$Y\text{пр} = -1{,}709855 + 0{,}022518\cdot 250 = 3{,}92 \text{ (ум. грош. од.)}.$$

З погляду прийнятих припущень отриманий прогноз є лише точковою оцінкою істинної складової $Y(X)$, тому необхідно знайти інтервальні оцінки для отриманого прогнозу, що враховують помилку прогнозу:

$$\Delta Y\text{пр} = t_{\text{табл}}\cdot\sigma_e\cdot\sqrt{\frac{1}{n}+\frac{(X\text{пр}-X)^2}{\sum_{i=1}^{n}(X_i-X)^2}} = 0{,}2840827;$$

$$3{,}92 - 0{,}2840827 \le Y\text{пр} \le 3{,}92 + 0{,}2840827;$$

$$3{,}6355 \le Y\text{пр} \le 4{,}2036.$$

Тобто з 95 % рівнем довірчої ймовірності прогнозне значення доходу банку, якщо величина залучених коштів дорівнює 250 ум. грош. од., буде знаходитись в інтервалі від 3,6355 до 4,2036 ум. грош. од.

2.3. Оцінювання параметрів множинної лінійної регресії МНК

Під час дослідження економічних явищ необхідно вивчати не тільки окремий вплив кожного фактора на результативну ознаку, але й ефект спільного впливу деякої групи факторів на досліджуваний показник.

Наприклад,

X_1

X_2 незалежні фактори Y залежний фактор

. (екзогенні) (ендогенний)

.

X_m

Найбільш часто передбачається, що між групою незалежних (екзогенних) змінних і залежним показником (ендогенною) змінною, існує лінійний зв'язок, який можна подати у вигляді

$$y = a_0 + a_1 \cdot x_1 + a_2 \cdot x_2 + \cdots + a_m \cdot x_m.$$

Для коректного використання МНК необхідно, щоб виконувалися усі умови відносно помилок моделі. За цих умов МНК дає незміщені, обґрунтовані й ефективні оцінки параметрів.

Слід скласти функцію $F(a_0, \ldots, a_m)$, рівну сумі квадратів відхилень e_i:

$$F(a_0, \ldots, a_m) = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - a_0 - a_1 x_{i1} - a_2 x_{i2} - \cdots - a_m x_{im})^2 \to min.$$

З необхідних умов мінімуму функції декількох змінних буде отримано систему рівнянь:

$$\frac{\partial F}{\partial a_i} = 0, i = 0,1, \ldots, m .$$

Узявши частинні похідні від правої частини вираження для функції F за параметрами a_i і перетворивши отримувані рівняння, перейдемо до наступної системи нормальних рівнянь:

$$
\begin{cases}
n \cdot a_0 + a_1 \cdot \sum_{i=1}^{n} x_{i1} + a_2 \cdot \sum_{i=1}^{n} x_{i2} + \dots + a_m \cdot \sum_{i=1}^{n} x_{im} = \sum_{i=1}^{n} y_i \\
a_0 \cdot \sum_{i=1}^{n} x_{i1} + a_1 \cdot \sum_{i=1}^{n} {x_{i1}}^2 + \dots + a_m \cdot \sum_{i=1}^{n} x_{i1} \cdot x_{im} = \sum_{i=1}^{n} x_{i1} \cdot y_i \\
\dots \\
a_0 \cdot \sum_{i=1}^{n} x_{im} + a_1 \cdot \sum_{i=1}^{n} x_{i1} \cdot x_{im} + \dots + a_m \cdot \sum_{i=1}^{n} {x_{im}}^2 = \sum_{i=1}^{n} x_{im} \cdot y_i
\end{cases}.
$$

Підчас запису системи рівнянь для знаходження параметрів рівняння множинної регресії можна керуватися таким простим правилом: перше рівняння виходить як сума *n* рівнянь регресії; друге і подальші – як сума *n* рівнянь регресії, усі члени якої помножені на x_1, потім на x_2 і т. д.

Якщо ввести позначення

$$
X = \begin{pmatrix} 1 & x_{11} & x_{12} & \dots & x_{1m} \\ 1 & x_{21} & x_{22} & \dots & x_{2m} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nm} \end{pmatrix}, X^T = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ x_{11} & x_{21} & x_{31} & \dots & x_{n1} \\ \dots & \dots & \dots & \dots & \dots \\ x_{1m} & x_{2m} & x_{3m} & \dots & x_{nm} \end{pmatrix},
$$

$$
a = \begin{pmatrix} a_0 \\ a_1 \\ \dots \\ a_m \end{pmatrix}, \ Y = \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix}.
$$

У матричній формі система нормальних рівнянь записується таким чином:

$$
X^T \cdot X \cdot a = X^T \cdot Y.
$$

Для того, щоб знайти вектор оцінок a, тобто щоб система мала рішення, необхідно, щоб матриця $X^T X$ була *неособливою (невиродженою)*, тобто $|X^T X| \neq 0$ Для цього вектори значень пояснюючих змінних (тобто стовпці матриці X) мають бути лінійно незалежні, тобто ранг матриці має дорівнювати числу її стовпців $r(X) = m + 1$.

Крім того, має бути виконана умова $n > m + 1$. Іншими словами число наявних спостережень кожної з пояснюючих змінних повинне, принаймні, на одиницю перевершувати число пояснюючих змінних. Таким чином, якщо матриця $X^T X$ не вироджена, то рішенням системи рівнянь є вектор $a = (a_0, a_1, \dots, a_m)^T$:

$$
a = (X^T \cdot X)^{-1} \cdot X^T \cdot Y.
$$

Як і у разі парної лінійної регресії, особливе значення для перевірки статистичної значущості множинної лінійної регресії мають *залишки* (різниці між істинними значеннями показника і значеннями, вичисленими за рівнянням лінійної регресії):

$$e_i = y_i - y_i,\ i = 1, .., n.$$

Приклад 2.3. Необхідно дослідити залежність ВВП (Y) (ум. грош. од.) від таких змінних: витрати трудових ресурсів (X_1) (ум. грош. од.), витрат основних фондів (X_2) (ум. грош. од.). Вихідні дані наведені у табл. 2.3.

Таблиця 2.3

Вихідні дані завдання

№ п/п	1	2	3	4	5	6	7	8	9	10
Y	51,2	54,7	65,6	53,3	72,7	75,7	85,6	91,5	96,7	102,6
X_1	9,8	10,1	10,9	10,4	12,1	12,8	13,2	13,6	12,3	13,5
X_2	6,8	10,5	12,2	12,8	13,4	13,8	14,4	14,6	16,5	18,2

Розглядається двофакторна лінійна регресійна модель: $y = a_0 + a_1x_1 + a_2x_2$. Оцінки параметрів можна одержати на основі МНК за таким алгоритмом:

1. Незалежні змінні записати у вигляді матриці X:

$$X = \begin{matrix} 1 & 9{,}8 & 6{,}8 \\ 1 & 10{,}1 & 10{,}5 \\ 1 & 10{,}9 & 12{,}2 \\ 1 & 10{,}4 & 12{,}8 \\ 1 & 12{,}1 & 13{,}4 \\ 1 & 12{,}8 & 13{,}8 \\ 1 & 13{,}2 & 14{,}4 \\ 1 & 13{,}6 & 14{,}6 \\ 1 & 12{,}3 & 16{,}5 \\ 1 & 13{,}5 & 18{,}2 \end{matrix}.$$

2. Обчислити матрицю $B = X^T \cdot X$ і вектор $X^T \cdot Y$, де X^T – транспонована матриця X, Y – вектор спостережень залежної змінної.

$$B = X^T X = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 9,8 & 10,1 & 10,9 & 10,4 & 12,1 & 12,8 & 13,2 & 13,6 & 12,3 & 13,5 \\ 6,8 & 10,5 & 12,2 & 12,8 & 13,4 & 13,8 & 14,4 & 14,6 & 16,5 & 18,2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 9,8 & 6,8 \\ 1 & 10,1 & 10,5 \\ 1 & 10,9 & 12,2 \\ 1 & 10,4 & 12,8 \\ 1 & 12,1 & 13,4 \\ 1 & 12,8 & 13,8 \\ 1 & 13,2 & 14,4 \\ 1 & 13,6 & 14,6 \\ 1 & 12,3 & 16,5 \\ 1 & 13,5 & 18,2 \end{pmatrix} =$$

$$= \begin{pmatrix} 10 & 118,7 & 133,2 \\ 118,7 & 1428,01 & 1614,86 \\ 133,2 & 1614,86 & 1864,18 \end{pmatrix}.$$

$$X^T Y = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 9,8 & 10,1 & 10,9 & 10,4 & 12,1 & 12,8 & 13,2 & 13,6 & 12,3 & 13,5 \\ 6,8 & 10,5 & 12,2 & 12,8 & 13,4 & 13,8 & 14,4 & 14,6 & 16,5 & 18,2 \end{pmatrix} \cdot \begin{pmatrix} 51,2 \\ 54,7 \\ 65,6 \\ 53,3 \\ 72,7 \\ 75,7 \\ 85,6 \\ 91,5 \\ 96,7 \\ 102,6 \end{pmatrix} =$$

$$= \begin{pmatrix} 749,6 \\ 9121,05 \\ 10455,32 \end{pmatrix}.$$

3. Обчислити зворотну матрицю $B^{-1} = (X^T X)^{-1}$.

$$B^{-1} = \frac{1}{detB} \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix}^T,$$

де *det B* – визначник матриці *B*;

A_{ij}– алгебричне доповнення елемента b_{ij}матриці B.

$det\ B = b_{11} \cdot b_{22} \cdot b_{33} + b_{13} \cdot b_{21} \cdot b_{32} + b_{12} \cdot b_{23} \cdot b_{31} - b_{13} \cdot b_{22} \cdot b_{31} - b_{11} \cdot b_{23} \cdot b_{32} -$
$- b_{12} \cdot b_{21} \cdot b_{33}$ = 5 529,93.

$$A_{ij} = (-1)^i + j \cdot M_{ij},$$

де M_{ij} – мінор для елементаb_{ij}, тобто визначник матриці, отриманої з вихідної шляхом викреслення i-го рядка й j-го стовпця.

Матриця B^{-1} буде мати вигляд:

$$B^{-1} = \frac{1}{5529{,}93} \begin{matrix} 52866{,}85 & -6060{,}114 & 1472{,}95 \\ -6060{,}114 & 889{,}56 & -337{,}76 \\ 1472{,}95 & -337{,}76 & 190{,}41 \end{matrix} =$$

$$= \begin{matrix} 9{,}5601 & -1{,}0959 & 0{,}2664 \\ -1{,}0959 & 0{,}1609 & -0{,}0611 \\ 0{,}2664 & -0{,}0611 & 0{,}0344 \end{matrix}.$$

4. Обчислити параметри моделі за формулою:

$$a = B^{-1} \cdot X^{T} \cdot Y = \begin{matrix} 9{,}5601 & -1{,}0959 & 0{,}2664 \\ -1{,}0959 & 0{,}1609 & -0{,}0611 \\ 0{,}2664 & -0{,}0611 & 0{,}0344 \end{matrix} \cdot \begin{matrix} 749{,}6 \\ 9121{,}05 \\ 10455{,}32 \end{matrix} =$$

$$= \begin{matrix} -44{,}385 \\ 7{,}174 \\ 2{,}567 \end{matrix}.$$

Таким чином, теоретична лінійна залежність між ВВП і витратами трудових ресурсів і основних фондів має вигляд:

$$y_i = -44{,}385 + 7{,}174 \cdot x_{i1} + 2{,}567 \cdot x_{i2}.$$

2.4. Оцінювання тісноти та значущості зв'язку між змінними в рівнянні множинної регресії

Оскільки оцінювання параметрів лінійної множинної економетричної моделі здійснювалося за даними спостережень, то природно, що ці оцінки будуть відхилятися від істинних значень відповідних параметрів. Оцінки середньоквадратичних відхилень параметрів обчислюються за формулою:

$$\sigma_{a_i} = \sqrt{\sigma_e^2 \cdot B_{jj}^{-1}}, \qquad \sigma_e = \sqrt{\frac{\sum_{i=1}^{n}(y_i - y_i)^2}{n - m - 1}},$$

де B_{jj}^{-1} – діагональний елемент матриці B^{-1}.

Щоб визначити істотність лінійного зв'язку між Y і X_j потрібно перевірити гіпотезу про рівність нулю коефіцієнта H_0: $a_j = 0$ для всіх $j = 1, 2, \dots, m$

проти відповідних альтернативних гіпотез $H_1 : a_j \neq 0$ для всіх $j = 1, 2, \ldots, m$. Для цього скористаємося критерієм Стьюдента для оцінки a_j.

$$t_{a_j} = \frac{a_j}{\sigma_{a_j}}.$$

Для оцінювання підібраної лінійної множинної моделі і ступеня її адекватності реальному економічному процесу використовують коефіцієнт множинної кореляції:

$$R = \sqrt{1 - \frac{\sum_{i=1}^{n} (Y_i - Y_i)^2}{\sum_{i=1}^{n} (Y_i - Y)^2}}.$$

Він характеризує тісноту зв'язку всіх пояснювальних змінних із залежною. Значення R лежить у діапазоні від 0 до 1. Чим ближче коефіцієнт R до 1, тим краще підібрана модель для опису залежності між досліджуваними економічними явищами.

Коефіцієнт детермінації R^2 також застосовують для оцінювання адекватності множинної регресії. Однак R^2 зростає під час додавання кожного регресора, хоча це не обов'язково відповідає зростанню якості моделі. Тому проводять "корекцію" R^2 на число ступенів свободи і розраховують скоректований або адаптований (*adjusted*) коефіцієнт детермінації за формулою:

$$R^2_{adj} = 1 - \frac{\sum_{i=1}^{n}(y_i - y_i)^2}{\sum_{i=1}^{n}(y_i - y)^2} \cdot \frac{n-1}{n-m-1} = 1 - \left(1 - R^2\right) \frac{n-1}{n-m-1}.$$

Чим більше число регресорів, включених до моделі, тим менше R^2_{adj} порівняно з R^2. На відміну від R^2 скоректований коефіцієнт детермінації може зменшуватися під час включення до моделі додаткової пояснювальної змінної, якщо вона не суттєво впливає на залежну змінну.

Під час аналізу та вивчення тенденцій розвитку досліджуваних економічних явищ множинну лінійну модель використовують для отримання прогнозних значень залежного фактора Y від деякого набору передбачуваних значень незалежних факторів X_1, X_2, ... X_m:

$$Y\text{пр} = a_0 + a_1 \cdot X_{1\text{пр}} + a_2 \cdot X_{2\text{пр}} + \ldots + a_m \cdot X_{m\text{пр}},.$$

У ході цього, оскільки на процес побудови моделі істотно впливає обсяг спостережень, необхідно використовувати не тільки точкову оцінку прогнозу Yпр, але і знайти довірчі інтервали:

$$Y\text{пр} - \Delta Y\text{пр} \le Y\text{пр} \le Y\text{пр} + \Delta Y\text{пр},$$

де $\Delta Y\text{пр} = t_{\text{табл}} \cdot \sigma_e \cdot \sqrt{X_{\text{пр}}^T \cdot B^{-1} \cdot X_{\text{пр}}}$;

$X_{\text{пр}}^T$ = (1, $X_{1\text{пр}}$, $X_{2\text{пр}}$, ... $X_{m\text{пр}}$).

Приклад 2.4. За даними прикладу 2.3 оцінити статистичну значущість підібраного рівняння регресії. Знайти прогнозне значення ВВП, якщо прогнозні значення витрат трудових ресурсів $X_{1\text{пр}}$ = 14 ум. грош. од., витрат основних фондів $X_{2\text{пр}}$ = 17 ум. грош. од.

Для оцінки статистичної значущості параметрів моделі використовуємо критерій Стьюдента. Проміжні розрахунки наведені в табл. 2.4.

Таблиця 2.4

Вихідні дані завдання й проміжні розрахунки

№ п/п	Y	X_1	X_2	Y	$e = Y - Y$	e^2	$(Y - \overline{Y})^2$	$(Y - \overline{Y})^2$
1	51,2	9,8	6,8	43,37	7,83	61,31	564,54	997,93
2	54,7	10,1	10,5	55,02	-0,32	0,10	410,47	397,60
3	65,6	10,9	12,2	65,13	0,47	0,22	87,61	96,63
4	53,3	10,4	12,8	63,08	-9,78	95,65	469,16	141,13
5	72,7	12,1	13,4	76,82	-4,12	16,97	5,11	3,46
6	75,7	12,8	13,8	82,86	-7,16	51,27	0,55	62,41
7	85,6	13,2	14,4	87,27	-1,67	2,79	113,21	151,54
8	91,5	13,6	14,6	90,66	0,84	0,71	273,57	246,49
9	96,7	12,3	16,5	86,21	10,49	110,04	472,63	126,56
10	102,6	13,5	18,2	99,18	3,42	11,70	763,97	586,61
Σ	749,6	118,7	133,2	749,6	0,00	350,75	3 160,82	2 810,36

Слід розрахувати дисперсію похибок:

$$\sigma_e{}^2 = \frac{\sum_{i=1}^{n}(y_i - y_i)^2}{n-m-1} = \frac{350{,}75}{10-2-1} = 50{,}107,$$

середньоквадратичні похибки оцінок параметрів:

$$\sigma_{a_0} = \sqrt{\sigma_e^2 \cdot B_{11}^{-1}} = \sqrt{50{,}107 \cdot 9{,}5601} = 21{,}887,$$

$$\sigma_{a_1} = \sqrt{\sigma_e^2 \cdot B_{22}^{-1}} = \sqrt{50{,}107 \cdot 0{,}1609} = 2{,}839,$$

$$\sigma_{a_2} = \sqrt{\sigma_e^2 \cdot B_{33}^{-1}} = \sqrt{50{,}107 \cdot 0{,}0344} = 1{,}313.$$

Для оцінювання істотності впливу факторів визначимо критерій Стьюдента:

$$t_{a_0} = \frac{a_0}{\sigma_{a_0}} = \frac{-44{,}382}{21{,}887} = -2{,}028; \quad t_{a_1} = \frac{a_1}{\sigma_{a_1}} = \frac{7{,}174}{2{,}839} = 2{,}527;$$

$$t_{a_2} = \frac{a_2}{\sigma_{a_2}} = \frac{2{,}567}{1{,}313} = 1{,}955.$$

Отримані значення треба порівняти із критичним значенням $t_{\text{табл}}$ для числа ступеня свободи $k = n - m - 1 = 7$ і рівня значущості $\alpha = 0{,}05$ ($t_{\text{табл}}$ = 1,895).

Оскільки значення $|t_{a_1}|$ і $|t_{a_2}|$ більше $t_{\text{табл}}$, дійдемо висновку, що отримані значення для коефіцієнтів a_1 і a_2 статистично значущі, тобто витрати трудових ресурсів та основних фондів істотно впливають на величину ВВП.

Оскільки $|t_{a_0}|$ > $t_{\text{табл}}$, то можна говорити про те, що вплив інших факторів, що не були враховані у моделі, також є суттєвим.

Для перевірки адекватності отриманої моделі необхідно обчислити коефіцієнт детермінації: $R^2 = 1 - \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2} = 1 - \frac{350{,}75}{3160{,}82} = 0{,}8836.$

У цьому випадку 88,36 % загальної зміни ВВП пояснюється змінами витрат трудових ресурсів та основних фондів, у той час, як на інші фактори доводиться лише 11,64 % зміни.

Скоректований коефіцієнт детермінації обчислимо за формулою:

$$R_{adj}^2 = 1 - \left(1 - R^2\right)\frac{n-1}{n-m-1} = 1 - \left(1 - 0{,}8836\right)\frac{9}{7} = 0{,}85.$$

Його значення досить високе, щоб стверджувати про адекватність моделі.

Коефіцієнт множинної кореляції є мірою тісноти зв'язку всіх пояснювальних змінних із залежною і визначається:

$$R = \sqrt{1 - \frac{\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2}{\sum_{i=1}^{n}\left(Y_i - \bar{Y}\right)^2}} = \sqrt{1 - \frac{350{,}75}{3\,160{,}82}} = 0{,}94.$$

Зв'язок між досліджуваними економічними явищами тісний, оскільки $R > 0{,}7$.

Перевіримо статистичну значущість моделі загалом (зв'язок між залежною змінною Y і незалежними змінними X_1 і X_2) за допомогою критерію Фішера:

$$F = \frac{\sum_{i=1}^{n}\left(\hat{Y}_i - \bar{Y}\right)^2}{m} : \frac{\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2}{n-m-1} = \frac{2\,810{,}36}{2} : \frac{350{,}75}{7} = 27{,}05.$$

Розраховане значення статистики І. Фішера необхідно порівняти з табличним для числа ступенів свободи k_1 = 2, k_2 = 10 – 2 – 1 = 7, рівня значущості α = 0,05 ($F_{0{,}05}$ (2; 7) = 19,4). Оскільки 27,05 > 19,4, то приймається гіпотеза, що побудована модель є статистично значущою, тобто зв'язок між залежною та пояснювальними змінними істотний.

Оскільки побудована множинна лінійна економетрична модель є адекватною і статистично значущою, то отримане рівняння залежності ВВП від витрат трудових ресурсів та основних фондів може бути використане для прогнозу. Точкова оцінка прогнозу має вигляд:

Yпр = –44,385 + 7,174·14 + 2,567·17 = 99,69 (ум. грош. од.).

Довірчі інтервали прогнозу:

$$\varDelta Y_{\text{пр}} = 1{,}895 \cdot 7{,}079 \cdot \sqrt{\begin{pmatrix} 1 & 14 & 17 \end{pmatrix} \cdot \begin{pmatrix} 9{,}5601 & -1{,}0959 & 0{,}2664 \\ -1{,}0959 & 0{,}1609 & -0{,}0611 \\ 0{,}2664 & -0{,}0611 & 0{,}0344 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 14 \\ 17 \end{pmatrix}} = 7{,}81.$$

Інтервальний прогноз має вигляд:

$$99{,}69 - 7{,}81 \le Y\text{пр} \le 99{,}69 + 7{,}81,$$
$$84{,}07 \le Y\text{пр} \le 107{,}5.$$

Таким чином, якщо витрати трудових ресурсів у прогнозному періоді складуть 14 ум. грош. од., а витрати основних фондів – 17 ум. грош. од., то ВВП може скласти від 84,07 ум. грош. од. до 107,5 ум. грош. од.

2.5. Стандартизована форма моделі множинної регресії

Розглянемо метод оцінювання параметрів множинної моделі в стандартизованому вигляді. Нехай $X(n\,x\,m)$ – матриця спостережень, що містить n значень, над m факторами, що описують деякий економічний процес (або явище). Рівняння множинної лінійної економетричної моделі може бути подано у стандартному вигляді за умови, що значення вихідних факторів необхідно центрувати і нормувати, тобто:

$$x_{i1} = \frac{X_{i1} - X}{\sigma_1},\ x_{i2} = \frac{X_{i2} - X}{\sigma_2},\ x_{i3} = \frac{X_{i3} - X}{\sigma_3}.$$

Рівняння в стандартному вигляді виглядатиме так:

$$y = \beta_1 \cdot x_1 + \beta_2 \cdot x_2 + \dots + \beta_m \cdot x_m.$$

Основні етапи оцінювання параметрів множинної лінійної економетричної моделі в стандартному вигляді пов'язані з використанням парних коефіцієнтів кореляції. Представимо їх більш докладно.

1. Обчислюємо матрицю парних коефіцієнтів кореляції для всіх ознак:

$$r_{yx_j} = \frac{(x_{ij}-x_j)(y_i-y)}{\overline{\overline{(x_{ij}-x_j)^2(y_i-y)^2}}} \text{ або } r_{x_k x_j} = \frac{(x_{ik}-x_k)(x_{ij}-x_j)}{\overline{\overline{(x_{ik}-x_k)^2(x_{ij}-x_j)^2}}}.$$

2. Складаємо систему рівнянь для визначення β-коефіцієнтів:

$$\begin{aligned} \beta_1 \quad + \beta_2 \cdot r_{12} + \beta_3 \cdot r_{13} + \ldots + \beta_m \cdot r_{1m} &= r_{1y} \\ \beta_1 \cdot r_{21} + \beta_2 + \beta_3 \cdot r_{23} + \ldots + \beta_m \cdot r_{3m} &= r_{2y} \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & \\ \beta_1 \cdot r_{m1} + \beta_2 \cdot r_{m2} + \beta_3 \cdot r_{m3} + \ldots + \beta_m &= r_{my}, \end{aligned}$$

де r_{1y} – коефіцієнт парної кореляції i-го фактора і результативної ознаки y.

Різними методами вирішуємо дану систему рівнянь і знаходимо коефіцієнти β_i $(i = 1, m)$. β_i показує, на яку частину свого квадратичного відхилення зміниться залежна змінна Y, якщо незалежна змінна x_i зміниться на одне значення свого середньоквадратичного відхилення.

3. Обчислимо коефіцієнти регресії:

$$a_j = \beta_j \cdot \frac{\sigma_y}{\sigma_j},$$

$$a_0 = y - \sum_{j=1}^{m} b_j \cdot x_j.$$

4. Множинний коефіцієнт кореляції $R_{1,2,\ldots,m}$ обчислюється таким чином:

$$R_{1,2,\ldots,m} = \sqrt{\beta_1 \cdot r_{1y} + \beta_2 \cdot r_{2y} + \cdots + \beta_m \cdot r_{my}}.$$

Перевірити істотність множинного коефіцієнта кореляції можна за допомогою перетворення І. Фішера:

$$z = \frac{1}{2} \cdot ln \frac{1+R}{1-R} .$$

Розподіл величини z близький до нормального і має середню $E(z)$ і середньоквадратичне відхилення σ_z:

$$\sigma_z = \frac{1}{\sqrt{n-m-2}}.$$

Слід порівняти величину $t = \frac{z}{\sigma_z}$ з критичним $t_{\text{табл}}$ для відповідних α та $k = n - m - 2$. Якщо $t > t_{\text{табл}}$, то можна вважати, що між залежним і вибраними незалежними факторами є лінійний статистичний зв'язок.

Контрольні запитання для самодіагностики

1. Сформулюйте особливості лінійної економетричної моделі.
2. Які вимоги висуваються до об'єму спостережень, необхідного для побудови багатофакторної регресійної моделі?
3. Поясніть суть методу найменших квадратів для оцінювання параметрів лінійних економетричних моделей.
4. Запишіть різні форми системи нормальних рівнянь для множинної лінійної моделі. Якими методами може бути вирішена система нормальних рівнянь у даному випадку?
5. У чому суть понять "незміщеність, обґрунтованість і ефективність оцінок"? Яким гіпотезам повинні задовольняти відхилення в моделі, щоб оцінки параметрів моделі, отримані за допомогою МНК, мали властивості незміщеності, обґрунтованості й ефективності?
6. У зв'язку з чим необхідно перевіряти статистичну значущість оцінок параметрів моделі?
7. У чому суть критерію Стьюдента? Як визначається статистична значущість оцінок параметрів моделі?
8. Як визначаються довірчі інтервали для оцінок параметрів моделі?
9. Що таке адекватність моделі? Методи визначення адекватності моделі.
10. У чому суть коефіцієнта множинної кореляції? Якими методами можна його розрахувати?
11. Як здійснюється розрахунок прогнозних значень за множинною економетричною моделлю?
12. Чим різняться рівняння регресії в натуральному і стандартизованому виглядах?

Тести

1. Регресійні рівняння описують:

а) структурний зв'язок між показниками економічних процесів;

б) функціональний зв'язок між економічними процесами;

в) кореляційний зв'язок між економічними показниками.

2. Для оцінювання параметрів економетричної моделі використовують:

а) критерій Стьюдента;

б) метод найменших квадратів;

в) критерій Фішера.

3. У випадку парної лінійної регресії від знака коефіцієнта кореляції:

а) залежить напрям кореляційного зв'язку факторів і показника;

б) не залежить напрям кореляційного зв'язку факторів і показника.

4. Оцінки параметрів моделі називаються незміщеними, якщо:

а) математичне сподівання оцінок параметрів збігається з істинними значеннями цих параметрів;

б) оцінки параметрів сходяться за імовірністю до істинних значень параметрів;

в) у класі лінійних оцінок оцінки параметрів моделі мають мінімальні дисперсії.

5. Однією з передумов методу найменших квадратів є:

а) дисперсія збурень є постійною величиною;

б) сума залишків моделі, відмінна від нуля.

6. Критерій Стьюдента використовується для оцінювання статистичної значущості:

а) параметрів моделі;

б) коефіцієнта кореляції;

в) як параметрів моделі, так і коефіцієнта кореляції.

7. Коефіцієнт детермінації вимірює:

а) варіацію незалежної змінної;

б) нахил лінії регресії;

в) перетинання лінії регресії;

г) загальну варіацію залежної перемінний, котра пояснюється регресією.

8. Щоб перевірити значущість окремого параметра, використовують:

а) F-тест;

б) t-тест;

в) χ^2-тест.

9. Для перевірки значущості одночасно всіх параметрів використовується:

а) F-тест;

б) t-тест;

в) χ^2-тест.

10. Оцінки параметрів моделі називаються ефективними, якщо:

а) математичне сподівання оцінок параметрів збігається з істинними значеннями цих параметрів;

б) оцінки параметрів сходяться за імовірністю до істинних значень параметрів;

в) у класі лінійних оцінок оцінки параметрів моделі мають мінімальні дисперсії.

11. Довірчі інтервали функції регресії визначаються за допомогою:

а) t-тесту та стандартної похибки моделі;

б) t-тесту та стандартної похибки оцінки параметрів моделі.

12. *Оцінки параметрів моделі називаються обґрунтованими, якщо:*

а) математичне сподівання оцінок параметрів збігається з істинними значеннями цих параметрів;

б) оцінки параметрів сходяться за імовірністю до істинних значень параметрів;

в) у класі лінійних оцінок оцінки параметрів моделі мають мінімальні дисперсії.

13. З урахуванням співвідношення між дивідендними виплатами (y, ум. грош. од.) і прибутком підприємства (x, ум. грош. од.): y = 10,2 + 0,5x, за умови збільшення прибутку підприємства на 1 ум. грош. од. можна очікувати такі додаткові дивідендні виплати:

а) 10,2;

б) 0,5;

в) 10,7.

14. Точковий прогноз – це:

а) прогнозні значення залежного фактора, знайдені за моделлю, у разі деякого набору передбачуваних значень незалежних факторів;

б) нижня і верхня межі довірчого інтервалу, в якому із заданою ймовірністю буде перебувати прогнозне значення залежної змінної.

15. Коефіцієнт детермінації розраховується за формулою:

а) $R^2 = 1 - \frac{\sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{n} (Y_i - \bar{Y})^2}$;

б) $R^2 = 1 - \frac{\sum_{i=1}^{n} (\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^{n} (Y_i - \bar{Y})^2}$.

Практичні завдання

1. Подані значення ціни на товар (грн) і обсяг пропозиції (тис. грн) (табл. 2.5). Необхідно оцінити параметри моделі за допомогою МНК. Перевірити статистичну значущість параметрів і моделі в цілому за допомогою критерію Стьюдента, коефіцієнта детермінації, критерію Фішера. Навести економічну інтерпретацію параметрів моделі, розрахувати коефіцієнт еластичності. Знайти точковий та інтервальний прогноз обсягу пропозиції, якщо ціна товару складе 26 грн.

Таблиця 2.5

Вихідні дані

№ п/п	1	2	3	4	5	6	7	8
X	12	16	17	18	22	24	28	30
Y	36	42	48	53	54	50	58	60

2. Перед маркетинговим відділом підприємства поставлено завдання проаналізувати залежність обсягу попиту Y (тис. грн) від ціни на товар X_1 (грн) І витрат на рекламу X_2 (тис. грн). У табл. 2.6 наведені вихідні дані. Необхідно: визначити параметри лінійної економетричної моделі за допомогою МНК; перевірити за допомогою критерію Стьюдента гіпотезу про суттєвість впливу витрат на рекламу і ціни на величину реалізованої продукції; оцінити статистичну значущість та адекватність моделі; знайти прогнозне значення обсягу реалізації, якщо ціна складе 10 грн, а витрати на рекламу – 5,6 тис. грн. Зробити економічні висновки.

Таблиця 2.6

Вихідні дані

ї

№ п/п	1	2	3	4	5	6	7	8
X_1	8,3	5,6	14,7	3,8	5,0	4,5	4,1	6,0
X_2	6,5	9,2	2,7	7,1	13,2	15,2	12,5	5,9
Y	16,5	21,6	7,5	18,6	29,0	29,7	30,0	15,6

3. Наведені значення числа кредитних договорів (X, тис. од) і доходу банку (Y, тис. грн) (табл. 2.7). Необхідно перевірити гіпотезу про суттєвий вплив числа кредитних договорів та інших факторів на дохід.

Таблиця 2.7

Вихідні дані

№ п/п	1	2	3	4	5	6	7	8	9	10
X	5,4	7,6	2,3	5,9	11,0	12,6	10,4	4,9	2,4	1,6
Y	13,7	18,0	6,2	15,5	24,1	24,8	25,0	13,0	8,1	6,7

4. Відомі дані про такі показники підприємства: випуск готової продукції підприємства – Y (млн грн), відсоток устаткування, що знаходиться на поточному ремонті – X_1 (%), витрати праці – X_2 (млн грн) (табл. 2.8). Необхідно визначити параметри лінійної моделі в натуральному і стандартизованому вигляді.

Таблиця 2.8

Вихідні дані

№ п/п	1	2	3	4	5	6	7	8
X_1	2,2	3,1	5,4	6,1	7,8	5,6	4,2	4,4
X_2	5,4	3,2	5,6	7,8	3,5	6,8	4,8	5,1
Y	38,7	16,8	34,3	52,0	10,0	44,7	29,4	31,8

5. У табл. 2.9 наведено вихідні дані показників діяльності банку.

Необхідно перевірити наявність лінійного взаємозв'язку між відповідними показниками діяльності банку. Побудувати лінійну багатофакторну економетричну модель. Провести оцінювання адекватності та статистичної значущості моделі. Розрахувати прогнозне значення залежної змінної та інтервал його зміни, якщо значення факторних ознак зміняться на 10 % щодо їх середнього значення.

Таблиця 2.9

Вихідні дані

№ п/п	Ставка відсотка за кредитами для фізичних осіб (X_1)	Сума кредитів юридичним особам (X_2)	Сума довгострокових кредитів (X_3)	Дохід банку (Y)
1	18	2 595	5 126	8 572
2	13	6 190	2 430	9 046
3	18	1 452	2 705	4 728
4	15	1 599	2 510	4 694
5	14	16 622	25 280	46 942
6	16	11 075	16 111	32 266
7	19	992	1 933	3 755
8	15	863	2 510	3 927
9	15	4 740	7 500	13 691
10	15	1 923	3 976	5 800
11	18	1 326	2 790	4 518
12	19	767	1 493	2 769
13	17	2 439	3 394	7 608
14	17	4 020	6 443	12 206
15	15	2 104	3 714	6 462
16	16	6 144	10 918	19 262
17	16	3 770	5 145	10 219
18	16	1 259	2 390	4 047
19	16	1 291	1 861	3 779
20	18	856	1 677	2 914
21	16	7 323	9 964	21 578
22	17	1 378	2 609	4 589
23	16	970	1 973	3 616
24	17	1 480	2 822	4 743

Ключові слова

Парна лінійна регресія. Множинна лінійна регресія. Метод найменших квадратів (МНК). Оцінювання значущості параметрів. Критерій Стьюдента. Оцінювання значущості моделі. Критерій Фішера. Ефективність оцінки. Обґрунтованість оцінки. Незміщеність оцінки. Довірчі інтервали. Коефіцієнт еластичності. Стандартна помилка прогнозу. Точковий прогноз. Інтервальний прогноз. Рівняння регресії в стандартизованому вигляді. Оцінювання параметрів через коефіцієнти парних кореляцій. Коефіцієнт детермінації. Коефіцієнт множинної кореляції. Коефіцієнт парної кореляції.

Розділ 3. Мультиколінеарність та її вплив на оцінювання параметрів моделі

3.1. Поняття мультиколінеарності. Вплив мультиколінеарності на характеристики множинної лінійної моделі.

3.2. Методи оцінювання ступеня мультиколінеарності. Метод Фаррара – Глобера.

3.3. Методи виключення мультиколінеарності.

3.1. Поняття мультиколінеарності. Вплив мультиколінеарності на характеристики множинної лінійної моделі

Однією з передумов класичної регресійної моделі виступає незалежність стовпців $X_j,\ j = 1, \dots, m$ вихідної матриці незалежних чинників моделі (регресорів). Число стовпців матриці регресорів дорівнює m, отже, число лінійно-незалежних стовпців також повинно дорівнювати m, тобто рангу матриці:

$$rank\ X\ = rank\ X^T X\ = m.$$

Якщо серед стовпців матриці є лінійно-залежні, то:

$$rank\ X\ = rank\ X^T X\ < m,$$

а, отже, визначник такої матриці буде дорівнювати нулю:

$$det\ X^T X\ = 0,$$

тобто матриця є виродженою і не має оберненої матриці. Знайти МНК-оцінки параметрів моделі у цьому випадку неможливо.

Тому під час побудови множинної регресії необхідно перевіряти відсутність тісних зв'язків між чинниками $X_1, X_2, \dots, X_m$.

Мультиколінеарність (multicollinearity) – наявність тісних лінійних зв'язків між включеними до множинної економетричної моделі екзогенними змінними.

Існує повна і часткова мультиколінеарність.

Повна (екстремальна) мультиколінеарність виникає тоді, коли всі регресори пов'язані лінійною залежністю, що не дозволяє розділити внесок окремої екзогенної змінної (X_i) в залежну змінну Y. Повна колінеарність призводить до невизначеності значень параметрів. Якщо розглянути 3-вимірний простір коефіцієнтів, то в цьому просторі вектор дійсних коефіцієнтів у даному випадку не єдиний, а є цілою прямою лінією! Будь-яка точка цієї прямої – істинний вектор коефіцієнтів.

Слід проілюструвати випадок повної мультиколінеарності на прикладі двофакторної множинної регресії.

Нехай специфікація економетричної моделі має вигляд:

$$Y = a_0 + a_1X_1 + a_2X_2 + \varepsilon. \tag{3.1}$$

І нехай між регресорами існує строга лінійна залежність:

$$X_2 = \alpha_0 + \alpha_1X_1. \tag{3.2}$$

Якщо підставити в рівняння (3.1) вираз (3.2), то буде отримано:

$$Y = a_0 + a_1X_1 + a_2\left(\alpha_0 + \alpha_1X_1\right) + \varepsilon,$$

або

$$Y = \left(a_0 + a_2\alpha_0\right) + (a_1 + a_2\alpha_1)X_1 + \varepsilon.$$

Введемо відповідні позначення для параметрів моделі, отримаємо специфікацію рівняння парної регресії вигляду:

$$Y = b_0 + b_1X_1 + \varepsilon, \tag{3.3}$$

де $b_0 = a_0 + a_2\alpha_0$,
$b_1 = a_1 + a_2\alpha_1$.

За оцінками параметрів b_0 і b_1 неможна однозначно визначити параметри регресії (3.1), оскільки в системі (3.3) невідомих більше, ніж початкових даних.

На практиці повна мультиколінеарність зустрічається рідко.

Часткова (недосконала, стохастична) мультиколінеарність характерна для випадків, коли частина екзогенних чинників $(X_1, X_2, \ldots, X_m)$ знаходиться в кореляційному зв'язку або утворює різні лінійні комбінації вигляду $X_i = \alpha_1 X_1 + \alpha_2 X_2 + \ldots + \alpha_m X_m$. Вона характеризується коефіцієнтами кореляції між незалежними змінними моделі r_{ij}. Фактори X_i і X_j можуть бути визнані колінеарними, якщо $r_{X_i,X_j} > 0{,}8$. Якщо між регресорами є високий ступінь мультиколінеарності, тобто $r_{ij} \approx 1$, то матриця $X^T X$ має повний ранг, але близька до виродженої, тобто $\det X^T X \approx 0$. Тому у разі поганої обумовленості матриці $X^T X^{-1}$ МНК-оцінки, оцінки залежної змінної Y можуть бути отримані, але виявляться незмістовними і нестійкими. Оцінка називається нестійкою, якщо невелика зміна вихідних даних (додавання або виключення одного-двох спостережень) буде призводити до значної її зміни. Нестійкість виражається у збільшенні статистичної невизначеності – дисперсії оцінок. Це означає, що конкретні результати оцінювання можуть сильно відрізнятися для різних вибірок незважаючи на те, що вибірки однорідні. Таким чином, оцінки параметрів виходять неточними, а значить складно буде дати інтерпретацію впливу тих чи інших факторів на результативну змінну.

Причини виникнення часткової мультиколінеарності:

1. Використання малої сукупності спостережень під час побудови моделі.

2. Досліджувані факторні ознаки характеризують одну і ту ж сторону явища або процесу (наприклад, показники обсягу виробленої продукції та середньорічної вартості основних фондів одночасно включати в модель не рекомендується, оскільки обидва характеризують розмір підприємства) [38].

3. Використання у якості факторних ознак таких, сумарне значення яких є постійною величиною (наприклад, коефіцієнт придатності і коефіцієнт зносу основних фондів).

4. Факторні ознаки, що є елементами один одного (наприклад, витрати на виробництво продукції і собівартість одиниці продукції).

5. Факторні ознаки, за економічним змістом дублюють одна одну (наприклад, прибуток і рентабельність продукції).

6. Факторні ознаки можуть мати загальний часовий тренд, щодо якого вони здійснюють малі коливання.

7. Значення однієї з факторних ознак є лаговим значенням другої, тобто між зв'язком показників існує зрушення в часі.

Наслідки часткової мультиколінеарності у регресійній моделі:

1. Збільшення дисперсій оцінок параметрів. Це розширює інтервальні оцінки і погіршує їх точність.

2. Оцінки мають великі стандартні похибки і малу значущість (тобто відбувається зменшення t-статистики), тоді як модель у цілому є значущою (спостерігається високе значення коефіцієнта детермінації і відповідної F-статистики) Це призводить до невиправданого висновку про статистичну незначущість регресорів.

3. Оцінки параметрів мають невиправдано великі значення або невірні знаки. Невелика зміна початкових даних (наприклад, додавання нових спостережень) призводить до істотної зміни оцінок параметрів моделі.

Таким чином, перший і третій наслідки мультиколінеарності не дозволяють проводити аналіз економічної проблеми на основі такого регресійного рівняння, але адекватність і статистична значущість моделі в цілому дозволяють її використання для короткострокового прогнозування.

3.2. Методи оцінювання ступеня мультиколінеарності. Метод Фаррара – Глобера

Для визначення наявності мультиколінеарності в регресійній моделі використовуються такі способи [40]:

1. Величина визначника матриці $X^T X$.

Якщо $det\ X^T X\ = 1$, то мультиколінеарність відсутня. Чим ближче $det\ X^T X$ до нуля, тим більший ступіньмультиколінеарності між включеними до моделі факторними ознаками. Цей показник вважають точковою мірою мультиколінеарності.

2. Мінімальне власне число матриці $X^T X$. Власні числа (характеристичні корені) λ_i знаходяться через рішення характеристичного рівняння матриці $X^T X$:

$$det\ X^T X - E\lambda\ = 0.$$

Мінімальне власне число визначається як:

$$\lambda_{min} = \min_i\ \lambda_i\ .$$

Значення λ впливають на дисперсії оцінок параметрів. Чим менше λ_{min}, тим сильніша мультиколінеарність.

3. Міра обумовленості матриці $X^T X$ *за Нейманом – Гольдштейном*:

$$\frac{\lambda_{max}}{\lambda_{min}},$$

де $\lambda_{max} = \max_i \lambda_i$ – максимальне власне число матриці $X^T X$.

Чим ближче $\frac{\lambda_{max}}{\lambda_{min}}$ до нескінченності, тим сильніша мультиколінеарність. Існує деяка шкала для даного відношення, що дозволяє провести якісне оцінювання ступеня мультиколінеарності за величиною міри Неймана – Гольдштейна: якщо це відношення більше 30, то мультиколінеарність середнього ступеня, якщо більше 100 – мультиколінеарність велика.

4. Максимальна зв'язаність. Для побудови цього показника розраховується регресія змінної i на решту незалежних змінних з номерами 1, 2, ..., i – 1, i + 1, ..., m, визначається коефіцієнт детермінації R_i^2 в регресії X_i на $X_1, X_2, \ldots, X_{i-1}, X_{i+1}, \ldots, X_m$. Як міра зв'язаності використовується величина $\max_i(R_i^2)$. Якщо значення $\max_i(R_i^2)$ близьке до 1, то сукупність незалежних змінних схильна до значного впливу ефекту мультиколінеарності.

5. *Метод Фаррара – Глобера (Farrar – Glauber Test).* Даний метод заснований на застосуванні трьох видів статистичних критеріїв:

1) усього масиву незалежних змінних (критерій χ^2);

2) кожної незалежної змінної з усіма іншими (F-критерій);

3) кожної пари незалежних змінних (t-критерій).

Порівнявши ці критерії з їх критичними значеннями, можна зробити висновок щодо наявності чи відсутності мультиколінеарності між незалежними змінними.

Розглянемо *покроковий алгоритм методу Фаррара – Глобера*.

1 крок. Обчислити парні коефіцієнти кореляції для вихідних екзогенних змінних $X_1, X_2, \ldots, X_m$за формулою:

$$r_{ij} = \frac{\sum_{k=1}^{n} (x_{ki} - x_i)(x_{kj} - x_j)}{\sqrt{\sum_{k=1}^{n} (x_{ki} - x_i)^2 \sum_{k=1}^{n} (x_{kj} - x_j)^2}},$$

де r_{ij}– парний коефіцієнт кореляції між екзогенними змінними X_i і X_j, що характеризує щільність зв'язку між ними;

x_{ki} – значення екзогенної змінної X_i для i-го спостереження ($k = \overline{1,n}$);

$\bar{x}_i,\ \bar{x}_j$ – середні значення відповідно для X_i і X_j.

Слід зауважити, якщо число екзогенних змінних більше двох, то на основі високих значень парних коефіцієнтів кореляції не можливо стверджувати, що отриманий зв'язок є явищем мультиколінеарності. Необхідна подальша перевірка.

2 крок. З отриманих значень парних коефіцієнтів кореляції побудувати кореляційну матрицю $R = (r_{ij})$, тобто матрицю моментів нормалізованої системи нормальних рівнянь:

$$R = (r_{ij}) = \begin{pmatrix} 1 & r_{12} & \dots & r_{1m} \\ r_{21} & 1 & \dots & r_{2m} \\ \dots & \dots & \dots & \dots \\ r_{m1} & r_{m2} & \dots & 1 \end{pmatrix}.$$

3 крок. Розрахувати визначник $\det(R)$ кореляційної матриці R та обчислити критерій χ^2 за формулою:

$$\chi^2 = -\left(n - 1 - \frac{1}{6}(2m + 5)\right)\ln(\det(R)).$$

Порівняти отримане значення χ^2 з табличним з 0,5m(m-1) ступенями свободи і рівні значущості α. Якщо $\chi^2 \geq \chi^2_{табл}$, то в масиві незалежних змінних існує загальна мультиколінеарність.

4 крок. Визначити матрицю похибок Z, обернену до матриці $R = (r_{ij})$:

$$Z = R^{-1} = \begin{pmatrix} z_{11} & z_{12} & \dots & z_{1m} \\ z_{21} & z_{22} & \dots & z_{2m} \\ \dots & \dots & \dots & \dots \\ z_{m1} & z_{m2} & \dots & z_{mm} \end{pmatrix}.$$

5 крок. Визначити множинні коефіцієнти кореляції $R_i\ (i = \overline{1,m})$, що характеризують тісноту зв'язку кожної змінної з іншими змінними:

$$R_i = \sqrt{1 - \frac{1}{z_{ii}}}.$$

6 крок. Перевірити статистичну значущість зв'язку кожної змінної з іншими змінними на основі F-критерію:

$$F_i = \frac{z_{ii} - 1 \quad n - m}{m - 1},$$

де z_{ii}– діагональні елементи матриці Z.

Значення F_i порівняти з табличними значеннями з $(m - 1)$ і $(n - m)$ ступенями свободи та рівні значущості α. Якщо $F_i \geq F_{\text{табл}}$, то відповідна i-та незалежна змінна мультиколінеарна з іншими.

7 крок. Знайти частинні коефіцієнти кореляції, які змінюються в межах $-1; 1$ і характеризують щільність зв'язку між двома змінними за умови, що інші змінні не впливають на цей зв'язок, тобто досліджується існування парної мультиколінеарності:

$$r_{ij}^{ч} = \frac{-z_{ij}}{\overline{z_{ii} \cdot z_{jj}}},$$

де z_{ij} – елементи матриці Z, що розміщені в i-му рядку та j-му стовпці, $i = 1, m$, $j = 1, m$;

z_{ii} і z_{jj} – діагональні елементи матриці Z.

Однак якщо порівняти конкретні числові значення частинних та парних коефіцієнтів, то можна побачити, що перші значно відрізняються від останніх. Тому на основі знання коефіцієнтів кореляції висновок про мультиколінеарність однозначно зробити неможливо.

8 крок. Перевірити статистичну значущість зв'язку між кожними двома змінними на основі розрахунку t-критерію за формулою:

$$t_{ij} = r_{ij}^{ч} \frac{\overline{n - m}}{\overline{1 - r_{ij}^{ч}}^{\,2}}.$$

Розрахункові значення t-статистик порівнюємо з табличними з $(n - m)$ ступенями свободи та рівнем значущості α. Якщо $t_{ij} \geq t_p(\alpha,\ n - m)$, то між незалежними змінними X_i і X_j існує мультиколінеарність.

Приклад 3.1. На рівень соціально-економічного розвитку регіонів впливає низка чинників, серед яких середньомісячна заробітна платня робітників (грн), чисельність працездатного населення (осіб на 1 000 населення), обсяг виробництва товарів народного вжитку на душу населення (грн). Перш ніж будувати множинну регресію, необхідно перевірити наявність мультиколінеарності у масиві екзогенних змінних. Вихідні дані наведені в табл. 3.1.

Таблиця 3.1

Вихідні дані

Регіон	Середньомісячна заробітна платня робітників, грн X_1	Чисельність працездатного населення, осіб на 1 000 населення X_2	Виробництво товарів народного вжитку, грн X_3
1	2 694	386,25	400,84
2	2 591	392,85	247,29
3	3 385	440,44	319,75
4	3 811	424,88	266,04
5	2 617	391,4	336,58
6	2 695	403,52	203,09
7	3 155	431,55	373,46
8	2 682	375,63	190,93
9	3 528	412,09	470,23
10	2 660	397,98	245,37
11	3 335	410,22	205,58
12	2 791	403,77	314,22
13	3 321	423,97	237,99
14	3 100	412,47	365,39
15	3 215	411,61	379,50
16	2 903	383,5	181,03
17	2 722	401,09	446,23
18	2 466	373,42	336,50
19	3 068	436,21	434,68
20	2 543	413,93	342,11

Перевіримо наявність мультиколінеарності за допомогою алгоритму Фаррара – Глобера.

1 крок. Розрахуємо парні коефіцієнти кореляції між факторами за формулою:

$$r_{ij} = \frac{\sum_{k=1}^{n} \left(X_{ki} - \bar{X}_i\right)\left(X_{kj} - \bar{X}_j\right)}{\sqrt{\sum_{k=1}^{n} \left(X_{ki} - \bar{X}_i\right)^2 \sum_{k=1}^{n} \left(X_{kj} - \bar{X}_j\right)^2}}.$$

Для цього знайдемо спочатку середні значення для кожного з регресорів:

$$\bar{X}_1 = \frac{\sum_{k=1}^{20} X_{k1}}{20} = \frac{59\,282}{20} = 2964{,}1;$$

$$\bar{X}_2 = \frac{\sum_{k=1}^{20} X_{k2}}{20} = \frac{8\,126{,}78}{20} = 406{,}339;$$

$$\bar{X}_3 = \frac{\sum_{k=1}^{20} X_{k3}}{20} = \frac{6\,296{,}81}{20} = 314{,}84.$$

Подальші розрахунки простіше вести у табличній формі, як це показано в табл. 3.2.

Отримаємо такі парні коефіцієнти кореляції між регресорами:

$$r_{12} = \frac{95\,021{,}5}{\sqrt{2\,700\,827{,}8 \cdot 6\,985{,}492}} \approx 0{,}6918,$$

$$r_{13} = \frac{61\,927{,}3}{\sqrt{2\,700\,827{,}8 \cdot 150\,410{,}52}} \approx 0{,}0972,$$

$$r_{23} = \frac{9\,507{,}28}{\sqrt{6\,985{,}492 \cdot 150\,410{,}52}} \approx 0{,}2933.$$

Як бачимо, жоден із цих коефіцієнтів не перевищує 0,8. Але говорити про відсутність мультиколінеарності зарано.

2 крок. Сформуємо кореляційну матрицю:

$$R = \left(r_{ij}\right) = \begin{pmatrix} 1 & 0{,}6918 & 0{,}0972 \\ 0{,}6918 & 1 & 0{,}2933 \\ 0{,}0972 & 0{,}2933 & 1 \end{pmatrix}.$$

Таблиця 3.2

Проміжні розрахунки

Регіон	$X_1 - X_1$	$X_2 - X_2$	$X_3 - X_3$	$(X_1 - X_1)^2$	$(X_2 - X_2)^2$	$(X_3 - X_3)^2$	$X_1 - X_1 \times (X_2 - X_2)$	$X_1 - X_1 \times (X_3 - X_3)$	$X_2 - X_2 \times (X_3 - X_3)$
1	-270,1	-20,089	86,00	72 954,01	403,568	7 396,000	5 426,04	-23 228,6	-1 727,65
2	-373,1	-13,489	-67,55	139 203,61	181,953	4 563,003	5 032,75	25 202,9	911,18
3	420,9	34,101	4,91	177 156,81	1 162,88	24,108	14 353,1	2 066,6	167,44
4	846,9	18,541	-48,80	717 239,61	343,769	2 381,440	15 702,4	-41 328,7	-904,80
5	-347,1	-14,939	21,74	120 478,41	223,174	472,628	5 185,33	-7 546	-324,77
6	-269,1	-2,819	-111,75	72 414,81	7,947	12 488,06	758,59	30 071,9	315,02
7	190,9	25,211	58,62	36 442,81	635,595	3 436,304	4 812,78	11 190,6	1477,87
8	-282,1	-30,709	-123,91	79 580,41	943,043	15 353,69	8 663,01	34 955	3 805,15
9	563,9	5,751	155,39	317 983,21	33,074	24 146,05	3 242,99	87 624,4	893,65
10	-304,1	-8,359	-69,47	92 476,81	69,873	4 826,081	2 541,97	21 125,8	580,70
11	370,9	3,881	-109,26	137 566,81	15,062	11 937,75	1 439,46	-40 524,5	-424,04
12	-173,1	-2,569	-0,62	29 963,61	6,600	0,384	444,69	107,3	1,59
13	356,9	17,631	-76,85	127 377,61	310,852	5 905,923	6 292,5	-27 427,8	-1 354,94
14	135,9	6,131	50,55	18 468,81	37,589	2 555,303	833,2	6 869,7	309,92
15	250,9	5,271	64,66	62 950,81	27,783	4 180,916	1 322,49	16 223,2	340,82
16	-61,1	-22,839	-133,81	3 733,21	521,620	17 905,12	1 395,46	8 175,8	3 056,09
17	-242,1	-5,249	131,39	58 612,41	27,552	17 263,33	1 270,78	-31 809,5	-689,67
18	-498,1	-32,919	21,66	248 103,61	1 083,66	469,156	16 396,9	-10 788,8	-713,03
19	103,9	29,871	119,84	10 795,21	892,277	14 361,63	3 103,6	12 451,4	3 579,74
20	-421,1	7,591	27,27	177 325,2	57,623	743,653	-3 196,57	-11 483,4	207,01
Сума	**0,0**	**0,0**	**0,0**	**2 700 827,8**	**6 985,492**	**150 410,52**	**95 021,50**	**61 927,3**	**9 507,28**

3 крок. Розраховуємо визначник кореляційної матриці $\det R = 0{,}4654$ та обчислюємо критерій χ^2 за формулою:

$$\chi^2 = -\left(n - 1 - \frac{1}{6}(2m + 5)\right) \ln \det R =$$
$$= -\left(20 - 1 - \frac{1}{6}(2 \cdot 3 + 5)\right) \ln 0{,}4654 \approx 13{,}13.$$

Табличне значення критерію χ^2 з числом ступенів свободи $k = 0{,}5m \cdot$ $\cdot(m-1) = 3$ і рівні значущості $\alpha = 0{,}95$ дорівнює $\chi^2_{\text{табл}}(\alpha; 0{,}5m(m-1)) =$ $= 7{,}81$. Оскільки $\chi^2 \geq \chi^2_{\text{табл}}$, то в масиві незалежних змінних існує загальна мультиколінеарність і дослідження необхідно продовжити.

4 крок. Визначаємо матрицю похибок Z, обернену до матриці $R = (r_{ij})$:

$$Z = R^{-1} = \begin{pmatrix} 1{,}9639 & -1{,}4253 & 0{,}2272 \\ -1{,}4253 & 2{,}1285 & -0{,}4858 \\ 0{,}2272 & -0{,}4858 & 1{,}1204 \end{pmatrix}.$$

5 крок. Визначаємо множинні коефіцієнти кореляції R_i, що характеризують тісноту зв'язку кожної змінної з іншими змінними:

$$R_1 = \sqrt{1 - \frac{1}{1{,}9639}} \approx 0{,}70;$$
$$R_2 = \sqrt{1 - \frac{1}{2{,}1285}} \approx 0{,}73;$$
$$R_3 = \sqrt{1 - \frac{1}{1{,}1204}} \approx 0{,}33.$$

Як бачимо, перший та другий регресори мають тісний зв'язок з іншими регресорами.

6 крок. Перевіряємо статистичну значущість зв'язку кожної змінної з іншими змінними на основі F-критерію:

$$F_1 = \frac{(1{,}9639 - 1)(20 - 3)}{3 - 1} \approx 8{,}19;$$

$$F_2 = \frac{(2{,}1285 - 1)(20 - 3)}{3 - 1} \approx 9{,}59;$$
$$F_3 = \frac{(1{,}1204 - 1)(20 - 3)}{3 - 1} \approx 1{,}02.$$

Табличне значення F-статистики з рівнем значущості $\alpha = 0{,}95$ і ступенями свободи $k_1 = (m - 1) = 3 - 1 = 2$ і $k_2 = (n - m) = 20 - 3 = 17$ дорівнює $F_{\text{табл}}(0{,}95; 2; 17) = 3{,}59$.

Оскільки розрахункові значення $F_1 > F_{\text{табл}}$ і $F_2 > F_{\text{табл}}$, то відповідно регресори X_1 і X_2 мультиколінеарні з іншими.

7 крок. Знаходимо частинні коефіцієнти кореляції:

$$r_{12}^{\text{ч}} = \frac{-(-1{,}4253)}{\sqrt{1{,}9639 \cdot 2{,}1285}} \approx 0{,}6971,$$
$$r_{13}^{\text{ч}} = \frac{-(0{,}2272)}{\sqrt{1{,}9639 \cdot 1{,}1204}} \approx -0{,}1532,$$
$$r_{23}^{\text{ч}} = \frac{-(-0{,}4858)}{\sqrt{2{,}1285 \cdot 1{,}1204}} \approx 0{,}3146.$$

Між регресорами X_1 і X_2 існує тісний зв'язок.

Цікаво порівняти отримані значення частинних і парних коефіцієнтів кореляції. Зазвичай перші значно менші за останні. Але в даному випадку частинні коефіцієнти, які характеризують щільність зв'язку між двома змінними за умови, що інші змінні не впливають на цей зв'язок, за абсолютним значенням більші, ніж парні, більш того для регресорів X_1 і X_3 відбулася навіть зміна напрямку зв'язку з прямого на зворотний. Таким чином, можна зробити висновок, що інші змінні досить суттєво впливають на зв'язок між досліджуваними показниками.

8 крок. Перевіряємо статистичну значущість зв'язку між кожними двома змінними на основі розрахунку t-критерію за формулою:

$$t_{12} = \left|r_{12}^{\text{ч}}\right| \sqrt{\frac{n - m}{1 - \left(r_{12}^{\text{ч}}\right)^2}} = 0{,}6971 \cdot \sqrt{\frac{20 - 3}{1 - (0{,}6971)^2}} \approx 4{,}0;$$
$$t_{13} = |-0{,}1532| \cdot \sqrt{\frac{20 - 3}{1 - (-0{,}1532)^2}} \approx 0{,}639;$$

$$t_{23} = 0{,}3146 \cdot \frac{\sqrt{20-3}}{\sqrt{1-(0{,}3146)^2}} \approx 1{,}366.$$

Розрахункові значення t-статистик порівнюємо з табличними при ступені свободи $k = n - m = 20 - 3 = 17$ та рівні значущості $\alpha = 0{,}95$. Оскільки $t_{12} \geq t_p(0{,}95; 17) = 2{,}11$, то між незалежними змінними X_1 і X_2 існує статистично значуща мультиколінеарність.

3.3. Методи виключення мультиколінеарності

1. Методи виключення змінних моделі.

Одним із найбільш поширених способів усунення мультиколінеарності є виключення незалежних чинників із моделі. Розглянемо найбільш поширені методи.

1.1. Метод додаткової регресії.

Для виявлення списку залежних регресорів будується додаткова регресія, тобто регресія кожного регресора $X_j (j = \overline{1, m})$ на регресори, що залишилися (так само, як під час визначення максимальної зв'язаності). На основі F-статистики перевіряється статистична значущість коефіцієнтів детермінації додаткових регресійних моделей:

$$F_j = \frac{R_j^2}{1 - R_j^2} \cdot \frac{n - m}{m - 1},$$

де n – число спостережень (обсяг вибірки);

m – число регресорів у первинній специфікації моделі.

Статистика F_j має розподіл Фішера з числом ступенів свободи $k_1 = (m - 1)$, $k_2 = (n - m)$. Якщо коефіцієнт статистично незначущий, то регресор не приводить до мультиколінеарності і його залишають у списку змінних моделі. Інакше його слід виключити з моделі.

1.2. Метод виключення одного з двох сильно зв'язаних чинників.

Для визначення кореляційно залежних чинників розраховують матрицю парних коефіцієнтів кореляції і визначають регресори, які мають значення коефіцієнтів, близькі до одиниці. Вибір регресора, який необхідно виключити з моделі, здійснюється за критерієм мінімального впливу регресора на незалежний чинник Y.

У лінійних моделях коефіцієнти кореляції між факторами можуть бути додатними і від'ємними. У першому випадку збільшення одного фактора супроводжується збільшенням іншого фактора. У другому випадку за умови підвищення одного фактора відбувається зниження іншого. Виходячи з цього, можна встановити припустиму і неприпустиму мультиколінеарності.

Неприпустима мультиколінеарність буде тоді, коли між факторами X_i і X_j існує значна додатна кореляція і при цьому вплив кожного фактора на кореляційний зв'язок з Y односпрямований, тобто збільшення обох факторів X_i і X_j веде до збільшення або зниження Y: r_{yi} $r_{yj} > 0$.

Іншими словами, обидва фактори діють на Y однаково і значна додатна кореляція між ними може дозволити виключити один із них.

Припустима мультиколінеарність така, за якої фактори діють на Y неоднаково. Тут можливі два випадки:

а) у разі значної додатної кореляції між факторами вплив кожного фактора на кореляційний зв'язок з Y різноспрямований, тобто збільшення одного фактора веде до зростання значень Y ($r_{yi} > 0$), а збільшення іншого фактора призводить до зменшення Y ($r_{yj} < 0$): r_{yi} $r_{yj} < 0$;

б) за умови значної від'ємної кореляції між факторами збільшення одного фактора супроводжується зменшенням іншого і це робить фактори різнозначними, тому можливий будь-який знак впливу факторів на функцію Y: r_{yi} $r_{yj} >< 0$.

У разі наявності неприпустимої мультиколінеарності виключення одного з факторів проводиться в наступній послідовності (цей же підхід корисний для перевірки наявності припустимої мультиколінеарності) [16]:

1) з двох факторів, пов'язаних значною кореляцією, виключається, насамперед, фактор на підставі теоретичних міркувань. Якщо такий підхід не дає результату, то виключається той фактор, якому відповідає менший коефіцієнт кореляції з функцією;

2) після видалення фактора економетричній моделі повинен відповідати більший коефіцієнт множинної кореляції, ніж до видалення фактора.

Тоді це підтверджує наявність неприпустимої мультиколінеарності між розглянутими факторами і правильність видалення одного з них.

1.3. Метод коефіцієнта детермінації.

Обчислюються коефіцієнти детермінації для регресії Yна регресори $X_1, X_2, \dots, X_m$ і розраховується їх статистична значущість на основі F-статистики. У регресії залишають той набір регресорів, який забезпечує максимальне значення коефіцієнта детермінації. Даний метод реалізований у статистичних пакетах прикладних програм у вигляді методу послідовного включення/виключення незалежних змінних моделі.

1.3.1. Метод включення факторів.

Метод полягає в тому, що в модель включаються фактори по одному в певній послідовності. На першому кроці в модель вводиться той фактор, який має найбільший коефіцієнт кореляції з залежною змінною. На другому і наступних кроках до моделі включається фактор, який має найбільший коефіцієнт кореляції із залишками моделі. Після включення кожного фактора до моделі розраховують її характеристики, і перевіряють адекватність моделі.

Побудова моделі закінчується, якщо модель перестає задовольняти певним умовам (наприклад, $k < n/3$, $S_{\varepsilon,k-1} - S_{\varepsilon,k} > \zeta$, де n – число спостережень; k – число включених до моделі факторних чинників; ζ – деяке задане мале число; $S_{\varepsilon,k}$ – середньоквадратична помилка; $S_{\varepsilon,k-1}$ – середньоквадратична помилка моделі, отримана на попередньому кроці і включає $k-1$ змінних)

1.3.2. Метод виключення факторів.

Метод полягає в тому, що до моделі включаються всі фактори. Потім після побудови рівняння регресії з моделі виключають фактор, коефіцієнт при якому незначущий і має найменше значення t-критерію. Після цього отримують нове рівняння регресії і знову проводять оцінювання значущості всіх решти коефіцієнтів регресії.

Процес виключення факторів продовжується до того часу, поки модель не стане відповідати певним умовам і всі коефіцієнти регресії будуть значущі.

2. Методи, що використовують зовнішню інформацію.

Під зовнішньою інформацією розуміються теоретичні обмеження (тобто деякі обмеження щодо параметрів моделі чи деякого зв'язку між ними) та зовнішні емпіричні оцінки (наприклад, отриманні з перехресних статистичних даних з інших вибірок).

Збільшення числа спостережень приводить до зменшення мультиколінеарності, що, у свою чергу, приводить до зменшення теоретичних дисперсій коефіцієнтів регресії.

3. Перехід до зміщених методів оцінювання.

В умовах мультиколінеарності дисперсії зміщених оцінок можуть виявитися дуже великими. Тому іноді доцільно відмовитися від вимоги незміщеності. Одним із таких підходів називається "рідж-регресія" (гребенева регресія). Він заснований на "коректуванні" МНК-оцінок за допомогою такого перетворення:

$$b_\tau = \left(X^T X + \tau E_m\right)^{-1} X^T Y = D^{-1} X^T Y,$$

де τ– деяке невелике додатне число;

E_m– одинична матриця розмірністю $m \times m$.

Додавання до діагональних елементів матриці $X^T X$ "гребеня" τ приводить до зміщеності оцінок. Але в результаті таких перетворень замість погано обумовленої матриці обертається добре обумовлена матриця $\left(X^T X + \tau E_m\right)$. Універсальних рекомендацій щодо значень τ немає, зазвичай використовують τ у діапазоні від 0,1 до 0,4.

4. Методи перетворення даних.

Замість абсолютних значень вихідних змінних можна перейти до використання, наприклад, приростів, темпів зростання, темпів приросту змінних.

Під час дослідження ефекту мультиколінеарності може виявитися, що це явище пов'язане з наявністю у складі змінних лінійних або нелінійних трендів.

$$Y_t = Y(t) + \eta_t, \quad t = \overline{1,T},$$
$$X_{it} = X_i(t) + \zeta_{it}, \quad i = \overline{1,m}.$$

У цьому випадку рекомендується спочатку виділити і виключити тренди, а потім визначити параметри регресії на залишках, тобто:

$$\eta_t = \varphi(\zeta_{1t}, \zeta_{2t}, \dots, \zeta_{mt}) + \varepsilon_t.$$

Після цього будуть виконуватися передумови застосування МНК, зокрема математичне сподівання похибок дорівнюватиме 0 ($M\varepsilon_t = 0$), дисперсія буде сталою ($D\varepsilon_t = \sigma^2$).

5. Метод головних компонент.

Якщо переліченими методами не вдається усунути мультиколінеарність, то для оцінювання параметрів багатовимірної моделі доцільно застосовувати метод головних компонент.

Алгоритм методу головних компонент містить такі кроки:

1 крок. Нормалізація змінних $X_1, X_2, \ldots, X_m$ регресійної моделі шляхом такого перетворення:

$$x_{ij}^* = \frac{x_{ij} - x_j}{\sigma_{x_j}},$$

де x_{ij} – значення змінної X_j для спостереження i;

n – кількість спостережень $(i = 1, n)$;

m – кількість незалежних змінних $(j = 1, m)$;

x_j – середнє арифметичне j-ї незалежної змінної;

σ_{x_j} – середньоквадратичне відхилення j-ї незалежної змінної.

2 крок. Побудова матриці X^*, елементами якої є нормалізовані незалежні змінні.

3 крок. Визначення кореляційної матриці (матриці моментів нормалізованої системи рівнянь) за формулою:

$$R = \frac{1}{n}\ X^{*T}X^* \ = \ \begin{matrix} 1 & r_{12} & \ldots & r_{1m} \\ r_{21} & 1 & \ldots & r_{2m} \\ \ldots & \ldots & \ldots & \ldots \\ r_{m1} & r_{m2} & \ldots & 1 \end{matrix} \ .$$

Недіагональні елементи матриці R характеризують щільність зв'язку однієї незалежної змінної з іншою, тобто є парними коефіцієнтами кореляції.

4 крок. Визначення характеристичних (власних) чисел матриці R, тобто визначення коренів $\lambda_1, \lambda_2, \ldots, \lambda_m$ з наступного рівняння:

$$det\ \ R - \lambda E \ \ = 0,$$

де E – одинична матриця розмірності $m \times m$.

5 крок. Ранжування значень $\lambda_1, \lambda_2, \ldots, \lambda_m$ за абсолютним внеском кожного головного компонента в загальну дисперсію.

6 крок. Розв'язання системи рівнянь такого виду:

$$R - \lambda E \;\; a = 0,$$

та обчислення власних векторів a_k $k = 1, m$ за умови, що вони відповідають таким співвідношенням:

$$a_k{}^T a_j = \begin{matrix} 1, & j = k, \\ 0, & j \neq k. \end{matrix}$$

7 крок. Визначення головних компонент векторів $F_k = Xa_k$, $(k = 1, m)$, які задовольняють умови:

$$\sum_{i=1}^{n} f_{ik} = 0,$$
$$\frac{1}{n} F_k{}^T F_k = \lambda_k, \qquad k = 1, m \;\; ;$$
$$F_j{}^T F_k = 0, \qquad j = 1, m \;\; , \qquad j \neq k.$$

8 крок. Визначення параметрів моделі $Y = FP$ за умови, що $P = F^{-1}Y$.

9 крок. Визначення параметрів моделі $Y = FB$ за умови, що $B = aP$.

Треба підкреслити той факт, що метод головних компонент доцільно застосовувати, по-перше, для оцінювання параметрів моделей з великою кількістю факторів, по-друге, для моделей, у яких незалежні змінні мають однакові одиниці вимірювання.

Контрольні запитання для самодіагностики

1. Дайте визначення поняття мультиколінеарності.
2. Поясніть різницю між повною і частковою мультиколінеарністями.
3. Чим відрізняються припустима і неприпустима мультиколінеарності?
4. Які наслідки наявності мультиколінеарності у моделі?
5. Назвіть основні ознаки наявності мультиколінеарності у моделі.
6. Назвіть методи оцінювання ступеня мультиколінеарності.

7. Для чого і яким чином застосовують метод Фаррара – Глобера?

8. Чим відрізняються парний і частинний коефіцієнти кореляції між факторними змінними?

9. Назвіть основні способи усунення негативного ефекту мультиколінеарності.

10. Опишіть метод додаткової регресії.

11. Як застосовується алгоритм виключення зайвих факторів?

12. У чому полягають методи, що використовують зовнішню інформацію?

13. Які методи перетворення даних можна застосовувати для подолання негативних наслідків мультиколінеарності?

14. У чому полягає метод рідж-регресії?

15. Охарактеризуйте метод головних компонентів.

Тести

1. Під мультиколінеарністю розуміється лінійна залежність:

а) ендогенної змінної з одним або декількома регресорами;

б) між двома або декількома регресорами;

в) між регресорами і похибками моделі.

2. За умови мультиколінеарності матриця $X^T X$ *:*

а) близька до виродженої;

б) є виродженою;

в) є невиродженою.

3. Наявність часткової мультиколінеарності призводить до зменшення:

а) МНК-оцінок параметрів моделі;

б) дисперсій МНК-оцінок параметрів моделі;

в) t-статистик МНК-оцінок параметрів моделі.

4. Частинний коефіцієнт кореляції змінюється в межах:

а) від 0 до 1;

б) від -1 до 0;

в) від -1 до 1.

5. Для усунення мультиколінеарності використовують:

а) метод оцінювання максимальної зв'язаності;

б) метод Фаррара – Глобера;

в) метод рідж-регресії.

6. У разі мультиколінеарності всі оцінки параметрів моделі або їх більша частина будуть статистично:

а) значущими за умови високого значення коефіцієнта множинної кореляції;

б) незначущими за умови високого значення коефіцієнта множинної кореляції;

в) значущими за умови низького значення коефіцієнта множинної кореляції.

7. Ступінь мультиколінеарності тим більший, чим більше:

а) визначник матриці коефіцієнтів системи нормальних рівнянь X^TX ;

б) мінімальне власне число матриці X^TX ;

в) міра Неймана – Гольдштейна для матриці X^TX .

8. Для визначення наявності мультиколінеарності використовують:

а) звичайний метод найменших квадратів;

б) метод Фаррара – Глобера;

в) метод регресії на головних компонентах і метод гребневої регресії.

9. Найбільш повну перевірку наявності мультиколінеарності дає змогу провести:

а) визначник матриці X^TX ;

б) метод Фаррара – Глобера;

в) метод оцінювання максимальної зв'язаності.

10. При значній від'ємній кореляції між двома факторами, що входять до складу множинної регресії, напрямок зв'язку між кожним з них і залежною змінною:

а) має значення;

б) не має значення.

Практичні завдання

1. На основі стандартизованих даних було побудовано матрицю X^TX . Потрібно розрахувати визначник цієї матриці і зробити висновок про наявність мультиколінеарності.

$$X^TX = \begin{matrix} 1 & 0,849 & 0,601 \\ 0,849 & 1 & 0,868 \\ 0,601 & 0,868 & 1 \end{matrix} .$$

2. На основі відомої кореляційної матриці між ендогенною та екзогенними змінними моделі визначить, чи є мультиколінеарність між X_1 і X_3 та між X_1 і X_4 припустимою.

Змінні	Y	X_1	X_2	X_3	X_4
Y	1	0,977	0,374	0,706	-0,867
X_1	0,977	1	0,390	0,764	-0,898
X_2	0,374	0,390	1	0,549	-0,283
X_3	0,706	0,764	0,549	1	-0,607
X_4	-0,867	-0,898	-0,283	-0,607	1

3. За даними матриці Z, оберненої до кореляційної матриці, потрібно визначити частинні коефіцієнти кореляції між регресорами та перевірити їхню статистичну значущість за допомогою t-статистики, якщо обсяг вибірки $n = 25$.

$$Z = \begin{vmatrix} 2{,}3236 & -1{,}8978 & -0{,}6634 \\ -1{,}8978 & 2{,}7226 & 0{,}9917 \\ -0{,}6634 & 0{,}9917 & 1{,}3620 \end{vmatrix}.$$

4. За вихідними даними, наведеними у табл. 3.3, дослідити наявність мультиколінеарності між факторними змінними за допомогою алгоритму Фаррара – Глобера. Зробити висновки.

Таблиця 3.3

Вихідні дані

X_1	X_2	X_3	Y
2	16	15	9,2
1	18	18	9,4
2	14	20	10,2
4	10	22	10,4
4	12	21	10,5
5	9	21	11,3
8	12	24	11,5
9	8	25	11,6
5	5	28	12,3

5. За даними, наведеними у табл. 3.3, необхідно побудувати регресійні моделі на основі методів включення і виключення факторів. Порівняти отримані результати. Зробити висновки.

Ключові слова

Мультиколінеарність. Повна (екстремальна) мультиколінеарність. Часткова (недосконала, стохастична) мультиколінеарність. Визначник матриці $X^T X$. Метод Фаррара – Глобера. Методи виключення змінних моделі. Ридж-регресія. Методи перетворення даних. Метод головних компонент.

Розділ 4. Узагальнений метод найменших квадратів

4.1. Гетероскедастичність в економетричних моделях та методи її визначення.

4.2. Узагальнений метод найменших квадратів (метод Ейткена).

4.1. Гетероскедастичність в економетричних моделях та методи її визначення

Застосування методу найменших квадратів для оцінювання параметрів класичної економетричної моделі передбачає, що повинні виконуватися певні передумови щодо випадкової величини ε_t.

У моделі

$$Y = f\ \ X_t\ \ + \varepsilon_t \tag{4.1}$$

випадкова складова ε_t є неспостережуваною величиною. Після оцінювання параметрів моделі можна розрахувати різниці між фактичними і теоретичними значеннями результативної ознаки Y і визначити оцінки випадкової складової $e_t = Y_t - Y_t$. Оскільки вони не є реальними випадковими залишками, їх можна вважати деякою вибірковою реалізацією невідомого залишку заданого рівняння, тобто ε_t.

Під час зміни специфікації моделі та додавання нових спостережень вибіркові оцінки залишків ε_t можуть мінятися. Тому до завдання регресійного аналізу входить не тільки побудова самої моделі, але й дослідження випадкових відхилень ε_t, тобто залишкових величин.

Метод найменших квадратів одержує оцінки параметрів моделі на основі мінімізації суми квадратів залишків. Тому дуже важливо досліджувати поведінку залишкових величин регресії ε_t. У темі 2 були наведені припущення щодо поведінки залишків ε_t: залишки є випадковими величинами і їхнє середнє значення дорівнює 0; вони мають однакову (постійну) дисперсію, не корелюють між собою і з регресорами, підкоряються нормальному розподілу. Це і є передумови використання МНК і побудови класичної економетричної моделі.

Виконання усіх передумов дає змогу отримати *незміщені, обґрунтовані й ефективні* МНК-оцінки параметрів. Ці властивості оцінок мають надзвичайно важливе практичне значення у використанні результатів регресії й кореляції.

Сутність узагальненої регресійної моделі полягає у тому, що коваріації і дисперсії незалежних змінних, а, як наслідок, і залишків, можуть бути довільними (таким чином узагальнена модель множинної регресії відрізняється від класичної тільки видом коваріаційної матриці).

Розглянемо випадок порушення передумов класичної регресійної моделі, що стосується сталості дисперсії кожної випадкової величини e_t.

Якщо дисперсія залишків стала для кожного спостереження, то це явище називається ***гомоскедастичністю*** (*homoscedasticity*):

$$M\left(\varepsilon_i^2\right) = D\left(\varepsilon_i\right) = \sigma_\varepsilon^2 = const,\ \ i = \overline{1,n}. \tag{4.2}$$

Коваріаційна матриця залишків має вигляд:

$$M\left(\varepsilon^T\varepsilon\right) = \begin{pmatrix} \sigma^2 & 0 & \ldots & 0 \\ 0 & \sigma^2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \sigma^2 \end{pmatrix} = \sigma_\varepsilon^2 E.$$

Якщо це припущення не задовольняється в якомусь окремому випадку, то наявна ***неоднорідність спостережень***, тобто гетероскедастичність (помилки e_i некорельовані, але мають несталу дисперсію).

Якщо дисперсія залишків змінюється для кожного спостереження або групи спостережень, то це явище називається ***гетероскедастичністю*** (*heteroscedasticity*):

$$M\left(\varepsilon_i^2\right) = D\left(\varepsilon_i\right) = \sigma_\varepsilon^2 \neq const,\ \ i = \overline{1,n}. \tag{4.3}$$

Коваріаційна матриця залишків для узагальненої економетричної моделі має вигляд:

$$M\left(\varepsilon^T\varepsilon\right) = \Omega = \sigma_\varepsilon^2 S.$$

Причинами гетероскедастичності, як правило, є:

1. Неоднорідність досліджуваних об'єктів. Наприклад, під час аналізу залежності попиту від доходу споживача з'ясовується, що чим більше дохід, тим більше індивідуальне значення попиту коливається щодо очікуваного значення.

2. Характер спостережень. Найчастіше гетероскедастичність притаманна моделям із вихідними даними з просторової вибірки, аніж із даними, поданими часовими рядами.

3. Неправильна специфікація моделі.

Розглянемо питання про доцільність припущення гомоскедастичності залишків моделі і про те, що відбувається, якщо це припущення не задовольняється.

Насамперед, варто зауважити, що сутність припущення про гомоскедастичність полягає в тому, що варіація кожної випадкової складової e_i навколо її математичного сподівання не залежить від значення факторів x:

$$\sigma_{e_i}^2 \neq f\left(x_1, x_2, \dots, x_m\right).$$

Форма гетероскедастичності залежить від знаків і значень коефіцієнтів у залежності:

$$\sigma_{e_i}^2 = f\left(x_1, x_2, \dots, x_m\right).$$

Оскільки e_i – неспостережувана випадкова величина, то справжня форма гетероскедастичності невідома.

Наслідки порушення припущення про гомоскедастичність:

1) отримані МНК-оцінки параметрів регресії будуть незміщеними, обґрунтованими, але неефективними (не будуть мати найменшої дисперсії порівняно з іншими оцінками даного параметра, вони не будуть, навіть, асимптотично ефективними);

2) оскільки зростає дисперсія залишків, то дисперсії оцінок параметрів також будуть зростати. Це призводить до зростання знайдених середньоквадратичних відхилень параметрів, тоді як t-статистики для цих параметрів будуть занижені, а отже, неможливо оцінити справжню значущість параметрів регресії;

3) через зростання дисперсії залишків суттєво збільшуються довірчі інтервали для прогнозних значень $Y_{\text{пр}}$, що знижує цінність отриманих результатів.

Варто зазначити, що якщо незважаючи на гетероскедастичність будуть використовуватися звичайні процедури перевірки гіпотез (t-критерій, F-критерій), то висновки можуть бути неправильними. Гетероскедастичність є суттєвою проблемою, а тому потрібно вміти з'ясовувати її наявність.

Як і в разі мультиколінеарності, єдиних правил виявлення гетероскедастичності немає, а є різноманітні тести (критерії), а саме:

1) непараметричний тест Гольдфельда – Квандта (графічний метод);

2) параметричний тест Гольдфельда – Квандта;

3) критерій μ;

4) тест Глейзера;

5) тест рангової кореляції Спірмана;

6) тест Уайта;

7) тест Парка та ін.

Слід зауважити, що інколи в ході проведення економетричних досліджень гетероскедастичність вгадується інтуїтивно або висувається як абсолютне припущення. Наприклад, вивчаючи витрати підприємства, можна помітити, що дисперсія залишків зростає відповідно до зростання доходу. Отже, перший крок до виявлення гетероскедастичності – глибокий аналіз змісту досліджуваної проблеми.

Під час проведення аналізу залишків моделі перевіряються такі гіпотези:

основна гіпотеза H_0, яка припускає сталість дисперсій випадкових помилок моделі регресії, тобто присутність у моделі гомоскедастичності;

альтернативна гіпотеза H_1, яка передбачає мінливість дисперсії випадкових помилок у різних спостереженнях, тобто присутність у моделі гетероскедастичності.

Розглянемо деякі з методів визначення гетероскедастичності.

У 1965 р. Стефан Гольдфельд (Stephen Goldfeld) і Річард Квандт (Richard Quandt) запропонували два тести для перевірки наявності гетероскедастичності: параметричний і непараметричний [42]. У даний час у літературних джерелах під тестом Гольдфельда – Квандта мають на увазі параметричний.

Непараметричний тест Гольдфельда – Квандта (*Goldfeld – Quandtnonparametrictest*) не спирається на припущення про нормальний розподіл залишків і базується на графічному зображенні і підрахунку кількості піків значень квадратів залишків регресійної моделі, побудованої на основі припущення про відсутність гетероскедастичності, після впорядкування (ранжування) спостережень за X_j. Якщо для всіх значень змінної X_j залишки розподіляються приблизно однаково, то дисперсія їх однорідна і гетероскедастичність відсутня, як на рис. 4.1. Якщо вона змінюється, то гетероскедастичність присутня (рис. 4.2).

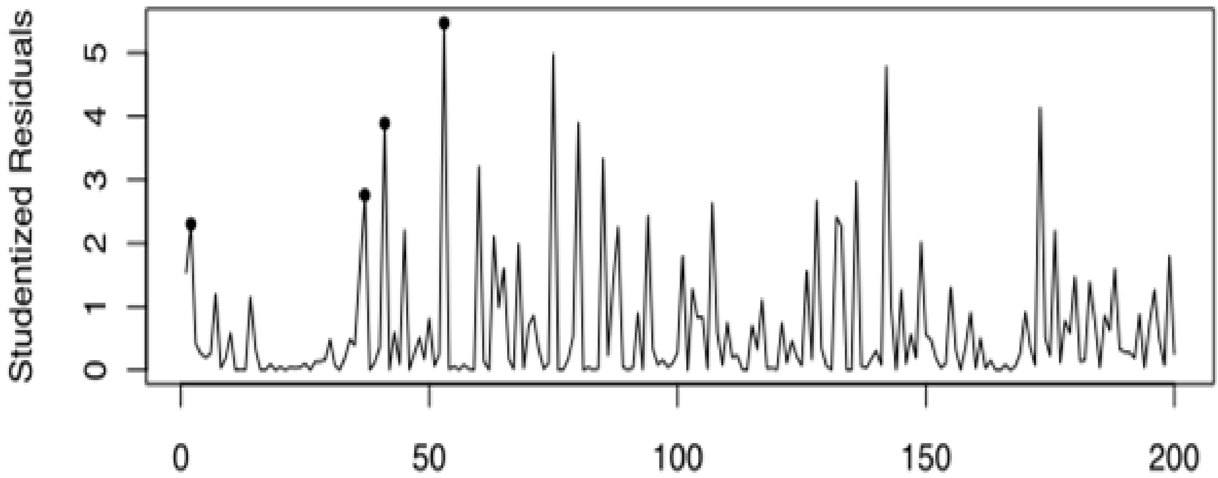


Рис. 4.1. **Непараметричний тест Гольдфельда – Квандта виявив ситуацію, схожу на гомоскедастичність**

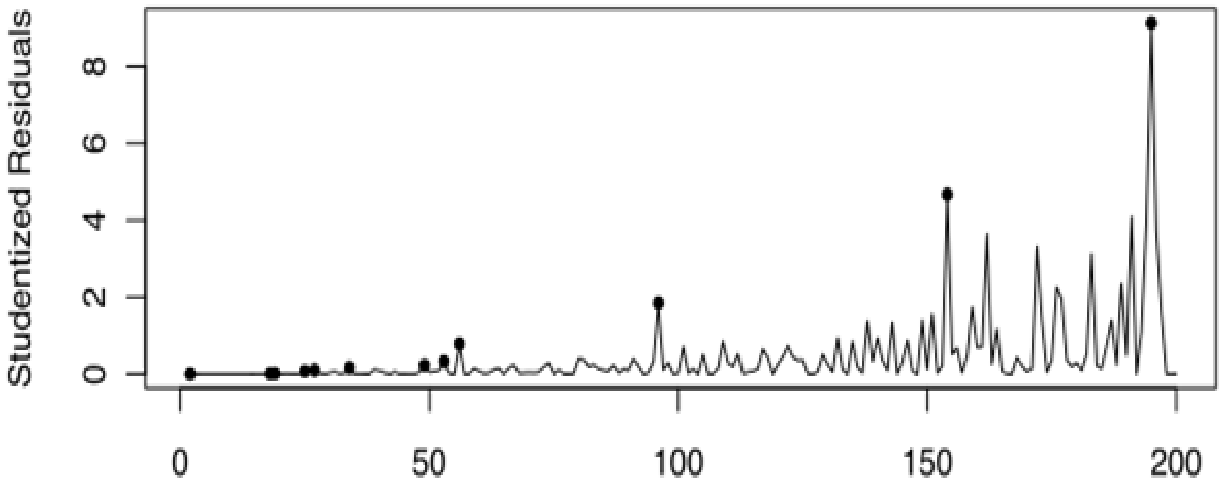


Рис. 4.2. **Непараметричний тест Гольдфельда – Квандта виявив ситуацію, схожу на гетероскедастичність**

Варто зазначити, що цей тест не цілком надійний для перевірки на гетероскедастичність. Однак він дуже простий і часто використовується для попереднього оцінювання наявності гетероскедастичності множини спостережень.

Параметричний тест Гольдфельда – Квандта (*Goldfeld – Quandtparametrictest*) – це статистичний тест, що дозволяє перевірити гетероскедастичність випадкових помилок регресійної моделі, застосовується до невеликих вибірок і припускає нормальний розподіл і незалежність випадкових величин e_i [21]. Він припускає впорядкованість за однією пояснювальною зміною, видалення точок даних у центрі і порівняння середніх відхилень із лівого і з правого боку, як це показано на рис. 4.3.

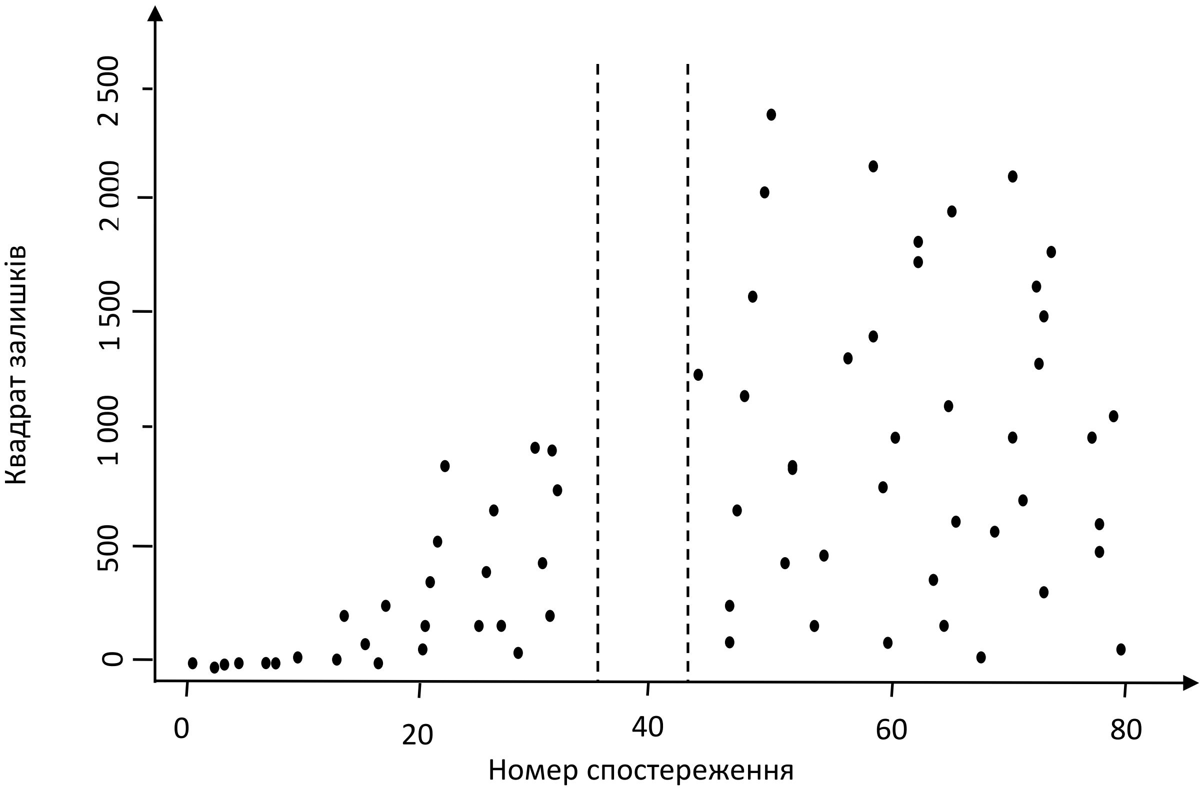


Рис. 4.3. **Результат роботи параметричного тесту Гольдфельда – Квандта**

Розглянемо алгоритм даного тесту:

1 крок. Спостереження (вихідні дані) впорядкувати відповідно до величини елементів вектора X_j, який може спричинити зміну дисперсії залишків.

2 крок. Відкинути c спостережень, які розміщені всередині векторів вихідних даних, де $c = \frac{4n}{15}$, n – кількість елементів вектора X_j.

3 крок. Побудувати дві моделі на основі звичайного МНК за двома створеними сукупностями спостережень обсягом $\frac{n-c}{2}$ за умови, що $\frac{n-c}{2} \geq m$, де m – кількість змінних.

4 крок. Знайти суму квадратів залишків S_1 і S_2 за першою і другою моделями:

$$S_1 = \sum_{i=1}^{n_1} e_{i1}^2, \qquad S_2 = \sum_{i=1}^{n_2} e_{i2}^2,$$

де e_{i1} і e_{i2} — залишки відповідно за першою і другою моделями.

5 крок. Розрахувати критерій:

$$F^* = \frac{S_1}{S_2}, \text{ якщо } S_1 > S_2 \text{ або } F^* = \frac{S_2}{S_1}, \text{ якщо } S_2 > S_1,$$

який у разі виконання гіпотези про гомоскедастичність відповідатиме F-розподілу з $\frac{n-c-2m}{2}$ і $\frac{n-c-2m}{2}$ ступенями свобод.

6 крок. Значення критерію F^* порівняти з табличним значенням F-критерію за умови вибраного рівня значущості α і відповідних ступеней свободи; якщо $F^* \leq F_{\text{табл}}$, то гетероскедастичність відсутня. В іншому випадку – чим більше значення F^*, тим більша гетероскедастичність залишків.

Зауваження. Обсяги підмножин не обов'язково мають бути однакового розміру (n_1 і n_2). Тоді, якщо $S_1 > S_2$, критерій буде розраховуватися за формулою:

$$F^* = \frac{S_1}{n_1} \cdot \frac{n_2}{S_2},$$

або, якщо $S_2 > S_1$:

$$F^* = \frac{S_2}{n_2} \cdot \frac{n_1}{S_1}.$$

Ступені свободи для F-критерію будуть відповідно $n_1 - m$; $n_2 - m$.

Параметричний тест Голдфельда – Квандта пропонує простий та інтуїтивно зрозумілий тест для діагностики гетероскедастичності для однофакторних і багатофакторних регресійних моделей. Однак існують деякі недоліки за умов певних специфікацій моделей або порівняно з іншими тестами. Так, цей тест вимагає, щоб дані були впорядковані за відомою пояснювальною змінною. Але якщо структура помилки залежить від невідомої змінної, або неспостережуваної, то цей тест не дає рекомендацій. Крім того, дисперсія помилки повинна бути монотонною функцією зазначеної пояснювальної змінної. Наприклад, коли зіштовхуються з квадратичною функцією, що відображає вплив регресора на дисперсії помилки, тест Голдфельда – Квандта може неправильно прийняти нульову гіпотезу про гомоскедастичність помилок.

Герберт Глейзер (Herbert Glejser) у своїй роботі у 1969 р. із викладенням тесту Глейзера проводить експеримент на невеликій вибірці даних для перевірки потужності та чутливості параметричного тесту Голдфельда – Квандта. Його результати показують обмежений успіх для тесту Голдфельда – Квандта тільки у випадках "чистої гетероскедастичності", коли дисперсії можуть бути описані як функції тільки основної пояснюючої змінної.

Приклад 4.1. Необхідно перевірити існування гетероскедастичності у моделі, побудованій на основі вихідних даних, наведених у табл. 4.1.

Таблиця 4.1

Вихідні дані

№ п/п	Питома вага покупних виробів (X_1)	Коефіцієнт змінності устаткування (X_2)	Фондоозброєність праці (X_3)	Продуктивність праці (Y)
1	2	3	4	5
1	0,4	1,37	6,4	9,26
2	0,26	1,49	7,8	9,38
3	0,4	1,44	9,76	12,11
4	0,5	1,42	7,9	10,81
5	0,4	1,35	5,35	9,35
6	0,19	1,39	9,9	9,87
7	0,25	1,16	4,5	8,17
8	0,44	1,27	4,88	9,12

Закінчення табл. 4.1

1	2	3	4	5
9	0,17	1,16	3,46	5,88
10	0,39	1,25	3,6	6,3
11	0,33	1,13	3,56	6,22
12	0,25	1,1	5,65	5,49
13	0,32	1,15	4,28	6,5
14	0,02	1,23	8,85	6,61
15	0,06	1,39	8,52	4,32
16	0,15	1,38	7,19	7,37
17	0,08	1,35	4,82	7,02
18	0,2	1,42	5,46	8,25
19	0,2	1,37	6,2	8,15
20	0,3	1,41	4,25	8,72
21	0,24	1,35	5,38	6,64
22	0,1	1,48	5,88	8,1
23	0,11	1,24	9,27	5,52
24	0,47	1,4	4,36	9,37
25	0,53	1,45	10,31	13,17

Для перевірки гіпотези про відсутність гетероскедастичності застосуємо параметричний тест Гольдфельда – Квандта.

1. Впорядкуємо значення всіх змінних за зростанням значення змінної X3. Результати впорядкування наведені у табл. 4.2.

2. Відкинемо c значень, які містяться всередині впорядкованого ряду:

$$c = \frac{4n}{15} = \frac{4 \cdot 25}{15} \approx 7.$$

У результаті отримано 2 сукупності спостережень з однаковим обсягом:

$$n_1 = \frac{n - c}{2} = \frac{25 - 7}{2} = 9, \qquad n_2 = \frac{n - c}{2} = \frac{25 - 7}{2} = 9.$$

3. Побудуємо дві економетричні моделі за двома щойно отриманими вибірками.

$$Y_I = -5{,}0653 + 4{,}7464X_1 + 4{,}3092X_2 + 1{,}3585X_3,$$
$$Y_{II} = -7{,}8453 + 11{,}6618X_1 + 5{,}7244X_2 + 0{,}6632X_3.$$

4. Визначимо залишки за цими двома моделями:

$$e_I = Y_I - Y_I, \qquad e_{II} = Y_{II} - Y_{II}.$$

Залишки і квадрати залишків наведено в табл. 4.2.

Таблиця 4.2

Впорядковані дані і результати розрахунків

№	X_1	X_2	X_3	Y	Група	$\hat{Y}$	e	e^2
9	0,17	1,16	3,46	5,88	1	5,4405	0,4395	0,1932
11	0,33	1,13	3,56	6,22		6,2065	0,0135	0,0002
10	0,39	1,25	3,6	6,3		7,0627	-0,7627	0,5817
20	0,3	1,41	4,25	8,72		8,2080	0,5120	0,2621
13	0,32	1,15	4,28	6,5		7,2233	-0,7233	0,5232
24	0,47	1,4	4,36	9,37		9,1212	0,2488	0,0619
7	0,25	1,16	4,5	8,17		7,2330	0,9370	0,8779
17	0,08	1,35	4,82	7,02		7,6796	-0,6596	0,4351
8	0,44	1,27	4,88	9,12		9,1251	-0,0051	0,0000
5	0,4	1,35	5,35	9,35		Сума		**2,9353**
21	0,24	1,35	5,38	6,64				
18	0,2	1,42	5,46	8,25				
12	0,25	1,1	5,65	5,49				
22	0,1	1,48	5,88	8,1				
19	0,2	1,37	6,2	8,15				
1	0,4	1,37	6,4	9,26				
16	0,15	1,38	7,19	7,37	2	6,5720	0,7980	0,6368
2	0,26	1,49	7,8	9,38		8,8891	0,4909	0,2410
4	0,5	1,42	7,9	10,81		11,3535	-0,5435	0,2954
15	0,06	1,39	8,52	4,32		6,4618	-2,1418	4,5872
14	0,02	1,23	8,85	6,61		5,2983	1,3117	1,7207
23	0,11	1,24	9,27	5,52		6,6836	-1,1636	1,3540
3	0,4	1,44	9,76	12,11		11,5354	0,5746	0,3302
6	0,19	1,39	9,9	9,87		8,8930	0,9770	0,9545
25	0,53	1,45	10,31	13,17		13,4734	-0,3034	0,0921
						Сума		**10,2119**

5. Обчислимо залишкові дисперсії і знайдемо їх співвідношення:

$$F^* = \frac{S_2}{S_1} = \frac{10{,}2119}{2{,}9353} = 3{,}497.$$

6. Порівняємо отримане значення F^* з критичним значенням F-критерію, за умови $n_1 - m = 9 - 3 = 6$ і $n_2 - m = 6$ ступенях свободи і рівні довіри 0,95 $F_{\text{табл}} = 4{,}28$. Оскільки $F^* < F_{\text{табл}}$, то гетероскедастичність відсутня.

Критерій μ – цестатистичний тест, що дозволяє перевірити гетероскедастичність випадкових помилок регресійної моделі тоді, коли вихідна сукупність спостережень досить велика.

Розглянемо відповідний алгоритм.

1 крок. Вихідні дані залежної змінної Y поділяються на k груп $r = 1, k$ відповідно до зміни рівня величини Y.

2 крок. За кожною групою даних обчислюється сума квадратів відхилень:

$$S_r = \sum_{i=1}^{n_r} (y_{ir} - y_r)^2,$$

де n_r – кількість спостережень r-ї групи;

n – загальна сукупність спостережень.

3 крок. Визначається сума квадратів відхилень у цілому за всією сукупністю спостережень:

$$\sum_{r=1}^{k} S_r = \sum_{r=1}^{k} \sum_{i=1}^{n_r} (y_{ir} - y_r)^2.$$

4 крок. Обчислюється параметр α:

$$\alpha = \prod_{r=1}^{k} \left(\frac{S_r}{n_r} \right)^{n_r/2} \Bigg/ \left(\frac{\sum_{r=1}^{k} S_r}{n} \right)^{n/2}.$$

5 крок. Обчислюється критерій:

$$\mu = -2\ln\alpha,$$

який наближено відповідатиме розподілу χ^2 з ступенями свободи $k-1$, коли дисперсія всіх спостережень однорідна. Тобто якщо значення μ не менше за табличне значення χ^2 за вибраного рівня довіри з ступенями свободи $k-1$, то спостерігається гетероскедастичність.

Приклад 4.2. Для даних із прикладу 4.1 перевіримо наявність гетероскедастичності згідно з критерієм μ.

Тест застосуємо до залежної змінної Y, яка поділена у табл. 4.3 за значеннями на 3 групи. Для визначення наявності гетероскедастичності використаємо алгоритм розрахунку критерію μ.

1. Дані залежної змінної були поділені на 3 групи, у кожній з яких міститься вісім або дев'ять спостережень, як наведено в табл. 4.3.

Таблиця 4.3

Вихідні дані

№ п/п	Група 1	Група 2	Група 3
1	4,32	6,64	9,12
2	5,49	7,02	9,26
3	5,52	7,37	9,35
4	5,88	8,1	9,37
5	6,22	8,15	9,38
6	6,3	8,17	9,87
7	6,5	8,25	10,81
8	6,61	8,72	12,11
9	–	–	13,17

2. Обчислимо суму квадратів відхилень індивідуальних значень кожної групи від свого середнього значення:

$$Y_{\mathrm{I}} = 5{,}855;\ Y_{\mathrm{II}} = 7{,}8025;\ \ Y_{\mathrm{III}} = 10{,}2711.$$

$$S_1 = \sum_{i=1}^{8} \left(Y_{i\mathrm{I}} - Y_{\mathrm{I}}\right)^2 = 38{,}2;$$

$$S_2 = \sum_{i=1}^{8} \left(Y_{i\text{II}} - \bar{Y}_{\text{II}}\right)^2 = 64{,}4162;$$

$$S_3 = \sum_{i=1}^{9} \left(Y_{i\text{III}} - \bar{Y}_{\text{III}}\right)^2 = 17{,}0383.$$

3. Знайдемо суму квадратів відхилень за всіма трьома групами:

$$\sum_{r=1}^{3} S_r = S_1 + S_2 + S_3 = 38{,}2 + 64{,}4162 + 17{,}0383 = 119{,}655.$$

4. Обчислимо параметр α:

$$\alpha = \frac{\left[\left(\frac{38{,}2}{8}\right)^4 \cdot \left(\frac{64{,}4162}{8}\right)^4 \cdot \left(\frac{17{,}0383}{9}\right)^{4{,}5}\right]}{\left(\frac{119{,}655}{25}\right)^{12{,}5}} = 0{,}072145.$$

5. Знайдемо критерій μ:

$$\mu = -2\ln\left(0{,}072145\right) = 5{,}258.$$

Цей критерій наближено задовольняє розподіл χ^2 з $k-1 = 2$ ступенями свободи. Порівняємо значення критерію μ з табличним значенням критерію χ^2 з $k-1 = 2$ ступенями свободи за рівня довіри 0,99, $\chi^2_{\text{табл}} = 9{,}21$. Оскільки $\mu < \chi^2_{\text{табл}}$, то дисперсія залишків буде залишатись сталою, тобто для даних табл. 4.3 спостерігається гомоскедастичність.

Тест рангової кореляції Спірмена – непараметричний статистичний тест, що дозволяє перевірити гетероскедастичність випадкових помилок регресійної моделі. Особливість тесту полягає в тому, що не конкретизується форма можливої залежності дисперсії випадкових помилок моделі від тієї чи іншої змінної.

За допомогою звичайного МНК оцінюється вихідна регресійна модель і визначаються залишки регресії $e_t = y_t - y_t$:

$$y_t = \quad {}_i\, a_i\, x_{it} + \varepsilon_t.$$

Далі ранжуються залишки e_t і змінна x_{jt}, від якої передбачається залежність дисперсії випадкових помилок, і визначається коефіцієнт рангової кореляції Спірмена:

$$\rho = 1 - \frac{6 \quad d_t^2}{n \;\; n^2 - 1},$$

де d_t– різниця рангів змінних e_t і x_{jt}.

Доведено, що за умови справедливості нульової гіпотези (відсутність гетероскедастичності, тобто в даному випадку – рівність нулю істинного значення коефіцієнта рангової кореляції Спірмена ρ) статистика $\rho \quad \overline{n-1}$ асимптотично (тобто за умови досить великого n) має стандартний нормальний розподіл $N \;\; 0{,}1$. Отже, якщо значення цієї статистики більше критичного значення цього розподілу (за даного рівня значущості), то гетероскедастичність визнається значущою. В іншому випадку гетероскедастичність незначуща (це не виключає можливої залежності дисперсії помилок від інших змінних, тому потрібно провести тест для всіх "підозрілих" змінних).

Тест Глейзера (*Glejser test*) – це статистичний тест, що дозволяє оцінити наявність певного виду гетероскедастичності випадкових помилок регресійній моделі[41]. Тест заснований на такій моделі можливої залежності стандартного відхилення випадкової помилки моделі σ_t від деякого фактора x_{jt}:

$$\sigma_t = a_0 + a_1 x_{jt}^{\gamma} + u_t.$$

Нульова гіпотеза полягає в рівності коефіцієнта a_1 нулю (відсутність гетероскедастичності даного виду). Якщо у тесті відкидається нульова гіпотеза, то гетероскедастичність даного виду визнається статистично значущою. Якщо нульова гіпотеза не відкидається, то швидше за все гетероскедастичності даного виду немає у моделі (проте, це не виключає можливості гетероскедастичності іншого виду).

Розглянемо алгоритм тесту Глейзера.

1 крок. За допомогою звичайного МНК оцінюється вихідна регресійна модель і визначаються залишки регресії $e_t = y_t - y_t$:

$$y_t = \sum_{i} a_i x_{it} + \varepsilon_{t.}$$

2 крок. Побудова допоміжних регресійних функцій, що характеризують залежність величини залишків за модулем від пояснюючої змінної x_{jt}, яка може зумовити зміну дисперсії залишків, виду:

$$e_t \ = a_0 + a_1 x_{jt}^{\gamma} + u_t,$$

де γ – величина, яка може набувати значення (-1; -0,5;.0,5; 1; 2).

3 крок. Для кожного значення γ перевіряється статистична значущість коефіцієнта a_1 за допомогою критерію Стьюдента або еквівалентного йому в даному випадку F-тесту на значущість допоміжної регресії в цілому. Якщо для деяких γ коефіцієнт a_1 визнається значущим (тестова статистика більше критичного значення), то гетероскедастичність даного виду визнається значущою і вибирається модель із тим значенням γ, для якого коефіцієнт a_1 найбільш значущий (з найбільшим значенням тестової статистики).

Перевага цього методу полягає в можливості розрізняти випадок чистої і змішаної гетероскедастичності. Можливі чотири випадки, які наведено у табл. 4.4.

Таблиця 4.4

Визначення гетероскедастичності залишків за тестом Глейзера

Оцінка	a_1 – статистично значуща	a_1 – статистично незначуща
a_0 – статистично значуща	чиста і змішана гетероскедастичність	змішана гетероскедастичність
a_0 – статистично незначуща	чиста гетероскедастичність	гетероскедастичність відсутня

Нульова гіпотеза про відсутність гетероскедастичності полягає в рівності коефіцієнта a_1нулю. Відхилення цієї гіпотези означає наявність гетероскедастичності зазначеного виду, прийняття нульової гіпотези означає, що гетероскедастичність даного виду відсутня.

Розглянемо алгоритм тесту Парка.

1 крок. За допомогою звичайного МНК оцінюється вихідна регресійна модель:

$$y_t = \sum_i a_i\, x_{it} + \varepsilon_t$$

і визначаються залишки регресії $e_t = y_t - y_t$.

2 крок. Далі також за допомогою звичайного МНК оцінюється така допоміжна регресія:

$$ln(e_t^2) = a_0 + a_1 \ln\ x_{jt}\ .$$

3 крок. Перевіряється статистична значущість коефіцієнта a_1 за допомогою t-критерію Стьюдента або еквівалентного в даному випадку F-тесту на значущість допоміжної регресії в цілому. Якщо коефіцієнт визнається значущим, то випадкові помилки моделі визнаються гетероскедастичними, а в іншому випадку гетероскедастичність даного виду вважається незначною (що не виключає можливість наявності гетероскедастичності іншого виду і в цьому випадку слід використовувати також й інші тести).

Одним із найважливіших питань оцінювання параметрів моделі з гетероскедастичністю, тобто коли дисперсії залишків визначаються $M\ e'e\ = \sigma_e^2 S$, є визначення матриці S.

Оскільки явище гетероскедастичності пов'язане лише з тим, що змінюються дисперсії залишків, а коваріація між ними відсутня, то матриця S має бути діагональною, а саме:

$$S = \begin{matrix} \frac{1}{\lambda_1} & 0 & 0 & \dots & 0 \\ 0 & \frac{1}{\lambda_2} & 0 & \dots & 0 \\ 0 & 0 & \frac{1}{\lambda_3} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \frac{1}{\lambda_n} \end{matrix}\ .$$

Як було зазначено, найчастіше за наявності гетероскедастичності для певних вихідних даних одна (або кілька) пояснювальних змінних можуть різко змінюватись від одного спостереження до іншого, тоді як залежна змінна має такі самі коливання, як і для попередніх спостережень.

А це означає, що дисперсія залишків, яка змінюватиметься від одного спостереження до іншого (чи для групи спостережень), може бути пропорційною до величини пояснювальної змінної X (або до її квадрата), яка зумовлює гетероскедастичність, або пропорційною до квадрата залишків.

Звідси в матриці S значення λ_i можна обчислити, користуючись, наприклад, такими гіпотезами:

- $M\ \ e'e\ \ = \sigma_e^2 x_{ij}$, тобто дисперсія залишків пропорційна до зміни пояснювальної змінноїx_{ij};
- $M\ \ e'e\ \ = \sigma_e^2 x_{ij}^2$, тобто зміна дисперсії пропорційна до зміни квадрата пояснювальної змінної x_{ij}^2;
- $M\ \ e'e\ \ = \sigma_e^2\ \ e_t\ \ ^2$, тобто дисперсія залишків пропорційна до зміни квадрата залишків за модулем.

Для першої гіпотези: $\lambda_i = 1\ \ x_{ij}$.

Для другої гіпотези: $\lambda_i = 1\ \ x_{ij}^2$

Для третьої гіпотези: $\lambda_i = \ \ e_t\ \ ^2$, або $\lambda_i = \ \ a_0 - a_1 x_{ij}\ \ ^2$, або $\lambda_i = \ \ a_0 - a_1 x_{ij}^{-1}\ \ ^2$.

4.2. Узагальнений метод найменших квадратів (метод Ейткена)

Узагальнена регресійна модель має таку специфікацію [16]:

$$Y = X\beta + \varepsilon, \tag{4.4}$$

де $Y = \ \ Y_1, Y_2, \dots, Y_n\ \ ^{\mathrm{T}}$ – $(n \cdot 1)$ вектор-стовпець значень ендогенної змінної;

$X\ \ n \times k$ – детермінована матриця регресорів повного рангу;

$\beta = \ \ \beta_1, \beta_2, \dots, \beta_k\ \ ^{\mathrm{T}}$– $k \cdot 1$ вектор-стовпець параметрів моделі;

$\varepsilon = \ \ \varepsilon_1, \varepsilon_2, \dots, \varepsilon_n\ \ ^{\mathrm{T}}$– $(n \cdot 1)$вектор-стовпець випадкових збурювань.

Щодо випадкових збурювань регресії, то приймаються такі передумови:

1) $E(\varepsilon) = 0$;

2) $M(\varepsilon'\varepsilon) = C_{\varepsilon\varepsilon} = \Omega = \begin{pmatrix} \sigma_1^2 & \Omega_{12} & \dots & \Omega_{1n} \\ \Omega_{21} & \sigma_2^2 & \dots & \Omega_{2n} \\ \dots & \dots & \dots & \dots \\ \Omega_{n1} & \Omega_{n2} & \dots & \sigma_n^2 \end{pmatrix} = \sigma_\varepsilon^2 S$ – автоковаріаційна матриця вектора збурювань, де $\Omega_{ts} = \Omega_{st} = \operatorname{cov}(\varepsilon_t, \varepsilon_s) \neq 0$.

Оцінювання параметрів узагальненої регресійної моделі

1.Звичайний (однокроковий) метод найменших квадратів (1МНК)

До системи рівнянь спостережень (4.4) можна застосувати звичайний метод найменших квадратів (*Ordinary Least Squares*, *OLS*), тоді одержимо оцінку параметрів моделі через вибіркові дані:

$$\beta_{OLS} = AY = (X^T X)^{-1} X^T Y.$$

Дана оцінка є незміщеною:

$$E(\beta_{OLS}) = E(AY) = AE(Y) = AE(X\beta + \varepsilon) = \beta + AE(\varepsilon) = \beta.$$

Автоковаріаційна матриця вектора оцінок параметрів, відповідно до передумов узагальненої регресійної моделі, визначається так:

$$C_{\beta\beta OLS} = AC_{\varepsilon\varepsilon}A^T = A\Omega A^T = (X^T X)^{-1} X^T \Omega X (X^T X)^{-1}.$$

Визначимо основні кількісні характеристики вектора залишків регресії:

$$E(e) = ME(\varepsilon) = 0,$$

$$C_{ee} = M\Omega M.$$

За тих умов, що в загальному випадку:

$$E(e'e) = tr(C_{ee}) = tr(M\Omega M) = tr(M\Omega) \neq \sigma^2(n - m),$$

оцінка параметра $s = \sigma$ є *зміщеною,* що призводить до неадекватності оцінки автоковариаційної матриці вектора β_{OLS}:

$$C_{\beta\beta OLS} = s^2 (X^T X)^{-1}.$$

Таким чином, наслідки застосування звичайного МНК до оцінювання параметрів узагальненої регресійної моделі такі ж, як і під час порушення третьої і четвертої передумов Гаусса – Маркова (див. розділ 2).

2. Узагальнений метод найменших квадратів (УМНК)

Для подолання цих наслідків під час оцінювання параметрів узагальненої регресійної моделі з гетероскедастичністю залишків використовується узагальнений МНК (*GeneralizedLeastSquares, GLS*), застосування якого базується на такій теоремі [27].

Теорема Ейткена. У класі лінійних незміщених оцінок вектора параметрів β узагальненої регресійної моделі оцінка

$$\beta^{*} = A^{*}Y = (X^{T}\Omega^{-1}X)^{-1}X^{T}\Omega^{-1}Y$$

є ефективною.

Перевіримо незміщеність оцінки Ейткена:

$$E(\beta^{*}) = E(A^{*}Y) = A^{*}E(Y) = A^{*}X\beta + A^{*}E(\varepsilon) = (X^{T}\Omega^{-1}X)^{-1}X^{T}\Omega^{-1}X\beta + 0 = \beta.$$

Нехай автоковаріаційна матриця для узагальненої регресії задається так: $M(e'e) = \sigma_e^2 S$.

Завдання полягає в знаходженні оцінок елементів вектора β в моделі. Для цього використовується матриця S, за допомогою якої коригується вихідна інформація.

Оскільки S – додатно визначена матриця, то вона може бути зображена як добуток PP', де матриця P є невиродженою, тобто:

$$S = PP', \tag{4.5}$$

коли

$$P^{-1}SP^{-1'} = E, \tag{4.6}$$

і

$$P^{-1'}P^{-1} = S^{-1}. \tag{4.7}$$

Помноживши рівняння (4.4) ліворуч на матрицю P^{-1}, отримаємо:

$$P^{-1}Y = P^{-1}X\beta + P^{-1}e. \tag{4.8}$$

Позначимо $\quad Y^{*} = P^{-1}Y;\ \ X^{*} = P^{-1}X;\ \ e^{*} = P^{-1}e.$

Тоді модель матиме вигляд:

$$Y^{*} = X^{*}\beta + e^{*}. \qquad (4.9)$$

Використовуючи (4.6), неважко показати, що $M\ \ e^{*\prime}e^{*}\ \ = \sigma^{2}E$, тобто модель (4.9) задовольняє умови (4.2), коли параметри моделі можна оцінити на основі звичайного МНК.

Звідси

$$\beta_{GLS} = \ \ X^{*T}X^{*}\ \ ^{-1}X^{*T}Y^{*} = \ \ X^{T}S^{-1}X\ \ ^{-1}X^{T}S^{-1}Y. \qquad (4.10)$$

Ця оцінка є незміщеною лінійною оцінкою вектора β, який має найменшу дисперсію і матрицю коваріацій:

$$var\ \ \beta\ \ = \sigma_{e}^{2}\ \ X^{*T}X^{*}\ \ ^{-1} = \sigma_{e}^{2}\ \ X^{T}S^{-1}X\ \ ^{-1}. \qquad (4.11)$$

Незміщену оцінку для дисперсії σ_{e}^{2} можна дістати так:

$$\frac{1}{n-m}\ \ Y^{*} - X^{*}\beta\ \ ^{T}\ \ Y^{*} - X^{*}\beta\ \ = \frac{1}{n-m}\ \ Y - X\beta\ \ ^{T}S^{-1}\ \ Y - X\beta\ \ =$$
$$= \frac{1}{n-m}e^{T}S^{-1}e.$$

Оцінка параметрів β, яку знайдено за допомогою (4.10), є оцінкою узагальненого методу найменших квадратів (методу Ейткена).

У разі заданої матриці S оцінку параметрів моделі можна обчислити згідно з (4.10), а стандартну помилку — згідно із (4.11). Тому можна сконструювати звичайні критерії значущості і довірчі інтервали для параметрів β_{j}.Визначивши залишки $e = Y - X\beta$ і помноживши ліворуч на матрицю P^{-1}, отримаємо:

$$P^{-1}e = P^{-1}Y - P^{-1}X\beta,$$

або $e^{*} = Y^{*} - X^{*}\beta$.

Звідси $Y^{*} = X^{*}\beta + e^{*}$.

Тоді $Y^{*T}Y^{*} = \ \ X^{*}\beta + e^{*}\ \ ^{T}\ \ X^{*}\beta + e^{*}\ \ $.

Оскільки $\beta_{GLS} = (X^{*T}X^*)^{-1}X^{*T}Y^* = (X^TS^{-1}X)^{-1}X^TS^{-1}Y$,

то $Y^TS^{-1}Y = \beta^TX^TS^{-1}Y + eS^{-1}e$.

Отже, було розділено загальну суму квадратів для (4.9) на суму квадратів регресії і залишкову. Таким чином, для побудови оптимальної оцінки в моделі (4.7) треба мінімізувати "узагальнену" суму квадратів відхилень. Згідно з цими даними дисперсійний аналіз буде виконано для перетворених вихідних даних. Крім того, коли незалежна змінна Y^* виміряна відносно початку відліку, а не у формі відхилення від середньої, то необхідно визначити її середнє значення і скористатись ним для корекції загальної суми квадратів і суми квадратів регресії.

Модель УМНК іноді специфікується у вигляді:

$$Y = X\beta + e,$$
$$M(e) = 0, M(e'e) = V,$$

де $V = \sigma_e^2 S$– відома симетрична додатно визначена матриця.

Тоді вираз для оцінювання параметрів згідно з методом Ейткена запишеться так:

$$\beta_{GLS} = (X^TV^{-1}X)^{-1}X^TV^{-1}Y,$$

а для її коваріаційної матриці так:

$$var(\beta) = \sigma_e^2 (X^TV^{-1}X)^{-1}.$$

Покажемо ефективність оцінок Ейткена β_{GLS} порівняно з оцінками параметрів β, отриманими звичайним методом найменших квадратів.

Розглянемо вектор:

$$b = \beta - \beta_{GLS} = AY - A^*Y = (A - A^*)Y.$$

Необхідно визначити автоковаріаційну матрицю цього вектора:

$$C_{bb} = (A - A^*)C_{YY}(A - A^*)^T = A\Omega A^T - A^*\Omega A^T - A\Omega A^{*T} + A^*\Omega A^{*T}.$$

З урахуванням того, що:

$$A^*\Omega A^T = A\Omega A^{*T} = A^*\Omega A^{*T},$$

$$C_{bb} = A\Omega A^T - A^*\Omega A^T = (X^TX)^{-1}X^T\Omega X(X^TX)^{-1} - (X^T\Omega^{-1}X)^{-1}$$
$$= C_{\beta\beta} - C_{\beta_{GLS}\beta_{GLS}}.$$

Діагональні елементи матриці C_{bb} є дисперсії, і, отже,

$$(C_{bb})_{dg} = (C_{\beta\beta} - C_{\beta_{GLS}\beta_{GLS}})_{dg} \geq 0,$$

таким чином:

$$var(\beta) \geq var(\beta_{GLS}).$$

Через незміщеність розглянутих оцінок спостерігається ефективність оцінок Ейткена порівняно з оцінками звичайного методу найменших квадратів.

Таким чином, за наявності гетероскедастичності для оцінювання параметрів моделі доцільно застосувати УМНК (метод Ейткена), оператор оцінювання якого має вигляд:

$$\beta_{GLS} = (X^TS^{-1}X)^{-1}X^TS^{-1}Y.$$

Вектор β у такому разі містить незміщену лінійну оцінку параметрів моделі, яка має найменшу дисперсію і матрицю коваріацій:

$$var(\beta) = \sigma_e^2(X^TS^{-1}X)^{-1}.$$

Контрольні запитання для самодіагностики

1. Дайте визначення поняття гомоскедастичності.
2. Дайте визначення поняття гетероскедастичності.
3. Яким чином гетероскедастичність впливає на властивості оцінок параметрів моделі?
4. Назвіть причини наявності гетероскедастичності у моделі.
5. У чому полягає різниця між класичною та узагальненою регресійними моделями?

6. Опишіть наслідки гетероскедастичності.

7. Поясніть різницю між параметричним і непараметричним тестами Гольдфельда – Квандта.

8. Назвіть основні кроки μ-критерію.

9. У чому полягає особливість тесту рангової кореляції Спірмена?

10. Назвіть основні кроки тесту Парка.

11. Наведіть переваги та недоліки методів оцінки наявності гетероскедастичності.

11. В якому випадку використовується узагальнений метод найменших квадратів?

12. Яким чином застосовується УМНК?

13. Які властивості мають оцінки параметрів регресії, отримані за допомогою методу Ейткена (УМНК)?

14. Доведіть ефективність УМНК-оцінок параметрів порівняно з МНК-оцінками.

Тести

1. МНК-оцінки параметрів узагальненої регресійної моделі:

а) зміщені;

б) незміщені;

в) стохастичні.

2. УМНК-оцінки параметрів узагальненої регресійної моделі:

а) зміщені;

б) незміщені;

в) стохастичні.

3. Основні джерела гетероскедастичності:

а) помилки специфікації;

б) помилки вимірювання;

в) характер спостережень.

4. У разі гетероскедастичності помилки моделі мають:

а) постійну дисперсію;

б) біноміальний розподіл;

в) експоненційний розподіл;

г) непостійну дисперсію.

5. У разі гомоскедастичності залишки моделі мають:

а) постійну дисперсію;

б) біноміальний розподіл;

в) експоненційний розподіл;

г) непостійну дисперсію.

6. У разі гетероскедастичності відхилень оцінки параметрів моделі, отримані за звичайним МНК, будуть:

а) незміщеними, обґрунтованими й ефективними;

б) зміщеними, обґрунтованими й ефективними;

в) незміщеними, необґрунтованими й ефективними;

г) незміщеними, обґрунтованими і неефективними;

д) зміщеними, необґрунтованими і неефективними.

7. Для знаходження УМНК-оцінок необхідно:

а) знати математичне очікування залишків;

б) знати коваріаційну матрицю вектора залишків;

в) довести непостійність дисперсії залишків;

г) знати закон розподілу залишків.

8. Для перевірки моделі на гетероскедастичність використовують:

а) метод Феррара – Глобера;

б) критерій Стьюдента;

в) міру Неймана – Гольдштейна;

г) тест Голдфельда – Квандта.

9. Чиста гетероскедастичність визначається:

а) однією змінною;

б) декількома змінними;

в) законом розподілу залишків.

10. Перевірити гіпотезу про змішану (чисту) гетероскедастичність можна, використовуючи:

а) метод Феррара – Глобера;

б) критерій μ;

в) параметричний тест Голдфельда – Квандта;

г) непараметричний тест Голдфельда – Квандта;

д) тест Глейзера.

Практичні завдання

1. За даними діяльності підприємства, наведеними у табл. 4.5, необхідно побудувати економетричну модель рентабельності і перевірити

гіпотезу про гомоскедастичність залишків на основі непараметричного тесту Гольдфельда – Квандта.

Таблиця 4.5

Вихідні дані

№ п/п	Продуктивність праці (X_1)	Фондовіддача (X_3)	Невиробничі витрати (X_4)	Рентабельність (Y)
1	9,26	1,45	17,72	13,26
2	9,38	1,3	18,39	10,16
3	12,11	1,37	26,46	13,72
4	10,81	1,65	22,37	12,85
5	9,35	1,91	28,13	10,63
6	9,87	1,68	17,55	9,12
7	8,17	1,94	21,92	25,83
8	9,12	1,89	19,52	23,39
9	5,88	1,94	23,99	14,68
10	6,3	2,06	21,76	10,05
11	6,22	1,96	25,68	13,99
12	5,49	1,02	18,13	9,68
13	6,5	1,85	25,74	10,03
14	6,61	0,88	21,21	9,13
15	4,32	0,62	22,97	5,37
16	7,37	1,09	16,38	9,86
17	7,02	1,6	13,21	12,62
18	8,25	1,53	14,48	5,02
19	8,15	1,4	13,38	21,18
20	8,72	2,22	13,69	25,17
21	6,64	1,32	16,66	19,4
22	8,1	1,48	15,06	21
23	5,52	0,68	20,09	6,57
24	9,37	2,3	15,98	14,19
25	13,17	1,37	18,27	15,81

2. На основі вихідних даних із табл. 4.5 необхідно перевірити існування гетероскедастичності залишків у моделі рентабельності на основі параметричного тесту Гольдфельда – Квандта.

3. На основі вихідних даних із табл. 4.5 перевірити існування гетероскедастичності залишків у моделі рентабельності на основі тесту Парка.

4. На основі вихідних даних із табл. 4.5 перевірити існування гетероскедастичності залишків у моделі рентабельності на основі параметричного тесту Глейзера.

5. На основі результатів дослідження гетероскедастичності залишків у моделі рентабельності з завдань 1 – 4 визначити матрицю S і оцінити параметри цієї моделі за допомогою узагальненого методу найменших квадратів.

Ключові слова

Гомоскедастичність. Гетероскедастичність. Дисперсія залишків. Матриця коваріацій параметрів. Непараметричний тест Гольдфельда – Квандта. Параметричний тест Гольдфельда – Квандта. Критерій μ. Тест Глейзера.Тест рангової кореляції Спірмена. Тест Парка. Узагальнений метод найменших квадратів (метод Ейткена).

Розділ 5. Побудова моделі з автокорельованими залишками

5.1. Автокореляція залишків. Методи перевірки автокореляції залишків.

5.2. Методи оцінювання параметрів моделі з автокорельованими залишками.

5.1. Автокореляція залишків. Методи перевірки автокореляції залишків

В економетричних дослідженнях надзвичайно важливо дотримування четвертої передумови регресійного аналізу, яка полягає в тому, що значення залишків ε_t розподілені незалежно один від одного. Однак часто виявляється, що припущення про те, що ε_t – випадкові незалежні (некорельовані) помилки з постійною дисперсією й нульовим математичним очікуванням, нездійсненні. Якщо вид функції обраний невдало, то не можна говорити, що відхилення від регресійної моделі є незалежними. У цьому випадку спостерігається концентрація додатних і від'ємних відхилень від регресії й можна сумніватися в їхньому випадковому характері. Якщо послідовні значення ε_t корелюють між собою, то має місце автокореляція помилок.

Автокореляція залишків – це наявність взаємозв'язку між послідовними елементами ряду залишків моделі.

Автокореляція залишків найчастіше спостерігається тоді, коли економетрична модель будується на основі часових рядів. Якщо існує кореляція між послідовними значеннями деякої пояснювальної змінної, то буде спостерігатись і кореляція послідовних значень залишків.

Наявність автокореляції може бути ***обумовлена*** багатьма ***причинами:***

1. *Помилки специфікації*:

1) якщо в моделі не врахований деякий істотний фактор, то його вплив відображається в залишках, унаслідок чого останні можуть виявитися автокорельованими;

2) якщо в моделі не враховано кілька несуттєвих факторів, взаємний вплив яких є істотним внаслідок збігу фаз і напрямів їхньої зміни;

3) якщо обрано неправильний тип моделі.

2. *Інерція* у зміні економічних показників. Чимало економічних показників (наприклад, інфляція, безробіття, ВНП тощо) мають певну циклічність, пов'язану з хвилеподібністю ділової активності. Дійсно, економічний підйом приводить до зростання зайнятості, скорочення інфляції, збільшення ВНП і т. д. Це зростання продовжується до тих пір, поки зміна кон'юнктури ринку та низки економічних характеристик не призведе до уповільнення зростання, потім зупинки і руху назад розглянутих показників. У будь-якому випадку ця трансформація відбувається не миттєво, а має певну інертність.

3. *Ефект павутини.* У багатьох виробничих та інших сферах економічні показники реагують на зміну економічних умов з запізненням (часовим лагом). Наприклад, пропозиція сільськогосподарської продукції реагує на зміну ціни з запізненням (рівним періоду дозрівання врожаю). Велика ціна сільськогосподарської продукції в минулому році викличе (швидше за все) її перевиробництво в поточному році, а отже, ціна на неї знизиться і т. д.

4. *Згладжування даних.* Найчастіше дані за деяким тривалим часовим періодом отримують усередненням даних за його складовими. Це може призвести до певного згладжування коливань, які були всередині розглянутого періоду, що, у свою чергу, може спричинити автокореляцію.

Якщо для оцінювання параметрів моделі з автокореляцією залишків застосувати МНК, то можливі такі *наслідки*:

1) оцінки параметрів моделі можуть бути незміщеними і обґрунтованими, але неефективними, тобто вибіркова дисперсія оцінки a_i може бути невиправдано великою;

2) статистичні критерій t- і F-статистик не можуть бути використані для верифікації моделі, бо їх розрахунок не враховує наявності корельованості залишків;

3) неефективність оцінок параметрів економетричної моделі зазвичай призводить до неефективних прогнозів, тобто прогнозні значення матимуть велику вибіркову дисперсію.

Для того, щоб з'ясувати наявність автокореляції, можна вдатися до графічного зображення, тобто нанести на графік послідовні величини ε_t (або ε_i) залежно від часу (або залежно від номера спостереження i, $i = \overline{1,n}$).

Якщо величина ε_t часто міняє знак, то говорять про *від'ємну автокореляцію*, якщо рідко – то про *додатну автокореляцію*.

Існують такі *методи перевірки наявності автокореляції залишків*:

1) критерій Дарбіна – Уотсона;

2) критерій фон Неймана;

3) нециклічний коефіцієнт автокореляції;

4) циклічний коефіцієнт автокореляції;

5) автокореляційна матриця;

6) спектральна щільність й ін.

Розглянемо деякі з цих методів.

Критерій Дарбіна – Уотсона для оцінювання автокореляції помилок застосовують найчастіше. Даний критерій заснований на перевірці гіпотези про існування автокореляції між сусідніми залишковими членами ряду. Статистика, що відповідає цьому критерію, позначається зазвичай як d або DW імає вигляд:

$$d = \frac{\sum_{t=2}^{n} (e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^2}.$$

Якщо автокореляція відсутня або незначна, то значення d приблизно дорівнює 2, при повній автокореляції значення d близько або до 0, або до 4.

Для d-статистики знайдені критичні границі, що дозволяють прийняти або відкинути гіпотезу про наявність автокореляції. Авторами цього критерію визначені верхні (d_L) і нижні (d_U) границі з 1; 2,5; 5 % рівнями значущості (рис. 5.1):

1. Якщо $0 < d \leq d_L$, то приймається гіпотеза про наявність додатної автокореляції.

2. Якщо $d_L < d < d_U$, то немає статистичних підстав ні прийняти, ні відкинути цю гіпотезу.

3. Якщо $d_U \leq d < 4 - d_U$, то приймається гіпотеза про відсутність автокореляції.

4. Якщо $4 - d_U < d < 4 - d_L$, то немає статистичних підстав ні прийняти, ні відкинути цю гіпотезу.

5. Якщо $4 - d_L < d$, то спостерігається від'ємна автокореляція.

Додатна автокореляція	Зона невизначеності	Автокореляція відсутня	Зона невизначеності	Від'ємна автокореляція
0 d_L	d_U	2 $4 - d_U$	$4 - d_L$	4

Рис. 5.1. **Перевірка гіпотези про наявність автокореляції залишків за критерієм Дарбіна – Уотсона**

У разі наявності в моделі лагової залежної змінної даний критерій непридатний, можна використовувати асимптотичний *h-тест Дарбіна*. Обидва ці теста призначені для перевірки автокореляції випадкових помилок першого порядку.

Критерій фон Неймана використовують для визначення автокореляції для залишків e_t, рідше. Статистика фон Неймана визначається за формулою:

$$Q = \frac{\sum_{t=2}^{n} \left(e_t - e_{t-1}\right)^2 \Big/ (n-1)}{\sum_{t=1}^{n} \left(e_t\right)^2 \Big/ n} = \frac{n}{n-1} \cdot \frac{\sum_{t=2}^{n} \left(e_t - e_{t-1}\right)^2}{\sum_{t=1}^{n} \left(e_t\right)^2}.$$

Значення статистики Дарбіна – Уотсона і статистики фон Неймана пов'язані між собою через співвідношення: $Q = \frac{n}{n-1} \cdot d$.

Якщо значення Q, що обчислюється за формулою, менше (або більше) деякого критичного значення, то говорять про додатну (відповідно про від'ємну) автокореляцію (рис. 5.2).

Додатна автокореляція	Автокореляція відсутня	Від'ємна автокореляція
0	$Q_{кр}^{+}$	$Q_{кр}^{-}$

Рис. 5.2. **Перевірка гіпотези про наявність автокореляції залишків за критерієм фон Неймана**

Так, якщо $Q_{факт} < Q_{кр}^{+}$, то існує додатна автокореляція, якщо $Q_{факт} > Q_{кр}^{-}$, то існує від'ємна автокореляція, а якщо $Q_{кр}^{+} < Q_{факт} < Q_{кр}^{-}$, то автокореляція залишків відсутня.

Тест Бройша (Бреуша) – Годфрі (тест серій) – більш універсальний асимптотичний тест для перевірки наявності автокореляції випадкових помилок більшого порядку. Цей тест є асимптотичним, тобто для достовірності висновків потрібен великий обсяг вибірки. У даному тесті випадкові помилки не обов'язково повинні бути нормально розподілені. Тест можна застосувати також і в авторегресійних моделях (на відміну від критерію Дарбіна – Уотсона).

Для перевірки автокореляції порядку p тест використовує допоміжну регресію МНК-залишків вихідної моделі на фактори цієї моделі і лагові значення залишків:

$$e_t = \sum_{j=1}^{m} \beta_j x_{tj} + \sum_{i=1}^{p} \alpha_i e_{t-i} + u_t.$$

Далі для цієї допоміжної регресії перевіряється гіпотеза про одночасне дорівнювання нулю всіх коефіцієнтів при лагових залишках. Перевірка здійснюється за допомогою відповідної LM-статистики: nR^2, де R^2 – коефіцієнт детермінації допоміжної моделі, а n – обсяг вибірки (цей обсяг вибірки на p менше обсягу вибірки для вихідної моделі, оскільки через лагові значення залишків у допоміжній регресії перші p спостережень не враховуються). Статистика тесту має асимптотичний розподіл $\chi^2(p)$. Якщо розраховане значення статистики перевищує критичне значення, то автокореляція визнається значущою, в іншому випадку вона незначуща.

Для тестування спільної гіпотези про рівність нулю всіх коефіцієнтів автокореляції до деякого порядку можна використовувати Q-тест Бокса – Пірса або Q-тест Л'юнга – Бокса.

Q-статистика Бокса – Пірса (*Box –PierceQ-statistics*) – статистичний критерій, призначений для знаходження автокореляції часових рядів. Замість тестування на випадковість кожного окремого коефіцієнта, він перевіряє на відміну від нуля відразу кілька коефіцієнтів автокореляції:

$$Q = n \sum_{k=1}^{m} \rho_k^2,$$

де n – число спостережень;

ρ_k – автокореляція k-го порядку;

m – число лагів, які перевіряються.

Однак на практиці цей критерій не рекомендується застосовувати, оскільки його вибіркові значення можуть значно відхилятися від розподілу χ^2. Замість нього застосовується Q-тест Л'юнга – Бокса, який дає більш якісні результати.

***Q*-тест Л'юнга – Бокса** може бути визначений таким чином. Висуваються дві конкуруючі гіпотези:

H_0: дані є випадковими (тобто є білим шумом).

H_1: дані не є випадковими.

Розраховується статистика:

$$Q = n(n+2)\sum_{k=1}^{m}\frac{\rho_k^2}{n-k},$$

де n – число спостережень;

ρ_k – автокореляція k-го порядку,

m – число лагів, які перевіряються.

Якщо $Q > \chi^2(m)$, де $\chi^2(m)$ – критичне значення статистики χ^2 з m ступенями свободи, то нульова гіпотеза відкидається і визнається наявність у часовому ряді автокореляції до m-го порядку. Q-тест Л'юнга – Бокса заснований на статистиці Бокса – Пірса, він має такий самий асимптотичний розподіл, але його розподіл ближче до χ^2 для кінцевих вибірок. Крім того, критерій не втрачає своєї спроможності навіть, якщо процес не має нормального розподілу (за наявності кінцевої дисперсії).

Для перевірки наявності автокореляції в динамічному ряді також використовують розрахунок автокореляційної матриці й автокореляційної функції, визначення спектральної щільності й т. д.

Найчастіше обмежуються аналізом ***автокореляційної матриці*** або ***матриці Лорана***:

$$P_n = \begin{pmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{n-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{n-2} \\ \dots & \dots & \dots & \dots & \dots \\ \rho_{n-1} & \rho_{n-2} & \rho_{n-3} & \dots & 1 \end{pmatrix},$$

де ρ_k – коефіцієнт автокореляції порядку k, тобто коефіцієнт кореляції між членами ряду залишків з лагом k, а саме між e_t і e_{t-k}:

$$\rho_k = \rho\,(e_t, e_{t-k}) = \frac{cov(e_t, e_{t-k})}{\sqrt{D(e_t)}\,\sqrt{D(e_{t-k})}}, k = 1, n-1.$$

Найбільш важливим із різних коефіцієнтів є коефіцієнт автокореляції першого порядку ρ_1 (його позначають ρ),що показує тісноту зв'язку між сусідніми рівнями $e_1, e_2, \dots, e_{n-1}$ і $e_2, e_3, \dots, e_n$.

Визначити оцінку цього коефіцієнта можна за формулами *нециклічного* або *циклічного коефіцієнта автокореляції*.

Зв'язок між коефіцієнтами автокореляції різного порядку можна визначити так: $\rho_k = \rho\,(e_t, e_{t-k}) = (\rho)^k$.

Нециклічний коефіцієнт автокореляції. Цей коефіцієнт виражає ступінь взаємозв'язку залишків кожного наступного значення з попереднім, а саме:

I ряд – $e_1, e_2, e_3, \dots, e_{n-1}$;

II ряд – $e_2, e_3, e_4, \dots, e_n$.

Він обчислюється за формулою:

$$r^* = \frac{\sum_{t=2}^{n}(e_t e_{t-1}) - \frac{1}{n-1}(\sum_{t=1}^{n} e_t)(\sum_{t=2}^{n} e_{t-1})}{\sqrt{\left[\sum_{t=1}^{n} e_t^2 - \frac{1}{n-1}(\sum_{t=1}^{n} e_t)^2\right]\left[\sum_{t=2}^{n} e_{t-1}^2 - \frac{1}{n-1}(\sum_{t=2}^{n} e_{t-1})^2\right]}}.$$

Коефіцієнт r^* може приймати значення в інтервалі (-1;+1). Від'ємні значення свідчать про від'ємну автокореляцію, додатні – про додатну. Значення, що містяться в деякій критичній області близько нуля, свідчать про відсутність автокореляції, тобто стверджують нульову гіпотезу про відсутність автокореляції залишків. Оскільки ймовірнісний розподіл r^* встановити важко, то на практиці замість r^* обчислюють циклічний коефіцієнт автокореляції r^0.

Циклічний коефіцієнт автокореляції. Він виражає ступінь взаємозв'язку "замкнених" рядів залишків:

I ряд – $e_1, e_2, e_3, \dots, e_{n-1}, e_n$;

II ряд – $e_2, e_3, e_4, \dots, e_n, e_1$.

Обчислюється за формулою:

$$r^0 = \frac{\sum_{t=2}^{n} e_t e_{t-1} + e_n e_1 - \frac{1}{n}(\sum_{t=1}^{n} e_t)^2}{\sum_{t=1}^{n} e_t^2 - \frac{1}{n}(\sum_{t=1}^{n} e_t)^2}.$$

Для досить довгих рядів вплив циклічних членів на величину коефіцієнта r^0 незначний, тому можна вважати, що ймовірнісний розподіл r^0 наближається до розподілу r^*. Якщо останній член ряду дорівнює першому, тобто $e_1 = e_n$, то нециклічний коефіцієнт автокореляції дорівнює циклічному. Очевидно, що коли залишки не містять тренда, то припущення про рівність $e_1 = e_n$ недалеке від реальності і циклічний коефіцієнт автокореляції наближається до нециклічного [27].

Фактично обчислене значення циклічного коефіцієнта автокореляції порівнюється з табличним для вибраного рівня значущості і довжини ряду n. Якщо $r_{\text{факт}} > 0$ і $r_{\text{факт}} \geq r_{\text{кр}}^{+}$, то існує додатна автокореляція. Якщо $r_{\text{факт}} < 0$ і $r_{\text{факт}} > r_{\text{кр}}^{-}$ і, то існує від'ємна автокореляція.

Припускаючи, що:

$$\sum_{t=1}^{n} e_t \approx \sum_{t=2}^{n} e_{t-1} \approx 0,$$

циклічний коефіцієнт автокореляції можна записати у вигляді:

$$r^0 = \frac{n}{n-1} \cdot \frac{\sum_{t=2}^{n} e_t e_{t-1}}{\sum_{t=1}^{n} e_t^{\;2}}.$$

Приклад 5.1. Для моделі залежності заощаджень від рівня доходу треба визначити наявність автокореляції помилок. Вихідні дані та проміжні розрахунки наведені в табл. 5.1.

За допомогою функції "Регресія" з "Аналізу даних" у MS Excel на основі застосування МНК визначимо оцінки параметрів парної лінійної економетричної моделі залежності заощаджень від доходу:

$$Y = 12{,}469 + 6{,}075 \cdot X.$$

Таблиця 5.1

Вихідні дані

Рік	Заощадження, Y	Дохід, X	Y	Залишки, e_t	$(e_t - e_{t-1})^2$	e_t^2	$e_t \cdot e_{t-1}$
1	18,6	0,95	18,24	0,36	–	0,13	–
2	18,6	0,82	17,45	1,15	0,62	1,32	0,414
3	19,7	1,04	18,79	0,91	0,06	0,83	1,0465
4	21,1	1,53	21,76	-0,66	2,49	0,44	-0,6006
5	22,9	1,94	24,25	-1,35	0,48	1,83	0,891
6	23,9	1,75	23,10	0,80	4,64	0,64	-1,08
7	25,2	1,99	24,56	0,64	0,02	0,41	0,512
8	14,6	0,42	15,02	-0,42	1,13	0,18	-0,2688
9	15,5	0,59	16,05	-0,55	0,02	0,31	0,231
10	16,7	0,90	17,94	-1,24	0,47	1,53	0,682
11	18,6	0,95	18,24	0,36	2,55	0,13	-0,4464
$\sum$	**215,4**	**12,88**	**215,4**	**0**	**12,47**	**7,75**	**1,3807**

У результаті розрахунку критерію Дарбіна – Уотсона отримано такий результат:

$$d = \frac{\sum_{t=2}^{n} (e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^2} = \frac{12{,}47}{7{,}75} = 1{,}608.$$

У критерії Дарбіна – Уотсона існують верхня й нижня границі. Для моделі з однією незалежною змінною й 11 спостереженнями нижня границя $d_L = 0{,}93$, верхня – $d_U = 1{,}32$ з рівнем значущості 0,05 %.

Оскільки розрахункове значення d потрапляє в інтервал $d_U \leq d < 4 - d_U$ $(1{,}32 \leq \ d < 2{,}68)$, то можна зробити висновок про відсутність автокореляції залишків першого порядку у досліджуваній моделі.

Розрахуємо статистику фон Неймана:

$$Q = \frac{n}{n-1} \cdot \frac{\sum_{t=2}^{n} (e_t - e_{t-1})^2}{\sum_{t=1}^{n} (e_t)^2} = \frac{11 \cdot 12{,}47}{10 \cdot 7{,}75} = 1{,}7699.$$

Критичні значення цього критерію для обсягу вибірки $n = 15$ і рівня значущості $\alpha = 0{,}05$ є: $Q_{\text{кр}}^{+} = 1{,}29$ і $Q_{\text{кр}}^{-} = 3{,}3$. Оскільки значення Q = 1,7699, тобто 1,29 $<$ Q $<$3,3, то автокореляція залишків відсутня.

Розрахуємо циклічний коефіцієнт автокореляції:

$$r^0 = \frac{n}{n-1} \cdot \frac{\sum_{t=2}^{n} e_t e_{t-1}}{\sum_{t=1}^{n} e_t^{\ 2}} = \frac{11 \cdot 1{,}3807}{10 \cdot 7{,}75} = 0{,}196.$$

Отримане значення $r_{\text{факт}} > 0$. Критичне значення для додатної автокореляції для обсягу вибірки $n = 15$ і рівня значущості $\alpha = 0{,}05$ є: $r_{\text{кр}}^{+} = 0{,}328$. Оскільки $0{,}196 < 0{,}328$, то автокореляція залишків відсутня.

Таким, чином, усі три критерії дозволили зробити однаковий висновок.

5.2. Методи оцінювання параметрів з автокорельованими залишками

Один із найбільш простих способів урахування корельованості залишків – є припущення про те, що ряд випадкових відхилень формує авторегресійних процес першого порядку, тобто:

$$e_t = \rho e_{t-1} + \varepsilon_t,$$

де ε_t – незалежні випадкові величини, що мають нормальний розподіл;

ρ – коефіцієнт авторегресії.

Методи оцінювання параметрів моделі з автокорельованими залишками поділяються на групи:

1. Якщо значення ρ відомо:
   - метод Ейткена (УМНК);
   - метод перетворення вихідної інформації.
2. Якщо значення ρ невідомо:
   - метод Кохрейна – Оркатта;
   - метод Хілдрета – Лу;
   - метод Дарбіна.

Розглянемо ці методи більш детально.

За наявності автокореляції поширеним методом оцінювання невідомих параметрів є **узагальнений метод найменших квадратів (метод Ейткена)**, який базується на скоригованій вихідній інформації з урахуванням коваріації залишків. Цей метод було розглянуто у п. 4.3. Отримані за допомогою цього методу оцінки є незміщеними та ефективними.

Формула для оцінювання параметрів моделі на основі методу Ейткена запишеться так:

$$A = X^T S^{-1} X \ ^{-1} X^T S^{-1} Y \tag{5.1}$$

або

$$A = X^T V^{-1} X \ ^{-1} X^T V^{-1} Y, \tag{5.2}$$

де A – вектор оцінок параметрів економетричної моделі;

X – матриця незалежних змінних;

X^T – матриця, транспонована до матриці X;

S^{-1} – матриця, обернена до матриці кореляції залишків;

V^{-1} – матриця, обернена до матриці V, де $V = \sigma_e^2 S$, а σ_e^2 – залишкова дисперсія;

Y – вектор залежних змінних.

Отже, щоб оцінити параметри моделі на основі методу Ейткена, треба сформувати матрицю S або V.

Матриця S має вигляд:

$$S = \begin{matrix} 1 & \rho & \rho^2 & \rho^3 & \rho^4 & \dots & \rho^{n-1} \\ \rho & 1 & \rho & \rho^2 & \rho^3 & \dots & \rho^{n-2} \\ \rho^2 & \rho & 1 & \rho & \rho^2 & \dots & \rho^{n-3} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \rho^{n-4} & \rho^{n-5} & \dots & 1 \end{matrix}.$$

У цій симетричній матриці ρ^k виражає коефіцієнт автокореляції k-го порядку для залишків e_t. Очевидно, що коефіцієнт автокореляції нульового порядку дорівнює 1.

Оскільки коваріація залишків ρ^k при $k > 2$ часто наближається до нуля, то матриця, обернена до матриці S, матиме такий вигляд:

$$S^{-1} = \frac{1}{1-\rho^2} \begin{matrix} 1 & -\rho & 0 & 0 & 0 & ... & 0 \\ -\rho & 1+\rho^2 & -\rho & 0 & 0 & ... & 0 \\ 0 & -\rho & 1+\rho^2 & -\rho & 0 & ... & 0 \\ ... & ... & ... & ... & ... & ... & ... \\ 0 & 0 & 0 & 0 & 0 & ... & 1 \end{matrix} .$$

Параметр ρ наближено можна знайти на основі залишків, якщо обчислити циклічний коефіцієнт кореляції r^0. На практиці, як правило, $\rho \approx r^0$, але r^0 коригується на величину зміщення:

$$r'_{adj} = \frac{n}{n-1} \cdot \frac{\sum_{t=2}^{n} e_t e_{t-1}}{\sum_{t=1}^{n} e_t^2} + \frac{m+1}{n},$$

де e_t– величина залишків у період t;

e_{t-1} – величина залишків у період $t-1$;

n – число спостережень;

$\frac{m+1}{n}$ – величина зміщення (m– кількість незалежних змінних).

Матриця $V = \sigma_e^2 S$, де σ_e^2– залишкова дисперсія, що визначається за формулою:

$$\sigma_e^2 = \frac{1}{n-m-1} e^T e,$$

де e^T– вектор, транспонований до вектора залишків e;

$n - m - 1$ – число ступенів свободи.

Дисперсія залишків з урахуванням зміщення обчислюється так:

$$\sigma_e^2 = \frac{1}{n-m-1} e^T e \; \frac{n - \frac{1+\lambda\rho}{1-\lambda\rho}}{n-1} .$$

Величину λ можна обчислити методом 1МНК за допомогою авторегресійного рівняння $x_t = \lambda x_{t-1} + \varepsilon_t$. Тоді:

$$\lambda = \frac{\sum_{t=2}^{n} x_t x_{t-1}}{\sum_{t=2}^{n} x_{t-1}^2},$$

де x_t взято як відхилення від свого середнього значення.

Розглянемо алгоритм методу Ейткена для оцінювання параметрів моделі з автокореляцією залишків.

1 крок. Оцінка параметрів моделі за методом 1МНК.

2 крок. Дослідження залишків на наявність автокореляції.

3 крок. Формування матриці коваріації залишків V або S.

4 крок. Обернення матриці V або S.

5 крок. Оцінка параметрів методом Ейткена, тобто згідно з (5.1), (5.2).

Приклад 5.2. Для підприємства "Альфа" необхідно побудувати модель залежності індексу зниження собівартості продукції (Y_t) від продуктивності праці (X_t). Вихідні дані зібрані за 10 кварталів $(t = 1,10)$ і наведені в табл. 5.2.

Таблиця 5.2

Вихідні дані

t	1	2	3	4	5	6	7	8	9	10
X_t	15,5	16,2	17,2	18,4	20,2	22,8	26	27,3	19,2	20
Y_t	17,4	18,7	20,4	21,6	23,1	25,9	30	31,2	21,6	22,5

1. Висуваємо припущення про існування лінійної залежності між індексом зниження собівартості продукції (Y_t) і продуктивністю праці (X_t):

$$Y_t = a_0 + a_1 X_t + e_t;$$

$$Y_t = a_0 + a_1 X_t;$$

$$e_t = Y_t - Y_t.$$

2. Визначимо МНК-оцінки параметрів моделі a_0, a_1 за допомогою функції "Регресія" з "Аналізу даних" у MS Excel, припускаючи що залишки e_t не корельовані:

$$a_0 = 0{,}19276; a_1 = 1{,}13645.$$

Економетрична модель має вигляд: $Y_t = 0{,}19276 + 1{,}1365X_t$.

3. Знайдемо розрахункові значення індексу зниження собівартості продукції на основі моделі і визначимо залишки e_t (табл. 5.3).

Таблиця 5.3

Розрахункові значення

t	Y	Y_t	e_t	$(e_t - e_{t-1})^2$	$e_t \cdot e_{t-1}$	e_t^2
1	17,4	17,8078	-0,4078	–	–	0,1663
2	18,7	18,6033	0,0967	0,2545	-0,0394	0,0094
3	20,4	19,7397	0,6603	0,3176	0,0639	0,4360
4	21,6	21,1035	0,4965	0,0268	0,3278	0,2465
5	23,1	23,1491	-0,0491	0,2977	-0,0244	0,0024
6	25,9	26,1039	-0,2039	0,0240	0,0100	0,0416
7	30,0	29,7405	0,2595	0,2147	-0,0529	0,0673
8	31,2	31,2179	-0,0179	0,0770	-0,0046	0,0003
9	21,6	22,0126	-0,4126	0,1558	0,0074	0,1702
10	22,5	22,9218	-0,4218	0,0001	0,1740	0,1779
Сума	**232,4**	**232,40**	**0,00**	**1,3682**	**0,4618**	**1,3180**

4. Перевіримо наявність автокореляції залишків у моделі.

У результаті розрахунку критерію Дарбіна – Уотсона отримано такий результат:

$$d = \frac{\sum_{t=2}^{n}(e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^2} = \frac{1{,}3682}{1{,}318} = 1{,}038.$$

Для моделі з однієї незалежної змінної й 10 спостережень нижня границя $d_l = 1{,}08$, верхня – $d_u = 1{,}32$ з рівнем значущості 0,05 %.

Оскільки розрахункове значення d потрапляє в інтервал $0 < d < d_l$ (0 ≤ d< 1,08), то можна зробити висновок про присутність додатної автокореляції залишків досліджуваній моделі.

Розрахуємо статистику фон Неймана:

$$Q = \frac{n}{n-1} \cdot \frac{\sum_{t=2}^{n} (e_t - e_{t-1})^2}{\sum_{t=1}^{n} (e_t)^2} = \frac{10}{9} \cdot 1{,}038 = 1{,}153.$$

Критичні значення цього критерію для обсягу вибірки $n = 10$ і рівня значущості $\alpha = 0{,}05$ є: $Q_{\text{кр}}^{+} = 1{,}18$ і $Q_{\text{кр}}^{-} = 3{,}61$.

Оскільки значення Q = 1,153, тобто $Q < Q_{\text{кр}}^{+}$ (1,153 < 1,18), то присутня додатна автокореляція залишків.

Розрахуємо циклічний коефіцієнт автокорреляції:

$$r^0 = \frac{n}{n-1} \cdot \frac{\sum_{t=2}^{n} e_t e_{t-1}}{\sum_{t=1}^{n} (e_t)^2} = \frac{10 \cdot 0{,}4618}{9 \cdot 1{,}318} = 0{,}389.$$

Отримане значення $r^0 > 0$.Критичне значення для додатної автокореляції для обсягу вибірки $n = 10$ і рівня значущості $\alpha = 0{,}05$ є: $r_{\text{кр}}^{+} = 0{,}36$. Оскільки 0,389 $>$ 0,36, то присутня додатна автокореляція залишків.

5. Використаємо метод Ейткена для оцінювання параметрів економетричної моделі з автокорельованими залишками. Оператор оцінювання запишеться так:

$$A = (X^T S^{-1} X)^{-1} X^T S^{-1} Y,$$

де S^{-1} – матриця, обернена до матриці S.

Матриця S – матриця коваріацій залишків, яка має вигляд:

$$S = \begin{pmatrix} 1 & \rho & \rho^2 & \rho^3 & \rho^4 & \rho^5 & \rho^6 & \rho^7 & \rho^8 & \rho^9 \\ \rho & 1 & \rho & \rho^2 & \rho^3 & \rho^4 & \rho^5 & \rho^6 & \rho^7 & \rho^8 \\ \rho^2 & \rho & 1 & \rho & \rho^2 & \rho^3 & \rho^4 & \rho^5 & \rho^6 & \rho^7 \\ \rho^3 & \rho^2 & \rho & 1 & \rho & \rho^2 & \rho^3 & \rho^4 & \rho^5 & \rho^6 \\ \rho^4 & \rho^3 & \rho^2 & \rho & 1 & \rho & \rho^2 & \rho^3 & \rho^4 & \rho^5 \\ \rho^5 & \rho^4 & \rho^3 & \rho^2 & \rho & 1 & \rho & \rho^2 & \rho^3 & \rho^4 \\ \rho^6 & \rho^5 & \rho^4 & \rho^3 & \rho^2 & \rho & 1 & \rho & \rho^2 & \rho^3 \\ \rho^7 & \rho^6 & \rho^5 & \rho^4 & \rho^3 & \rho^2 & \rho & 1 & \rho & \rho^2 \\ \rho^8 & \rho^7 & \rho^6 & \rho^5 & \rho^4 & \rho^3 & \rho^2 & \rho & 1 & \rho \\ \rho^9 & \rho^8 & \rho^7 & \rho^6 & \rho^5 & \rho^4 & \rho^3 & \rho^2 & \rho & 1 \end{pmatrix}.$$

6. Щоб сформувати матрицю S, необхідно визначити величину ρ, яка характеризує взаємозв'язок між послідовними членами ряду залишків. Припустимо, що залишки описуються авторегресійною моделлю першого степеня $e_t = \rho e_{t-1} + \varepsilon_t$. Тоді ρ можна оцінити через формулу скоректованого циклічного коефіцієнта автокореляції:

$$\rho \approx r'_{adj} = r^0 + \frac{m+1}{n} = 0{,}3893 + \frac{2}{10} = 0{,}5893.$$

Отже, матриця S матиме вигляд:

$$S = \left| \begin{array}{cccccccccc}
1 & 0{,}5893 & 0{,}3473 & 0{,}2047 & 0{,}1206 & 0{,}0711 & 0{,}0419 & 0{,}0247 & 0{,}0145 & 0{,}0086 \\
0{,}5893 & 1 & 0{,}5893 & 0{,}3473 & 0{,}2047 & 0{,}1206 & 0{,}0711 & 0{,}0419 & 0{,}0247 & 0{,}0145 \\
0{,}3473 & 0{,}5893 & 1 & 0{,}5893 & 0{,}3473 & 0{,}2047 & 0{,}1206 & 0{,}0711 & 0{,}0419 & 0{,}0247 \\
0{,}2047 & 0{,}3473 & 0{,}5893 & 1 & 0{,}5893 & 0{,}3473 & 0{,}2047 & 0{,}1206 & 0{,}0711 & 0{,}0419 \\
0{,}1206 & 0{,}2047 & 0{,}3473 & 0{,}5893 & 1 & 0{,}5893 & 0{,}3473 & 0{,}2047 & 0{,}1206 & 0{,}0711 \\
0{,}0711 & 0{,}1206 & 0{,}2047 & 0{,}3473 & 0{,}5893 & 1 & 0{,}5893 & 0{,}3473 & 0{,}2047 & 0{,}1206 \\
0{,}0419 & 0{,}0711 & 0{,}1206 & 0{,}2047 & 0{,}3473 & 0{,}5893 & 1 & 0{,}5893 & 0{,}3473 & 0{,}2047 \\
0{,}0247 & 0{,}0419 & 0{,}0711 & 0{,}1206 & 0{,}2047 & 0{,}3473 & 0{,}5893 & 1 & 0{,}5893 & 0{,}3473 \\
0{,}0145 & 0{,}0247 & 0{,}0419 & 0{,}0711 & 0{,}1206 & 0{,}2047 & 0{,}3473 & 0{,}5893 & 1 & 0{,}5893 \\
0{,}0086 & 0{,}0145 & 0{,}0247 & 0{,}0419 & 0{,}0711 & 0{,}1206 & 0{,}2047 & 0{,}3473 & 0{,}5893 & 1
\end{array} \right|.$$

Треба знайти зворотну матрицю:

$$S^{-1} = \left| \begin{array}{cccccccccc}
1{,}5321 & -0{,}9029 & 0{,}0000 & 0{,}0000 & 0{,}0000 & 0{,}0000 & 0{,}0000 & 0{,}0000 & 0{,}0000 & 0{,}000 \\
-0{,}903 & 2{,}0642 & -0{,}9029 & 0{,}0000 & 0{,}0000 & 0{,}0000 & 0{,}0000 & 0{,}0000 & 0{,}0000 & 0{,}000 \\
0{,}0000 & -0{,}9029 & 2{,}0642 & -0{,}9029 & 0{,}0000 & 0{,}0000 & 0{,}0000 & 0{,}0000 & 0{,}0000 & 0{,}000 \\
0{,}0000 & 0{,}0000 & -0{,}9029 & 2{,}0642 & -0{,}9029 & 0{,}0000 & 0{,}0000 & 0{,}0000 & 0{,}0000 & 0{,}000 \\
0{,}0000 & 0{,}0000 & 0{,}0000 & -0{,}9029 & 2{,}0642 & -0{,}9029 & 0{,}0000 & 0{,}0000 & 0{,}0000 & 0{,}000 \\
0{,}0000 & 0{,}0000 & 0{,}0000 & 0{,}0000 & -0{,}9029 & 2{,}0642 & -0{,}9029 & 0{,}0000 & 0{,}0000 & 0{,}000 \\
0{,}0000 & 0{,}0000 & 0{,}0000 & 0{,}0000 & 0{,}0000 & -0{,}9029 & 2{,}0642 & -0{,}9029 & 0{,}0000 & 0{,}000 \\
0{,}0000 & 0{,}0000 & 0{,}0000 & 0{,}0000 & 0{,}0000 & 0{,}0000 & -0{,}9029 & 2{,}0642 & -0{,}9029 & 0{,}000 \\
0{,}0000 & 0{,}0000 & 0{,}0000 & 0{,}0000 & 0{,}0000 & 0{,}0000 & 0{,}0000 & -0{,}9029 & 2{,}0642 & -0{,}903 \\
0{,}0000 & 0{,}0000 & 0{,}0000 & 0{,}0000 & 0{,}0000 & 0{,}0000 & 0{,}0000 & 0{,}0000 & -0{,}9029 & 1{,}532
\end{array} \right|.$$

Як можна бачити, вона має вигляд:

$$S^{-1}=\frac{1}{1-\rho^2}\begin{pmatrix} 1 & -\rho & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -\rho & 1+\rho^2 & -\rho & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\rho & 1+\rho^2 & -\rho & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\rho & 1+\rho^2 & -\rho & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\rho & 1+\rho^2 & -\rho & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\rho & 1+\rho^2 & -\rho & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\rho & 1+\rho^2 & -\rho & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\rho & 1+\rho^2 & -\rho & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\rho & 1+\rho^2 & -\rho \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\rho & 1 \end{pmatrix}.$$

Слід виконати такі розрахунки:

$$X^TS^{-1}=\begin{vmatrix} 0{,}629 & 0{,}258 & 0{,}258 & 0{,}258 & 0{,}258 & 0{,}258 & 0{,}258 & 0{,}258 & 0{,}258 & 0{,}629 \\ 9{,}121 & 3{,}915 & 4{,}264 & 4{,}213 & 4{,}497 & 5{,}350 & 8{,}434 & 15{,}542 & -3{,}075 & 13{,}306 \end{vmatrix}.$$

$$X^TS^{-1}X=\begin{matrix} 3{,}32557699 & 65{,}5664469 \\ 65{,}5664469 & 1419{,}12424 \end{matrix}.$$

$$X^TS^{-1}X^{\ -1}=\begin{matrix} 3{,}37526716 & -0{,}15594426 \\ -0{,}15594426 & 0{,}0079096 \end{matrix}.$$

$$X^TS^{-1}Y=\begin{matrix} 74{,}8465369 \\ 1623{,}22681 \end{matrix}.$$

$$A=\begin{matrix} 3{,}37526716 & -0{,}15594426 \\ -0{,}15594426 & 0{,}0079096 \end{matrix}\quad\begin{matrix} 74{,}8465369 \\ 1623{,}22681 \end{matrix}=\begin{matrix} -0{,}5058393 \\ 1{,}1671937 \end{matrix}.$$

$$a_0=-0{,}50584;\quad a_1=1{,}16719.$$

Економетрична модель має вигляд:

$$Y_t^*=-0{,}50584+1{,}16719X_t.$$

7. Знайдемо розрахункові значення Y_t^* на основі побудованої економетричної моделі та визначимо нові залишки u_t (табл. 5.4).

Таблиця 5.4

Розрахункові значення

t	Y	Y_t^*	u_t	$u_t - u_{t-1}\ {}^2$	$u_t \cdot u_{t-1}$	u_t^2
1	17,4	17,5857	-0,1857	–	–	0,0345
2	18,7	18,4027	0,2973	0,2333	-0,0552	0,0884
3	20,4	19,5699	0,83011	0,2839	0,2468	0,6891
4	21,6	20,9705	0,62948	0,0403	0,5225	0,3962
5	23,1	23,0715	0,02853	0,3611	0,0180	0,0008
6	25,9	26,1062	-0,2062	0,0551	-0,0059	0,0425
7	30,0	29,8412	0,1588	0,1332	-0,0327	0,0252
8	31,2	31,3585	-0,1585	0,1007	-0,0252	0,0251
9	21,6	21,9043	-0,3043	0,0212	0,0482	0,0926
10	22,5	22,838	-0,338	0,0011	0,1029	0,1143
Сума	**232,4**	**231,65**	–	**1,22992**	**0,8194**	**1,5087**

8. Обчислимо критерій Дарбіна – Уотсона:

$$d = \frac{1{,}2299}{1{,}5087} = 0{,}815.$$

Порівнявши його з критичним значенням з $n = 10$ і $\alpha = 0{,}05$, коли $d_{\text{факт}} < d_l$, можна дійти висновку, що звільнення від автокореляції залишків не відбулося. Це означає, що вихідна гіпотеза, коли залишки описуються авторегресійною схемою першого порядку, не виконується. Якщо залишки описуються авторегресійною схемою вищого порядку, то доцільно виконати оцінювання параметрів моделі методом Кохрейна – Оркатта або Дарбіна, які будуть розглянуті далі.

Метод перетворення вихідної інформації. Випадок, коли залишки задовольняють авторегресійну модель першого порядку, допускає альтернативний підхід до пошуку оцінок параметрів моделі за допомогою двокрокової процедури:

1) перетворення вихідної інформації із застосуванням для цього параметра ρ;

2) застосування 1МНК для оцінювання параметрів на основі перетворених даних.

Для цього треба знайти матрицю перетворення B, щоб модель:

$$BY = BXA + Be$$

мала скалярну дисперсійну матрицю $M(Be'eB') = \sigma_e^2 E$.

Розглянемо матрицю B_1 розміром $n \times n$:

$$B_1 = \begin{pmatrix} \sqrt{1-\rho^2} & 0 & 0 & 0 & \dots & 0 \\ -\rho & 1 & 0 & 0 & \dots & 0 \\ 0 & -\rho & 1 & 0 & \dots & 0 \\ 0 & 0 & -\rho & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & 1 \end{pmatrix}.$$

Безпосередньо множенням легко переконатись, що $M(B_1 e' e B_1{}') = \sigma_e^2 E$.

А це означає, що можна застосувати 1МНК до перетворених даних $B_1 Y$ і $B_1 X$, які мають вигляд:

$$B_1 Y = \begin{pmatrix} \sqrt{1-\rho^2}\, y_1 \\ y_2 - \rho y_1 \\ y_3 - \rho y_2 \\ y_4 - \rho y_3 \\ \dots \\ y_n - \rho y_{n-1} \end{pmatrix};$$

$$B_1 X = \begin{pmatrix} \sqrt{1-\rho^2} & \sqrt{1-\rho^2}\, x_1^1 & \dots & \sqrt{1-\rho^2}\, x_1^m \\ 1-\rho & x_2^1 - \rho x_1^1 & \dots & x_2^m - \rho x_1^m \\ \dots & \dots & \dots & \dots \\ 1-\rho & x_n^1 - \rho x_n^1 & \dots & x_n^m - \rho x_n^m \end{pmatrix}.$$

Неважко показати, що застосування 1МНК до даних $B_1 Y$ і $B_1 X$ дає таку саму оцінку параметрів моделі, як і метод Ейткена.

У загальному випадку, коли немає інформації ні про порядок авторегресійної моделі, ні про значення параметрів у ній, а через це не можна застосувати ні метод Ейткена, ні метод перетворення вихідної інформації, в економетричній літературі пропонуються наближені методи Кохрейна – Оркатта, Хілдрета – Лу і Дарбіна [27].

Метод Кохрейна – Оркатта (*Cochrane – Orcutt*)**.** Нехай задано економетричну модель [21]:

$$y_t = a_0 + a_1 x_t + e_t.$$

За допомогою звичайного МНК оцінюються її параметри і визначаються залишки регресії $e_t = y_t - y_t$. Відносно них висувається гіпотеза про присутність автокореляції, що описується авторегресійною схемою першого порядку:

$$e_t = \rho e_{t-1} + \varepsilon_t, \quad \rho \ < 1. \tag{5.3}$$

У якості приблизного значення ρ береться його МНК-оцінка з регресії (5.3).

Перетворюємо вихідну інформацію за допомогою ρ, і дістанемо:

$$y_t - \rho y_t = a_0 \ \ 1 - \rho \ \ + a_1 \ \ x_t - \rho x_{t-1} \ \ + \varepsilon_t. \tag{5.4}$$

У цій моделі залишки ε_t мають нульове математичне сподівання і постійну дисперсію. Для перетвореної моделі (5.4) буде отримано нові МНК-оцінки параметрів a_0, a_1 і ρ і новий вектор залишків.

Процедура триває доти, доки такі значення параметрів a_0, a_1 і ρ не будуть відрізнятись менше, ніж на задану величину.

Проведені дослідження показали, що в результаті застосування методу Кохрейна – Оркатта завжди буде знайдено глобальний оптимум і цей алгоритм забезпечує порівняно добру збіжність.

Часто пропонується альтернативний підхід до використання цього ітеративного методу.

На відміну від попереднього, у ньому подальші ітерації припиняються тоді, коли на основі критерію Дарбіна – Уотсона робиться висновок про відсутність автокореляції залишків.

Коли ітеративний процес припиняється, то виконується перевірка значущості параметрів за допомогою останньої економетричної моделі. У такому разі звичайні формули дадуть обґрунтовані оцінки дисперсій залишків.

Метод Хілдрета – Лу (*Hildreth – Lu*). У даному методі виконується прямий пошук значення коефіцієнта автокореляції, яке мінімізує суму квадратів залишків перетвореної моделі. А саме задаються значення ρ з можливого інтервалу (-1; 1) з деяким кроком (наприклад, 0,1 або 0,05). Для кожного з них виконуються авторегресійні перетворення:

$$y_t - \rho y_{t-1} = a_0 (1 - \rho) + \sum_{i=1}^{m} a_i (x_{it} - \rho x_{it-1}) + \varepsilon_t,$$

оцінюється модель звичайним МНК і знаходиться сума квадратів залишків. Вибирається той коефіцієнт автокореляції, для якого ця сума квадратів мінімальна. Далі в околиці знайденої точки будується сітка з більш дрібним кроком і процедура повторюється заново.

Метод Дарбіна (*Durbin*). Дарбін запропонував просту двокрокову процедуру, яка також дає оцінки параметрів, вони асимптотично мають той самий вектор середніх і ту саму матрицю дисперсій, що й оцінки методу найменших квадратів.

1 крок. Підставити значення залишків, яке підпорядковане авторегресійній моделі першого порядку $e_t = \rho e_{t-1} + \varepsilon_t$, до економетричної моделі $y_t = a_0 + a_1 x_t + e_t$. Тоді отримаємо $y_t = a_0 + a_1 x_t + \rho e_{t-1} + \varepsilon_t$, де $e_{t-1} = y_{t-1} - a_0 - a_1 x_{t-1}$.

Звідси:

$$y_t = a_0 (1 - \rho) + \rho y_{t-1} + a_1 x_t - a_1 \rho x_{t-1} + \varepsilon_t,$$

де ε_t має скалярну матрицю дисперсій.

Тобто y_{t-1} включається до числа регресорів, а ρ – до числа оцінюваних параметрів.

Згідно з 1МНК визначаються параметри цієї моделі, куди входить і коефіцієнт ρ. У результаті обчислень наявне $\rho = r$.

2 крок. Значення $\rho = r$ використовується для перетворення змінних $(y_t - r y_{t-1})$ і $(x_t - r x_{t-1})$, а 1МНК застосовується до перетворених даних. Коефіцієнт при $(x_t - r x_{t-1})$ є оцінкою параметра a_1, а вільний член, поділений на $-r$, оцінює параметр a_0.

Метод Дарбіна дуже просто поширюється на випадок кількох незалежних змінних і для автокореляції вищих порядків.

Нехай задано модель:

$$y_t = a_0 + a_1 x_{1t} + a_2 x_{2t} + \cdots + a_m x_{mt} + e_t, \qquad (5.5)$$

де $e_t = \rho_1 e_{t-1} + \rho_2 e_{t-2} + \varepsilon_t$.

Підставивши значення e_t в (5.5), отримаємо:

$$y_t = \rho_1 y_{t-1} + \rho_2 y_{t-2} + a_1 x_{1t} + \cdots + a_m x_{mt} - \rho_1 a_1 x_{1,t-1} - \ldots$$
$$-\rho_1 a_m x_{m,t-1} - \rho_2 a_1 x_{1,t-2} - \cdots - \rho_2 a_m x_{m,t-2} + \varepsilon_t.$$

Застосувавши 1МНК, обчислимо параметри цієї моделі. Коефіцієнти $\rho_1 = r_1$ і $\rho_2 = r_2$ використаємо для перетворення даних:

$$y_t - r_1 y_{t-1} - r_2 y_{t-2}\ , \quad x_{jt} - r_1 x_{j,t-1} - r_2 x_{j,t-2}\ , \qquad j = 1, m, \qquad t = 1, n.$$

Знову застосуємо 1МНК для цих перетворених даних і знайдемо оцінки параметрів моделі a_0, a_j $j = 1, m$.

У відомих роботах із проблем оцінювання параметрів моделей з автокорельованими залишками було доведено такі результати [27]:

1МНК дає менш ефективні оцінки порівняно з іншими методами, якщо сукупність спостережень $n = 20$ одиниць, а $\rho > 0{,}3$.

Якщо $\rho < 0{,}3$, то зниження ефективності оцінок 1МНК порівняно зі складнішими процедурами невелике.

Метод Дарбіна забезпечує найкраще оцінювання для ширшого кола параметрів порівняно з іншими методами.

Контрольні запитання для самодіагностики

1. Дайте визначення понять автокореляції, автокореляції залишків.
2. Чим відрізняється додатна автокореляція від від'ємної?
3. Що є причиною виникнення автокореляції залишків у регресійній моделі?
4. Які наслідки автокореляції залишків?
5. Які існують методи оцінювання автокореляції залишків?
6. Як визначається оцінка невідомого коефіцієнта автокореляції залишків?
7. Як визначається порядок авторегресійної схеми?

8. У чому полягає особливість використання методу Ейткена для оцінювання параметрів моделі з автокорельованими залишками?

9. Як формується матриця S у методі Ейткена для оцінювання параметрів моделі з автокорельованими залишками?

10. Які оцінки параметрів дає метод перетворення вихідної інформації порівняно з методом Ейткена?

11. Чим відрізняється оцінювання параметрів моделі з автокорельованими залишками за методами Кохрейна – Оркатта, Хілдрета – Лу, Дарбіна?

12. Які переваги у застосуванні методу Дарбіна для оцінювання параметрів моделі з автокорельованими залишками?

Тести

1. Автокореляція – це кореляція між:

а) відповідними за моментом часу значеннями двох часових рядів;

б) членами одного і того ж часового ряду;

в) кореляція між членами часових рядів.

2. Автокореляція залишків – це:

а) постійна дисперсія залишків для кожного спостереження;

б) зміна дисперсії залишків для груп спостережень;

в) наявність зв'язку між послідовними значеннями залишків.

3. Основні джерела автокореляції:

а) помилки специфікації;

б) помилки вимірювання;

в) характер спостережень.

4. За допомогою якого тесту можна визначити присутність автокореляції:

а) Дарбіна – Уотсона;

б) Феррара – Глобера;

в) Глейзера.

5. У моделі може мати місце значна автокореляція відхилень, якщо за критерієм Дарбіна – Уотсона:

а) $d \approx 2$;

б) $d \approx 0$;

в) $d \approx 0$ або $d \approx 4$.

6. Стверджують про наявність від'ємної автокореляції відхилень економетричної моделі, якщо за критерієм Дарбіна – Уотсона:

а) $d \approx 0$ або $d < d_l$;

б) $d \approx 2$ або $d_l < d < d_u$;

в) $d \approx 4$ або $d > (4 - d_l)$.

7. Оцінки параметрів, отримані за допомогою методу найменших квадратів (1МНК) у разі автокореляції відхилень, будуть:

а) незміщеними, обґрунтованими та ефективними;

б) зміщеними, обґрунтованими і неефективними;

в) незміщеними, необґрунтованими і ефективними;

г) незміщеними, обґрунтованими і неефективними;

д) зміщеними, необґрунтованими і неефективними.

8. Які з методів оцінювання параметрів моделі з автокорельованими залишками припускає використання відомої їх коваріаційної матриці:

а) метод Кохрейна – Оркатта;

б) метод Хілрета – Лу;

в) метод Ейткена;

г) метод Дарбіна.

9. Які з методів оцінювання параметрів моделі з автокорельованими залишками застосовуються у разі невідомої коваріаційної матриці відхилень:

а) метод Ейткена, метод Хілрета – Лу;

б) метод Кохрейна – Оркатта, метод Хілрета – Лу, метод Дарбіна;

в) метод Дарбіна – Уотсона, метод Глейзера, метод Ейткена;

г) метод Ейткена, метод Голдфельда – Квандта;

д) метод Глейзера, метод Хілрета – Лу.

10. За допомогою якого методу можна позбавитися від негативних наслідків автокореляції залишків високого порядку у моделі:

а) Дарбіна;

б) Фон Неймана;

в) Глейзера.

Практичні завдання

1. За даними з табл. 5.5 побудувати лінійну економетричну модель і перевірити наявність автокореляції залишків за критерієм Дарбіна – Уотсона.

Таблиця 5.5

Вихідні дані

t	1	2	3	4	5	6	7	8	9	10
X_t	10	12	13,4	16,5	18,4	12,5	14,6	13,4	19,2	20
Y_t	30	35	40	44	55	42	48	40	41	52

2. За даними з табл. 5.5 перевірити наявність автокореляції залишків за критерієм фон Неймана.

3. Визначити циклічний коефіцієнт автокореляції для даних із табл. 5.5.

4. За даними з табл. 5.6 побудувати лінійну економетричну модель і перевірити наявність автокореляції залишків за критерієм Дарбіна – Уотсона і фон Неймана. Оцінити параметри моделі на основі методу Ейткена (УМНК).

Таблиця 5.6

Вихідні дані

t	1	2	3	4	5	6	7	8	9	10
X_t	10	9,8	13	14	15,1	18,2	20,2	23,5	25	20
Y_t	27	15	15	15	16	16	16	17	18	12

5. За даними з табл. 5.6 оцінити на основі методу Дарбіна параметри моделі:

$$y_t = a_0 + a_1 x_t + e_t,$$

де $e_t = \rho_1 e_{t-1} + \rho_2 e_{t-2} + \varepsilon_t$.

Порівняйте отримані результати з результатами із завдання 4.

Ключові слова

Автокореляція залишків. Від'ємна автокореляція. Додатна автокореляція. Критерій Дарбіна – Уотсона. Критерій фон Неймана. Тест Бройша (Бреуша) – Годфрі (тест серій). Q-тест Л'юнга – Бокса. Q-статистика Бокса – Пірса. Нециклічний коефіцієнт автокореляції. Циклічний коефіцієнт автокореляції. Метод перетворення вихідної інформації. Метод Кохрейна – Оркатта. Метод Хілдрета – Лу. Метод Дарбіна.

Розділ 6. Емпіричні методи кількісного аналізу на основі статистичних рівнянь

6.1. Нелінійні однофакторні економетричні моделі, їх властивості. Методи оцінювання параметрів нелінійних моделей.

6.2. Еластичність функцій однієї та багатьох змінних.

6.3. Виробничі функції, їх класифікація та основні властивості.

6.4. Виробнича функція Кобба – Дугласа, особливості побудови та аналізу.

6.5. Основні характеристики виробничих функцій, їх геометрична та економічна інтерпретації.

6.1. Нелінійні однофакторні економетричні моделі, їх властивості. Методи оцінювання параметрів нелінійних моделей

У багатьох практичних випадках моделювання економічних залежностей на основі лінійних рівнянь дає досить задовільні результати і може використовуватися для аналізу та прогнозування. Однак через багатогранність та складність економічних процесів досить часто між реальними економічними явищами можливі нелінійні взаємозв'язки. Існує досить великий клас економетричних моделей, за допомогою яких можна вивчати нелінійні зв'язки між різними факторами. Вибір форми залежності повинен ґрунтуватися на основі змістовного аналізу досліджуваних процесів, а також за результатами аналізу взаємозв'язку змінних, що входять у модель. Виділяють *два класи нелінійних регресійних моделей*:

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

2. Нелінійні як відносно включених в аналіз пояснюючих змінних, так і за оцінюваними параметрами – квазілінійні (степенева функція, показникові) [36].

Слід розглянути основні типи нелінійних економетричних моделей, які найбільш часто використовуються для моделювання економічних процесів:

1) поліном m-го ступеня $y = a_0 + a_1 x + \cdots + a_m x^m$;

2) показникова функція $y = a_0 a_1^x$;

3) степенева (мультиплікативна) функція $y = a_0 x^{a_1}$;

4) модифікована експонента $y = k + a_0 e^{a_1 x}$;

5)гіперболічна залежність $y = a_0 + \frac{a_1}{x}$;

6) крива Гомперця $y = a_0 a_1^{a_2^x}$;

7) логістична крива $y = \frac{k}{1+a_1 a_2^x}$.

Для оцінювання параметрів нелінійних моделей використовують два основні підходи [36]:

- перший основний прийом, за допомогою якого спрощується процес оцінювання параметрів нелінійної моделі – це лінеаризація. Лінеаризація – перехід від нелінійних зв'язків (показникової, степеневої тощо) до лінійного зв'язку за допомогою різних перетворень, що дозволяє надалі використовувати звичайний метод найменших квадратів (МНК). Отже, для оцінювання параметрів виконується заміна змінних, логарифмування відповідних частин рівняння або комбінований підхід для моделей нелінійних за змінними та параметрами;
- другий підхід використовується, коли підібрати відповідні перетворення не вдається, у даному випадку використовують методи нелінійної оптимізації на основі вихідних змінних.

Варто розглянути геометричне подання основних нелінійних функцій (рис. 6.1 – 6.9) [15].

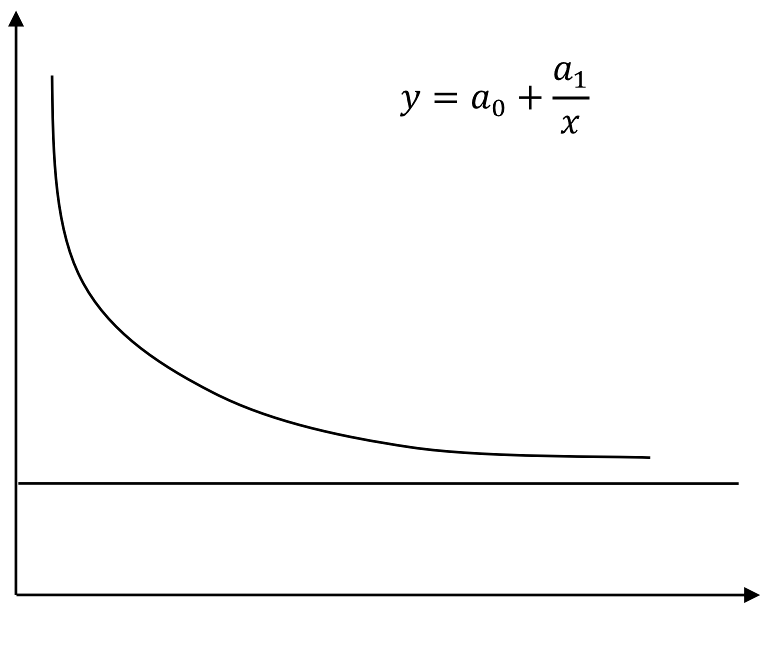
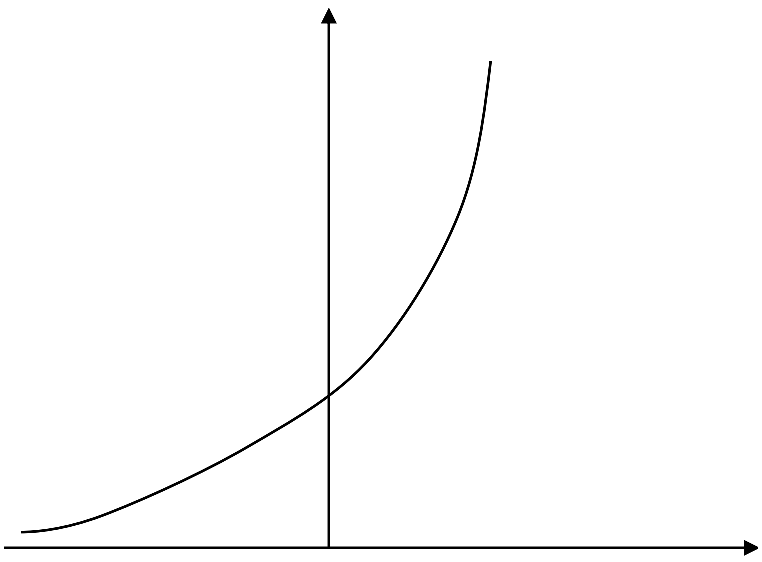


Рис. 6.1. **Гіперболічний зв'язок**



$$y = a_0 \cdot e^{a_1 x}$$

Рис. 6.2. **Експоненціальний зв'язок**

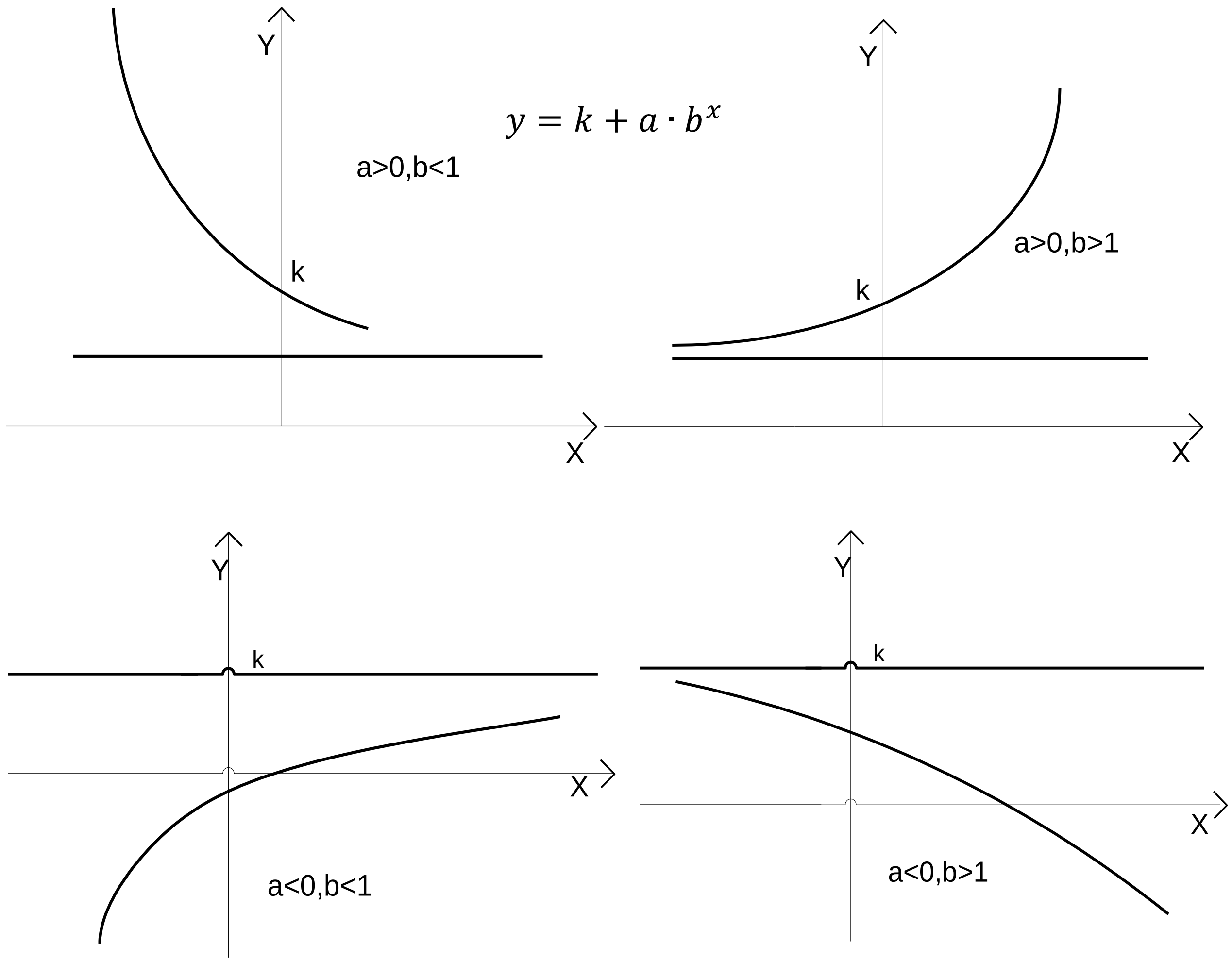


Рис. 6.3. **Модифікована експонента**

$$y = k \cdot a^{b^x}$$

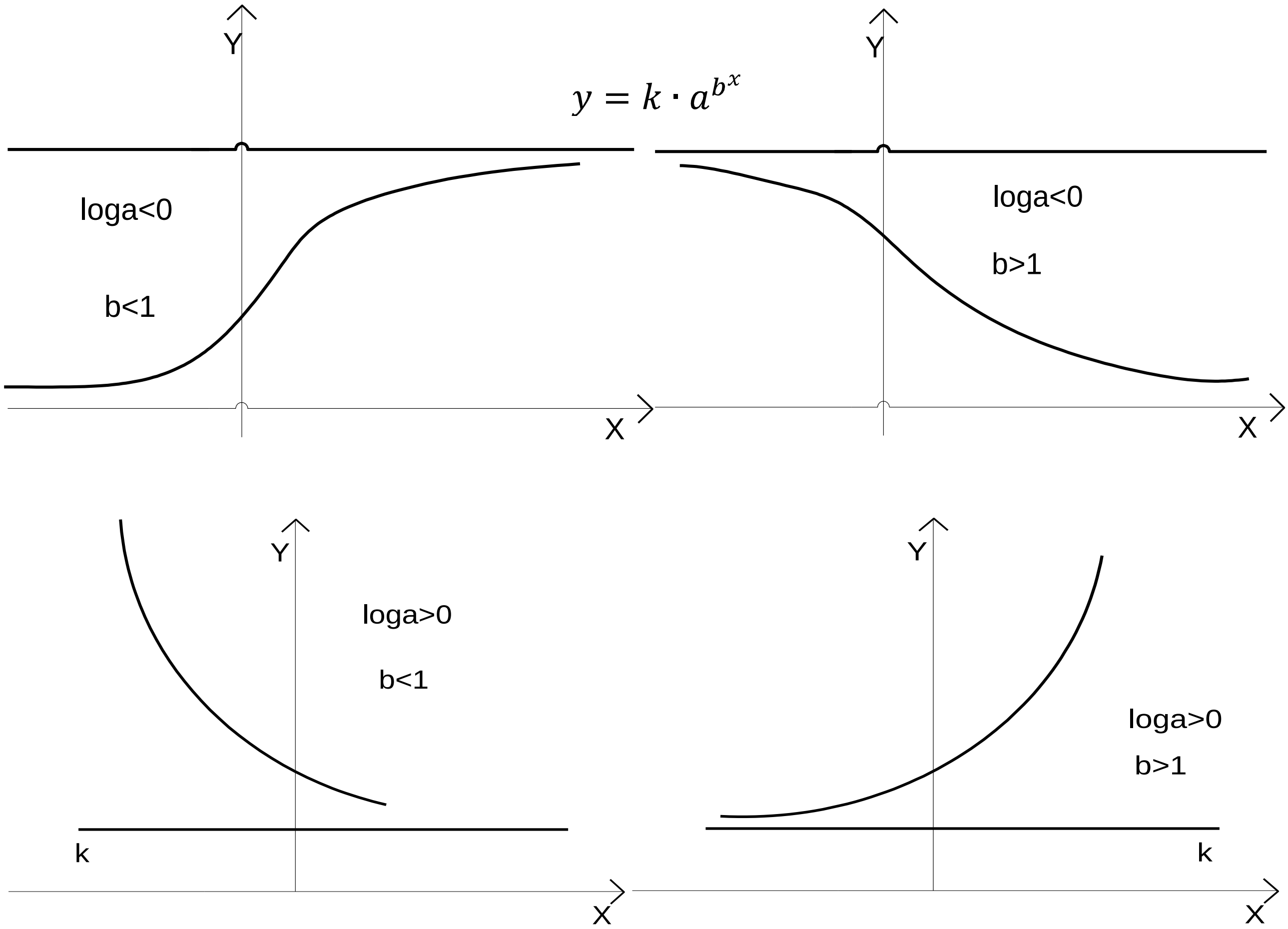


Рис. 6.4. **Крива Гомперця**

$$y = \frac{k}{1 + b \cdot e^{-ax}}$$

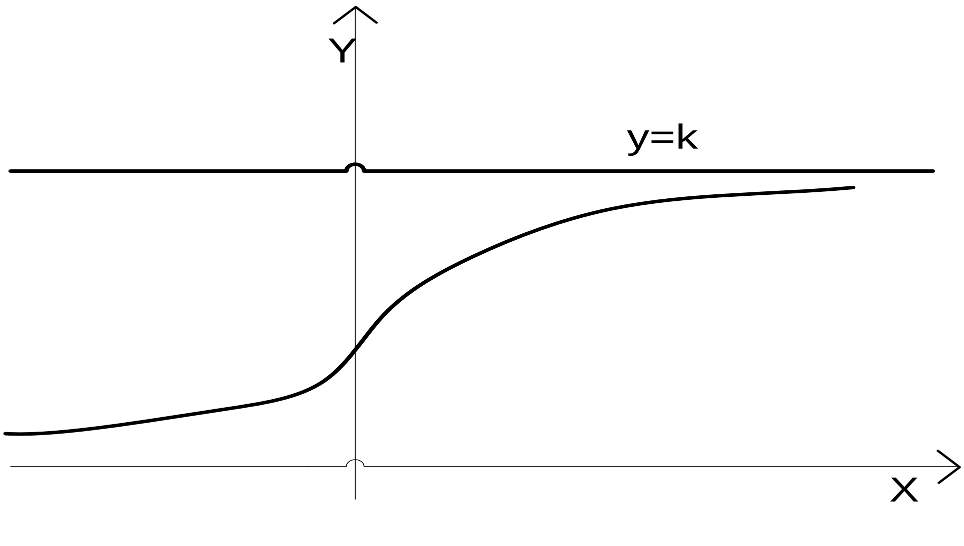


Рис. 6.5. **Логістична крива**

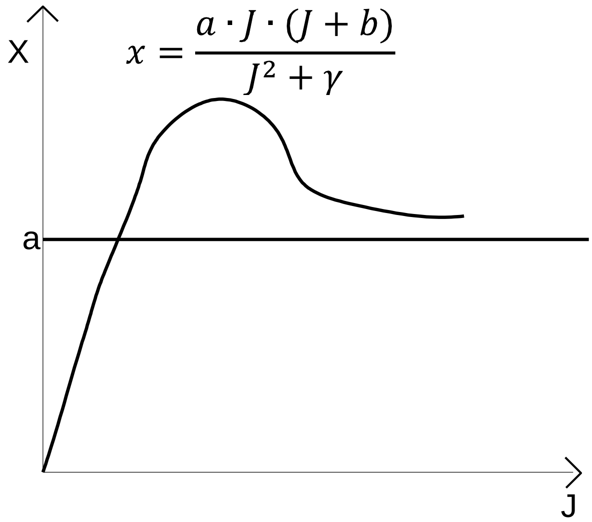


Рис. 6.6. **Функція Торнквіста.**
Попит на малоцінні товари залежно від доходу

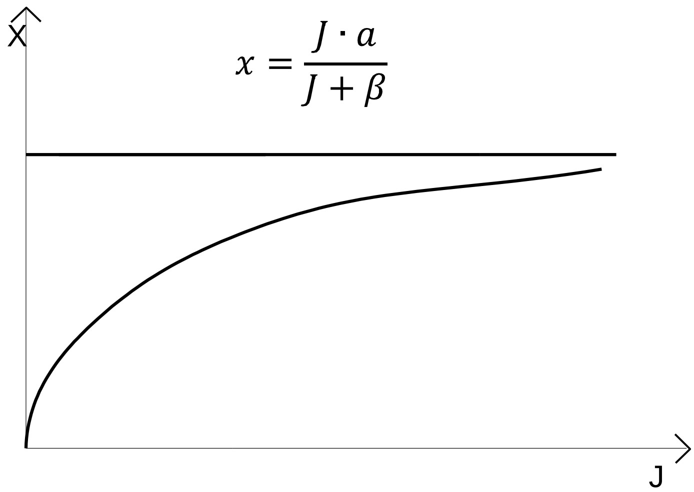


Рис. 6.7. **Функція Торнквіста.**
Попит на товари першої необхідності залежно від доходу

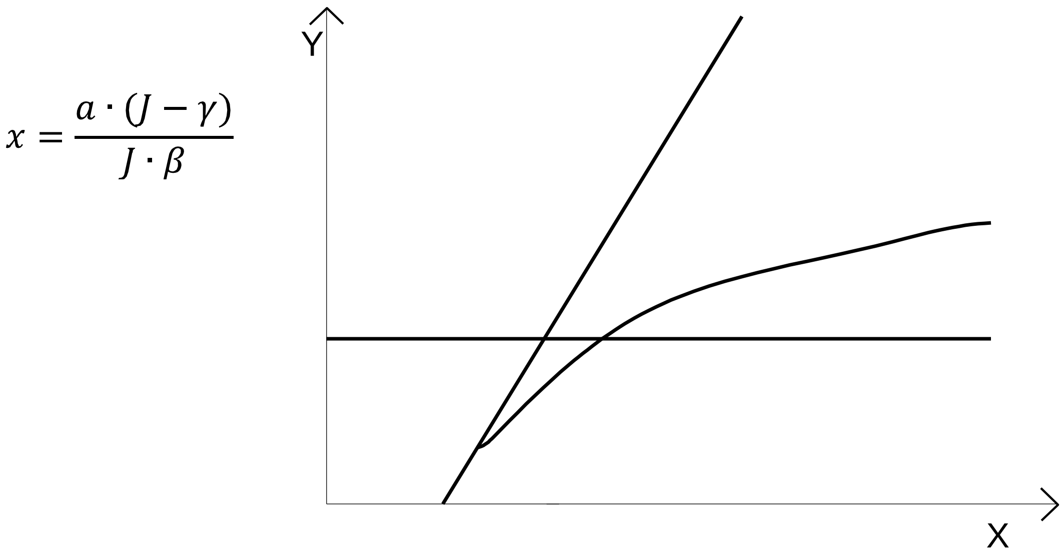


Рис. 6.8. **Функція Торнквіста.**
Попит на товари другої необхідності залежно від доходу

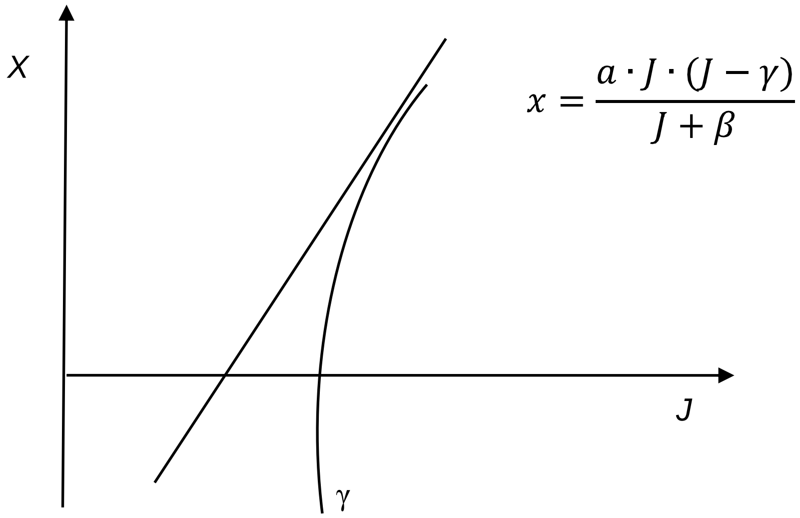


Рис. 6.9. **Функція Торнквіста.**
Попит на предмети розкоші залежно від доходу

Слід розглянути *процедуру лінеаризації* деяких розглянутих функцій [1].

1. Найчастіше в економіці використовують поліноміальні функції другого й третього ступенів. Функції більш високих ступенів мають занадто велику кількість перегинів, які неможливо інтерпретувати з погляду економічної сторони явищ. Такі функції легко лінеаризуються за допомогою переходу до нових змінних. Варто проілюструвати процедуру лінеаризації на прикладі полінома другого ступеня. Регресійна модель має вигляд:

$$y = a_0 + a_1 x + \ a_2 x^2 + \varepsilon. \tag{6.1}$$

Позначивши $x_1' = x,\ x_2' = x^2$ і підставивши нові змінні в модель (6.1) буде одержано $y = a_0 + a_1 x_1' + \ a_2 x_2' + \varepsilon$. Параметри моделі легко одержати за допомогою МНК.

2. Зв'язок показниковий (експонентний): $y = a_0 \cdot e^{a_1 x}$. Лінеаризація в цій моделі здійснюється так: $y' = lny$, тоді перетворена модель має такий вигляд: $y' = lna_0 + a_1 \cdot x = a_0' + a_1 \cdot x$.

3. Степенева (мультиплікативна) функція: $y = a_0 \cdot x^{a_1}$. Лінеаризація в даній моделі здійснюється в такий спосіб: $y' = lny$, тоді перетворена модель має вигляд: $y' = lna_0 + a_1 \cdot lnx = a_0' + a_1 \cdot x'$.

4. Модифікована показникова: $y = k + a \cdot b^x$. Ця функція має горизонтальну асимптоту, виду $y = k$, її графік прагне до асимптоти при $x \to \infty$, або $x \to -\infty$.

Відмінна риса модифікованої показникової функції в тому, що:

$y\ \ x+1\ \ -y\ \ x\ \ =U_{x_1}=k+a\cdot b^x-k-a\cdot b^{x-1}=a\cdot b^{x-1}\cdot\ \ b-1$, тоді

$$\frac{U_{x_2}}{U_{x_1}}=\frac{U_{x_3}}{U_{x_2}}=......=b=const.$$

Тоді модель на логарифмах цих відносин буде лінійна відносно x: $logU_x=loga+\log\ \ b-1\ \ +\ \ x-1\ \ \cdot logb$. Шляхом заміни залежної змінної й параметрів одержуємо: $z=I_0+I_1\cdot\ \ x-1$.

5. Зв'язок гіперболічний: $y=a_0+\frac{a_1}{x}$. Оцінка параметрів проводиться для моделі, у якій здійснюється заміна: $x'=\frac{1}{x}$. Таким чином, отримана лінійна модель: $y'=a_0+a_1x'$.

6. Крива Гомперця: $y=a_0\cdot a_1^{a_2^x}$. Після логарифмування моделі $y=a_0\cdot a_1^{a_2^x}$ буде одержано

$$lny=lna_0+a_2^xlna_1. \qquad (6.2)$$

Модель (6.2) є модифікованою показниковою функцією, спосіб лінеаризації якої був наведений. Слід перейти до перших приростів і другий раз прологарифмувати модель:

$$\begin{gathered}\Delta lny=lny\ \ x\ \ -lny\ \ x-1\ \ =lna_0+a_2^xlna_1-lna_0-a_2^{x-1}lna_1=\\=a_2^{x-1}\ \ a_2-1\ \ lna_1\\\rightarrow\\ln\Delta lny=xlna_2-lna_2+\ln\ \ a_2-1\ \ +lnlna_1.\end{gathered} \qquad (6.3)$$

Провівши заміни $z=ln\Delta lny,\ \ B=lna_2,\ \ A=-lna_2+\ln\ \ a_2-1\ \ +lnlna_1$ у моделі (6.3) буде одержано лінійну модель: $z=A+Bx$, параметри якої A і B можуть бути отримані за допомогою МНК.

Після їхнього оцінювання за допомогою МНК параметри початкової нелінійної моделі $y=a_0a_1^{a_2^x}$ одержимо в такий спосіб:

$$\begin{gathered}B=lna_2\rightarrow a_2=e^B;\\A=-lna_2+\ln\ \ a_2-1\ \ +lnlna_1\rightarrow\\A=-B+\ln\ \ e^B-e^0\ \ +lnlna_1=-B+lne^{B-0}+lnlna_1\rightarrow\\A=-B+B+lnlna_1\rightarrow\\A=lnlna_1\rightarrow a_1=e^{e^A}.\end{gathered}$$

Параметр a_0 можна одержати, підставляючи в $y = a_0 a_1^{a_2^x}$ значення $y = y$ й $x = x$.

7. Логістична крива $y = \frac{k}{1+a_1 a_2^x}$. Логістична крива центрально симетрична щодо крапки перегину. За умови $x \to -\infty$ ордината прагне до нуля, а при $x \to \infty$, то ордината прагне до асимптоти. Логістична крива схожа з кривою Гомперця. Обидві криві характеризують зростання із мінливим відхиленням приросту до ординати. Відмінність укладається в тому, що в кривій Гомперця постійні відносини перших різниць логарифмів послідовно рівновіддалених одна від одної ординат, а в логістичної кривої незмінні відносини перших різниць зворотних значень.

Щоб оцінити параметри такої моделі, слід або встановити параметр k рівним певній константі виходячи з передумов теорії досліджуваного процесу, а потім знаходити інші параметри, або знаходити параметри моделі у формі:

$$y = \frac{k}{1+a_1 a_2^x} = \frac{1}{b_0+b_1 a_2^x}, \tag{6.4}$$

де $b_0 = \frac{1}{k}$, $b_1 = \frac{a_1}{k}$.

Процедура лінеаризації (6.4) полягає в такому: ввести нову змінну $y^* = \frac{1}{y}$. Модель набуває виду модифікованої показникової функції $y^* = b_0 + b_1 a_{2,}^x$ процедура лінеаризації якої була наведена.

Слід розглянути особливості побудови нелінійної економетричної моделі на такому прикладі [24]. Необхідно здійснити аналіз взаємозв'язку між рівнем витрат на захист інтелектуальної власності та прибутком підприємства (табл. 6.1). Передбачається, що для опису сукупності може бути використана степенева функція такого вигляду: $y = a_0 x^{a_1}$. Для лінеаризації буде застосовано процедуру логарифмування:

$$lny = lna_0 + a_1 lnx.$$

Таблиця 6.1

Вихідні дані

Період	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
X	1,20	2,40	3,60	4,80	6,00	7,20	8,40	9,60
Y	4,81	5,21	6,14	6,62	7,12	7,97	8,76	9,53
Період	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
X	10,80	12,00	13,20	14,40	15,60	16,80	18,00	19,20
Y	9,95	10,19	9,95	10,20	10,33	11,09	11,12	11,50

Треба зробити заміни: $y^* = lny,\ a_0^* = lna_0,\ x^* = lnx$. Отримаємо лінійну форму моделі: $y^* = a_0^* + a_1 x^*$. Оскільки перетворена залежність є лінійною, її параметри можуть бути знайдені за допомогою МНК за такою формулою: $a = (X^T X)^{-1} \cdot XY$.

Послідовність розрахунку параметрів моделі $y^* = a_0^* + a_1 x^*$ наведено далі. Перетворені значення показників подано в табл. 6.2.

Таблиця 6.2

Перетворені значення показників

X	Y	$y^* = lny$	$x^* = lnx$
1,2	4,81	1,571	0,182
2,4	5,21	1,651	0,875
3,6	6,14	1,815	1,281
4,8	6,62	1,890	1,569
6	7,12	1,963	1,792
7,2	7,97	2,076	1,974
8,4	8,76	2,170	2,128
9,6	9,53	2,254	2,262
10,8	9,95	2,298	2,380
12	10,19	2,321	2,485
13,2	9,95	2,298	2,580
14,4	10,2	2,322	2,667
15,6	10,33	2,335	2,747
16,8	11,09	2,406	2,821
18	11,12	2,409	2,890
19,2	11,5	2,442	2,955

Слід розрахувати матриці для знаходження параметрів моделі:

$$B = X^{*T} \times X^T = \begin{matrix} 16 & 33{,}589 \\ 33{,}589 & 79{,}855 \end{matrix}; \qquad B^{-1} = \begin{matrix} 0{,}534 & -0{,}225 \\ -0{,}225 & 0{,}107 \end{matrix};$$

$$X^{*T} \times Y^* = \begin{matrix} 34{,}221 \\ 75{,}094 \end{matrix};$$

$$a = B^{-1} \cdot X^{*T} Y^* = \begin{matrix} 0{,}534 & -0{,}225 \\ -0{,}225 & 0{,}107 \end{matrix} \cdot \begin{matrix} 34{,}221 \\ 75{,}094 \end{matrix} = \begin{matrix} 1{,}407 \\ 0{,}348 \end{matrix}.$$

Виходячи з наведених розрахунків, параметри моделі $y^* = a_0^* + a_1 x^*$ дорівнюють $a_0^* = 1{,}407,\ a_1 = 0{,}348$. Тобто модифікована модель має вигляд: $y^* = 1{,}407 + 0{,}348x^*$. Далі виконується зворотний перехід до нелінійної форми моделі:

$$a_0^* = ln a_0 \rightarrow a_0 = e^{a_0^*} = e^{1{,}407} = 4{,}086.$$

Таким чином, степенева модель залежності прибутку підприємства від рівня витрат на захист інтелектуальної власності має вигляд:

$$y = 4{,}086 \cdot x^{0{,}3489}.$$

Графік моделі з коефіцієнтом детермінації наведено на рис. 6.10.

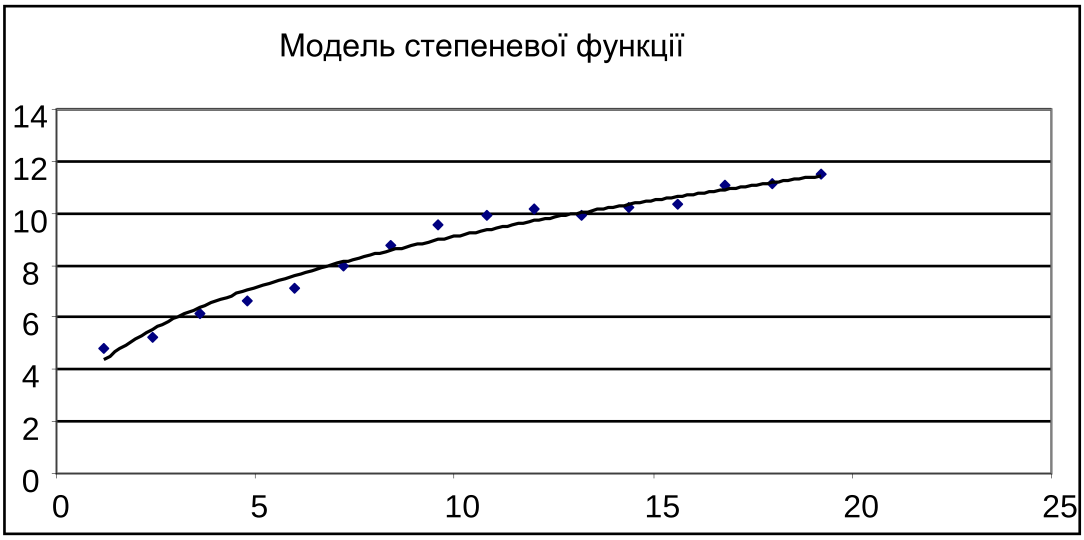


Рис. 6.10. **Модель степеневої функції**

Для побудованої моделі коефіцієнт детермінації становить 0,97, що дозволяє зробити висновок про адекватність моделі, високу точність прогнозу й можливість застосування моделі для аналізу й прогнозування досліджуваних процесів.

6.2. Еластичність функцій однієї та багатьох змінних

Для дослідження економічних об'єктів та процесів використовують диференціальне числення. Такі дослідження характеризують не стан (як середнє значення чи сумарна величина), а процеси зміни об'єкта.

Таким чином, похідна виступає як швидкість зміни деякого економічного об'єкта за часом або відносно іншого об'єкта. В економіці не завжди можна використовувати диференціальне числення в зв'язку з дискретністю економічних показників у часі (річні, квартальні та ін.). Одним із прикладів використання диференціального числення є еластичність [37].

Еластичністю функції $E_x(y)$ називається границя відношення відносного приросту функції y до відносного приросту змінної x (при $\Delta x \to 0$)

$$E_x(y) = \lim_{\Delta x \to 0}\left(\frac{\Delta y}{y} : \frac{\Delta x}{x}\right) = \frac{x}{y} \cdot y'. \quad (6.5)$$

Еластичність функції наближено відображає, на скільки відсотків зміниться функція $y = f(x)$ у разі зміни незалежної змінної x на 1 %.

Властивості еластичності функції:

1. Еластичність функції дорівнює добутку незалежної змінної на темп зміни функції T_y, тобто $E_x(y) = x \cdot T_y$.

2. Еластичність добутку (частки) двох функцій дорівнює сумі (різниці) еластичностей цих функцій: $E_x(uv) = E_x(u) + E_x(v)$, $E_x\left(\frac{u}{v}\right) = E_x(u) - E_x(v)$.

3. Еластичність взаємообернених величин – взаємообернені величини: $E_x(y) = \frac{1}{E_x(y)}$.

Еластичність функції застосовується під час аналізу попиту та пропозиції. Наприклад, еластичність попиту відносно ціни (або доходу) – це коефіцієнт, що наближено відображає, на скільки відсотків зміниться попит (обсяг пропозиції) у процесі зміни ціни (або доходу) на 1 %.

Якщо функція залежить від декількох змінних, то під час розрахунків потрібно вказувати, за якою змінною обчислюється еластичність. Якщо розглядати попит як функцію двох змінних (ціна товару та дохід покупця) $Q_d = f(P, I)$, то можна обчислити дві еластичності: еластичність попиту за ціною $E_P(f(P, I)) = f_P'(P, I)\frac{P}{f(P,I)}$ і еластичність попиту за доходом $E_I(f(P, I)) = f_P'(P, I)\frac{I}{f(P,I)}$.

Для змістовного економічного аналізу результатів побудованих моделей досить поширеним є розрахунок таких характеристик, як середня та гранична продуктивність та еластичність. Економічні характеристики деяких нелінійних функцій наведені в табл. 6.3.

Таблиця 6.3

Економічні характеристики деяких нелінійних функцій

Функції	Рівняння $Y = f(x)$	Середня продуктивність $\frac{y}{x}$	Гранична продуктивність $\frac{\Delta y}{\Delta x}$	Коефіцієнт еластичності $E = \frac{\partial y}{\partial x} \cdot \frac{x}{y}$
Лінійна	$y = a_0 + a_1x_1$	$\frac{a_0}{x} + a_1$	a_1	$\frac{a_1x}{a_0 + a_1x}$
Квадратична	$y = a_0 + a_1x - a_2x^2$	$\frac{a_0}{x} + a_1 - a_2x$	$a_1 - 2a_2x$	$\frac{a_1 - 2a_2x \;\; x}{a_0 + a_1x - a_2x^2}$
Кубічна	$y = a_0 + a_1x + a_2x^2 - a_3x^3$	$\frac{a_0}{x} + a_1 - a_2x - a_3x^2$	$a_1 + 2a_2x - 3a_3x^2$	$\frac{a_1 + 2a_2x - 3a_3x^2 \;\; x}{a_0 + a_1x + a_2x^2 - a_3x^3}$
Гіперболічна	$y = a_0 + \frac{a_1}{x}$	$\frac{a_0}{x} + \frac{a_1}{x_2}$	$-\frac{a_1}{x^2}$	$-\frac{a_1}{a_0x + a_1}$
Степенева	$y = a_0x^{a_1}$	$a_0x^{a_1-1}$	$a_0a_1x^{a_1-1}$	a_1
Показникова	$y = a_0 - ka_1^x$	$\frac{a_0 - ka_1^x}{x}$	$-ka_1^x lna_1$	$-\frac{ka_1^x lna_1}{a_0 - ka_1^x}$
Експоненційна	$y = a_0e^{a_1x}$	$\frac{a_0e^{a_1x}}{x}$	$a_0a_1e^{a_1x}$	a_1x

6.3. Виробничі функції, їх класифікація та основні властивості

У сфері виробництва під час аналізу кількісного співвідношення показника і чинників у ролі показника можуть виступати: об'єм продукції, що випускається, прибуток, товарообіг, рентабельність, собівартість одиниці продукції фондовіддача й ін. Чинниками, які впливають на перераховані показники, можуть виступати: робоча сила, продуктивність суспільної праці, рівень розвитку науки і техніки, утворення і тому подібне.

Виробнича функція – це функція, незалежна змінна якої набуває значення об'ємів ресурсу (чинник виробництва), що витрачається, а залежна змінна – значення об'ємів продукції, що випускається [15; 40]:

$$y = f\ \ x, a\ \ . \tag{6.6}$$

У формулі (6.6) $x\ \ x \geq 0$ і $y\ \ y \geq 0$ – числові величини, a – вектор параметрів моделі, тобто Y є функцією від однієї змінної X. У зв'язку з цим виробнича функція (ВФ) називається однoресурсною або однофакторною ВФ і область її визначення – множина невід'ємних чисел. Запис $Y = f(x)$ означає, що якщо ресурс витрачається або використовується в кількості X одиниць, то продукція випускається в кількості $Y = f(x)$ одиниць. На рис. 6.11 наведений приклад однофакторної виробничої функції.

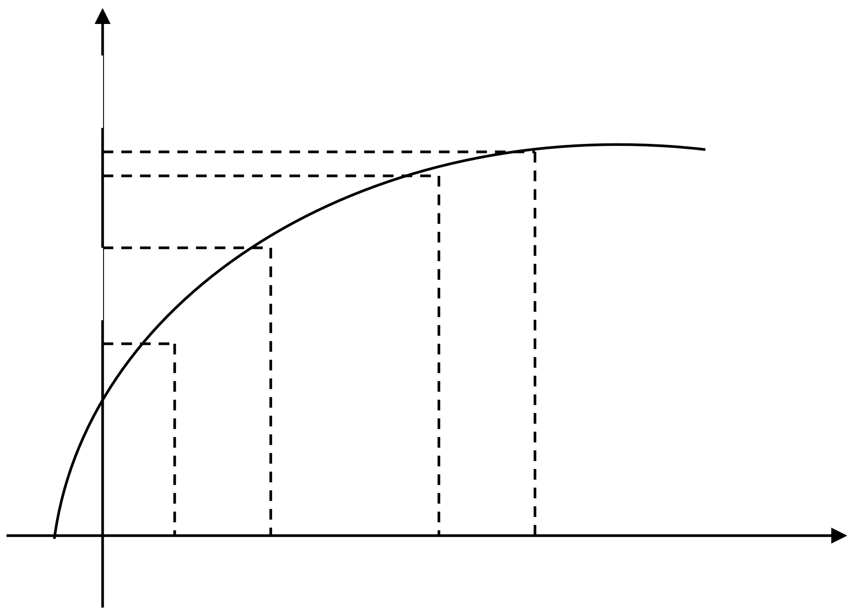


Рис. 6.11. **Приклад однофакторної виробничої функції**

На графіку (рис. 6.11) видно, що зі зростанням величини ресурсу, що витрачається, X об'єм випуску Y зростає, проте у ході цього кожна додаткова одиниця ресурсу дає все менший приріст об'єму продукції, що випускається. Відмічену обставину відображає фундаментальне положення економічної теорії як закон спадаючої ефективності.

ВФ можуть мати різні області використання. Принцип "витрати-випуск" знаходить свою реалізацію як на мікро-, так і на макроекономічному рівні. На мікроекономічному рівні в ролі виробничої системи виступає окреме підприємство або фірма, на макроекономічному рівні – галузь, міжгалузевий комплекс, регіон і т. д. Крім того, під час побудови ВФ можуть бути використані як часові ряди *(time series),* так і дані просторового типу *(cross-section data)* [12].

Точне тлумачення понять ресурсу, що витрачається (використовуваного), і продукції, що випускається, а також вибір одиниць їх вимірювання залежать від:

- характеру і масштабу виробничої системи;
- особливостей вирішуваних завдань;
- наявності початкових умов.

На мікроекономічному рівні витрати і випуск можуть вимірюватися як у натуральних, так і у вартісних показниках. На макроекономічному рівні витрати і випуск вимірюються, як правило, у вартісних показниках і є вартісними (ціннісні) агрегатами, тобто сумарні величини добутків об'ємів ресурсів, що витрачаються, і продуктів, що випускаються, на їх ціни.

Виробнича функція декількох змінних – це функція, незалежні змінні якої набувають значення обсягів ресурсів, що витрачаються, а значення залежної змінної характеризує обсяг випуску продукції:

$$Y = F \ x,\ a \ = \quad x_1, x_2, \ldots, x_n,\ a \ . \tag{6.7}$$

У формулі (6.7) $y \quad y \geq 0$ – скалярна величина, $x,\ a$ – векторні величини.

Відповідно до законів економічної теорії $x_1, x_2, \ldots, x_n \geq 0$, отже, областю визначення багатофакторної ВФ є множина n-вимірних векторів X, всі координати $x_1, x_2, \ldots, x_n$ яких – додатні числа.

Для окремого підприємства (фірми), що випускає однорідний продукт, ВФ $f(x_1, x_2, \dots, x_n, a)$ може зв'язувати об'єм випуску з витратами робочого часу за різними видами трудової діяльності, різними видами сировини, комплектуючими виробами, енергією, основним капіталом. ВФ такого типу характеризують *технологію підприємства, що діє*.

Під час побудови ВФ для регіону або країни в цілому як величина річного випуску Y частіше беруть сукупний продукт регіону або країни, що обчислюється зазвичай у незмінних цінах, а не в поточних, а як ресурси розглядають основний капітал K, живий труд – L, об'єм використовуваних природних ресурсів, технічний прогрес. У зв'язку з цим, отримують *дво-, три- та багатофакторні ВФ* [11].

За врахуванням фактора часу виділяють:

- статичну ВФ, параметри і змінні якої не залежать від часу:

$$y = f(x_1, x_2, \dots, x_n, a);$$

- динамічну ВФ, змінні і параметри якої залежать від часу, а чинник часу включений у модель як незалежна величина:

$$y(t) = f(x_1(t), x_2(t), \dots, x_n(t), a(t)).$$

ВФ може мати лінійний і нелінійний характер.

Лінійна адитивна виробнича функція може бути наведена таким чином:

$$Y = a_0 + a_1 x_1 + a_2 x_2 + \dots + a_n x_n,$$

а функція вигляду

$$Y = a_0 x_1^{a_1} x_2^{a_2}$$

є нелінійною мультиплікативною виробничою функцією.

Виділення суттєвих видів ресурсів (чинників виробництва) і вибір аналітичної форми функції називається *специфікацією ВФ.* Вибір аналітичної форми ВФ ґрунтується, перш за все, на теоретичних міркуваннях, які повинні явно враховувати особливості взаємозв'язку між конкретними ресурсами (у разі мікроекономічного рівня) або економічними закономірностями (у разі макроекономічного рівня).

Перетворення реальних і експертних даних у модельну інформацію, тобто розрахунок чисельних значень параметрів ВФ на базі статистичних даних за допомогою регресійного або кореляційного аналізу, називається *параметризацією ВФ* [11].

Перевірка істинності (адекватності) виробничої функції називається її *верифікацією.* На специфікацію і параметризацію в процесі вдосконалення ВФ впливають результати верифікації ВФ. Необхідно зазначити, що оцінювання параметрів ВФ зазвичай проводиться за допомогою методу найменших квадратів (МНК). Слід розглянути основні властивості ВФ (табл. 6.4).

Таблиця 6.4

Основні властивості виробничої функції

№ п/п	Аналітичне представлення	Змістовна інтерпретація
1	$f(0,0) = 0$	Основна властивість означає, що без витрат ресурсів неможливий випуск продукції
1а	$f(x_1, 0) = f(0, x_2) = 0$	Дана властивість характеризує той факт, що неможливо отримати випуск продукції у разі відсутності хоча б одного ресурсу
2	$f\ 1 \geq f(0)$	Властивість означає, що зі зростанням витрат хоча б одного ресурсу об'єм випуску зростає
2а	$x > 0 \Rightarrow \dfrac{\partial f(x)}{\partial x_i} > 0$ $(i = 1, 2)$	Дана властивість підтверджує, що із зростанням витрат хоча б одного ресурсу об'єм випуску зростає, оскільки перша частинна похідна позитивна
3	$x > 0 \Rightarrow \dfrac{\partial^2 f(x)}{\partial x_i^2} \leq 0$ $(i = 1, 2)$	Властивість означає, що зі зростанням витрат одного ресурсу у разі незмінної кількості іншого ресурсу величина приросту випуску продукції на кожну додаткову одиницю будь-якого ресурсу не зростає (закон спадаючої ефективності)
3а	$x > 0 \Rightarrow \dfrac{\partial^2 f(x)}{\partial x_1 \partial x_2} \geq 0$ $x = (x_1, x_2)$	Ця властивість означає, що у ході зростання одного ресурсу гранична ефективність іншого ресурсу зростає
4	$f(tx_1, tx_2) = t^p f(x_1, x_2)$	Дана властивість означає, що ВФ є однорідною функцією ступеня $p > 0$.За умови $p > 1$ зі зростанням масштабу виробництва в t разів, тобто з переходом від вектора X до вектора tx, об'єм випуску продукції зростає в t^p разів, оскільки спостерігається зростання ефективності виробництва від зростання масштабу виробництва. За умови $p = 1$ є постійна ефективність виробництва у ході зростання його масштабу

Необхідно зазначити, що для нелінійної двофакторної виробничої функції всі перераховані властивості справедливі, тоді як для лінійної ВФ властивість 4 не виконується.

6.4. Виробнича функція Кобба – Дугласа, особливості побудови та аналізу

Одним із найширше використовуваних для моделювання результатів розвитку окремого регіону або країни є функція Кобба – Дугласа. Виходячи з наведених класифікацій виробничих функцій (ВФ), вона належить до мультиплікативних, багатофакторних, нелінійних ВФ і має вигляд:

$$Y = a_0 \cdot L^{a_1} \cdot K^{a_2},$$

де Y – об'єм випуску або доходу;

$L\ (X_1)$ – кількість одиниць витраченоЇ праці за певний період часу;

$K\ (X_2)$ – об'єм використаного за деякий період основного капіталу;

$a_0,\ a_1,\ a_2$ – параметри моделі [15; 40].

Для того, щоб провести параметризацію моделі зазвичай переходять від мультиплікативної її форми до аддитивної за допомогою операції логарифмування. Для ВФ Кобба – Дугласа наявні такі перетворення:

$$y = a_0 x_1^{a_1} x_2^{a_2};$$
$$lny = w,\ \ lnx_1 = v_1,\ \ lnx_2 = v_2;$$
$$w = lna_0 + a_1 v_1 + a_2 v_2.$$

У лінійному вигляді завдання знаходження параметрів може бути вирішена або за допомогою методу найменших квадратів (МНК), або в матричному вигляді.

Параметри ВФ розраховувалися різними авторами за різними методиками, тому вони мають певні відмінності, однак можна виділити, що у всіх авторів параметр a_2 вище за параметр a_1. Також цікавий той факт, що майже у всіх авторів сума параметрів виявилася близькою до одиниці.

Одним з істотних недоліків наведеного виду функції Кобба – Дугласа є те, що вона не відображає вплив на зростання об'єму виробництва прогресивних змін у складі робочої сили й організації виробництва. Технологічний прогрес як важливу рушійну силу економічного розвитку не можна, очевидно, ігнорувати під час побудови моделей суспільного виробництва. Проте технічний прогрес – явище складне і багатогранне й адекватно відобразити його за допомогою однієї або декількох змінних величин надзвичайно важко. Тому як "перше наближення" у виробничі функції дослідники використовують експоненціальну тенденцію, залежну від часу як показник впливу науково-технічного прогресу. Для цього вводиться множник, який характеризує НТП e^{pt}, де параметр $p\ (p > 0)$ характеризує темп приросту випуску продукції з урахуванням впливу НТП. Тоді ВФ Кобба – Дугласа має вигляд [15]:

$$y(t) = e^{pt} f(x_1(t), x_2(t))(t = 0, 1, \dots , \qquad T).$$

Подана ВФ – простий приклад динамічної ВФ, оскільки вона включає нейтральний, тобто не матеріалізований ні в одному з чинників виробництва науково-технічний прогрес. Науково-технічний прогрес є нейтральним у тому сенсі, що він не змінює відносну ефективність обох ресурсів, підвищуючи віддачу кожного з них рівною мірою. У складніших випадках технічний прогрес може впливати безпосередньо на продуктивність праці або капіталовіддачу:

$$Y\ t\ = f\ A\ t\ ,\ L\ t\ ,\ K\ t \quad \text{або } Y(t) = f(A\ t\ ,\ K\ t\ ,\ L(t)).$$

У даному випадку він є відповідно *працезберігаючим* або *капіталозберігаючим* НТП.

Приклад 6.1. Слід розглянути приклад побудови, оцінювання параметрів та адекватності виробничої функції Кобба – Дугласа. Відомі такі вихідні дані обсягу виробленої продукції залежно від витрат трудових ресурсів та вкладеного капіталу для 15 вибіркових регіонів. Вихідні дані наведені в табл. 6.5.

Таблиця 6.5

Вихідні дані

Регіони	Трудові ресурси (тис. осіб) (L, X_1)	Об'єм капіталу (тис. грн) (K, X_{21})	Обсяг виробленої продукції (млн грн) Y
1	2	3	4
1	16,88	21,84	18,22
2	16,05	22,71	17,53
3	14,74	20,87	16,35
4	15,8	25,68	17,94

Закінчення табл. 6.5

1	2	3	4
5	17,72	20,9	19,24
6	15,29	26,68	17,66
7	16,2	22,99	18,21
8	16,81	23,68	18,44
9	17,68	25,54	19,76
10	16,45	22,2	17,85
11	17,17	26,02	19,27
12	17,9	27,94	20,93
13	16,74	26,6	19,04
14	14,53	24,87	16,75
15	17,2	25,47	19,56

Y як результативний фактор ВФ слід подати у вигляді такої мультиплікативної нелінійної функції від факторів виробництва:

$$Y = a_0 \cdot X_1^{a_1} \cdot X_2^{a_2},$$

де $a_0,\ a_1,\ a_2$ – параметри моделі, які потрібно оцінити.

Оскільки вихідна залежність $Y = (X_1,\ X_2)$ є нелінійною функцією, то щоб оцінити параметри моделі, необхідно за допомогою перетворення лінеаризувати дану функцію, тобто привести її до лінійного вигляду шляхом заміни змінних.

$$lnY = ln\ \ A_0 \cdot X_1^{a_1} \cdot X_2^{a_2}\ \ = lnA_0 + a_1 \cdot lnX_1 + a_2 \cdot lnX_2.$$

Перейдемо до нових змінних:

$$lnY = Z, \qquad lnA_0 = a_0, lnX_1 = Z_1, \qquad lnX_2 = Z_2.$$

$$Z = a_0 + a_1 \cdot Z_1 + a_2 \cdot Z_2,$$

де Z – результативна ознака;

Z_1, Z_2 – незалежні ознаки;

відповідно $a_0,\ a_1,\ a_2$ – параметри моделі.

Оскільки перетворена залежність Z є лінійною від $Z_1 Z_2$, то для оцінювання параметрів можна застосувати метод найменших квадратів (МНК). Значення лінеаризованих змінних наведено в табл. 6.6.

Таблиця 6.6

Значення лінеаризованих змінних

Регіони	$Z_1 = lnX_1$	$Z_2 = lnX_2$	$Z = lnY$
1	2,826	3,084	2,903
2	2,776	3,123	2,864
3	2,691	3,038	2,794
4	2,760	3,246	2,887
5	2,875	3,040	2,957
6	2,727	3,284	2,871
7	2,785	3,135	2,902
8	2,822	3,165	2,915
9	2,872	3,240	2,984
10	2,800	3,100	2,882
11	2,843	3,259	2,959
12	2,885	3,330	3,041
13	2,818	3,281	2,947
14	2,676	3,214	2,818
15	2,845	3,238	2,973

Система нормальних рівнянь для визначення невідомих параметрів моделі a_0, a_1, a_2 має такий вигляд:

$$10 \cdot a_0 + a_1 \cdot \sum_{i=1}^{n} Z_{i_1} + a_2 \cdot \sum_{i=1}^{n} Z_{i_2} = \sum_{i=1}^{n} Z_i \,;$$

$$a_0 \cdot \sum_{i=1}^{n} Z_{i_1} + a_1 \cdot \sum_{i=1}^{n} (Z_{i_1})^2 + a_2 \cdot \sum_{i=1}^{n} Z_{i_2} \cdot Z_{i_1} = \sum_{i=1}^{n} Z_i \cdot Z_{i_1};$$

$$a_0 \cdot \sum_{i=1}^{n} Z_{i_2} + a_{21} \cdot \sum_{i=1}^{n} Z_{i_1} \cdot Z_{i_2} + a_2 \cdot \sum_{i=1}^{n} (Z_{i_2})^2 = \sum_{i=1}^{n} Za \cdot Z_{i_2}.$$

Проміжні розрахунки для визначення параметрів моделі наведено в табл. 6.7.

Таблиця 6.7

Проміжні розрахунки для визначення параметрів моделі

Регіони	$(Z_{i_1})^2$	$(Z_{i_2})^2$	$Z_{i_1}Z_{i_2}$	$Z_iZ_{i_1}$	$Z_iZ_{i_2}$
1	7,987	9,509	8,715	8,203	8,951
2	7,705	9,752	8,668	7,949	8,943
3	7,239	9,231	8,175	7,518	8,490
4	7,618	10,535	8,958	7,968	9,370
5	8,264	9,240	8,738	8,500	8,989
6	7,438	10,784	8,956	7,831	9,429
7	7,756	9,829	8,731	8,082	9,098
8	7,964	10,015	8,931	8,225	9,223
9	8,251	10,499	9,307	8,570	9,668
10	7,842	9,611	8,681	8,071	8,934
11	8,084	10,620	9,265	8,412	9,642
12	8,322	11,089	9,607	8,773	10,127
13	7,940	10,764	9,245	8,303	9,667
14	7,162	10,328	8,600	7,543	9,057
15	8,094	10,481	9,210	8,459	9,627
Сума	117,664	152,288	133,788	122,407	139,216

Отже, отримані параметри моделі мають такі значення:

a_0 = *-0,289;* $A_0 = e^{a_0}$ = *0,749;* a_1 = *0,872;* a_2 = *0,239.*

Таким чином, побудована функція Кобба – Дугласа буде мати такий вигляд: $Y = 0{,}749 \cdot X_1^{0{,}872} \cdot X_2^{0{,}239}$.

Слід провести оцінювання адекватності моделі на основі коефіцієнта детермінації та критерію Фішера. Розрахунки для оцінювання адекватності наведені в табл. 6.8.

Таблиця 6.8

Розрахунки для оцінювання адекватності моделі

Регіони	$Y_{\imath}$	$Y_{\imath} - Y_i$	$Y_{\imath} - Y_i\ ^2$	$Y_i - Y_{\imath}\ ^2$	$Y_{\imath} - Y_{\imath}\ ^2$
1	2	3	4	5	6
1	2,911626	-0,009106	8,2919E-05	0,0001	2,13E-06
2	2,876984	-0,01307	0,00017084	0,0024	0,001303
3	2,782541	0,011687	0,00013658	0,0141	0,017042

Закінчення табл. 6.8

1	2	3	4	5	6
4	2,892656	-0,005623	3,1615E-05	0,0007	0,000417
5	2,94347	0,013521	0,00018282	0,0019	0,000923
6	2,873166	-0,001864	3,4755E-06	0,0017	0,001594
7	2,888025	0,013946	0,0001945	0,0001	0,000628
8	2,927326	-0,012804	0,00016394	0,0000	0,000203
9	2,989399	-0,005739	3,2941E-05	0,0050	0,005824
10	2,893027	-0,011024	0,00012152	0,0010	0,000402
11	2,96832	-0,00977	9,5454E-05	0,0021	0,003051
12	3,021641	0,019542	0,0003819	0,0164	0,011784
13	2,951467	-0,004925	2,4252E-05	0,0011	0,001473
14	2,811918	0,00648	4,1991E-05	0,0090	0,010235
15	2,964738	0,008749	7,6538E-05	0,0036	0,002668
Сума	2,913		**0,00174127**	**0,05929**	**0,057549**

Треба розрахувати коефіцієнт детермінації моделі:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} \left(Y_i - \hat{Y}_i\right)^2}{\sum_{i=1}^{n} \left(Y_i - \bar{Y}\right)^2} = 1 - \frac{0{,}00174}{0{,}0593} = 0{,}971.$$

Для оцінювання статистичної значущості моделі слід застосувати критерій Фішера:

$$F = \frac{\dfrac{\sum_{i=1}^{n} \left(\hat{Y}_i - \bar{Y}\right)^2}{m}}{\dfrac{\sum_{i=1}^{n} \left(Y_i - \hat{Y}_i\right)^2}{n-m-1}} = \frac{0{,}0575}{2} \Big/ \frac{0{,}00174}{7} = 198{,}3.$$

Отже, високий коефіцієнт детермінації, який наближається до одиниці, та досить значне значення критерію Фішера, яке перевищує табличне значення, свідчить про значну адекватність та статистичну значущість моделі, що підтверджує правильність обраного взаємозв'язку між досліджуваними змінними. Таким чином, отримані результати можуть бути використані для ґрунтовної інтерпретації між досліджуваними чинниками та прогнозування тенденцій розвитку.

6.5. Основні характеристики виробничих функцій, їх геометрична та економічна інтерпретація

Для застосування моделей виробничих функцій в практиці управління на різних рівнях ієрархії дослідники все більш опираються на визначення основних характеристик, які є ґрунтовною основою розробки та прийняття управлінських рішень [11; 15; 27; 40]. Слід розглянути процедуру їх оцінювання, економічну та геометричну інтерпретацію.

1. *Середня продуктивність ресурсів*.

Середня продуктивність праці показує середню кількість продукції на одиницю витраченої праці, розраховується за формулою:

$$A_1 = \frac{y}{x_1} = a_0 \cdot x_1^{a_1 - 1} \cdot x_2^{a_2}.$$

Середня фондовіддача (капіталовіддача), показує обсяг продукції в розрахунку на одиницю використаних виробничих фондів:

$$A_2 = \frac{y}{x_1} = a_0 \cdot x_1^{a_1} \cdot x_2^{a_2 - 1}.$$

Геометричний зміст даної характеристики – середня продуктивність ресурсу дорівнює тангенсу кута нахилу хорди, проведеної до осі абсцис (рис. 6.12), тобто $A = tga, \alpha = arctgA$.

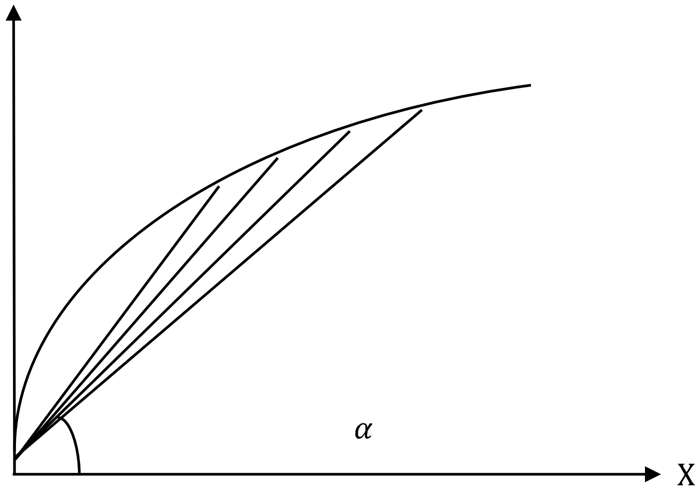


Рис. 6.12. **Графік середньої продуктивності ресурсу**

2. *Гранична (маржинальна) продуктивність ресурсів*.

Гранична продуктивність праці показує скільки додаткових одиниць продукції приносить додаткова одиниця витраченої праці, розраховується відповідно до формули:

$$M_1 = \frac{\partial y}{\partial x_1} = a_0 \cdot a_1 \cdot x_1^{a_1-1} \cdot x_2^{a_2}.$$

Гранична фондовіддача (капіталовіддача) показує, скільки додаткових одиниць продукції приносить додаткова одиниця основних фондів.

$$M_2 = \frac{\partial y}{\partial x_2} = a_0 \cdot a_2 \cdot x_1^{a_1} \cdot x_2^{a_2-1}.$$

Геометричний зміст даної характеристики – гранична ефективність (продуктивність) ресурсу дорівнює тангенсу кута нахилу дотичної, проведеної до графіка ВФ, у точці x_0 до осі абсцис (рис. 6.13), тобто

$$M = tg\varphi, \varphi = arctgM.$$

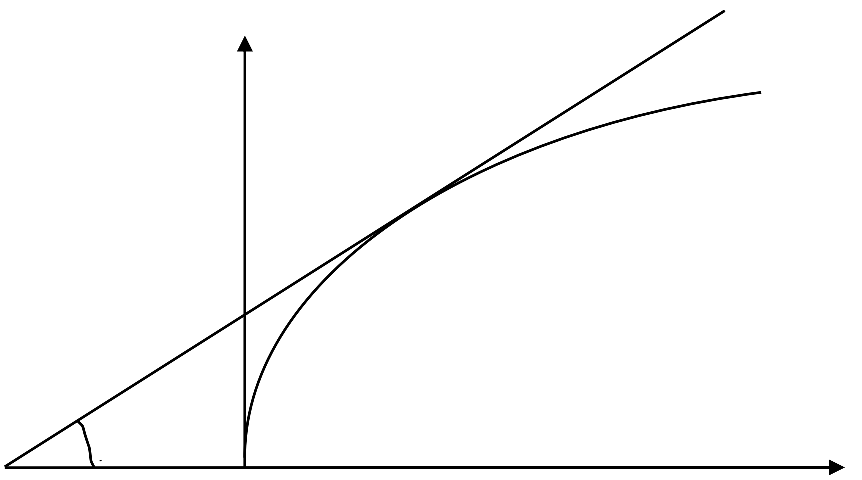


Рис. 6.13. **Графік граничної ефективності ресурсу**

3. *Еластичність випуску продукції* за факторами виробництва.

Еластичність випуску продукції за витратами праці показує, на скільком відсотків збільшиться випуск продукції під час збільшення витрат праці на 1 %, може бути визначена таким чином:

$$E_1 = \frac{\partial y}{\partial x_1} \cdot \frac{x_1}{y}; \qquad E_1 = \frac{M_1}{A_1}; \qquad E_1 = a_1.$$

Еластичність випуску продукції за витратами виробничих фондів показує, на скількох відсотків збільшиться випуск продукції у ході збільшення основних фондів на 1 %.

$$E_2 = \frac{\partial y}{\partial x_2} \cdot \frac{x_2}{y}; \qquad E_2 = \frac{M_2}{A_2}; \qquad E_2 = a_2.$$

Сумарна еластичність за витратами (праці і капіталу). Сумарна еластичність за витратами показує ефект одночасного пропорційного збільшення обсягу ресурсів праці й основних фондів.

$$E = E_1 + E_2 = a_1 + a_2.$$

Еластичність ВФ у точці $C(x_0,\ y_0)$ по модулю дорівнює відношенню відстаней по дотичній від точці C з координатами $(x_0,\ f(x_0))$ до точки перетину з осями Y і X (рис. 6.14).

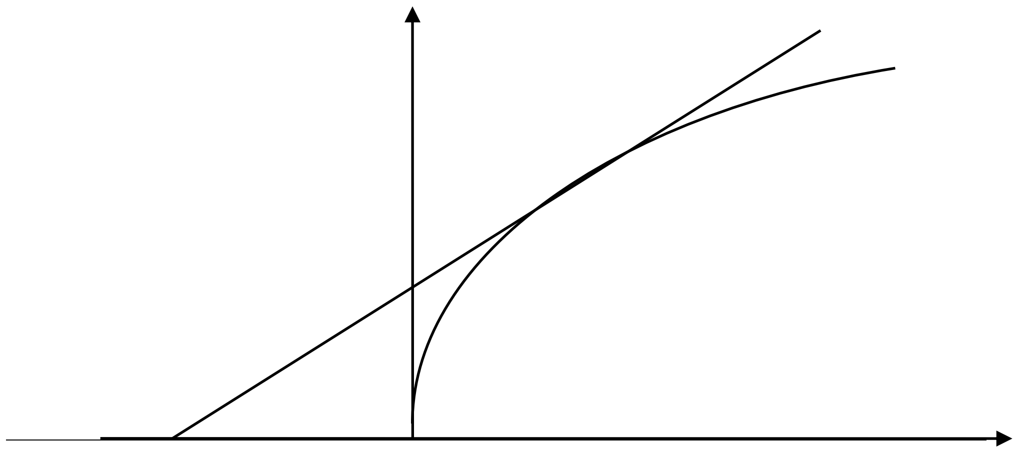


Рис. 6.14. **Графік еластичності виробничої функції**

4. *Фондоозброєність (капіталоозброєність) праці* показує, скільки у середньому капіталу доводиться на одиницю затрачуваної праці і розраховується за формулою:

$$FT = \frac{x_2}{x_1} = a_0^{-\frac{1}{a_2}} \cdot y_0^{\frac{1}{a_2}} \cdot x_1^{-1-\frac{a_1}{a_2}}.$$

На основі даних співвідношень можна дослідити залежність між продуктивністю праці та її капіталоозброєністю. Якщо сума показників у ВФ Кобба – Дугласа $Y = a_0 \cdot {X_1}^{a1} \cdot {X_2}^{a2}$ дорівнює одиниці $(a_1 + a_2 = 1)$, то:

$$\frac{y}{x_1} = \frac{a_0 \cdot x_1^{a_1} \cdot x_2^{a_2}}{x_1} = \frac{a_0 \cdot x_2^{a_2}}{x_1^{1-a_1}} = a_0 \cdot \left(\frac{x_2}{x_1}\right)^{a_2},$$

або якщо перейти до нових позначень

$$z = \frac{Y}{x_1},\ \ k = \frac{x_2}{x_1},\ \text{то}$$

буде одержано таку залежність:

$$Z = a_0 \cdot K^{a_2}.$$

Оскільки $0 < a_2 < 1$, то з формули випливає, що продуктивність праці Z зростає повільніше, ніж капіталоозброєність.

5. *Фондомісткість продукції* показує, який обсяг капіталу витрачається на одиницю випуску продукції і розраховується за формулою:

$$F = \frac{x_2}{Y} = \frac{1}{a_0}\left(\frac{x_2}{x_1}\right)^{a_1}.$$

Із наведених співвідношень видно, що вплив зростання фондоозброєності праці на продуктивність праці і фондомісткість пов'язаний з величиною параметра a_1. Якщо $a_1 < 0{,}5$, то зі зростанням фондоозброєності порівняно швидко зростає продуктивність праці і досить повільно збільшується фондомісткість продукції. За умови $a_1 > 0{,}5$ ситуація змінюється: зростання фондоозброєності праці досить різко збільшує фондомісткість, а темпи зростання продуктивності праці спадають. Так, якщо $a_1 = 0{,}6$ збільшенню фондоозброєності праці на 10 % відповідає зростанню фондомісткості на 6 %, а продуктивності праці – лише на 4 %.

6. *Потреби у витратах ресурсів.* Виробнича функція дозволяє розрахувати потребу в одному з ресурсів у разі заданого обсягу виробництва й величини іншого ресурсу.

$$X_1 = \frac{Y}{a_0 \cdot x_2^{a_2}}^{\frac{1}{a_1}}$$

– потреба у витратах праці за умови відомих значень обсягу випуску і витрат капіталу.

$$X_2 = \frac{Y}{a_0 \cdot x_1^{a_1}}^{\frac{1}{a_2}}$$

– потреба у витратах капіталу за умови відомих значеннях обсягу випуску і витратах праці.

7. *Ізокванти виробничої функції*. Ізокванти ВФ (криві рівного випуску продукту, криві байдужості для виробників) – це лінії рівня $q = f\ \ x_1, x_2\ $, $(q > 0)$, що становлять множину точок, у яких ВФ набуває значення, рівного q. Ізокванти є різними наборами (співвідношення) використовуваних ресурсів, що забезпечують однаковий обсяг випуску продукції. Графік ізоквант виробничої функції наведений на рис. 6.15.

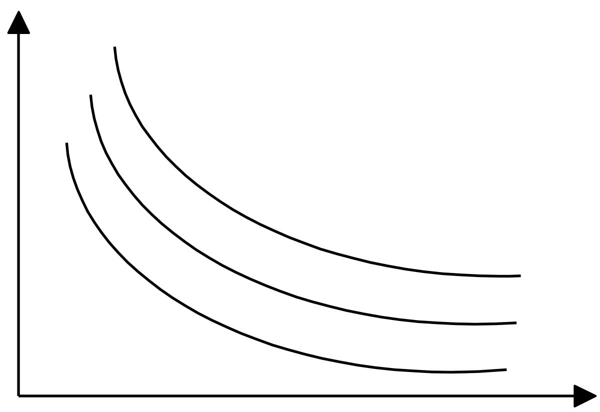


Рис. 6.15. **Ізокванти виробничої функції**

Ізокванти мають такі властивості:

1. Ізокванти не можуть перетинатися.

2. Кожна наступна ізокванта, що проходить далі від початку координат, відображає більшу величину випуску, ніж попередня. Сукупність ізоквант утворює карту ізоквант.

3. Ізокванти мають негативний нахил.

4. Гранична норма заміщення одного ресурсу іншим зменшується під час руху вздовж ізокванти.

5. Ізокванти увігнуті щодо початку координат.

6. Графічне відображення ізоквант визначається поєднанням взаємозамінності і взаємодоповнюваності ресурсів, що використовуються.

7. У разі жорсткої взаємозамінності ресурсів (досконалої субституції), ізокванта набуває лінійного виду. У разі жорсткої взаємодоповнюваності ресурсів (комплементарності), ізокванта зводиться до точки.

8. *Ізокоста* – лінія, що демонструє комбінації факторів виробництва, які можна купити за однакову загальну суму грошей (рис. 6.16).

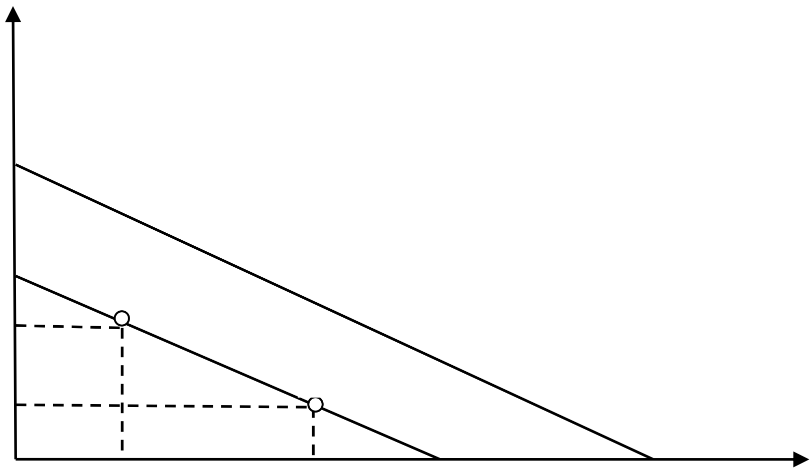


Рис. 6.16. **Графік ізокост (TC_1, TC_2)**

Ізокосту інакше називають лінією рівних витрат. Ізокости є паралельними прямими, оскільки допускається, що фірма може придбати будь-яку бажану кількість факторів виробництва за незмінним цінами. Нахил ізокости виражає відносні ціни факторів виробництва. Кожна точка на лінії ізокости характеризується одними і тими ж загальними витратами. Ці лінії прямі, оскільки факторні ціни мають негативний нахил і паралельні.

Сумістивши ізокванти та ізокости, можна визначити оптимальну позицію фірми. Точка, в якій ізокванта дотикається (але не перетинає) ізокосту, означає найбільш дешеву за вартістю комбінацію чинників, необхідних для випуску певного обсягу продукції.

9. *Ізокліналь* виробничої функції. Для заданої ізокванти можна побудувати ізокліналь ВФ – це лінія, яка з'єднує початок координат і точки на ізоквантах ВФ, для яких рівними будуть граничні норми заміщення ресурсів (рис. 6.17).

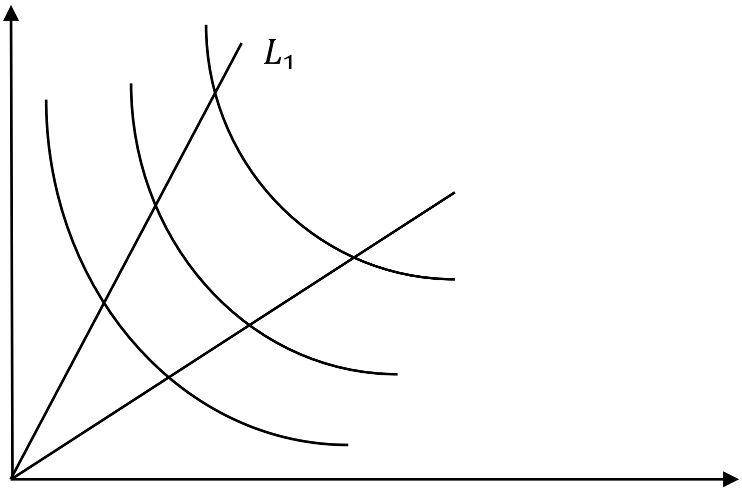


Рис. 6.17. **Графік ізокліналей (L_1, L_2) ВФ Кобба – Дугласа**

10. *Гранична норма заміни i-го ресурсу j-м ресурсом*.

Гранична норма заміни розраховується таким чином:

$$R_{ij} = -\frac{\Delta x_j}{\Delta x_i}, \qquad R'_{ij} = -\frac{\partial x_j}{\partial x_i}.$$

Гранична норма заміни ресурсів R_{ij}показує, на скільки одиниць збільшуються витрати j-го ресурсу (у разі незмінного фіксованого випуску продукції), якщо витрати i-го ресурсу зменшаться на одну одиницю (i – номер *ресурсу, що заміняється*, а j – номер *ресурсу, яким заміняють*). Для двофакторної виробничої функції справедлива рівність:

$$R_{12} = \frac{E_1}{E_2} \cdot \frac{x_2}{x_1}.$$

Гранична норма заміни ресурсу збігається з тангенсом кута нахилу f до осі абсцис дотичної, проведеної до ізокванти ВФ у точці $(x_0, f(x_0))$. (рис. 6.18).

Якщо $x_1 = L$ (де L – витрати праці), а $x_2 = K$ (де K – витрати капіталу), то відношення

$$\frac{x_2}{x_1} = \frac{K}{L}$$

називається *капіталоозброєністю праці*. В цьому випадку гранична норма заміщення основного капіталу працею дорівнює відношенню еластичності випуску за капіталом і працею, поділеному на капіталоозброєність праці.

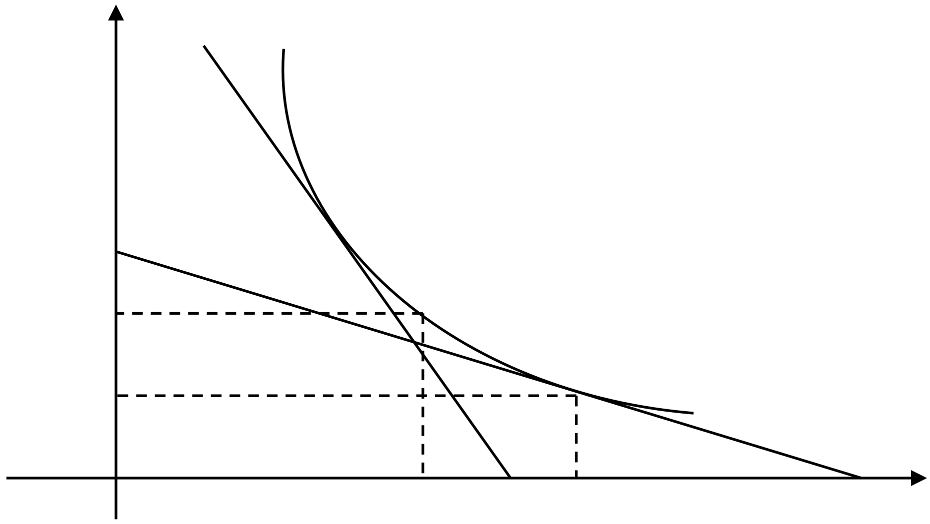


Рис. 6.18. **Ілюстрація граничної норми заміщення ресурсу**

11. *Еластичність заміщення факторів (ресурсів):*

$$\sigma_{ij} = \frac{\partial \, \frac{x_j}{x_i}}{\frac{x_j}{x_i}} : \frac{\partial R_{ij}}{R_{ij}}.$$

Еластичність заміщення ресурсів має такий економічний зміст: вона приблизно показує, на скільки відсотків повинно змінитися відношення ресурсів (у разі незмінного фіксованого випуску продукції), щоб у ході цього гранична норма заміщення R_{ij} змінилася на 1 %.

Еластичність заміни ресурсів має таку геометричну інтерпретацію для випадку ВФ Кобба – Дугласа. Оскільки ізокліналі це промені, то X_2/X_1 – це тангенс кута нахилу ізокліналі до осі абсцис. Величина σ_{12} показує, на скільки % необхідно повернути ізокліналь (тобто змінити $tg\xi$), щоб $tg\varphi$ змінився на 1 % (рис. 6.19).

Слід розглянути завдання оцінювання основних характеристик виробничої функції за відомими початковими умовами [24].

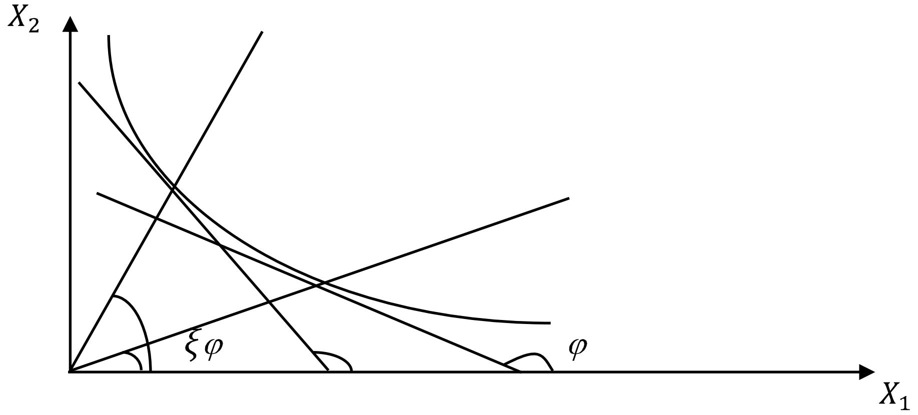


Рис. 6.19. **Визначення еластичності заміни ресурсів для ВФ**

Приклад 6.2. Для наведеної виробничої функції Кобба – Дугласа необхідно розрахувати основні характеристики (середні та граничні продуктивності ресурсів, еластичність випуску продукції за факторами та сумарну еластичність, фондоозброєність та фондомісткість ресурсів, побудувати ізокванти виробничої функції та ізокліналь, розрахувати граничні норми заміщення ресурсів у заданій точці на ізокванті. Зробити висновки.

Нехай задана модель такого вигляду:

$$Y = 2 \cdot {X_1}^{0,6} \cdot {X_2}^{0,4}; \quad X_1(0) = 100,\ X_2(0) = 100\ ;\ Y = 800, \quad X_1 = 100, X_2\ .$$

1. Розрахувати значення обсягу випуску продукції для початкових умов, якщо $X_1(0) = 100,\ X_2(0) = 100$:

$$Y = 2 \cdot 100^{0,6} \cdot 100^{0,4} = 200.$$

2. Розрахувати середню продуктивність ресурсів.

Середня продуктивність праці показує середню кількість продукції на одиницю витраченої праці, розраховується за формулою:

$$A_1 = \frac{y}{x_1} = a_0 \cdot x_1^{a_1 - 1} \cdot x_2^{a_2}; \quad A_1 = \frac{200}{100} = 2 \cdot 100^{-0,4} \cdot 100^{0,4} = 2.$$

Середня фондовіддача (капіталовіддача), показує обсяг продукції в розрахунку на одиницю використаних виробничих фондів:

$$A_2 = \frac{y}{x_1} = a_0 \cdot x_1^{a_1} \cdot x_2^{a_2 - 1}; \quad A_{21} = \frac{200}{100} = 2 \cdot 100^{0,6} \cdot 100^{-0,6} = 2.$$

3. Розрахувати граничну продуктивність ресурсів.

Гранична продуктивність праці показує скільки додаткових одиниць продукції приносить додаткова одиниця витраченої праці, розраховується відповідно до формули:

$$M_1 = \frac{\partial y}{\partial x_1} = 2 \cdot 0{,}6 \cdot 100^{-0{,}4} \cdot 100^{0{,}4} = 1{,}2.$$

Гранична фондовіддача (капіталовіддача) показує, скільки додаткових одиниць продукції приносить додаткова одиниця основних фондів.

$$M_2 = \frac{\partial y}{\partial x_2} = 2 \cdot 0{,}4 \cdot 100^{0{,}6} \cdot 100^{-0{,}6} = 0{,}8.$$

4. Еластичність випуску продукції за факторами виробництва.

Еластичність випуску продукції за витратами праці показує, на скільки відсотків збільшиться випуск продукції у разі збільшення витрат праці на 1 %, може бути визначена таким чином:

$$E_1 = \frac{M_1}{A_1}; \qquad E_1 = \frac{1{,}2}{2} = 0{,}6.$$

Еластичність випуску продукції за витратами виробничих фондів показує, на скільки відсотків збільшиться випуск продукції у разі збільшення основних фондів на 1 %.

$$E_2 = \frac{M_2}{A_2}; \qquad E_2 = \frac{0{,}8}{2} = 0{,}4.$$

Сумарна еластичність за витратами (праці і капіталу).

Сумарна еластичність за витратами показує ефект одночасного пропорційного збільшення обсягу ресурсів праці й основних фондів.

$$E = E_1 + E_2 = a_1 + a_2 = 0{,}6 + 0{,}4 = 1.$$

5. Фондоозброєність (капіталоозброєність) праці показує, скільки (у середньому) капіталу приходиться на одиницю затрачуваної праці і розраховується за формулою:

$$\frac{x_2}{x_1} = a_0^{-\frac{1}{a_2}} \cdot y_0^{\frac{1}{a_2}} \cdot x_1^{-1-\frac{a_1}{a_2}} = \frac{100}{100} = 2^{-2{,}5} \cdot 200^{2{,}5} \cdot 100^{-2{,}5} = 1.$$

6. Фондомісткість продукції показує, який обсяг капіталу витрачається на одиницю випуску продукції і розраховується за формулою:

$$F = \frac{x_2}{Y} = \frac{1}{a_0}\ \frac{x_2}{x_1}^{\ a_1}; \qquad F = \frac{100}{200} = \frac{1}{2}\ \frac{100}{100}^{\ 0,6} = \frac{1}{2}.$$

7. Ізокванти виробничої функції.

Побудувати ізокванту ВФ для обсягу виробництва $Y = 800$, $Y = 500$. Змінюючи значення обсягу затрачуваного капіталу, розрахувати потребу у витратах праці й одержати такі комбінації:

$$X_1 = \ \frac{800}{2 \cdot x_2^{0,4}}^{\ \frac{1}{0,6}}; \qquad X_1 = \ \frac{500}{2 \cdot x_2^{0,4}}^{\ \frac{1}{0,6}}.$$

Розрахунки для побудови ізоквант наведено в табл. 6.9.

Таблиця 6.9

Розрахунок потреб у ресурсах для побудови ізоквант ВФ

X_2	100	150	200	250	300	350	400	450	500	550
$X_1\ (Y = 800)$	1 008	769,2	635	547,19	484,57	437,2	400	369,79	344,71	323,5
$X_1\ (Y = 500)$	460,5	351,43	290,1	250	221,39	199,8	182,8	168,95	157,49	147,8

Графіки ізоквант ВФ наведено на рис. 6.20.

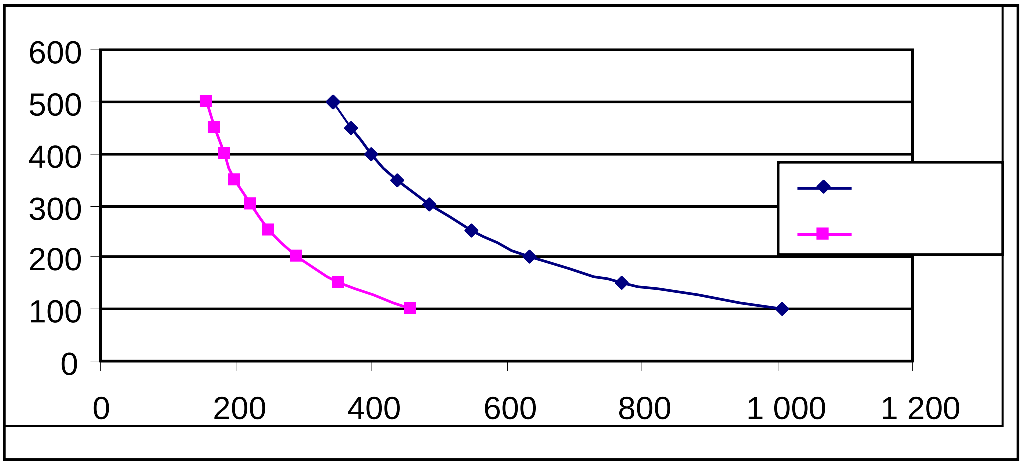


Рис. 6.20. **Ізокванти виробничої функції для $Y = 800$ та $Y = 500$**

8. Гранична норма заміни i-го ресурсу j-м ресурсом розраховується для двофакторної виробничої функції таким чином:

$$R_{12} = \frac{E_1}{E_2} \cdot \frac{x_2}{x_1}; \qquad R_{12} = \frac{0{,}6}{0{,}4} \cdot \frac{700}{300} = 3{,}5.$$

Гранична норма заміни ресурсу збігається з тангенсом кута нахилу f до осі абсцис дотичної, проведеної до ізокванти ВФ у точці $(x_0,\ f(x_0))$.

$$R_{12} = tg\ 180 - \alpha\ = tg74° = 3{,}5.$$

Для дослідження факторів зростання економіки виділяють екстенсивні фактори зростання (за рахунок збільшення витрат ресурсів, тобто збільшення масштабу виробництва) та інтенсивні фактори зростання (за рахунок підвищення ефективності використання ресурсів). Розглянемо завдання визначення масштабу та ефективності виробництва на прикладі.

Приклад 6.3. На основі вихідних даних динаміки підприємства за 20 років була побудована ВФ Кобба – Дугласа для валового прибутку, виду $Y = 2{,}24 \cdot L^{0{,}404} \cdot K^{0{,}803}$. Відомо, що випуск продукції за досліджуваний час збільшився в 2,82 рази, основні виробничі фонди за цей же період збільшилися в 2,88 раза, а число зайнятих на підприємстві в 1,93 раза. Необхідно розрахувати за даною функцією масштаб та ефективність виробництва.

Рішення. Щоб виразити масштаб та ефективність виробництва за допомогою ВФ, необхідно зробити перехід до відносних (безрозмірних) показників. У відносних показниках мультиплікативна ВФ може бути записана таким чином:

$$\frac{Y}{Y_0} = \frac{K}{K_0}^{a_1} \cdot \frac{L}{L_0}^{a_2},$$

де Y_0, L_0, K_0 – значення випуску і витрат фондів та праці за базовий рік.

Дана форма легко приводиться до початкового вигляду:

$$Y = \frac{Y_0}{K_0^{a_1} \cdot L_0^{a_2}} \cdot K^{a_1} \cdot L^{a_2} = A \cdot K^{a_1} \cdot L^{a_2}.$$

Таким чином, коефіцієнт A, який дорівнює: $A = \frac{Y_0}{K_0^{a_1} \cdot L_0^{a_2}}$ інтерпретується як коефіцієнт, який порівнює ресурси з витратами. Якщо позначити випуск і ресурси в відносних (безрозмірних) одиницях вимірювання через $Y,\ L,\ K$, то ВФ можна записати таким чином:

$$Y = K^{a_1} \cdot L^{a_2}.$$

Знайти ефективність економіки за ВФ на основі частинних показників ефективності: продуктивності праці та капіталу. Таким чином, узагальнений показник економічної ефективності є зважене середнє геометричне частинних показників економічної ефективності:

$$E = \left(\frac{Y}{K}\right)^{a} \cdot \left(\frac{Y}{L}\right)^{1-a},$$

а отже, роль вагових коефіцієнтів виконують відносні еластичності:

$$a = \frac{a_1}{a_1 + a_2}, \quad 1 - a = \frac{a_2}{a_1 + a_2},$$

тобто частинні ефективності формують узагальнену ефективність із такими ж пріоритетами, з якими входять у ВФ відповідні ресурси.

Оскільки масштаб виробництва M проявляється в обсягах витрачених ресурсів, середній розмір використаних ресурсів (тобто масштаб виробництва) можна визначити за формулою:

$$M = K^{a} \cdot L^{1-a}.$$

Із даних співвідношень випливає, що випуск Y є добутком економічної ефективності та масштаба виробництва:

$$Y = E \cdot M.$$

Слід знайти відносні еластичності за фондами та трудовими ресурсами:

$$a = \frac{a_1}{a_1 + a_2} = \frac{0{,}404}{0{,}404 + 0{,}803} = 0{,}3347, \quad 1 - a = \frac{a_2}{a_1 + a_2} = 0{,}6653.$$

Визначити частинні ефективності ресурсів:

$$E_K = \frac{Y}{K} = \frac{2{,}82}{2{,}88} = 0{,}98, \quad E_L = \frac{Y}{KL} = \frac{2{,}82}{1{,}93} = 1{,}46.$$

Отже, узагальнений показник ефективності як середнє геометричне частинних, буде мати такий вигляд:

$$E = E_K^a \cdot E_L^{1-a} = 0{,}98^{0{,}3347} \cdot 1{,}46^{0{,}6653} = 1{,}278.$$

Масштаб виробництва встановлюємо як середнє геометричне темпів зростання ресурсів:

$$M = K^a \cdot \ L^{1-a} = 2{,}88^{0{,}3347} \cdot 1{,}93^{0{,}6653} = 2{,}207.$$

Таким чином, можна зробити такий висновок, що зростання виробництва на підприємстві за 20 років в 2,82 раза відбулося за рахунок зростання масштаба виробництва в 2,207 рази і за рахунок підвищення ефективності виробництва в 1,278 рази $(2{,}82 = 1{,}273 \cdot 2{,}207)$.

Контрольні запитання для самодіагностики

1. Що таке лінеаризація?
2. Які типи нелінійних функцій вам відомі?
3. Наведіть покроковий алгоритм побудови нелінійної моделі.
4. Наведіть приклади функцій нелінійних за факторами та за параметрами.
5. Дайте визначення виробничої функції.
6. Охарактеризуйте види виробничих функцій.
7. Які властивості має виробнича функція?
8. Подайте основні напрями використання виробничих функцій на макро- та макрорівні.
9. Що описує модель Кобба – Дугласа – Тимбергена?
10. Яким чином визначаються оцінки параметрів функції Кобба – Дугласа?
11. Які характеристики використовуються для оцінювання адекватності нелінійних моделей?
12. Назвіть основні характеристики виробничої функції.
13. Дайте геометричну інтерпретацію характеристик виробничих функцій.

14. Що таке ізокванта та ізокліналь виробничої функції?

15. Що таке еластичність функції однієї та багатьох змінних?

16. Як за допомогою виробничої функції визначити масштаб та ефективність виробництва?

Тести

1. Виробнича функція – це:

а) функція, що оптимізує витрати виробництва;

б) функція, незалежна змінна якої набуває значення об'ємів виробництва, а залежні змінні – об'ємів ресурсу, що витрачається;

в) функція, залежна змінна якої набуває значення об'ємів виробництва, а незалежні змінні – об'ємів ресурсів, що витрачаються.

2. Виробнича функція Кобба – Дугласа належить до виду:

а) однофакторних економетричних моделей;

б) множинних лінійних моделей;

в) множинних нелінійних моделей.

3. Величина, яка показує, на скільки відсотків збільшиться випуск продукції під час збільшення витрат капіталу на 1 % називається:

а) граничним продуктом капіталу;

б) еластичністю обсягів виробництва за капіталом;

в) граничною нормою заміни капіталу іншими ресурсами.

4. Геометричне місце точок факторів виробничої функції Кобба – Дугласа, для яких обсяги випущеної продукції залишаються незмінним за умови різних обсягів витрат ресурсів, називається:

а) ізокліналлю;

б) ізоквантою;

в) ізокостою.

5. Ефект одночасного, пропорційного збільшення об'ємів ресурсів праці і капіталу відображає:

а) еластичність випуску продукції за працею;

б) еластичність випуску продукції за капіталом;

в) сумарна еластичність за витратами ресурсів.

6. Виробнича функція Кобба –Дугласа має:

а) постійну граничну норму заміни праці капіталом;

б) постійну потребу ресурсів;

в) постійну еластичність заміщення праці капіталом.

7. ВФ Кобба – Дугласа має вигляд$Y = 2\,L^{0,3} \cdot K^{0,7}$. Середня продуктивність праці в точці ($L_0 = 100,\ K_0 = 100$) дорівнює:

а) 2;

б) 0,6;

в) 0,3.

8. ВФ Кобба – Дугласа має вигляд$Y = 2\,L^{0,3} \cdot K^{0,7}$. Капіталоозброєність дорівнює 20. Чому буде дорівнювати середня продуктивність праці:

а) 16,28;

б) 0,6;

в) 1,4?

9. Для однорідної ВФ Кобба – Дугласа першого ступеня середня продуктивність ресурсу:

а) більше граничної продуктивності ресурсу;

б) менше граничної продуктивності ресурсу;

в) дорівнює граничної продуктивності ресурсу.

10. Зміна обсягу виробництва продукції за рахунок зміни капіталу на одиницю у разі незмінних значень інших факторів виробництва є:

а) граничним продуктом капіталу;

б) еластичністю обсягу виробництва за капіталом;

в) граничною нормою заміни капіталом інших ресурсів.

11. ВФ Кобба – Дугласа повинна задовольняти такі умови:

а) $Y(L,K) \geq 0,\ Y(L,0) = 0,\ Y(0,K) = 0$;

б) $Y(L,K) > 0,\ Y(L,0) > 0,\ Y(0,K) > 0$;

в) $Y(L,K) > 0,\ Y(L,0) < 0,\ Y(0,K) < 0$.

12. Перетворення, яке дозволяє лінеаризувати експонентну функцію вигляду $y = a_0 \cdot e^{a_1 x}$:

а) заміна $x' = 1x$;

б) $lny = lna_0 + a_1 \cdot x$;

в) $lny = a_0 + a_1 \cdot x + a_2 \cdot x_2$;

г) $lny = lna_0 + a_1 \cdot lnx$.

13. Для моделювання процесів із перегинами використовується:

а) модифікована експонента;

б) крива Гомперця;

в) логістична крива;

г) гіперболічна крива.

14. До класу S-кривих належить:

а) модифікована експонента;

б) крива Гомперця;

в) функція Торнквіста;

г) логістична крива;

д) гіперболічна крива.

15. Економетрична модель вигляду $y = a_0 + a_1 \cdot x_1^{b_1} + a_2 \cdot x_2^{b_2} + e$ є:

а) статичною регресійною лінійною моделлю;

б) функцією Кобба – Дугласа;

в) нелінійною багатофакторною моделлю.

Практичні завдання

1. На основі даних про зростання прибутку на підприємстві (млн грн) та обсяг випуску продукції (тис. шт.), що наведені в табл. 6.10, необхідно визначити параметри нелінійної моделі,припускаючи, що в якості моделі можна використовувати поліном 2-го ступеня. Визначити параметри моделі за допомогою методу найменших квадратів, оцінити адекватність моделі; привести інтерпретацію отриманих результатів.

Таблиця 6.10

Вихідні дані

№ п/п	1	2	3	4	5	6	7	8	9
X	1	1,5	2	3	3,5	4	4,5	5	5,5
Y	2,75	3,31	4	5,75	6,81	8	9,31	10,75	12,31

2. Відомі дані про дохід підприємства (млн грн) у зв'язку із зростанням випуску продукції (тис. шт.) (табл. 6.11).

Таблиця 6.11

Вихідні дані

№ п/п	1	2	3	4	5	6	7	8	9	10
X	1	1,5	2	3	3,5	4	4,5	5	5,5	6
Y	3,69	2,14	1,67	1,35	1,28	1,24	1,22	1,20	1,19	1,19

Необхідно, використовуючи лінеаризацію, скласти систему нормальних рівнянь для моделі вигляду:

$$y = a_0\, x^{a_1} \cdot e^{a2/x}$$

і визначити її параметри за допомогою МНК. Схематично побудувати графік функції, дослідити її властивості. Здійснити зворотній перехід до параметрів нелінійної моделі; оцінити якість моделі; привести інтерпретацію отриманих результатів.

3. Дослідити виробничий процес у регіоні, за допомогою виробничої функції, що описується залежністю між обсягом виробленої продукції, обсягом трудових ресурсів та об'ємом капіталу. У табл. 6.12 наведені дані про показники діяльності регіону. Необхідно оцінити параметри моделі, її адекватність та статистичну значущість, привести інтерпретацію отриманих результатів.

Таблиця 6.12

Вихідні дані

Регіони	Трудові ресурси (тис. осіб)	Обсяг капіталу (млн грн)	Обсяг виробленої продукції (млн грн)	Регіони	Трудові ресурси (тис. осіб)	Обсяг капіталу (млн грн)	Обсяг виробленої продукції (млн грн)
1	8,32	14,61	24,42	14	5,97	17,64	25,07
2	7,49	15,48	25,03	15	8,64	18,24	29,16
3	6,18	13,64	21,62	16	8,22	15,1	25,37
4	7,24	18,45	27,94	17	6,1	13,4	20,12
5	9,16	13,67	24,18	18	6,3	18,45	25,78
6	6,73	19,45	28,17	19	8,12	17,31	26,97
7	7,64	15,76	25,56	20	9,44	19,71	32,01
8	8,25	16,45	26,83	21	8,5	17,45	28,26
9	9,12	18,31	29,31	22	7,39	15,97	25,44
10	7,89	14,97	24,68	23	9,64	19,24	31,95
11	8,61	18,79	29,54	24	6,16	12,67	20,23
12	9,34	20,71	31,91	25	5,49	16,48	23,14

4. Для наведеної виробничої функції Кобба – Дугласа розрахувати основні характеристики (середні та граничні продуктивності ресурсів, побудувати їх графіки, знайти граничні продукти праці та капіталу, розрахувати еластичність випуску продукції за факторами та сумарну еластичність, фондоозброєність та фондомісткість ресурсів, побудувати ізокванти виробничої функції та ізокліналь, розрахувати граничні норми заміщення ресурсів у заданій точці на ізокванті). Зробити висновки.

$$Y = 4\,L^{0,8} \cdot K^{0,2}; \quad L_0 = 100,\ K_0 = 100\ ;$$
$$\Delta L = -2,\ \Delta K = -3;\ Y = 300,\quad L = 100,\ K\ .$$

5. На основі вихідних даних динаміки підприємства за 20 років була побудована ВФ Кобба – Дугласа вигляду $Y = 2{,}34\,L^{-0{,}306} \cdot K^{0{,}805}$. Відомо, що випуск продукції за досліджуваний час збільшився в 2,95 разів, основні виробничі фонди в 2,4 раза, а число зайнятих – у 1,5 разів. Необхідно розрахувати за даною функцією масштаб та ефективність виробництва.

Ключові слова

Виробнича функція. Виробнича функція Кобба – Дугласа. Гранична продуктивність праці. Гранична фондовіддача. Гранична норма заміщення. Еластичність. Еластичність випуску продукції за факторами виробництва. Ізокванта. Ізокліналь. Лінеаризація. Нелінійна модель. Продуктивність праці. Середня продуктивність ресурсів. Фондовіддача. Фондоозброєність.

Розділ 7. Економетричні моделі динаміки

7.1. Основні поняття та види динамічних рядів.
7.2. Моделі трендів.
7.3. Моделі згладжування часових рядів.
7.4. Метод характеристик.

7.1. Основні поняття та види динамічних рядів

Статистичний опис руху в часі економічних явищ здійснюється за допомогою динамічних (часових) рядів. Під ***динамічним рядом*** розуміється послідовність значень деякого процесу, який протікає в часі [13; 14]. Прикладами часових рядів можуть бути фінансові індекси, щоденні курси валют, котирування акцій, річні обсяги продажів, квартальні обсяги виробництва, ділова активність тощо, тобто змінні, значення яких змінюються з часом. Звичайно елементи часового ряду (члени, рівні ряду) нумерують відповідно до номера моменту часу, до якого вони відносяться (наприклад $Y_1,\ Y_2,\ \dots,\ Y_n$), і порядок рівнів ряду відіграє важливу роль у подальших дослідженнях. Рівні часового ряду одержують, як правило, або в результаті безпосереднього вимірювання ординат досліджуваного процесу через певні проміжки часу, або у процесі усереднювання за певний період часу (наприклад, середня ціна продажів за день, середній прибуток за рік).

Розрізняють два види часових рядів [26; 38]. Вимірювання деяких величин (температури, напруги тощо) проводиться безперервно, принаймні, теоретично. При цьому спостереження можна фіксувати у вигляді графіка. Але навіть у тому випадку, коли величини, що вивчаються, реєструються безперервно, практично під час їх обробки використовуються тільки ті значення, які відповідають дискретній множині моментів часу. Отже, якщо час вимірюється безперервно, часовий ряд називається ***безперервним***, якщо ж час фіксується дискретно, тобто через фіксований інтервал, то часовий ряд є ***дискретним*** [38]. Частіше доводиться мати справу із дискретними часовими рядами, які отримують двома способами:

вибіркою з безперервних часових рядів через регулярні проміжки часу (наприклад, чисельність населення, величина власного капіталу фірми, обсяг грошової маси, курс акцій) – такі часові ряди є ***моментними***;

накопиченням змінної протягом деякого періоду часу (наприклад, обсяг виробництва будь-якого виду продукції, кількість опадів, обсяг імпорту) – у цьому випадку часові ряди є ***інтервальними***.

Залежно від наявності основної тенденції досліджуваного процесу динамічні ряди поділяються на ***стаціонарні*** та ***нестаціонарні*** [25; 27; 34]. Якщо математичне сподівання значень досліджуваного показника та дисперсія постійні, не залежать від часу, то процес є стаціонарним, отже і ряд динаміки є стаціонарним. Однак, економічні процеси зазвичай не є стаціонарними та характеризуються певною тенденцію зміни показника у часі. На рис. 7.1 та 7.2 наведено приклад стаціонарного та нестаціонарного ряду динаміки.

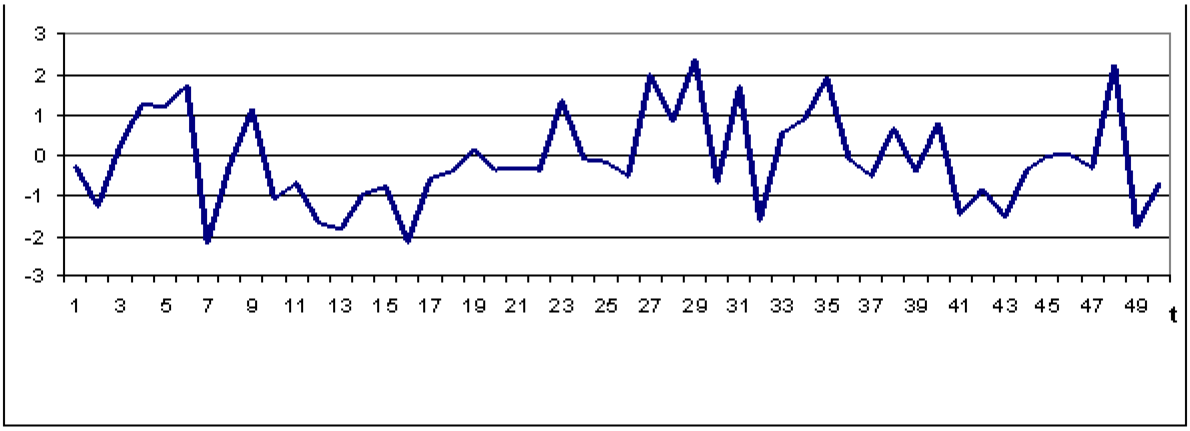


Рис. 7.1. **Приклад стаціонарного ряду динаміки**

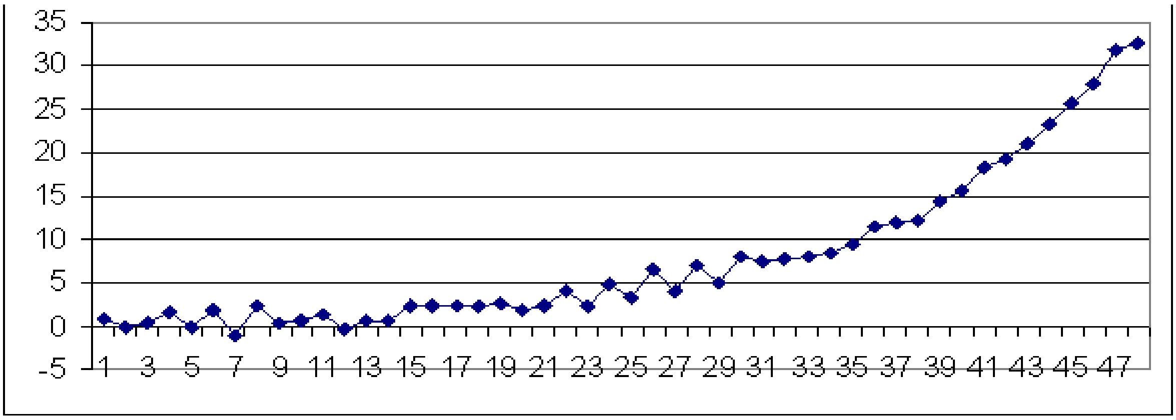


Рис. 7.2. **Приклад нестаціонарного ряду динаміки**

Оцінка та аналіз різних типів часових рядів, дослідження їх поведінки та розробка прогнозів є однією з головних складових прикладних економетричних досліджень в сучасній економіці знань. Таким чином можна сформулювати *основні цілі дослідження динамічного ряду* [1]:

опис характерних особливостей ряду;

з'ясування механізму, що породжує динамічний ряд;

підбір статистичної моделі, що описує динамічний ряд;

прогноз майбутніх значень ряду на основі минулих спостережень;

управління процесом, що породжує часовий ряд.

Практична реалізація перерахованих цілей виконується в результаті таких етапів:

1) графічне подання динамічного ряду;

2) виділення й усунення тренда, сезонних і циклічних складових (вирівнювання, згладжування);

3) виділення і усунення низько- або високочастотних складових (вирівнювання, згладжування);

4) дослідження випадкової складової динамічного ряду, що залишилася після згладжування і фільтрації;

5) побудова математичної моделі для опису випадкової складової і перевірка її адекватності;

6) прогнозування майбутнього розвитку процесу, описаного динамічним рядом;

7) дослідження взаємозв'язків між різними динамічними рядами.

Для вирішення перерахованих завдань використовуються такі методи і моделі [1; 26; 34]:

методи кореляційного аналізу, що дозволяють виявити періодичні залежності і їх лаги всередині одного процесу (автокореляція) або між декількома процесами (кроскореляція);

методи декомпозиції для дослідження ізольованих та комплексних (багатовимірних) рядів динаміки та факторів еволюційного й осциляторного характеру;

методи спектрального аналізу, що націлені на виявлення періодичних і квазіперіодичних складових динамічного ряду;

методи перетворення динамічних рядів (згладжування і фільтрація) з метою усунення високочастотних або сезонних складових;

моделі авторегресії і ковзного середнього (для опису і прогнозування випадкової складової динамічного ряду).

Динаміка рядів соціально-економічних явищ і процесів у загальному випадку формується під впливом систематичних та випадкових факторів

[10; 16; 21; 27]. До систематичних факторів відносять фактори еволюційної та осцилятивної дії. Фактори еволюційного характеру – це зміни, що визначають деякі загальні вектори розвитку, багаторічну еволюцію та характеризують певні сталі закономірності. Фактори осцилятивного характеру – це циклічні (кон'юнктурні) та сезонні коливання. Нерегулярні коливання для соціально-економічних процесів умовно поділяють на дві групи: фактори непередбачуваної дії (війни, екологічні, технологічні катастрофи, фінансові кризи) та дійсно випадкові коливання, що є результатом дії великої кількості відносно несуттєвих другорядних факторів.

Отже, специфікація моделі динамічного ряду, як правило, включає:

систематичну складову – детерміновану послідовність $d_t \;\; t \geq 1$, елементи якої є функцією часу;

випадкову (іррегулярну) складову $\varepsilon_t \;\; t \geq 1$ [11].

Загальний вигляд моделі динамічного ряд з систематичною та випадковою складовою можна представити таким чином:

$$Y_t = d_t + \varepsilon_t, \quad \mathrm{t} = 1, \ldots, \mathrm{n}; \tag{7.1}$$

$$Y_t = d_t \cdot \varepsilon_t, \quad \mathrm{t} = 1, \ldots, \mathrm{n}, \tag{7.2}$$

де n – кількість спостережень;

d_t– детермінована складова;

ε_t– випадкова складова (випадкові перешкоди), що впливає на значення рівнів.

Як відомо, залежно від взаємозв'язку між основними компонентами динамічного ряду розрізняють адитивну та мультиплікативну моделі часового ряду [15; 27; 38; 40]. Модель, в якій всі компоненти ряду представлені як сума цих складових є адитивною, а добутком – мультиплікативною. Аналітично рівняння моделей декомпозиції можна представити таким чином:

– адитивна модель:

$$Y_t = T + C + S + R; \tag{7.3}$$

– мультиплікативна модель:

$$Y_t = T \cdot C \cdot S \cdot R; \tag{7.4}$$

– змішана (мультиплікативно-адитивна) модель:

$$Y_t = T \cdot C \cdot S + R, \tag{7.5}$$

де T – трендова складова;

C – циклічна складова;

S – сезонна складова;

R – випадкова складова.

Адитивна модель характеризується головним чином тим, що характер циклічних та сезонних коливань (флуктуацій) залишається постійним у часі, якщо амплітуда коливань непостійна, то використовується мультиплікативна модель. Приклад динамічного ряду за адитивною та мультиплікативною моделлю часового ряду наведено на рис. 7.3 та 7.4.

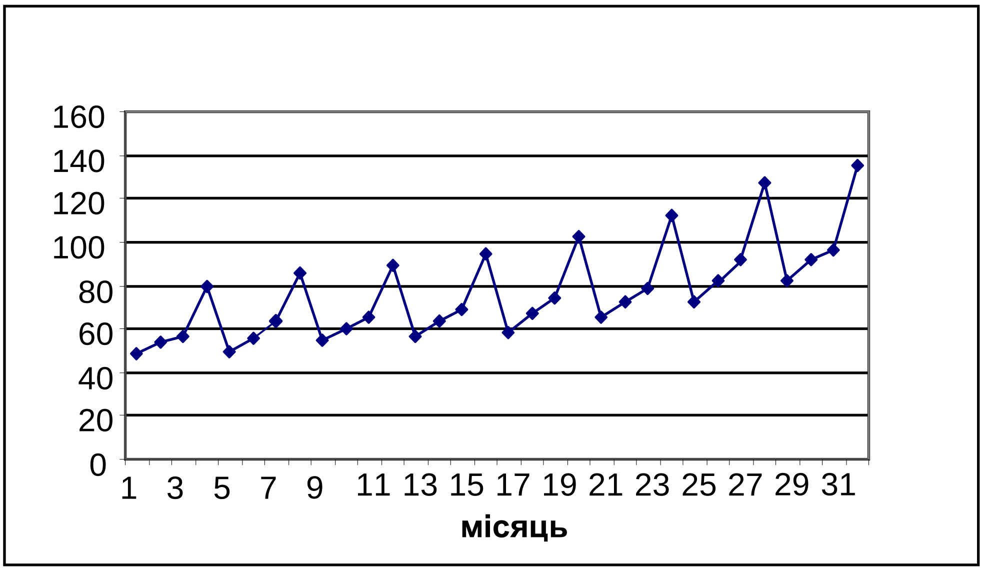


Рис. 7.3. **Приклад адитивної моделі часового ряду**

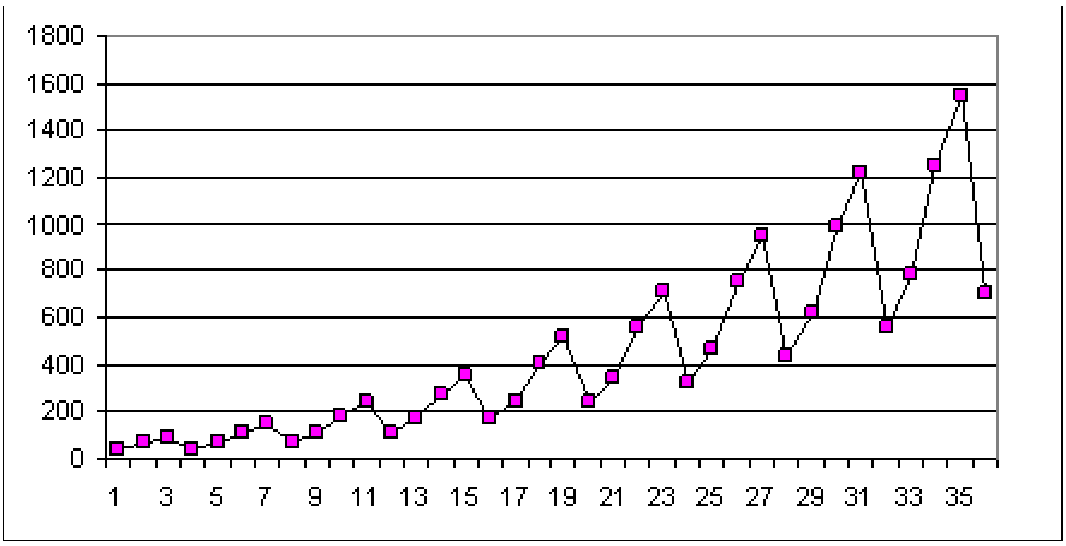
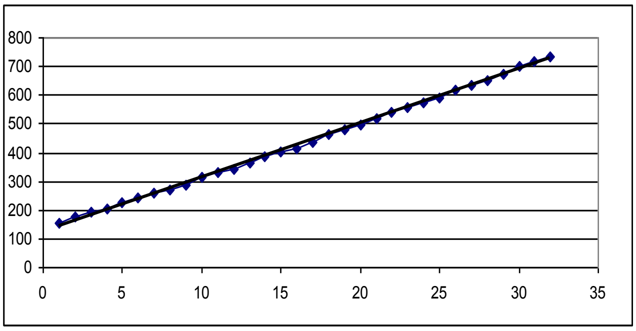
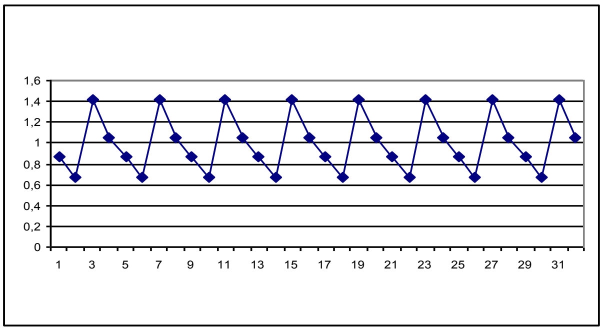


Рис. 7.4. **Приклад мультиплікативної моделі часового ряду**

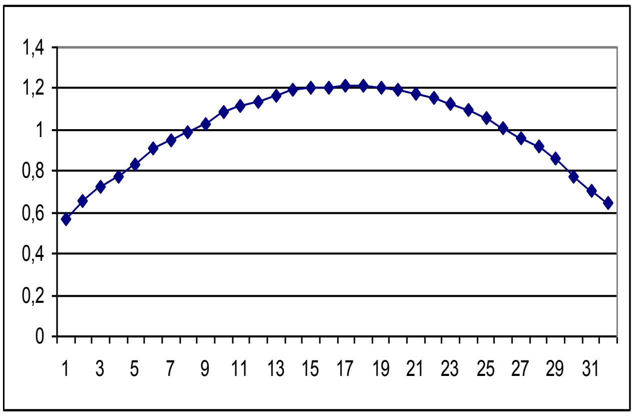
Отже, слід більш детально розглянути основні виділені складові (компоненти) часового ряду, графічне зображення яких наведено на рис. 7.5.



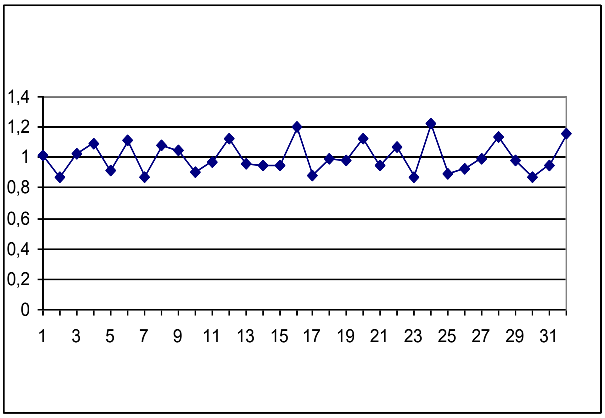
а) трендова компонента



б) сезонна компонента



в) циклічна компонента



г) випадкова компонента

Рис. 7.5. **Приклад основних компонент часового ряду**

T* – *трендова складова (тренд) – це складова, яка плавно змінюється з часом та показує загальну тенденцію (зростає або спадає, але не повторюється регулярним чином), описує вплив довготривалих чинників, ефект яких виявляється поступово (наприклад, зростання населення, зростання споживання, зміна структури споживання, економічний розвиток та ін.).

У соціально-економічних рядах динаміки зазвичай розрізняють основну тенденцію за трьома наступними видами [1; 38; 39]:

1) тенденція середнього рівня аналітично подається за допомогою математичної функції, навколо якої коливаються фактичні рівні досліджуваного явища до поступово спадаючої або зростаючої величини. В даному випадку значення тренду в деякі моменти часу дорівнюють математичному сподіванню динамічного ряду;

2) тенденція дисперсії характеризує тенденцію зміни відхилень між емпіричними рівнями та детермінованою компонентою ряду;

3) тенденція автокореляції (автоковаріації) характеризує зміну взаємозв'язку між окремими рівнями ряду динаміки, тобто зміна величини кореляції між поточними та попередніми значеннями ряду. Особливість полягає у складності візуального визначення взаємозв'язку між значеннями динамічного ряду.

Перед визначенням типу тренда та його аналізом необхідно перевірити гіпотезу на його наявність. Це питання буде розглянуто більш детально.

***C* – циклічна складова** – це складова, що описує тривалі періоди відносного підйому і спаду. Вона складається з циклів, які міняються за амплітудою і протяжністю, отже, її сутність в тому, що значення досліджуваного показника впродовж деякого часу зростає, досягаючи певного максимуму, потім знижується, досягаючи певного мінімуму, потім знову змінюється до попереднього рівня. На циклічну складову впливають такі чинники (що складно ідентифікуються формальними методами) як зростання і виснаження ресурсів, тривало діючі несприятливі погодні умови, зміни у фінансовій і податковій політиці та ін.) Для аналізу цієї складової часового ряду зазвичай залучається додаткова інформація про інші часові ряди, наприклад, про перенасиченість ринку, прийняття законів про податкові пільги тощо.

Для моделювання циклічної компоненти використовується певна періодична функція з періодами, що кратні циклам, в аналітичний вираз функції повинні бути включені гармоніки (тригонометричні функції), періодичність яких зумовлена змістовною сутністю задачі.

S **– сезонна складова** – це складова, що визначається коливаннями, які періодично повторюються за деякий певний період кожного року, місяця, тижня чи дня, отже відображує повторюваність процесу протягом не досить тривалого часу. Призначена для опису поведінки, що регулярно змінюється протягом заданого періоду (наприклад, обсяг продажів шампанського наприкінці грудня кожного року, обсяг перевезень пасажирів зранку та ввечері, попит на морозиво в літні місяці і т. д.).

Відзначимо, що динамічний ряд не завжди містить сезонну чи циклічну компоненти. Перевірку на наявність чи відсутність сезонних коливань проводять за допомогою відповідних критеріїв (дисперсійного, гармонічного аналізу) або візуально при побудові графіків. Значення сезонної компоненти розраховуються методом ковзної середньої і побудовою адитивної чи мультиплікативної моделі. В основі моделювання сезонних коливань є побудова сезонних хвиль та розрахунок індексів сезонності.

$\boldsymbol{R}$ **– випадкова складова** – це складова, що обумовлює стохастичну природу елементів часового ряду та відображує вплив випадкових факторів, що не піддаються обліку та реєстрації [25; 26].

Для опису випадкової складової часового ряду використовуються стохастичні (ймовірнісні) моделі, зокрема моделі стаціонарних випадкових процесів з дискретним часом (стаціонарні послідовності).

Випадковий характер нерегулярної компоненти (ε_t) досліджують на основі наступних критеріальних оцінок [27; 38]:

1) умова випадковості коливань рівнів випадкової компоненти (критерій піків і оцінка гіпотези про незалежність величини від часу);

2) умова нормальності розподілу залишкової компоненти (перевірка асиметрії і ексцесу часового ряду, гіпотез та оцінка середньоквадратичних відхилень, побудова гістограми розподілу);

3) умова рівності математичного очікування значень випадкової компоненти нулю (критерій Стьюдента);

4) умова незалежності значень ряду залишкової компоненти між собою (оцінка автокореляції на основі критерію Дарбіна – Уотсона, циклічного і нециклічного коефіцієнта автокореляції).

Загальний висновок про випадковий характер залишкової компоненти можна зробити тільки в тому випадку, коли всі критерії перевірки властивостей дають позитивні результати.

7.2. Моделі трендів

Згладжування (вирівнювання) рівнів часового ряду виконується за допомогою спеціально підібраних функцій (тренда), які описують закономірності розвитку в часі досліджуваних економічних явищ. Вибір тієї або іншої функції у якості тренда є найважливішим етапом аналізу часового ряду, оскільки помилки на даному етапі можуть призвести до досить серйозних наслідків, особливо під час прогнозування рівнів ряду. Основним інструментом, що дозволяє віддати перевагу тій або іншій моделі, є прирости рівнів ряду, а також деякі їх перетворення [10; 20]. В якості простих ***моделей тренда*** в аналізі часових рядів використовуються поліноми, експоненти та логістичні криві [25; 26].

Поліноміальний тренд має вигляд:

$$Y_t = b_0 + b_1 \cdot t + b_2 \cdot t^2 + \dots + b_n \cdot t^n, \quad (7.6)$$

де t – незалежна змінна (час);

$b_i,\ i = 1, \dots, n$ – параметри моделі, яким, за умови невеликих значень i, можна дати конкретну інтерпретацію. Наприклад, b_0 – рівень ряду у початковий момент часу ($t = 0$), b_1 – швидкість зростання; b_2 – прискорення зростання, b_3 – зміна прискорення.

Тренд, що описується ***поліномом першого ступеня***, має вигляд:

$$Y_t = b_0 + b_1 \cdot t.$$

Дана модель може бути використана для опису тенденції динамічного ряду, в якому рівні з часом або рівномірно зростають, або рівномірно спадають.

Поліном другого ступеня визначається виразом:

$$Y_t = b_0 + b_1 \cdot t + \ b_2 \cdot t^2.$$

Поліном другого ступеня використовується для опису тенденції, в якій прирости рівнів ряду з часом змінюються рівномірно, тобто або рівноприскорено зростають, або рівноприскорено спадають. Залежно від знаку параметра b_2 змінюється форма кривої, так за умови $b_2 > 0$ гілки параболи спрямовані вгору, і парабола має мінімальне значення, за умови $b_2 < 0$ гілки параболи спрямовані вниз, і парабола має максимальне значення. Параметри $b_0,\ b_1$ визначають положення даної кривої в просторі.

Проста модель ***експоненціального (показового)*** тренда має вигляд:

$$Y_t = a \cdot b^t, \quad (7.7)$$

де a, b – параметри моделі.

Модель (7.5) застосовується для опису рівнів ряду з ***постійними темпами зростання і приросту*** ("лавиноподібні процеси"), на які не впливають різного роду обмеження. Характерною особливістю моделі є те, що прирости рівнів ряду залежать від величини самої функції.

Однією з різновидів показового тренда є логарифмічна парабола:

$$Y_t = a \cdot b^t \cdot c^{t^2}, \quad (7.8),$$

де a, b, c – параметри моделі.

Назва кривої випливає з формули, яка одержана в результаті лінеаризації специфікації (7.5):

$$lnY_t = lna + t \cdot lnb + t^2 \cdot lnc. \quad (7.9)$$

Темп приросту кривої є лінійною функцією часу:

$$\tau_{\text{пр}} = lnb + 2t \cdot lnc. \quad (7.10)$$

Для моделювання процесів, що мають насичення, використовуються криві,які мають асимптоту, відмінну від нуля. Зокрема, ***модифікована експонента***, відмінна від експоненти (13.5) тільки доданком k, має горизонтальну асимптоту $Y = k$ за умови $t \to \infty$, $t \to -\infty$. Якщо $a < 0$ асимптота находиться нижче кривої, то

$$Y_t = k + a \cdot b^t. \quad (7.11)$$

Процеси (страхові, демографічні, пов'язані з науково-технічним прогресом), у розвитку яких можна виділити чотири стадії, наприклад, приріст незначний, приріст збільшується, приріст зменшується, приріст незначний, моделюються трендом S-подібної форми. До кривих S-подібної форми відносяться: крива Гомперца, крива Перла – Ріда, логістична крива.

Рівняння ***кривої Гомперца*** має вигляд:

$$Y_t = k \cdot a^{b^t}. \quad (7.12)$$

Крива Перла – Ріда визначається виразом вигляду:

$$\frac{1}{Y_t} = k + a \cdot b^t. \quad (7.13)$$

Логістичну криву звичайно подають у такому вигляді:

$$Y_t = \frac{k}{1 + b \cdot e^{f(t)}} = \frac{k}{1 + b \cdot e^{-at}}. \quad (7.14)$$

7.3. Моделі згладжування динамічних рядів

У ході вирішення завдань згладжування рівнів часового ряду основною проблемою є вибір форми кривої, яка на практиці розв'язується в рамках емпіричного підходу. Проте остаточне рішення про вибір тієї або іншої аналітичної моделі тренда залишається за змістовним аналізом. Існує декілька підходів для визначення форми кривої [1; 26; 36].

1. *Графічний метод.* Припустимі результати виходять тільки за умови відносно простої конфігурації тренда.

2. *Метод послідовних різниць.* Використовується у процесі вибору кривих поліноміального типу. Порядок різниць приймається за ступінь полінома.

3. *Критеріальний підхід*. Найбільш адекватною є така крива, яка мінімізує значення обраного критерію і відповідає змістовному поданню досліджуваного процесу. У якості критерію, як правило, обирається сума квадратів відхилень фактичних значень рівнів ряду від розрахункових, одержаних під час вирівнювання.

4. *Метод характеристик приросту.* Порівнюються характеристики зміни приростів рівнів часового ряду з відповідними характеристиками обраної кривої. Метод включає попередню обробку часового ряду, яка складається з трьох етапів:

- згладжування ряду за ковзною середньою;
- визначення середніх приростів;
- визначення ряду похідних характеристик приросту.

Слід розглянути докладніше кожен із етапів.

Прості ковзні середні. Треба позначити рівні часового ряду як y_t, $\mathrm{t} = 1, \ldots, \mathrm{n}$. Для кожного поточного значення y_t можна розрахувати середнє значення за деяким інтервалом (інтервал згладжування), що включає $m < n$ послідовних членів ряду. Наприклад, вважаючи довжину інтервалу згладжування $m = 2p + 1$ за умови непарного m, буде одержано середню на момент t:

$$y_t = \frac{1}{m} \sum_{i=t-p}^{t+p} y_i, \tag{7.15}$$

де $p = \frac{m-1}{2}$, i – порядковий номер рівня на інтервалі згладжування;

y_i – фактичне значення рівня ряду на момент i [10; 20; 26].

Середня, що визначається за формулою (7.13), одержала назву ***простої ковзної середньої***. Зазвичай за умови практичних розрахунків довжина інтервалу згладжування приймається рівною 3, 5, 7 рокам.

Зважені ковзні середні. Прості ковзні середні є досить грубим засобом для згладжування, оскільки в результаті такого згладжування зникають деякі особливості розвитку процесу. З цієї причини частіше застосовуються ***зважені ковзні середні (ЗКС)*** [10; 20; 26].

У даному методі кожному рівню ряду в межах інтервалу згладжування приписується вага, залежна від відстаней від члена ряду до середини інтервалу згладжування. Система ваг визначається, виходячи з таких міркувань. Для кожних m рівнів ряду із зрушенням у часі на один крок підбирають поліноми вигляду:

$$y_i = a + bi + ci^2 + \cdots, \tag{7.16}$$

де i– порядковий номер рівня у межах інтервалу згладжування.

Центральна ордината параболи (7.14) приймається за згладжене значення відповідного рівня фактичного ряду. І оскільки відлік часу в межах інтервалу згладжування проводиться від його середини, наприклад, $-2, -1, 0, +1, +2$, то згладжене значення рівня ряду для моменту $i = 0$, відповідно до (7.14) дорівнює параметру підібраної параболи. Звідси, зокрема, випливає і формула (7.13) для згладжування за простою ковзною середньою. У цьому випадку передбачається апроксимація поліномом першого ступеня, а оцінка значення параметра a рівняння прямої є середньою вибірковою величиною.

Середні прирости. Під час аналізу динаміки часового ряду, зокрема, методом характеристик приросту, необхідно визначати середні прирости тренда (середньої швидкості зміни тренда). Для вирішення цього завдання вимагається знати значення самого тренда. Для подолання цієї трудності в якості значень тренда можливо використання, наприклад, значень ковзної середньої.

Нехай m перших рівнів ряду апроксимуються поліномом першого ступеня:

$$y_t = a + b \cdot t,$$

якщо відлік часу ведеться від середини інтервалу, то, як було показано, параметр a є середньою вибірковою з рівнів ряду на інтервалі згладжування, а параметр b характеризує приріст тренда, поданого даною прямою, тобто деяким чином усереднений показник для рівнів ряду (і, отже, їх приростів), охоплених інтервалом згладжування. Параметр b визначається за формулою:

$$b = \frac{{}_{-p}^{+p} t \cdot y_t}{{}_{-p}^{+p} t^2}. \tag{7.17}$$

Наприклад, для $m = 3,\ p = 1$ вираз набуває вигляду:

$$u_t = \frac{-1 \cdot y_{t-1} + 0 \cdot y_t + 1 \cdot y_{t+1}}{(-1)^2 + 0^2 + 1^2} = \frac{-y_{t-1} + y_{t+1}}{2}; \tag{7.18}$$

для $m = 5,\ p = 2$ одержано:

$$u_t = \frac{-2 \cdot y_{t-1} - y_{t-1} + y_{t+1} + 2 \cdot y_{t+2}}{10}. \tag{7.19}$$

Характеристики зміни приростів. У ході вибору форми кривої в ролі тренда досліджуваного часового ряду були розглянуті різні перетворення приростів рівнів ряду. Для кожної з розглянутих кривих існує перетворення, яке є лінійним відносно часу t і становить, відповідно, її характеристичну властивість.

Попереднім етапом виділення тренда із даних часового ряду є перевірка гіпотези про існування тенденції у досліджуваному процесі. З цією метою розроблено досить простих критеріїв, заснованих на кореляціях рангів, зворотних точках тощо, проте найбільш надійні результати можна отримати шляхом застосування критеріїв Фішера, Стьюдента та методу Фостера – Стюарта [10; 15; 26; 38].

У процесі визначення тренда в дисперсії використовується F-критерій Фішера. З цією метою вихідний динамічний ряд $Y_1,\ Y_2, \ldots,\ Y_n$ розподіляють на дві групи 1 і 2, обсягом n_1 і n_2, для кожної з груп обчислюють дисперсії S_1^2, S_2^2. Перевірка наявності трендової складової здійснюється за допомогою F-критерію Фішера таким чином: формується статистика:

$$F = \frac{S_1^2}{S_2^2}, \text{ якщо } S_1^2 > S_2^2 \text{ або } F = \frac{S_2^2}{S_1^2}, \text{ якщо } S_2^2 > S_1^2.$$

Обчислене значення F-статистики порівнюється з її критичним (табличним) значенням F_α для заданого рівня значущості α і $k_1 = n_1 - 1$ і $k_2 = n_2 - 1$ ступенів свободи.

Якщо $F < F_\alpha$, то гіпотеза про наявність тренда в дисперсії не підтверджується і можна вважати, що такий тренд у динамічному ряді відсутній.

Після аналізу наявності тренда у дисперсії треба перейти до аналізу тренда наявності тренда у середньому, використовуючи *метод порівняння середніх*.

З цією метою слід обчислити значення t_p за формулою:

$$t_p = \frac{y_1 - y_2}{\sqrt{(n_1 - 1)\cdot S_1^2 + (n_2 - 1)\cdot S_2^2}} \cdot \sqrt{\frac{n_1 \cdot n_2 \cdot (n_1 + n_2 - 2)}{n_1 + n_2}}.$$

Далі треба шукати критичне значення $t_\alpha(n_1 + n_2 - 2)$ для заданого рівня значущості α. Якщо $t_p > t_\alpha$, то можна сказати, що гіпотеза про наявність тренда середнього в ряді спостережень підтверджується, тобто можна сказати, що статистичні розрахунки підтверджують наявність тренда в даному ряді. За умови цього, якщо $y_2 > y_1$, то тренд зростаючий.

Слід розглянути використання *методу Фостера – Стюарта* для цього ж завдання.

Обчислити значення U_i та I_i за формулами:

$U_i = 1$, якщо y_i більше всіх попередніх, 0 – у протилежному випадку;
$I_i = 1$, якщо y_i менше всіх попередніх, 0 – у протилежному випадку.

Далі обчислюємо величини $S_i = U_i + I_i$ та $d_i = U_i - I_i$, а також $S = \sum_{i=1}^{n} S_i, d = \sum_{i=1}^{n} d_i$.

Розрахувати для кожного з цих показників значення критерію Стьюдента за формулами:

$$td = \frac{d - 0}{\delta_2}, \qquad ts = \frac{S - \mu}{\delta_1},$$

де μ, δ_1, δ_2 – табличні значення.

Важливо врахувати, що під час використання методу Фостера – Стюарта можливо виникнення таких ситуацій:

а) $td > t_\alpha$, $ts > t_\alpha$;

б) $td < t_\alpha$, $ts > t_\alpha$;

в) $td > t_\alpha$, $ts < t_\alpha$;

г) $td < t_\alpha$; $ts < t_\alpha$.

Щоб правильно зробити висновок про наявність тренда під час використання цього методу, треба проаналізувати поведінку величин S і d в різних випадках.

1. Якщо ряд монотонно зростає (спадає), то $S = n - 1$ і $d = n - 1$ і за умови достатньої кількості спостережень буде отримано $td > t_\alpha$ і $ts > t_\alpha$, тобто у випадку а) можна зробити висновок про те, що гіпотеза про наявність тренда не відкидається.

2. Якщо ряд змінюється таким чином, що відбуваються коливання показника, в цьому випадку $S \to n-1$, а $d \to 0$, тобто у випадку б) можна зазначити, що наявний тренд у дисперсії або в середньому.

3. Якщо $td > t_\alpha$, а $ts < t_\alpha$, то підтвердити або спростувати гіпотезу про наявність тренда в дисперсії або в середньому не можна (випадок в).

4. У випадку $td < t_\alpha$ і $ts < t_\alpha$ можна підтвердити гіпотезу про відсутність тренда в середньому і в дисперсії (випадок г).

Для прикладу дослідження реальних економічних процесів слід побудувати модель декомпозиції часового ряду, на основі оцінювання та аналізу основних складових часового ряду.

Приклад 7.1. Для наведених значень ряду динаміки продажів необхідно побудувати модель декомпозиції часового ряду, виділити трендову, циклічну сезонну та випадкову складові. Побудувати проноз за кварталами на основі трендової, циклічної та сезонної компоненти. Оцінити якість побудованої моделі [24].

Графік зміни обсягу продажів товару наведено на рис. 7.6.

Рис. 7.6. **Динаміка обсягу продажів товару**

Основне завдання, що виникає під час аналізу часових рядів – визначення наявності тренда. Існують різні методи, що дозволяють визначити наявність тренда: критерій Фішера для визначення тренда в дисперсії; критерій Стьюдента для визначення наявності тренда у середньому.

Метод Фішера використовуються для тестування тренда у дисперсії рівнів ряду. Відповідно до даних методів вихідний ряд $y_1, y_2, \ldots y_n$ розбивається на 2 сукупності:

$y_1, y_2, \ldots y_k;$

$y_{k+1}, y_{k+2}, \ldots y_n.$

Для кожної із сукупностей визначається середнє і дисперсія:

$$F_{\text{теор}} = \frac{S_1^2}{S_2^2} = \frac{488{,}78}{203{,}684} = 2{,}39.$$

$F_{\text{табл}}\ 0{,}05, 14, 14\ = 2{,}48$, а отже $F_{\text{теор}} < F_{\text{табл}}$, тренд у дисперсії відсутній.

Отже, якщо не виявлено тренда у дисперсії рівнів ряду, то ряд тестується на наявність тренда у середньому за допомогою критерію Стьюдента, розрахункове значення якого визначається за формулою:

$$t_{\text{теор}} = \frac{65{,}3763 - 88{,}125}{\sqrt{15 \cdot 203{,}684 + 15 \cdot 488{,}79}} \cdot \sqrt{\frac{16 \cdot 16 \cdot\ 16 + 16 - 2}{16 + 16}} = 3{,}45.$$

$t_{\text{табл}}\ 0{,}95; 30\ = 2{,}042$, а отже $t_{\text{теор}} > t_{\text{табл}}$, тренд у середньому існує.

Слід подати часовий ряд у мультиплікативній формі:

$$Y = T \cdot C \cdot S \cdot R.$$

Варто розглянути алгоритм оцінювання та аналізу складових часового ряду.

1. Оскільки дані про обсяг продажів товару наведені за ряд років за кожним кварталом, то для первісного аналізу необхідно усунути вплив сезонних факторів, провівши згладжування часового ряду методом ковзної середньої з лагом $m = 4$.

2. Для отриманих згладжених значень слід провести повторне згладжування за методом ковзної середньої з лагом $m = 2$. За допомогою даної процедури відбувається часткове усунення міжрічних розходжень. Позначити розраховані значення як (СМА) або Y_{Φ} – центровані ковзні середні. На основі даної характеристики виділяється трендова компонента. Результати розрахунку для даних етапів наведені в табл. 7.1.

Таблиця 7.1

Динаміка обсягу продажів та характеристики ряду

Роки	Квар-тали	$t(x)$	Обсяги продажів ($Y_{\text{н}}$)	Ковзні середні (S_s)	Центровані ковзні середні (СМА)	Коефіцієнт зміни обсягу продажів (K)
1	2	3	4	5	6	7
2007	1	1	48,6	59,805		
	2	2	54,2	60,03		
	3	3	56,62	60,48	59,9175	0,944966
	4	4	79,8	62,2	60,255	1,324371

Закінчення табл. 7.1

1	2	3	4	5	6	7
2008	1	5	49,5	63,625	61,34	0,806978
	2	6	56	64,925	62,9125	0,890125
	3	7	63,5	65,9	64,275	0,987942
	4	8	85,5	66,275	65,4125	1,30709
2009	1	9	54,7	67,15	66,0875	0,827691
	2	10	59,9	67,725	66,7125	0,897883
	3	11	65	68,725	67,4375	0,963855
	4	12	89	69,625	68,225	1,304507
2010	1	13	57	70,925	69,175	0,823997
	2	14	63,9	71,35	70,275	0,909285
	3	15	68,6	72,15	71,1375	0,96433
	4	16	94,2	73,55	71,75	1,312892
2011	1	17	58,7	75,7	72,85	0,805765
	2	18	67,1	77,35	74,625	0,899162
	3	19	74,2	78,65	76,525	0,969618
	4	20	102,8	79,675	78	1,317949
2012	1	21	65,3	82,025	79,1625	0,824886
	2	22	72,3	83,725	80,85	0,894249
	3	23	78,3	86,15	82,875	0,944796
	4	24	112,2	89,475	84,9375	1,320971
2013	1	25	72,1	93,3	87,8125	0,821068
	2	26	82	95,85	91,3875	0,897278
	3	27	91,6	98,35	94,575	0,968543
	4	28	127,5	99,625	97,1	1,313079
2014	1	29	82,3	101,475	98,9875	0,831418
	2	30	92		100,55	0,914968
	3	31	96,7			
	4	32	134,9			

3. Центровані ковзні середні містять еволюторну тенденцію (або тренд), графік яких наведено на рис. 7.2. Виділити тренд і розрахувати параметри лінійної залежності можна за допомогою методу найменших квадратів:

$$T = a + bx,$$

де $a,\ b$ – параметри лінійної залежності;

x – порядковий номер відповідного кварталу ($x = 1^*, 2 \ldots n$).

Оскільки центровані ковзні середні обсягів продажу товару містять у собі тренд, то можна вважати ці значення фактичними даними ($Y_{ф}$). Розрахувати параметри:

$$b = \frac{{}_{i=1} X\cdot Y_{ф} - 28\cdot X\cdot Y}{{}_{I=1}^{28} X^{2} - 28\cdot X^{2}}\ ,\ \ a = Y_{ф} - b\cdot X.$$

У результаті розрахунків було отримано $a\ =\ 55{,}01$; $b\ =\ 1{,}44$.

Отже, лінійна тенденція (тренд) можуть бути наведені як:

$$\mathrm{T}\ =\ 55{,}01\ +\ 1{,}44\cdot\ t.$$

Графік тренда наведено на рис. 7.7. Значення обсягу продажів, отриманих на основі тренда, позначити як Trend $Y_{т}$. Значення трендової компоненти наведено в табл. 7.3.

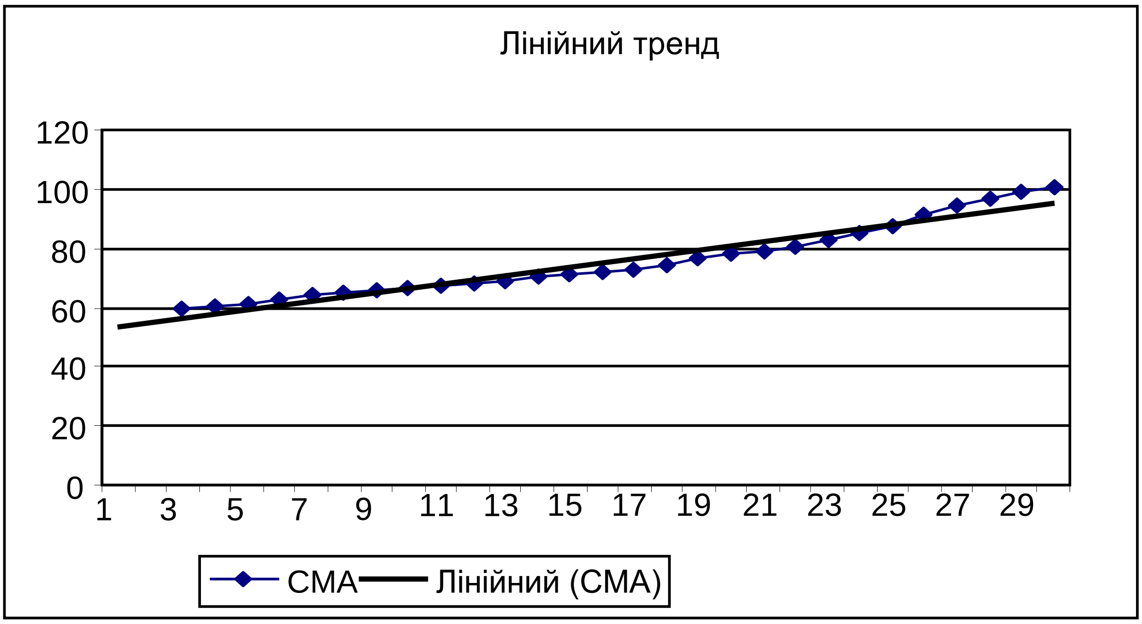


Рис. 7.7. **Лінійний тренд**

4. Визначити значення циклічної компоненти, що розраховуються за формулою:

$$C = \frac{CMA}{T} = \frac{Y_{Ф}}{Y_{T}}.$$

Графічне подання значень циклічної складової наведено на рис. 7.8, що відображає тенденцію зниження (підвищення) обсягу продажів за часовий період, пов'язаних із впливом економічних циклів (спад або зростання економічних показників). Розрахункові значення циклічної компоненти наведено в табл. 7.4.

Рис. 7.8. **Циклічна компонента**

5. Необхідно визначити сукупний вплив сезонної і випадкової складових на динаміку обсягу продажів товару.

Позначити цей вплив через змінну, яку назвати коефіцієнтом зміни обсягу продажів (K).

$$K = S \cdot R = \frac{T \cdot C \cdot S \cdot R}{T \cdot C} = \frac{Y_{н}}{CMA},$$

де $Y_{н}$ – значення обсягу продажів, наведених в умові завдання.

Значення даного коефіцієнта наведені в табл. 7.4.

6. Після цього проводиться розрахунок сезонних складових за такою схемою (табл. 7.2).

1) групуються за роками і кварталами коефіцієнти зміни обсягу продажів;

2) у стовпцях, де наведені коефіцієнти зміни обсягу продажів, знаходяться максимальне і мінімальне значення, що умовно викреслюються;

3) розраховується сума за кожним стовпцем, без максимальних і мінімальних значень коефіцієнтів зміни обсягу;

4) розраховується модифіковане середнє (MC):

$$MC = \frac{\Sigma}{n},$$

де n – кількість елементів у кожному стовпці, за винятком мінімального і максимального елементів;

5) розраховується скореговане модифіковане середнє (СМС) таким чином:

- розраховується сума модифікованих середніх, яка дорівнює 3,997834;
- визначається скореговане модифіковане середнє:

$$S_1 = 0{,}820924 \cdot \frac{4}{3{,}997834} = 0{,}821369; \quad S_3 = 0{,}962262 \cdot \frac{4}{3{,}997834} = 0{,}962784;$$

$$S_2 = 0{,}899571 \cdot \frac{4}{3{,}997834} = 0{,}900059; \quad S_4 = 1{,}315076 \cdot \frac{4}{3{,}997834} = 1{,}315789;$$

- розраховується сезонний індекс:

Сезонний індекс = Модифіковане середнє ·100 %.

Значення сезонних індексів наведені у табл. 7.2.

Таблиця 7.2

Оцінка сезонної компоненти

Роки	1	2	3	4
2007			0,944966	1,324371
2008	0,806978	0,890125	0,987942	1,30709
2009	0,827691	0,897883	0,963855	1,304507
2010	0,823997	0,909285	0,96433	1,312892
2011	0,805765	0,899162	0,969618	1,317949
2012	0,824886	0,894249	0,944796	1,320971
2013	0,821068	0,897278	0,968543	1,313079
2014	0,831418	0,914968		
Сума	4,104618	4,497857	4,811312	6,575381
Модифіковане середнє	0,820924	0,899571	0,962262	1,315076
Скореговане модифіковане середнє	0,821368	0,900059	0,962784	1,315789
Сезонний індекс	82,13685	90,00588	96,27839	131,5789

Графік сезонної компоненти наведено на рис. 7.9.

Рис. 7.9. **Сезонна компонента**

7. Провести розрахунок значень випадкової складової (R) таким чином:

$$R = \frac{S \cdot R}{S} = \frac{K}{S},$$

де K – коефіцієнт зміни обсягу продажів;

S – сезонна складова.

Під час розрахунку значень випадкової складової (R) варто використовувати порядок кварталів 1, 2, 3, 4 (як у вихідних даних) і відповідні їм S_1, S_2, S_3, S_4. Значення випадкової компоненти наведено в табл. 7.3.

Таблиця 7.3

Складові часового ряду

t	Тренд Trend (Y_T)	Циклічна складова, C	Сезонна складова, S	Випадкова складова, R	Y теор.	Похибки, e
1	2	3	4	5	6	7
1	56,63	1,058052	0,962784	0,981493	57,6876	-1,0676
2	58,06	1,037806	1,315789	1,006523	79,28285	0,517145
3	59,49	1,031098	0,821368	0,982479	50,38274	-0,88274
4	60,92	1,032707	0,900059	0,988963	56,62495	-0,62495
5	62,35	1,030874	0,962784	1,026131	61,88293	1,617066
6	63,78	1,025596	1,315789	0,993389	86,06904	-0,56904
7	65,21	1,013457	0,821368	1,007697	54,28219	0,41781

Закінчення табл. 7.3

1	2	3	4	5	6	7
8	66,64	1,001088	0,900059	0,997582	60,04517	-0,14517
9	68,07	0,990708	0,962784	1,001113	64,92774	0,072262
10	69,5	0,981655	1,315789	0,991426	89,76969	-0,76969
11	70,93	0,975257	0,821368	1,0032	56,81817	0,181834
12	72,36	0,971186	0,900059	1,010251	63,25163	0,648367
13	73,79	0,964053	0,962784	1,001606	68,49004	0,109962
14	75,22	0,953869	1,315789	0,997798	94,40785	-0,20785
15	76,65	0,950424	0,821368	0,981003	59,83669	-1,13669
16	78,08	0,955751	0,900059	0,999004	67,16689	-0,06689
17	79,51	0,962458	0,962784	1,007098	73,67704	0,522963
18	80,94	0,963677	1,315789	1,001642	102,6315	0,168473
19	82,37	0,96106	0,821368	1,004282	65,02158	0,278417
20	83,8	0,964797	0,900059	0,993545	72,76976	-0,46976
21	85,23	0,972369	0,962784	0,981317	79,79071	-1,49071
22	86,66	0,980123	1,315789	1,003939	111,7598	0,440188
23	88,09	0,99685	0,821368	0,999634	72,12642	-0,02642
24	89,52	1,020861	0,900059	0,99691	82,25412	-0,25412
25	90,95	1,039857	0,962784	1,005982	91,05529	0,544714
26	92,38	1,051093	1,315789	0,997941	127,7631	-0,26309
27	93,81	1,055191	0,821368	1,012235	81,30521	0,994786
28	95,24	1,055754	0,900059	1,016564	90,50091	1,499086

Графік випадкової компоненти наведено на рис. 7.10.

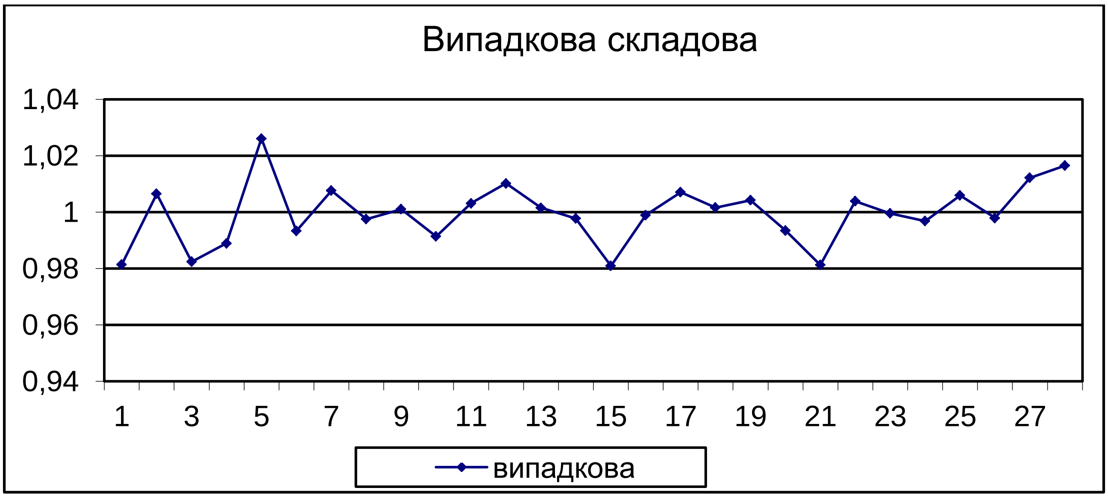


Рис. 7.10. **Випадкова складова**

8. Розглянути прогнозування складових на основі досліджень даного часового ряду.

На підставі проведених розрахунків тренд для даного часового ряду можна подати в такому вигляді:

$$T = 55{,}010 + 1{,}44 \cdot X,$$

де X – порядковий номер кварталу.

Розрахувати можливі значення обсягу продажів у 2000 році за кварталами на основі тренда:

1 квартал 2017 (41) $Y_T = 55{,}010 + 1{,}44 \cdot 41 = 113{,}83.$
2 квартал 2017 (42) $Y_T = 55{,}010 + 1{,}44 \cdot 42 = 115{,}26.$
3 квартал 2017 (43) $Y_T = 55{,}010 + 1{,}44 \cdot 43 = 116{,}69.$
4 квартал 2017 (44) $Y_T = 55{,}010 + 1{,}44 \cdot 44 = 118{,}12.$

Визначити можливі значення обсягу продажів з урахуванням впливу сезонної компоненти:

1 квартал 2017 $Y_{TS} = Y_T \cdot S_1 = 113{,}83 \cdot 0{,}821 = 93{,}49638.$
2 квартал 2017 $Y_{TS} = Y_T \cdot S_2 = 115{,}26 \cdot 0{,}900 = 103{,}7408.$
3 квартал 2017 $Y_{TS} = Y_T \cdot S_3 = 116{,}69 \cdot 0{,}9627 = 112{,}3473.$
4 квартал 2017 $Y_{TS} = Y_T \cdot S_4 = 118{,}12 \cdot 1{,}3158 = 155{,}421.$

Визначити можливі значення обсягів продажу, припускаючи, що крім урахованих у моделі факторів діє і циклічна складова.

Припускаючи періодичність дії циклічної складовий на основі візуального аналізу графіка розподілу даної складової ($t = 27$ кварталів), визначимо значення циклічної складової:

$$C_{41} = C_{41-27} = C_{14} = 0{,}953869; \quad C_{42} = C_{42-27} = C_{15} = 0{,}950424;$$
$$C_{43} = C_{43-27} = C_{16} = 0{,}955751; \quad C_{44} = C_{44-27} = C_{17} = 0{,}962458.$$

Таким чином,
1 квартал 2017 $Y_{TSC} = Y_{TS} \cdot C_{14} = 93{,}49638 \cdot 0{,}953869 = 89{,}18326.$
2 квартал 2017 $Y_{TSC} = Y_{TS} \cdot C_{15} = 103{,}7408 \cdot 0{,}950424 = 98{,}59773.$
3 квартал 2017 $Y_{TSC} = Y_{TS} \cdot C_{16} = 112{,}3473 \cdot 0{,}955751 = 107{,}3759.$
4 квартал 2017 $YY_{TSC} = Y_{TS} \cdot C_{17} = 155{,}421 \cdot 0{,}962458 = 149{,}5861.$

Прогнозні значення складових часового ряду та прогнозне значення обсягу продажів наведено у табл. 7.4.

Таблиця 7.4

Прогнозування складових ряду динаміки

Період	Прогноз тренда (T)	Прогноз сезонної компоненти (S)	Прогноз циклічної компоненти (C)	$Y_{\text{теор}}(T \cdot S \cdot C)$
41	113,83	0,821368	0,953869	89,18326
42	115,26	0,900059	0,950424	98,59773
43	116,69	0,962784	0,955751	107,3759
44	118,12	1,315789	0,962458	149,5861

9. Важливим етапом прогнозування соціально-економічних явищ є оцінювання точності і надійності прогнозів.

Емпіричною мірою точності прогнозу слугує величина його помилки, що визначається як різниця між прогнозними і фактичними значеннями досліджуваного показника. Даний підхід можливий тільки в двох випадках:

а) період попередження відомий, уже закінчився і дослідник має необхідні фактичні значенням прогнозованого показника;

б) будується ретроспективний прогноз, тобто розраховуються прогнозні значення показника для періоду часу, за який уже наявні фактичні значення. Це робиться з метою перевірки розробленої методики прогнозування.

Для оцінювання точності прогнозу й адекватності моделі використовується ряд статистичних критеріїв [25; 26]. Теоретичне значення обсягу продажів та похибки моделі наведені в табл. 7.3. На основі даної інформації слід оцінити адекватність моделі та якість прогнозу (табл. 7.5).

Таблиця 7.5

Оцінювання якості прогнозу

Критерій якості прогнозу	Формула розрахунку	Значення помилок за моделлю
1	2	3
Середня помилка	$m.e. = \frac{{}_{t=1}^{n} e_t}{n}$	0,001369

Закінчення табл. 7.5

1	2	3
Середня абсолютна помилка	$m.a.e. = \frac{\sum_{t=1}^{n} e_t}{n}$	0,570993
Сума квадратів помилок	$s.s.e. = \sum_{t=1}^{n} e_t^2$	14,79841
Середньоквадратична помилка	$m.s.e. = \sqrt{\frac{\sum_{t=1}^{n} e_t^2}{n}}$	0,72699
Середньовідсоткова помилка	$m.p.e. = \frac{1}{n} \sum_{t=1}^{n} \frac{e_t}{y_t} \cdot 100\ \%$	-0,03304
Середня абсолютна відсоткова помилка	$m.a.p.e. = \frac{1}{n} \sum_{t=1}^{n} \frac{e_t}{y_t} \cdot 100\ \%$	0,806345

Якщо значення середньої абсолютної процентної помилки знаходиться в інтервалі:

$0 < m.a.p.e. < 10\ \%$, модель забезпечує високу точність прогнозу.

$10\% < m.a.p.e. < 20\ \%$, модель забезпечує задовільну точність прогнозу.

$m.a.p.e. > 20\ \%$, модель не адекватна.

Отже, на основі оцінювання та аналізу моделі декомпозиції часового ряду, можна зробити висновок, що побудована модель адекватна та забезпечує високу точність прогнозу.

7.4. Метод характеристик

Для підбору виду функції, яка може розглядатися в якості моделі тренда часового ряду, слід використовувати *метод характеристик* [25; 26]. Цей метод ґрунтується на тому, що найбільш типові нелінійні функції можна розпізнати за окремими розрахунковими характеристиками ряду вихідних даних. Якщо деяка характеристика для ряду вихідних даних постійна (наприклад, постійні прирости або темпи зростання), то найбільш прийнятною для моделювання цього ряду буде відповідна їй функція.

Лінійна функція, наприклад, має постійні перші прирости $u_t^{\ 1} = y_{t+1} - y_t = const$, експоненті відповідають постійні темпи зростання $\tau_t = \frac{y_{t+1}}{y_t}$ і т. д. Алгоритм методу характеристик включає такі кроки:

Крок 1. Вихідний ряд рівнів згладжується за допомогою ковзної середньої y_t.

Крок 2. Для згладженого ряду розраховуються такі характеристики:

2.1. Перші прирости: $u_t^{(1)} = y_{t+1} - y_t$.

2.2. Другі прирости (прирости приростів): $u_t^{\ 2} = u_{t+1}^{1} - u_t$.

2.3. Треті прирости:$u_t^{(3)} = u_{t+1}^{(2)} - u_t^{(2)}$тощо.

Якщо будь-які з цих приростів постійні, то тренд можна описати за допомогою полінома відповідного ступеня:

2.4. Темпи зростання:

$$\tau_t = \frac{y_{t+1}}{y_t}.$$

Якщо $\tau_t = const$, то тренд ряду описується показниковою функцією.

2.5. Темпи зростання приростів:

$$v_t = \frac{u_{t+1}^{(1)}}{u_t^{(1)}}.$$

Якщо $v_t = const$, то за криву зростання в моделі можна взяти модифіковану експоненту.

2.6. Зворотні значення рівнів $z_t = \frac{1}{y_t}$ й перші різниці для них:

$$w_t^{(1)} = z_{t+1} - z_t.$$

Якщо $w_t^{(1)} = const$, то тренд можна подати зворотною функцією:

$$y_t = \frac{1}{a_0 + a_1 t}.$$

2.7. Логарифми рівнів $S_t = ln y_t$ й перші різниці для них $q_t^{(1)} = S_{t+1} - S_t$. Якщо $q_t^{(1)}$ буде наближатися до нуля під час зростання t, можна розглянути у якості кривої степеневу (мультиплікативну) функцію.

2.8. Темпи зростання приростів логарифмів рівнів: $p_t = \frac{q_{t+1}^{(1)}}{q_t^{(1)}}$. Якщо величини $p_t = const$, то за криву зростання можна взяти функцію Гомперца.

2.9. Темпи зростання перших приростів зворотних величин $\mu_t = \frac{w_{t+1}^{(1)}}{w_t^{(1)}}$. Якщо величини $\mu = const$, то за криву зростання можна взяти логістичну криву.

У табл. 7.6 наведено систему характерних змін показників для різних видів кривих, що дозволяє спростити процедуру вибору тренда.

Таблиця 7.6

Система характерних змін показників для різних видів кривих

Показники	Характер зміни в часі	Види кривої
u_t	Майже однакові	Пряма
u_t	Лінійно змінюються	Парабола другого ступеня
${u_t}^2$	Лінійно змінюються	Парабола третього ступеня
u_t/y_t	Майже однакові	Експонента
u_t/y_t	Лінійно змінюються	Логарифмічна парабола
$\ln(u_t)$	Лінійно змінюються	Модифікована експонента
$\ln(u_t/y_t)$	Лінійно змінюються	Крива Гомперца
$\mathrm{n}(u_t/{y_t}^2)$	Лінійно змінюються	Логістична крива

Крок 3. Після розрахунку всіх характеристик, слід оцінити за допомогою коефіцієнтів варіації однорідність кожного ряду характеристик. Найменші значення коефіцієнта варіації відповідають тим нелінійним функціям, які є найбільш імовірними для опису нелінійного тренда.

Приклад 7.2. Слід розглянути метод характеристик на прикладі. Необхідно підібрати криву зростання для опису тренда за рядом спостережень величини прибутку виробничого об'єднання (табл. 7.7). Оцінити параметри нелінійної трендової моделі та розрахувати прогнозні значення прибутку виробничого об'єднання на два періоди вперед.

Таблиця 7.7

Вихідні дані, млн грн

t	1	2	3	4	5	6	7	8
$Y(t)$	4,01	4,34	5,12	5,52	5,93	6,64	7,30	7,94
t	9	10	11	12	13	14	15	16
$Y(t)$	8,29	8,49	8,29	8,50	8,61	9,24	9,27	9,58

Спочатку за вихідними даними обчислимо ковзну середню з лагом $m = 5$ за формулою:

$$y_t = \frac{y_{t-2} + y_{t-1} + y_t + y_{t+1} + y_{t+2}}{5}.$$

Результати розрахунків наведені в табл. 7.8.

Таблиця 7.8

Вихідні дані часового ряду й розраховані характеристики

№ п/п	y_t	y_t	$u_t^{(1)}$	$u_t^{(2)}$	$u_t^{(3)}$	τ_t	v_t	z_t	$w_t^{(1)}$	S_t	$wq_t^{(1)}$	p_t	μ_t
1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	4,01												
2	4,34												
3	5,12	4,984	0,526	0,066	-0,094	1,106	1,125	0,21	-0,029	1,606	0,100	1,017	0,586
4	5,52	5,51	0,592	-0,028	0,018	1,107	0,953	0,181	-0,017	1,707	0,102	0,866	0,824
5	5,93	6,102	0,564	-0,01	-0,032	1,092	0,982	0,164	-0,014	1,809	0,088	0,903	0,786
6	6,64	6,666	0,554	-0,042	-0,14	1,083	0,924	0,15	-0,011	1,897	0,080	0,858	0,909
7	7,3	7,22	0,512	-0,182	0,092	1,071	0,645	0,139	-0,01	1,977	0,069	0,610	0,500
8	7,94	7,732	0,33	-0,09	-0,016	1,043	0,727	0,129	-0,005	2,045	0,042	0,702	0,800
9	8,29	8,062	0,24	-0,106	0,162	1,03	0,558	0,124	-0,004	2,087	0,029	0,546	0,250
10	8,49	8,302	0,134	0,056	-0,09	1,016	1,418	0,12	-0,001	2,116	0,016	1,391	3,000
11	8,29	8,436	0,19	-0,034	0,136	1,023	0,821	0,119	-0,003	2,133	0,022	0,805	1,000

Закінчення табл. 7.8

1	2	3	4	5	6	7	8	9	10	11	12	13	14
12	8,5	8,626	0,156	0,102		1,018	1,654	0,116	-0,003	2,155	0,018	1,615	0,667
13	8,61	8,782	0,258			1,029		0,113	-0,002	2,173	0,029		
14	9,24	9,04						0,111		2,202			
15	9,27												
16	9,58												
V			0,493	-3,219	26,693	0,034	0,348		-0,943	0,098		0,361	0,813

Потім обчислити всі названі характеристики ряду: $u_t^{(1)}$, $u_t^{(2)}$, $u_t^{(3)}$,τ_t, v_t, z_t, $w_t^{(1)}$, S_t, $q_t^{(1)}$, p_t, μ_t (табл. 7.8).

Оскільки ознакою наявності певного тренда є сталість значень відповідної йому характеристики, то слід оцінити сталість за допомогою коефіцієнтів варіації. Коефіцієнт варіації розраховуються за формулою:

$$V = \frac{\sigma}{x},$$

де σ – середньоквадратичне значення характеристики;

x – середньоарифметичне значення характеристики.

Розрахунок коефіцієнтів варіації для всіх характеристик наведено в останньому рядку табл. 7.8. За результатами розрахунків за коефіцієнтом варіації можна зробити висновок, що такі характеристики, як τ_t, v_t, S_t приблизно постійні, тобто тренд цього процесу можна описати такими нелінійними функціями:

1) показниковою функцією $y = a_0 a_1^t$;
2) модіфікованою експоненційною функцією $y_t = k + a_0 e^{a_1 t}$;
3) степеневою (мультиплікативною) залежністю $y = a_0 t^{a_1}$.

Для вибору остаточного варіанту кривої зростання необхідно зробити розрахунки за обраними кривими і обрати ту, яка приводить до мінімальних похибок.

Для оцінювання параметрів нелінійних моделей використовується процедура лінеаризації, тобто нелінійні моделі приводять шляхом деяких перетворень до лінійного вигляду, та вже для модифікованих моделей використовується МНК.

Для лінеаризації застосувати процедуру логарифмування:

$$lny = lna_0 + tlna_1.$$

Зробити заміни: $y^* = lny$, $a_0^* = lna_0$, $a_1^* = lna_1$.

Буде отримано лінійну форму моделі: $y^* = a_0^* + a_1^* \cdot t$.

Оскільки перетворена залежність є лінійною, її параметри можуть бути знайдені за допомогою МНК:

$$a = (X^T X)^{-1} YX.$$

Послідовність розрахунку параметрів отриманої моделі $y^* = a_0^* + a_1^* \cdot t$ наведена на рис. 7.11.

I.

$y^* = lny$
1,39
1,47
1,63
1,71
1,78
1,89
1,99
2,07
2,12
2,14
2,12
2,14
2,15
2,22
2,23
2,26

II.

X	
1	1
1	2
1	3
1	4
1	5
1	6
1	7
1	8
1	9
1	10
1	11
1	12
1	13
1	14
1	15
1	16

IV.

$X^T X$	
16	136
136	1496

V.

$(X^T X)^{-1}$	
0,28	-0,03
-0,03	0,00

VI.

YX
31,30
284,85

VII.

a^*
1,49
0,06

III.

X^T															
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16

Рис. 7.11. **Процедура розрахунку параметрів моделі**

Виходячи з наведених на рис. 7.11 розрахунків параметри моделі $y^* = a_0^* + a_1^* \cdot t$ дорівнюють $a_0^* = 1{,}49$, $a_1^* = 0{,}06$. Тобто модифікована модель має вигляд: $y^* = 1{,}49 + 0{,}06t$.

Прогнозні значення прибутку виробничого об'єднання дорівнюють:

$$y_{17} = 4{,}44 \cdot 1{,}06^{17} = 11{,}96;$$
$$y_{18} = 4{,}44 \cdot 1{,}06^{18} = 12{,}67.$$

Контрольні запитання для самодіагностики

1. Що таке динамічний ряд?
2. У чому особливості побудови моделей динаміки?
3. Які складові можуть бути виділені в часовому ряді?
4. Дайте визначення "тренд". Які основні види трендів?
5. Охарактеризуйте моделі декомпозиції часового ряду.
6. Які особливості побудови авторегресійних моделей?
7. У чому суть гармонійного і спектрального аналізу?
8. Опишіть алгоритм побудови моделі декомпозиції часового ряду.
9. Які методи перевірки моделі на наявність тренда вам відомі.
10. Що таке сезонні коливання. Як визначити сезонну компоненту.
11. Наведіть переваги та недоліки експоненційного згладжування.
12. Які методи виявлення форми тренда вам відомі. Наведіть приклади.
13. У чому особливості прогнозування за трендовими моделями.
14. Які коефіцієнти використовуються для оцінювання адекватності динамічних моделей.
15. Проінтерпретуйте значення параметрів за основними кривими зростання.
16. Наведіть методи перевірки випадкової компоненти.

Тести

1. Детермінована та випадкова складові динамічного ряду є величинами:

а) спостережуваними;

б) неспостережуваними;

в) стохастичними.

2. Модель розкладання динамічного ряду на детерміновану і випадкову складові має форми:

а) адитивну;

б) мультиплікативну;

в) структурну.

3. Мультиплікативна динамічна модель може бути наведена в такому вигляді:

а) $y_t = T + C + S + R$;

б) $y_t = T \cdot C \cdot S \cdot R$;

в) $y_t = T \cdot C \cdot S + R$.

4. У разі адитивного характеру сезонності:

а) амплітуда коливань змінюється в часі пропорційно рівню тренда;

б) амплітуда періодичних коливань приблизно постійна і не залежить від рівня тренда.

5. Ступінь полінома моделі, яка використовується для опису тенденції, в якій прирости рівнів ряду з часом змінюються рівномірно:

а) перший;

б) другий;

в) третій.

6. Тренд із специфікацією виду $Y_t = b_0 + b_1 \cdot t + b_1 \cdot t^2 + \cdots + b_n \cdot t^n$ є:

а) поліноміальним;

б) експоненціальним;

в) логістичним;

г) квадратичним.

7. Тренд із специфікацією вигляду $Y_t = a \cdot b^t$ є:

а) поліноміальним;

б) експоненціальним;

в) логістичним;

г) логарифмічним.

8. Які тести застосовують на наявність (відсутність) тренда у рівнях динамічного ряду:

а) Дарбіна – Уотсона;

б) Голдфельдта – Квандта;

в) Фішера;

г) Стьюдента.

9. Для тестування ряду на наявність тренда в дисперсії використовується:

а) критерій Стьюдента;

б) критерій Фішера;

в) критерій Дікі – Фуллера;

г) критерій Дарбіна – Уотсона.

10. Що характеризує коефіцієнт параболічного тренда a_1:

а) середня зміна аналізованого явища від одного періоду (моменту) до другого періоду (моменту) часу;

б) середнє прискорення зміни аналізованого явища від одного періоду (моменту) до другого періоду (моменту) часу;

в) середній вирівняний рівень ряду для періоду (моменту) часу, прийнятого за початок відліку;

г) постійний ланцюговий темп зміни рівнів часового ряду.

11. Параметри трендової моделі можуть бути оцінені на основі:

а) методу характеристик;

б) методу найменших квадратів;

в) методу Стьюдента;

г) методу Фішера.

12. Рівняння тренда має такий вигляд $y = 32{,}5 - 4{,}6t$. На скільки в середньому за рік у досліджуваному періоді змінюється ознака:

а) збільшується на 32,5;

б) збільшується на 4,6;

в) зменшується на 4,6;

г) зменшується на 32,5?

13. Складова рівнів часового ряду, призначена для опису регулярних змін протягом заданого періоду, називається:

а) сезонною складовою;

б) трендовою складовою;

в) циклічною складовою.

Практичні завдання

1. Для поданих вихідних даних (табл. 7.9) значень обсягів продажу фармацевтичних препаратів за 8 років у поквартальному аспекті необхідно побудувати модель декомпозиції часового ряду. Виділіть трендову, циклічну, сезонну та випадкову складові. Оцініть адекватність побудованої моделі. Побудуйте проноз за кварталами на 2015 рік.

Таблиця 7.9

Вихідні дані

Рік	2007				2008				2009				2010			
1	2				3				4				5			
Кв	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
$Y(t)$	74	115	125	71	80	123	147	55	88	126	162	74	100	152	181	80

Закінчення табл. 7.9

1	2	3	4	5												
Рік	2011				2012				2013				2014			
Кв	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
$Y(t)$	116	171	203	91	125	194	233	103	142	218	263	119	164	257	312	133

2. Для досліджуваного динамічного ряду виробництва продукції на підприємстві (табл. 7.10) за два роки у помісячному аспекті необхідно дослідити ряд виробництва продукції на наявність тренда за методами Стьюдента, Фішера та Фостера – Стьюдента. Зробити висновки про стаціонарність чи не стаціонарність ряду.

Таблиця 7.10

Вихідні дані

Період (t)	1	2	3	4	5	6	7	8	9	10	11	12
$Y(t)$	40,68	68,76	91,76	45,03	67,68	115,9	152,7	73,83	110	184	240	114,23
Період (t)	13	14	15	16	17	18	19	20	21	22	23	24
$Y(t)$	168,81	279,6	359,61	171	248,7	405,3	514,94	240,6	348	562	708	327,31

3. У табл. 7.11 наведені спостереження за величиною прибутку виробничого підприємства за 4 роки у поквартальному аспекті.

Таблиця 7.11

Вихідні дані, млн грн

Період (t)	1	2	3	4	5	6	7	8
$Y(t)$	20,05	21,7	25,6	27,6	29,7	33,2	36,5	39,7
Період (t)	9	10	11	12	13	14	15	16
$Y(t)$	41,45	42,5	41,45	42,5	43,1	46,2	46,35	47,9

На основі даних, що наведені в табл. 7.11 необхідно на основі методу характеристик здійснити вибір трендової моделі, яка найкраще описує тенденцію динаміки за досліджуваною змінною; оцінити параметри моделі, її адекватність; привести інтерпретацію отриманих результатів. Розрахувати прогнозні значення прибутку на 5 рік.

4. На основі наведених даних динаміки зміни показника курсу акцій підприємства (табл. 7.12), необхідно: побудувати графік вихідних даних; провести згладжування за простою та зваженою ковзною середньою з різною довжиною інтервалів; оцінити точність отриманих варіантів прогнозної моделі та обрати найкращий з них за критерієм середньої абсолютної процентної помилки.

Таблиця 7.12

Вихідні дані, грн

t	1	2	3	4	5	6	7	8	9	10
y_t	10,8	11	10,82	10,5	10,35	10,35	10,15	9,77	9,57	9,32
t	11	12	13	14	15	16	17	18	19	20
y_t	9,59	9,32	10,5	9,85	9,85	9,85	9,85	9,90	9,85	9,85

5. Відомі дані про зростання прибутку підприємства (млн грн) у зв'язку з розширенням випуску продукції. Необхідно перевірити за допомогою тесту припущення щодо поліноміальної моделі взаємозалежності. Визначити параметри нелінійної моделі за допомогою МНК, припускаючи, що в якості моделі можна використовувати поліном 2-го ступеня, оцінити її адекватність; привести інтерпретацію отриманих результатів. Розрахувати прогнозні значення прибутку на 2 попередні роки.

Таблиця 7.13

Вихідні дані

t	1	2	3	4	5	6	7	8	9	10
Y	2,5	3,1	4,1	5	6,5	7,1	8,5	9,2	10	11,5

Ключові слова

Динамічна економетрична модель. Динамічний ряд. Адитивна модель часового ряду. Мультиплікативна модель часового ряду. Випадкова складова. Стохастична складова. Тренд. Циклічна складова. Сезонна складова. Метод Фостера – Стюарта. Стаціонарний динамічний ряд. Тенденція автокореляції. Тенденція дисперсії. Тенденція середнього рівня.

Розділ 8. Моделі розподіленого лага

8.1. Загальна характеристика та класифікація моделей з лаговими змінними.

8.2. Обґрунтування величини лага. Моделі з поліноміальними лагами.

8.3. Моделі з геометричними лагами.

8.4. Інструментальні змінні.

8.1. Загальна характеристика та класифікація моделей з лаговими змінними

Припущення про те, що на поточне значення залежної змінної впливають тільки поточні значення факторних ознак є справедливим, якщо використовуються просторові дані. У динамічних моделях залежна змінна y_t може бути пов'язана зі значеннями пояснювальних змінних x не тільки в момент часу t, але і з їх значеннями в попередні моменти часу. Так, наприклад, величина попиту на житлову нерухомість визначається величиною доходу не тільки поточного періоду, а й попередніх років:

$$y_t = f(x_t, x_{t-1}, \dots, x_{t-k}) + \varepsilon_t,$$

де y_t – величина попиту на нерухомість в t-й період часу;

$x_t, x_{t-1}, \dots, x_{t-k}$ – дохід у період часу t;

$t-1, \dots, t-k$ – лагові змінні.

Лагові змінні – пояснюючі змінні, взяті в моделі регресії з запізненням у часі. *Лаг* – величина інтервалу запізнювання. Прикладами урахування лага в моделі є дослідження залежності випуску продукції від інвестицій, споживання товарів від доходу, обсягу продажу від бюджету витрат на рекламу і т. д.

У якості лагової може розглядатися не тільки факторна, а й залежна змінна. Зокрема, на величину дивідендних виплат впливає не тільки значення прибутку корпорацій у поточний період, але і величина дивідендних виплат у попередній період часу: $y_t = f(x_t, y_{t-1}) + \varepsilon_t$, де y_t – величина дивідендних у t-й період часу, x_t – прибуток у t-й період часу, y_{t-1} – дивідендні виплати в період часу $t-1$ (лагова змінна).

Моделі з лаговими змінними можна поділити на такі основні класи [37]:

1) моделі з розподіленими лагами:

$y_t = f(x_t, x_{t-1}, \dots, x_{t-k}) + \varepsilon_t$;

2) моделі з лаговими залежними змінними (моделі авторегресії):

$y_t = f(y_{t-1}, \dots, y_{t-k}) + \varepsilon_t$;

3) моделі з лаговими залежними і незалежними змінними (авторегресійні моделі з розподіленими лагами):

$y_t = f(y_{t-1}, \dots, y_{t-k}, x_t, x_{t-1}, \dots, x_{t-k}) + \varepsilon_t$.

Моделі розподіленого лага включає такі основні типи моделей:

модель із кінцевим числом лагів: $y_t = c + a_0 x_t + a_1 x_{t-1} + \dots + a_k x_{t-k} + \varepsilon_t$;

модель із нескінченним числом лагів: $y_t = c + a_0 x_t + a_1 x_{t-1} + a_2 x_{t-2} + \dots + \varepsilon_t$.

Більш поширеними в економічних дослідженнях є моделі з кінцевим числом лагів. Коефіцієнти $a_j,\ j = 1,2,\dots,k$ називаються коефіцієнтами лага, а послідовність $a = \{a_j : j = 1,2,\dots,k\}$ – структурою лага. Коефіцієнт a_0 називають короткостроковим мультиплікатором, оскільки він характеризує зміну результівної змінної під час зміни факторної змінної на 1 у фіксований момент часу t. Будь-яку суму коефіцієнтів $\sum_{j=0}^{m} a_j$, де $m < k$ називають проміжним мультиплікатором, а суму всіх лагових коефіцієнтів $\sum_{j=0}^{k} a_j$ – довгостроковим мультиплікатором, який характеризує загальну зміну результуючої змінної під час зміни факторної ознаки на 1 через k періодів часу. Якщо виконується умова: $a_j > 0,\ j = 1,2,\dots,k$, то можна визначити *нормовані коефіцієнти лага* $b_j = \frac{a_j}{\sum_{j=0}^{k} a_j}$, *які показують пропорцію довгострокового вплив*у, що припадає на певні період часу. Послідовність коефіцієнтів $b = \{b_j : j = 1,2,\dots,k\}$ називається *нормованою структурою лага для моделі*. Нормовані коефіцієнти можна використовувати як вагові для розрахунку *середньої величини лага*, яка показує середній інтервал часу, протягом якого буде відбуватися зміна залежної змінної під впливом змінної, що пояснює, в момент часу t: $\bar{j} = \sum_{j=0}^{k} j b_j$. Чим менше значення середнього лага, тим більш чутливою є результуюча змінна до зміни факторної ознаки. Високі значення середнього лага показують, що вплив факторної ознаки на результуючу змінну будуть проявлятися з плином тривалого періоду часу. Характеристикою структури лага є *медіанний* лаг, який показує тривалість періоду, необхідного для реалізації половини загальної зміни результуючої ознаки після збільшення значень факторної змінної на 1. Для медіанного лага справедлива рівність: $\sum_{j=0}^{M_e} b_j = 0{,}5$.

Розглянемо економічну інтерпретацію параметрів моделі з кінцевим числом лагів на прикладі моделі залежності обсягу інвестицій в економіку країни (y_t) від величини оцінки її інвестиційної привабливості (x_t), яка здійснюється міжнародними рейтинговими агентствами:

$$y_t = 2{,}091 + 0{,}6x_t + 1{,}1x_{t-1} + 0{,}8x_{t-2}.$$

Аналіз коефіцієнтів лага дозволяє зробити висновок, що збільшення рейтингу інвестиційної привабливості країни на 1 призводить до припливу інвестицій в поточному періоді, рівному 0,6 ум. грош. од. (короткостроковий мультиплікатор); через два роки збільшення інвестицій складе 1,7 ум. грош. од. (проміжний мультиплікатор); через три роки – 2,5 ум. грош. од. (довгостроковий мультиплікатор). Нормована структура лага має такий вигляд: $b = \ 0{,}24; \ 0{,}44; 0{,}32$. Тобто на поточний рік припадає 24 % впливу факторної ознаки, через рік – 44 %, через два роки – 32 %. Середній лаг становитиме 1,08 роки.

8.2. Обґрунтування величини лага. Моделі з поліноміальними лагами

Однією з проблем під час побудови моделей розподіленого лага є обґрунтування величини лага. Теоретично побудову моделі розподіленого лага можна узагальнити на будь-яку кількість незалежних змінних x_{t-k}. Проте практично реалізація такого підходу ускладнена у зв'язку з великим числом факторів і невеликою довжиною часового ряду. Для обґрунтування величини лага використовується *взаємна кореляційна функція*, яка характеризує тісноту зв'язку елементів вектора залежної змінної y_t з елементом вектора незалежної змінної x_t, зрушеним один щодо іншого на часовий лаг τ. Для побудови взаємної кореляційної функції розглядається безліч коефіцієнтів кореляції між рівнями часових рядів x_t та y_t, x_{t-1} та y_t,..., x_{t-k} та y_t. Для різних значень $\tau = 0,1, \ldots, k$ на основі взаємної кореляційної функції можна отримати $k+1$ значення r_τ:

$$r_\tau = \frac{(n-\tau) \sum_{t=1}^{n-\tau} y_t x_{t-\tau} - \sum_{t=1}^{n-\tau} y_t \sum_{t=1}^{n-\tau} x_{t-\tau}}{(n-\tau) \sum_{t=1}^{n-\tau} y_t^2 - \left(\sum_{t=1}^{n-\tau} y_t\right)^2 \ (n-\tau) \sum_{t=1}^{n-\tau} x_{t-\tau}^2 - \left(\sum_{t=1}^{n-\tau} x_{t-\tau}\right)^2}.$$

Величина лага визначається за максимальним значенням коефіцієнта кореляції. Якщо серед безлічі значень r_τ є кілька, величина яких за модулем наближається до 1, то це означає, що запізнювання впливу змінної x_t відбувається протягом певного проміжку часу, тобто спостерігається кілька часових лагів для взаємопов'язаних часових рядів.

Розглянемо побудову взаємної кореляційної функції на прикладі.

Приклад 8.1. Необхідно визначити величину лага у впливі темпу зростання ВРП регіону (x_t) на темп зростання ВРП сусіднього регіону (y_t). Вихідні дані наведені в табл. 8.1.

Таблиця 8.1

Вихідні дані

t	1	2	3	4	5	6	7	8	9	10
y_t	1,12	1,02	0,97	1,05	1,02	1,02	1,06	1,03	1,06	1,08
x_t	1,05	1,04	0,95	1,11	1,02	1,00	1,07	1,06	1,07	1,10
t	11	12	13	14	15	16	17	18	19	
y_t	1,14	1,09	1,07	1,07	1,05	1,01	1,05	1,12	1,22	
x_t	1,09	1,04	1,08	1,07	0,99	1,06	1,11	1,15	1,15	

Слід розрахувати коефіцієнт кореляції за умови $\tau = 0$. Парний коефіцієнт кореляції r_0 дорівнює 0,71. Послідовно зрушуючи рівні часового ряду x_t на один часовий інтервал, буде отримано серію коефіцієнтів (табл. 8.2).

Таблиця 8.2

Значення взаємної кореляційної функції

Величина лага τ	1	2	3	4	5
Коефіцієнт кореляції r_τ	0,75	0,38	0,12	-0,23	0,65

Найбільше значення коефіцієнта взаємної кореляції $r_\tau = 0{,}75$ відповідає лагу τ = 1. Тобто найбільший вплив підвищення рівня ділової активності в прилеглому регіоні буде спостерігатися через 1 рік. Модель розподіленого лага в цьому випадку може бути подана таким чином:

$$y_t = c + a_0 x_t + a_1 x_{t-1} + \varepsilon_t \, .$$

Другим підходом до обґрунтування величини лага є порівняння якості моделей з різним набором лагових змінних. У цьому випадку вибір здійснюється на користь моделі, для якої всі коефіцієнти регресії при лагових змінних будуть статистично значущі за t-критерієм Стьюдента.

Оскільки факторні ознаки моделі $y_t = f\ (x_t, x_{t-1}, \dots, x_{t-k})\ + \varepsilon_t$ відображають одне і те ж явище з деяким запізненням, то можна чекати сильної кореляційної залежності між ними. Проблему побудови моделей з лаговими змінними вирішують перетворенням цих змінних у нові відносно незалежні змінні. Найбільш поширеним методом оцінювання параметрів моделей із незалежними лаговими змінними є метод *Ширлі Алмон*.

В основі методу Ширлі Алмон лежить гіпотеза, що лагові коефіцієнти регресії апроксимуються поліномом відповідного ступеня від величини лага:

$$a_j = P_r\ (j)\ = b_0 + b_1 j + b_2 j^2 + \dots + b_r j^r.$$

У ході цього апріорі висувається припущення про ступінь полінома. Як правило, використовується многочлен невисокого ступеня ($r \leq 4$).

Наприклад, апроксимуємо параметри регресії з максимальним лагом $k\ =\ 5$

$$y_t = a_0 x_t + a_1 x_{t-1} + \dots + a_5 x_{t-5} + \varepsilon_t,$$

багаточленом третього ступеня $r\ =\ 3$, тобто

$$a_j = P_3\ (j)\ = b_0 + b_1 j + b_2 j^2 + b_3 j^3.$$

Виразимо параметри регресії a_j через параметри багаточлена $b_j (j = 0{,}3)$.

$$\begin{gathered} a_0 = P_3\ (0)\ = b_0, \\ a_1 = P_3\ (1)\ = b_0 + b_1 + b_2 + b_3, \\ a_2 = P_3\ (2)\ = b_0 + 2b_1 + 4b_2 + 8b^3, \\ a_3 = P_3\ (3)\ = b_0 + 3b_1 + 9b_2 + 27b_3, \\ a_4 = P_3\ (4)\ = b_0 + 4b_1 + 16b_2 + 24b_3, \\ a_5 = P_3\ (5)\ = b_0 + 5b_1 + 25b_2 + 125b_3. \end{gathered}$$

Після підстановки в рівняння регресії параметрів, виражених через коефіцієнти многочлена, і винесення за дужки оцінок коефіцієнтів многочлена, одержимо:

$$y_t = b_0(x_t + x_{t-1} + x_{t-2} + x_{t-3} + x_{t-4} + x_{t-5}) + \ b_1(x_{t-1} + 2x_{t-2} +$$
$$+ 3x_{t-3} + 4x_{t-4} + 5x_{t-5}) + b_2(x_{t-1} + 4x_{t-2} + 9x_{t-3} + 16x_{t-4} +$$
$$+ 25x_{t-5}) + b_3 \ \ x_{t-1} + 8x_{t-2} + 27x_{t-3} + 64x_{t-4} + 125x_{t-5} \ \ + \varepsilon_t.$$

Слід розглядати доданки в дужках при b_0, b_1, b_2, b_3 як нові змінні z. Модель з розподіленими лагами набуває вигляду:

$$y_t = b_0 z_{0t} + b_1 z_{1t} + b_2 z_{2t} + b_3 z_{3t} + \varepsilon_t \ ,$$

де $z_{0t} = x_t + x_{t-1} + x_{t-2} + x_{t-3} + x_{t-4} + x_{t-5}$,

$z_{1t} = x_{t-1} + 2x_{t-2} + 3x_{t-3} + 4x_{t-4} + 5x_{t-5}$,

$z_{2t} = x_{t-1} + 4x_{t-2} + 9x_{t-3} + 16x_{t-4} + 25x_{t-5}$,

$z_{3t} = x_{t-1} + 8x_{t-2} + 27x_{t-3} + 64x_{t-4} + 125x_{t-5}$.

Оцінки параметрів при перетворених змінних z визначаються МНК. Далі на основі параметрів b_0, b_1, b_2, b_3здійснюється перехід до оцінювання параметрів a_0, a_1, a_2, a_3, a_4, a_5. Якість оцінок коефіцієнтів b_i залежить від дисперсій і кореляцій між перетвореними факторами регресії.

У матричному вигляді можна записати, що $a = Hb$,

де

$$H = \begin{matrix} 1 & 0 & 0 & ... & 0 \\ 1 & 1 & 1 & ... & 1 \\ 1 & 2 & 4 & ... & 2^r \\ ... & ... & ... & ... & ... \\ 1 & k & k^2 & ... & k^r \end{matrix} \ .$$

Стандартна помилка коефіцієнтів регресії при лагових змінних визначається за формулою:

$$\sigma_{a_j} = \sigma_e \ \ \overline{H_i^T (z^T z)^{-1} H_i} \ .$$

Далі за допомогою критерію Стьюдента оцінюється статистична значущість коефіцієнтів. Якість моделі оцінюється через коефіцієнт детермінації R^2.

Як одиничний випадок оцінювання параметрів регресії з лаговими факторами методом Ширлі Алмон у роботах Джонстоном розглядається метод, який полягає у заміні многочлена інтерполяційною формулою Лагранжа. Многочлен $P_r\ j$ однозначно визначається його значеннями $P_r\ j_0\ ,\ P_r\ j_1\ ,\ P_r\ j_2\ ,\ \ldots, P_r\ j_r$ у $r+1$ точках $j_0,\ j_1,\ldots\ j_r$. Значення многочлена в довільній точці j можна знайти за допомогою інтерполяційної формули Лагранжа:

$$P_r\ j\ = \frac{j-j_1\quad j-j_2\ \ \ldots\ \ j-j_r}{j_0-j_1\quad j_0-j_2\ \ \ldots.(j_0-j_r)} P_r\ j_0\ + \frac{j-j_0\quad j-j_2\ \ \ldots\ \ j-j_r}{j_1-j_0\quad j_1-j_2\ \ \ldots.(j_1-j_r)} P_r\ j_1\ + \\ + \cdots + \frac{j-j_0\quad j-j_1\ \ \ldots\ \ j-j_{r-1}}{j_r-j_0\quad j_r-j_1\ \ \ldots.(j_r-j_{r-1})} P_r\ j_r\ .$$

Варто припустити, що значення многочлена $P_r\ j$ в точках $j_0,\ j_1,\ldots\ j_r$ відомі, тоді за формулою можна обчислити значення многочлена відповідно в точках 0,1, ..., 5. Після чого вихідна регресія запишеться у вигляді:

$$y_t = P_3\ 0\ x_t + P_3\ 1\ x_{t-1} + \cdots + P_3\ 5\ x_{t-5} + \varepsilon_t\ .$$

Оскільки значення многочлена $P_r\ j_0\ ,\ P_r\ j_1\ , P_r\ j_2\ ,\ldots, P_r\ j_r$ у точках $j_0, j_1, \ldots\ j_r$ невідомі, то завдання полягає в знаходженні оцінок їх значення на основі статистичних даних.

Для ілюстрації методу Алмон слід розглянути такий приклад.

Приклад 8.2. Необхідно побудувати модель залежності обсягу виробництва (y_t) від величини інвестицій (x_t). Вихідні дані наведені в табл. 8.3.

Таблиця 8.3

Вихідні дані

x

Роки	y_t	x_t	*Роки*	y_t	x_t
1	67,6	39,7	11	102,3	61,5
2	76,0	41,6	12	116,9	67,3
3	77,8	43,1	13	127,0	69,7
4	75,1	40,9	14	136,3	75,4
5	79,1	45,3	15	145,6	80,3
6	80,7	46,2	16	152,5	79,3
7	82,4	46,3	17	153,7	83,9
8	87,3	49,7	18	161,6	93,0
9	90,1	52,5	19	181,3	107,1
10	95,1	56,0	20	220,7	123,1

Передбачається, що обсяг виробництва (y_t) залежить від величини інвестицій (x_t) у поточному й трьох попередніх роках, тобто модель має вигляд:

$$y_t = \mathrm{c} + a_0 x_t + a_1 x_{t-1} + a_2 x_{t-2} + a_3 x_{t-3} + \varepsilon_t \ .$$

Також передбачається, що порядок апроксимуючого багаточлена дорівнює 2.

Представимо параметри моделі у вигляді багаточлена другого ступеня:

$$a_j = P_2 \ \ j \ \ = b_0 + b_1 j + b_2 j^2.$$

Відповідно одержимо:

$$\begin{gathered} a_0 = P_2 \ \ 0 \ \ = b_0, \\ a_1 = P_2 \ \ 1 \ \ = b_0 + b_1 + b_2, \\ a_2 = P_2 \ \ 2 \ \ = b_0 + 2b_1 + 4b_2, \\ a_3 = P_2 \ \ 3 \ \ = b_0 + 3b_1 + 9b_2 \ . \end{gathered}$$

Треба підставити ці значення у вихідне рівняння:

$$\begin{gathered} y_t = \mathrm{c} + b_0 x_t + \ \ b_0 + b_1 + b_2 \ \ x_{t-1} + \ \ b_0 + 2b_1 + 4b_2 \ \ x_{t-2} + \\ + \ \ b_0 + 3b_1 + 9b_2 \ \ x_{t-3} = \ \mathrm{c} + b_0(x_t + x_{t-1} + x_{t-2} + x_{t-3}) + \\ + b_1 \ \ x_{t-1} + 2x_{t-2} + 3x_{t-3} \ \ + b_2 \ \ x_{t-1} + 4x_{t-2} + 9x_{t-3} \ \ + \varepsilon_t. \end{gathered}$$

Розрахувати значення змінних z_{it} (табл. 8.4).

Таблиця 8.4

Значення перетворених факторів

Роки	z_{0t}	z_{1t}	z_{2t}	*Роки*	z_{0t}	z_{1t}	z_{2t}
1	165,35	245,48	567,02	10	254,48	358,32	817,34
2	170,96	251,96	587,83	11	273,90	388,79	892,43
3	175,55	256,48	596,94	12	292,73	416,68	959,85
4	178,79	259,61	595,79	13	304,72	440,20	1 009,09
5	187,54	274,72	639,08	14	318,92	466,22	1 079,43
6	194,76	280,94	650,79	15	336,52	483,45	1 124,02
7	204,56	290,92	668,32	16	363,29	498,64	1 142,12
8	219,72	310,11	713,22	17	407,12	544,77	1 234,08
9	237,36	331,15	758,45				

Використовуючи отримані z_{it}, буде одержано результати регресії:

$$y_t = -7{,}14 + 0{,}44z_{0t} + 0{,}6z_{1t} - 0{,}29z_{2t},$$

$$t_c = -3{,}58,\ t_{b_0} = 3{,}99,\ t_{b_1} = 1{,}87,\ t_{b_2} = -2{,}59,\ R^2 = 0{,}9967.$$

Усі параметри рівняння статистично значущі ($t_{\text{табл}}$ = 1.77 при k = 13). Значення коефіцієнта детермінації R^2 = 0,9967 дозволяє зробити висновок про хорошу якість моделі.

Відповідні оцінки параметрів вихідного рівняння визначаються в такий спосіб:

$$a_0 = b_0 = 0{,}44,$$
$$a_1 = b_0 + b_1 + b_2 = 0{,}44 + 0{,}6 - 0{,}29 = 0{,}75,$$
$$a_2 = b_0 + 2b_1 + 4b_2 = 0{,}44 + 2 \cdot 0{,}6 - 4 \cdot 0{,}29 = 0{,}49,$$
$$a_3 = b_0 + 3b_1 + 9b_2 = 0{,}44 + 3 \cdot 0{,}6 - 9 \cdot 0{,}29 = -0{,}35\,.$$

Таким чином, отримана модель розподіленого лага має вигляд:

$$y_t = -7{,}14 + 0{,}44x_t + 0{,}75x_{t-1} + 0{,}49x_{t-2} - 0{,}35x_{t-3}.$$

Розподіл лагових коефіцієнтів регресії наведений на рис. 8.1.

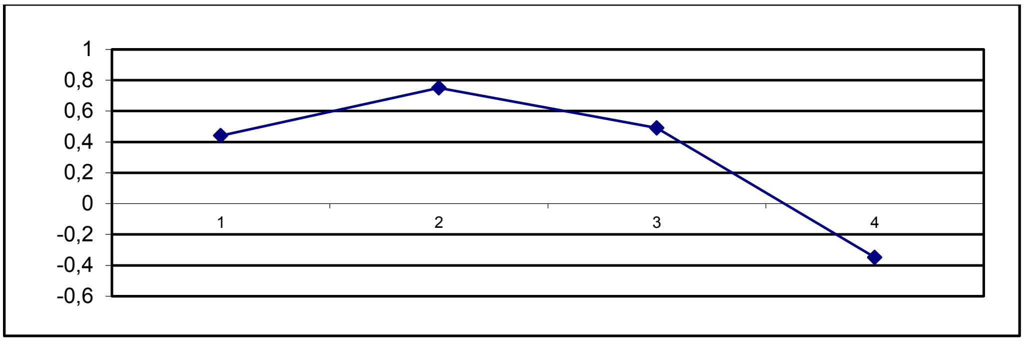


Рис. 8.1. **Розподіл лагових коефіцієнтів регресії**

Якщо до вихідних даних прикладу застосувати МНК, то результати регресійного аналізу такі:

$$y_t = -10{,}79 + 0{,}64x_t + 1{,}19x_{t-1} + 0{,}67x_{t-2} - 0{,}48x_{t-3},$$

$$t_c = -3{,}44,\ t_{a_0} = 2{,}83,\ t_{a_1} = 2{,}97,\ t_{a_2} = 1{,}6,\ t_{a_3} = -1{,}6, R^2 = 0{,}9968.$$

Незважаючи на те, що коефіцієнт детермінації вище, коефіцієнти лагових змінних x_{t-2} і x_{t-3} статистично незначущі.

8.3. Моделі з геометричними лагами

Для моделі з нескінченним числом лагових змінних, виду:

$$y_t = c + a_0 x_t + a_1 x_{t-1} + a_2 x_{t-2} + \cdots + \varepsilon_t$$

оцінювання параметрів неможливе без якого-небудь припущення щодо поведінки коефіцієнтів при лагових змінних. Одним із припущень є припущення про те, що лагові коефіцієнти зменшуються в геометричній прогресії: $a_k = a_0 \lambda^k$. З урахуванням цього припущення модель має вигляд:

$$y_t = c + a_0 x_t + a_0 \lambda x_{t-1} + a_0 \lambda^2 x_{t-2} + \cdots + \varepsilon_t. \qquad (8.1)$$

Наявність у моделі нескінченного числа лагових змінних не дозволять провести оцінювання параметрів. Тому використовується перетворення Л. Койка, яке полягає в тому, що в модель вводиться оператор зсуву на один період:

$$y_{t-1} = c + a_0 x_{t-1} + a_0 \lambda x_{t-2} + a_0 \lambda^2 x_{t-3} + \cdots + \varepsilon_{t-1}.$$

Помножуючи рівняння на λ буде одержано:

$$\lambda y_{t-1} = \lambda c + a_0 \lambda x_{t-1} + a_0 \lambda^2 x_{t-2} + a_0 \lambda^3 x_{t-3} + \cdots + \lambda \varepsilon_{t-1}\,. \qquad (8.2)$$

Відняти отримане рівняння (8.2) з вихідного (8.1):

$$y_t - \lambda y_{t-1} = \ 1 - \lambda\ c + a_0 x_t + \varepsilon_t - \lambda \varepsilon_{t-1}\,.$$

Після перетворення буде отримано:

$$y_t = \ 1 - \lambda\ c + a_0 x_t + \lambda y_{t-1} + u_t\,. \qquad (8.3)$$

Тобто від моделі рівняння з розподіленими лагами з нескінченним їх числом Л. Койк перейшов до моделі авторегресії, для якої потрібно оцінити параметри $c,\ \lambda,\ a_0$. Далі зі співвідношення $a_k = a_0\lambda^k$ визначаються параметри вихідної моделі (8.1). Перетворення Л. Койка приводить до суттєвих спрощень. Фактично проблема мультиколінеарності факторів усувається шляхом заміни x_{t-1}, x_{t-2} ...на y_{t-1}.

Модель Л. Койка дозволяє аналізувати короткостроковий і довгостроковий мультиплікатори. Короткостроковим мультиплікатором є параметр a_0, а довгостроковим – сума коефіцієнтів регресії, що є сумою геометричної прогресії:

$$\sum_{j=0}^{\infty} b_j = b_0 + b_0\,\lambda + b_0\lambda^2 + \cdots = b_o\left(1 + \lambda + \lambda^2 + \lambda^3 + \cdots\right) = \frac{b_0}{1-\lambda}.$$

Прикладом розподілу Л. Койка є *модель адаптивних очікувань*. Процес адаптивних очікувань – це процедура коригування очікувань, коли в кожний період часу реальне значення змінної порівнюється з її очікуваним значенням. Якщо реальне значення виявляється більше, то значення, очікуване в наступному періоді, коригується в бік його підвищення, якщо менше – у бік зменшення. Передбачається, що розмір коригування пропорційний різниці між реальними та очікуваними значеннями змінної.

Таким чином, якщо розглядається змінна x, а x_t^e – її значення, очікуване в період часу t, то

$$x_{t+1}^e - x_t^e = \lambda\left(x_t - x_t^e\right)\ (0 \le \lambda \le 1).$$

Цей вираз може бути подано у вигляді:

$$x_{t+1}^e = \lambda x_t + (1-\lambda)x_t^e\ (0 \le \lambda \le 1). \qquad (8.4)$$

Таким чином, значення змінної, очікуване в $t+1$ період часу, формується як зважене середнє її реального й очікуваних значень у поточному періоді. Чим більше величина λ, тим швидше очікуване значення адаптується до попередніх реальних значень. Слід зазначити, що процес адаптивних очікувань спрямований у майбутнє.

Припустимо, що залежна змінна y_t, пов'язана з очікуваним значенням x в році $t+1$:

$$y_t = \alpha + \beta x_{t+1}^e + u_t. \qquad (8.5)$$

У рівнянні (8.5) результуюча змінна виражена через величину x_{t+1}^{e}, що не спостерігається і яку необхідно замінити спостережуваними змінними. Запишемо вираз (8.4) для $t-1$ періоду:

$$x_t^e = \lambda x_{t-1} + (1-\lambda)x_{t-1}^e. \tag{8.6}$$

Замінити величину x_t^e у виразі (8.4) на рівняння (8.6):

$$x_{t+1}^e = \lambda x_t + \lambda(1-\lambda)x_{t-1} + (1-\lambda)^2 x_{t-1}^e.$$

Повторивши процедуру заміни, отримаємо:

$$x_{t+1}^e = \lambda\,(x_t + (1-\lambda)\,x_{t-1} + (1-\lambda)^2 x_{t-2} + \cdots).$$

Підставивши отриманий вираз у (8.5) і замінивши $(1-\lambda)$ на γ, отримаємо:

$$y_t = \alpha + \beta\,\lambda\cdot(x_t + \gamma x_{t-1} + \gamma^2 x_{t-2} + \cdots) + u_t,$$

звідки видно, що значення y_t визначається поточним і минулими значеннями x_t з лагами, котрі підпорядковуються розподілу Койка.

У моделі Л. Койка випадкова помилка $u_t = \varepsilon_t - \lambda\varepsilon_{t-1}$ корельована зі змінною y_{t-1}. Тому оцінювання параметрів моделі МНК дає зміщені і неспроможні оцінки. Замість МНК можуть бути застосовані інструментальні змінні.

8.4. Інструментальні змінні

Пояснимо використання методу інструментальних змінних (ІЗ). Цей метод також має велике значення під час оцінювання параметрів моделей, що складаються з декількох рівнянь.

Метод інструментальних змінних полягає у частковій заміні непридатної змінної, що пояснює, такою змінною, що не корельована з випадковим членом. Обмежимося випадком парної регресії:

$$y = \alpha + \beta x + u$$

і припустимо, що з якої-небудь причини x має випадкову складову, залежну від u. Будемо також припускати, що у великих вибірках $V_{ar(x)}$ прагне до кінцевої межі σ_x^2. У цих умовах безпосереднє застосування МНК для побудови регресійної залежності y від x призвело б до неспроможних оцінок параметрів.

Тепер припустимо, що можна знайти іншу змінну z, яка корельована з x, але не коррельована з u. Покажемо, що засноване на використанні інструментальних змінних оцінювання параметра, що визначається як:

$$b_{\text{І3}} = \frac{cov(z,y)}{cov\ (z,x)}$$

є спроможною за умови, що у разі зростаючого числа спостережень $cov(z,y)$ прагне до кінцевої, відмінної від нуля межі, яку слід позначити як σ_{xz}. Це означає, що у великих вибірках $b_{\text{І3}}$ прагне до істинного значення β. Перед цим корисно порівняти $b_{\text{І3}}$ з оцінкою МНК, яку позначимо як $b_{\text{мнк}}$:

$$b_{\text{мнк}} = \frac{cov(x,y)}{var(x)} = \frac{cov(x,y)}{cov(x,x)}.$$

Оцінка ІЗ у парному регресійному аналізі виходить шляхом підстановки інструментальної змінної z замість x в чисельнику і замість одного x (але не обох) в знаменнику.

Можна записати вираз для $b_{\text{І3}}$ таким чином:

$$b_{\text{І3}} = \frac{cov\ \ z,y}{cov\ \ \ z,x} = \frac{cov\ \ z,\ \ \alpha+\beta x+u}{cov\ \ \ z,x} = \frac{cov\ \ z,a}{cov\ \ \ z,x} + \frac{cov\ \ z,\beta x}{cov\ \ \ z,x} + \frac{cov\ \ z,u}{cov\ \ \ z,x} =$$
$$= \beta + \frac{cov(z,u)}{cov\ (z,x)},$$

оскільки $cov(z,a)$ дорівнює нулю (a є постійною) і $cov(z,\beta x)$ дорівнює $\beta cov(z,x)$. Таким чином, можна помітити, що оцінка за методом інструментальних змінних дорівнює істинному значенню плюс помилка, рівна $\frac{cov(z,u)}{cov\ (z,x)}$. У великих вибірках помилка зникає, оскільки

$$plimb_{\text{І3}} = \beta + \frac{plimcov(z,u)}{plimcov\ (z,x)} = \beta + \frac{0}{\sigma_{x,z}} = \beta,$$

за умови, що змінна z дійсно розподілена незалежно від u. Отже, на великих вибірках $b_{\text{І3}}$ буде прагнути до істинного значення β.

Майже нічого не можна зазначити про розподіл оцінки $b_{\text{І3}}$ на малих вибірках, але під час збільшення n її розподіл буде прагнути до нормального з математичним очікуванням β і дисперсією:

$$pop.var\ \ b_{\text{І3}}\ \ \to \frac{\sigma_u^2}{npop.var\ \ x} \cdot \frac{1}{r_{x,z}^2},$$

де $r_{x,z}^2$ – вибірковий коефіцієнт кореляції між x і z.

Порівняти отриманий вираз із дисперсією оцінки МНК:

$$pop.var\ \ b_{\text{МНК}}\ \ \to \frac{\sigma_u^2}{nvar\ \ x}.$$

Основна відмінність полягає у тому, що дисперсія $b_{\text{І3}}$ множиться на $\frac{1}{r_{x,z}^2}$. Чим тісніше кореляція між x і z, тим менше буде цей коефіцієнт а, отже, тим меншою буде дисперсія $b_{\text{І3}}$. Отже, якщо постає проблема вибору між декількома можливими інструментальними змінними, то слід вибрати найбільш тісно корельовану з x, тому що за інших рівних умов вона дасть найбільш ефективні оцінки. Разом із тим було б небажано використовувати інструментальну змінну, яка б повністю корелювала з x, навіть якби її вдалося знайти, бо тоді вона автоматично стала б корельованою також і з u, і як і раніше було отримано б неспроможні оцінки. У цьому випадку потрібна інструментальна змінна, найбільш тісно корельована з x, але без кореляції з u.

Нехай розглядається модель авторегресії вигляду:

$$y_t = a + b_0 x_t + c y_{t-1} + u_t.$$

Застосування для оцінювання параметрів рівняння МНК можливо, якщо виконується передумова МНК щодо автокореляції залишків. У разі наявності в правій частині лагової залежної змінної може мати місце автокореляція залишків. Крім того, може мати місце залежність між y_{t-1} і u_t, тобто порушується припущення про гомоскедастичність залишків. Тому замість лагової залежної змінної y_{t-1} слід ввести в модель інструментальну змінну z_t, яка тісно корельована з лаговою змінною y_{t-1}, але не корельована залишками u_t, тобто перейти до моделі вигляду:

$$y_t = a + b_0 x_t + c z_t + u_t.$$

У якості z_t візьмемо оцінку y_{t-1}, тобто y_{t-1}, отриману за регресією:

$$y_{t-1} = A + Bx_{t-1}.$$

Далі застосовуємо МНК до моделі:

$$y_t = a + cA + b_0 x_t + cBx_{t-1} + u_t.$$

Слід зазначити, що застосування розглянутої інструментальної змінної може призвести до колінеарності факторів. Однак якщо колінеарність факторів не привела до великих стандартних помилок оцінок, неправильних знаків у коефіцієнтів регресії, то застосування інструментальної змінної можна вважати можливим.

Контрольні запитання для самодіагностики

1. Які причини визначають лагові ефекти в економетричних моделях?

2. Назвіть основні класи моделей з лаговими змінними.

3. Наведіть приклад моделі з розподіленими лагами. Яким чином інтерпретуються лагові коефіцієнти?

4. У чому статистична складність оцінювання параметрів з урахуванням лагових ефектів звичайними методами?

5. У якому випадку можна використовувати МНК для оцінювання параметрів моделі з розподіленими лагами?

6. У чому полягає ідея методу Ширлі Алмон?

7. Яким чином оцінюються параметри лагових моделей у методі Джонстона?

8. У якому випадку доцільно використовувати метод Койка?

9. У чому суть моделі "адаптивних очікувань"?

10. Наведіть перетворення Койка.

11. Наведіть переваги та недоліки методів оцінки параметрів лагових моделей.

12. У якому випадку використовуються інструментальні змінні?

Тести

1. Моделі, в яких на залежну змінну роблять вплив не тільки поточні значення факторної змінної, але і її значення в попередні періоди часу називаються:

а) дистрибутивно-лаговими моделями;

б) авторегресійними моделями;

в) авторегресійними моделями з розподіленими лагами.

2. Модель $y_t = c + b_1 y_{t-1} + b_2 y_{t-2} + a_0 x_t + a_1 x_{t-1} + a_2 x_{t-2} + \varepsilon_t$ *належить до класу:*

а) моделей з розподіленими лагами;

б) моделей з лаговими залежними змінними (моделей авторегресії);

в) моделей з лаговими залежними і незалежними змінними (авторегресійних моделей з розподіленими лагами).

3. Структурою лага називається:

а) послідовність коефіцієнтів $a = \ a_j: j = 1,2,\dots,k$, де k – величина лага;

б) послідовність коефіцієнтів $b = \ b_j: j = 1,2,\dots,k$, де $b_j = \frac{a_j}{\sum_{j=0}^{k} a_j}$;

в) усі відповіді правильні.

4. Довгостроковий мультиплікатор показує:

а) пропорцію довгострокового впливу, що припадає на певний період часу;

б) середній інтервал часу, протягом якого буде відбуватися зміна залежної змінної;

в) загальну зміну результуючої змінної у ході зміни факторної ознаки на 1 через k періодів часу.

5. Короткостроковий мультиплікатор характеризує:

а) зміну результуючої змінної у ході зміни факторної змінної на 1 у фіксований момент часу t;

б) тривалість періоду, необхідного для реалізації половини загальної зміни результуючої ознаки після збільшення значень факторної змінної на 1.

6. Для обґрунтування величини лага застосовується:

а) взаємна кореляційна функція;

б) порівняння критеріїв якості моделей з різним набором лагових змінних;

в) всі відповіді правильні.

7. Якщо взаємна кореляційна функція набуває таких значень $r_0 = 0{,}86, r_1 = 0{,}44, r_2 = 0{,}36,\ r_3 = 0{,}45$:

а) величина лага дорівнює 0;

б) величина лага дорівнює 1;

в) величина лага дорівнює 3.

8. В основі методу Ширлі Альмон лежить гіпотеза:

а) про поліноміальний розподіл лагових коефіцієнтів;

б) про геометричний розподіл лагових коефіцієнтів.

9. Статистична значущість параметрів моделі з розподіленим лагом оцінюється за допомогою:

а) стандартної помилки моделі;

б) критерію Фішера;

в) критерію Стьюдента.

10. Оцінювання параметрів моделі з розподіленим лагом стандартними методами ускладнене через наявність:

а) автокореляції;

б) гетероскедастичності;

в) мультиколінеарності.

11. Для оцінювання параметрів моделі з поліноміальними лагами використовується:

а) метод Ширлі Алмон;

б) метод Койка;

в) інструментальні змінні;

г) усі відповіді правильні.

12. Моделі, які містять лагові значення залежної змінної, – це:

а) авторегресійні моделі;

б) моделі адаптивних очікувань;

в) усі відповіді правильні.

13. У методі Койка для усунення мультиколінеарності здійснюється:

а) перетворенням вихідної моделі в авторегресійну;

б) формування штучних змінних, які є лінійною комбінацією вихідної системи чинників, але не перебувають у кореляційного зв'язку між собою;

в) усі відповіді правильні.

14. Для оцінювання параметрів моделі з нескінченним числом лагів, в якій лагові коефіцієнти зменшуються в геометричній прогресії, застосовується:

а) метод Ширлі Алмон;

б) метод Джонстона;

в) метод Койка.

15. Для оцінювання лагових коефіцієнтів в умовах автокореляції помилок використовується:

а) метод Ширлі Алмон;

б) метод Джонстона;

в) метод Койка;

г) метод Фішера;

д) інструментальні змінні.

Практичні завдання

1. Подано ряди динаміки обсягів продажів двох дочірніх підприємств (табл. 8.5). Необхідно для даних рядів знайти значення коефіцієнтів кореляції без урахування часового лага і з урахуванням лага. Зробити висновки.

Таблиця 8.5

Вихідні дані

t	1	2	3	4	5	6	7	8	9	10
Z_t	3	3,99	4,41	4,73	4,99	4,73	4,41	4,00	3,00	2,00
Z_{0t}	3,25	3,70	3,86	3,96	3,96	3,70	3,50	3,26	2,74	2,29

2. Подано ряди динаміки індексів ділової активності фондових ринків (табл. 8.6). Необхідно обґрунтувати величину лага. Побудувати модель розподіленого лага. Оцінити параметри моделі МНК методом Ширлі Алмон. Провести оцінювання статистичної значущості параметрів за допомогою критерію Стьюдента. Зробити висновки.

Таблиця 8.6

Вихідні дані

t	1	2	3	4	5	6	7	8	9	10
X_t	3,9	4,2	4,5	4,8	5,1	5,4	5,7	6,0	6,0	6,6
Y_t	12,6	13,2	13,8	14,4	15,0	15,9	16,8	17,7	18,6	19,5

3. На основі даних про величину інвестицій, податкове навантаження (табл. 8.7) побудувати модель адаптивних очікувань.

Таблиця 8.7

Вихідні дані

t	1	2	3	4	5	6	7	8	9	10
X_t	32	23	35	27	22	16	22	17	11	7
Y_t	90	150	250	446	184	292	454	716	270	436

4. На основі даних про обсяг продажів і величину витрат на рекламу (табл. 8.8) оцінити параметри моделі з лаговими змінними:

$$y_t = a + b_0 x_t + c y_{t-1} + u_t.$$

Для оцінювання параметрів моделі використовувати МНК, інструментальну змінну. Порівняти якість моделі, оціненої МНК, і моделі з інструментальною змінною. Зробити висновки.

Таблиця 8.8

Вихідні дані

t	1	2	3	4	5	6	7	8	9	10	11	12
X_t	52	54	59	51	70	67	62	63	71	69	77	113
Y_t	109	108	110	112	115	118	121	124	125	127	131	134

5. У табл. 8.9 подано дані динаміки інвестиційних вкладень та обсягу продукції, що випускається промисловим підприємством. Побудувати модель з розподіленим лагом, визначити всі її характеристики, провести оцінювання адекватності. Дати економічну інтерпретацію результатів моделювання.

Таблиця 8.9

Вихідні дані

t	1	2	3	4	5	6	7	8	9	10
X_t	92	104	115	104	110	104	115	121	127	133
Y_t	58	104	104	92	104	115	138	138	115	121
t	11	12	13	14	15	16	17	18	19	20
X_t	127	133	138	150	156	161	173	179	179	173
Y_t	127	144	156	156	173	161	184	184	173	184

Ключові слова

Лаг. Лагові змінні. Моделі з лаговими змінними. Модель з розподіленим лагом. Короткостроковий мультиплікатор. Довгостроковий мультиплікатор. Медіанний лаг. Взаємна кореляційна функція. Метод Ширлі Алмон. Метод Джонстона. Метод Койка. Модель адаптивних очікувань. Інструментальні змінні.

Розділ 9. Економетричні моделі на основі системи структурних рівнянь

9.1. Сутність структурного моделювання.

9.2. Системи одночасних рівнянь: класифікація, ідентифікація, специфікація.

9.3. Методи оцінки параметрів структурних рівнянь: НМНК, ДМНК, ТМНК.

9.1. Сутність структурного моделювання

В основі економічного аналізу діяльності підприємства лежить виявлення, оцінювання та прогнозування впливу факторів на зміну результативних показників. Функціонування будь-якої системи, а також підсистем здійснюється в умовах складної взаємодії комплексу чинників внутрішнього і зовнішнього порядку. Всі фактори впливають на систему і визначають її поведінку, перебувають у взаємозв'язку і взаємозумовленості (сутність структурного моделювання). Зв'язок економічних явищ – спільне зміна двох або більше явищ. Серед багатьох форм закономірних зв'язків явищ важливу роль відіграє причинний, сутність якої полягає в породженні одного явища іншим (основа структурного моделювання – причинно-наслідкові зв'язки) [19].

Кількісна характеристика взаємопов'язаних явищ здійснюється за допомогою ознак (показників). Ознаки, що характеризують причину, називаються факторними (незалежними, екзогенними); ознаки, що характеризують наслідок, називаються результативними (залежними). Сукупність факторних і результативних ознак, пов'язаних причинно-наслідковими зв'язками називають факторною системою (структурною моделлю). Під час вивчення зв'язків при дослідженні соціально-економічних та фінансових процесів вирішується кілька завдань [18]:

- установлення факту наявності або відсутності зв'язку між досліджуваними показниками;
- вимірювання тісноти зв'язку;
- установлення невипадкового характеру виявлених зв'язків;
- кількісне оцінювання впливу зміни факторів на зміну результативного показника;
- виділення найбільш значущих чинників, що визначають поведінку результативного показника.

Застосування методів структурного моделювання в складних соціально-економічних системах пояснюється такими причинами [27]:

1) необхідно вивчити вплив факторів, за якими не можна побудувати жорстко детерміновану факторну модель;

2) необхідно вивчити вплив складних чинників, які не піддаються об'єднанню;

3) необхідно обґрунтувати вплив факторів, які не можуть бути виражені одним кількісним показником (сукупні показники, інтегральні).

Однак, процес моделювання може ускладнюватися тим, що для побудови адекватних моделей необхідна наявність певної сукупності факторів, а також достатній обсяг спостережень.

Структурне моделювання має сенс тільки в тому випадку, якщо виділені фактори піддаються хоча б мінімального управлінню, тобто прямому чи непрямому впливу з боку управлінського персоналу. Розрахунки заради розрахунків деякою мірою безглузді. Структурні моделі будуються саме для того, щоб зрозуміти механізм взаємозв'язку тих чи інших сторін досліджуваних складних об'єктів, спробувати визначити ключові фактори, якими можна усвідомлено управляти, тим самим впливаючи на кінцеві результати. Моделювання структурними рівняннями – це всеосяжна і надзвичайно потужна техніка аналізу складних об'єктів, що включає велику кількість методів із різних областей економіко-математичного моделювання [18; 19].

Отже, побудова комплексної економетричної моделі є багатоетапним процесом послідовного удосконалення та розширення моделі на основі статистичного аналізу та експериментальної верифікації результатів. Очевидно, що визначення чисельних оцінок моделі повинен передувати теоретичний, економічний, економетричний аналізи, відбір змінних та зв'язків між ними, попередня обробка статистичних рядів. Рівняння і змінні можуть бути подані в декількох варіантах. Експериментальна перевірка варіантів може привести до розгляду нових змінних, нових рівнянь і нової їх верифікації (оцінювання якості отриманих результатів). Наступний експериментальний аналіз моделі в цілому знову може внести корективи і викликати необхідність додаткових перевірок. Комплексна економетрична модель є сукупністю рівнянь, що описують зв'язки між деякими показниками. Співвідношення між деякими змінними можуть мати стохастичний і детермінований характер. Стохастичні зв'язки реалізуються з деякою мірою вірогідності і описуються регресійними рівняннями.

Детерміновані співвідношення виражаються тотожністю і не містять випадкових величин. Для багаторозмірних моделей отримання чисельних оцінок параметрів і вивчення їх властивостей – це досить складне завдання.

Загальний процес побудови та застосування комплексної економетричної моделі наведено на рис. 9.1.

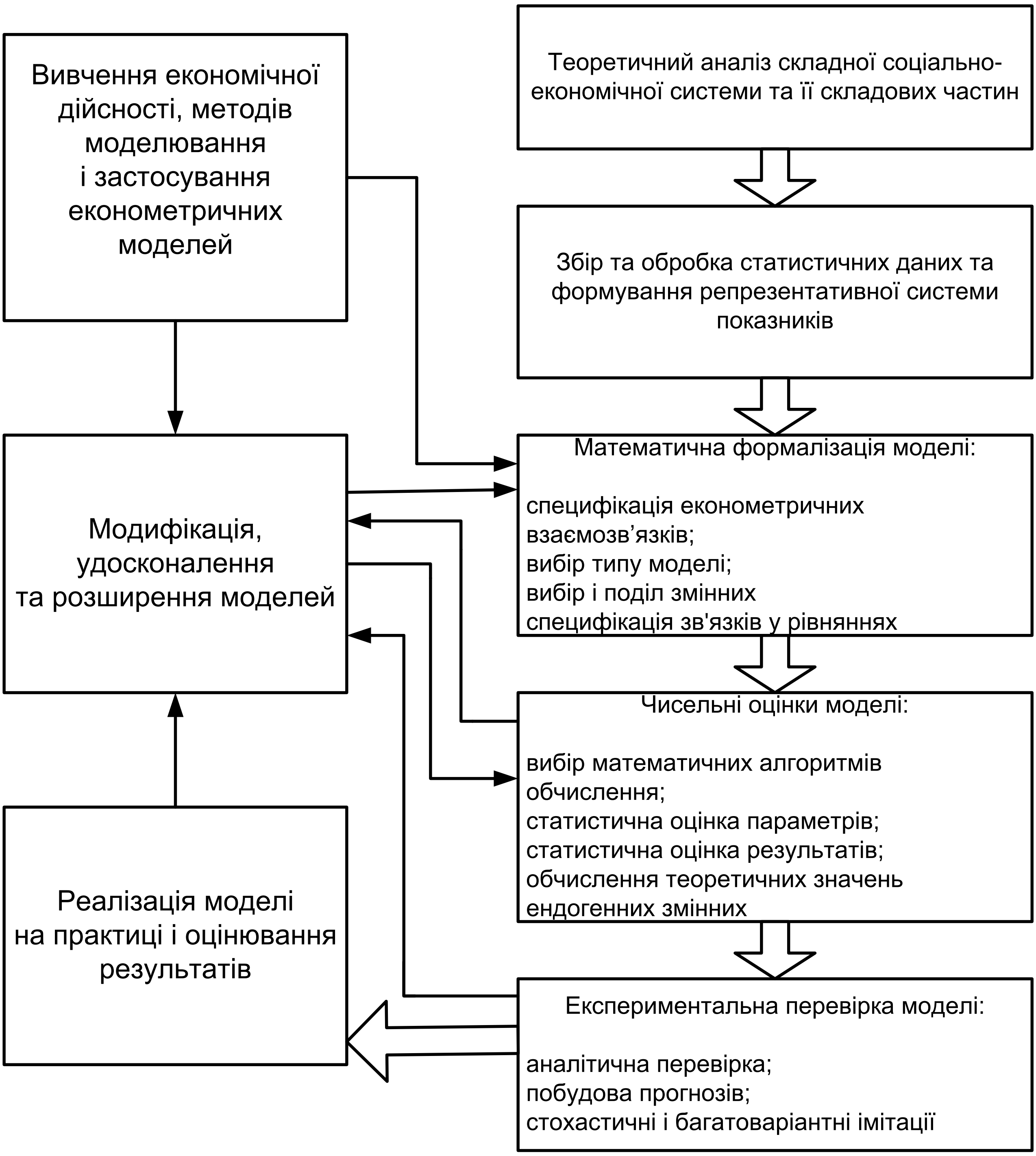


Рис. 9.1. **Процес побудови та застосування комплексної економетричної моделі**

Побудова системи структурних рівнянь дозволяє глибше вивчити причини зв'язку, що лежать в основі варіації результуючих змінних. При цьому відбувається визначення й оцінка непрямих (опосередкованих) і безпосередніх (прямих) впливів ознак. Саме тому системи структурних рівнянь часто інтерпретуються як статистичні описи причинно-наслідкових зв'язків, як причинні моделі, які пояснюють механізм формування варіації вихідних характеристик системи (результативних ознак). При використанні апарату кореляційно-регресійного аналізу структурне моделювання реалізує спробу подолати непрямий характер вивчення зв'язків цим методом,то більш точно підійти до визначення та вимірювання причинних (безпосередніх) зв'язків між змінними.

Отже, якщо головна задача стоїть в уточненні характеру зв'язків змінних, то ефективним є застосування "путьового аналізу" (path analysis, p-аналіз), в основу якого покладене активне використання графа зв'язків. Метод "путьового аналізу" (або путьових коефіцієнтів) запропонований у 1920-х рр. американським генетиком С. Райтом [31].

Головним припущенням методу є положення про адитивність і лінійність зв'язків між змінними. Використання лінійних залежностей між усіма змінними робить p-аналіз спеціальним додатком регресійного аналізу, при якому коефіцієнти регресії інтерпретуються в термінах причинно-наслідкових відносин.

"Путьовий аналіз" дозволяє провести декомпозицію кореляції та визначити наступні поняття, такі, як "повний (сукупний) зв'язок", "сукупний вплив", "прямий вплив", "непрямий вплив". Якщо коефіцієнт кореляції (r_{ij}) нульового порядку розглядати як ступінь вимірювання повного зв'язку двох змінних, то мірою сукупного впливу j-ї змінної на i-у змінну (q_{ij}) буде її частина, яка не залежить ні від спільних для них змінних – причин, ні від кореляції між загальними для j-ї та i-ї змінними причинами (компоненти "помилкової кореляції"), ні від наявності в моделі апріорної кореляції зумовлених змінних – входів, що, наприклад, не досліджуються в повному обсязі.

Таким чином ми можемо розкласти повний зв'язок двох змінних на чотири складові з урахуванням асиметрії впливу: сукупний вплив (причинний вплив) j-ї змінної на i-у на дві компоненти, що вимірюють ефект помилкової кореляції, і на випадкову компоненту. У свою чергу, сукупний вплив може бути розкладено на дві складові з урахуванням того, яким чином він здійснюється – безпосередньо або через інші змінні.

Прямий вплив однієї змінної на іншу вимірюється коефіцієнтом (p_{ij}); в цьому випадку в ланцюзі між пояснюючою та пояснювальною змінними немає проміжних ланок. Непрямий вплив – це вплив тих складових сукупного впливу однієї змінної на іншу, який утворюється при врахуванні ефекту передачі впливу за посередництвом змінних, специфікованих в моделі як проміжні ланки в причинному ланцюзі, який пов'язує досліджувані змінні. Але, постулюючи зв'язки між змінними, важко уникнути суб'єктивності. Два дослідники, маючи одні й ті ж дані, можуть отримати дві різні діаграми для аналізу. Так що *path*-аналіз – це не досить досконалий метод аналізу, який запрограмований заздалегідь, оскільки він вимагає творчого підходу.

Досліджуючі розглянуті питання Х. Блейлок запропонував формальний прийом, заснований на ідеях Г. Саймона про хибну кореляції і каузальну впорядкованість (процедура Саймона – Блейлока). Зміст процедури Саймона – Блейлока полягає в гіпотезі про повну специфіковану лінійну рекурсивну причинну модель, оцінку її параметрів, а потім використання цих значень для відтворення емпіричної кореляційної матриці. Основна ідея полягає в тому, що модель, яка не відтворює емпіричних кореляцій, повинна бути відкинута.

Доцільність застосування процедури Саймона – Блейлока очевидна, коли відомий причинний пріоритет серед змінних. Якщо існують дві гіпотези, які постулюють різні причинні ланцюги (структури графа), то, використовуючи процедуру Саймона – Блейлока, можна відтворити емпіричні кореляції і відкинути ті каузальні ланцюги, де неузгодженість занадто велика. Таким чином можливо порівнювати різні теорії.

Інший рекурентний прийом – викреслення зв'язків у надмірно пов'язаному графі з метою вивчення поведінки системи та її елементів у нових умовах. Стійкість системи може означати вірність гіпотези. Рішення про виключення того чи іншого зв'язку в моделі може бути прийняте або на основі критерію статистичної значущості, або на основі довільно встановленого порогового критерію величини коефіцієнта причинного впливу. Перевіркою правильності гіпотез і коректності моделі підтверджується при випробуваннях на контрольних даних.

Використання p-аналізу в соціально-економічних дослідженнях пов'язане з рядом труднощів:

1) насамперед не завжди можна вважати, що лінійна залежність повністю задовольняє та відображує всі можливі причинно-наслідкові зв'язки в реальних структурах;

2) слід враховувати, що p-аналіз розроблений для кількісних змінних. Структурні моделі та "путьовий аналіз" ілюструють єдність теоретичного (якісного) і формального математичного (кількісного) підходів;

3) значущість результатів аналізу визначається, в першу чергу, правильністю побудови логічного каркасу структурної моделі – максимально пов'язаного графа зв'язків;

4) застосування "путьового аналізу" доцільно при обмеженому числі змінних (приблизно при 10 змінних).

9.2. Системи одночасних рівнянь: класифікація, ідентифікація, специфікація

Складні соціально-економічні системи, та процеси, що в них протікають, для більш досконального вивчення доцільно представити за допомогою системи взаємопов'язаних (одночасних) рівнянь. Залежно від характеру взаємозв'язків між ендогенними та екзогенними досліджуваними змінними система рівнянь може бути представлена по-різному, оскільки одні і ті ж змінні системи можуть одночасно розглядатися як результуючі в одному рівнянні, так як і факторні в другому рівнянні. Отже, розрізняють три основних види систем рівнянь [40]:

1) системи незалежних рівнянь (система вирішена відносно ендогенних змінних) – для даної системи кожна залежна змінна Y розглядається як функція одного і того ж набору факторів X;

2) системи рекурсивних рівнянь (рекурсивні системи) – для даної системи кожна залежна змінна Y включається в кожне послідуюче рівняння в якості факторної змінної поряд з сукупністю відповідного набору факторів X;

3) системи взаємозалежних рівнянь (система невирішена відносно ендогенних змінних) – у даній системі одні і ті ж змінні одночасно розглядаються як залежні в одних рівняннях та незалежні в інших.

Аналітично розглянуті системи рівнянь можуть бути представлені таким чином:

1) системи незалежних рівнянь

$$\begin{gathered} Y_1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m + \varepsilon_1, \\ Y_2 = a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m + \varepsilon_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ Y_n = a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m + \varepsilon_n. \end{gathered}$$

Особливості:

- кожна результативна ознака є функцією однієї певної сукупності факторів;
- набір факторів у кожному рівнянні може варіювати залежно від досліджуваної проблеми;
- кожне рівняння системи може розглядатися самостійно;
- так як фактори, включені до моделі, повністю не пояснюють зміну залежних змінних, то в рівняння включається вільний член a_0;
- в кожному рівнянні наявна випадкова помилка.

В якості прикладу такої моделі можна подати модель економічної ефективності певного виробництва, такого вигляду:

$$Y_1 = a_{01} + a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \varepsilon_1$$
$$Y_2 = a_{02} + a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \varepsilon_2\text{'}$$

де Y_1 – рентабельність виробництва;
Y_2 – собівартість виробництва;
x_1 – витрати на інновації;
x_2 – вартість ОВФ;
x_3 – загальні витрати праці.

2) системи рекурсивних рівнянь (рекурсивні системи):

$$Y_1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m + \varepsilon_1,$$
$$Y_2 = b_{21}y_1 + a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m + \varepsilon_2,$$
$$Y_3 = b_{31}y_1 + b_{32}y_2 + a_{31}x_1 + a_{32}x_2 + \cdots + a_{3m}x_m + \varepsilon_3.$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$Y_n = b_{n1}y_1 + b_{n2}y_2 + \ldots. + b_{nn-1}y_{n-1} + a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m + \varepsilon_{n.}$$

Особливості:

- кожне рівняння системи може розглядатися самостійно;
- процес оцінювання здійснюється поетапно;
- використання МНК дозволяє отримати незміщені та обґрунтовані оцінки параметрів;
- не є досить розповсюдженими в економічних дослідженнях.

В якості прикладу такої моделі можна представити модель дослідження продуктивності трудових ресурсів і рівня фондовіддачі, такого вигляду:

$$Y_1 = a_{01} + a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \varepsilon_1,$$
$$Y_2 = a_{02} + b_{21}y_1 + a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \varepsilon_2,$$

де Y_1 –продуктивність праці;

Y_2 – фондовіддача;

x_1 – фондоозброєність праці;

x_2 – енергоозброєність праці;

x_3 – рівень кваліфікації персоналу.

3) системи взаємозалежних рівнянь, взаємопов'язаних (спільних) рівнянь (система невирішена відносно ендогенних змінних):

$$\begin{gathered} Y_1 = b_{12}y_2 + b_{13}y_3 + \cdots + b_{1n}y_n + \\ +a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m + \varepsilon_1 \\ Y_2 = b_{21}y_1 + b_{23}y_3 + \cdots + b_{2n}y_n + \\ +a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m + \varepsilon_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ Y_n = b_{n1}y_1 + b_{n2}y_2 + \cdots + b_{nn-1}y_{n-1} + \\ +a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m + \varepsilon_n \end{gathered}.$$

Особливості:

- одні й ті ж залежні змінні в одних рівняннях входять в ліву частину, а в інших рівняннях – в праву;
- змінні одночасно розглядаються як залежні так і незалежні;
- кожне рівняння системи не може розглядатися самостійно;
- для визначення параметрів необхідні спеціальні методи оцінювання;
- найбільш розповсюджені в економічних дослідженнях.

Приклад. 9.1. Маркетинговою компанією було проведене дослідження ринку молочної продукції в певному регіоні. На основі отриманих даних методом найменших квадратів було побудовано наступну систему рівнянь, яка описує загальний взаємозв'язок попиту та пропозицію на молочну продукцію:

$$g_t = 0{,}74 + 0{,}56\, p_{t-1};$$
$$p_t = 0{,}67 - 0{,}12\, g_t.$$

У поданій моделі g_t – логарифм обсягу пропозиції молочних продуктів у певний період t; p_t – логарифм ціни період t; p_{t-1} – логарифм ціни період $t-1$. Необхідно визначити коефіцієнти еластичності при пояснювальних змінних та зробити висновки.

Перше рівняння характеризує коефіцієнт еластичності пропозиції молочних продуктів від ціни на неї в минулому періоді, тобто при збільшенні ціни на 1 % у періоді $t-1$, обсяг пропозиції період t збільшиться на 0,56 %. Друге рівняння характеризує еластичність ціни продукції від обсягу пропозиції молочних продуктів на ринку, тобто якщо ціна продукції підвищиться на 1 %, то попит на неї впаде на 8,33 % ($k_{\text{ел}} = \frac{1}{-0{,}12} = -8{,}33\ \%$).

Використання систем одночасних рівнянь повинно ґрунтуватися на знанні основних законів економіки на макро-, мікро- та мезорівні. Найбільш розповсюджені моделі дослідження попиту та пропозиції, макроекономічне моделювання механізмів функціонування економіки на прикладі конкретних держав, аналіз функцій витрат та виробничих функцій.

Отже, структурна форма моделі відображує реальний економічний об'єкт або процес та визначає як зміна будь-якої екзогенної змінної визначає значення ендогенних змінних [38].

Структурні рівняння моделі поділяються на поведінкові рівняння та тотожності:

– поведінкові рівняння описують взаємодію між екзогенними та ендогенними змінними (регресійні рівняння);

– тотожності встановлюють співвідношення між ендогенними змінними, не містять випадкових складових та структурних коефіцієнтів моделі.

Для визначення структурних коефіцієнтів моделі необхідно структурну форму моделі перетворити у приведену форму. Розглянемо сутність структурної та приведеної форми моделі, а також особливості чисельної оцінки параметрів та пов'язані з цим проблеми ідентифікації. Основними складовими приведеної та структурної форми є ендогенні та екзогенні змінні.

Структурна форма одночасних рівнянь містить в якості пояснюючих змінних як ендогенні, так і екзогенні змінні та характеризують реальну структуру взаємозв'язку змінних. Загальну структурну форму моделі можна подати таким чином [38]:

$$Y_1 = \mathrm{c}_{10} + b_{12}y_2 + b_{13}y_3 + \cdots + b_{1n}y_n + a_{11}x_1 + \ \ldots + a_{1n}x_m + \ \varepsilon_1,$$
$$Y_2 = \mathrm{c}_{20} + b_{21}y_1 + b_{23}y_3 + \cdots + b_{2n}y_n + a_{21}x_1 + \ \ldots + a_{2n}x_m + \ \varepsilon_2,$$
$$Y_n = \mathrm{c}_{n0} + b_{n1}y_1 + b_{n2}y_2 + \cdots + b_{nn-1}y_{n-1} + a_{nm}x_1 + \ \ldots + a_{nm}x_m + \ \varepsilon_n,$$

де c_{i0}–вільний член рівняння моделі;

b_{ij}–коефіцієнти ендогенних змінних;

a_{ij}–коефіцієнти екзогенних змінних, $i = 1, n$, $j = 1, m$;

ε_i–випадкова складова (похибка) i-го рівняння ($i = 1, n$).

Приведена форма одночасних рівнянь містить в якості пояснюючих змінних тільки екзогенні змінні. Приведена форма необхідна для отримання прогнозних значень ендогенних змінних і для отримання розрахункових значень ендогенних змінних та отримання незміщених оцінок параметрів структурної форми одночасних рівнянь. Загальний вигляд приведеної форми моделі можна подати таким чином [38]:

$$Y_1 = \alpha_{10} + \alpha_{11}x_1 + \ldots + \alpha_{1m}x_m + \delta_1,$$
$$Y_2 = \alpha_{20} + \alpha_{21}x_1 + \ldots + \alpha_{2m}x_m + \delta_2,$$
$$\ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots$$
$$Y_n = \alpha_{n0} + \alpha_{n1}x_1 + \ldots + \alpha_{nm}x_m + \delta_n,$$

де α_{i0} – вільний член рівняння системи;

α_{ij} – коефіцієнти передвизначених змінних, що є функцією коефіцієнтів структурної форми моделі $i = 1, n$, $j = 1, m$;

δ_i – випадкова складова (похибка) i-го рівняння приведеної форми моделі ($i = 1, n$).

Передвизначені змінні системи одночасних рівнянь – це екзогенні та лагові ендогенні змінні цієї системи.

Перехід від приведеної форми моделі до структурної форми пов'язаний з рішенням проблеми ідентифікації. Отже, ідентифікація – це можливість чисельної оцінки параметрів структурних рівнянь за оцінками коефіцієнтів приведених рівнянь, тобто встановлення відповідності між приведеною та структурною формами моделі. Встановлення єдиної відповідності між приведеною та структурною формами моделі становить головну задачу ідентифікації [9; 27].

Розрізняють три класи структурних моделей відповідно до необхідної умови ідентифікації [18; 27; 40]:

- строгоідентифіковані;
- надідентифіковані;
- недоідентифіковані системи.

Рівняння моделі є строгоідентифікованим, якщо кількість ендогенних змінних даного рівняння на одиницю більше кількості передвизначених змінних системи (екзогенних), які не входять в дане рівняння. Виконання умови ідентифікованості моделі перевіряється для кожного рівняння системи окремо, модель є строгоідентифікованою в цілому, коли строго ідентифіковане кожне рівняння системи. Якщо хоча б одне рівняння в системі не є строгоідентифікованим, а надідентифіковане або недоідентифіковане, то відповідно маємо надідентифіковану або недоідентифіковану модель системи рівнянь.

Слід розглянути основні співвідношення для визначення необхідних та достатніх умов ідентифікації. Необхідна умова ідентифікації – "умова порядку" – недостатня, але обов'язкова умова ідентифікації. Основні формули розрахунку необхідної умови ідентифікації, за якими можна визначити тип структурної моделі мають такий вигляд:

1) строгоідентифіковані системи:

$$(n + m) - (n_i + m_i) = n - 1;$$

2) надідентифіковані системи:

$$(n + m) - (n_i + m_i) > n - 1;$$

3) недоідентифіковані системи:

$$(n + m) - (n_i + m_i) < n - 1,$$

де n – загальна кількість ендогенних змінних у системі;

m – загальна кількість екзогенних змінних у системі;

n_i – кількість ендогенних змінних в i-му рівнянні системи;

m_i – кількість екзогенних змінних в i-му рівнянні системи.

Достатня умова ідентифікації – "умова рангу" – полягає в наступному – визначник матриці, складеної з коефіцієнтів при змінних, відсутніх у досліджуваному рівнянні, не дорівнює нулю, а ранг цієї матриці не менше числа ендогенних змінних системи без одиниці. Формулу розрахунку достатньої умови ідентифікації можна подати таким чином:

$$\Delta^* \neq 0 \text{ та } M^* = n - 1,$$

де Δ^* – визначник матриці коефіцієнтів змінних системи (матриці A), відсутніх у даному рівнянні;

M^* – ранг матриці A;

n – загальна кількість ендогенних змінних у системі.

Якщо виконується дана умова, то рівняння моделі є строгоідентифікованим.

На підставі комбінації необхідної та достатньої умови ідентифікації системи рівнянь – умов порядку та рангу можна сформулювати загальні принципи ідентифікації:

1. Якщо $(m - m_i) > n_i - 1$і ранг матриці A дорівнює $M^* = n - 1$, то рівняння надідентифіковане.

2. Якщо $(m - m_i) = n_i - 1$ і ранг матриці A дорівнює $M^* = n - 1$, то рівняння строгоідентифіковане.

3. Якщо $(m - m_i) < n_i - 1$ і ранг матриці A менше $M^* = n - 1$, то рівняння недоідентифіковане.

Для більш змістовної інтерпретації розглянутих моделей за ідентифікацією розглянемо їх характерні особливості [37; 38]:

- строгоідентифікована модель – всі структурні коефіцієнти однозначно визначаються через приведені коефіцієнти, характерно для систем, в яких кількість рівнянь для визначення коефіцієнтів структурних рівнянь в точності дорівнює кількості цих коефіцієнтів;

- недоідентифікована модель – структурні коефіцієнти неможливо визначити через приведені коефіцієнти, в даному випадку, система, що пов'язує коефіцієнти структурних рівнянь з коефіцієнтами приведених рівнянь буде несумісна, характерно для систем, в яких кількість рівнянь менше кількості коефіцієнтів структурних рівнянь;

- надідентифікована модель – структурні коефіцієнти, виражені через приведені коефіцієнти мають два або більше числових значень, отже, за розрахунками, можна отримати декілька варіантів значень структурних рівнянь, характерно для систем, в яких кількість рівнянь для визначення коефіцієнтів структурних рівнянь більше кількості коефіцієнтів, що визначаються.

Слід розглянути завдання визначення ідентифікованості моделі системи рівнянь з використанням необхідної та достатньої умови ідентифікації системи рівнянь.

Приклад 9.2. Нехай ми маємо систему рівнянь взаємозв'язку ендогенних та екзогенних змінних такого вигляду:

$$\begin{aligned}
Y_{1t} &= \beta_{10} + \beta_{12} y_{2t} + \beta_{13} y_{3t} + \gamma_{11} x_{1t} + \varepsilon_{1t}, \\
Y_{2t} &= \beta_{20} + \beta_{23} y_{3t} + \gamma_{21} x_{1t} + \gamma_{22} x_{2t} + \varepsilon_{2t}, \\
Y_{3t} &= \beta_{30} + \beta_{31} y_{1t} + \gamma_{31} x_{1t} + \gamma_{32} x_{2t} + \varepsilon_{3t}, \\
Y_{4t} &= \beta_{40} + \beta_{41} y_{1t} + \beta_{42} y_{2t} + \gamma_{43} x_{3t} + \varepsilon_{4t},
\end{aligned}$$

де y_{it} – ендогенні змінні системи;

x_{it} – екзогенні змінні системи;

β_{i0} – вільний член рівняння моделі;

β_{ij} – коефіцієнти при ендогенних змінних;

γ_{ij} – коефіцієнти при екзогенних змінних, $i = 1, n$, $j = 1, m$;

ε_i – випадкова складова (похибка) i-го рівняння ($i = 1, n$).

Для полегшення розрахунків запишемо систему у вигляді таблиці (табл. 9.1). Використовуючи дану таблицю (табл. 9.1), слід перевірити умову порядку для кожного окремого рівняння.

Таблиця 9.1

Коефіцієнти системи рівнянь

Рівняння системи	1	y_{1t}	y_{2t}	y_{3t}	y_{4t}	x_{1t}	x_{2t}	x_{3t}
1	$-\beta_{10}$	1	$-\beta_{12}$	$-\beta_{13}$	0	$-\gamma_{11}$	0	0
2	$-\beta_{20}$	0	1	$-\beta_{23}$	0	$-\gamma_{21}$	$-\gamma_{22}$	0
3	$-\beta_{30}$	$-\beta_{31}$	0	1	0	$-\gamma_{31}$	$-\gamma_{32}$	0
4	$-\beta_{40}$	$-\beta_{41}$	$-\beta_{42}$	0	1	0	0	$-\gamma_{43}$

Отже, для даної задачі: $n = 4$ – загальна кількість ендогенних змінних у системі, $m = 3$ – загальна кількість екзогенних змінних у системі. Розрахунок за умовами ідентифікації наведено в табл. 9.2.

Таблиця 9.2

Перевірка ідентифікації за умовою порядку

Рівняння системи	Кількість екзогенних змінних $(m - m_i)$	Кількість ендогенних змінних $n_i - 1$	Ідентифікація рівняння
1	$(3-1)\ =\ 2$	$(3-1)\ =\ 2$	строгоідентифіковане
2	$(3-2)\ =\ 1$	$(2-1)\ =\ 1$	строгоідентифіковане
3	$(3-2)\ =\ 1$	$(2-1)\ =\ 1$	строгоідентифіковане
4	$(3-1)\ =\ 2$	$(3-1)\ =\ 2$	строгоідентифіковане

За умовою порядку кожне рівняння є строгоідентифіковане. Далі необхідно провести перевірку за умовою рангу. Наприклад, розглянемо перше рівняння, в яке входять змінні y_{1t}, y_{2t}, y_{3t} та x_{1t}. Для перевірки цього рівняння необхідно побудувати відповідну матрицю коефіцієнтів для змінних y_{4t}, x_{2t}таx_{3t}, які включені в інші рівняння моделі, крім першого. Матриця буде мати вигляд:

$$A = \begin{matrix} 0 & -\gamma_{22} & 0 \\ 0 & -\gamma_{32} & 0 \\ 1 & 0 & -\gamma_{43} \end{matrix}.$$

Визначник $\mathrm{A} = 0$, оскільки визначник матриці дорівнює нулю, то ранг матриці $\mathrm{A} < 3$. Тому рівняння (1) не задовольняє умову рангу, тобто не є строгоідентифікованим.

На основі розглянутих особливостей та характеристик моделей структурних рівнянь побудовано агреговану схему класифікації даних моделей, яка наведена на рис. 9.2.

Докладна класифікація має важливе значення, тому що різні типи структурних економетричних моделей, як правило, вимагають різних методів оцінки параметрів, перевірки гіпотез, а також методів прогнозних розрахунків. Для класифікації різних типів структурних економетричних моделей використовуються такі основні три критерії [18; 19]:

1) форма матриці структурних коефіцієнтів загальних залежних змінних;

2) форма дисперсійно-коваріаційної матриці вектора помилок $\varepsilon(t)$;

3) наявність або відсутність автокореляції помилок у рівняннях.

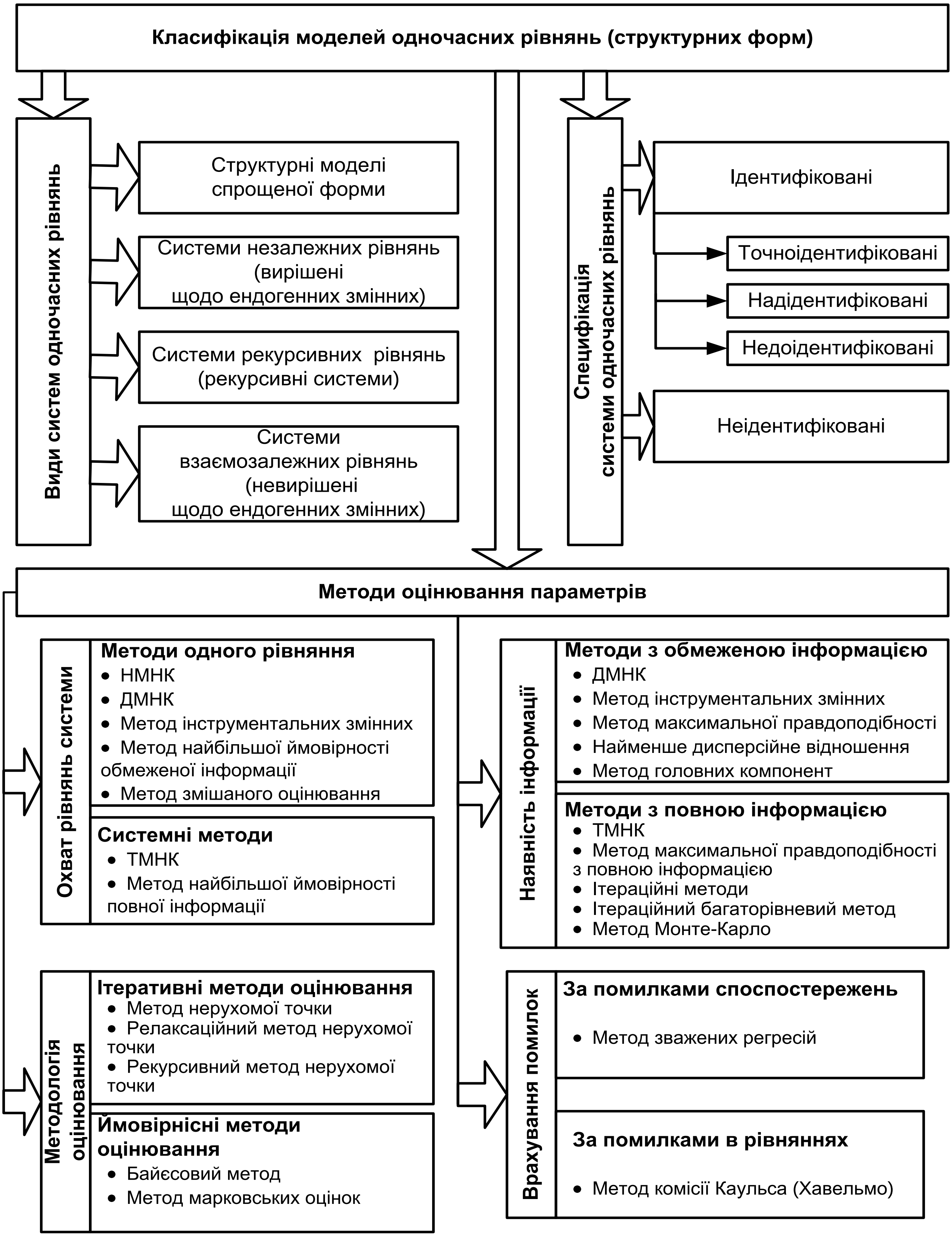


Рис. 9.2. **Основні критерії класифікації моделей структурних форм**

Класифікації різних типів структурних економетричних моделей за всіма критеріями наведено на рис. 9.3.

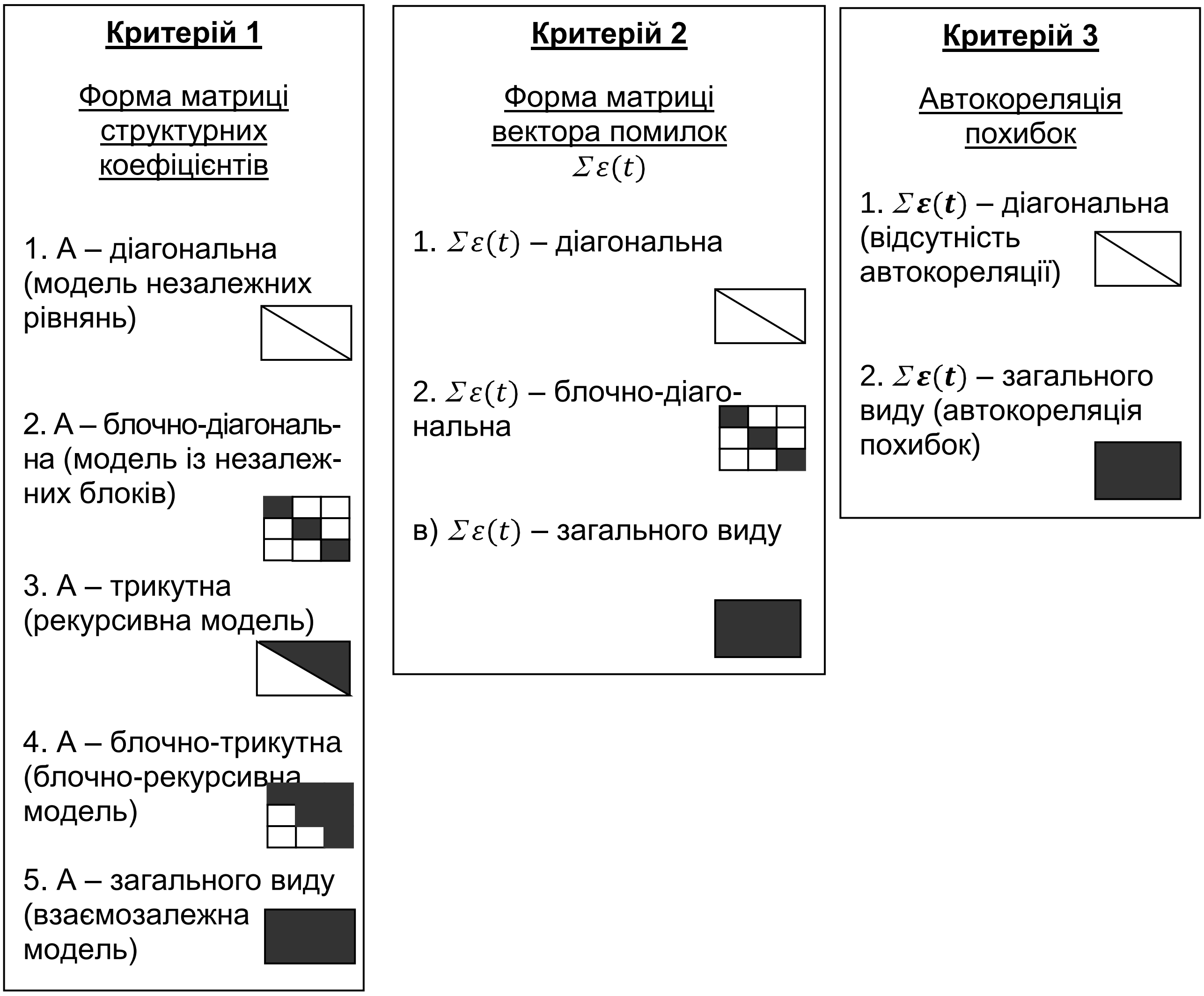


Рис. 9.3. **Критерії класифікації різних типів структурних економетричних моделей**

9.3. Методи оцінювання параметрів структурних рівнянь

Для отримання якісних оцінок параметрів системи одночасних рівнянь необхідно правильно вибрати метод оцінювання. Вибір метода визначається умовами системи, обмеженнями та агрегацією певних критеріїв. Досить суттєвим фактором під час вибору алгоритму оцінки параметрів є простота його реалізації. Найбільш класичними та розповсюдженими методами оцінки параметрів структурних форм є непрямий метод найменших квадратів (НМНК) та двокроковий метод найменших квадратів (ДМНК) [19; 27; 36]. Розглянемо їх сутність більш докладно.

Непрямий МНК полягає у такому:

- застосовується тільки в випадку строгоідентифікованої структурної моделі;

• складають наведену форму моделі і визначають чисельні значення параметрів кожного її рівняння звичайним МНК;

• шляхом алгебраїчних перетворень переходять від наведеної форми до рівнянь структурної форми моделі, отримуючи тим самим чисельні оцінки структурних параметрів;

• отримані оцінки параметрів структурної форми моделі є незміщеними та обґрунтованими;

• є досить чутливим до порушення умов незалежності факторів між собою, в системі рівнянь це є причиною необґрунтованості оцінок структурних коефіцієнтів та неможливості коректно представити економічну інтерпретацію результатів.

Загальний алгоритм застосування НМНК містить такі кроки [27]:

Крок 1. Перевіряється умова ідентифікованості для кожного рівняння. Якщо кожне рівняння строгоідентифіковане, то виконується перехід до кроку 2.

Крок 2. Структурна форма моделі перетворюється до приведеної форми.

Система одночасових структурних рівнянь в матричному вигляді запишеться так:

$$Y = A\,Y + B\,X + \varepsilon.$$

Якщо кожне рівняння системи розв'язати відносно залежної змінної Y, то одержимо приведену форму моделі, яка має такий вигляд:

$$Y = R\,X\ +\ \delta,$$

де залишки δ є лінійною комбінацією залишків ε.

Зв'язок між коефіцієнтами структурної та приведеної форми моделі можна записати таким чином:

$$R = (E - A)^{-1} \cdot B = -A^{-1} \cdot B,\ AR\ +\ B\ =\ 0.$$

Крок 3. Оцінка параметрів кожного рівняння приведеної форми моделі за класичним МНК.

Оцінка параметрів моделі на основі системи одночасових рівнянь за МНК даватиме зміщення, яке дорівнюватиме:

$$\frac{1 - a_1 \ \sigma^2 / m_{ss}}{1 + \sigma^2 / m_{ss}},$$

де m_{ss} – момент другого порядку залежної змінної, який прямує до деякої константи.

Крок 4. Перетворення оцінок параметрів приведеної форми до структурних коефіцієнтів.

Розрахунок оцінок параметрів рівнянь структурної форми отримаємо на основі співвідношення:

$$A\,R = -\ B,$$

де A і B – параметри структурних рівнянь;

R – матриця оцінок параметрів приведеної форми моделі.

Двокроковий МНК полягає у такому:

- застосовується в випадку надідентифікованої структурної моделі двох типів: якщо всі рівняння системи надідентифіковані та якщо система містить поряд з надідентифікованими строгоідентифіковані рівняння;
- для надідентифікованих рівнянь ДМНК на відміну від НМНК визначає єдині оцінки параметрів моделі;
- у випадку, коли рівняння моделі строгоідентифіковані, то НМНК і ДМНК дають однакову оцінку параметрів;
- основна ідея методу – на основі приведеної форми моделі отримати для надідентифікованого рівняння теоретичні значення ендогенних змінних, що знаходяться в правій частині рівняння;
- метод є двокроковим, оскільки два рази застосовується МНК: перший раз у процесі визначення приведеної форми моделі і знаходження на її основі оцінок теоретичних значень ендогенних змінних, і на другому кроці стосовно до структурного надідентифікованого рівняння у ході визначення структурних коефіцієнтів моделі за даними теоретичних (розрахункових) значень ендогенних змінних;
- перший етап методу (етап побудови приведених рівнянь) застосовується тільки для конкретних рівнянь, що дозволяє мінімізувати обсяг розрахунків;

- застосування ДМНК буде ефективним тільки в тому випадку, коли коефіцієнт детермінації приведених рівнянь, отриманих на першому кроці буде досить значним. При низькому значенні коефіцієнта детермінації використання ДМНК не є ефективним, так як у цьому випадку інструментальна змінна практично не відповідає істинному значенню змінної, яку замінює;
- використання методу інструментальних змінних як складової частини ДМНК дозволяє отримати обґрунтовані оцінки й оцінки стандартних відхилень для вибірок зі значними обсягами спостережень та факторів.

Загальний алгоритм застосування ДМНК містить такі кроки [27]:

Крок 1. Перевіряється умова ідентифікованості для кожного рівняння системи. Якщо кожне рівняння не є строгоідентифікованим, то виконується перехід до кроку 2.

Крок 2. Структурна форма моделі перетворюється до приведеної форми.

Крок 3. Оцінка параметрів кожного рівняння приведеної форми моделі за класичним МНК.

Крок 4. У правій частині надідентифікованого рівняння структурної моделі вибираються ендогенні змінні і розраховуються їх теоретичні значення за відповідними наведеними рівняннями.

Крок 5. За допомогою МНК на основі фактичних значень екзогенних і теоретичних значень ендогенних змінних оцінюються параметри надідентифікованого рівняння структурної форми.

Отже, систему рівнянь для обчислення оцінок двокроковим методом найменших квадратів аналітично можна представити таким чином:

$$\begin{matrix} Y_1' X(X'X)^{-1}X'Y_1 & Y_1' X_1 \\ X_1' Y_1 & X_1' X_1 \end{matrix} \quad \begin{matrix} A \\ B \end{matrix} = \begin{matrix} Y_1' X(X'X)^{-1}X'Y \\ X_1' Y \end{matrix},$$

де Y – вектор залежної або ендогенної змінної;

Y_1 – матриця поточних ендогенних змінних, які входять у праву частину рівняння;

X – матриця всіх пояснювальних або екзогенних змінних;

X_1 – матриця пояснювальних або екзогенних змінних даного рівняння;

A – вектор оцінок структурних параметрів, які стосуються змінних матриці Y_1;

B – вектор оцінок структурних параметрів, які стосуються змінних матриці X_1.

Оператор оцінювання ДМНК запишеться так:

$$\begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} Y_1' X(X'X)^{-1}X'Y_1 & Y_1' X_1 \\ X_1' Y_1 & X_1' X_1 \end{pmatrix}^{-1} \begin{pmatrix} Y_1' X(X'X)^{-1}X'Y \\ X_1' Y \end{pmatrix}.$$

Слід розглянути реалізацію методу ДМНК на такому прикладі.

Приклад 9.3. Нехай є такі дані, числові значення яких наведено в табл. 9.3 [24].

Таблиця 9.3

Вихідні дані

№ п/п	X_1	X_2	X_3	Y_1	Y_2
1	564,2	349,0	104,3	70,2	50,3
2	568,7	349,2	94,2	70,4	54,8
3	572,6	349,9	102,8	72,3	55,6
4	576,3	353,5	98,7	73,9	60,4
5	580,2	355,5	99,8	75,3	62,2
6	584,4	355,8	100,5	76,9	64,6
7	588,1	357,2	112,8	77,3	64,7
8	591,3	361,8	106,7	78,1	66,7
9	595,2	366,2	100,3	79,4	70,5
10	598,4	368,9	105,0	81,1	73,6
11	603,0	370,9	105,6	82,7	74,3
12	606,2	371,0	106,2	83,3	76,5

Подані змінні мають такий економічний зміст: Y_1 – загальний обсяг експорту продукції; Y_2– загальний обсяг імпорту продукції; X_1 – середній товарообіг торгівлі країн, з якими підтримуються зовнішньоекономічні відносини; X_2 – національний дохід країни; X_3 – індекс споживчих цін.

Слід скласти систему одночасних рівнянь:

$$Y_1 = b_{11}y_2 + a_{10} + a_{11}x_1 + \varepsilon_1,$$
$$Y_2 = b_{21}y_1 + a_{20} + a_{22}x_2 + a_{23}x_3 + \varepsilon_2.$$

Для вибору методу оцінювання невідомих параметрів системи одночасних рівнянь треба застосувати умову ідентифікації. Для даної системи одночасних рівнянь буде одержано:

$$n = 2, \quad m = 3, \quad n_1 = 2, \quad n_2 = 2, \quad m_1 = 1, \quad m_2 = 2.$$

Для рівнянь системи за загальною формулою ідентифікації маємо:

$$(n + m) - (n_i + m_i) = n - 1;$$
$$2 + 3 - 2 + 1 > 2 - 1;$$
$$2 + 3 - 2 - 2 = 2 - 1.$$

Умова строгої ідентифікації для першого рівняння не виконана, тобто неможливо оцінити параметри структурної форми, використовуючи НМНК. Для оцінювання параметрів структурної форми надідентифікованої системи використовуємо двокроковий метод найменших квадратів (ДМНК). Отже, наведена форма моделі буде мати такий вигляд:

$$Y_1 = c_{10} + c_{11}x_1 + c_{12}x_2 + c_{13}x_3 + \delta_1,$$
$$Y_2 = c_{20} + c_{21}x_1 + c_{22}x_2 + c_{23}x_3 + \delta_2.$$

Оцінки параметрів для прийнятої форми одержано, використовуючи метод МНК для кожного рівняння регресії:

$$Y_1 = -114{,}436 + 0{,}297\,x_1 + 0{,}043\,x_2 + 0{,}017\,x_3 + \delta_1,$$
$$Y_2 = -287{,}819 + 0{,}656\,x_1 - 0{,}034\,x_2 - 0{,}188\,x_3 + \delta_2.$$

Далі слід знайти теоретичні значення залежної змінної за першим і другим рівняннями регресії (табл. 9.4).

Таблиця 9.4

Теоретичні значення змінних моделі

№ п/п	Y_1	Y_2
1	69,940	50,523
2	71,109	55,368
3	72,447	56,283
4	73,629	59,357
5	74,893	61,638
6	76,165	64,251
7	77,538	64,315
8	78,580	67,402
9	79,816	71,012
10	80,964	72,133
11	82,427	74,968
12	83,392	76,950

Далі слід виконати другий етап двокрокового методу найменших квадратів (ДМНК). Підставити в структурну форму отримані прогнозні значення Y_1 та Y_2. Оцінювання параметрів моделі слід провести для кожного рівняння структурної форми. З урахуванням отриманих оцінок структурна форма економетричної моделі буде мати такий вигляд:

$$Y_1 = -\,0{,}073\, y_2 - 135{,}123 + 0{,}37\, x_1 + \varepsilon_1,$$
$$Y_2 = 2{,}208\, y_1 - 35{,}122 - 0{,}129\, x_2 - 0{,}227\, x_3 + \varepsilon_2,$$

де ε_1, ε_2 – випадкові помилки структурної економетричної моделі.

Таким чином, за допомогою ДМНК отримані оцінки параметрів структурної форми економетричної моделі, модель є адекватною та статистично значущою та підтверджує взаємозв'язок між досліджуваними показниками: експортом, імпортом, національним доходом, товарообігом зовнішньої торгівлі та індексом споживчих цін.

Розглянуті методи оцінювання параметрів – непрямий і двокроковий методи найменших квадратів застосовуються для оцінки параметрів кожного окремого рівняння моделі. ***Трикроковий метод найменших квадратів***, запропонований Зельнером і Гейлом, призначений для одночасної оцінки параметрів всіх рівнянь моделі, його використання за певних обставин є більш ефективним, ніж двокроковий метод [27].

Трикроковий метод найменших квадратів забезпечує кращу порівняно з двокроковим методом асимптотичну ефективність оцінок лише в тому разі, коли матриця залишків $\Sigma = M(\varepsilon\varepsilon')$ не є діагональною, тобто коли залишки, які входять в різні рівняння моделі, корелюють між собою.

Щоб застосувати трикроковий метод найменших квадратів на практиці необхідне виконання таких вимог:

1) усі тотожності, які входять в систему рівнянь, треба виключити для знаходження оцінок параметрів;

2) кожне неідентифіковане рівняння також треба виключити з системи;

3) якщо система рівнянь, що залишилась, має строго ідентифіковані і надідентифіковані рівняння, то трикроковий метод оцінки доцільно застосовувати до кожної з цих груп;

4) для групи надідентифікованих рівнянь оцінки параметрів знаходяться на основі співвідношення:

$$\delta = \frac{A}{B} = \begin{array}{ccc} S_{11}^{2}\ {}^{-1}Z_1'\ X'X\ {}^{-1}X'Z_1 & \dots & S_{1r}^{2}\ {}^{-1}Z_1'\ X'X\ {}^{-1}X'Z_r \\ \dots & \dots & \dots \\ S_{r1}^{2}\ {}^{-1}Z_r'\ X'X\ {}^{-1}X'Z_1 & \dots & S_{rr}^{2}\ {}^{-1}Z_r'\ X'X\ {}^{-1}X'Z_r \end{array}{}^{-1} \times$$

$$\times \begin{array}{c} {}_{j=1}^{r}\ S_{1j}^{2}\ {}^{-1}Z_1'X\ X'X\ {}^{-1}X'Y_j \\ \dots \\ {}_{j=1}^{r}\ S_{rj}^{2}\ {}^{-1}Z_r'X\ X'X\ {}^{-1}X'Y_j \end{array},$$

де $\delta = \frac{A}{B}$ – оцінки параметрів моделі;

Z_s – змінні моделі, які знаходяться в правій частині s-го рівняння, $Z_s = \ Y_s X_s\ , (s = 1, r)$;

S_{sj}^{2} – дисперсії залишків для кожного рівняння, які є наближеною оцінкою σ_{sj}^{2};

5) якщо група надідентифікованих рівнянь має тільки одне рівняння, то трикроковий метод перетворюється на двокроковий;

6) якщо матриця коваріацій $\Sigma = M(uu')$ для структурних залишків блочно-діагональна, то вся процедура оцінювання на основі трикрокового методу найменших квадратів може бути застосована окремо до кожної групи рівнянь, які відповідають одному блоку.

Практичне застосування ТМНК обмежено через велику кількість розрахунків та досить значущою чутливістю результатів оцінювання до помилок специфікації.

Серед методів з обмеженою інформацією, що застосовуються для оцінювання параметрів у взаємозалежних системах, виділяють ***метод максимальної правдоподібності*** [36; 37]. Цей метод виходить з передумов, відмінних від тих, які використовують розглянуті вищі методи, однак тісно пов'язаний з ними. У методах, заснованих на принципі максимальної правдоподібності, будується функція ймовірності для залежної змінної, що залежить від значень невідомих параметрів, і шукаються ті значення параметрів, при яких вона досягає певного максимуму. Якщо всі умови, необхідні для отримання незміщених та обґрунтованих оцінок за методом найменших квадратів, виконуються, а помилки розподілені за нормальним законом, то можна показати, що максимум функції правдоподібності збігається з мінімумом суми квадратів відхилень. Застосування методу максимальної правдоподібності в загальному випадку є досить трудомістким.

У разі строгоідентіфікованих систем результати методу максимальної правдоподібності з обмеженою інформацією збігаються не тільки з оцінками за ДМНК, але і з оцінками НМНК. Якщо система надідентифікована, то оцінки методу максимальної правдоподібності асимптотично наближені до оцінок за ДМНК. Однак для малих вибірок метод максимальної правдоподібності зазвичай дає кращі результати. До недоліків методу максимальної правдоподібності слід віднести велику кількість розрахунків, чутливість результатів до мультиколінеарності і специфікації моделі.

Метод максимальної правдоподібності з повною інформацією – теоретично найбільш досконалий метод оцінювання параметрів взаємозалежних моделей. На відміну від методу максимальної правдоподібності з обмеженою інформацією він враховує специфікацію всіх рівнянь моделі і зв'язки між помилками [36; 37]. Аналогічно ТМНК параметри всіх рівнянь оцінюються одночасно.

У методі максимальної правдоподібності з повною інформацією, так само як і в ТМНК використовується оцінка матриці, оберненої до коваріаційної матриці помилок. Обчислені таким чином оцінки параметрів при зростанні числа спостережень асимптотично наближаються до ітераційних оцінок за методом максимальної правдоподібності з повною інформацією. Аналогічний результат оцінювання маємо також і для ТМНК. Основна перевага даного методу полягає в тому, що його легко пристосувати до нелінійних моделей.

Методи з повною інформацією дають більш ефективні оцінки параметрів, ніж методи з обмеженою інформацією. Це особливо важливо для надідентифікованих моделей. Однак, з іншого боку, методи з повною інформацією вимагають великої кількості розрахунків, результати чутливі до мультиколінеарності і помилок специфікації. Саме тому застосування їх на практиці без використання сучасних програмних засобів є досить трудомістким.

Контрольні запитання для самодіагностики

1. У чому особливості систем одночасних рівнянь?
2. Які види систем одночасних рівнянь ви знаєте?
3. Що таке структурна форма моделі?
4. Що таке приведена форма моделі?
5. У чому полягає проблема ідентифікації структурних моделей?
6. Які критерії можуть бути використані під час ідентифікації структурних моделей?

7. Які необхідні та достатні умови ідентифікації?

8. Яка система є строго ідентифікованою?

9. Які методи можуть бути використані під час оцінювання параметрів систем одночасних рівнянь?

10. Наведіть приклади комплексних економетричних моделей.

11. Наведіть алгоритм непрямого методу найменших квадратів (НМНК), особливості його застосування.

12. Коли застосовується двокроковий метод найменших квадратів (ДМНК)?

13. Чи можна виконувати оцінювання параметрів моделі окремо для групи строгоідентифікованих та надідентифікованих рівнянь?

14. У чому полягають недоліки використання наведеної форми моделі?

15. У якому випадку оцінки параметрів за НМНК співпадають з ДМНК?

Тести

1. Системами економетричних рівнянь є:

а) системи одночасних рівнянь;

б) системи рекурсивних рівнянь;

в) системи незалежних рівнянь;

г) усі пераховані варіанти.

2. Система одночасних рівнянь відрізняється від інших видів економетричних систем тим, що в ній:

а) ендогенна змінна одного рівняння знаходиться в іншому рівнянні системи в якості фактора;

б) одні й ті ж ендогенні змінні системи в одних рівняннях знаходяться в лівій частині, а в інших рівняннях – у правій частині;

в) кожна ендогенна змінна є функцією однієї і тієї ж сукупності екзогенних змінних.

3. МНК дозволяє отримати обґрунтовані і незміщені оцінки параметрів системи:

а) рекурсивних рівнянь;

б) одночасних рівнянь;

в) незалежних рівнянь.

4. Ендогенна змінна – це:

а) залежна змінна, яка визначається всередині системи;

б) факторна змінна;

в) зумовлена змінна;

г) структурна змінна.

5. Екзогенні змінні моделі характеризуються тим, що:

а) визначаються попередніми моментами часу;

б) є незалежними і визначаються поза системою;

в) є ідентифікованими;

г) є залежними і визначаються усередині системи.

6. Використовуючи необхідний критерій ідентифікації, можна зазначити, що система недоідентифікована, якщо:

а) $(n + m) - (n_i + m_i) > n - 1$;

б) $(n + m) - (n_i + m_i) = n - 1$;

в) $(n + m) - (n_i + m_i) < n - 1$.

7. Використовуючи необхідний критерій ідентифікації можна зазначити, що система надідентифікована, якщо:

а) $(n + m) - (n_i + m_i) > n - 1$;

б) $(n + m) - (n_i + m_i) = n - 1$;

в) $(n + m) - (n_i + m_i) < n - 1$.

8. Для оцінювання параметрів строгоідентифікованої системи використовують:

а) узагальнений метод найменших квадратів (метод Ейткена);

б) непрямий метод найменших квадратів (НМНК);

в) двокроковий метод найменших квадратів (ДМНК).

9. Для оцінювання параметрів надідентифікованої системи використовують:

а) узагальнений метод найменших квадратів (метод Ейткена);

б) непрямий метод найменших квадратів (КМНК);

в) двокроковий метод найменших квадратів (ДМНК).

10. Рекурсивна форма економетричної моделі – це система регресійних рівнянь, в яких:

а) одні й ті ж змінні в одних рівняннях системи входять у ліву частину, а в інших – у праву;

б) залежна змінна попереднього рівняння виступає у вигляді незалежної змінної наступного рівняння;

в) залежні змінні одних рівнянь не виступають в якості незалежних змінних інших рівнянь, тобто система вже вирішена щодо ендогенних змінних.

11. Ідентифікація – це:

а) єдиність відповідності між наведеною та структурною формами моделі;

б) метод оцінювання параметрів системи рівнянь;

в) оцінювання адекватності моделі.

12. Економетрична модель подана в такому вигляді:

$$y_1 = a_1x_1 + a_2y_3 + e_1$$
$$y_2 = a_3x_4 + a_4y_3 + e_2$$
$$y_3 = a_5x_2 + a_6y_1 + e_3$$
$$y_4 = a_7x_3 + a_8y_2 + e_4$$

Наведена система є:

а) системою, яка не вирішена щодо ендогенних змінних;

б) рекурсивною системою;

в) системою, вирішеною щодо ендогенних змінних.

13. Економетрична модель подана в такому вигляді:

$$y_1 = a_1x_1 + a_2x_3 + e_1$$
$$y_2 = a_3x_4 + a_4y_1 + e_2$$
$$y_3 = a_5x_2 + a_6y_2 + e_3$$
$$y_4 = a_7x_3 + a_8y_3 + e_4$$

Наведена система є:

а) системою, яка не вирішена щодо ендогенних змінних;

б) рекурсивною системою;

в) системою, вирішеною щодо ендогенних змінних.

14. Економетрична модель подана в такому вигляді:

$$y_1 = a_1x_1 + a_2x_3 + e_1$$
$$y_2 = a_3x_4 + a_4y_1 + e_2$$
$$y_3 = a_5x_2 + a_6y_2 + e_3$$
$$y_4 = a_7x_3 + a_8y_3 + e_4$$

Наведена система є:

а) строгоідентифікованою;

б) недоідентифікованою;

в) надідентифікованою.

15. Структурна форма економетричної моделі – це система регресійних рівнянь, в яких:

а) одні й ті ж змінні в одних рівняннях системи входять у ліву частину, а в інших – у праву;

б) залежна змінна попереднього рівняння виступає у вигляді незалежної змінної наступного рівняння;

в) залежні змінні одних рівнянь не виступають в якості незалежних змінних інших рівнянь, тобто система вже передвизначена щодо ендогенних змінних.

Практичні завдання

1. Необхідно розглянути модель, яка складається з рівняння пропозиції та рівняння попиту на продукцію сільського господарства. Рівняння специфікуються на основі таких функцій:

$$g_t = a_{10} + a_{11} p_{t-1} + \varepsilon_{1t};$$
$$p_t = a_{20} + a_{21} g_t + \varepsilon_{2t}.$$

Перше рівняння моделі – це рівняння пропозиції: її кількість на ринку в період t залежить від ціни в періоді $t-1$, відповідно (p_{t-1}).

Випадкова змінна ε_{1t} характеризує залишки, на величину яких можуть впливати ті чинники, якими знехтували: витрати, технологічні зміни, тощо.

Друге рівняння моделі – це рівняння попиту. Ціна p_t на ринку в період t залежить від пропозиції (g_t) в цьому ж періоді. Змінна ε_{2t} характеризує залишки в цьому рівнянні.

Необхідно визначити ідентифікованість моделі та запропонувати методи оцінки параметрів.

2. Маркетинговою компанією було проведене дослідження ринку цукру. На основі отриманих даних методом найменших квадратів було побудовано наступну систему рівнянь, яка описує взаємозв'язок попиту та пропозицію на досліджуваний товар:

$$g_t = 0{,}55 + 0{,}53_{11} p_{t-1};$$
$$p_t = 1{,}73 - 0{,}09 g_t.$$

У поданій моделі g_t – логарифм пропозиції цукру в певний період t, p_t – логарифм ціни період t, p_{t-1} – логарифм ціни період $t-1$.

Необхідно визначити коефіцієнти еластичності при пояснювальних змінних та зробити висновки.

3. Для представленої системи рівнянь,що наведена такою залежністю:

$$\begin{gathered} y_{1,t} = c_{10} + b_{14} y_{4,t} + b_{12} y_{1,t-1} + \varepsilon_1 \\ y_{2,t} = c_{20} + b_{23} y_{3,t} + b_{22} y_{2,t-1} + \varepsilon_2 \\ y_{3,t} = c_{30} + b_{34} y_{4,t} + a_{31} x_{1,t} + \varepsilon_3 \\ y_{4,t} = y_{1,t} + y_{2,t} + x_{2,t} \end{gathered},$$

де $y_{1,t}$ – витрати на споживання в період t;

$y_{1,t-1}$ – витрати на споживання в період $t-1$;

$y_{2,t}$ – обсяги інвестицій в період t;

$y_{2,t-1}$ – обсяги інвестицій в період $t-1$;

$y_{3,t}$ – відсоткова ставка в період t;

$y_{4,t}$ – загальний дохід в період t;

$x_{1,t}$ – грошова маса в період t;

$x_{2,t}$ – загальні витрати держави в період t.

Необхідно дослідити взаємозв'язок змінних у системі одночасних рівнянь та виділити ендогенні, екзогенні змінні та лагові змінні. Визначити ідентифікованість моделі та запропонувати методи оцінювання параметрів.

4. Для представленої системи рівнянь у приведеній формі, що наведена такою залежністю:

$$Y_1 = a_{10} + a_{11}x_1 + a_{12}x_2 + \varepsilon_1;$$
$$Y_2 = a_{20} + a_{21}x_1 + a_{22}x_2 + \varepsilon_2;$$
$$Y_1 = a_{30} + a_{31}x_1 + a_{32}x_2 + \varepsilon_3.$$

Необхідно визначити відповідну їй структурну форму.

5. Для побудованої системи одночасних рівнянь, що містить регресійні рівняння продуктивності прибутку та вартості основних виробничих фондів, виду:

$$y_1 = 0{,}14y_2 + 1{,}23x_1 + \varepsilon_1;$$
$$y_2 = 0{,}55y_1 + 1{,}34x_2 + \varepsilon_2.$$

Необхідно привести структурну модель до наведеної форми, проаналізувати отримані результати.

6. Використовуючи результати спостережень, наведені в табл. 9.5, необхідно оцінити структурні коефіцієнти моделі, що наведена такою залежністю:

$$Y_1 = b_{11}y_2 + a_{10} + a_{11}x_1 + a_{12}x_2 + a_{13}x_4 + \varepsilon_1$$
$$Y_2 = b_{21}y_1 + a_{20} + a_{22}x_2 + a_{23}x_3 + a_{24}x_4 + \varepsilon_2,$$

де y_1– продуктивність праці;

y_2 – заробітна плата;

x_1 – фондомісткість продукції;

x_2 – плинність робочої чинності;

x_3 – рівень витрат робочого часу;

x_4 – стаж працюючих.

Необхідно оцінити параметри структурної моделі та представити змістовну інтерпретацію результатів моделювання.

Таблиця 9.5

Вихідні дані

y_1	y_2	x_1	x_2	x_3	x_4
116,18	2294,55	138,36	3,27	2,71	13,72
118,55	2378,18	142,36	3,09	2,82	15,09
111,82	2190,91	138,36	2,91	3,02	13,72
114,00	2294,55	140,36	2,73	3,22	16,46
120,73	2556,36	134,55	2,55	3,22	19,20
122,91	2609,09	128,73	2,18	3,33	21,95
127,45	2712,73	128,73	2,00	3,42	27,44
116,18	2316,36	119,27	2,00	3,62	27,44
134,18	2816,36	138,36	2,00	3,73	28,82
134,18	2765,45	138,36	1,27	3,73	30,18
138,55	2869,09	142,36	1,27	3,42	35,67
143,09	2921,82	144,18	1,09	4,02	27,44
145,27	2921,82	146,18	1,09	4,22	32,93
149,82	3025,45	153,82	0,91	4,33	35,67
149,82	2972,73	157,64	1,09	4,73	38,42
138,55	2847,27	161,45	1,09	4,84	39,78
140,91	2900,00	161,45	1,09	4,84	42,53

Ключові слова

Система одночасних рівнянь. Система рекурсивних рівнянь. Система незалежних рівнянь. Структурна форма моделі. Наведена форма моделі. Ідентифікована модель. Надідентифікована модель. Неідентифікована модель. Двокроковий метод найменших квадратів. Непрямий метод найменших квадратів.

Розділ 10. Лабораторний практикум

Лабораторні роботи призначені для закріплення теоретичного і практичного матеріалу, набуття навичок роботи з пакетами прикладних програм (ППП), що забезпечують побудову й дослідження різних типів моделей, а також для розширення знань студентів у галузі застосування ПЕОМ для економічних розрахунків, прогнозування й аналізу діяльності економічних систем.

Для виконання лабораторних робіт пропонується використати ППП Statistica 6.0. Даний пакет містить множину статистичних методів, що підтримують рішення різних економетричних задач. ППП Statistica призначений для роботи в середовищі *Windows*. Під час розробки лабораторних робіт передбачалося, що студент ознайомлений з основними принципами та прийомами роботи в середовищі *Windows*.

Кожна лабораторна робота розглянута на прикладі рішення конкретної задачі з докладними коментарями й рисунками. Лабораторні роботи рекомендується виконувати послідовно, оскільки дії та прийоми, загальні для всіх робіт, будуть указуватися один раз. Крім того, послідовне виконання дозволяє краще засвоїти й закріпити матеріал навчальної дисципліни.

Лабораторні роботи стосуються основних тем дисципліни і ґрунтуються на теоретичному матеріалі відповідної теми, а також попередніх тем. Кожна робота містить мету й завдання для виконання, рекомендації до виконання.

Лабораторна робота 1. Варіаційні ряди та їх статистичні характеристики

Мета – закріплення теоретичного й практичного матеріалу, набуття навичок роботи в модулі Basic Statistics/Tables.

Завдання – провести аналіз варіаційного ряду для вибіркових даних у модулі *Basic Statistics*/*Tables* ППП *Statistica:*

1. Розрахувати статистичні характеристики ряду (середнє, дисперсію, середнє квадратичне відхилення, моду, медіану, розмах варіації, коефіцієнти асиметрії та ексцесу).

2. Побудувати гістограму та полігон розподілу випадкової величини, зробити висновки щодо характеру закону розподілу.

3. За допомогою критеріїв Пірсона та Колмогорова – Смірнова перевірити гіпотезу про нормальний закон розподілу.

4. Зробити висновки щодо угрупування об'єктів за величиною відповідного показника.

Для рішення й аналізу задач даного типу в ППП *Statistica* передбачений модуль *Basic Statistics*/*Tables* (*Основні статистики й таблиці*). Слід розглянути порядок роботи в даному модулі.

1. Запуск *Statistica* і підготовка даних.

У меню програм вибрати програму *Statistica*, після її запуску виберіть у меню пункт *File/New для* підготовки власних даних. З'явиться діалогове вікно, у якому необхідно вказати кількість змінних *(number of variables)* і кількість випадків *(number of cases).* Після уведення слід натиснути кнопку вікна *OK* (рис. 10.1.1).

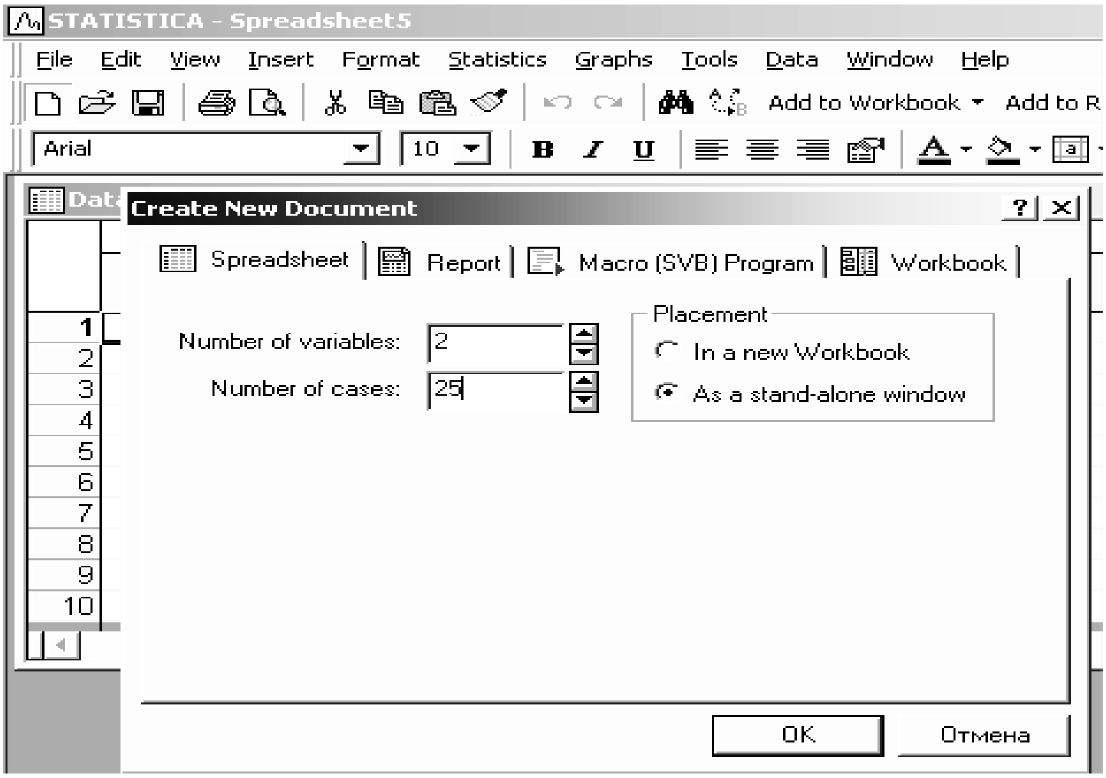


Рис. 10.1.1. **Визначення кількості змінних і спостережень**

З'явиться порожнє поле, що становить таблицю розміром 25 х 2: 25 спостережень, 2 змінні (рис. 10.1.2).

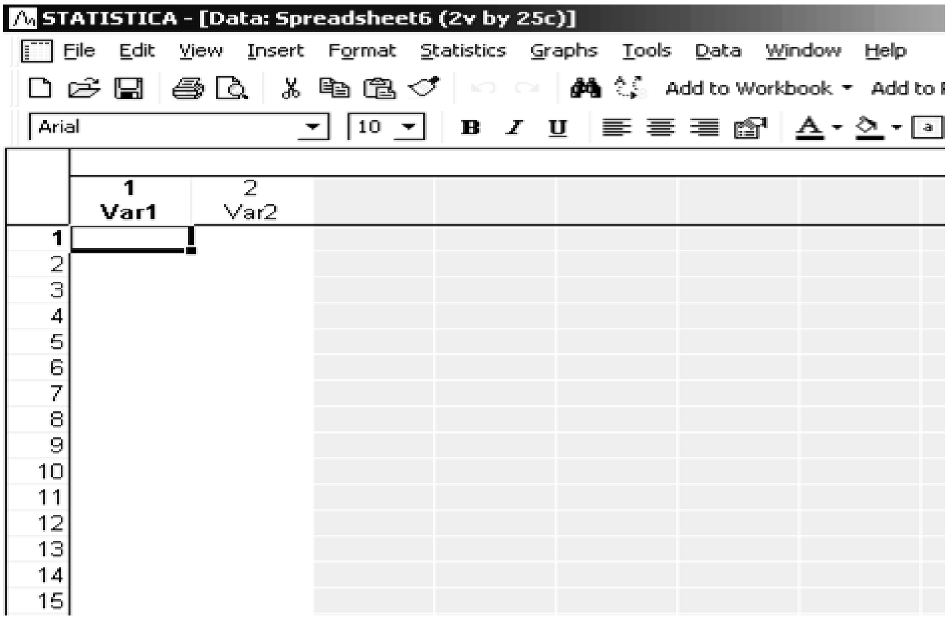


Рис. 10.1.2. **Поле даних**

Кожен елемент даних, тобто значення показника, займає одну комірку поля даних. Після заповнення всіх комірок поля даних буде одержано таблицю, наведену на рис. 10.1.3.

	1 Bank	2 % dohod
1	Приватбанк	1187477
2	Проминвест	793821
3	Аваль	876148
4	Ощадбанк	389719
5	Укрсоцбанк	459234
6	Укрсиббанк	451074
7	Укрэксим	328131
8	Райффайзенбанк	209010
9	Надра	273945
10	Брокбизнесбанк	167741
11	Укрпромбанк	232158
12	Финансы и кредит	175292
13	Первый укр. международный банк	111185
14	Хрещатик	70674
15	Форум	145468

Рис. 10.1.3. **Вихідні дані**

2. Розрахунки.

Розрахувати основні статистичні характеристики ряду (середнє, дисперсію, середнє квадратичне відхилення, моду, медіану, розмах, коефіцієнти асиметрії й ексцесу).

Щоб почати обчислювальні процедури, необхідно ввійти в позицію меню *Statistic/Basic Statistics/Tables* (рис. 10.1.4). Після підтвердження вибору модуля з'явиться діалогове вікно, що дозволяє задати напрям аналізу *Descriptive statistics (описові статистики),* подане на рис. 10.1.5.

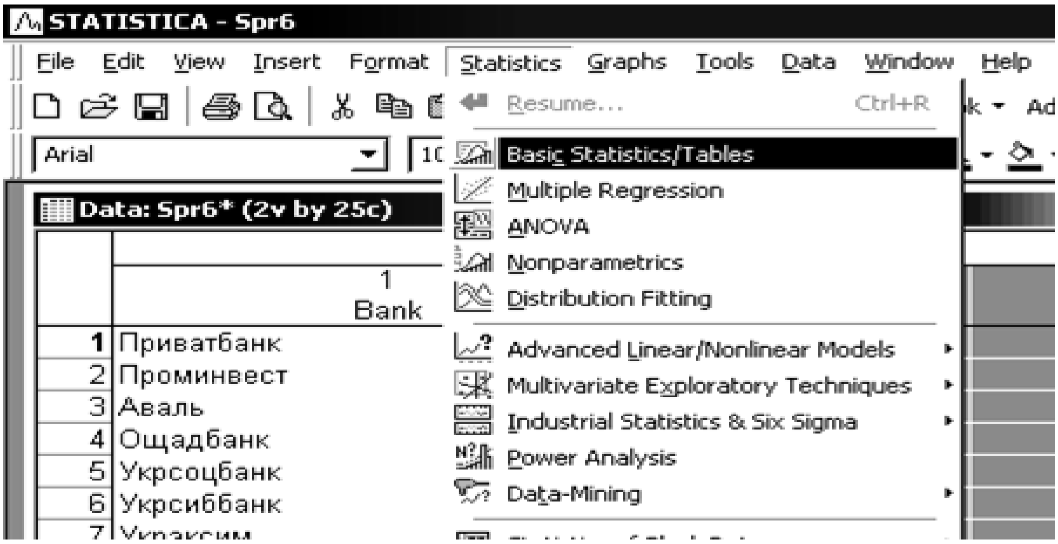


Рис. 10.1.4. **Вибір модуля**

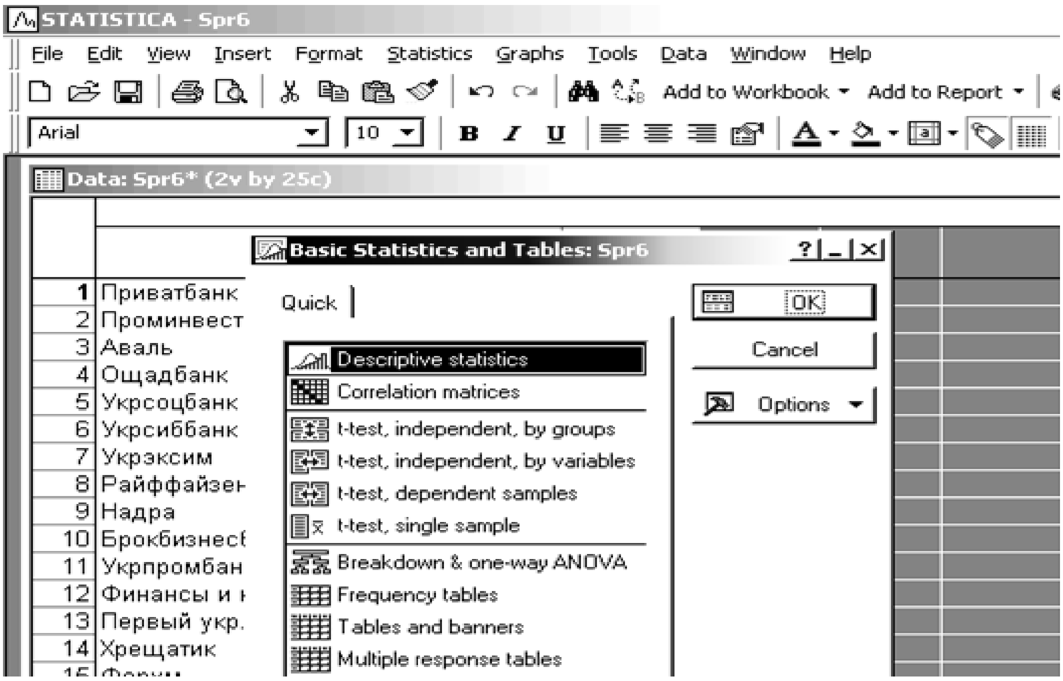


Рис. 10.1.5. **Вибір напряму аналізу**

Після вибору напряму аналізу з'явиться стартова панель модуля, де необхідно задати вихідні параметри: *Variable* (*змінні*) і відповідний набір процедур для подальшого аналізу (рис. 10.1.6).

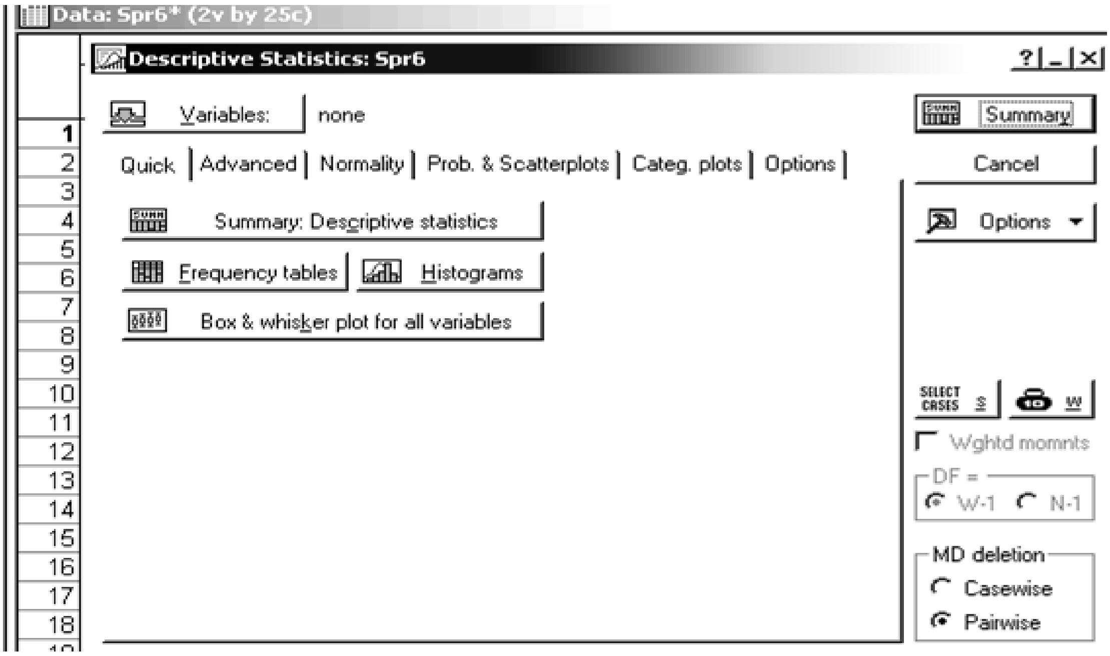


Рис. 10.1.6. **Стартова панель модуля**

Ініціювати кнопку *Variable (змінні)* і у вікні, що з'явилося, вказати показники, за якими здійснюється аналіз. Після зазначення змінних підтвердити свій вибір натисканням кнопки *ОК* (рис. 10.1.7).

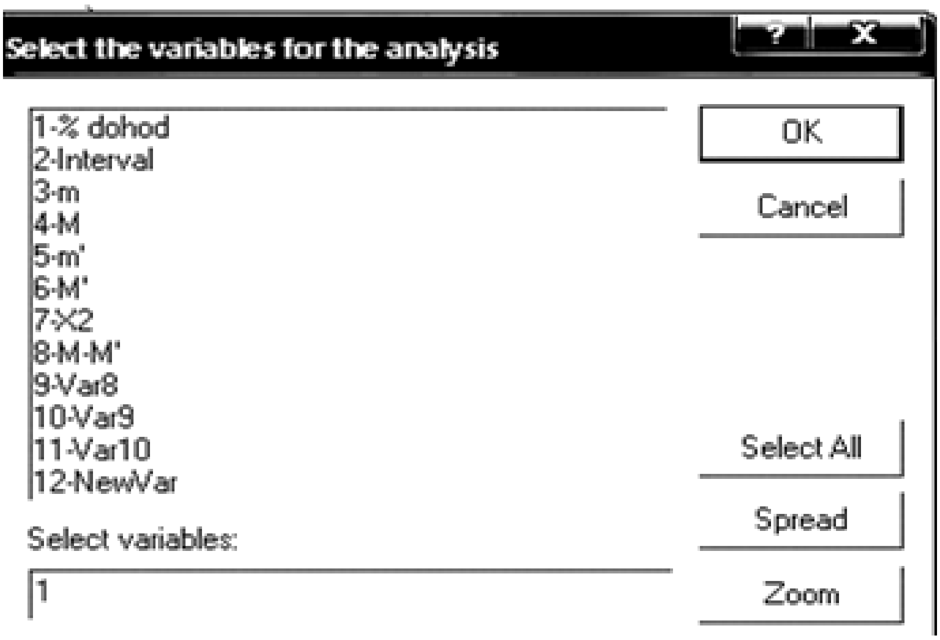


Рис. 10.1.7. **Вибір змінних для аналізу**

Далі після ініціювання вкладки *Advanced* необхідно виділити основні статистики для розрахунку (рис. 10.1.8).

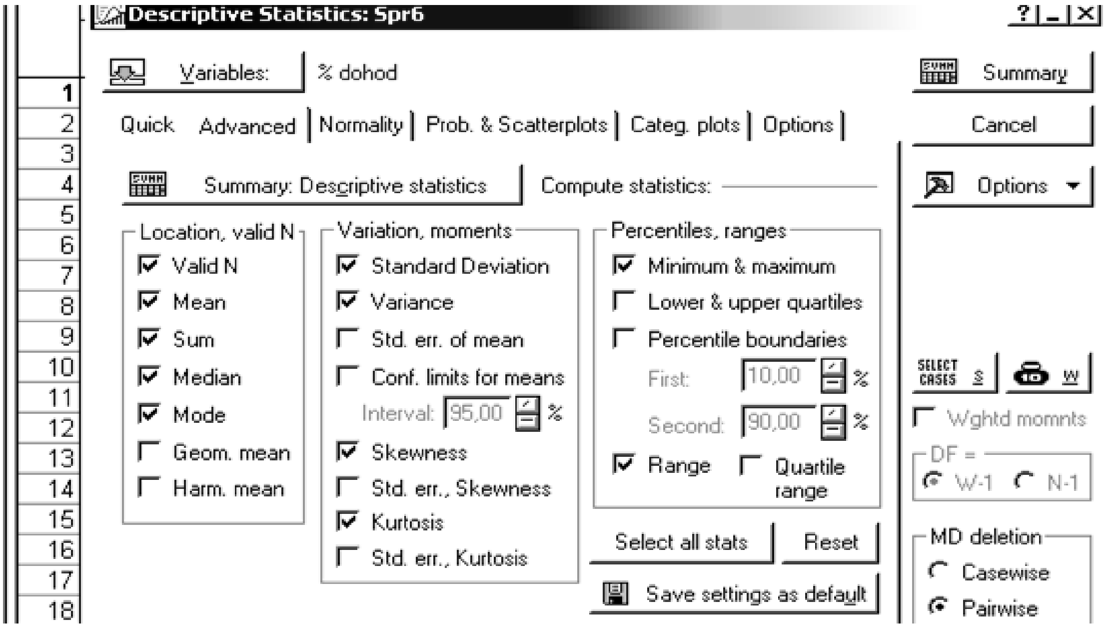


Рис. 10.1.8. **Вибір описових статистик**

Такими є: *Valid N* (*число спостережень*), *Mean* (*середнє*), *Sum* (*сума значень*), *Median* (*медіана*), *Mode* (*мода*), *Standard Deviation* (*середнє квадратичне відхилення*), *Variance* (*дисперсія*), *Skewness* (*коефіцієнт асиметрії*), *Kurtosis* (*коефіцієнт ексцесу*), *Min & Max* (*мінімум і максимум*), *Range* (*розмах вибірки*). Результати розрахунку описових статистик для даної вибірки буде одержано натисканням клавіші *Summary.* Результати наведені на рис. 10.1.9.

Variable	Descriptive Statistics (Spr6)									
	Valid N	Mean	Median	Sum	Minimum	Maximum	Range	Variance	Std.Dev.	Skewness	Kurtosis
% dohod	25	270766,5	145468,0	6769163	35241,00	1187477	1152236	8,370020E+10	289309,9	1,999460	3,715184

Рис. 10.1.9. **Описові статистики**

3. Графічне подання вибіркових даних та їх групування.

Побудувати гістограму й полігон розподілу випадкової величини та провести угрупування вибірки.

Для наочності подання досліджуваної сукупності побудувати полігон розподілу. Для цього необхідно зайти в меню *Graphs/2D Graphs/ Scaterplots* (рис. 10.1.10),вибрати змінні (рис. 10.1.11), задати параметри графіка (рис. 10.1.12) і побудувати полігон розподілу випадкової величини (рис. 10.1.13).

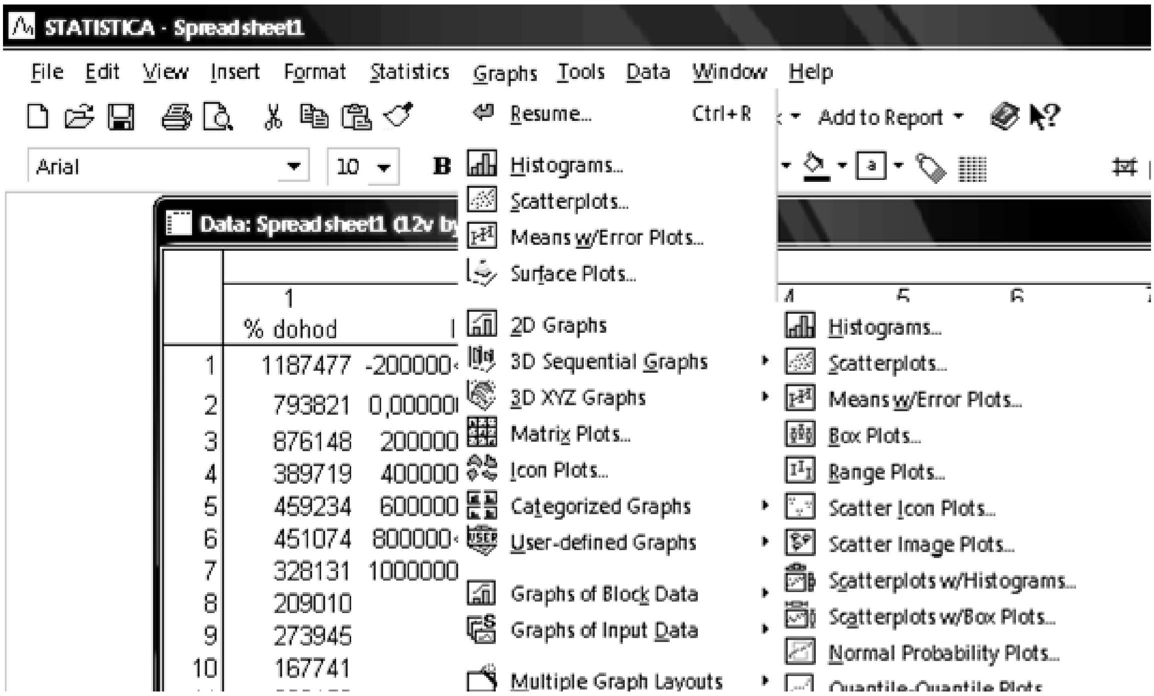


Рис. 10.1.10. **Вибір типу графіка**

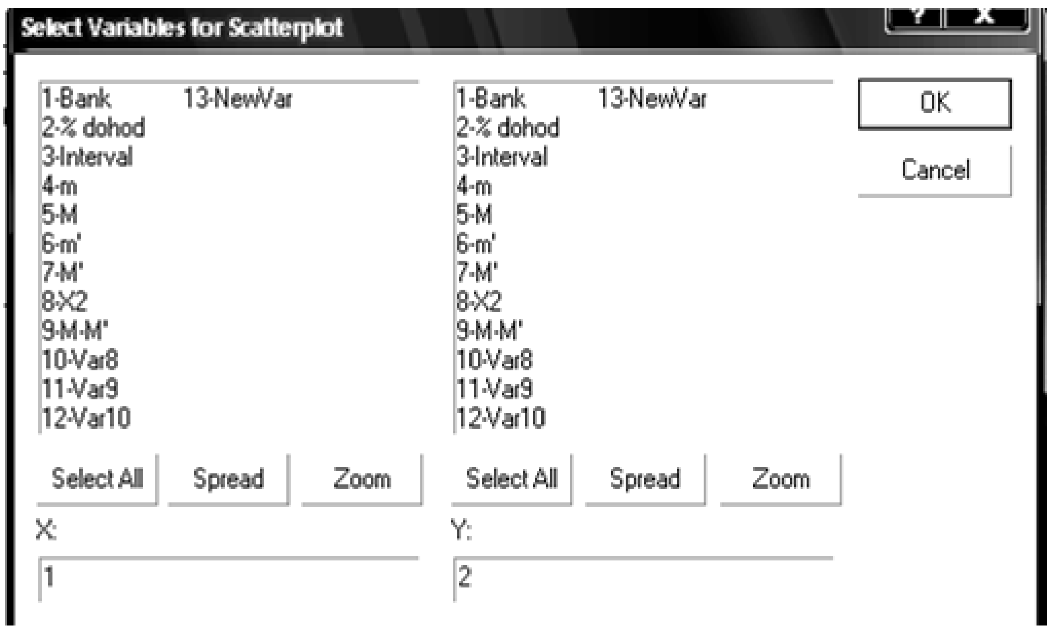


Рис. 10.1.11. **Вибір змінних для побудови графіка**

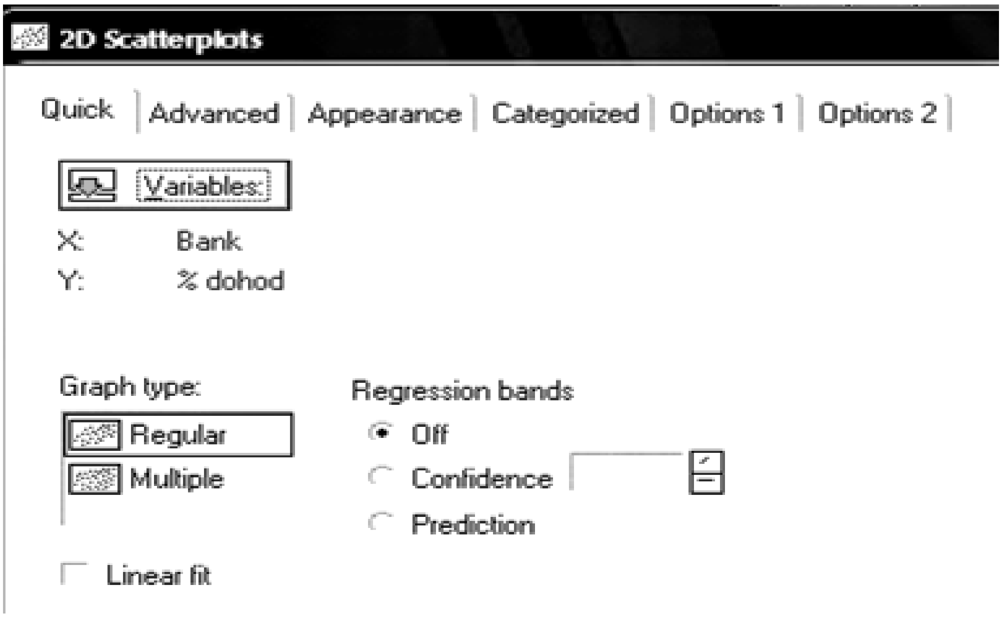


Рис. 10.1.12. **Вибір параметрів графіка**

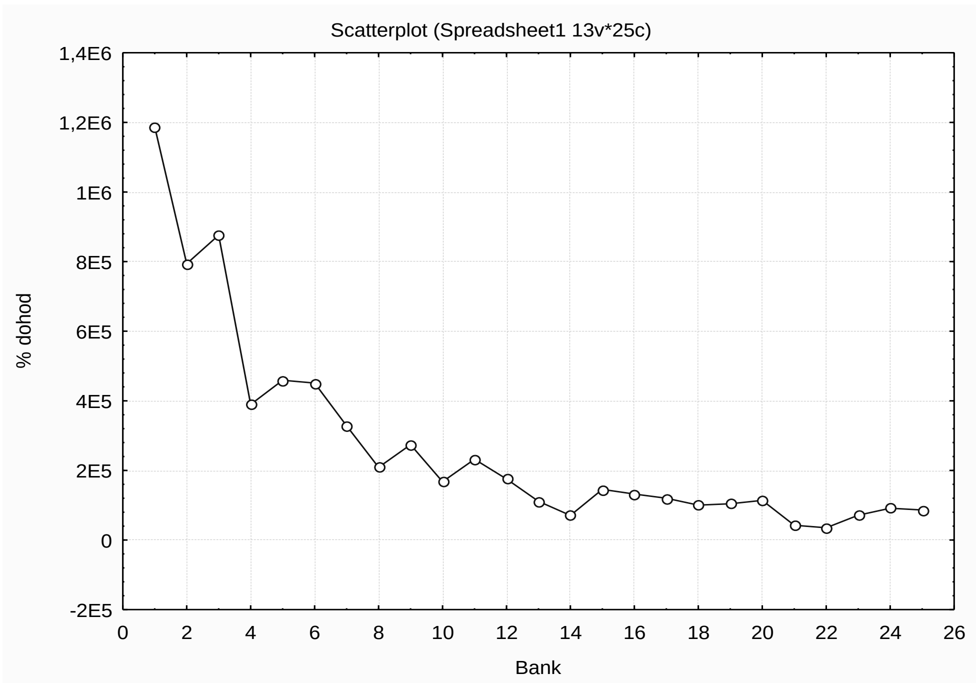


Рис. 10.1.13. **Полігон розподілу випадкової величини**

Подальший аналіз здійснюється в рамках перевірки вибірки на нормальний закон розподілу. Для проведення угрупування вибірки в стартовій панелі модуля вибрати вкладку *Normality*, де можна задавати бажану кількість інтервалів і критерій Колмогорова – Смірнова для тестування вибірки (рис. 10.1.14). Ініціювавши клавішу *Frequency tables* (*Таблиці частот*), буде одержано наступну таблицю (рис. 10.1.15).

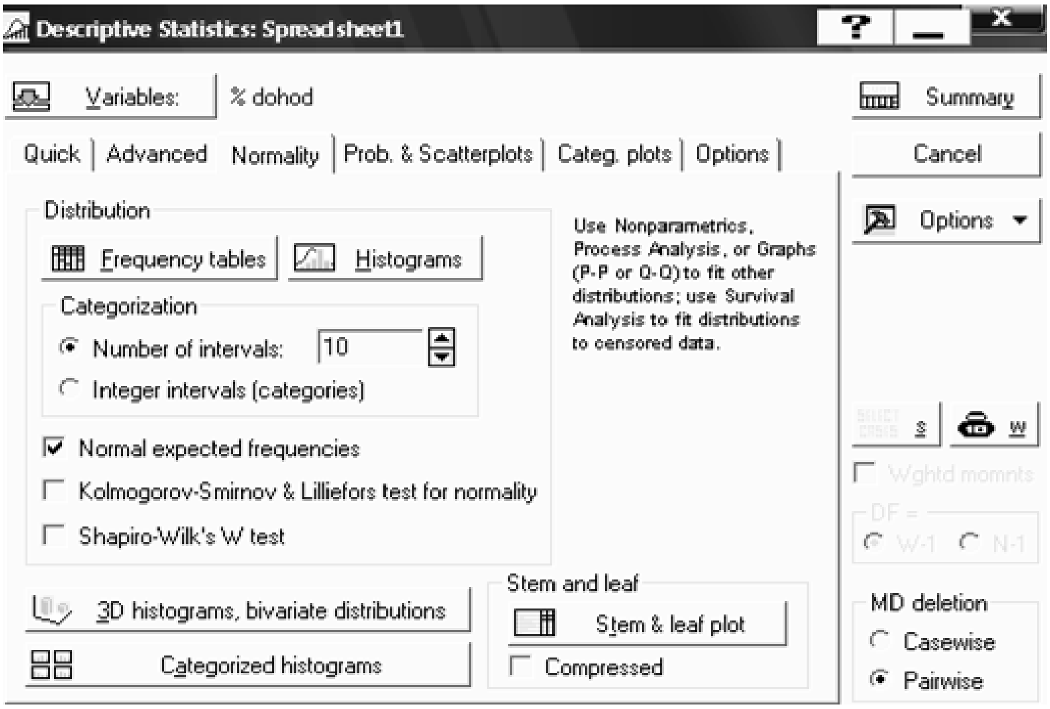


Рис. 10.1.14. **Вибір параметрів угрупування випадкової величини**

Category	Frequency table: % dohod (Spr6) K-S d=,23308, p<,15 ; Lilliefors p<,01 Count	Cumulative Count	Percent of Valid	Cumul % of Valid	% of all Cases	Cumulative % of All
-200000,<x<=0,000000	0	0	0,00000	0,0000	0,00000	0,0000
0,000000<x<=200000,0	15	15	60,00000	60,0000	60,00000	60,0000
200000,0<x<=400000,0	5	20	20,00000	80,0000	20,00000	80,0000
400000,0<x<=600000,0	2	22	8,00000	88,0000	8,00000	88,0000
600000,0<x<=800000,0	1	23	4,00000	92,0000	4,00000	92,0000
800000,0<x<=1000000,	1	24	4,00000	96,0000	4,00000	96,0000
1000000,<x<=1200000,	1	25	4,00000	100,0000	4,00000	100,0000
Missing	0	25	0,00000		0,00000	100,0000

Expected Count	Cumulative Expected	Percent Expected	Cumulative % Expected
4,366527	4,36653	17,46611	17,46611
5,718006	10,08453	22,87202	40,33813
6,726783	16,81132	26,90713	67,24526
4,999658	21,81097	19,99863	87,24389
2,347086	24,15806	9,38835	96,63224
0,695495	24,85355	2,78198	99,41422
0,129962	24,98352	0,51985	99,93407

Рис. 10.1.15. **Результат угрупування вибірки**

Як видно, вихідна сукупність із 25 банків розподілена на 7 інтервалів, у кожному інтервалі розраховані такі характеристики: *Count* (*частота*), *Cumulative Count* (*накопичена частота*), *Percent of Valid* (*% від загальної частоти*), *Cumul % of Valid* (*накопичений % від загальної частоти*), *% of all Cases* (*% від загального числа спостережень*), *Cumulative % of all Cases* (*накопичений % від загального числа спостережень*), *Expected Count* (*теоретична частота*), *Cumulative Expected* (*накопичена теоретична частота*), *% Expected* (*% від загальної теоретичної частоти*), *Cumulative % Expected* (*накопичений % від загальної теоретичної частоти*).

Ініціювавши клавішу *Histograms* (вкладиш *Normality*), буде одержано таку гістограму розподілу з накладеною кривою нормального закону розподілу (рис. 10.1.16).

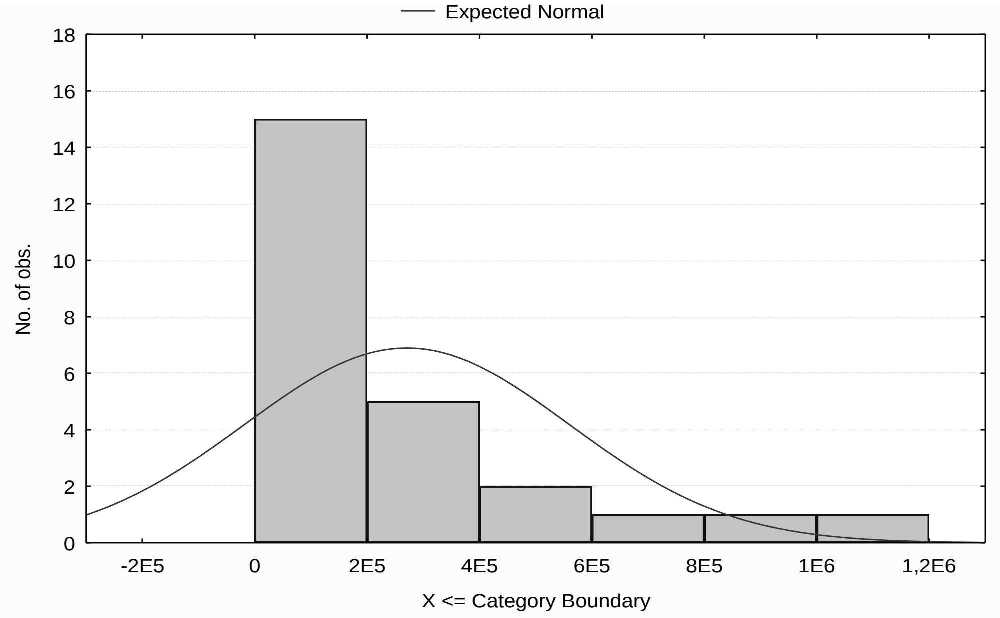


Рис. 10.1.16. **Гістограма розподілу**

4. Аналіз характеру закону розподілу випадкової величини.

Подальший аналіз вибірки передбачає розрахунок критерію Пірсона та Колмогорова – Смірнова для формування відповідних висновків про характер закону розподілу. Для визначення характеру закону розподілу та його відповідності нормальному закону слід дослідити за допомогою графіків порівняння емпіричних, теоретичних і накопичених частот. Вихідні дані для побудови графіків та розраховані значення критерію Пірсона та Колмогорова – Смірнова наведені на рис. 10.1.17. Для побудови графіків інтервальних значень частоти розподілу досліджуваної сукупності необхідно зайти в меню *Graphs/2D Graphs/Scaterplots* вибрати змінні та задати параметри графіка (рис. 10.1.18). На рис. 10.1.19, 10.1.20 наведені графіки порівняння емпіричних і теоретичних частот та накопичених емпіричних та теоретичних частот, які дозволяють зробити висновки про відповідність нормальному закону розподілу й визначити розбіжність частот у кожному з досліджуваних інтервалів.

	1 % dohod	2 Interval	3 m	4 M	5 m'	6 M'	7 X2	8 M-M'
1	1187477	-200000<x<=0,000000	0	0	4,366527	4,36653	4,36653	-4,36653
2	793821	0,000000<x<=200000	15	15	5,718006	10,08453	15,06739	4,91547
3	876148	200000<x<=400000	5	20	6,726783	16,81132	0,44327	3,18868
4	389719	400000<x<=600000	2	22	4,999658	21,81097	1,79971	0,18903
5	459234	600000<x<=800000	1	23	2,347086	24,15806	0,77315	-1,15806
6	451074	800000<x<=1000000	1	24	0,695495	24,85355	0,13332	-0,85355
7	328131	1000000<x<=1200000	1	25	0,129962	24,98352	5,82450	0,01648
8	209010					**X^2**	**28,4078616**	
9	273945						**Kolmogorov**	**0,98309**

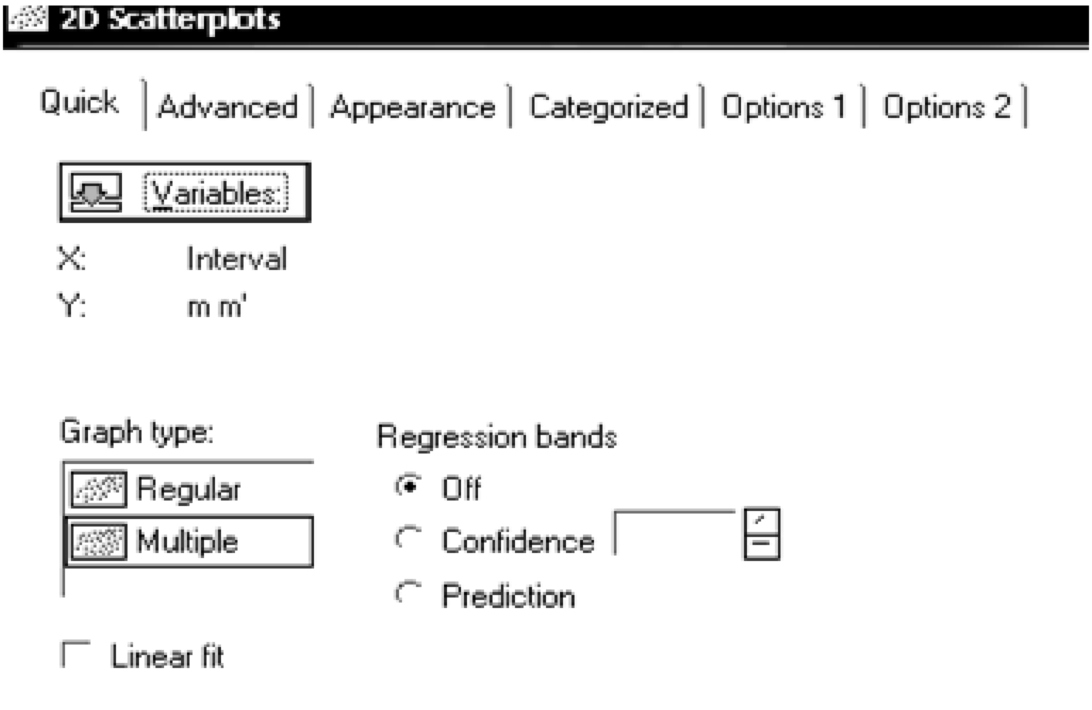


Рис. 10.1.18. **Вибір змінних для побудови графіків**

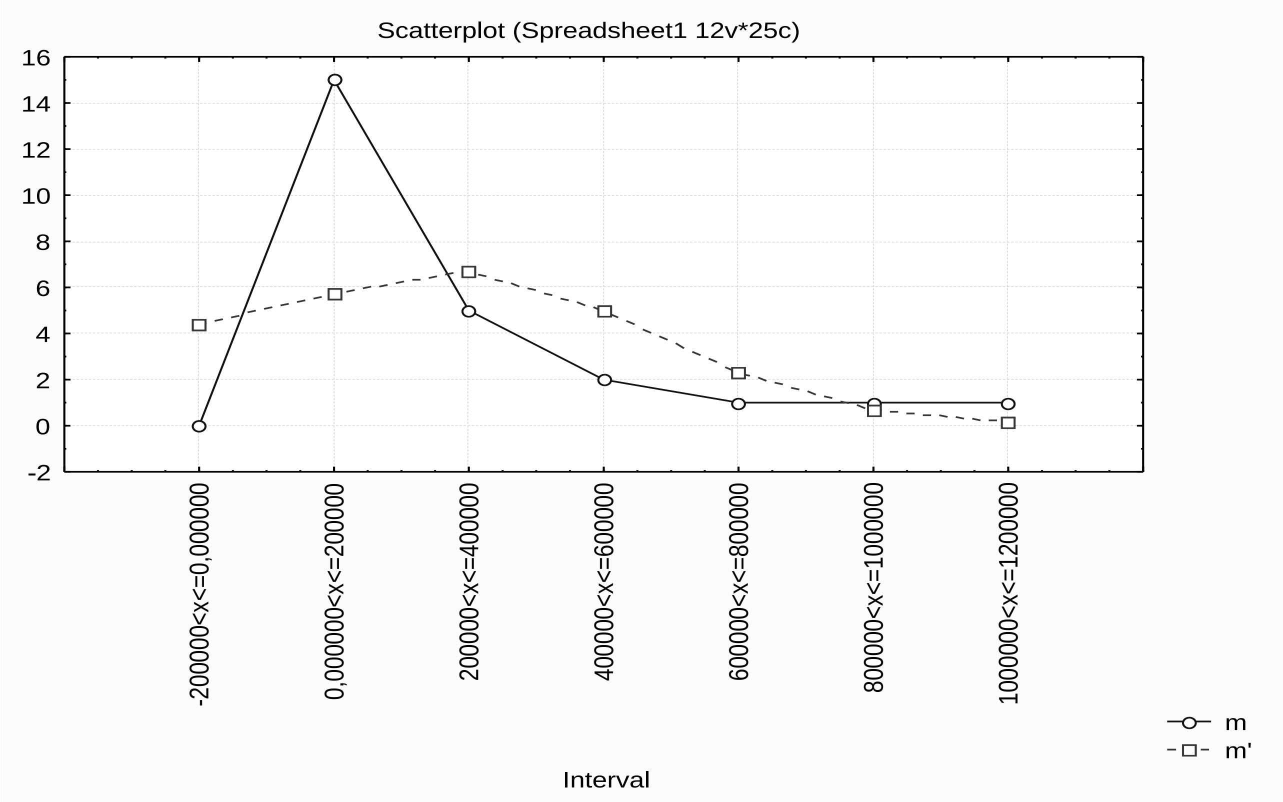


Рис. 10.1.19. **Графік порівняння емпіричних та теоретичних частот**

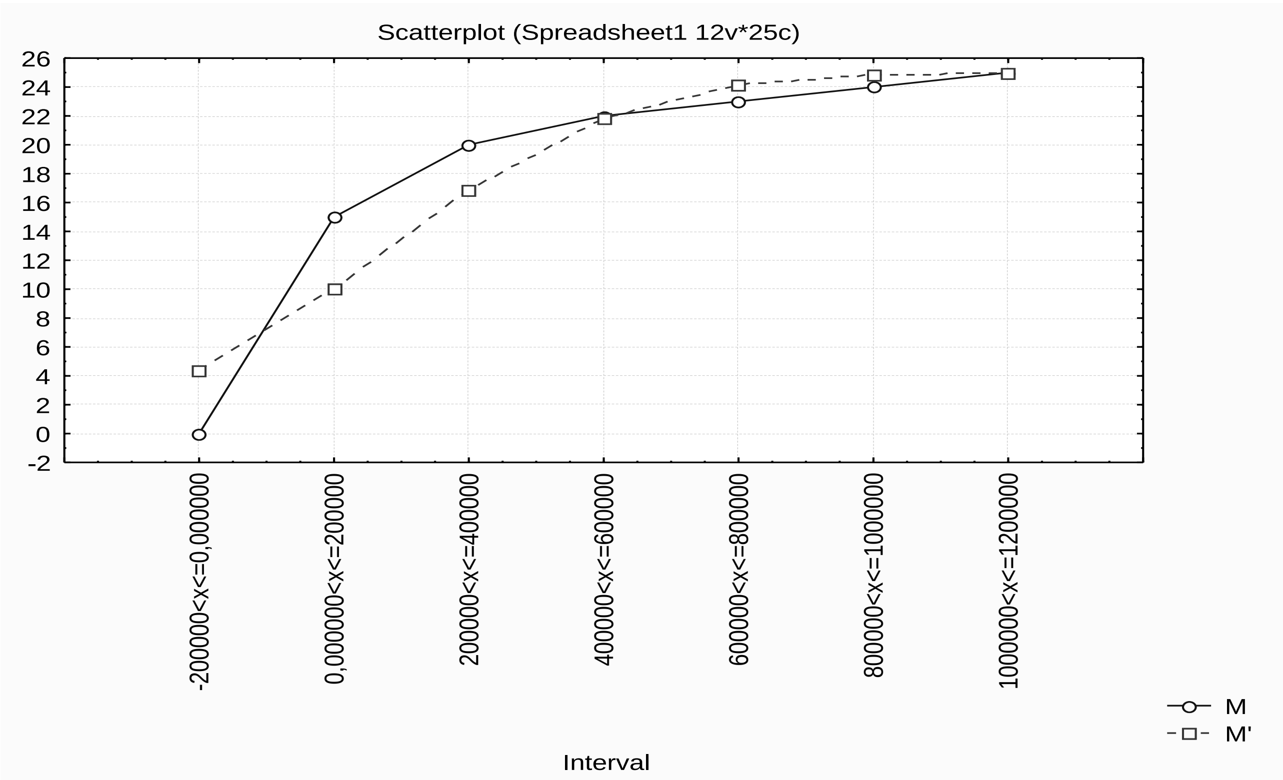


Рис. 10.1.20. **Графік порівняння накопичених емпіричних та теоретичних частот**

Зробити висновки про угрупування даних об'єктів за величиною показника відсоткового доходу. Порівняти отримані значення з табличними за відповідними критеріями й зробити висновки про характер закону розподілу.

Лабораторна робота 2. Побудова й аналіз простої лінійної економетричної моделі

Мета – закріплення теоретичного й практичного матеріалу, набуття навичок побудови й аналізу простих економетричних моделей у модулі *Multiple Regression*.

Завдання – перевірити наявність лінійного зв'язку між відповідними показниками в модулі *Multiple Regression* ППП *Statistica*:

1. Побудувати лінійну економетричну модель та визначити всі її характеристики (параметри моделі, середні квадратичні відхилення параметрів моделі, дисперсію та середнє квадратичне відхилення похибок моделі, коефіцієнти кореляції та детермінації).

2. Перевірити статистичну значущість параметрів та коефіцієнта кореляції за допомогою критерію Стьюдента. Перевірити адекватність моделі за критерієм Фішера.

3. Розрахувати теоретичні значення залежної змінної та похибки моделі, побудувати графік лінійної функції з довірчими інтервалами, побудувати гістограму і графік розподілу похибок, згрупувати дані за значеннями похибок, дати економічну інтерпретацію отриманому групуванню.

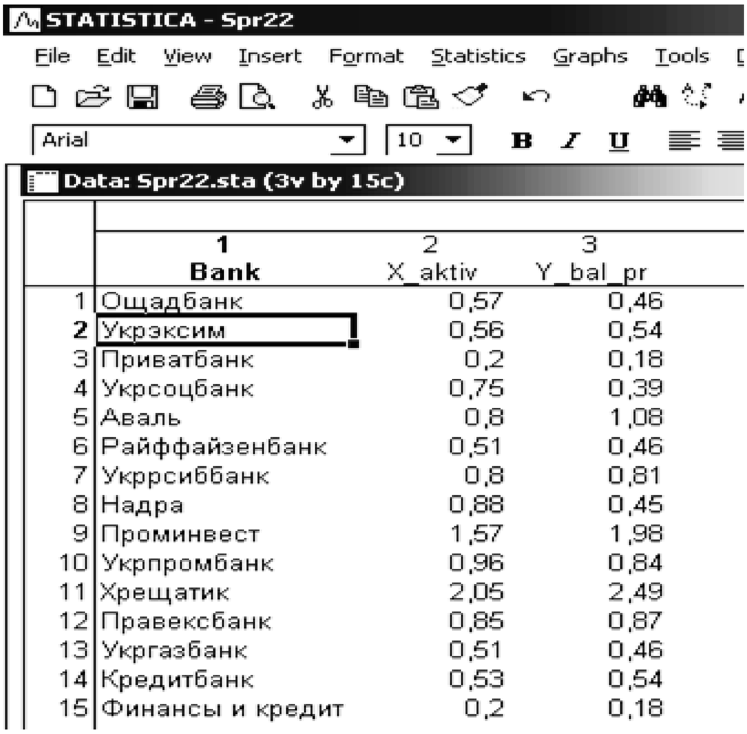
4. Розрахувати прогнозні значення залежної змінної та довірчі інтервали, якщо відомі дані незалежного показника.

5. Зробити висновки відносно адекватності побудованої моделі, дати економічну інтерпретацію отриманої залежності і можливості її теоретичного використання.

Для побудови й аналізу простих лінійних економетричних моделей у ППП *Statistica* передбачений модуль *Multiple Regression* (*Множинна регресія*). Слід розглянути порядок роботи в даному модулі.

1. Запуск *Statistica* і підготовка даних.

У меню програм вибрати програму *Statistica*, після її запуску вибрати у меню пункт *File/New для* підготовки власних даних. З'явиться діалогове вікно, в якому необхідно вказати кількість змінних *(Number of variables)* і кількість випадків *(Number of Cases).* Після уведення натиснути кнопку вікна *OK.* Після заповнення всіх комірок поля даних буде одержано таблицю, аналогічну поданій на рис. 10.2.1.



STATISTICA - Spr22

Data: Spr22.sta (3v by 15c)

	1 Bank	2 X_aktiv	3 Y_bal_pr
1	Ощадбанк	0,57	0,46
2	Укрэксим	0,56	0,54
3	Приватбанк	0,2	0,18
4	Укрсоцбанк	0,75	0,39
5	Аваль	0,8	1,08
6	Райффайзенбанк	0,51	0,46
7	Укррсиббанк	0,8	0,81
8	Надра	0,88	0,45
9	Проминвест	1,57	1,98
10	Укрпромбанк	0,96	0,84
11	Хрещатик	2,05	2,49
12	Правексбанк	0,85	0,87
13	Укргазбанк	0,51	0,46
14	Кредитбанк	0,53	0,54
15	Финансы и кредит	0,2	0,18

Рис. 10.2.1. **Вихідні дані**

2. Розрахунки.

Щоб приступити до обчислювальних процедур, необхідно вибрати позицію меню *Statistics/Multiple Regression* (рис. 10.2.2). Після підтвердження вибору модуля перед вами з'явиться стартова панель даного модуля, де необхідно задати змінні для аналізу (рис. 10.2.3).

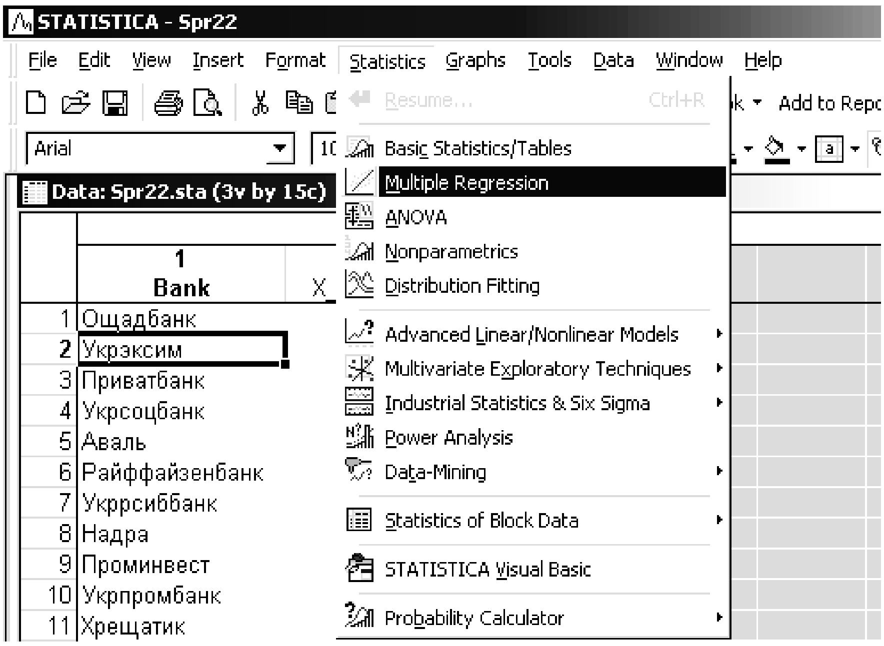


Рис. 10.2.2. **Вибір модуля**

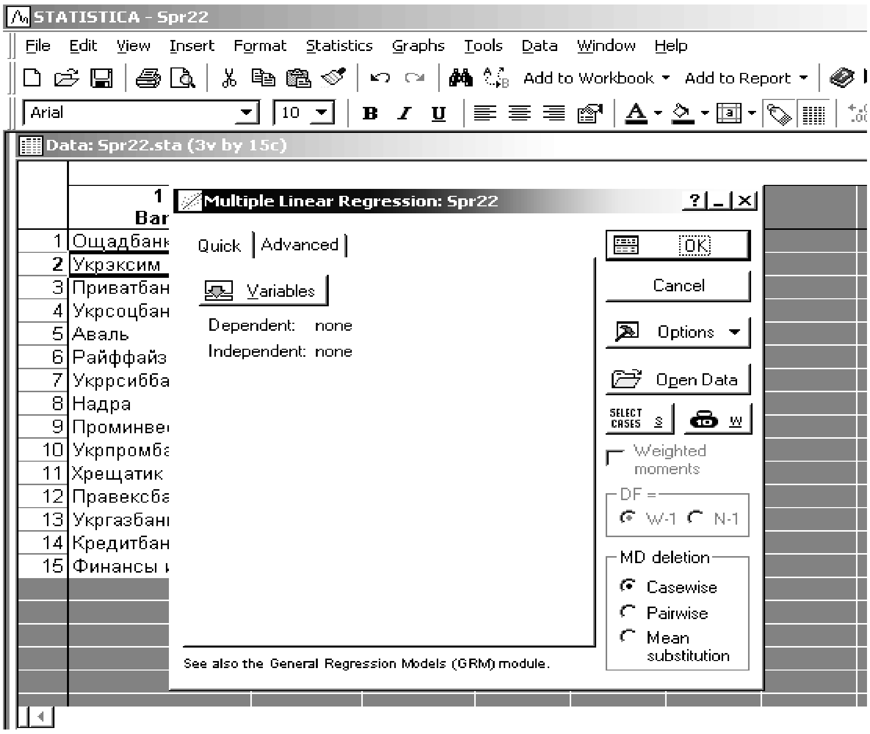


Рис. 10.2.3. **Стартова панель модуля**

Ініціювати кнопку *Variables (змінні)* і у вікні, що з'явилося, вказати *Dependent* (залежну) і *Independent* (незалежну) змінні для побудови простої регресійної моделі. Вибір змінних наведений на рис. 10.2.4. Після вказівки змінних підтвердити свій вибір натисканням кнопки *ОК* (рис. 10.2.5).

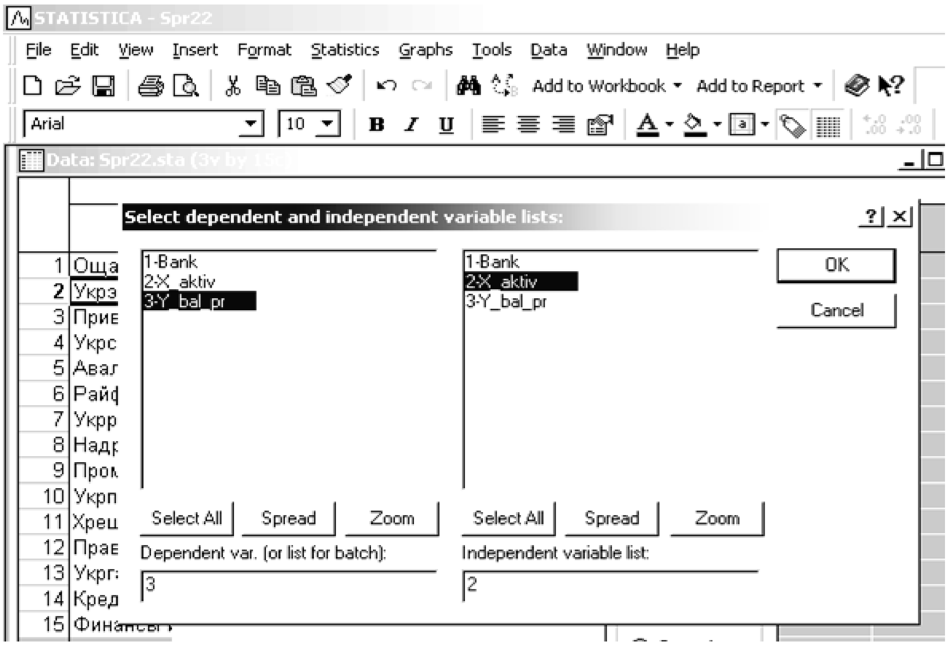


Рис. 10.2.4. **Вибір змінних для аналізу**

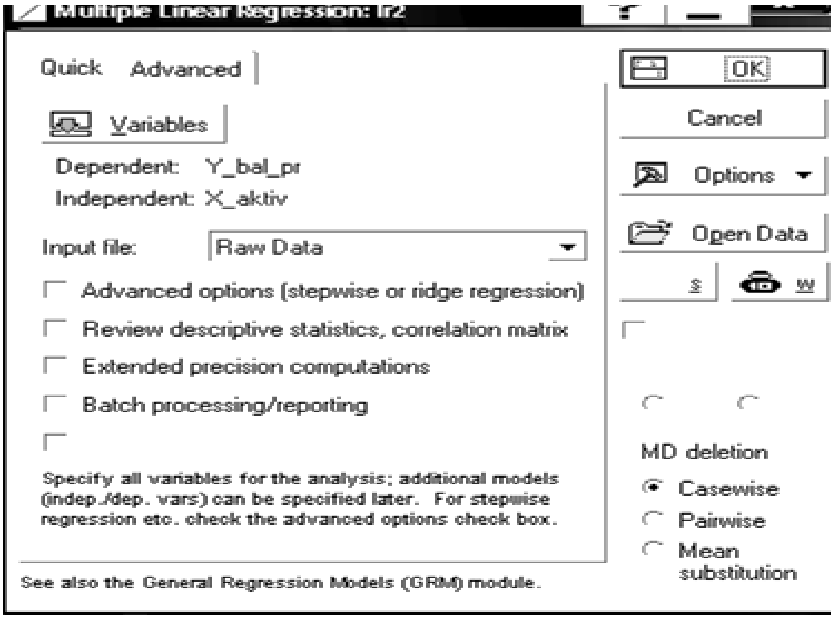


Рис. 10.2.5. **Підтвердження вибору змінних**

3. Побудова моделі, визначення її характеристик, перевірка її адекватності та статистичної значущості.

Побудувати лінійну економетричну модель і визначити всі її характеристики. Результати побудови лінійної економетричної моделі будуть подані в діалоговому вікні (рис. 10.2.6). У верхній частині вікна подана основна інформація моделі, у нижній частині знаходяться функціональні кнопки, що дозволяють всебічно розглянути результати аналізу.

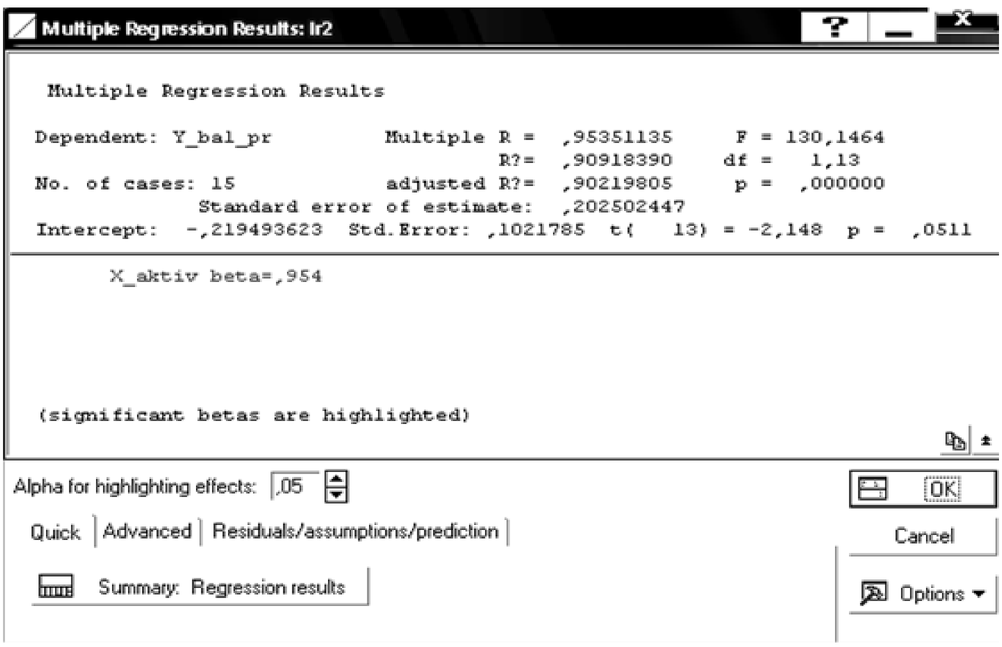


Рис. 10.2.6. **Вікно результатів регресійного аналізу**

Ініціювавши кнопку *Summary: Regression results* (на вкладці *Quick*) (*результати регресійного аналізу*), визначити найважливіші характеристики моделі й ступінь її адекватності (рис. 10.2.7).

	Regression Summary for Dependent Variable: Y_bal_pr (Spr22) R= ,95351135 R?= ,90918390 Adjusted R?= ,90219805 F(1,13)=130,15 p<,00000 Std.Error of estimate: ,20250					
N=15	Beta	Std.Err. of Beta	B	Std.Err. of B	t(13)	p-level
Intercept			-0,219494	0,102178	-2,14814	0,051130
X_aktiv	0,953511	0,083581	1,279592	0,112164	11,40817	0,000000

Рис. 10.2.7. **Результати регресійного аналізу**

Проаналізувати отримані результати моделі:

1) аналіз адекватності:

R = 0.9535 – коефіцієнт множинної кореляції (у випадку простої лінійної регресії дорівнює модулю коефіцієнта парної кореляції);

R^2 = 0.9091 – коефіцієнт детермінації моделі;

Adjusted R^2 = 0.9021 – скорегований коефіцієнт детермінації на число спостережень і число параметрів моделі;

F (1,13) = 130.15 – критерій Фішера статистичної значущості моделі з числом ступенів свободи та рівнем значущості *p*;

Std.Error of estimate = 0.2025 – середнє квадратичне відхилення помилок моделі; це статистика – міра розсіву досліджуваних значень відносно регресійної прямої;

2) аналіз параметрів та їх статистичної значущості:

Beta (*a*1) = (0.9535) – стандартизовані значення коефіцієнтів регресії (ваги) – оцінюються за стандартизованими даними з вибірковим середнім, яке дорівнює нулю і стандартним відхиленням, яке дорівнює одиниці;

Std.Error of Beta(*a*1) = 0.08358 – середнє квадратичне відхилення параметрів моделі для стандартизованих коефіцієнтів регресії;

B (*a*1, *a*2) = (-0.2194; 1.2795) – нестандартизовані параметри моделі, а отже модель має вигляд: *Y = -0.2194 + 1.2796x;*

Std.Error of B = (0.10; 0.11) – середнє квадратичне відхилення параметрів моделі;

t(13) = (-2.14; 11.4) – значущість параметрів за критерієм Стьюдента;

p-level = (0.051; 0.000) – рівень значущості критерію Стьюдента.

Побудувати графік лінійної функції з довірчими інтервалами. Для цього в меню *Graphs/Scatterplots* необхідно вказати змінні, лінію рівня й довірчі інтервали (рис. 10.2.8). Ініціювавши кнопку *ОК,* буде одержано такий графік (рис. 10.2.9).

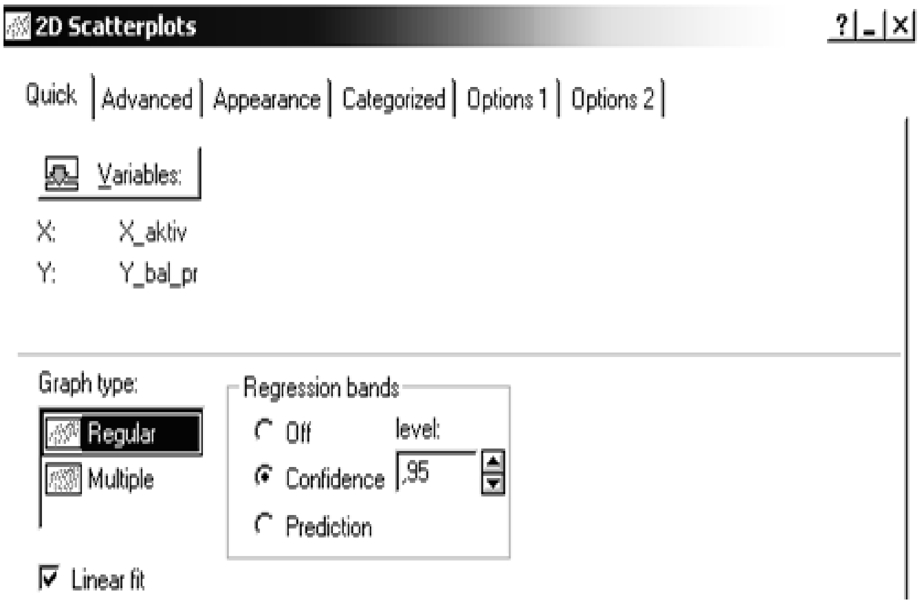


Рис. 10.2.8. **Завдання параметрів графіка**

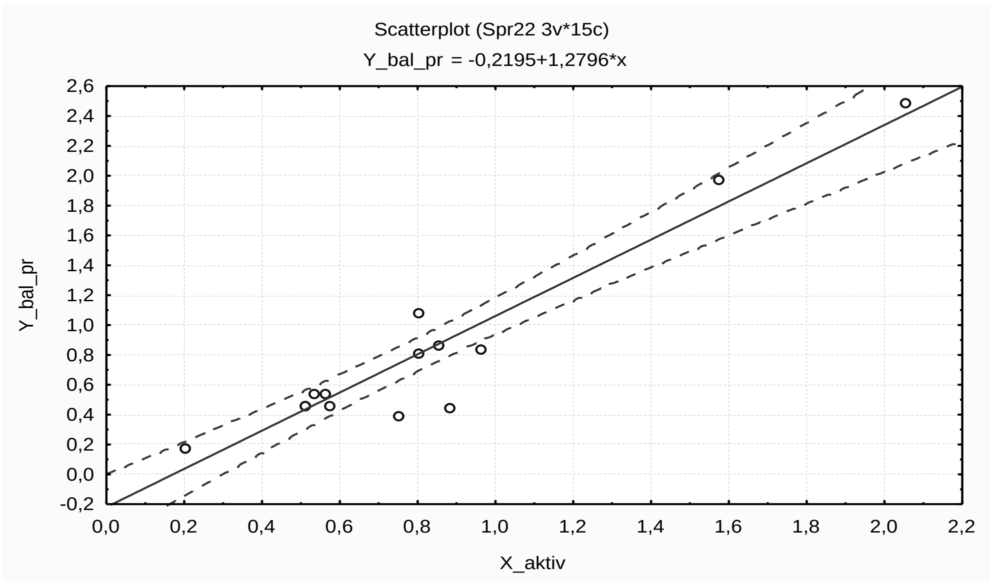


Рис. 10.2.9. **Графік лінійної залежності**

Для перевірки гіпотези про значущість регресійної моделі використовується дисперсійний аналіз. Для цього необхідно ініціювати кнопку *Advanced / ANOVA* у нижній частині інформаційного вікна (рис. 10.2.10).

Результати дисперсійного аналізу для досліджуваної моделі наведені на рис. 10.2.11. У даній таблиці наведено суму квадратів відхилень за регресією *(Sums of Squares Regress),* суму квадратів відхилень похибок *(Sums of Squares Residual),* дисперсію похибок *(Mean Squares Residual)* та критерій Фішера.

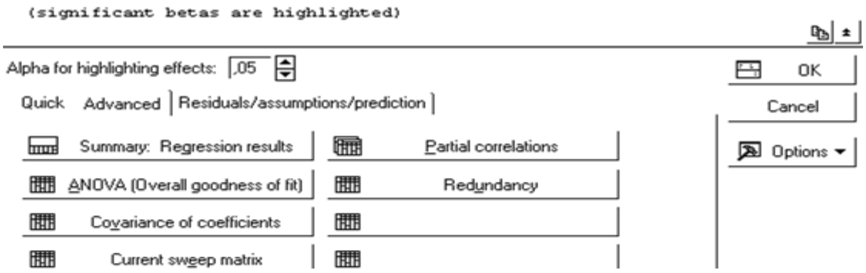


Рис. 10.2.10. **Вибір вікна дисперсійного аналізу**

Effect	Analysis of Variance; DV: Y_bal_pr (Spreads Sums of Squares	df	Mean Squares	F	p-level
Regress.	5,336946	1	5,336946	130,1464	0,000000
Residual	0,533094	13	0,041007		
Total	5,870040				

Рис. 10.2.11. **Таблиця дисперсійного аналізу**

Для розрахунку основних характеристик (середні значення та середні квадратичні відхилення) залежної та незалежної змінних та матриці коефіцієнтів парних кореляцій необхідно ініціювати опцію *Means & standard deviations* та *Correlations,* які знаходяться в меню *Descriptive statistics* (вкладка *Advanced*) (рис. 10.2.12).

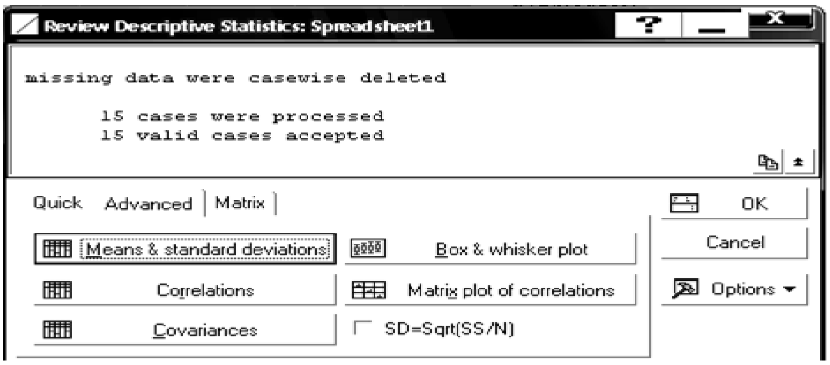


Рис. 10.2.12. **Вікно вибору розрахунку основних статистик**

Основні статистики та матриця коефіцієнтів парної кореляції наведені на рис. 10.2.13.

Variable	Means and Standard D		
	Means	Std.Dev.	N
X_aktiv	0,78266	0,48251	15
Y_bal_pr	0,78200	0,64752	15

Variable	Correlations (lr2)	
	X_aktiv	Y_bal_pr
X_aktiv	1,00000	0,95351
Y_bal_pr	0,95351	1,00000

Рис. 10.2.13. **Основні статистики та матриця коефіцієнтів парної кореляції**

4. Аналіз помилок.

Розрахувати теоретичні значення залежної змінної й помилки моделі. Побудувати гістограму й графік розподілу помилок на нормальному ймовірнісному папері.

Щоб розрахувати й проаналізувати залишки, у нижній частині інформаційного вікна результатів регресійного аналізу є опція *Perform residual analysis* (*всебічний аналіз залишків*) (рис. 10.2.14). Ініціювавши дану опцію, буде одержано меню для аналізу помилок моделі (рис. 10.2.15).

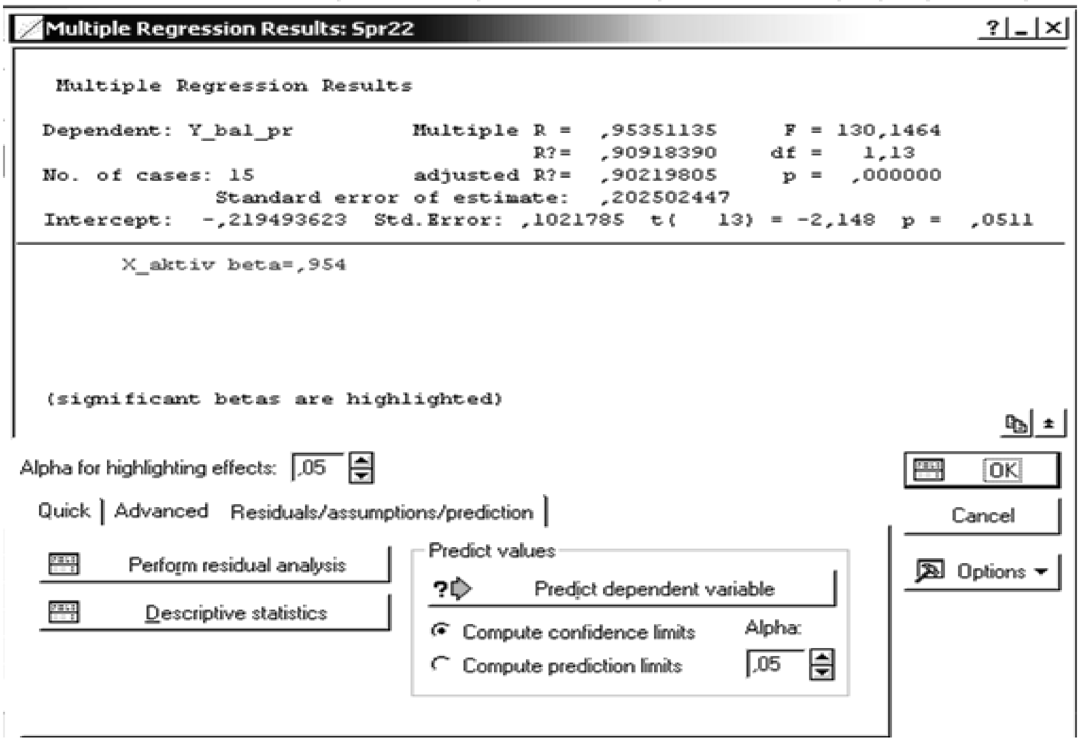


Рис. 10.2.14. **Вікно вибору аналізу похибок**

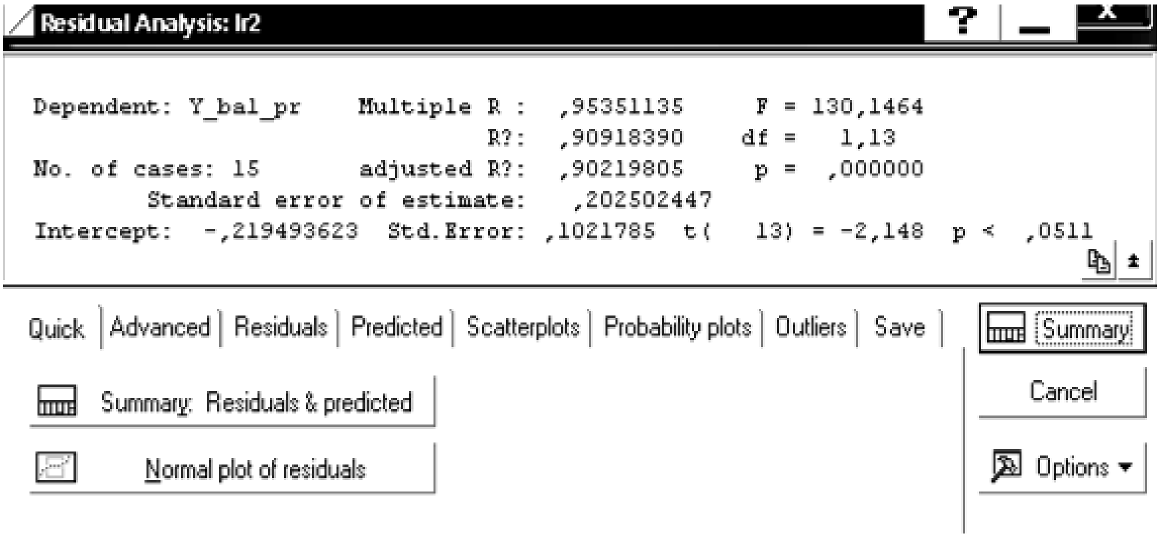


Рис. 10.2.15. **Меню аналізу помилок моделі**

Графік розсіву помилок моделі в діапазоні $\pm 3\sigma$ можна отримати, ініціювавши опцію *Residuals/Casewise plot of residuals* (рис. 10.2.16). Даний графік аналізує властивість сталості дисперсії помилок.

Кнопка аналізу помилок *Quik*/*Summary: Residuals & Predicted* відображає спостережувані значення залежної змінної (*Observed value),* теоретичні значення залежної змінної *(Predicted value)* і помилки моделі *(Residual)* як різницю спостережуваних і теоретичних значень (рис. 10.2.17).

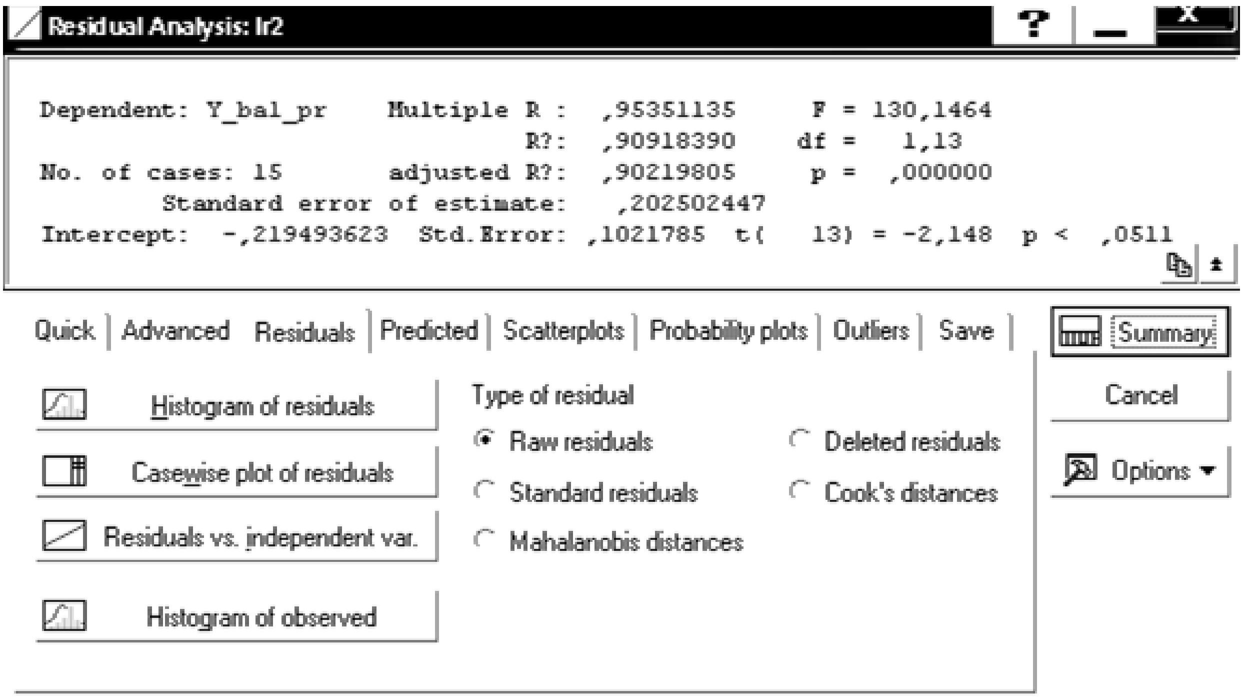


Рис. 10.2.16. **Аналіз дисперсії помилок моделі**

Case	Raw Residuals -3s . . 0 . . +3s	Raw Residual (Spreadsheet1) Dependent variable: Y_bal_pr Observed Value	Predicted Value	Residual
1	. . . *. . . .	0,460000	0,509873	-0,049874
2	. . . .* . . .	0,540000	0,497078	0,042922
3	. . . . *. . .	0,180000	0,036425	0,143575
4	. . *	0,390000	0,740200	-0,350200
5	. . . . . * . .	1,080000	0,804180	0,275820
6	. . . .* . . .	0,460000	0,433098	0,026902
7	. . . * . . .	0,810000	0,804180	0,005820
8	. *.	0,450000	0,906547	-0,456547
9	. . . . * . .	1,980000	1,789465	0,190535
10	. . .*	0,840000	1,008914	-0,168914
11	. . . . * . . .	2,490000	2,403669	0,086331
12	. . . * . . .	0,870000	0,868159	0,001841
13	. . . .* . . .	0,460000	0,433098	0,026902
14	. . . . * . . .	0,540000	0,458690	0,081310
15	. . . . *. . .	0,180000	0,036425	0,143575
Minimum	. *.	0,180000	0,036425	-0,456547
Maximum	. . . . . * . .	2,490000	2,403669	0,275820
Mean	. . . * . . .	0,782000	0,782000	0,000000
Median	. . . .* . . .	0,540000	0,740200	0,026902

Рис. 10.2.17. **Аналіз помилок моделі**

Для перевірки гіпотези про нормальність розподілу похибок необхідно побудувати графік розподілу помилок на нормальному ймовірнісному папері *(Normal plot of residuals),* який наведено на рис. 10.2.18.

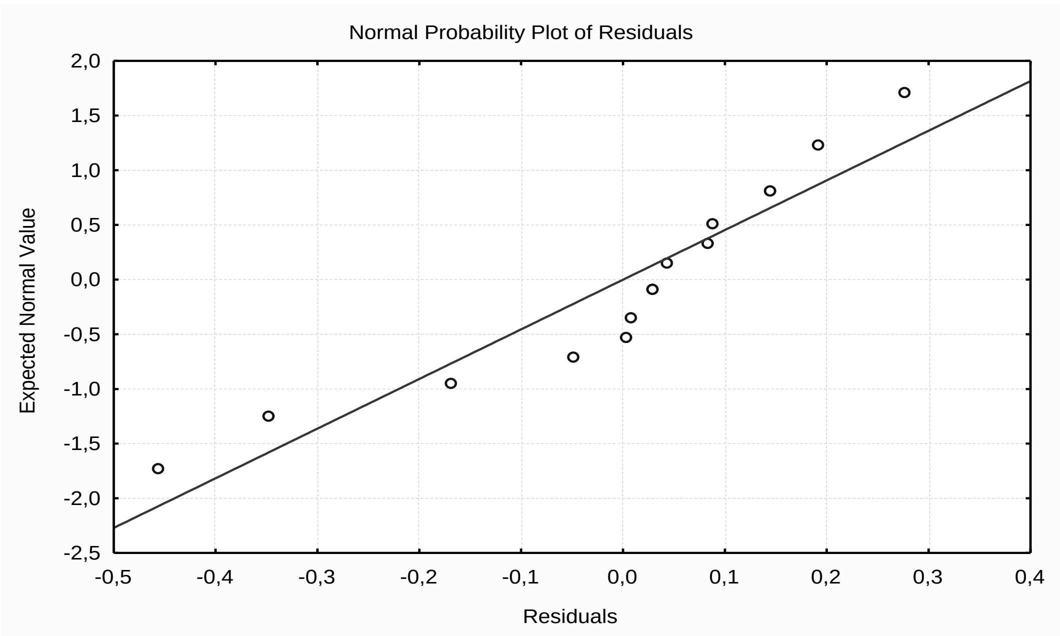


Рис. 10.2.18. **Графік розподілу помилок на нормальному ймовірнісному папері**

Оскільки одна з основних гіпотез щодо випадкової величини говорить, що помилки повинні бути розподілені за нормальним законом, слід подати гістограму розподілу помилок *(Residuals/Histogram plot of residuals)* і проаналізувати її (рис. 10.2.19). Із побудованого графіка видно, що розподіл помилок не відповідає кривій нормального закону, очевидно через малу кількість спостережень.

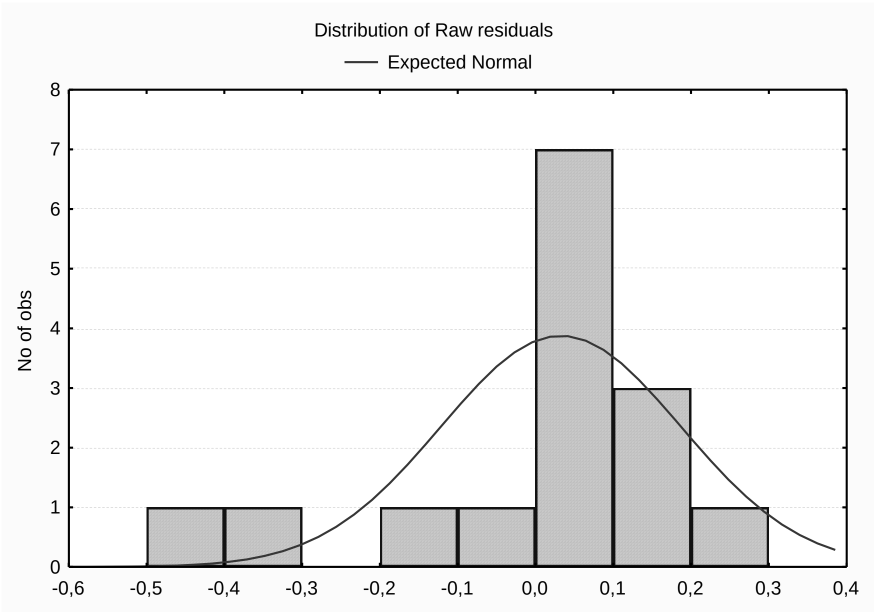


Рис. 10.2.19. **Гістограма розподілу помилок**

5. Розрахунок прогнозу значень залежної змінної та довірчих інтервалів зміни.

Оскільки модель є адекватною, її параметри значущі, то за моделлю можна скласти прогноз. Щоб розрахувати прогнозні значення залежної змінної, у нижній частині вікна результатів регресійного аналізу є опція *Predict dependent variable* (прогнозування залежної змінної) (див. рис. 10.2.14). Ініціювавши дану опцію, необхідно вказати значення незалежної змінної, для якої необхідно спрогнозувати залежну величину (рис. 10.2.20).

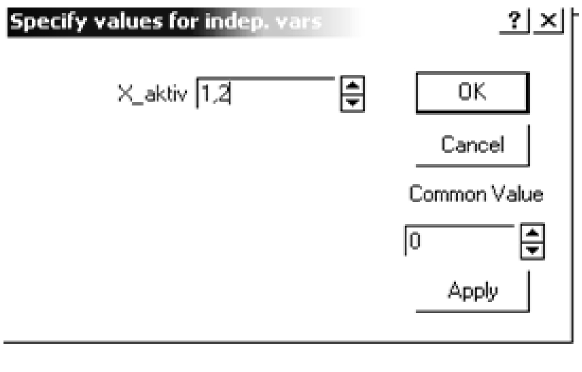


Рис. 10.2.20. **Значення незалежної змінної**

Результати прогнозування подаються у вигляді таблиці, у якій зазначені коефіцієнти моделі й порядок розрахунків (рис. 10.2.21).

Variable	Predicting Values for (Spr22) variable: Y_bal_pr		
	B-Weight	Value	B-Weight * Value
X_aktiv	1,279592	1,200000	1,535510
Intercept			-0,219494
Predicted			1,316016
-95,0%CL			1,164405
+95,0%CL			1,467627

Рис. 10.2.21. **Результати прогнозу**

Прогнозне значення залежної змінної (*Predicted*) = 1,3160; і довірчі інтервали для прогнозного значення:

$$1{,}164405 \le 1{,}316016 \le 1{,}467627.$$

Проведено всебічний аналіз однофакторної лінійної економетричної моделі залежності прибутку найбільших банків України від значення чистих активів їх діяльності.

Лабораторна робота 3. Побудова й аналіз множинної лінійної економетричної моделі

Мета – закріплення теоретичного й практичного матеріалу за темою "Множинна регресія", набуття навичок побудови й аналізу багатофакторних економетричних моделей у модулі *Multiple Regression*.

Завдання – необхідно перевірити наявність лінійного множинного зв'язку між відповідними показниками в модулі *Multiple Regression* ППП *Statistica*:

1. Побудувати лінійну багатофакторну економетричну модель (включити всі відповідні фактори) і визначити всі її характеристики (параметри моделі, середні квадратичні відхилення параметрів моделі, дисперсію та середнє квадратичне відхилення похибок моделі, коефіцієнти множинної кореляції і детермінації).

2. Перевірити статистичну значущість параметрів моделі, коефіцієнта множинної кореляції. Перевірити адекватність моделі за допомогою критерію Фішера.

3. Навести таблиці з теоретичними значеннями залежного показника та значень похибок моделі. Побудувати графік лінійної функції з довірчими інтервалами. Знайти прогнозне значення залежної змінної та довірчі інтервали, якщо відомі дані про майбутні значення незалежних показників.

4. Перевірити модель на наявність мультиколінеарності. Навести матрицю парних кореляцій для факторних ознак. За методом Фаррара – Глобера оцінити суттєвість мультиколінеарності.

5. Навести результати дослідження моделі за критерієм Дарбіна – Уотсона і нециклічного коефіцієнта автокореляції. Зробити висновки щодо наявності автокореляції. Побудувати гістограму і графік розподілу похибок. Навести групування даних за значеннями похибок, дати економічну інтерпретацію.

6. Зробити висновки щодо адекватності побудованої багатофакторної моделі, дати економічну інтерпретацію моделі в цілому.

7. Виключити з моделі фактори, які найменш впливають на залежну змінну або взаємозалежні між собою (використовувати результати критерію Стьюдента та коефіцієнти парних кореляцій). Визначити всі вказані характеристики побудованих моделей, зробити висновки щодо їх адекватності.

8. Побудувати і дати інтерпретацію моделей, побудованих на основі методів покрокового включення і покрокового виключення змінних.

9. Якщо в моделі присутня мультиколінеарність, то для оцінювання параметрів використати метод рідж-регресії. Визначити всі характеристики моделі. Навести графіки зміни значень оцінок параметрів рідж-моделі залежно від значення параметра. Оцінити ступінь зсуву оцінок параметрів.

10. Зробити порівняльний аналіз побудованих моделей. Визначити найбільш адекватну й економічно інтерпретовану модель.

Для побудови й всебічного аналізу множинних лінійних економетричних моделей у ППП *Statistica* передбачений модуль *Multiple Regression* (*множинна регресія*). Слід розглянути порядок роботи в даному модулі під час побудови багатофакторної моделі.

1. Запуск програми *Statistica* і підготовка даних.

У меню програм вибрати програму *Statistica*, після її запуску вибрати у меню пункт *File/New для* підготовки власних даних. Після введення натисніть кнопку вікна *OK.* Після заповнення всіх комірок поля даних буде одержано таблицю, аналогічну наведеній на рис. 10.3.1.

	1 X1	2 X2	3 X3	4 Y
1	0,45	0,83	0,69	0,13
2	0,81	0,8	0,83	0,77
3	0,46	0,66	0,16	0,25
4	1,08	1,11	0,65	0,5
5	0,39	0,67	0,34	0,12
6	0,18	0,19	0,82	0,22
7	0,54	0,35	5,36	0,82
8	0,46	0,23	0,54	0,91
9	0,87	0,81	1,94	0,96
10	0,71	0,9	0,22	0,16
11	0,07	0,13	0,83	0,11
12	0,14	0,24	0,35	0,02
13	0,03	0,05	0,66	0,01
14	0,136	0,176	0,18	0,043
15	0,073	0,064	0,16	0,969
16	0,105	0,066	0,65	0,172
17	0,15	0,182	0,34	0,108
18	0,189	0,359	0,82	0,064
19	9,155	1,234	0,79	4,654
20	4,744	0,339	0,43	0,076
21	63,898	16,332	0,15	25,152
22	7,314	1,004	0,03	1,087
23	0,85	0,453	0,126	0,024
24	0,184	2,635	0,445	0,914
25	0,01	0,031	0,087	0,001

Рис. 10.3.1. **Вихідні дані**

На рис. 10.3.1 *x1*, *x2*, *x3* – факторні незалежні змінні, *y* – результуюча залежна змінна.

2. Побудова лінійної багатофакторної економетричної моделі.

Щоб почати обчислювальні процедури, необхідно ввійти в позицію меню *Statistics/Multiple Regression*. Після підтвердження вибору модуля з'явиться стартова панель даного модуля, де необхідно задати змінні для аналізу.

Ініціюйте кнопку *Variables (змінні)* і у вікні, що з'явилося, вказати *Dependent* (залежну) і *Independent* (незалежну) змінні для побудови багатофакторної регресійної моделі. Вибір змінних поданий на рис. 10.3.2. Після вказівки змінних підтвердити свій вибір натисканням кнопки *OK.*

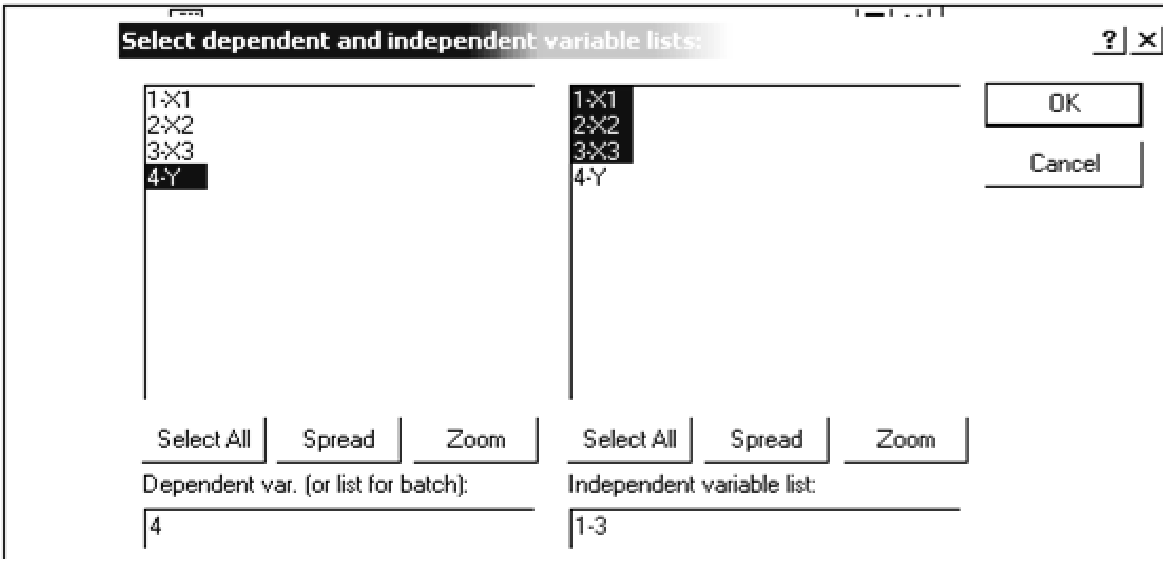


Рис. 10.3.2. **Вибір змінних для аналізу**

Побудувати лінійну багатофакторну економетричну модель і визначити всі її характеристики. Перевірити статистичну значущість параметрів моделі та адекватність моделі за критерієм Фішера. Результати побудови моделі наведені на рис. 10.3.3.

	Regression Summary for Dependent Variable: Y (Spr38) R= ,99329417 R?= ,98663330 Adjusted R?= ,98472378 F(3,21)=516,69 p<0,0000 Std.Error of estimate: ,61913					
N=25	Beta	Std.Err. of Beta	B	Std.Err. of B	t(21)	p-level
Intercept			-0,159728	0,165543	-0,964874	0,345589
X1	0,747946	0,119810	0,293843	0,047069	6,242776	0,000003
X2	0,253338	0,119732	0,396141	0,187223	2,115880	0,046472
X3	0,036745	0,025417	0,175450	0,121359	1,445707	0,163019

Рис. 10.3.3. **Результати багатофакторного регресійного аналізу**

Для перевірки гіпотези про значущість регресійної моделі використовується дисперсійний аналіз. Для цього необхідно ініціювати кнопку *Advanced /ANOVA*. Результати дисперсійного аналізу для досліджуваної моделі наведені на рис. 10.3.4. У даній таблиці наведено суму квадратів відхилень за регресією *(Sums of Squares Regress),* суму квадратів відхилень похибок *(Sums of Squares Residual),* дисперсію похибок *(Mean Squares Residual)* та критерій Фішера *(F)*.

Effect	Analysis of Variance; DV: Y (Spreadsheet17) Sums of Squares	df	Mean Squares	F	p-level
Regress.	591,6353	1	591,6353	1283,945	0,000000
Residual	10,5983	23	0,4608		
Total	602,2336				

Рис. 10.3.4. **Таблиця дисперсійного аналізу**

3. Перевірка моделі на наявність мультиколінеарності.

Наступним кроком дослідження багатофакторної регресійної моделі є перевірка наявності мультиколінеарності в моделі.

Одним зі способів перевірки моделі на мультиколінеарність є розрахунок матриці парних кореляцій. У меню аналізу моделі (рис. 10.3.5) ініціюйте кнопку *Descriptive statistics/Correlations* (*Описові статистики/кореляція*). Матриця коефіцієнтів парних кореляцій наведена на рис. 10.3.6.

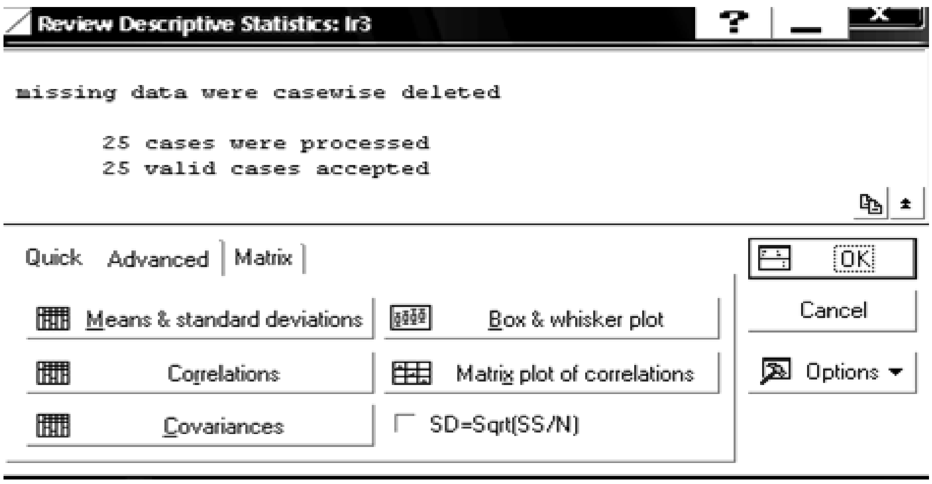


Рис. 10.3.5. **Вікно вибору оцінки кореляційної залежності**

Variable	Correlations (Spr38) X1	X2	X3	Y
X1	1,000000	0,977544	-0,120649	0,991162
X2	0,977544	1,000000	-0,115205	0,980255
X3	-0,120649	-0,115205	1,000000	-0,082679
Y	0,991162	0,980255	-0,082679	1,000000

Рис. 10.3.6. **Матриця коефіцієнтів парних кореляцій**

Ініціювавши опцію *Matrix plot of correlations (кореляційні графіки),* буде отримано гістограми і діаграми розсіювання досліджуваних змінних в моделі (рис. 10.3.7).

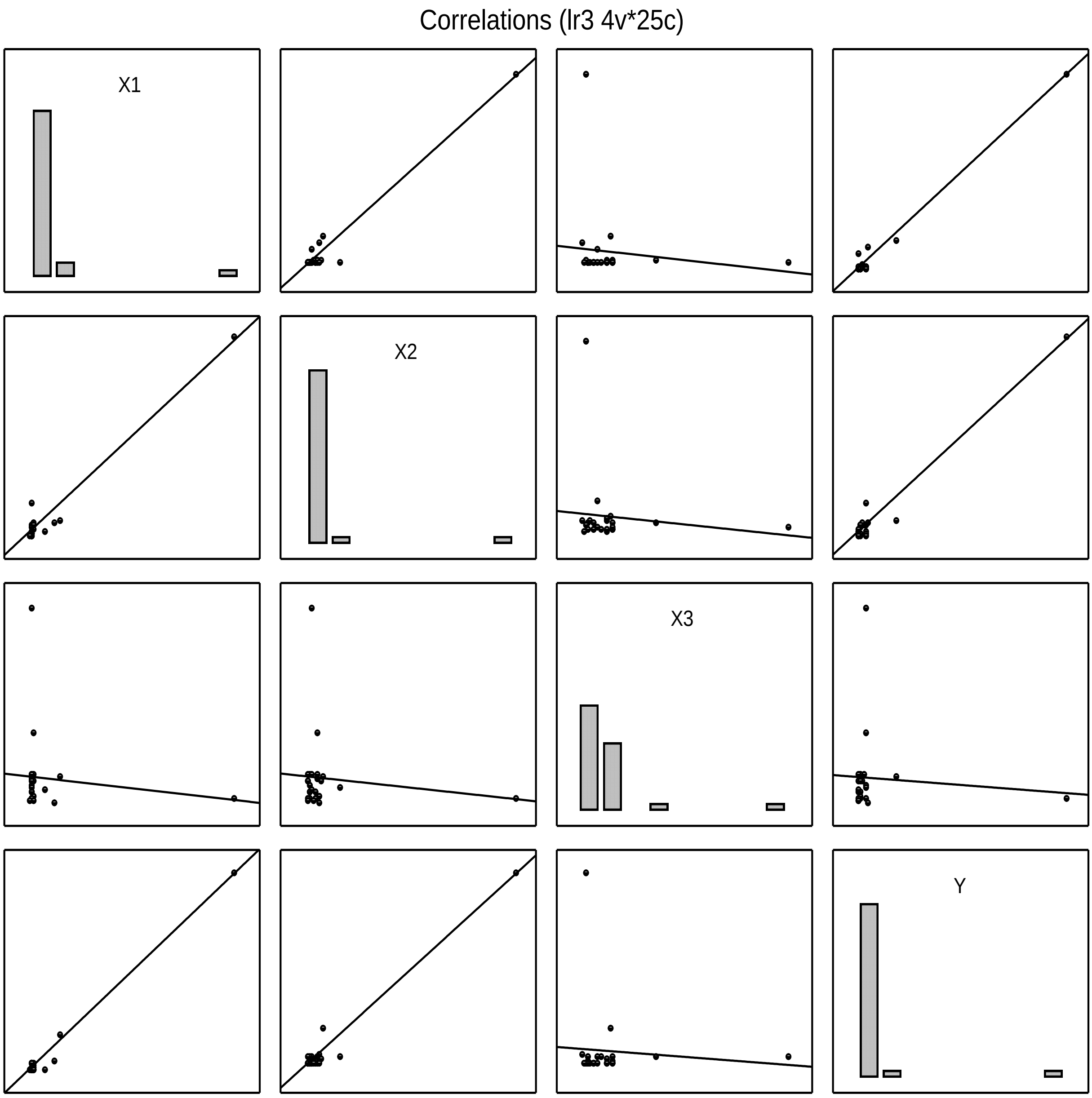


Рис. 10.3.7. **Гістограми і діаграми розсіювання досліджуваних змінних**

Коефіцієнти парних кореляцій та діаграми розсіювання свідчать про сильний ступінь лінійного взаємозв'язку між досліджуваними парами змінних.

Більш детальний аналіз передбачає розрахунок частинних кореляцій та розрахунок критерію максимальної спряженості для кожної з незалежних змінних. Ініціювавши опцію *Partial correlations (частинні кореляції)* (рис. 10.3.8), буде отримана така таблиця (рис. 10.3.9). Дана таблиця містить значення стандартизованих коефіцієнтів регресії (*betain)*; частинних кореляцій *(partial correlations)*, які відображають ступінь впливу кожної незалежної змінної на результуючу, за умови, що інші змінні не впливають на даний зв'язок; півчастинних кореляцій *(Semipartial correlations)*; коефіцієнт детермінації *(Rsquare)* між даною змінною та іншими незалежними змінними, які входять у рівняння регресії, що відображує міру максимальної спряженості; толерантність моделі *(tolerance),* яка розраховується як *(1 – Rsquare);* значення критерію Стьюдента *(t)* для перевірки гіпотези про значущість частинних коефіцієнтів кореляції з числом ступенів волі; рівень значущості *(p-level); і*мовірність відхилення гіпотези про значущість частинних коефіцієнтів кореляції.

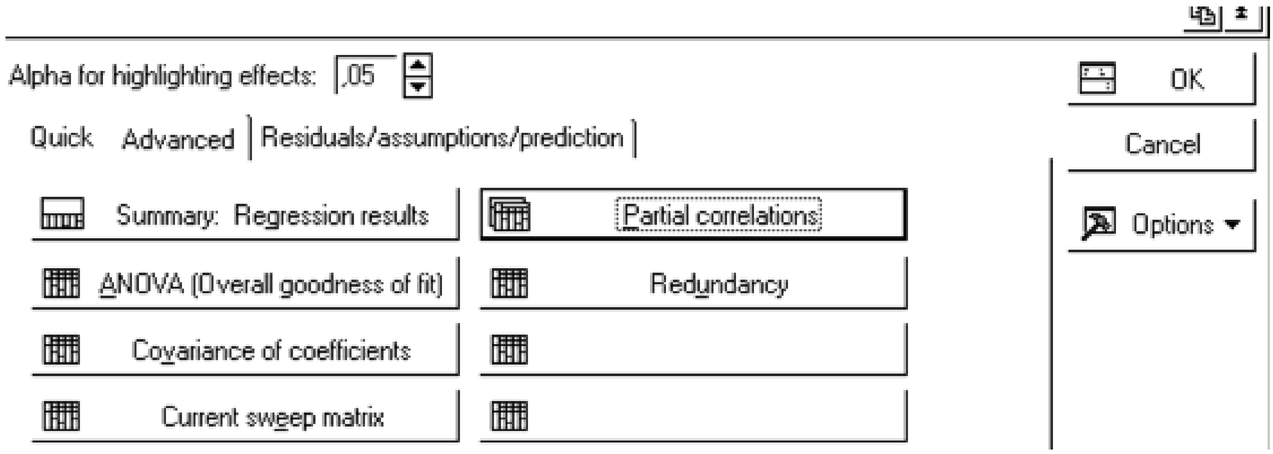


Рис. 10.3.8. **Вікно вибору оцінки значущості незалежних змінних**

Variable	Variables currently in the Equation; DV: Y (Spreadsheet17)						
	Beta in	Partial Cor.	Semipart Cor.	Tolerance	R-square	t(21)	p-level
X1	0,747946	0,806125	0,157500	0,044342	0,955658	6,242776	0,000003
X2	0,253338	0,419196	0,053382	0,044400	0,955600	2,115880	0,046472
X3	0,036745	0,300862	0,036474	0,985276	0,014724	1,445707	0,163019

Рис. 10.3.9. **Значущість змінних у рівнянні регресії**

Оцінювання ступеня впливу незалежних змінних на результуючий показник можливо виконати натисканням клавіші *Redundancy (збитковість)*, результати якої наведено на рис. 10.3.10.

Variable	Redundancy of Independent Variables; DV: R-square column contains R-square of resp variable with all other independent variables			
	Toleran.	R-square	Partial Cor.	Semipart Cor.
X1	0,044342	0,955658	0,806125	0,157500
X2	0,044400	0,955600	0,419196	0,053382
X3	0,985276	0,014724	0,300862	0,036474

Рис. 10.3.10. **Оцінювання надміру незалежних змінних**

Для оцінювання ступеня мультиколінеарності за алгоритмом Фаррара – Глобера використовуються частинні коефіцієнти кореляції між факторними змінними та їх статистична значущість, для їх розрахунку необхідно дослідити модель без залежної змінної, зробивши одну з факторних змінних залежною, та визначити дані характеристики (рис. 10.3.11).

Variable	Variables currently in the Equation; DV: X1 (lr3)						
	Beta in	Partial Cor.	Semipart Cor.	Tolerance	R-square	t(22)	p-level
X2	0,976606	0,977242	0,970104	0,986728	0,013272	21,60825	0,000000
X3	-0,008139	-0,038366	-0,008085	0,986728	0,013272	-0,18008	0,858735

Variable	Variables currently in the Equation; DV: X2 (lr3)						
	Beta in	Partial Cor.	Semipart Cor.	Tolerance	R-square	t(22)	p-level
X1	0,977879	0,977242	0,970736	0,985444	0,014556	21,60825	0,000000
X3	0,002775	0,013073	0,002755	0,985444	0,014556	0,06132	0,951656

Рис. 10.3.11. **Оцінювання рівня зв'язку між незалежними змінними**

Таким чином, значення коефіцієнтів частинних кореляцій дорівнює: $r_{12} = 0{,}977242,\ r_{13} = -0{,}038366,\ r_{23} = 0{,}013073$. Значущість часткових коефіцієнтів кореляції за допомогою t-критерію Стьюдента: $t_{12} = 21{,}60825$, $t_{13} = -0{,}018008, t_{23} = 0{,}06132$. Значення критеріїв t_{ij}порівняти з табличним за умови $(n - m)$ ступенях вільності і рівні значущості α. Якщо $t_{ij} > t_{\text{табл}}$, то між незалежними змінними x_i та x_j існує мультиколінеарність. $t_{\text{табл}}\ 0{,}05; 22\ = 2{,}07$, отже $t_{12} > t_{\text{табл}}$, то можна зробити висновок, що між змінними x_1 та x_2 існує тісний лінійний зв'язок (мультиколінеарність).

4. Розрахунок та дослідження помилок моделі.

Для проведення даного аналізу необхідно розрахувати теоретичні значення залежної змінної й помилки моделі, навести результати дослідження моделі за критерієм Дарбіна – Уотсона й нециклічного коефіцієнта автокореляції та зробити висновки про наявність автокореляції.

Щоб розрахувати й проаналізувати залишки, у нижній частині вікна результатів регресійного аналізу є опція *Perform residual analysis* (всебічний аналіз залишків). Ініціювавши дану опцію, буде одержано меню для аналізу помилок моделі (див. лаб. роб. № 2).

Ініціювавши клавішу аналізу помилок *Summary: Residuals & Predicted,* буде одержано таблицю зі спостережуваними значеннями залежної змінної *(observed value)*, теоретичними значеннями залежної змінної *(predicted value)* і помилками моделі*(residual)* та графіком розсіювання помилок моделі в діапазоні $\pm 3\sigma$, який аналізує властивість сталості дисперсії помилок (рис. 10.3.12).

	Raw Residuals	Raw Residual (lr3) Dependent variable: Y		
Case	-3s . . 0 . . +3s	Observed Value	Predicted Value	Residual
1	. . . *	0,13000	0,42236	-0,29236
2	. . . . * . . .	0,77000	0,54082	0,22918
3	. . . * . . .	0,25000	0,26496	-0,01496
4	. . . *	0,50000	0,71138	-0,21138
5	. . . *. . . .	0,12000	0,27994	-0,15994
6	. . . .* . . .	0,22000	0,11230	0,10770
7	. . . *	0,82000	1,07801	-0,25801
8	. . . . .* . .	0,91000	0,16130	0,74870
9	. . . . * . . .	0,96000	0,75716	0,20284
10	. . . *	0,16000	0,44403	-0,28403
11	. . . * . . .	0,11000	0,05796	0,05204
12	. . . * . . .	0,02000	0,03789	-0,01789
13	. . . * . . .	0,01000	-0,01531	0,02531
14	. . . .* . . .	0,04300	-0,01846	0,06146
15	. . . . . * . .	0,96900	-0,08485	1,05385
16	. . . .* . . .	0,17200	0,01131	0,16069
17	. . . .* . . .	0,10800	0,01610	0,09190
18	. . . *. . . .	0,06400	0,18189	-0,11789
19	. . . . . . * .	4,65400	3,15785	1,49615
20	. *.	0,07600	1,44400	-1,36800
21	. . . * . . .	25,15200	25,11234	0,03966
22	. *	1,08700	2,39243	-1,30543
23	. . . *	0,02400	0,29160	-0,26760
24	. . . *. . . .	0,91400	1,01625	-0,10224
25	. . . .* . . .	0,00100	-0,12924	0,13024
Minimum	. *.	0,00100	-0,12924	-1,36800
Maximum	. . . . . . * .	25,15200	25,11234	1,49615
Mean	. . . * . . .	1,52976	1,52976	0,00000
Median	. . . * . . .	0,17200	0,27994	0,02531

Рис. 10.3.12. **Аналіз помилок моделі та графік їх розсіювання**

У меню аналізу помилок, ініціювавши кнопку *Advanced*/*Durbin-Watson statistic* (статистика Дарбіна – Уотсона), буде одержано значення автокореляції помилок моделі за критерієм Дарбіна – Уотсона та значення нециклічного коефіцієнта автокореляції (рис. 10.3.13).

	Durbin-Watson d (Spr3 and serial correlation of	
	Durbin-Watson d	Serial Corr.
Estimate	2,464231	-0,238982

Рис. 10.3.13. **Автокореляція помилок моделі**

Значення отриманих коефіцієнтів порівнюються з табличними значеннями і робляться висновки про наявність автокореляції. Для подальшого всебічного аналізу помилок необхідно побудувати гістограму й графік розподілу помилок (рис. 10.3.14, 10.3.15).

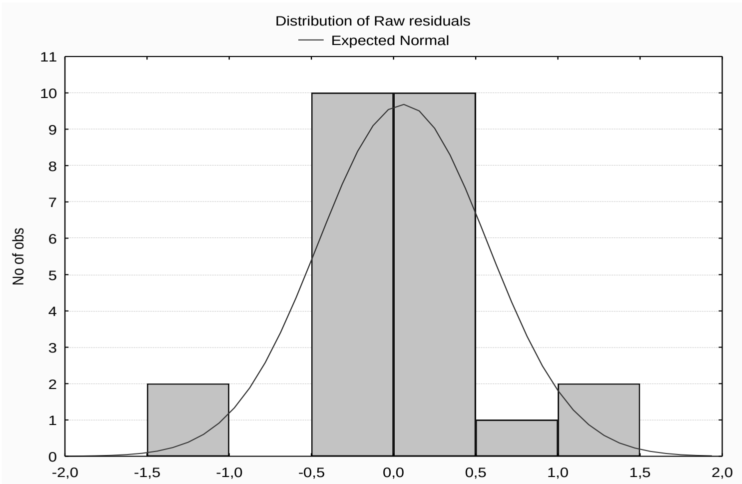


Рис. 10.3.14. **Гістограма розподілу помилок**

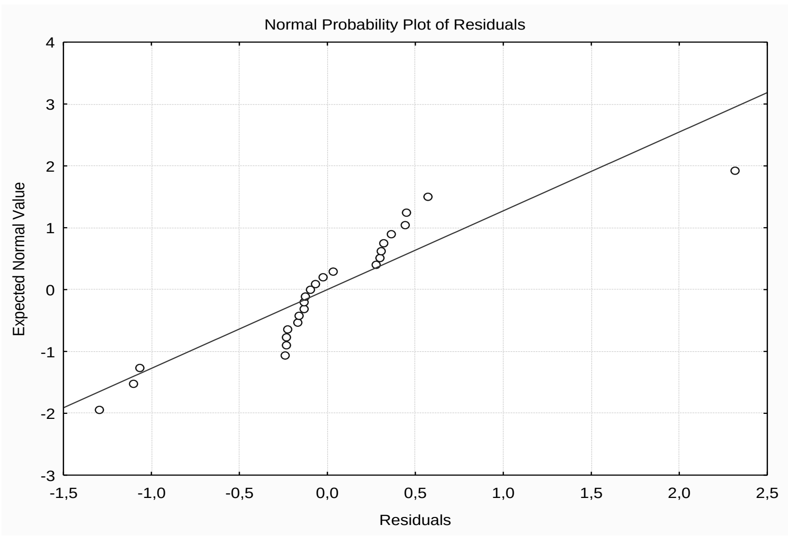


Рис. 10.3.15. **Графік розподілу помилок на нормальному ймовірнісному папері**

За отриманими результатами необхідно згрупувати об'єкти за додатними та від'ємними значеннями помилок, дати економічну інтерпретацію даному групуванню.

5. Побудова моделей за методами покрокового включення й покрокового виключення незалежних змінних.

В умовах мультиколінеарності незалежних змінних ефективним методом оцінювання параметрів економетричних моделей є реалізація покрокової регресії, яка передбачає оцінювання параметрів моделі через коефіцієнти кореляцій.

У модулі *Multiple Regression* реалізований метод покрокового включення змінних *(forward stepwise)* і метод покрокового виключення *(backward stepwise).* Вибір методів здійснюється на стартовій панелі у меню *Advanced* ініціюванням опції *Advanced options (stepwise or ridge regression)* (рис. 10.3.16).

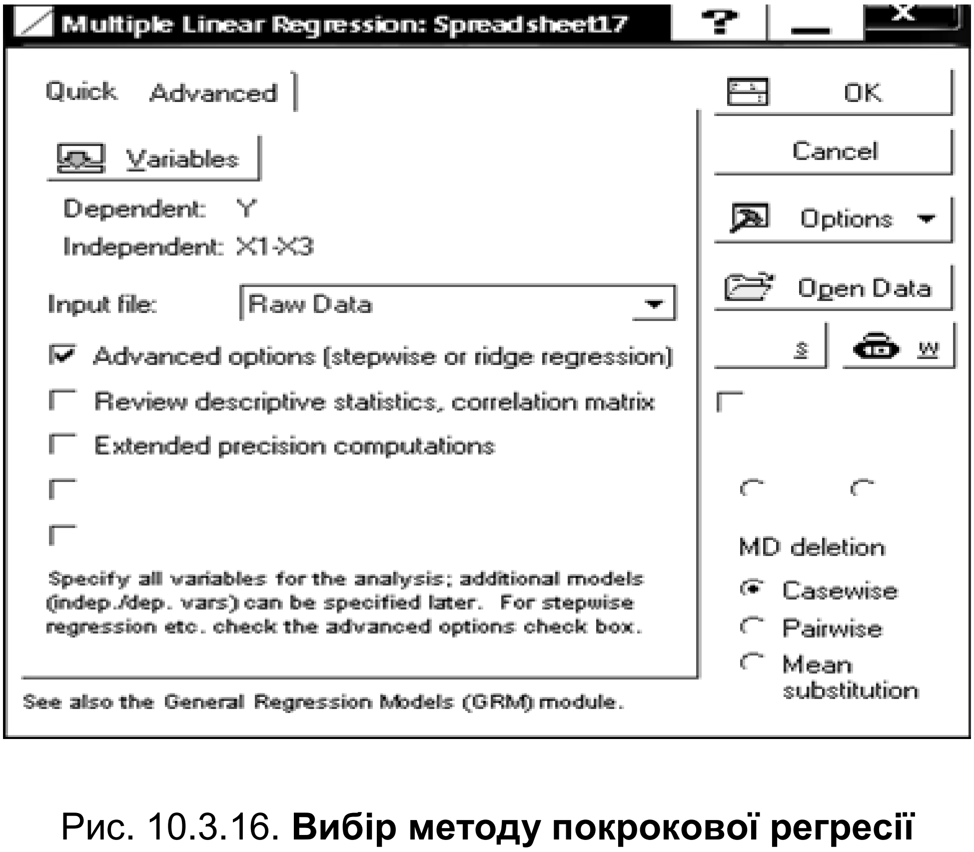


Рис. 10.3.16. **Вибір методу покрокової регресії**

Вибір методу оцінювання, порогові значення F-критерію включення або виключення, послідовність подання результатів вибирається у вкладці *Stepwise* (рис. 10.3.17).

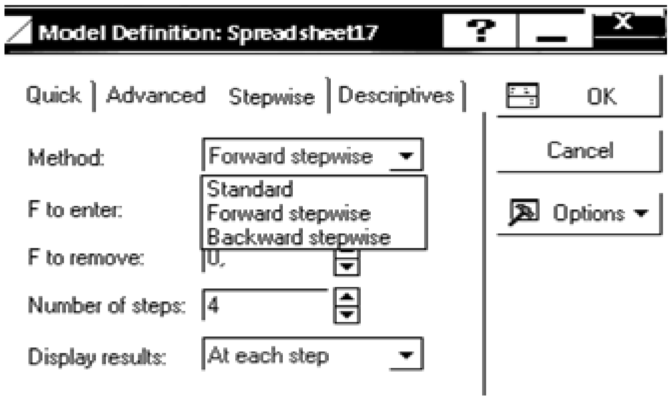


Рис. 10.3.17. **Вибір методів оцінювання параметрів**

Послідовність етапів реалізації алгоритму покрокового включення *(Forward stepwise)* подано на рис. 10.3.18.

```
MultipleRegressionResults (Step   0)
Dependent: Y              Multiple R = 0,00000000     F = 0,000000
                           R?= 0,00000000    df =   0,24
  No. of cases: 25          adjusted R?= 0,00000000     p = -0,00000
              Standard error of estimate: 5,009297816
       Step 0:  No variables in the regression equation

   Multiple Regression Results (Step   1)
  Dependent: Y              Multiple R =  ,99116180     F = 1283,945
                           R?=  ,98240171    df =   1,23
  No. of cases: 25          adjusted R?=  ,98163656     p = 0,000000
              Standard error of estimate:  ,678818769
  Intercept:   ,081243426  Std.Error: ,1416544  t(   23) = ,57353  p =  ,5719
            X1 beta=,991

   Multiple Regression Results (Step   2)
  Dependent: Y              Multiple R =  ,99262428     F = 737,4498
                           R?=  ,98530296    df =   2,22
  No. of cases: 25          adjusted R?=  ,98396686     p = 0,000000
              Standard error of estimate:  ,634287475
  Intercept:  -,030737796  Std.Error: ,1428533  t(   22) = -,2152  p =  ,8316
            X1 beta=,741           X2 beta=,256

   Multiple Regression Results (step   3, final solution)
   no other F to enter exceeds specified limit
  Dependent: Y              Multiple R =  ,99329417     F = 516,6896
                           R?=  ,98663330    df =   3,21
  No. of cases: 25          adjusted R?=  ,98472378     p = 0,000000
              Standard error of estimate:  ,619134289
  Intercept:  -,159727778  Std.Error: ,1655426  t(   21) = -,9649  p =  ,3456
            X1 beta=,748           X2 beta=,253           X3 beta=,037
```

Рис. 10.3.18. **Реалізація моделі покрокового включення змінних**

Після закінчення процедури покрокової регресії в нижній інформаційній частині вікна стане активною опція *Stepwise regression summary (Результати покрокової регресії)* (рис. 10.3.19).

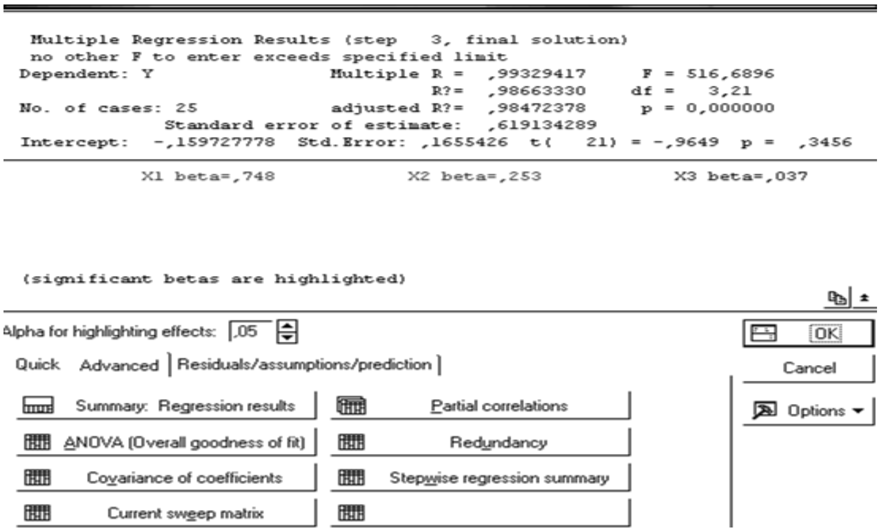


Рис. 10.3.19. **Вікно вибору результатів покрокової регресії**

Натиснувши дану опцію, буде отримано таблицю результатів покрокової регресії, наведену на рис. 10.3.20, де відображено адекватність моделі на кожному з етапів та зміну характеристик моделі для кожного кроку.

	Summary of Stepwise Regression; DV: Y (lr3)						
Variable	Step +in/-out	Multiple R	Multiple R-square	R-square change	F - to entr/rem	p-level	Variables included
X1	1	0,991162	0,982402	0,982402	1283,945	0,000000	1
X2	2	0,992624	0,985303	0,002901	4,343	0,048989	2
X3	3	0,993294	0,986633	0,001330	2,090	0,163019	3

Рис. 10.3.20. **Результати покрокової регресії включення**

Аналіз результатів регресійної моделі методом покрокового включення змінних моделі (адекватність, статистична значущість) отримано ініціюванням опції *Summary: Regression results,* яку наведено на рис. 10.3.21.

	Regression Summary for Dependent Variable: Y (Spreadsheet17) R= ,99329417 R?= ,98663330 Adjusted R?= ,98472378 F(3,21)=516,69 p<0,0000 Std.Error of estimate: ,61913					
N=25	Beta	Std.Err. of Beta	B	Std.Err. of B	t(21)	p-level
Intercept			-0,159728	0,165543	-0,964874	0,345589
X1	0,747946	0,119810	0,293843	0,047069	6,242776	0,000003
X2	0,253338	0,119732	0,396141	0,187223	2,115880	0,046472
X3	0,036745	0,025417	0,175450	0,121359	1,445707	0,163019

Рис. 10.3.21. **Регресійна модель методом покрокового включення**

Послідовність етапів реалізації алгоритму покрокового виключення *(Backward stepwise)* наведено на рис. 10.3.22.

```
MultipleRegressionResults (Step   0)
Dependent: Y               Multiple R =  ,99329417     F = 516,6896
                            R?=  ,98663330    df =   3,21
  No. of cases: 25           adjusted R?=  ,98472378     p = 0,000000
            Standard error of estimate:  ,619134289
  Intercept:  -,159727778  Std.Error: ,1655426  t(   21) = -,9649  p =  ,3456
           X1 beta=,748           X2 beta=,253           X3 beta=,037

 Multiple Regression Results (Step   1)
  Dependent: Y               Multiple R =  ,99262428     F = 737,4498
                            R?=  ,98530296    df =   2,22
  No. of cases: 25           adjusted R?=  ,98396686     p = 0,000000
            Standard error of estimate:  ,634287475
  Intercept:  -,030737796  Std.Error: ,1428533  t(   22) = -,2152  p =  ,8316
           X1 beta=,741           X2 beta=,256

 Multiple Regression Results (step   2, final solution)
 no other F to remove is less than specified limit
  Dependent: Y               Multiple R =  ,99116180     F = 1283,945
                            R?=  ,98240171    df =   1,23
  No. of cases: 25           adjusted R?=  ,98163656     p = 0,000000
            Standard error of estimate:  ,678818769
  Intercept:   ,081243426  Std.Error: ,1416544  t(   23) = ,57353  p =  ,5719
                X1 beta=,991
```

Рис. 10.3.22. **Реалізація моделі покрокового виключення змінних**

Натиснувши опцію *Stepwise regression summary,* буде отримана таблиця результатів покрокової регресії виключення для кожного етапу, яка подана на рис. 10.3.23.

	Summary of Stepwise Regression; DV: Y (lr3)						
Variable	Step +in/-out	Multiple R	Multiple R-square	R-square change	F - to entr/rem	p-level	Variables included
X3	-1	0,992624	0,985303	-0,001330	2,090069	0,163019	2
X2	-2	0,991162	0,982402	-0,002901	4,342880	0,048989	1

Рис. 10.3.23. **Результати покрокової регресії виключення**

Регресійна модель методом покрокового виключення змінних подана на рис. 10.3.24.

	Regression Summary for Dependent Variable: Y (Spreadsheet1 R= ,99116180 R?= ,98240171 Adjusted R?= ,98163656 F(1,23)=1283,9 p<0,0000 Std.Error of estimate: ,67882					
N=25	Beta	Std.Err. of Beta	B	Std.Err. of B	t(23)	p-level
Intercept			0,081243	0,141654	0,57353	0,571850
X1	0,991162	0,027661	0,389395	0,010867	35,83218	0,000000

Рис. 10.3.24. **Регресійна модель методом покрокового виключення**

Аналіз побудованих моделей дозволить вибрати найбільш адекватну модель із точки зору опису реальних економічних процесів, їх взаємозв'язків та цілей вирішення задач.

6. Побудова моделі на основі методу рідж-регресії звільнення від мультиколінеарності.

Одним із методів, який дозволяє скорегувати матрицю незалежних змінних у разі наявності в моделі мультиколінеарності, є метод рідж-регресії.

У модулі *Multiple Regression* реалізацію даного методу можна здійснити на стартовій панелі у меню *Advanced* ініціюванням опції *Advanced options (stepwise or ridge regression)* (див. рис. 10.3.16).

Для вибору методу рідж-регресії і параметра зсуву λ у меню *Advanced* необхідно ініціювати опцію *Ridge regression; lambda* (рис. 10.3.25).

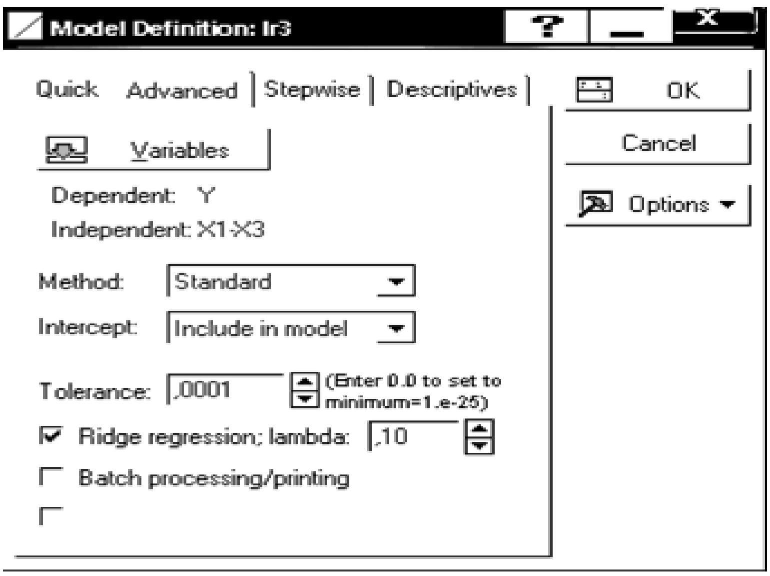


Рис. 10.3.25. **Вибір методу рідж-регресії і параметра зсуву λ**

Результати регресійних моделей методом рідж-регресії (адекватність, статистична значущість) з різними значеннями параметра зсуву λ будуть отримані ініціюванням опції *Summary: Regression results,* що наведено на рис. 10.3.26, 10.3.27.

	Ridge Regression Summary for Dependent Variable: Y (lr3) l=,10000 R= ,96780176 R?= ,93664024 Adjusted R?= ,92758884 F(3,21)=103,48 p<,00000 Std.Error of estimate: 1,3480					
N=25	Beta	Std.Err. of Beta	B	Std.Err. of B	t(21)	p-level
Intercept			-0,127295	0,342878	-0,371254	0,714167
X1	0,521135	0,114292	0,204737	0,044902	4,559671	0,000171
X2	0,430861	0,114231	0,673730	0,178621	3,771846	0,001120
X3	0,027120	0,052697	0,129493	0,251613	0,514651	0,612171

Рис. 10.3.26. **Рідж-регресія за умови $\lambda = 0,1$**

	Ridge Regression Summary for Dependent Variable: Y (lr3) l=,90000 R= ,82181444 R?= ,67537898 Adjusted R?= ,62900454 F(3,21)=14,564 p<,00002 Std.Error of estimate: 3,0511					
N=25	Beta	Std.Err. of Beta	B	Std.Err. of B	t(21)	p-level
Intercept			0,395520	0,709551	0,557422	0,583133
X1	0,348421	0,105264	0,136883	0,041355	3,309961	0,003331
X2	0,336603	0,105246	0,526340	0,164571	3,198262	0,004321
X3	-0,000981	0,090430	-0,004685	0,431777	-0,010850	0,991446

Рис. 10.3.27. **Рідж-регресія за умови $\lambda = 0,9$**

Дослідження ступеня зміни параметрів моделі з різними значеннями параметра зсуву λ наведено в табл. 10.3.1.

Таблиця 10.3.1

Параметри моделі рідж-регресії

λ	0	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	0,99
a0	-0,160	-0,13	-0,039	0,043	0,117	0,183	0,244	0,299	0,349	0,396	0,434
a1	0,294	0,205	0,188	0,177	0,168	0,161	0,154	0,148	0,142	0,137	0,133
a2	0,396	0,674	0,671	0,651	0,629	0,606	0,585	0,564	0,545	0,526	0,511
a3	0,175	0,129	0,097	0,072	0,052	0,036	0,023	0,012	0,003	-0,005	-0,010

Графік гребневого сліду оцінок параметрів, який відображує ступінь зміщення оцінок регресійної моделі, наведено на рис. 10.3.28.

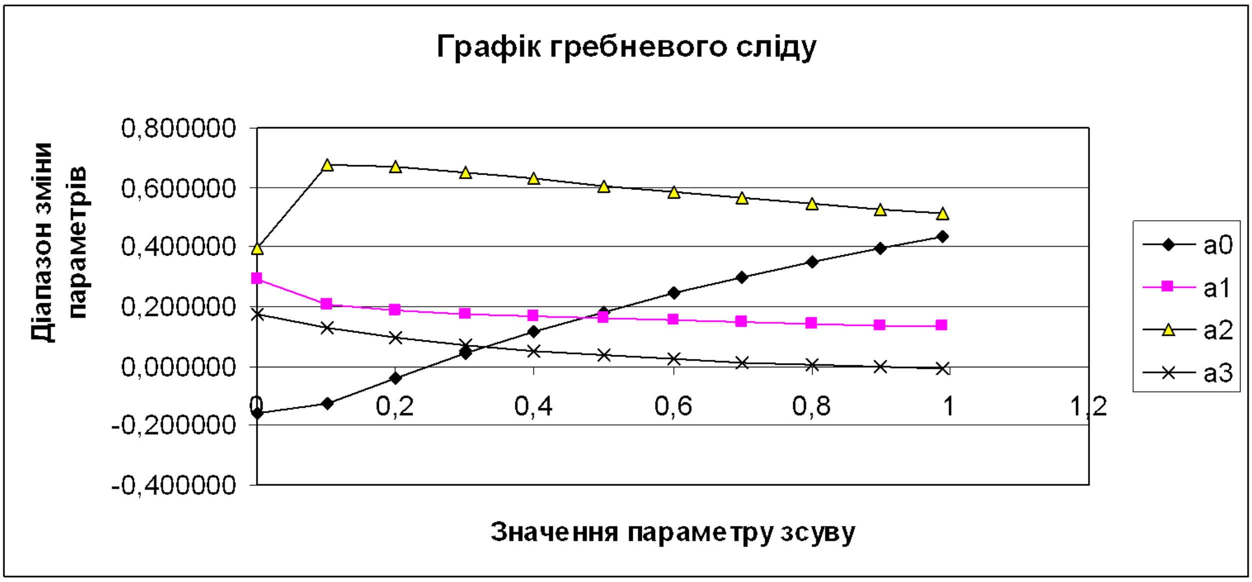


Рис. 10.3.28. **Графік гребеневого сліду оцінок параметрів**

7. Прогнозування.

Оскільки побудовані моделі є адекватними, їх параметри статистично значущі, то за моделлю можна скласти прогнози. Щоб розрахувати прогнозні значення залежної змінної, у нижній частині вікна результатів регресійного аналізу є опція *Predict dependent variable* (прогнозування залежної змінної), ініціювавши дану опцію необхідно вказати значення незалежних змінних, для яких необхідно спрогнозувати залежну величину (див. лаб. роб. 2).

Результати прогнозування для трифакторної та однофакторної економетричних моделей наведені на рис. 10.3.29, 10.3.30.

Прогнозне значення залежної змінної (*predicted*) для трифакторної моделі = 1,5315; і довірчі інтервали для прогнозного значення:

$$1{,}2740 \leq 1{,}5315 \leq 1{,}7891.$$

	Predicting Values for (lr3) variable: Y		
Variable	B-Weight	Value	B-Weight * Value
X1	0,293843	3,720000	1,093096
X2	0,396141	1,200000	0,475369
X3	0,175450	0,700000	0,122815
Intercept			-0,159728
Predicted			1,531552
-95,0%CL			1,274027
+95,0%CL			1,789077

Рис. 10.3.29. **Результати прогнозу для трифакторної моделі**

Variable	Predicting Values for (lr3) variable: Y		
	B-Weight	Value	B-Weight * Value
X1	0,389395	3,720000	1,448548
Intercept			0,081243
Predicted			1,529791
-95,0%CL			1,248942
+95,0%CL			1,810640

Рис. 10.3.30. **Результати прогнозу для однофакторної моделі**

Прогнозне значення залежної змінної (*predicted*) для однофакторної моделі = 1,5298; і довірчі інтервали для прогнозного значення:

$$1{,}2489 \leq 1{,}5298 \leq 1{,}8106.$$

Таким чином, можна зробити висновки, що прогнози за моделями майже не відрізняються, однак довірчі інтервали зміни залежної змінної для однофакторної моделі є більш широкими.

Лабораторна робота 4. Побудова й аналіз множинної нелінійної регресії Кобба – Дугласа

Мета – закріплення теоретичного і практичного матеріалу за темою "Нелінійна регресія", набуття навичок побудови й аналізу нелінійних виробничих функцій у модулі *Nonlinear Estimation*.

Завдання – необхідно перевірити наявність лінійного та нелінійного зв'язків між обсягом виробництва продукції та наявними виробничими ресурсами у *модулях Multiple Regressioni Nonlinear Estimation* ППП *Statistica:*

1. Побудувати лінійну багатофакторну модель. Визначити всі її характеристики (знайти параметри моделі за допомогою МНК, середні квадратичні відхилення параметрів моделі, дисперсію та середнє квадратичне відхилення похибок моделі, коефіцієнти множинної кореляції і детермінації).

2. Перевірити значущість економетричної моделі за допомогою критерію Фішера.

3. Перевірити наявність автокореляції залишків за допомогою критерію Дарбіна – Уотсона та коефіцієнта нециклічної автокореляції. Навести гістограму і графік розподілу похибок. Зробити висновки щодо наявності автокореляції.

4. Зробити висновки щодо адекватності лінійної багатофакторної економетричної моделі.

5. Перевірити існування нелінійного зв'язку між обсягом виробництва і величиною виробничих ресурсів шляхом побудови виробничої функції Кобба – Дугласа.

6. Навести гістограму і графік розподілу похибок. Зробити висновки щодо наявності автокореляції похибок.

7. Зробити висновки щодо адекватності нелінійної економетричної моделі.

8. Визначити характеристики виробничої функції. Знайти можливі комбінації виробничих ресурсів за умови фіксованих рівнів виробництва. Навести графіки ізоквант.

Як розглядувалося раніше, для побудови й всебічного аналізу множинних лінійних економетричних моделей у ППП *Statistica* передбачений модуль *Multiple Regression* (*множинна регресія*), для побудови нелінійних економетричних моделей передбачено модуль *Nonlinear Estimation* (*нелінійне оцінювання*).

1. Запуск *Statistica* і підготовка даних.

У меню програм вибрати програму *Statistica*, після її запуску вибрати у меню пункт *File/New* для підготовки власних даних. Після уведення натиснути кнопку вікна *OK.* Після заповнення всіх комірок поля даних буде одержано таблицю, аналогічну поданій на рис. 10.4.1, де *L* – чисельність робочої сили (тис. осіб), *K* – основні фонди (тис. грн), *Y* – обсяг виробленої продукції (млн грн).

	1 L	2 K	3 Y
1	0,14	0,16	0,11
2	0,09	0,21	0,22
3	0,3	0,26	0,3
4	0,38	0,18	0,56
5	5,07	3,8	7,71
6	6,14	2,19	3,5
7	1,53	1,13	6,78
8	0,71	1,42	1,44
9	0,7	0,63	0,84
10	0,55	0,45	0,76
11	0,47	0,3	2,03
12	0,27	0,25	0,81
13	0,14	0,16	0,11
14	0,09	0,21	0,22
15	0,3	0,26	0,3
16	0,38	0,18	0,56
17	0,87	1	1
18	3,79	0,65	0,85
19	0,53	0,77	0,56
20	0,05	0,55	0,17
21	0,121	0,165	0,109
22	0,271	0,092	0,116
23	0,189	0,197	0,192
24	0,026	0,108	0,173
25	0,142	0,212	0,028

Рис. 10.4.1. **Вихідні дані для побудови виробничої функції**

Таким чином, досліджується класична виробнича функція між факторами виробництва (дані про чисельність робітників, вартість основних фондів підприємства) та обсягом виробленої продукції.

Слід розглянути порядок роботи під час побудови багатофакторних виробничих функцій.

2. Побудова лінійної багатофакторної виробничої функції.

Побудову й аналіз багатофакторної лінійної виробничої функції необхідно проводити в модулі *Multiple Regression* так, як це було розглянуто в попередніх лабораторних роботах. Характеристики лінійної виробничої функції наведені на рис. 10.4.2.

	Regression Summary for Dependent Variable: Y (Spreadsheet9) R= ,83173856 R?= ,69178903 Adjusted R?= ,66376985 F(2,22)=24,690 p<,00000 Std.Error of estimate: 1,1488					
N=25	Beta	Std.Err. of Beta	B	Std.Err. of B	t(22)	p-level
Intercept			-0,060536	0,289925	-0,208798	0,836527
L	-0,007511	0,207505	-0,009274	0,256206	-0,036198	0,971451
K	0,837897	0,207505	2,007017	0,497037	4,037966	0,000550

Рис. 10.4.2. **Модель лінійної виробничої функції**

Таким чином, лінійна виробнича функція буде мати наступний вигляд: $\hat{Y} = -0{,}06054 - 0{,}009274 \cdot L + 2{,}007 \cdot K$.

Для побудови графіка залежності факторів виробництва та обсягу виробленої продукції необхідно в меню *Graphs/2D Graphs/Scaterplots* вибрати змінні та задати параметри графіка (рис. 10.4.3) й побудувати графік (рис. 10.4.4).

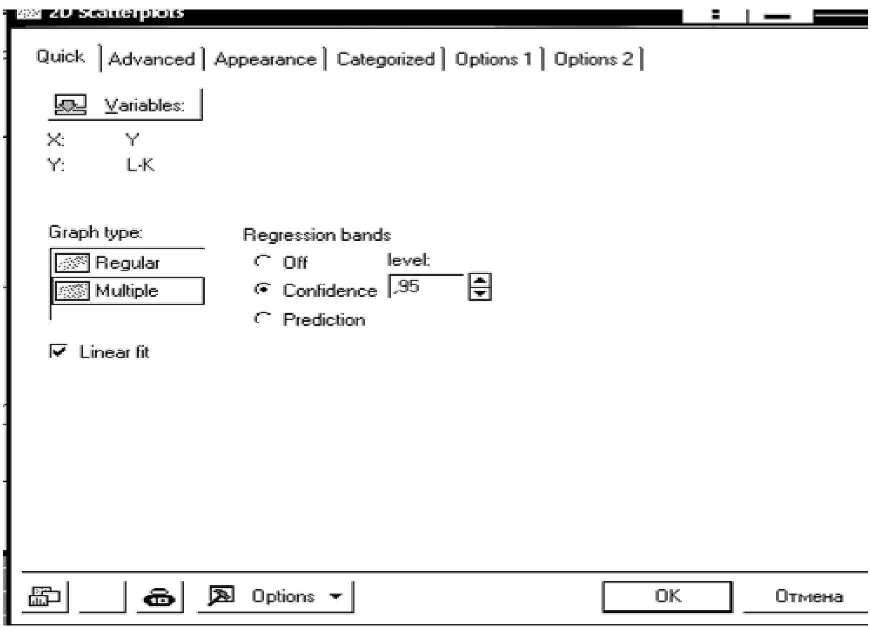


Рис. 10.4.3. **Задавання параметрів графіка**

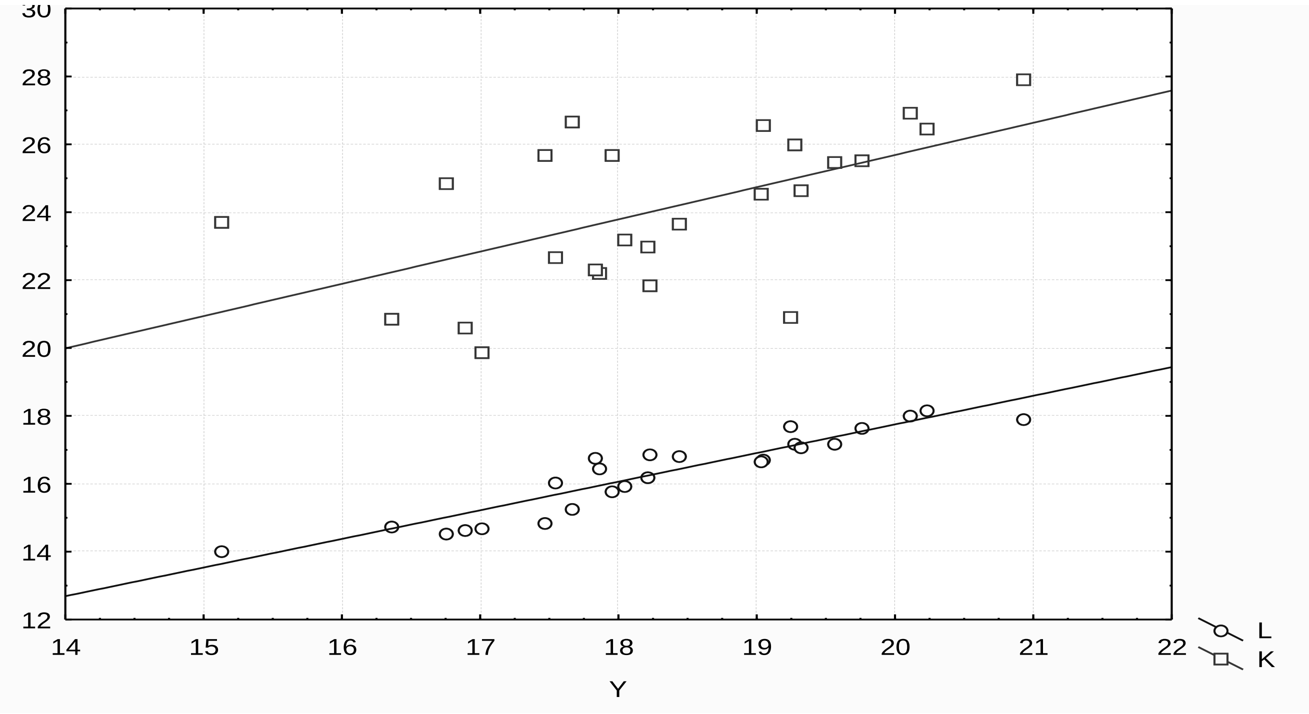


Рис. 10.4.4. **Графік залежності факторів виробництва**

На рис. 10.4.5 наведені результати дисперсійного аналізу лінійної виробничої функції.

	Analysis of Variance; DV: Y (Spreadsheet15)				
Effect	Sums of Squares	df	Mean Squares	F	p-level
Regress.	40,20125	2	20,10063	121,7467	0,000000
Residual	3,63224	22	0,16510		
Total	43,83350				

Рис. 10.4.5. **Таблиця дисперсійного аналізу**

Усебічний аналіз лінійної багатофакторної функції щодо її адекватності необхідно провести так, як це було розглянуто в лабораторній роботі 3.

3. Побудова нелінійної виробничої функції Кобба – Дугласа.

Слід перевірити існування нелінійного зв'язку між обсягом виробництва й величиною виробничих ресурсів шляхом побудови виробничої функції Кобба – Дугласа та провести аналіз адекватності нелінійної економетричної моделі.

Перевірку існування нелінійного зв'язку між обсягом виробництва й величиною виробничих ресурсів на основі побудови виробничої функції Кобба – Дугласа буде проведено в модулі *Advanced Linear/Nonlinear Models* (*додаткові лінійні/нелінійні моделі*). Вибір модуля *Nonlinear Estimation* (*нелінійне оцінювання*) наведено на рис. 10.4.6.

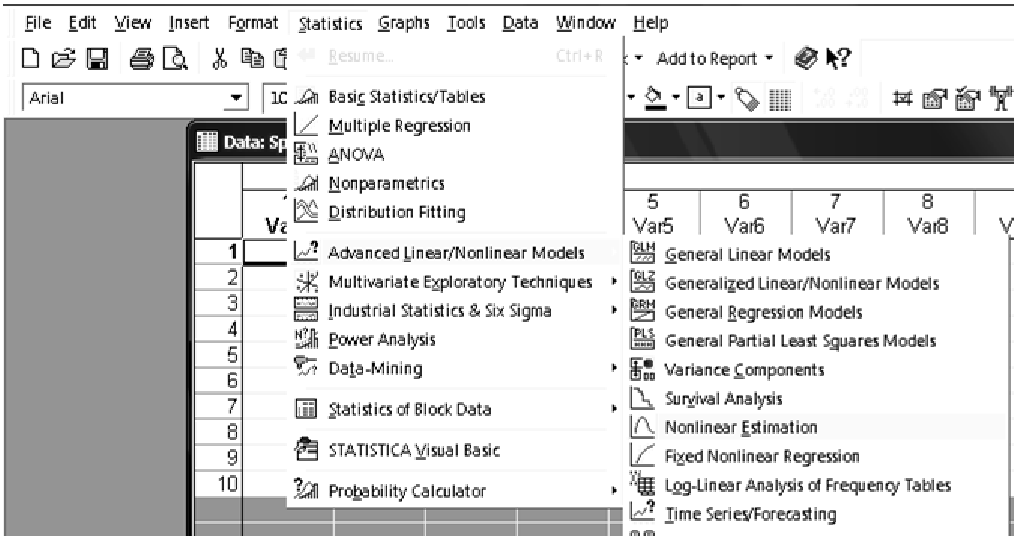


Рис. 10.4.6. **Вибір модуля *Nonlinear Estimation***

Вид стартової панелі модуля наведений на рис. 10.4.7. У вікні модуля подано такі види нелінійної оцінки:

User – specified regression, least squares (визначена користувачем регресія з похибками за методом найменших квадратів);

User – specified regression, custom loss function (визначена користувачем регресія з заданою функцією похибок);

Quick Logit regression (логіт регресія);

Quick Probit regression (пробіт регресія);

Exponential growth regression (регресія експоненційного зростання);

Piecewise linear regression (шматочно-лінійна регресія).

Для вирішення поставленої задачі необхідно вибрати опцію *User – specified regression, least squares* (рис. 10.4.7).

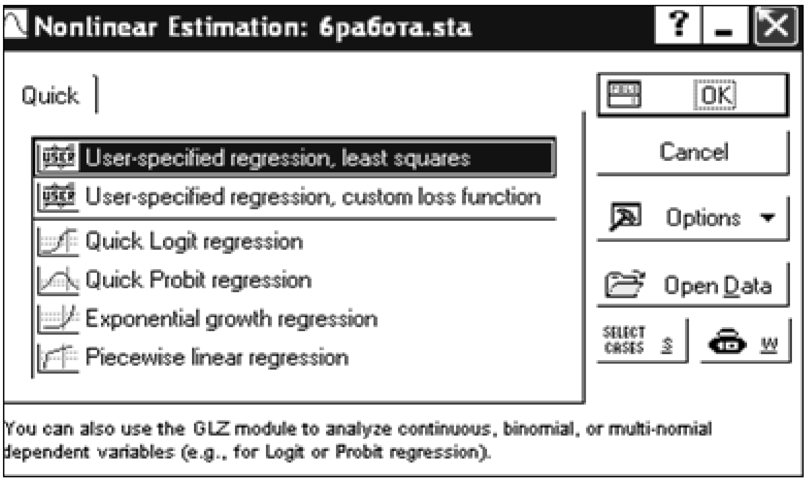


Рис. 10.4.7. **Стартова панель модуля *Nonlinear Estimation***

Після чого у вікні *Function to be estimated* задати вид оцінюваної функції, тобто вписати математичну модель Кобба – Дугласа виду: $Y = a_0 L^{a1} K^{a2}$, як показано на рис. 10.4.8.

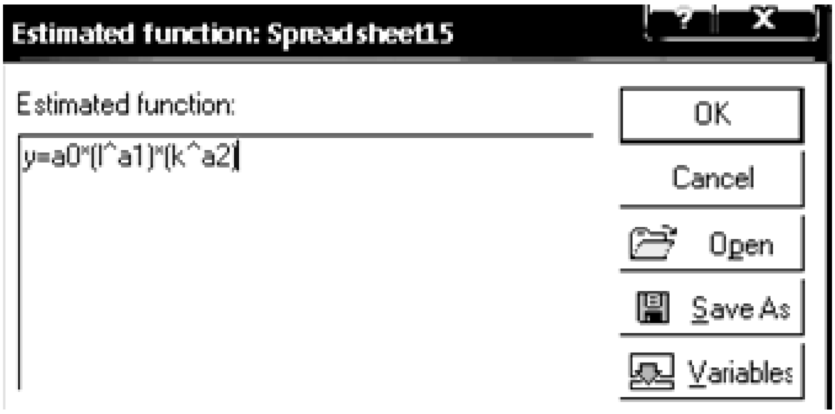


Рис. 10.4.8. **Визначення виду функції оцінювання**

У наступному вікні необхідно задати метод оцінювання параметрів нелінійної функції *Levenberg-Marquardt* (Левенберга – Марквардта) або *Gauss-Newton (Гауса – Ньютона)*, у разі необхідності на вкладці *Advanced* можна визначити параметри ітераційної процедури та початкові значення оцінюваних параметрів і запустити процедуру оцінювання (рис. 10.4.9).

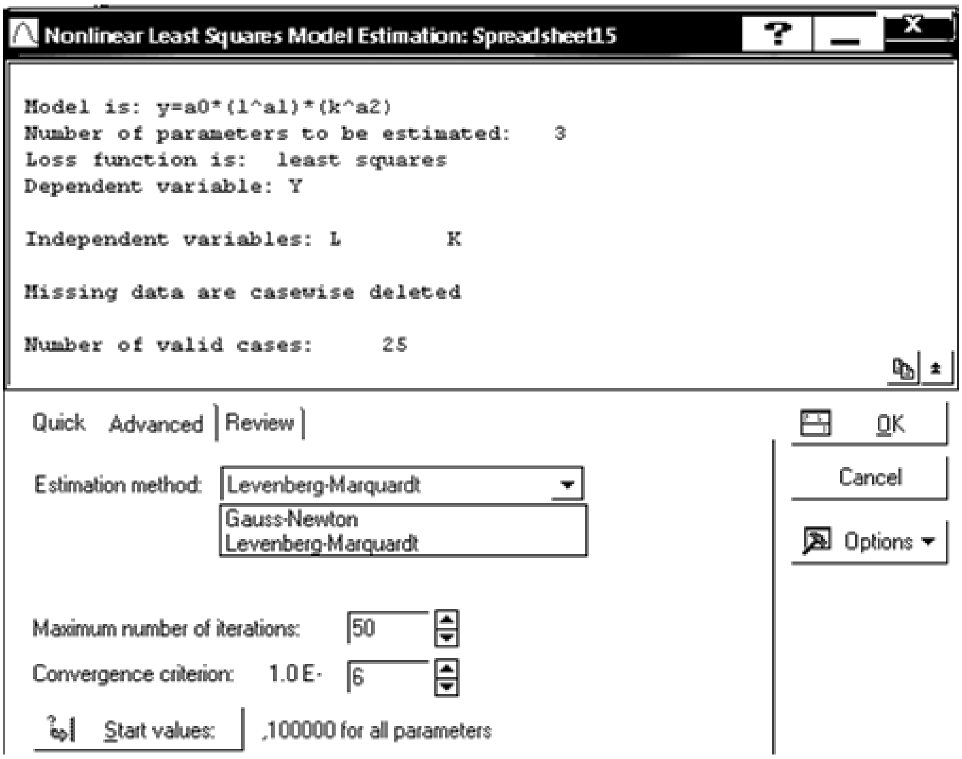


Рис. 10.4.9. **Вибір методу оцінювання параметрів моделі**

Результати побудови виробничої функції Кобба – Дугласа можна проаналізувати за допомогою вікна, наведеного на рис. 10.4.10.

У даному вікні подані результати моделі (вид досліджуваної моделі, залежна та незалежні змінні, вид функції оцінювання), у нижній інформаційній частині наведені опції для всебічного аналізу моделей.

Якість отриманої моделі оцінюється за допомогою значень: *Final value of Loss function* (*кінцеве значення функції втрат – сума квадратів похибок моделі*), *Proportion of variance accounted for* (*відсоток поясненої дисперсії*), *R* (*коефіцієнт множинної кореляції*).

Оцінки параметрів моделі можна одержати, ініціювавши опцію *Quick/ Summary: Parameter estimates* (*результат: параметри моделі*) (рис. 10.4.10). Результат оцінювання параметрів наведений на рис. 10.4.11.

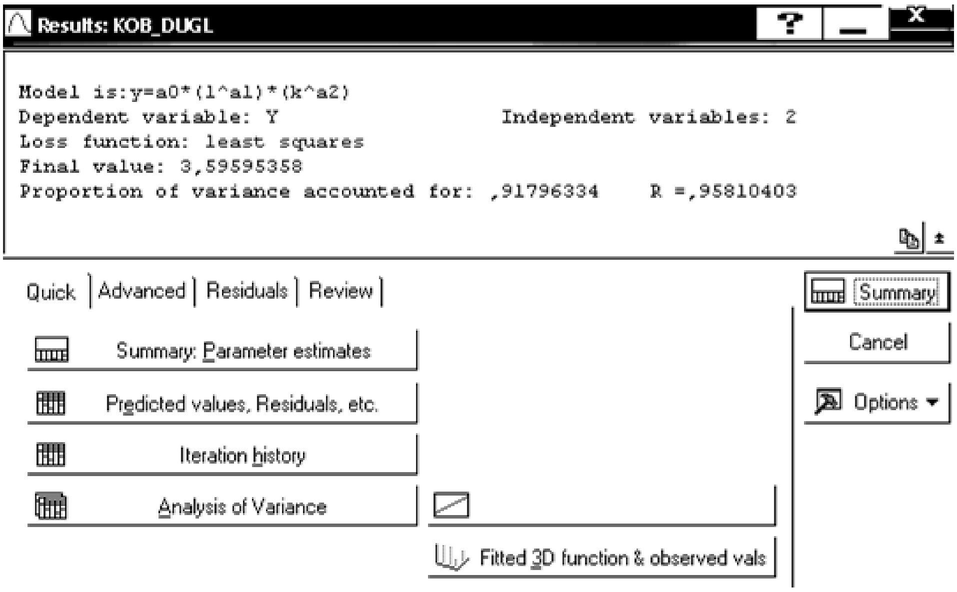


Рис. 10.4.10. **Результати побудови нелінійної функції**

	Model is: y=a0*(l^a1)*(k^a2) (KOB_DUGL) Dep. Var. : Y Level of confidence: 95.0% (alpha=0.050)					
	Estimate	Standard error	t-value df = 22	p-level	Lo. Conf Limit	Up. Conf Limit
a0	1,015082	0,196381	5,16894	0,000035	0,607813	1,422351
a1	0,831944	0,066939	12,42837	0,000000	0,693121	0,970767
a2	0,179076	0,052450	3,41424	0,002485	0,070302	0,287850

Рис. 10.4.11. **Результат оцінювання параметрів моделі**

У даному вікні наведені оцінки параметрів моделі *(Estimates),* значення середнє квадратичних відхилень параметрів (*Standard error*), статистики

Стьюдента(*t-value*), рівень значущості статистики Стьюдента *(p-level),* нижня й верхня межі інтервальних оцінок параметрів (*Lo. Conf Limit; Up. Conf Limit*).

Таким чином, виробнича функція Кобба – Дугласа буде мати такий вигляд: $\hat{Y} = 1{,}0151 \cdot L^{0{,}832} \cdot K^{0{,}1791}$. Інтервальні оцінки параметрів можуть змінюватись у межах:

$$0{,}6078 \le a_0 \le 1{,}4224; \;\; 0{,}6931 \le a_1 \le 0{,}9708; \;\; 0{,}0703 \le a_0 \le 0{,}2879.$$

Ініціюючи клавішу *Iteration history* (історія ітерацій), буде отримано результати покрокової процедури оцінки параметрів. Параметри моделі, які мінімізують критерій мінімуму суми квадратів похибок *(loss function – корінь суми квадратів похибок),* отримано за 18 ітерацій (рис. 10.4.12).

	Model is: Y=a0*(L^a1)*(K^a2) (KOB_DUGI Dep. Var. : Y			
	Loss Function	a0	a1	a2
1	90,88450	0,100000	0,100000	0,100000
2	86,61525	0,162661	0,324647	0,297077
3	82,21830	0,197060	0,400325	0,363446
4	74,07371	0,238798	0,476075	0,429868
5	57,39187	0,293465	0,557900	0,501593
6	23,82586	0,362687	0,642153	0,575360
7	3,45091	0,467805	0,725975	0,519059
8	3,00184	0,504927	0,741867	0,476206
9	2,77441	0,586740	0,777673	0,394912
10	2,41607	0,676775	0,814787	0,317968
11	2,24122	0,783303	0,843207	0,246994
12	1,95950	0,848843	0,852135	0,216425
13	1,92320	0,874916	0,853537	0,206579
14	1,91323	0,930557	0,847682	0,191915
15	1,90420	0,987293	0,837301	0,182518
16	1,89661	1,014538	0,831927	0,179142
17	1,89630	1,015079	0,831944	0,179077
18	1,89630	1,015082	0,831944	0,179076

Рис. 10.4.12. **Ітераційна процедура визначення параметрів**

Дисперсійний аналіз побудованої моделі буде отримано ініціюванням *Analysis of Variance (аналіз варіації)* (рис. 10.4.13). У таблиці дисперсійного аналізу наведено суму квадратів відхилень за регресією *(Sums of SquaresRegress),* суму квадратів похибок моделі *(Sums of Squares Residual),* дисперсію похибок *(Mean Squares Residual)* та критерій Фішера *(F-value)*.

Effect	Model is: Y=a0*(L^a1)*(K^a2) (KOB_DUGL) Dep. Var. : Y 1 Sum of Sqares	2 DF	3 Mean Squares	4 F-value	5 p-value
Regression	8422,006	3,00000	2807,335	17175,24	0,00
Residual	3,596	22,00000	0,163		
Total	8425,602	25,00000			
Corrected Total	43,833	24,00000			
Regression vs.Corrected Total	8422,006	3,00000	2807,335	1537,09	0,00

Рис. 10.4.13. **Результати дисперсійного аналізу**

Для подальшого аналізу необхідно отримати теоретичні значення залежної змінної та похибки моделі, ініціювавши клавішу *(Predicted values, Residuals)* на вкладці *Quick.* Результати наведені на рис. 10.4.14.

	Model is: y=a0*(L^a1)*(K^a2) (KOB Dep. Var. : Y Observed	Predicted	Residuals
1	18,22000	18,51119	-0,291194
2	17,53000	17,87537	-0,345367
3	16,35000	16,40290	-0,052898
4	17,94000	18,03605	-0,096054
5	19,24000	19,12316	0,116838
6	17,66000	17,67086	-0,010860
7	18,21000	18,05382	0,156184
8	18,44000	18,71646	-0,276463
9	19,76000	19,78501	-0,025006
10	17,85000	18,17116	-0,321163
11	19,27000	19,37352	-0,103524
12	20,93000	20,31371	0,616287
13	19,04000	19,04406	-0,004055
14	16,75000	16,72527	0,024725
15	19,56000	19,32760	0,232404
16	17,82000	18,49325	-0,673249
17	16,88000	16,29501	0,584987
18	17,46000	17,13878	0,321217
19	19,02000	18,71514	0,304857
20	20,11000	20,27532	-0,165316
21	19,32000	19,08862	0,231383
22	18,04000	17,85077	0,189229
23	20,22000	20,39817	-0,178171
24	17,01000	16,24533	0,764668
25	15,13000	16,12579	-0,995790

Рис. 10.4.14. **Теоретичні значення залежної змінної та похибки моделі**

Для порівняння емпіричних (досліджуваних) та теоретичних (розрахованих за моделлю) значень залежної змінної слід побудувати графік. Вибір опцій графічного аналізу отриманих результатів та типу графіка подані на рис. 10.4.15. На рис. 10.4.16 наведено графік порівняння теоретичних та емпіричних значень залежної змінної.

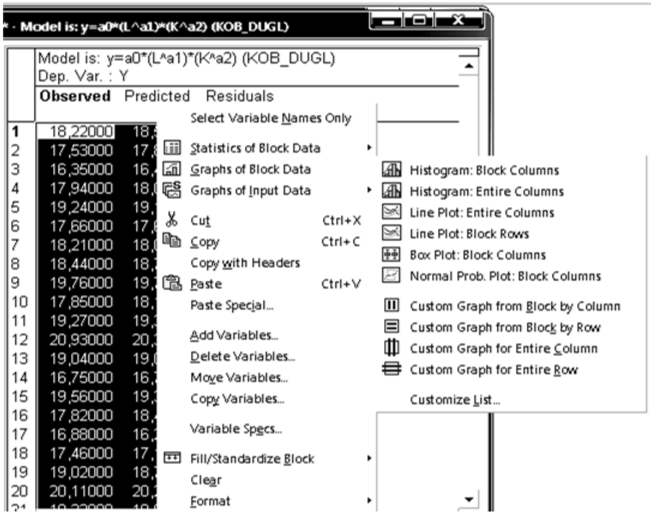


Рис. 10.4.15. **Вибір опцій графічного аналізу змінних**

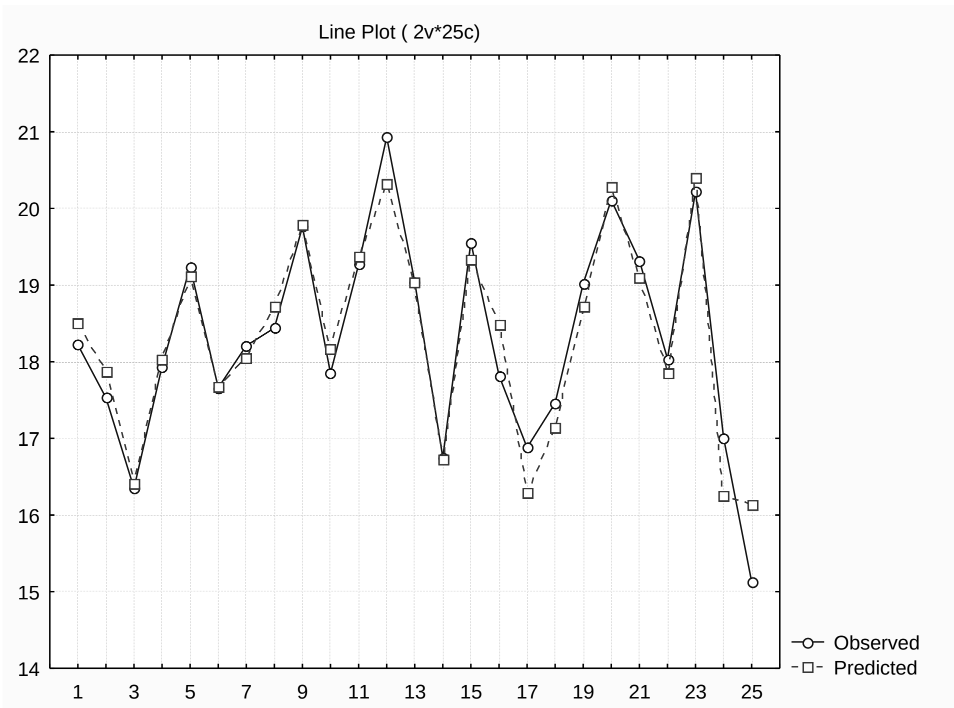


Рис. 10.4.16. **Графік теоретичних та емпіричних значень залежної змінної**

Графік виробничої функції у тривимірному просторі (залежність факторів виробництва і обсягу виробленої продукції) буде одержано ініціюванням опції *Fitted 3Dfunction & observed vals* (див. рис. 10.4.10). Графік наведено на рис. 10.4.17.

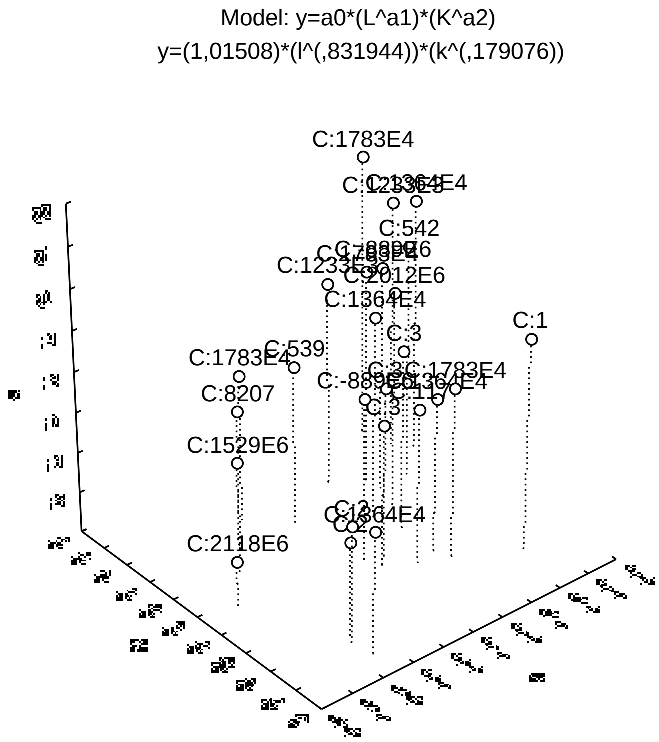


Рис. 10.4.17. **Графік виробничої функції у тривимірному просторі**

Усебічний аналіз помилок моделі можна отримати в опціях *Residuals (похибки)* (рис. 10.4.18).

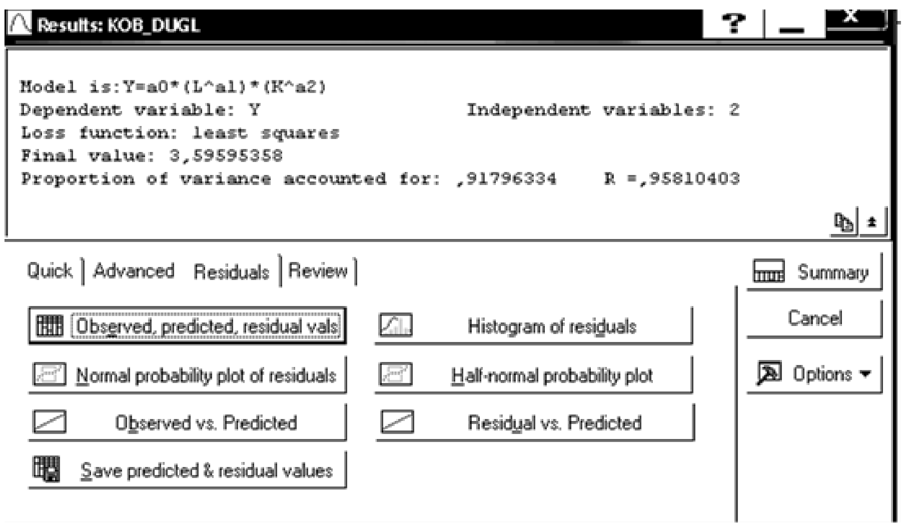


Рис. 10.4.18. **Меню аналізу помилок моделі**

Гістограма помилок та графік розподілу помилок на нормальному ймовірнісному папері наведено на рис. 10.4.19, 10.4.20.

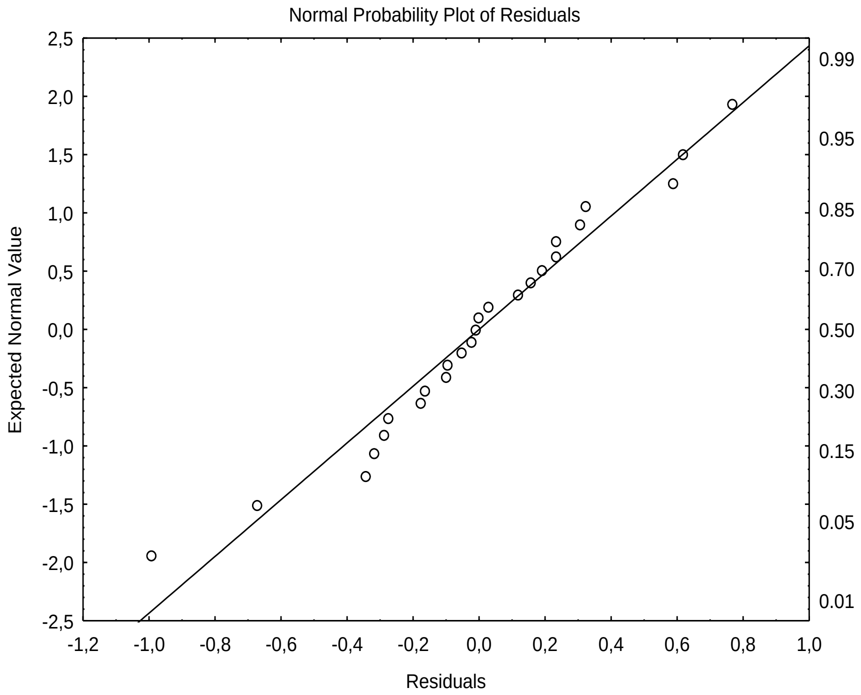


Рис. 10.4.19. **Графік розподілу помилок на нормальному ймовірнісному папері**

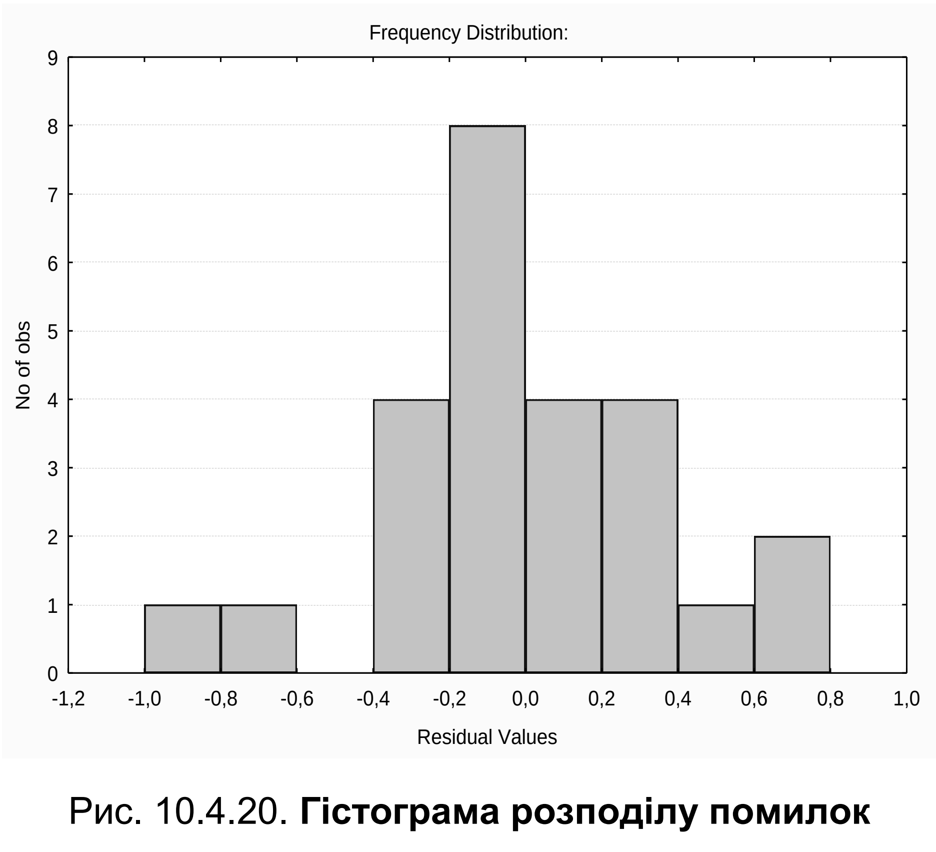


Рис. 10.4.20. **Гістограма розподілу помилок**

4. Визначення основних характеристик виробничої функції.

Розрахунок характеристик виробничої функції можливий шляхом задавання формули розрахунку в області специфікації змінної *Long name* (рис. 10.4.21).

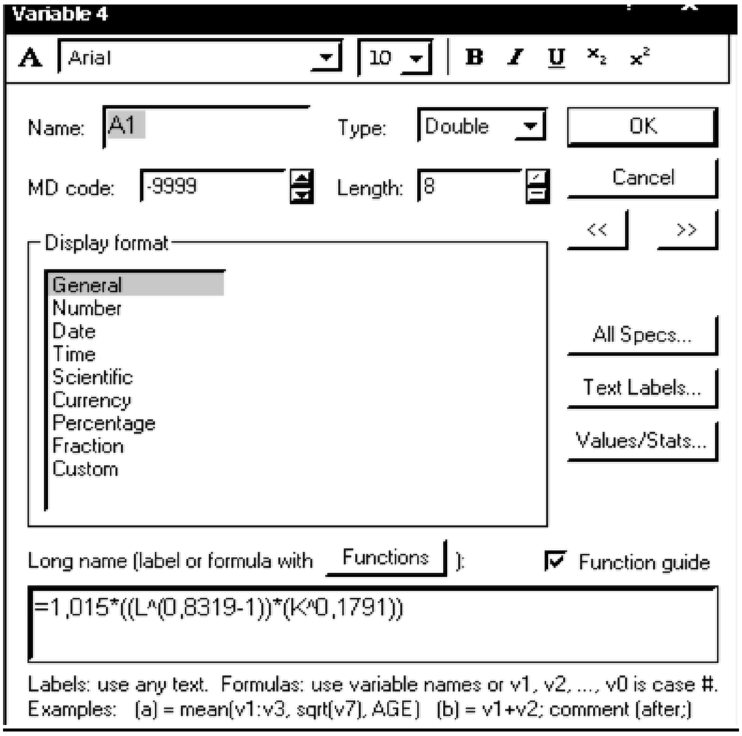


Рис. 10.4.21. **Розрахунок характеристик виробничої функції**

Для подальшого аналізу необхідно визначити такі характеристики (див. розділ 6):

1) середня продуктивність праці:

$$A_1 = \frac{y}{x_1} = a_0 \cdot x_1^{a_1 - 1} \cdot x_2^{a_2};$$

2) середня фондовіддача (капіталовіддача):

$$A_2 = \frac{y}{x_1} = a_0 \cdot x_1^{a_1} \cdot x_2^{a_2 - 1};$$

3) гранична продуктивність праці:

$$M_1 = \frac{\partial y}{\partial x_1} = a_0 \cdot a_1 \cdot x_1^{a_1 - 1} \cdot x_2^{a_2};$$

4) гранична фондовіддача (капіталовіддача):

$$M_2 = \frac{\partial y}{\partial x_2} = a_0 \cdot a_2 \cdot x_1^{a_1} \cdot x_2^{a_2 - 1};$$

5) еластичність випуску продукції за витратами праці:

$$E_1 = \frac{\partial y}{\partial x_1} \cdot \frac{x_1}{y}; \qquad E_1 = \frac{M_1}{A_1};$$

6) еластичність випуску продукції за витратами виробничих фондів:

$$E_2 = \frac{\partial y}{\partial x_2} \cdot \frac{x_2}{y}; \qquad E_2 = \frac{M_2}{A_2};$$

7) сумарна еластичність за витратами (праці і капіталу):

$$E = E_1 + E_2;$$

8) фондоозброєність (капіталоозброєність) праці:

$$FT = \frac{x_2}{x_1} = a_0^{-\frac{1}{a_2}} \cdot y_0^{\frac{1}{a_2}} \cdot x_1^{-1-\frac{a_1}{a_2}};$$

9) фондомісткість продукції:

$$F = \frac{x_2}{Y} = \frac{1}{a_0}\ \frac{x_2}{x_1}^{\ a_1};$$

10) потреба у витратах праці та капіталу:

$$X_1 = \ \frac{Y}{a_0 \cdot x_2^{a_2}}^{\ \frac{1}{a_1}}; X_2 = \ \frac{Y}{a_0 \cdot x_1^{a_1}}^{\ \frac{1}{a_2}};$$

11) гранична норма заміни i-го ресурсу j-м ресурсом:

$$R_{ij} = -\frac{\Delta x_j}{\Delta x_i}, \qquad R'_{ij} = -\frac{\partial x_j}{\partial x_i}; \quad R_{12} = \frac{E_1}{E_2} \cdot \frac{x_2}{x_1}.$$

Результати розрахунку основних характеристик подані на рис. 10.4.22.

	1 L	2 K	3 Y	4 A1	5 A2	6 M1	7 M2	8 E1	9 E2	10 E	11 FT	12 F
1	16,88	21,84	18,22	1,096	0,847	0,912	0,152	0,8319	0,1791	1,01	1,294	1,199
2	16,05	22,71	17,53	1,114	0,787	0,926	0,141	0,8319	0,1791	1,01	1,415	1,295
3	14,74	20,87	16,35	1,113	0,786	0,926	0,141	0,8319	0,1791	1,01	1,416	1,276
4	15,8	25,68	17,94	1,141	0,702	0,95	0,126	0,8319	0,1791	1,01	1,625	1,431
5	17,72	20,9	19,24	1,079	0,915	0,898	0,164	0,8319	0,1791	1,01	1,179	1,086
6	15,29	26,68	17,66	1,156	0,662	0,961	0,119	0,8319	0,1791	1,01	1,745	1,511
7	16,2	22,99	18,21	1,114	0,785	0,927	0,141	0,8319	0,1791	1,01	1,419	1,262
8	16,81	23,68	18,44	1,113	0,79	0,926	0,142	0,8319	0,1791	1,01	1,409	1,284
9	17,68	25,54	19,76	1,119	0,775	0,931	0,139	0,8319	0,1791	1,01	1,445	1,293
10	16,45	22,2	17,85	1,104	0,818	0,919	0,147	0,8319	0,1791	1,01	1,35	1,244
11	17,17	26,02	19,27	1,128	0,744	0,939	0,133	0,8319	0,1791	1,01	1,515	1,35
12	17,9	27,94	20,93	1,135	0,727	0,944	0,13	0,8319	0,1791	1,01	1,561	1,335
13	16,74	26,6	19,04	1,137	0,716	0,946	0,128	0,8319	0,1791	1,01	1,589	1,397
14	14,53	24,87	16,75	1,151	0,672	0,957	0,12	0,8319	0,1791	1,01	1,712	1,485
15	17,2	25,47	19,56	1,124	0,759	0,935	0,136	0,8319	0,1791	1,01	1,481	1,302
16	16,78	22,33	17,82	1,102	0,828	0,917	0,148	0,8319	0,1791	1,01	1,331	1,253
17	14,66	20,63	16,88	1,111	0,79	0,925	0,141	0,8319	0,1791	1,01	1,407	1,222
18	14,86	25,68	17,46	1,153	0,667	0,959	0,12	0,8319	0,1791	1,01	1,728	1,471
19	16,68	24,54	19,02	1,122	0,763	0,933	0,137	0,8319	0,1791	1,01	1,471	1,29
20	18	26,94	20,11	1,126	0,753	0,937	0,135	0,8319	0,1791	1,01	1,497	1,34
21	17,06	24,68	19,32	1,119	0,773	0,931	0,139	0,8319	0,1791	1,01	1,447	1,277
22	15,95	23,2	18,04	1,119	0,769	0,931	0,138	0,8319	0,1791	1,01	1,455	1,286
23	18,2	26,47	20,22	1,121	0,771	0,932	0,138	0,8319	0,1791	1,01	1,454	1,309
24	14,72	19,9	17,01	1,103	0,816	0,918	0,146	0,8319	0,1791	1,01	1,352	1,17
25	14,05	23,71	15,13	1,148	0,68	0,955	0,122	0,8319	0,1791	1,01	1,688	1,567

Рис. 10.4.22. **Результат розрахунку характеристик виробничої функції**

5. Побудова ізоквант виробничої функції.

Для побудови ізоквант необхідно знайти можливі комбінації виробничих ресурсів за умови фіксованих рівнів виробництва. Для цього необхідно розрахувати значення показника витрат одного з ресурсів у разі відомих витрат другого ресурсу. Побудувати ізокванти для таких прогнозованих значень рівня виробництва: ($Y_{пр}$ = 26,5 млн грн, $Y_{пр}$ = 20,2 млн грн, $Y_{пр}$ = 28,3 млн грн, $Y_{пр}$ = 25,1 млн грн), залишаючи значення показника витрат трудових ресурсів попередніми за формулою:

$$X_2 = \frac{Y}{a_0 \cdot x_1^{a_1}}^{\frac{1}{a_2}}.$$

Розрахунок потреби у витратах капіталу та значення граничної норми заміни i-го ресурсу j-м ресурсом подано на рис. 10.4.23.

За отриманим даними будуються графіки ізоквант виробничої функції. Для цього вибрати в пункті меню *Graphs* підпункт *Scatterplots* (*Точкові графіки*).

	1 L	2 K	3 Y	4 K1	5 K2	6 K3	7 K4	8 R_1_2	9 R_2_1
1	16,88	21,84	18,22	162,0095	35,58604	233,8277	119,6537	6,009741	0,166397
2	16,05	22,71	17,53	204,763	44,97703	295,5337	151,2297	6,572304	0,152154
3	14,74	20,87	16,35	304,0943	66,79554	438,8981	224,5917	6,576586	0,152055
4	15,8	25,68	17,94	220,2522	48,37928	317,8891	162,6693	7,549418	0,132461
5	17,72	20,9	19,24	129,2926	28,39965	186,6076	95,49031	5,478455	0,182533
6	15,29	26,68	17,66	256,5122	56,34394	370,2231	189,4495	8,105016	0,12338
7	16,2	22,99	18,21	196,1039	43,07502	283,036	144,8344	6,591731	0,151705
8	16,81	23,68	18,44	165,167	36,2796	238,3849	121,9857	6,54319	0,152831
9	17,68	25,54	19,76	130,657	28,69933	188,5767	96,49795	6,709871	0,149034
10	16,45	22,2	17,85	182,639	40,11741	263,6022	134,8898	6,268485	0,159528
11	17,17	26,02	19,27	149,685	32,87891	216,0398	110,5513	7,039025	0,142065
12	17,9	27,94	20,93	123,3633	27,09724	178,0497	91,11113	7,250182	0,137928
13	16,74	26,6	19,04	168,3996	36,98965	243,0505	124,3731	7,380771	0,135487
14	14,53	24,87	16,75	325,0533	71,39928	469,1482	240,0712	7,95034	0,125781
15	17,2	25,47	19,56	148,4762	32,61339	214,2951	109,6585	6,87822	0,145386
16	16,78	22,33	17,82	166,5431	36,58186	240,371	123,002	6,181193	0,161781
17	14,66	20,63	16,88	311,8793	68,50554	450,1341	230,3414	6,536433	0,152989
18	14,86	25,68	17,46	292,8546	64,3267	422,6759	216,2906	8,026972	0,12458
19	16,68	24,54	19,02	171,2317	37,61175	247,1381	126,4648	6,833671	0,146334
20	18	26,94	20,11	120,212	26,40505	173,5015	88,7837	6,951854	0,143847
21	17,06	24,68	19,32	154,221	33,87526	222,5865	113,9014	6,719573	0,148819
22	15,95	23,2	18,04	210,7946	46,30188	304,239	155,6843	6,756205	0,148012
23	18,2	26,47	20,22	114,1977	25,08399	164,8211	84,34182	6,755509	0,148027
24	14,72	19,9	17,01	306,0182	67,21813	441,6748	226,0126	6,279438	0,15925
25	14,05	23,71	15,13	379,9446	83,45637	548,3725	280,6117	7,83846	0,127576

Рис. 10.4.23. **Результат розрахунку для побудови ізоквант**

Під час задавання характеристик графіка вибрати *Graph type – Multiple* (*складний*), зніти оцінку *Linear fit*, у якості відображених змінних на графіку, вибираємо по осі *X* значення трудових ресурсів (*L),* а по осі *Y* – розраховані значення потреби в основних фондах (*К1 – К4*), як це подано на рис. 10.4.24.

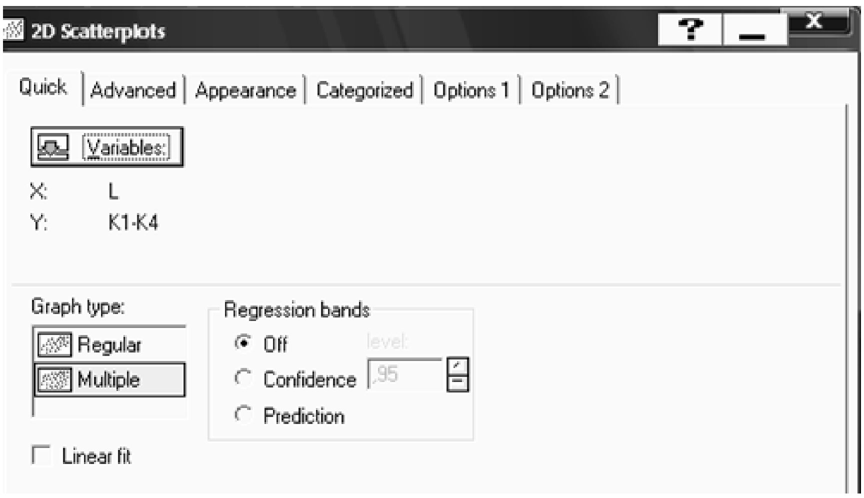


Рис. 10.4.24. **Визначення характеристик графіка**

У результаті буде одержано графік ізоквант для заданих значень рівня виробництва (рис. 10.4.25).

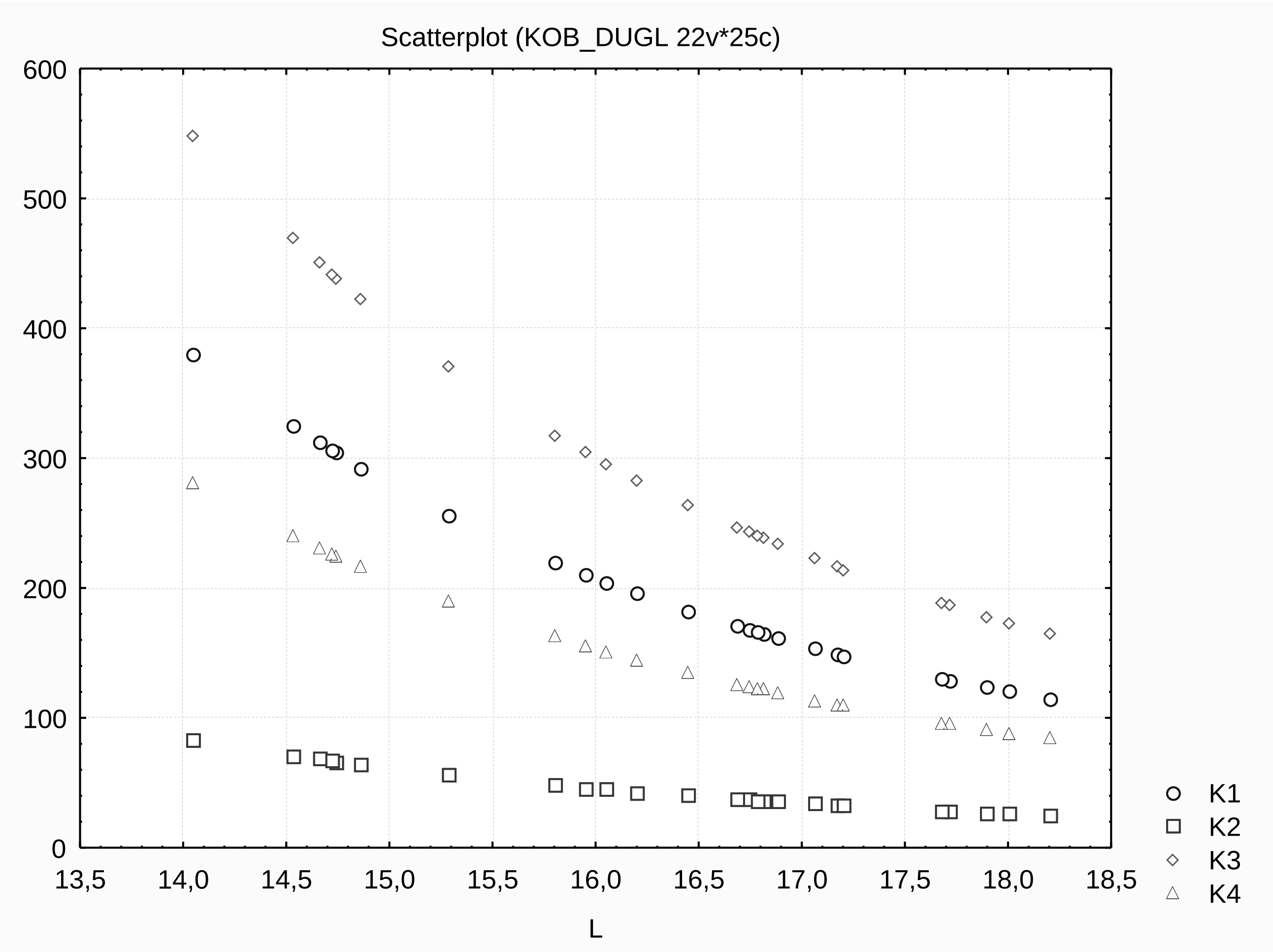


Рис. 10.4.25. **Графік ізоквант виробничої функції**

Ізокванта виробничої функції – це лінії рівня $q = f(L, K), (q > 0)$, що становить собою безліч точок, у яких ВФ набуває значення, рівного q (рівень виробництва). Таким чином, ізокванти є різним набором (співвідношенням) використовуваних ресурсів, що забезпечують однаковий обсяг випуску продукції. На даному графіку можна також провести ізокліналь у будь-якій точці та розрахувати граничні норми заміщення ресурсів.

Лабораторна робота 5. Побудова й аналіз економетричних моделей динаміки

Мета – закріплення теоретичного й практичного матеріалу за темою "Економетричні моделі динаміки", набуття навичок побудови й аналізу економетричних моделей динаміки в модулі *Time Series/Forecasting*.

Завдання – необхідно побудувати моделі декомпозиції за даними часового ряду об'єму продажів у модулі *Time Series/Forecasting* ППП *Statistica*:

1. Навести графік динаміки показника й проаналізувати характер зміни значень показника у часі.

2. Перевірити наявність тренда в середньому та дисперсії за методами Стьюдента та Фішера.

3. Провести декомпозицію часового ряду на такі складові частини: трендово-циклічну, сезонну і випадкову, використовуючи мультиплікативну модель часового ряду. Навести таблицю результатів розрахунку цих складових часового ряду. Виділити, якщо є, тренд із трендово-циклічної складової. Оцінити параметри тренда за допомогою МНК. Навести графіки тренда й трендово-циклічної складової. Проаналізувати циклічну складову, навести її графік. Проаналізувати сезонні індекси, зробити висновки про вплив сезонних факторів на показник. Проаналізувати розподіл і характеристики випадкової величини.

4. Провести декомпозицію часового ряду на такі складові частини: трендово-циклічну, сезонну і випадкову, використовуючи аддитивну модель часового ряду.

5. Навести оцінки якості моделей часових рядів (середня помилка, середня абсолютна помилка, середнє квадратичне відхилення помилок, середня відсоткова помилка, середня абсолютна процентна помилка). Виконати порівняльний аналіз моделей і визначити найбільш адекватну з них.

6. Розрахувати за допомогою моделі декомпозиції часового ряду прогнозні значення показника на 2 роки уперед за кварталами.

7. Навести економічну інтерпретацію моделей.

Для побудови й всебічного аналізу економетричних моделей динаміки у ППП *Statistica* передбачений модуль *Time Series/Forecasting (часові ряди/прогнозування*), де крім моделей декомпозиції можлива побудова різних типів моделей.

Вихідні дані динаміки об'єму продажів для побудови моделей декомпозиції наведені на рис. 10.5.1.

1. Дослідження характеру поведінки показника та перевірка наявності тренда.

Для побудови графіка вихідних даних вибрати в пункті меню *Graphs* підпункт *Scatterplots (Точкові графіки)*. Під час задавання характеристик графіка вибрати *Graph type – Regular* (*простій*), поставити оцінку *Linear fit* (*лінійний тренд*), у якості змінних, відображених на графіку, вибрати по осі *X період часу*, а по осі *Y – об'єм продажу*. Результат наведений на рис. 10.5.2.

	1 YEAR	2 Quater	3 t	4 Y
1	2000	1	1	48,6
2		2	2	54,2
3		3	3	56,62
4		4	4	79,8
5	2001	1	5	49,5
6		2	6	56
7		3	7	63,5
8		4	8	85,5
9	2002	1	9	54,7
10		2	10	59,9
11		3	11	65
12		4	12	89
13	2003	1	13	57
14		2	14	63,9
15		3	15	68,6
16		4	16	94,2
17	2004	1	17	58,7
18		2	18	67,1
19		3	19	74,2
20		4	20	102,8
21	2005	1	21	65,3
22		2	22	72,3
23		3	23	78,3
24		4	24	112,2
25	2006	1	25	72,1
26		2	26	82
27		3	27	91,6
28		4	28	127,5
29	2007	1	29	82,3
30		2	30	92
31		3	31	96,7
32		4	32	134,9

Рис. 10.5.1. **Вихідні дані динаміки об'єму продажів**

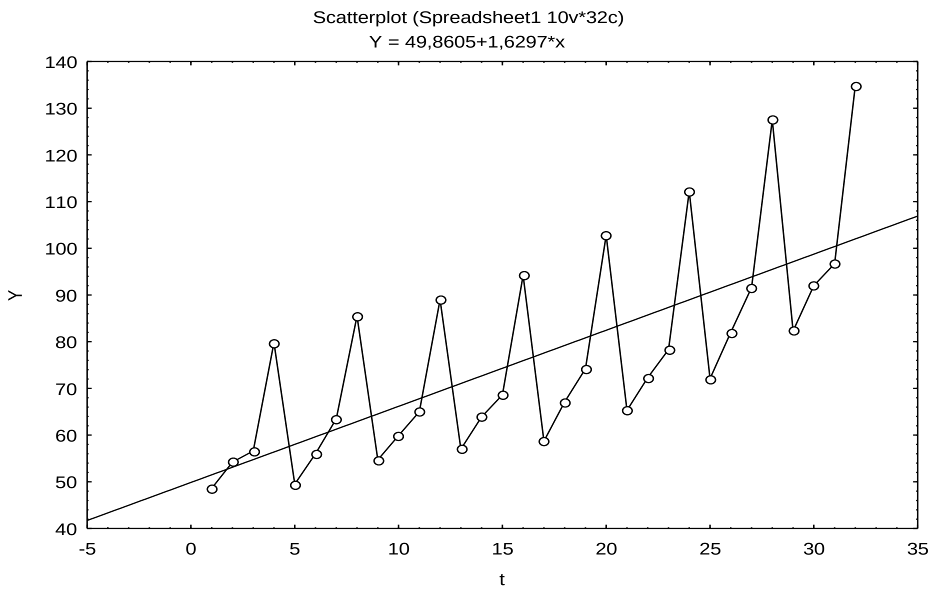


Рис. 10.5.2. **Графік вихідних даних динаміки об'єму продажів**

Основне завдання, що постає під час аналізу часових рядів – визначення наявності тренда. Існують різні методи, що дозволяють визначити наявність тренда: критерій Фішера для визначення тренда в дисперсії; критерій Стьюдента для визначення наявності тренда у середньому (див. розділ 7).

Значення критерію Фішера дорівнює:

$$F_{\text{теор}} = \frac{488{,}79}{203{,}684} = 2{,}39.$$

$F_{\text{табл}}(0{,}05;\ 14;\ 14) = 2{,}48$, отже $F_{\text{теор}} < F_{\text{табл}}$, тренд у дисперсії відсутній.

Якщо немає тренда у дисперсії рівнів ряду, то ряд тестується на наявність тренда у середньому за допомогою критерію Стьюдента:

$$t_{\text{теор}} = \frac{|65{,}3763 - 88{,}125|}{\sqrt{15 \cdot 203{,}684 + 15 \cdot 488{,}79}} \cdot \sqrt{\frac{16 \cdot 16 \cdot (16 + 16 - 2)}{16 + 16}} = 3{,}45.$$

$t_{\text{табл}}$(0,95;30)=2,042, а отже $t_{\text{теор}} > t_{\text{табл}}$, і тренд у середньому існує.

Аналіз часових рядів здійснюється в модулі *Advanced Linear/Nonlinear Models/Time Series/Forecasting (часові ряди/прогнозування*) (рис. 10.5.3). Стартова панель модуля наведена на рис. 10.5.4.

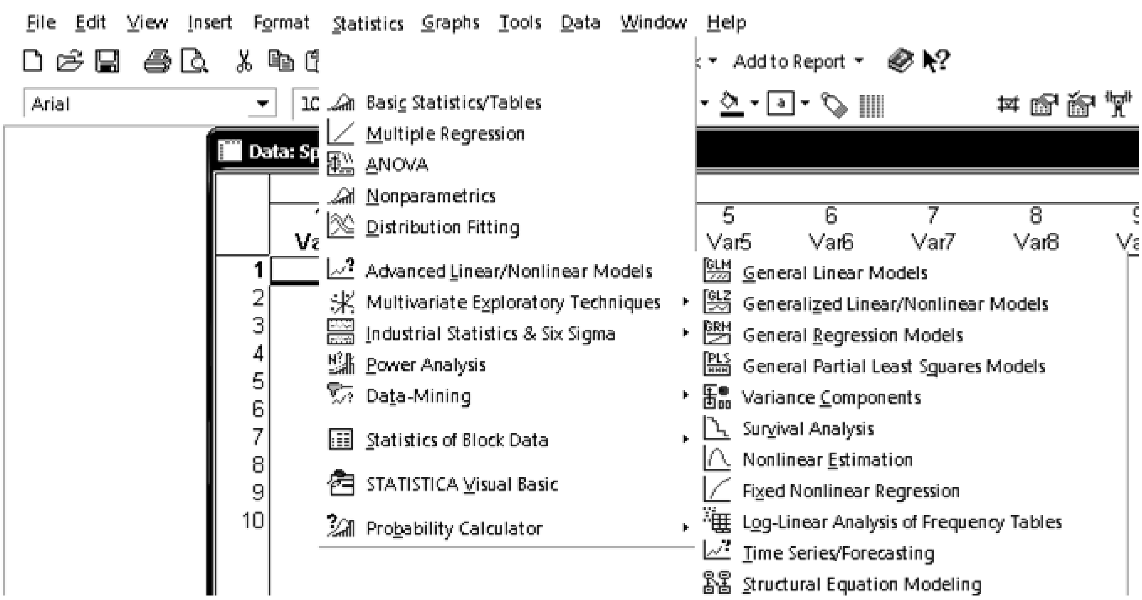


Рис. 10.5.3. **Вибір модуля *Time Series/Forecasting***

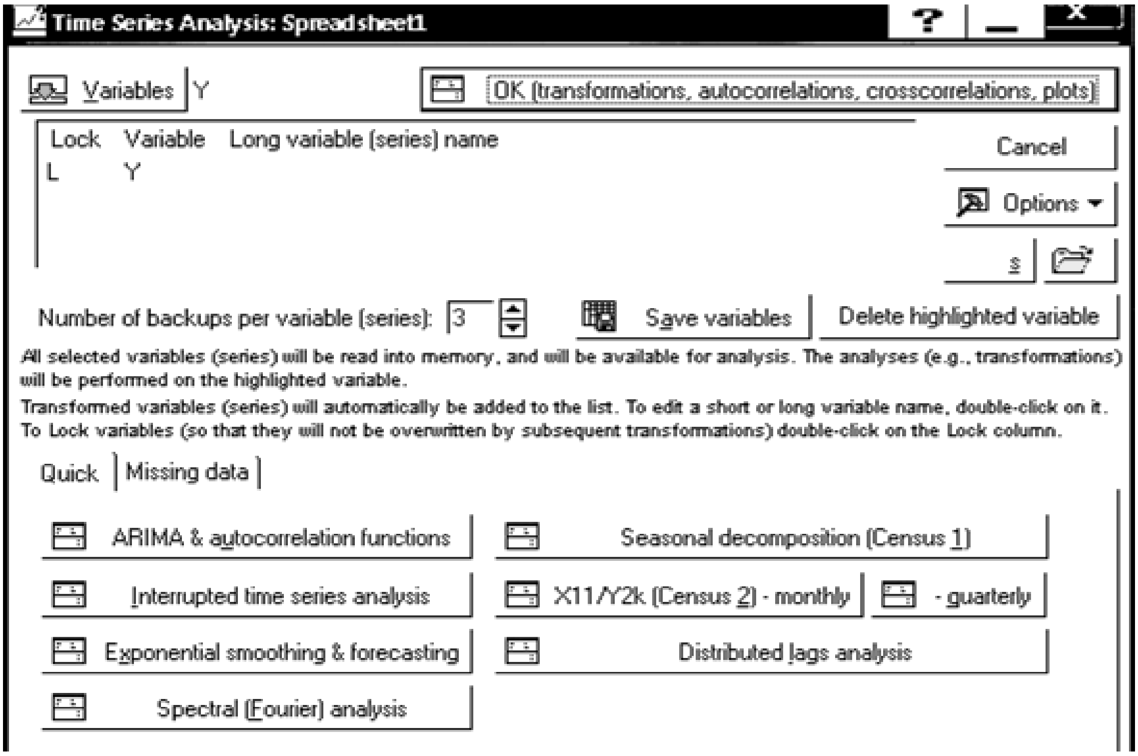


Рис. 10.5.4. **Стартова панель модуля *Time Series/Forecasting***

У даному модулі реалізовані такі методи аналізу часових рядів:

ARIMA & autocorrelation function – моделі авторегресії і проінтегрованої ковзної середньої;

Interrupted time series analysis – моделі аналізу перерваних часових рядів (моделі інтервенції);

Exponential smoothing & forecasting – моделі експоненційного згладжування та прогнозування;

Spectral (Fourier) analysis – моделі спектрального аналізу Фур'є;

Seasonal Decomposition (Census 1) – моделі сезонної декомпозиції;

X11 / Y2k (Census 2) monthly/quarterly – спеціальні моделі сезонної декомпозиції;

Distributed Lags Analysis – моделі аналізу розподілених лагів.

2. Мультиплікативна декомпозиція часового ряду.

Слід провести декомпозицію часового ряду на такі складові: трендово-циклічну, сезонну й випадкову, використовуючи мультиплікативну модель часового ряду. Для цього вибрати опцію *Seasonal Decomposition* (*сезонна декомпозиція*) у стартовій панелі модуля (див. рис. 10.5.4). Ініціювавши дану опцію, буде отримано таке вікно для вибору типу аналізу (рис. 10.5.5).

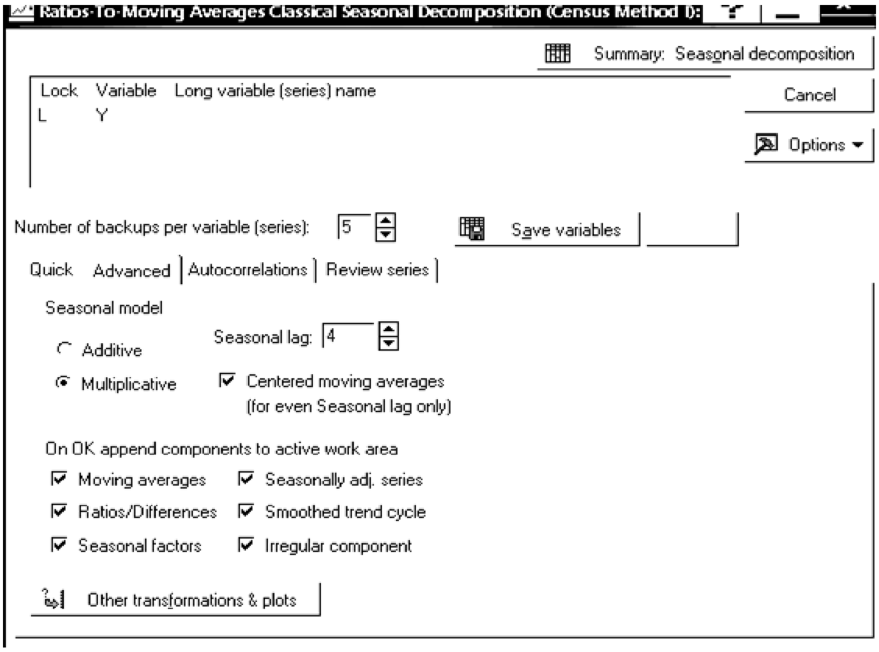


Рис. 10.5.5. **Параметри моделі сезонної декомпозиції**

У даному вікні у вкладці *Advanced* необхідно задати параметри сезонної декомпозиції:

змінна для аналізу – *Y* (о*б'єм продажів*);

вид моделі (*Seasonal model*) – *Multiplicative* (*Мультиплікативна*);

лаг сезонного компонента (задається число сезонних індексів) (*Seasonal lag*) – 4 (квартальні дані);

центрування ковзної середньої (*Centered moving averages*).

У цьому ж вікні вибрати додаткові компоненти, які необхідно розрахувати (див. рис. 10.5.5), що повинні бути відображені в робочій області:

Moving averages – ковзні середні;

Ratios/Differences – відношення/різниці (якщо модель мультиплікативна, то береться відношення, адитивна – різниця вихідного ряду та тренда);

Seasonal factors – сезонні індекси;

Seasonallyadj. series – скорегований ряд на сезонність (без сезонної компоненти);

Smoothed trend cycle – згладжена трендово-циклічна складова;

Irregular component – нерегулярна (випадкова) складова.

На рис. 10.5.6 наведені результати моделі сезонної декомпозиції об'єму продажів, які отримано, ініціюванням опції *Summary: Seasonal Decomposition* у верхній частині вікна на рис. 10.5.5. Таким чином, наявні такі компоненти:

Seasonal factors – сезонні індекси;

Smoothed trend cycle – згладжена трендово-циклічна складова;

Irregular component – випадкова складова.

Case	Seasonal Decomposition: Multipl. season (4); Centered means (Spreadsheet1) Y						
	Y	Moving Averages	Ratios	Seasonal Factors	Adjusted Series	Smoothed Trend-c.	Irreg. Compon.
1	48,6000			82,1508	59,1595	59,2184	0,999006
2	54,2000			90,0212	60,2080	59,3887	1,013796
3	56,6200	59,9175	94,4966	96,2948	58,7986	59,7294	0,984417
4	79,8000	60,2550	132,4371	131,5332	60,6691	60,2811	1,006436
5	49,5000	61,3400	80,6978	82,1508	60,2550	61,2512	0,983737
6	56,0000	62,9125	89,0125	90,0212	62,2076	62,7435	0,991459
7	63,5000	64,2750	98,7942	96,2948	65,9434	64,3434	1,024866
8	85,5000	65,4125	130,7090	131,5332	65,0026	65,4235	0,993566
9	54,7000	66,0875	82,7691	82,1508	66,5848	66,2538	1,004997
10	59,9000	66,7125	89,7883	90,0212	66,5399	66,7175	0,997338
11	65,0000	67,4375	96,3855	96,2948	67,5011	67,4310	1,001038
12	89,0000	68,2250	130,4507	131,5332	67,6635	68,2539	0,991350
13	57,0000	69,1750	82,3997	82,1508	69,3846	69,3542	1,000438
14	63,9000	70,2750	90,9285	90,0212	70,9833	70,3865	1,008479
15	68,6000	71,1375	96,4330	96,2948	71,2396	71,0842	1,002186
16	94,2000	71,7500	131,2892	131,5332	71,6169	71,7510	0,998131
17	58,7000	72,8500	80,5765	82,1508	71,4539	72,7740	0,981860
18	67,1000	74,6250	89,9162	90,0212	74,5380	74,4893	1,000653
19	74,2000	76,5250	96,9618	96,2948	77,0551	76,3882	1,008731
20	102,8000	78,0000	131,7949	131,5332	78,1552	78,0449	1,001413
21	65,3000	79,1625	82,4886	82,1508	79,4879	79,3079	1,002270
22	72,3000	80,8500	89,4249	90,0212	80,3144	80,6668	0,995631
23	78,3000	82,8750	94,4796	96,2948	81,3128	82,4916	0,985711
24	112,2000	84,9375	132,0971	131,5332	85,3016	85,0517	1,002939
25	72,1000	87,8125	82,1068	82,1508	87,7654	88,0574	0,996684
26	82,0000	91,3875	89,7278	90,0212	91,0897	91,2538	0,998201
27	91,6000	94,5750	96,8543	96,2948	95,1246	94,3742	1,007952
28	127,5000	97,1000	131,3079	131,5332	96,9337	97,1890	0,997373
29	82,3000	98,9875	83,1418	82,1508	100,1816	99,3727	1,008140
30	92,0000	100,5500	91,4968	90,0212	102,1982	100,8103	1,013767
31	96,7000			96,2948	100,4208	101,7262	0,987168
32	134,9000			131,5332	102,5596	102,1842	1,003675

Рис. 10.5.6. **Результат мультиплікативної сезонної декомпозиції**

Отримані результати необхідно скопіювати в таблицю з вихідними даними, користуючись контекстним меню (*Copy with Headers/Paste*) для подальшого їх аналізу та побудови графіків.

Для побудови графіка трендово-циклічної складової необхідно, перебуваючи в таблиці вихідних даних, вибрати в пункті меню *Graphs* підпункт *Scatterplots(точкові графіки)*. Під час задавання характеристик графіка вибрати *Graph type – Regular* (*простий*), поставити оцінку *Linear* *fit* (*Лінійний тренд*), у якості змінних, відображених на графіку, вибрати по осі *X період часу (t)*, а по осі *Y – Smoothed trend cycle (TC)*. Результат наведений на рис. 10.5.7.

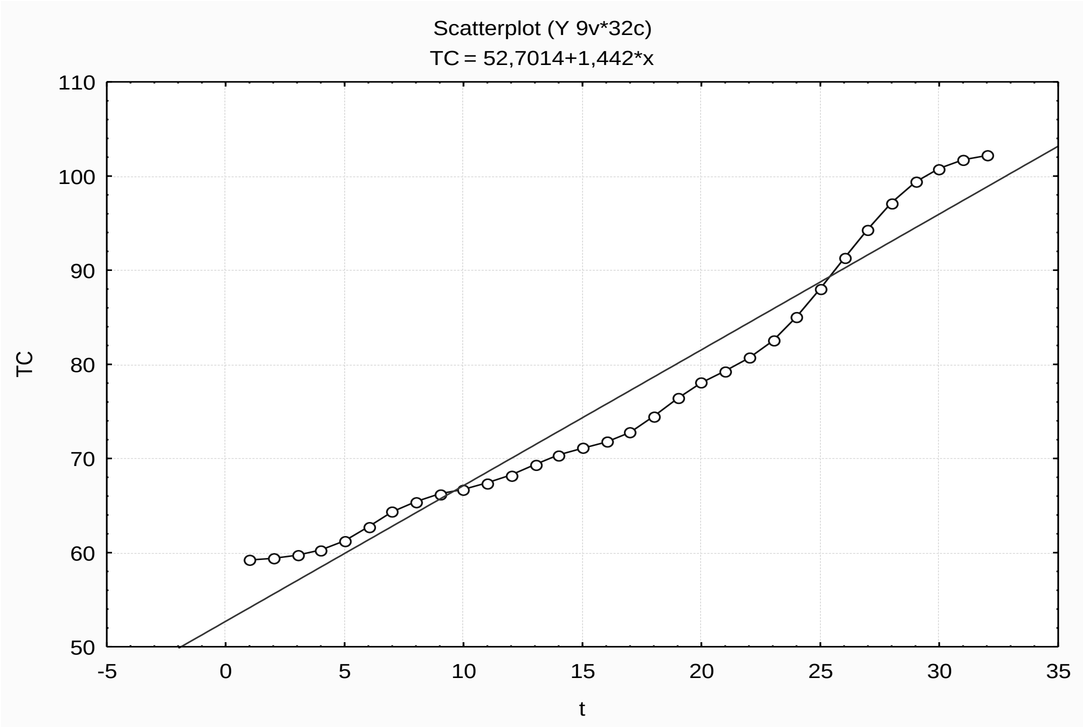


Рис. 10.5.7. **Графік трендово-циклічної складової**

Наступним етапом аналізу є визначення виду тренда та його виокремлення із трендово-циклічної складової. Із аналізу графіка трендово-циклічної компоненти можна передбачити існування лінійного тренда. Його виокремлення здійснюється шляхом побудови лінійної однофакторної моделі виду $TC = a_0 + a_1 t$ й оцінювання параметрів тренда за допомогою МНК у модулі *Multiple Regression* (*множинна регресія*) так, як це було описано раніше.

У якості залежної змінної необхідно задати *Smoothed trend cycle (TC)*, а незалежної – *період часу (t)*. Результат побудови моделі тренда наведено на рис. 10.5.8.

	Regression Summary for Dependent Variable: TC (Y) R= ,97368890 R?= ,94807008 Adjusted R?= ,94633909 F(1,30)=547,70 p<0,0000 Std.Error of estimate: 3,2181					
N=32	Beta	Std.Err. of Beta	B	Std.Err. of B	t(30)	p-level
Intercept			52,70136	1,164974	45,23822	0,000000
t	0,973689	0,041605	1,44195	0,061614	23,40303	0,000000

Рис. 10.5.8. **Оцінювання параметрів лінійного тренда**

Таким чином, модель тренда буде мати такий вигляд:

$$T = 52{,}7014 + 1{,}442t.$$

Значення трендової складової (T) знаходяться як теоретичні значення за побудованою моделлю тренда (ініціювавши клавішу аналізу помилок *Summary: Residuals & Predicted,*теоретичні значення залежної змінної *(Predicted value)*). Графік тренда будується аналогічно до трендово-циклічної складової. Результат побудови наведений на рис. 10.5.9.

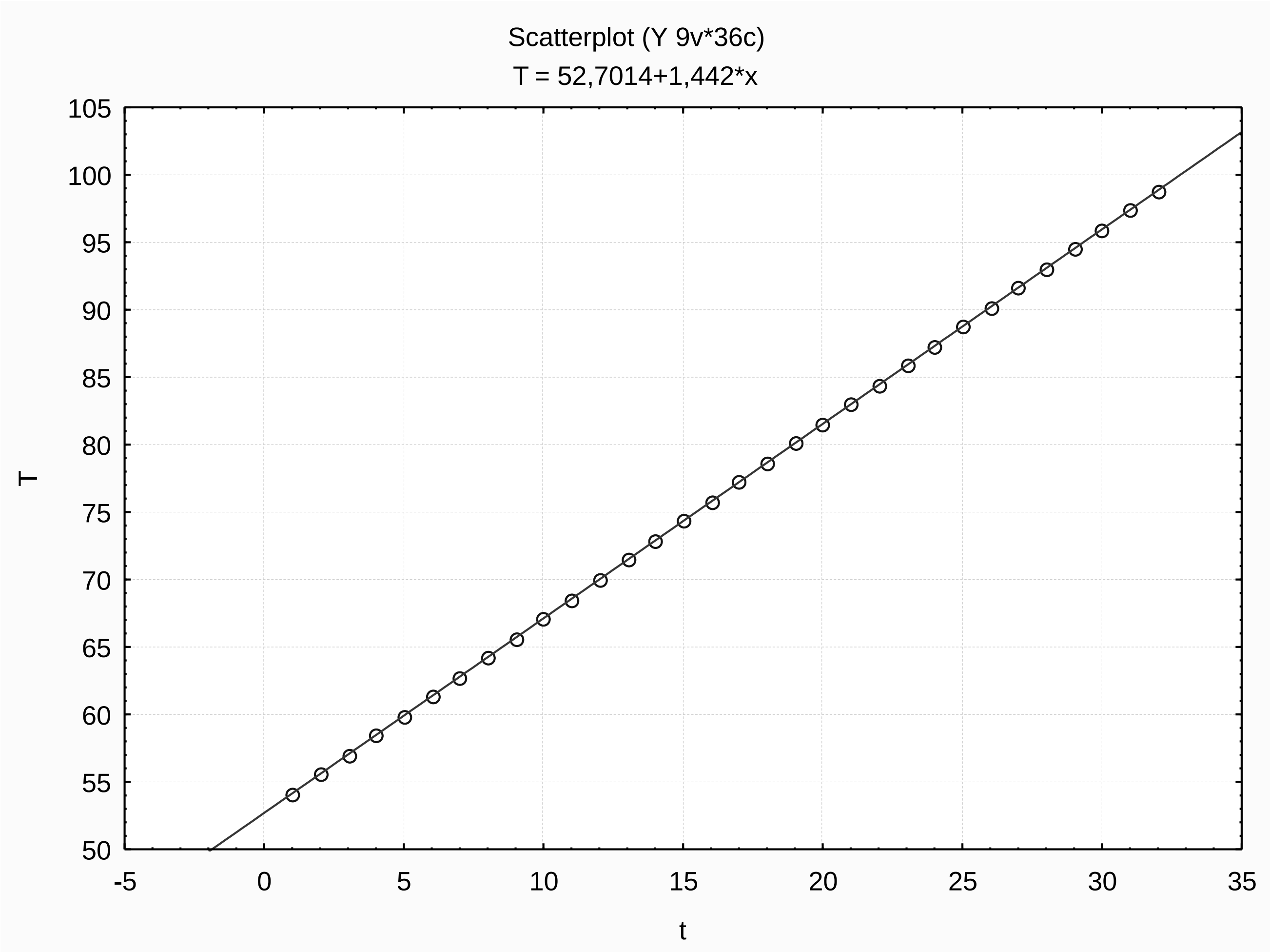


Рис. 10.5.9. **Графік трендової складової**

Розрахунок значень циклічної складової (*C*) можливий шляхом задавання формули розрахунку *C = TC/T* у області специфікації змінної *Long name* (рис. 10.5.10). Графік циклічної складової наведений на рис. 10.5.11.

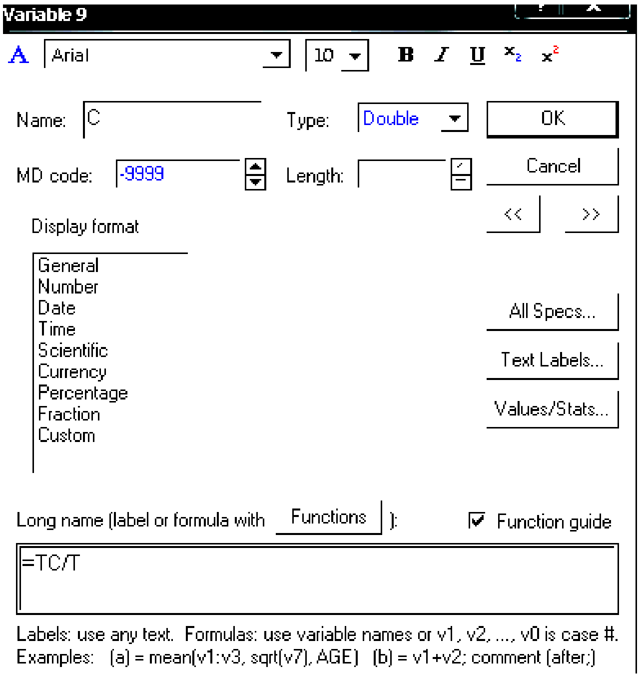


Рис. 10.5.10. **Розрахунок значень циклічної складової**

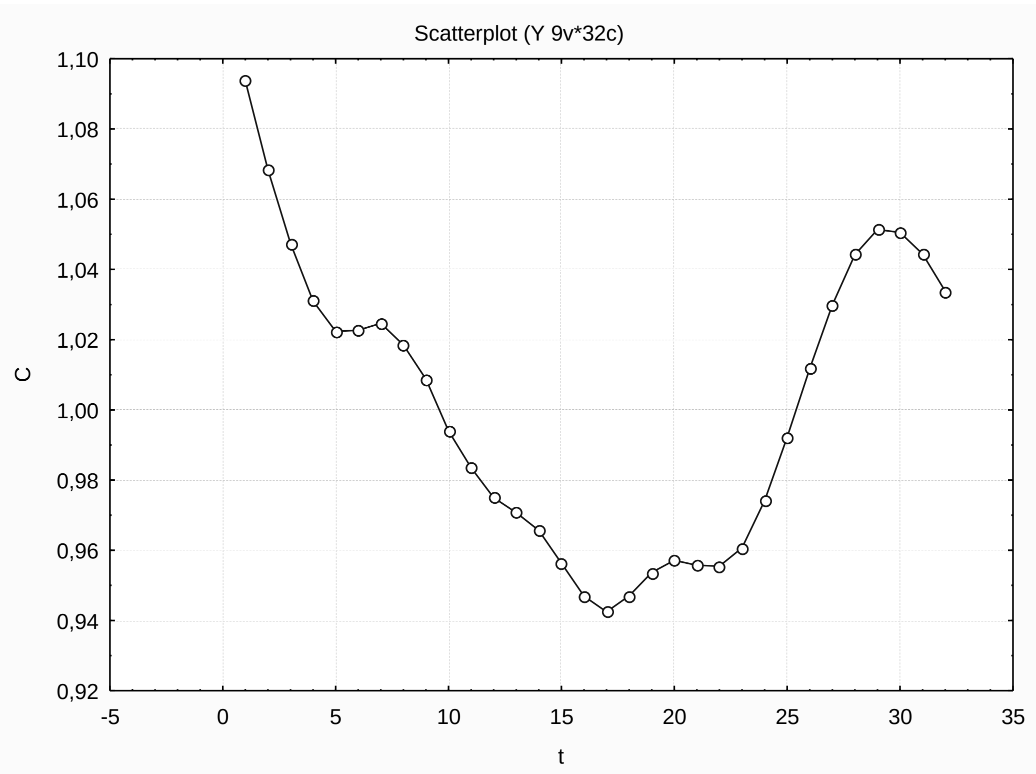


Рис. 10.5.11. **Графік циклічної складової**

Проаналізувавши графік циклічної складової, можна зробити висновок, що циклічна складова має період, рівний 28, це необхідно для подальшого прогнозування поведінки компонент.

Графіки сезонної складової (*Seasonal factors*) і випадкової складової (*Irregular component*) наведені на рис. 10.5.12, 10.5.13.

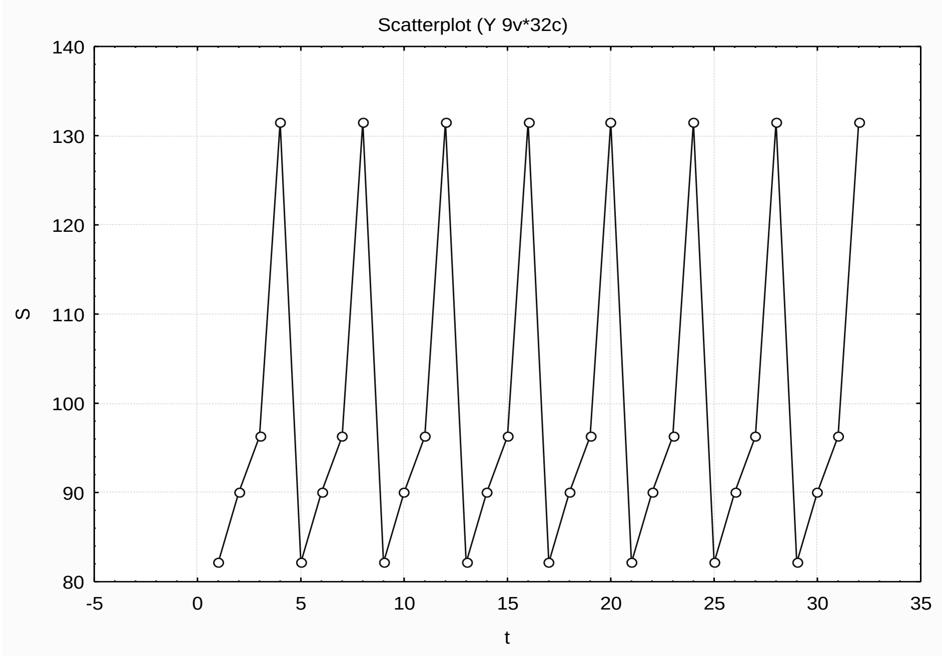


Рис. 10.5.12. **Графік сезонної складової**

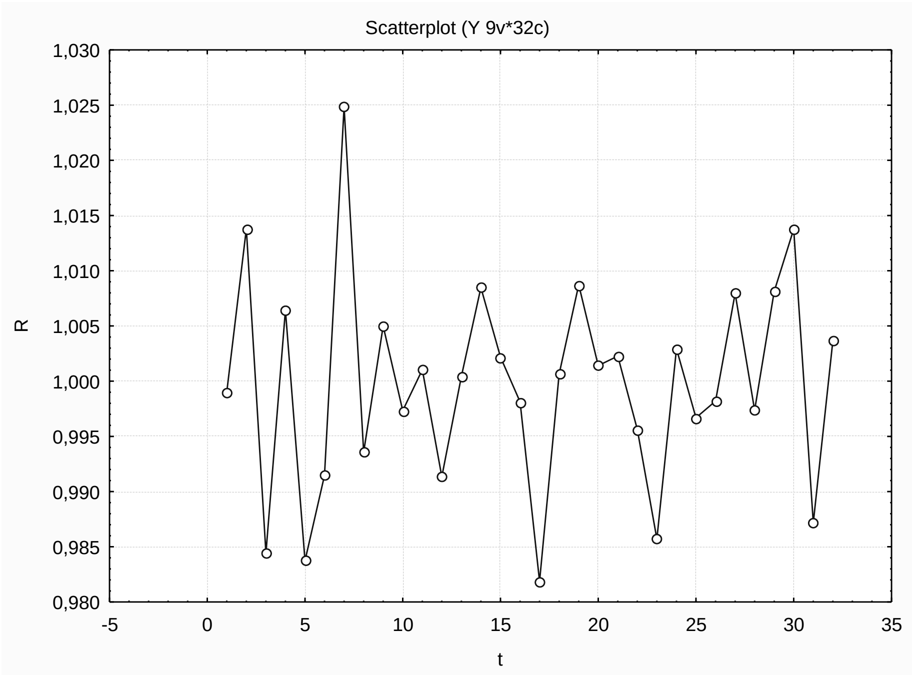


Рис. 10.5.13. **Графік випадкової складової**

Для перевірки випадкової компоненти на нормальний закон розподілу необхідно виділити стовпець *Irregular component (R)*, натискаючи праву кнопку, викликати контекстне меню, вибрати *Graphs of Input Data/ Histogram (R)/NormalFit*та *Graphs of Input Data*/*Probability Plot (R)*/*Normal Probability*, як це подано на рис. 10.5.14. Гістограма випадкової компоненти та графік випадкової компоненти на нормальному ймовірнісному папері подано на рис. 10.5.15, 10.5.16.

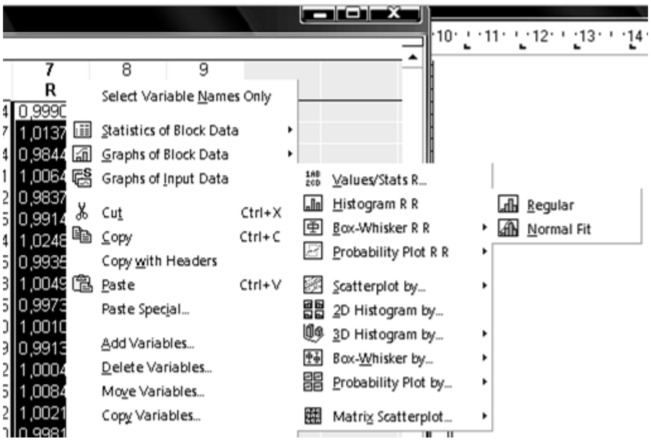


Рис. 10.5.14. **Побудова графіків аналізу випадкової компоненти**

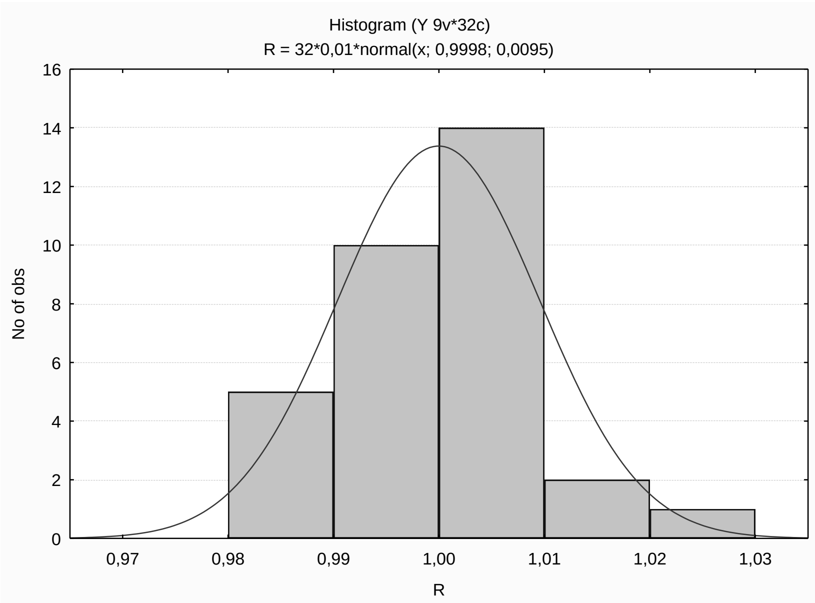


Рис. 10.5.15. **Гістограма випадкової компоненти**

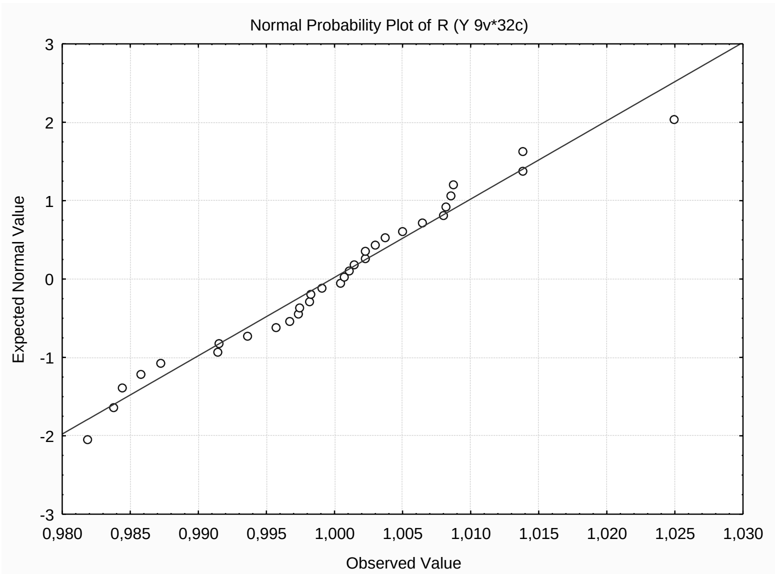


Рис. 10.5.16. **Графік випадкової компоненти на нормальному ймовірнісному папері**

На рис. 10.5.17 наведені всі компоненти моделі мультиплікативної декомпозиції, теоретичні значення об'ємів продажів та помилки за даною моделлю. На рис. 10.5.18 наведено графік порівняння вихідних даних і теоретичних значень об'ємів продажів.

4 Y	5 S	6 TC	7 R	8 TREND	9 C	10 Y_teor	11 E
48,6	82,1508	59,2184	0,999006	54,14331	1,093734	48,64838	-0,04838
54,2	90,0212	59,3887	1,013796	55,58526	1,068426	53,46243	0,737568
56,62	96,2948	59,7294	0,984417	57,02721	1,047384	57,51629	-0,89629
79,8	131,5332	60,2811	1,006436	58,46916	1,03099	79,28971	0,51029
49,5	82,1508	61,2512	0,983737	59,91111	1,022367	50,31833	-0,81833
56	90,0212	62,7435	0,991459	61,35306	1,022662	56,48241	-0,48241
63,5	96,2948	64,3434	1,024866	62,79501	1,024657	61,95929	1,540708
85,5	131,5332	65,4235	0,993566	64,23696	1,018472	86,05365	-0,55365
54,7	82,1508	66,2538	1,004997	65,67892	1,008752	54,42802	0,271983
59,9	90,0212	66,7175	0,997338	67,12086	0,993991	60,0599	-0,1599
65	96,2948	67,4310	1,001038	68,56281	0,983493	64,93257	0,067428
89	131,5332	68,2539	0,991350	70,00477	0,974989	89,77654	-0,77654
57	82,1508	69,3542	1,000438	71,44672	0,970712	56,97506	0,024936
63,9	90,0212	70,3865	1,008479	72,88866	0,965672	63,36278	0,537223
68,6	96,2948	71,0842	1,002186	74,33062	0,956324	68,45035	0,149651
94,2	131,5332	71,7510	0,998131	75,77257	0,946926	94,3764	-0,1764
58,7	82,1508	72,7740	0,981860	77,21452	0,942492	59,78446	-1,08446
67,1	90,0212	74,4893	1,000653	78,65647	0,947021	67,0562	0,043805
74,2	96,2948	76,3882	1,008731	80,09842	0,953679	73,5578	0,6422
102,8	131,5332	78,0449	1,001413	81,54037	0,957132	102,6549	0,145057
65,3	82,1508	79,3079	1,002270	82,98232	0,95572	65,15207	0,147926
72,3	90,0212	80,6668	0,995631	84,42427	0,955494	72,61726	-0,31726
78,3	96,2948	82,4916	0,985711	85,86622	0,960699	79,43505	-1,13505
112,2	131,5332	85,0517	1,002939	87,30817	0,974155	111,8713	0,328737
72,1	82,1508	88,0574	0,996684	88,75012	0,992194	72,33985	-0,23985
82	90,0212	91,2538	0,998201	90,19207	1,011772	82,14777	-0,14777
91,6	96,2948	94,3742	1,007952	91,63402	1,029903	90,87738	0,72262
127,5	131,5332	97,1890	0,997373	93,07597	1,04419	127,8359	-0,33585
82,3	82,1508	99,3727	1,008140	94,51792	1,051363	81,63545	0,664548
92	90,0212	100,8103	1,013767	95,95987	1,050546	90,75063	1,249371
96,7	96,2948	101,7262	0,987168	97,40182	1,044397	97,95702	-1,25702
134,9	131,5332	102,1842	1,003675	98,84377	1,033795	134,4061	0,49388

Рис. 10.5.17. **Значення виділених компонент ряду динаміки**

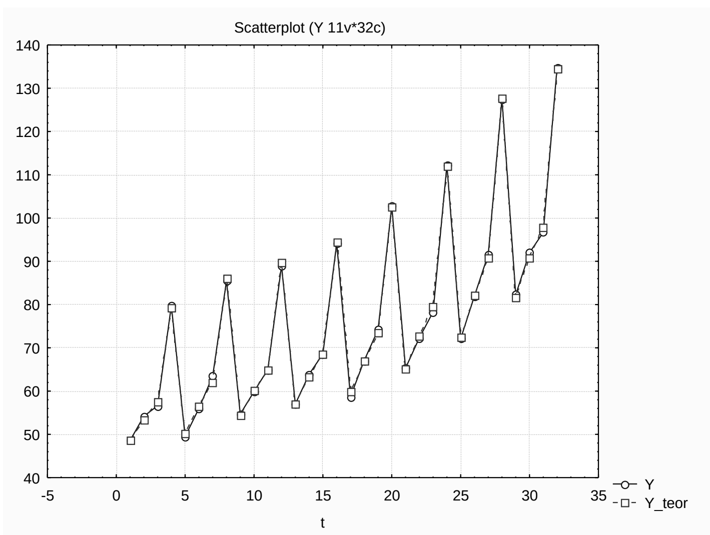


Рис. 10.5.18. **Графік вихідних даних і теоретичних значень *Y***

3. Адитивна декомпозиція часового ряду.

Слід провести декомпозицію часового ряду на такі складові: трендово-циклічну, сезонну й випадкову, використовуючи адитивну модель часового ряду. Ініціювавши дану опцію, буде отримано наступне вікно для вибору типу аналізу (рис. 10.5.19).

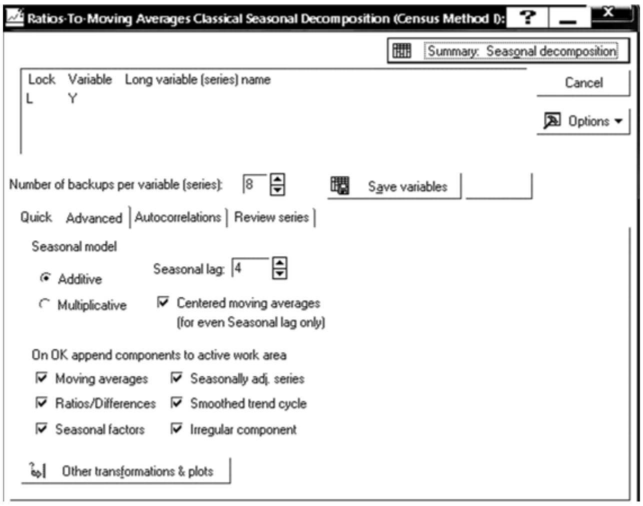


Рис. 10.5.19. **Параметри моделі адитивної сезонної декомпозиції**

На рис. 10.5.20 наведені результати моделі адитивної сезонної декомпозиції.

	Seasonal Decomposition: Additive season (4); Centered means (Y_add) Y						
Case	Y	Moving Averages	Diffrncs	Seasonal Factors	Adjusted Series	Smoothed Trend-c.	Irreg. Compon.
1	48,6000			-13,5618	62,1618	61,6629	0,49891
2	54,2000			-7,6043	61,8043	61,0544	0,74988
3	56,6200	59,9175	-3,2975	-2,5771	59,1971	59,8375	-0,64032
4	79,8000	60,2550	19,5450	23,7432	56,0568	59,7885	-3,73175
5	49,5000	61,3400	-11,8400	-13,5618	63,0618	61,5313	1,53048
6	56,0000	62,9125	-6,9125	-7,6043	63,6043	62,9894	0,61492
7	63,5000	64,2750	-0,7750	-2,5771	66,0771	64,4752	1,60190
8	85,5000	65,4125	20,0875	23,7432	61,7568	65,0063	-3,24952
9	54,7000	66,0875	-11,3875	-13,5618	68,2618	66,3291	1,93270
10	59,9000	66,7125	-6,8125	-7,6043	67,5043	66,8005	0,70381
11	65,0000	67,4375	-2,4375	-2,5771	67,5771	67,4530	0,12413
12	89,0000	68,2250	20,7750	23,7432	65,2568	67,8952	-2,63841
13	57,0000	69,1750	-12,1750	-13,5618	70,5618	69,3291	1,23270
14	63,9000	70,2750	-6,3750	-7,6043	71,5043	70,4116	1,09270
15	68,6000	71,1375	-2,5375	-2,5771	71,1771	71,1419	0,03524
16	94,2000	71,7500	22,4500	23,7432	70,4568	71,6063	-1,14952
17	58,7000	72,8500	-14,1500	-13,5618	72,2618	72,7846	-0,52286
18	67,1000	74,6250	-7,5250	-7,6043	74,7043	74,6338	0,07048
19	74,2000	76,5250	-2,3250	-2,5771	76,7771	76,5530	0,22413
20	102,8000	78,0000	24,8000	23,7432	79,0568	78,1174	0,93937
21	65,3000	79,1625	-13,8625	-13,5618	78,8618	79,1291	-0,26730
22	72,3000	80,8500	-8,5500	-7,6043	79,9043	80,7449	-0,84063
23	78,3000	82,8750	-4,5750	-2,5771	80,8771	82,6530	-1,77587
24	112,2000	84,9375	27,2625	23,7432	88,4568	85,3285	3,12825
25	72,1000	87,8125	-15,7125	-13,5618	85,6618	87,5735	-1,91175
26	82,0000	91,3875	-9,3875	-7,6043	89,6043	91,1894	-1,58508
27	91,6000	94,5750	-2,9750	-2,5771	94,1771	94,5308	-0,35365
28	127,5000	97,1000	30,4000	23,7432	103,7568	97,8396	5,91714
29	82,3000	98,9875	-16,6875	-13,5618	95,8618	98,6402	-2,77841
30	92,0000	100,5500	-8,5500	-7,6043	99,6043	100,4449	-0,84063
31	96,7000			-2,5771	99,2771	103,3461	-4,06893
32	134,9000			23,7432	111,1568	104,7966	6,36014

Рис. 10.5.20. **Результат адитивної сезонної декомпозиції**

Аналіз побудованої моделі, виділення трендової та циклічної складової, побудову графіків основних компонент слід провести аналогічно, як це було розглянуто для мультиплікативної моделі.

На рис. 10.5.21 наведено графік виокремленої циклічної складової адитивної моделі.

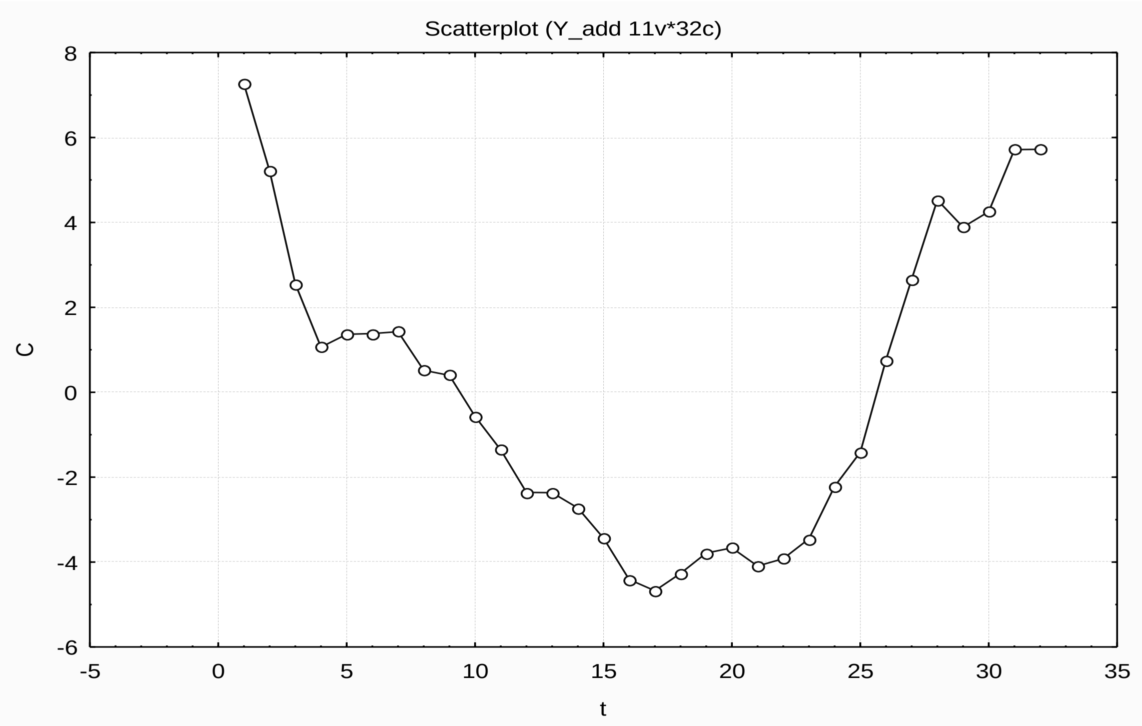


Рис. 10.5.21. **Графік циклічної складової адитивної моделі**

4. Прогнозування.

Для прогнозування динаміки об'ємів продажів необхідно вибрати з побудованих моделей найбільш адекватну. Оцінка якості моделей часових рядів визначається за такими критеріями адекватності: середня помилка, середня абсолютна помилка, среднє квадратичне відхилення помилок, середня відсоткова помилка, середня абсолютна відсоткова помилка. Порівняння даних моделей за виокремленими критеріями наведено в табл. 10.5.1.

Таблиця 10.5.1

Оцінка якості моделей

Критерії якості прогнозу	Мультиплікативна модель	Адитивна модель
1	2	3
Середня помилка	-0,00473	0,012569
Середня абсолютна помилка	0,522097	1,659735

Закінчення табл. 10.5.1

1	2	3
Сума квадратів помилок	13,9139	167,9041
Середньоквадратична помилка	0,659401	2,290634
Середня відсоткова помилка	-0,02942	-0,01816
Середня абсолютна відсоткова помилка	0,71955	1,9813

Отже, на основі оцінки та аналізу моделі декомпозиції часового ряду можна зробити висновок, що побудовані моделі є адекватними та забезпечують високу точність прогнозу. Але мультиплікативна модель є найбільш адекватною, тому для неї і буде побудовано прогноз.

Перш ніж здійснити прогнозування величини об'єму продажів на 2 роки (8 кварталів) уперед за допомогою моделі декомпозиції часового ряду, необхідно виконати ряд дій:

1) додати 8 нових спостережень після останнього;

2) у стовпці даних *Період часу(t1)*вписати відповідні порядкові числівники (продовжуючи ряд);

3) у стовпці *Seasonal Factors (S_pr)* вписати відповідні значення сезонних індексів для відповідних номерів кварталу року;

4) у стовпці прогнозування циклічної компоненти *(C_pr)* вписати відповідні значення циклічної складової з урахуванням періоду циклу, рівного 28, наприклад, *C(33) = C(33 – 28) = C(5)* і так на вісім кроків уперед*;*

5) у стовпці трендової компоненти *(TREND_pr)* задати перерахування даних (*Vars/Recalculate*) відповідно до моделі отриманого тренда $T = 52{,}7014 + 1{,}442 \cdot t$;

6) додати нову змінну (*Y_pr),* де будуть розраховуватися прогнозні значення об'єму продажів.

Тоді розрахувати прогнозні значення показника об'єму продажів на 8 кварталів уперед можливо шляхом завдання в області специфікації змінної *Long Name*, як це було описано раніше, у моделі такого вигляду:

$$Y_{pr} = (S_{pr}/100) \cdot TREND_{pr} \cdot C_{pr}.$$

На рис. 10.5.22 подано графік динаміки прогнозних значень об'ємів продажів.

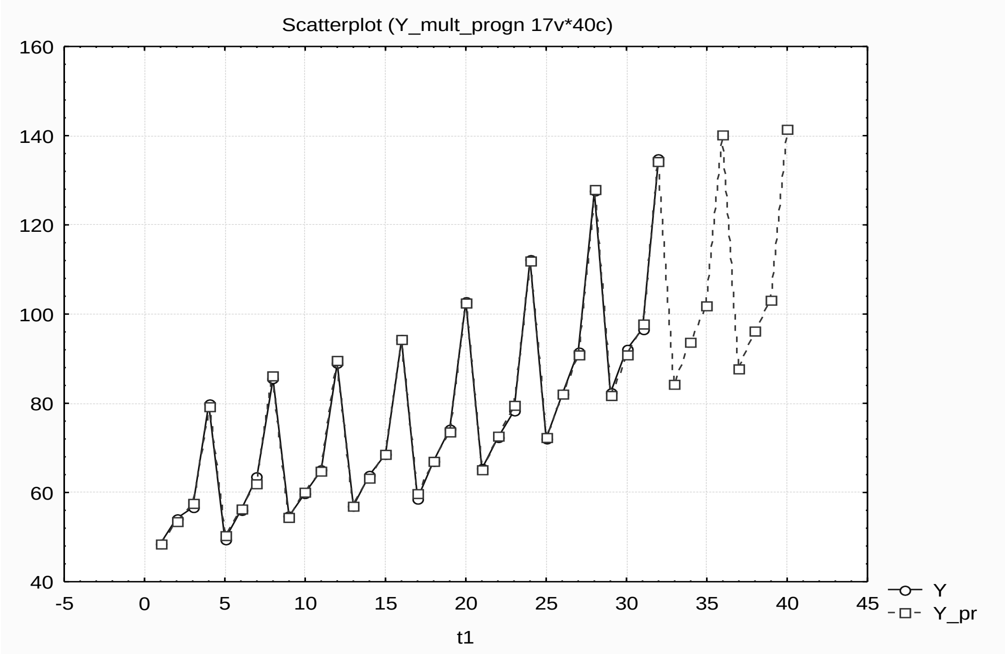


Рис. 10.5.22. **Графік прогнозних значень об'ємів продажів**

Результати розрахунку прогнозних значень показника об'єму продажів наведено на рис. 10.5.23.

12 t1	13 S_pr	14 TREND_pr	15 C_pr	16 Y_pr
1	82,1508	54,1434	1,093734	48,64846
2	90,0212	55,5854	1,068426	53,46257
3	96,2948	57,0274	1,047384	57,51648
4	131,5332	58,4694	1,03099	79,29004
5	82,1508	59,9114	1,022367	50,31857
6	90,0212	61,3534	1,022662	56,48273
7	96,2948	62,7954	1,024657	61,95968
8	131,5332	64,2374	1,018472	86,05424
....				
30	90,0212	95,9614	1,050540	90,75208
31	96,2948	97,4034	1,044397	97,95861
32	131,5332	98,8454	1,033795	134,4083
33	82,1508	100,2874	1,022367	84,22969
34	90,0212	101,7294	1,022662	93,65339
35	96,2948	103,1714	1,024657	101,7983
36	131,5332	104,6134	1,018472	140,1431
37	82,1508	106,0554	1,008752	87,88794
38	90,0212	107,4974	0,993991	96,18892
39	96,2948	108,9394	0,983493	103,1713
40	131,5332	110,3814	0,974989	141,5569

Рис. 10.5.23. **Фрагмент таблиці розрахунку прогнозних значень**

Отже, на основі оцінки та аналізу моделей декомпозиції часового ряду, можна зробити висновок, що побудована модель є адекватною та забезпечує високу точність прогнозу.

Лабораторна робота 6. Побудова моделі розподіленого лага

Мета – закріплення теоретичного й практичного матеріалу за темою "Моделі розподіленого лага", набуття навичок побудови й аналізу економетричних лагових моделей у модулі *Distributed Lags Analysis*.

Завдання – побудувати моделі розподілених лагів за даними часових рядів капітальних вкладень та чистої продукції у модулі *Distributed Lags Analysis* ППП *Statistica*:

1. Визначити порядок лагової моделі.

2. Побудувати лагову модель методом Ширлі Алмон, визначити всі її характеристики і провести оцінювання її адекватності.

3. Навести порівняльну характеристику моделей з різним числом лагів запізнення.

4. Побудувати лагову модель методом МНК, визначити всі її характеристики і провести оцінювання її адекватності. Результати порівняти з моделлю Алмон.

5. За найбільш адекватною моделлю побудувати прогнози та їх графіки.

6. Дати економічну інтерпретацію результатів моделювання.

Для побудови моделей розподіленого лага в *ППП Statistica* є модуль *Time Series/Forecasting/Distributed Lags Analysis* (часові ряди/прогнозування/аналіз розподілених лагів). Розглянемо порядок розрахункових процедур для побудови лагових моделей.

1. Підготовка даних і запуск процедури розрахунків.

Вихідні дані для побудови моделі наведені на рис. 10.6.1.

	1 X t	2 Y t
1	18,4	11,5
2	20,7	20,7
3	23	20,7
4	20,7	18,4
5	21,9	20,7
6	20,7	23
7	23	27,6
8	24,2	27,6
9	25,3	23
10	26,5	24,2
11	25,3	25,3
12	26,5	28,8
13	27,6	31,1
14	29,9	31,1
15	31,1	34,5
16	32,2	32,2
17	34,5	36,8
18	35,7	36,8
19	35,7	34,5
20	34,5	36,8

Рис. 10.6.1. **Вихідні дані**

На рис. 10.6.1 X_t – капітальні вкладення; Y_t – чиста продукція досліджуваного об'єкта.

Щоб приступити до обчислювальних процедур, необхідно ввійти в позицію меню *Statistics/ Advanced Linear – Nonlinear Models/Time Series/ Forecasting* (рис. 10.6.2).

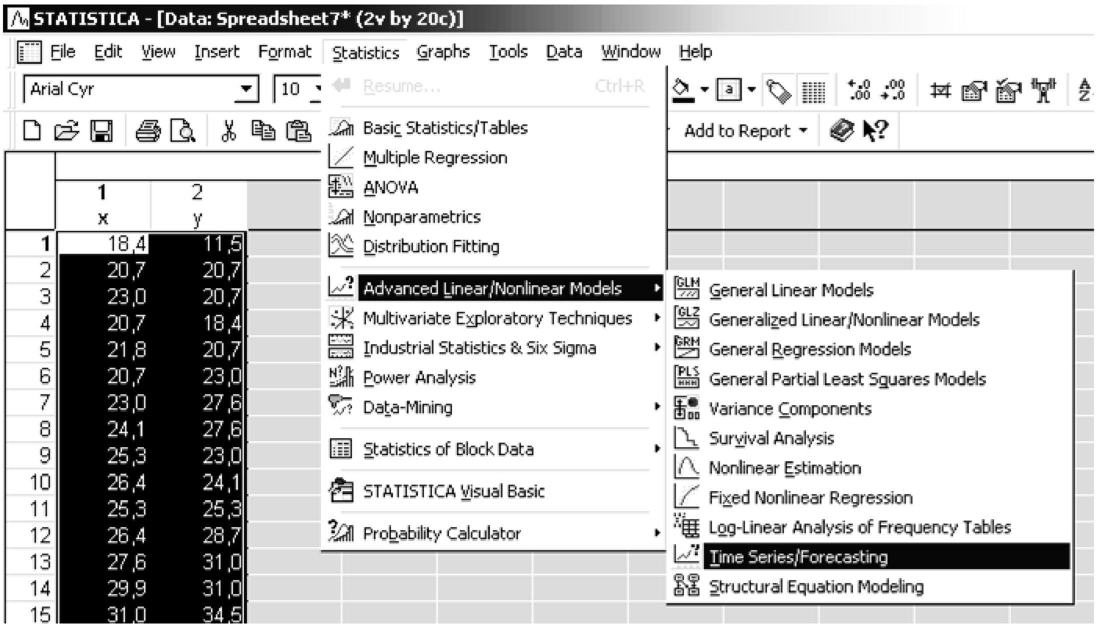


Рис. 10.6.2. **Вибір модуля *Time Series/Forecasting***

Після підтвердження вибору модуля з'явиться стартова панель модуля (рис. 10.6.3).

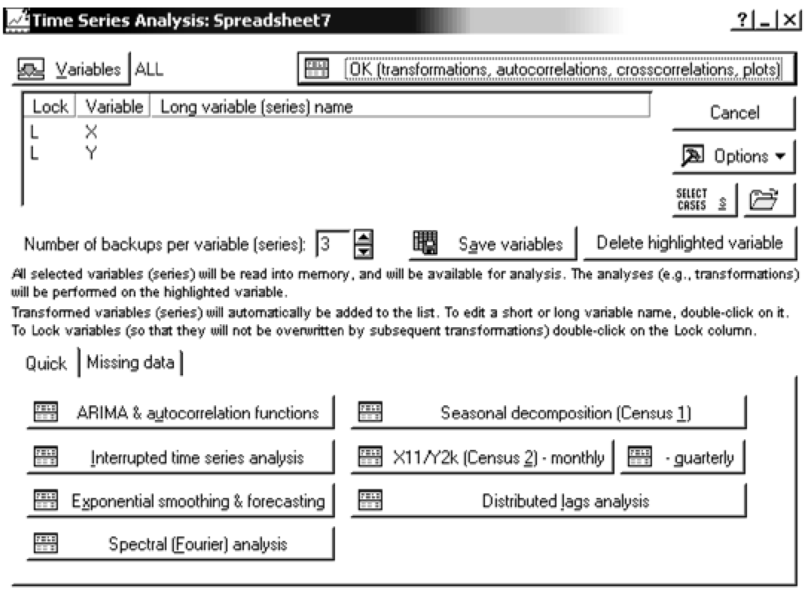


Рис. 10.6.3. **Стартова панель модуля *Time Series/Forecasting***

На стартовій панелі модуля можливо здійснити вибір типу досліджуваних динамічних моделей. Ініціювавши модуль Distributed Lags Analysis (аналіз розподілених лагів) на екрані з'явиться стартова панель вибраного модуля (рис. 10.6.4).

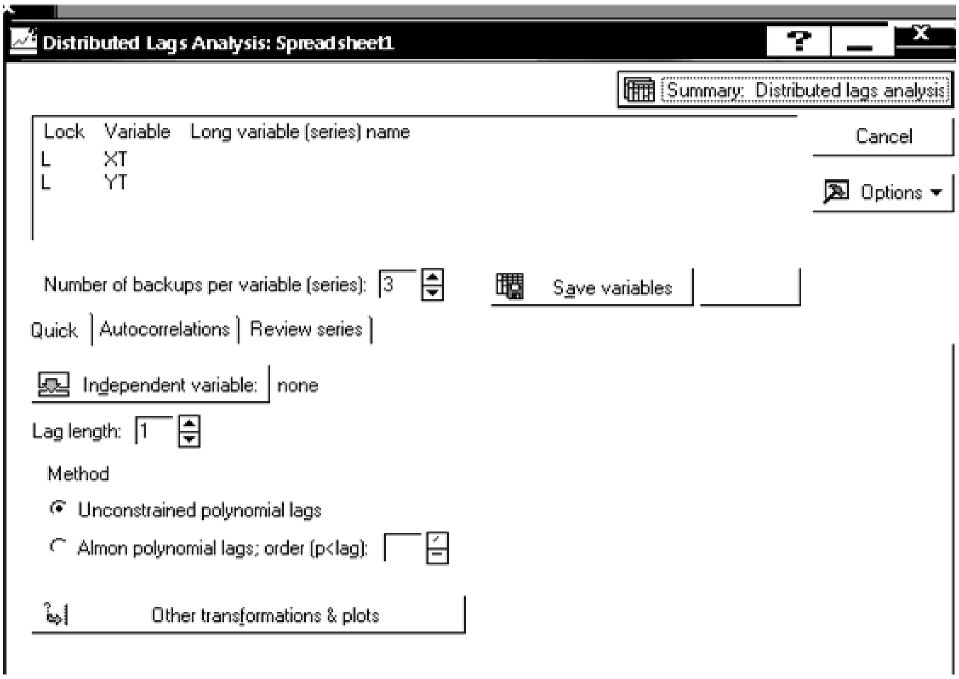


Рис. 10.6.4. **Стартова панель модуля *Distributed Lags Analysis***

На рис. 10.6.5 наведено діалог вибору залежної та незалежної змінної (залежна змінна обирається у верхній частині діалогового вікна, а незалежна – ініціюванням клавіші *Independent variable*).

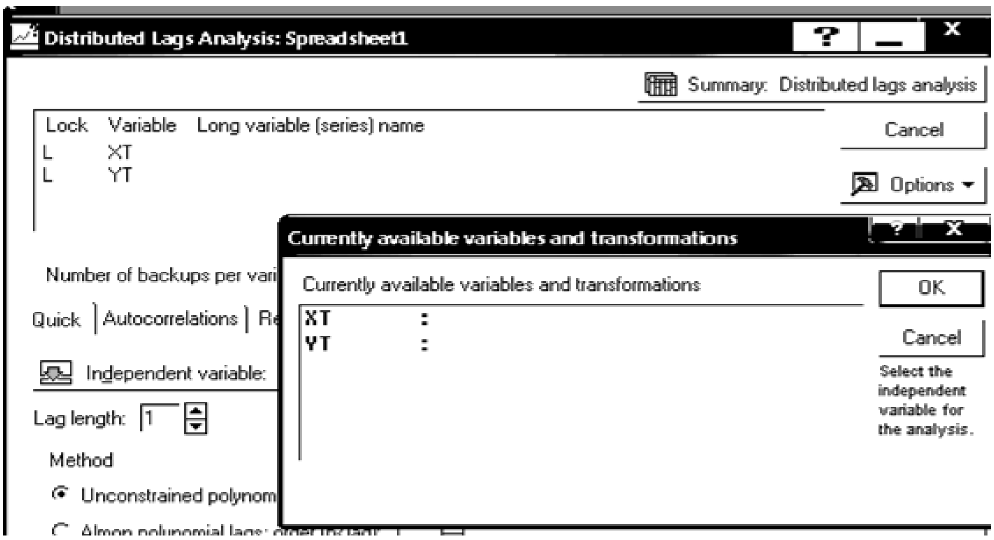


Рис. 10.6.5. **Вибір змінних для аналізу**

2. Визначення порядку лагової моделі.

Для обґрунтування величини лага запізнення доцільно використовувати взаємну кореляційну функцію. Ця функція характеризує тісноту зв'язку кожного елемента вектора залежної змінної Y_t зелементом вектора пояснювальної змінної X_t , зсунутим один відносно одного на часовий лаг. Вихідні дані для визначення взаємозв'язку між економічними даними (капітальні вкладення та чиста продукція) подані на рис. 10.6.6.

	1 X t	2 Y t	3 X_1	4 Y_1	5 X_2	6 Y_2	7 X_3	8 Y_3
1	18,4	11,5	18,4	20,7	18,4	20,7	18,4	18,4
2	20,7	20,7	20,7	20,7	20,7	18,4	20,7	20,7
3	23	20,7	23	18,4	23	20,7	23	23
4	20,7	18,4	20,7	20,7	20,7	23	20,7	27,6
5	21,9	20,7	21,9	23	21,9	27,6	21,9	27,6
6	20,7	23	20,7	27,6	20,7	27,6	20,7	23
7	23	27,6	23	27,6	23	23	23	24,2
8	24,2	27,6	24,2	23	24,2	24,2	24,2	25,3
9	25,3	23	25,3	24,2	25,3	25,3	25,3	28,8
10	26,5	24,2	26,5	25,3	26,5	28,8	26,5	31,1
11	25,3	25,3	25,3	28,8	25,3	31,1	25,3	31,1
12	26,5	28,8	26,5	31,1	26,5	31,1	26,5	34,5
13	27,6	31,1	27,6	31,1	27,6	34,5	27,6	32,2
14	29,9	31,1	29,9	34,5	29,9	32,2	29,9	36,8
15	31,1	34,5	31,1	32,2	31,1	36,8	31,1	36,8
16	32,2	32,2	32,2	36,8	32,2	36,8	32,2	34,5
17	34,5	36,8	34,5	36,8	34,5	34,5	34,5	36,8
18	35,7	36,8	35,7	34,5	35,7	36,8		
19	35,7	34,5	35,7	36,8				
20	34,5	36,8						

Рис. 10.6.6. **Дані з запізненням (лагами)**

Для побудови кореляційної матриці необхідно зайти в модуль *Basic Statistics/Tables* і вибрати напрям аналізу *Correlation matrices* (рис. 10.6.7).

На рис. 10.6.8 подана матриця кореляцій між відповідними часовими лагами.

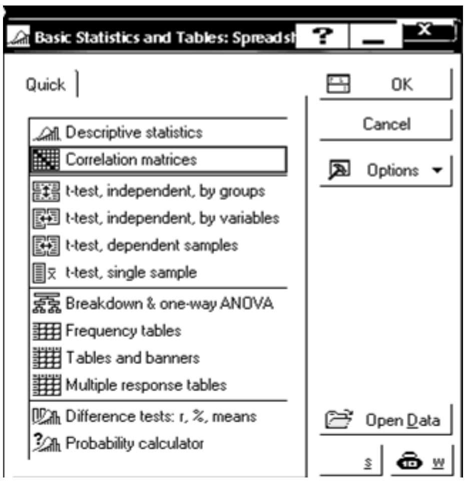


Рис. 10.6.7. **Розрахунок кореляційної матриці**

Correlations (Spreadsheet1)
Marked correlations are significant at $p < ,05$
N=17 (Casewise deletion of missing data)

Variable	X t	Y t	X_1	Y_1	X_2	Y_2	X_3	Y_3
X t	1,00	**0,91**	1,00	0,87	1,00	0,86	1,00	0,90
Y t	0,91	1,00	0,91	0,87	0,91	0,81	0,91	0,83
X_1	1,00	0,91	1,00	**0,87**	1,00	0,86	1,00	0,90
Y_1	0,87	0,87	0,87	1,00	0,87	0,90	0,87	0,83
X_2	1,00	0,91	1,00	0,87	1,00	**0,86**	1,00	0,90
Y_2	0,86	0,81	0,86	0,90	0,86	1,00	0,86	0,90
X_3	1,00	0,91	1,00	0,87	1,00	0,86	1,00	**0,90**
Y_3	0,90	0,83	0,90	0,83	0,90	0,90	0,90	1,00

Рис. 10.6.8. **Кореляційна матриця**

Найбільше значення коефіцієнта кореляції за модулем визначає зрушення або часовий лаг. Якщо серед множини значень є кілька величин, які наближаються до одиниці, то це означає, що запізнення впливу змінної X_t відбувається протягом певного проміжку часу і в результаті наявно кілька часових лагів для двох взаємопов'язаних часових рядів. Визначивши величину часового лага, можна будувати економетричну модель розподіленого лага.

3. Побудова лагової моделі і оцінювання її адекватності.

У групі опцій *Method (Метод)* необхідно вибрати метод оцінювання регресійних коефіцієнтів.

Оскільки під час оцінки параметрів моделі розподіленого лага, як правило, виникає проблема мультиколінеарності, варто скористатися методом *Алмон (Almon)* і вибрати опцію *Almon Polynomial lags (поліноміальні лаги Алмон)* у групі опцій *Method* (рис. 10.6.9). Задати порядок полінома *(order p < lag)* рівним 2.

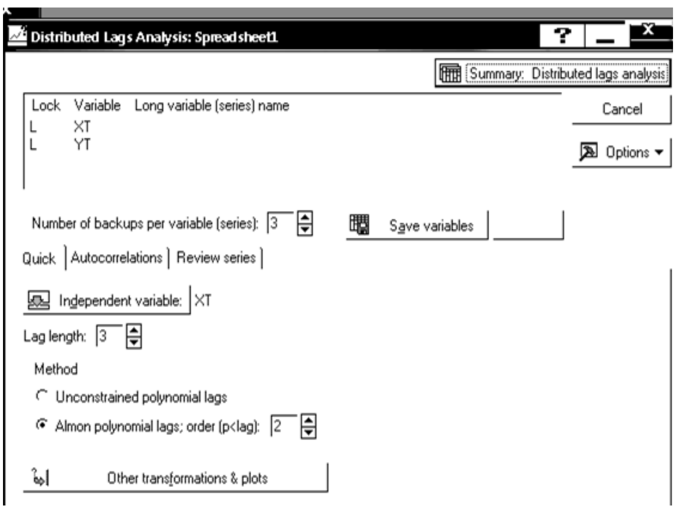


Рис. 10.6.9. **Завдання параметрів для побудови лагової моделі**

Опція *Unconstrained Polynomial lags (звичайні лаги)*проводить оцінювання без використання поліноміальних лагів.

Опція *Lag Length (довжина лага)* дозволяє задати величину зрушення одного ряду щодо іншого. Задати в цій опції число 3.

Група опцій *Review and plot variables (перегляд змінних й побудува графіків)* дозволяє всебічно переглянути дані в електронних таблицях і побудувати графіки, зокрема, побудувати графіки двох змінних зі списку в різних шкалах.

Зробити установки, як подано на рис. 10.6.9. У якості залежної змінної обрана змінна Y_t (чиста продукція). Вона висвітлена в інформаційному полі у верхній частині панелі. Натиснути кнопку *Summary: Distributed Lags Analysis (підсумки: Аналіз розподілених лагів).* У таблицях, що з'явилися на екрані, подано всебічний аналіз даних моделей.

На рис. 10.6.10 наведені критерії якості лагової моделі: коефіцієнт множинної кореляції (*R*), коефіцієнт детермінації (*R-square*), оцінки коефіцієнтів регресії зі штучними змінними *(alpha coeff.)*, їх стандартні помилки *(StandartError)*, значущість за критерієм Стьюдента *(t)* та рівень значущості (*p*). Аналіз значень критерію Стьюдента моделі зі штучними змінними дозволяє зробити висновок про статистичну значущість її параметрів, що дає можливість перейти до коефіцієнтів моделі розподіленого лага.

	Almon Polyn. Distr.Lags; Alpha Coefficients (Spreadsheet1) Indep: XT Dep: YT Lag: 3 Polyn. order: 2 R=,9988 R-square=,9977 N:17			
poly- nomial	Alpha Coeff.	Standard Error	t(14)	p
0	1,49875196617	0,274133891175	5,46722610526	0,000082957944
1	-3,40404866901	0,640618312944	-5,31369241283	0,000109417255
2	1,10818791136	0,211095125100	5,24970868388	0,000122912251

Рис. 10.6.10. **Оцінки коефіцієнтів регресії зі штучними змінними**

Результати дисперсійного аналізу для досліджуваної моделі наведені на рис. 10.6.11. У даній таблиці наведено суму квадратів відхилень за регресією *(Sums of Squares Regress),* суму квадратів відхилень похибок *(SumsofSquaresResidual),* дисперсію похибок *(Mean Squares Residual)* та критерій Фішера *(F)*.

Effect	Almon Polyn. Distr.Lags; Analysis of Varianc Indep: XT Dep: YT Lag: 3 Polyn. order: 2 R=,9988 R-square=,99				
	Sums of Squares	df	Mean Square	F	p
Regress.	14783,87	4	3695,966	1398,552	0,000000
Residual	34,36	13	2,643		
Total	14818,22				

Рис. 10.6.11. **Таблиця дисперсійного аналізу**

На рис. 10.6.12 подані оцінки коефіцієнтів моделі розподіленого лага з усіма характеристиками. Таким чином, лагова модель, побудована методом Алмон, має вигляд:

$$y_t = 1{,}498752x_t - 0.797109x_{t-1} - 0{,}876594x_{t-2} + 1{,}260297x_{t-3}.$$

Lag	Almon Polyn. Distr.Lags; Regression Coefficients (Spreadsheet1) Indep: XT Dep: YT Lag: 3 Polyn. order: 2 R=,9988 R-square=,9977 N:17			
	Regressn Coeff.	Standard Error	t(13)	p
0	1,498751966170	0,274133891175	5,46722610526	0,000108002889
1	-0,797108791476	0,220312097717	-3,61808906428	0,003121989038
2	-0,876593726402	0,218439062296	-4,01298978849	0,001475555159
3	1,260297161392	0,292123097158	4,31426742238	0,000840759301

Рис. 10.6.12. **Коефіцієнти лагової моделі**

Результати побудови лагової моделі з використанням опції *Unconstrained Polynomiallags (Звичайні лаги)*(оцінювання без використання поліноміальних лагів за звичайним МНК), наведено на рис. 10.6.13, 10.6.14. Побудована модель має вигляд:

$$y_t = 1{,}515922x_t - 0.846824x_{t-1} - 0{,}826284x_{t-2} + 1{,}242502x_{t-3}.$$

Lag	Polyn. Distr. Lags; Regression Coefficients (Spreadsheet1) Indep: XT Dep: YT Lag: 3 R=,9988 R-square=,9977 N:17			
	Regressn Coeff.	Standard Error	t(13)	p
0	1,515922161018	0,304776613626	4,97387953421	0,000254440616
1	-0,846824475286	0,392045485912	-2,16001588009	0,050032259619
2	-0,826283786103	0,394027814929	-2,09701892810	0,056115347134
3	1,242502205024	0,323628920626	3,83928050256	0,002048750800

Рис. 10.6.13. **Коефіцієнти лагової моделі**

	Polyn. Distr. Lags; Analysis of Variance (Spr Indep: XT Dep: YT Lag: 3 R=,9988 R-square=,9977 N:17				
Effect	Sums of Squares	df	Mean Square	F	p
Regress.	14783,93	4	3695,982	1401,177	0,000000
Residual	34,29	13	2,638		
Total	14818,22				

Рис. 10.6.14. **Таблиця дисперсійного аналізу**

4. Інтерпретація результатів моделювання і побудова прогнозів.

Як видно, побудовані моделі за різними методами майже не різняться. Аналіз адекватності моделей дозволяє зробити висновок про наявність лага у впливі капітальних інвестицій на чисту продукцію, рівного 1, 2, 3 рокам. Графік досліджуваних і прогнозних значень чистої продукції для моделі з лагом запізнення, рівного 3 рокам, подано на рис. 10.6.15.

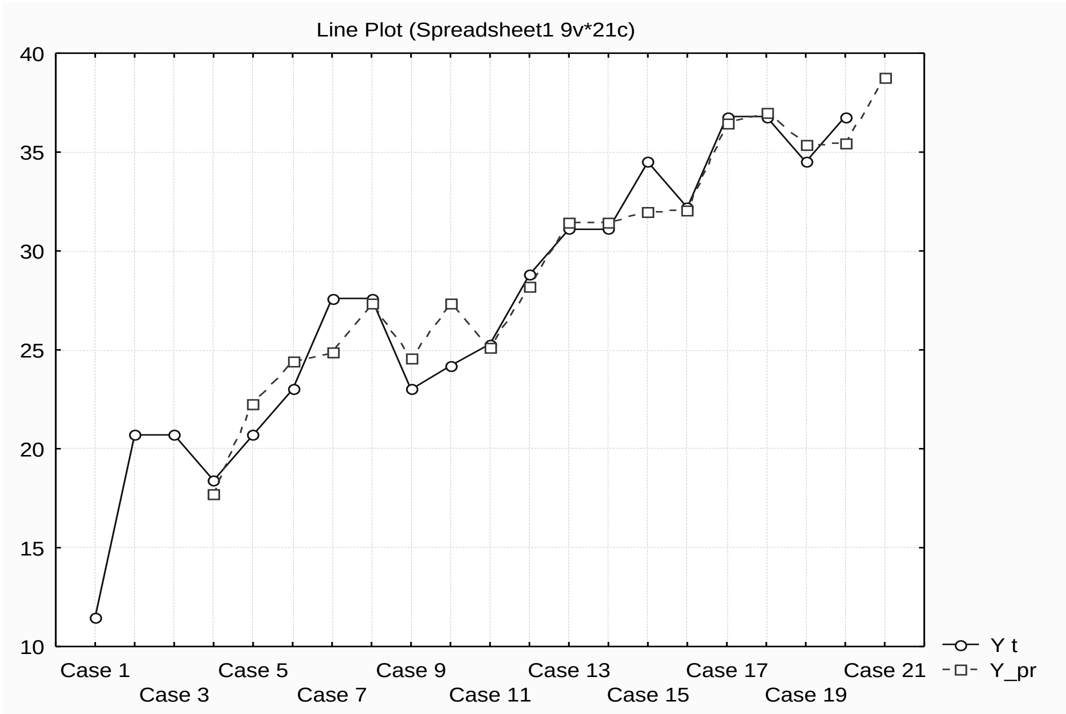


Рис. 10.6.15. **Графік досліджуваних і прогнозних значень чистої продукції**

Поданий графік підтверджує якість побудованої моделі, тобто вплив об'єму капітальних інвестицій на динаміку чистої продукції підприємства є зрушеним у часі. Подальший аналіз передбачає дослідження особливостей часових лагів, а саме розрахунок середнього лага та його медіанного значення для економічної інтерпретації моделей.

Лабораторна робота 7. Побудова й аналіз систем одночасових рівнянь

Мета – закріплення теоретичного й практичного матеріалу за темою "Економетричні моделі на основі системи структурних рівнянь", набуття навичок побудови й аналізу різних видів систем одночасних рівнянь у модулі *Structural equation modeling*.

Завдання – побудувати складні моделі соціально-економічних процесів за допомогою системи структурних рівнянь у модулі *Structural equation modeling* ППП *Statistica*:

1. Дослідити типи причинно-наслідкових зв'язків між досліджуваними змінними.
2. Визначити вид оцінюваної моделі і побудувати діаграму причинно-наслідкових зв'язків.
3. Побудувати вибраний тип моделі.
4. Дослідити якість побудованої моделі.
5. Оцінити параметри моделі, їх характеристики і статистичну значущість.
6. Побудувати рекурентні типи систем структурних рівнянь, вибрати найбільш адекватну й економічно інтерпретовану модель.
7. Розрахувати теоретичні значення залежних змінних і побудувати прогнози.

Для побудови моделей систем одночасових рівнянь у *ППП Statistica* передбачено модуль *Statistics/Advanced Linear/Nonlinear Models/Structural equation modeling* (Моделювання структурними рівняннями). Слід розглянути порядок розрахункових процедур для побудови структурної моделі.

1. Підготовка даних і запуск процедури розрахунків.

Використовуючи результати спостереження, наведені на рис. 10.7.1, де Y1 – експорт, Y2 – імпорт, X1 – середній товарообіг торгівлі країн, з якими підтримуються зовнішньоекономічні відносини, X2 – національний дохід, побудувати модель аналізу макроекономічних показників за допомогою такої системи структурних рівнянь:

$$\begin{gathered} y_1 = b_1 y_2 + a_{10} + a_{11} x_1 + a_{12} x_2 + \varepsilon_1 \\ y_2 = b_2 y_1 + a_{20} + a_{21} x_1 + a_{22} x_2 + \varepsilon_2 \end{gathered}.$$

	1 X1	2 X2	3 Y1	4 Y2
1	565,54	385,54	90,05	67,2
2	569,09	390,05	91,2	67,39
3	573,22	394,08	92,26	71,52
4	577,54	398,78	93,6	75,94
5	581,76	401,47	95,23	78,72
6	586,37	403,39	95,52	82,94
7	590,3	405,5	95,9	86,02
8	594,53	408,48	96,29	90,43
9	598,56	410,4	98,02	91,39
10	602,59	413,47	99,46	93,41
11	606,14	413,95	101,18	97,44
12	610,08	415,39	102,72	100,8

Рис. 10.7.1. **Вихідні дані**

Побудова комплексної економетричної моделі здійснюється в модулі *Structural equation modeling* (моделювання структурними рівняннями). Щоб розпочати обчислювальні процедури, необхідно ввійти в позицію меню *Statistics/ Advanced Linear/Nonlinear Models/Structural equation modeling* (рис. 10.7.2).

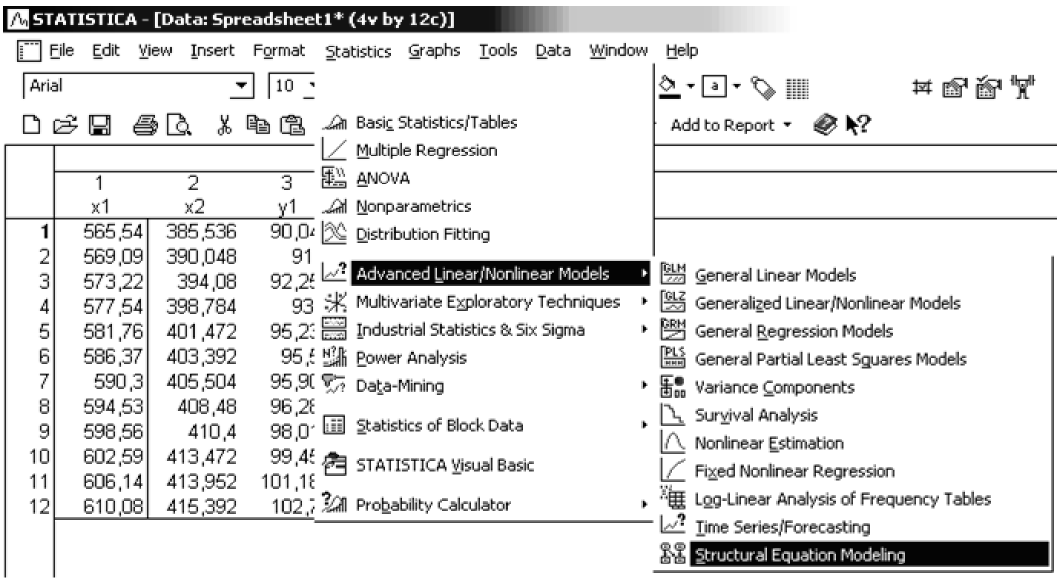


Рис. 10.7.2. **Вибір модуля**

Після підтвердження вибору модуля з'явиться стартова панель модуля (рис. 10.7.3). Функціональні клавіші *Path Tool (засоби зв'язків)* і *Path Wizards (конструктор зв'язків)* дозволяють редагувати опис моделей і створювати нові описи в діалоговому режимі. Опція *Set Parameters* (*установити параметри)* дозволяє вибрати метод оцінювання.

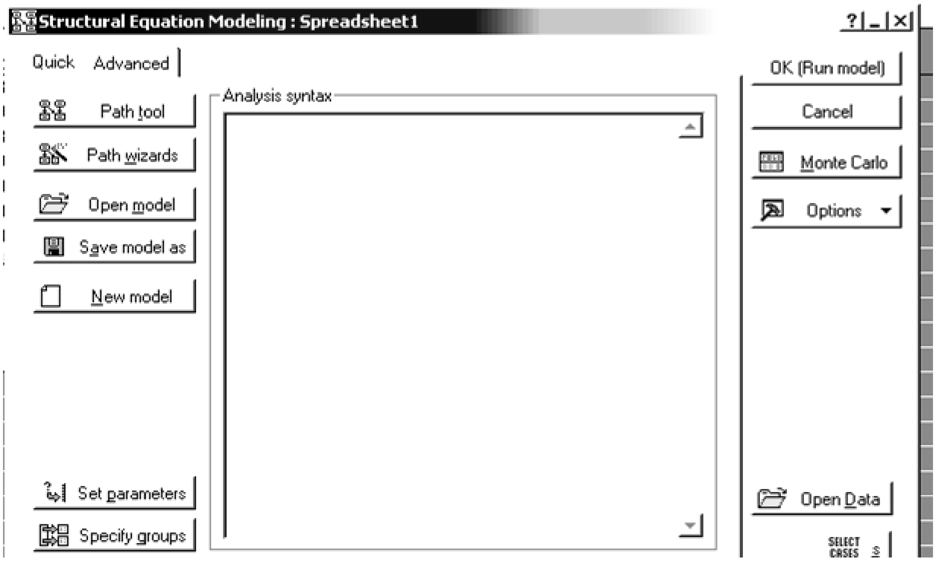


Рис. 10.7.3. **Стартова панель модуля Structural equation modeling**

У*SEPATH* розрізняють 4 типи змінних:

manifest endogenous – явні ендогенні;

manifest exogenous – явні екзогенні;

latent endogenous – скриті ендогенні;

latent exogenous – скриті екзогенні.

Після вибору функціональної клавіші *Path Wizards (конструктор зв'язків)* на екрані з'явиться діалогове вікно *Define Exogenous Variables (Визначити екзогенні змінні)* (рис. 10.7.4).

Екзогенні змінні задаються в такій послідовності: у рядку вводиться ім'я схованої змінної, далі за допомогою функціональної клавіші *Vars (Змінні)* вибираються явні змінні, що пов'язані із заданою екзогенною змінною.

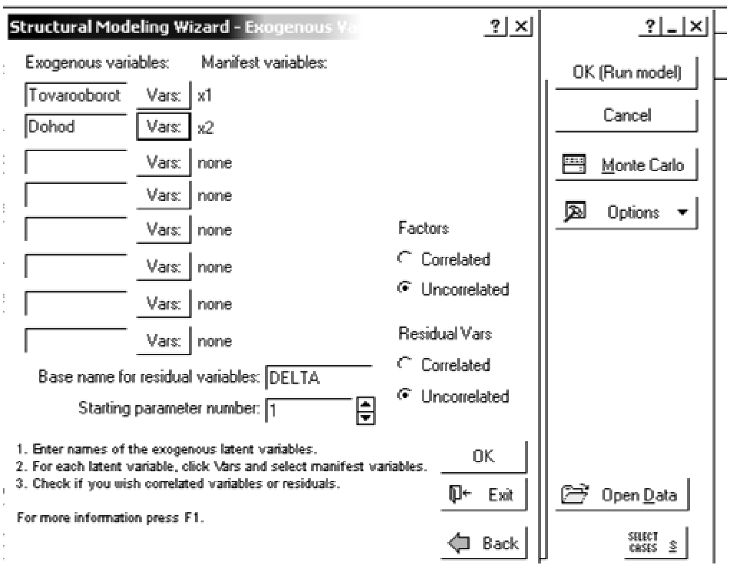


Рис. 10.7.4. **Вибір екзогенних змінних**

Вибір екзогенних змінних здійснюється аналогічним чином у вікні аналізу *Define Endogenous Variables (Визначити ендогенні змінні)* (рис. 10.7.5).

Після вибору змінних здійснюється перехід до діалогу *Define Structural Equation Paths (визначити зв'язки структурних рівнянь)*. У даному діалоговому вікні встановлюються зв'язки між обраними змінними. Для того щоб задати зв'язок, необхідно вибрати у вікні змінну зі списку (*From – із)*, що показує звідки йде зв'язок, потім вибрати змінну в списку (*To – в)*, що показує, куди йде зв'язок, ініціювати кнопку *Add (додати).* Заданий таким чином зв'язок буде доданий до існуючих. Аналогічно задаються всі зв'язки моделі (рис. 10.7.6).

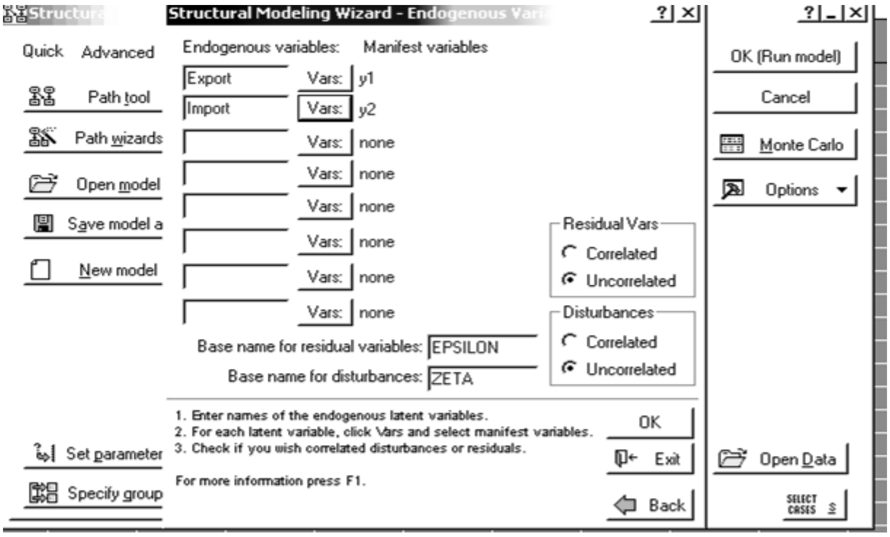


Рис. 10.7.5. **Вибір ендогенних змінних**

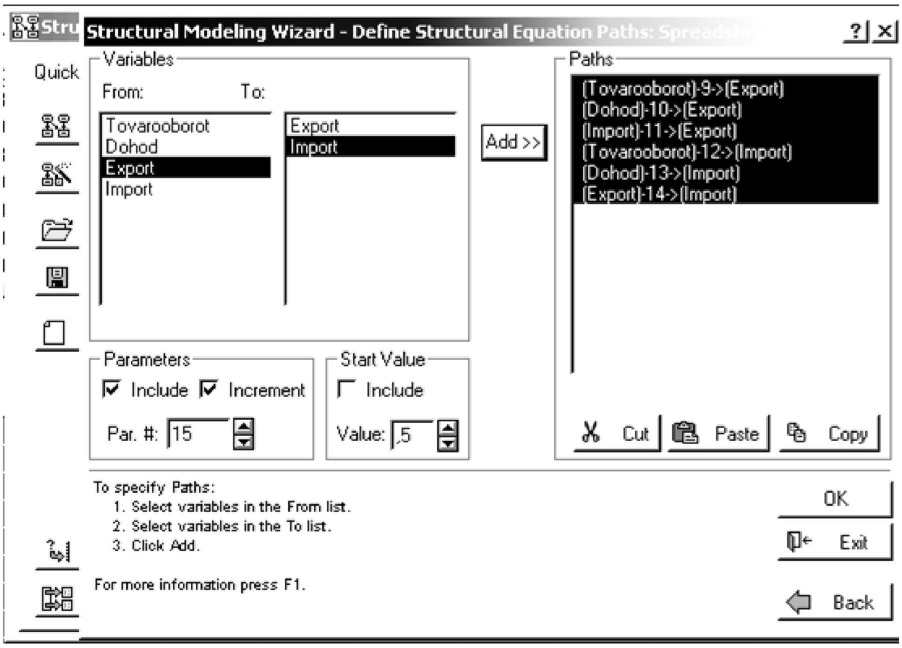


Рис. 10.7.6. **Визначення зв'язків між змінними**

Завершальне вікно діалогу (рис. 10.7.7) дозволяє приєднати модель до існуючої програми (*Append this model to existing program)* або замінити існуючу програму на програму з новою моделлю *(Replace existing program with new model)*.

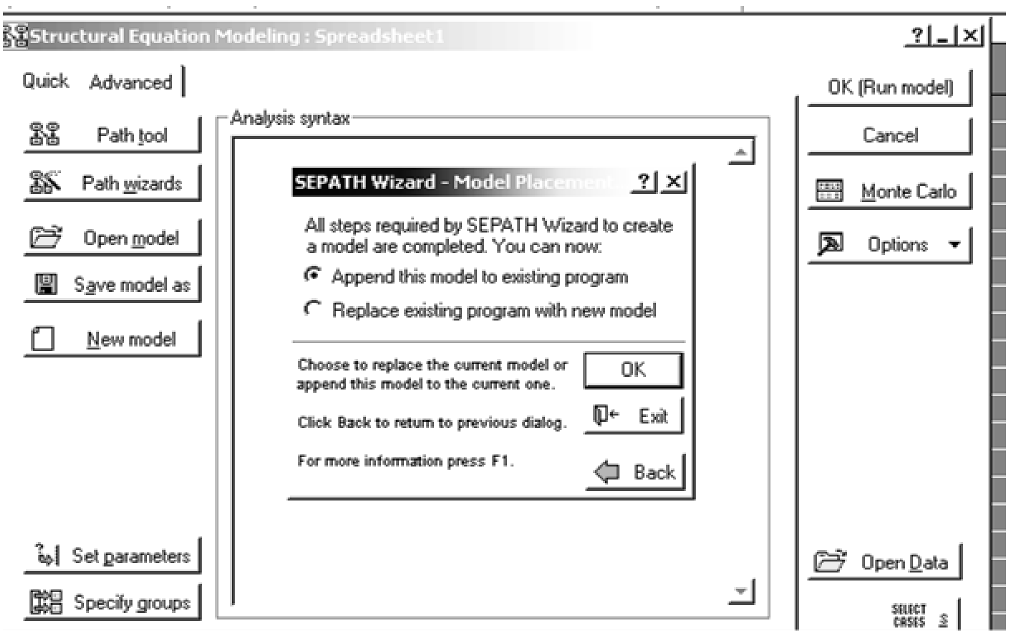


Рис. 10.7.7. **Завершальне вікно побудови моделі**

Опція *Set Parameters* (установити параметри)дозволяє встановити параметри аналізу: тип даних (Data to Analyze), опції виводу (Output Options), метод оцінювання параметрів (Discrepancy Functions) (рис. 10.7.8).

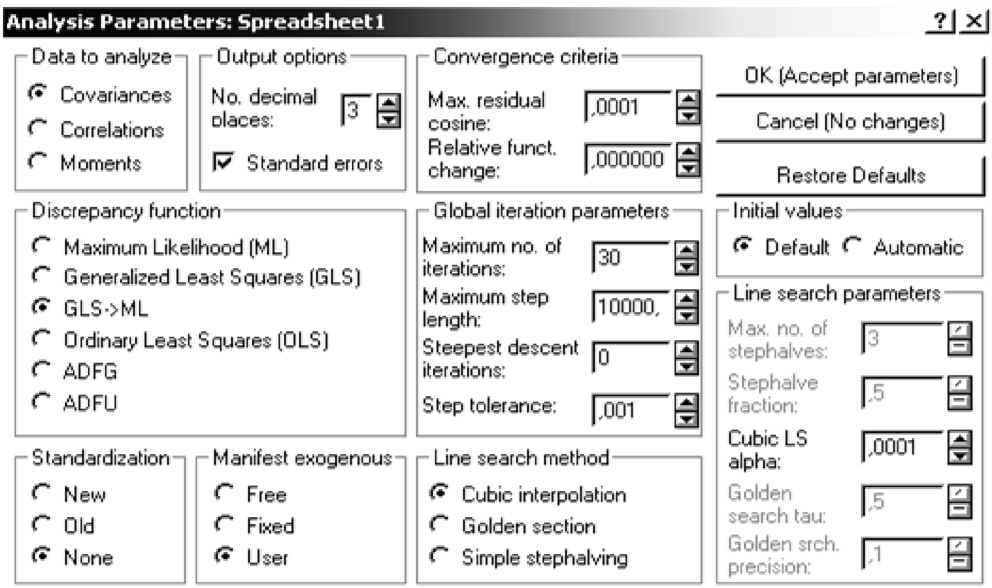


Рис. 10.7.8. **Вибір методу оцінювання**

Система пропонує такі процедури оцінювання невідомих параметрів, доступні під час вибору відповідних опцій:

Maximum Likelhood (ML) – метод максимальної правдоподібності;

Generalized Least Square (GLS) – узагальнений метод найменших квадратів;

GLS $\rightarrow$ ML – подвійна обчислювальна процедура, де спочатку використовується процедура оцінювання узагальненим МНК, потім процедура оцінювання методом максимальної правдоподібності;

Ordinary Least Square (OLS) – звичайний метод найменших квадратів.

Після того як модель записана й параметри аналізу встановлені, можна зробити обчислення. Функціональна клавіша*OK (Run Model) (оцінити поточну модель)* (див. рис. 10.7.7) дозволяє запустити процедуру оцінювання.

Функціональна частина вікна результатів оцінювання (рис. 10.7.9) дозволяє вибрати такі напрями аналізу:

1) *Noncentraility-based indicates* – індекси нецентральності;

2) *Other Single Sample Indices* – інші вибіркові індекси;

3) Model summary – підсумкова модель.

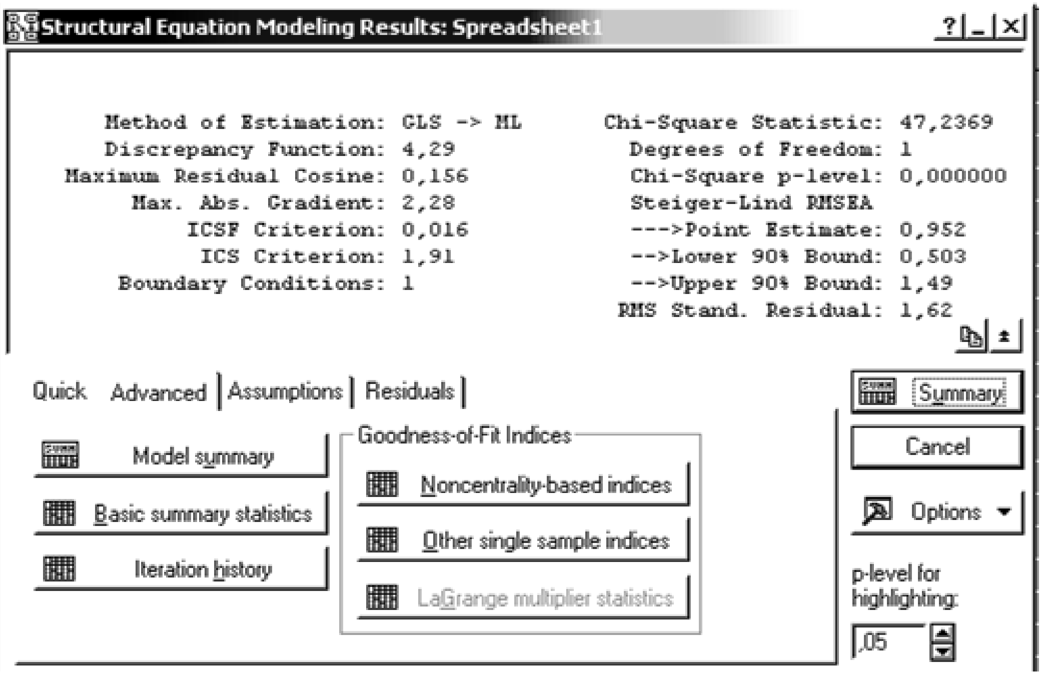


Рис. 10.7.9. **Вікно результатів структурного моделювання**

Вибір першого напряму дозволяє переглянути існуючі індекси нецентральності (рис. 10.7.10), які показують ступінь адекватності моделі.

	Noncentrality Fit Indices (Spreadsheet1		
	Lower 90% Conf. Bound	Point Estimate	Upper 90% Conf. Bound
Population Noncentrality Parameter	0,253	0,906	2,234
Steiger-Lind RMSEA Index	0,503	0,952	1,495
McDonald Noncentrality Index	0,327	0,636	0,881
Population Gamma Index	0,472	0,688	0,888
Adjusted Population Gamma Index	-4,276	-2,119	-0,122

Рис. 10.7.10. **Індекси якості підгонки моделі**

У таблиці послідовно подані: нижня межа 90 % довірчого інтервалу, точкова оцінка індексу, верхня межа 90 % довірчого інтервалу. У таблиці наведені такі індекси:

Steiger-Lind RMSEA Index – індекс Штейгера – Лінда. Близькість індексу до 0 вказує на якість підгонки;

McDonald's Index of Noncentrality – індекс нецентральності Мак-Доналдса. Якісна підгонка відповідає значенням цього індексу більше 0,95.

Вибір другого напряму дає можливість переглянути інші вибіркові індекси, аналіз значень яких дозволяє відібрати найкращий варіант моделі (рис. 10.7.11).

	Single S
	Value
Joreskog GFI	0,665
Joreskog AGFI	-2,345
Akaike Information Criterion	5,931
Schwarz's Bayesian Criterion	6,327
Browne-Cudeck Cross Validation Index	7,294
Independence Model Chi-Square	133,095
Independence Model df	6,000
Bentler-Bonett Normed Fit Index	0,645
Bentler-Bonett Non-Normed Fit Index	-1,178
Bentler Comparative Fit Index	0,636
James-Mulaik-Brett Parsimonious Fit Index	0,108
Bollen's Rho	-1,129
Bollen's Delta	0,650

Рис. 10.7.11. **Індекси якості підгонки моделі**

У даній таблиці подані такі критерії:

Akaike Information Criterion – інформаційний критерій Акайка;

Schwarz's Bayesian Criterion – критерій Шварца.

Серед конкуруючих варіантів моделі перевага віддається варіанту з найменшими значеннями даних критеріїв.

Вибір третього напряму дозволяє переглянути таблицю з результатами оцінювання (рис. 10.7.12).

	Model Estimates (Spreadsheet1)			
	Parameter Estimate	Standard Error	T Statistic	Prob. Level
(Tovarooborot)-1->[x1]	14,828	3,164	4,687	0,000
(Dohod)-2->[x2]	9,747	2,078	4,690	0,000
(DELTA1)-3-(DELTA1)	0,165	0,000		
(DELTA2)-4-(DELTA2)	0,000	0,000		
(EPSILON1)-5-(EPSILON1)	0,733	0,340	2,157	0,031
(EPSILON2)-6-(EPSILON2)	0,301	0,446	0,676	0,499
(ZETA1)-7-(ZETA1)	0,000	0,000		
(ZETA2)-8-(ZETA2)	0,732	0,000		
(Tovarooborot)-9->(Export)	10,114	2,172	4,657	0,000
(Dohod)-10->(Export)	-4,209	0,938	-4,488	0,000
(Import)-11->(Export)	-0,173	0,000		
(Tovarooborot)-12->(Import)	9,015	1,951	4,621	0,000
(Dohod)-13->(Import)	0,424	0,355	1,194	0,232
(Export)-14->(Import)	0,524	0,000		

Рис. 10.7.12. **Оцінки параметрів моделі**

У даній таблиці подані такі характеристики моделі:

Parameter Estimate – параметри моделі;

Standard Error – середнє квадратичне відхилення параметрів моделі;

T-Statistic – значущість параметрів за критерієм Стьюдента;

Prob. Level – рівень значущості критерію Стьюдента.

Глосарій

Автокореляція залишків – це наявність взаємозв'язку між послідовними елементами ряду залишків моделі.

Авторегресійна модель – це клас динамічних економетричних моделей, які містять в якості факторних змінних лагові значення ендогенних змінних.

Адитивна модель часового ряду – це модель, у якій всі компоненти ряду динаміки подані як сума цих складових, а саме: $yt = f\ t\ + + g(t) + h(t) + \varepsilon t$; застосовують у випадку, коли амплітуда сезонних коливань із часом не змінюється.

Верифікація означає перевірку вірогідності, точності й адекватності моделі об'єктивним закономірностям, реальним соціально-економічним процесам.

Взаємна кореляційна функція – це функція, яка характеризує тісноту зв'язку елементів вектора залежної змінної y_t з елементом вектора незалежної змінної x_t, зрушеним один щодо іншого на часовий лаг τ. Для побудови взаємної кореляційної функції розглядається множина коефіцієнтів кореляції між рівнями часових рядів x_t, та y_t, x_{t-1} та $y_{t,\ldots\ldots,}$ x_{t-k}, та y_t. Для різних значень $\tau = 0, 1, \ldots, k$ на основі взаємної кореляційної функції можна отримати $n + 1$ значення r_τ.

Випадкова складова (випадкові перешкоди) – це складова динамічного ряду, що відображає вплив випадкових чинників на значення його рівнів.

Виробнича функція – це функція, незалежна змінна якої набуває значення об'ємів ресурсу (чинник виробництва), що витрачається або який використовується, а залежна змінна – значення об'ємів продукції, що випускається, та має такий загальний вигляд: $y = f(x,\ a)$.

Виробнича функція Кобба – Дугласа – це мультиплікативна нелінійна модель залежності обсягу виробництва від витрат факторів виробництва, ресурсів (праці і капіталу), що його створюють. Загальний вигляд функції: $Y = a_0 K^{a_1} L^{a_2}$, де Y– об'єм випуску або доходу; K – об'єм використовуваного капіталу або фондів; L – кількість трудових ресурсів; a_0 – технологічний коефіцієнт, нейтральний НТП; a_1 – коефіцієнт еластичності праці; a_2 – коефіцієнт еластичності капіталу.

Гетероскедастичність – це явище, у разі якого дисперсія залишків змінюється для кожного спостереження або групи спостережень.

Гомоскедастичність – це явище сталої дисперсії залишків для кожного спостереження.

Гранична норма заміщення i-го ресурсу (чинника виробництва) j-м показує, на скільки одиниць збільшаться витрати j-го ресурсу (у разі фіксованого випуску продукції), якщо витрати i-го ресурсу зменшаться на одиницю.

Гранична продуктивність праці – це прирощення обсягу продукції, викликане використанням додаткової одиниці праці у разі фіксованих інших умов.

Гранична фондовіддача – це прирощення додаткових одиниць продукції, викликане використанням додаткової одиниці основних фондів при фіксованих інших умовах.

Двокроковий метод найменших квадратів (ДМНК) – це один із способів вирішення систем одночасних рівнянь, що застосовується як для ідентифікованих, так і для надідентифікованих моделей.

Динамічна економетрична модель – це економетрична модель, значення змінних якої розглядаються у динаміці.

Динамічний ряд – це послідовність значень деякого процесу, який протікає в часі. Прикладами часових рядів можуть бути фінансові індекси, щоденні курси валют, котирування акцій, річні обсяги продажів, квартальні обсяги виробництва, ділова активність тощо, тобто змінні, значення яких змінюються з часом.

Екзогенні змінні – це вхідні змінні, зовнішні, незалежні змінні, регресори або факторні ознаки, значення яких задаються за межами економетричної моделі.

Економетрична модель – це особливий клас економіко-математичних моделей, в яких дослідник вирішує цілий ряд завдань: вибір форми математичної залежності, що описує поведінку економічного об'єкта на основі системи спостережень; оцінювання параметрів даної моделі різними методами (метод найменших квадратів, метод максимальної правдоподібності та ін.); перевірка статистичної значущості моделі. У загальному вигляді економетрична модель може бути подана як система лінійних рівнянь: $BY = AX + E$, де B – матриця коефіцієнтів ендогенних змінних;

Y – вектор ендогенних змінних; A – матриця коефіцієнтів екзогенних змінних; X – вектор екзогенних змінних; E – вектор випадкових збурень (помилок, відхилень).

Економетрія – це один із напрямів економіко-математичних методів аналізу, що полягає в статистичному вимірюванні (оцінюванні) параметрів математичних виразів, що характеризують деяку економічну концепцію про взаємозв'язок і розвиток об'єкта, явища, що необхідно для отримання конкретних економічних висновків на основі економетричних моделей.

Еластичність випуску продукції за факторами виробництва показує, на скільки відсотків збільшиться випуск продукції під час збільшення витрат праці або капіталу на 1 %.

Еластичність – це міра чутливості однієї змінної (наприклад, обсягів випуску або доходу) до зміни іншої (наприклад, капіталу або праці), що показує, на скільки відсотків зміниться перший показник у разі зміни іншого на 1 %.

Ендогенні змінні – це вихідні характеристики, внутрішні, залежні, результуючі змінні або ознаки, значення яких визначаються на основі рівняння регресії.

Етапи побудови економетричної моделі**:** якісний аналіз (постановка мети аналізу, визначення сукупності, визначення результативних і факторних ознак, вибір періоду, за який проводиться аналіз, вибір методу аналізу); попередній аналіз модельованої сукупності (перевірка однорідності сукупності, виключення аномальних спостережень, уточнення необхідного обсягу ознак, установлення законів розподілу ознак); побудова економетричної моделі (встановлення переліку чинників, розрахунок оцінок параметрів рівнянь регресії, перебір конкуруючих варіантів моделі); оцінювання адекватності моделі (перевірка статистичної суттєвості рівняння залежності в цілому і його окремих параметрів; перевірка відповідності формальних властивостей оцінок завданням дослідження); економічна інтерпретація і практичне використання моделі.

Ефективність оцінок. Оцінки ефективні, якщо їх дисперсія мінімальна.

Зумовлені змінні (передвизначені) – це лагові та поточні екзогенні, а також – лагові ендогенні змінні моделі.

Ідентифікація моделі – це проведення статистичного аналізу моделі та оцінювання якості її параметрів; установлення відповідності між наведеної та структурної формами моделі.

Ідентифікована модель – це різновид структурної моделі системи одночасних рівнянь, в якій всі структурні коефіцієнти однозначно визначаються через наведені коефіцієнти.

Ізокванта – це геометричне місце точок у просторі ресурсів, у яких різні поєднання виробничих ресурсів дають одну і ту ж кількість продукції, що випускається.

Ізокліналь – це лінія, що сполучає початок координат (точку О (0,0)) і точки на ізоквантах ВФ, для яких рівними будуть граничні норми заміни ресурсів.

Коефіцієнт детермінації – це критерій оцінювання адекватності моделі, відношення поясненої (через передбачувану регресійну модель) варіації результативної ознаки до всієї варіації в цілому, тобто вимірює загальну варіацію залежної змінної, яка пояснюється регресією:

$R^2 = 1 - \frac{{}_{i=1}^{n}\ Y_i - Y_i\ ^2}{{}_{i=1}^{n}\ Y_i - Y\ ^2}$. Таким чином, чим ближче коефіцієнт до 1, тим більше адекватна модель для опису конкретного економічного явища.

Коефіцієнт множинної кореляції – це міра тісноти взаємозв'язку всіх факторних ознак із залежною змінною. Визначається за формулою:

$R = \overline{1 - \frac{{}_{i=1}^{n}\ Y_i - Y_i\ ^2}{{}_{i=1}^{n}\ Y_i - Y\ ^2}}$. Чим ближче коефіцієнт R до 1, тем краще підібрана модель для опису залежності. Зв'язок між досліджуваними економічними явищами вважається достовірним, якщо $R > 0{,}7$.

Коефіцієнт парної кореляції – це оцінка тісноти взаємозв'язку прямої або зворотної залежності між парами змінних: результуючої змінної й факторної ознаки x_j; а також парами незалежних змінних x_i і x_j. Розраховуються за формулою: $r_{yxi} = \frac{x_i - x_i\ \ y_i - y_i}{\overline{x_i - x_i\ ^2\ \ y_i - y_i\ ^2}}$ або $r_{x_i x_j} =$

$= \frac{x_i - x_i\ \ x_j - x_j}{\overline{x_i - x_i\ ^2\ \ x_j - x_j\ ^2}}$. Чим ближче коефіцієнт за модулем до 1, тем тісніше ступінь лінійної залежності між ознаками.

Коінтеграція – це причинно-наслідковий зв'язок у рівнях двох або більше часових рядів, що виражається в збігу або протилежній спрямованості їх тенденцій і випадкових коливань.

Критерій μ – це статистичний тест, що дозволяє перевірити гетероскедастичність випадкових помилок регресійної моделі тоді, коли вихідна сукупність спостережень досить велика.

Критерій Дарбіна – Уотсона – це найпоширеніший критерій, який застосовується для оцінювання автокореляції (першого порядку) між сусідніми залишковими членами ряду.

Критерій Стьюдента (t-критерій) використовується під час перевірки гіпотези про значущість параметрів моделі, парних, частинних коефіцієнтів кореляції. Якщо $t_{\text{теор}} \leq t_{\text{табл}}$ то з імовірністю помилки α приймається гіпотеза H_0: відповідна факторна ознака незначно впливає на результативну ознаку. В іншому випадку, тобто у разі $t_{\text{теор}} > t_{\text{табл}}$, приймається гіпотеза H_1: відповідна факторна ознака значно впливає на результативну ознаку. У даному критерії $t_{\text{табл}}$ – значення порогове (критичне) для критерію Стьюдента, яке береться з таблиць для відповідного рівня значущості α і кількості ступенів свободи k.

Критерій Фішера (F-критерій). За допомогою цього критерію перевіряється існування зв'язку між залежної й незалежною змінними: $F = \frac{R^2}{1-R^2} \cdot \frac{n-m-1}{m}$. *Розраховане значення статистики Фішера необхідно* порівняти з табличним $F_{\text{табл}}$ для числа ступенів свободи $k_1 = m$, $k_2 = m - n - 1$, рівня значущості α. За умови $F > F_{\text{табл}}$ побудова регресійна модель відповідає реальній дійсності.

Критерій фон Неймана – це критерій, який застосовується для оцінювання автокореляції залишків першого порядку**.**

Лаг – це величина інтервалу запізнювання.

Лагові змінні – це пояснюючі змінні, взяті в моделі регресії з запізненням у часі.

Лінеаризація – це перехід від нелінійних зв'язків (гіперболічної, показової, степеневої) до лінійного зв'язку за допомогою різних перетворень, що дозволяє надалі використовувати звичайний метод найменших квадратів (МНК).

Медіанний лаг – це лаг, який показує тривалість періоду, необхідного для реалізації половини загальної зміни результуючої ознаки після збільшення значень факторної змінної на 1. Для медіанного лагу справедлива наступна рівність: ${}_{j=0}^{M_e} \alpha_j = 0,5$.

Метод виключення факторів – це один із методів виключення мультиколінеарності, який полягає в тому, що до моделі спочатку включаються всі фактори. Потім із моделі виключають фактор, коефіцієнт при якому статистично незначущий і має найменше значення t-критерію. Після цього отримують нове рівняння регресії і знову проводять оцінювання значущості всіх решти коефіцієнтів регресії.

Метод включення факторів – це один із методів виключення мультиколінеарності і полягає в тому, що в модель включаються фактори по одному в певній послідовності.

Метод головних компонентів – це один з методів виключення мультиколінеарності, заснований на переході від вихідного простору колінеарних факторних змінних до ортогонального простору нових штучних змінних – головних компонентів.

Метод Дарбіна – це двокроковий метод оцінювання параметрів узагальненої регресійної моделі з автокореляцією залишків, який забезпечує найкращу оцінку для ширшого кола параметрів порівняно з іншими методами.

Метод Койка – це метод оцінювання параметрів моделі з нескінченним числом лагів. Одним із припущень методу є припущення про те, що лагові коефіцієнти зменшуються в геометричній прогресії: $a_k = a_0\lambda^k$.

Метод Кохрейна – ***Оркатта*** – це ітеративний метод оцінювання параметрів узагальненої регресійної моделі з автокореляцією залишків.

Метод найменших квадратів (МНК) – це метод визначення невідомих параметрів моделі за умови мінімізації функціонала F. Функціонал F поданий сумою квадратів відхилень реальних значень результативної (залежної) змінної від її передбачуваних теоретичних значень. Теоретичні значення визначаються на підставі обраної дослідником моделі, що відображає залежність результативної змінної від факторних ознак.

Метод перетворення вихідної інформації – це метод, який застосовується для оцінювання параметрів моделі за допомогою двокрокової процедури, коли залишки задовольняють авторегресійну модель першого порядку.

Метод рідж-регресії – це один із методів виключення мультиколінеарності, заснований на покращенні обумовленості матриці X^TX шляхом додавання до її діагональних елементів "гребеня" τ, що, однак, призводить до зміщеності МНК-оцінок параметрів.

Метод Фаррара – ***Глобера*** – це метод визначення наявності мультиколінеарності у масиві регресорів, заснований на застосуванні трьох видів статистичних критеріїв: 1) усього масиву незалежних змінних (критерій χ^2); 2) кожної незалежної змінної з усіма іншими (F-критерій); 3) кожної пари незалежних змінних (t-критерій).

Метод Фостера – ***Стюарта*** – це критерій перевірки на наявність тренда, який тестує не тільки середнє значення, а й дисперсію рівнів часового ряду. Часто цей метод використовують у разі детального аналізу часового ряду і побудови за ним прогнозів.

Метод Хілдрета – Лу – це ітеративний метод оцінювання параметрів узагальненої регресійної моделі з автокореляцією залишків.

Метод Ширлі Алмон – це метод оцінювання параметрів моделі з кінцевим числом лагів. В основі методу Ширлі Алмон лежить гіпотеза, що лагові коефіцієнти регресії апроксимуються поліномом відповідного ступеня від величини лага: $a_j = P_r\ j\ \ = b_0 + b_1 j + b_2 j^2 + \cdots + b_r j^r$.

Модель з розподіленим лагом– це модель, в якій залежна змінна y_t пов'язана зі значеннями пояснювальних зміннихx не тільки в момент часу t, але і їх значеннями в попередні моменти часу. Однофакторна модель розподіленого лага має вигляд: $y_t\ =\ f(x_t,\ x_{t-1},\ \ldots\ ,\ \ x_{t-k})\ +\ \varepsilon_t$.

Мультиколінеарність – це наявність тісних лінійних зв'язків між включеними до множинної економетричної моделі екзогенними змінними.

Мультиплікативна модель часового ряду – це модель, в якій всі компоненти ряду динаміки подані як добуток цих складових, а саме: $y_t\ =\ f(t)\ +\ g(t)\ +\ h(t)\ + \varepsilon_t$; модель застосовують у випадку, якщо збільшується амплітуда коливань відповідно до тренда.

Наведена форма моделі – це один зі способів запису системи одночасних рівнянь, у якому кожна ендогенна змінна визначена у вигляді лінійної функції від усіх зумовлених змінних.

Надідентифікована модель – це різновид структурної моделі системи одночасних рівнянь, у якій структурні коефіцієнти, виражені через наведені коефіцієнти, мають два і більше числових значень.

Незміщеність оцінок. Оцінки параметрів моделі будуть незміщеними, якщо математичне очікування їх вибіркових значень, знайдених під час багаторазового повторення вибірки, не відрізняється від дійсного значення, тобто $M(a) = a$.

Неідентифікована модель – це різновид структурної моделі системи одночасних рівнянь, у якій структурні коефіцієнти неможливо знайти за наведеними коефіцієнтами моделі.

Нелінійна модель – це економіко-математична модель, що відображає стан або функціонування системи (нелінійної системи, стохастичної системи) таким чином, що всі або деякі взаємозв'язки в ній є нелінійними, тобто не задовольняють умовам лінійності.

Непараметричний тест Гольдфельда – Квандта (*Goldfeld-Quandt nonparametric test*) не спирається на припущення про нормальний розподіл залишків і базується на графічному зображенні і підрахунку кількості піків значень квадратів залишків регресійної моделі, побудованої на основі припущення про відсутність гетероскедастичності, після впорядкування (ранжування) спостережень за X_j. Якщо для всіх значень змінної X_j залишки розподіляються приблизно однаково, то дисперсія їх однорідна і гетероскедастичність відсутня. Якщо вона змінюється, то гетероскедастичність присутня.

Неприпустима мультиколінеарність – це коли між факторами X_i і X_j існує значна додатна кореляція і у ході цього вплив кожного фактора на кореляційний зв'язок із Y односпрямований, тобто збільшення обох факторів X_i і X_j приводить до збільшення або зниження Y: $r_{yi} \quad r_{yj} > 0$.

Іншими словами, обидва фактори діють на Y однаково і значна додатна кореляція між ними може дозволити виключити один із них.

Непрямий метод найменших квадратів (НМНК) – це один зі способів вирішення систем одночасних рівнянь, заснований на отриманні обґрунтованих і незміщених оцінок параметрів структурної форми моделі за оцінками параметрів наведеної форми.

Нормовані коефіцієнти лага – це коефіцієнти, які показують пропорцію довгострокового впливу факторної змінної, що припадає на певний період часу.

Обґрунтованість оцінок. Оцінки будуть обґрунтованими, якщо у разі дуже малої величини ε справедливим є твердження $\lim_{n\to\infty} P \quad a -$ $-a \ < \varepsilon \ = 1$.

Параметризація – це визначення виду економічної моделі, вираз у математичній формі взаємозв'язку між її змінними, формулювання вихідних передумов і обмежень моделі.

Параметричний тест Гольдфельда – Квандта (*Goldfeld-Quandt parametric test*) – це статистичний тест, що дозволяє перевірити гетероскедастичність випадкових помилок регресійної моделі, застосовується до невеликих вибірок і припускає нормальний розподіл і незалежність випадкових величин e_i. Він припускає впорядкованість за однією пояснювальною змінною, видалення точок даних у центрі і порівняння середніх відхилень із лівого і з правого боку.

Повна(екстремальна) мультиколінеарність виникає тоді, коли всі регресори позв'язані лінійною залежністю, що не дозволяє розділити внесок окремої екзогенної змінної (X_i) в залежну змінну Y.

Припустима мультиколінеарність – це така, за якої фактори діють на функцію Y неоднаково.

Продуктивність праці – це ступінь ефективності використання трудових ресурсів. Продуктивність праці вимірюється кількістю продукції, випущеної працівником за одиницю часу.

Регресія – це функція, що відображає статистичний зв'язок між ознаками.

Сезонна складова – це складова динамічного ряду, призначена для опису поведінки, що регулярно змінюється протягом заданого періоду (наприклад, обсяг продажів шампанського наприкінці грудня кожного року, обсяг перевезень пасажирів зранку та ввечері, попит на морозиво в літні місяці тощо).

Середня величина лага – показує середній інтервал часу, протягом якого буде відбуватися зміна залежної змінної під впливом пояснювальної змінної x у момент часу t.

Середня продуктивність ресурсів – характеризує середню кількість продукції на одиницю витраченої праці чи використаних виробничих фондів.

Система незалежних рівнянь – це різновид систем економетричних рівнянь, в якій кожна результативна ознака є функцією однієї і тієї ж сукупності факторів; набір факторів у кожному рівнянні системи може змінюватися залежно від досліджуваного явища.

Система одночасних рівнянь – це різновид економетричних моделей, що складається з тотожностей і регресійних рівнянь, у яких поряд з факторними ознаками включені результативні ознаки з інших рівнянь системи.

Система рекурсивних рівнянь – це різновид систем економетричних рівнянь, у якій результативна ознака одного рівняння системи в кожному наступному рівнянні є фактором поряд з однією і тією ж сукупністю факторів.

Стаціонарний динамічний ряд – це ряд, характер якого не змінюється з часом, а отже, це ряд x_t , $t = 1, \ldots, n$, якщо для будь-якого m ($m < n$) сумісний розподіл ймовірностей випадкових величин $X_{t_1} \ldots X_{t_m}$ такий самий, як і для $X_{t_{1+\tau}} \ldots X_{t_{m+\tau}}$ за умови будь-яких $t_1, \ldots t_m$і τ, таких, що $1 \leq t_1, \ldots t_m \leq n$ і $1 \leq t_1, +\tau, \ldots t_m + \tau \leq n$.

Стохастична складова – це випадкова змінна в економетричній моделі $Y = f(X, u)$, яка має розподіл із математичним сподіванням, що дорівнює нулю, і сталою дисперсією σ_u^2. Це дає змогу розглядати змінну u як стохастичне збурення (помилку, відхилення) моделі.

Структура лага – це послідовність лагових коефіцієнтів $a = a_j: j = 1,2, \ldots, k$ у моделі розподіленого лага $y_t = f(x_t, x_{t-1}, \ldots, x_{t-k}) + \varepsilon_t$.

Структурна форма моделі – це один із способів запису системи одночасних рівнянь, який відображає реальний економічний об'єкт або явище і показує, як зміна будь якої ендогенної змінної визначає значення ендогенної змінної моделі.

Тенденція автокореляції – це вид тенденції часового ряду, який характеризує зв'язок між окремими рівнями ряду динаміки.

Тенденція дисперсії – це вид тенденції часового ряду, яка характеризує напрям зміни відхилень між емпіричними рівнями і детермінованою компонентою ряду.

Тенденція середнього рівня – це вид тенденції часового ряду, який виражається зазвичай за допомогою математичного рівняння функції, навколо якої варіюють фактичні рівні досліджуваного явища. Рівняння тенденції має вигляд: $Y_t = f_t + \varepsilon$. Сенс цієї функції полягає в тому, що значення тренда в окремі моменти часу виступають як математичне сподівання ряду динаміки.

Тест Бройша (Бреуша) – Годфрі (тест серій) – це універсальний асимптотичний тест для перевірки наявності автокореляції випадкових похибок більшого порядку. У даному тесті випадкові помилки не обов'язково повинні бути нормально розподілені. Тест можна застосувати також і в авторегресійних моделях (на відміну від критерію Дарбіна – Уотсона).

Тест Глейзера (*Glejser test*) – це статистичний тест, що дозволяє оцінити наявність певного виду гетероскедастичності випадкових помилок регресійній моделі. Тест заснований на наступній моделі можливої залежності стандартного відхилення випадкової помилки моделі σ_t від деякого фактора x_{jt}: $\sigma_t = a_0 + a_1 x_{jt}^{\gamma} + u_t$.

Тест Парка (*Park test*) – це параметричний статистичний тест, використовуваний для перевірки певного виду гетероскедастичності випадкових помилок регресійній моделі. У даному тесті передбачається можливість залежності дисперсії випадкової помилки σ_t^2 моделі від значень деякого фактора x_{jt} такого вигляду: $ln(\sigma_t^2) = ln(\sigma^2) + a_1 \ln\ x_{jt}\ + u_t$.

Тест рангової кореляції Спірмена – це непараметричний статистичний тест, що дозволяє перевірити гетероскедастичність випадкових помилок регресійної моделі. Особливість тесту полягає в тому, що не конкретизується форма можливої залежності дисперсії випадкових помилок моделі від тієї чи іншої змінної.

Тотожність – це один із різновидів структурних рівнянь моделі, яка встановлює співвідношення між ендогенними змінними та не містить випадкових складових і структурних коефіцієнтів.

Тренд – це еволюторна складова динамічного ряду, яка плавно змінюється з часом (зростає або спадає, але не повторюється регулярним чином), описує вплив довготривалих чинників, ефект яких виявляється поступово (наприклад, зростання населення, зростання споживання, зміна структури споживання, економічний розвиток тощо).

Узагальнений метод найменших квадратів (метод Ейткена) – це метод оцінювання параметрів узагальненої економетричної моделі з гетероскедастичністю або автокореляцією залишків, оператор оцінювання якого має вигляд: $a_{GLS} = \ X^{\mathrm{T}}S^{-1}X\ ^{-1}X^{\mathrm{T}}S^{-1}Y$ або $a_{GLS} =$ $= \ X^{\mathrm{T}}V^{-1}X\ ^{-1}X^{\mathrm{T}}V^{-1}Y$, де $V = \sigma_e^2 S$.

Фондовіддача – це показник, що характеризує рівень ефективності використання основних виробничих фондів підприємства чи галузі та визначає кількість продукції, виробленої на одну грошову одиницю виробничих основних фондів.

Фондоозброєність – це показник, що характеризує оснащеність працівників підприємств або галузей виробництва основними виробничими фондами. Визначаться як відношення середньорічної вартості основних виробничих фондів до середньорічної облікової чисельності працівників або робітників.

Циклічна складова – це складова динамічного ряду, яка описує тривалі періоди відносного підйому і спаду. Вона складається з циклів, які міняються за амплітудою і протяжністю. На циклічну складову впливають такі чинники (що складно ідентифікуються формальними методами), як зростання і виснаження ресурсів, тривало діючі несприятливі погодні умови, зміни у фінансовій і податковій політиці і т. д.) Для аналізу цієї складової часового ряду зазвичай залучається додаткова інформація про інші часові ряди, наприклад, про перенасиченість ринку, прийняття законів про податкові пільги і т. д.

Часткова (недосконала, стохастична) мультиколінеарність характерна для випадків, коли частина екзогенних чинників $(X_1,\ X_2,\ldots,\ X_m)$ знаходиться в кореляційному зв'язку або утворює різні лінійні комбінації вигляду $X_i = \alpha_1 X_1 + \alpha_2 X_2 + \ldots + \alpha_m X_m$.

Предметний покажчик

Використана література

1. Бабешко Л. О. Основы эконометрического моделирования / Л. О. Бабешко. – М. : КомКнига, 2006. – 432 с.

2. Бернд Э. Р. Практика эконометрики: классика и современность : учебник для студентов вузов, обучающихся по специальностям 060000 экономики и управления / Э. Р. Бернд; пер. с англ. под ред. проф. С. А. Айвазяна. – М. : ЮНИТИ-ДАНА, 2005. – 863 с.

3. Боровиков В. П. STATISTICA: искусство анализа данных на компьютере. Для профессионалов / В. П. Боровиков. – СПб. : Питер, 2001. – 656 с.

4. Боровиков В. П. Популярное введение в программу STATISTICA / В. П. Боровиков. – М. : Компьютер Пресс, 1998. – 194 с.

5. Боровиков В. П. Прогнозирование в системе STATISTICA в среде Windows / В. П. Боровиков, Г. И. Ивченко. – М. : Финансы и статистика, 1997. – 268 с.

6. Валентинов В. А. Эконометрика: практикум / В. А. Валентинов. – М. : Дашков и К^{о}, 2007. – 436 с.

7. Грабовецький Б. Є. Економічне прогнозування і планування : навч. посіб. / Б. Є. Грабовецький. – К. : Центр навчальної літератури, 2003. – 188 с.

8. Гур'янова Л. С. Лабораторний практикум з навчальної дисципліни "Економетрія" для студентів напряму підготовки "Економічна кібернетика" денної форми навчання / Л. С.Гур'янова, О. А.Сергієнко. – Х. : Вид. ХНЕУ, 2009. – 96 с.

9. Доугерти К. Введение в эконометрику / К. Доугерти ; пер. с англ. – М. : ИНФРА-М, 1997. – 402 с.

10. Дуброва Т. А. Статистические методы прогнозирования : учеб. пособ. для вузов / Т. А. Дуброва. – М. : ЮНИТИ-ДАНА, 2003. – 206 с.

11. Економіко-математичне моделювання : навч. посіб. / Т. С. Клебанова, О. В. Раєвнєва, С. В. Прокопович. – Х. : ВД "ІНЖЕК", 2010. – 352 с.

12. Замков О. О. Эконометрические методы в макроэкономичесом анализе / О. О. Замков. – М. : ГУ ВШЭ, 2001. – 280 с.

13. Иванов В. В. Анализ временных рядов и прогнозирование экономических показателей / В. В. Иванов. – Х. : ХНУ, 1999. – 230 с.

14. Клебанова Т. С. Методы прогнозирования : учеб. пособ. / Т. С. Клебанова,В. В. Иванов, Н. А. Дубровина. – Х. : Изд. ХГЭУ, 2002. – 372 с.

15. Клебанова Т. С. Эконометрия : учеб. пособ. / Т. С. Клебанова, Н. А. Дубровина, Е. В. Раевнева. – 2-е изд., испр. – Х. : ИД "ИНЖЭК", 2005. – 160 с.

16. Кремер Н. Ш. Эконометрика : учебник для вузов / Н. Ш. Кремер, Б. А. Путко. – М. : ЮНИТИ-ДАНА, 2002. – 311 с.

17. Лугінін О.Є. Економетрія : навч. посіб. / О.Є. Лугінін. – 2-ге вид., перероб. та доп. – К. : Центр учбової літератури, 2008. – 278 с.

18. Лук'яненко І. Г. Сучасні економетричні методи у фінансах / І. Г. Лук'яненко, Ю. О. Городніченко. – К. : Літера ЛТД, 2002. – 352 с.

19. Лук'яненко І. Економетрика / І. Лук'яненко, Л. Краснікова. – К. : Товариство "Знання", КОО, 1998. – 494 с.

20. Лукашин Ю.П. Адаптивные методы краткосрочного прогнозирования / Ю.П. Лукашин. – М. : Статистика, 1988. – 247 с.

21. Магнус Я. Р. Эконометрика. Начальный курс : учебник / Я. Р. Магнус, П. К. Катышев, А. А. Пересецкий. – 6-е изд., перераб. и доп. – М. : Дело, 2004. – 576 с.

22. Малыгин. Д. Е. Разработка и исследование макромоделей налогообложения : монография / Д. Е Малыгин. – Тамбов : Изд. Тамб. гос. техн. ун-та, 2009. – 88 с.

23. Марно В. Путеводитель по современной эконометрике / В. Марно. – М. : Научная книга, 2008. – 616 с.

24. Методичні рекомендації до виконання практичних завдань з навчальної дисципліни "Економетрика" для студентів напряму підготовки 6.030502 "Економічна кібернетика" денної форми навчання / Т. С. Клебанова, Л. С. Гур'янова, О. А. Сергієнко та ін. – Х. : Вид. ХНЕУ, 2012. – 136 с.

25. Методы и модели прогнозирования социально-экономических процессов : учеб. пособ. /Т. С. Клебанова, В. А. Курзенев, В. Н. Наумов и др. – СПб. : Изд. СЗИУ РАНХ и ГС, 2012. – 566 с.

26. Моделі і методи соціально-економічного прогнозування / В. М. Геєць, Т.С. Клебанова, О. І. Черняк та ін. – Х. : ВД "ІНЖЕК", 2005. – 396 с.

27. Наконечний С. І. Економетрія : підручник / С. І. Наконечний, Т. О. Терещенко, Т.П. Романюк. – 4-те вид., доп. та перероб. – К. : КНЕУ, 2006. – 528 с.

28. Орлов. А. Н. Эконометрика : учеб. пособ. для вузов / А. Н. Орлов. – М. : Изд. "Экзамен", 2002. – 576 с.

29. Пашута М. Т. Прогнозування та макроекономічне планування : навч. посіб. / М. Т. Пашута. – К. : МАУП, 1998. – 192 с.

30. Сценарные модели сбалансированного социально-экономического развития регионов : монографія / под ред. Т. С. Клебановой, О. В. Мозенкова. – Бердянск : Издатель Ткачук А. В., 2013. – 328 с.

31. Тихомиров Н. П. Эконометрика / Н. П. Тихомиров, Е. Ю. Дорохина. – М. : Изд. "Экзамен", 2003. – 512 с.

32. Уотшем Т. Количественныеметоды в финансах / Т. Уотшем, К. Паррамоу ; пер. с англ. под ред. М. Р. Ефимовой. – М. : Финансы, ЮНИТИ, 1999. – 527с.

33. Холод Б.И. Комплекс эконометрических моделей построения возможных вариантов развития региона : монография / Б. И. Холод, В. А. Ткаченко, Л. Д. Гармидер ; Днепропетр. ун-т экономики и права. – Днепропетровск : Монолит, 2007. – 324 с.

34. Черняк О. І. Динамічна економетрика / О. І. Черняк, А. В. Ставицький. – К. : КВІЦ, 2000. – 120 с.

35. Чураков Е. П. Математические методы обработки экспериментальных данных в экономике : учеб. пособ. / Е. П. Чураков. – Финансы и статистика, 2004. – 240 с.

36. Эконометрика / под ред. д-ра экон. наук, проф. М. С. Мхитаряна. – М. : Проспект, 2009. – 384 с.

37. Эконометрика : учебник / под ред. И.И. Елисеевой. – М. : Проспект, 2009. – 228 с.

38. Эконометрика : учеб. пособ. в схемах и таблицах / Н. М. Гореева, Л. Н. Демидова, Л. М. Клизогуб и др. ; под ред. С. А. Орехова. – М. : Эксмо, 2008. – 224 с.

39. Эконометрия / В. И. Суслов, Н. М. Ибрагимов, Л. П. Талышева и др. – Новосибирск : СО РАН, 2005. – 744 с.

40. Эконометрия на персональном компьютере / Т. С. Клебанова, Н. А. Дубровина, А. В. Милов и др. – Х. : Изд. ХГЭУ, 2002. – 208 с.

41. Glejser H. A New Test for Heteroskedasticity / H.Glejser // Journal of the American Statistical Association. – 1969. – Vol. 64. – P. 316–323.

42. Goldfeld S. M. Some Tests for Homoscedasticity / S. M. Goldfeld, R. E. Quandt // Journal of the American Statistical Association. – 1965. – Vol. 60, Issue 310. – P. 539–547.

43. White H. A Heteroscedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity / H. White // Econometrica. – 1980. – Vol. 48, No. 4. – P. 817–838.

Додатки

Додаток А

Таблиця А.1

Критерій λ Колмогорова – Смірнова

λ	P(λ)	λ	P(λ)	λ	P(λ)	λ	P(λ)	λ	P(λ)	λ	P(λ)
0,3	1,00	0,55	0,9228	0,8	0,5441	1,1	0,1777	1,6	0,0120	2,1	0,0003
0,35	0,9997	0,6	0,8643	0,85	0,4653	1,2	0,1122	1,7	0,0062	2,2	0,0001
0,4	0,9972	0,65	0,7920	0,9	0,3927	1,3	0,0681	1,8	0,0032	2,3	0,0001
0,45	0,9874	0,7	0,7112	0,95	0,3275	1,4	0,0397	1,9	0,0015	2,4	0000
0,5	0,9639	0,75	0,6275	1,00	0,2700	1,5	0,0222	2,0	0,0007	2,5	0000

Умовні позначення:

λ – значення критерію; P(λ) – ймовірність.

Таблиця А.2

Таблиця ймовірностей $P(\chi^2)$ для критерія Пірсона χ^2

χ^2 \ k	1	2	3	4	5	6	7	8	9	10
1	2	3	4	5	6	7	8	9	10	11
1	0,3173	0,6065	0,8013	0,91	0,9626	0,9856	0,9948	0,9982	0,9994	0,9998
2	1574	3679	5724	7318	8491	9197	9598	9810	9915	9963
3	0833	2231	3916	5578	7000	8088	8850	9344	9643	9814
4	0455	1353	2615	4060	5494	6767	7798	8571	9114	9473
5	0254	0821	1718	2873	4159	5438	6600	7576	8343	8912
6	0143	0498	1116	1991	3062	4232	5398	6472	7399	8153
7	0081	0302	0719	1359	2206	3208	4289	5366	6371	7254
8	0047	0183	0460	0916	1562	2381	3326	4335	5341	6288
9	0027	0111	0293	0611	1091	1736	2527	3423	4373	5321
10	0016	0067	0186	0404	0752	1247	1886	2650	3505	4405
11	0009	0041	0117	0266	0514	0884	1386	2017	2757	3575
12	0005	0025	0074	0174	0348	0620	1006	1512	2133	2851
13	0003	0015	0046	0113	0234	0430	0721	1119	1626	2237
14	0002	0009	0029	0073	0156	0296	0512	0818	1223	1730
15	0001	0006	0018	0047	0104	0203	0360	0591	0909	1321
16	0001	0003	0011	0030	0068	0138	0251	0424	0669	0996
17	0000	0002	0007	0019	0045	0093	0174	0301	0487	0744
18		0001	0004	0012	0029	0062	0120	0212	0352	0550
19		0001	0003	0008	0019	0042	0082	0149	0252	0403
20		0000	0002	0005	0013	0028	0056	0103	0179	0293
21			0001	0003	0008	0018	0038	0071	0126	0211
22			0001	0002	0005	0012	0025	0049	0089	0151
23			0000	0001	0003	0008	0017	0034	0062	0107
24				0001	0002	0005	0011	0023	0043	0076
25				0001	0001	0003	0008	0016	0030	0053
26				0000	0001	0002	0005	0010	0020	0037
27					0001	0001	0003	0007	0014	0026
28					000	0001	0002	0005	0010	0018
29						0001	0001	0003	0006	0012
30						0000	0001	0002	0004	0009

Закінчення табл. А.2

Таблиця ймовірностей $\mathrm{P}(\chi^2)$ для критерія Пірсона χ^2 (продовження)

χ^2 \ k	11	12	13	14	15	16	17	18	19	20
1	2	3	4	5	6	7	8	9	10	11
1	0,9999	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000
2	9985	0,9994	0,9998	0,9999	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000
3	9907	9955	9979	9991	0,9996	0,9998	0,9999	1,0000	1,0000	1,0000
4	9699	9834	9912	9955	9977	9989	9995	0,9998	0,9999	1,0000
5	9312	9580	9752	9858	9921	9958	9978	9989	9994	0,9997
6	8734	9161	9462	9665	9797	9881	9932	9962	9979	9989
7	7991	8576	9022	9347	9576	9733	9835	9901	9942	9967
8	7133	7851	8436	8893	9238	9489	9665	9786	9867	9919
9	6219	7029	7729	8311	8775	9134	9403	9597	9735	9829
10	5304	6160	6939	7622	8197	8666	9036	9319	9539	9682
11	4433	5289	6108	6860	7526	8095	8566	8944	9238	9462
12	3626	4457	5276	6063	6790	7440	8001	8472	8856	9161
13	2933	3690	4478	5265	6023	6728	7362	7916	8386	8774
14	2330	3007	3738	4497	5255	5987	6671	7291	7837	8305
15	1825	2414	3074	3782	4514	5246	5955	6620	7226	7764
16	1411	1912	2491	3134	3821	4530	5238	5925	6573	7166
17	1079	1496	1993	2562	3189	3856	4544	5231	5899	6530
18	0816	1157	1575	2068	2627	3239	3888	4557	5224	5874
19	0611	0885	1231	1649	2137	2687	3285	3918	4568	5218
20	0453	0671	0952	1301	1719	2202	2742	3328	3946	4579
21	0334	0504	0729	1016	1368	1785	2263	2794	3368	3971
22	0244	0375	0554	0786	1078	1432	1847	2320	2843	3405
23	0177	0277	0417	0603	0841	1137	1498	1906	2373	2888
24	0127	0203	0311	0458	0651	0895	1194	1550	1962	2424
25	0091	0148	0231	0346	0499	0698	0947	1249	1605	2014
26	0016	0028	0047	0076	0119	0180	0263	0374	0518	0699
27	0,9999	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000
28	9985	0,9994	0,9998	0,9999	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000
29	9907	9955	9979	9991	0,9996	0,9998	0,9999	1,0000	1,0000	1,0000
30	9699	9834	9912	9955	9977	9989	9995	0,9998	0,9999	1,0000

Продовження додатка А

Таблиця А.3

Таблиця розподілу Пірсона (χ^2-розподілу)

k \ α	0,999	0,995	0,99	0,98	0,95	0,9	0,8	0,75	0,70	0,50
1	0,0000016	0,000039	0,00016	0,00063	0,0039	0,02	0,06	0,10	0,15	0,46
2	0,0020	0,0100	0,02010	0,04041	0,1026	0,21	0,45	0,58	0,71	1,39
3	0,0243	0,0717	0,11483	0,18483	0,3518	0,58	1,01	1,21	1,42	2,37
4	0,0908	0,2070	0,29711	0,42940	0,7107	1,06	1,65	1,92	2,20	3,36
5	0,2102	0,4117	0,55430	0,75189	1,1455	1,61	2,34	2,68	3,00	4,35
6	0,3811	0,6757	0,87209	1,13442	1,6354	2,20	3,07	3,46	3,83	5,35
7	0,5985	0,9893	1,23904	1,56429	2,1673	2,83	3,82	4,26	4,67	6,35
8	0,8571	1,3444	1,64650	2,03248	2,7326	3,49	4,59	5,07	5,53	7,34
9	1,1519	1,7349	2,08790	2,53238	3,3251	4,17	5,38	5,90	6,39	8,34
10	1,4787	2,1559	2,55821	3,05905	3,9403	4,87	6,18	6,74	7,27	9,34
11	1,8339	2,6032	3,05348	3,60869	4,5748	5,58	6,99	7,58	8,15	10,34
12	2,2142	3,0738	3,57057	4,17829	5,2260	6,30	7,81	8,44	9,03	11,34
13	2,6172	3,5650	4,10692	4,76545	5,8919	7,04	8,63	9,30	9,93	12,34
14	3,0407	4,0747	4,66043	5,36820	6,5706	7,79	9,47	10,17	10,82	13,34
15	3,4827	4,6009	5,22935	5,98492	7,2609	8,55	10,31	11,04	11,72	14,34
16	3,9416	5,1422	5,81221	6,61424	7,9616	9,31	11,15	11,91	12,62	15,34
17	4,4161	5,6972	6,40776	7,25500	8,6718	10,09	12,00	12,79	13,53	16,34
18	4,9048	6,2648	7,01491	7,90622	9,3905	10,86	12,86	13,68	14,44	17,34
19	5,4068	6,8440	7,63273	8,56704	10,1170	11,65	13,72	14,56	15,35	18,34
20	5,9210	7,4338	8,26040	9,23670	10,8508	12,44	14,58	15,45	16,27	19,34
21	6,4467	8,0337	8,89720	9,91456	11,5913	13,24	15,45	16,34	17,18	20,34
22	6,9830	8,6427	9,54249	10,60003	12,3380	14,04	16,31	17,24	18,10	21,34
23	7,5292	9,2604	10,19572	11,29260	13,0905	14,85	17,19	18,14	19,02	22,34
24	8,0849	9,8862	10,85636	11,99182	13,8484	15,66	18,06	19,04	19,94	23,34
25	8,6493	10,5197	11,52398	12,69727	14,6114	16,47	18,94	19,94	20,87	24,34
26	9,2221	11,1602	12,19815	13,40858	15,3792	17,29	19,82	20,84	21,79	25,34
27	9,8028	11,8076	12,87850	14,12542	16,1514	18,11	20,70	21,75	22,72	26,34
28	10,3909	12,4613	13,56471	14,84748	16,9279	18,94	21,59	22,66	23,65	27,34
29	10,9861	13,1211	14,25645	15,57448	17,7084	19,77	22,48	23,57	24,58	28,34
30	11,5880	13,7867	14,95346	16,30617	18,4927	20,60	23,36	24,48	25,51	29,34

Умовні позначення:
α – рівень значущості; k – число ступенів свободи.

Закінчення додатка А

Таблиця А.4

Таблиця розподілу Пірсона (χ^2-розподілу) (продовження)

k \ α	**0,30**	**0,25**	**0,20**	**0,10**	**0,05**	**0,025**	**0,02**	**0,01**	**0,0050**	**0,001**
1	1,07	1,32	1,64	2,71	3,84	5,02	5,41	6,63	7,88	10,83
2	2,41	2,77	3,22	4,61	5,99	7,38	7,82	9,21	10,60	13,82
3	3,66	4,11	4,64	6,25	7,81	9,35	9,84	11,34	12,84	16,27
4	4,88	5,39	5,99	7,78	9,49	11,14	11,67	13,28	14,86	18,47
5	6,06	6,63	7,29	9,24	11,07	12,83	13,39	15,09	16,75	20,52
6	7,23	7,84	8,56	10,64	12,59	14,45	15,03	16,81	18,55	22,46
7	8,38	9,04	9,80	12,02	14,07	16,01	16,62	18,48	20,28	24,32
8	9,52	10,22	11,03	13,36	15,51	17,53	18,17	20,09	21,95	26,12
9	10,66	11,39	12,24	14,68	16,92	19,02	19,68	21,67	23,59	27,88
10	11,78	12,55	13,44	15,99	18,31	20,48	21,16	23,21	25,19	29,59
11	12,90	13,70	14,63	17,28	19,68	21,92	22,62	24,72	26,76	31,26
12	14,01	14,85	15,81	18,55	21,03	23,34	24,05	26,22	28,30	32,91
13	15,12	15,98	16,98	19,81	22,36	24,74	25,47	27,69	29,82	34,53
14	16,22	17,12	18,15	21,06	23,68	26,12	26,87	29,14	31,32	36,12
15	17,32	18,25	19,31	22,31	25,00	27,49	28,26	30,58	32,80	37,70
16	18,42	19,37	20,47	23,54	26,30	28,85	29,63	32,00	34,27	39,25
17	19,51	20,49	21,61	24,77	27,59	30,19	31,00	33,41	35,72	40,79
18	20,60	21,60	22,76	25,99	28,87	31,53	32,35	34,81	37,16	42,31
19	21,69	22,72	23,90	27,20	30,14	32,85	33,69	36,19	38,58	43,82
20	22,77	23,83	25,04	28,41	31,41	34,17	35,02	37,57	40,00	45,31
21	23,86	24,93	26,17	29,62	32,67	35,48	36,34	38,93	41,40	46,80
22	24,94	26,04	27,30	30,81	33,92	36,78	37,66	40,29	42,80	48,27
23	26,02	27,14	28,43	32,01	35,17	38,08	38,97	41,64	44,18	49,73
24	27,10	28,24	29,55	33,20	36,42	39,36	40,27	42,98	45,56	51,18
25	28,17	29,34	30,68	34,38	37,65	40,65	41,57	44,31	46,93	52,62
26	29,25	30,43	31,79	35,56	38,89	41,92	42,86	45,64	48,29	54,05
27	30,32	31,53	32,91	36,74	40,11	43,19	44,14	46,96	49,64	55,48
28	31,39	32,62	34,03	37,92	41,34	44,46	45,42	48,28	50,99	56,89
29	32,46	33,71	35,14	39,09	42,56	45,72	46,69	49,59	52,34	58,30
30	33,53	34,80	36,25	40,26	43,77	46,98	47,96	50,89	53,67	59,70

Умовні позначення:
α – рівень значущості; k – число ступенів свободи.

Додаток Б

Таблиця Б.1

Таблиця значень щільності ймовірностей (функції Гауса)

$f\ t\ = \frac{1}{2\pi} e^{\frac{-t^2}{2}}$ **для стандартного нормального закону розподілу**

1	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
0,0	0,3989	0,3989	0,3989	0,3988	0,3986	0,3984	0,3982	0,3980	0,3977	0,3973
0,1	3970	3965	3961	3956	3951	3945	3939	3932	3925	3918
0,2	3910	3902	3894	3885	3876	3867	3857	3847	3836	3825
0,3	3814	3802	3790	3778	3765	3752	3739	3725	3712	3697
0,4	3683	3668	3653	3637	3621	3605	3589	3572	3555	3538
0,5	3521	3503	3485	3467	3448	3429	3410	3391	3372	3352
0,6	3332	3312	3292	3271	3251	3230	3209	3187	3166	3144
0,7	3123	3101	3097	3056	3034	3011	2989	2966	2943	2920
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685
0,9	2661	2637	2613	2589	2565	2541	2516	2492	2468	2444
1,0	2420	2396	2371	2347	2323	2299	2275	2251	2227	2203
1,1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965
1,2	1942	1919	1895	1872	1849	1826	1804	1781	1758	1736
1,3	1714	1691	1669	1647	1626	1604	1582	1561	1539	1518
1,4	1497	1476	1456	1435	1415	1394	1374	1354	1334	1315
1,5	1295	1276	1257	1238	1219	1200	1182	1163	1145	1127
1,6	1109	1092	1074	1057	1040	1023	1006	0989	0973	0957
1,7	0940	0925	0909	0893	0878	0863	8048	0833	0818	0804
1,8	0790	0775	0761	0748	0734	0721	0707	0694	0681	0669
1,9	0656	0644	0632	0620	0608	0596	0584	0573	0562	0551
2,0	0540	0529	0519	0508	0498	0488	0478	0468	0459	0449
2,1	0440	0431	0422	0413	0404	0396	0387	0379	0371	0363
2,2	0855	0347	0339	0332	0325	0317	0310	0303	0297	0290
2,3	0283	0277	0270	0264	0258	0252	0246	0241	0235	0229
2,4	0224	0219	0213	0203	0203	0198	0194	0189	0184	0180
2,5	0175	0171	0167	0163	0158	0154	0151	0147	0143	0139
2,6	0136	0132	0129	0126	0122	0119	0116	0113	0110	0107
2,7	0104	0101	0099	0096	0093	0091	0088	0086	0084	0081
2,8	0079	0077	0075	0073	0071	0069	0067	0065	0063	0061
2,9	0060	0058	0056	0055	0053	0051	0050	0048	0047	0046
3,0	0044	0043	0042	0040	0039	0038	0037	0036	0035	0034
4,0	0001	0001	0001	0000	0000	0000	0000	0000	0000	0000

Додаток В

Таблиця В.1

Таблиця функції значень стандартизованого

нормального розподілу $\Phi\ z\ = \frac{\mathbf{1}}{\overline{\mathbf{2\pi}}}\ {}_{-\infty}^{z}\, \boldsymbol{e}^{\frac{-t^2}{2}}\boldsymbol{dt}$

z	0	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,500	0,504	0,508	0,512	0,516	0,520	0,524	0,528	0,532	0,536
0,1	0,540	0,544	0,548	0,552	0,556	0,560	0,564	0,568	0,571	0,575
0,2	0,579	0,583	0,587	0,591	0,595	0,599	0,603	0,606	0,610	0,614
0,3	0,618	0,622	0,626	0,629	0,633	0,637	0,641	0,644	0,648	0,652
0,4	0,655	0,659	0,663	0,666	0,670	0,674	0,677	0,681	0,684	0,688
0,5	0,692	0,695	0,699	0,702	0,705	0,709	0,712	0,716	0,719	0,722
0,6	0,726	0,729	0,732	0,736	0,739	0,742	0,745	0,749	0,752	0,755
0,7	0,758	0,761	0,764	0,767	0,770	0,773	0,776	0,779	0,782	0,785
0,8	0,788	0,791	0,794	0,797	0,800	0,802	0,805	0,808	0,811	0,813
0,9	0,816	0,819	0,821	0,824	0,826	0,829	0,832	0,834	0,837	0,839
1	0,841	0,844	0,846	0,849	0,851	0,853	0,855	0,858	0,860	0,862
1,1	0,864	0,867	0,869	0,871	0,873	0,875	0,877	0,879	0,881	0,883
1,2	0,885	0,887	0,889	0,891	0,893	0,894	0,896	0,898	0,900	0,902
1,3	0,903	0,905	0,907	0,908	0,910	0,912	0,913	0,915	0,916	0,918
1,4	0,919	0,921	0,922	0,924	0,925	0,927	0,928	0,929	0,931	0,932
1,5	0,933	0,935	0,936	0,937	0,938	0,939	0,941	0,942	0,943	0,944
1,6	0,945	0,946	0,947	0,948	0,950	0,951	0,952	0,953	0,954	0,955
1,7	0,955	0,856	0,957	0,958	0,959	0,960	0,961	0,962	0,963	0,963
1,8	0,964	0,965	0,966	0,966	0,967	0,968	0,969	0,969	0,097	0,971
1,9	0,971	0,972	0,973	0,973	0,974	0,974	0,975	0,976	0,976	0,977
2	0,977	0,978	0,978	0,979	0,979	0,980	0,980	0,981	0,981	0,982
2,1	0,982	0,983	0,983	0,983	0,984	0,984	0,985	0,985	0,985	0,986
2,2	0,986	0,986	0,987	0,987	0,988	0,988	0,988	0,988	0,989	0,989
2,3	0,989	0,990	0,990	0,990	0,990	0,991	0,991	0,991	0,991	0,992
2,4	0,992	0,992	0,992	0,993	0,993	0,993	0,993	0,993	0,993	0,994
2,5	0,994	0,994	0,994	0,994	0,995	0,995	0,995	0,995	0,995	0,995
2,6	0,995	0,996	0,996	0,996	0,996	0,996	0,996	0,996	0,996	0,996
2,7	0,997	0,997	0,997	0,997	0,997	0,997	0,100	0,997	0,997	0,997
2,8	0,997	0,998	0,998	0,998	0,998	0,998	0,998	0,998	0,998	0,998
2,9	0,998	0,998	0,998	0,998	0,998	0,998	0,999	0,999	0,999	0,999
3	0,999	0,999	0,999	0,999	0,999	0,999	0,999	0,999	0,999	0,999

Додаток Д

Таблиця Д.1

Таблиця значення функції (інтеграл ймовірностей)

$$\Phi(z) = \int_{-t}^{t} f(x)\,dx = \frac{2}{\sqrt{2\pi}} \int_{0}^{t} e^{\frac{-x^2}{2}}\,dt$$

	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
0,0	0,0000	0,008	0,016	0,0239	0,0319	0,0399	0,0478	0,0558	0,0638	0,0717
0,1	0,0797	0,0876	0,055	0,1034	0,1113	0,1192	0,1271	0,135	0,1428	0,1507
0,2	0,1585	0,1663	0,1741	0,1819	0,1897	0,9774	0,2051	0,2128	0,2205	0,2282
0,3	0,2358	0,2434	0,251	0,2586	0,2661	0,2737	0,2812	0,2886	0,296	0,3035
0,4	0,3108	0,3128	0,3255	0,3328	0,3401	0,3473	0,3545	0,3616	0,3688	0,3759
0,5	0,3829	0,3899	0,3969	0,4039	0,4108	0,4177	0,4245	0,4313	0,4381	0,4448
0,6	0,4515	0,4581	0,4647	0,4713	0,4778	0,4843	0,4907	0,4971	0,5035	0,5098
0,7	0,5161	0,5223	0,5285	0,5346	0,5407	0,5467	0,5527	0,5587	0,5646	0,5705
0,8	0,5763	0,5821	0,5878	0,5935	0,5991	0,6047	0,6102	0,6157	0,6211	0,6265
0,9	0,6319	0,6372	0,6424	0,6476	0,6528	0,6579	0,6629	0,6679	0,6729	0,6778
1,0	0,6825	0,6875	0,6923	0,697	0,7017	0,7063	0,7109	0,7154	0,7199	0,7243
1,1	0,7287	0,733	0,7373	0,7415	0,7457	0,7499	0,754	0,758	0,762	0,766
1,2	0,7699	0,7737	0,7775	0,7813	0,785	0,7887	0,7923	0,7959	0,7994	0,8029
1,3	0,8064	0,8098	0,8132	0,8165	0,8198	0,8282	0,8262	0,8293	0,8324	0,8355
1,4	0,8385	0,8415	0,8444	0,8473	0,8501	0,8529	0,8557	0,8584	0,8611	0,8638
1,5	0,8664	0,869	0,8715	0,874	0,8764	0,8789	0,8812	0,8836	0,8859	0,8882
1,6	0,8904	0,8926	0,8948	0,8969	0,899	0,9011	0,9031	0,9051	0,907	0,909
1,7	0,9109	0,9127	0,9146	0,9164	0,9181	0,9199	0,9216	0,9233	0,9249	0,9265
1,8	0,9281	0,9297	0,9312	0,9327	0,9342	0,9357	0,9371	0,9385	0,9399	0,9412
1,9	0,9426	0,9439	0,9451	0,9446	0,9476	0,9488	0,95	0,9512	0,9523	0,9534
2,0	0,9545	0,9556	0,9566	0,9576	0,9586	0,9596	0,9606	0,9616	0,9625	0,9634
2,1	0,9643	0,9651	0,966	0,9668	0,9666	0,9684	0,9692	0,97	0,9707	0,9715
2,2	0,9722	0,9729	0,9736	0,9743	0,9749	0,9756	0,9762	0,9768	0,9774	0,978
2,3	0,9786	0,9791	0,9797	0,9802	0,9801	0,9812	0,9817	0,9822	0,9827	0,9832
2,4	0,9836	0,9841	0,9845	0,9849	0,9853	0,9857	0,9861	0,9865	0,9869	0,9872
2,5	0,9876	0,9879	0,9883	0,9886	0,9889	0,9892	0,9895	0,9898	0,9901	0,9904
2,6	0,9907	0,991	0,9912	0,9915	0,9917	0,992	0,9922	0,9924	0,9926	0,9928
2,7	0,9931	0,9933	0,9935	0,9937	0,9939	0,994	0,9942	0,9944	0,9946	0,9947
2,8	0,9949	0,9951	0,9952	0,9953	0,9955	0,9956	0,9958	0,9959	0,996	0,9961
2,9	0,9963	0,9964	0,9965	0,9966	0,9967	0,9968	0,9969	0,997	0,9971	0,9972
3,0	0,9973	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,998
3,1	0,9981	0,9982	0,9982	0,9983	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986
3,2	0,9986	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,999	0,999
3,3	0,999	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,4	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995	0,9995
3,5	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997	0,9997
3,6	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998	0,9998	0,9998
3,7	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998
3,8	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
3,9	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
4,0	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999

Додаток Е

Таблиця Е.1

Двосторонні квантилі розподілу Стьюдента $t_{\alpha}(k)$

k \ α	**0,20**	**0,40**	**0,50**	**0,60**	**0,80**	**0,90**	**0,95**	**0,98**	**0,99**
1	0,325	0,727	1,000	1,376	3,078	6,314	12,706	31,821	63,657
2	0,289	0,262	0,816	1,061	1,886	2,920	4,303	6,965	9,925
3	0,277	0,584	0,765	0,978	1,638	2,353	3,182	4,541	5,841
4	0,271	0,569	0,741	0,941	1,533	2,132	2,776	3,747	4,604
5	0,267	0,559	0,727	0,920	1,476	2,015	2,571	3,365	4,032
6	0,265	0,553	0,718	0,906	1,440	1,943	2,447	3,143	3,707
7	0,263	0,549	0,711	0,896	1,415	1,895	2,365	2,998	3,499
8	0,262	0,546	0,706	0,889	1,379	1,860	2,306	2,896	3,355
9	0,261	0,543	0,703	0,883	1,383	1,833	2,262	2,821	3,250
10	0,260	0,542	0,700	0,879	1,372	1,812	2,228	2,764	3,169
11	0,260	0,540	0,697	0,876	1,363	1,796	2,201	2,718	3,106
12	0,259	0,539	0,695	0,873	1,356	1,782	2,179	2,681	3,055
13	0,259	0,538	0,694	0,870	1,350	1,771	2,160	2,650	3,012
14	0,258	0,537	0,692	0,868	1,345	1,761	2,145	2,624	2,977
15	0,258	0,536	0,691	0,866	1,341	1,753	2,131	2,602	2,947
16	0,258	0,535	0,690	0,865	1,337	1,746	2,120	2,583	2,921
17	0,257	0,534	0,689	0,863	1,333	1,740	2,110	2,567	2,898
18	0,257	0,534	0,688	0,862	1,330	1,734	2,101	2,552	2,878
19	0,257	0,533	0,688	0,861	1,328	1,729	2,093	2,539	2,861
20	0,257	0,533	0,687	0,860	1,325	1,725	2,086	2,528	2,845
21	0,257	0,532	0,686	0,859	1,323	1,721	2,080	2,518	2,831
22	0,256	0,532	0,686	0,858	1,321	1,717	2,074	2,508	2,819
23	0,256	0,532	0,685	0,858	1,319	1,714	2,069	2,500	2,807
24	0,256	0,531	0,685	0,857	1,318	7,711	2,064	2,492	2,797
25	0,256	0,531	0,684	0,856	1,316	1,708	2,060	2,485	2,787
26	0,256	0,531	0,684	0,856	1,315	1,706	2,056	2,479	2,779
27	0,256	0,531	0,684	0,855	1,314	1,703	2,052	2,473	2,771
28	0,256	0,530	0,683	0,855	1,313	1,701	2,048	2,467	2,763
29	0,256	0,530	0,683	0,854	1,311	1,699	2,045	2,462	2,756
30	0,256	0,530	0,683	0,854	1,310	1,697	2,042	2,457	2,750
40	0,255	0,529	0,681	0,851	1,303	1,684	2,021	2,423	2,704
60	0,254	0,527	0,679	0,848	1,296	1,671	2,000	2,390	2,660
100	0,254	0,526	0,677	0,845	1,290	1,660	1,984	2,364	2,626
200	0,254	0,525	0,676	0,843	1,286	1,652	1,972	2,345	2,601
∞	0,253	0,524	0,675	0,842	1,282	1,645	1,960	2,326	2,576

Умовні позначення:
α – рівень значущості; k – число ступенів свободи.

Додаток Ж

Таблиця Ж.1

Таблиця розподілу Фішера для $\alpha = 0,05$ (5 %)

$k2$ \ $k1$	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
1	161	200	216	225	230	234	237	239	241	242	243
2	18,5	19,0	19,2	19,2	19,3	19,3	19,4	19,4	19,4	19,4	19,4
3	10,10	9,55	9,28	9,20	9,01	8,94	8,89	8,85	8,81	8,79	8,76
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,94
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,70
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,03
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,60
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,31
9	5,12	4,26	3,68	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,10
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,94
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,82
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,72
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	5,67	2,63
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,57
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,51
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,46
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,41
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,37
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,34
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,31
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,28
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30	2,26
23	4,48	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27	2,24
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	2,22
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24	2,20
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	2,13
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	2,04
50	4,03	3,18	2,79	2,56	2,40	2,29	2,20	2,13	2,07	2,03	1,99
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99	1,95
70	3,98	3,13	2,74	2,50	2,35	2,23	2,14	2,07	2,02	1,97	1,93
80	3,96	3,11	2,72	2,49	2,33	2,21	2,13	2,06	2,00	1,95	1,91
90	3,95	3,10	2,71	2,47	2,32	2,20	2,11	2,04	1,99	1,94	1,90
100	3,94	3,09	2,70	2,46	2,31	2,19	2,10	2,03	1,97	1,93	1,89
120	3,92	3,07	2,68	2,45	2,29	2,18	2,09	2,02	1,96	1,91	1,87
140	3,91	3,06	2,67	2,44	2,28	2,16	2,08	2,01	1,95	1,90	1,86
160	3,90	3,05	2,66	2,43	2,27	2,16	2,07	2,00	1,94	1,89	1,85
180	3,89	3,05	2,65	2,42	2,26	2,15	2,06	1,99	1,93	1,88	1,84
200	3,88	3,04	2,65	2,42	2,26	2,14	2,06	1,98	1,93	1,88	1,84
∞	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,88	1,83	1,79

Закінчення табл. Ж.1

Таблиця розподілу Фішера для $\alpha = 0,05$ (5 %)

$k2$ \ $k1$	12	13	14	15	20	30	40	50	100	200	∞
1	244	245	245	246	248	250	251	252	253	254	254
2	19,4	19,4	19,4	19,4	19,4	19,5	19,5	19,5	19,5	19,5	19,5
3	8,74	8,73	8,71	8,70	8,66	8,62	8,60	8,58	8,55	8,54	8,53
4	5,91	5,89	5,87	5,86	5,80	5,75	5,72	5,70	5,66	5,65	5,63
5	4,68	4,66	4,64	4,62	4,56	4,50	4,46	4,44	4,41	4,39	4,37
6	4,00	3,98	3,96	3,94	3,87	3,81	3,77	3,75	3,71	3,69	3,67
7	3,57	3,55	3,53	3,51	3,44	3,38	3,34	3,32	3,27	3,25	3,23
8	3,28	3,26	3,24	3,22	3,15	3,08	3,04	3,02	2,97	2,96	2,93
9	3,07	3,05	3,03	3,01	2,94	2,86	2,83	2,80	2,76	2,73	2,71
10	2,91	2,98	2,86	2,85	2,77	2,70	2,66	2,64	2,59	2,56	2,54
11	2,79	2,76	2,74	2,72	2,65	2,57	2,53	2,51	2,46	2,43	2,40
12	2,69	2,66	2,64	2,62	2,54	2,47	2,43	2,40	2,36	2,32	2,30
13	2,60	2,58	2,55	2,53	2,46	2,38	2,34	2,31	2,26	2,23	2,21
14	2,53	2,51	2,48	2,46	2,39	2,31	2,27	2,24	2,19	2,16	2,13
15	2,48	2,45	2,42	2,40	2,33	2,25	2,20	2,18	2,12	2,10	2,07
16	2,42	2,40	2,37	2,35	2,28	2,19	2,15	2,12	2,07	2,04	2,01
17	2,38	2,35	2,33	2,31	2,23	2,15	2,10	2,08	2,02	1,99	1,96
18	2,34	2,31	2,29	2,27	2,19	2,11	2,10	2,08	1,98	1,95	1,92
19	2,31	2,28	2,26	2,23	2,16	2,07	2,03	2,00	1,94	1,91	1,88
20	2,28	2,25	2,22	2,20	2,12	2,04	1,99	1,97	1,91	1,88	1,84
21	2,25	2,22	2,20	2,18	2,10	2,01	1,96	1,94	1,88	1,84	1,81
22	2,23	2,20	2,17	2,15	2,07	1,98	1,94	1,91	1,85	1,82	1,78
23	2,20	2,18	2,15	2,13	2,05	1,96	1,91	1,88	1,82	1,79	1,76
24	2,18	2,15	2,13	2,11	2,03	1,94	1,89	1,86	1,80	1,77	1,73
25	2,16	2,14	2,11	2,09	2,01	1,92	1,87	1,84	1,78	1,50	1,71
30	2,09	2,06	2,04	2,01	1,93	1,84	1,79	1,76	1,70	1,66	1,62
40	2,00	1,97	1,95	1,92	1,84	1,74	1,69	1,66	1,63	1,60	1,51
50	1,95	1,92	1,89	1,87	1,78	1,69	1,63	1,60	1,59	1,55	1,44
60	1,92	1,89	1,86	1,84	1,75	1,65	1,59	1,56	1,48	1,44	1,39
70	1,89	1,86	1,84	1,81	1,72	1,62	1,57	1,53	1,45	1,40	1,35
80	1,88	1,84	1,82	1,79	1,70	1,60	1,54	1,51	1,43	1,38	1,32
90	1,86	1,83	1,80	1,78	1,69	1,59	1,53	1,49	1,41	1,36	1,30
100	1,85	1,82	1,79	1,77	1,68	1,57	1,52	1,48	1,39	1,34	1,28
120	1,83	1,80	1,78	1,75	1,66	1,55	1,50	1,46	1,37	1,32	1,25
140	1,82	1,79	1,76	1,74	1,65	1,54	1,48	1,44	1,35	1,30	1,23
160	1,81	1,78	1,75	1,73	1,64	1,53	1,47	1,43	1,34	1,28	1,21
180	1,81	1,77	1,75	1,72	1,63	1,52	1,46	1,42	1,33	1,27	1,20
200	1,80	1,77	1,74	1,72	1,62	1,52	1,46	1,41	1,32	1,26	1,19
∞	1,75	1,72	1,69	1,67	1,57	1,46	1,39	1,35	1,24	1,17	1,00

Умовні позначення:

$k1$, $k2$ – число ступенів свободи чисельника та знаменника.

Продовження додатка Ж

Таблиця Ж.2

Таблиця розподілу Фішера для $\alpha = 0,01\ (1\ \%)$

$k2 \backslash k1$	1	2	3	4	5	6	7	8	9	10	11
1	4052,2	4999,5	5403,4	5624,6	5763,6	5858,9	5928,4	5981,1	6022,5	6055,9	6083,3
2	98,50	99,00	99,17	99,25	99,30	99,33	99,36	99,37	99,39	99,40	99,41
3	34,12	30,82	29,46	28,71	28,24	27,91	27,67	27,49	27,35	27,23	27,13
4	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66	14,55	14,45
5	16,26	13,27	12,06	11,39	10,97	10,67	10,46	10,29	10,16	10,05	9,96
6	13,75	10,92	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87	7,79
7	12,25	9,55	8,45	7,85	7,46	7,19	6,99	6,84	6,72	6,62	6,54
8	11,26	8,65	7,59	7,01	6,63	6,37	6,18	6,03	5,91	5,81	5,73
9	10,56	8,02	6,99	6,42	6,06	5,80	5,61	5,47	5,35	5,26	5,18
10	10,04	7,56	6,55	5,99	5,64	5,39	5,20	5,06	4,94	4,85	4,77
11	9,65	7,21	6,22	5,67	5,32	5,07	4,89	4,74	4,63	4,54	4,46
12	9,33	6,93	5,95	5,41	5,06	4,82	4,64	4,50	4,39	4,30	4,22
13	9,07	6,70	5,74	5,21	4,86	4,62	4,44	4,30	4,19	4,10	4,02
14	8,86	6,51	5,56	5,04	4,69	4,46	4,28	4,14	4,03	3,94	3,86
15	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80	3,73
16	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89	3,78	3,69	3,62
17	8,40	6,11	5,18	4,67	4,34	4,10	3,93	3,79	3,68	3,59	3,52
18	8,29	6,01	5,09	4,58	4,25	4,01	3,84	3,71	3,60	3,51	3,43
19	8,18	5,93	5,01	4,50	4,17	3,94	3,77	3,63	3,52	3,43	3,36
20	8,10	5,85	4,94	4,43	4,10	3,87	3,70	3,56	3,46	3,37	3,29
21	8,02	5,78	4,87	4,37	4,04	3,81	3,64	3,51	3,40	3,31	3,24
22	7,95	5,72	4,82	4,31	3,99	3,76	3,59	3,45	3,35	3,26	3,18
23	7,88	5,66	4,76	4,26	3,94	3,71	3,54	3,41	3,30	3,21	3,14
24	7,82	5,61	4,72	4,22	3,90	3,67	3,50	3,36	3,26	3,17	3,09
25	7,77	5,57	4,68	4,18	3,85	3,63	3,46	3,32	3,22	3,13	3,06
30	7,72	5,53	4,64	4,14	3,82	3,59	3,42	3,29	3,18	3,09	3,02
40	7,68	5,49	4,60	4,11	3,78	3,56	3,39	3,26	3,15	3,06	2,99
50	7,64	5,45	4,57	4,07	3,75	3,53	3,36	3,23	3,12	3,03	2,96
60	7,60	5,42	4,54	4,04	3,73	3,50	3,33	3,20	3,09	3,00	2,93
70	7,56	5,39	4,51	4,02	3,70	3,47	3,30	3,17	3,07	2,98	2,91
80	7,31	5,18	4,31	3,83	3,51	3,29	3,12	2,99	2,89	2,80	2,73
90	7,17	5,06	4,20	3,72	3,41	3,19	3,02	2,89	2,78	2,70	2,63
100	7,08	4,98	4,13	3,65	3,34	3,12	2,95	2,82	2,72	2,63	2,56
120	7,01	4,92	4,07	3,60	3,29	3,07	2,91	2,78	2,67	2,59	2,51
140	6,96	4,88	4,04	3,56	3,26	3,04	2,87	2,74	2,64	2,55	2,48
160	6,93	4,85	4,01	3,53	3,23	3,01	2,84	2,72	2,61	2,52	2,45
180	6,90	4,82	3,98	3,51	3,21	2,99	2,82	2,69	2,59	2,50	2,43
200	6,85	4,79	3,95	3,48	3,17	2,96	2,79	2,66	2,56	2,47	2,40
∞	6,82	4,76	3,92	3,46	3,15	2,93	2,77	2,64	2,54	2,45	2,38

Продовження додатка Ж

Закінчення табл. Ж.2

Таблиця розподілу Фішера для $\alpha = 0,01$ (1 %)

$k2$ \ $k1$	12	13	14	15	20	30	40	50	100	200	∞
1	6106,3	6125,8	6142,7	6157,3	6208,7	6260,6	6286,8	6302,5	6334,1	6349,9	6362,7
2	99,42	99,42	99,43	99,43	99,45	99,47	99,47	99,48	99,49	99,49	99,50
3	27,05	26,98	26,92	26,87	26,69	26,50	26,41	26,35	26,24	26,18	26,14
4	14,37	14,31	14,25	14,20	14,02	13,84	13,75	13,69	13,58	13,52	13,47
5	9,89	9,82	9,77	9,72	9,55	9,38	9,29	9,24	9,13	9,08	9,03
6	7,72	7,66	7,60	7,56	7,40	7,23	7,14	7,09	6,99	6,93	6,89
7	6,47	6,41	6,36	6,31	6,16	5,99	5,91	5,86	5,75	5,70	5,66
8	5,67	5,61	5,56	5,52	5,36	5,20	5,12	5,07	4,96	4,91	4,87
9	5,11	5,05	5,01	4,96	4,81	4,65	4,57	4,52	4,41	4,36	4,32
10	4,71	4,65	4,60	4,56	4,41	4,25	4,17	4,12	4,01	3,96	3,92
11	4,40	4,34	4,29	4,25	4,10	3,94	3,86	3,81	3,71	3,66	3,61
12	4,16	4,10	4,05	4,01	3,86	3,70	3,62	3,57	3,47	3,41	3,37
13	3,96	3,91	3,86	3,82	3,66	3,51	3,43	3,38	3,27	3,22	3,18
14	3,80	3,75	3,70	3,66	3,51	3,35	3,27	3,22	3,11	3,06	3,02
15	3,67	3,61	3,56	3,52	3,37	3,21	3,13	3,08	2,98	2,92	2,88
16	3,55	3,50	3,45	3,41	3,26	3,10	3,02	2,97	2,86	2,81	2,76
17	3,46	3,40	3,35	3,31	3,16	3,00	2,92	2,87	2,76	2,71	2,66
18	3,37	3,32	3,27	3,23	3,08	2,92	2,84	2,78	2,68	2,62	2,58
19	3,30	3,24	3,19	3,15	3,00	2,84	2,76	2,71	2,60	2,55	2,50
20	3,23	3,18	3,13	3,09	2,94	2,78	2,69	2,64	2,54	2,48	2,43
21	3,17	3,12	3,07	3,03	2,88	2,72	2,64	2,58	2,48	2,42	2,37
22	3,12	3,07	3,02	2,98	2,83	2,67	2,58	2,53	2,42	2,36	2,32
23	3,07	3,02	2,97	2,93	2,78	2,62	2,54	2,48	2,37	2,32	2,27
24	3,03	2,98	2,93	2,89	2,74	2,58	2,49	2,44	2,33	2,27	2,22
25	2,99	2,94	2,89	2,85	2,70	2,54	2,45	2,40	2,29	2,23	2,18
30	2,96	2,90	2,86	2,81	2,66	2,50	2,42	2,36	2,25	2,19	2,14
40	2,93	2,87	2,82	2,78	2,63	2,47	2,38	2,33	2,22	2,16	2,11
50	2,90	2,84	2,79	2,75	2,60	2,44	2,35	2,30	2,19	2,13	2,08
60	2,87	2,81	2,77	2,73	2,57	2,41	2,33	2,27	2,16	2,10	2,05
70	2,84	2,79	2,74	2,70	2,55	2,39	2,30	2,25	2,13	2,07	2,02
80	2,66	2,61	2,56	2,52	2,37	2,20	2,11	2,06	1,94	1,87	1,82
90	2,56	2,51	2,46	2,42	2,27	2,10	2,01	1,95	1,82	1,76	1,70
100	2,50	2,44	2,39	2,35	2,20	2,03	1,94	1,88	1,75	1,68	1,62
120	2,45	2,40	2,35	2,31	2,15	1,98	1,89	1,83	1,70	1,62	1,56
140	2,42	2,36	2,31	2,27	2,12	1,94	1,85	1,79	1,65	1,58	1,51
160	2,39	2,33	2,29	2,24	2,09	1,92	1,82	1,76	1,62	1,55	1,48
180	2,37	2,31	2,27	2,22	2,07	1,89	1,80	1,74	1,60	1,52	1,45
200	2,34	2,28	2,23	2,19	2,03	1,86	1,76	1,70	1,56	1,48	1,40
∞	2,31	2,26	2,21	2,17	2,01	1,84	1,74	1,67	1,53	1,45	1,37

Умовні позначення:

$k1$, $k2$ – число ступенів свободи чисельника та знаменника.

Продовження додатка Ж

Таблиця Ж.3

Таблиця розподілу Фішера для $\alpha = 0,1\ (10\ \%)$

$k2$ \ $k1$	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
1	39,86	49,50	53,59	55,83	57,24	58,20	58,91	59,44	59,86	60,19	60,47
2	8,53	9,00	9,16	9,24	9,29	9,33	9,35	9,37	9,38	9,39	9,40
3	5,54	5,46	5,39	5,34	5,31	5,28	5,27	5,25	5,24	5,23	5,22
4	4,54	4,32	4,19	4,11	4,05	4,01	3,98	3,95	3,94	3,92	3,91
5	4,06	3,78	3,62	3,52	3,45	3,40	3,37	3,34	3,32	3,30	3,28
6	3,78	3,46	3,29	3,18	3,11	3,05	3,01	2,98	2,96	2,94	2,92
7	3,59	3,26	3,07	2,96	2,88	2,83	2,78	2,75	2,72	2,70	2,68
8	3,46	3,11	2,92	2,81	2,73	2,67	2,62	2,59	2,56	2,54	2,52
9	3,36	3,01	2,81	2,69	2,61	2,55	2,51	2,47	2,44	2,42	2,40
10	3,29	2,92	2,73	2,61	2,52	2,46	2,41	2,38	2,35	2,32	2,30
11	3,23	2,86	2,66	2,54	2,45	2,39	2,34	2,30	2,27	2,25	2,23
12	3,18	2,81	2,61	2,48	2,39	2,33	2,28	2,24	2,21	2,19	2,17
13	3,14	2,76	2,56	2,43	2,35	2,28	2,23	2,20	2,16	2,14	2,12
14	3,10	2,73	2,52	2,39	2,31	2,24	2,19	2,15	2,12	2,10	2,07
15	3,07	2,70	2,49	2,36	2,27	2,21	2,16	2,12	2,09	2,06	2,04
16	3,05	2,67	2,46	2,33	2,24	2,18	2,13	2,09	2,06	2,03	2,01
17	3,03	2,64	2,44	2,31	2,22	2,15	2,10	2,06	2,03	2,00	1,98
18	3,01	2,62	2,42	2,29	2,20	2,13	2,08	2,04	2,00	1,98	1,95
19	2,99	2,61	2,40	2,27	2,18	2,11	2,06	2,02	1,98	1,96	1,93
20	2,97	2,59	2,38	2,25	2,16	2,09	2,04	2,00	1,96	1,94	1,91
21	2,96	2,57	2,36	2,23	2,14	2,08	2,02	1,98	1,95	1,92	1,90
22	2,95	2,56	2,35	2,22	2,13	2,06	2,01	1,97	1,93	1,90	1,88
23	2,94	2,55	2,34	2,21	2,11	2,05	1,99	1,95	1,92	1,89	1,87
24	2,93	2,54	2,33	2,19	2,10	2,04	1,98	1,94	1,91	1,88	1,85
25	2,92	2,53	2,32	2,18	2,09	2,02	1,97	1,93	1,89	1,87	1,84
30	2,91	2,52	2,31	2,17	2,08	2,01	1,96	1,92	1,88	1,86	1,83
40	2,90	2,51	2,30	2,17	2,07	2,00	1,95	1,91	1,87	1,85	1,82
50	2,89	2,50	2,29	2,16	2,06	2,00	1,94	1,90	1,87	1,84	1,81
60	2,89	2,50	2,28	2,15	2,06	1,99	1,93	1,89	1,86	1,83	1,80
70	2,88	2,49	2,28	2,14	2,05	1,98	1,93	1,88	1,85	1,82	1,79
80	2,84	2,44	2,23	2,09	2,00	1,93	1,87	1,83	1,79	1,76	1,74
90	2,81	2,41	2,20	2,06	1,97	1,90	1,84	1,80	1,76	1,73	1,70
100	2,79	2,39	2,18	2,04	1,95	1,87	1,82	1,77	1,74	1,71	1,68
120	2,78	2,38	2,16	2,03	1,93	1,86	1,80	1,76	1,72	1,69	1,66
140	2,77	2,37	2,15	2,02	1,92	1,85	1,79	1,75	1,71	1,68	1,65
160	2,76	2,36	2,15	2,01	1,91	1,84	1,78	1,74	1,70	1,67	1,64
180	2,76	2,36	2,14	2,00	1,91	1,83	1,78	1,73	1,69	1,66	1,64
200	2,75	2,35	2,13	1,99	1,90	1,82	1,77	1,72	1,68	1,65	1,63
∞	2,74	2,34	2,12	1,99	1,89	1,82	1,76	1,71	1,68	1,64	1,62

Продовження додатка Ж

Закінчення табл. Ж.3

Таблиця розподілу Фішера для $\alpha = 0,1\ (10\ \%)$

$k2$ \ $k1$	12	13	14	15	20	30	40	50	100	200	∞
1	60,71	60,90	61,07	61,22	61,74	62,26	62,53	62,69	63,01	63,17	63,30
2	9,41	9,41	9,42	9,42	9,44	9,46	9,47	9,47	9,48	9,49	9,49
3	5,22	5,21	5,20	5,20	5,18	5,17	5,16	5,15	5,14	5,14	5,13
4	3,90	3,89	3,88	3,87	3,84	3,82	3,80	3,80	3,78	3,77	3,76
5	3,27	3,26	3,25	3,24	3,21	3,17	3,16	3,15	3,13	3,12	3,11
6	2,90	2,89	2,88	2,87	2,84	2,80	2,78	2,77	2,75	2,73	2,72
7	2,67	2,65	2,64	2,63	2,59	2,56	2,54	2,52	2,50	2,48	2,47
8	2,50	2,49	2,48	2,46	2,42	2,38	2,36	2,35	2,32	2,31	2,30
9	2,38	2,36	2,35	2,34	2,30	2,25	2,23	2,22	2,19	2,17	2,16
10	2,28	2,27	2,26	2,24	2,20	2,16	2,13	2,12	2,09	2,07	2,06
11	2,21	2,19	2,18	2,17	2,12	2,08	2,05	2,04	2,01	1,99	1,98
12	2,15	2,13	2,12	2,10	2,06	2,01	1,99	1,97	1,94	1,92	1,91
13	2,10	2,08	2,07	2,05	2,01	1,96	1,93	1,92	1,88	1,86	1,85
14	2,05	2,04	2,02	2,01	1,96	1,91	1,89	1,87	1,83	1,82	1,80
15	2,02	2,00	1,99	1,97	1,92	1,87	1,85	1,83	1,79	1,77	1,76
16	1,99	1,97	1,95	1,94	1,89	1,84	1,81	1,79	1,76	1,74	1,72
17	1,96	1,94	1,93	1,91	1,86	1,81	1,78	1,76	1,73	1,71	1,69
18	1,93	1,92	1,90	1,89	1,84	1,78	1,75	1,74	1,70	1,68	1,66
19	1,91	1,89	1,88	1,86	1,81	1,76	1,73	1,71	1,67	1,65	1,64
20	1,89	1,87	1,86	1,84	1,79	1,74	1,71	1,69	1,65	1,63	1,61
21	1,87	1,86	1,84	1,83	1,78	1,72	1,69	1,67	1,63	1,61	1,59
22	1,86	1,84	1,83	1,81	1,76	1,70	1,67	1,65	1,61	1,59	1,57
23	1,84	1,83	1,81	1,80	1,74	1,69	1,66	1,64	1,59	1,57	1,55
24	1,83	1,81	1,80	1,78	1,73	1,67	1,64	1,62	1,58	1,56	1,54
25	1,82	1,80	1,79	1,77	1,72	1,66	1,63	1,61	1,56	1,54	1,52
30	1,81	1,79	1,77	1,76	1,71	1,65	1,61	1,59	1,55	1,53	1,51
40	1,80	1,78	1,76	1,75	1,70	1,64	1,60	1,58	1,54	1,52	1,50
50	1,79	1,77	1,75	1,74	1,69	1,63	1,59	1,57	1,53	1,50	1,48
60	1,78	1,76	1,75	1,73	1,68	1,62	1,58	1,56	1,52	1,49	1,47
70	1,77	1,75	1,74	1,72	1,67	1,61	1,57	1,55	1,51	1,48	1,46
80	1,71	1,70	1,68	1,66	1,61	1,54	1,51	1,48	1,43	1,41	1,38
90	1,68	1,66	1,64	1,63	1,57	1,50	1,46	1,44	1,39	1,36	1,33
100	1,66	1,64	1,62	1,60	1,54	1,48	1,44	1,41	1,36	1,33	1,30
120	1,64	1,62	1,60	1,59	1,53	1,46	1,42	1,39	1,34	1,30	1,27
140	1,63	1,61	1,59	1,57	1,51	1,44	1,40	1,38	1,32	1,28	1,25
160	1,62	1,60	1,58	1,56	1,50	1,43	1,39	1,36	1,30	1,27	1,24
180	1,61	1,59	1,57	1,56	1,49	1,42	1,38	1,35	1,29	1,26	1,22
200	1,60	1,58	1,56	1,55	1,48	1,41	1,37	1,34	1,28	1,24	1,20
∞	1,59	1,57	1,55	1,54	1,47	1,40	1,36	1,33	1,26	1,22	1,19

Умовні позначення:

$k1,\ k2$ – число ступенів свободи чисельника та знаменника.

Закінчення додатка Ж

Таблиця Ж.4

Таблиця Z-перетворення Фішера

(значення $Z = \frac{1}{2} \cdot ln\frac{1+r}{1-r}$)

	0	1	2	3	4	5	6	7	8	9
0,0	0,000	0,010	0,020	0,030	0,040	0,050	0,060	0,070	0,080	0,090
0,1	0,100	0,110	0,121	0,131	0,141	1,151	0,161	0,172	0,182	0,192
0,2	0,203	0,213	0,224	0,234	0,245	0,255	0,266	0,277	0,288	0,299
0,3	0,310	0,321	0,332	0,343	0,354	0,365	0,377	0,388	0,400	0,412
0,4	0,424	0,436	0,448	0,460	0,472	0,485	0,497	0,510	0,523	0,536
0,5	0,549	0,563	0,576	0,590	0,604	0,618	0,633	0,648	0,663	0,678
0,6	0,693	0,709	0,725	0,741	0,758	0,775	0,793	0,811	0,829	0,848
0,7	0,867	0,887	0,908	0,929	0,951	0,973	0,996	1,020	1,045	1,071
0,8	1,099	1,127	1,157	1,188	1,221	1,256	1,293	1,333	1,376	1,422
0,9	1,472	1,528	1,589	1,658	1,738	1,832	1,946	2,092	2,298	2,647
1,0	2,647	2,700	2,759	2,826	2,903	2,995	3,106	3,250	3,453	3,800

Додаток З

Таблиця З.1

Критичні значення коефіцієнтів рангової кореляції Спірмена

n \ α	**0,05**	**0,01**	n \ α	**0,05**	**0,01**	n \ α	**0,05**	**0,01**
5	0,94		**17**	0,48	0,62	**29**	0,37	0,48
6	0,85		**18**	0,47	0,6	**30**	0,36	0,47
7	0,78	0,94	**19**	0,46	0,58	**31**	0,36	0,46
8	0,72	0,88	**20**	0,45	0,57	**32**	0,36	0,45
9	0,68	0,83	**21**	0,44	0,56	**33**	0,34	0,45
10	0,64	0,79	**22**	0,43	0,54	**34**	0,34	0,44
11	0,61	0,76	**23**	0,42	0,53	**35**	0,33	0,43
12	0,58	0,73	**24**	0,41	0,52	**36**	0,33	0,43
13	0,56	0,7	**25**	0,4	0,51	**37**	0,33	0,43
14	0,54	0,68	**26**	0,39	0,5	**38**	0,32	0,41
15	0,52	0,66	**27**	0,38	0,49	**39**	0,32	0,41
16	0,5	0,64	**28**	0,38	0,48	**40**	0,31	0,4

Умовні позначення:
α – рівень значущості; n – кількість спостережень у вибірці.

Додаток И

Таблиця И.1

Критичні значення коефіцієнтів лінійної кореляції Пірсона

n \ α	**0,1**	**0,05**	**0,01**	n \ α	**0,1**	**0,05**	**0,01**	n \ α	**0,1**	**0,05**	**0,01**
5	0,805	0,878	0,959	**32**	0,296	0,349	0,449	**59**	0,216	0,256	0,418
6	0,729	0,811	0,917	**33**	0,291	0,344	0,442	**60**	0,214	0,254	0,414
7	0,669	0,754	0,875	**34**	0,287	0,339	0,436	**61**	0,213	0,252	0,411
8	0,621	0,707	0,834	**35**	0,283	0,334	0,43	**62**	0,211	0,25	0,408
9	0,582	0,666	0,798	**36**	0,279	0,329	0,424	**63**	0,209	0,248	0,405
10	0,549	0,632	0,765	**37**	0,275	0,325	0,418	**64**	0,207	0,246	0,402
11	0,521	0,602	0,735	**38**	0,271	0,32	0,413	**65**	0,206	0,244	0,399
12	0,497	0,576	0,708	**39**	0,276	0,316	0,408	**66**	0,204	0,242	0,396
13	0,476	0,553	0,684	**40**	0,264	0,312	0,403	**67**	0,203	0,24	0,393
14	0,458	0,532	0,661	**41**	0,26	0,308	0,398	**68**	0,201	0,239	0,39
15	0,441	0,514	0,641	**42**	0,257	0,304	0,393	**69**	0,2	0,237	0,388
16	0,426	0,497	0,623	**43**	0,254	0,301	0,389	**70**	0,198	0,235	0,385
17	0,412	0,482	0,606	**44**	0,251	0,297	0,384	**80**	0,185	0,22	0,361
18	0,4	0,468	0,59	**45**	0,248	0,294	0,38	**90**	0,174	0,207	0,341
19	0,389	0,456	0,575	**46**	0,246	0,291	0,469	**100**	0,165	0,197	0,324
20	0,378	0,444	0,561	**47**	0,243	0,288	0,465	**110**	0,158	0,187	0,31
21	0,369	0,433	0,549	**48**	0,24	0,285	0,46	**120**	0,151	0,179	0,297
22	0,36	0,423	0,537	**49**	0,238	0,282	0,456	**130**	0,145	0,172	0,285
23	0,352	0,413	0,526	**50**	0,235	0,279	0,451	**140**	0,14	0,166	0,275
24	0,344	0,404	0,515	**51**	0,233	0,276	0,447	**150**	0,135	0,16	0,266
25	0,337	0,396	0,505	**52**	0,231	0,273	0,443	**200**	0,117	0,139	0,231
26	0,33	0,388	0,436	**53**	0,228	0,271	0,439	**250**	0,104	0,124	0,207
27	0,323	0,381	0,487	**54**	0,226	0,268	0,435	**300**	0,095	0,113	0,189
28	0,317	0,374	0,479	**55**	0,224	0,266	0,432	**350**	0,088	0,105	0,175
29	0,311	0,367	0,471	**56**	0,222	0,263	0,428	**400**	0,082	0,098	0,164
30	0,306	0,361	0,463	**57**	0,22	0,261	0,424	**500**	0,074	0,088	0,147
31	0,301	0,355	0,456	**58**	0,218	0,259	0,421				

Умовні позначення:
α – рівень значущості; n – кількість спостережень у виборці.

Додаток К

Таблиця К.1

Критерій Дарбіна – Уотсона для $\alpha = 0,05$ (5 %)

	$m=1$		$m=2$		$m=3$		$m=4$		$m=5$	
n	dl	du	dl	du	dl	du	dl	du	dl	du
6	0,61	1,40								
7	0,70	1,36	0,47	1,90						
8	0,76	1,33	0,56	1,78	0,37	2,29				
9	0,82	1,32	0,63	1,70	0,46	2,13	0,30	2,59		
10	0,88	1,32	0,70	1,64	0,53	2,02	0,38	2,41	0,24	2,81
11	0,93	1,32	0,76	1,60	0,60	1,93	0,44	2,28	0,32	2,65
12	0,97	1,33	0,81	1,58	0,66	1,86	0,51	2,18	0,38	2,51
13	1,01	1,34	0,86	1,56	0,72	1,82	0,57	2,09	0,45	2,39
14	1,05	1,35	0,91	1,55	0,77	1,78	0,63	2,03	0,51	2,30
15	1,08	1,36	0,95	1,54	0,81	1,75	0,69	1,96	0,56	2,22
16	1,11	1,37	0,98	1,54	0,86	1,73	0,73	1,94	0,62	2,16
17	1,13	1,38	1,02	1,54	0,90	1,71	0,78	1,90	0,66	2,10
18	1,16	1,39	1,05	1,54	0,93	1,70	0,82	1,87	0,71	2,06
19	1,18	1,40	1,07	1,54	0,97	1,69	0,96	1,85	0,75	2,02
20	1,20	1,41	1,10	1,54	1,00	1,68	0,89	1,83	0,79	1,99
21	1,22	1,42	1,13	1,54	1,03	1,67	0,93	1,81	0,83	1,96
22	1,24	1,43	1,15	1,54	1,05	1,66	0,96	1,80	0,86	1,94
23	1,26	1,44	1,17	1,54	1,08	1,66	0,99	1,79	0,90	1,92
24	1,27	1,45	1,19	1,55	1,10	1,66	1,01	1,78	0,93	1,90
25	1,29	1,45	1,21	1,55	1,12	1,65	1,04	1,77	0,95	1,89
26	1,30	1,46	1,23	1,55	1,14	1,65	1,06	1,76	0,98	1,87
27	1,32	1,47	1,24	1,56	1,16	1,65	1,08	1,75	0,10	1,86
28	1,33	1,78	1,26	1,56	1,18	1,65	1,10	1,75	1,03	1,85
29	1,34	1,48	1,27	1,56	1,20	1,65	1,12	1,74	1,05	1,84
30	1,35	1,49	1,28	1,57	1,21	1,65	1,14	1,74	1,07	1,83
35	1,40	1,52	1,34	1,58	1,28	1,65	1,22	1,73	1,16	1,80
40	1,44	1,54	1,39	1,60	1,34	1,66	1,29	1,72	1,23	1,79
45	1,43	1,57	1,43	1,62	1,38	1,67	1,34	1,72	1,29	1,78
50	1,50	1,59	1,46	1,63	1,42	1,67	1,38	1,72	1,34	1,77
55	1,53	1,60	1,49	1,64	1,45	1,68	1,41	1,72	1,37	1,77
60	1,55	1,62	1,51	1,65	1,48	1,69	1,44	1,73	1,41	1,77
65	1,57	1,63	1,54	1,66	1,50	1,70	1,47	1,73	1,44	1,77
70	1,58	1,64	1,55	1,67	1,53	1,70	1,49	1,74	1,46	1,77
75	1,60	1,65	1,57	1,68	1,54	1,71	1,52	1,79	1,49	1,77
80	1,61	1,66	1,59	1,69	1,56	1,72	1,53	1,74	1,51	1,77
85	1,62	1,67	1,60	1,70	1,58	1,72	1,55	1,75	1,53	1,77
90	1,64	1,68	1,61	1,70	1,59	1,73	1,57	1,75	1,54	1,78
95	1,65	1,69	1,62	1,71	1,60	1,73	1,58	1,76	1,56	1,78
100	1,65	1,69	1,63	1,72	1,61	1,74	1,59	1,76	1,57	1,78

Закінчення табл. К.1

Критерій Дарбіна – Уотсона для $\alpha = 0,05$ (5 %)

	$m = 6$		$m = 7$		$m = 8$		$m = 9$		$m = 10$	
n	dl	du	dl	du	dl	du	dl	du	dl	du
11	0,12	2,89								
12	0,16	2,67	0,11	3,06						
13	0,21	2,49	0,14	2,64	0,09	3,18				
14	0,26	2,35	0,18	2,67	0,12	2,96	0,08	3,29		
15	0,30	2,24	0,23	2,53	0,16	2,82	0,11	3,10	0,07	3,37
16	0,35	2,15	0,27	2,42	0,20	2,68	0,14	2,94	0,09	3,20
17	0,39	2,08	0,31	2,31	0,24	2,57	0,18	2,81	0,13	3,05
18	0,44	2,02	0,37	2,24	0,28	2,47	0,22	2,70	0,16	2,93
19	0,48	1,96	0,40	2,17	0,32	2,36	0,26	2,60	0,20	2,81
20	0,52	1,92	0,44	2,11	0,36	2,31	0,29	2,51	0,23	2,71
21	0,55	1,88	0,47	2,06	0,40	2,24	0,33	2,43	0,27	2,63
22	0,57	1,85	0,51	2,02	0,44	2,19	0,37	2,37	0,30	2,55
23	0,62	1,82	0,55	1,98	0,47	2,14	0,40	2,31	0,34	2,48
24	0,65	1,79	0,58	1,94	0,51	2,10	0,44	2,26	0,38	2,42
25	0,68	1,78	0,61	1,92	0,54	2,06	0,47	2,21	0,41	2,37
26	0,71	1,76	0,64	1,89	0,57	2,03	0,51	2,17	0,44	2,31
27	0,74	1,74	0,67	1,87	0,60	2,00	0,54	2,13	0,47	2,27
28	0,76	1,73	0,70	1,85	0,66	1,97	0,57	2,10	0,50	2,23
29	0,79	1,72	0,72	1,83	0,69	1,95	0,60	2,07	0,53	2,19
30	0,81	1,71	0,76	1,82	0,69	1,93	0,62	2,04	0,56	2,16
31	0,83	1,7	0,77	1,8	0,71	1,91	0,65	2,02	0,59	2,13
32	0,86	1,69	0,79	1,78	0,73	1,89	0,67	2,0	0,62	2,1
33	0,88	1,68	0,82	1,78	0,76	1,87	0,7	1,98	0,64	2,0
34	0,9	1,68	0,84	1,77	0,78	1,68	0,72	1,96	0,67	2,06
35	0,91	1,67	0,86	1,76	0,80	1,85	0,74	1,94	0,69	2,04
40	1,00	1,65	0,95	1,72	0,90	1,80	0,84	1,88	0,79	1,96
45	1,07	1,64	1,02	1,71	0,97	1,77	0,93	1,83	0,86	1,90
50	1,12	1,64	1,06	1,69	1,04	1,75	1,00	1,81	0,96	1,60
55	1,17	1,64	1,13	1,69	1,10	1,73	1,06	1,79	1,02	1,84
60	1,21	1,64	1,18	1,68	1,14	1,73	1,11	1,77	1,07	1,82
65	1,25	1,64	1,22	1,68	1,19	1,72	1,15	1,76	1,12	1,80
70	1,28	1,65	1,25	1,68	1,22	1,72	1,19	1,75	1,16	1,79
75	1,31	1,65	1,28	1,68	1,26	1,72	1,23	1,75	1,20	1,79
80	1,34	1,65	1,31	1,68	1,29	1,71	1,26	1,75	1,23	1,78
85	1,36	1,66	1,34	1,69	1,31	1,71	1,29	1,74	1,26	1,77
90	1,38	1,66	1,36	1,69	1,34	1,71	1,31	1,74	1,29	1,77
95	1,40	1,67	1,38	1,69	1,36	1,72	1,34	1,74	1,31	1,77
100	1,42	1,67	1,40	1,69	1,38	1,72	1,36	1,74	1,34	1,77

Умовні позначення:
n – число спостережень; m – число факторів;
$dl,\ du$ – критичні значення інтервалів статистики Дарбіна – Уотсона.

Продовження додатка К

Таблиця К.2

Критерій Дарбіна – Уотсона для $\alpha = 0,01$ (1 %)

	$m = 1$		$m = 2$		$m = 3$		$m = 4$		$m = 5$	
n	dl	du	dl	du	dl	du	dl	du	dl	du
6	0,39	1,14								
7	0,44	1,04	0,29	1,68						
8	0,5	1,00	0,35	1,49	0,23	2,1				
9	0,55	1,00	0,41	1,39	0,28	1,88	0,18	2,43		
10	0,6	1,00	0,47	1,33	0,34	1,73	0,23	2,19	0,15	2,69
11	0,65	1,01	0,52	1,29	0,4	1,64	0,29	2,03	0,19	2,45
12	0,7	1,02	0,57	1,27	0,45	1,58	0,34	1,91	0,24	2,28
13	0,74	1,04	0,62	1,26	0,5	1,53	0,39	1,83	0,29	2,15
14	0,78	1,05	0,66	1,25	0,55	1,49	0,46	1,76	0,34	2,05
15	0,81	1,07	0,70	1,25	0,59	1,46	0,49	1,7	0,39	1,97
16	0,84	1,09	0,74	1,25	0,63	1,45	0,53	1,66	0,44	1,9
17	0,87	1,11	0,77	1,26	0,67	1,43	0,57	1,63	0,48	1,85
18	0,9	1,12	0,81	1,26	0,71	1,42	0,61	1,6	0,52	1,8
19	0,93	1,13	0,84	1,27	0,74	1,42	0,65	1,58	0,56	1,77
20	0,95	1,15	0,87	1,27	0,77	1,41	0,69	1,57	0,6	1,74
21	0,98	1,16	0,89	1,28	0,8	1,41	0,72	1,55	0,63	1,71
22	1	1,17	0,92	1,28	0,83	1,41	0,75	1,54	0,67	1,69
23	1,02	1,19	0,94	1,29	0,86	1,41	0,78	1,53	0,7	1,67
24	1,04	1,2	0,96	1,3	0,88	1,41	0,81	1,53	0,73	1,66
25	1,06	1,21	0,98	1,31	0,91	1,41	0,83	1,52	0,76	1,65
26	1,07	1,22	1,0	1,31	0,93	1,41	0,86	1,52	0,78	1,64
27	1,09	1,23	1,02	1,32	0,95	1,41	0,88	1,52	0,81	1,63
28	1,10	1,24	1,04	1,33	0,97	1,42	0,90	1,51	0,83	1,62
29	1,12	1,25	1,05	1,33	0,99	1,42	0,92	1,51	0,86	1,61
30	1,13	1,26	1,07	1,34	1,01	1,42	0,94	1,51	0,88	1,61
35	1,20	1,31	1,14	1,37	1,09	1,44	1,03	1,51	0,97	1,59
40	1,21	1,34	1,2	1,4	1,15	1,46	1,1	1,51	1,04	1,59
45	1,29	1,38	1,25	1,42	1,2	1,47	1,16	1,53	1,11	1,59
50	1,32	1,4	1,29	1,45	1,25	1,49	1,21	1,54	1,16	1,59
55	1,36	1,43	1,32	1,47	1,28	1,51	1,25	1,55	1,21	1,59
60	1,38	1,45	1,35	1,48	1,32	1,52	1,28	1,56	1,25	1,6
65	1,41	1,47	1,38	1,50	1,35	1,53	1,32	1,57	1,26	1,6
70	1,43	1,49	1,40	1,52	1,37	1,55	1,34	1,58	1,31	1,61
75	1,45	1,50	1,42	1,53	1,4	1,56	1,37	1,59	1,34	1,62
80	1,47	1,52	1,44	1,54	1,42	1,57	1,39	1,6	1,36	1,62
85	1,48	1,53	1,46	1,55	1,44	1,58	1,41	1,6	1,39	1,63
90	1,5	1,54	1,47	1,65	1,45	1,59	1,43	1,61	1,41	1,64
95	1,51	1,55	1,49	1,57	1,47	1,6	1,45	1,62	1,43	1,64
100	1,52	1,56	1,5	1,58	1,48	1,6	1,46	1,63	1,44	1,65

Закінчення табл. К.2

Критерій Дарбіна – Уотсона для $\alpha = 0,01$ (1 %)

	$m = 6$		$m = 7$		$m = 8$		$m = 9$		$m = 10$	
n	dl	du	dl	du	dl	du	dl	du	dl	du
11	0,12	2,89								
12	0,16	2,67	0,11	3,06						
13	0,21	2,49	0,14	2,64	0,09	3,18				
14	0,26	2,35	0,18	2,67	0,12	2,96	0,08	3,29		
15	0,3	2,24	0,23	2,53	0,16	2,82	0,11	3,1	0,07	3,37
16	0,35	2,15	0,27	2,42	0,2	2,68	0,14	2,94	0,09	3,2
17	0,39	2,08	0,31	2,31	0,24	2,57	0,18	2,81	0,13	3,05
18	0,44	2,02	0,37	2,24	0,28	2,47	0,22	2,7	0,16	2,93
19	0,48	1,96	0,4	2,17	0,32	2,36	0,26	2,6	0,2	2,81
20	0,52	1,92	0,44	2,11	0,36	2,31	0,29	2,51	0,23	2,71
21	0,55	1,88	0,47	2,06	0,4	2,24	0,33	2,43	0,27	2,63
22	0,57	1,85	0,51	2,02	0,44	2,19	0,37	2,37	0,3	2,55
23	0,62	1,82	0,55	1,98	0,47	2,14	0,4	2,31	0,34	2,48
24	0,65	1,79	0,58	1,94	0,51	2,1	0,44	2,26	0,38	2,42
25	0,68	1,78	0,61	1,92	0,54	2,06	0,47	2,21	0,41	2,37
26	0,71	1,76	0,64	1,89	0,57	2,03	0,51	2,17	0,44	2,31
27	0,74	1,74	0,67	1,87	0,6	2,0	0,54	2,13	0,47	2,27
28	0,76	1,73	0,7	1,85	0,66	1,97	0,57	2,1	0,5	2,23
29	0,79	1,72	0,72	1,83	0,69	1,95	0,6	2,07	0,53	2,19
30	0,81	1,71	0,76	1,82	0,69	1,93	0,62	2,04	0,56	2,16
31	0,83	1,7	0,77	1,8	0,71	1,91	0,65	2,02	0,59	2,13
32	0,86	1,69	0,79	1,78	0,73	1,89	0,67	2,0	0,52	2,1
33	0,88	1,68	0,82	1,78	0,76	1,87	0,7	1,98	0,64	2,0
34	0,90	1,68	0,84	1,77	0,78	1,86	0,72	1,96	0,67	2,06
35	0,91	1,67	0,86	1,76	0,8	1,85	0,74	1,94	0,69	2,04
40	1,0	1,65	0,95	1,72	0,9	1,8	0,84	1,88	0,79	1,96
45	1,07	1,64	1,02	1,71	0,97	1,77	0,93	1,83	0,86	1,90
50	1,12	1,64	1,06	1,69	1,04	1,75	1,0	1,81	0,96	1,96
55	1,17	1,64	1,13	1,69	1,1	1,73	1,06	1,79	1,02	1,84
60	1,21	1,64	1,18	1,68	1,14	1,73	1,11	1,77	1,07	1,82
65	1,25	1,64	1,22	1,68	1,19	1,72	1,15	1,76	1,12	1,80
70	1,28	1,65	1,25	1,68	1,22	1,72	1,19	1,75	1,16	1,79
75	1,31	1,65	1,28	1,68	1,26	1,72	1,23	1,75	1,20	1,79
80	1,34	1,65	1,31	1,68	1,29	1,71	1,26	1,75	1,23	1,78
85	1,36	1,66	1,34	1,69	1,31	1,71	1,29	1,74	1,26	1,77
90	1,38	1,66	1,36	1,69	1,34	1,71	1,31	1,74	1,29	1,77
95	1,4	1,67	1,38	1,69	1,36	1,72	1,34	1,74	1,31	1,77
100	1,42	1,67	1,4	1,69	1,38	1,72	1,36	1,4	1,34	1,77

Умовні позначення:
n – число спостережень; m – число факторів;
$dl,\ du$ – критичні значення інтервалів статистики Дарбіна – Уотсона.

Додаток Л

Таблиця Л.1

Критичні значення Q (відношення фон Неймана), для яких $P\left(\frac{\delta^2}{s^2} < Q\right) = 0$

n	Q	n	Q
4	0,7811	**12**	0,0743
5	0,4775	**15**	0,0468
6	0,3215	**20**	0,0259
7	0,2311	**25**	0,0164
8	0,1740	**30**	0,0113
9	0,1357	**40**	0,0063
10	0,1088	**50**	0,0040
11	0,0891	**60**	0,0028

Умовні позначення:
n – число спостережень.

Таблиця Л.2

Критичні значення статистики Неймана (Q)

Число спостережень	Додатна автокореляція		Від'ємна автокореляція	
	$\alpha = 0,05$	$\alpha = 0,01$	$\alpha = 0,05$	$\alpha = 0,01$
10	1,18	0,84	3,61	3,26
15	1,29	0,99	3,30	2,99
20	1,37	1,10	3,12	2,84
25	1,42	1,17	2,99	2,74
30	1,47	1,24	2,90	2,67

Умовні позначення:
α – рівень значущості.

Таблиця Л.3

Критичні значення циклічного коефіцієнта автокореляції ($r0$)

Число спостережень	Додатна автокореляція		Від'ємна автокореляція	
	$\alpha = 0,05$	$\alpha = 0,01$	$\alpha = 0,05$	$\alpha = 0,01$
10	0,360	0,525	0,564	0,705
15	0,328	0,475	0,462	0,597
20	0,299	0,432	0,399	0,524
25	0,276	0,398	0,356	0,473
30	0,257	0,370	0,325	0,433

Умовні позначення:
α – рівень значущості.

Таблиця Л.4

Критичні значення інтеграційної статистики Дарбіна – Уотсона для оцінювання стаціонарних часових рядів

n	$\alpha = 0,01$		$\alpha = 0,05$		$\alpha = 0,1$	
	IDW_L	IDW_U	IDW_L	IDW_U	IDW_L	IDW_U
24	1,0177	1,0581	1,1976	1,2724	1,3119	1,3995
89	1,434	1,494	1,5599	1,6029	1,6345	1,6634
119	1,5082	1,5686	1,6334	1,6601	1,6917	1,7097
151	1,5548	1,6121	1,6721	1,6983	1,7256	1,7429

Умовні позначення:
n – число спостережень; α – рівень значущості;
IDW_L, IDW_U – критичні значення інтервалів.

Закінчення додатка Л

Таблиця Л.5

Критичні значення інтеграційної статистики Дарбіна – Уотсона для оцінювання нестаціонарних часових рядів

$\boldsymbol{n}$	$\boldsymbol{\alpha = 0,01}$		$\boldsymbol{\alpha = 0,05}$		$\boldsymbol{\alpha = 0,1}$	
	$\boldsymbol{IDW_L}$	$\boldsymbol{IDW_U}$	$\boldsymbol{IDW_L}$	$\boldsymbol{IDW_U}$	$\boldsymbol{IDW_L}$	$\boldsymbol{IDW_U}$
24	0,9409	0,9563	0,9266	0,9326	0,8361	0,839
89	0,4309	0,4415	0,2976	0,2989	0,2421	0,2436
119	0,3255	0,3318	0,225	0,2262	0,1814	0,1821
151	0,259	0,2622	0,1764	0,1756	0,1413	0,1419

Умовні позначення:
n – число спостережень; α – рівень значущості;
IDW_L, IDW_U – критичні значення інтервалів.

Таблиця Л.6

Односторонні критичні значення статистики Діккі – Фулера (DF)

Рівень значущості	**Розмір вибірки**			
	25	**50**	**100**	**∞**
Модель виду $\Delta Y_t = b_1 \cdot Y_{t-1} + e_{1t}$				
$\boldsymbol{\alpha = 0,01}$	-2,66	-2,62	-2,6	-2,58
$\boldsymbol{\alpha = 0,025}$	-2,26	-2,25	-2,24	-2,23
$\boldsymbol{\alpha = 0,05}$	-1,95	-1,95	-1,95	-1,95
Модель виду $\Delta Y_t = a_0 + b_1 \cdot Y_{t-1} + e_{1t}$				
$\boldsymbol{\alpha = 0,01}$	-3,75	-3,58	-3,51	-3,43
$\boldsymbol{\alpha = 0,025}$	-3,33	-3,22	-3,17	-3,12
$\boldsymbol{\alpha = 0,05}$	-3,00	-2,93	-2,89	-2,86
Модель виду $\Delta Y_t = a_0 + a_1 t + b_1 \cdot Y_{t-1} + e_{1t}$				
$\boldsymbol{\alpha = 0,01}$	-4,38	-4,15	-4,04	-3,96
$\boldsymbol{\alpha = 0,025}$	-3,95	-3,80	-3,69	-3,66
$\boldsymbol{\alpha = 0,05}$	-3,60	-3,50	-3,45	-3,41

Таблиця Л.7

Критичні значення статистик для перевірки гіпотези в методі Фостера – Стюарта ($\alpha = 0,05$)

$\boldsymbol{n}$	$\boldsymbol{\mu}$	$\boldsymbol{\sigma^1}$	$\boldsymbol{\sigma^2}$	$\boldsymbol{n}$	$\boldsymbol{\mu}$	$\boldsymbol{\sigma^1}$	$\boldsymbol{\sigma^2}$
10	3,858	1,288	1,964	**60**	7,36	2,201	2,713
15	4,636	1,521	2,153	**65**	7,519	2,236	2,742
20	5,195	1,677	2,279	**70**	7,666	2,268	2,769
25	5,632	1,791	2,373	**75**	7,803	2,297	2,793
30	5,99	1,882	2,447	**80**	7,931	2,324	2,816
35	6,294	1,956	2,509	**85**	8,051	2,349	2,837
40	6,557	2,019	2,561	**90**	8,165	2,373	2,857
45	6,79	2,072	2,606	**95**	8,273	2,395	2,876
50	6,998	2,121	2,645	**100**	8,375	2,416	2,894
55	7,187	2,163	2,681				

Умовні позначення:
n – число спостережень; μ – середнє значення величини S;
σ^1 – стандартна похибка величини S; σ^2 – стандартна похибка величини d.

Додаток М

Таблиця М.1

Відсоткові критичні значення критерію Смірнова – Грабса (T)

Число спостережень	Рівень значущості			Число спостережень	Рівень значущості		
	0,9	**0,95**	**0,99**		**0,9**	**0,95**	**0,99**
3	1,412	1,414	1,414	**27**	2,749	2,913	3,239
4	1,689	1,71	1,728	**28**	2,764	2,929	3,258
5	1,869	1,917	1,972	**29**	2,778	2,944	3,275
6	1,996	2,067	2,161	**30**	2,792	2,958	3,291
7	2,093	2,182	2,31	**31**	2,805	2,972	3,307
8	2,172	2,273	2,431	**32**	2,818	2,985	3,322
9	2,238	2,349	2,532	**33**	2,83	2,998	3,337
10	2,294	2,414	2,616	**34**	2,842	3,01	3,351
11	2,343	2,47	2,689	**35**	2,853	3,022	3,364
12	2,387	2,519	2,753	**36**	2,864	3,033	3,377
13	2,426	2,563	2,809	**37**	2,874	3,044	3,389
14	2,761	2,602	2,859	**38**	2,885	3,055	3,401
15	2,494	2,638	2,905	**39**	2,894	3,065	3,413
16	2,523	2,67	2,946	**40**	2,904	3,075	3,424
17	2,551	2,701	2,983	**41**	2,913	3,084	3,435
18	2,577	2,728	3,017	**42**	2,922	3,094	3,445
19	2,601	2,754	3,049	**43**	2,931	3,103	3,455
20	2,623	2,779	3,079	**44**	2,94	3,112	3,465
21	2,644	2,801	3,106	**45**	2,948	3,12	3,474
22	2,664	2,823	3,132	**46**	2,956	3,129	3,483
23	2,683	2,843	3,156	**47**	2,964	3,137	3,492
24	2,701	2,862	3,179	**48**	2,972	3,145	3,501
25	2,718	2,88	3,2	**49**	2,98	3,152	3,51
26	2,734	2,897	3,22	**50**	2,987	3,16	3,518

Закінчення додатка М

Таблиця М.2

Критичні значення $C_α$-оцінки для L- та L'-критеріїв Тітьєна та Мура ($\alpha = 0,05$)

Число спостережень	1	2	3	4	5	6	7	8	9	10
3	0,003									
4	0,051	0,001								
5	0,125	0,018								
6	0,203	0,055	0,01							
7	0,273	0,106	0,032							
8	0,326	0,146	0,064	0,022						
9	0,372	0,194	0,099	0,045						
10	0,418	0,233	0,129	0,07	0,034					
11	0,045	0,27	0,162	0,098	0,054					
12	0,489	0,305	0,196	0,125	0,076	0,042				
13	0,517	0,337	0,224	0,15	0,098	0,06				
14	0,54	0,363	0,25	0,174	0,122	0,079	0,05			
15	0,556	0,387	0,276	0,197	0,14	0,097	0,066			
20	0,639	0,484	0,377	0,299	0,238	0,188	0,15	0,115	0,088	0,066
25	0,696	0,55	0,45	0,374	0,312	0,262	0,222	0,184	0,154	0,126
30	0,73	0,599	0,506	0,434	0,376	0,327	0,283	0,245	0,212	0,183
35	0,762	0,642	0,554	0,482	0,424	0,376	0,334	0,297	0,264	0,235
40	0,784	0,672	0,588	0,523	0,468	0,421	0,378	0,342	0,31	0,28
45	0,802	0,696	0,618	0,556	0,502	0,456	0,417	0,382	0,35	0,32
50	0,82	0,722	0,646	0,588	0,535	0,49	0,45	0,414	0,383	0,356

Таблиця М.3

Критичні значення $C_α$-оцінки для E-критерію Тітьєна та Мура ($\alpha = 0,05$)

Число спостережень	1	2	3	4	5	6	7	8	9	10
3	0,001									
4	0,025	0,001								
5	0,081	0,01								
6	0,146	0,034	0,004							
7	0,208	0,065	0,016							
8	0,265	0,099	0,034	0,01						
9	0,314	0,137	0,057	0,021						
10	0,356	0,172	0,083	0,037	0,014					
11	0,386	0,204	0,107	0,055	0,026					
12	0,424	0,234	0,133	0,073	0,039	0,018				
13	0,455	0,262	0,156	0,092	0,053	0,028				
14	0,484	0,293	0,179	0,112	0,068	0,039	0,021			
15	0,509	0,317	0,206	0,134	0,084	0,052	0,03			
20	0,597	0,416	0,302	0,221	0,163	0,119	0,085	0,059	0,041	0,028
25	0,652	0,493	0,381	0,298	0,236	0,186	0,146	0,114	0,089	0,068
30	0,698	0,549	0,443	0,364	0,298	0,246	0,203	0,116	0,137	0,112
35	0,732	0,596	0,495	0,417	0,351	0,298	0,254	0,214	0,181	0,164
40	0,758	0,629	0,534	0,458	0,395	0,343	0,297	0,259	0,223	0,195
45	0,778	0,658	0,567	0,492	0,433	0,381	0,337	0,299	0,263	0,233
50	0,797	0,684	0,599	0,529	0,468	0,417	0,373	0,334	0,299	0,268

Додаток Н

Таблиця Н.1

G-розподіл

k \ m	**Рівень значущості**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
2	**0,05**	0,998	0,975	0,939	0,906	0,877	0,853	0,838
	0,01	0,999	0,995	0,979	0,959	0,937	0,917	0,809
3	**0,05**	0,967	0,871	0,798	0,746	0,707	0,677	0,653
	0,01	0,993	0,942	0,883	0,834	0,903	0,761	0,734
4	**0,05**	0,906	0,768	0,684	0,629	0,59	0,56	0,537
	0,01	0,968	0,864	0,781	0,721	0,676	0,641	0,613
5	**0,05**	0,841	0,684	0,598	0,544	0,507	0,478	0,456
	0,01	0,928	0,789	0,696	0,633	0,588	0,553	0,526
6	**0,05**	0,781	0,616	0,532	0,48	0,445	0,418	0,398
	0,01	0,883	0,722	0,626	0,564	0,52	0,487	0,461
7	**0,05**	0,727	0,561	0,48	0,431	0,397	0,373	0,354
	0,01	0,838	0,664	0,569	0,508	0,466	0,435	0,411
8	**0,05**	0,68	0,516	0,438	0,391	0,36	0,336	0,319
	0,01	0,795	0,615	0,521	0,463	0,423	0,393	0,37
9	**0,05**	0,639	0,478	0,403	0,358	0,329	0,307	0,29
	0,01	0,754	0,573	0,481	0,425	0,387	0,359	0,338
k \ m	**Рівень значущості**	**8**	**9**	**10**	**16**	**36**	**144**	**∞**
2	**0,05**	0,816	0,801	0,788	0,734	0,66	0,518	0,5
	0,01	0,882	0,867	0,854	0,795	0,7	0,606	0,5
3	**0,05**	0,633	0,617	0,603	0,547	0,475	0,403	0,33
	0,01	0,711	0,691	0,674	0,606	0,515	0,423	0,33
4	**0,05**	0,518	0,502	0,488	0,437	0,372	0,309	0,25
	0,01	0,59	0,57	0,554	0,488	0,406	0,325	0,25
5	**0,05**	0,439	0,424	0,412	0,365	0,307	0,251	0,2
	0,01	0,504	0,485	0,47	0,409	0,335	0,254	0,2
6	**0,05**	0,382	0,368	0,357	0,314	0,261	0,212	0,167
	0,01	0,44	0,423	0,408	0,353	0,286	0,223	0,167
7	**0,05**	0,338	0,326	0,315	0,276	0,228	0,183	0,143
	0,01	0,391	0,375	0,362	0,311	0,249	0,193	0,143
8	**0,05**	0,304	0,293	0,283	0,246	0,202	0,162	0,15
	0,01	0,352	0,337	0,325	0,278	0,221	0,17	0,125
9	**0,05**	0,277	0,266	0,257	0,223	0,182	0,145	0,11
	0,01	0,321	0,307	0,295	0,251	0,199	0,152	0,11

Умовні позначення:

$m = n - 1$; n – число вибіркових даних; k – число досліджуваних сукупностей.

Додаток П

Таблиця П.1

Основні статистичні функції MS Excel

Позначення	Значення функції
1	2
FРАСП	Повертає F-розподіл імовірності (ступінь відхилення) для двох сукупностей даних
FРАСПОБР	Повертає зворотне значення для F-розподілу ймовірностей
ZTEST	Повертає двостороннє р-значення z-тесту
БЕТАОБР	Повертає зворотну функцію до інтегральної функції щільності бета-ймовірності
БЕТАРАСП	Повертає інтегральну функцію щільності бета-ймовірності
БИНОМРАСП	Повертає окреме значення біноміального розподілу
ВЕЙБУЛЛ	Повертає розподіл Вейбула
ВЕРОЯТНОСТЬ	Повертає ймовірність того, що значення діапазону знаходиться усередині заданих меж
ГАММАОБР	Повертає зворотний гамма-розподіл
ГАММАРАСП	Повертає гамма-розподіл
ГИПЕРГЕОМЕТ	Повертає гіпергеометричний розподіл
ДИСП	Оцінює дисперсію за вибіркою
ДИСПР	Обчислює дисперсію для генеральної сукупності
ДОВЕРИТ	Повертає довірчий інтервал для середнього генеральної сукупності
КВАДРОТКЛ	Повертає суму квадратів відхилень даних від середнього за вибіркою
КВПИРСОН	Повертає квадрат коефіцієнта кореляції Пірсона за даними
КОВАР	Повертає коваріацію, середнє попарних добутків відхилень
КОРРЕЛ	Повертає коефіцієнт кореляції між двома множинами даних
ЛОГНОРМОБР	Повертає зворотний логарифмічний нормальний розподіл
ЛОГНОРМРАСП	Повертає інтегральний логарифмічний нормальний розподіл
МАКС	Повертає максимальне значення зі списку аргументів
МИН	Повертає мінімальне значення зі списку аргументів
НОРМАЛИЗАЦИЯ	Повертає нормалізоване значення
НОРМОБР	Повертає зворотний нормальний розподіл
НОРМРАСП	Повертає нормальну функцію розподілу
НОРМСТОБР	Повертає зворотне значення стандартного нормального розподілу
НОРМСТРАСП	Повертає стандартний нормальний інтегральний розподіл
ПЕРЕСТ	Повертає кількість перестановок для заданого числа об'єктів
ПИРСОН	Повертає коефіцієнт кореляції Пірсона
ПУАССОН	Повертає розподіл Пуассона
СРГАРМ	Повертає середнє гармонійне сукупності даних із додатних чисел
СРГЕОМ	Повертає середнє геометричне елементів масиву додатних чисел
СРЗНАЧ	Повертає середнє (арифметичне) аргументів
СРОТКЛ	Повертає середнє абсолютних значень відхилень даних від середнього

Закінчення додатка П

Закінчення табл. П.1

1	2
СТАНДОТКЛОН	Оцінює стандартне відхилення за вибіркою
СТАНДОТКЛОНП	Обчислює стандартне відхилення за генеральною сукупністю
СТЬЮДРАСП	Повертає t-розподіл Стьюдента
СТЬЮДРАСПОБР	Повертає зворотний t-розподіл Стьюдента
ТТЕСТ	Повертає ймовірність, відповідну до t-тесту Стьюдента
ФТЕСТ	Повертає результат F-тесту, однобічну ймовірність подібності двох сукупностей
ХИ2ОБР	Повертає значення зворотне до однобічної ймовірності розподілу χ^2
ХИ2РАСП	Повертає однобічну ймовірність розподілу χ^2
ХИ2ТЕСТ	Повертає тест на незалежність: значення розподілу χ^2 для статистичного розподілу і відповідного числа ступенів свободи
ЧАСТОТА	Повертає розподіл частот у вигляді вертикального масиву
ЭКСПРАСП	Повертає експонентний розподіл

Зміст

НАВЧАЛЬНЕ ВИДАННЯ

Гур'янова Лідія Семенівна
Клебанова Тамара Семенівна
Сергієнко Олена Андріанівна
Прокопович Світлана Валеріївна

ЕКОНОМЕТРИКА

Навчальний посібник
для студентів напряму підготовки
"Економічна кібернетика"
всіх форм навчання

Відповідальний за випуск *Клебанова Т. С.*

Відповідальний редактор *Оленич М. М.*

Редактор *Бутенко В. О.*

Коректор *Лященко О. Г.*

План 2015 р. Поз. № 41-П.
Підп. до друку 24.12.2015 р. Формат 60x90 1/16. Папір офсетний. Друк цифровий.
Ум. друк. арк. 24,0. Обл.-вид. арк. 30,0. Тираж 400 пр. Зам. № 267.

Видавець і виготівник – ХНЕУ ім. С. Кузнеця, 61166, м. Харків, просп. Леніна, 9-А

Свідоцтво про внесення суб'єкта видавничої справи до Державного реєстру
ДК № 4853 від 20.02.2015 р.