МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ЧЕРНІГІВСЬКИЙ ДЕРЖАВНИЙ ТЕХНОЛОГІЧНИЙ УНІВЕРСИТЕТ

ЕКОНОМЕТРИКА

Методичні вказівки
до практичних занять
для студентів галузі знань 0305 «Економіка і підприємництво»
за напрямами підготовки
6.030505 «Управління персоналом і економіка праці»,
6.030508 “Фінанси і кредит”

ЗАТВЕРДЖЕНО
на засіданні кафедри
математичного моделювання
та інформатики
протокол № 4 від 27.11.2012 р.

Чернігів ЧДТУ 2012

Економетрика. Методичні вказівки до практичних занять для студентів галузі знань 0305 «Економіка і підприємництво» за напрямами підготовки 6.030505«Управління персоналом і економіка праці», 6.030508 “Фінанси і кредит”/ Укл.: Ткач Ю.М.- Чернігів: ЧДТУ, 2012. - 117 с.

Укладач: Ткач Юлія Миколаївна, кандидат педагогічних наук, доцент

Відповідальний за випуск: Ткач Юлія Миколаївна, завідувач кафедри математичного моделювання та інформатики кандидат педагогічних наук, доцент

Рецензент: Мехед Д.Б, канд. пед. наук, доцент кафедри математичного моделювання та інформатики Чернігівського державного технологічного університету

ВСТУП

Економетричне моделювання широко застосовується в різноманітних економічних дослідженнях. Опанування курсу „Економетрика” надає студентам навички логічного мислення, формує здатність аналізувати економічні явища, знаходити взаємозв'язок між ними.

Мета вивчення дисципліни – ознайомлення студентів з методами досліджень, тобто методами перевірки, обґрунтування, оцінювання кількісних закономірностей та якісних тверджень (гіпотез) в мікро- та макроекономіці на основі аналізу статистичних даних. Знання, здобуті студентами під час вивчення економетрики, широко застосовуються в менеджменті, маркетингу, фінансовій справі тощо.

Завдання курсу полягають у наступному:

• опанування методів побудови та оцінювання економетричних моделей;

• набуття практичних навичок кількісного вимірювання взаємозв'язків між економічними показниками;

• визначення критеріїв для перевірки гіпотези щодо якостей економічних показників та форм їх зв'язку;

• поглиблення теоретичних знань в галузі математичного моделювання економічних процесів та явищ;

• використання результатів економетричного аналізу для прогнозування та прийняття обґрунтованих економічних рішень

Зв'язок з іншими навчальними дисциплінами. Базовими для курсу „Економетрика” є дисципліни економічного циклу, такі як „Економічна теорія”, „Мікроекономіка”, „Макроекономіка”. Математичною основою курсу є дисципліна „Теорія ймовірностей та математична статистика”, «Статистика».

Після вивчення дисципліни студент має

знати:

- термінологію, основні поняття та означення дисципліни;
- методологічні основи економетричного моделювання;
- принципи та методи побудови економеричних моделей (парних та багатофакторних);
- методику статичної оцінки значущості побудованих економетричних моделей;
- принципи розв’язування типових задач за наявності мультиколінеарності, гетероскедастичності та автокореляції.

вміти :

- класифікувати математичні моделі економетрики;
- будувати рівняння парної та множинної регресії;
- оцінювати якість побудованої економетричної моделі;
- виявляти мультиколінеарність, гетероскедастичність, автокореляцію та будувати економетричні моделі при наявності цих явищ.

1 ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ
ЗМІСТОВИЙ МОДУЛЬ 1
ЗАГАЛЬНІ ПРИНЦИПИ ПОБУДОВИ ЕКОНОМЕТРИЧНИХ МОДЕЛЕЙ

Тема 1. Вступ до курсу

Предмет і задачі навчальної дисципліни «Економетрика». Необхідність вивчення економетрики. Мета курсу. Зв'язок з іншими дисциплінами. Застосування економетричних досліджень в економіці.

Тема 2. Проста лінійна регресія

Структура моделі та основні припущення при її побудові. Оцінювання моделі. Метод найменших квадратів. Надійні інтервали оцінок. Числові критерії адекватності моделі. Коефіцієнт детермінації. Інші методи оцінювання моделі та їхнє практичне значення. Перевірка статистичних гіпотез. Перевірка моделі на адекватність. Прогнозування за допомогою простої лінійної регресії. Моделі, які зводяться до моделі множинної лінійної регресії. Приклади застосування простої лінійної регресії.

ЗМІСТОВИЙ МОДУЛЬ 2
БАГАТОФАКТОРНА РЕГРЕСІЯ

Тема 1. Множинна регресія

Структура моделі та основні припущення при її побудові. Оцінювання моделі. Метод найменших квадратів. Коефіцієнт детермінації. Властивості параметрів моделі. Дисперсія моделі. Перевірка гіпотез. Прогнозування за допомогою лінійної регресії. Моделі, що зводяться до моделі множинної лінійної регресії. Приклади застосування множинної лінійної регресії. Інтерпретація коефіцієнтів регресії.

Тема 2. Мультиколінеарність

Мультиколінеарність регресії. Методи визначення мультиколінеарності. Шляхи позбавлення мультиколінеарності. Приклади оцінювання регресії з мульколінеарними змінними.

Тема 3. Модель лінійної регресії з гетероскедастичними збуреннями

Модель лінійної регресії з гетероскедастичними збуреннями, її структура та основні припущення. Виявлення гетероскедастичності. Зважений метод найменших квадратів. Методи позбавлення гетероскедастичності. Приклад аналізу лінійної регресії з гетероскедастичними збуреннями.

Тема 4. Модель лінійної регресії з автокорельованими збуреннями

Виявлення автокореляції. Оцінювання моделі. Узагальнений метод найменших квадратів. Методи позбавлення автокореляції. Приклад аналізу лінійної регресії з автокорельованими збуреннями.

Тема 5. Вступ до теорії часових рядів

Основні визначення. Порядок аналізу часових рядів. Адитивна та мультиплікативна моделі часових рядів. Міри точності прогнозів. Лаговий оператор. Стаціонарність часових рядів. Функція автокореляції. Стабільність моделі. Методи згладжування часових рядів.

Таблиця 1.1 - Структура навчальної дисципліни

Назви змістових модулів і тем	Кількість годин											
	денна форма						Заочна форма					
	усього	у тому числі					усього	у тому числі				
		л	п	лаб	інд	с.р.		л	п	лаб	інд	с.р.
1	2	3	4	5	6	7	8	9	10	11	12	13
Змістовий модуль 1. ЗАГАЛЬНІ ПРИНЦИПИ ПОБУДОВИ ЕКОНОМЕТРИЧНИХ МОДЕЛЕЙ												
Тема 1. **Вступ до курсу**	2	2										
Тема 2. **Проста лінійна регресія**	4	4										
2.1. Метод найменших квадратів	4		2			2/1						
2.2. Аналіз якості побудованої моделі та оцінок її параметрів	8		4			4/2						
Разом за змістовим модулем 1	18	6	6			6/3						
Змістовий модуль 2. БАГАТОФАКТОРНА РЕГРЕСІЯ												
Тема 1. **Множинна регресія**	20	4	4			12/6						
Тема 2. **Мультиколінеарність**	18	2	4			12/6						
Тема 3. **Модель лінійної регресії з гетероскедастичними збуреннями**	16	2	2			12/6						
Тема 4. **Модель лінійної регресії з автокорельованими збуреннями**	16	2	2			12/6						
Тема 5. **Вступ до теорії часових рядів**	20	2				18/9						
Разом за змістовим модулем 2	90	12	12			66/33						
Усього годин	108	18	18			72/36						

Методи навчання

При викладанні навчальної дисципліни використовуються інформаційно-ілюстративний та проблемний методи навчання.

Методи контролю

Для визначення рівня засвоєння студентами навчального матеріалу використовуються такі методи контролю:

- поточний контроль (усне та письмове опитування, тестування);
- оцінка за самостійну роботу (оцінювання рефератів, контрольних робіт, розрахунково-графічних робіт);
- проміжний контроль;
- підсумковий контроль (іспит).

Для діагностики знань використовується модульно-рейтингова система зі 100-бальною шкалою оцінювання.

2 ЗМІСТ ПРАКТИЧНИХ ЗАНЯТЬ

2.1 Метод найменших квадратів

Відведений час: 2 год.

Мета: навчати студентів складати та розв'язувати системи нормальних рівнянь для методу найменших квадратів; записувати рівняння парної лінійної регресії.

Завдання для практичного заняття:

1. Пригадайте основні теоретичні питання теми.
2. Орієнтовні запитання та завдання:
   - запишіть рівняння парної лінійної регресії;
   - розкрийте сутність і особливості використання методу найменших квадратів (МНК)?
   - підготуйте вихідну інформацію для розв'язування задач методом найменших квадратів;
   - запишіть систему нормальних рівнянь за методом найменших квадратів;
   - знайдіть розв'язок системи рівнянь.
3. Виконайте індивідуальне завдання.

Короткі теоретичні відомості

Економетрія – це один з напрямків економіко-математичних методів аналізу, що полягає в статистичному вимірі (оцінюванні) параметрів математичних виразів, які характеризують деяку економічну концепцію про взаємозв'язок і розвиток об'єкта, явища, що необхідно для одержання конкретних економічних висновків на основі економетричних моделей.

Економетрика – це наука, що досліджує кількісні закономірності й взаємозалежності в економіці за допомогою методів математичної статистики. Основа цих методів – кореляційно-регресійний аналіз.

Економетрія виникла як спроба передбачити поводження товарного та грошового ринків з врахуванням випадкових економічних явищ (наприклад, у випадку коливання попиту або цін).

Термін економетрія запропонував львівський вчений П. Чомпа, опублікувавши у Львові в 1910 році книгу «Нариси економетрії і природної теорії бухгалтерії, яка грунтується на політичній економії».

Засновники економетрії (Р. Фріш, Е. Шумпетер, Я. Тірберчен) намагались поєднати економічну теорію з математичними та статистичними методами.

За рубежем перші праці з економетрії, що належали Муру, вийшли друком протягом 1914–1917 рр. У 1928 році було опубліковано роботи Ч. Кобба і П. Дугласа про виробничу функцію. Ця функція ввійшла в економетрію як класичний приклад і досі є важливим інструментом економічного аналізу.

В останні 30 років розроблялись нові методи розв'язування та дослідження економетричних моделей з врахуванням зв'язків між економічними параметрами (мультиколінеарності), у випадку із запізнювальними (лаговими) змінними, методи застосування обчислювальної техніки.

Структура дисципліни

Економетрію можна поділити на дві частини:

- економетричні методи;
- економетричні моделі.

Економетричні методи умовно можна розбити на 4 групи:

- оцінювання параметрів класичної економетричної моделі;
- методи оцінювання параметрів узагальненої економетричної моделі;
- методи оцінювання параметрів динамічної економетричної моделі, їх верифікація (перевірка значущості);
- методи оцінювання параметрів економетричних моделей, які побудовані на основі систем одночасних структурних рівнянь.

Серед засобів математико-статистичного інструментарію економетрії центральне місце займає **регресійний аналіз**.

Під регресією розуміємо односторонню стохастичну залежність однієї випадкової змінної від іншої або декількох випадкових змінних. У цьому тлумаченні регресія використовується для дослідження та оцінки залежностей між економічними явищами, які породжені, як правило, сукупністю причин.

Розглядаючи причино-наслідкові зв'язки необхідно виявити дію суттєвих, звільнившись при цьому від впливу другорядних причин та елементів випадковості.

Математичне розв'язання зводиться до отримання функції регресії. За допомогою методів математичної статистики можна дослідити залежність між такими економічними показниками як, наприклад, національний дохід, капітальні вкладення, трудові ресурси тощо. Явища, які підлягають дослідженню, повинні бути кількісно варійованими величинами.

Перш ніж застосовувати математико-статистичний апарат явище треба проаналізувати зі змістовної точки зору і вирішити, яку змінну розглядати як залежну (тобто змінну, яка підлягає поясненню за допомогою функції регресії), а які змінні під час аналізу вважаються незалежними (пояснюючими). Всі ці змінні повинні бути трактовані економічною теорією.

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

- вибір форми математичної залежності, що описує поводження економічного об'єкта на основі системи спостережень;
- оцінка параметрів даної моделі різними методами (метод

найменших квадратів, метод максимальної правдоподібності та ін.);

- перевірка статистичної значущості моделі.

Часто ***економетричну модель у загальному вигляді*** представляють як систему лінійних рівнянь:

$$BY=AX+E,$$

де B – матриця коефіцієнтів при ендогенних (залежних) змінних; Y – вектор ендогенних змінних (спостережуване значення залежної змінної); A – матриця коефіцієнтів при екзогенних (пояснюючих) змінних; X – вектор екзогенних змінних; AX – пояснена частина, що залежить від значень пояснюючих змінних; E – вектор випадкових збурень (помилок, відхилень).

Економетричні моделі включають досить широкий клас різних економіко-математичних моделей. Приведемо одну із **класифікацій економетричних моделей**.

1. ***За способом математичного подання*** економетричні моделі можна умовно розділити на прості й складні. *Прості* економетричні моделі представлені одним рівнянням, однією залежністю, *складні* – декількома рівняннями, декількома залежностями.

2. ***За кількістю факторних ознак,*** що включають у модель, прості економетричні моделі можна розділити на однофакторні й багатофакторні. *Однофакторні моделі* містять одну незалежну ознаку, *багатофакторні* – ряд незалежних ознак. Однофакторні й багатофакторні моделі можуть бути представлені *лінійними* та *нелінійними* функціями.

3. ***Складні економетричні моделі*** можуть бути представлені трьома видами систем одночасних рівнянь залежно від форми включення до правої частини ендогенних змінних. Звичайно виділяють три типи систем:

1) системи, що розв'язані відносно ендогенних змінних;

2) рекурсивні системи;

3) системи, що не розв'язані відносно ендогенних змінних.

4. ***Залежно від наявності (відсутності) у моделі фактора часу*** розрізняють *динамічні* й *статичні моделі.*

Побудова будь-якої економетричної моделі, незалежно від того, на якому рівні і для яких показників вона будується, здійснюється як послідовність певних кроків.

Крок 1. Знайомство з економічною теорією, висунення гіпотези взаємозв'язку. Чітка постановка задачі.

Крок 2. Специфікація моделі. Використовуючи всі ті форми функцій, які можуть бути застосовані для вивчення взаємозв'язків, необхідно сформулювати теоретичні уявлення і прийняті гіпотези у вигляді математичних рівнянь. Ці рівняння встановлюють зв'язки між основними визначальними змінними за припущення, що всі інші змінні є випадковими.

Крок 3. Формування масивів вихідної інформації згідно з метою та завданнями дослідження.

Крок 4. Оцінка параметрів економетричної моделі методом найменших квадратів, що дає змогу проаналізувати залишки і відповісти на запитання: чи не суперечить специфікація моделі передумовам “класичної” моделі лінійної регресії?

Крок 5. Якщо деякі передумови моделі не виконуються, то для продовження аналізу треба замінювати специфікацію або застосовувати інші методи оцінювання параметрів.

Крок 6. Проведення аналізу вірогідності моделі та визначення прогнозу за побудованою моделлю.

Лінійна парна регресія

Загальний вигляд такої моделі:

$$Y = f(x, e),$$

де Y - залежна змінна (результативна ознака); x - незалежна змінна (фактор); e — стохастична складова (випадкові збудення).

Аналітична форма цієї моделі може бути різною залежно від економічної сутності зв’язків. Найбільш поширена форма залежності – лінійна:

$$Y = b_0 + b_1 \cdot x_i + e_i.$$

де b_0 – *постійна регресії (константа)*, економічна інтерпретація якої можлива не завжди і яка визначає точку перетину прямої регресії з віссю ординат (рисунок 2.1); b_1 – *коефіцієнт регресії*, відображає нахил лінії, вздовж якої розсіяні дані спостереженя. Цей показник може бути трактований як такий, що характеризує коливання змінної y, викликане зміною значення x на одиницю (з точки зору математики цей показник являє собою тангенс кута нахилу φ (рисунок 2.1). Знак b_1 визначає напрям цієї зміни. Якщо коефіцієнт регресії додатній, то зі зростанням значення x зростає y; якщо коефіцієнт регресії від’ємний - зростання значення x призводить до зменшення y; e_i - *випадкові збурення*.

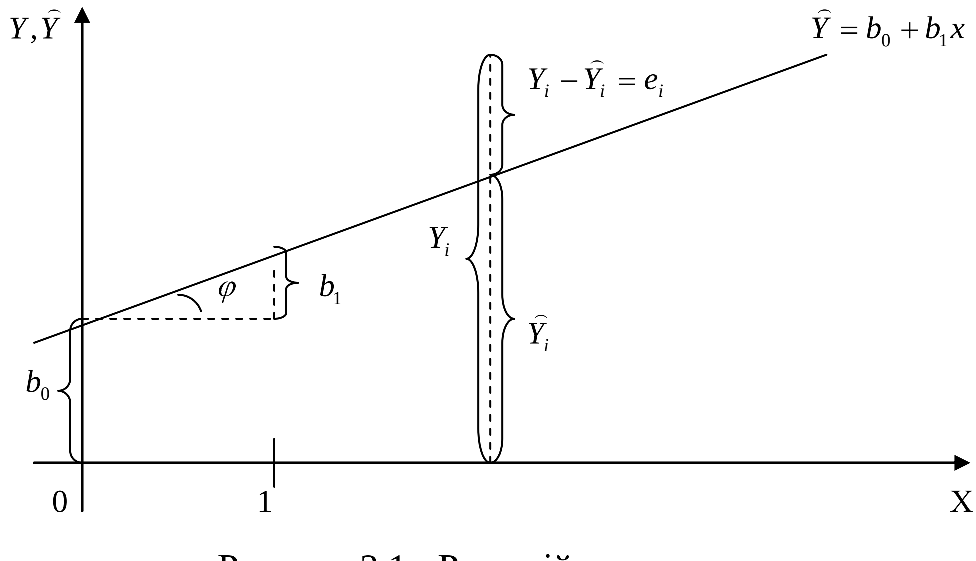


Рисунок 2.1 - Регресійна пряма

Причини, які спонукають появу випадкових збудників ε_i в рівняннях можуть бути такими:

1. Слід пам'ятати, що будь-яка регресійна модель є в певній мірі спрощенням реальної ситуації, яка в дійсності являє собою складне переплетіння різних факторів, багато з яких практично неможливо врахувати в моделі. Так, наприклад, попит на товар буде визначатися як його ціною, так і ціною на ті товари, які можуть його заміняти, ціною на супроводжуючі товари, доходами споживачів та їх вподобаннями тощо. Однак в цьому переліку не враховуються традиції як релігійні, так і національні, особливості кліматичних умов і багато інших факторів. При цьому ще виникає проблема визначення факторів, які за певних умов будуть домінуючими, а якими можна знехтувати. В ряді випадків існують фактори, які не можна використати в моделі тому, що для них проблематично одержати необхідні статистичні дані. Наприклад, величина заощаджень родини визначається не лише доходами її членів, а й їхнім здоров'ям, інформацію про що в цивілізованих країнах тримають в таємниці. Окрім цього, багато факторів мають випадковий характер (погода, стихійні лиха), які посилюють неоднозначність.

2. Неправильно вибрана функціональна залежність. Це може трапитися внаслідок недостатнього дослідження процесу, який підлягає моделюванню. Так, виробнича функція, яка описує залежність Y від одного фактора X може бути виражена лінійним співвідношенням:

$$Y = b_0 + b_1 X \ ,$$

хоча насправді, при більш ретельному дослідженні, стане відомо, що співвідношення між Y та X матиме нелінійний характер, наприклад:

$$Y = b_0 X^{b_1} \ .$$

Вибір форм функціональної залежності між змінними називають *специфікацією* моделі.

3. Можуть бути невірно вибрані пояснюючі змінні.

4. В багатьох моделях залежність між факторами має складну форму зв'язку між цілими комплексами подібних величин. Так, при дослідженні залежності попиту Y в якості пояснюючої вибирають змінну, яка уособлює складну комбінацію індивідуальних запитів, що мають на неї певний вплив поряд із факторами, які враховані в моделі. Здійснюється так зване агрегування пояснюючих змінних, що може бути однією із причин появи в моделі випадкового збудника ε_i .

5. Можуть бути допущені помилки при аналізі та обробці статистичних даних, які також сприятимуть появі ε_i .

6. Як правило, будь-яка статистична інформація є обмеженою і, крім цього, більшість моделей описуються неперервними функціями, але при цьому використовуються вибіркові дані, які мають дискретну структуру.

7. Слід також зважити на наявність людського фактора, який в тій чи

іншій мірі обов'язково є присутнім в будь-якому економічному процесі, але врахувати який в моделі поки що практично неможливо. В певних ситуаціях цей фактор може навіть якісну модель деформувати до примітивного рівня.

Аналітична форма економетричної моделі може бути різною залежно від економічної сутності зв'язків. Найбільш поширені форми залежностей:

1) $Y = b_0 + b_1 X$ - лінійна;

2) $Y = b_0 X^{b_1}$ - степенна;

3) $Y = b_0 \cdot b_1^X$ - показникова;

4) $Y = b_0 + \frac{b_1}{X}$ - гіперболічна та інші.

де b_0, b_1 — невідомі параметри моделі.

Неважко переконатись, що наведені нелінійні форми залежностей за допомогою елементарних перетворень приводяться до лінійних.

Останнє з цих співвідношень є лінійним відносно $\frac{1}{X}$, а друге та третє можна звести до лінійної форми для перетворених змінних, якщо взяти логарифми від виразів в обох частинах кожного з рівнянь. Тоді одержимо:

$$2)\ \ln Y = \ln b_0 + b_1 \ln X.$$

$$3)\ \ln Y = \ln b_0 + X \ln b_1;$$

Специфікацію моделі у випадку парної регресії можна здійснити використовуючи графічне зображення емпіричних даних (точок (x_i, y_i)) на кореляційному полі в декартовій системі координат. Тобто побудувати так звану діаграму розсіювання (рисунок 2.2 а, б, в).

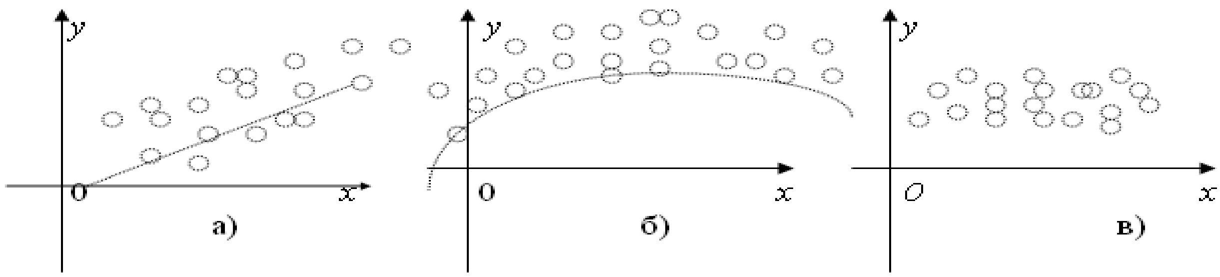


Рисунок 2.2 - Кореляційне поле

За рисунком 2.2,а можна припустити, що зв'язок між y та x є лінійним $Y = b_0 + b_1 x$; за рисунком 2,б - залежність близька до параболічної $Y = b_0 + b_1 x + b_2 x^2$, а за рисунком 2,в - явної залежності між y та x не спостерігається.

Для множинної регресії виявлення залежності між змінними значно ускладнюється. Тоді може спрацювати інтуїція.

Метод найменших квадратів (1МНК)

Нехай рівняння регресії має вигляд:

$Y = b_0 + b_1 x$, де b_0, b_1 - невідомі параметри моделі.

Відповідно до **методу найменших квадратів** невідомі параметри b_0 й b_1 вибираються таким чином, щоб сума квадратів відхилень емпіричних значень y_i від значень y_i, що знайдені за рівнянням регресії, була мінімальною:

$$S = \sum_{i=1}^{n} (\hat{y}_i - y_i)^2 = \sum_{i=1}^{n} (b_0 + b_1 x_i - y_i)^2 \to \min .$$

На підставі необхідної умови екстремуму функції двох змінних $S = S(b_0, b_1)$ прирівнюємо до нуля її частинні похідні, тобто

$$\begin{cases} \dfrac{\partial S}{\partial b_0} = 2\sum_{i=1}^{n} (b_0 + b_1 x_i - y_i) = 0; \\ \dfrac{\partial S}{\partial b_1} = 2\sum_{i=1}^{n} (b_0 + b_1 x_i - y_i) x_i = 0, \end{cases}$$

звідки після перетворень одержимо систему нормальних рівнянь для визначення параметрів лінійної системи:

$$\begin{cases} b_0 n + b_1 \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i; \\ b_0 \sum_{i=1}^{n} x_i + b_1 \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i. \end{cases}$$

Тепер, розділивши обидві частини рівняння на *n*, одержимо *систему нормальних рівнянь* у вигляді:

$$\begin{cases} b_0 + b_1 \bar{x} = \bar{y}; \\ b_0 \bar{x} + b_1 \overline{x^2} = \overline{xy}, \end{cases}$$

де відповідні середні визначаються за формулами

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}; \qquad \overline{xy} = \frac{\sum_{i=1}^{n} x_i y_i}{n};$$

$$\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n}; \qquad \overline{x^2} = \frac{\sum_{i=1}^{n} x_i^2}{n}.$$

Підставляючи значення $b_0 = \bar{y} - b_1 \bar{x}$ з першого рівняння системи у рівняння регресії, одержимо

$$b_1 = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\overline{x^2} - \bar{x}^2} = \frac{K_{XY}}{\sigma_x^2},$$

де b_1 – вибірковий коефіцієнт регресії; K_{XY} – вибірковий кореляційний момент або вибіркова кореляція; σ_x^2 – вибіркова дисперсія змінної *X*:

$$K_{XY} = \overline{xy} - \bar{x} \cdot \bar{y}, \; \sigma_x^2 = \overline{x^2} - \bar{x}^2 .$$

Відзначимо, що вільний член рівняння (b_0) не завжди має реальний зміст.

Якщо значення змінних великі, то розрахунки можуть бути проведені у відхиленнях від середніх величин.

$$\begin{matrix} x_i' = x_i - \bar{x} \\ y_i' = y_i - \bar{y} \end{matrix}; \quad \hat{y}_i = b_1 \cdot x_i';$$

$$b_1 = \frac{\overline{x'y'}}{\overline{x'^2}};$$

$$\frac{\hat{y} - \bar{y}}{\sigma_y} = \frac{b_1}{\sigma_y} \cdot \frac{(x - \bar{x})}{\sigma_x} \cdot \sigma_x .$$

Після обчислення коефіцієнтів рівняння регресії b_0 та b_1 можна визначити теоретичне значення результуючої $\hat{y}$ (у рівняння регресії підставляють відповідні значення x_i).

Альтернативний спосіб обчислення параметрів парної економеричної моделі:

$$b_1 = \frac{\sum_i (y_i - \bar{y})(x_i - \bar{x})}{\sum_i (x_i - \bar{x})^2}; \quad b_0 = \bar{y} - b_1 \bar{x}.$$

Коефіцієнт еластичності – границя відношення зміни у відсотках однієї ознаки при зміні на один відсоток іншої:

$$K_E = \lim_{\Delta x \to 0} \frac{\frac{\Delta y}{y}}{\frac{\Delta x}{x}} = \frac{\partial y}{\partial x} \cdot \frac{x}{y}$$

Тобто коефіцієнт еластичності показує на скільки відсотків зміниться результуюча y при зміні фактора x на 1%.

Оскільки під час побудови економетричних моделей використовуються сукупності спостережень, то розраховують середній коефіцієнт еластичності:

$$K_E = f'(\bar{x}) \frac{\bar{x}}{\bar{y}} = b_1 \frac{\bar{x}}{\bar{y}}.$$

Середній коефіцієнт еластичності показує, на скільки відсотків в середньому за сукупністю зміниться результуюча y від своєї величини при зміні фактора x на 1% від свого значення.

Приклад розв’язування задачі

Приклад. Побудувати економетричну модель, яка характеризує залежність витрат на одиницю продукції від рівня фондомісткості продукції. Зробити економічні висновки. Вихідні дані наведені в таблиці 2.1.

Таблиця 2.1 – Вихідні дані

№ п/п	1	2	3	4	5	6	7	8	9	10
Y_i	50	40	65	55	45	42	56	60	64	65
X_i	90	75	120	100	80	78	110	115	115	125

Розв'язання. Нехай залежність між витратами на одиницю продукції і рівнем фондомісткості описується прямою лінією:

$$Y = b_0 + b_1 X + \varepsilon,$$

де Y — витрати на одиницю продукції; X — рівень фондомісткості; ε — залишки.

Розрахункові значення витрат на одиницю продукції можна знайти, скориставшись такою моделлю: $\widehat{Y} = \widehat{b}_0 + \widehat{b}_1 X$.

Щоб оцінити параметри моделі $\widehat{b}_0$ і $\widehat{b}_1$ методом 1МНК, запишемо систему нормальних рівнянь:

$$\begin{cases} n\widehat{b}_0 + \widehat{b}_1 \sum_{i=1}^{10} x_i = \sum_{i=1}^{10} y_i; \\ \widehat{b}_0 \sum_{i=1}^{10} x_i + \widehat{b}_1 \sum_{i=1}^{10} x_i^2 = \sum_{i=1}^{10} x_i y_i. \end{cases}$$

Коефіцієнти для цих рівнянь системи знаходимо за таблицею 2.2:

Таблиця 2.2 – Додаткові обчислення

№ п/п	Y_i	X_i	X_i^2	X_iY_i
1	50	90	8100	4500
2	40	75	5625	3000
3	65	120	14400	7800
4	55	100	10000	5500
5	45	80	6400	3600
6	42	78	6084	3276
7	56	110	12100	6160
8	60	115	13225	6900
9	64	115	13225	7350
10	65	125	15625	8125
Σ	542	1008	104784	56221

$$\begin{cases} 10\widehat{b}_0 + 1008\widehat{b}_1 = 542 \\ 1008\widehat{b}_0 + 104784\widehat{b}_1 = 56221. \end{cases}$$

Розв'язком системи є параметри $\widehat{b}_0 = 3{,}8$; $\widehat{b}_1 = 0{,}5$.

Економетрична модель має вигляд

$$y = 3{,}8 + 0{,}5x + \varepsilon$$

Скористаємось альтернативним способом обчислення параметрів за допомогою відхилень середніх арифметичних.

Доповнимо таблицю 2.2 до таблиці 2.3.

Таблиця 2.3 – Доповнена таблиця

№ п/п	Y_i	X_i	X_i^2	X_iY_i	$X_i-\overline{X}$	$Y_i-\overline{Y}$	$(X_i-\overline{X})^2$	$(X_i-\overline{X})*(Y_i-\overline{Y})$
1	50	90	8100	4500	-10,8	-4,2	116,64	45,36
2	40	75	5625	3000	-25,8	-14,2	665,64	336,36
3	65	120	14400	7800	29,2	10,8	368,64	207,36
4	55	100	10000	5500	-0,8	0,8	0,64	-0,64
5	45	80	6400	3600	-20,8	-9,2	432,64	181,36
6	42	78	6084	3276	-22,8	-12,2	519,64	278,16
7	56	110	12100	6160	9,2	1,8	84,64	16,56
8	60	115	13225	6900	14,2	5,8	201,64	82,36
9	64	115	13225	7350	14,2	9,8	201,64	139,16
10	65	125	15625	8125	24,2	10,8	585,64	261,26
Σ	542	1008	104784	56221			3376	1587,5

Дістанемо

$$\widehat{b_1} = \frac{\sum_{i=1}^{10}(Y_i-\overline{Y})(X_i-\overline{X})}{\sum_{i=1}^{10}(X-\overline{X})^2} = \frac{1587,5}{3776} \approx 0,5.$$

$$\widehat{b_0} = \overline{Y} - b_1\overline{X} = 54,2 - 0,5\cdot 100,8 = 3,8.$$

Зауважимо, що оцінки параметрів моделі згідно з методом 1МНК є досить чутливими до точності розрахунків та адекватності аналітичної форми моделі. Оскільки вільний член моделі $\hat{a}_0 = 3,8 \neq 0$, то рівень витрат на одиницю продукції не є строго пропорційним до рівня фондомісткості. Кількісна оцінка параметра $\hat{a}_1 = 0,5$ показує, що граничне збільшення витрат при зростанні фондомісткості продукції на 1 грн. становить 0,5 грн. Тобто, коли рівень фондомісткості збільшиться на одиницю, то витрати на одиницю продукції збільшаться на 0, 5 одиниць.

Еластичність витрат щодо фондомісткості продукції визначається коефіцієнтом еластичності:

$$E_{y/x} = \frac{\partial Y}{\partial X} : \frac{\overline{Y}}{\overline{X}} = 0,5 : \frac{54,2}{100,8} = 0,5{:}0,54 = 0,93.$$

Значення цього коефіцієнта слід тлумачити так: при збільшенні фондомісткості продукції на 1 % витрати на одиницю гранично зростуть на 0,93 %.

Завдання для індивідуальної та самостійної роботи студентів

Побудувати економетричну модель, яка характеризує залежність витрат на одиницю продукції від рівня фондомісткості продукції. Зробитит економічні висновки. Вихідні дані наведені в таблиці 2.4.

Таблиця 2.4 - Дані про роботу 10 однотипних підприємств

1	Y	18,9	23,2	21,3	23,5	24,9	24,6	26,8	27,7	30,4	29,3
	x	245	255	242	262	269	275	271	266	291	296
2	Y	23,5	26,6	26,6	30,5	30,9	33,2	35,2	35,3	38,2	42,1

Продовження таблиці 2.4

	x	92	97	108	111	102	117	120	113	130	141
3	*Y*	21,5	23,7	23,2	24,6	27,5	25,4	27,6	28,9	28,5	30,8
	x	125	136	128	151	152	135	157	158	173	173
4	*Y*	13,2	15,9	19,1	17,6	18,8	21,1	21,9	23,6	23	23,2
	x	194	210	231	231	232	237	235	254	252	242
5	*Y*	13,4	16,7	18,5	19,7	21,1	21,9	24,2	26,7	27,3	31,3
	x	256	271	269	269	279	272	279	278	288	293
6	*Y*	23,4	25,2	25,2	25,3	29	28	28,9	28,1	31,9	29,3
	x	134	147	144	143	161	153	166	161	164	161
7	*Y*	17,4	19,1	18,9	19,7	20,4	21,9	21	25,6	26,8	26,2
	x	292	303	303	307	296	321	309	337	333	340
8	*Y*	14,2	15,8	17,6	21,9	19,7	21,4	26,2	24,6	27,9	31,2
	x	207	221	219	238	232	227	245	234	240	264
9	*Y*	18,3	19,5	19,7	21,1	19	20,2	21	21,4	21,8	22,4
	x	238	237	243	252	239	250	257	258	264	268
10	*Y*	16,9	19,3	23,3	21,6	23,9	25,9	28,7	29,2	30,7	34,3
	x	151	156	161	160	168	186	195	192	179	195
11	*Y*	26,6	26,6	29,9	29,6	33,3	31,8	32,8	37,3	37	38,4
	x	293	288	311	307	317	310	328	326	333	338
12	*Y*	22,7	25,8	28,1	28,9	31,4	33	31,8	37,2	38	39,6
	x	209	216	218	228	234	226	230	235	244	246
13	*Y*	22,7	24,8	26	26,4	27,1	27,8	30,6	31,5	31,4	32,8
	x	215	234	241	246	241	231	249	258	262	267
14	*Y*	13,4	12,9	14,4	14,5	13,8	14,3	15,5	15,1	16,4	16,9
	x	269	274	288	283	285	287	303	298	309	311
15	*Y*	14,1	14,9	16,7	17,7	19,5	20,8	20,9	22,5	24,2	22,7
	x	212	220	218	224	225	229	231	237	250	237
16	*Y*	15,7	19,1	19	20,9	21,4	21,4	20,7	22,8	24,2	25,3
	x	171	191	189	204	205	193	197	203	206	230
17	*Y*	22,3	22,7	23,6	24,3	26	24,9	24,7	26,3	26,9	26,4
	x	179	182	178	190	196	188	195	201	207	200
18	*Y*	24,3	27,2	28,2	28,1	29,3	32,4	34,2	36,2	37,1	38,7
	x	170	181	177	188	176	184	195	198	201	201
19	*Y*	31,5	32,5	32,6	34,9	35,4	36,4	36,3	40,2	39,7	42
	x	129	142	140	146	149	147	149	165	161	162
20	*Y*	25,3	24,8	26,5	28,6	27,6	29,9	31,2	32,2	33,2	33,1
	x	253	251	260	266	270	273	280	277	281	278
21	*Y*	26,2	24,2	27,1	29,1	29,9	30,6	32	32,7	34,8	35,6
	x	135	125	132	134	142	131	147	149	151	161
22	*Y*	22,5	26,3	28,1	27,3	29,9	31	31,1	33	35,5	36
	x	272	296	292	295	303	301	301	304	316	323
23	*Y*	15,3	16,2	18,7	20,3	20,9	24,3	24,9	27,8	29,6	30,1
	x	141	142	155	152	149	163	154	160	179	176
24	*Y*	14,9	14,6	13,5	15,7	17,4	18,1	21,9	22,2	20,9	23
	x	290	277	281	288	285	296	299	313	303	313
25	*Y*	15,4	16,3	17,4	18,1	18,2	20,8	22,2	21,3	23,1	25,2
	x	262	270	276	267	279	286	283	285	294	301
26	*Y*	11,2	14,8	15,4	16,7	16,8	21,1	19,8	21,9	23,8	24,2

Продовження таблиці 2.4

	x	188	199	196	203	208	221	216	211	226	225
27	Y	18,9	23,2	21,3	23,5	24,9	24,6	26,8	27,7	30,4	29,3
	x	245	255	242	262	269	275	271	266	291	296
28	Y	23,5	26,6	26,6	30,5	30,9	33,2	35,2	35,3	38,2	42,1
	x	92	97	108	111	102	117	120	113	130	141
29	Y	21,5	23,7	23,2	24,6	27,5	25,4	27,6	28,9	28,5	30,8
	x	125	136	128	151	152	135	157	158	173	173
30	Y	13,2	15,9	19,1	17,6	18,8	21,1	21,9	23,6	23	23,2
	x	194	210	231	231	232	237	235	254	252	242

Дані про роботу підприємств необхідно вибрати з таблиці відповідно до номера варіанта. Номер варіант визначається за вказівкою викладача.

2.2 Аналіз якості побудованої моделі та оцінок її параметрів

Відведений час: 4 год.

Мета: формувати вміння складати та розв’язувати системи нормальних рівнянь за методом найменших квадратів; записувати рівняння парної лінійної регресії, навчати здійснювати аналіз якості побудованої моделі та оцінок її параметрів.

Завдання для практичного заняття:

1. Пригадайте основні теоретичні питання теми.
2. Орієнтовні запитання та завдання:
   - що означає перевірити значущість рівняння регресії?
   - назвіть показники, за допомогою яких, здійснюють перевірку адекватності побудованох моделі.
3. Виконайте індивідуальне завдання.

Короткі теоретичні відомості

Оцінка значущості рівняння регресії

Перевірити значущість рівняння регресії – значить установити, чи відповідає математична модель експериментальним даним і чи досить включених у рівняння пояснюючих змінних (однієї або декількох) для опису результуючих факторів.

Щільність зв’язку явищ, що вивчаються, можна оцінити за допомогою лінійного коефіцієнту парної кореляції r_{xy} $(-1 \le r_{xy} \le 1)$:

$$r_{xy} = \frac{\frac{1}{n}\sum_i (x_i - \overline{x})(y_i - \overline{y})}{\sigma_x \sigma_y} = \frac{\overline{yx} - \overline{y} \cdot \overline{x}}{\sqrt{(\overline{x^2} - \overline{x}^2) \cdot (\overline{y^2} - \overline{y}^2)}} .$$

Коефіцієнт кореляції має наступні *властивості*:

1) він приймає значення на відрізку [–1; 1], тобто $-1 \le r \le 1$. Чим

ближче $|r|$ до 1, тим тісніше кореляційний зв'язок.

2) при $|r| = 1$ кореляційний зв'язок стає функціональним. При цьому всі спостережувані значення лежать на одній лінії.

3) при $r = 0$ кореляційний зв'язок відсутній та лінія регресії паралельна осі x.

При $r > 0$ ($b_1 > 0$) кореляційний зв'язок називається *прямим*.

При $r < 0$ ($b_1 < 0$) кореляційний зв'язок називається *зворотним*.

Коефіцієнт кореляції для нелінійної регресії ρ_{xy} ($0 \le \rho_{xy} \le 1$):

$$\rho_{xy} = \sqrt{1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \overline{y})^2}}.$$

де y_i - фактичні значення результуючої, $\hat{y}_i$ - теоретичні (розрахункові) значення результуючої.

Загальну якість рівняння регресії оцінюють через розрахунок *коефіцієнта детермінації*.

Коефіцієнт детермінації показує, яка частина зміни залежної змінної обумовлена зміною незалежного фактора ($0 \le R^2 \le 1$).

Чим ближче R^2 до 1, тим краще рівняння регресії наближає (апроксимує) експериментальні дані.

У випадку парної регресії $R^2 = r^2$, де r – коефіцієнт кореляції.

Коефіцієнт детермінації обчислюється за формулою:

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

Середня похибка апроксимації – середнє відносне відхилення розрахункових значень від фактичних:

$$\overline{A} = \frac{1}{n} \sum \left| \frac{y_i - \hat{y}_i}{y_i} \right| \cdot 100\%.$$

Побудоване рівняння регресії вважається задовільним, якщо значення $\overline{A}$ не перевищує 10-12 %.

Чим вищій показник детермінації або чим менша середня похибка апроксимації, тим краще побудована модель описує вихідні дані.

Оцінка значущості всього рівняння регресії в цілому здійснюється за допомогою F-критерія Фішера.

F-критерій Фішера полягає у тому, що проводиться перевірка гіпотези H_0 про статистичну значущість рівняння регресії. Для цього виконується порівняння фактичного $F_{факт}$ та табличного (критичного) $F_{табл}$ значень F-критерія Фішера.

$F_{факт}$ визначається за формулою:

$$F_{факт}=\frac{\sum\frac{(\hat{y}_i-\bar{y})^2}{m}}{\sum\frac{(y_i-\hat{y}_i)^2}{n-m-1}}=\frac{r_{xy}^2}{1-r_{xy}^2}\cdot\frac{n-m-1}{m},$$

де n- кількість одиниць сукупності, m- кількість параметрів при змінних (для лінійної регресії m=1).

Для нелінійної регресії замість r_{xy}^2 використовують R^2.

$F_{табл}$ - максимально можливе значення критерію під впливом випадкових факторів при ступенях вільності $k_1=m$, $k_2=n-m-1$ та рівні значущості α. Рівень значущості α - ймовірність відкинути правильну гіпотезу при умові, що вона є вірна. Зазвичай α надають значення 0,05 або 0,01. Це означає, що у 5% або 1% випадків ми можемо помилитися, а у 95% або 99% випадків (рівень довіри) наші висновки будуть правильними.

Якщо $F_{табл}<F_{факт}$, то гіпотеза H_0 відхиляється та визнається статистична значимість та надійність рівняння регресії.

Якщо $F_{табл}>F_{факт}$, то гіпотеза H_0 не відхиляється та визнається статистична незначимість та ненадійність рівняння регресії.

Для оцінки статистичної значущості коефіцієнтів лінійної регресії та лінійного коефіцієнта парної кореляції r_{xy} застосовують t-критерій Стьюдента та розраховують довірчі інтервали кожного з показників.

t-критерій Стьюдента полягає у тому, що висувається гіпотеза H_0 про випадкову природу показників, тобто про незначне їх відхилення від нулю. Фактичне значення критерія $t_{факт}$ для коефіцієнтів регресії та коефіцієнта кореляції r_{xy} розраховується шляхом співставлення їх значень із величиною стандартної похибки:

$$t_{b_0}=\frac{b_0}{m_{b_b}};\ t_{b_1}=\frac{b_1}{m_{b_1}};\ t_r=\frac{r_{xy}}{m_{r_{xy}}}.$$

Стандартні похибки обчислюється за формулами:

$$m_{b_1}=\sqrt{\frac{\frac{\sum(y_i-\hat{y}_i)^2}{n-2}}{\sum(x_i-\bar{x})^2}};\ m_{b_0}=\sqrt{\frac{\sum(y_i-\hat{y}_i)^2}{n-2}\cdot\frac{\sum x_i^2}{n\sum(x_i-\bar{x})^2}};\ m_{r_{xy}}=\sqrt{\frac{1-r_{xy}^2}{n-2}}.$$

Порівнюючи фактичне та табличне значення для t-критерія роблять відповідні висновки.

$t_{табл}$- максимально можливе значення критерія під впливом випадкових факторів для $k=n-2$ ступенів вільності та рівні значущості α.

Якщо $t_{табл}<t_{факт}$, то гіпотеза H_0 відхиляється (тобто коефіцієнти рівняння регресії та коефіцієнт кореляції r_{xy} не випадково відмінні від нуля та

сформовані під впливом систематично діючого фактора x).

Якщо $t_{табл} > t_{факт}$, то гіпотеза H_0 не відхиляється та визнається випадкова природа формування b_0, b_1 та r_{xy}.

Довірчі інтервали

Довірчі інтервали визначають межі, в яких лежать точні значення визначених показників, із заданим ступенем достовірності.

Для розрахунку довірчих інтервалів для параметрів b_0 та b_1 рівняння лінійної регресії визначають граничну похибку Δ для кожного показника:

$$\Delta_{b_0} = t_{табл} \cdot m_{b_0};\ \Delta_{b_1} = t_{табл} \cdot m_{b_1}.$$

$t_{табл}$ – це табличне значення t-критерія Стьюдента для $k = n - 2$ ступенів вільності та заданого рівня значущості α.

Тоді довірчі інтервали обчислюються:

$$\gamma_{b_0} = b_0 \pm \Delta_{b_0};\ \gamma_{b_1} = b_1 \pm \Delta_{b_1}.$$

Якщо в межі довірчого інтервалу потрапляє нуль (тобто нижня границя від'ємна, а верхня – додатна), то параметр, що оцінюється приймається за нуль.

Приклад розв'язування задачі

Приклад. З метою вивчення роботи 10 однотипних підприємств проведені статистичні дослідження залежності їх прибутку (*Y*, млн грн) від вартості основних фондів (*x*, млн грн).

Необхідно:

1. Побудувати діаграму розсіювання вихідних даних. За її характером записати вигляд шуканої залежності.

2. Обчислити коефіцієнт кореляції між *Y* та *x* та зробити висновок щодо наявності залежності між цими величинами.

3. Невідомі коефіцієнти функції регресії оцінити методом найменших квадратів. Побудувати графік отриманої функції на діаграмі розсіювання.

4. Проаналізувати якість побудованої моделі та оцінок її параметрів.

5. Обчислити прогнозне значення прибутку десятого підприємства, якщо у наступному році передбачається збільшення величини його основних фондів на 5% від їх попереднього значення.

Дані про роботу підприємств наведені в таблиці 2.5.

Таблиця 2.5 - Показники роботи 10 однотипних підприємств

Y	33.0	33.1	33.6	34.0	34.7	34.9	35.0	36.1	36.0	37.0
x	240	244	246	248	250	252	254	260	264	270

Розв'язання

1. Вибір формули зв'язку змінних називається *специфікацією* рівняння регресії. У випадку однієї пояснюючої змінної вибір формули, зазвичай, здійснюється згідно з графічним зображенням реальних статистичних даних у вигляді точок у декартовій системі координат, що називається *кореляційним полем (діаграмою розсіювання)*.

Аналізуючи діаграму розсіювання, обирають вигляд функції регресії. Вона може бути лінійною ($\hat{y} = b_0 + b_1 x$), логарифмічною ($\hat{y} = b_0 + b_1 \ln x$), показниковою ($\hat{y} = b_0 \cdot b_1^{\,x}$), степеневою ($\hat{y} = b_0 \cdot x^{b_1}$) та ін.

Кореляційне поле для нашого прикладу зображено на рисунке 2.3

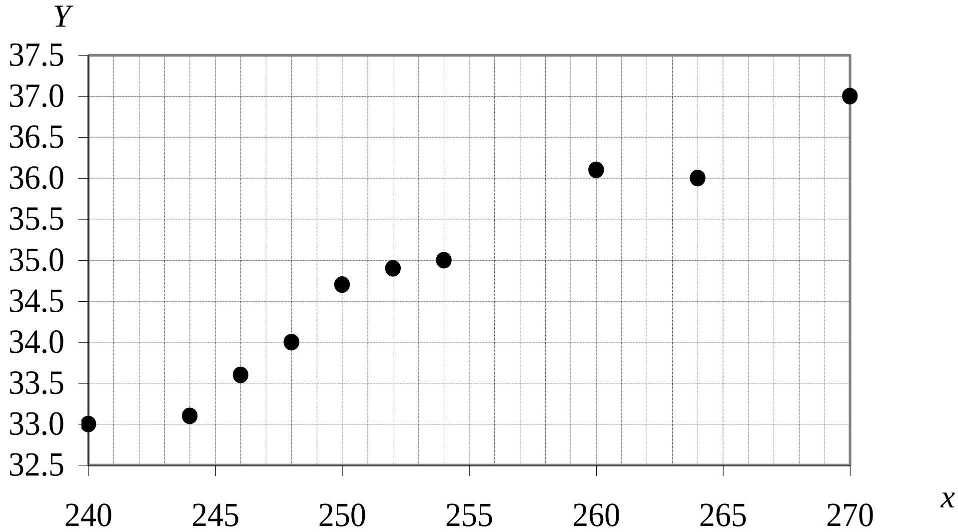


Рисунок 2.3. Кореляційне поле для даних з таблиці 2.5

Виходячи з положення точок на кореляційному полі, припускаємо, що залежність між Y та x – лінійна:

$$\hat{y} = b_0 + b_1 x.$$

2. Коефіцієнт кореляції r служить індикатором наявності додатного (змінні змінюються в одному напрямку, $r > 0$) або від'ємного (змінні змінюються в різних напрямках, $r < 0$) зв'язку, тобто дозволяє виявити якісні залежності між досліджуваними величинами. Окрім цього, за його допомогою можна визначити силу залежності між показником y та фактором x.

Вибірковий коефіцієнт кореляції r між величинами y та x обчислюється за формулою

$$r = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sqrt{\overline{x^2} - \bar{x}^2} \cdot \sqrt{\overline{y^2} - \bar{y}^2}},$$

де $\bar{z} = \frac{1}{n}\sum_{i=1}^{n} z_i$ – середнє значення величини z, n – обсяг вибірки.

Обчислимо вибірковий коефіцієнт кореляції за даними прикладу. Для наочності обчислень побудуємо таблиця 2.6.

Таблиця 2.6 - Додаткові обчислення

i	x	Y	x^2	xy	y^2
1	240	33.0	57600	7920.0	1089.00
2	244	33.1	59536	8076.4	1095.61
3	246	33.6	60516	8265.6	1128.96
4	248	34.0	61504	8432.0	1156.00
5	250	34.7	62500	8675.0	1204.09
6	252	34.9	63504	8794.8	1218.01
7	254	35.0	64516	8890.0	1225.00
8	260	36.1	67600	9386.0	1303.21
9	264	36.0	69696	9504.0	1296.00
10	270	37.0	72900	9990.0	1369.00
Сума	2528	347.4	639872	87933.8	12084.88
Середнє	252.8	34.74	63987.2	8793.38	1208.488

Тоді

$$r = \frac{8793{,}38 - 252{,}8 \cdot 34{,}74}{\sqrt{63987{,}2 - (252{,}8)^2} \cdot \sqrt{1208{,}488 - (34{,}74)^2}} = 0{,}9795.$$

Аналізуючи значення r, можна зазначити, що між величинами y та x існує додатний зв'язок ($r > 0$), тобто зі збільшенням величини основних фондів спостерігається збільшення прибутку даної фірми.

Отже, між прибутком фірми та величиною основних фондів існує значущий лінійний зв'язок.

3. Рівняння лінійної регресії, що характеризує залежність прибутку y фірми від величини x будемо шукати у вигляді:

$$\hat{y} = b_0 + b_1 x + u\,,$$

де b_0 і b_1 – невідомі коефіцієнти; u – випадковий член, що характеризує дію на показник y неврахованих факторів, помилок в даних, помилок округлень і т. д.

Використовуючи метод найменших квадратів (МНК), для знаходження оцінок (наближених значень) b_0 і b_1 (невідомих коефіцієнтів рівняння регресії) скористаємося системою рівнянь

$$\begin{cases} nb_0 + b_1 \sum\limits_{i=1}^{n} x_i = \sum\limits_{i=1}^{n} y_i\,, \\ b_0 \sum\limits_{i=1}^{n} x_i + b_1 \sum\limits_{i=1}^{n} x_i^2 = \sum\limits_{i=1}^{n} x_i y_i\,. \end{cases}$$

Поділивши рівняння системи на n, отримаємо:

$$\begin{cases} b_0 + b_1 \bar{x} = \bar{y}, \\ b_0 \bar{x} + b_1 \overline{x^2} = \overline{xy}. \end{cases} \Rightarrow \begin{cases} b_0 = \bar{y} - b_1 \bar{x} \\ b_1 = \dfrac{\overline{xy} - \bar{x} \cdot \bar{y}}{\overline{x^2} - \bar{x}^2} = r \cdot \dfrac{\sqrt{\overline{y^2} - \bar{y}^2}}{\sqrt{\overline{x^2} - \bar{x}^2}}. \end{cases}$$

Отже (з таблиці 2.6),

$$b_1 = \frac{8793{,}38 - 252{,}8 \cdot 34{,}74}{63987{,}2 - (252{,}8)^2} = 0{,}140,$$

$$b_0 = 34{,}74 - 0{,}140 \cdot 252{,}8 = \text{-}0{,}644,$$

Тоді рівняння лінійної регресії набирає вигляду:

$$\hat{y} = -0{,}644 + 0{,}14x.$$

Побудувавши пряму рівняння регресії, на діаграмі розсіювання отримаємо рисунок 2.4.

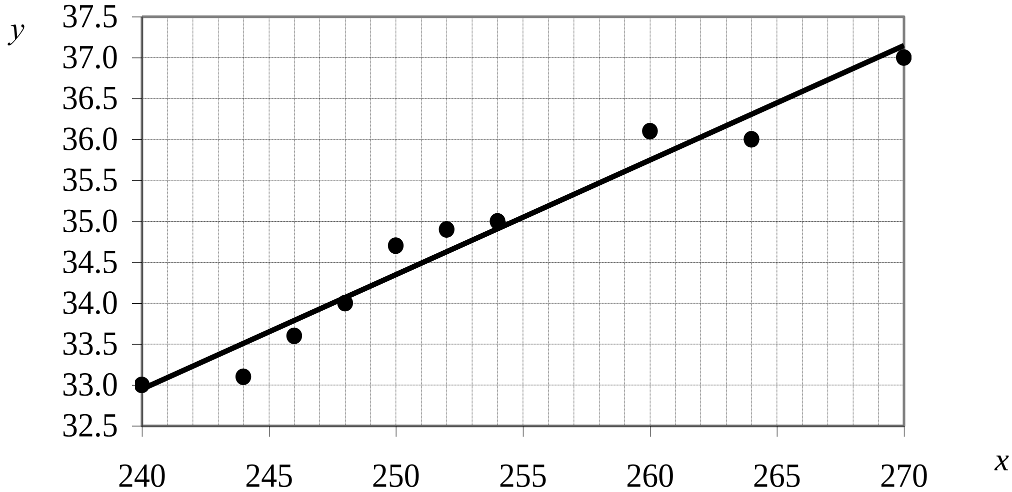


Рисунок 2.4 - Рівняння прямої регресії на діаграмі розсіювання

4. Мірою загальної якості рівняння регресії (його відповідності статистичним даним) є коефіцієнт детермінації R^2, який обчислюється за формулою

$$R^2 = 1 - \frac{\sum \hat{u}_i^2}{\sum (y_i - \bar{y})^2} = 1 - \frac{0{,}656}{16{,}204} = 0{,}9595.$$

Для обчислення коефіцієнта детермінації використано дані з таблиці 2.7.

Таблиця 2.7 - Обчислення випадкових відхилень

i	y_i	$\hat{y}_i$ =-0,644+0,140 x_i	$\hat{u}_i = y_i - \hat{y}_i$	$\hat{u}_i^2$	$(y_i - \bar{y})^2$
1	33,0	32,948	0,052	0,003	3,028
2	33,1	33,508	-0,408	0,167	2,690
3	33,6	33,788	-0,188	0,035	1,300
4	34,0	34,068	-0,068	0,005	0,548
5	34,7	34,348	0,352	0,124	0,002
6	34,9	34,628	0,272	0,074	0,026
7	35,0	34,908	0,092	0,008	0,068
8	36,1	35,748	0,352	0,124	1,850
9	36,0	36,308	-0,308	0,095	1,588
10	37,0	37,147	-0,147	0,022	5,108
$\sum$	347,4	347,400	0,000	0,656	16,204

Високе значення коефіцієнта детермінації R^2 свідчить про високу якість побудованого рівняння регресії.

Звідси $r^2 = 0{,}9595$, $r = 0{,}795$. Тобто, зв’язок із змінними моделі тісний.

Оцінимо значущість побудованого рівняння регресії в цілому за допомогою F-критерію:

$$F_{факт} = \frac{r_{xy}^2}{1 - r_{xy}^2} \cdot \frac{n - m - 1}{m} = \frac{0{,}9595}{1 - 0{,}9595} \cdot \frac{10 - 1 - 1}{1} = 189{,}1.$$

де n - кількість одиниць сукупності, m - кількість параметрів при змінних (для лінійної регресії m =1).

$F_{табл}$ при ступенях вільності $k_1 = m = 1$, $k_2 = n - m - 1 = 8$ та рівні значущості $\alpha = 0.05$ дорівнює 5,32.

Оскільки $F_{табл} < F_{факт}$, то визнаємо статистичну значимість та надійність побудованого рівняння регресії.

Оцінимо статистичну значущість коефіцієнтів лінійної регресії та лінійного коефіцієнта парної кореляції за допомогою t – критерія Стьюдента.

Обчислимо спочатку стандартні похибки, використовуючі додаткові обчислення (дивись таблицю 2.8):

$$m_{b_0} = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n - 2} \cdot \frac{\sum x_i^{\,2}}{n \sum (x_i - \bar{x})^2}} = \sqrt{\frac{0{,}656}{8} \cdot \frac{639872}{10 \cdot 793{,}6}} = 2{,}57;$$

$$m_{b_1} = \sqrt{\frac{\frac{\sum (y_i - \hat{y}_i)^2}{n - 2}}{\sum (x_i - \bar{x})^2}} = \sqrt{\frac{\frac{0{,}656}{8}}{793{,}6}} = 0{,}01;$$

$$m_{r_{xy}} = \sqrt{\frac{1 - r_{xy}^2}{n - 2}} = 0{,}071.$$

Таблиця 2.8 - Допоміжні обчислення

i	$\hat{u}_i^{\,2} = (y_i - \hat{y}_i)^2$	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
1	0,003	-12,8	163,84
2	0,167	-8,8	77,44
3	0,035	-6,8	46,24
4	0,005	-4,8	23,04
5	0,124	-2,8	7,84
6	0,074	-0,8	0,64
7	0,008	1,2	1,44
8	0,124	7,2	51,84
9	0,095	11,2	125,44
10	0,022	17,2	295,84
$\sum$	0,656		793,6

Співставимо значення параметрів рівняння регресії та коефіцієнта кореляції:

$$t_{b_0} = \frac{b_0}{m_{b_b}} = \frac{0,644}{2,57} = 0,25;$$

$$t_{b_1} = \frac{b_1}{m_{b_1}} = \frac{0,14}{0,01} = 14;$$

$$t_r = \frac{r_{xy}}{m_{r_{xy}}} = \frac{0,9795}{0,071} = 13,8.$$

Порівняємо фактичне та табличне значення для t-критерія.

$t_{табл}$ для $k = n - 2 = 8$ ступенів вільності та рівні значущості $\alpha = 0,05$ дорівнює 2,31.

Оскільки:

- $t_{табл} < t_{факт}$ для r_{xy} та b_1, то ці коефіцієнти є статистично значущі.
- $t_{табл} > t_{факт}$ для b_0, то визнається випадкова природа формування цього показника.

Отже, коефіцієнт b_1 є значущим, тобто кут нахилу отриманої прямої, знайдений з великою точністю, і розкид значень Y навколо лінії регресії відносно малий. При цьому коефіцієнт b_0 незначущий.

Визначимо довірчі інтервали для коефіцієнтів рівняння регресії.

Обчислимо граничну похибку Δ для кожного показника:

$\Delta_{b_0} = t_{табл} \cdot m_{b_0} = 2,31 \cdot 2,57 = 5,94$; $\Delta_{b_1} = t_{табл} \cdot m_{b_1} = 2,31 \cdot 0,01 = 0,0231$.

$t_{табл}$ – це табличне значення t-критерія Стьюдента для $k = n - 2$ ступенів вільності та заданого рівня значущості α дорівнює 2,31.

Тоді довірчі інтервали обчислюються:

$\gamma_{b_0} = -0,644 \pm 5,94$; $\gamma_{b_1} = 0,14 \pm 0,0231$.

Оскільки параметр не може бути одночасно і додатнім, і від'ємним, то коефіцієнт b_0 треба прийняти за нуль.

5. Очікуване прогнозне значення основних фондів десятого підприємства у наступному році дорівнює: $x_{10}^{*} = 1,05 \cdot x_{10} = 283,5$, тоді

$$\hat{y}(283,5) = -0,644 + 0,140 \cdot 283,5 = 39,04.$$

Завдання для індивідуальної та самостійної роботи студентів

Приклад. З метою вивчення роботи 10 однотипних підприємств проведені статистичні дослідження залежності їх прибутку (Y, млн. грн.) від вартості основних фондів (x, млн. грн.).

Необхідно:

1. Побудувати діаграму розсіювання вихідних даних. За її характером записати вигляд шуканої залежності.

2. Обчислити коефіцієнт кореляції між Y та x та зробити висновок щодо наявності залежності між цими величинами.

3. Невідомі коефіцієнти функції регресії оцінити методом найменших квадратів. Обчислити коефіцієнт детермінації лінійної моделі. Побудувати графік отриманої функції на діаграмі розсіювання.

4. Обчислити середньоквадратичні відхилення невідомих коефіцієнтів та перевірити їх значущість. Зробити висновки.

5. Обчислити прогнозне значення прибутку десятого підприємства, якщо у наступному році передбачається збільшення величини його основних фондів на 5% від їх попереднього значення.

Дані про роботу 10 підприємств наведені в таблиці 2.9.

Таблиця 2.9 - Дані про роботу 10 однотипних підприємств

1	*Y*	13,2	12,2	13,8	13,9	15,6	15,6	16	17,3	18,1	19,5
	x	181	171	177	179	190	186	185	189	198	208
2	*Y*	26,5	30,3	28,9	32,8	34,6	33,9	38,4	41,6	42,9	43,8
	x	282	282	274	285	300	286	291	305	307	312
3	*Y*	28,4	30	32	31,1	30,4	32,6	33,1	33,7	36,5	38,6
	x	258	267	272	269	263	269	270	271	280	284
4	*Y*	22,1	24,2	25,4	28,1	27,7	29,9	32,6	33,1	33,7	34,2
	x	117	129	134	136	132	141	146	139	140	150
5	*Y*	13,1	14,8	17,7	21,9	22,3	26,1	27,2	27,3	29,4	31,8
	x	239	235	252	252	244	266	265	263	266	275
6	*Y*	20,5	22	24	23,9	24,6	27,4	29,1	29,1	31,1	32,8
	x	131	129	133	135	145	151	157	143	156	165
7	*Y*	13,2	15,4	16,4	17	18,2	20,6	22,9	22,8	23,4	26,1
	x	244	247	266	261	279	291	282	292	291	310
8	*Y*	13,8	15	13	18	19,5	18,8	19,6	20,6	23,1	24
	x	202	209	196	210	222	220	219	217	228	240
9	*Y*	17,4	20,8	23,1	22,4	24,7	25,9	27,5	28,4	31,7	32
	x	256	266	272	263	273	274	279	287	292	285
10	*Y*	22,6	22,4	22,6	23	25,3	26,3	26,6	27,7	29,8	30,9
	x	227	232	234	239	251	250	259	255	273	268
11	*Y*	26,8	25,7	28,4	28,2	28,5	30,1	29,9	32,1	32,3	34,3
	x	189	187	189	186	184	200	192	205	208	217
12	*Y*	11	12,4	12,7	13,8	15,2	17,7	18	18,9	20,4	22
	x	291	294	292	293	296	308	310	311	318	312
13	*Y*	18	18,8	20	20,4	23,2	25,7	27	27,8	31,6	32,4
	x	161	171	167	161	168	177	185	176	186	185
14	*Y*	20,6	20,4	23,3	24,6	26,7	28,9	30,1	32,6	33,5	34,2
	x	196	200	201	215	221	226	217	240	245	242
15	*Y*	26,6	27,1	28,1	29,4	30,4	32,1	31,7	32,9	32,8	34,3
	x	240	237	258	260	263	267	263	272	274	277
16	*Y*	21,9	22	22,2	24,1	25,6	26,8	27,9	29,5	29,2	30,7
	x	196	187	187	194	194	202	204	208	211	208
17	*Y*	12,3	15,7	17,3	17	19,3	18,9	20,8	24,2	23,4	22,7
	x	206	213	220	217	232	230	245	250	239	256
18	*Y*	11,6	13,7	14,7	16,7	17	18,9	21,9	23	22,8	23,9
	x	189	196	197	202	202	201	213	222	208	211
19	*Y*	25,5	26,8	28,5	29,9	30,1	31,8	33,3	35,3	38,2	38,4
	x	245	246	247	258	251	257	279	267	290	290
20	*Y*	13,9	17	15	16,7	19,7	20,6	20,8	20,1	20,3	21,4
	x	226	244	231	232	247	243	255	247	257	260

Продовження таблиці 2.9

21	*Y*	18,9	23,2	21,3	23,5	24,9	24,6	26,8	27,7	30,4	29,3
	x	245	255	242	262	269	275	271	266	291	296
22	*Y*	23,5	26,6	26,6	30,5	30,9	33,2	35,2	35,3	38,2	42,1
	x	92	97	108	111	102	117	120	113	130	141
23	*Y*	21,5	23,7	23,2	24,6	27,5	25,4	27,6	28,9	28,5	30,8
	x	125	136	128	151	152	135	157	158	173	173
24	*Y*	13,2	15,9	19,1	17,6	18,8	21,1	21,9	23,6	23	23,2
	x	194	210	231	231	232	237	235	254	252	242
25	*Y*	13,4	16,7	18,5	19,7	21,1	21,9	24,2	26,7	27,3	31,3
	x	256	271	269	269	279	272	279	278	288	293
26	*Y*	23,4	25,2	25,2	25,3	29	28	28,9	28,1	31,9	29,3
	x	134	147	144	143	161	153	166	161	164	161
27	*Y*	17,4	19,1	18,9	19,7	20,4	21,9	21	25,6	26,8	26,2
	x	292	303	303	307	296	321	309	337	333	340
28	*Y*	14,2	15,8	17,6	21,9	19,7	21,4	26,2	24,6	27,9	31,2
	x	207	221	219	238	232	227	245	234	240	264
29	*Y*	18,3	19,5	19,7	21,1	19	20,2	21	21,4	21,8	22,4
	x	238	237	243	252	239	250	257	258	264	268
30	*Y*	16,9	19,3	23,3	21,6	23,9	25,9	28,7	29,2	30,7	34,3
	x	151	156	161	160	168	186	195	192	179	195

Дані про роботу підприємств необхідно вибрати з таблиці відповідно до номера варіанта. Номер варіант визначається за вказівкою викладача.

2.3 Побудова множинної економетричної моделі

Відведений час: 4 год.

Мета: формувати навички будувати багатофакторні економетричні моделі та здійснювати аналіз якості побудованої моделі та оцінок її параметрів.

Завдання для практичного заняття:

1. Пригадайте основні теоретичні питання теми.
2. Орієнтовні запитання та завдання:
   - як визначається лінійна множинна економетрична модель?
   - при яких умовах можна застосувати метод найменших квадратів для оцінки параметрів моделі ?
   - за якою формулою знаходять оцінки параметрів моделі?
   - як можна оцінити статистичну значущість побудованої моделі?
3. Виконайте індивідуальне завдання.

Короткі теоретичні відомості

На будь-який економічний показник y, як правило, впливає не один, а декілька факторів (регресорів) $x_1, x_{2,} ..., x_m$. Так, наприклад, попит населення на певний товар буде визначатися не тільки ціною на нього, але й цінами на його замінники, доходами споживачів й іншими факторами.

У цьому випадку маємо справу з множинною (багатофакторною) лінійною моделлю (регресією), що описує взаємний зв'язок між залежною змінною y та факторами $x_1, x_{2,} ..., x_m$ і яку можна подати у такому вигляді:

$$y = f(x_1, x_2, x_3 ... x_m, \varepsilon),$$

де y - залежна (результуюча) змінна; $x_j, (j = \overline{1,m})$ - незалежні змінні; ε - стохастична складова.

Загальний запис теоретичної лінійної множинної регресії може бути зроблений у такому вигляді:

$$\begin{gathered} y_1 = b_0 + b_1 x_{11} + b_2 x_{12} + ... + b_j x_{1j} + ... + b_m x_{1m} + \varepsilon_1, \\ y_2 = b_0 + b_1 x_{12} + b_2 x_{22} + ... + b_j x_{2j} + ... + b_m x_{2m} + \varepsilon_2, \\ \vdots \\ y_i = b_0 + b_1 x_{i1} + b_2 x_{i2} + ... + b_j x_{ij} + ... + b_m x_{im} + \varepsilon_3, \\ \vdots \\ y_n = b_0 + b_1 x_{n1} + b_2 x_{n2} + ... + b_j x_{nj} + ... + b_m x_{nm} + \varepsilon_n, \end{gathered}$$

де $b_j \ (j = \overline{1,m})$ – теоретичні коефіцієнти регресії (часткові коефіцієнти) або параметри теоретичної регресії, які характеризують реакцію залежної змінної $y_i \ (i = \overline{1,n})$ на зміну кожного фактора $X_j \ (j = \overline{1,m})$;

b_0 – вільний член, який визначає значення y_i за умови, коли значення факторів дорівнюють нулеві;

$x_{ij} \quad (i = \overline{1,n}, \ j = \overline{1,m})$ – значення X_j-го фактора при i-ому спостереженні;

ε_i – випадковий збудник при i-ому спостереженні.

Для однозначного визначення параметрів b_j моделі необхідно, щоб виконувалась нерівність

$n \geq m + 1,$

де n – число спостережень;

m – число факторів в моделі.

У векторно-матричній формі теоретичну модель можна подати так:

$$\vec{Y} = X \cdot \vec{A} + \vec{\varepsilon},$$

де

$$\vec{Y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_i \\ \vdots \\ y_n \end{pmatrix}, \quad X = \begin{pmatrix} 1 & x_{11} & x_{12} & ... & x_{1j} & ... & x_{1m} \\ 1 & x_{21} & x_{22} & ... & x_{2j} & ... & x_{2m} \\ \vdots & & \vdots & & \vdots & & \vdots \\ 1 & x_{i1} & x_{i2} & ... & x_{ij} & ... & x_{im} \\ \vdots & & \vdots & & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & ... & x_{nj} & ... & x_{nm} \end{pmatrix}, \quad \vec{A} = \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_j \\ \vdots \\ b_m \end{pmatrix}, \quad \vec{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_i \\ \vdots \\ \varepsilon_n \end{pmatrix}.$$

Передумови застосування методу найменших квадратів (1мнк)

Нехай економетрична модель у матричній формі має вигляд

$$Y = XA + u,$$

де Y - вектор значень залежної змінної;

X - матриця незалежних змінних розміром $n \times m$ (n - число спостережень, m - кількість незалежних змінних);

A - вектор оцінок параметрів моделі;

u - вектор залишків.

Щоб застосувати 1МНК для оцінки параметрів моделі, необхідне виконання таких умов:

1) математичне сподівання залишків дорівнює нулю.

Коли математичне сподівання залишків не дорівнює нулю, то це означає, що існує систематичний вплив на залежну змінну, а до модельної специфікації не введено всіх основних незалежних змінних. Якщо ця передумова не виконується, то йдеться про помилку специфікації.

Зауважимо, що коли економетрична модель має вільний член, то майже завжди за рахунок його значення можна скоригувати рівняння так, щоб математичне сподівання залишків дорівнювало нулю. Отже, для таких моделей перша умова практично виконуватиметься завжди.

2) значення u_i вектора залишків u незалежні між собою і мають постійну дисперсію.

Ця умова передбачає наявність сталої дисперсії залишків. Цю властивість називають ***гомоскедастичністю***. Проте вона може виконуватись лише тоді, коли залишки ***u*** є помилками вимірювання. Якщо залишки акумулюють загальний вплив змінних, які не враховані в моделі, то звичайно дисперсія залишків не може бути сталою величиною, вона змінюється для окремих груп спостережень. У такому разі йдеться про явище ***гетероскедастичності***, яке впливає на методи оцінювання параметрів.

3) незалежні змінні моделі не пов'язані із залишками.

Передбачаєтсья незалежність між залишками і пояснювальними змінними, яка порушується насамперед тоді, коли економетрична модель будується на базі одночасових структурних рівнянь або має лагові змінні. Тоді для оцінювання параметрів моделі використовуються, як правило, дво- або трикроковий метод найменших квадратів.

4) незалежні змінні моделі утворюють лінійно незалежну систему векторів, або, іншими словами, незалежні змінні не повинні бути мультиколінеарними.

Зауважимо, що під поняттям «вектор» будемо розуміти «матриця-вектор», тобто матриця-рядок або матриця-стовпець.

Таким чином всі пояснювальні змінні, які входять до економетричної моделі, мають бути незалежними між собою. Проте очевидно, що в економіці дуже важко вирізнити такий масив незалежних (пояснювальних) змінних, які були б зовсім не пов'язані між собою. Тоді щоразу необхідно з'ясовувати, чи не впливатиме залежність пояснювальних змінних на оцінку параметрів моделі. Це явище називають ***мультиколінеарністю*** змінних, що призводить до ненадійності оцінки параметрів моделі, робить їх чутливими до вибраної специфікації моделі та до конкретного набору даних. Знижується рівень довіри до результатів верифікації моделей з допомогою 1МНК.

Отже, явище мультиколінеарності з усіх точок зору є дуже небажаним. Але воно досить поширене.

Етапи побудови множинної лінійної економетричної моделі

1. *Вибір усіх можливих факторів, які впливають на результуючу змінну.* Відбувається оцінювання відповідного явища або процесу з точки зору економічної теорії. Приміром, від чого може залежати валовий внутрішній продукт (від величини основних та оборотних фондів, величини інвестицій, кількість зайнятого населення тощо).

2. *Аналіз та вимірювання факторних ознак*. Здійснюється збір статистичних даних. Тобто проводяться спостереження, вимірювання, кількісний аналіз відібраних факторів тощо.

3. *Економетричний аналіз факторів*. Розраховується матриця коефіцієнтів парної кореляції

$$R = \begin{pmatrix} 1 \;\; r_{y1} \;\; r_{y2} .. r_{yk} \\ r_{1y} \;\; 1 \;\; r_{12} .. r_{1k} \\ r_{2y} \;\; r_{21} \;\; 1 .. r_{2k} \\ \dots\dots\dots\dots\dots \\ r_{ky} \;\; r_{k1} \;\; r_{k2} \dots 1 \end{pmatrix}$$

Ця матриця є симетричною, тобто коефіцієнти кореляції між результуючою змінною y та факторною ознакою x_j рівні між собою ($r_{yj} = r_{jy}$, $j = \overline{1,k}$), коефіцієнти кореляції між i-м та j-м факторами теж рівні ($r_{ij} = r_{ji}$, $i = \overline{1,k}$, $j = \overline{1,k}$).

Якщо значення коефіцієнта парної кореляції між факторами наближене до одиниці, то це свідчить про тісний зв’язок між ними. У цьому випадку один із факторів необхідно вилучити з розгляду (бажано залишити фактор, який є вагомішим з економічної точки зору або той, що сильніше корелює із результуючою змінною y).

4. *Вибір методу оцінювання параметрів та побудова моделі*. Визначається метод побудови множинної лінійної економетричної моделі. Найчастіше застосовують метод найменших квадратів (1МНК) або його модифікації (наприклад, узагальнений метод найменших квадратів або метод Ейткена).

<u>Оператор оцінювання за 1МНК</u>

$$\hat{A} = (X'X)^{-1} X'Y$$

де X - матриця незалежних змінних, X' - транспонована матриця X, $(X'X)^{-1}$ - матриця обернена до матриці $X'X$, $X'Y$ - добуток транспонованої матриці X' та матриці-стовпця Y, $\hat{A}$ - матриця стовпець шуканих параметрів економетричної моделі.

5. *Перевірка моделі на адекватність.* Перевіряється множинна лінійна економетрична модель на точність за такою ж схемою, як і перевіряється парна лінійна модель: обчислюються стандартні похибки моделі, коефіцієнти множинної детермінації та кореляції, проводять діагностику моделі.

6. *Аналіз отриманих результатів.* Аналізуються отримані параметри економетричної моделі з економічної точки зору.

Аналіз ступеня адекватності побудованої моделі та вибіркових даних

На початковому етапі правильність виконаних розрахунків можна перевірити, порівнюючи середні значення фактичні $\bar{y}$ та розрахункові $\bar{\hat{y}}$. Якщо $\bar{y} = \bar{\hat{y}}$, то попередні розрахунки є вірними.

Визначимо ступінь адекватності моделі та статистичних даних через відхилення між фактичними значеннями та значеннями, обчисленими за моделлю:

$$e_i = y_i - \hat{y}_i \ .$$

Середнє значення $\bar{e} = 0$, отже, розбіжностей не існує, модель адекватна.

Проведемо дисперсійний аналіз побудованої моделі.

Для дисперсій виконується рівність:

$$\sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 + \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \ ,$$

де $\sum_{i=1}^{n}(y_i - \bar{y})^2$ - загальна сума квадратів відхилень,

$\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2$ - сума квадратів відхилень, обумовлена регресією («факторна»),

$\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$ - остаточна сума квадратів відхилень.

Коефіцієнт множинної детермінації:

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

Перевіримо статистичну значущість коефіцієнта множинної детермінації R^2 за критерієм Фішера:

$$F = \frac{R^2}{1-R^2} \cdot \frac{n-m-1}{m};$$

Знаходимо табличне значення критерією Фішера F_{α,k_1,k_2} при рівні значущості α (0,05 або 0,01) та ступенях свободи $k_1=m$ і $k_2=n-m-1$.

Якщо $F > F_{\alpha,k_1,k_2}$, то статистична гіпотеза $H_0 : R^2 = 0$ відхиляється, отже, всі фактори мають вплив на залежну змінну y.

Обчислимо коваріаційну матрицю.

Оцінки коваріаційної матриці $\operatorname{cov}(\vec{A}\cdot(\vec{A})') = \operatorname{var}(\hat{A}) = \hat{\sigma}_u^2 (X'X)^{-1}$ використовуються для знаходження стандартних помилок та обчислення довірчих інтервалів оцінок параметрів $\widehat{b}_j$. Вони використовуються й при перевірці їх статистичної значущості. На головній діагоналі матриці $\operatorname{var}(\hat{A})$ містяться оцінки дисперсій $\sigma_{b_j}^2$ j-ї оцінки параметрів, що ж до елементів $\sigma_{b_j b_k}$ ($j \neq k$), які розміщені поза головною діагоналлю, то вони є оцінками коваріації між $\widehat{b}_j$ і $\widehat{b}_k$.

Отже,

$$\operatorname{var}(\widehat{A}) = \sigma_u^2 (X'X)^{-1} = \begin{pmatrix} \sigma_{b_1}^2 & \sigma_{b_1 b_2} & \dots & \sigma_{b_1 b_j} & \dots & \sigma_{b_1 b_m} \\ \sigma_{b_2 b_1} & \sigma_{b_2}^2 & \dots & \sigma_{b_2 b_j} & \dots & \sigma_{b_2 b_m} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \sigma_{b_j b_1} & \sigma_{b_j b_2} & \dots & \sigma_{b_j}^2 & \dots & \sigma_{b_j b_m} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \sigma_{b_m b_1} & \sigma_{b_m b_2} & \dots & \sigma_{b_m b_j} & \dots & \sigma_{b_m}^2 \end{pmatrix},$$

де $\hat{\sigma}_u^2$ — незміщена оцінка дисперсії залишків;

Дисперсії залишків:

$$\sigma_u^2 = \frac{Y'Y - A'X'Y}{n-m},$$

де n – число спостережень;
m – число факторів в моделі.

Дисперсії оцінок параметрів знаходяться на головній діагоналі отриманої матриці.

Середньоквадратичні відхилення оцінок параметрів дорівнюють стандартним помилкам оцінки параметра моделі:

$$S_{\widehat{b}_0^*} = \sqrt{\sigma_{\widehat{b}_0}^2}$$

$$S_{\widehat{b}_1} = \sqrt{\sigma_{\widehat{b}_1}^2}$$

$$S_{\widehat{b}_2} = \sqrt{\sigma_{\widehat{b}_2}^2}$$

Як і у випадку простої регресії, квадрат *коефіцієнта множинної кореляції* дорівнює коефіцієнту множинної детермінації. Тобто, має місце рівність:

$$R = \sqrt{R^2}\ .$$

Безпосередньо коефіцієнт множинної кореляції можна обчислити за формулою:

$$r_{y\hat{y}} = R = \frac{\sum_{i=1}^{n}(y_i - \bar{y})(\hat{y}_i - \bar{\bar{y}})}{\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2 \sum_{i=1}^{n}(\hat{y}_i - \bar{\bar{y}})^2}}\ .$$

Перевіримо статистичну значущість коефіцієнта кореляції.

Для цього використаємо t-критерій Стьюдента:

$$t = \frac{R\sqrt{n-m}}{\sqrt{1-R^2}}\,,$$

де R^2 коефіцієнт детермінації моделі, R -коефіцієнт кореляції ($R = \sqrt{R^2}$).

Розраховане за формулою фактичне значення t-критерію зіставляємо його з табличним $t_{табл}$ (знаходимо за статистичними таблицями для відповідного рівня значимості α та числа ступенів вільності $k = n - m - 1$).

Якщо $t > t_{табл}$, то можна зробити висновок про значимість коефіцієнта кореляції.

Перевірка гіпотези про значимість оцінок параметрів моделі (b_0 , b_1 *та* b_2).

Для цього використаємо t-критерію Стьюдента:

$$t_j = \frac{|\hat{b}_j|}{\sqrt{\sigma_e^2 c_{jj}}}\,,$$

де $\hat{b}_j$ - параметри матриці $\hat{A}$ (параметри побудованої економетричної моделі), σ_e^2 - дисперсія залишків, c_{jj} - діагональний елемент матриці $(X'X)^{-1}$ (знаменник являє собою стандартну помилку оцінки параметра моделі $S_{\hat{b}_j} = \sqrt{\sigma_e^2 c_{jj}}$).

Обчислене значення t-критерію зіставляємо його з табличним $t_{табл}$.

Якщо $t > t_{табл}$, то оцінка значимості відповідного параметру моделі є достовірною.

На підставі t-критерію Стьюдента та стандартних помилок розрахуємо *довірчі інтервали* для параметрів $\hat{a}_j$:

$$\hat{b}_j - t_j\sqrt{\sigma_e^2 c_{jj}} \le b_j \le \hat{b}_j + t_j\sqrt{\sigma_e^2 c_{jj}}\ .$$

Для випадку множинної регресії вводиться поняття часткового коефіцієнту еластичності K_{E_j}.

Частковий коефіцієнт еластичності – границя відношення зміни у відсотках *Y* при зміні на один відсоток одного з регресорів X_j:

$$K_{E_i} = \lim_{\Delta x_i \to 0} \frac{\frac{\Delta y}{y}}{\frac{\Delta x_o}{x_j}} = \frac{\partial y}{\partial x_j} \cdot \frac{x_j}{y}.$$

В даному випадку K_{E_i} визначає еластичність впливу обраного регресора X_j на залежну змінну *Y*.

Приклад розв'язування задачі

Приклад. Необхідно провести дослідження залежності ціни автомобіля (y) від таких характеристик як вік авто (x_1) та його об'єм двигуна (x_2) на основі вибіркових даних, наведених в таблиці 2.10.

Таблиця 2.10 – Вихідні дані

№ з/п	Ціна автомобіля (у, тис. у.о.)	Вік автомобіля (x_1, років)	Об'єм двигуна автомобіля (x_2, $дм^3$)
1.	11	6	3
2.	3,2	10	1,3
3.	8,7	10	1,8
4.	1,6	16	1,8
5.	17	2	2,4
6.	18,9	4	4
7.	15,8	3	2
8.	18	6	3
9.	19	3	4
10.	6	6	1,6
11.	13	6	2,6
12.	22,7	5	2,1
13.	13,9	3	2
14.	10,9	3	1,5
15.	9,5	8	2,6
16.	28	3	2,2
17.	6	8	1,8
18.	5,8	9	1,6
19.	2,8	11	1,6
20.	26	1	2,2
21.	10,5	6	2
22.	16,9	7	2,3
23.	3,4	9	1,5
24.	14	1	1,4
25.	23	1	1,8
Середнє	13,024	5,88	2,164

Припустимо, що між ціною та означеними технічними характеристиками існує лінійна залежність:

$$Y = b_0 + b_1 X_1 + b_2 X_2 + \varepsilon$$

Необхідно:

1) визначити $\hat{b}_i$ для залежності між досліджуваним фактором y (ціною автомобіля) та пояснюючими змінними x_1 (вік автомобіля) і x_2 (об'єм двигуна);

2) проаналізувати ступінь адекватності побудованої моделі та вибіркових даних.

Розв'язання

Обчислимо коефіцієнти парної кореляції:

$r_{yx_1} = -0{,}77887 \approx -0{,}8$,

$r_{yx_2} = 0{,}452719 \approx 0{,}5$,

$r_{x_1x_2} = -0{,}26042 \approx -0{,}3$.

Складемо матрицю коефіцієнтів парної кореляції:

$$R = \begin{pmatrix} 1\ r_{y1}\ r_{y2}..r_{yk} \\ r_{1y}\ 1\ r_{12}..r_{1k} \\ r_{2y}\ r_{21}\ 1..r_{2k} \\ \dots\dots\dots\dots \\ r_{ky}\ r_{k1}\ r_{k2}...1 \end{pmatrix} = \begin{pmatrix} 1 & -0{,}8 & 0{,}5 \\ -0{,}8 & 1 & -0{,}3 \\ 0.5 & -0{,}3 & 1 \end{pmatrix}.$$

Оскільки $|r_{yx_1}|$ наближене до одиниці, то це свідчить про тісний кореляційний зв'язок між y та x_1. Значення $|r_{yx_2}|$ свідчить про менш тісний зв'язок y та x_2. Але в цілому ці фактори суттєво впливають на результуючу змінну.

Абсолютна величина парного коефіцієнта кореляції між факторними ознаками x_1 та x_2 дорівнює 0,3. Це значення ближче до нуля, тобто між x_1 та x_2 дуже слабкий зв'язок.

Отже, можна будувати економетричну модель із даними факторними ознаками.

Оцінки параметрів $\hat{b}_i$ визначимо за методом найменших квадратів.

Запишемо теоретичну модель у векторно-матричному вигляді:

$$\vec{Y} = X\vec{A} + \vec{\varepsilon}\ .$$

Сформуємо матрицю X, першим стовпчиком якої будуть елементи, значення яких дорівнюють одиниці, іншими – значення пояснюючих змінних, та вектор $\vec{Y}$, елементи якого складаються із значень залежної змінної.

$$X = \begin{pmatrix} 1 & 6 & 3 \\ 1 & 10 & 1,3 \\ 1 & 10 & 1,8 \\ 1 & 16 & 1,8 \\ 1 & 2 & 2,4 \\ 1 & 4 & 4 \\ 1 & 3 & 2 \\ 1 & 6 & 3 \\ 1 & 3 & 4 \\ 1 & 6 & 1,6 \\ 1 & 6 & 2,6 \\ 1 & 5 & 2,1 \\ 1 & 3 & 2 \\ 1 & 3 & 1,5 \\ 1 & 8 & 2,6 \\ 1 & 3 & 2,2 \\ 1 & 8 & 1,8 \\ 1 & 9 & 1,6 \\ 1 & 11 & 2,2 \\ 1 & 1 & 2 \\ 1 & 6 & 2,3 \\ 1 & 7 & 2 \\ 1 & 9 & 1,5 \\ 1 & 1 & 1,4 \\ 1 & 1 & 1,8 \end{pmatrix} \qquad Y = \begin{pmatrix} 11 \\ 3,2 \\ 8,7 \\ 1,6 \\ 17 \\ 18,9 \\ 15,8 \\ 18 \\ 19 \\ 6 \\ 13 \\ 22,7 \\ 13,9 \\ 10,9 \\ 9,5 \\ 28 \\ 6 \\ 5,8 \\ 2,8 \\ 26 \\ 10,5 \\ 16,9 \\ 3,4 \\ 14 \\ 23 \end{pmatrix}$$

Транспонуємо матрицю X:

$$X' = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 6 & 10 & 10 & 16 & 2 & 4 & 3 & 6 & 3 & 6 & 6 & 5 & 3 & 3 & 8 & 3 & 8 & 9 & 11 & 1 & 6 & 7 & 9 & 1 & 1 \\ 3 & 1,3 & 1,8 & 1,8 & 2,4 & 4 & 2 & 3 & 4 & 1,6 & 2,6 & 2,1 & 2 & 1,5 & 2,6 & 2,2 & 1,8 & 1,6 & 2,2 & 2 & 2,3 & 2 & 1,5 & 1,4 & 1,8 \end{pmatrix}$$

Перемножимо матрицю X' спочатку на матрицю X, потім на вектор $\vec{Y}$:

$$X'X = \begin{pmatrix} 25 & 147 & 54,1 \\ 147 & 1189 & 301,6 \\ 54,1 & 301,6 & 129,45 \end{pmatrix}; \qquad X'\vec{Y} = \begin{pmatrix} 325,6 \\ 1400,4 \\ 762,95 \end{pmatrix}.$$

Обчислимо матрицю, обернену до $X'X$. Ця матриця існує, якщо визначник матриці $X'X$ не дорівнює нулю: $det(X'X) \neq 0$. Тому спочатку розрахуємо визначник матриці $X'X$:

$$det(X'X)=93643,75.$$

Тоді:

$$(X'X)^{-1} = \begin{pmatrix} 0,67227 & -0,02897 & -0,21347 \\ -0,02897 & 0,00330 & 0,00441 \\ -0,21347 & 0,00441 & 0,08667 \end{pmatrix}.$$

Використовуючи формулу $\hat{A} = (X'X)^{-1} X'Y$ знаходимо статистичні оцінки параметрів моделі за методом найменших квадратів:

$$\widehat{A} = \begin{pmatrix} 0,67227 & -0,02897 & -0,21347 \\ -0,02897 & 0,00330 & 0,00441 \\ -0,21347 & 0,00441 & 0,08667 \end{pmatrix} \cdot \begin{pmatrix} 325,6 \\ 1400,4 \\ 762,95 \end{pmatrix} = \begin{pmatrix} 15,4608 \\ -1,4417 \\ 2,7914 \end{pmatrix}.$$

Таким чином, $\widehat{b}_0 = 15,4608; \widehat{b}_1 = -1,4417; \widehat{b}_2 = 2,7914$.

Отже, залежність ціни автомобіля від його віку та об'єму двигуна має вигляд:

$$\widehat{y} = 15,4608 - 1,4417x_1 + 2,7914x_2.$$

Аналіз ступеня адекватності побудованої моделі та вибіркових даних

Обчислимо теоретичні значення $\vec{Y}$ за формулою:

$$\vec{Y} = X\vec{A},$$

результати розрахунків наведено в таблиці 2.11.

Таблиця 2.11 – Допоміжні обчислення

№	$\widehat{y}_i$	$e_i = y_i - \widehat{y}_i$	e_i^2	$\widehat{y}_i - \bar{y}$	$(\widehat{y}_i - \bar{y})^2$	$y_i - \bar{y}$	$(y_i - \bar{y})^2$
1.	15,185	-4,185	17,511	2,161	4,668	-2,024	4,097
2.	4,672	-1,472	2,167	-8,352	69,752	-9,824	96,511
3.	6,068	2,632	6,928	-6,956	48,387	-4,324	18,697
4.	-2,583	4,183	17,493	-15,607	243,563	-11,424	130,508
5.	19,277	-2,277	5,184	6,253	39,097	3,976	15,809
6.	20,86	-1,96	3,84	7,836	61,396	5,876	34,527
7.	16,718	-0,918	0,844	3,694	13,649	2,776	7,706
8.	15,185	2,815	7,926	2,161	4,668	4,976	24,761
9.	22,301	-3,301	10,899	9,277	86,068	5,976	35,713
10.	11,277	-5,277	27,843	-1,747	3,053	-7,024	49,337
11.	14,068	-1,068	1,141	1,044	1,09	-0,024	0,001
12.	14,114	8,586	73,718	1,09	1,188	9,676	93,625
13.	16,718	-2,818	7,944	3,694	13,649	0,876	0,767
14.	15,323	-4,423	19,56	2,299	5,284	-2,124	4,511
15.	11,185	-1,685	2,838	-1,839	3,384	-3,524	12,419
16.	17,277	10,723	114,989	4,253	18,086	14,976	224,281
17.	8,951	-2,951	8,711	-4,073	16,586	-7,024	49,337
18.	6,951	-1,151	1,326	-6,073	36,877	-7,224	52,186
19.	4,068	-1,268	1,608	-8,956	80,212	-10,224	104,53
20.	20,16	5,84	34,103	7,136	50,925	12,976	168,377
21.	12,393	-1,893	3,584	-0,631	0,398	-2,524	6,371
22.	11,789	5,111	26,124	-1,235	1,526	3,876	15,023
23.	6,672	-3,272	10,708	-6,352	40,345	-9,624	92,621
24.	17,927	-3,927	15,422	4,903	24,04	0,976	0,953
25.	19,044	3,956	15,653	6,02	36,236	9,976	99,521
Сума	325,6	0	438,061		904,125		1342,186

Правильність виконаних розрахунків можна перевірити, порівнюючи середні значення $\bar{y}$ та $\bar{\widehat{y}}$, де:

$$\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n} = 13{,}024; \quad \bar{\bar{y}} = \frac{\sum_{i=1}^{n} \hat{y}_i}{n} = 13{,}024,$$

оскільки $\bar{y} = \bar{\bar{y}}$, то попередні розрахунки є вірними.

Для подальшого статистичного дослідження моделі використаємо результати табл.3.11.

Визначимо ступінь адекватності моделі та статистичних даних через відхилення між фактичними значеннями та значеннями, обчисленими за моделлю. Запишемо їх як елементи вектора $\vec{e}$:

$$e_i = y_i - \hat{y}_i \ .$$

Середнє значення $\bar{e} = 0$, отже розбіжностей не існує, модель адекватна.

Обчислимо коефіцієнт множинної детермінації:

$$R^2 = \frac{\sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} = 1 - \frac{\sum_{i=1}^{n} e_i^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2};$$

$$R^2 = \frac{904{,}1246}{1342{,}1856} = 0{,}6736.$$

Звідси, коефіцієнт множинної кореляції буде:

$$R = \sqrt{R^2} = \sqrt{0{,}6736} = 0{,}82 .$$

Отримано досить високий коефіцієнт множинної кореляції. Це свідчить про те, що зв’язок між ціною автомобіля, його віком та об’ємом двигуна досить тісний.

Обчислимо спостережене значення критерію Фішера за формулою:

$$F = \frac{R^2}{1 - R^2} \cdot \frac{n - m - 1}{m} = \frac{0{,}6736}{1 - 0{,}6736} \cdot \frac{22}{2} = 22{,}7032 .$$

$F_{табл}(\alpha = 0{,}05; k_1 = 2; k_2 = 22) = 3{,}44$.

Оскільки $F > F_{\alpha, k_1, k_2}$, то всі фактори мають вплив на залежну змінну.

Знайдемо тепер незміщену оцінку для дисперсії залишків σ_u^2:

$$\sigma_u^2 \approx 19{,}9119 .$$

Коваріаційна матриця оцінок параметрів дорівнює:

$$\mathrm{cov}(\hat{A} \cdot (\hat{A})') = \sigma_u^2 \cdot (X'X)^{-1};$$

$$\mathrm{cov}(\hat{A} \cdot (\hat{A})') = 19{,}9119 \cdot \begin{pmatrix} 0{,}67227 & -0{,}02897 & -0{,}21347 \\ -0{,}02897 & 0{,}00330 & 0{,}00441 \\ -0{,}21347 & 0{,}00441 & 0{,}08667 \end{pmatrix} = \begin{pmatrix} 13{,}3861 & -0{,}5768 & -4{,}2505 \\ -4{,}2505 & 0{,}0658 & 0{,}0878 \\ -4{,}2505 & 0{,}0878 & 1{,}7257 \end{pmatrix}$$

Отже, дисперсії оцінок параметрів можна записати як:

$$S_{\hat{b}_0}^2 = 13{,}3861; \quad S_{\hat{b}_1}^2 = 0{,}0658, \quad S_{\hat{b}_2}^2 = 1{,}7257 .$$

Середньоквадратичні відхилення оцінок параметрів дорівнюють:

$$S_{\hat{b}_0} = \sqrt{S^2_{\hat{b}_0}} = 3{,}6587;$$

$$S_{\hat{b}_1} = \sqrt{S^2_{\hat{b}_1}} = 0{,}2565;$$

$$S_{\hat{b}_2} = \sqrt{S^2_{\hat{b}_2}} = 1{,}3137.$$

Перевіримо статистичну значущість параметрів $\hat{b}_0$, $\hat{b}_1$ та $\hat{b}_2$. Для цього розрахуємо t –критерій Стьюдента:

$$t_{\hat{b}_0} = \frac{15{,}4608}{3{,}6587} = 4{,}2258;$$

$$t_{\hat{b}_1} = \frac{1{,}4417}{0{,}2565} = 5{,}6206;$$

$$t_{\hat{b}_2} = \frac{2{,}7914}{1{,}3137} = 2{,}1249.$$

Знайдемо табличне значення критерію для ступенів вільності $n - m = 25 - 3 - 1 = 22$ та рівня значущості $\alpha = 0{,}05$: $t_{табл} = 2{,}074$.

Оскільки $t > t_{табл}$, то відповідні оцінки параметрів моделі є значущими.

Перевіримо на значущість коефіцієнт кореляції. Для цього використаємо t-критерій Стьюдента:

$$t = \frac{R\sqrt{n-m}}{\sqrt{1-R^2}} = 6{,}732,$$

де R^2 коефіцієнт детермінації моделі, R -коефіцієнт кореляції ($R = \sqrt{R^2}$).

Табличне значення критерію для ступенів вільності $n - m = 25 - 3 - 1 = 22$ та рівня значущості $\alpha = 0{,}05$: $t_{табл} = 2{,}074$.

Оскільки $t > t_{табл}$, то робимо висновок про його статистичну значущість.

Довірчі інтервали для коефіцієнтів регресії b_j :

$$15{,}4608 - 3{,}6587 \cdot 2{,}074 < b_0 < 15{,}4608 + 3{,}6587 \cdot 2{,}074;$$

$$-1{,}4417 - 0{,}2565 \cdot 2{,}074 < b_1 < -1{,}4417 + 0{,}2565 \cdot 2{,}074;$$

$$2{,}7914 - 1{,}3137 \cdot 2{,}074 < b_2 < 2{,}7914 + 1{,}3137 \cdot 2{,}074;$$

або

$$7{,}8726 < b_0 < 23{,}0489;$$

$$-1{,}9737 < b_1 < -0{,}9097;$$

$$0{,}0669 < b_2 < 5{,}5160.$$

На основі побудованої залежності можна зробити висновок, що ціна автомобіля, який міг би мати технічні характеристики, рівні середнім значенням $\bar{x}_1$ та $\bar{x}_2$, дорівнювала б завдяки $b_0 = 15{,}4608$ середньому значенню залежної змінної – $\bar{y}$ =13,024 тис.дол. Збільшення віку автомобіля на один рік зменшує ціну на 1,4417 тис.дол., а збільшення об’єму двигуна на 1дм^3 призводить до збільшення ціни на 2,7914 тис.дол. Фактори, включені в модель, пояснюють “поведінку” ціни на 67,4%. Звичайно, це можна пояснити тим, що на формування ціни на автомобіль мають вплив також інші фактори (престиж марки, тип двигуна, оснащення салону, колір тощо), які в даному прикладі не розглядались.

Визначимо часткові коефіцієнти еластичності за формулою:

$$K_{E_i} = b_j \frac{\bar{x}_j}{\bar{y}}$$

Одержимо:

$$K_{E_1} \approx -0{,}084,$$

$$K_{E_2} \approx 0{,}16.$$

Отже, $K_{E_1} = -0{,}84$ інформує про те, що при збільшенні першого регресора (вік автомобіля) на *k* відсотків, значення залежної змінної моделі *Y* (ціна автомобіля) зменшиться на $8{,}4 \cdot k$ відсотків, а при збільшенні другого регресора (об’єм двигуна автомобіля) на *k* відсотків, *Y* збільшиться на $16 \cdot k$ відсотків.

Завдання для індивідуальної та самостійної роботи студентів

Приклад. Необхідно провести дослідження залежності ціни автомобіля (*Y*) від таких характеристик як вік авто (X_1) та його об’єм двигуна (X_2) на основі вибіркових даних, наведених в табл.

Необхідно:

1) визначити $\widehat{b_i}$ для залежності між досліджуваним фактором *y* (ціною автомобіля) та пояснюючими змінними x_1 (вік автомобіля) і x_2 (об’єм двигуна);
2) проаналізувати ступінь адекватності побудованої моделі та вибіркових даних.

Дані про роботу підприємств наведені в таблиці 2.12.

Таблиця 2.12 - Показники роботи 10 однотипних підприємств

1	Y	13,2	12,2	13,8	13,9	15,6	15,6	16	17,3	18,1	19,5
	x_1	181	171	177	179	190	186	185	189	198	208
	x_2	53	76	80	65	66	65	94	94	76	66
2	Y	26,5	30,3	28,9	32,8	34,6	33,9	38,4	41,6	42,9	43,8
	x_1	282	282	274	285	300	286	291	305	307	312
	x_2	97	93	86	86	98	87	81	83	84	86
3	Y	28,4	30	32	31,1	30,4	32,6	33,1	33,7	36,5	38,6
	x_1	258	267	272	269	263	269	270	271	280	284
	x_2	77	87	103	107	78	77	75	75	77	69
4	Y	22,1	24,2	25,4	28,1	27,7	29,9	32,6	33,1	33,7	34,2
	x_1	117	129	134	136	132	141	146	139	140	150
	x_2	84	97	99	93	79	93	90	72	77	88
5	Y	13,1	14,8	17,7	21,9	22,3	26,1	27,2	27,3	29,4	31,8
	x_1	239	235	252	252	244	266	265	263	266	275
	x_2	85	73	92	72	50	83	88	67	69	74

Продовження таблиці 2.12

6	Y	20,5	22	24	23,9	24,6	27,4	29,1	29,1	31,1	32,8
	x_1	131	129	133	135	145	151	157	143	156	165
	x_2	91	78	79	76	96	93	92	66	73	77
7	Y	13,2	15,4	16,4	17	18,2	20,6	22,9	22,8	23,4	26,1
	x_1	244	247	266	261	279	291	282	292	291	310
	x_2	82	66	85	75	85	78	63	69	70	69
8	Y	13,8	15	13	18	19,5	18,8	19,6	20,6	23,1	24
	x_1	202	209	196	210	222	220	219	217	228	240
	x_2	80	83	64	60	77	67	68	54	55	92
9	Y	17,4	20,8	23,1	22,4	24,7	25,9	27,5	28,4	31,7	32
	x_1	256	266	272	263	273	274	279	287	292	285
	x_2	84	75	77	90	81	83	78	74	82	102
10	Y	22,6	22,4	22,6	23	25,3	26,3	26,6	27,7	29,8	30,9
	x_1	227	232	234	239	251	250	259	255	273	268
	x_2	89	67	72	61	57	71	39	74	60	74
11	Y	26,8	25,7	28,4	28,2	28,5	30,1	29,9	32,1	32,3	34,3
	x_1	189	187	189	186	184	200	192	205	208	217
	x_2	99	98	84	80	66	90	78	79	85	78
12	Y	11	12,4	12,7	13,8	15,2	17,7	18	18,9	20,4	22
	x_1	291	294	292	293	296	308	310	311	318	312
	x_2	78	74	61	60	59	68	74	65	62	43
13	Y	18	18,8	20	20,4	23,2	25,7	27	27,8	31,6	32,4
	x_1	161	171	167	161	168	177	185	176	186	185
	x_2	78	86	85	64	67	75	85	65	65	63
14	Y	20,6	20,4	23,3	24,6	26,7	28,9	30,1	32,6	33,5	34,2
	x_1	196	200	201	215	221	226	217	240	245	242
	x_2	106	98	81	110	99	87	65	83	90	66
15	Y	26,6	27,1	28,1	29,4	30,4	32,1	31,7	32,9	32,8	34,3
	x_1	240	237	258	260	263	267	263	272	274	277
	x_2	81	74	103	95	89	79	73	71	70	69
16	Y	21,9	22	22,2	24,1	25,6	26,8	27,9	29,5	29,2	30,7
	x_1	196	187	187	194	194	202	204	208	211	208
	x_2	96	70	74	71	70	72	76	73	86	61
17	Y	12,3	15,7	17,3	17	19,3	18,9	20,8	24,2	23,4	22,7
	x_1	206	213	220	217	232	230	245	250	239	256
	x_2	93	82	76	74	83	91	87	70	64	90
18	Y	11,6	13,7	14,7	16,7	17	18,9	21,9	23	22,8	23,9
	x_1	189	196	197	202	202	201	213	222	208	211
	x_2	95	97	96	100	95	82	89	106	83	75
19	Y	25,5	26,8	28,5	29,9	30,1	31,8	33,3	35,3	38,2	38,4
	x_1	245	246	247	258	251	257	279	267	290	290
	x_2	87	89	80	87	76	76	98	66	84	76
20	Y	13,9	17	15	16,7	19,7	20,6	20,8	20,1	20,3	21,4
	x_1	226	244	231	232	247	243	255	247	257	260
	x_2	108	103	103	84	78	64	86	88	92	89
21	Y	18,9	23,2	21,3	23,5	24,9	24,6	26,8	27,7	30,4	29,3
	x_1	245	255	242	262	269	275	271	266	291	296
	x_2	108	85	83	87	86	94	69	67	74	93

Продовження таблиці 2.12

22	Y	23,5	26,6	26,6	30,5	30,9	33,2	35,2	35,3	38,2	42,1
	x_1	92	97	108	111	102	117	120	113	130	141
	x_2	72	66	84	74	66	74	72	65	75	71
23	Y	21,5	23,7	23,2	24,6	27,5	25,4	27,6	28,9	28,5	30,8
	x_1	125	136	128	151	152	135	157	158	173	173
	x_2	99	86	83	100	83	77	82	79	95	74
24	Y	13,2	15,9	19,1	17,6	18,8	21,1	21,9	23,6	23	23,2
	x_1	194	210	231	231	232	237	235	254	252	242
	x_2	91	98	89	98	93	85	74	90	89	73
25	Y	13,4	16,7	18,5	19,7	21,1	21,9	24,2	26,7	27,3	31,3
	x_1	256	271	269	269	279	272	279	278	288	293
	x_2	85	119	96	95	111	87	92	89	106	100
26	Y	23,4	25,2	25,2	25,3	29	28	28,9	28,1	31,9	29,3
	x_1	134	147	144	143	161	153	166	161	164	161
	x_2	90	105	102	94	99	95	120	112	89	96
27	Y	17,4	19,1	18,9	19,7	20,4	21,9	21	25,6	26,8	26,2
	x_1	292	303	303	307	296	321	309	337	333	340
	x_2	96	90	80	83	49	85	68	65	48	67
28	Y	14,2	15,8	17,6	21,9	19,7	21,4	26,2	24,6	27,9	31,2
	x_1	207	221	219	238	232	227	245	234	240	264
	x_2	100	105	94	100	105	83	94	81	67	91
29	Y	18,3	19,5	19,7	21,1	19	20,2	21	21,4	21,8	22,4
	x_1	238	237	243	252	239	250	257	258	264	268
	x_2	84	107	95	111	87	86	75	87	81	84
30	Y	26,6	27,1	28,1	29,4	30,4	32,1	31,7	32,9	32,8	34,3
	x_1	240	237	258	260	263	267	263	272	274	277
	x_2	81	74	103	95	89	79	73	71	70	69

Дані необхідно вибрати з таблиці відповідно до номера варіанта.
Номер варіант визначається за вказівкою викладача.

2.4 Мультиколінеарність

Відведений час: 4 год.

Мета: навчати досліджувати сукупність на наявність мультиколінеарності та формувати первинні навчики усувати її.

Завдання для практичного заняття:
1. Пригадайте основні теоретичні питання теми.
2. Орієнтовні запитання та завдання:
   - що означає мультиколінеарність факторів ?
   - як впливає наявність мультнколінеарності на оцінку параметрів моделі ?
   - які статистичні критерії використовуються для виявлення мультиколінеарності ?
   - які кроки треба здійснти за алгоритмом Фаррара-Глобера ?

$-$які існують способи усунення мультиколінеарності ?

3. Виконайте індивідуальне завдання.

Короткі теоретичні відомості

Мультиколінеарністю називається висока залежність (корельованість) різних факторів.

Якщо між факторами є тісний зв'язок, то визначник матриці $(X'X)$ дуже малий. Це призводить до великих помилок при оцінці параметрів рівняння регресії. В остаточному підсумку виходить незначуще рівняння регресії, що не має реального змісту.

Основні наслідки мультиколінеарності:

- падає точність оцінок параметрів, яка виявляється в зростанні помилок деяких оцінок, в значному збільшені дисперсії оцінок параметрів;
- оцінки деяких параметрів стають незначущими;
- оцінки деяких параметрів стають чутливими до обсягів сукупності спостережень.

Тому при побудові економетричної моделі потрібно визначити існування мультиколінеарності та усунути її.

При наявності мультиколінеарності факторів доцільно звернути увагу і на специфікацію моделі. Іноді заміна однієї функції іншою, не суперечить інформації, дозволяє усунути мультиколінеарність.

У випадку економетричних моделей з великого кількісно факторів, коли не вдається позбутися мультиколінеарності, параметри моделі оцінюють методом головних компонент.

Ознаки мультиколінеарності

1. Якщо серед парних коефіцієнтів кореляції незалежних змінних є такі, рівень яких наближається або дорівнює множинному коефіцієнту кореляції, то це свідчить про можливість існування мультиколінеарності. Інформацію про парну залежність може дати симетрична матриця коефіцієнтів парної кореляції, або кореляції нульового порядку:

$$r=\begin{pmatrix} r_{X_1X_1} & r_{X_1X_2} & r_{X_1X_3} & \cdots & r_{X_1X_K} \\ r_{X_2X_1} & r_{X_2X_2} & r_{X_2X_3} & \cdots & r_{X_2X_K} \\ r_{X_3X_1} & r_{X_3X_2} & r_{X_3X_3} & \cdots & r_{X_3X_K} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ r_{X_KX_1} & r_{X_KX_2} & r_{X_KX_3} & \cdots & r_{X_KX_K} \end{pmatrix}.$$

Але, якщо в моделі фігурує більше двох незалежних змінних, вивчення питання про мультиколінеарність не може обмежуватись інформацією, що дає ця матриця. Явище мультиколінеарності ні в якому разі не зводиться тільки до існування парної кореляції між незалежними змінними.

Більш загальна перевірка передбачає визначення визначника (детермінанта) матриці r, який називається детермінантом кореляції і позначається $|r|$. Числові значення детермінанта кореляції знаходяться на множині: $|r| \in [0,1]$.

2. Якщо $|r| = 0$, то існує повна мультиколінеарність, якщо $|r| = 1$- мультиколінеарність відсутня, чим ближче $|r|$ до нуля, тим певніше можна стверджувати, що між незалежними змінними існує мультиколінеарність. Незважаючи на те, що на числове значення $|r|$ впливає дисперсія незалежних змінних, цей показник можна вважати точковою мірою тісноти мультиколінеарності.

3. Якщо в економетричній моделі одержано мале значення параметра $\hat{a}_k$ при високому рівні коефіцієнта детермінації і при цьому F-критерій суттєво відрізняється від нуля, то це також свідчить про наявність мультиколінеарності.

4. Якщо коефіцієнт детермінації R^2, що розрахований для регресійних залежностей між однією незалежною змінною та іншими, має значення, яке близьке до одиниці, то можна говорити про наявність мультиколінеарності.

5. Якщо при побудові економетричної моделі на основі покрокової регресії включення нової незалежної змінної суттєво змінює оцінку параметрів моделі при незначному підвищенні (або зниженні) коефіцієнтів кореляції чи детермінації, то ця змінна, очевидно, знаходиться в лінійній залежності від інших, які введені в модель раніше.

Всі ці методи виявлення мультиколінеарності мають один загальний недолік: жоден із них не проводить чіткої межі між тим, що треба вважати «суттєвою» мультиколінеарністю, яку треба враховувати, і тим, коли мультиколінеарністю можна знехтувати.

Для виявлення мультиколінеарності можна використати критерій (алгоритм) Феррара-Глобера. Цей алгоритм має три види статистичних критеріїв, згідно з якими перевіряється мультиколінеарність всього масиву незалежних змінних (χ^2 - «хі» - квадрат); кожної незалежної змінної з рештою змінних (F-критерій); кожної пари незалежних змінних (t-критерій).

Алгоритм включає такі кроки:

1. *Стандартизація (нормалізація) змінних.*

Позначимо вектори незалежних змінних економетричної моделі через $X_1, X_2, X_3, \ldots, X_m$. Елементи стандартизованих векторів розрахуємо за формулою:

$$X_{ik}^{*} = \frac{X_{ik} - \overline{X}_k}{\sqrt{n\sigma_{X_k}^2}},$$

де n - число спостережень, $(i = \overline{1,n})$;

m - число незалежних змінних, $(k = \overline{1,m})$;

$\overline{X}_k$ - середня арифметична k - ії незалежної змінної;

$\sigma^2_{X_k}$ - дисперсія k-ї незалежної змінної.

2.*Визначення кореляційної матриці* (елементами цієї матриці є коефіцієнти парної лінійної кореляції між незалежними змінними):

$$R = X^{*\prime}\, X^{*},$$

де X^* - матриця стандартизованих незалежних змінних;

$X^{*\prime}$ - матриця, транспонована до матриці X^*.

3. *Визначення критерію χ^2 (хі-квадрат):*

$$\chi^2 = -\left[n - 1 - \frac{1}{6}(2m+5)\right] \ln|R| \;,$$

де $|R|$ - визначник кореляційної матриці R.

Значення цього критерію порівнюється з табличним при $\frac{1}{2}m\,(m-1)$ ступенях свободи і рівні значущості α. Якщо $\chi^2_{\text{факт}} < \chi^2_{\text{табл}}$, то в масиві незалежних змінних не існує мультиколінеарності.

4. *Визначення оберненої матриці*:

$$C = R^{-1}\,.$$

5. *Обчислення F-критеріїв:*

$$F_k = (c_{kk} - 1)\,\frac{n-m}{m-1}\;,$$

де c_{kk} — діагональні елементи матриці C. Фактичні значення критеріїв F_k порівнюються з табличними при $n-m$ і $m-1$ ступенях свободи і рівні значущості α. Якщо $F_{k\ \text{факт}} > F_{\text{табл}}$, відповідна k-та незалежна змінна мультиколінеарна з іншими.

6. *Обчислення часткових коефіцієнтів кореляції:*

$$r_{jk} = \frac{-c_{jk}}{\sqrt{c_{jj}\cdot c_{kk}}}\,,$$

де c_{jk} – елемент матриці С, який розміщений j-ому рядку і k-ому стовпці; c_{jj} та c_{kk} - діагональні елементи матриці С.

7. *Обчислення t-критеріїв*:

$$t_{kj} = \frac{r_{kj}\sqrt{n-m}}{\sqrt{1-r_{kj}^2}}\,.$$

Фактичні значення критеріїв t_{kj} порівнюються з табличними при $n-m$ ступенях свободи і рівні значущості α. Якщо $t_{kj\ \text{факт}} > t_{\text{табл}}$, між незалежними змінними X_k і X_j існує мультиколінеарність.

Приклад розв'язування задачі

Приклад. Дані про роботу фірми представлені у таблиці 2.13: прибуток фірми (y, у.г.о.) та фактори, від яки він залежить- фондовіддача (x_1, у.г.о.), продуктивність праці (x_2, у.г.о.), питомі інвестиції (x_3, у.г.о.).

Таблиця 2.13 - Показники роботи фірми

№ місяця	Прибуток фірми y, у.г.о.	Фондовіддача x_1, у.г.о.	Продуктивність праці x_2, у.г.о.	Питомі інвестиції x_3, у.г.о.
1	40	12	5	15
2	45	17	7	18
3	40	13	6	16
4	43	14	7	17
5	48	16	6	20
6	39	15	5	15
7	42	14	6	16
8	45	17	9	18
9	38	12	5	19
10	48	18	10	20
11	50	20	11	22
12	48	17	10	21
13	49	18	12	21
14	45	19	8	20
15	49	20	9	22
16	52	22	14	23
17	54	24	15	24
18	51	21	13	20
19	55	25	16	24
20	56	27	18	25

Необхідно: 1) перевірити масив незалежних змінних на наявність мультиколінеарності за алгоритмом Феррара-Глобера;

2) при наявності мультиколінеарності змінних знайти істинні оцінки параметрів моделі;

3) побудувати модель, позбавлену від колінеарності.

*Розв'язання***.** Перевіримо масив даних на мультиколінеарність за алгоритмом Феррара-Глобера.

1. Нормалізуємо незалежні змінні (дивись таблиця 2.14)

Таблиця 2.14 - Нормалізовані незалежні змінні

	X_1^*	X_2^*	X_3^*
	-1,41989	-1,15996	-1,58071
	-0,24643	-0,65563	-0,59276
	-1,1852	-0,9078	-1,25139
$\mathbf{X}^*=$	-0,9505	-0,65563	-0,92208
	-0,48112	-0,9078	0,065863
	-0,71581	-1,15996	-1,58071
	-0,9505	-0,9078	-1,25139
	-0,24643	-0,1513	-0,59276

Продовження таблиці 2.14

	-1,41989	-1,15996	-0,26345
	-0,01173	0,100866	0,065863
	0,45765	0,353032	0,72449
	-0,24643	0,100866	0,395176
	-0,01173	0,605198	0,395176
	0,222958	-0,40347	0,065863
	0,45765	-0,1513	0,72449
	0,927034	1,10953	1,053804
	1,396419	1,361696	1,383117
	0,692342	0,857364	0,065863
	1,631111	1,613862	1,383117
	2,100496	2,118194	1,712431

2. Визначимо кореляційну матрицю

Транспонуємо нормалізовану матрицю X^*

$$\mathbf{X}^{*T} = \left|\begin{array}{cccccccccc} -1,4198 & -0,2464 & -1,185 & -0,951 & -0,4811 & -0,716 & -0,951 & -0,246 & -1,4199 & -0,0117 \\ -1,1599 & -0,6556 & -0,908 & -0,656 & -0,9078 & -1,16 & -0,908 & -0,151 & -1,16 & 0,1009 \\ -1,5807 & -0,5928 & -1,251 & -0,922 & 0,0659 & -1,581 & -1,251 & -0,593 & -0,2635 & 0,0659 \end{array}\right.$$

Продовження матриці X^{*T}

$$\left.\begin{array}{cccccccccc} 0,4576 & -0,2464 & -0,012 & 0,223 & 0,458 & 0,927 & 1,3964 & 0,692 & 1,631 & 2,1 \\ 0,353 & 0,1009 & 0,6052 & -0,403 & -0,151 & 1,1095 & 1,3617 & 0,857 & 1,614 & 2,118 \\ 0,7245 & 0,3952 & 0,3952 & 0,0659 & 0,724 & 1,0538 & 1,3831 & 0,066 & 1,383 & 1,712 \end{array}\right|$$

Знайдемо добуток матриць X^{*T} та X^*

$$X^{*T} \times X^* = \begin{vmatrix} 19 & 17,896 & 16,941 \\ 17,896 & 19 & 16,642 \\ 16,941 & 16,642 & 19 \end{vmatrix}$$

Запишемо кореляційну матрицю R.

Для цього необхідно кожний елемент матриці $X^{*T} \times X^*$ домножити на $\frac{1}{n-1}$, де n - кількість спостережень (у нашому випадку n=20)

$$R = \begin{vmatrix} 1 & 0,942 & 0,892 \\ 0,942 & 1 & 0,876 \\ 0,892 & 0,876 & 1 \end{vmatrix}$$

3. Визначимо критерій χ^2 (хі-квадрат). Для цього розрахуємо визначник кореляційної матриці: $\det(R) = 0,0218$

Логарифм визначника: $\ln\det(r) = -3,825$.

Оскільки визначник кореляційної матриці наближається до 0, то в масиві пояснюючих змінних може існувати мультиколінеарність.

Перевіримо весь масив змінних на наявність колінеарності за критерієм χ^2.

Критерій розраховується за формулою:

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

де m – кількість незалежних змінних.

Обчислимо:

$$\chi^2 = -\left\{20 - 1 - \frac{2\times 3+5}{6}\right\}\times(-3.825) = 65.663$$

Розраховане значення χ_p^2 порівняємо з табличним для $\nu = \frac{m(m-1)}{2} = \frac{3(3-1)}{2} = 3$ ступенів свободи та рівня значущості $\alpha = 0.05$: $\chi^2_{табл} = 7,81$

Оскільки $\chi_p^2 > \chi^2_{табл}$ то в масиві незалежних змінних (фондовіддача, продуктивність праці, питомі інвестиції) існує мультиколінеарність.

4. Знайдемо обернену кореляційну матрицю.

$$C = \begin{vmatrix} 10,671 & -7,376 & -3,055 \\ -7,376 & 9,393 & -1650 \\ -3,055 & -1,650 & 5,169 \end{vmatrix} \quad \text{структура матриці } C = \begin{vmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{2} & C_{33} \end{vmatrix}$$

5. Перевіримо на колінеарність окремі змінні з іншими за F-критеріями.

Критерій розраховується за формулою:

$$F = (C_{ii} - 1)\left(\frac{n-m}{m-1}\right),$$

де C_{ii} - діагональний елемент матриці С (матриці , оберненої до кореляційної матриці R).

Обчислимо F-критерій для фактору x_1 (фондовіддача):

$$F_1 = (10,671 - 1)\left(\frac{20-3}{3-1}\right) = 82,205$$

Обчислимо F-критерій для фактору x_2 (продуктивність праці):

$$F_2 = (9,393 - 1)\left(\frac{20-3}{3-1}\right) = 71,34$$

Обчислимо F-критерій для фактору x_3 (питомі інвестиції):

$$F_3 = (5,169 - 1)\left(\frac{20-3}{3-1}\right) = 35,436$$

Розраховане значення F-критерію порівняємо з табличним для $\nu_1 = n - m = 20 - 3 = 17$ та $\nu_2 = m - 1 = 3 - 1 = 2$ ступенів свободи та рівня значущості $\alpha = 0.05$: $F_{табл} = 3,59$

Оскільки $F_1 > F_{табл}$, $F_2 > F_{табл}$, $F_3 > F_{табл}$, то кожна із незалежних змінних (фондовіддача, продуктивність праці, питомі інвестиції) колінеарна з іншими.

6. Визначимо частинні коефіцієнти кореляції $r_{ij,k}$

Коефіцієнти розраховуються за формулою:

$$r_{ij,k} = \frac{-C_{ij}}{\sqrt{C_{ii}C_{jj}}}$$

Обчислимо коефіцієнт кореляції $r_{ij,k}$ для факторів x_1 (фондовіддача) та

x_2 (продуктивність праці):

$$r_{12,3} = \frac{-(-7,376)}{\sqrt{10,671 \times 9,393}} = 0,737$$

Обчислимо коефіцієнт кореляції $r_{ij,k}$ для факторів x_1 (фондовіддача) та x_3 (питомі інвестиції):

$$r_{13,2} = \frac{-(-3,055)}{\sqrt{10,671 \times 5,169}} = 0,411$$

Обчислимо коефіцієнт кореляції $r_{ij,k}$ для факторів x_2 (продуктивність праці) та x_3 (питомі інвестиції):

$$r_{23,1} = \frac{-(-1,650)}{\sqrt{9,393 \times 5,169}} = 0,237$$

На основі частинних коефіцієнтів кореляції можна стверджувати:

- між фондовіддачею та продуктивністю праці існує тісний зв'язок, якщо не враховувати вплив питомих інвестицій;
- між фондовіддачею та питомими інвестиціями існує помірний зв'язок, якщо не враховувати вплив продуктивністі праці;
- між продуктивністю праці та питомими інвестиціями існує слабкий зв'язок, якщо не враховувати вплив фондовіддачі;

7. Перевіримо на колінеарність кожну пару змінних за допомогою *t*-критерію

Критерії розраховуються за формулою:

$$t_{ij} = \frac{r_{ij,k}\sqrt{n-m}}{\sqrt{1-r_{ij,k}^2}}$$

Обчислимо $\underline{t}$ -критерій для фактору x_1 (фондовіддача):

$$t_{12} = \frac{0,737 \times \sqrt{20-3}}{\sqrt{1-0,737^2}} = 4,496$$

Обчислимо $\underline{t}$ -критерій для фактору x_2 (продуктивність праці):

$$t_{13} = \frac{0,411 \times \sqrt{20-3}}{\sqrt{1-0,411^2}} = 1,859$$

Обчислимо $\underline{t}$ -критерій для фактору x_3 (питомі інвестиції):

$$t_{23} = \frac{0,237 \times \sqrt{20-3}}{\sqrt{1-0,237^2}} = 1,006$$

Розраховане значення *t*-критерію порівняємо з табличним для $\nu = n - m = 20 - 3 = 17$ ступенів свободи та рівня значущості $\alpha = 0.05$: $t_{табл} = 1.740$

Оскільки $t_{12} > t_{табл}$, то фондовіддача та продуктивність праці колінеарні між собою; $t_{13} > t_{табл}$, то фондовіддача та питомі інвестиції колінеарні між собою; $t_{23} < t_{табл}$,то продуктивність праці та питомі інвестиції не колінеарні між собою.

Отже, розрахунки показали, що в масиві незалежних змінних існує колінеарність між факторами x_1 та x_2 а також між x_1 та x_3 Тому для визначення незміщених, ефективних та обґрунтованих оцінок параметрів (істинних оцінок) необхідно позбутися наслідків колінеарності.

Визначимо істинні оцінки параметрів моделі $Y = f(X_1 X_2 X_3)$

Нормалізуємо фактор y - прибуток фірми (дивись таблиця 2.15).

Таблиця 2.15- Нормалізація залежної змінної

№ з/п	Прибуток фірми Y, у.г.о.	Y^*
1.	40	-1,273
2.	45	-0,344
3.	40	-1,273
4.	43	-0,715
5.	48	0,2136
6.	39	-1,458
7.	42	-0,901
8.	45	-0,344
9.	38	-1,644
10.	48	0,2136
11.	50	0,5852
12.	48	0,2136
13.	49	0,3994
14.	45	-0,344
15.	49	0,3994
16.	52	0,9567
17.	54	1,3283
18.	51	0,771
19.	55	1,514
20.	56	1,6998

Тоді вхідні дані для побудови «нормалізованої моделі» можна записати у вигляді таблиці 2.16.

Таблиця 2.16 – Вихідні дані

№ місяця	*Y**	*X₁**	*X₂**	*X₃**
1	-1,273	-1,41989	-1,15996	-1,58071
2	-0,344	-0,24643	-0,65563	-0,59276
3	-1,273	-1,1852	-0,9078	-1,25139
4	-0,715	-0,9505	-0,65563	-0,92208
5	0,2136	-0,48112	-0,9078	0,065863
6	-1,458	-0,71581	-1,15996	-1,58071
7	-0,901	-0,9505	-0,9078	-1,25139
8	-0,344	-0,24643	-0,1513	-0,59276
9	-1,644	-1,41989	-1,15996	-0,26345

Продовження таблиці 2.16

10	0,2136	-0,01173	0,100866	0,065863
11	0,5852	0,45765	0,353032	0,72449
12	0,2136	-0,24643	0,100866	0,395176
13	0,3994	-0,01173	0,605198	0,395176
14	-0,344	0,222958	-0,40347	0,065863
15	0,3994	0,45765	-0,1513	0,72449
16	0,9567	0,927034	1,10953	1,053804
17	1,3283	1,396419	1,361696	1,383117
18	0,771	0,692342	0,857364	0,065863
19	1,514	1,631111	1,613862	1,383117
20	1,6998	2,100496	2,118194	1,712431

Визначимо «нормалізовані» оцінки параметрів економетричної моделі за допомогою методу найменших квадратів. Отримаємо:

$b_1^* = 0{,}382$, $b_2^* = 0{,}382$, $b_3^* = 0{,}382$,

Визначимо істинні оцінки параметрів моделі $Y = f(X_1 X_2 X_3)$

Істинні оцінки параметрів моделі визначаються за формулами:

$$b_1 = b_1^* \times \left(\frac{\sigma_Y}{\sigma_{X_1}}\right),\; b_2 = b_2^* \times \left(\frac{\sigma_Y}{\sigma_{X_2}}\right),\; b_3 = b_3^* \times \left(\frac{\sigma_Y}{\sigma_{X_3}}\right),$$

$$b_0 = \overline{Y} - b_1\overline{X}_1 - b_2\overline{X}_2 - b_3\overline{X}_3$$

Тоді,

$$b_1 = 0{,}382 \times \left(\frac{5{,}38}{4{,}26}\right) = 0{,}482 \qquad b_2 = 0{,}302 \times \left(\frac{5{,}38}{3{,}97}\right) = 0{,}409 \qquad b_3 = 0{,}310 \times \left(\frac{5{,}38}{3{,}04}\right) = 0{,}549$$

$$b_0 = 46{,}85 - 0{,}482 \times 18{,}05 - 0{,}409 \times 9{,}60 - 0{,}549 \times 19{,}80 = 23{,}353$$

Отже, модель прибутку без колінеарності має вигляд:

$$y = 23{,}353 + 0{,}482x_1 + 0{,}409x_2 + 0{,}549x_3$$

Завдання для індивідуальної та самостійної роботи студентів

Приклад. Нехай на витрати обігу впливають: обсяг вантажообороту (x_1), запаси вантажообороту (x_2) та трудомісткість його одиниці (x_3). Щоб побудувати економетричну модель цієї залежності на основі методу 1МНК, необхідно бути впевненим, що між факторами вантажообороту, запасів та трудомісткості не існує мультиколінеарності. Треба дослідити наявність мультиколінеарності між цими факторами на основі даних, що наведені в таблиці 2.17.

Таблиця 2.17 - Показники роботи 10 однотипних підприємств

1	x_1	14,3	17,2	17,0	17,8	16,3	17,3	16,9	14,8	19,6	11,4
	x_2	49,8	45,2	41,4	41,7	38,7	40,6	33,6	13,9	32,5	52,5
	x_3	2,68	2,21	2,24	2,15	2,41	2,25	2,13	2,56	1,9	3,0
2	x_1	17,8	16,3	17,2	16,8	14,9	19,6	11,5	17,2	19,5	12,5
	x_2	41,8	38,7	40,5	34,5	13,9	31,5	52,5	37,0	43,6	48,3
	x_3	2,17	2,4	2,24	2,13	2,55	1,9	3,0	2,22	1,96	2,82

Продовження таблиці 2.17

3	x_1	16,9	14,5	19,5	11,5	17,1	19,6	12,5	16,5	16,0	16,1
	x_2	33,7	13,8	31,5	52,5	37,0	43,6	48,3	44,7	35,7	49,3
	x_3	2,13	2,56	1,91	3,0	2,23	1,96	2,82	2,29	2,33	2,31
4	x_1	19,5	12,4	17,2	19,5	12,6	16,5	16,1	16,0	16,2	18,0
	x_2	43,6	48,3	44,7	35,7	49,2	41,3	43,8	48,5	42,3	43,0
	x_3	1,95	2,83	2,29	2,33	2,31	2,54	2,11	2,42	2,38	2,46
5	x_1	15,6	13,5	15,3	14,9	15,1	16,1	16,7	15,4	17,1	16,8
	x_2	40,4	38,9	36,6	41,4	32,2	31,4	32,6	38,7	44,3	39,3
	x_3	2,11	2,78	2,17	2,15	2,11	1,97	1,96	2,12	2,02	2,13
6	x_1	15,1	16,1	16,7	15,5	17,2	16,9	17,0	16,2	15,0	18,0
	x_2	32	32,4	32,6	38,7	44,3	39,3	40,4	41,5	45,2	50,2
	x_3	2,12	1,98	1,96	2,15	2,02	2,05	2,02	2,13	2,14	1,9
7	x_1	16,9	16,2	15,5	18,2	17,3	17,1	16,4	16,7	14,2	17,2
	x_2	39,2	41	41,3	45,2	50,2	51,6	48,0	48,6	49,8	45,0
	x_3	2,02	2,13	2,14	1,89	2,48	1,94	1,93	1,96	2,57	2,21
8	x_1	17,2	17,1	16,5	16,8	14,5	17,2	17,1	17,9	16,2	17,3
	x_2	51,8	50,4	48,0	48,6	49,8	45,0	40,4	41,7	38,8	40,6
	x_3	2,48	1,95	1,92	1,96	2,68	2,22	2,23	2,15	2,41	2,25
9	x_1	11,7	18,3	18,2	15,6	17,4	13,8	15,0	18,6	16,2	15,7
	x_2	44,2	42,4	38,9	36,4	35,2	45,2	35,2	41,6	42,2	40,4
	x_3	2,85	1,81	1,88	2,25	1,98	2,48	2,25	1,85	2,15	2,25
10	x_1	18,4	15,6	17,5	13,8	15,0	18,7	16,2	15,7	17,9	15,4
	x_2	38,7	36,5	35,2	45,5	35,2	41,6	42,2	40,4	47,3	47,1
	x_3	1,89	2,25	1,98	2,48	2,25	1,85	2,15	2,25	1,9	2,2
11	x_1	13,8	15,0	18,6	16,2	15,7	17,9	15,3	16,3	17,7	16,8
	x_2	45,5	35,2	41,6	42,3	40,4	47,4	47,1	43,2	39,1	38,2
	x_3	2,46	2,25	1,85	2,15	2,25	1,9	2,2	2,09	1,87	2,0
12	x_1	18,8	16,4	16,1	17,8	15,5	16,3	17,8	16,8	17,5	16,7
	x_2	42,1	42,3	40,4	47,4	47,1	43,2	39,0	38,2	37,3	36,7
	x_3	1,86	2,15	2,25	1,9	2,2	2,1	1,86	2,0	2,48	2,02
13	x_1	13,8	14,8	16,9	16,8	14,8	18,0	17,5	15,7	15,3	14,9
	x_2	48,0	46,4	42,3	39,4	42,3	40,1	39,4	39,1	43,2	44,7
	x_3	2,45	2,3	2,0	2,05	2,23	1,89	2,0	2,21	2,01	2,31
14	x_1	16,8	17,0	14,8	17,9	16,9	19,7	14,0	17,1	18,2	17,4
	x_2	40,4	42,3	40,1	39,4	39,1	43,2	44,5	45,7	37,8	46,4
	x_3	1,88	2,0	2,21	2,01	2,31	2,16	1,74	2,25	1,87	1,82
15	x_1	19,7	14,0	17,2	18,3	17,4	16,1	18,8	17,9	17,6	15,7
	x_2	37,8	46,4	48,2	49,6	46,4	42,6	49,4	40,1	39,4	39,1
	x_3	1,74	2,25	1,88	1,82	1,9	1,98	1,77	1,89	2,0	2,21

Продовження таблиці 2.17

16	x_1	32,1	31,0	32,4	33,4	31,2	34,8	35,4	33,0	34,8	33,3
	x_2	48,0	42,1	42,3	43,7	42,8	41,8	30,8	44,4	51,2	54,6
	x_3	2,12	2,2	2,11	2,08	2,21	1,88	1,91	2,0	1,9	1,99
17	x_1	35,4	33,0	34,8	33,3	36,1	38,3	30,6	32,1	37,6	34,8
	x_2	30,0	44,4	51,2	54,6	57,6	53,2	57,6	58,3	55,7	55,7
	x_3	1,9	2,0	1,91	1,99	1,54	1,74	2,23	2,14	1,84	2,01
18	x_1	30,6	32,1	37,6	35,8	34,2	34,4	32,5	33,4	37,8	35,8
	x_2	57,6	58,3	55,7	55,7	56,4	60,4	34,4	33,9	34,2	38,0
	x_3	2,23	2,14	1,84	2,01	2,04	2,02	1,98	1,95	1,7	2,1
19	x_1	34,2	37,2	38,2	29,4	37,2	34,5	35,0	43,7	31,9	37,3
	x_2	30,3	31,1	28,5	36,7	38,7	45,7	41,5	39,0	42,7	39,8
	x_3	1,75	1,86	1,82	1,89	2,01	1,75	1,6	1,69	1,78	2,48
20	x_1	40,2	39,4	43,7	38,4	38,8	39,9	30,1	31,7	37,2	39,7
	x_2	66,5	63,0	65,5	64,2	61,6	63,2	75,5	75,5	71,0	67,2
	x_3	1,95	1,62	2,24	1,88	1,68	1,62	2,15	2,03	1,95	1,63
21	x_1	16,8	17,0	14,8	17,9	16,9	19,7	14,0	17,1	18,2	17,4
	x_2	40,4	42,3	40,1	39,4	39,1	43,2	44,5	45,7	37,8	46,4
	x_3	1,88	2,0	2,21	2,01	2,31	2,16	1,74	2,25	1,87	1,82
22	x_1	19,7	14,0	17,2	18,3	17,4	16,1	18,8	17,9	17,6	15,7
	x_2	37,8	46,4	48,2	49,6	46,4	42,6	49,4	40,1	39,4	39,1
	x_3	1,74	2,25	1,88	1,82	1,9	1,98	1,77	1,89	2,0	2,21
23	x_1	15,6	13,5	15,3	14,9	15,1	16,1	16,7	15,4	17,1	16,8
	x_2	40,4	38,9	36,6	41,4	32,2	31,4	32,6	38,7	44,3	39,3
	x_3	2,11	2,78	2,17	2,15	2,11	1,97	1,96	2,12	2,02	2,13
24	x_1	15,1	16,1	16,7	15,5	17,2	16,9	17,0	16,2	15,0	18,0
	x_2	32	32,4	32,6	38,7	44,3	39,3	40,4	41,5	45,2	50,2
	x_3	2,12	1,98	1,96	2,15	2,02	2,05	2,02	2,13	2,14	1,9
25	x_1	16,9	16,2	15,5	18,2	17,3	17,1	16,4	16,7	14,2	17,2
	x_2	39,2	41	41,3	45,2	50,2	51,6	48,0	48,6	49,8	45,0
	x_3	2,02	2,13	2,14	1,89	2,48	1,94	1,93	1,96	2,57	2,21
26	x_1	11,7	18,3	18,2	15,6	17,4	13,8	15,0	18,6	16,2	15,7
	x_2	44,2	42,4	38,9	36,4	35,2	45,2	35,2	41,6	42,2	40,4
	x_3	2,85	1,81	1,88	2,25	1,98	2,48	2,25	1,85	2,15	2,25
27	x_1	18,4	15,6	17,5	13,8	15,0	18,7	16,2	15,7	17,9	15,4
	x_2	38,7	36,5	35,2	45,5	35,2	41,6	42,2	40,4	47,3	47,1
	x_3	1,89	2,25	1,98	2,48	2,25	1,85	2,15	2,25	1,9	2,2
28	x_1	13,8	15,0	18,6	16,2	15,7	17,9	15,3	16,3	17,7	16,8
	x_2	45,5	35,2	41,6	42,3	40,4	47,4	47,1	43,2	39,1	38,2
	x_3	2,46	2,25	1,85	2,15	2,25	1,9	2,2	2,09	1,87	2,0

Продовження таблиці 2.17

29	x_1	18,8	16,4	16,1	17,8	15,5	16,3	17,8	16,8	17,5	16,7
	x_2	42,1	42,3	40,4	47,4	47,1	43,2	39,0	38,2	37,3	36,7
	x_3	1,86	2,15	2,25	1,9	2,2	2,1	1,86	2,0	2,48	2,02
30	x_1	13,8	14,8	16,9	16,8	14,8	18,0	17,5	15,7	15,3	14,9
	x_2	48,0	46,4	42,3	39,4	42,3	40,1	39,4	39,1	43,2	44,7
	x_3	2,45	2,3	2,0	2,05	2,23	1,89	2,0	2,21	2,01	2,31

Дані необхідно вибрати з таблиці відповідно до номера варіанта.
Номер варіант визначається за вказівкою викладача.

2.5 Модель лінійної регресії з гетероскедастичними збуреннями

Відведений час: 2 год.

Мета: навчати здійснювати перевірку гетероскедастичності залишків (відхилень); формувати первинні навички знаходження оцінок параметрів у випадку гетероскедастнчності відхилень (оцінювати параметри моделі узагальненим методом найменших квадратів *(методом Ейткена))*.

Завдання для практичного заняття:

1. Пригадайте основні теоретичні питання теми.
2. Орієнтовні запитання та завдання:
   - дайте означення гетероскедастичності.
   - як впливає наявність гетероскедастичності на оцінку параметрів моделі ?
   - які існують методи визначення гетероскедастичності?
   - суть тесту Голдфельда-Квандт.
   - за якою формулою знаходять оцінки параметрів моделі при використані методу Ейткена ?
3. Виконайте індивідуальне завдання.

Короткі теоретичні відомості

Якщо дисперсія залишків в економетричному моделюванні змінюється для кожного спостереження або для груп спостережень, то це явище називається *гетероскедастичністю*.

Гетероскедастичність призводить до того, що оцінки параметрів моделі більше не є кращими оцінками, або не є оцінками з мінімальною дисперсією, тобто вони не володіють властивістю ефективності.

Наявність гетероскедастичності спричиняє порушення властивостей оцінок параметрів моделі при розрахунку їх за методом 1МНК. Тому завжди виникає необхідність вивчати це явище, і, якщо воно існує, для оцінки параметрів моделі використовувати узагальнений метод найменших квадратів (метод Ейткена).

Для визначення гетероскедастичності застосовуються чотири критерії:

1) $F_{критерій}$ μ;

2) параметричний тест Гольдфельда—Квандта;

3) непараметричний тест Гольдфельда—Квандта;

4) тест Глейсера.

Розглянемо параметричний *тест Голдфельда-Квандта*. Його застосовують у випадку коли сукупність спостережень невелика.

При проведенні перевірки за цим тестом припускають, що стандартне відхилення (σ_i) розподілу ймовірностей ε_i пропорційно значенню X у цьому спостереженні. Також припускається, що збурення розподілено нормально і не підтверджено автокореляцію.

Тест Голдфельда-Квандта складається з таких кроків:

1) усі *n* спостережень упорядковуємо за зростанням X;
2) відкидаємо c середніх спостережень ($c \approx \frac{n \cdot 4}{15}$), при чому бажано вибирати таке значення c, яке дозволяє мати два підмасива однокової довжини;
3) будуємо дві економетричні моделі на основі 1МНК за двома створеними сукупностями спостережень, знаходимо суми квадратів залишків у кожній з побудованих регресій (S_1 та S_2 відповідно);
4) розрахуємо критерій *R*:

$$R = \frac{S_2}{S_1},$$

який при виконанні гіпотези про гомоскедастичність буде відповідати *F*-розподілу з $(n_1 - c - 2m)/2$, $(n_2 - c - 2m)/2$ ступенями свободи. Це означає, що розраховане значення R порівнюється з табличним значенням F-критерію при ступенях свободи $(n_1 - c - 2m)/2$ і $(n_2 - c - 2m)/2$ і вибраному рівні довіри. Якщо $R \le F_{табл}$, то гетероскедастичність відсутня.

Метод Голдфельда-Квандта можна застосовувати і при припущенні, що σ_i обернено пропорційно значенню X. При цьому використовується така ж процедура, але тестовою статистикою буде відношення $F = S_1/S_2$.

При наявності в масиві даних гетероскедастичності для оцінки параметрів моделі метод найменших квадратів (1МНК) застосовувати не можна. У такому випадку оцінка параметрів моделі здійснюється на основі узагальненого методу найменших квадратів ***(методу Ейткена).***

За методом Ейткена оператор оцінювання має вигляд:

$$\hat{A} = (X'S^{-1}X)^{-1}X'S^{-1}Y,$$

При цьому матрицю S^{-1} визначаємо користуючись гіпотезою: $\lambda_i = \dfrac{1}{x_i}$,

тобто

$$S^{-1} = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \lambda_4 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & \lambda_5 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & \lambda_{14} \end{pmatrix}$$

Приклад розв'язування задачі

Приклад. Проведено опитування 18 родин. У таблиці 2.18 наведено дані про величину їх заощаджень (y) та доходів (x). Оскільки величина доходів дуже сильно відрізняється від величини заощаджень, то необхідно виконати перевірку масиву даних на наявність гетероскедастичності.

Таблиця 2.18 - Результати опитування

№	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
y	2,3	2,2	2,08	2,2	2,1	2,32	2,45	2,5	2,2	2,5	3,1	2,5	2,82	3,04	2,7	3,94	3,1	3,99
x	15	15	16	17	17	18	19	20	20	22	64	68	72	80	85	90	95	100

Розв'язання

Упорядкуємо дані в таблиці 2.18 за зростанням x (дивись таблицю 2.19). Обчислимо значення c: $c = \dfrac{18 \cdot 4}{15} = 4{,}8$. Для того, щоб після відкидання c спостережень залишилось два однакових підмасива нехай $c \approx 4$.

Таблиця 2.19 - Упорядковані дані та додаткові обчислення

№ спостереження	Заощадження (y)	Дохід (x)	$\hat{y}$ розрахункові	Залишки (е)	Квадрат залишків (e^2)
1	2	3	4	5	6
1	2,3	15	2,16	2,3 –2,16=0,14	0,020
2	2,2	15	2,16	2,2 – 2,16=0,04	0,002
3	2,08	16	2,20	2,08 – 2,2=-0,12	0,015
4	2,2	17	2,25	2,2 –2,25=-0,05	0,002
5	2,1	17	2,25	2,1 –2,25=-0,15	0,022
6	2,32	18	2,29	2,32 –2,29=0,03	0,001
7	2,45	19	2,34	2,45 –2,34=0,11	0,012
8	2,5	20			S_1=0,07453

Продовження таблиці 2.19

9	2,2	20			
10	2,5	22			
11	3,1	64			
12	2,5	68	2,53	2,5 –2,53=-0,03	0,001
13	2,82	72	2,68	2,82 –2,68=0,14	0,019
14	3,04	80	2,99	3,04 –2,99=0,05	0,002
15	2,7	85	3,18	2,7 –3,18=-0,48	0,234
16	3,94	90	3,38	3,94 –3,38=0,56	0,318
17	3,1	95	3,57	3,1 –3,57=-0,47	0,220
18	3,99	100	3,76	3,99 –3,76=0,23	0,052
					S_2=0,84581

Побудуємо регресії для кожного з підмасивів за допомогою 1МНК (допоміжні обчислення таблиця 2.20).

Таблиця 2.20 – Перший підмасив

№	y	x	$y \cdot x$	x^2
1	2,3	15	34,5	225
2	2,2	15	33	225
3	2,08	16	33,28	256
4	2,2	17	37,4	289
5	2,1	17	35,7	289
6	2,32	18	41,76	324
7	2,45	19	46,55	361
Сума	15,65	117	262,19	1969

$$\begin{cases} 7b_0 + 117b_1 = 15{,}65 \\ 117b_0 + 1969b_1 = 262 \end{cases}$$

Звідси, $b_0 = 1{,}475$, $b_1 = 0{,}045$.

Отже, $\hat{y} = 0{,}045 + 1{,}475x$.

Розрахуємо теоретичні значення результуючої змінної та занесемо їх у таблицю 2.19 - 4 стовпчик (допоміжні обчислення таблиця 2.21).

Таблиця 2.21 – Другий підмасив

№	y	x	$y \cdot x$	x^2
12	2,5	68	170	4624
13	2,82	72	203,04	5184
14	3,04	80	243,2	6400
15	2,7	85	229,5	7225
16	3,94	90	354,6	8100
17	3,1	95	294,5	9025
18	3,99	100	399	10000
Сума	22,09	590	1893,84	50558

$$\begin{cases} 7b_0 + 590b_1 = 22{,}09 \\ 590b_0 + 50558b_1 = 1894 \end{cases}$$

Звідси, $b_0 = -0{,}093$, $b_1 = 0{,}038$.

Отже, $\hat{y} = -0{,}093 + 0{,}038x$.

Розрахуємо теоретичні значення результуючої змінної та занесемо їх у таблицю (4 стовпчик).

Заповнимо стовпчики 5 та 6 у таблиці 2.19 (залишки та квадрат залишків) та знайдемо суми квадратів залишків по кожному із підмасивів.

Отримали, що S_1=0,07453 , S_2=0,84581.

Тоді, $R = \frac{S_2}{S_1} = 11{,}3483$. Табличне значення F-критерію дорівнює 11.

Оскільки $R > F_{табл}$, то гетероскедастичність присутня.

Отже, метод найменших квадратів для оцінки параметрів моделі застосовувати не можна.

Використаємо метод Ейткена:

$$\hat{A} = (X'S^{-1}X)^{-1}X'S^{-1}Y \ .$$

Запишемо матрицю змінних:

$$X' = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 15 & 15 & 16 & 17 & 17 & 18 & 19 & 68 & 72 & 80 & 85 & 90 & 95 & 100 \end{pmatrix}.$$

Визначимо матрицю S^{-1} ($\lambda_i = \frac{1}{x_i}$), тобто

$$S^{-1} = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \lambda_4 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & \lambda_5 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & \lambda_{14} \end{pmatrix}$$

Визначимо добутки матриць:

$$X'S^{-1} = \begin{pmatrix} 0{,}0667 & 0{,}0667 & 0{,}0625 & 0{,}0589 & 0{,}0589 & 0{,}0555 & 0{,}0526 & 0{,}05 & 0{,}05 & 0{,}0154 & 0{,}0156 & 0{,}0148 & 0{,}0139 & 0{,}0125 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

$$X'S^{-1}X = \begin{pmatrix} 0{,}6672 & 18 \\ 18 & 833 \end{pmatrix};$$

Знайдемо обернену матрицю:

$$(X'S^{-1}X)^{-1} = \begin{pmatrix} 3{,}5934 & -0{,}0776 \\ -0{,}0776 & 0{,}0029 \end{pmatrix};$$

та добуток:

$$X'S^{-1}Y = \begin{pmatrix} 1,5998 \\ 48,04 \end{pmatrix}.$$

Обчислимо вектор оцінок параметрів моделі:

$$\hat{A} = \begin{pmatrix} 3,5934 & -0,0776 \\ -0,0776 & 0,0029 \end{pmatrix} * \begin{pmatrix} 1,5998 \\ 48,04 \end{pmatrix} = \begin{pmatrix} 2,0187 \\ 0,0141 \end{pmatrix}.$$

Звідси $\widehat{b}_0 = 2,0187$; $\widehat{b}_1 = 0,0141$.

Економетрична модель витрат на харчування запишеться так:

$$\widehat{y} = 2,0187 + 0,0141x \ .$$

Параметр моделі $\widehat{b}_1 = 0,0141$ свідчить про те, що збільшення загальних затрат на одиницю сприятиме граничному зростанню витрат на харчування на 0,014 одиниць.

Далі проводиться економічний аналіз характеристик економетричної моделі. Обчислюється:

- коефіцієнт детермінації, $R^2 = 0,722$ (це означає, що на 72,2% варіація витрат на харчування залежить від варіації загальних затрат);

- коефіцієнт кореляції, $R = \sqrt{R^2} = 0,85$ (свідчить про досить тісний зв'язок витрат на харчування і загальних затрат);

- залишкова дисперсія, $\sigma_U^2 = 0,083$ (показує, що розрахункові значення витрат на харчування дуже близькі до фактичних);

- матриця коваріацій оцінок параметрів моделі:

$$\text{var}(\hat{A}) = \sigma_U^2 (X'S^{-1}X)^{-1} = 0,083 \begin{pmatrix} 3,5934 & -0,0776 \\ -0,0776 & 0,0029 \end{pmatrix} = \begin{pmatrix} 0,298 & -0,0064 \\ -0,0064 & 0,0002 \end{pmatrix}$$

(діагональні елементи цієї матриці є дисперсіями оцінок параметрів моделі, інші елементи характеризують коваріацію між оцінками);

- стандартні помилки оцінок параметрів

$$S_{\hat{a}_j} = \sqrt{\text{var}(\hat{A})} = \sigma_U \sqrt{c_{jj}};$$
$$S_{\hat{a}_0} = \sqrt{0,298} = 0,546 \ ;$$
$$S_{\hat{a}_1} = \sqrt{0,0002} = 0,014 \ .$$

- довірчі інтервали оцінок (для побудови довірчих інтервалів оцінок параметрів моделі знайдемо t-критерій при ступенях свободи $n - m = 12$ і рівні довіри $\alpha = 0,05$: $t_{\text{крит}} = 2,179$):

$$\hat{a}_0 - t_{0,05} S_{\hat{a}_0} \le \hat{a}_0 \le \hat{a}_0 + t_{0,05} S_{\hat{a}_0}$$

$$2{,}0187 - 1{,}102 \le a_0 \le 2{,}0187 + 1{,}102$$

$$0{,}917 \le a_0 \le 3{,}121$$

$$\hat{a}_1 - t_{0,05} S_{\hat{a}_1} \le \hat{a}_1 \le \hat{a}_1 + t_{0,05} S_{\hat{a}_1}$$

$$0{,}016 \le a_1 \le 0{,}045$$

Рівень стандартних помилок та довірчі інтервали оцінок параметрів моделі свідчать про те, що отримані оцінки є неефективними та зміщеними.

Завдання для індивідуальної та самостійної роботи студентів

Приклад. Нехай треба побудувати економетричну модель, яка характеризує залежність обсягу споживання (на душу населення) від ціни (за одиницю продукції) та доходу (на душу населення). Для побудови цієї моделі використовується вихідна сукупність даних, яка включає 10 спостережень (таблиця 2.22). Ці дані наведені в табл. Для того, щоб правильно вибрати метод для оцінки параметрів моделі, необхідно перевірити, чи властива гетероскедастичність для наведених вихідних даних.

Таблиця 2.22 - Вихідна сукупність даних

1	Y	75	67	56	59	51	58	60	65	45	40
	x_1	5	6	7	8	8	10	7	8	9	10
	x_2	100	110	112	115	120	135	120	130	110	100
2	Y	55	87	56	59	51	58	60	65	70	50
	x_1	15	16	17	18	18	20	17	18	19	15
	x_2	160	170	112	115	120	135	145	100	110	120
3	Y	50	45	55	50	60	70	80	85	75	90
	x_1	10	12	9	10	8	6	8	10	12	7
	x_2	70	62	62	60	55	50	82	65	74	87
4	Y	55	57	56	59	51	58	60	50	45	65
	x_1	5	6	7	8	9	10	8	9	15	8
	x_2	100	110	112	115	120	135	140	120	110	130
5	Y	50	45	55	50	60	70	52	47	55	50
	x_1	10	12	9	10	8	16	10	12	9	10
	x_2	60	62	65	60	55	50	70	62	62	60
6	Y	50	45	55	5060	70	55	57	56	59	51
	x_1	10	12	9	10	8	6	5	6	7	8
	x_2	60	62	62	60	55	50	20	22	20	25
7	Y	50	45	55	50	60	70	50	45	55	50
	x_1	10	12	9	10	18	6	15	12	19	10
	x_2	70	62	72	60	55	60	60	82	62	60
8	Y	50	45	55	50	60	70	50	45	55	50
	x_1	10	12	19	10	8	16	10	12	9	10
	x_2	90	62	82	60	85	50	70	62	62	80

Продовження таблиці 2.22

9	Y	55	57	56	59	51	58	60	50	45	65
	x_1	10	12	9	10	8	6	8	10	12	7
	x_2	70	62	62	60	55	50	82	65	74	87
10	Y	100	120	130	125	140	150	155	160	100	120
	x_1	30	35	40	30	45	30	55	60	30	35
	x_2	90	75	85	90	105	93	97	100	80	75
11	Y	200	220	230	225	240	250	255	260	200	220
	x_1	30	35	40	35	45	50	55	62	30	38
	x_2	80	75	85	90	95	93	97	100	80	75
12	Y	110	130	140	125	140	160	155	180	120	120
	x_1	20	25	30	20	35	20	45	50	20	25
	x_2	20	95	85	100	95	103	97	100	90	85
13	Y	100	120	130	125	140	150	155	160	100	120
	x_1	30	35	40	30	45	50	55	60	30	35
	x_2	8	17	18	19	19	19	20	9	7	28
14	Y	100	120	130	125	140	150	155	160	150	140
	x_1	30	35	40	30	45	50	55	60	65	70
	x_2	28	27	25	20	35	40	45	50	120	125
15	Y	300	280	350	340	330	320	310	300	320	280
	x_1	25	20	30	30	28	28	25	24	27	22
	x_2	25	24	26	27	27	25	25	26	24	25
16	Y	350	280	350	340	300	320	320	280	380	340
	x_1	26	22	30	30	29	28	25	24	23	21
	x_2	15	14	16	17	17	15	16	14	16	15
17	Y	400	380	350	360	430	420	310	400	350	380
	x_1	25	20	30	28	28	25	24	25	20	30
	x_2	35	34	36	37	37	35	36	34	35	34
18	Y	320	280	350	340	330	320	310	300	310	290
	x_1	15	10	20	20	18	18	15	14	15	10
	x_2	10	11	13	12	14	11	12	9	10	11
19	Y	50	45	55	50	60	70	52	47	55	50
	x_1	10	12	9	10	8	16	10	12	9	10
	x_2	60	62	65	60	55	50	70	62	62	50
20	Y	100	120	130	125	140	150	155	160	100	120
	x_1	30	35	40	30	45	30	55	60	30	35
	x_2	90	75	85	90	105	93	97	100	80	75
21	Y	226	224	226	230	253	263	266	277	298	309
	x_1	227	232	234	239	251	250	259	255	273	268
	x_2	89	67	72	61	57	71	39	74	60	74
22	Y	268	257	284	282	285	301	299	321	323	343
	x_1	189	187	189	186	184	200	192	205	208	217
	x_2	99	98	84	80	66	90	78	79	85	78
23	Y	110	124	127	138	152	177	180	189	204	220
	x_1	291	294	292	293	296	308	310	311	318	312
	x_2	78	74	61	60	59	68	74	65	62	43
24	Y	181	188	200	204	232	257	272	278	316	324
	x_1	161	171	167	161	168	177	185	176	186	185
	x_2	78	86	85	64	67	75	85	65	65	63

Продовження таблиці 2.22

25	Y	206	204	233	246	267	289	301	326	335	342
	x_1	196	200	201	215	221	226	217	240	245	242
	x_2	106	98	81	110	99	87	65	83	90	66
26	Y	266	271	281	294	304	321	317	329	328	343
	x_1	240	237	258	260	263	267	263	272	274	277
	x_2	81	74	103	95	89	79	73	71	70	69
27	Y	219	222	222	241	256	268	279	295	292	307
	x_1	196	187	187	194	194	202	204	208	211	208
	x_2	96	70	74	71	70	72	76	73	86	61
28	Y	123	157	173	175	193	189	208	242	234	227
	x_1	206	213	220	217	232	230	245	250	239	256
	x_2	93	82	76	74	83	91	87	70	64	90
29	Y	116	137	147	167	175	189	219	231	228	239
	x_1	189	196	197	202	202	201	213	222	208	211
	x_2	95	97	96	100	95	82	89	106	83	75
30	Y	255	268	285	299	301	318	333	353	382	384
	x_1	245	246	247	258	251	257	279	267	290	290
	x_2	87	89	80	87	76	76	98	66	84	76

Дані необхідно вибрати з таблиці відповідно до номера варіанта.
Номер варіант визначається за вказівкою викладача.

2.6 Модель лінійної регресії з автокорельованими збуреннями

Відведений час: 2 год.

Мета: навчати здійснювати перевірку автокореляції відхилень, формувати первинні навички оцінювання параметрів моделі у випадку автокореляції.

Завдання для практичного заняття:

1. Пригадайте основні теоретичні питання теми.
2. Орієнтовні запитання та завдання:
   - дайте означення автокореляції.
   - які причини виникнення автокореляції відхилень ?
   - як визначають наявність та відсутність автокореляції відхилень ?
   - які методи оцінки параметрів моделі використовують у випадку автокореляції відхилень?
   - за якою формулою знаходять коефіцієнт кореляції відхилень ρ ?
   - за якою формулою знаходять оцінки параметрів моделі у випадку автокореляції відхилень?
3. Виконайте індивідуальне завдання.

Короткі теоретичні відомості

В економетричних дослідженнях часто зустрічаються такі випадки, коли дисперсія залишків є постійною, але спостерігається їх коваріація. Це явище має назву автокореляції залишків.

Автокореляція має місце тоді, коли залишки не є залежними один від одного, оскільки поточні значення Y знаходяться під впливом минулих значень. Залежність між залишками можна описати таким чином: приміром, залишки e_t знаходяться під впливом залишків з попереднього періоду часу e_{t-1} та деякого поточного значення випадкової змінної u_t. Тоді залишок e_t буде описуватись наступною авторегресійною функцією першого порядку:

$$e_t = \rho e_{t-1} + u_t ,$$

де ρ - характеризує силу зв'язку величини залишків у період t від величини залишків у період $t-1$.

Автокореляція відхилень може бути наслідком кореляції між послідовними значеннями деякого фактора x_i, великих похибок при одержанні даних, помилкової специфікації форми залежності між змінними або відсутності в рівнянні регресії деякого суттєвого фактора.

При оцінювані параметрів економетричної моделі методом найменших квадратів без врахування наявності автокореляції відхилень можливі такі **наслідки**:

- оцінки параметрів моделі будуть зміщеними, неефективними;
- неефективність оцінок параметрів приводить до прогнозу, який може мати велику вибіркову дисперсію;
- в дисперсійному аналізі не можна застосувати статистичні критерії t (Ст'юдента) та F (Фішера).

Найчастіше наявність автокореляції відхилень перевіряють за критеріям Дарбіна - Уотсона, застосовують також критерій Неймана, циклічний або нециклічний коефіцієнти автокореляції.

Розглянемо *критерий Дарбіна-Уотсона*. Відповідно до цього критерію необхідно обчислити d-статистику (або DW):

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

Критерій Дарбіна—Уотсона може приймати значення на множині $DW \in [0,4]$.

Якщо критерій Дарбіна-Уотсона дорівнює двом, то не автокореляції існує, якщо - нулю, то має місце абсолютна додатні автокореляція, а якщо – чотирьом, то має місце абсолютно від'ємна автокорреляція. Існує таблиця (див додатки), яка містить критичні значення: нижнє d_L та верхнє d_u (або d_H и d_B). Обчислене значення d-порівнюють з табличним (при заданому рівні значущості α, відповідно до кількості спостережень в сукупності n та кількості незалежних змінних моделі m). Для d < 2 керуються наступним правилом:

– якщо $d < d_L$, то має місце додатна автокореляція;

– якщо $d > d_u$, то автокореляції відсутня;

– якщо $d_L < d < d_u$, то нічого певного сказати не можна.

Якщо розраховане значення d більше двох (d > 2), то перевіці підлягає величина величина (4 – d). Висновки роблять аналогічні (тільки автокореляція буде від'ємна).

Автокореляція може виникнути через те, що не всі важливі фактори введено у модель або неправильно вибрано вид рівняння регресії.

Для оцінювання параметрів економетричної моделі, з автокорельованими відхиленнями існує декілька методів: загальний метод найменших квадратів для випадку автокореляції і Ейткена), метод перетворення вихідної інформації та наближені методи Дарбіна і Кочрена – Орката.

Метод Ейткена. Оператор оцінювання запишеться так:

$$\widehat{A} = (X'S^{-1}X)^{-1}X'S^{-1}Y,$$

де

$$S^{-1} = \frac{1}{1-\rho^2}\begin{pmatrix} 1 & -\rho & 0 & 0 & 0 & \cdots & 0 \\ -\rho & 1+\rho^2 & -\rho & 0 & 0 & \cdots & 0 \\ 0 & -\rho & 1+\rho^2 & -\rho & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 1 \end{pmatrix}.$$

На практиці для розрахунку ρ використовуються співвідношення:

$$\rho \approx r = \frac{\sum_{t=2}^{n} u_t \ u_{t-1}}{\sum_{t=1}^{n} u_t^2} \quad \text{або} \quad \rho \approx r' = \frac{\sum_{t=2}^{n} u_t \ u_{t-1}}{\sum_{t=1}^{n} u_t^2} \cdot \frac{n}{n-1}$$

Після визначення оцінок параметрів рівняння регресії бажано ще раз застосувати критерій Дарбіна—Уотсона (з метою перевірки на наявність автокореляції). У випадку, якщо ми не звільнились від автокореляції залишків, то це означає, що вихідна гіпотеза, що залишки описуються авторегресійною схемою першого порядку, не дотримується. Тоді залишки описуються авторегресійною схемою більш високого порядку, тобто доцільно виконати оцінку параметрів моделі методом Кочрена—Оркатта або Дарбіна.

Приклад розв'язування задачі

Приклад. Перевірити наявність автокореляції залишків (масив даних у таблиці 2.23), використовуючи критерій Дарбіна—Уотсона.

Таблиця 2.23 – Вихідні дані

№	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**
y	**2,3**	**2,2**	**2,08**	**2,2**	**2,1**	**2,32**	**2,45**	**2,5**	**2,2**	**2,5**	**3,1**	**2,5**	**2,82**	**3,04**	**2,7**	**3,94**	**3,1**	**3,99**
x	**15**	**15**	**16**	**17**	**17**	**18**	**19**	**20**	**20**	**22**	**64**	**68**	**72**	**80**	**85**	**90**	**95**	**100**

Розв'язання. За допомогою методу 1МНК побудуємо економетричну модель (див. попередня практична).

У результаті отримали економетричну модель витрат на харчування у вигляді:

$$\hat{y} = 1{,}999 + 0{,}014x \cdot$$

Для подальшого застосування критерія Дарбіна—Уотсона проведемо допоміжні обчислення (таблиця 2.24).

Таблиця 2.24 - Допоміжні обчислення

y	x	$\hat{y}$ розрах.	Залишки (e)	$(e_i - e_{i-1})^2$	$e_i^{\ 2}$
1	2	3	4	5	6
2,3	15	2,21607	0,08393		0,00704
2,2	15	2,21607	-0,01607	0,01	0,00026
2,08	16	2,23055	-0,15055	0,01808	0,02267
2,2	17	2,24503	-0,04503	0,01114	0,00203
2,1	17	2,24503	-0,14503	0,01	0,02103
2,32	18	2,2595	0,0605	0,04224	0,00366
2,45	19	2,27398	0,17602	0,01335	0,03098
2,5	20	2,28846	0,21154	0,00126	0,04475
2,2	20	2,28846	-0,08846	0,09	0,00782
2,5	*22*	2,31741	0,18259	0,07347	0,03334
3,1	64	2,92546	0,17454	6,5E-05	0,03046
2,5	68	2,98337	-0,48337	0,43284	0,23364
2,82	72	3,04128	-0,22128	0,06869	0,04896
3,04	80	3,15709	-0,11709	0,01085	0,01371
2,7	85	3,22948	-0,52948	0,17006	0,28035
3,94	90	3,30187	0,63813	1,36332	0,40721
3,1	95	3,37425	-0,27425	0,83245	0,07522
3,99	100	3,44664	0,54336	0,66849	0,29524
Сума				3,81631	1,55134

$$d = \frac{\sum_{t=2}^{18} (e_t - e_{t-1})^2}{\sum_{t=1}^{18} e_t^2} = \frac{3{,}81631}{1{,}55134} = 2{,}46. \qquad 4 - \mathrm{d} = 4 - 2{,}46 = 1{,}54.$$

З таблиці Б.3 Дарбина-Уотсона при рівні значущост 0,05, при n = 18 і кількістю незалежних змінних k = 1, визначимо граничні значення:

d_L = 1,16, d_u = 1,39.

Оскільки 4 – d = 1,54 > d_u = 1,39, то з ймовірністю 0,95 можно стверджувати, що автокореляція залишків, побудованої моделі, не існує.

Завдання для індивідуальної та самостійної роботи студентів

Приклад. Перевірити наявність автокореляції залишків (масив даних у таблиці 2.25), використовуючи критерій Дарбіна—Уотсона.

Таблиця 2.25 – Вихідні дані

1	Y	100	120	130	125	140	150	155	160	100	120
	x_1	30	35	40	30	45	30	55	60	30	35
	x_2	90	75	85	90	105	93	97	100	80	75
2	Y	226	224	226	230	253	263	266	277	298	309
	x_1	227	232	234	239	251	250	259	255	273	268
	x_2	89	67	72	61	57	71	39	74	60	74
3	Y	268	257	284	282	285	301	299	321	323	343
	x_1	189	187	189	186	184	200	192	205	208	217
	x_2	99	98	84	80	66	90	78	79	85	78
4	Y	110	124	127	138	152	177	180	189	204	220
	x_1	291	294	292	293	296	308	310	311	318	312
	x_2	78	74	61	60	59	68	74	65	62	43
5	Y	181	188	200	204	232	257	272	278	316	324
	x_1	161	171	167	161	168	177	185	176	186	185
	x_2	78	86	85	64	67	75	85	65	65	63
6	Y	206	204	233	246	267	289	301	326	335	342
	x_1	196	200	201	215	221	226	217	240	245	242
	x_2	106	98	81	110	99	87	65	83	90	66
7	Y	266	271	281	294	304	321	317	329	328	343
	x_1	240	237	258	260	263	267	263	272	274	277
	x_2	81	74	103	95	89	79	73	71	70	69
8	Y	219	222	222	241	256	268	279	295	292	307
	x_1	196	187	187	194	194	202	204	208	211	208
	x_2	96	70	74	71	70	72	76	73	86	61
9	Y	123	157	173	175	193	189	208	242	234	227
	x_1	206	213	220	217	232	230	245	250	239	256
	x_2	93	82	76	74	83	91	87	70	64	90
10	Y	116	137	147	167	175	189	219	231	228	239
	x_1	189	196	197	202	202	201	213	222	208	211
	x_2	95	97	96	100	95	82	89	106	83	75
11	Y	75	67	56	59	51	58	60	65	45	40
	x_1	5	6	7	8	8	10	7	8	9	10
	x_2	100	110	112	115	120	135	120	130	110	100
12	Y	55	87	56	59	51	58	60	65	70	50
	x_1	15	16	17	18	18	20	17	18	19	15
	x_2	160	170	112	115	120	135	145	100	110	120
13	Y	50	45	55	50	60	70	80	85	75	90
	x_1	10	12	9	10	8	6	8	10	12	7
	x_2	70	62	62	60	55	50	82	65	74	87
14	Y	55	57	56	59	51	58	60	50	45	65
	x_1	5	6	7	8	9	10	8	9	15	8
	x_2	100	110	112	115	120	135	140	120	110	130
15	Y	50	45	55	50	60	70	52	47	55	50
	x_1	10	12	9	10	8	16	10	12	9	10
	x_2	60	62	65	60	55	50	70	62	62	60

Продовження таблиці 2.25

16	Y	50	45	55	5060	70	55	57	56	59	51
	x_1	10	12	9	10	8	6	5	6	7	8
	x_2	60	62	62	60	55	50	20	22	20	25
17	Y	50	45	55	50	60	70	50	45	55	50
	x_1	10	12	9	10	18	6	15	12	19	10
	x_2	70	62	72	60	55	60	60	82	62	60
18	Y	50	45	55	50	60	70	50	45	55	50
	x_1	10	12	19	10	8	16	10	12	9	10
	x_2	90	62	82	60	85	50	70	62	62	80
19	Y	55	57	56	59	51	58	60	50	45	65
	x_1	10	12	9	10	8	6	8	10	12	7
	x_2	70	62	62	60	55	50	82	65	74	87
20	Y	100	120	130	125	140	150	155	160	100	120
	x_1	30	35	40	30	45	30	55	60	30	35
	x_2	90	75	85	90	105	93	97	100	80	75
21	Y	200	220	230	225	240	250	255	260	200	220
	x_1	30	35	40	35	45	50	55	62	30	38
	x_2	80	75	85	90	95	93	97	100	80	75
22	Y	110	130	140	125	140	160	155	180	120	120
	x_1	20	25	30	20	35	20	45	50	20	25
	x_2	20	95	85	100	95	103	97	100	90	85
23	Y	100	120	130	125	140	150	155	160	100	120
	x_1	30	35	40	30	45	50	55	60	30	35
	x_2	8	17	18	19	19	19	20	9	7	28
24	Y	100	120	130	125	140	150	155	160	150	140
	x_1	30	35	40	30	45	50	55	60	65	70
	x_2	28	27	25	20	35	40	45	50	120	125
25	Y	300	280	350	340	330	320	310	300	320	280
	x_1	25	20	30	30	28	28	25	24	27	22
	x_2	25	24	26	27	27	25	25	26	24	25
26	Y	350	280	350	340	300	320	320	280	380	340
	x_1	26	22	30	30	29	28	25	24	23	21
	x_2	15	14	16	17	17	15	16	14	16	15
27	Y	400	380	350	360	430	420	310	400	350	380
	x_1	25	20	30	28	28	25	24	25	20	30
	x_2	35	34	36	37	37	35	36	34	35	34
28	Y	320	280	350	340	330	320	310	300	310	290
	x_1	15	10	20	20	18	18	15	14	15	10
	x_2	10	11	13	12	14	11	12	9	10	11
29	Y	50	45	55	50	60	70	52	47	55	50
	x_1	10	12	9	10	8	16	10	12	9	10
	x_2	60	62	65	60	55	50	70	62	62	50
30	Y	100	120	130	125	140	150	155	160	100	120
	x_1	30	35	40	30	45	30	55	60	30	35
	x_2	90	75	85	90	105	93	97	100	80	75

Дані необхідно вибрати з таблиці відповідно до номера варіанта.
Номер варіант визначається за вказівкою викладача.

3 ПРИКЛАДИ ВИКОНАННЯ ЗАВДАНЬ ЗА ДОПОМОГОЮ ТАБЛИЧНОГО ПРОЦЕСОРУ EXCEL

3.1 ПОБУДОВА ПАРНОЇ ЕКОНОМЕТРИЧНОЇ МОДЕЛІ ІЗ ВИКОРИСТАННЯМ МОЖЛИВОСТЕЙ ТАБЛИЧНОГО ПРОЦЕСОРУ EXCEL

Запустити табличний процесор **Excel.** Це можна зробити кількома способами:
а) натиснути кнопку **Пуск** та в переліку **Все программы** вибрати назву **Microsoft Excel**;
б) якщо на робочому столі є ярлик **Microsoft Excel,** то двічі клацніть на ноьому лівою кнопкою миші.

При запуску **Excel** автоматично виводить на екран нову робочу книгу з умовним ім’ям **Книга 1**. Це ім’я з’являєтсья праворуч від імені **Microsoft Excel,** його можна змінити на те ім’я, яке Вам потрібно.

Розв’яжемо задачу, запропоновану у практичній, за допомогою табличного процесору **Excel**.

Приклад. Побудувати економетричну модель, яка характеризує залежність витрат на одиницю продукції від рівня фондомісткості продукції. Зробити економічні висновки. Вихідні дані наведені в таблиці 3.1.

Таблиця 3.1- Вихідні дані

№ п/п	1	2	3	4	5	6	7	8	9	10
Y_i	50	40	65	55	45	42	56	60	64	65
X_i	90	75	120	100	80	78	110	115	115	125

Розв’язування. Вихідні дані необхідно розмістити в стовпцях А і В. При цьому перший рядок відводиться для назв стовпців. Останній рядок призначений для обчислення автосум значень відповідних стовпців.

Для отримання якісних графіків вихідні дані необхідно відсортувати. Для чого виділяються значення фактора х (без назви !!!), що знаходяться у колонці А. На панелі інструментів натискається кнопка сортування за зростанням (А Я ↓). У діалоговому вікні (рисунок 3.1) необхідно натиснути кнопку **[Сортировка]**.

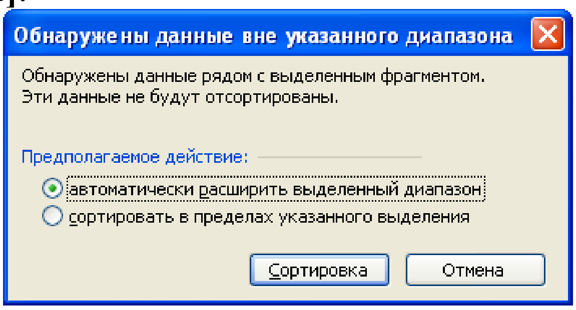


Рисунок 3.1 – Діалогове вікно

В результаті значення фактора будуть відсортовані за зростанням. При цьому відповідність між значеннями фактора x і результату y зберігаються.

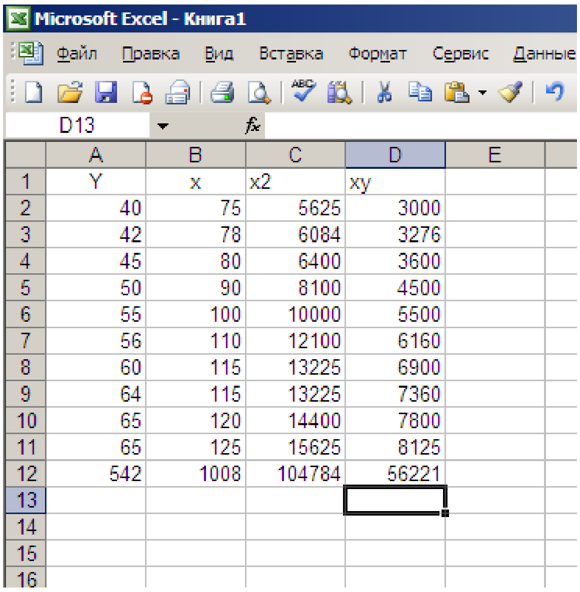
Розглянемо процес обчислення параметрів лінійного рівняння регресії $\hat{y} = b_0 + b_1 x$, а також оцінку його точності.

Для знаходження параметрів b_0 і b_1 необхідно скласти і вирішити наступну систему:

$$\begin{cases} b_0 n + b_1 \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i; \\ b_0 \sum_{i=1}^{n} x_i + b_1 \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i. \end{cases}.$$

де n – кількість спостережень (n=10).

Таким чином, для складання системи необхідно обчислити наступні значення сум: $\sum_{i=1}^{n} x_i$, $\sum_{i=1}^{n} y_i$, $\sum_{i=1}^{n} x_i^2$ и $\sum_{i=1}^{n} x_i \cdot y_i$ (рисунок 3.2).



	A	B	C	D	E
1	Y	x	x2	xy	
2	40	75	5625	3000	
3	42	78	6084	3276	
4	45	80	6400	3600	
5	50	90	8100	4500	
6	55	100	10000	5500	
7	56	110	12100	6160	
8	60	115	13225	6900	
9	64	115	13225	7360	
10	65	120	14400	7800	
11	65	125	15625	8125	
12	542	1008	104784	56221	
13					
14					
15					
16					

Рисунок 3.2 – Приклад зображення відповідних обчислень

Суми $\sum_{i=1}^{n} x_i$, $\sum_{i=1}^{n} y_i$ обчислюються відповідно у комірках з використанням кнопки [**Автосумма**] ([Σ]) на панелі инструментів.

Квадрати значень фактора x^2 обчислюються у стовпчику C. Для цього у комірку C2 заноситься формула: = A2^2. Дана формула розповсюджується на комірки, що лежать нижче, шляхом протягування за допомогою миші. У останній комірці стовпчику C обчислюється сума значень цього стовбця, тобто $\sum_{i=1}^{n} x_i^2$.

Добуток $x*y$ обчислюється у стовбці D. Для цього у комірку D2 вводиться формула: =A2*B2. Дана формула розповсюджується на комірки, що лежать нижче, шляхом протягування за допомогою миші. У останній комірці стовпчику D обчислюється сума значень цього стовпця, тобто $\sum_{i=1}^{n} x_i \cdot y_i$.

На основі проведених обчислень складається система нормальних рівнянь:

$$\begin{cases} 10\hat{b}_0 + 1008\hat{b}_1 = 542 \\ 1008\hat{b}_0 + 104784\hat{b}_1 = 56221. \end{cases}$$

Цю систему необхідно розв'язати з метою визначення параметрів b_0 і b_1.

Розв'язком системи є параметри $\hat{b}_0 = 3{,}8$; $\hat{b}_1 = 0{,}5$.

Отже, економетрична модель має вигляд

$$y = 3{,}8 + 0{,}5x.$$

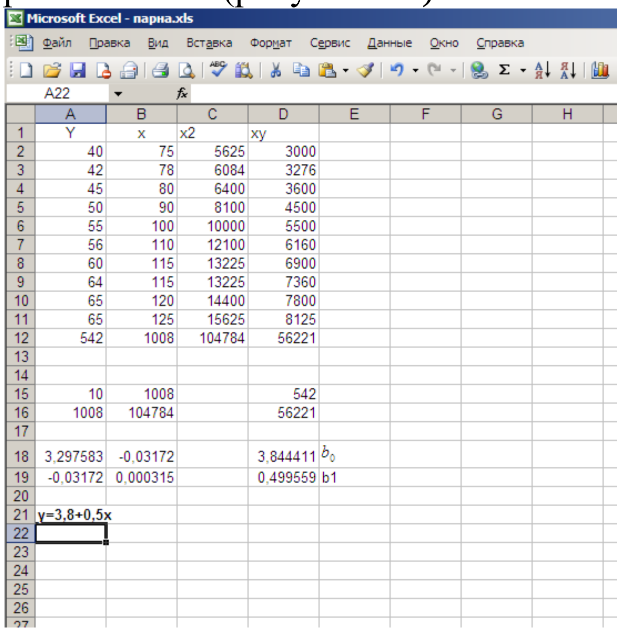
*Розглянемо інший можливий варіант побудови економетричної моделі за допомогою можливостей **Excel**.*

З математики відомо, що розв'язати систему лінійних рівнянь можна матричним способом. Для цього необхідно знайти обернену матрицю та помножити її на вектор-стовпець правих частин рівнянь системи.

Реалізуємо даний підхід в **Excel**. При цьому розв'язувати будемо в два етапи:

1. Визначимо матрицю обернену до матриці коефіцієнтів лівих частин рівнянь системи, за допомогою функції **МОБР**.
1. Визначимо розв'язок системи множенням оберненої матриці на матрицю правих частин рівнянь системи за допомогою **МУМНОЖ**.

Занесемо параметри системи в комірки, при цьому в осередках A15: B16 буде знаходиться матриця лівих частин рівнянь, а в осередках D15: D16 - матриця-стовпець правих частин (рисунок 3.3).



	A	B	C	D	E
1	Y	x	x2	xy	
2	40	75	5625	3000	
3	42	78	6084	3276	
4	45	80	6400	3600	
5	50	90	8100	4500	
6	55	100	10000	5500	
7	56	110	12100	6160	
8	60	115	13225	6900	
9	64	115	13225	7360	
10	65	120	14400	7800	
11	65	125	15625	8125	
12	542	1008	104784	56221	
13					
14					
15	10	1008		542	
16	1008	104784		56221	
17					
18	3,297583	-0,03172		3,844411	$\hat{b}_0$
19	-0,03172	0,000315		0,499559	b1
20					
21	y=3,8+0,5x				

Рисунок 3.3 - Приклад зображення відповідних обчислень

У комірку A18 введемо функцію **МОБР** (Категорія: Математичні). Ця функція обчислює матрицю обернену до даної. Як аргумент функції задамо масив комірок A15: B16 (рисунок 3.4) і натиснемо ОК.

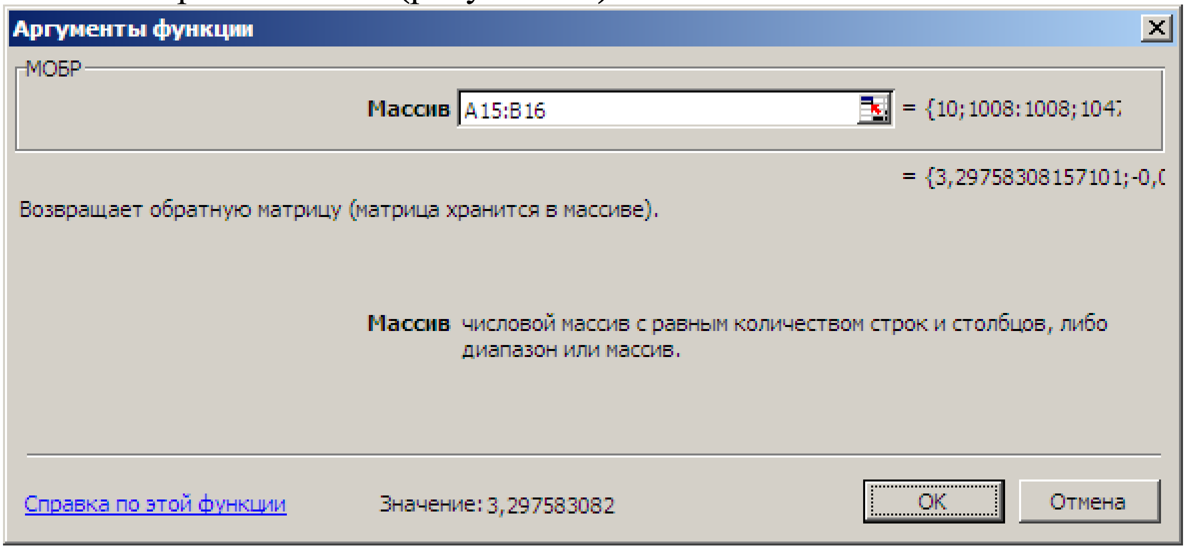


Рисунок 3.4 – Відповідне вікно функції **Аргументы функции**

У результаті в комірці A18 буде отримано значення 3,297583. Оскільки результатом повинна бути матриця розмірністю 2x2, то необхідно здійснити наступні додаткові дії. Виділяємо діапазон комірок A18: B19, на клавіатурі натискаємо клавішу F2 (при цьому в комірці A18 з'явиться введена раніше формула). Далі натискається комбінація трьох клавіш: **Ctrl + Shift + Enter** (утримуючи клавіші Ctrl і Shift, останньою натискається клавіша Enter). У результаті в комірках A18: B19 з'явиться матриця обернена даній.

Для отримання остаточного розв'язку необхідно отриману матрицю помножити на вектор-стовпець правих частин рівнянь системи. Для цього скористаємося функцією **МУМНОЖ** (Категорія: Математичні). Введемо її в клітинку D18, задамо аргументи (при цьому першим масивом повинна бути обов'язково зворотна матриця) і натиснемо ОК (рисунок 3.5).

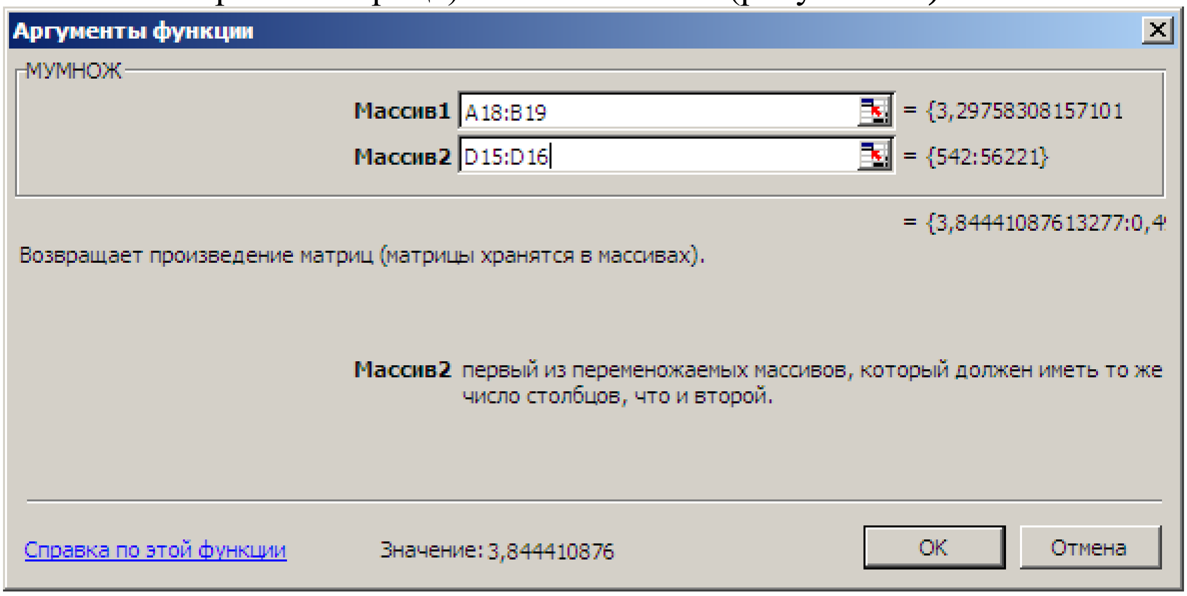


Рисунок 3.5 – Відповідне вікно функції **Аргументы функции**

У результаті в комірці D8 буде отримано значення 3,844411.

Оскільки рішення системи має складаються з двох значень (значення b_0 і b_1), то необхідно здійснити наступні додаткові дії. Виділяємо діапазон комірок D18: D19, натискаємо клавішу F2. Далі натискається комбінація трьох клавіш: Ctrl + Shift + Enter. У результаті в комірках D18: D19 виходить остаточне рішення: $\hat{b}_0 = 3{,}8$; $\hat{b}_1 = 0{,}5$.

Економетрична модель має вигляд

$$y = 3{,}8 + 0{,}5x + \varepsilon$$

Для побудови графіка і оцінки якості побудованого рівняння регресії необхідно за побудованим рівнянням розрахувати теоретичні значення результативної змінної $\hat{y}$ (Стовпець F). У комірці F2 вводимо формулу: = $ D $ 18 + $ D $ 19 * B2 (знаки $ відображають, те що дана клітинка є зафіксованою; фіксування здійснюється за допомогою клавіші F4).

Дана формула поширюється на нижче лежачі комірки шляхом протягування за допомогою миші.

Для побудови графіка необхідно викликати Майстер діаграм, наприклад через головне меню: Вставка / Діаграма ... Тип і вид діаграми вибирається відповідно до рисунка 3.6.

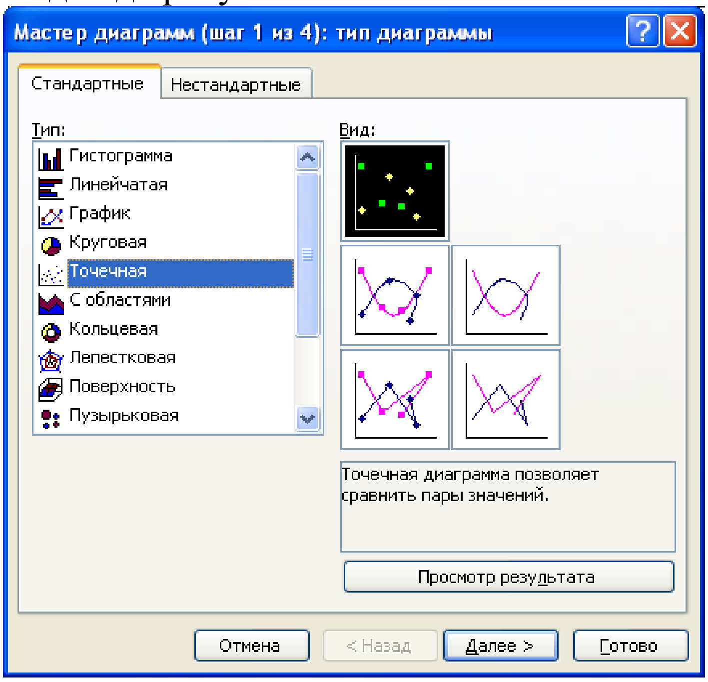


Рисунок 3.6 – Вікно **Мастер диаграмм** при виборі типу діаграми

Натискається кнопка Далее. У вікні вибирається вкладка Ряд. Натискається кнопка Добавить. У полі, що з'явилось, вводяться посилання на значення фактора x і результативної ознаки y (рисунок 3.7).

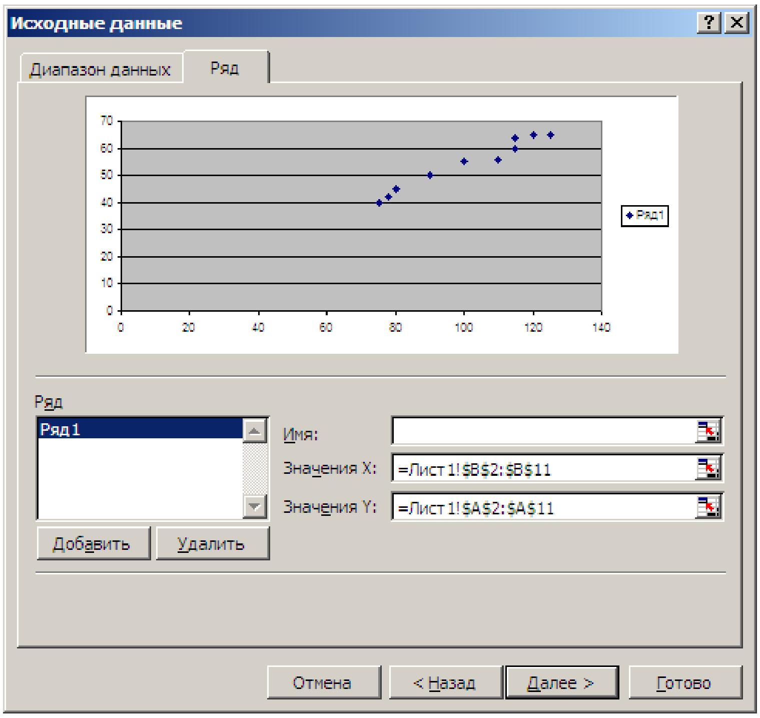


Рисунок 3.7 - Вікно **Мастер диаграмм** при введенні вихідних даних

У вікні перегляду з'являється точковий графік, що відображає вихідні дані - Поле кореляції. Щоб отримати на одному графіку і поле кореляції і лінію регресії необхідно натиснути кнопку Добавить та у відповідні поля занести посилання на клітинки з даними для побудови лінії регресії (рисунок 3.8). Натиснути Далее.

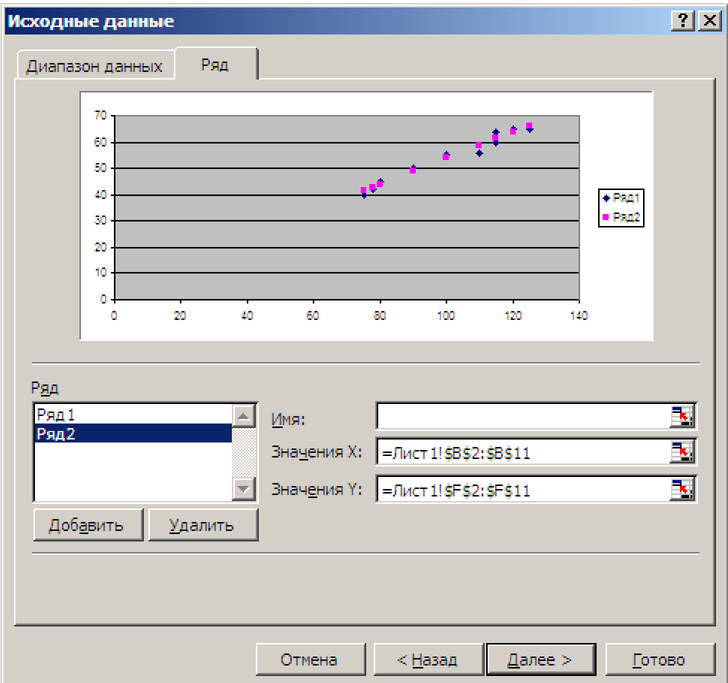


Рисунок 3.8 - Вікно **Мастер диаграмм** при введенні декількох рядів

Значення фактора беруться ті ж самі, що і при побудові поля кореляції, а значення результативного ознаки беруться не вихідні, а розрахункові $\hat{y}$. Далі натискається кнопку Готово.

Для форматування графіка або його окремих компонентів необхідно зробити подвійне клацання мишею по елементу, що цікавить. Наприклад, можна лінію регресії зробити дійсно лінією, а не крапками. Для цього робиться подвійне клацання мишею на рожевих точках і у вікні вибирається тип, колір і товщина лінії, прибираються маркери. Аналогічно можна відформатувати осі графіка, наприклад, нанести назви, змінити діапазон зміни і так далі.

РІЗНІ ВИДИ РІВНЯНЬ РЕГРЕСІЇ

1. Лінійне рівняння регресії

$$\overset{\frown}{y} = b_0 + b_1 \cdot x$$

Система нормальних рівнянь у загальному вигляді:

$$\begin{cases} b_0 \cdot n + b_1 \cdot \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i; \\ b_0 \cdot \sum_{i=1}^{n} x_i + b_1 \cdot \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i \cdot y_i \end{cases}$$

2. Гіперболічне рівняння регресії

$$\overset{\frown}{y} = b_0 + \frac{b_1}{x}$$

Система нормальних рівнянь у загальному вигляді:

$$\begin{cases} b_0 \cdot n + b_1 \cdot \sum_{i=1}^{n} \frac{1}{x_i} = \sum_{i=1}^{n} y_i; \\ b_0 \cdot \sum_{i=1}^{n} \frac{1}{x_i} + b_1 \cdot \sum_{i=1}^{n} \frac{1}{x_i^2} = \sum_{i=1}^{n} \frac{y_i}{x_i} \end{cases}$$

3. Рівняння регресії степенної функції

$$\overset{\frown}{y} = b_0 \cdot x^{b_1}$$

Система нормальних рівнянь у загальному вигляді:

$$\begin{cases} n \cdot \lg b_0 + b_1 \cdot \sum_{i=1}^{n} \lg x_i = \sum_{i=1}^{n} \lg y_i; \\ \lg b_0 \cdot \sum_{i=1}^{n} \lg x_i + b_1 \cdot \sum_{i=1}^{n} (\lg x_i)^2 = \sum_{i=1}^{n} \lg x_i \cdot \lg y_i \end{cases}$$

4. Показательное уравнение регрессии

$$\overset{\frown}{y} = b_0 \cdot b_1^x$$

Система нормальних рівнянь у загальному вигляді::

$$\begin{cases} n \cdot \lg b_0 + \lg b_1 \cdot \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} \lg y_i; \\ \lg b_0 \cdot \sum_{i=1}^{n} x_i + \lg b_1 \cdot \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i \cdot \lg y_i \end{cases}$$

5. Параболічне рівняння регресії

$$\widehat{y} = b_0 + b_1 \cdot x + b_2 \cdot x^2$$

Система нормальних рівнянь у загальному вигляді:

$$\begin{cases} b_0 \cdot n + b_1 \cdot \sum_{i=1}^{n} x_i + b_2 \cdot \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} y_i; \\ b_0 \cdot \sum_{i=1}^{n} x_i + b_1 \cdot \sum_{i=1}^{n} x_i^2 + b_2 \cdot \sum_{i=1}^{n} x_i^3 = \sum_{i=1}^{n} x_i \cdot y_i \\ b_0 \cdot \sum_{i=1}^{n} x_i^2 + b_1 \cdot \sum_{i=1}^{n} x_i^3 + b_2 \cdot \sum_{i=1}^{n} x_i^4 = \sum_{i=1}^{n} x_i^2 \cdot y_i \end{cases}$$

Перевірити значущість рівняння регресії необхідно за допомогою обчислення відповідних показників. Зокрема, треба обчислити

- *лінійний коефіцієнт парної кореляції* r_{xy} $(-1 \le r_{xy} \le 1)$:

$$r_{xy} = \frac{\frac{1}{n}\sum_i (x_i - \overline{x})(y_i - \overline{y})}{\sigma_x \sigma_y} = \frac{\overline{yx} - \overline{y} \cdot \overline{x}}{\sqrt{(\overline{x^2} - \overline{x}^2) \cdot (\overline{y^2} - \overline{y}^2)}}.$$

- *коефіцієнт детермінації* $(0 \le R^2 \le 1)$:

$$R^2 = r^2 \text{ або } R^2 = \frac{\sum_{i=1}^{n} (\widehat{y}_i - \overline{y})^2}{\sum_{i=1}^{n} (y_i - \overline{y})^2} = 1 - \frac{\sum_{i=1}^{n} (y_i - \widehat{y})^2}{\sum_{i=1}^{n} (y_i - \overline{y})^2}.$$

- *середню похибку апроксимації*:

$$\overline{A} = \frac{1}{n}\sum \left| \frac{y_i - \widehat{y}_i}{y_i} \right| \cdot 100\%.$$

- *F-критерій Фішера:*

$$F_{факт} = \frac{\sum \frac{(\widehat{y}_i - \overline{y})^2}{m}}{\sum \frac{(y_i - \widehat{y}_i)^2}{n-m-1}} = \frac{r_{xy}^2}{1 - r_{xy}^2} \cdot \frac{n-m-1}{m},$$

де n - кількість одиниць сукупності, m - кількість параметрів при змінних (для лінійної регресії m =1).

- *t-критерій Стьюдента*:

$$t_{b_0}=\frac{b_0}{m_{b_b}};\ t_{b_1}=\frac{b_1}{m_{b_1}};\ t_r=\frac{r_{xy}}{m_{r_{xy}}},$$

де $m_{b_1}=\sqrt{\frac{\frac{\sum(y_i-\hat{y}_i)^2}{n-2}}{\sum(x_i-\bar{x})^2}}$; $m_{b_0}=\sqrt{\frac{\sum(y_i-\hat{y}_i)^2}{n-2}\cdot\frac{\sum x_i^{\ 2}}{n\sum(x_i-\bar{x})^2}}$; $m_{r_{xy}}=\sqrt{\frac{1-r_{xy}^2}{n-2}}$.

- *довірчі інтервали:*

$\gamma_{b_0}=b_0\pm\Delta_{b_0}$; $\gamma_{b_1}=b_1\pm\Delta_{b_1}$,

де $\Delta_{b_0}=t_{табл}\cdot m_{b_0}$; $\Delta_{b_1}=t_{табл}\cdot m_{b_1}$.

Примітка. Обчислити всі ці показники можна аналогічно використовуючи можливості табличного процесору Excel. Після обчислення необхідно зробити висновки.

3.2 ПОБУДОВА БАГАТОФАКТОРНОЇ ЕКОНОМЕТРИЧНОЇ МОДЕЛІ ЗА ДОПОМОГОЮ МОЖЛИВОСТЕЙ ТАБЛИЧНОГО ПРОЦЕСОРА EXCEL

Приклад. Необхідно провести дослідження залежності ціни автомобіля (Y) від таких характеристик як вік авто (X_1) та його об'єм двигуна (X_2) на основі вибіркових даних, наведених в таблиці 3.2

Таблиця 3.2 – Вихідні дані

№	Ціна автомобіля (у, тис. у.о.)	Вік автомобіля (x_1, років)	Об'єм двигуна автомобіля (x_2, $дм^3$)
1.	11	6	3
2.	3,2	10	1,3
3.	8,7	10	1,8
4.	1,6	16	1,8
5.	17	2	2,4
6.	18,9	4	4
7.	15,8	3	2
8.	18	6	3
9.	19	3	4
10.	6	6	1,6
11.	13	6	2,6
12.	22,7	5	2,1
13.	13,9	3	2
14.	10,9	3	1,5
15.	9,5	8	2,6
16.	28	3	2,2
17.	6	8	1,8
18.	5,8	9	1,6
19.	2,8	11	1,6
20.	26	1	2,2
21.	10,5	6	2
22.	16,9	7	2,3
23.	3,4	9	1,5
24.	14	1	1,4
25.	23	1	1,8
Середнє	13,024	5,88	2,164

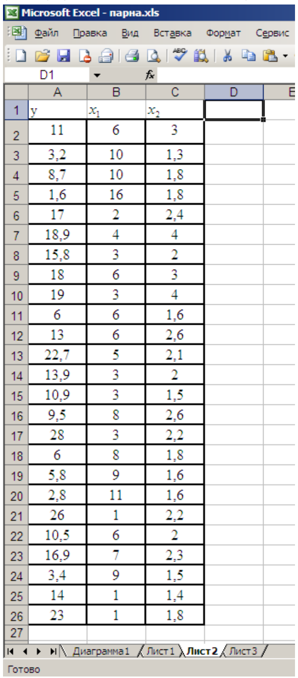
Припустимо, що між ціною та означеними технічними характеристиками існує лінійна залежність:

$$Y = b_0 + b_1 X_1 + b_2 X_2 + \varepsilon$$

Необхідно:

3) визначити $\widehat{b_i}$ для залежності між досліджуваним фактором y (ціною автомобіля) та пояснюючими змінними x_1 (вік автомобіля) і x_2 (об'єм двигуна);
4) проаналізувати ступінь адекватності побудованої моделі та вибіркових даних.

Розв'язання. Введить вихідні дані так як це показано на рисунку 3.9.



y	x_1	x_2
11	6	3
3,2	10	1,3
8,7	10	1,8
1,6	16	1,8
17	2	2,4
18,9	4	4
15,8	3	2
18	6	3
19	3	4
6	6	1,6
13	6	2,6
22,7	5	2,1
13,9	3	2
10,9	3	1,5
9,5	8	2,6
28	3	2,2
6	8	1,8
5,8	9	1,6
2,8	11	1,6
26	1	2,2
10,5	6	2
16,9	7	2,3
3,4	9	1,5
14	1	1,4
23	1	1,8

Рисунок 3.9 – Приклад введення даних

У головному меню послідовно виберіть **Сервис/Надстройки**. Встановість позначку (галочку) навпроти **Пакет анализа** (рисунок 3.10), після чого клацніть на **ОК**.

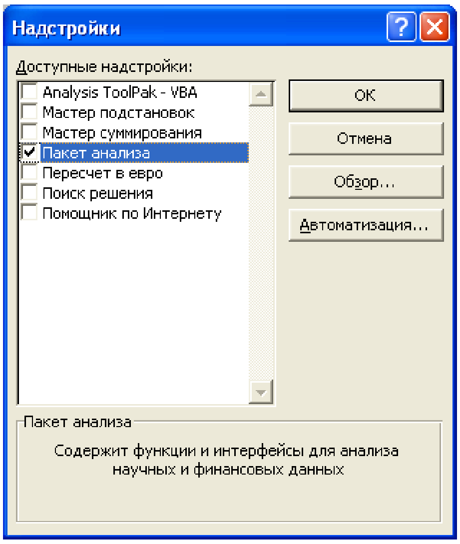


Рисунок 3.10 – Вікно **Надстойки**

У головному меню виберіть **Сервис/Анализ данных/Описательная статистика** (рисунок 3.11), після чого клацніть **ОК.**

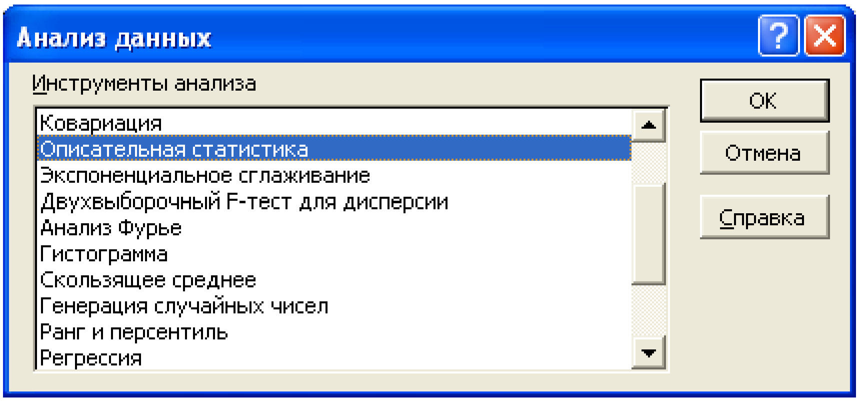


Рисунок 3.11 – Вікно **Анализ данных** при виборі функції Описательная статистика

Заповність діалогове вікно вводу даних та параметрів вводу (рисунок 3.12).

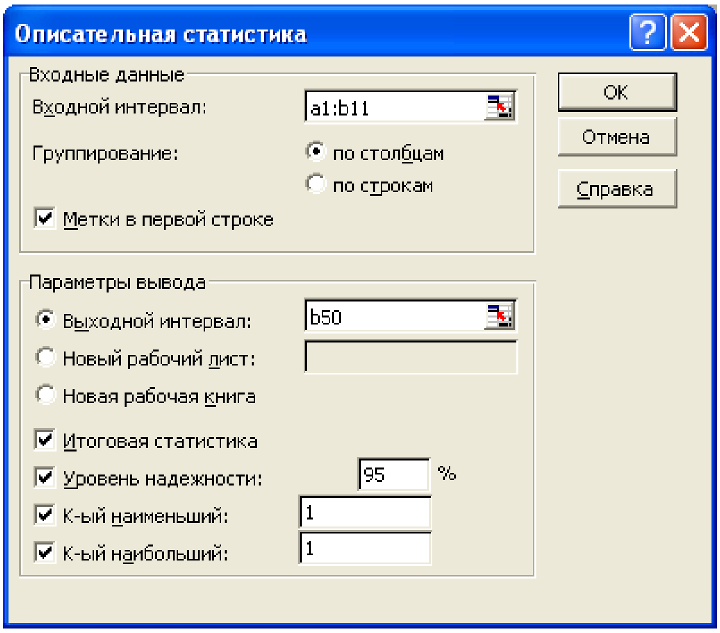


Рисунок 3.12 – Вікно меню **Описательная статистика**

Входной интервал – це діапазон , який містить дані, що аналізуються (це може бути один або декілька стовпців, рядків);

Группирование – за стовпцями або за рядками (треба вказати додатково);

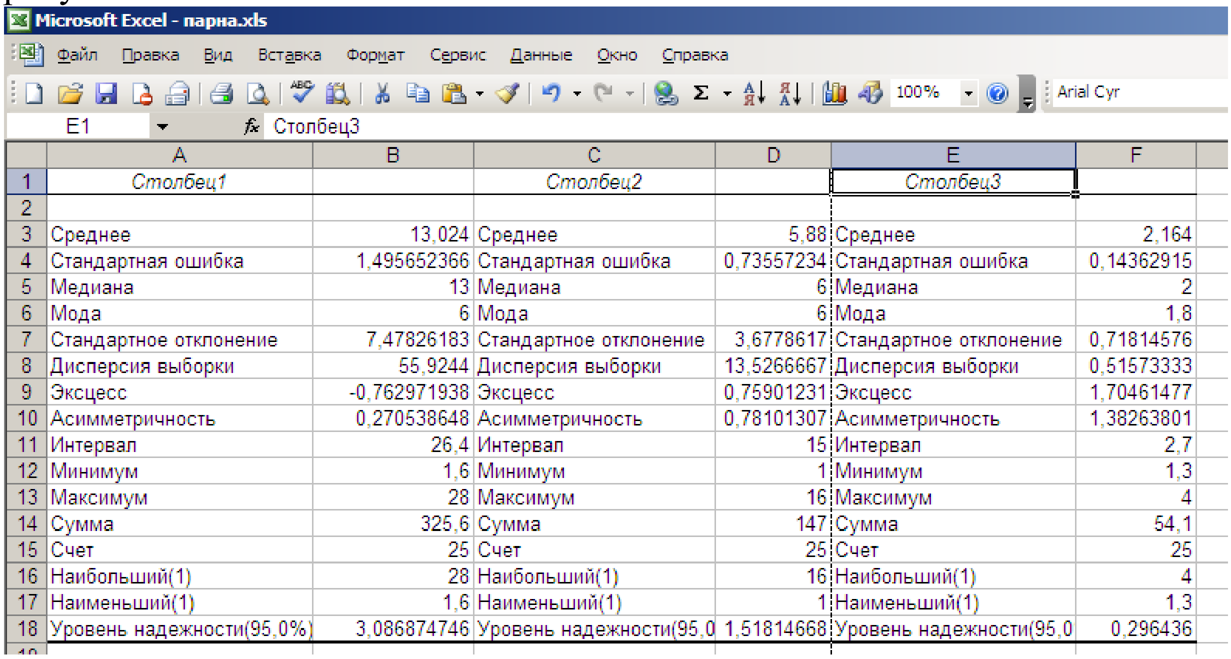
Метки – флажок, який вказує, чи містить перший рядок назву стовпців чи ні;

Выходной интервал – достатньо вказати ліву верхню комірку майбутнього дапазону;

Новый рабочий лист – можно задати довільне ім’я нового листка.

Якщо необхідно отримати додаткову інформацію *Итоговой статистики, Уровня надежности, к - го наибольшего и к - го наименьшего значений,* встановить відповідні позначки у діалоговому вікні, після чого клацніть **ОК**.

Результати обчислень відповідних показників представлені на рисунке 3.13.



	A	B	C	D	E	F
1	Столбец1		Столбец2		Столбец3	
2						
3	Среднее	13,024	Среднее	5,88	Среднее	2,164
4	Стандартная ошибка	1,495652366	Стандартная ошибка	0,73557234	Стандартная ошибка	0,14362915
5	Медиана	13	Медиана	6	Медиана	2
6	Мода	6	Мода	6	Мода	1,8
7	Стандартное отклонение	7,47826183	Стандартное отклонение	3,6778617	Стандартное отклонение	0,71814576
8	Дисперсия выборки	55,9244	Дисперсия выборки	13,5266667	Дисперсия выборки	0,51573333
9	Эксцесс	-0,762971938	Эксцесс	0,75901231	Эксцесс	1,70461477
10	Асимметричность	0,270538648	Асимметричность	0,78101307	Асимметричность	1,38263801
11	Интервал	26,4	Интервал	15	Интервал	2,7
12	Минимум	1,6	Минимум	1	Минимум	1,3
13	Максимум	28	Максимум	16	Максимум	4
14	Сумма	325,6	Сумма	147	Сумма	54,1
15	Счет	25	Счет	25	Счет	25
16	Наибольший(1)	28	Наибольший(1)	16	Наибольший(1)	4
17	Наименьший(1)	1,6	Наименьший(1)	1	Наименьший(1)	1,3
18	Уровень надежности(95,0%)	3,086874746	Уровень надежности(95,0	1,51814668	Уровень надежности(95,0	0,296436

Рисунок 3.13 – Приклад зображення обчислень

За даними рисунка 3.13 знаходимо, що:

1) вибіркове середнє значення y дорівнює 13,024, x_1 - 5,88, x_2 - 2,164;
2) середнє квадратичне відхилення y дорівнює 7,48, x_1 - 3,68, x_2 - 0,72;
3) дисперсія y дорівнює 55,9, x_1 - 13,5, x_2 - 0,52;
4) і т.д.

У головному меню виберіть **Сервис/Анализ данных/Регрессия** (рисунок 3.14), після цього клацніть **ОК**.

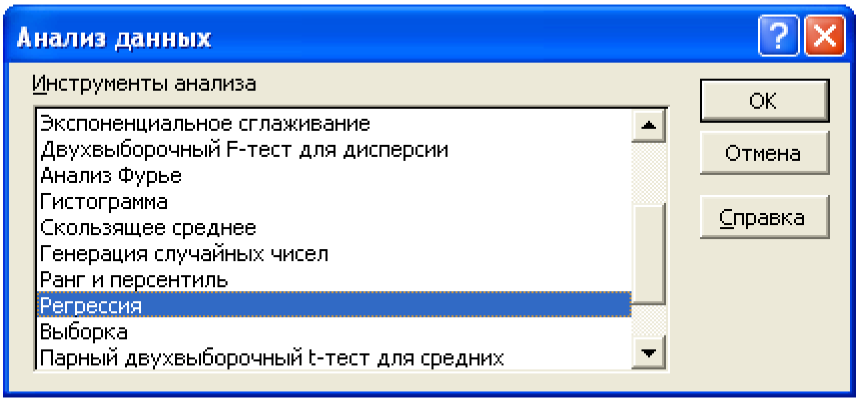


Рисунок 3.14 - Вікно **Анализ данных** при виборі функції Регресія

Заповніть діалогове вікно вводу параметрів інструменту «Регрессия» (рисунок 3.15).

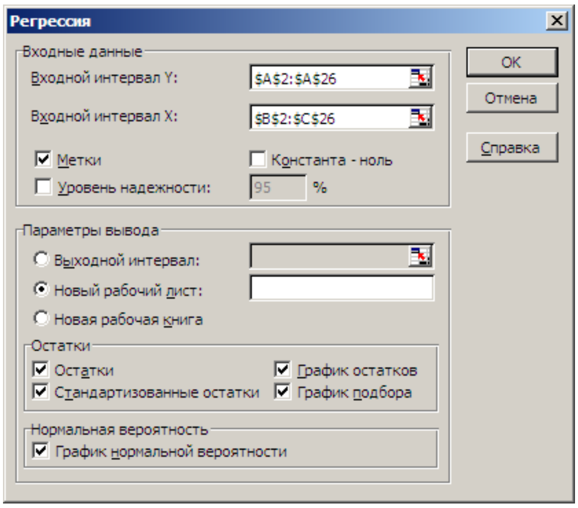


Рисунок 3.15 – Вікно налагодження функції **Регрессия**

Входной интервал Y – це діапазон, який містить дані результативної ознаки, що аналізуються;

Входной интервал X– це діапазон, який містить дані факторів незалежної ознаки;

Метки – флажок, який вказує, чи містить перший рядок назву стовпців або ні;

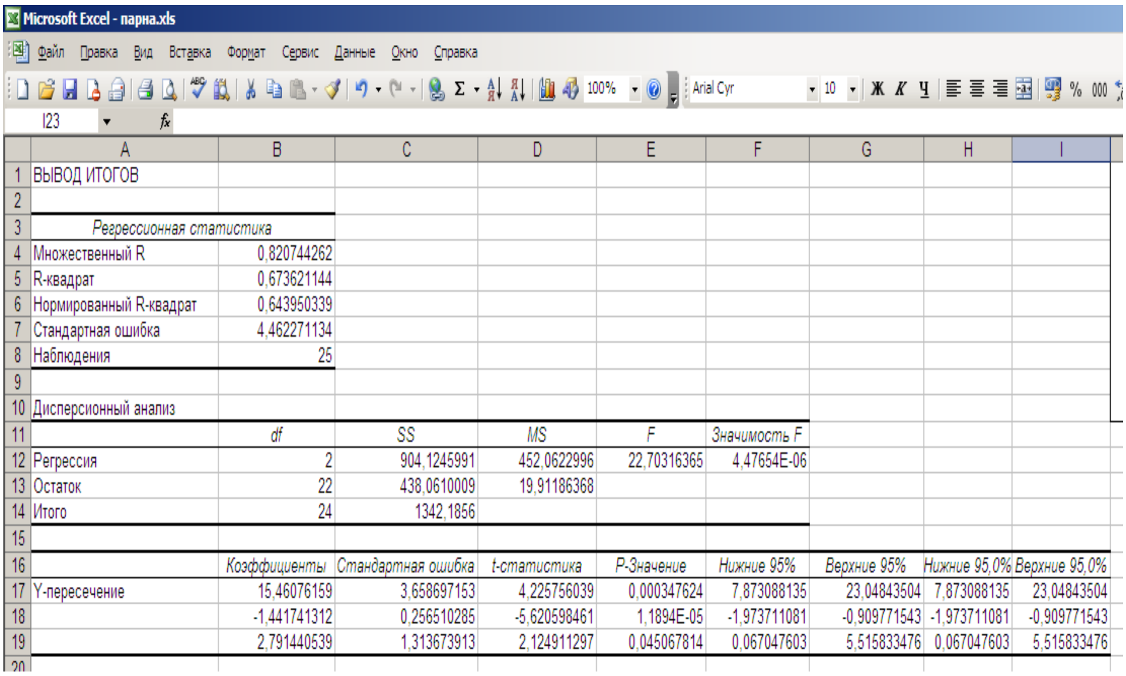
Константа – нуль – флажок, який вказує наявність або відсутність вільного члена у рівнянні;

Выходной интервал – тут досить вказати ліву верхню комірку майбутнього діапазону;

Новый рабочий лист – можна задати довільне ім'я нового листка.

Якщо необхідно отримати додаткову інформацію за залишками і графіками залишкі, то встановіть відповідні флажки в діалоговому вікні, після чого клацніть на кнопці **ОК**.

Результати регресійного аналізу представлені на рисунках 3.16, 3.17, 3.18.



ВЫВОД ИТОГОВ								
Регрессионная статистика								
Множественный R	0,820744262							
R-квадрат	0,673621144							
Нормированный R-квадрат	0,643950339							
Стандартная ошибка	4,462271134							
Наблюдения	25							
Дисперсионный анализ								
	df	SS	MS	F	Значимость F			
Регрессия	2	904,1245991	452,0622996	22,70316365	4,47654E-06			
Остаток	22	438,0610009	19,91186368					
Итого	24	1342,1856						
	Коэффициенты	Стандартная ошибка	t-статистика	P-Значение	Нижние 95%	Верхние 95%	Нижние 95,0%	Верхние 95,0%
Y-пересечение	15,46076159	3,658697153	4,225756039	0,000347624	7,873088135	23,04843504	7,873088135	23,04843504
	-1,441741312	0,256510285	-5,620598461	1,1894E-05	-1,973711081	-0,909771543	-1,973711081	-0,909771543
	2,791440539	1,313673913	2,124911297	0,045067814	0,067047603	5,515833476	0,067047603	5,515833476

Рисунок 3.16 – Приклад зображення виводу регресійного аналізу

За рисунком 3.16 робимо такі висновки:

Регресійна статистика («Регрессионная статистика»)

a) R = 0,821 – коефіцієнт кореляції;

b) R^2 = 0,674 – не скоригований коефіцієнт детермінації (без урахування кількості ступенів вільності) він оцінює долю варіації результуючої Y за рахунок введення в модель факторів. У нашому прикладі ці доля складає 67,4% . Це свідчить про досить високий ступінь обумовленості варіації результуючої від варіації факторів (тобто, це свідчить про досить тісний зв’язок фактору та результуючої);

c) Нормований $\overline{R^2}$ = 0,64 – скоректований коефіцієнт детермінації, що визначає тісноту зв’язку з урахуванням кількості ступенів вільності загальної та залишкової дисперсій. Він дає оцінку, яка не залежить від кількості незалежних змінних (факторів) моделі, і тому може порівнюватись з різними моделями з різною кількості факторів.

Зв’язок між не скоригованим та скоригованим коефіцієнтами детермінації визначається за формулою:

$$\overline{R^2} = 1 - (1 - R^2)\frac{(n-1)}{(n-p-1)},$$

d) де n – кількість спостережень, p - кількість незалежних змінних.

Обидва коефіцієнта (R^2, $\overline{R^2}$) вказують на досить високу детермінованість результуючої Y у моделі із заданими факторами x_1, x_2.

Стандартна похибка залишків s_e = 4,46.

Дисперсійний аналіз («Дисперсионный аналіз») включає п’ять стовпців:

1) ***(df)*** – *ступені вільності*, тобто кількість свободи незалежної варійованої ознаки. Кількість ступенів вільності пов’язана із кількістю одиниць сукупності n і кількістю оцінок, які за нею визначаються:

для регресії $df = p =$; для залишку $df = n - (p+1) = 22$; загальна $df = n - 1 = 24$.

2) ***(SS)*** – *суми квадратів відхилень*:

$\sum_{i=1}^{n}\left(Y_i - \overline{Y}\right)^2$ =904,125 - регресія (пояснена);

$\sum_{i=1}^{n}\left(Y_i - \widehat{Y}_i\right)^2 = 438{,}061$ - залишкова (не пояснена);

$\sum_{i=1}^{n}\left(Y - \overline{Y}\right)^2 = 1342{,}1856$ - загальна (залежної змінної).

3) ***(MS)*** – дисперсія на одну ступінь вільності:

$$S_p^2 = \frac{\sum_{i=1}^{n}\left(\hat{Y}_i - \overline{Y}\right)^2}{p} = 452{,}062$$ - регресій (пояснена);

$$S_e^2 = \frac{\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2}{n-(p+1)} = 19{,}912$$ - залишків (не пояснена);

4) **(F)** – фактичне значення F-критерія:

$$F = \frac{S_p^2}{S_e^2} = \frac{\dfrac{\sum\left(\hat{Y}_i - \overline{Y}\right)^2}{p}}{\dfrac{\sum\left(Y_i - \hat{Y}_i\right)^2}{n-(p+1)}} = \frac{\dfrac{\sum\left(\hat{Y}_i - \overline{Y}\right)^2}{p\sum\left(Y_i - \overline{Y}\right)^2}}{\dfrac{\sum\left(Y_i - \hat{Y}_i\right)^2}{n-(p+1)\sum\left(Y_i - \overline{Y}\right)^2}} = \frac{\dfrac{R^2}{p}}{\dfrac{1-R^2}{n-(p+1)}} = \frac{R^2\left[n-(p+1)\right]}{\left(1-R^2\right)p} = 22{,}7;$$

5) **(Значущість F)** – рівень значущості F-критерія α = 4,47654E-06 = 0,00000447.

Коефіцієнт детермінації R^2 дорівнює 0,6736, що свідчить про досить сильну залежність між незалежними змінними та результуючою. Можна використати F-статистику, для того щоб визначити чи є цей результат випадковим (саме значення R^2). Величина α застосовується для позначення ймовірності помилкового висновку про те, що має місце досить сильний зв'язок.

Припустимо, що насправді взаємозв'язок між змінними відсутній. Ми просто обрали дуже рідкісні спостереження, для яких статистичний аналіз показав сильний взаємозв'язок.

Оцінку надійності рівняння регресії у цілому і коефіцієнта детермінації дає F-критерій Фішера. Його розрахункове (фактичне) значення F = 22,7 порівнюється із табличним, яке при рівні значущості α =0,05 та при кількості ступенів вільності $v_1 = p = 2$, $v_2 = n-(p+1) = 22$ становить 3,44. Якщо фактичне значення перевищує табличне, то з ймовірністю 0,95 гіпотеза про ненадійність рівняння відхиляється і стверджується статистична значущість моделі та коефіцієнта детермінації. Або, що те саме, з таблиці дисперсійного аналізу слідує, що ймовірність випадково отримати таке значення F-критерія (F = 22,7) становить 0,0000447, що не перевищує гранично допустимого значення значущості α = 0,05. Отже, отримане значення є невипадковим, воно сформувалось під впливом суттєвих факторів, тобто, статистична значущість всієї моделі та коефіцієнта детермінації підтверджується.

Оцінка параметрів моделі та їх значущості («Оценка параметров модели и их значимость»)

Цей блок результатів містить 9 стовпців.

Перший та другий – назва та величина оцінок параметрів моделі: *Y – пересечение* – b_0 = 15,46076159; змінна x_1 – b_1 =-1,441741312; змінна x_2 - b_2 = 2,791440539.

Таким чином, отримали модель: $\hat{y} = 15,4608 - 1,4417x_1 + 2,7914x_2$.

Величини b_j $(j = 1,2)$ показують, наскільки зміниться результуюча, якщо збільшити значення деякого фактора на одиницю (при незмінній величині решти факторів). Наприклад,

- збільшення x_1 (вік автомобіля) на 1 рік, при незмінних інших факторах, буде сприяти зменшенню y (ціни автомобіля) на 1,44 (тис. у.о.);
- збільшення x_2 (об'єм двигуна автомобіля) на 1 $дм^3$, а інші фактори незмінні, буде сприяти збільшенню y (ціни автомобіля) на 2,89 (тис. у.о.).

Порівнювати силу впливу окремих факторів на результуючу співставляючі величини коефіцієнтів не варто, оскільки ці коефіцієнти залежать від власних одиниць вимірювання. З метою виявлення найбільш впливових показників (факторів) треба перейти до стандартизованого (нормалізованого) рівняння, у якому одиницею вимірювання впливу всіх факторів є середнє квадратичне відхилення.

Третій стовпчик містить стандартні похибки оцінок параметрів моделі:

S_{b_0} = 3,658697153; S_{b_1} = 0,256510285; S_{b_2} = 1,313673913.

Ці показники показують, яка частка значення даної характеристики сформувалася під впливом випадкових факторів. Стандартні похибки використовуються для розрахунку t-критерію Стьюдента, значення якого для різних оцінок представлені в четвертому стовпці.

Четвертий стовпчик – t-критерій:

t_0 = 4,225756039; t_1 = -5,620593461; t_2 = 2,124911297.

Якщо значення t-критерія більше 2–3, то можна зробити висновок про значущість даного параметра, яки формується під впливом невипадкових величин. Для більш ґрунтовних висновків використаємо результати, які знаходяться у п'ятому стовпчику.

П'ятий стовпчик – рівень значущості – показник ймовірності випадкових значень параметрів регресії:

α_0 = 0,000347624; α_1 = 0,000011894; α_2 = 0,045067814.

Якщо α_j менше за прийнятий нами рівень значущості (зазвичай 0,1; 0,05 або 0,01), то роблять висновок про невипадковість даного значення оцінки, тобто про те, що оцінки параметрів достовірні (статистично значущі). В іншому випадку гіпотеза про випадковість значень коефіцієнтів регресії приймається.

Для нашої задачі α_j менші за прийнятий нами рівень значущості α = 0,05, то ми робимо висновок про те, що оцінки параметрів моделі достовірні.

У випадку, якщо виникає протиріччя, а саме F-критерій стверджує про достовірність моделі у цілому, а t-критерій – недостовірність окремих оцінок, зазвичай дослідження оцінок параметрів моделі продовжується на наявність мультиколінеарності.

Решта чотири стовпчики з ймовірністю 95% визначають верхні та нижні границі оцінок параметрів моделі, тобто дозволяють здійснити інтервальне оцінювання параметрів (довірчі інтервали).

Інтервальне оцінювання параметрів b_j $(j=1,2)$ виконується наступним чином:

$$b_j \pm S_{b_j} \cdot t_\alpha, \qquad \text{або} \qquad b_j - S_{b_j} \cdot t_\alpha \le \beta_j \le b_j + S_{b_j} \cdot t_\alpha.$$

Так, наприклад, b_0 :

для α = 0,05 та 25-(2+1)=22 ступені вільності табличне значення t_α = 2,0739, тоді $S_{b_0} \cdot t_\alpha$ = 3,658697153· 2,0739 = 7,5877720256;

$$15{,}46076159 - 7{,}5877720256 \le b_0 \le 15{,}46076159 + 7{,}5877720256;$$

$$7{,}873088135 \le b_0 \le 23{,}04843504.$$

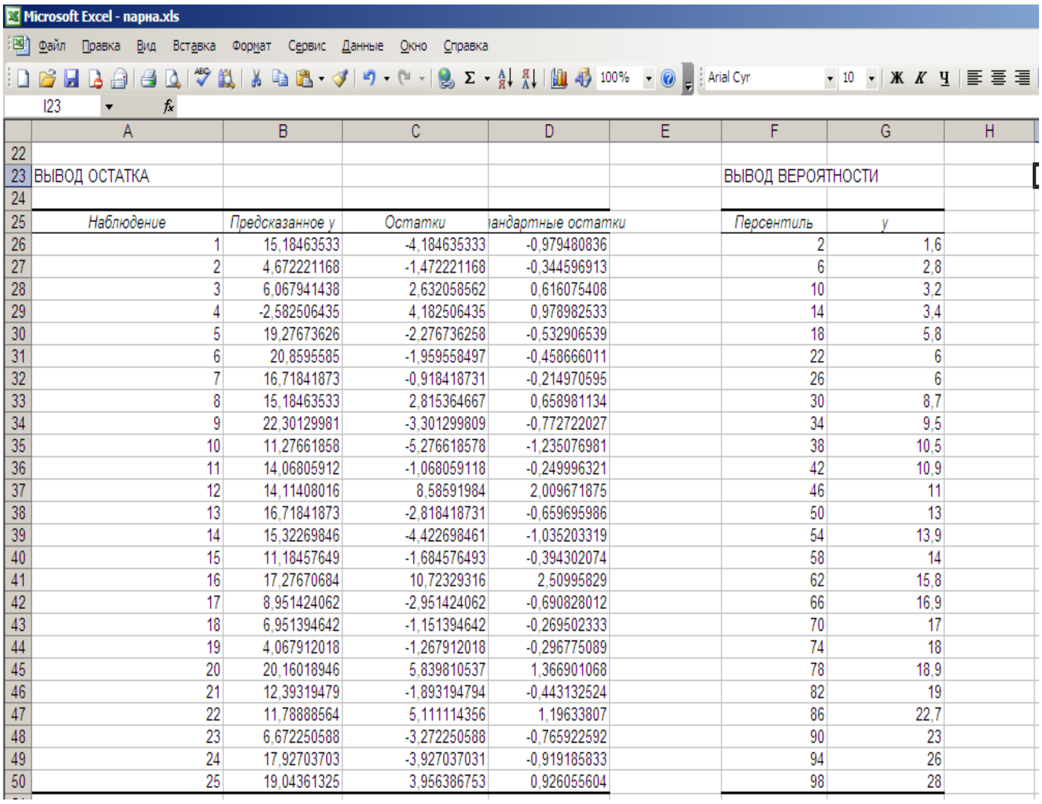
Аналогічно:

$$-1{,}973711081 \le b_1 \le -0{,}909771543$$

$$0{,}067047603 \le b_2 \le 5{,}515833476.$$

Таким чином, з ймовірністю 95%, збільшення віку автомобіля на одну одиницю забезпечить зменшення ціни автомобіля в межах від 1,973711081 до 0,909771543. Аналогічно інтерпретуються решта довірчих інтервалів.

Якщо довірчий інтервал деякої оцінки параметру регресії включає нуль, то це свідчить про недостовірність цього параметру.



Microsoft Excel - парна.xls

Файл Правка Вид Вставка Формат Сервис Данные Окно Справка

	A	B	C	D	E	F	G
23	ВЫВОД ОСТАТКА					ВЫВОД ВЕРОЯТНОСТИ	
25	Наблюдение	Предсказанное y	Остатки	андартные остатки		Персентиль	y
26	1	15,18463533	-4,184635333	-0,979480836		2	1,6
27	2	4,672221168	-1,472221168	-0,344596913		6	2,8
28	3	6,067941438	2,632058562	0,616075408		10	3,2
29	4	-2,582506435	4,182506435	0,978982533		14	3,4
30	5	19,27673626	-2,276736258	-0,532906539		18	5,8
31	6	20,8595585	-1,959558497	-0,458666011		22	6
32	7	16,71841873	-0,918418731	-0,214970595		26	6
33	8	15,18463533	2,815364667	0,658981134		30	8,7
34	9	22,30129981	-3,301299809	-0,772722027		34	9,5
35	10	11,27661858	-5,276618578	-1,235076981		38	10,5
36	11	14,06805912	-1,068059118	-0,249996321		42	10,9
37	12	14,11408016	8,58591984	2,009671875		46	11
38	13	16,71841873	-2,818418731	-0,659695986		50	13
39	14	15,32269846	-4,422698461	-1,035203319		54	13,9
40	15	11,18457649	-1,684576493	-0,394302074		58	14
41	16	17,27670684	10,72329316	2,50995829		62	15,8
42	17	8,951424062	-2,951424062	-0,690828012		66	16,9
43	18	6,951394642	-1,151394642	-0,269502333		70	17
44	19	4,067912018	-1,267912018	-0,296775089		74	18
45	20	20,16018946	5,839810537	1,366901068		78	18,9
46	21	12,39319479	-1,893194794	-0,443132524		82	19
47	22	11,78888564	5,111114356	1,19633807		86	22,7
48	23	6,672250588	-3,272250588	-0,765922592		90	23
49	24	17,92703703	-3,927037031	-0,919185833		94	26
50	25	19,04361325	3,956386753	0,926055604		98	28

Рисунок 3.17 - Приклад зображення виводу регресійного аналізу

На рисунку 3.17 представлена інформація про залишки, а саме

Вивід залишків («Вывод остатков»)

У цьому блоці результатів наводяться розрахункові значення залежної змінної та залишків, які визначаються як різниця між фактичними значеннями та розрахунковими.

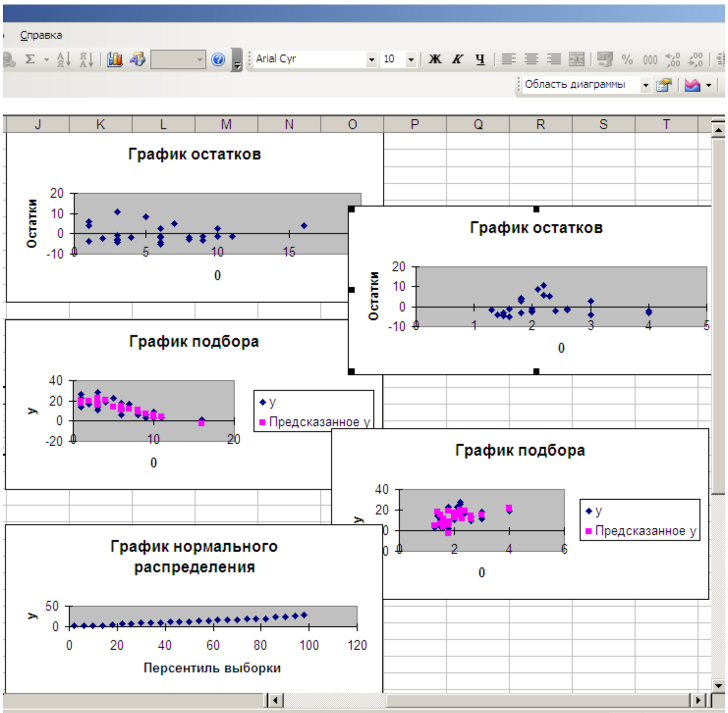


Рисунок 3.18 - Приклад зображення виводу регресійного аналізу

На рисунку 3.18 представлені графіки «остатков», «подбора» и «нормального распределения», тобто функція **Анализ данных/Регрессия** дозволяє вивести графіки залежності результативної ознаки від кожного з факторів, а також графіки залишків для парних залежностей. Для виводу цих графіків необхідно у вікні **Регрессия** встановити відповідні позначки (дивись рисунок 3.15).

У головному меню виберіть **Файл/Сохранить как**. У діалоговому вікні **Сохранение документа**, що з'явилось (рисунок 4.19**)**, вкажіть ім'я файлу (zadanie№…..), вкажіть в якій папці буде зберігатись файл (приміром, STATISTICA), після чого натисніть **Сохранить.**

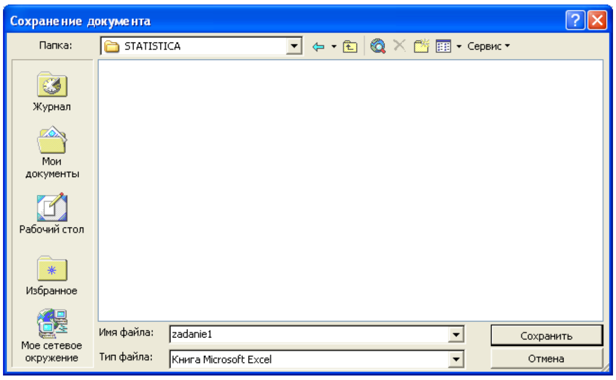


Рисунок 3.19 – Вікно **Сохранение документа**

3.3 ПОБУДОВА БАГАТОФАКТОРНОЇ ЛІНІЙНОЇ ЕКОНОМЕТРИЧНОЇ МОДЕЛІ НА ОСНОВІ МАТРИЧНОГО ОПЕРАТОРА 1МНК

Матричний оператор 1МНК має вигляд:

$$A = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y},$$

де

$$X = \begin{pmatrix} 1 & 6 & 3 \\ 1 & 10 & 1{,}3 \\ 1 & 10 & 1{,}8 \\ 1 & 16 & 1{,}8 \\ 1 & 2 & 2{,}4 \\ 1 & 4 & 4 \\ 1 & 3 & 2 \\ 1 & 6 & 3 \\ 1 & 3 & 4 \\ 1 & 6 & 1{,}6 \\ 1 & 6 & 2{,}6 \\ 1 & 5 & 2{,}1 \\ 1 & 3 & 2 \\ 1 & 3 & 1{,}5 \\ 1 & 8 & 2{,}6 \\ 1 & 3 & 2{,}2 \\ 1 & 8 & 1{,}8 \\ 1 & 9 & 1{,}6 \\ 1 & 11 & 2{,}2 \\ 1 & 1 & 2 \\ 1 & 6 & 2{,}3 \\ 1 & 7 & 2 \\ 1 & 9 & 1{,}5 \\ 1 & 1 & 1{,}4 \\ 1 & 1 & 1{,}8 \end{pmatrix} \qquad Y = \begin{pmatrix} 11 \\ 3{,}2 \\ 8{,}7 \\ 1{,}6 \\ 17 \\ 18{,}9 \\ 15{,}8 \\ 18 \\ 19 \\ 6 \\ 13 \\ 22{,}7 \\ 13{,}9 \\ 10{,}9 \\ 9{,}5 \\ 28 \\ 6 \\ 5{,}8 \\ 2{,}8 \\ 26 \\ 10{,}5 \\ 16{,}9 \\ 3{,}4 \\ 14 \\ 23 \end{pmatrix},$$

$\mathbf{X}'$ - транспонована матриця **X**.

Для транспонування матриці **X** виконаємо наступні дії:

1) виділіть область порожніх комірок, яка складається з (p+1) рядків та n стовпців, для виведення результату (p – кількість незалежних змінних), n – кількість спостережень;
2) активізуйте **Мастер функций** одним із можливих способів:
   a) у головному меню виберіть **Вставка/Функция**;
   b) на панелі інструментів **Стандартная** клацніть на **Вставка функции**;
3) у вікні виберіть Категорію **Ссылки и массивы**, Функцію – **ТРАНСП** (рисунок 3.20) клацніть на кнопці **ОК**;

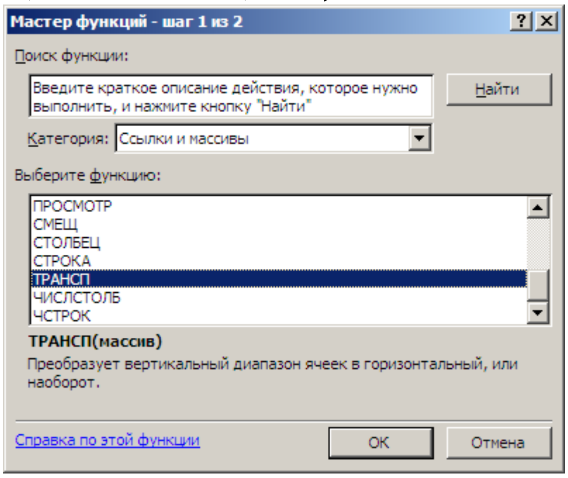


Рисунок 3.20 – Вікно **Мастер функции** при виборі функції **ТРАНСП**

4) у рядку **Массив** вкажіть діапазон комірок, у яких міститься матриця **X**. Клацніть **ОК.**
5) у лівій верхній комірці виділенної області з'явиться перший елемент результуючого масиву даних; для того, щоб вивести всі дані, клацніть <F2>, а потім – комбінацію клавиш <Ctrl>+<Shift>+<Enter>.
   Результат представлено на рисунке 3.21.

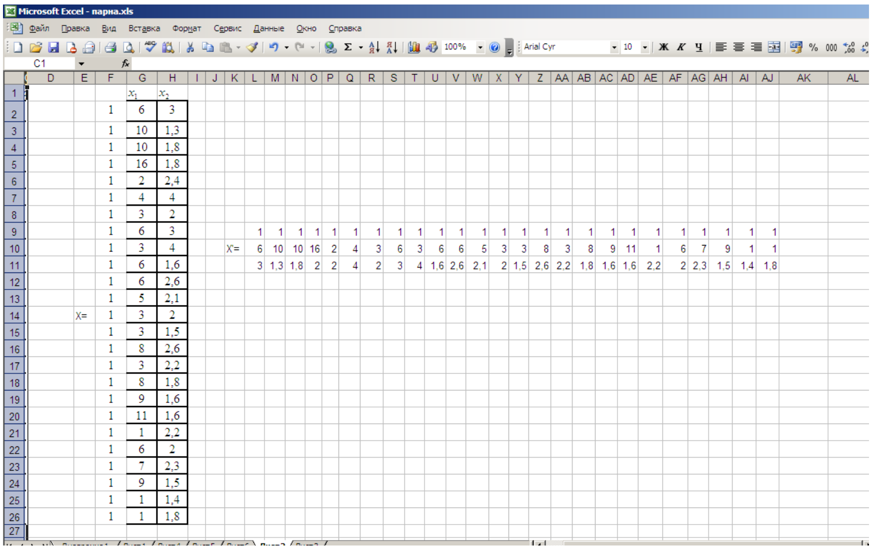


Рисунок 3.21 – Результат транспонування масиву даних

Добуток матриць (**X'X**) знаходимо за допомогою **Мастера функций,** використовуючи Категорію **Статистические,** функцію **МУМНОЖ**:

1) виділіть область порожніх комірок, яка складаєстяь з (р+1) рядків та (р+1) стовпців для виведення результату;
2) у вікні **МУМНОЖ** у рядку **Массив 1** вкажіть діапазон комірок, у яких міститься матриця **X'** (перший співмножник), а у рядку **Массив 2** – матриця **X** (другий співмножник). Клацніть **ОК;**
3) у лівій верхній комірці виділеної області з’явиться перший елемент результуючого масиву даних; для того, щоб вивести всі дані, клацніть <F2>, а потім – комбінацію клавиш <Ctrl>+<Shift>+<Enter>.

Результат множення матриць представлений на рисунке 3.22.

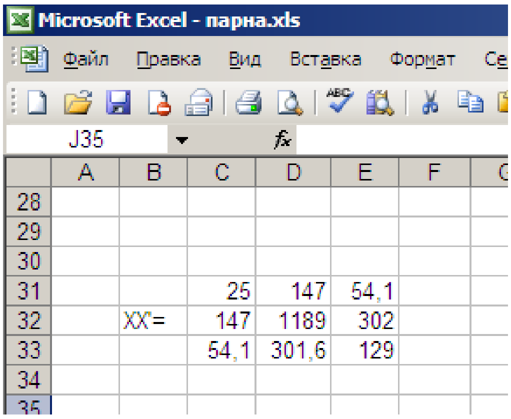


Рисунок 3.22 – Результат множення матриць

Аналогічно знайдемо обернену матрицю $(X'X)^{-1}$ та добутки матриць (**X'Y**), $(\mathbf{X'X})^{-1}$ (**X'Y**) (за допомогою функції МОБР та функції МУМНОЖ), рисунок 3.23:

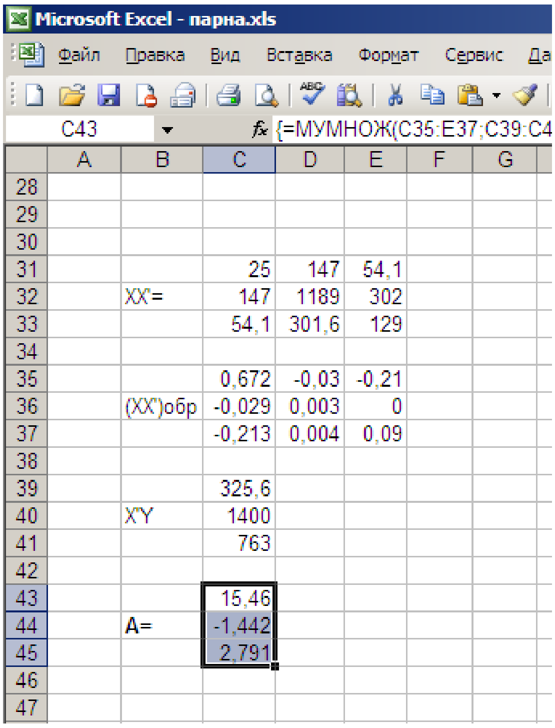


Рисунок 3.23 – Результат добутку відповідних матриць

Таким чином, отримали економетричну модель:

$$\hat{y} = 15{,}46 - 1{,}442x_1 + 2{,}791x_2 \, .$$

Далі, використовуючи відомі вже нам можливості табличного процесора Excel, необхідно обчислити відповідні показники, за допомогою яких здійснюється аналіз ступеня адекватності побудованої моделі.

Табличні значення критеріїв можна знайти у таблицях критичних значень (див. додатки) або можна найти в **Excel.** Наприклад, знайти критичні значення для t-критерія Стьюдента можна звернувшись до **Мастеру функций**, категорія **Статистические,** функція **СТЬЮДРАСПОБР** (рисунок 3.24).

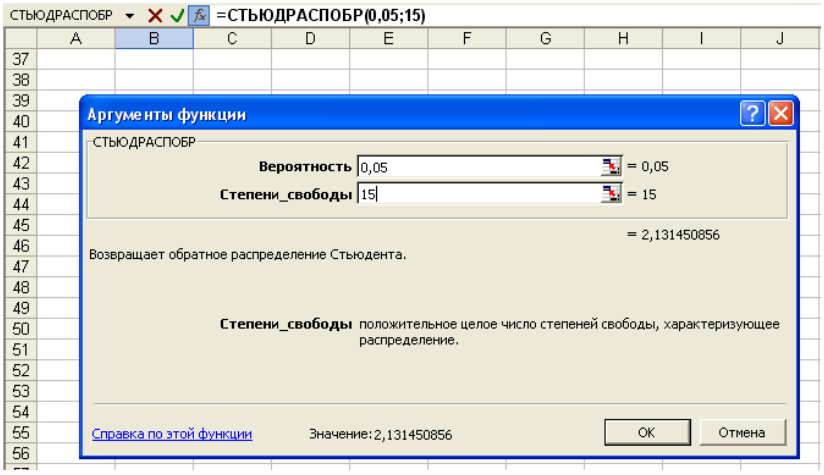


Рисунок 3.24 – Вікно знаходження табличного значення t-критерія Стьюдента

У діалоговому вікні цієї функції необхідно ввести ймовірність на рівні значущості (0,05 або 0,01) та кількість ступенів вільності.

3.4 ВИКОРИСТАННЯ СТАНДАРТНОЇ ПРОГРАМИ «ЛИНЕЙН»

Вбудована статистична функція **ЛИНЕЙН** визначає коефіцієнти лінійної регресії.

Порядок виконання дій:

1) введіть вихідні дані;
2) виділіть область порожніх комірок, яка складається з 5 рядків та (p + 1) стовпців (де p – кількість незалежних змінних) для виведення результатів регресійної статистики;
3) активізуйте **Мастер функций** будь-яким способом:
   a) у головному меню виберіть **Вставка/Функция**;
   b) на панелі інструментів **Стандартная** клацніть на **Вставка функции** $\boxed{f_x}$
4) у вікні, що відкрилось, виберіть Категорію **Статистические**, Функцію –

ЛИНЕЙН. Клацніть **ОК**;

5) заповніть аргументи функції (рисунок 3.25):

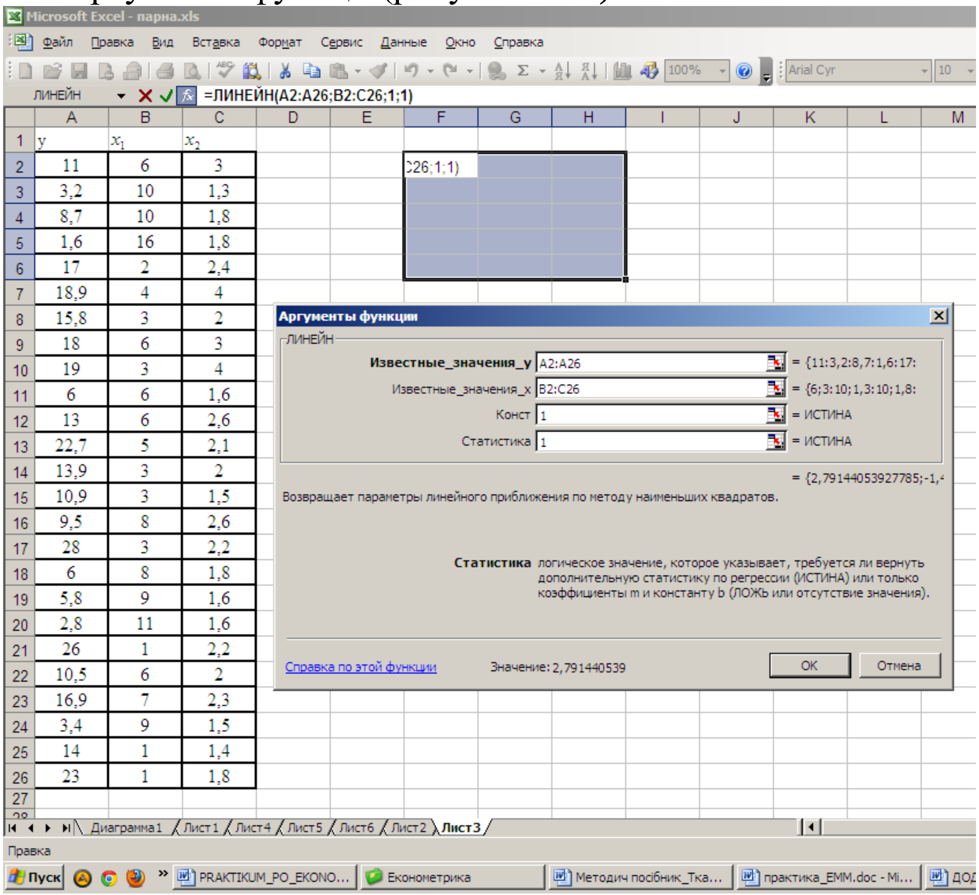


Рисунок 3.25 – Введення даних

- *Известные значения У* – діапазон, якйи містить дані, що характеризують результуючу ознаку;
- *Известные значения Х* – діапазон, який містить дані, що описують всі незалежні змінні;
- *Константа* – логічне значення, яке вказує на наявність або відсутність вільного члена у рівнянні; якщо *Константа* = 1, то вільний член розраховується звичайним чином, якщо *Константа* = 0, то вільний член = 0;
- *Статистика* – логічне значення, яке вказує, чи виводити додаткову інформацію регресійного аналізу (так чи ні). Якщо *Статистика* = 1, то додаткова інформація виводиться, якщо *Статистика* = 0, то виводяться тільки оцінки параметрів рівняння. Клацніть **ОК**;
- у лівій верхній комірці виділеної області з’явиться перший елемент результуючого масиву даних; для того, щоб вивести всі дані, клацніть <F2>, а потім – комбінацію клавиш <Ctrl>+<Shift>+<Enter>.

Додаткова статистика буде виводитись у такому вигляді (таблиця 3.3).

Таблиця 3.3 – Додаткова статистика

Значення b_p	Значення b_{p-1}	. . .	Значення b_2	Значення b_1	Значення b_0
Стандартна помилка оцінки b_p (s_{b_p})	Стандартна помилка оцінки b_{p-1} ($s_{b_{p-1}}$)	. . .	Стандартна помилка оцінки b_2 (s_{b_2})	Стандартна помилка оцінки b_1 (s_{b_1})	Стандартна помилка оцінки b_0 (s_{b_0})
Коефіцієнт детермінації R^2	Стандартна помилка моделі (залишків) S_e	#Н/Д	#Н/Д	#Н/Д	#Н/Д
F-статистика	Кількість ступенів вільності n-(p+1)	#Н/Д	#Н/Д	#Н/Д	#Н/Д
Регресійна сума квадратів	Залишкова сума квадратів	#Н/Д	#Н/Д	#Н/Д	#Н/Д

Функція **ЛИНЕЙН** також може бути використана для розрахунку оцінок параметрів моделей, які за допомогою перетворень зводяться до лінійного вигляду.

Для нашого прикладу результат застосування функції **ЛИНЕЙН** представлений на рисунке 3.26:

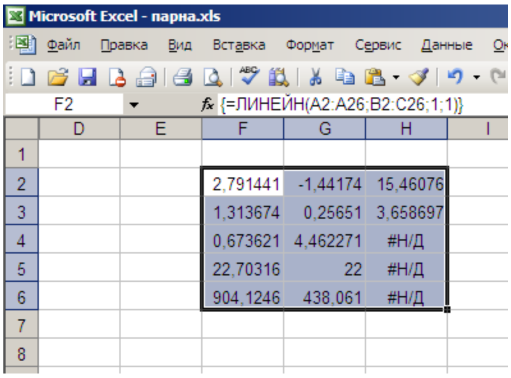


Рисунок 3.26 – Результат застосування функції **ЛИНЕЙН**

Таким образом, отримали економетричну модель:

$$\hat{y} = 15{,}46 - 1{,}442x_1 + 2{,}791x_2\ .$$

За рисунком 3.26 також можна визначити, що коефіцієнт детермінації дорівнює R^2= 0,673621; F-статистика = 22,7, ступенів вільності – 22 і т.д.

3.5 ДОСЛІДЖЕННЯ МУЛЬТИКОЛІНЕАРНОСТІ

(за алгоритмом Фаррара-Глобера)

Приклад. У таблиці 3.4 задані дані з діяльності фірми. Прибуток фірми (y, у.г.о.) та фактори, від яких він залежить - фондовіддача (x_1, у.г.о.), продуктивність праці (x_2, у.г.о.), питомі інвестиції (x_3, у.г.о.).

Таблиця 3.4 – Вихідні дані

№ місяця	Прибуток фірми y, у.г.о.	Фондовіддача x_1, у.г.о.	Продуктивність праці x_2, у.г.о.	Питомі інвестиції x_3, у.г.о.
1	40	12	5	15
2	45	17	7	18
3	40	13	6	16
4	43	14	7	17
5	48	16	6	20
6	39	15	5	15
7	42	14	6	16
8	45	17	9	18
9	38	12	5	19
10	48	18	10	20
11	50	20	11	22
12	48	17	10	21
13	49	18	12	21
14	45	19	8	20
15	49	20	9	22
16	52	22	14	23
17	54	24	15	24
18	51	21	13	20
19	55	25	16	24
20	56	27	18	25

Завдання: 1) перевірити масив незалежних змінних на наявність мультиколінеарності за алгоритмом Фаррара-Глобера;

2) при наявності мультиколінеарності змінних знайти істинні оцінки параметрів моделі;

3) побудувати модель, позбавлену від колінеарності.

Розв'язання.

Перевірка масиву даних на мультиколінеарність за алгоритмом Фаррара-Глобера включає в себе наступні кроки:

1. Стандартизація (нормалізація) змінних.
2. Визначення кореляційної матриці.
3. Визначення критерію χ^2 (хі-квадрат).
4. Визначення оберненої матриці.
5. Обчислення F-критеріїв.
6. Обчислення часткових коефіцієнтів кореляції.
7. Обчислення t-критеріїв.

Здійснимо всі ці кроки із використанням можливостей табличного процесора **Excel**.

Нормалізація змінних передбачає додаткові обчислення. А саме, треба визначити середні значення та стандартні відхилення змінних x_1, x_2, x_3.

Для цього використаємо функції **СРЗНАЧ** та **СТАНДОТКЛОН Мастер функций** в категорії «Статистические» (таблиця 3.5)

Таблиця 3.5 - Допоміжні розрахунки

	x_1	x_2	x_3
«СРЗНАЧ»	18,05	9,60	19,80
«СТАНДОТКЛОН»	4,26	3,97	3,04

Нормалізація (стандартизація) незалежних змінних

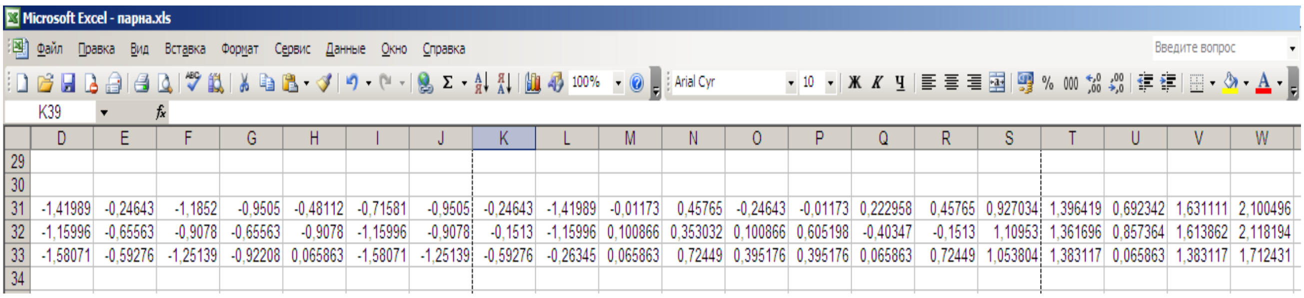
Для цього використаємо функцію **НОРМАЛИЗАЦИЯ**, категорія «Статистические» (таблиця 3.6).

Таблиця 3.6 - Нормалізовані незалежні змінні

	X_1^*	X_2^*	X_3^*
	-1,41989	-1,15996	-1,58071
	-0,24643	-0,65563	-0,59276
	-1,1852	-0,9078	-1,25139
	-0,9505	-0,65563	-0,92208
	-0,48112	-0,9078	0,065863
	-0,71581	-1,15996	-1,58071
	-0,9505	-0,9078	-1,25139
	-0,24643	-0,1513	-0,59276
	-1,41989	-1,15996	-0,26345
X^* =	-0,01173	0,100866	0,065863
	0,45765	0,353032	0,72449
	-0,24643	0,100866	0,395176
	-0,01173	0,605198	0,395176
	0,222958	-0,40347	0,065863
	0,45765	-0,1513	0,72449
	0,927034	1,10953	1,053804
	1,396419	1,361696	1,383117
	0,692342	0,857364	0,065863
	1,631111	1,613862	1,383117
	2,100496	2,118194	1,712431

Транспонування нормалізованої матриці X^*

Для цього використаємо функцію **ТРАНСП**, категорія «Ссылки и массивы» або «Полный алфавитный перечень» (рисунок 3.27):



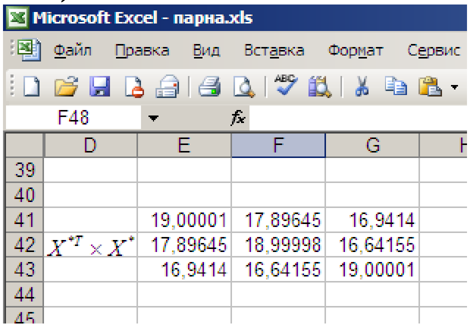
	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
29																				
30																				
31	-1,41989	-0,24643	-1,1852	-0,9505	-0,48112	-0,71581	-0,9505	-0,24643	-1,41989	-0,01173	0,45765	-0,24643	-0,01173	0,222958	0,45765	0,927034	1,396419	0,692342	1,631111	2,100496
32	-1,15996	-0,65563	-0,9078	-0,65563	-0,9078	-1,15996	-0,9078	-0,1513	-1,15996	0,100866	0,353032	0,100866	0,605198	-0,40347	-0,1513	1,10953	1,361696	0,857364	1,613862	2,118194
33	-1,58071	-0,59276	-1,25139	-0,92208	0,065863	-1,58071	-1,25139	-0,59276	-0,26345	0,065863	0,72449	0,395176	0,395176	0,065863	0,72449	1,053804	1,383117	0,065863	1,383117	1,712431
34																				

Рисунок 3.27 – Результа використання функції **ТРАНСП**

Обчислення добутку матриць X^{*T} та X^*

Для цього використаємо функцію **МУМНОЖ**, категорія

«Математичні» (рисунок 3.28).



	D	E	F	G
39				
40				
41		19,00001	17,89645	16,9414
42	$X^{*T}\times X^{*}$	17,89645	18,99998	16,64155
43		16,9414	16,64155	19,00001
44				

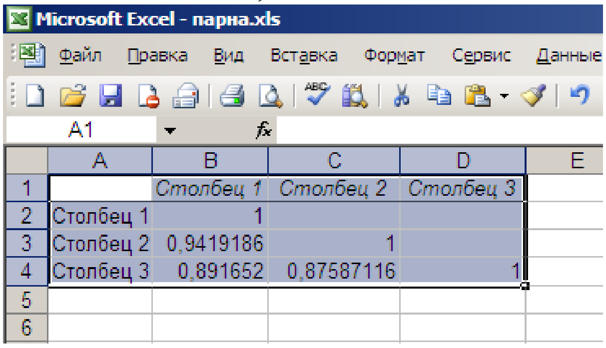
Рисунок 3.28 – Результат використання функції **МУМНОЖ**

Визначення кореляційної матриці r

Для цього необхідно кожний елемент матриці $X^{*T}\times X^{*}$ домножити на $\frac{1}{n-1}$, де ***n***- кількість спостережень (у нашому випадку ***n*** =20)

$$r = \begin{vmatrix} 1 & 0{,}942 & 0{,}892 \\ 0{,}942 & 1 & 0{,}876 \\ 0{,}892 & 0{,}876 & 1 \end{vmatrix}$$

Визначити кореляційну матрицю також можна за допомогою програми **Сервис/Анализ данных/Корреляция**. У результаті виведуться нижньо трикутна кореляційна матриця (рисунок 3.29), яка є симетричною. Відтворити елементи, яких не вистачає, можна шляхом копіювання.



	A	B	C	D
1		Столбец 1	Столбец 2	Столбец 3
2	Столбец 1	1		
3	Столбец 2	0,9419186	1	
4	Столбец 3	0,891652	0,87587116	1

Рисунок 3.29 – Кореляційна матриця

Розрахунок визначника кореляційної матриці det(r)

Для цього використаємо функцію **МОПРЕД**, категорія «Математичні»:

$$\det(r) = 0.0218$$

Оскільки визначник кореляційної матриці *наближається до 0*, то в масиві пояснюючих змінних <u>може</u> існувати мультиколінеарність.

Перевірка всього масиву змінних на наявність колінеарності за критерієм χ^2

Критерій розраховується за формулою:

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

де m – кількість незалежних змінних.

Визначимо логарифм визначника кореляційної матриці (функція **LN**) :

$$\ln \det(r) = -3.825$$

Тоді,

$$\chi^2 = -\left\{20 - 1 - \frac{2 \times 3 + 5}{6}\right\} \times (-3.825) = 65.663$$

Розраховане значення χ^2 порівняємо з табличним для $\nu = \frac{m(m-1)}{2} = \frac{3(3-1)}{2} = 3$ ступенів вільності та рівня значущості $\alpha = 0.05$ (див. табл. «Процентилі χ^2-розподілу»): $\chi^2_{табл} = 7{,}81$.

Табличне значення χ^2 можна також знайти в категорії Статистические **Мастера функций** Excel, функция **ХИ2ОБР** (рисунок 3.30).

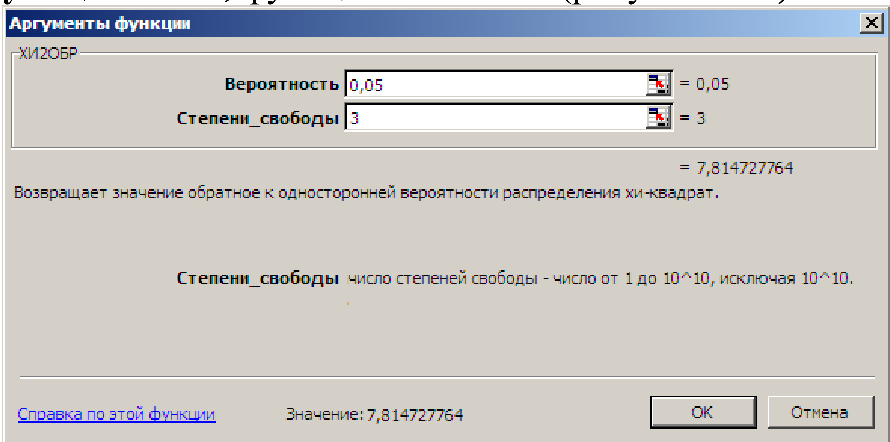


Рисунок 3.30 – Визначення табличного значення χ^2

Оскільки $\chi^2 > \chi^2_{табл}$, то в масиві незалежних змінних (фондовіддача, продуктивність праці, питомі інвестиції) існує мультиколінеарність.

Перевірка на колінеарність окремих змінних з іншими за F-критеріями.

Критерій розраховується за формулою:

$$F = (C_{ii} - 1)\left(\frac{n - m}{m - 1}\right),$$

де C_{ii} - діагональний елемент матриці С (матриці , оберненої до кореляційної матриці r)

Знайдемо обернену матрицю C за допомогою функції **МОБР** категорії «Математичні»:

$$C = \begin{vmatrix} 10{,}671 & -7{,}376 & -3{,}055 \\ -7{,}376 & 9{,}393 & -1650 \\ -3{,}055 & -1{,}650 & 5{,}169 \end{vmatrix} \qquad \text{структура матриці } C = \begin{vmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{2} & C_{33} \end{vmatrix}$$

F-критерій для фактору x_1 (фондовіддача):

$$F_1 = (10{,}671 - 1)\left(\frac{20 - 3}{3 - 1}\right) = 82{,}205$$

F-критерій для фактору x_2 (продуктивність праці):

$$F_2 = (9{,}393 - 1)\left(\frac{20 - 3}{3 - 1}\right) = 71{,}34$$

F-критерій для фактору x_3 (питомі інвестиції):

$$F_3 = (5{,}169 - 1)\left(\frac{20-3}{3-1}\right) = 35{,}436$$

Розраховане значення F-критерію порівняємо з табличним для $\nu_1 = n - m = 20 - 3 = 17$ та $\nu_2 = m - 1 = 3 - 1 = 2$ ступенів вільності та рівня значущості $\alpha = 0.05$ (див. табл. «F-розподіл»): $F_{табл} = 3{,}59$

Оскільки $F_1 > F_{табл}$, $F_2 > F_{табл}$, $F_3 > F_{табл}$ то кожна із незалежних змінних (фондовіддача, продуктивність праці, питомі інвестиції) колінеарна з іншими.

Визначення частинних коефіцієнтів кореляції $r_{ij,k}$

Коефіцієнти розраховуються за формулою:

$$r_{ij,k} = \frac{-C_{ij}}{\sqrt{C_{ii}C_{jj}}}$$

Коефіцієнт кореляції $r_{ij,k}$ для факторів x_1 (фондовіддача) та x_2 (продуктивність праці):

$$r_{12,3} = \frac{-(-7{,}376)}{\sqrt{10{,}671 \times 9{,}393}} = 0{,}737$$

Коефіцієнт кореляції $r_{ij,k}$ для факторів x_1 (фондовіддача) та x_3 (питомі інвестиції):

$$r_{13,2} = \frac{-(-3{,}055)}{\sqrt{10{,}671 \times 5{,}169}} = 0{,}411$$

Коефіцієнт кореляції $r_{ij,k}$ для факторів x_2 (продуктивність праці) та x_3 (питомі інвестиції):

$$r_{23,1} = \frac{-(-1{,}650)}{\sqrt{9{,}393 \times 5{,}169}} = 0{,}237$$

На основі частинних коефіцієнтів кореляції можна стверджувати:

- між фондовіддачею та продуктивністю праці існує тісний зв'язок, якщо не враховувати вплив питомих інвестицій;
- між фондовіддачею та питомими інвестиціями існує помірний зв'язок, якщо не враховувати вплив продуктивності праці;
- між продуктивністю праці та питомими інвестиціями існує слабкий зв'язок, якщо не враховувати вплив фондовіддачі.

Перевірка на колінеарність кожної пари змінних за допомогою t-критерія

Критерії розраховуються за формулою:

$$t_{ij} = \frac{r_{ij,k}\sqrt{n-m}}{\sqrt{1 - r_{ij,k}^2}}$$

$\underline{t}$ -критерій для фактору x_1 (фондовіддача):

$$t_{12} = \frac{0{,}737 \times \sqrt{20-3}}{\sqrt{1 - 0{,}737^2}} = 4{,}496$$

t -критерій для фактору x_2 (продуктивність праці):

$$t_{13} = \frac{0{,}411 \times \sqrt{20-3}}{\sqrt{1-0{,}411^2}} = 1{,}859$$

t -критерій для фактору x_3 (питомі інвестиції):

$$t_{23} = \frac{0{,}237 \times \sqrt{20-3}}{\sqrt{1-0{,}237^2}} = 1{,}006$$

Розраховане значення t-критерію порівняємо з табличним для $\nu = n - m = 20 - 3 = 17$ ступенів вільності та рівня значущості $\alpha = 0{,}05$ (див. табл. «Процентилі t-розподілу»): $t_{табл} = 1{,}740$.

Оскільки $t_{12} > t_{табл}$, то фондовіддача та продуктивність праці колінеарні між собою; $t_{13} > t_{табл}$, то фондовіддача та питомі інвестиції колінеарні між собою; $t_{23} < t_{табл}$, то продуктивність праці та питомі інвестиції не колінеарні між собою.

Розрахунки показали, що в масиві незалежних змінних існує колінеарність між факторами x_1 та x_2, а також між x_1 та x_3. Тому для визначення незміщених, ефективних та обґрунтованих оцінок параметрів (істинних оцінок) необхідно позбутися наслідків колінеарності.

Визначення істинних оцінок параметрів моделі $Y = f(X_1 X_2 X_3)$

Нормалізація фактору Y – прибуток фірми

За допомогою функцій **СРЗНАЧ**, **СТАНДОТКЛОН** та **НОРМАЛИЗАЦИЯ** здійснимо допоміжні обчислення (таблиця 3.7, 3.8).

Таблиця 3.8 - Допоміжні обчислення

	y
«СРЗНАЧ»	46,85
«СТАНДОТКЛОН»	5,38

Таблиця 3.8 - Нормалізація залежної змінної

Прибуток фірми Y, у.г.о.	Y^*
40	-1,273
45	-0,344
40	-1,273
43	-0,715
48	0,2136
39	-1,458
42	-0,901
45	-0,344
38	-1,644
48	0,2136
50	0,5852
48	0,2136
49	0,3994
45	-0,344
49	0,3994
52	0,9567

Продовження таблиці 3.8

54	1,3283
51	0,771
55	1,514
56	1,6998

Тоді, масив нормалізованих змінних для визначення оцінок параметрів нормалізованої моделі буде наступним (таблиця 3.9):

Таблиця 3.9 - Вхідні дані для побудови «нормалізованої моделі»

№ місяця	*Y**	X_1*	X_2*	X_3*
1	-1,273	-1,41989	-1,15996	-1,58071
2	-0,344	-0,24643	-0,65563	-0,59276
3	-1,273	-1,1852	-0,9078	-1,25139
4	-0,715	-0,9505	-0,65563	-0,92208
5	0,2136	-0,48112	-0,9078	0,065863
6	-1,458	-0,71581	-1,15996	-1,58071
7	-0,901	-0,9505	-0,9078	-1,25139
8	-0,344	-0,24643	-0,1513	-0,59276
9	-1,644	-1,41989	-1,15996	-0,26345
10	0,2136	-0,01173	0,100866	0,065863
11	0,5852	0,45765	0,353032	0,72449
12	0,2136	-0,24643	0,100866	0,395176
13	0,3994	-0,01173	0,605198	0,395176
14	-0,344	0,222958	-0,40347	0,065863
15	0,3994	0,45765	-0,1513	0,72449
16	0,9567	0,927034	1,10953	1,053804
17	1,3283	1,396419	1,361696	1,383117
18	0,771	0,692342	0,857364	0,065863
19	1,514	1,631111	1,613862	1,383117
20	1,6998	2,100496	2,118194	1,712431

Визначення «нормалізованих» оцінок параметрів.

Для цього скористаємося функцією **ЛИНЕЙН** без визначення параметру b_0^*

(*при заповненні робочого поля функції в вікні* ***КОНСТ*** *позначити «ЛОЖЬ» або «0»*)

b_3^*	b_2^*	b_1^*	
0,3103	0,3024	0,382	0
0,1491	0,2009	0,214	#Н/Д
0,9269	0,2858	#Н/Д	#Н/Д
71,878	17	#Н/Д	#Н/Д
17,612	1,3885	#Н/Д	#Н/Д

Істинні оцінки параметрів моделі визначаються за формулами:

$$b_1 = b_1^* \times \left(\frac{\sigma_Y}{\sigma_{X_1}} \right),\ b_2 = b_2^* \times \left(\frac{\sigma_Y}{\sigma_{X_2}} \right),\ b_3 = b_3^* \times \left(\frac{\sigma_Y}{\sigma_{X_3}} \right),$$

$$b_0 = \overline{Y} - b_1 \overline{X}_1 - b_2 \overline{X}_2 - b_3 \overline{X}_3$$

Для зручності зведемо всі необхідні для розрахунків дані в допоміжну таблицю 3.9.

Таблиця 3.9 - Допоміжні дані

Показник				
Середнє значення	6,85	8,05	,60	9,80
Середнє квадратичне відхилення, σ	,38	,26	,97	,04

Істинні параметри:

$$b_1 = 0{,}382 \times \left(\frac{5{,}38}{4{,}26} \right) = 0{,}482,\ b_2 = 0{,}302 \times \left(\frac{5{,}38}{3{,}97} \right) = 0{,}409,\ b_3 = 0{,}310 \times \left(\frac{5{,}38}{3{,}04} \right) = 0{,}549.$$

Звідси, $b_0 = 46{,}85 - 0{,}482 \times 18{,}05 - 0{,}409 \times 9{,}60 - 0{,}549 \times 19{,}80 = 23{,}353$

Модель прибутку без колінеарності має вигляд:

$$Y = 23.353 + 0.482x_1 + 0.409x_2 + 0.549x_3.$$

Оцінка достовірності побудованої моделі проведемо за допомогою визначення коефіцієнту кореляції та коефіцієнту детермінації.

Коефіцієнт кореляції визначимо за допомогою функції **КОРРЕЛ**, категорія «Статистичні»

$$r_{YY_{tr}} = 0{,}9628.$$

Коефіцієнт детермінації $R^2 = (0{,}9628)^2 = 0{,}9269$

Примітка. Дослідження масиву на наявність гетероскедастичності та автокореляції здійснюється за допомогою відповідних тестів (алгоритмів, методів тощо). Всі необхідні обчислення та перетворення можна виконувати у табличному процесорі Excel, використовуючі при цьому розглянуті вище операції та функції.

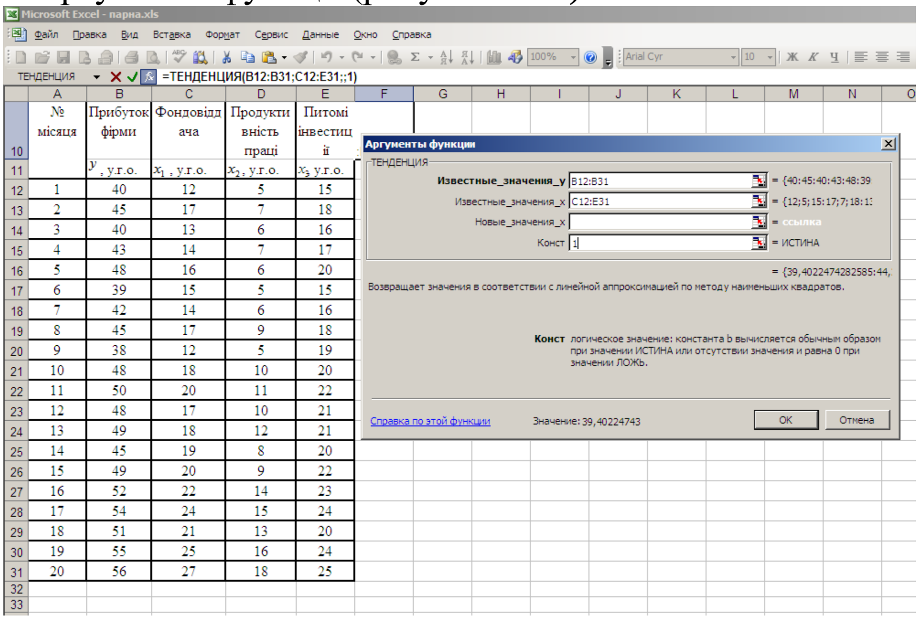
3.6 ТОЧКОВИЙ ПРОГНОЗ

Виконати прогноз прибутку фірми на майбутні, приміром, чотири місяці можна за допомогою функції **ТЕНДЕНЦИЯ.**

Для цього треба виконати наступні дії:

1) виділіть область порожніх комірок, яка складається з одного стовпчика та 4 рядків (за кількістю прогнозних періодів);
2) активізуйте **Мастер функций;**

3) виберіть Категорию **Статистические**, функцію **ТЕНДЕНЦИЯ.** Клацніть на **ОК**;
4) заповніть аргументи функції (рисунок 3.31):



№ місяця	Прибуток фірми y, у.г.о.	Фондовіддача x_1, у.г.о.	Продуктивність праці x_2, у.г.о.	Питомі інвестиції x_3 у.г.о.
1	40	12	5	15
2	45	17	7	18
3	40	13	6	16
4	43	14	7	17
5	48	16	6	20
6	39	15	5	15
7	42	14	6	16
8	45	17	9	18
9	38	12	5	19
10	48	18	10	20
11	50	20	11	22
12	48	17	10	21
13	49	18	12	21
14	45	19	8	20
15	49	20	9	22
16	52	22	14	23
17	54	24	15	24
18	51	21	13	20
19	55	25	16	24
20	56	27	18	25

Рисунок 3.31 – Приклад точкового прогнозу

- *Известные значения У* – діапазон, який містить дані, що характеризують результуючу;
- *Известные значения Х* – діапазон, який містить дані, що описують всі незалежні змінні;
- *Новые значения Х* – діапазон комірок, який містить значення змінних, які необхідно підставити у рівняння для визначення очікуваних значень y;
- *Константа* – логічне значення, яке вказує на наявність або відсутність вільного члена у рівнянні (*Константа* = 1). Клацніть на **ОК.**

5) у лівій верхній комірці виділеної області з’явиться перший елемент результуючого масиву даних; для того, щоб вивести всі дані, клацніть <F2>, а потім – комбінацію клавиш <Ctrl>+<Shift>+<Enter>.

Зауважимо, що інтервальний прогноз є більш реалістичним, ніж точковий. Оскільки вказує не одне значення, а діапазон значень, якому із вказаною ймовірністю повинно належати майбутнє значення результуючої.

РЕКОМЕНДОВАНА ЛІТЕРАТУРА

Навчальна література

1. Доугерти К. Ведение в эконометрику: учебник / К. Догерти. – М. : ИНФРА-М, 2004. – 432 с.

2. Елисеева И. И. Практикум по эконометрике / И. И. Елисеева, С. В. Курышева, Н. М. Гордеенко; под ред. чл.-корр. РАН И. И. Елисеевой. – М. : Финансы и статистика, 2002. – 191 с.

3. Елисеева И. И. Эконометрика: учебник для вузов / И. И. Елисеева, С. В. Курышева, Т. В. Костеева; под ред. чл.-корр. РАН И. И. Елисеевой. – М. : Финансы и статистика, 2002. – 342 с.

4. Клебанова Т. С. Эконометрия : учебное пособие / Т. С. Клебанова, Н. А. Дубровина, Е. В. Раевнева. – Харьков : ИНЖЭК, 2005. – 160 с.

5. Кремер Н. Ш. Эконометрика : учебник для вузов / Н. Ш. Кремер, Б. А. Путко; под ред. проф. Н. Ш. Кремера. – М. : ЮНИТИ-ДАНА, 2005. – 311 с.

6. Лугінін О. Є. Економетрія: навчальний посібник / О. Є. Лугінін, С. В. Білоусова, О. М. Білоусов. – К. : Центр навчальної літератури, 2005. – 252 с.

7. Наконечний С. І. Економетрія : підручник / С. І. Наконечний, Т. О. Терещенко, Т. П. Романюк. – К. : КНЕУ, 2005. – 520 с.

8. Скітер І.С. Економетрія. Навчально – методичний посібник / Укладач Скітер І.С . – Чернігів: ЧДІЕУ, 2003. - 66 с.

9. Лук'яненко І.Г. Економетрика: Підручник / І.Г.Лук'яненко, Л.І.Краснікова. – К.: "Знання", КОО, 1998. – 493 с.

6. Лук'яненко І.Г. Економетрика: Практикум з використанням комп'ютера / І.Г.Лук'яненко, Л.І.Краснікова. – К.: "Знання", КОО, 1998. – 217 с.

7. Магнус Я.Р. и др. Эконометрика: начальный курс / Я.Р.Магнус. – М., 2000. - 400 с.

8. Наконечний С.І. Економетрія: Навч.-метод. посібник для самост. вивч. дисц./ С.І.Наконечний, Т.О.Терещенко. – К.: КНЕУ, 2001. - 192 с.

10. Здрок В.В. Економетрія: Підручник / В.В.Здрок, Т.Я.Лагоцький. – К.: Знання, 2010. – 541 с.

11.Сеньо П.С. теорія ймовірностей та математична статистика: Підручник / П.С.Сеньо. – К.: Центр навчальної літератури , 2004. – 448 с.

Методична література

12.Браславец М. Е. Кибернетика / М. Е.Браславец, Т. Ф.Гуревич. - К.: Вища школа, 1977. - 325 с.

13.Виханский О. С. Стратегическое управление / О. С.Виханский. - М.: Гардарики, 1999. - 296 с.

14.Кальянов Г. Н. CASE структурный системный аналіз / Г.Н. Кальянов. - М.: Лори, 1996. - 242 с.

15.Марка Д. А. Методология стуктурного анализа и проектирования / Д. А.Марка, К. Мак-Гоуэн.Пер. с англ. - М.: 1993. - 240 с.

16.Прогнозування і розробка програм: Метод. посібник / За ред. В. Ф. Беседіна. - К.: Наук. світ, 2000. - 468 с.

17.Проектирование информационных систем с использованием CASE-технологий: Учеб. пособие / Санкт-Петербургский гос. ун-т водных коммуникаций. - СПб.: СПГУВК, 2000. — 172 с.

18.Скрипченко И.А. Анализ данных в Microsoft Excel: учебное пособие / И.А.Скрипченко. – Иркутск: Изд-во ИрГТУ, 1999. – 324 с.

Допоміжна література

19.Акофф Р. Л. Планирование в больших экономических системах / Р. Л. Акофф. Пер. с англ. - М.: Сов. радио, 1972. - 223 с.

20.Андрейчиков А. В. Анализ, синтез, планирование решений в экономике / А.В.Андрейчиков, О.Н.Андрейчикова. - М.: Финансы и статистика, 2000. - 368 с.

21.Бронштейн И. Н., Семендяев К. А. Справочник по математике для инженеров и учащихся втузов. -М.: Наука. – 1981.- 721 с.

22.Гибсон Дж. Л. Организации: поведение, структура, процессы: Пер. с англ. / Гибсон Дж. Л., Иванцевич Д. М., Донелли Д. Х.-мл. - М.: ИНФРА-М, 2000. - 662 с.

23.Карагодова О. О. Дослідження операцій: Навч. Посібник / О.О. Карагодова. – К.:ЦУЛ, 2007. – 256 с.

24.Кирьянов Д. Самоучитель MathCad 11 / Д. Кирьянов. – С.-П.: БХВ, 2003. - 538 с.

25.Косоруков О. А. Исследование операций: Учебник / О.А. Косоруков. – М.: Экзамен, 2003. – 448 с.

26.Лавренов С.М. Excel: Сборник примеров и задач / С.М. Лавренов. – М.: ФиС, 2003. – 336 с.

27.Макарова Н. В. Статистика в Excel: Учебное пособие / Н.В. Макарова. – М.: ФиС, 2002. – 368 с.

28.Шмойлова Р. А. Практикум по теории статистики / Р.А. Шмойлова. – М.: ФиС, 2003. – 416 с.

ДОДАТОК А - ЕЛЕМЕНТИ МАТРИЧНИХ ПЕРЕТВОРЕНЬ

ОЗНАЧЕННЯ МАТРИЦІ. ОСНОВНІ ВИДИ МАТРИЦЬ

Розглянемо множину $m \times n$ дійсних чисел, записаних у вигляді прямокутної таблиці з m рядків і n стовпців:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{pmatrix}.$$

Матрицею називається таблиця упорядкованих чисел, яка складається з m рядків і n стовпців.

Позначаються матриці літерами *A, B, C* тощо.

Числа a_{ij} називаються її **елементами**. Індекси i та j елемента a_{ij} позначають відповідно номер рядка та стовпця, на перетині яких міститься даний елемент. Наприклад, елемент a_{23} міститься в другому рядку і третьому стовпці.

Розглянемо матрицю

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{32} \end{pmatrix},$$

яка має два рядки (m = 2) і три стовпці (n = 3), тобто розміром 2 × 3. Загалом, якщо матриця m рядків має рядків і n стовпців, розмір такої матриці дорівнює (m x n).

Якщо в матриці A кількість рядків m дорівнює кількості стовпців n (m = n) ,її називають **квадратною** порядку m (або n). Якщо m ≠ n, то матриця A є прямокутною розміром (m x n).

Матриця-стовпець є прямокутна матриця порядку m x 1:

$$A = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \dots \\ a_{mj} \end{pmatrix}.$$

Матриця-рядок є прямокутна матриця порядку 1x n:

$$A = \begin{pmatrix} a_{i1} & a_{i2} & \dots & a_{in} \end{pmatrix}.$$

Матриця, усі елементи якої дорівнюють нулю, називається **нульовою**:

$$A = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}.$$

Розглянемо квадратну матрицю порядку n x n.

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{pmatrix}.$$

Елементи $a_{11}, a_{22} \dots a_{nn}$ утворюють головну діагональ матриці A; елементи $a_{1n}, a_{2,\, n-1} \dots a_{n1}$ — побічну діагональ матриці A.

Квадратна матриця, в якої всі елементи, крім елементів головної діагоналі, дорівнюють нулю, називається **діагональною**, тобто

$$A=\begin{pmatrix} a_{11} & 0 & 0 & \ldots & 0 \\ 0 & a_{22} & 0 & \ldots & 0 \\ 0 & 0 & a_{33} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & a_{nn} \end{pmatrix}.$$

Скорочено: квадратна матриця $A=(a_{ij})$ є діагональною, якщо $a_{ij}=0$ для $i \neq j$ та $a_{ij} \neq 0$ для $i=j$.

Квадратна матриця E_n є **одиничною** n-го порядку, якщо всі елементи головної діагоналі дорівнюють одиниці, а всі інші елементи — нулю, тобто

$$A=E_n=\begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 \end{pmatrix}.$$

Скорочено: квадратна матриця $I=(a_{ij})$ є одиничною, якщо $a_{ij}=0$ для $i \neq j$ та $a_{ij}=1$ для $i=j$.

Квадратна матриця $A=(a_{ij})$ є **трикутною**, якщо всі її елементи над головною діагоналлю ($a_{ij}=0$,коли i < j) або під цією діагоналлю ($a_{ij}=0$, коли i > j) дорівнюють нулю:

$$A=\begin{pmatrix} a_{11} & 0 & 0 & \ldots & 0 \\ a_{21} & a_{22} & 0 & \ldots & 0 \\ a_{31} & a_{32} & a_{33} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & a_{n3} & \ldots & a_{nn} \end{pmatrix} \quad \text{або} \quad A=\begin{pmatrix} a_{11} & a_{12} & a_{13} & \ldots & a_{1n} \\ 0 & a_{22} & a_{23} & \ldots & a_{2n} \\ 0 & 0 & a_{33} & \ldots & a_{3n} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & a_{nn} \end{pmatrix}.$$

Якщо в матриці А поміняємо місцями відповідно елементи рядків на елементи стовпців (або навпаки), дістанемо **транспоновану матрицю** (позначається A′ або):

ЕЛЕМЕНТАРНІ ДІЇ НАД МАТРИЦЯМИ

Дві матриці $A=(a_{ij})$ та $B=(b_{ij})$ одного й того самого порядку (m x n) вважаються **рівними**, якщо всі відповідні елементи цих матриць рівні між собою, тобто

$$A=B \ \Leftrightarrow a_{ij}=b_{ij} \ (i=\overline{1,m};\ j=\overline{1,n}).$$

Отже, матриці різних порядків завжди не рівні між собою.

Матриці можна додавати, віднімати, множити матрицю на число та матриці на матрицю.

Додавання і віднімання виконуються лише для матриць одного й того самого порядку. Якщо $A=(a_{ij})$ і $B=(b_{ij})$ мають порядок m Ч n, то

Скорочено: $C = (c_{ij}) = (a_{ij} \pm b_{ij})$.

Очевидно,що

$$A + B = B + A; \quad (A + B)' = A' + B'; \; A + (B - A) = B.$$

При додаванні матриць A, B, і C одного й того самого порядку справджується закон асоціативності:

$$(A + B) + C = A + (B + C).$$

Добутком скаляра λ на матрицю $A = (a_{ij})$ порядку (m x n) називається матриця, елементи якої дорівнюють λa_{ij}, тобто

$$\lambda A = \begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \ldots & \lambda a_{1n} \\ \lambda a_{21} & \lambda a_{22} & \ldots & \lambda a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ \lambda a_{m1} & \lambda a_{m2} & \ldots & \lambda a_{mn} \end{pmatrix} = (\lambda a_{ij})$$

При множенні матриці A на скаляр λ виконуються такі закони:

а) $\lambda A = A\lambda$;

б) $(\lambda A)' = \lambda A'$;

в) $\lambda(A + B) = \lambda A + \lambda B$;

г) $(\lambda + \gamma)A = \lambda A + \gamma B$;

д) $(\lambda\gamma)A = \lambda(\gamma A) = \gamma(\alpha A)$.

Дві матриці A і B можна помножити одна на одну, тобто визначити $C = AB$, коли кількість стовпців матриці A дорівнює кількості рядків матриці B.

Правило множення двох матриць A на B: кожний елемент матриці c_{ij} є сумою добутків відповідних елементів i-го рядка матриці A на елементи j-го стовпця матриці B, тобто

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj} \; .$$

Приклад. Знайти добуток $C = AB$, коли

$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{32} \end{pmatrix}$ - порядок 2 x 3;

$B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{pmatrix}$ - порядок 3 x 2.

Добуток цих двох матриць існує, оскільки кількість стовпців матриці A дорівнює трьом, стільки ж рядків має матриця B, тобто виконується умова множення двох матриць. Перемноживши ці матриці, дістанемо:

$$C = A \cdot B = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} & a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} & a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}$$

Порядок матриці C, яка є добутком A і B, дорівнює 2 x2.

При множенні матриць діють такі закони:

а) $AB \neq BA$, тобто добуток матриць не є комутативним.

Нехай $A=\begin{pmatrix}1 & 3\\ 2 & 4\end{pmatrix}$; $B=\begin{pmatrix}3 & 4\\ 2 & 1\end{pmatrix}$; $AB=\begin{pmatrix}9 & 7\\ 14 & 12\end{pmatrix}$; $BA=\begin{pmatrix}11 & 25\\ 4 & 10\end{pmatrix}$.

Отже, $AB \neq BA$.

б) $(AB)C = A(BC)$;

в) $(A+B)C = AC + BC$;

г) $C\,(A+B) = CA + CB$;

д) $\alpha(AB) = (\alpha A)\cdot B = A(\alpha B)$;

е) $AE = EA = A$, де E - одинична матриця того самого порядку, що й матриця A;

є) $(AB)' = B'A'$;

ж) $(ABC)' = C'B'A'$.

Як окремий випадок добуток матриці розміру 1 хp (вектор-рядок) на матрицю порядку p х 1 (вектор-стовпець) дає скаляр, а саме:

$$A = (a_1, a_2, a_3 \ldots a_p); \quad B = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \ldots \\ b_p \end{pmatrix},$$

$$C = (a_1 b_1 + a_2 b_2 + a_3 b_3 + \ldots + a_p b_p)$$

Якщо вектор $A = \begin{pmatrix} a_1 \\ a_2 \\ \ldots \\ a_n \end{pmatrix}$,

то

$$A'A = (a_1 \quad a_2 \quad \ldots \quad a_n)\begin{pmatrix} a_1 \\ a_2 \\ \ldots \\ a_n \end{pmatrix} = \sum_{i=1}^{n} a_i^2 = a_1^2 + a_2^2 + \ldots + a_n^2$$

і

$$AA' = \begin{pmatrix} a_1 \\ a_2 \\ \ldots \\ a_n \end{pmatrix}(a_1 \quad a_2 \quad \ldots \quad a_n) = \begin{pmatrix} a_1^2 & a_1a_2 & \ldots & a_1a_n \\ a_2a_1 & a_2^2 & \ldots & a_2a_n \\ \ldots & \ldots & \ldots & \ldots \\ a_na_1 & a_na_2 & \ldots & a_n^2 \end{pmatrix}.$$

Детермінантом (визначником) квадратної матриці А порядку n називається алгебраїчна сума n членів, кожний з яких містить n співмножників, узятих по одному з кожного рядка (стовпця) матриці**.**

Позначається:

$$det\, A \text{ або } |A| = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix}$$

Властивості визначників.

1. При транспонуванні матриці її визначник не змінюється, тобто $|A| = |A'|$.

2. Якщо всі елементи рядка (стовпця) матриці дорівнюють нулю, то її визначник також дорівнює нулю.

3. При перестановці двох будь-яких стовпців (рядків) визначника його знак змінюється на протилежний, а абсолютна величина не змінюється.

4. Визначник з двома однаковими стовпцями (рядками) дорівнює нулю.

5. При множенні якого-небудь стовпця (рядка) на довільне число $\lambda \neq 0$ значення визначника множиться на це саме число.

6. Спільний множник всіх елементів стовця (рядка) можна винести за знак визначника.

7. Якщо два стовпці (рядки) визначника пропорційні, то визначник дорівнює нулю.

8. Визначник не зміниться, якщо до будь-якого стовпця (рядка) додати елементи другого стовпця (рядка), попередньо помноживши їх на відмінний від нуля множник.

Розглянемо визначник матриці n-го порядку

$$|A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1,j} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2,j} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3,j} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & a_{i3} & \dots & a_{ij} & \dots & a_{in} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nj} & \dots & a_{nn} \end{vmatrix}$$

Викреслимо в ньому i-й рядок і j-й стовпець, на перетині яких міститься елемент a_{ij}. У результаті залишиться визначник матриці (n – 1)-го порядку

$$M_{ij} = \det\left|A^*\right| = \begin{vmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1,j-1} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2,j-1} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3,j-1} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{i-1,1} & a_{i-1,2} & a_{i-1,3} & \dots & a_{i-1,j-1} & \dots & a_{i-1,n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{n,j-1} & \dots & a_{nn} \end{vmatrix}.$$

Визначник матриці (n – 1)-го порядку, в якій викреслені і-й рядок та j-й стовпець, називається **мінором елемента** a_{ij} і позначається M_{ij}.

Мінор M_{ij}, який береться зі знаком $(-1)^{i+j}$ (i -номер рядка; j - номер стовпця елемента a_{ij}), є **алгебраїчним доповненням** цього елемента, тобто

$$A_{ij} = (-1)^{i+j} M_{ij}.$$

Визначник дорівнює сумі попарних добутків елементів будь-якого його стовпця (рядка) на відповідні їх алгебраїчні доповнення:

$$|A| = a_{11}A_{11} + a_{21}A_{21} + a_{31}A_{31} + \ldots + a_{n1}A_{n1} =$$
$$= a_{21}A_{21} + a_{22}A_{22} + a_{23}A_{23} + \ldots + a_{2n}A_{2n} =$$
$$= \ldots = a_{1j}A_{1j} + a_{2j}A_{2j} + a_{3j}A_{3j} + \ldots + a_{nj}A_{nj} =$$
$$= a_{i1}A_{i1} + a_{i2}A_{i2} + a_{i3}A_{i3} + \ldots + a_{in}A_{in}.$$

Ця властивість дає змогу розкласти визначник за елементами стовпця (рядка). Нехай потрібно знайти визначник $|A|$. Розкладемо його за елементами другого рядка:

$$|A|=\begin{vmatrix}1 & 2 & -1\\ 3 & 0 & 2\\ 4 & -2 & 5\end{vmatrix}=3\cdot(-1)^3\begin{vmatrix}2 & -1\\ -2 & 5\end{vmatrix}+0\cdot(-1)^4\begin{vmatrix}1 & -1\\ 4 & 5\end{vmatrix}+2\cdot(-1)^5\begin{vmatrix}1 & 2\\ 4 & -2\end{vmatrix}=$$
$$=-24+20=-4.$$

Зауважимо, що детермінант другого порядку обчислюють відніманням добутку елементів побічної від добутку елементів головної діагоналі:

$$\det\begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}=\begin{vmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{vmatrix}=a_{11}a_{22}-a_{21}a_{12}.$$

Обчислити детермінант матриці A розміром 3 Ч 3 можна згідно з властивістю (18) або за правилом Сарруса (Sarrus):

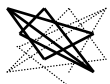
$$\begin{array}{c}\qquad\qquad\qquad\qquad + \quad + \quad + \quad - \quad - \quad - \\ |A|=\begin{vmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{vmatrix}=\begin{matrix}a_{11} & a_{12} & a_{13} & a_{11} & a_{12}\\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22}\\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32}\end{matrix}= \\ \qquad\qquad\qquad\qquad - \quad - \quad - \quad + \quad + \quad + \end{array}$$

$$=a_{11}a_{22}a_{33}+a_{21}a_{32}a_{13}+a_{12}a_{23}a_{31}-a_{31}a_{22}a_{13}-a_{11}a_{23}a_{32}-a_{21}a_{12}a_{33}.$$

Визначимо $|A|$ за правилом Сарруса:

$$|A|=\begin{vmatrix}1 & 2 & -1\\ 3 & 0 & 2\\ 4 & -2 & 5\end{vmatrix}=0+6+16-0+4-30=-4.$$

Правило Сарруса часто називають правилом трикутника. Далі поступово проілюструємо обчислення всіх членів визначника $|A|$.



Члени зі знаком «плюс»:

$$\overbrace{\begin{vmatrix}1 & 2 & -1\\ 3 & 0 & 2\\ 4 & -2 & 5\end{vmatrix}}^{0};\qquad \overbrace{\begin{vmatrix}1 & 2 & -1\\ 3 & 0 & 2\\ 4 & -2 & 5\end{vmatrix}}^{6};\qquad \overbrace{\begin{vmatrix}1 & 2 & -1\\ 3 & 0 & 2\\ 4 & -2 & 5\end{vmatrix}}^{16}.$$

Члени зі знаком «мінус»:

$$\overbrace{\begin{vmatrix}1 & 2 & -1\\ 3 & 0 & 2\\ 4 & -2 & 5\end{vmatrix}}^{0};\qquad \overbrace{\begin{vmatrix}1 & 2 & -1\\ 3 & 0 & 2\\ 4 & -2 & 5\end{vmatrix}}^{-4};\qquad \overbrace{\begin{vmatrix}1 & 2 & -1\\ 3 & 0 & 2\\ 4 & -2 & 5\end{vmatrix}}^{-30}.$$

Рангом матриці А називається найвищий порядок відмінного від нуля мінора цієї матриці.

Якщо $rg\ A = r$, то це означає, що серед мінорів матриці є, зрештою, хоча б один мінор r-го порядку, відмінний від нуля, тоді коли всі мінори вищого порядку: $r+1$, $r+2$ і т.д. дорівнюють нулю.

Нехай потрібно знайти ранг матриці A:

$$A=\left(\begin{vmatrix}1 & 2 & -1\\ 3 & 0 & 2\\ 4 & -2 & 5\\ 2 & 4 & -2\end{vmatrix}\right).$$

Згідно з означенням рангу матриці його значення не може перевищувати 3.

Обчислимо один з визначників третього порядку матриці A:

$$|A| = \begin{vmatrix} 1 & 2 & -1 \\ 3 & 0 & 2 \\ 4 & -2 & 5 \end{vmatrix} = -4 \neq 0.$$

Отже, ранг матриці A дорівнює 3.

ОБЕРНЕНА МАТРИЦЯ

Для кожної невиродженої матриці А порядку n x n існує однозначно визначена обернена матриця A^{-1} того самого порядку, така що виконується рівність

$$A^{-1} \cdot A = A \cdot A^{-1} = E_n,$$

E_n — одинична матриця порядку *n*.

Отже, якщо визначник (det A) не дорівнює нулю, то матриця А має обернену:

$$A^{-1} = \frac{1}{|A|} J,$$

де $|A|$ — визначник матриці A; J — так звана приєднана до A матриця. Вона складається з алгебраїчних доповнень A_{ji} до елементів a_{ij} матриці A, а саме:

$$J = \begin{pmatrix} A_{11} & A_{21} & A_{31} & \dots & A_{n1} \\ A_{12} & A_{22} & A_{32} & \dots & A_{n2} \\ \dots & \dots & \dots & \dots & \dots \\ A_{1n} & A_{2n} & A_{3n} & \dots & A_{nn} \end{pmatrix}.$$

Отже,

$$A^{-1} = \frac{1}{|A|} \begin{pmatrix} A_{11} & A_{21} & A_{31} & \dots & A_{n1} \\ A_{12} & A_{22} & A_{32} & \dots & A_{n2} \\ \dots & \dots & \dots & \dots & \dots \\ A_{1n} & A_{2n} & A_{3n} & \dots & A_{nn} \end{pmatrix} = \frac{1}{|A|} (A_{ij})'.$$

Наведемо основні властивості оберненої матриці:

а) $(AB)^{-1} = B^{-1}A^{-1}$;

б) $(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$;

в) $(A')^{-1} = (A^{-1})'$;

г) $|A^{-1}| = \frac{1}{|A|}$;

д) $A^{-1}A = E$;

е) $|E| = |A^{-1}A| = |A^{-1}||A|$.

Приклад. Знайти матрицю A^{-1}, обернену до матриці A:

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 3 & 0 & 2 \\ 4 & -2 & 5 \end{pmatrix}.$$

$$|A| = -4$$

Визначимо приєднану матрицю J:

$$A_{11} = \begin{vmatrix} 0 & 2 \\ -2 & 5 \end{vmatrix} = 4; \; A_{12} = -\begin{vmatrix} 3 & 2 \\ 4 & 5 \end{vmatrix} = -7; \; A_{13} = \begin{vmatrix} 3 & 0 \\ 4 & -2 \end{vmatrix} = -6;$$

$$A_{21} = -\begin{vmatrix} 2 & -1 \\ 2 & 5 \end{vmatrix} = -8; \; A_{22} = \begin{vmatrix} 1 & -1 \\ 4 & 5 \end{vmatrix} = 9; \; A_{23} = -\begin{vmatrix} 1 & 2 \\ 4 & -2 \end{vmatrix} = 10;$$

$$A_{31}=\begin{vmatrix}2 & -1\\ 0 & 2\end{vmatrix}=4;\ A_{32}=-\begin{vmatrix}1 & -1\\ 3 & 2\end{vmatrix}=-5;\ A_{33}=\begin{vmatrix}1 & 2\\ 3 & 0\end{vmatrix}=-6;$$

Отже,

$$J=\begin{pmatrix}4 & -8 & 4\\ -7 & 9 & -5\\ -6 & 10 & -6\end{pmatrix}.$$

Тоді

$$A^{-1}=\frac{1}{-4}\begin{pmatrix}4 & -8 & 4\\ -7 & 9 & -5\\ -6 & 10 & -6\end{pmatrix}=\begin{pmatrix}-1 & 2 & -1\\ \frac{7}{4} & -\frac{9}{4} & \frac{5}{4}\\ \frac{3}{2} & -\frac{5}{2} & \frac{3}{2}\end{pmatrix}.$$

Перевіримо, чи справді матриця A^{-1} є оберненою до матриці A. Знайдемо $A^{-1}A$ або AA^{-1}; у результаті має бути одинична матрицю E. Отже, перевіряється рівність $A^{-1}A = AA^{-1} = E$:

$$\begin{pmatrix}-1 & 2 & -1\\ \frac{7}{4} & -\frac{9}{4} & \frac{5}{4}\\ \frac{3}{2} & -\frac{5}{2} & \frac{3}{2}\end{pmatrix}\begin{pmatrix}1 & 2 & -1\\ 3 & 0 & 2\\ 4 & -2 & 5\end{pmatrix}=\begin{pmatrix}-1+6-4 & -2+0+2 & 1+4-5\\ \frac{7}{4}-\frac{27}{4}+\frac{20}{4} & \frac{14}{4}+0-\frac{10}{4} & -\frac{7}{4}-\frac{18}{4}+\frac{25}{4}\\ -\frac{3}{2}-\frac{15}{2}+\frac{12}{2} & \frac{6}{2}+0-\frac{6}{2} & -\frac{3}{2}-\frac{10}{2}+\frac{15}{2}\end{pmatrix}=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}.$$

ДОДАТОК В - Статистичні таблиці

Таблиця Б.1 - Розподіл Стьюдента (*t – критерій)*

Кількість ступенів вільності	Рівень значущості $\alpha=0,10$	Рівень значущості $\alpha=0,05$	Рівень значущості $\alpha=0,01$
1	6,3138	12,706	63,657
2	2,9200	4,3027	9,9248
3	2,3534	3,1825	5,8409
4	2,1318	2,7764	4,6041
5	2,0150	2,5706	4,0321
6	1,9432	2,4469	3,7074
7	1,8946	2,3646	3,4995
8	1,8595	2,3060	3,3554
9	1,8331	2,2622	3,2498
10	1,8125	2,2281	3,1693
11	1,7959	2,2010	3,1058
12	1,7823	2,1788	3,0545
13	1,7709	2,1604	3,0123
14	1,7613	2,1448	2,9768
15	1,7530	2,1315	2,9467
16	1,7459	2,1199	2,9208
17	1,7396	2,1098	2,8982
18	1,7341	2,1009	2,8784
19	1,7291	2,0930	2,8609
20	1,7247	2,0860	2,8453
21	1,7207	2,0796	2,8314
22	1,7171	2,0739	2,8188
23	1,7139	2,0687	2,8073
24	1,7109	2,0639	2,7969
25	1,7081	2,0595	2,7874
26	1,7056	2,0555	2,7787
27	1,7033	2,0518	2,7707
28	1,7011	2,0484	2,7633
29	1,6991	2,0452	2,7564
30	1,6973	2,0423	2,7500
40	1,6839	2,0211	2,7045
60	1,6707	2,0003	2,6603
120	1,6577	1,9799	2,6174
∞	1,6449	1,9600	2,5758

Таблиця Б.2 - Критерій Дарбіна - Уотсона ($\alpha = 0{,}05$)

n	$k' = 1$		$k' = 2$		$k' = 3$		$k' = 4$		$k' = 5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	1,08	1,36	0,95	1,54	0,82	1,75	0,69	1,97	0,56	2,2
16	1,10	1,37	0,98	1,54	0,86	1,73	0,74	1,93	0,62	2,1
17	1,13	1,38	1,02	1,54	0,90	1,71	0,78	1,90	0,67	2,10
18	1,16	1,39	1,05	1,53	0,93	1,69	0,82	1,87	0,71	2,06
19	1,18	1,40	1,08	1,53	0,97	1,68	0,86	1,85	0,75	2,02
20	1,20	1,41	1,10	1,54	1,00	1,68	0,90	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	1,46	1,21	1,55	1,12	1,66	1,04	1,77	0,95	1,89
26	1,30	1,47	1,22	1,55	1,14	1,65	1,06	1,76	0,98	1,88
27	1,32	1,48	1,24	1,56	1,16	1,65	1,08	1,76	1,01	1,86
28	1,33	1,48	1,26	1,56	1,18	1,65	1,10	1,75	1,03	1,85
29	1,34	1,49	1,27	1,56	1,20	1,65	1,12	1,74	1,05	1,84
30	1,35	1,50	1,28	1,57	1,21	1,21	1,65	1,14	1,07	1,83
31	1,36	1,50	1,30	1,57	1,23	1,65	1,16	1,74	1,09	1,83
32	1,37	1,51	1,31	1,57	1,24	1,65	1,18	1,73	1,11	1,82
33	1,38	1,51	1,32	1,58	1,26	1,65	1,19	1,73	1,13	1,81
34	1,39	1,52	1,33	1,58	1,27	1,65	1,21	1,73	1,15	1,81
35	1,40	1,52	1,34	1,58	1,28	1,65	1,22	1,73	1,16	1,80
36	1,41	1,52	1,35	1,59	1,29	1,65	1,24	1,73	1,18	1,80
37	1,42	1,53	1,36	1,59	1,31	1,66	1,25	1,72	1,19	1,80
38	1,43	1,54	1,37	1,59	1,32	1,66	1,26	1,72	1,21	1,79
39	1,43	1,54	1,38	1,60	1,33	1,66	1,27	1,72	1,22	1,79
40	1,44	1,54	1,39	1,60	1,34	1,66	1,29	1,72	1,23	1,79
45	1,48	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	1,38	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	1,52	1,70	1,49	1,74	1,46	1,77
75	1,60	1,65	1,57	1,68	1,54	1,71	1,51	1,74	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	1,57	1,72	1,55	1,75	1,52	1,77
90	1,63	1,68	1,61	1,70	1,59	1,73	1,57	1,75	1,54	1,78
95	1,64	1,69	1,62	1,71	1,60	1,73	1,58	1,75	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

Примітки. 1. n-кількістьспостережень; k' - кількість пояснюючих змінних.

Таблиця Б.3 - Критерій Дарбіна - Уотсона **(**$\alpha = 0{,}01$**)**

n	$k' = 1$		$k' = 2$		$k' = 3$		$k' = 4$		$k' = 5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	1,81	1,07	0,70	1,25	0,59	1,46	0,49	1,70	0,39	1,96
16	0,84	1,09	0,74	1,25	0,63	1,44	0,53	1,66	0,44	1,90
17	0,87	1,10	0,77	1,25	0,67	1,43	0,57	1,63	0,48	1,85
18	0,90	1,12	0,80	1,26	0,71	1,42	0,61	1,60	0,52	1,80
19	0,93	1,13	0,83	1,26	0,74	1,41	0,65	1,58	0,56	1,77
20	0,95	1,15	0,86	1,27	0,77	1,41	0,68	1,57	0,60	1,74
21	0,97	1,16	0,89	1,27	0,80	1,41	0,72	1,55	0,63	1,71
22	1,00	1,17	0,91	1,28	0,83	1,40	0,75	1,54	0,66	1,69
23	1,02	1,19	0,94	1,29	0,86	1,40	0,77	1,53	0,70	1,67
24	1,04	1,20	0,96	1,30	0,88	1,41	0,80	1,53	0,72	1,66
25	1,05	1,21	0,98	1,30	0,90	1,41	0,83	1,52	0,75	1,65
26	1,07	1,22	1,00	1,31	0,93	1,41	0,85	1,52	0,78	1,64
27	1,09	1,23	1,02	1,32	0,95	1,41	0,88	1,51	0,81	1,63
28	1,10	1,24	1,04	1,32	0,97	1,41	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	0,85	1,61
30	1,13	1,26	1,07	1,34	1,01	1,42	0,94	1,51	0,88	1,61
31	1,15	1,27	1,08	1,34	1,02	1,42	0,96	1,51	0,90	1,60
32	1,16	1,28	1,10	1,35	1,04	1,43	0,98	1,51	0,92	1,60
33	1,17	1,29	1,11	1,36	1,05	1,43	1,00	1,51	0,94	1,59
34	1,18	1,30	1,13	1,36	1,07	1,43	1,01	1,51	0,95	1,95
35	1,19	1,31	1,14	1,37	1,08	1,44	1,03	1,51	0,97	1,59
36	1,21	1,32	1,15	1,38	1,10	1,44	1,04	1,51	0,99	1,59
37	1,22	1,32	1,16	1,38	1,11	1,45	1,06	1,51	1,00	1,59
38	1,23	1,33	1,18	1,39	1,12	1,45	1,07	1,52	1,02	1,58
39	1,24	1,34	1,19	1,39	1,14	1,45	1,09	1,52	1,03	1,58
40	1,25	1,34	1,20	1,40	1,15	1,46	1,10	1,52	1,05	1,58
45	1,29	1,38	1,24	1,42	1,20	1,48	1,16	1,53	1,11	1,58
50	1,32	1,40	1,28	1,45	1,24	1,49	1,20	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	1,60
65	1,41	1,47	1,38	1,50	1,35	1,53	1,31	1,57	1,28	1,61
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	1,39	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	1,60	1,36	1,62
85	1,48	1,53	1,46	1,55	1,43	1,58	1,41	1,60	1,39	1,63
90	1,50	1,54	1,47	1,56	1,45	1,59	1,43	1,61	1,41	1,64
95	1,51	1,55	1,49	1,57	1,47	1,60	1,45	1,62	1,42	1,64
100	1,52	1,56	1,50	1,58	1,48	1,60	1,46	1,63	1,44	1,65

Примітки. 1. n-кількістьспостережень; k' - кількість пояснюючих змінних.

Таблиця Б.4 - χ^2-розподіл

df	0,5	1	2,5	5	10	90	95	97,5	99	99,5
1	0,000039	0,00016	0,00098	0,0039	0,0158	2,71	3,84	5,02	6,63	7,88
2	0,0100	0,0201	0,0506	0,1026	0.2107	4,61	5,99	7,38	9,21	10,61
3	0,0717	0,115	0,216	0,352	0,584	6,25	7,81	9,35	11,34	12,84
4	0,207	0,297	0,484	0,711	1,064	7,78	9,49	11,14	13,28	14,86
5	0,412	0,554	0,831	1,15	1,61	9,24	11,07	12,83	15,09	16,75
6	0,646	0,872	1,24	1,64	2,20	10,64	12,59	14,45	16,81	18,55
7	0,989	1,24	1,69	2,17	2,83	12,02	14,07	16,01	18,48	20,28
8	1,34	1,65	2,18	2,73	3,49	13,36	15,51	17,53	20,09	21,96
9	1,73	2,09	2,70	3,33	4,17	14,68	16,92	19,02	21,67	23,59
10	2,16	2,56	3,25	3,94	4,87	15,99	18,31	20,48	23,21	25,19
11	2,60	3,05	3,82	4,57	5,58	17,27	19,68	21,92	24,73	26,76
12	3,07	3,57	4,40	5,23	6,30	18,55	21,03	23,34	26,22	28,30
13	3,57	4,11	5,01	5,89	7,04	19,81	22,36	24,74	27,69	29,82
14	4,07	4,66	5,63	6,57	7,79	21,06	23,68	26,12	29,14	31,32
15	4,60	5,23	6,26	7,26	8,55	22,31	25,00	27,49	30,58	32,80
16	5,14	5,81	6,91	7,96	9,31	23,54	26,30	28,85	32,00	34,27
18	6,26	7,01	8,23	9,39	10,86	25,99	28,87	31,53	34,81	37,16
20	7,43	8,26	9,59	10,85	12,44	28,41	31,41	34,17	37,57	40,00
24	9,89	10,86	10,40	13,85	15,66	33,20	36,42	39,36	42,98	45,56
30	13,79	14,95	16,79	18,49	20,60	40,26	43,77	46,98	50,89	53,67
40	20,71	22,16	24,43	26,51	29,05	51,81	55,76	59,34	63,69	66,77
60	35,53	37,48	40,48	43,19	46,46	74,40	79,08	83,30	88,38	91,95
120	83,85	86,92	91,58	95,70	100,62	140,23	146,57	152,21	158,95	163,64

Таблиця Б.5 - Розподіл Фішера (F-розподіл), $\alpha = 0{,}05$

Ступінь вільності знаменника	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	161	200	216	225	230	234	237	239	241	242	244	246	248	249	250	251	252	253	254
2	18,5	19,0	19,2	19,2	19,3	19,3	19,4	19,4	19,4	19,4	19,4	19,4	19,4	19,5	19,5	19,5	19,5	19,5	19,5
3	10,1	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,74	8,70	8,66	8,64	8,62	8,59	8,57	8,55	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,91	5,86	5,80	5,77	5,75	5,72	5,69	5,66	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,68	4,62	4,56	4,53	4,50	4,46	4,43	4,40	4,37
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,00	3,94	3,87	3,84	3,81	3,77	3,74	3,70	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,57	3,51	3,44	3,41	3,38	3,34	3,30	3,27	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,28	3,22	3,15	3,12	3,08	3,04	3,01	2,97	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,07	3,01	2,94	2,90	2,86	2,83	2,79	2,75	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,91	2,85	2,77	2,74	2,70	2,66	2,62	2,58	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,79	2,72	2,65	2,61	2,57	2,53	2,49	2,45	2,40
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,69	2,62	2,54	2,51	2,47	2,43	2,38	2,34	2,30
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,64	2,60	2,53	2,46	2,42	2,38	2,34	2,30	2,25	2,21
14	4,60	3,74	3,34	3,11	2,96	2,82	2,76	2,70	2,65	2,60	2,53	2,46	2,39	2,35	3,31	2,27	2,22	2,18	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,48	2,40	2,33	2,29	2,25	2,20	2,16	2,11	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	254	2,49	2,42	2,35	2,28	2,24	2,19	2,15	2,11	2,06	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,38	2,31	2,23	2,19	2,15	2,10	2,06	2,01	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,34	2,27	2,19	2,15	2,11	2,06	2,02	1,97	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,31	2,23	2,16	2,11	2,07	2,03	1,98	1,93	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,28	2,20	2,12	2,08	2,04	1,99	1,95	1,90	1,84
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,25	2,18	2,10	2,05	2,01	1,96	1,92	1,87	1,81
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30	2,23	2,15	2,07	2,03	1,98	1,94	1,89	1,84	1,78
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27	2,20	2,13	2,05	2,01	1,96	1,91	1,86	1,81	1,76
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	2,18	2,11	2,03	1,98	1,94	1,89	1,84	1,79	1,73
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24	2,16	2,09	2,01	1,96	1,92	1,87	1,82	1,77	1,71
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	2,09	2,01	1,93	1,89	1,84	1,79	1,74	1,68	1,62
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	2,00	1,92	1,84	1,79	1,74	1,69	1,64	1,58	1,51
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99	1,92	1,84	1,75	1,70	1,65	1,59	1,53	1,47	1,39
120	3,92	3,07	2,68	2,45	2,29	2,18	2,09	2,02	1,96	1,91	1,83	1,75	1,66	1,61	1,55	1,50	1,43	1,35	1,25
∞	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,88	1,83	1,75	1,67	1,57	1,52	1,46	1,39	1,32	1,22	1,00

Таблиця Б.6 - Розподіл Фішера (F-розподіл), $\alpha = 0{,}01$

Ступінь вільності знамен-ника	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	4,052	5,000	5,403	5,625	5,764	5,859	5,928	5,982	6,023	6,056	6,106	6,157	6,209	6,235	6,261	6,287	6,313	6,339	6,366
2	98,5	99,0	99,2	99,2	99,3	99,3	99,4	99,4	99,4	99,4	99,4	99,4	99,4	99,5	99,5	99,5	99,5	99,5	99,5
3	34,1	30,8	29,5	28,7	28,2	27,9	27,7	27,5	27,3	27,2	72,1	26,9	26,7	26,6	26,5	26,4	26,3	26,2	26,1
4	21,2	18,0	16,7	16,0	15,5	15,2	15,0	14,8	14,7	14,5	14,4	14,2	14,0	13,9	13,8	13,7	13,7	13,6	13,5
5	16,3	13,3	12,1	11,4	11,0	10,7	10,5	10,3	10,2	10,1	9,89	9,72	9,55	9,47	9,38	9,29	9,20	9,11	9,02
6	13,7	10,9	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87	7,72	7,56	7,40	7,31	7,23	7,14	7,06	6,97	6,88
7	12,2	9,55	8,45	7,85	7,46	7,19	6,99	6,84	6,72	6,62	6,47	6,31	6,16	6,07	5,99	5,91	5,82	5,74	4,64
8	11,3	8,65	7,59	7,01	6,63	6,37	6,18	6,03	5,91	5,81	5,67	5,52	5,36	5,28	5,20	5,12	5,03	4,95	4,86
9	10,6	8,02	6,99	6,42	6,06	5,80	5,61	5,47	5,35	5,26	5,11	4,96	4,81	4,73	4,65	4,57	4,48	4,40	4,31
10	10,0	7,56	6,55	5,99	5,64	5,39	5,20	5,06	4,94	4,85	4,71	4,56	4,41	4,33	4,25	4,17	4,08	4,00	3,91
11	9,65	7,21	6,22	5,67	5,32	5,07	4,89	4,74	4,63	4,54	4,40	4,25	4,10	4,02	3,94	3,86	3,78	3,69	3,60
12	9,33	6,93	5,95	5,41	5,06	4,82	4,64	4,50	4,39	4,30	4,16	4,01	3,86	3,78	3,70	3,62	3,54	3,45	3,36
13	9,07	6,70	5,74	5,21	4,86	4,62	4,44	4,30	4,19	4,10	3,96	3,82	3,66	3,59	3,51	3,43	3,34	3,25	3,17
14	8,86	6,51	5,56	5,04	4,70	4,46	4,28	4,14	4,03	3,94	3,80	3,66	3,51	3,43	3,35	3,27	3,18	3,09	3,00
15	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80	3,67	3,52	3,37	3,29	3,21	3,13	3,05	2,96	2,87
16	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89	3,78	3,69	3,55	3,41	3,26	3,18	3,10	3,02	2,93	2,84	2,75
17	8,40	6,11	5,19	4,67	4,34	4,10	3,93	3,79	3,68	3,59	3,46	3,31	3,16	3,08	3,00	2,92	2,83	2,75	2,65
18	8,29	6,01	5,09	4,58	4,25	4,01	3,84	3,71	3,60	3,51	3,37	3,23	3,08	3,00	2,92	2,84	2,75	2,66	2,57
19	8,19	5,93	5,01	4,50	4,17	3,94	3,77	3,63	3,52	3,43	3,30	3,15	3,00	2,92	2,84	2,76	2,67	2,58	2,49
20	8,10	5,85	4,94	4,43	4,10	3,87	3,70	3,56	3,46	3,37	3,23	3,09	2,94	2,86	2,78	2,69	2,61	2,52	2,42
21	8,02	5,78	4,87	4,37	4,04	3,81	3,64	3,51	3,40	3,31	3,17	3,03	2,88	2,80	2,72	2,64	2,55	2,46	2,36
22	7,95	5,72	4,82	4,31	3,99	3,76	3,59	3,45	3,35	3,26	3,12	2,98	2,83	2,75	2,67	2,58	2,50	2,40	2,31
23	7,88	5,66	4,76	4,26	3,94	3,71	3,54	3,41	3,30	3,21	3,07	2,93	2,78	2,70	2,62	2,54	2,45	2,35	2,26
24	7,82	5,61	4,72	4,22	3,90	3,67	3,50	3,36	3,26	3,17	3,03	2,89	2,74	2,66	2,58	2,49	2,40	2,31	2,21
25	7,77	5,57	4,68	4,18	3,86	3,63	3,46	3,32	3,22	3,13	2,99	2,85	2,70	2,62	2,53	2,45	2,36	2,27	2,17
30	7,56	5,39	4,51	4,02	3,70	3,47	3,30	3,17	3,07	2,98	2,84	2,70	2,55	2,47	2,39	2,30	2,21	2,11	2,01
40	7,31	5,18	4,31	3,83	3,51	3,29	3,12	2,99	2,89	2,80	2,66	2,52	2,37	2,29	2,20	2,11	2,02	1,92	1,80
60	7,08	4,98	4,13	3,65	3,34	3,12	2,95	2,82	2,72	2,63	2,50	2,35	2,20	2,12	2,03	1,94	1,84	1,73	1,60
120	6,85	4,79	3,95	3,48	3,17	2,96	2,79	2,66	2,56	2,47	2,34	2,19	2,03	1,95	1,86	1,76	1,66	1,53	1,38
∞	6,63	4,61	3,78	3,32	3,02	2,80	2,64	2,51	2,41	2,32	2,18	2,04	1,88	1,79	1,70	1,59	1,47	1,32	1,00

ЗМІСТ