



НАРОДНА УКРАЇНСЬКА АКАДЕМІЯ

ЕКОНОМЕТРИКА

Видавництво НУА

НАРОДНА УКРАЇНСЬКА АКАДЕМІЯ

ЕКОНОМЕТРИКА

Навчальний посібник
для студентів факультету «Бізнес-управління»

Харків
Видавництво НУА
2021

УДК 330.43(076.5+075.8+075.4)

Д18

*Затверджено на засіданні
кафедри інформаційних технологій і математики
Народної української академії
(протокол № 6 від 01.02.2021)*

Р е ц е н з е н т канд. техн. наук В. А. Кірвас.

Данилевич С. Б., Дьячкова О. В.

Д18 Економетрика : навч. посібник для студентів факультету «Бізнес-управління» / С. Б. Данилевич; О. В. Дьячкова ; Нар. укр. акад. [каф. інформ. технологій і математики]. – Харків. Вид-во НУА, 2021. – 60 с.

У посібнику розглянуто зміст розділу «Економетрика» навчальної дисципліни «Економіко-математичні методи і моделі». Видання містить програму і тематичний план, вказано форми контролю та оцінки модулів у балах, наведено методичні рекомендації щодо вивчення тем і виконання завдань, питання для самоперевірки, словник основних термінів, список рекомендованої літератури.

УДК 004.383.1(072+078.5)

ЗАГАЛЬНІ ПОЛОЖЕННЯ

Розділ «Оптимізаційні методи та моделі. Економетрика» навчальної дисципліни «Економіко-математичні методи і моделі» вивчається згідно з навчальним планом підготовки фахівців освітньо-кваліфікаційного рівня «бакалавр».

Предмет економетрики – кількісна оцінка залежності між економічними величинами, тобто розробка моделей та їхнє використання в наукових дослідженнях та управлінні економічними системами.

Мета дисципліни «Оптимізаційні методи та моделі. Економетрика» – навчити підходам до формування економетричних моделей різних рівнів і змісту, принципам економетричного дослідження.

Основні завдання економетрики:

- побудова економетричної моделі;
- оцінка параметрів побудованої моделі, при яких обрана модель найбільш адекватна реальним даним;
- перевірка якості знайдених параметрів моделі і самої моделі в цілому;
- використання побудованих моделей для пояснення поведінки досліджуваних економічних показників, прогнозування, осмисленого проведення економічної політики.

ОРГАНІЗАЦІЙНО-МЕТОДИЧНІ ВКАЗІВКИ

Курс забезпечує підготовку здобувачів вищої освіти за спеціальністю 051 – «Економіка». Він складається з модулів:

модуль 1 «Парна регресія»

модуль 2 «Множинний регресійний аналіз»;

модуль 3 «Моделювання динамічних процесів».

В результаті вивчення дисципліни здобувач має знати:

- мету і завдання економетрики;
- область і ступінь її застосування;
- основні положення, теоретичні основи і методи;
- сутність методу найменших квадратів (МНК) та умови його застосування до оцінювання параметрів лінійної моделі парної регресії;
- інформаційні системи підтримки економетричних досліджень і розрахунків;
- основи регресійного аналізу;
- алгоритм перевірки адекватності регресійної моделі в цілому та значущості її параметрів;
- основи статистичного оцінювання і аналізу точності параметрів рівняння регресії;
- передумови для застосування класичних регресійних моделей;
- методи подолання мультиколінеарності;
- особливості побудови моделей з якісними змінними (dummy-змінними);
- основи аналізу економетричних моделей, що представляють собою системи одночасних рівнянь;

- основні елементи часових рядів;
- основи аналізу і прогнозування часових рядів;

вміти:

- будувати економіко-математичні моделі;
- виконувати постановку та формалізацію задачі економетричного моделювання;
- вирішувати типові завдання економетрики (дослідження розвитку економічних процесів і прогнозування їх динаміки, вибір факторів при побудові математико-статистичних моделей, вибір і побудова математико-статистичної моделі, здійснення модельних експериментів, аналіз отриманих результатів);
- використовувати основні прийоми економетричного дослідження емпіричних даних;
- здійснювати прогнозування за моделями часових рядів і визначати точність прогнозу;
- самостійно працювати з навчально-методичною літературою та комп'ютерними програмами при аналізі застосування економетричних методів.

мати уявлення:

- про основні статистичні інструменти, методи і способи обробки економічних даних;
- про задачі, які можна розв'язувати завдяки застосуванню економетричних моделей;
- щодо статистичних властивостей МНК-оцінок параметрів рівняння множинної регресії, довірчий інтервал для лінії регресії;
- про джерела виникнення мультиколінеарності, її наслідки;
- про автокореляцію рівнів часових рядів.

Набуття компетентностей:

інтегральна:

здатність розв'язувати складні спеціалізовані завдання та практичні проблеми в економічній сфері, які характеризуються комплексністю та невизначеністю умов, що передбачає застосування теорій та методів економічної науки.

загальні:

ЗК8. Здатність до пошуку, оброблення та аналізу інформації з різних джерел.

ЗК11. Здатність приймати обґрунтовані рішення.

фахові (спеціальні) компетентності (СК):

СК4. Здатність пояснювати економічні та соціальні процеси і явища на основі теоретичних моделей, аналізувати і змістовно інтерпретувати отримані результати.

СК7. Здатність застосовувати комп'ютерні технології та програмне забезпечення з обробки даних для вирішення економічних завдань, аналізу інформації та підготовки аналітичних звітів.

СК11. Здатність обґрунтовувати економічні рішення на основі розуміння

закономірностей економічних систем і процесів та із застосуванням сучасного методичного інструментарію.

СК16. Здатність аналізувати та моделювати оцінку економічного ризику для різних професійних сфер та видів економічної діяльності.

Програмні результати, а саме:

ПРН10. Проводити аналіз функціонування та розвитку суб'єктів господарювання, визначати функціональні сфери, розраховувати відповідні показники які характеризують результативність їх діяльності.

ПРН12. Застосовувати набуті теоретичні знання для розв'язання практичних завдань та змістовно інтерпретувати отримані результати.

ПРН13. Ідентифікувати джерела та розуміти методологію визначення і методи отримання соціально-економічних даних, збирати та аналізувати необхідну інформацію, розраховувати економічні та соціальні показники.

ПРН17. Виконувати міждисциплінарний аналіз соціально-економічних явищ і проблем в одній або декількох професійних сферах з врахуванням ризиків та можливих соціально-економічних наслідків.

ПРН25. Розробляти та впроваджувати оптимальні рішення щодо управління розвитком суб'єктів економічної діяльності на основі використання сучасного економіко-математичного інструментарію та цифрових технологій.

Навчальний матеріал базується на теоретичних знаннях і практичних навичках, отриманих в курсах інформатики, теорії ймовірностей, математичної статистики, твердих знаннях у галузі економічної теорії.

Отримані в результаті оволодіння даного модуля вміння і навички можуть бути застосовані в подальшому при вивченні професійно-орієнтованих і спеціалізованих дисциплін, при виконанні розрахункових завдань, лабораторних, курсових і дипломних робіт.

Дисципліна має практичну спрямованість. Для досягнення необхідного рівня практичних знань і умінь передбачається проведення практичних занять у комп'ютерних аудиторіях ХГУ «НУА», самостійне виконання лабораторних і контрольних завдань. Для індивідуального вивчення здобувачам надаються список тем по модулю і питання до них, список навчально-методичної та довідкової літератури, завдання для самостійної роботи.

Загальний обсяг модулів:

	Найменування теми	Кількість годин		
		Усього	Практичні заняття	Самостійна робота
1.	Парна регресія			
1.1.	Лінійна регресія	18	6	12
1.2.	Нелінійна регресія	18	6	12

2.	Множинний регресійний аналіз			
2.1.	Багатофакторні моделі	18	6	12
2.2.	Особливі випадки	18	8	12
3.	Моделювання динамічних процесів			
3.1.	Автокореляція	16	4	12
3.2.	Моделювання динамічних процесів	18	6	12
	Всього за семестр	106	36	72

Перевірка теоретичних знань і практичних навичок передбачає:

- поточний контроль під час занять, при перевірці виконання практичних робіт;
- тестування з використанням комп'ютерних навчально-контролюючих програм;
- поточний контроль у вигляді контрольних завдань.

З метою забезпечення максимальної ефективності практичного освоєння матеріалу дисципліни на кожному занятті і під час індивідуальної роботи кожен студент групи забезпечується в комп'ютерному класі індивідуальним робочим місцем.

Академічні успіхи здобувачів визначаються за рейтинговою Європейською кредитно-трансферною та акумулюючою системою (ECTS), що виглядає наступним чином:

Кількість набраних балів	За чотирирівневою шкалою	За дворівневою шкалою
90 – 100	Відмінно	Зараховано
80 – 89	Добре	
70 – 79	Добре	
60 – 69	Задовільно	
50 – 59	Задовільно	
25 – 49	Незадовільно	Не зараховано (повторне вивчення курсу)
0 – 24	Незадовільно (повторне вивчення курсу)	

Загальна сума балів за модуль дисципліни складається з балів, отриманих за кожну тему. Максимально можливу кількість балів з кожної теми вказано нижче у таблиці «Зміст практичних занять». Максимальна кількість балів за тему виставляється при безпомилковому і своєчасному виконанні практичних робіт і здачі тестів із використанням комп'ютерних контролюючих програм на оцінку «відмінно». Сума максимальних балів по всіх роботах модуля становить 100 балів.

Для успішного освоєння і здачі модулів необхідно набрати за семестр не менш ніж 50 балів за рейтинговою 100-бальною шкалою за умови виконання та успішного складання всіх передбачених програмою робіт. Отримана кількість балів за модуль враховується при академічній атестації студента з дисципліни «Економіко-математичні методи і моделі» за семестр.

Академічна атестація враховує результати освоєння і здачі матеріалу модулів дисципліни, а також підсумковий контроль, що проводиться в екзаменаційну сесію. Підсумкова атестаційна оцінка складається з двох компонент – кількості балів за модулі (внесок 70%), залікова оцінка (внесок 30%). Максимально можлива підсумкова кількість балів для атестації – 100 балів.

ЗМІСТ МОДУЛЯ 1

«Парна регресія»

Тема 1.1. Лінійна регресія.

Основи економетричного моделювання. Статистична база. Принципи побудови економетричних моделей. Парна лінійна регресія. Метод найменших квадратів. Інтерпретація рівняння регресії. Коефіцієнт детермінації. Перевірка на значимість коефіцієнтів регресії і адекватність моделі в цілому.

Тема 1.2. Нелінійна регресія.

Перетворення змінних (базисна процедура). Нелінійні моделі, що зводяться до лінійних. Логарифмічні перетворення. Коефіцієнт еластичності. Випадкові помилки. Порівняння моделей.

Тематичний план модуля 1

Найменування теми		Кількість годин		
		Усього	Практичні заняття	Самостійна робота
1.1.	Лінійна регресія	18	6	12
1.2.	Нелінійна регресія	18	6	12
	Усього за семестр	36	12	24

ЗМІСТ МОДУЛЯ 2

«Множинний регресійний аналіз»

Тема 2.1. Багатофакторні моделі.

Теорема Гауса – Маркова. Етапи і методи побудови моделей множинної регресії. Специфікація змінних у рівняннях регресії. Багатофакторні лінійні моделі. Множинна регресія в нелінійних моделях. Властивості коефіцієнтів множинної регресії. Якість оцінки.

Тема 2.2. Особливі випадки.

Узагальнені економетричні моделі. Особливі випадки в економетричних моделях (гетероскедастичності, мультиколінеарності, автокореляції). Фіктивні

змінні. Множинні сукупності фіктивних змінних. Фіктивні коефіцієнти для коефіцієнта нахилу, тест Чоу.

Тематичний план модуля 2

Найменування теми		Кількість годин		
		Усього	Практичні заняття	Самостійна робота
2.1.	Багатофакторні моделі	18	6	12
2.2.	Особливі випадки	18	8	12
Усього за семестр		36	14	24

ЗМІСТ МОДУЛЯ 3

«Моделювання динамічних процесів»

Тема 3.1. Автокореляція.

Економетричні моделі динаміки. Автокореляція рівнів часового ряду. Автокореляція в залишках. Критерій Дарбіна – Уотсона. Оцінювання коефіцієнтів регресії при наявності автокореляції в залишках.

Тема 3.2. Моделювання динамічних процесів.

Моделювання тенденції, сезонних і циклічних коливань. Використання фіктивних змінних за наявності структурних змін.

Тематичний план модуля 3

Найменування теми		Кількість годин		
		Усього	Практичні заняття	Самостійна робота
3.1.	Автокореляція	16	4	12
3.2.	Моделювання динамічних процесів	18	6	12
Усього за семестр		34	10	24

МОДУЛЬ 1. ПАРНА РЕГРЕСІЯ

Тема 1.1. Лінійна регресія

Кореляційно-регресійний аналіз є основним у вивченні взаємозв'язків явищ. Цей метод містить два складники – кореляційний аналіз і регресійний аналіз.

Кореляційний аналіз – розділ математичної статистики, що вивчає тісноту зв'язку (за допомогою розрахунку коефіцієнтів кореляції) між змінними без їх розділення на факторні і результативні.

Регресійний аналіз – розділ математичної статистики, що вивчає форму залежності між факторними і результативними змінними.

Діаграма розсіювання (кореляційне поле) – графічне подання кореляційного зв'язку (де осі відповідають обом змінним).

Регресійні моделі:

парна (проста) регресія – регресія між двома змінними – y і x , тобто модель вигляду $y = F(x)$, де y – результативна ознака; x – ознака-чинник. Парну регресію будують, якщо є домінуючий чинник, який використовують як пояснюючу змінну;

множинна регресія – регресія результативної ознаки з двома і більшою кількістю чинників, тобто модель вигляду $y = F(x_1, x_2, \dots, x_k)$.

Регресійні моделі з одним рівнянням:

1. Лінійна: $y = a + bx + \varepsilon$.
2. Степенева: $y = a \cdot x^b \cdot \varepsilon$.
3. Гіперболічна: $y = a + bx^{-1} + \varepsilon$.
4. Поліноміальна: $y = a + bx + cx^2 + dx^3 + \dots + \varepsilon$.
5. Експоненційна: $y = e^{a+bx} \times \varepsilon$

та інші.

Специфікація моделі полягає у запису певного типу моделі на основі відповідної теорії взаємозв'язку між змінними.

Статистичні характеристики – показники, що характеризують послідовності $x = (x_1, x_2, \dots, x_n)$ і $y = (y_1, y_2, \dots, y_n)$ (див. Додаток 1).

Завдання за темою 1.1. Лінійна регресія

Завдання 1. Аналіз залежності між змінними

Створіть нову книгу *Excel*. У першому рядку вкажіть своє прізвище, групу, дату виконання. Файл збережіть під назвою «*Прізвище ЕММ1*» у своїй папці.

1. Порівняйте (письмово) діяльність двох менеджерів, контролюючих залишки товарів в двох ваших магазинах. Хто з них краще управляє запасами або обоє працюють однаково?

Стоки останніх шести тижнів:

магазин 1	33	31	32	36	31	31
магазин 2	22	34	58	52	10	21

Розрахуйте і порівняйте середні значення стоків. Як вони характеризують роботу менеджерів?

Розрахуйте і порівняйте стандартні відхилення. Як вони характеризують роботу менеджерів?

Розрахуйте і порівняйте коефіцієнти варіації. Як вони характеризують роботу менеджерів?

Формули для розрахунків наведено в Додатку 1. Зробіть висновки про роботу менеджерів.

2. Дана таблиця споживання м'яса:

Поголів'я великої рогатої худоби (млн т)	Обсяг виробництва молока (тис. т)	Обсяг виробництва м'яса (тис. т)	Споживання м'яса (у кг на душу населення за рік)
57	1,49	8,37	69
54,7	1,38	8,26	60
52,2	1,29	7,51	69
48,9	1,1	6,8	57
43,3	0,99	5,79	55
39,7	0,9	5,33	51
35,1	0,88	4,85	50

2.1. Проаналізуйте споживання м'яса:

а) Розрахуйте середнє, вибіркoву дисперсію, стандартне відхилення, коефіцієнт варіації. Інтерпретуйте ці показники і зробіть висновок (письмово) про однорідність даних.

б) Запустіть *Пакет аналізу / Описова статистика* (виберіть підсумкову статистику). Порівняйте (усно) результати з вашими розрахунками.

2.2. З'ясуйте, чи існує зв'язок між поголів'ям КРХ і виробництвом молока (письмово). Для цього:

а) Розрахуйте коваріацію і коефіцієнт кореляції. Інтерпретуйте ці показники.

б) Побудуйте діаграму розсіяння.

в) Викличте *Пакет аналізу / Коваріація*, а потім *Кореляція* для тих же даних. Порівняйте результати з вашими розрахунками.

2.3. Проведіть аналіз залежності між собою усіх початкових даних за допомогою *Пакету аналізу / Кореляція*.

Зробіть письмові висновки.

Завдання 2. Статистична база економетричних моделей

1. Вибрати вихідні дані для свого варіанту з Додатка 2. Скопіювати їх до файлу на аркуш «База».
2. За вихідними даними побудувати кореляційне поле за допомогою *Майстер діаграм / Точкова*. За змінну X візьміть споживчі витрати світу, а Y – витрати будь-якої країни.
3. Для даних X і Y розрахувати середні значення.
4. Створити нові стовпці: $(X - X_{\text{сеп}})^2$, $(Y - Y_{\text{сеп}})^2$, $(X - X_{\text{сеп}})(Y - Y_{\text{сеп}})$. Розрахувати відповідні значення у цих стовпцях.
5. Знайти середні по цих стовпцях, позначити їх відповідно як дисперсії X і Y , коваріацію.
6. Знайти коефіцієнт кореляції: $r = \text{cov}(X, Y) / (D(X) \cdot D(Y))^{1/2}$.
7. Зробити висновок про кореляційний зв'язок між X і Y .
8. Результат збережіть у своїй папці і скопіюйте в папку *\$control*

Завдання 3. Побудова лінійної економетричної моделі

1. Скопіювати Ваш файл попередньої роботи *Прізвище ЕММ1.xlsx*. Перейменувати аркуш «База» у «ЕММ1». Додати новий аркуш «ЕММ2», вказати на ньому знову прізвище, групу, поточну дату виконання роботи. Файл зберегти під назвою *Прізвище ЕММ2* в своїй папці.
2. Зробити копію аркуша «ЕММ1», назвати його «Лінійний тренд».
 - 2.1. Для наявної точкової діаграми побудувати лінійний тренд. Задати відображення рівняння на діаграмі. Налаштувати масштаб відображення шкали на діаграмі так, щоб обидві лінії займали всю область побудови діаграми. Задати для ліній кольори: синій і червоний.
 - 2.2. Нижче діаграми вписати (або скопіювати) рівняння отриманого лінійного тренду.
3. Зробити копію аркуша «ЕММ1», назвати його «Метод МНК».
 - 3.1. Видалити з нього всі дані, крім вихідних (стовпців X і Y). Скопіювати з презентації *ЕММ2.pptx* формули (зображення) розрахунку коефіцієнтів лінійної регресії.
 - 3.2. Розрахувати за цими формулами значення коефіцієнтів.
 - 3.3. Нижче вписати рівняння лінійної регресії з отриманими коефіцієнтами.
4. Нижче діаграми вписати (або скопіювати) рівняння отриманого лінійного тренду.
5. Додати новий аркуш «Лінійна». Скопіювати на нього вихідні дані (стовпці X і Y) з попереднього аркуша.
 - 5.1. Розрахувати коефіцієнти лінійної регресії за допомогою вбудованої статистичної функції *ЛИНЕЙН*. (Підказка. Не забудьте, що це функція масивів, вводиться за допомогою *Ctrl + Shift + Enter*). Кожен отриманий коефіцієнт забезпечте пояснювальною написом.
 - 5.2. Нижче напишіть рівняння лінійної регресії з отриманими коефіцієнтами.
6. Додати новий аркуш «Пакет аналізу». Скопіювати на нього вихідні дані (стовпці X і Y) з попереднього аркуша.
 - 6.1. За допомогою інструменту аналізу даних *Регресія* проведіть регресійний аналіз даних. Результати розташуйте нижче на тому ж аркуші.
 - 6.2. Осередки, що містять розраховані значення коефіцієнтів, виділіть жовтим кольором. Забезпечте їх приміткою з описом кожного коефіцієнта.
 - 6.3. Нижче напишіть рівняння лінійної регресії з отриманими коефіцієнтами.
 - 6.4. Знайдіть серед результатів осередок, що містить значення кореляції, виділіть її червоним кольором. Забезпечте її також приміткою з описом.
7. Порівняйте (письмово, нижче) отримані значення коефіцієнтів регресії і коефіцієнта кореляції з результатами на аркуші «ЕММ1».
8. Результат збережіть у своїй папці і скопіюйте в папку *\$control*.

Завдання 4. Оцінка значущості параметрів і адекватності лінійної економетричної моделі

1. Скопіювати Ваш файл попередньої роботи *Прізвище ЕММ2.xlsx*. Зберегти його під назвою *Прізвище ЕММ3.xlsx*. На аркуші «Пакет аналізу» забезпечити примітками осередки:
 - розрахованих коефіцієнтів b_0, b_1 ;
 - значення кореляції;
 - коефіцієнта детермінації;
 - $t(b_0), t(b_1)$;
 - F - критерій Фішера.
2. Написати висновок за коефіцієнтом кореляції і коефіцієнтом детермінації.
3. Оцінити значущість параметрів. Написати висновок.
4. Оцінити адекватність моделі в цілому (значущість коефіцієнта детермінації). Написати висновок.
5. Розрахувати $y_{оцін.}$. На одному рисунку побудувати кореляційне поле даних (x, y) і пряму $(x, y_{оцін.})$.
6. Результат збережіть у своїй папці і скопіюйте в папку *\$control*.

Тема 1.2. Нелінійна регресія

Нелінійна регресія – окремий випадок регресійного аналізу, в якому регресійна модель, що розглядається, – це функція, залежна від параметрів і від однієї або декількох вільних змінних. Залежність від параметрів передбачається нелінійною.

Вибір виду регресії

Для парної регресії застосовують три способи вибору виду математичної функції:

- графічний
- аналітичний (виходячи з теорії досліджуваного взаємозв'язку);
- експериментальний

Види нелінійних моделей

- регресії, нелінійні за включеними до аналізу пояснювальними змінними, але лінійні за оцінюваними параметрами;
- регресії, нелінійні за оцінюваними параметрами.

Моделі, нелінійні за змінними

Приклади моделей, нелінійних за змінними

- Поліноміальна:

$$y = b_0 + b_1x + b_2x^2 + \varepsilon.$$

$$y = b_0 + b_1x + b_2x^2 + b_3x^3 + \varepsilon.$$

Параболу 2-го ступеня зазвичай застосовують, коли на деякому інтервалі змінюється характер зв'язку змінних – наприклад, прямий зв'язок замінюється зворотним.

- Гіперболічна: $y = b_0 + b_1/x + \varepsilon$.

Гіперболу використовують для характеристики зв'язку питомих витрат сировини, матеріалів, палива від обсягу випуску, часу обігу товарів від товарообігу, приросту зарплати від рівня безробіття (крива Філіпса), витрат на непродуктивні товари від доходів або загальної суми витрат (криві Енгеля) тощо.

- Квадратного кореня: $y' = b_0 + b_1 \cdot (x)^{1/2} + \varepsilon$.
- Напівлогарифмічна: $y' = b_0 + b_1 \cdot \ln x + \varepsilon$.

Лінеаризація моделей, нелінійних за змінними

Регресія, нелінійна за змінними, але лінійна за параметрами, може бути зведена до лінійної регресії шляхом заміни змінних.

Заміною $z_i = \varphi(x_i)$, $i = 1, \dots, n$ нелінійна парна регресія $y = b_0 + b_1 \cdot \varphi(x)$ приводиться до лінійної парної регресії $y = b_0 + b_1 \cdot z$.

$$y' = b_0 + b_1 \cdot \varphi(x) \quad \Leftrightarrow \quad z = \varphi(x) \quad \Leftrightarrow \quad y' = b_0 + b_1 \cdot z.$$

Моделі, нелінійні за параметрами

Деякі моделі, нелінійні за параметрами, також можуть бути зведені до лінійних. Вони наводяться до лінійного вигляду за допомогою перетворень – наприклад, логарифмування.

Приклади нелінійних за параметрами моделей, внутрішньо лінійних

- Степенева: $y = b_0 x^{b_1} \cdot \varepsilon$

Широко використовується, тому що коефіцієнт b має чітке економічне пояснення – це коефіцієнт еластичності.

- Експоненційна: $y = e^{b_0 + b_1 x} \cdot \varepsilon$.
- Показникова: $y = b_0 \cdot b_1^x \cdot \varepsilon$.
- Логістична: $y = \frac{b_0}{1 + b_1 e^{-b_2 x + \varepsilon}}$
- Зворотна: $y = \frac{1}{b_0 + b_1 x + \varepsilon}$.

Лінеаризація моделей, нелінійних за параметрами, але внутрішньо лінійних

Нелінійна парна регресія іноді може бути приведена до лінійної за допомогою логарифмування:

етап 1) логарифмування:

$$y = b_0 \cdot b_1^x \quad \Leftrightarrow \quad \ln y = \ln b_0 + x \ln b_1 \quad (\text{для показникової})$$

$$y = b_0 \cdot x^{b_1} \quad \Leftrightarrow \quad \ln y = \ln b_0 + b_1 \ln x \quad (\text{для степеневі})$$

етап 2) послідовна заміна величин і параметрів:

$$\ln y = y_1, \ln b_0 = a_0, \ln b_1 = a_1 \quad \Leftrightarrow \quad y_1 = a_0 + a_1 x \quad (\text{для показникової})$$

$$\ln x = z \quad \Leftrightarrow \quad y_1 = a_0 + b_1 z \quad (\text{для степеневі})$$

етап 3) потенціювання для визначення специфікації моделі:

$$y' = e^{\ln y'} = e^{y' \cdot 1} = e^{a_0 + a_1 x} \quad (\text{для показникової})$$

$$y' = e^{\ln y'} = e^{y' \cdot 1} = e^{a_0 + b_1 z} \quad (\text{для степеневий})$$

Приклади нелінійних за параметрами моделей і внутрішньо нелінійних

Не приводяться до лінійного вигляду:

$$y = b_0 + b_1 x^{b_2}; \quad y = 1 - 1/(1 - x \cdot b)$$
 тощо.

Коефіцієнт еластичності

Коефіцієнт еластичності показує, наскільки зміниться $y = f(x)$ при зміні x на 1%. Для парної нелінійної регресії його розраховують як відносну зміну y до відносної зміни x :

$$e = \left(\frac{dy}{y} \right) / \left(\frac{dx}{x} \right) = y'(x) \cdot \frac{x}{y}$$

Завдання за темою 1.2. Нелінійна регресія

Завдання 1. Оцінка параметрів економетричних моделей, нелінійних за змінними

1. Скопіювати Ваш файл попередньої роботи *ЕММ3.xlsx*. Зберегти його під назвою *ЕММ4.xlsx*. Перейменувати в копії аркуш «Пакет аналізу» у «Лінійна».
2. Додати аркуш «Гіпербола». Скопіювати дані y і x . Додати новий заголовок $1/x$ та розрахувати дані в ньому. Оцінити лінійну модель для даних y і $1/x$ на тому ж аркуші. Написати короткі висновки (за зразком лінійної регресії).
3. Новий аркуш назвати «Парабола». Скопіювати дані y і x . Додати новий заголовок x^2 і розрахувати відповідні значення. Оцінити лінійну модель для даних y і x^2 на тому ж аркуші. Написати короткі висновки.
4. Новий аркуш назвати «Корінь». Скопіювати дані y і x . Додати новий заголовок $x^{1/2}$. Оцінити лінійну модель для даних y і $x^{1/2}$ на тому ж аркуші. Написати короткі висновки.
5. Назвати новий аркуш «Логарифм». Скопіювати дані x і y . Додати новий заголовок. Оцінити лінійну модель для даних x і y на тому ж аркуші. Написати короткі висновки.
6. Для кожної моделі розрахувати $y_{\text{оцін}}$ і побудувати графіки: кореляційне поле даних (x, y) і $(x, y_{\text{оцін}})$.
7. Порівняти лінійну і нелінійні моделі. Сформулювати висновки на аркуші «Порівняння».
8. Результат зберегти у своїй папці, копію відправити в папку \$control.

Завдання 2. Оцінка параметрів економетричних моделей, нелінійних за параметрами

1. Скопіювати Ваш файл попередньої роботи *ЕММ4.xlsx*. Зберегти його під ім'ям *ЕММ5.xlsx*.
2. Назвати новий аркуш «Показникова». Скопіювати вихідні дані y і x . Додати новий заголовок $\ln y$, розрахувати $\ln y$. Оцінити лінійну модель $\ln y = a_0 + a_1 x$ для даних $\ln y$ і x на тому ж аркуші. Написати короткі висновки.
3. Додати новий заголовок $\ln y_{\text{оцін}}$. Розрахувати $\ln y_{\text{оцін}} = a_0 + a_1 x$.
4. Додати новий заголовок $\ln y_{\text{оцін}}$. Розрахувати $y_{\text{оцін}} = \exp(\ln y_{\text{оцін}})$.
5. Додати новий заголовок $(y - y_{\text{оцін}})^2$. Розрахувати $(y - y_{\text{оцін}})^2$. Знайти суму цього стовпчика. Проаналізувати її значення (письмово).
6. Побудувати графіки: кореляційне поле даних (x, y) і $(x, y_{\text{оцін}})$.
7. Назвати новий аркуш «Степенева». Скопіювати дані y и x . Додати нові заголовки $\ln y$ і $\ln x$, розрахувати $\ln y$ і $\ln x$. Оцінити лінійну модель $\ln y = a_0 + b_1 \ln x$ для даних $\ln y$ і $\ln x$ на тому ж аркуші. Написати короткі висновки.
8. Додати новий заголовок $\ln y_{\text{оцін}}$. Розрахувати $\ln y_{\text{оцін}} = a_0 + a_1 \ln x$.
9. Додати новий заголовок $\ln y_{\text{оцін}}$. Розрахувати $y_{\text{оцін}} = \exp(\ln y_{\text{оцін}})$.
10. Додати новий заголовок $(y - y_{\text{оцін}})^2$. Розрахувати $(y - y_{\text{оцін}})^2$. Знайти суму цього стовпчика. Проаналізувати її значення (письмово).
11. Побудувати графіки: кореляційне поле даних (x, y) і $(x, y_{\text{оцін}})$.
12. Перемістити створений раніше аркуш «Порівняння» напри кінець книги. Додати на цей аркуш таблицю для порівняння даних:

	A	B	C	D	E	F	G
1	Прізвище	Ім'я	Дата				
2	Порівняння моделей						
3	Лінійна	F	t(b₀)	t(b₁)	R²	SSE	
4	Гіпербола						
5	Парабола						
6	Корінь						
7	Логарифм						
8	Показникова						
9	Степенева						
10							

Заповнити цю таблицю (за зразком):

	A	B	C	D	E	F	G
1	Прізвище Ім'я Дата						
2	Порівняння моделей						
3	Лінійна	F	t(b₀)	t(b₁)	R²	SSE	
4	Гіпербола	адекв.	не знач.	знач.	=Лінійна!B17	=Лінійна!C25	
5	Парабола	адекв.	знач.	знач.			
6	Корінь						
7	Логарифм						
8	Показникова						
9	Степенева						
10							

13.Зробити загальні висновки за побудованими моделями, вказати найкращу з них (виписати цю модель), обґрунтувати її вибір (письмово).

14.Результат зберегти у своїй папці, копію відправити в папку \$control.

МОДУЛЬ 2. МНОЖИННИЙ РЕГРЕСІЙНИЙ АНАЛІЗ

Тема 2.1. Багатофакторні моделі

Мета – навчитися будувати, оцінювати і порівнювати моделі з великим числом факторів, визначати при цьому вплив кожного з них окремо, а також сукупний їхній вплив на модельований показник.

Множинна регресія – регресія результативної ознаки з двома або більшою кількістю чинників, тобто модель вигляду

$$y = F(x_1, x_2, \dots, x_k).$$

Для вибору специфікації моделі необхідно провести:

- 1) відбір факторів для побудови моделі;
- 2) вибір рівняння моделі.

Вимоги до факторів моделі:

- фактори повинні бути кількісними (якщо чинники якісні, слід придумати для них кількісні відповідності);
- не повинна спостерігатися висока кореляція (і тим паче точний функціональний зв'язок) між пояснюючими факторами.

Відбір факторів для моделі множинної регресії:

Фактори повинні пояснювати варіацію результативної ознаки. Частку поясненої ними варіації показує коефіцієнт детермінації R^2 . Вплив інших (не врахованих в моделі) факторів оцінюється як $1-R^2$. Якщо до моделі додати ще один фактор, то коефіцієнт детермінації повинен зрости. Якщо ж це не так, то доданий фактор не поліпшив модель і від нього треба відмовитися. (Зайві фактори призводять до статистичної незначущості параметрів регресії за t -критерієм Стьюдента.)

Зазвичай кількість пояснюючих чинників рекомендують вибирати в 6-7 разів менше, ніж кількість спостережень. Інакше параметри рівняння регресії і рівняння в цілому будуть статистично незначущі.

Мультиколінеарність

Дві змінні вважаються колінеарними, якщо коефіцієнт їхньої кореляції $|r_{xy}| \geq 0,7$ (між цими змінними існує лінійна залежність). Для множинної регресії фактори вважаються мультиколінеарними, якщо коефіцієнт їх кореляції $|r_{x_i, x_j}| \geq 0,7$. Для їх виявлення розраховують матрицю парних кореляцій (наприклад, за допомогою модуля *Кореляція* пакету аналізу *MS Excel*).

Мультиколінеарність особливо часто має місце при аналізі макроекономічних даних (наприклад, при аналізі доходів, виробництва). Кореляція між факторами існує завжди; важливо, щоб вона не була занадто сильною.

Якщо модель побудована із включенням мультиколінеарних факторів, у ній можуть бути знаки або великі значення, що не відповідають економічній теорії, всі змінні можуть виявитися незначущими, а вся модель в цілому при цьому значущою тощо. Отримувані оцінки – наприклад, коефіцієнти кореляції – виявляються нестабільними як в сенсі статистичної значущості, так і за величиною і знаком, отже, вони ненадійні.

Мультиколінеарні фактори мають бути замінені, перетворені або один із них виключений. При цьому виборі перевага віддається не тому фактору, що тісніше пов'язаний із результативною ознакою, а тому, що при тісному зв'язку з результативною ознакою найменш тісно пов'язаний із іншими факторами.

Способи усунення мультиколінеарності

- можна збільшити обсяг вибірки (чим більше даних, тим менше дисперсія оцінок МНК). Однак зазвичай додаткові дані нелегко знайти;
- замінити мультиколінеарні фактори або перетворити їх – додати, розділити, замінити їх середнім тощо, але при цьому отримані факторні змінні повинні мати економічний сенс;
- часто виключають змінні, висококорельовані з іншими. Однак якщо будь-які чинники було включено до моделі на основі рекомендацій економічної теорії, їх не слід виключати.

При цьому перевагу в моделі віддають не тому фактору, що тісніше пов'язаний з результативною ознакою, а тому, що при тісному зв'язку з результативною ознакою найменш тісно пов'язаний з іншими факторами.

Методи побудови моделей множинної регресії:

- метод включення;
- метод виключення.

Метод включення: будують модель з однієї пояснюючої змінною. Потім додають ще один фактор, будують нову модель, і обидві моделі порівнюють. Якщо модель покращилася, додають ще один фактор. Якщо модель погіршилася – доданий фактор замінюють на інший. Після перебору різних комбінацій пояснюючих чинників отримують найкращу модель.

Метод виключення: будують і досліджують модель з усіма факторами. Потім видаляють один фактор, будують і досліджують нову модель і порівнюють із попередньою. Вибирають кращу модель, і процес триває. Для вибору фактору, що видаляється:

- будують матрицю парних коефіцієнтів кореляції;
- вибирають два фактори з найбільшим коефіцієнтом парної кореляції;
- із цих факторів вибирають той, який має менший коефіцієнт кореляції з результируючою ознакою.

Порівняння моделей:

- якщо кількість факторів у моделях однакова, то порівнюють за коефіцієнтом детермінації;
- якщо до моделі включено різну кількість чинників, то порівнюють за нормованим коефіцієнтом детермінації.

Вибір рівняння моделі

Моделі множинної регресії (як і парної регресії) можуть бути лінійними і нелінійними – як за змінними, так і за параметрами. Для оцінки деяких нелінійних моделей їх зводять до лінійних (в тому числі за допомогою логарифмування).

Часто використовують лінійну і степеневу моделі регресії. Вони дозволяють дати хорошу інтерпретацію коефіцієнтів регресії.

Наприклад, у лінійної регресії

$$y = b_0 + b_1 x_1 + b_2 x_2 + \dots + b_n x_n$$

параметр b_i при x_i характеризує середню зміну результату y при зміні фактору x_i на одиницю і при незмінному значенні інших факторів, закріплених на середньому рівні.

Оцінка побудованої моделі

Якість побудованої моделі оцінюють за коефіцієнтом детермінації R^2 . Для лінійної регресії він збігається з квадратом коефіцієнта множинної кореляції. Його коректують, щоб врахувати вплив числа факторів в моделі.

Скоригований (нормований) коефіцієнт детермінації:

$$R_{adj}^2 = 1 - (1 - R^2) \frac{n - 1}{n - m - 1} = R^2 - \frac{m}{n - m - 1} (1 - R^2)$$

(n – кількість факторів, m – кількість спостережень).

Значущість рівняння множинної регресії в цілому оцінюють за критерієм Фішера:

$$F = R^2(n - m - 1) / ((1 - R^2)m).$$

Оцінку значущості коефіцієнтів регресії здійснюють (як і в парній регресії) за t -критерієм Стьюдента.

Для лінійної регресії параметр b_i при x_i показує, на скільки одиниць зміниться величина залежної змінної y при зміні значення i -й описовій змінної x_i на

одну одиницю при інших рівних умовах (всі інші описові змінні дорівнюють своїм середнім значенням).

Завдання за темою 2.1. Багатофакторні моделі

Завдання 1. Побудова багатофакторної моделі

1. Створіть нову книгу Excel *Прізвище ЕММ5*. Укажіть прізвище, групу, назву роботи, дату виконання.

Реалізуйте приклад, що розглянуто нижче.

Необхідно дослідити для 12 муніципальних і приватних автотранспортних підприємств залежність чистого прибутку за рік (Y) від кількості автобусів (X_1 , шт.), маршрутних таксі (X_2 , шт.) і форми власності підприємства (див. табл.).

Порядковий № підприємства	Прибуток, млн грн.	Кількість автобусів, шт.	Кількість маршрутних таксі, шт.	Форма власності підприємства
	Y	X_1	X_2	
1	2241	48	18	муніципальне
2	3818	58	16	приватне
3	2905	42	25	муніципальне
4	3237	58	27	муніципальне
5	4150	52	30	приватне
6	5727	56	42	приватне
7	1328	27	21	муніципальне
8	2905	52	21	муніципальне
9	5561	75	37	приватне
10	3652	48	35	приватне
11	4648	58	33	приватне
12	3486	48	21	муніципальне

1. Уведіть фіктивну змінну X_3 : 0 – муніципальне АТП, 1 – приватне АТП.

2. Перевірте фактори на колінеарність (для цього обчисліть матрицю парних коефіцієнтів кореляції за допомогою модуля *Кореляція* пакета аналізу Excel):

	Y	X_1	X_2	X_3
Y	1			
X_1	0,804124189	1		
X_2	0,773166796	0,443934029	1	
X_3	0,781020866	0,550469012	0,633441653	1

Жоден із коефіцієнтів кореляції між факторами не перевищує за модулем 0,8, що свідчить про відсутність колінеарності. Можна

3. Побудуйте трьохфакторну модель.

В отриманій регресії значущість коефіцієнта b_0 (вільного члена) 0,20492 перевищує 0,05. Виключіть його із розгляду і побудуйте регресію ще раз, задавши в діалоговому вікні *Регресія* опцію *Константа – нуль*.

Проаналізуйте значення коефіцієнтів детермінації ($R^2 = 0,98733324$; $R^2 = 0,899098534$). Зробіть висновки.

Оцініть отримане рівняння лінійної регресії ($Y_{\text{розрах}} = 38,26177111 \cdot x_1 + 44,8077191 \cdot x_2 + 938,5459397 \cdot x^3$).

Зробіть (письмово) висновки про смислове значення кожного коефіцієнта. (Кожен додатковий автобус приносить в середньому 38,262 млн прибутку, кожне маршрутне таксі – 938,546 млн, при однаковому складі автопарків прибуток приватного ($x_3=1$) більше в середньому на 44,808.)

Випишіть із отриманого рівняння ще два – окремо для кожного типу підприємства (підставив $x_3=0$ і $x_3=1$).

Результат збережіть у своїй папці і скопіюйте в папку *\$control*.

Завдання 2. Прогнозування на підставі багатofакторної моделі

Будівельна компанія має намір за 2 роки здійснити будівництво 16-поверхового житлового будинку на 192 квартири загальною площею 11600 кв. м. Оцініть вартість будівництва на підставі даних за аналогічними завершеними будівництвами в тому ж районі. Відомі:

y – вартість будівництва (тис. у. о.);

x_1 – загальна площа (кв. м);

x_2 – кількість квартир у будинку;

x_3 – кількість поверхів;

x_4 – тривалість будівництва (міс.).

Вихідні дані надано у таблиці.

№	x_1	x_2	x_3	x_4	y
1	39596	616	22	56	12998
2	7227	192	16	20	7899,5
3	19227	480	20	34	9184,8
4	20728	288	24	35	11289
5	10056	192	12	23	5984,8
6	14612	336	14	29	7124,2
7	23721	448	16	38	9898,6
8	22795	560	20	37	11116
9	35720	480	20	51	13257
10	3522	112	14	16	5686,1
21	20356	264	22	34	9095,6

№	x_1	x_2	x_3	x_4	y
11	14088	384	16	28	7888,8
12	4553	144	18	17	7539,7
13	7045	224	14	20	6233,5
14	41605	672	24	58	13754
15	27976	616	22	43	11226
16	5869	96	12	19	6216,4
17	20212	480	20	35	10725
18	19856	336	12	34	7052,9
19	28490	480	20	43	9211,4
20	6981	128	16	20	7378,4
22	3984	96	12	17	5914,5

Передбачається, що існує лінійна залежність між незалежними змінними (x_1, x_2, x_3 і x_4) та залежною змінною (y).

1. Організуйте дані у вигляді однієї вертикальної таблиці.

2. Побудуйте матрицю коефіцієнтів кореляції, перевірте наявність або відсутність мультиколінеарності.

3. Визначте коефіцієнти лінійного багатофакторного рівняння регресії методом виключення. (Зверніть увагу, що при включенні відразу всіх пояснюючих чинників у модель майже всі параметри регресії виявляться незначущими – хоча є сильна кореляція усіх факторів з величиною y .)

Для кожного етапу побудови регресії визначайте значущість коефіцієнтів і адекватність моделі. Укажіть коротко ваші дії на кожному наступному етапі. Використовуйте для аналізу нормований коефіцієнт детермінації.

4. Запишіть вид отриманої багатофакторної залежності. Поясніть економічний сенс параметрів моделі.

5. Дайте оцінку вартості майбутнього будівництва вказаного будинку за допомогою отриманої моделі.

Завдання 3. Аналіз даних за допомогою багатофакторної моделі

Провести дослідження причин і виявити фактори, що мають найбільший вплив на збитки виробництва шоколадних виробів кондитерських підприємств. Несуттєві факторні ознаки виключити зворотним методом покрокової регресії.

Дані підприємства (організуйте їх у вигляді однієї таблиці):

x_1 – відсоток реалізованої за місяць шоколадної продукції (%);

x_2 – вартість 1 тони сировини, зокрема какао-бобів (грн.);

x_3 – витрата сировини (какао-бобів) на одну тонну кондитерських виробів (%);

x_4 – витрати на 1 грн. виробленої продукції (грн.);

x_5 – витрати на електроенергію (грн.);

y – збитки підприємства від даного виробництва.)

№ міс	1	2	3	4	5	6	7	8	9	10
y	93376	72009	58908	70653	65531	100742	69252	86457	68589	66770
x_1	47,6	64,9	83,2	63,1	70,6	46,9	67,1	47	70	71,1
x_2	5000	4160	3000	4100	3900	5100	4000	4800	3900	3900
x_3	0,78	0,68	0,653	0,68	0,678	0,78	0,68	0,67	0,67	0,67
x_4	0,9	0,76	0,699	0,74	0,71	0,9	0,73	0,85	0,72	0,7
x_5	1144	623	594	676	589	1171	667	884	598	658

№ міс	11	12	13	14	15	16	17	18	19	20
y	90170	75987	79888	71673	90106	85259	91735	89593	50848	73742
x_1	48,6	46,9	52,1	65,4	48,9	48	49,9	49,3	88,7	61,2
x_2	4900	4180	4300	4150	4900	4500	5000	4900	2600	4170

x_3	0,78	0,67	0,67	0,68	0,78	0,68	0,78	0,78	0,55	0,67
x_4	0,87	0,78	0,8	0,75	0,89	0,82	0,9	0,86	0,689	0,77
x_5	908	625	825	642	862	833	1055	872	928	672

№ міс	21	22	23	24	25	26	27	28	29	30
y	80502	83589	91959	79944	67005	83238	100484	100398	110071	110534
x_1	49,7	51,1	50,1	53,1	69	46,9	50,5	48,6	47,8	49,6
x_2	4200	4700	4900	4600	3900	4400	5200	5300	5400	5500
x_3	0,68	0,66	0,78	0,66	0,67	0,66	0,78	0,78	0,78	0,78
x_4	0,79	0,84	0,88	0,83	0,71	0,81	0,9	0,9	0,9	0,9
x_5	642	860	914	811	595	805	1134	1110	1131	1147

Тема 2.2. Особливі випадки

Фіктивні змінні

Щоб додати до розгляду якісні дані (освіта, стать, район, сезон, тип продукції або підприємства тощо), треба перетворити їх у кількісні змінні, які називають **фіктивними**.

Зазвичай фіктивна змінна має вигляд:

$D = 1$, якщо фактор зачення₁;

$D = 0$, якщо фактор зачення₂.

Якщо фактор може мати k варіантів значень, слід вводити $k-1$ фіктивних змінних, що приймають значення 1 або 0 – так щоб у сукупності набори цих 1 і 0 описували всі варіанти значень факторної ознаки.

Значення якісного фактору, при якому всі фіктивні змінні рівні 0, називають базовим значенням.

Фіктивні змінні можуть входити в регресійну модель як адитивні (окремими складовими):

$$y = b_0 + b_1 \cdot x + b_2 \cdot D + \varepsilon \quad (\text{адитивна фіктивна змінна})$$

так і мультиплікативні:

$$y = b_0 + b_1 \cdot x + b_2 \cdot D + b_3 \cdot x \cdot D + \varepsilon \quad (\text{мультиплікативна фіктивна змінна})$$

Мультиплікативні фіктивні змінні впливають не тільки на вільний член, але і на коефіцієнт нахилу регресії і дозволяють врахувати вплив якісних факторів на кількісні.

Тест Чоу

Щоб з'ясувати, чи треба вводити до отриманої моделі регресії додаткові фіктивні змінні, застосовують метод (або тест) Чоу. Його використовують у тих випадках, коли вихідні дані можна було б розділити на частини (підвибірки), або, навпаки, коли можна було б об'єднати дві вибірки даних в одну і розглядати

єдину модель регресії. Тоді тест Чоу дозволяє перевірити, чи краще якість загальної моделі регресії, ніж якості часткових моделей підвибірок, чи ні.

Суть тесту Чоу:

1. Висувають гіпотезу про те, що краще якість загальної моделі. Її перевіряють за допомогою F-критерію Фішера: розраховують значення критерію

$$F_{\text{розрах}} = \frac{SSE - (SSE_1 + SSE_2)}{SSE_1 + SSE_2} \frac{n-2m-2}{m+1},$$

де SSE , SSE_1 , SSE_2 – залишкові суми квадратів для всієї вихідної вибірки і для обох її частин відповідно;

$m+1$ – кількість факторів;

$n-2m-2$ – кількість спостережень вибірки.

2. Порівнюють отриманий результат із критичним значенням F критерію $F_{кр}$ – його визначають за таблицею розподілу Фішера з рівнем значущості α (зазвичай 0,05) і двома ступенями вільності $m+1$ і $n-2m-2$.

Якщо $F_{\text{розрах}} \leq F_{кр}$, то основну гіпотезу приймають, базисна модель є оптимальною і розбивати загальну регресію на підвибірки не має сенсу.

Якщо ж навпаки, $F_{\text{розрах}} > F_{кр}$, то основну гіпотезу відхиляють і необхідно будувати дві моделі. Побудову регресії для кожної підвибірки можна провести, додавши до загальної моделі набір фіктивних змінних, що визначають розбиття вибірки на частини.

Умови для побудови регресійних моделей

Умови Гауса – Маркова для побудови класичних регресійних моделей, лінійних за коефіцієнтами, таких що містять адитивну випадкову складову:

- (1) випадкова складова має нульове середнє значення;
- (2) випадкова складова має постійну дисперсію (*гомоскедастичність*);
- (3) значення випадкової складової не корельовані між собою (відсутня *автокореляція*);
- (4) пояснюючі змінні не корельовані з випадковою складовою;
- (5) регресійна модель правильно специфікована, лінійна за коефіцієнтами і містить адитивну випадкову складову.

Для множинної лінійної регресії необхідна також умова:

- пояснюючі змінні не є лінійними функціями інших пояснюючих змінних (відсутня *мультиколінеарність*);

Для перевірки статистичних гіпотез і побудови інтервальних оцінок важливе також виконання умови:

- випадкова складова розподілена нормально.

Якщо виконано умови (1) – (5), то оцінки, отримані методом найменших квадратів, є незміщеними, обґрунтованими й ефективними (теорема Гауса – Маркова).

У такому випадку ці оцінки, отримані за МНК, називають *BLUE* – *best linear unbiased estimators* (найкращі лінійні незміщені оцінки).

Гомоскедастичність і гетероскедастичність

Гомоскедастичність – сталість дисперсії випадкової складової, незалежно від спостереження. Порушення цієї умови називається гетероскедастичністю.

Розрізняють *справжню* і *помилкову* (змішану) гетероскедастичність.

Помилкова гетероскедастичність виникає при помилковій специфікації моделі, у разі невключення до неї істотно впливаючих змінних.

Тестування моделей на гетероскедастичність є однією з необхідних процедур при побудові регресійних моделей. Гетероскедастичність приводить до неефективності оцінок, отриманих за допомогою методу найменших квадратів. У цьому випадку виявляється зміщеною і необґрунтованою класична оцінка коваріаційної матриці МНК-оцінок параметрів і статистичні висновки про якість отриманих оцінок можуть бути неадекватними. При виявленні гетероскедастичності найчастіше слід перевизначити змінні.

Тест рангової кореляції Спірмена

Передбачається, що дисперсія випадкового члена буде або збільшуватися, або зменшуватися при збільшенні фактору x , і тому в регресії, що оцінюється за допомогою МНК, абсолютні величини залишків e і значення x будуть корельовані.

Алгоритм тесту:

1. Розраховують рівняння регресії і обчислюють залишки (різницю між фактичними і розрахунковими значеннями результативної ознаки).

2. Дані впорядковують за зростанням змінної x , від якої передбачають залежність дисперсії залишків. Кожному значенню x ставлять у відповідність порядковий номер (*rang* x).

3. Знову впорядковують вибірку, тепер за зростанням залишків. Кожному значенню залишків привласнюють ранг (порядковий номер, *rang* ε).

4. Обчислюють коефіцієнт рангової кореляції Спірмена між рангами фактору x і залишків ε .

Коефіцієнт рангової кореляції Спірмена – це кількісна оцінка статистичного вивчення зв'язку між явищами, яка показує, як відрізняється отримана при спостереженні сума квадратів різниць між рангами від випадку відсутності зв'язку:

де d_i – різниця рангів x і ε_i , n – кількість спостережень у вибірці.

5. Знайдений коефіцієнт рангової кореляції перевіряють на значущість. Для цього обчислюють статистичний критерій Стюдента:

де k – число пояснюючих змінних у моделі.

6. Порівнюють $t_{\text{фактич}}$ із $t_{\text{кр}}$ для рівня значущості α (зазвичай 0,05) і числа ступенів свободи $(n-k-1)$. Критичне значення $t_{\text{кр}}$ може бути отримано:

- зі спеціальної таблиці « t -розподіл Стюдента» (на перетині стовпців $\alpha = 0,05$ (або $\gamma = 0,95$) та $(n-k-1)$ (див. Додаток 4) або

– за допомогою вбудованої функції MS Excel СТЬЮДЕНТ.ОБР($\gamma; n - k - 1$) (де $\gamma = 1 - \alpha$).

Якщо $t_{\text{факт}} \geq t_{\text{кр}}$, то приймають гіпотезу про наявність гетероскедастичності.

Якщо $t_{\text{факт}} < t_{\text{кр}}$, то вважають, що в досліджуваній моделі присутня гомоскедастичність.

Тест Голдфельда – Квандта

Припускають пропорційність дисперсій відхилень і значень пояснюючої змінної x , усі n спостережень упорядковують за величиною x і розбивають на три підвибірки розмірностей. Оцінюють окремі регресії для першої і третьої підвбірок. Якщо припущення вірне, то дисперсія регресії за першою підвбіркою буде істотно менше дисперсії регресії за третьою підвбіркою.

Тест використовують для значних за обсягом вибірок. Кількість спостережень n має принаймні в два рази перевищувати кількість змінних m .

Алгоритм тесту Голдфельда – Квандта:

1. Упорядковують усі n спостережень за змінною x , від якої передбачають залежність дисперсії залишків (або найбільш суттєвою для моделі).

2. Розбивають усю вибірку на три підвбірки і виключають серединну частину (другу підвбірку) з k спостережень. Зазвичай k приймають від $n/4$ до $n/3$ спостережень. (Тоді перша і друга підвбірки міститимуть по $(n-k)/2$ значень кожна.)

Якщо вихідне припущення правильне, то дисперсія регресії по першій підвбірці буде істотно менше дисперсії регресії по третій підвбірці.

3. Оцінюють окремі регресії для першої і третьої підвбірок і обчислюють суми квадратів їх залишків.

4. Для порівняння відповідних дисперсій обчислюють відповідну F-статистику: $F_{\text{факт}}$, де $S_{1,2}$ – суми квадратів відповідних залишків.

5. Порівнюють отримане значення $F_{\text{факт}}$ із табличним значенням критерію Фішера $F_{\text{табл}}$ (з обраним рівнем значущості – наприклад, 0,05 – і числом ступенів свободи). Табличне значення $F_{\text{табл}}$ може бути:

- отримано з довідкової таблиці розподілу Фішера (див. Додаток 3) або
- обчислено за допомогою функції Excel F.ОБР ($\gamma; (n-k)/2-m; (n-k)/2-m$) (де $\gamma = 1 - \alpha$).

Якщо $F_{\text{факт}} \leq F_{\text{табл}}$, то приймають гіпотезу про наявність гомоскедастичності. Якщо $F_{\text{факт}} > F_{\text{табл}}$, то вважають, що в моделі спостерігається гетероскедастичність.

Завдання за темою 2.2. Особливі випадки

Завдання 1. Модель із фіктивними змінними

Міністерство інфраструктури досліджує вартість залізничних вантажоперевезень (Ціна, у. о.) у залежності від відстані (S , км), при цьому порівнюють

розцінки трьох регіональних управлінь залізниці (А, В, С). Є статистичні дані про вантажоперевезення за 36 договорами (замовленнями) (див. табл.).

Регіон	С	В	В	С	А	А	С	С	В
<i>S</i>	614	403	344	533	544	585	423	282	438
<i>Ціна</i>	705,9	470,2	380,8	596,3	634,4	636,3	539,4	334	472,4

Регіон	А	С	В	С	В	А	В	В	В
<i>S</i>	573	333	333	622	401	334	442	283	355
<i>Ціна</i>	635,6	416,8	355,5	680,8	468,7	401,1	512,1	305,1	409,7

Регіон	С	А	В	А	В	В	А	С	С
<i>S</i>	29	53	29	58	66	67	63	59	51
<i>Ціна</i>	36	69	47	73	64	60	62	71	62
Регіон	А	С	А	А	В	С	В	С	А
<i>S</i>	67	63	33	33	42	67	33	23	28
<i>Ціна</i>	70	71	52	63	48	71	46	25	55

1. Створіть нову книгу Excel *Прізвище ЕММ6*. Укажіть прізвище, групу, дату виконання роботи. Скопіюйте до цієї книги вихідні дані і перетворіть їх до однієї вертикальної таблиці. Відсортуйте дані за фірмами, що надають транспортні послуги.

2. Побудуйте кореляційне поле для змінних *Ціна* і *S*, зображуючи точки, що відповідають різним фірмам, різними символами.

3. Оцініть рівняння регресії $Ціна = b_0 + b_1 \cdot S + \varepsilon$. Що воно відображає? Оцініть якість побудованої регресії (зробіть висновок).

4. Створіть два нових (фіктивних) стовпця D_1 і D_2 . Надайте їм значення

$$D_1 = \begin{cases} 1, \text{ якщо регіон } A, \\ 0, \text{ в іншому випадку} \end{cases} \quad D_2 = \begin{cases} 1, \text{ якщо регіон } B, \\ 0, \text{ в іншому випадку} \end{cases}$$

5. Оцініть рівняння регресії $Ціна = b_0 + b_1 \cdot S + c_1 \cdot D_1 + c_2 \cdot D_2 + \varepsilon$,

де D_1, D_2 – фіктивні змінні, що відображають відмінність трьох фірм.

5.1. Надайте інтерпретацію побудованої регресії. Оцініть якість побудованої регресії (зробіть висновки).

5.2. На підставі отриманого рівняння випишіть рівняння для кожної фірми окремо (підставивши відповідні значення D_1 і D_2).

6. Створіть та обчисліть два нових стовпця $S \cdot D_1$ та $S \cdot D_2$.

6.1. Оцініть рівняння регресії $Ціна = b_0 + b_1 \cdot S + c_1 \cdot (S \cdot D_1) + c_2 \cdot (S \cdot D_2) + \varepsilon$.

6.2. Надайте інтерпретацію побудованого рівняння. Що відображає добуток змінних $(S \cdot D_1)$ та $(S \cdot D_2)$? Оцініть якість побудованої регресії (висновок).

7. Оцініть рівняння регресії

$$\text{Ціна} = b_0 + b_1 \cdot S + c_1 D_1 + c_2 D_2 + c_3 \cdot S \cdot D_1 + c_4 \cdot S \cdot D_2 + \varepsilon.$$

7.1. Оцініть якість побудованої регресії (висновок).

7.2. Надайте інтерпретацію побудованого рівняння. За необхідності побудуйте модель тільки зі значущими параметрами.

8. Яка з моделей краще для відображення досліджуваної залежності і чому?

9. Нанесіть на кореляційне поле графіки, що відповідають різним фірмам, виходячи з найкращої моделі (вихідні та теоретичні дані).

10. Виконане завдання скопіюйте в папку \$control.

Завдання 2. Тест Голдфелда – Квандта

Дана таблиця для оцінки залежності рівня продажів торгових точок (Y) від чисельності потенційних покупців (X). Необхідно за методом Голдфелда – Квандта перевірити модель на гетероскедастичність.

Дохід, Y	Чисельність потенційних покупців, X
215838	195132
475464	202752
197158	249978
610075	165747
764135	147550
182518	242420
494200	694045
482793	251220
196992	417200
216104	151408
579152	267684
493713	831660
316692	185853
307704	200882
309972	118386
635150	419700
667020	343480
501372	449682
487544	286930
403782	740420
305874	458080
548110	158799
749420	814000
196776	118002

Дохід, Y	Чисельність потенційних покупців, X
281582	380480
405040	147600
279034	567830
345708	388250
546996	467688
210134	250248
547488	919765
234292	302285
490614	325325

1. Упорядкуйте дані за зростанням змінної X .
2. Виберіть число C центральних спостережень змінної і виключіть їх із вибірки. Число C приймають від однієї четвертої до однієї третини загального числа спостережень. Виберіть, наприклад, число $C = 9$ – тобто виключіть дев'ять центральних спостережень, залишивши дві вибірки по дванадцять спостережень.
3. Для отриманих вибірок побудуйте рівняння регресії, знайдіть залишки і розрахуйте їхні суми квадратів.
4. Розрахуйте значення критерію F .
5. Порівняйте отримане значення з табличним значенням критерію Фішера з рівнем значущості 0,05 використовуючи функцію Excel $F.ОБР$.
6. Зробіть висновок про наявність гетероскедастичності помилок моделі.

Завдання 3. Тест Спірмена

1. На аркуші Excel «Тест Спірмена» за вихідними даними завдання 2 перевірити наявність гетероскедастичності для того ж фактору за допомогою тесту рангової кореляції Спірмена.
2. Для визначення табличного значення t -критерію Стьюдента використовуйте функцію Excel $СТЬЮДЕНТ.ОБР$.
3. Виконане завдання скопіюйте в папку \$control.

МОДУЛЬ 3. МОДЕЛЮВАННЯ ДИНАМІЧНИХ ПРОЦЕСІВ

Тема 3.1. Автокореляція

Часовий ряд – це сукупність значень якого-небудь показника за кілька послідовних моментів (періодів) часу.

Автокореляція рівнів часового ряду – це кореляційна залежність між послідовними рівнями часового ряду.

Коефіцієнт автокореляції 1-го порядку – це коефіцієнт кореляції між рівнями вихідного часового ряду y_t і рівнями того ж ряду, зсунутими на один момент часу y_{t-1} . При цьому загальне число пар спостережень, за якими проводиться розрахунок, так само $(n-1)$.

Коефіцієнт автокореляції порядку τ – це коефіцієнт кореляції між рядами y_t і $y_{t-\tau}$:

$$r_\tau = \frac{\text{cov}(y_t, y_{t-\tau})}{\sqrt{\text{var}(y_t)\text{var}(y_{t-\tau})}}$$

Кількість періодів τ , за якими розраховують коефіцієнт автокореляції, називають *лагом* (іноді зсувом, запізнюванням).

Значення коефіцієнта автокореляції, близьке по модулю до 1, свідчить про наявність у часовому ряді тенденції (тісного зв'язку між поточними і попередніми рівнями часового ряду).

Набір коефіцієнтів автокореляції з послідовними значеннями лагів задає автокореляційну функцію (ACF) часового ряду. Її графік називають корелограмою (уздовж осі абсцис – часовий лаг, уздовж осі ординат – значення коефіцієнтів автокореляції). Найбільше значення коефіцієнта $r_{t,t-\tau}$ дозволяє визначити лаг (періоди коливання ряду).

Моделі часових рядів, у яких значення часового ряду в даний момент лінійно залежать від попередніх значень цього ж ряду, називають **авторегресійними**. Авторегресійний процес p -го порядку AR (p) відображає залежність від значень за p попередніх періодів:

$$y_t = b_0 + b_1 y_{t-1} + \dots + b_p y_{t-p} + \varepsilon_t$$

Автокореляція в залишках – кореляційна залежність між залишками в послідовні моменти часу.

Критерій Дарбіна – Уотсона

Критерій Дарбіна – Уотсона виявлення автокореляції залишків:

1. За побудованим рівнянням регресії обчислюють залишки.
2. Обчислюють статистику DW Дарбіна – Уотсона

$$DW = \frac{\sum_{t=2}^n (\varepsilon_t - \varepsilon_{t-1})^2}{\sum_{t=1}^n \varepsilon_t^2}$$

3. За таблицею критичних точок розподілу Дарбіна – Уотсона (див. Додаток 5) знаходять нижню d_L і верхню d_U межі DW-статистики – для заданого рівня значущості α , кількості спостережень n і кількості пояснюючих змінних m .

4. Перевіряють, до якого з 5 інтервалів потрапила розрахована статистика DW:



Якщо DW потрапляє до інтервалу $[d_U; 4-d_L]$, то автокореляція відсутня. Значення DW у межах $[0; d_L]$ і $[4-d_L; 4]$ підтверджують наявність автокореляції (відповідно позитивну і негативну). Для інтервалів $[d_L; d_U]$ і $[4-d_U; 4-d_L]$ гіпотеза про відсутність автокореляції не може бути ні прийнята, ні відхилена.

Можна оцінити значення DW-статистики приблизно, враховуючи що $DW \approx 2 \cdot (1 - r_{\varepsilon_t \varepsilon_{t-1}})$, отже, $0 \leq DW \leq 4$. При $r_{\varepsilon_t \varepsilon_{t-1}} = 0$ (кореляція відсутня) $DW \approx 2$. Приблизно можна вважати, що автокореляція відсутня, якщо $1,5 < DW < 2,5$.

Критерій Дарбіна – Уотсона не застосовують для авторегресійних моделей часових рядів із лагом в 1 період. Він вимагає великого розміру вибірки, наявності вільного члена в моделі; передбачається, що залишки задовольняють схемі 1-го порядку $\varepsilon_t = k \cdot \varepsilon_{t-1} + v_t$

Якщо автокореляцію залишків не вдається усунути, можливо, вона викликана внутрішніми властивостями ряду. У такому випадку для її усунення може бути використано метод ковзних середніх.

Завдання за темою 3.1. Автокореляція

Завдання 1. Моделювання тренду часового ряду

Є дані про обсяги споживання електроенергії жителями регіону (Y_t) за 16 кварталів:

№ кварталу	Y_t	№ кварталу	Y_t	№ кварталу	Y_t	№ кварталу	Y_t
1	6,0	5	7,2	9	8,0	13	9,0
2	4,4	6	4,8	10	5,6	14	6,6
3	5,0	7	6,0	11	6,4	15	7,0
4	9,0	8	10,0	12	11,0	16	10,8

1. Побудуйте графік вихідних даних.
2. Чи є, на Ваш погляд, висхідна чи спадна тенденція? Чи є явно виражені коливання обсягу продажів у залежності від кварталу?
3. Додайте лінію лінійного тренду, відобразіть рівняння тренду і коефіцієнт детермінації.
4. Отримайте рівняння лінійної регресії за допомогою надбудови *Аналіз даних*.
5. Запишіть оцінене рівняння. Перевірте значущість рівняння і параметрів.

6. Проінтерпретуйте параметри (із застереженням на значущість або незначущість).

Завдання 2. Авторегресійна модель

Побудуйте авторегресійну модель (регресію ряду на самого себе з лагом). Для цього:

1. Розрахуйте значення автокореляційної функції ACF часового ряду (тобто коефіцієнти кореляції між рівнями ряду та їхніми попередніми значеннями з лагом 1, 2 ...). (Підказка. Можна використовувати *Пакет аналізу* або функцію Excel). Побудуйте корелограму (графік функції ACF).

2. Знайдіть найбільше за модулем значення і побудуйте для відповідного йому лага лінійну регресію. Оцініть значущість коефіцієнта і регресії в цілому.

3. Побудуйте на діаграмі графіки вихідного і змодельованого ряду даних. Оцініть візуально, чи задовільно модель описує дані?

4. Зробіть висновок про якість побудованої моделі.

Тема 3.2. Моделювання динамічних процесів

Моделі, що включають поточні і деякі попередні за часом значення пояснюючих змінних, а також час t , називають **динамічними моделями**.

Класи динамічних моделей:

- з лагами (містять тільки пояснюючі змінні в попередні моменти часу) – найчастіше їх можна звести до рівняння множинної регресії;
- авторегресійні (включають також і значення залежних змінних у попередні моменти часу).

Кожен рівень часового ряду можна вважати сформованим із трьох компонент: *трендової* (T), *циклічної* (S) і *випадкової* (E). Залежно від виду зв'язку між цими компонентами може бути побудована:

- адитивна модель часового ряду: $Y(t) = T(t) + S(t) + E(t)$;
- мультиплікативна модель часового ряду: $Y(t) = T(t) \cdot S(t) \cdot E(t)$.

Побудова аналітичної функції для моделювання тенденції (тренду) часового ряду називають *аналітичним вирівнюванням часового ряду*.

Виділення циклічної компоненти має сенс у разі схильності даних до циклічних (тижневих, місячних, квартальних, річних тощо) коливань. Циклічну компоненту часто називають *сезонною*.

Способи моделювання сезонності:

- авторегресійна модель;
- модель із фіктивними змінними;
- модель змінного (ковзного) середнього – *адитивна* модель;
- модель змінного (ковзного) середнього – *мультиплікативна* модель тощо.

Авторегресія часового ряду

1. Обчислюють значення автокореляційної функції ACF часового ряду (тобто коефіцієнти кореляції між рівнями ряду та їхніми попередніми значеннями з лагом 1, 2 ...) – для кожної пари спостережень (y_t, y_{t-1}); загальна кількість таких пар ($n-1$).

2. Визначають найбільше значення функції і відповідний йому лаг p . Для візуального відображення будують корелограму (графік автокореляційної функції ACF).

3. Для даного лага будують регресію вигляду $y_t = b_0 + b_1 y_{t-p} + \varepsilon_t$.

Застосування фіктивних змінних при дослідженні сезонних коливань

Число сезонних фіктивних змінних у регресійній моделі завжди має бути на одиницю менше сезонів всередині періоду, тобто має дорівнювати величині $L-1$, якщо L сезонів. Наприклад, для поквартальних даних необхідні 3 фіктивні змінні, а загальний вигляд регресійної моделі буде одним із різновидів адитивної моделі часового ряду:

$$y_t = b_0 + b_1 t + b_2 D_1 + b_3 D_2 + b_4 D_3 + \varepsilon_t.$$

Підставивши набори значень фіктивних змінних, отримаємо часткові моделі регресії для кожного сезону.

Вибір між адитивної і мультиплікативної моделями:

- побудувати графік ряду динаміки (за віссю абсцис – тимчасові проміжки, за віссю ординат – рівні ряду);

- проаналізувати наявність сезонної компоненти: якщо сезонні коливання мають приблизно однакову амплітуду, то для аналізу часового ряду використовують модель із адитивною компонентою. Якщо амплітуда сезонних коливань зменшується або збільшується, то для побудови краще підходить модель з мультиплікативною компонентою.

Зазвичай сезонну компоненту обчислюють і виключають із ряду (тобто ряд коригують на сезонність), потім виявляють тенденцію, тобто трендову компоненту.

При побудові прогнозу спочатку знаходять точковий прогноз трендової компоненти, а потім застосовують сезонну поправку. Точковий прогноз випадкової компоненти в адитивній моделі приймається таким, що дорівнює нулю.

Етапи побудови адитивної і мультиплікативної моделей:

1) вирівнювання (згладжування) вихідного ряду методом змінного (ковзного) середнього;

2) обчислення скоригованих значень сезонної компоненти S ;

3) усунення сезонної компоненти з вихідних рівнів ряду і отримання вирівняних даних: для адитивної моделі – віднімання сезонної компоненти з вихідних значень; для мультиплікативної – ділення вихідних значень на сезонну складову.

4) побудова тренду T на основі отриманих даних;

5) обчислення значень відповідно $(T + S)$ або $(T \cdot S)$;

6) обчислення абсолютних та / або відносних помилок;

7) обчислення середньоквадратичної помилки – для обґрунтування відповідності моделі вихідним даним або для вибору найкращої з множини моделей.

1) Методи згладжування (вирівнювання) часового ряду

Методи ідентичні й для адитивної, і для мультиплікативної моделей: метод усереднення за двома половинами; метод укрупнення інтервалів; метод експоненційної середньої; метод ковзної середньої.

Метод ковзної середньої заснований на розрахунку середніх рівнів ряду за певний період. Якщо цей період включає L сезонів, треба усереднити (згладити) кожні L послідовних значень ряду. У тому випадку, якщо L виявилось парним, після цього отримані середні треба центрувати (тобто знайти середні для кожних двох послідовних отриманих результатів).

2) Розрахунок сезонних компонент

В адитивній моделі:

Сезонну компоненту в адитивній моделі розраховують як різницю між вихідним значенням ряду і знайденим ковзним середнім (центрованим для парних L), що йому відповідає.

Отримані сезонні оцінки треба усереднити для кожного сезону.

Сума отриманих сезонних відхилень має дорівнювати нулю – тому що в моделях із сезонною компонентою передбачається, що сезонні впливи за період взаємно погашаються. Для цього треба скоригувати знайдені усереднені сезонні відхилення: відняти з кожного $1/L$ від їхньої суми.

У мультиплікативній моделі:

Сезонну компоненту в мультиплікативній моделі обчислюють як частку між вихідним значенням ряду і знайденим ковзним середнім (центрованим для парних L), що йому відповідає.

Отримані сезонні оцінки треба усереднити для кожного сезону.

Сума отриманих сезонних відхилень повинна дорівнювати числу сезонів у періоді. Інакше треба провести їх коригування, помноживши кожне на коефіцієнт: частка від L , поділеного на їхню суму.

3) Усунення сезонної компоненти

Знайдені сезонні (скориговані) компоненти треба усунути з вихідних даних.

Для адитивної моделі: відняти з вихідного значення; для мультиплікативної – розділити на сезонну компоненту.

Отримані оцінки можна використовувати для побудови тренду (попередньо можна побудувати для них кореляційне поле і візуально визначити форму залежності).

До знайденого тренду необхідно додати (для адитивної моделі) або помножити (для мультиплікативної) скориговані сезонні відхилення.

На закінчення слід обчислити помилки, зробити висновки, у разі необхідності – побудувати прогноз. Оскільки сезонним коливанням зазвичай відповідають невеликі тимчасові інтервали, то можливо лише короткострокове прогнозування за моделями з сезонною компонентою.

Завдання за темою 3.2. Моделювання динамічних процесів

Завдання 1. Фіктивні змінні

За даними завдання 1 теми 3.1 побудуйте модель із сезонністю вигляду:

$$y_t = b_0 + b_1 t + b_2 D_1 + b_3 D_2 + b_4 D_3 + \varepsilon_t$$

де D_i – фіктивна змінна, що приймає значення 1, якщо спостереження відноситься до i -го кварталу, і 0 в іншому випадку ($i = 1, 2, 3, 4$).

1. Додайте фіктивні змінні D_1, \dots, D_4 .

Оцініть модель $y_t = b_0 + b_1 t + b_2 D_1 + b_3 D_2 + b_4 D_3 + b_5 D_4$. Вочевидь, ви отримаєте результати, що вказують на помилку, тому що в моделі присутня повна мультиколінеарність: $D_1 + \dots + D_4 = 1 = \text{const}$.

2. Видаліть із моделі одну із фіктивних змінних D_i та оцініть модель $y_t = b_0 + b_1 t + b_2 D_1 + b_3 D_2 + b_4 D_3$.

3. Запишіть оцінене рівняння регресії. Перевірте його значущість (в цілому і за окремими параметрами). Якщо є незначущий фактор, виключіть його із рівняння та переоцініть регресію.

4. Побудуйте та оцініть візуально графік залишків. Зробіть висновок про якість побудованої моделі.

5. На основі останньої оціненої моделі запишіть чотири рівняння – окремо для кожного кварталу. З огляду на 95%-довірчі інтервали для коефіцієнтів регресії, зробіть висновок про те, в які квартали споживання більше.

6. Побудуйте прогноз на наступні два квартали.

Завдання 2. Ковзне середнє – адитивна модель

За даними завдання 1 побудуйте адитивну модель часового ряду. Для цього:

1. Згладьте ряд ковзним середнім із лагом 4.

2. Отримайте скориговані оцінки сезонних компонент окремо для кожного кварталу.

3. Виключіть сезонну компоненту і побудуйте лінійну регресію для трендової компоненти.

4. Запишіть підсумкове рівняння регресії. За його допомоги побудуйте прогноз на наступні 2 квартали.

Завдання 3. Ковзне середнє – мультиплікативна модель

Аналогічно до попереднього завдання, для вихідних даних завдання 1 побудуйте мультиплікативну модель часового ряду. Для цього:

1. Згладьте ряд ковзним середнім із лагом 4.

2. Отримайте скориговані оцінки сезонних компонент окремо для кожного кварталу.

3. Виключіть сезонну компоненту і побудуйте лінійну регресію для трендової компоненти.

4. Випишіть підсумкове рівняння регресії. За його допомоги побудуйте прогноз на наступні 2 квартали. Порівняйте 4 останні отримані моделі.

СПИСОК НАВЧАЛЬНО-МЕТОДИЧНОЇ ЛІТЕРАТУРИ

Базова література

1. Михайленко, С. В., Свищева, Е. В. (2012). *Оптимизационные методы и модели*. Харків: Изд-во НУА, 184 с.
2. Козьменко, О. В. (2014). *Економіко-математичні методи та моделі (економетрика)*. Суми: Університетська книга, 406 с.
3. Кузьмичов, А. І., Медведєв, М. Г. (2017). *Економетрія. Моделювання за-собами MS Excel*. Київ: Вид-во «Ліра-К», 212 с.
4. Лугінін, О. Є., Фомішина, В. М. та Дудченко, О. М. (2014). *Економетрика*. Херсон: ОЛДІ-ПЛЮС, 320 с.
5. Доля, В. Т. (2010). *Економетрія*. Харків: ХНАМГ, 171 с.

Додаткова література

6. Данилевич, С. Б., Дьячкова, О. В. (2016) *Економетрика*. Харків: Изд-во НУА. – 76 с.
7. Доугерти, К. (2009). *Введение в эконометрику*. Москва: Инфра-М, 479 с.
8. Елисеєва, И. И. (ред.) *Економетрика* (2014). Москва: Издательство Юрайт, 453 с.
9. Елисеєва, И. И., Курьшева, С. В. и Костева, Т. В. (2005) *Економетрика*. Москва: Финансы и статистика, 576 с.
10. Вовк, Л. В. (2017). *Математичний інструментарій моделювання економічних процесів*. Київ: Видавництво Ліра-К, 252 с.
11. Елисеєва, И. И. (ред.) *Практикум по эконометрике*. М.: Финансы и статистика, 2003, 192 с.
12. Єгоршин, О. О., Малярець, Л. М. (2011) *Лабораторний практикум з економетрики навчальної дисципліни "Економіко–математичні методи та моделі"*. Харків: Вид. ХНЕУ, 148 с.
13. Бегун, С. І. (2014). *Економетрика*. Луцьк: вид-во ПП Іванюк В.П., 60 с.
14. Малярець, Л. М., Железнякова, Э. Ю. и Норик, Л. А. (2014). *Економетрика в примерах и задачах для иностранных студентов*. Харків: Изд. ХНЭУ им. С. Кузнеця, 268 с.

Додаткові джерела:

1. Методичні матеріали на навчальному сервері НУА у папці *S:\\$tasks*.

ОСНОВНІ ФОРМУЛИ

Середнє арифметичне:
$$\bar{x} = \sum_{i=1}^n \frac{x_i}{n}.$$

Дисперсія:
$$D(x) = \sigma_x^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2,$$

де \bar{x} – середнє арифметичне ряду; n – кількість спостережень.

Середнє квадратичне відхилення:

$$\sigma_x = \sqrt{\sigma_x^2} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}}.$$

Коефіцієнт варіації:

$$V = \frac{\sigma_x}{\bar{x}} \times 100\%.$$

Коваріація:

$$\text{cov}(x, y) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}).$$

Коефіцієнт кореляції:

$$r = \frac{\text{cov}(X, Y)}{\sqrt{D(X) \cdot D(Y)}}.$$

Загальний вигляд економетричної моделі:

$$y = f(x_1, x_2, \dots, x_p, \varepsilon),$$

де y – залежна (ендогенна) змінна; x_1, x_2, \dots, x_p – незалежні (екзогенні) змінні, що визначають поведінку y ; ε – випадкова помилка для генеральної сукупності.

Узагальнена лінійна регресійна модель:

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

де $\beta_0, \beta_1, \beta_2, \dots, \beta_p$ – оцінювані параметри моделі.

Вибіркова лінійна регресійна модель:

$$y = b_0 + b_1 x_1 + \dots + b_p x_p + e.$$

де b_0, b_1, \dots, b_p – оцінки невідомих параметрів узагальненої регресійної моделі, e – випадкова помилка для вибірки.

Однофакторна вибіркова лінійна регресійна модель

$$y = b_0 + b_1 x + e,$$

де e – нормально розподілена випадкова величина з параметрами $(0,1)$.

Оцінка параметрів, обчислених методом найменших квадратів:

$$b_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{\text{cov}(x, y)}{\sigma_x^2}, \quad b_0 = \bar{y} - b_1 \bar{x}$$

Коефіцієнт детермінації (R^2) характеризує ступінь близькості кореляційного зв'язку до точного, функціонального. Обчислюється за формулою:

$$R^2 = \frac{SSR}{SST},$$

де $SSR = \sum (\hat{y}_i - \bar{y})^2$, $SST = \sum (y_i - \bar{y})^2$.

Для лінійної залежності: $R^2 = r_{xy}^2$, де r_{xy} – коефіцієнт кореляції.

Для оцінки адекватності прийнятої економетричної моделі експериментальним даним використовують **F -критерій (F -тест) Фішера:**

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

де n – число спостережень, m – число пояснюючих чинників. Для парної регресії $m=1$, отже, величина F :

$$F_{1, n-2} = \frac{MSR}{MSE} = \frac{SSR/1}{SSE/(n-2)} = (n-2) \frac{SSR}{SSE}.$$

Перевірка моделі на адекватність за допомогою F -критерію Фішера передбачає виконання певних етапів:

1. Висувають нульову гіпотезу H_0 про те, що $R^2 = 0$.
2. Обчислюють величину F – для парної регресії відношення:

$$F = \frac{R^2}{1 - R^2} (n - 2) = \frac{SSR}{SSE} (n - 2),$$

де $SSR = \sum (\hat{y}_i - \bar{y})^2$ $SSE = \sum (y_i - \hat{y}_i)^2$

3. Задають рівень значущості α . Наприклад, якщо вважають, що можлива помилка буде 0,05 (або 5%), це означає, що ми можемо помилитися не більш ніж в 5% випадків, а в 95% випадків наші висновки будуть правильними.
4. За статистичними таблицями F -розподілу Фішера з $(1, n-2)$ ступенями свободи і рівнем значущості α знаходять критичне значення $F_{кр}$.
5. Умова адекватності моделі: $F > F_{кр}$.

Якщо обчислене значення більше критичного $F > F_{кр}$, то відкидають гіпотезу H_0 про незначущість коефіцієнта детермінації з ризиком помилитися не більш ніж у 5% випадків і визнають статистичну значимість рівняння регресії. Тому, якщо $F > F_{кр}$, то зв'язок між розглянутими ознаками є і результати спостережень не суперечать припущенню. Таким чином, побудована регресійна модель адекватна реальній дійсності.

Якщо $F < F_{кр}$, то гіпотезу H_0 не відкидають, рівняння регресії вважають статистично незначущим.

Перевірка адекватності побудованого рівняння регресії поряд із перевіркою значущості коефіцієнта детермінації включає і перевірку значущості (значущої відмінності від нуля) кожного коефіцієнта (параметра) регресії. Для цієї перевірки застосовують ***t*-критерій (*t*-тест) Стьюдента:**

$$t_{n-m-1} = \frac{b_i}{S_{b_i}}$$

де n – число спостережень, m – число пояснюючих чинників, S_{b_i} – стандартна помилка параметра. Для парної регресії $m=1$, число ступенів вільності $n-2$, отже:

Етапи проведення *t*-тесту Стьюдента для **перевірки значущості параметрів** регресії:

1. Висувають нульову гіпотезу H_0 : параметр $b_i = 0$.
2. Обчислюють значення *t*-критерія Стьюдента – для парної регресії:

$$t_{n-2} = b_i / S_{b_i}, \quad i = 0, 1,$$

де $S_{b_1} = \frac{S_{ост}}{\sigma_x \sqrt{n}}$ – стандартна помилка параметра b_1 ,

$$S_{b_0} = \frac{S_{ост} \sqrt{\sum x_i^2}}{\sigma_x n} \text{ – стандартна помилка параметра } b_0,$$

$$S_{ост}^2 = \frac{SSE}{n-2} \text{ – стандартна помилка регресії.}$$

3. Визначають критичне значення $t_{кр}$ при певному рівні значущості (зазвичай $\alpha=0,05$) і кількості ступенів вільності $n-m-1$.
4. Умова значущості параметра регресії: $|t_b| > t_{кр}$

Обчислене значення порівнюють із критичним. Якщо $|t_b| > t_{кр}$, то гіпотезу H_0 відхиляють, коефіцієнт b визнають статистично значущим.

Якщо $|t_b| < t_{кр}$, то коефіцієнт b визнають статистично незначущим.

Довірчий інтервал для параметрів b_i : $b_i \pm S_{b_i} \cdot t_{кр}$

Прогнозне значення $y_{прог}$ для парної регресії при значенні $x = x_{прог}$:

$$y_{прог} = b_0 + b_1 \cdot x_{прог}$$

Помилка прогнозу: $S_y = S_{ост} \sqrt{1 + \frac{1}{n} + \frac{(x_{прог} - \bar{x})^2}{n \cdot \sigma_x^2}}$

Довірчий інтервал для прогнозу $y_{прог}$: $y_{прог} \pm S_y \cdot t_{табл}$

Коефіцієнт кореляції є випадковою величиною (тобто має імовірнісний характер), отже, необхідно перевіряти його значущість. Для цього використовують **t -критерій (t -тест) Стюдента** – для парної регресії:

$$t_{n-2} = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}.$$

Етапи проведення t -тесту Стюдента для **перевірки значущості коефіцієнта кореляції** (гіпотетичної лінійної залежності між x і y):

1. Висувають нульову гіпотезу про те, що залежності немає.

2. Обчислюють t -статистику для r : $t = r \sqrt{\frac{n-2}{1-r^2}}$.

3. Вибравши рівень значущості (наприклад, 5%), знаходять критичне значення $t_{кр}$ з $(n-2)$ ступенями вільності.

4. Умова значущості коефіцієнта кореляції: $|t| > t_{кр}$.

Якщо величина t за абсолютною величиною перевищує критичне значення $t_{кр}$, відхиляють нульову гіпотезу і приймають висновок про лінійну залежність (пряму або зворотну).

Якщо $|t| < t_{кр}$, то коефіцієнт кореляції визнають статистично незначущим.

Придатність моделі:

1. Модель за F -критерієм Фішера в цілому адекватна і всі коефіцієнти регресії значущі \Rightarrow модель можна використовувати для прийняття рішень і прогнозів.

2. Модель за F -критерієм Фішера адекватна, але частина коефіцієнтів регресії незначущі \Rightarrow модель придатна для прийняття деяких рішень, але не для прогнозування.

3. Модель за F -критерієм Фішера адекватна, але всі коефіцієнти регресії незначущі \Rightarrow на основі такої моделі не приймають рішення і не здійснюють прогнозів.

При економічній інтерпретації рівняння регресії широко використовують часткові коефіцієнти еластичності. **Еластичність** показує, на скільки відсотків у середньому зміниться значення результативної ознаки, якщо фактор зміниться на один відсоток. Для парної лінійної регресії еластичність визначається наступним чином:

$$E = (dy/dx) \cdot (x/y) = b_1 \cdot x/y.$$

Значення еластичності в будь-якій точці залежить не тільки від значення b_1 , але і від значень x і y в даній точці.

Нелінійна однофакторна регресія

Основні типи кривих:

експоненційна: $y = \alpha \cdot e^x$,

степенева: $y = \alpha \cdot x^\beta$,

зворотна: $y = b_0 + b_1 \frac{1}{x}$,

квадратична: $y = b_0 + b_1 x + b_2 x^2$,

модифікована експонента: $y = \alpha \cdot \beta^x + \gamma$,

крива Гомперца: $y = e^{\alpha \beta^x + \gamma}$,

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

Форми експоненційної кривої:

$y = \alpha \cdot \beta^x$, основна форма $\beta > 0$.

$y = \alpha \cdot e^{b_1 x}$, β замінюємо на e^{b_1} , де $b_1 = \ln \beta$.

$y = \alpha(1-r)^x$, β замінюємо на $(1-r)$, де $r = \beta - 1$.

$y = e^{b_0 + b_1 x}$, α замінюємо на e^{b_0} , β на e^{b_1} , де $b_0 = \ln \alpha$, $b_1 = \ln \beta$.

$y = 10^{b_0 + b_1 x}$, α замінюємо на 10^{b_0} , β на 10^{b_1} , де $b_0 = \log \alpha$, $b_1 = \log \beta$.

Перетворення для експоненційної кривої:

$y = \alpha \cdot \beta^x$, $\ln y = \ln \alpha + x \ln \beta$.

$y = \alpha \cdot e^{b_1 x}$, $\ln y = \ln \alpha + x b_1$.

$y = \alpha(1-r)^x$, $\ln y = \ln \alpha + x \ln(1-r)$.

$y = e^{b_0 + b_1 x}$, $\ln y = b_0 + b_1 x$.

$y = 10^{b_0 + b_1 x}$, $\log y = b_0 + b_1 x$.

Перетворення для степеневих функцій:

$y = \alpha \cdot x^\beta$, $\ln y = \ln \alpha + \beta \ln x$.

Перетворення для зворотної функції:

$y = b_0 + b_1 \frac{1}{x}$, $y = b_0 + b_1 z$.

Перетворення для квадратичної функції:

$y = b_0 + b_1 x + b_2 x^2$, $y = b_0 + b_1 z_1 + b_2 z_2$.

Додаток 2

ВАЛОВИЙ ВНУТРІШНІЙ ПРОДУКТ (ВВП) І СПОЖИВЧІ ВИТРАТИ (СП), 1970–2013¹

(млрд. дол., у цінах 1970 р.)

Рік	США		Китай		Японія		Німеччина		Велика Британія		Франція	
	ВВП	СП	ВВП	СП	ВВП	СП	ВВП	СП	ВВП	СП	ВВП	СП
1970	1075,9	901,9	91	59,3	209,1	124	215	154,8	129,7	100,1	148,9	106,4
1971	1167,8	923,4	98,1	62,9	236,2	134,1	249	162,6	146,7	104,5	166,6	112,7
1972	1282,4	964,8	111,6	66,4	312,7	147	298,7	171,8	168,6	110,4	204,3	117,6
1973	1428,5	1001,4	136,1	70,6	424,9	158,4	396,9	180,6	190,3	117	265,4	124,2
1974	1548,8	1012,2	141,5	72,8	471,6	161,1	443,6	185,1	205,6	118,2	286,5	131,1
1975	1688,9	1028,6	160,3	76,8	512,9	176,3	488,8	192,2	245,7	116,6	362	133,7
1976	1877,6	1071,6	150,9	80,3	576,4	183,5	517,8	200,1	236,1	118,4	373,4	140,1
1977	2086	1110,5	171,5	83	709,4	192	598,2	207,5	266,7	117,8	411,5	145,1
1978	2356,6	1153,5	214,2	89,1	996,7	201,4	737,7	212,4	337,9	122,1	508,2	151
1979	2632,1	1183,9	263,2	99,3	1037,5	215,6	878	221,9	439,2	126,9	615,9	156,6
1980	2862,5	1204,2	306,5	109	1087	222,6	946,7	228,5	563,3	125,8	703,5	161,3
1981	3211	1217,3	293,9	117,6	1201,5	230	797,4	232,9	535,3	126,9	617,6	167,7
1982	3345	1230,9	295,4	127	1116,8	242,6	773,6	230,8	510,3	129,2	586,8	173,4
1983	3638,1	1298,3	314,6	140,6	1218,1	253,4	767,8	234	484	134,8	561,9	175,3
1984	4040,7	1365,2	317,4	160,6	1294,6	261,6	722,4	239,4	456,1	137,5	532,7	178,3
1985	4346,7	1445,4	309,1	182,5	1384,5	273,2	729,8	242,9	482,6	141	555,2	182,5
1986	4590,2	1511,1	304,3	195,6	2051,1	279,8	1042,3	243,4	592,9	148,9	774,6	184,8
1987	4870,2	1568,7	329,9	213,7	2485,2	292,9	1293,3	248,9	728,2	156	938,3	190,7
1988	5252,6	1626,7	413,4	239,2	3015,4	307,6	1395,9	255,7	884,8	165,4	1023,5	196
1989	5657,7	1678,9	459,8	251	3017,1	321,5	1393,7	262,6	893,4	169	1030,1	202,1

¹ <http://www.be5.biz/makroekonomika/index.html>

Рік	США		Китай		Японія		Німеччина		Велика Британія		Франція	
	ВВП	СП	ВВП	СП	ВВП	СП	ВВП	СП	ВВП	СП	ВВП	СП
1990	5979,6	1728	404,5	252,6	3103,7	338,8	1764,9	272,1	1059,6	168,8	1275,3	207,9
1991	6174	1736,5	424,1	275,4	3536,8	348	1861,9	283,3	1108,5	169,5	1275,5	211
1992	6539,3	1795,5	499,9	314,6	3852,8	356,9	2123,2	291,3	1149,9	172,5	1408,7	216,2
1993	6878,7	1840,9	641,1	340,6	4415	364,2	2068,5	291,3	1036,2	177,2	1330,2	218,8
1994	7308,8	1897,5	582,7	378,3	4850,3	374,8	2205,9	297,1	1122,6	182	1401,6	222,3
1995	7664,1	1946,9	757	418,9	5333,9	385,2	2590,3	301,5	1235,6	183	1609,8	226,1
1996	8100,2	2009,5	892	469,5	4706,2	397,4	2501,6	306,8	1304,6	187,6	1614,3	231,3
1997	8608,5	2078	985	510,8	4324,3	403,3	2215,9	310,9	1438,7	193,8	1460,7	233,7
1998	9089,2	2171,8	1045,2	557,1	3914,6	401,8	2239,9	314,1	1529,1	202,5	1510,8	239,1
1999	9660,6	2286	1100,8	614	4432,6	410,2	2196,6	322,4	1558,4	212,6	1500,2	246,3
2000	10284,8	2400,4	1192,8	679,3	4731,2	418,1	1947,2	331,9	1548,6	220,6	1368,4	255,8
2001	10621,8	2465,3	1317,2	725	4159,9	428,8	1947,9	337,9	1529,2	229,1	1382,2	261,3
2002	10977,5	2539,7	1455,6	768	3980,8	436,6	2076,4	336,9	1674,4	236,6	1500,4	265,7
2003	11510,7	2623,3	1650,5	806	4302,9	442,5	2502,2	338,7	1943,8	246,7	1848,1	270,4
2004	12274,9	2714,5	1944,7	848,8	4655,8	450,5	2815,6	339,1	2298,1	254,6	2124,2	277,1
2005	13093,7	2795,9	2287,2	920,7	4571,9	459,9	2857,6	342,9	2412,1	261,9	2203,6	283,2
2006	13855,9	2870,8	2793,2	993,8	4356,7	467,3	2998,5	350,2	2582,8	268,3	2324,9	288,5
2007	14477,6	2938,3	3504,4	1108,5	4356,3	473,6	3435,5	350,3	2963,3	274,3	2663	293,5
2008	14718,6	2988,4	4547,3	1189,7	4849,2	477,4	3747,1	356,5	2791,9	277,6	2923,6	295,9
2009	14418,7	2947,5	5105,5	1298,7	5035,1	469,4	3412,8	353,6	2308,9	272	2693,7	296,3
2010	14964,4	3011	5949,8	1423,8	5495,4	485,2	3412	360,2	2407,9	274,3	2646,8	301,6
2011	15517,9	3056,3	7314,4	1585,9	5905,6	493,5	3752,1	369,6	2591,8	275,7	2862,7	305,5
2012	16163,2	3087,8	8229,4	1752,2	5937,9	503,7	3533,2	372,8	2614,9	279,1	2686,7	306,1

Рік	США		Китай		Японія		Німеччина		Велика Британія		Франція	
	ВВП	СП	ВВП	СП	ВВП	СП	ВВП	ВВП	СП	ВВП	СП	ВВП
1970	35,2	28,1	113,1	84	87,8	66,9	61,5	55	45,2	30,6	39,8	29,1
1971	39,3	31,6	124,3	87,3	99,1	69,7	65,9	56,9	50,5	32,2	45,2	30,6
1972	50,6	35,3	144,8	91,6	112,9	73,5	71,7	56,3	59,4	33	57,4	32,8
1973	71,7	39,6	175	97,2	131,1	76,5	85,5	56,8	85,8	33,8	76,3	35,2
1974	91,9	44	199	100,7	160,2	78,3	96,6	58,8	102,3	36,6	94,4	37,8
1975	108,1	44,4	227	101,2	173,6	82,9	100,4	62	109,1	38,1	111,3	38,3
1976	127,7	49,8	224,1	106,2	206,3	86,4	101,2	61,3	117,5	39,8	115	40,8
1977	148,7	51,7	256,8	108,9	211,3	90,6	117,4	66	116,5	40,9	128,5	41,7
1978	168,8	54	314,1	111,4	218,3	94,2	135,8	70,1	135,9	41,7	155,8	41,9
1979	186,1	58,2	392,5	118,9	242,7	95,3	150,3	66	150,4	42,4	208,2	42,5
1980	191,1	67,1	475,8	125,4	273,4	97,6	184,8	71,7	173,5	44,3	225,8	44,5
1981	225,6	58,6	429,4	128,8	305,7	99,8	197,1	74,9	202,1	45,8	196,7	45,3
1982	238,8	61	426	129,7	313,4	100	201,2	77,5	192,1	47,1	190,1	45,8
1983	170,4	61,3	441,7	130,6	340,3	103,1	219,6	83,8	192,6	47,7	165,8	46,4
1984	175,5	62,7	436,6	134,6	353,9	107,3	217,5	86,7	206,6	49,5	167	46,3
1985	187,4	64,9	450,9	138,6	363	113,6	226,5	89,4	182,2	52,3	175,4	47,4
1986	225,3	72,4	638,5	142	375,7	118,9	248,1	95	191,6	53,9	243,8	48,3
1987	247,1	71,1	803,4	147,5	429	122,6	274,6	98,8	228	55,1	309,2	51,3
1988	276	68,7	888,9	153,5	506	126,8	303,8	105,7	288,7	56,1	364,9	53,3
1989	375,5	70,8	925,9	159,3	563,6	131,1	300,7	109,7	320	59	402,4	56,4
1990	402,1	74,6	1177,8	161,5	592	135	326,8	113,2	323,8	61,5	520,5	58,4
1991	374,2	79,5	1242,5	164,8	608,2	137,2	289,7	113,6	329,6	63,8	559,9	60,4
1992	358,4	78,3	1316,3	167,8	590,1	139,8	290,9	118	326,1	65,8	612	62,2
1993	402,2	81,1	1061,8	165	574,8	142,2	284	123,6	317,3	67,6	509,4	61,9

Рік	США		Китай		Японія		Німеччина		Велика Британія		Франція	
	ВВП	СП	ВВП	СП	ВВП	СП	ВВП	ВВП	СП	ВВП	СП	ВВП
1994	573,7	85,6	1096	168,2	576	144,8	325,3	129,7	362,6	70	514,7	62,8
1995	769	91,5	1171,3	169,9	602	145,5	369,2	137,8	392,3	72,4	596,2	63,8
1996	839,7	94,9	1309,5	171,7	627	147,3	389,2	150,7	435,8	74,5	622,1	65,1
1997	871,2	98,1	1239,5	176,4	651	153,1	422,6	155,9	437,2	78,2	572,5	67
1998	843,8	98,4	1266,8	180,4	631,4	159,8	425,3	169,7	389,8	83,1	600,7	69,7
1999	587,1	98,7	1249	184,9	674,3	165,2	453,4	184,1	426,6	85,8	617,6	73
2000	644,7	101,1	1142,2	192,3	739,5	169,5	467,8	189,4	409,1	88,2	580,3	76,8
2001	554,2	102,1	1162,8	194,8	732,7	174,2	483	200,4	389,9	90,7	608,9	78,9
2002	506	103,6	1267,1	194,7	752,5	181,7	504,9	203,4	435,3	94,2	686,3	80,3
2003	552,4	103,5	1570,3	196,6	887,8	185	591,3	214,3	559	97,4	883,8	82,2
2004	663,7	106,3	1799,2	199,2	1018,4	186,9	715,5	222,9	678,6	100,1	1044,6	85,7
2005	882	111,3	1853,5	203	1164,2	190,9	837,5	240,4	762,4	101,5	1130,8	88,9
2006	1089,3	116	1943,4	206,9	1310,8	196,3	947,9	258,2	819	104,8	1236,4	92,1
2007	1366,9	122,7	2204	207,9	1457,9	201,8	1206,1	280,4	986,5	108,3	1441,4	95,7
2008	1653,5	127,4	2392	207,9	1542,6	204,7	1294,1	297,3	1055,8	107,8	1593,4	97,8
2009	1620,2	132,1	2186,1	202,1	1370,8	213,2	1338,2	324,9	1011	112,2	1454,3	95,6
2010	2143	139,5	2126,6	206	1614,1	217	1704,8	349,3	1290,3	112,3	1384,8	97,1
2011	2476,7	143,6	2278,2	206,4	1788,7	219,1	1930,5	378,2	1532,9	116,1	1454,5	97,7
2012	2248,8	150,4	2091,8	200,3	1832,7	222,9	1892,6	398,3	1578	121,4	1322,5	95,7

F-РОЗПОДІЛ ФІШЕРА

(рівень значущості $\alpha = 0,05$)

df ₂ / df ₁	1	2	3	4	5	6	7	8	9	10
1	161,448	199,5	215,707	224,583	230,162	233,986	236,768	238,883	240,543	241,882
2	18,5128	19	19,1643	19,2468	19,2964	19,3295	19,3532	19,371	19,3848	19,3959
3	10,128	9,5521	9,2766	9,1172	9,0135	8,9406	8,8867	8,8452	8,8123	8,7855
4	7,7086	6,9443	6,5914	6,3882	6,2561	6,1631	6,0942	6,041	5,9988	5,9644
5	6,6079	5,7861	5,4095	5,1922	5,0503	4,9503	4,8759	4,8183	4,7725	4,7351
6	5,9874	5,1433	4,7571	4,5337	4,3874	4,2839	4,2067	4,1468	4,099	4,06
7	5,5914	4,7374	4,3468	4,1203	3,9715	3,866	3,787	3,7257	3,6767	3,6365
8	5,3177	4,459	4,0662	3,8379	3,6875	3,5806	3,5005	3,4381	3,3881	3,3472
9	5,1174	4,2565	3,8625	3,6331	3,4817	3,3738	3,2927	3,2296	3,1789	3,1373
10	4,9646	4,1028	3,7083	3,478	3,3258	3,2172	3,1355	3,0717	3,0204	2,9782
11	4,8443	3,9823	3,5874	3,3567	3,2039	3,0946	3,0123	2,948	2,8962	2,8536
12	4,7472	3,8853	3,4903	3,2592	3,1059	2,9961	2,9134	2,8486	2,7964	2,7534
13	4,6672	3,8056	3,4105	3,1791	3,0254	2,9153	2,8321	2,7669	2,7144	2,671
14	4,6001	3,7389	3,3439	3,1122	2,9582	2,8477	2,7642	2,6987	2,6458	2,6022
15	4,5431	3,6823	3,2874	3,0556	2,9013	2,7905	2,7066	2,6408	2,5876	2,5437
16	4,494	3,6337	3,2389	3,0069	2,8524	2,7413	2,6572	2,5911	2,5377	2,4935
17	4,4513	3,5915	3,1968	2,9647	2,81	2,6987	2,6143	2,548	2,4943	2,4499
18	4,4139	3,5546	3,1599	2,9277	2,7729	2,6613	2,5767	2,5102	2,4563	2,4117
19	4,3807	3,5219	3,1274	2,8951	2,7401	2,6283	2,5435	2,4768	2,4227	2,3779
20	4,3512	3,4928	3,0984	2,8661	2,7109	2,599	2,514	2,4471	2,3928	2,3479
21	4,3248	3,4668	3,0725	2,8401	2,6848	2,5727	2,4876	2,4205	2,366	2,321
22	4,3009	3,4434	3,0491	2,8167	2,6613	2,5491	2,4638	2,3965	2,3419	2,2967

df₂ / df₁	1	2	3	4	5	6	7	8	9	10
23	4,2793	3,4221	3,028	2,7955	2,64	2,5277	2,4422	2,3748	2,3201	2,2747
24	4,2597	3,4028	3,0088	2,7763	2,6207	2,5082	2,4226	2,3551	2,3002	2,2547
25	4,2417	3,3852	2,9912	2,7587	2,603	2,4904	2,4047	2,3371	2,2821	2,2365
26	4,2252	3,369	2,9752	2,7426	2,5868	2,4741	2,3883	2,3205	2,2655	2,2197
27	4,21	3,3541	2,9604	2,7278	2,5719	2,4591	2,3732	2,3053	2,2501	2,2043
28	4,196	3,3404	2,9467	2,7141	2,5581	2,4453	2,3593	2,2913	2,236	2,19
29	4,183	3,3277	2,934	2,7014	2,5454	2,4324	2,3463	2,2783	2,2229	2,1768
30	4,1709	3,3158	2,9223	2,6896	2,5336	2,4205	2,3343	2,2662	2,2107	2,1646
40	4,0847	3,2317	2,8387	2,606	2,4495	2,3359	2,249	2,1802	2,124	2,0772
60	4,0012	3,1504	2,7581	2,5252	2,3683	2,2541	2,1665	2,097	2,0401	1,9926
120	3,9201	3,0718	2,6802	2,4472	2,2899	2,175	2,0868	2,0164	1,9588	1,9105
INF	3,8415	2,9957	2,6049	2,3719	2,2141	2,0986	2,0096	1,9384	1,8799	1,8307

df₂/df₁	12	15	20	24	30	40	60	120	INF
1	243,906	245,95	248,013	249,052	250,095	251,143	252,196	253,253	254,314
2	19,4125	19,4291	19,4458	19,4541	19,4624	19,4707	19,4791	19,4874	19,4957
3	8,7446	8,7029	8,6602	8,6385	8,6166	8,5944	8,572	8,5494	8,5264
4	5,9117	5,8578	5,8025	5,7744	5,7459	5,717	5,6877	5,6581	5,6281
5	4,6777	4,6188	4,5581	4,5272	4,4957	4,4638	4,4314	4,3985	4,365
6	3,9999	3,9381	3,8742	3,8415	3,8082	3,7743	3,7398	3,7047	3,6689
7	3,5747	3,5107	3,4445	3,4105	3,3758	3,3404	3,3043	3,2674	3,2298
8	3,2839	3,2184	3,1503	3,1152	3,0794	3,0428	3,0053	2,9669	2,9276
9	3,0729	3,0061	2,9365	2,9005	2,8637	2,8259	2,7872	2,7475	2,7067
10	2,913	2,845	2,774	2,7372	2,6996	2,6609	2,6211	2,5801	2,5379
11	2,7876	2,7186	2,6464	2,609	2,5705	2,5309	2,4901	2,448	2,4045

df₂/df₁	12	15	20	24	30	40	60	120	INF
12	2,6866	2,6169	2,5436	2,5055	2,4663	2,4259	2,3842	2,341	2,2962
13	2,6037	2,5331	2,4589	2,4202	2,3803	2,3392	2,2966	2,2524	2,2064
14	2,5342	2,463	2,3879	2,3487	2,3082	2,2664	2,2229	2,1778	2,1307
15	2,4753	2,4034	2,3275	2,2878	2,2468	2,2043	2,1601	2,1141	2,0658
16	2,4247	2,3522	2,2756	2,2354	2,1938	2,1507	2,1058	2,0589	2,0096
17	2,3807	2,3077	2,2304	2,1898	2,1477	2,104	2,0584	2,0107	1,9604
18	2,3421	2,2686	2,1906	2,1497	2,1071	2,0629	2,0166	1,9681	1,9168
19	2,308	2,2341	2,1555	2,1141	2,0712	2,0264	1,9795	1,9302	1,878
20	2,2776	2,2033	2,1242	2,0825	2,0391	1,9938	1,9464	1,8963	1,8432
21	2,2504	2,1757	2,096	2,054	2,0102	1,9645	1,9165	1,8657	1,8117
22	2,2258	2,1508	2,0707	2,0283	1,9842	1,938	1,8894	1,838	1,7831
23	2,2036	2,1282	2,0476	2,005	1,9605	1,9139	1,8648	1,8128	1,757
24	2,1834	2,1077	2,0267	1,9838	1,939	1,892	1,8424	1,7896	1,733
25	2,1649	2,0889	2,0075	1,9643	1,9192	1,8718	1,8217	1,7684	1,711
26	2,1479	2,0716	1,9898	1,9464	1,901	1,8533	1,8027	1,7488	1,6906
27	2,1323	2,0558	1,9736	1,9299	1,8842	1,8361	1,7851	1,7306	1,6717
28	2,1179	2,0411	1,9586	1,9147	1,8687	1,8203	1,7689	1,7138	1,6541
29	2,1045	2,0275	1,9446	1,9005	1,8543	1,8055	1,7537	1,6981	1,6376
30	2,0921	2,0148	1,9317	1,8874	1,8409	1,7918	1,7396	1,6835	1,6223
40	2,0035	1,9245	1,8389	1,7929	1,7444	1,6928	1,6373	1,5766	1,5089
60	1,9174	1,8364	1,748	1,7001	1,6491	1,5943	1,5343	1,4673	1,3893
120	1,8337	1,7505	1,6587	1,6084	1,5543	1,4952	1,429	1,3519	1,2539
INF	1,7522	1,6664	1,5705	1,5173	1,4591	1,394	1,318	1,2214	1

t-РОЗПОДІЛ СТЬЮДЕНТА

df	0,2	0,1	0,05	0,04	0,03	0,025	0,02	0,01	0,005	0,0005
1	1,376	3,078	6,314	7,916	10,579	12,706	15,895	31,821	63,657	636,619
2	1,061	1,886	2,920	3,320	3,896	4,303	4,849	6,965	9,925	31,599
3	0,978	1,638	2,353	2,605	2,951	3,182	3,482	4,541	5,841	12,924
4	0,941	1,533	2,132	2,333	2,601	2,776	2,999	3,747	4,604	8,610
5	0,920	1,476	2,015	2,191	2,422	2,571	2,757	3,365	4,032	6,869
6	0,906	1,440	1,943	2,104	2,313	2,447	2,612	3,143	3,707	5,959
7	0,896	1,415	1,895	2,046	2,241	2,365	2,517	2,998	3,499	5,408
8	0,889	1,397	1,860	2,004	2,189	2,306	2,449	2,896	3,355	5,041
9	0,883	1,383	1,833	1,973	2,150	2,262	2,398	2,821	3,250	4,781
10	0,879	1,372	1,812	1,948	2,120	2,228	2,359	2,764	3,169	4,587
11	0,876	1,363	1,796	1,928	2,096	2,201	2,328	2,718	3,106	4,437
12	0,873	1,356	1,782	1,912	2,076	2,179	2,303	2,681	3,055	4,318
13	0,870	1,350	1,771	1,899	2,060	2,160	2,282	2,650	3,012	4,221
14	0,868	1,345	1,761	1,887	2,046	2,145	2,264	2,624	2,977	4,140
15	0,866	1,341	1,753	1,878	2,034	2,131	2,249	2,602	2,947	4,073
16	0,865	1,337	1,746	1,869	2,024	2,120	2,235	2,583	2,921	4,015
17	0,863	1,333	1,740	1,862	2,015	2,110	2,224	2,567	2,898	3,965
18	0,862	1,330	1,734	1,855	2,007	2,101	2,214	2,552	2,878	3,922
19	0,861	1,328	1,729	1,850	2,000	2,093	2,205	2,539	2,861	3,883
20	0,860	1,325	1,725	1,844	1,994	2,086	2,197	2,528	2,845	3,850
21	0,859	1,323	1,721	1,840	1,988	2,080	2,189	2,518	2,831	3,819
22	0,858	1,321	1,717	1,835	1,983	2,074	2,183	2,508	2,819	3,792
23	0,858	1,319	1,714	1,832	1,978	2,069	2,177	2,500	2,807	3,768
24	0,857	1,318	1,711	1,828	1,974	2,064	2,172	2,492	2,797	3,745
25	0,856	1,316	1,708	1,825	1,970	2,060	2,167	2,485	2,787	3,725
26	0,856	1,315	1,706	1,822	1,967	2,056	2,162	2,479	2,779	3,707
27	0,855	1,314	1,703	1,819	1,963	2,052	2,158	2,473	2,771	3,690
28	0,855	1,313	1,701	1,817	1,960	2,048	2,154	2,467	2,763	3,674
29	0,854	1,311	1,699	1,814	1,957	2,045	2,150	2,462	2,756	3,659
30	0,854	1,310	1,697	1,812	1,955	2,042	2,147	2,457	2,750	3,646
31	0,853	1,309	1,696	1,810	1,952	2,040	2,144	2,453	2,744	3,633
32	0,853	1,309	1,694	1,808	1,950	2,037	2,141	2,449	2,738	3,622

df	0,2	0,1	0,05	0,04	0,03	0,025	0,02	0,01	0,005	0,0005
33	0,853	1,308	1,692	1,806	1,948	2,035	2,138	2,445	2,733	3,611
34	0,852	1,307	1,691	1,805	1,946	2,032	2,136	2,441	2,728	3,601
35	0,852	1,306	1,690	1,803	1,944	2,030	2,133	2,438	2,724	3,591
36	0,852	1,306	1,688	1,802	1,942	2,028	2,131	2,434	2,719	3,582
37	0,851	1,305	1,687	1,800	1,940	2,026	2,129	2,431	2,715	3,574
38	0,851	1,304	1,686	1,799	1,939	2,024	2,127	2,429	2,712	3,566
39	0,851	1,304	1,685	1,798	1,937	2,023	2,125	2,426	2,708	3,558
40	0,851	1,303	1,684	1,796	1,936	2,021	2,123	2,423	2,704	3,551
41	0,850	1,303	1,683	1,795	1,934	2,020	2,121	2,421	2,701	3,544
42	0,850	1,302	1,682	1,794	1,933	2,018	2,120	2,418	2,698	3,538
43	0,850	1,302	1,681	1,793	1,932	2,017	2,118	2,416	2,695	3,532
44	0,850	1,301	1,680	1,792	1,931	2,015	2,116	2,414	2,692	3,526
45	0,850	1,301	1,679	1,791	1,929	2,014	2,115	2,412	2,690	3,520
46	0,850	1,300	1,679	1,790	1,928	2,013	2,114	2,410	2,687	3,515
47	0,849	1,300	1,678	1,789	1,927	2,012	2,112	2,408	2,685	3,510
48	0,849	1,299	1,677	1,789	1,926	2,011	2,111	2,407	2,682	3,505
49	0,849	1,299	1,677	1,788	1,925	2,010	2,110	2,405	2,680	3,500
50	0,849	1,299	1,676	1,787	1,924	2,009	2,109	2,403	2,678	3,496
60	0,848	1,296	1,671	1,781	1,917	2,000	2,099	2,390	2,660	3,460
70	0,847	1,294	1,667	1,776	1,912	1,994	2,093	2,381	2,648	3,435
80	0,846	1,292	1,664	1,773	1,908	1,990	2,088	2,374	2,639	3,416
90	0,846	1,291	1,662	1,771	1,905	1,987	2,084	2,368	2,632	3,402
100	0,845	1,290	1,660	1,769	1,902	1,984	2,081	2,364	2,626	3,390
120	0,845	1,289	1,658	1,766	1,899	1,980	2,076	2,358	2,617	3,373
140	0,844	1,288	1,656	1,763	1,896	1,977	2,073	2,353	2,611	3,361
180	0,844	1,286	1,653	1,761	1,893	1,973	2,069	2,347	2,603	3,345
200	0,843	1,286	1,653	1,760	1,892	1,972	2,067	2,345	2,601	3,340
500	0,842	1,283	1,648	1,754	1,885	1,965	2,059	2,334	2,586	3,310
1000	0,842	1,282	1,646	1,752	1,883	1,962	2,056	2,330	2,581	3,300

ГРАНИЧНІ ЗНАЧЕННЯ ДЛЯ СТАТИСТИК ДАРБІНА – УОТСОНА

(рівень значущості $\alpha = 0,05$)

n	m = 1		m = 2		m = 3		m = 4		m = 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	2,21
16	1,10	1,37	0,98	1,54	0,86	1,73	0,74	1,93	0,62	2,15
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	1,45	1,21	1,55	1,12	1,66	1,04	1,77	0,95	1,89
26	1,30	1,46	1,22	1,55	1,14	1,65	1,06	1,76	0,98	1,88
27	1,32	1,47	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	1,48	1,27	1,56	1,20	1,65	1,12	1,74	1,05	1,84
30	1,35	1,49	1,28	1,57	1,21	1,65	1,14	1,74	1,07	1,83
31	1,36	1,50	1,30	1,57	1,23	1,65	1,16	1,74	1,09	1,83
32	1,37	1,50	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	1,51	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

ГЛОСАРІЙ

Автокореляція – кореляція між часовою змінною і лаговою змінною, складеною від тієї ж змінної.

Авторегресія – регресія залежної тимчасової змінної від лагової змінної, складеної від тієї ж змінної.

Гетероскедастичність – неоднорідність щодо дисперсії.

Гомоскедастичність – однорідність щодо дисперсії.

Екзогенні (зовнішні, незалежні) змінні – змінні, значення яких визначаються поза даної моделі, для якої вони є заданими.

Економетрика – розділ науки, що вивчає конкретні кількісні і якісні взаємозв'язки економічних об'єктів і процесів за допомогою математичних і статистичних методів і моделей.

Експоненціальна аналітична залежність – залежність, яка визначається співвідношенням $y = e^{a \cdot x_1^{b_1} \cdot x_2^{b_2} \cdot \dots \cdot x_n^{b_n} \cdot \varepsilon}$.

Екстраполяція – прогноз, отримання розрахункових значень за умови, що значення аргументу виходять за межі області визначення функції.

Ендогенні (внутрішні, залежні) змінні – змінні, значення яких визначаються всередині даної моделі.

Коефіцієнт детермінації R^2 – являє собою квадрат коефіцієнта кореляції, показує, яка частина дисперсії результативної ознаки пояснена рівнянням регресії.

Коефіцієнт еластичності E , – показник, що відображає, на скільки відсотків в у середньому зміниться результативний результативна ознака у при зміні фактору на 1% від свого номінального значення.

Кореляційне поле – сукупність точок на координатній площині, що зображують відображують спостереження.

Кореляційний аналіз – розділ математичної статистики, що вивчає наявність і тісноту зв'язку між змінними без поділу змінних на залежні і пояснюючі.

Кореляційний зв'язок – залежність, при якій зміна середнього значення результативної ознаки обумовлено обумовлена зміною факторних ознак.

Критерій Стьюдента (t -критерій) – застосовується для оцінки статистичної значущості коефіцієнтів отриманого рівняння регресії або коефіцієнта кореляції.

Критерій Фішера (F -критерій) – застосовується для оцінки статистичної значущості отриманого рівняння регресії в цілому.

Лаг – затримка. Зазвичай лаг розглядають між причиною і наслідком.

Лінійна аналітична залежність – залежність, яка що визначається співвідношенням $y = a + b_1 \cdot x_1 + b_2 \cdot x_2 + \dots + b_n \cdot x_n + \varepsilon$.

Лінійний коефіцієнт кореляції – характеризує тісноту лінійного зв'язку між

змінними і дорівнює $r_{xy} = \frac{\frac{1}{n} \sum_i (x_i + \bar{x})(y_i + \bar{y})}{\sigma_x \sigma_y} = \frac{\text{cov}(x, y)}{\sigma_x \sigma_y}$.

- Метод найменших квадратів (МНК)** – метод оцінки параметрів рівняння регресії, заснований на мінімізації суми квадратів відхилень розрахункових значень (за рівнянням регресії) залежною залежною змінною змінної від її спостережних значень.
- Мультиколінеарності** – множинна множинний лінійна лінійний зв'язок.
- Показникова аналітична залежність** – залежність, яка визначається співвідношенням $\hat{y} = a \cdot b^x$.
- Предмет економетрики** – кількісна оцінка взаємозв'язку між випадковими подіями, ознаками, показниками, чинниками змінних економічних об'єктів.
- Регресійний аналіз** – розділ математичної статистики, що вивчає форму залежності характеристик стохастичного процесу від одного або декількох факторів.
- Регресія** – залежність середнього значення якої-небудь величини від деякої іншої величини або від декількох величин.
- Регресія множинна** – залежність з декількома залежними змінними.
- Регресія парна** – залежність між двома змінними.
- Результативний Результативна ознака** – ознака, що змінюється під дією факторних ознак.
- Степенева аналітична залежність** – залежність, яка що визначається співвідношенням $y = a \cdot x_1^{b_1} \cdot x_2^{b_2} \cdot \dots \cdot x_n^{b_n} \cdot \varepsilon$.
- Тенденція** – це напрям розвитку, схильність, спрямованість.
- Тест Голдфельда – Квандта** – застосовується для перевірки залишків регресії на гетероскедастичність.
- Тест рангової кореляції Спірмена** – застосовується для перевірки залишків регресії на гетероскедастичність. Перевіряє наявність монотонної залежності між дисперсією помилки і величиною фактору.
- Тренд** – це основна тенденція змінення певного процесу.
- Факторна ознака** – ознака, під впливом якої змінюється інша ознака (результативна ознака).
- Фіктивні змінні** – змінні, які приймають лише значення 0 або 1, що вказує на відсутність або наявність якогось ефекту, який, як очікується, змінює результат. Застосовуються для врахування в моделі впливу якісного фактору.
- Функціональний зв'язок** – зв'язок, при якому певному значенню факторної ознаки відповідає одне і тільки одне значення результативної ознаки.
- Хибна кореляція** – високе значення коефіцієнтів кореляції при відсутності істинної залежності між змінними.
- Часовий ряд** (ряд динаміки, динамічний ряд) – упорядкована в часі послідовність чисельних, що характеристик, що описують рівні розвитку досліджуваного явища в послідовні моменти або періоди часу.

КОРОТКИЙ АНГЛО-УКРАЇНСЬКИЙ СЛОВНИК ЕКОНОМЕТРИЧНИХ ТЕРМІНІВ

Adjusted R^2 – скоригований R^2 .

Autocorrelation function (ACF) – автокореляційна функція.

Autoregressive model order p , AR (p) – авторегресійна модель порядку p .

Autoregressive integrated moving average model (ARIMA) – інтегрована модель авторегресії і ковзного середнього.

Autoregressive moving average model (ARMA) – модель авторегресії і ковзного середнього.

Best linear unbiased estimator (BLUE) – найкраща (з мінімальною дисперсією) оцінка в класі незміщених лінійних оцінок.

Box–Jenkins model = ARIMA – модель Бокса–Дженкінса = інтегрована модель авторегресії і ковзного середнього.

Central limit theorem (CLT) – центральна гранична теорема.

Classical normal regression (CNR) – модель класичної регресії, в якій помилки мають спільний нормальний розподіл.

Classical regression (CR) – модель регресії, в якій помилки незалежні, однаково розподілені, мають нульове середнє значення і постійну дисперсію.

Conditional distribution – умовний розподіл.

Conditional expectation – умовне середнє, умовне математичне очікування.

Confidence interval – довірчий інтервал.

Consistent estimator – заможна оцінка.

Convergence in distribution (law) – збіжність за розподілом .

Convergence in probability – збіжність за ймовірністю.

Correlation – кореляція.

Correlogram – графік (вибіркової) автокореляційної функції.

Correlation coefficient – коефіцієнт кореляції.

Covariance – коваріація.

Cross–section data – дані, що не мають тимчасової природи, порядок їх розташування не існує.

Curve fitting – підгонка кривої.

Density function – щільність розподілу.

Dependent (endogenous) variable – залежна (ендогенна) змінна.

Determination coefficient (R–squared) – коефіцієнт детермінації R^2 .

Distribution – розподіл.

Distribution function – функція розподілу.

Dummy variable – фіктивна незалежна змінна, що приймає, як правило, два значення – 0 або 1.

Dummy trap – ситуація, коли сума кількох фіктивних змінних, включених до регресії, дорівнює константі, також включеної до регресії.

Endogenous (dependent) variable – ендогенна (залежна) змінна.

Estimation – величина оцінки при заданих вибіркових значеннях.

Estimator – метод оцінювання, функція вибіркових значень.

Exogenous (independent) variable – екзогенна (незалежна) змінна, регресор.

Expectation (mean) – середнє значення, математичне сподівання.

Explanatory variables – пояснюючі змінні, регресори, незалежні змінні.

Explained (unexplained) variance – пояснена (непояснена) дисперсія.

Exponential smoothing – експоненціальне згладжування.

Fitted value – прогнозне значення.

First order condition (FOC) – необхідні умови екстремуму.

Generalized least square estimations (GLS) – узагальнений метод найменших квадратів.

Goodness of fit – якість наближення даних моделлю.

Heteroscedasticity – гетероскедастичність.

Homoscedasticity – гомоскедастичність.

Idempotent matrix – ідемпотентна матриця.

Independent (exogenous) variable – незалежна (екзогенна) змінна.

Indirect least squares – непрямий метод найменших квадратів.

Instrumental variable (IV) – інструментальна змінна.

Instrumental variables estimator (IV-estimator) – метод оцінювання за допомогою інструментальних змінних.

Intercept – вільний член в рівнянні регресії.

Joint distribution – спільний розподіл.

Lagged variable – лагірована (відстала) змінна (для часових рядів), змінна з запізненням, значення якої в попередні моменти часу входять у рівняння для поточного моменту часу.

Law of large numbers (LLN) – закон великих чисел.

Level of statistical significance (significance level) – рівень значущості статистичного критерію.

Likelihood function – функція правдоподібності.

Line of the best fit – лінія, що найкраще підходить.

Linear regression model – лінійна регресійна модель.

Loss function – функція втрат, що вимірює ступінь відхилення даних від «бажаних» значень.

Marginal distribution – маргінальний розподіл, тобто розподіл однієї або декількох компонент випадкового вектору.

Maximum likelihood method (ML) – метод максимальної правдоподібності.

Maximum likelihood estimator – оцінювання за допомогою методу максимальної правдоподібності.

Maximum likelihood estimation – оцінка максимальної правдоподібності.

Mean absolute deviation – середнє абсолютне відхилення.

Mean absolute percentage error – середнє відносне відхилення.

Mean-square error – середньоквадратична помилка.

Model specification – специфікація моделі.

Moving average – ковзне середнє.

Moving average model, MA (q) – модель змінного середнього (q).

Multicollinearity – мультиколінеарність.

Multiple regression model – багатовимірна регресійна модель.

Normal (Gaussian) distribution – нормальний (гаусів) розподіл.

OLS-estimator, OLS-estimation – оцінювання за допомогою методу найменших квадратів і значення цих оцінок.

Omitted variables – пропущені змінні (незалежні змінні, не включені до моделі).

Ordinary least squares method (OLS) – метод найменших квадратів, МНК.

Outliers – викиди (дані, що мають великі відхилення від більшості інших).

Partial autocorrelation function (PACF) – часткова автокореляційна функція.

Partial correlation coefficient – частковий коефіцієнт кореляції.

Reduced form of the model – приведена форма моделі.

Residuals – залишки.

Sample – вибірка.

Sample mean (variance, covariance, moment etc.) – вибіркове середнє (дисперсія, Коваріація, момент тощо).

Seemingly unrelated regression (SUR) – система зовні не пов'язаних між собою рівнянь.

Serial correlation – (для часових рядів) наявність кореляції між помилками, що відносяться до різних моментів часу.

Significance level – рівень значущості.

Simultaneous equations – одночасні рівняння.

Slopes – коефіцієнти при незалежних змінних у рівнянні регресії.

Standard deviation – стандартне відхилення (корінь із дисперсії).

Stationary time series – стаціонарний часовий ряд (статистичні властивості якого не залежать від часу).

Testing hypothesis – перевірка гіпотез.

Time series data – часові ряди, параметри системи, які спостерігаються в послідовні моменти часу.

Two stage least squares (ISLES) – двокроковий метод найменших квадратів.

Unbiased estimator – незміщене оцінювання (оцінка).

Under-, overestimation – оцінка параметра, що має негативне (відповідно позитивне) зміщення.

Variance – дисперсія.

Variance (covariance) matrix – матриця коваріацій випадкового вектору.

Weighted least squares – зважений метод найменших квадратів.

Додаток 8

КОНТРОЛЬНІ ПИТАННЯ

1. Що є предметом вивчення економетрики?
2. Зв'язок економетрики з іншими науками.
3. Статистична база економетричних моделей.
4. У чому відмінність між функціональною та стохастичною залежністю?
5. Що собою являє кореляційний зв'язок?
6. Якими статистичними методами досліджуються функціональні і кореляційні зв'язки?
7. Які основні завдання вирішують за допомогою кореляційного і регресійного аналізу?
8. Дайте визначення моделі.
9. Етапи економетричного аналізу.
10. Охарактеризуйте основні проблеми та правила побудови однофакторної лінійної регресійної моделі.
11. У чому полягає значення рівняння регресії?
12. Що характеризують коефіцієнти регресії?
13. Метод визначення параметрів рівняння регресії.
14. Передумови використання методу найменших квадратів.
15. Навіщо потрібна перевірка адекватності регресійної моделі?
16. Як здійснюється перевірка значущості коефіцієнтів?
17. Якими показниками вимірюється тіснота кореляційного зв'язку?
18. Яке значення має обчислення коефіцієнта детермінації?
19. Лінійні коефіцієнти кореляції і детермінації, їх зміст і значення.
20. Перевірка обґрунтованості показників тісноти зв'язку як необхідна умова поширення висновків за результатами вибірки на всю генеральну сукупність. Як вона здійснюється?
21. Як економічно охарактеризувати однофакторну регресійну модель?
22. Який економічний сенс мають коефіцієнти еластичності?
23. Коли регресійна модель вважається лінійною? Нелінійною?

24. Основні проблеми та правила побудови багатofакторної кореляційної моделі.
25. Сутність і призначення парних і часткових коефіцієнтів кореляції.
26. Сутність і значення сукупного коефіцієнта множинної кореляції і сукупного коефіцієнта детермінації.
27. Як перевірити адекватність в цілому? Значущість коефіцієнта регресії? Які критерії для цього можна використовувати?
28. Як економічно інтерпретувати багатofакторну регресійну модель?
29. Побудова довірчих інтервалів.
30. Виробнича функція Кобба – Дугласа.
31. Які висновки можна зробити на основі коефіцієнтів еластичності виробничої функції?
32. У чому причини мультиколінеарності?
33. Тестування на наявність мультиколінеарності та методи її виключення.
34. Коли виникає гетероскедастичність? Тести на гетероскедастичність та методи її усунення.
35. Авторегресійні моделі.
36. У чому полягає автокореляція? Тест на автокореляцію.
37. Дистрибутивно-лагові моделі.
38. Система одночасних рівнянь.
39. Методи оцінювання невідомих параметрів у системах одночасних рівнянь.
40. Область застосування непрямого методу найменших квадратів і двокрокового методу найменших квадратів.

ЗМІСТ

ЗАГАЛЬНІ ПОЛОЖЕННЯ.....	3
ЗМІСТ МОДУЛЯ 1 «Парна регресія».....	7
ЗМІСТ МОДУЛЯ 2 «Множинний регресійний аналіз»	7
ЗМІСТ МОДУЛЯ 3 «Моделювання динамічних процесів»	8
МОДУЛЬ 1. ПАРНА РЕГРЕСІЯ.....	8
Тема 1.1. Лінійна регресія	8
<i>Завдання за темою 1.1. Лінійна регресія.....</i>	<i>9</i>
Завдання 1. Аналіз залежності між змінними.....	9
Завдання 2. Статистична база економетричних моделей	10
Завдання 3. Побудова лінійної економетричної моделі	11
Завдання 4. Оцінка значущості параметрів і адекватності лінійної економетричної моделі.....	12
Тема 1.2. Нелінійна регресія	12
<i>Завдання за темою 1.2. Нелінійна регресія</i>	<i>14</i>
Завдання 1. Оцінка параметрів економетричних моделей, нелінійних за змінними.....	14
Завдання 2. Оцінка параметрів економетричних моделей, нелінійних за параметрами	15
МОДУЛЬ 2. МНОЖИННИЙ РЕГРЕСІЙНИЙ АНАЛІЗ	16
Тема 2.1. Багатофакторні моделі	16
<i>Завдання за темою 2.1. Багатофакторні моделі.....</i>	<i>19</i>
Завдання 1. Побудова багатофакторної моделі.....	19
Завдання 2. Прогнозування на підставі багатофакторної моделі	20
Завдання 3. Аналіз даних за допомогою багатофакторної моделі ..	21
Тема 2.2. Особливі випадки	22
<i>Завдання за темою 2.2. Особливі випадки.....</i>	<i>25</i>
Завдання 1. Модель із фіктивними змінними.....	25
Завдання 2. Тест Голдфельда – Квандта	27
Завдання 3. Тест Спірмена	28
МОДУЛЬ 3. МОДЕЛЮВАННЯ ДИНАМІЧНИХ ПРОЦЕСІВ.....	28
Тема 3.1. Автокореляція.....	28

<i>Завдання за темою 3.1. Автокореляція</i>	30
Завдання 1. Моделювання тренда часового ряду	30
Завдання 2. Авторегресійна модель	31
Тема 3.2. Моделювання динамічних процесів	31
<i>Завдання за темою 3.2. Моделювання динамічних процесів</i>	34
Завдання 1. Фіктивні змінні	34
Завдання 2. Ковзне середнє – адитивна модель	34
Завдання 3. Ковзне середнє – мультиплікативна модель	34
СПИСОК НАВЧАЛЬНО-МЕТОДИЧНОЇ ЛІТЕРАТУРИ	35
ДОДАТКИ	36

Навчальне видання

Данилевич Сергій Борисович
Дьячкова Ольга Володимирівна

ЕКОНОМЕТРИКА

Навчальний посібник
для студентів факультету «Бізнес-управління»,
які навчаються за напрямом підготовки
6.030504 – «Економіка підприємства

В авторській редакції

Підписано до друку 25.02.2021. Формат 60×84/16.
Папір офсетний. Гарнітура «Таймс».
Ум. друк. арк. 3,5. Обл.-вид. арк. 2,83.
Тираж 50 екз. Зам. № .

План 2020/2021 навч. р., поз. 3 в переліку робіт кафедри.

Видавництво
Народної української академії
Свідоцтво № 1153 від 16.12.2002.

Надруковано у видавництві
Народної української академії
Україна, 61000, Харків, МСП, вул. Лермонтовська, 27.