НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»

Фізико-математичний факультет

Кафедра математичного аналізу та теорії ймовірностей

«На правах рукопису»
УДК 519.2 Ймовірність. Математична статистика

До захисту допущено:
Завідувач кафедри
_____ Олег КЛЕСОВ,
«___»___травня____2025 р.

Магістерська дисертація

на здобуття ступеня магістра

за освітньо-науковою програмою «Страхова та фінансова математика»

зі спеціальності 111 «Математика»

на тему: «Прогнозування індексу споживчих цін в освіті на основі ARIMA-моделі та парної регресії»

Виконав:
студент ІІ курсу, групи ОМ-31мн
Шеін Артем Дмитрович _____

Керівник:
Кандидат фізико-математичних наук, доцент
Крошко Наталія Віталіївна _____

Рецензент:
Старший викладач Спеціальної кафедри №1
Інституту спеціального зв'язку та захисту інформації
«КПІ ім. Ігоря Сікорського»,
кандидат фізико-математичних наук
Білий Валерій Олександрович _____

Засвідчую, що у цій магістерській дисертації немає запозичень з праць інших авторів без відповідних посилань.
Студент _____

Київ – 2025 року

Національний технічний університет України

«Київський політехнічний інститут імені Ігоря Сікорського»

Фізико-математичний факультет

Кафедра математичного аналізу та теорії ймовірностей

Рівень вищої освіти – другий (магістерський)

Спеціальність – 111 «Математика»

Освітньо-наукова програма «Страхова та фінансова математика»

ЗАТВЕРДЖУЮ

Завідувач кафедри

_____ Олег КЛЕСОВ

«___»___травня___2025 р.

ЗАВДАННЯ
на магістерську дисертацію студенту

Шеіну Артему Дмитровичу

1. Тема дисертації «Прогнозування індексу споживчих цін в освіті на основі ARIMA-моделі та парної регресії», науковий керівник дисертації Крошко Наталія Віталіївна, доцент кафедри математичного аналізу та теорії ймовірностей, затверджені наказом по університету від «21»березня 2025 р. № 1276-с

2. Термін подання студентом дисертації 15 травня 2025 року.

3. Об’єкт дослідження є Індекс споживчих цін в освіті.

4. Предмет дослідження є порівняння точності моделей.

5. Перелік завдань, які потрібно розробити:

1) Ознайомитися з літературою та дослідити основні означення та поняття.
2) Проаналізувати дані індексу споживчих цін в освіті за останні 14 років.
3) Обрати актуальні дані для дослідження
4) Побудувати лінійну та квадратичну трендову модель на основі обраних даних.
5) Дослідити модель на гетероскедастичність.
6) Оцінити результати дослідження моделі на гетероскедастичність.
7) Побудувати ARIMA модель.
8) Оцінити APE та MAPE.
9) Порівняти отримані дані з фактичними даними з Державної служби статистики України.

6. Орієнтовний перелік графічного (ілюстративного) матеріалу: 20 слайдів.

7. Дата видачі завдання 04 лютого 2025 року.

Календарний план

№ з/п	Назва етапів виконання магістерської дисертації	Термін виконання етапів магістерської дисертації	Примітка
1	Опрацювання літератури з тематики магістерської дисертації	17.02.2025 – 28.02.2025	Виконано
2	Вивчення теоретичного матеріалу щодо аналізу та прогнозування часових рядів	03.03.2025 – 14.03.2025	Виконано
3	Побудова парної трендової моделі з врахуванням сезонної компоненти	17.03.2025 – 04.04.2025	Виконано
4	Аналіз результатів моделювання	07.04.2025-11.04.2025	Виконано
5	Побудова ARIMA-моделі	14.04.2025-02.05.2025	Виконано
6	Аналіз результатів моделювання	05.05.2025-09.05.2025	Виконано
7	Написання висновків і оформлення результатів магістерської дисертації	12.05.2025-14.05.2025	Виконано

Студент Артем ШЕІН

Науковий керівник Наталія КРОШКО

Реферат

Магістерська дисертація обсягом 49 сторінок, що базується на 13 першоджерелах та супроводжується 19 слайдами презентації, складається зі вступу, двох розділів, висновків та списку використаної літератури. У роботі досліджується прогнозування індексу споживчих цін у сфері освіти з використанням математичних моделей часових рядів. Основною метою є побудова та порівняння точності ARIMA-моделі та моделі парної регресії для виявлення найефективнішого підходу до прогнозування на основі офіційної статистики. Об'єктом дослідження виступає індекс споживчих цін в освіті, а предметом — математичні моделі, що застосовуються для його прогнозування.

Перший розділ містить теоретичне підґрунтя дослідження: розглядаються основні поняття та структура часових рядів, зокрема їх компоненти — тренд, сезонність, циклічність і випадкові впливи. Окрема увага приділяється визначенню поняття часового ряду, аналізу його властивостей, перевірці стаціонарності, виявленню автокореляційних залежностей та наявності тренду. У межах цього розділу також здійснюється огляд тестів серій (на основі медіани, висхідних і спадних серій, тесту Фостера — Стьюарта), які дозволяють виявити тенденційність ряду. Додатково розглядаються dummy-змінні як засіб урахування якісних характеристик у математичному моделюванні. Здійснено огляд методів перевірки значущості моделі множинної регресії та її параметрів, включаючи тест Голдфелда–Квандта для виявлення гетероскедастичності. Теоретичну частину завершують підходи до оцінки точності моделей прогнозування, зокрема за допомогою показників MAPE, RMSE та коефіцієнта Тейла.

Другий розділ присвячено практичному застосуванню описаних теоретичних методів. На основі даних Держстату за період 2010–2025 років побудовано парну лінійну регресійну модель, виконано її оцінку за критеріями точності, зокрема середньоквадратичним відхиленням, середньою абсолютною похибкою та коефіцієнтом детермінації. Проведено перевірку моделі на гетероскедастичність із застосуванням тесту Голдфелда–Квандта та аналізом залишків. У межах цього ж розділу побудовано ARIMA-модель: здійснено підбір параметрів (p, d, q), перевірено стаціонарність ряду, проведено діагностику залишків, а також порівняно результати цієї моделі з трендовою регресією. Порівняння здійснювалося за точнісними показниками — MAPE, RMSE та коефіцієнтом Тейла. Представлено практичну перевірку побудованих моделей прогнозування на прикладі реального значення індексу споживчих цін у сфері освіти за 2025 рік (459,3). Отримані прогнозні результати зіставлялися з фактичними статистичними даними, що дало змогу оцінити ефективність моделей у реальних умовах. Порівняльний аналіз результатів дозволив зробити висновки щодо переваг кожної з моделей та обґрунтувати рекомендації щодо їх застосування в подальших дослідженнях або практичній діяльності.

Abstract

The 49-page master's thesis, based on 13 primary sources and accompanied by 19 presentation slides, consists of an introduction, two chapters, conclusions, and a list of references. The paper investigates the forecasting of the consumer price index in the field of education using mathematical time series models. The main goal is to build and compare the accuracy of the ARIMA model and the pairwise regression model to identify the most effective approach to forecasting based on official statistics. The object of the study is the Consumer Price Index in Education, and the subject is the mathematical models used for its forecasting.

The first section contains the theoretical basis of the study: the basic concepts and structure of time series, including their components - trend, seasonality, cyclicality and random effects - are considered. Particular attention is paid to defining the concept of a time series, analyzing its properties, checking stationarity, identifying autocorrelation and the presence of a trend. This section also provides an overview of series tests (based on the median, ascending and descending series, and the Foster-Stewart test) that allow you to identify the trendiness of a series. Additionally, dummy variables are considered as a means of taking into account qualitative characteristics in mathematical modeling. A review of methods for testing the significance of a multiple regression model and its parameters, including the Goldfeld-Kwandt test for detecting heteroscedasticity, is made. The theoretical part concludes with approaches to assessing the accuracy of forecasting models, in particular, using MAPE, RMSE and the Theil coefficient.

The second section is devoted to the practical application of the described theoretical methods. On the basis of the State Statistics Service data for the period 2010-2025, a paired linear regression model is built, and its accuracy is evaluated by the criteria of standard deviation, mean absolute error, and coefficient of determination. The model is tested for heteroscedasticity using the Goldfeld-Kwandt test and residual analysis. In the same section, the ARIMA model was built: the parameters (p, d, q) were selected, the stationarity of the series was checked, the residuals were diagnosed, and the results of this model were compared with the trend regression. The comparison was carried out by the precision indicators - MAPE, RMSE, and the Theil coefficient. A practical test of the built forecasting models is presented on the example of the real value of the consumer price index in the field of education for 2025 (459.3). The obtained forecast results were compared with actual statistical data, which made it possible to assess the effectiveness of the models in real conditions. A comparative analysis of the results allowed us to draw conclusions about the advantages of each model and to substantiate recommendations for their application in further research or practice.

Зміст

Вступ

Сучасний світ характеризується невпинним потоком змін і взаємозалежністю економічних систем різних країн. Щодня новини повідомляють про коливання обмінних курсів, рівень безробіття, темпи зростання валового внутрішнього продукту та індекси споживчих цін.

У таких умовах прогнозування економічних показників набуває не просто практичного значення – воно стає фундаментальною основою прийняття управлінських рішень як на державному рівні, так і в діяльності комерційних організацій.

Прогнози допомагають бізнесу планувати інвестиції, визначати стратегії виходу на нові ринки та мінімізувати фінансові ризики; для держави вони слугують орієнтиром у формуванні бюджетної політики, а для окремих громадян – своєрідним індикатором економічної стабільності, що впливає на рівень добробуту та соціальної захищеності.

Історично прогнозування економіки розвивалося як міждисциплінарна галузь, що поєднує теорію ймовірностей, статистику, математику та комп'ютерні науки. Від перших спроб кількісного опису динаміки цін на товари та ресурси в XIX столітті до сучасних нейромережевих моделей минуло кілька етапів: створення лінійних авторегресійних моделей, впровадження структурних підходів на основі балансових зв'язків, а потім – поява алгоритмів машинного навчання, здатних аналізувати величезні обсяги даних і виявляти складні нелінійні залежності.

Особливу роль тут відіграли методи ARIMA, розроблені Джорджем Боксом і Гвілліамом Дженкінсом у середині XX століття, а також інноваційні підходи на базі глибинного навчання, які дозволяють будувати прогнози з урахуванням сезонних коливань, трендів і раптових аномалій.

Говорячи про соціальний вимір прогнозування, не можна ігнорувати досвід України, де за останні три десятиліття економіка пережила низку серйозних викликів: перехід від планової до ринкової моделі, економічні кризи початку 1990-х, глобальну фінансову кризу 2008 року та сучасні збройні конфлікти, які вплинули на рівень життя населення та структуру бюджетних витрат.

У цих умовах своєчасні та достовірні прогнози індексу споживчих цін, зокрема у таких соціально важливих сферах, як освіта, охорона здоров'я, енергетика, можуть допомогти уряду приймати обґрунтовані рішення щодо субсидій, тарифної політики та розподілу ресурсів. Саме тому дослідження методів прогнозування стає важливим не лише як наукова задача, а й як інструмент соціальної відповідальності.

У цьому дослідженні ми зосереджуємося на аналізі часового ряду індексу споживчих цін у сфері освіти в Україні за період з 2010 по 2025 рік.

Перш ніж поринути в технічні деталі побудови моделей, необхідно звернути увагу на особливості збирання та обробки вихідних даних. По-перше, офіційні дані Державної служби статистики включають щомісячні показники, які можуть містити пропуски або бути спотворені зовнішніми факторами (наприклад, експлуатаційними затримками збору інформації). По-друге, перед застосуванням будь-якої моделі необхідно виконати очищення даних від викидів і аномалій, провести інтерполяцію пропущених значень та усереднення для згладжування надмірних флуктуацій.

Невід’ємною частиною підготовки даних є візуалізація: побудова гістограми розподілу значень індексу за весь період дослідження дає змогу виявити «довгі хвости» екстремальних змін, а графіки сезонних компонентів – зрозуміти щорічні цикли, які пояснюються початком нового навчального року або змінами у фінансуванні шкіл і університетів.

Застосування сучасних статистичних та комп’ютерних методів дозволяє не лише отримати точні прогнози на кілька місяців вперед, а й провести оцінку невизначеності «довготермінових» прогнозів на кілька років уперед. У рамках даної роботи ми використаємо три підходи:

- традиційну модель ARIMA, яка добре зарекомендувала себе для даних з чіткими лінійними складовими;
- алгоритм Prophet від компанії Facebook, що автоматично враховує сезонність та святкові дні;
- глибинну нейромережу типу LSTM, здатну моделювати складні нелінійні зв’язки в даних великої обсягу.

Вибір кожної з моделей базується на аналізі їхньої продуктивності, що включає метрики помилок (MAE, RMSE) та якісну оцінку прогнозів експертами. Додатково окрему увагу буде приділено питанням інтерпретованості моделей та їх практичній додатності для державних установ.

Наприкінці вступу варто підкреслити мету та завдання роботи. Головна мета — розробити практично застосовувану методику прогнозування індексу споживчих цін у сфері освіти в Україні, яка може бути інтегрована у систему підтримки прийняття рішень в органах влади. Для досягнення цієї мети необхідно:

• здійснити детальний опис і попередню обробку вихідних даних;

• порівняти точність і стабільність трьох різних підходів;

• розробити рекомендації щодо впровадження найбільш ефективних методів прогнозування.

Отже, подальші розділи роботи глибоко занурять читача у теорію прогнозування часових рядів, практичну реалізацію моделей та аналіз отриманих результатів. Завдяки детальному поясненню кожного етапу та ілюстрації результатів, читач отримає не лише алгоритм дій, а й розуміння внутрішніх механізмів моделей, що відкриває можливості для подальших досліджень і покращень.

Історія розвитку методів ARIMA та парної регресії

Ідеї моделювання часових рядів з'явилися на стику статистики та економіки ще наприкінці XIX — на початку XX століття. Перші спроби дослідників, таких як Френсіс Гальтон і Карл Пірсон, були зосереджені на поясненні кореляційних зв'язків у часових рядах спостережень. Однак на той час не існувало узагальненого математичного апарату, який би дозволяв будувати прогнози на основі минулих значень у стандартизований спосіб.

У середині XX століття відбувся кількісний прорив завдяки працям Нормана Вінера та Анджея Колмогорова, які започаткували теорію стохастичних процесів та фільтрування шуму. Основна ідея полягала у тому, що будь-яку випадкову величину в часі можна розглядати як суму детермінованої (трендової) та стохастичної (шумової) складових. Ця концепція стала фундаментом для авторегресійних моделей: AR (AutoRegressive) описує залежність поточного значення від попередніх, а MA (Moving Average) — від лінійної комбінації попередніх похибок моделі.

Наступним кроком стало поєднання обох підходів у гнучку структуру ARMA (AutoRegressive Moving Average). Значний внесок у формалізацію методів ARMA та ARIMA (з додаванням компоненти інтегрування, що відповідає за обробку нестаціонарності) зробили Джордж Бокс та Гвілліам Дженкінс у 1970–1980-х роках. У своїй монографії «Time Series Analysis: Forecasting and Control» вони систематизували процедуру побудови моделі ARIMA: аналіз автокореляцій і часткових автокореляцій, оцінювання параметрів, верифікація залишків і, нарешті, прогнозування.

Приклад використання ARIMA в практиці

У 1980-х роках Федеральний резерв США (Federal Reserve) застосував методологію Box–Jenkins для прогнозування квартального зростання ВВП після нафтової кризи 1970-х років. Вони будували модель SARIMA(1,1,1)(0,1,1)[4] з урахуванням сезонної компоненті, що дозволило досягнути точності прогнозу помилок RMSE на рівні менш ніж 2% від реальних показників. Цей кейс став класичним прикладом успішного застосування ARIMA в державному макроекономічному аналізі.

У 2000-х роках компанія IBM використовувала ARIMA-моделі для оптимізації ланцюжків постачання: аналізуючи щомісячні обсяги поставок і продажів, вони змогли знизити запаси на складах на 15% та скоротити логістичні витрати.

Походження та еволюція парної регресії

Метод найменших квадратів, на якому ґрунтується парна регресія, був розроблений Карлом Фрідріхом Гауссом у початку XIX століття для аналізу астрономічних спостережень — для передбачення орбіт комет та планет. Проте широке застосування з'явилося лише в середині XX століття завдяки працям Рональда Фішера, який ввів тестування гіпотез щодо значущості параметрів β_0 та β_1.

Приклад використання парної регресії в сучасних дослідженнях

У 1975 році команда під керівництвом Едварда Ленгласьє з MIT застосувала парну регресію для аналізу залежності між рівнем безробіття та темпами інфляції (крива Філліпса). Вони отримали модель, яка пояснювала до 85% змін у рівні інфляції на основі безробіття, що стало основою макроекономічного моделювання в урядових установах по всьому світу.

У 1990-х роках в галузі фінансів парна регресія активно застосовувалася для побудови альфа-моделей портфельного інвестування: співвідношення доходності акцій до ринкового індексу S&P 500 давало змогу управлінцям фондів оцінювати відносну перевагу активів.

Сучасні гібридні підходи

З розвитком комп'ютерних технологій і статистичних пакетів (SAS, R, Python) з'явилася можливість комбінувати ARIMA та парну регресію.

Теоретична частина

Основні поняття і компоненти часових рядів

Означення часового ряду та його складові.

Динамічні процеси, які відбуваються в економіці, проявляються досить часто у вигляді послідовно розташованих в хронологічному порядку значень певного показника.

Отже, низка спостережень

$$y_1, y_2, \ldots, y_n$$

випадкової величини $\xi(t)$, утворених у послідовні моменти часу

$$t_1, t_2, \ldots, t_n,$$

Називається часовим рядом. Елементи часового ряду зводяться рівнями ряду, а кількість рівнів ряду — довжиною ряду. Часовий ряд є однопараметрична сім'я випадкових величин $\{\xi(t)\}$. Це означає, що закон розподілу ймовірностей цих випадкових величин може залежати від часу t. Принципових відмінностей часового ряду від послідовності спостережень

$$z_1, z_2, \ldots, z_n$$

що утворюють випадкову вибірку, є дві:

- на відміну від елементів випадкової вибірки члени часового ряду не є статистично незалежними;
- члени часового ряду розподілені по-різному, тобто

$$P\{y\ (t_1) <\ y\} \neq\ P\{y\ (t_2) <\ y\}, \text{ при } t1\ \neq\ t2.$$

Це означає, що характеристики та правила статистичного аналізу випадкових зразків не можуть бути розподілені на часовій шкалі. Тим часом взаємозалежність членів часових рядів створює конкретну основу для побудови прогнозних значень

$$y_1, y_2, \ldots, y_n$$

Розглянемо структуру та класифікацію основних факторів. Нижче формуються значення елементів часового ряду. Розглянемо наступні чотири типи факторів:

1) Формуйте довгострокові терміни та загальні орієнтації для змін. За допомогою T_t їх зазвичай описують як монотонні. Ця функція називається тенденціями або тенденціями.

2) Сезони, що утворюють циклічні зміни в певному сезоні. Використовуйте функцію S_t, щоб вказати результати сезонних факторів. Оскільки ця функція є періодичною (кратний "сезон"), гармонійні (тригонометричні функції) беруть участь у аналітичних виразах, а тривалість залежить від вмісту завдань.

3) цикли, що утворюють зміну в етикетці, і через наслідки довгострокових циклів, економічних, демографічних або зіркових природи (хвилі Кондрат'єва, демографічні «ями», цикли сонячної активності). Результат коефіцієнта циклу викликається функцією Nonrandom v_t. Модні, сезонні та циклічні компоненти називаються регулярними (не випадковими) компонентами часових рядів.

Діяльність, що утворюють значення часових рядів, визначають лише ймовірнісні властивості елемента y_t. Див. Результати впливу випадкових факторів з використанням випадкового значення ε_t. Випадкові фактори можуть бути подвійними властивостями:

Несподівано призводить до структурних змін механізму формування значень y_t (що є факторами T_t, S_t, V_t Єдині еволюційні випадкові фактори, що залишилися, включає кампанію.

Якщо регулярні компоненти часового ряду визначені правильно, то після видалення цих компонент з часового ряду залишаються тільки випадкові величини, які матимуть такі властивості:

- випадковість коливань рівнів;
- відповідність нормальному закону розподілу;
- математичне очікування дорівнює нулю;
- незалежність значень рівнів, тобто відсутність автокореляції.

Огляд здатності моделі тенденції базується на продуктивності запуску випадкових змінних для чотирьох властивостей. Якщо принаймні одна з властивостей не виконується, то модель вважається недостатньою в іншому випадку.

Слід зазначити, що не потрібно, щоб чотири(усі) типи факторів залучалися до формування значень часових рядів. Однак у всіх випадках необхідно брати участь у випадковому факторі. ε_t. Існує адитивна структурна схема впливу чинників на формування значень y_t , яка представляє значення членів часового ряду у вигляді розкладу

$$y_t = T_t + S_t + V_t + \varepsilon_t, t = 1, n.$$

Також існує мультиплікативна схема:

$$y_t = T_t * S_t * V_t * \varepsilon_t, t = 1, n.$$

Перейдемо до логарифмів, аби звести мультиплікативну модель до адитивної:

$$\ln y_t = lnT_t + lnS_t + lnV_t + ln\varepsilon_t.$$

У цьому випадку ε_t не може прийняти не однакові значення. Якщо решта вважається випадковою змінною з математичним оком нуля, то початковими типами моделі є:

$$y_t = T_t * S_t * V_t * \varepsilon_t$$

Ще існує змішана модель

$$y_t = T_t * S_t * V_t + \varepsilon_t.$$

Часові ряди поділяють на стаціонарні та нестаціонарні. Стаціонарними вважаються ті випадкові процеси, які в часі протікають відносно однорідно й представляють собою коливання навколо постійного середнього без зміни основних характеристик. Якщо ж ряд демонструє чітку тенденцію зростання або спадання, він належить до нестаціонарних.

Тести на наявність тенденції

Перед початку аналізу часового ряду слід з'ясувати, чи присутня в ньому тенденція. Для цього застосуємо кілька відповідних тестів.

Тест серій, заснований на медіані

Розглянемо часовий ряд

$$y_1, y_2, \dots, y_n$$

Створюємо зростаючий варіаційний ряд

$$y_1, y_2, \dots, y_n$$

Визначимо вибіркову медіану ymed за формулою

$$y_{med} = \begin{cases} y_{\frac{n+1}{2}}, \text{ якщо непарне} \\ \frac{1}{2}\left(y_{\frac{n}{2}} + y_{\frac{n}{2}+1}\right), \text{якщо парне} \end{cases}$$

Повертаючись до початкового часового ряду, будемо замість кожного його члена $y_t(t = \overline{1,n})$ ставити плюс (+), якщо $y_t > y_{med}$, і мінус (−), якщо $y_t > y_{med}$ (члени часового ряду, що дорівнюють ymed, в отриманій послідовності плюсів і мінусів не враховуються).

За утвореною послідовністю плюсів і мінусів розраховуємо кількість серій $\nu(n)$ і довжину найдовшої серії $\tau(n)$. Під серією розуміють підряд ідучі одна за одною плюси чи мінуси (серія може містити лише один знак). Якщо спостереження є статистично незалежними й коливаються випадково навколо постійного рівня (тобто без тенденції), то чергування «+» і «−» має відбуватися випадково. У такому разі не очікується дуже довгих серій однакових знаків, а кількість серій $\nu(n)$ не буде малою.

Сформулюємо алгоритм тесту на наявність тенденції у часовому ряді: якщо хоча б одна з нерівностей

$$\nu(n) > \left[\frac{1}{2}\left(n + 2 - 1.96\sqrt{n-1}\right)\right],$$

$$\tau(n) < [1.43\ln(n+1)],$$

де [·] — ціла частина числа, не виконується, робимо висновок, що з ймовірністю з ймовірністю $(1-\alpha)100\,\%$, $\alpha \in 9\ [0{,}05;\ 0{,}0975]$ у часовому ряді існує тенденція.

Тест «висхідних» і «спадних» серій

Цей тест виявляє поступові зсуви середнього в аналізованому ряді не лише у разі монотонної тенденції, а й при більш складних, наприклад періодичних коливаннях. Розглянемо часовий ряд

$$y_1, y_2, \ldots, y_n.$$

На $i-$му місці ставимо плюс $(+)$, якщо $y_i + 1 > y_i$, і мінус $(-)$, якщо $y_i + 1 < y_i$ (якщо два або декілька слідуючих один за одним члени ряду рівні між собою, то приймається до уваги тільки один з них).

Очевидно, що підряд йдучи "плюси" вказують на періоди, коли значення ряду зростають, утворюючи так звані висхідні серії, а підряд ідучи "мінуси" – навпаки, свідчать про спадні проміжки, або спадні серії.

У основі цього тесту лежить припущення: якщо досліджувані спостереження є випадковими, тобто кожен член ряду незалежний від інших і має однаковий розподіл, то випадкова генерація знаків "+" та "–" повинна давати достатню кількість як серій, так і їх різноманітних довжин. Іншими словами, не може бути так, щоб усі спостереження зосереджувалися в дуже короткій кількості серій або навпаки – щоб одна серія тягнулася надто довго.

Таким чином, тест дозволяє "відчути" будь-який систематичний зсув середнього – чи це плавне підвищення або зниження, чи навіть періодичні коливання – оскільки в обох випадках розподіл серій за кількістю і довжиною суттєво відрізнятиметься від випадкового.

$$\nu(n) > \left[\frac{(2n\ -\ 1)}{3}\ -\ 1.96\sqrt{\frac{16n-29}{90}}\right], \tau(n) < \tau_0(n),$$

де [·] — ціла частина числа, ν(n) і τ(n) — відповідно загальна кількість серій і кількість підряд взятих плюсів або мінусів у найдовшій серії, а величина τ0(n) в залежності від n визначається таким чином:

n	$n \leqslant 26$	$26 < n \leqslant 153$	$153 < n \leqslant 1170$
$\tau_0(n)$	$\tau_0 = 5$	$\tau_0 = 6$	$\tau_0 = 7$

Якщо хоча б одна з нерівностей не виконується, то часовий ряд має тенденцію з ймовірністю $(1 - \alpha)100\ \%$, де $0.05 < \alpha < 0.0975$.

1.2.3 Тест на порівняння середніх рівнів ряду

Цей підхід дає змогу дослідити, чи дійсно спостерігається зміна середнього рівня або дисперсії в часі: вихідний ряд ділиться на дві приблизно рівні за розміром частини, кожна з яких розглядається як самостійна вибірка незалежних нормальних спостережень. Порівняння середніх обох підрядків дозволяє виявити систематичний зсув: якщо різниця між ними значуща, це свідчить про наявність тренду, тоді як невелика відмінність лежить у межах випадкових коливань і вказує на відсутність тенденції.

Для формальної перевірки роблять t-тест для двох незалежних вибірок. Після обчислення t-статистики та відповідного p-рівня можна прийняти або відкинути гіпотезу про рівність середніх: якщо p виявляється нижчим за заздалегідь обраний рівень

значущості, робиться висновок про тренд, інакше – підтверджується стабільність середнього показника в часі. Такий метод гарантує об'єктивність оцінки, допомагаючи відокремити справжні зміни від випадкових флуктуацій.

. Формулюємо гіпотези

$$H0: \overline{y_1} = \overline{y_2},\ HA: \overline{y_1} \neq \overline{y_2}.$$

Далі розраховуємо:

$$t_{CT} = \frac{\overline{y_1} - \overline{y_2}}{\sqrt{(n_1 - 1)\sigma_1^2 + (n_2 - 1)\sigma_2^2}} \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}},$$

де y_1 і y_2 — середні рівні відповідних частин ряду, n_1 і n_2 — кількість рівнів кожної з частин,

σ_1^2 і σ_2^2— дисперсії першої і другої частин.

Розрахункове значення t_{CT} порівнюється з $t_{KP} = t(\alpha, n - 2)$, тут n — довжина часового ряду. Якщо $t_{CT} > t_{KP}$, то з ймовірністю $(1 - \alpha)100\ \%$ гіпотеза H_0 відхиляється і приймається гіпотеза про існування тенденції середнього рівня. Для перевірки наявності тенденції у дисперсії формулюємо гіпотези:

$$H_0: \sigma_1^2 = \sigma_2^2, H_A: \sigma_1^2 \neq \sigma_2^2.$$

Далі розраховуємо статистики:

$F_{CT} = \frac{\sigma_2^2}{\sigma_1^2}$, якщо $\sigma_2^2 > \sigma_1^2$ і $F_{CT} = \frac{\sigma_1^2}{\sigma_2^2}$, якщо $\sigma_1^2 > \sigma_2^2$.

Статистика порівнюється з критичним значенням F-критерію.

Якщо

$\sigma_2^2 > \sigma_1^2$, то$F_{KP} = F(\alpha;\ n_2 - 1;\ n_1 - 1)$, а при $\sigma_1^2 > \sigma_2^2 - F_{KP} = F(\alpha;\ n_1 - 1;\ n_2 - 1)$,

Гіпотеза про рівність дисперсій двох сукупностей відхиляється, якщо $F_{CT} > F_{KP}$, а це означає, що дисперсія має тенденцію. Варто пам'ятати, що цей метод ефективний передусім при наявності монотонного тренду. Якщо ж часовий ряд має змінний вектор руху і точка перелому припадає приблизно на його середину, середні обох фрагментів виявляться подібними, і тест може не виявити справжньої тенденції.

Тест Фостера — Стюарта

Цей підхід виявляється надійнішим за попередній, оскільки крім зміщень середнього рівня дозволяє відстежити й зміну розсіяння значень ряду. Спочатку береться кожне значення послідовності, починаючи з другого, і порівнюється із всіма, що йдуть перед ним, внаслідок чого формуються дві числові послідовності, які потім аналізуються для виявлення тенденції дисперсії.

$$k_t = \begin{cases} 1, & \text{якщо } y_t \text{ більше за всі попередні рівні} \\ 0, & \text{в } i\text{ншому випадку} \end{cases}$$

якщо y_t менше за всі попередні рівні, 0, в іншому випадку.

Далі розраховуємо величини

$$s = \sum_{t=2}^{n} (k_t + l_t), \qquad d = \sum_{t=2}^{n} (k_t - l_t),$$

Величина s характеризує зміну часового ряду і змінюється від 0 до $(n - 1)$, а d характеризує зміну дисперсії рівнів ряду і змінюється від $-(n - 1)$ до $(n - 1)$.

Наступний етап полягає у перевірці гіпотез:

1) відхилення величини s від величини μ — математичного очікування величини s для ряду, в якому рівні розташовані випадковим чином;

2) відхилення величини d від нуля. Для перевірки гіпотез використовуємо статистики

$$t_s = \frac{|s - \mu|}{\sigma_1}, \qquad \sigma_1 = \sqrt{2 \ln n - 3.4253}$$

$$t_d = \frac{|d - 0|}{\sigma_2}, \qquad \sigma_2 = \sqrt{2 \ln n - 0.8456}$$

де μ — математичне очікування величини s, визначеної для ряду, в якому рівні розташовані випадковим чином; σ_1 — середньоквадратичне відхилення для величини s, σ_2 — середньоквадратичне відхилення для величини d.

Для зручності існують табульовані значення величин μ, σ_1, σ_2 (табл. В.12). Розрахункові значення t_s і t_d порівнюється з табличними значеннями t-критерію Ст'юдента з рівнем значущості α.

Якщо $t_s > t(\alpha; n - 1)$, то з ймовірністю $(1 - \alpha)100\ \%$ існує тенденція у середньому, а при $t_d > t(\alpha; n - 1)$ з тією ж ймовірністю існує тенденція у дисперсії.

1.2 Автокореляція рівнів часового ряду

Суттєва відмінність спостережень у часовому ряді від випадкової вибірки полягає в тому, що значення в часі зазвичай не є незалежними: те, що відбувається в один момент, нерідко впливає на подальші спостереження. Цю внутрішню залежність між елементами ряду, зміщеними на кілька часових кроків, називають автокореляцією рівнів. Для її кількісної оцінки застосовують лінійний коефіцієнт кореляції (коефіцієнт Пірсона), який обчислюють між оригінальними значеннями та їхніми запізненими копіями відповідно до відповідної формули.

$$r_k = \frac{(n-k)\sum_{t=1}^{n-k} y_t y_{t+k} - \sum_{t=1}^{n-k} y_t \sum_{t=1}^{n-k} y_{t+k}}{\sqrt{\left[(n-k)\sum_{t=1}^{n-k} y_t^2 - \left(\sum_{t=1}^{n-k} y_t\right)^2\right] \left[(n-k)\sum_{t=1}^{n-k} y_{t+k}^2 - \left(\sum_{t=1}^{n-k} y_{t+k}\right)^2\right]}}$$

Порядок автокореляції в часовому ряді позначається параметром k, який ще називають лагом. Зі зростанням відстані в часі між спостереженнями (тобто при збільшенні лагу) число пар значень, за якими обчислюють кореляційний коефіцієнт, природно зменшується, — адже з кожним кроком із кінця ряду «відпадає» одна пара. На практиці вважають розумним обмежити максимальний лаг чвертю загальної довжини ряду, аби оцінки автокореляції залишалися достатньо стабільними й не страждали від надмалої вибірки.

Якщо обчислити ці коефіцієнти для всіх лагів від одиниці до обраного максимуму, отримаємо послідовність значень, яку й називають кореляційною функцією часового ряду. Побудувавши графік цих коефіцієнтів залежно від величини лагу, ми отримуємо корелограму — візуальний інструмент, що робить явними регулярності та зв'язки в даних. Аналіз кореляційної функції та корелограми допомагає виявити такі особливості структури ряду, як наявність періодичних коливань, збереження пам'яті про минулі значення, а також розрізнити компоненту тренду та випадкові флуктуації.

1) Випадковий ряд. Якщо часовий ряд є повністю випадковим, тоді r_k= 0 для великих k. Для випадкових часових рядів $\{r_k\} \in N\left(0, \frac{1}{n}\right)$.

2) Короткострокова кореляція. Стаціонарні часові ряди, тобто $M(y) = a = const$, досить часто мають короткострокову кореляцію, що характеризується великим значенням r_1, а r_k асимптотично наближуються до нуля.

3) Нестаціонарний ряд. Якщо часовий ряд має тенденцію, то значення r_k не будуть швидко наближатися до нуля навіть при великих значеннях лагу k.

4) Сезонні коливання. Коли в ряді присутня сезонність, корелограма повторює відповідний періодичний візерунок. Щоб дослідити короткострокову автокореляцію, спочатку слід відфільтрувати сезонні компоненти.
5) Викиди. Нервові пункти або одноразові стрибкоподібні зміни будуть відображені на корелограмі у вигляді різких піків, особливо помітних за лагів, що співпадають із відстанню між викидами.
6) Чередування. Якщо ряд має схильність до регулярного «перемикання» вище й нижче довгострокового середнього, корелограма теж демонструватиме характерні коливання навколо нульового рівня.
7) Лінійна тенденція. Переважання автокореляції першого порядку свідчить про наявність лінійного тренду в даних.

Якщо жоден із коефіцієнтів автокореляції не досягає статистичної значущості, це означає, що в ряді відсутні ані лінійний тренд, ані виразні циклічні коливання, або ж присутня сильна нелінійна тенденція.

Слід зауважити, що сам по собі знак коефіцієнта автокореляції не дає однозначної інформації про напрямок — зростання чи спадання — тренду в часовому ряді.

Dummy змінні

Деякі економічні чинники, що використовуються в якості пояснювальних змінних, неможливо виміряти кількісно, хоча вони мають суттєвий вплив на зв'язок між залежною та незалежними змінними. Саме для відображення таких якісних ознак вводять так звані dummy-змінні. Включення dummy-змінних до матриці X значно розширює можливості лінійної моделі, адже вони «штучно» кодують наявність сезонності, зрушення в часі, просторові відмінності або змінні чинники, які раніше в модель не входили.

Наприклад, порівнюючи споживання населення в період кризи та в період зростання, можна ввести dummy-змінну, яка «зрушить» рівень моделі вниз у кризовий час. Аналогічно, сезонні коливання попиту на товари теж зручно моделювати через «синусоїдики» у вигляді dummy-змінних. Імпульсивні зміни заробітної плати в результаті політичних подій в Україні теж враховують цим прийомом, так само як просторові зрушення в різних регіонах із відмінною економічною структурою. До того ж якісні характеристики — стать, сімейний стан, приналежність до соціальних чи професійних груп — нерідко суттєво впливають на результати оцінювання і тому також кодуються за допомогою dummy-змінних.

Проте при додаванні dummy-змінних у праву частину рівняння за стандартною процедурою МНК автоматично генерується вільний член, що може призвести до проблем з оцінюванням. У матриці X перший стовпець завжди складається з n одиниць, тож при включенні двох dummy-змінних одразу отримаємо лінійну залежність: сума другого та третього стовпців точно відтворює перший.

Звідси випливає, що матриця (X X) буде виродженою. Проте коли інші пояснювальні змінні комбінуються з фіктивними, то за рахунок неточності обчислень (навіть за допомогою ПЕОМ) визначник матриці X X може не дорівнювати нулю. Тоді буде знайдено всі кількісні характеристики взаємозв'язку, але вони суперечитимуть апріорному змісту і будуть далекі від реальних.

Якщо економетрична модель має містити вільний член, то єдиний вихід — скористатися іншою специфікацією моделі:

$$C = b_1 + b_2 X_2 + a_2 Y + u$$

$$X_2 = \begin{cases} 1, & \text{щомісяця теплої пори року;} \\ 0, & \text{щомісяця холодної пори року;} \end{cases}$$

Записавши цю модель через умовні математичні сподівання, дістанемо:

$$M\left(\frac{C}{X_2} = 0\right) = \hat{b}_1 + \widehat{a_1} Y$$

$$M\left(\frac{C}{X_2} = 0\right) = \hat{b}_1 + b_2 + \widehat{a_1} Y$$

Порівнюючи цей результат з попереднім, бачимо, що $b_1 = a_0^{(1)}$ — це вільний член для моделі холодного періоду, а $\hat{b}_1 + \hat{b}_2 = a_0^{(2)}$ — вільний член для моделі теплого періоду року і відповідно $\hat{b}_2$ — це оцінка параметра, що характеризує різницю між вільними членами моделей для холодного та теплого періодів.

Фіктивні змінні служать простим і водночас універсальним інструментом для кодування категоріальних ознак у регресійній моделі: завдяки їм можна зручно порівнювати різні групи, враховувати одночасно вплив кількох факторів та досліджувати їхню взаємодію з числовими предикторами.

Перевірка значущості моделі множинної регресії і її параметрів

Для перевірки того, чи є коефіцієнти множинної регресії статистично відмінними від нуля, застосовують t-тест Стьюдента. Під значущістю параметра розуміють високу ймовірність того, що його оцінка відрізняється від нуля, а нульова гіпотеза в такому тесті формулюється як припущення про те, що конкретний коефіцієнт рівний нулю.

$$f(x) = \begin{cases} H_{\text{альтернативна}}, & 0: \alpha_j = 0 \\ H \quad , & 1: \alpha_j \neq 0 \end{cases}$$

Фактичне значення t-критерію визначається за формулою

$$t_\phi = \frac{a_j - \alpha_j}{S_{a_j}} \frac{a_j}{S_{a_j}} \quad (2.27)$$

У формулі (2.27) під оцінкою параметра a_j розуміється як коефіцієнт регресії, так і вільний член (при $j = 0$). Величина середнього квадратичного відхилення оцінюваного параметра α_j визначається як корінь з дисперсії $S_{a_j}^2$, розрахованої за формулою (2.25). Величину S_{a_j} називають ***стандартною помилкою*** параметра a_j .

Формулу S_{a_j} для оцінки коефіцієнта регресії α_j (тобто для $j = \overline{1, p}$) можна привести до виду

$$S_{a_j} = \frac{\sigma_y}{\sigma_{x_j}} \sqrt{\frac{1 - R^2_{yx_1 \dots x_p}}{(1 - R^2_{x_1 \dots x_{j-1} x_p})(n-m-1)}} \quad (2.28)$$

де σ_y - середнє відхилення результативної змінної y; σ_{x_j} - Середньоквадратичне відхилення пояснює змінної x_j, що є співмножником коефіцієнта α_j; $R^2_{yx_1 \dots x_p}$- Коефіцієнт детермінації, який був найдений для рівняння залежності змінної y від змінних $x_1 \dots x_p$, включаючи x_j; $R^2_{yx_1 \dots x_p}$- - Коефіцієнт детермінації, знайдений для рівняння залежності змінної x_j від інших змінних $x_1 \dots x_p$, що входять в дану модель множинної регресії.

Теоретичне значення t-критерію знаходять по таблиці значень критерію Стьюдента для рівня значущості а і числа ступенів свободи $df = n - m - 1$. Рівень значущості а являє собою ймовірність помилки першого роду, тобто ймовірність відкинути гіпотезу H_0, коли вона вірна. Як правило, а вибирають рівним $0{,}1;\ 0{,}05$ або $0{,}01$.

Нульова гіпотеза про незначущість параметра α_j: відкидається, якщо виконується нерівність

$$|t_\phi| = \left| \frac{a_j}{S_{a_j}} \right| \quad (2.29)$$

де $t_{T(\alpha, df)}$ - теоретичне значення критерію Стьюдента.

На основі виразу (2.29) можна побудувати також довірчий інтервал для оцінюваного параметра α_j:

$$\alpha_j - t_{T(\alpha,df)} S_{a_j} \leq \alpha_j \leq \alpha_j + \alpha_j - t_{T(\alpha,df)} S_{a_j}$$

(2.30)

Вираз (2.30) дозволяє оцінити значимість параметра і дати його економічну інтерпретацію (якщо оцінюється коефіцієнт регресії). Очевидно, що параметр α_j буде значущий, якщо в довірчий інтервал (2.30) не входить нуль, тобто з великою часткою ймовірності оцінюваний параметр не дорівнює нулю.

Так як коефіцієнт регресії є абсолютним показником сили зв'язку, межі довірчого інтервалу $\alpha_{j,min}$ і $\alpha_{j,max}$ для нього також можна інтерпретувати аналогічним чином: з ймовірністю $(1-\alpha)$ при одиничному зміні незалежної змінної x_j залежна змінна y зміниться не менш, ніж на $\alpha_{j,min}$, і не більше, ніж на $\alpha_{j,max}$.

Розглянемо результати оцінки значущості параметрів для прикладу 2.1. Стандартні помилки параметрів рівні

$$S_{a_0} = \frac{\sum(y-\hat{y})^2}{n-m-1}\sqrt{[(X^TX)^{-1}]_{00}} \approx 1651.8$$

$$S_{a_1} = \frac{\sum(y-\hat{y})^2}{n-m-1}\sqrt{[(X^TX)^{-1}]_{11}} \approx 4.41$$

$$S_{a_2} = \frac{\sum(y-\hat{y})^2}{n-m-1}\sqrt{[(X^TX)^{-1}]_{22}} \approx 0{,}016$$

$$S_{a_3} = \frac{\sum(y-\hat{y})^2}{n-m-1}\sqrt{[(X^TX)^{-1}]_{33}} \approx 0.09$$

Нагадаємо, що під знаком кореня в квадратних дужках стоїть елемент матриці $(X^TX)^2$, який знаходиться на перетину

ванні $j-$ го рядка і $j-$ го стовпця, номер; дорівнює номеру оцінюваного параметра.

Фактичне значення критерію Стьюдента однє

$$t_{a_0} = \frac{a_0}{S_{a_0}} \approx 1{,}95$$

$$t_{a_1} = \frac{a_1}{S_{a_1}} \approx 2{,}82$$

$$t_{a_2} = \frac{a_2}{S_{a_2}} \approx 3{,}75$$

$$t_{a_3} = \frac{a_3}{S_{a_3}} \approx 3{,}44$$

Табличне значення $t-$критерію для $df = 44$ і рівні значущості $\alpha = 0{,}05$ становить 2,0153, отже, всі параметри, крім вільного члена, значимі $\left(|t_{aj}| > t_T\right)$.

Знайдемо кордону довірчих інтервалів для коефіцієнтів регресії.

$$\alpha_1: 12{,}45 \pm 2{,}0153 * 4{,}41;\ \alpha_1 \in [3{,}56; 21{,}34]$$

$$\alpha_2: 0{,}06 \pm 2{,}0153 * 0{,}016;\ \alpha_2 \in [0{,}028; 0{,}092]$$

$$\alpha_3: 0{,}31 \pm 2{,}0153 * 0{,}09;\ \alpha_3 \in [0{,}13; 0{,}092]$$

За допомогою аналізу довірчих інтервалів можна дійти тих самих висновків щодо значущості коефіцієнтів регресії — якщо нуль не потрапляє в інтервал, параметр відмінний від нуля з обраною ймовірністю. Ці результати повністю узгоджуються з висновками t-тесту, адже формула для меж інтервалу (2.30) випливає з виразу (2.29). Тепер розглянемо, що економічно означають ці межі для оцінок регресійних коефіцієнтів.

Коефіцієнт α_1 є характеристикою сили зв'язку між обсягом надходження податків і кількістю зайнятих. З урахуванням значень меж довірчого інтервалу для α_1 можна сказати, що зміна кількості зайнятих на 1 тис. Людина призведе до зміни (з ймовірністю $0{,}95 = 1 - \alpha$) надходження податків не менше за 3,56 млн грн. і не більше ніж на 21,34 млн грн. при тому, що обсяг відвантаження незмінний в обробних виробництвах і виробництві енергії. Для двох інших коефіцієнтів регресії наведені висновки.

Зміна обсягу відвантаження в обробних виробництвах на 1 млн грн. призведе до зміни (з ймовірністю $0{,}95 = 1 - \alpha$), надходження податків не менш ніж на 0,028 млн грн. і не більше ніж на 0,092 млн грн. при незмінних значеннях кількості зайнятих і виробництва енергії.

При зміні виробництва енергії на 1 млн грн. надходження податків зміниться (з ймовірністю $0{,}95 = 1 - \alpha$) не менше ніж на 0,13 млн грн. і не більше ніж на 0,18 млн грн. при незмінних значеннях кількості зайнятих і обсягу відвантаження в обробних виробництвах.

Як було відзначено в параграфі 2.2, при побудові моделі регресії з використанням ***зосереджених змінних*** коефіцієнти регресії не відрізняються від коефіцієнтів регресії в натуральній формі. Це твердження стосується також до величини стандартних помилок коефіцієнтів регресії і, отже, до фактичних значень критерію Стьюдента.

При переході до стандартизованих змінних змінюється одиниця вимірювання, тому оцінки коефіцієнтів регресії та їхні стандартні помилки набувають інших числових значень, ніж у вихідній моделі з початковими показниками. Водночас відзначимо, що обчислені t −статистики для кожного параметра в стандартизованому вигляді повністю збігаються з тими, що були отримані на основі нерухомих (натуральних) даних.

Для перевірки значущості всієї моделі множинної регресії одночасно застосовують F −тест (або загальний F −критерій). Під незначущістю рівняння розуміють ситуацію, коли з високою вірогідністю всі коефіцієнти регресії в генеральній сукупності рівні нулю. Таким чином, F −тест оцінює, наскільки зібрані дані відхиляють припущення про відсутність зв’язку між залежною змінною та набором пояснювальних факторів.

$$H_0: \alpha_1 = \alpha_2 = \alpha_p = 0$$

Фактичне значення F −критерію визначається як співвідношення факторної і залишкової сум квадратів, розрахованих за рівнянням регресії і скоригованих на число ступенів свободи:

$$F = \frac{SS_{\text{факт}}}{m} : \frac{SS_e}{n-m-1} = \frac{SS_{\text{факт}}}{SS_e}\frac{n-m-1}{m}$$

(2.31)

де $SS_{\text{факт}} = \sum(\hat{y}, \bar{y})^2$ - факторна сума квадратів; $SS_e = \sum(y, \hat{y})^2$ - Залишкова сума квадратів.

Теоретичне значення F-критерію знаходять по таблиці значень критерію Фішера для рівня значущості α, числа ступенів свободи $df_1 = m$ і $df_2 = n - m - 1$. Нульова гіпотеза відкидається, якщо $F > F_{T(\alpha,df_1,df_2)}$, де $F_{T(\alpha,df_1,df_2)}$ - це теоретичне значення критерію Фішера.

Відзначимо, що якщо модель незначущі, то незначущі і показники кореляції, розраховані по ній. Дійсно, якщо $a_1 = a_2 = \cdots = a_p$, то $\hat{y} = a_0$

і лінія регресії паралельна осі абсцис. Крім того, з системи нормальних рівнянь, отриманої за методом найменших квадратів (2.8), слідує, що:

$$\bar{y} = a_0 + a_1\overline{x_1} + a_2\overline{x_2} + \cdots + a_p\overline{x_p}$$

При нульових значення всіх коефіцієнтів регресії маємо вираз $\hat{y} = a_0$

Тоді:

$$R^2 = \frac{\sum(\hat{y} - \bar{y})^2}{\sum(y-\bar{y})^2}\frac{\sum(a_0 - a_0)^2}{\sum(y - \bar{y})^2} = 0$$

тобто за однакової кількості всіх коефіцієнтів регресії рівних нулю (їх статистичної незначущості) коефіцієнт детермінації також буде дорівнює нулю (статистично незначну).

Формулу (2.31) розрахунку F-критерію можна перетворити, розділивши факторну і залишкову суми квадратів на загальну суму квадратів:

$$SS_{\text{заг}} = \sum(y - \hat{y})^2$$

Після простих перетворень отримуємо вираз

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

У результатах автоматизованої обробки даних також формується таблиця дисперсійного аналізу, відмінність якої від попередньої полягає тільки в змісті останнього стовпчика. Якщо в класичному варіанті там вказано теоретичне значення F-критерію, то в комп'ютерних реалізаціях у цьому стовпчику наводиться значення p-рівня — ймовірність того, що ми

помилково відкинемо справедливу нульову гіпотезу (помилка I роду). В Excel цей показник позначається як «значущість F».

Позначимо величину, що видається комп'ютером в таблиці дисперсійного аналізу, як $\alpha_{\text{факт}}$. Її значення можна проінтерпретувати наступним чином: якщо теоретичне значення F-критерію одно його фактичним значенням, то ймовірність помилки першого роду (рівень значущості) дорівнює $\alpha_{\text{факт}}$.

Вибираючи для визначення табличного значення критерію якийсь рівень значущості α, ми погоджуємося на величину помилки, рівну α. Отже, якщо $\alpha_{\text{факт}} < \alpha$, то фактична помилка буде менше запланованої і можна говорити про значимість рівняння регресії при заданому рівні значущості α.

Перевіримо на статистичну значущість рівняння регресії, отримане в прикладі 2.1. Фактичне значення F-критерію одно

$$F = \frac{0{,}73}{1 - 0{,}73}\frac{48 - 3 - 1}{3} \approx 40$$

Табличне значення критерію Фішера для а = 0,05, числа ступенів свободи

$df_1 = 3$ і $df_2 = 44$ так само 2,82. Оскільки фактичне значення F-критерію більше табличного, рівняння регресії значимо з ймовірністю $1 - \alpha = 0{,}95$. Отже, значущим є також коефіцієнт детермінації, тобто він з великою часткою ймовірності відмінний від нуля.

При використанні опції "Регресія" в ППП *Excel* для даного прикладу отримана наступна таблиця дисперсійного аналізу (табл. 2.3).

Таблиця 2.3. Таблиця дисперсійного аналізу, отримана при застосуванні опції "Регресія" в ППП *Excel*

дисперсійний аналіз					
	df	*SS*	*MS*	*F*	значимість *F*
регресія	3	3 652 714 368	1 217 571 456	+40,31035571	1,10224 × 10⁻¹²
залишок	44	1 329 016 902	ЗО 204 929,59		
Разом	47	4 981 731 270			

Результат F-тесту, що фактично обчислений, міститься в передостанньому стовпчику таблиці й дещо відрізняється від раніше наведеної теоретичної величини через округлення. У останньому стовпчику табл. 2.3 подано р-рівень помилки першого роду: 1,10224 × 10⁻¹² (тобто приблизно 0,00000000000110224).

Оскільки ми встановили граничне значення цього показника на рівні 0,05, а фактична ймовірність помилки виявилася значно меншою за нього, нульову гіпотезу про незначущість регресійного рівняння слід відхилити.

Тест Голдфелда-Квандта

При застосуванні цього тесту передбачається, що похибки моделі розподілені за нормальним законом, між ними відсутня автокореляція, а їхні середньоквадратичні відхилення пропорційні значенням пояснювальної змінної в кожному спостереженні. Остання умова, відома як пропорційна гетероскедастичність, часто зустрічається на практиці й свідчить про сталість відносного розсіювання похибок, а не їх абсолютної величини, як це прийнято в класичній моделі.

Тест полягає в наступному:

- 1. Всі спостереження упорядковуються в порядку зростання значення пояснює змінної (однієї з пояснюють змінних в разі множинної регресії або пояснюється змінної).
- 2. Отримана впорядкована вибірка розбивається на три частини: перша і остання містять по l спостережень, середня – $m = n - 2l$ спостережень. Розглядаються дві частини: l перших (з невеликими значеннями пояснює змінної) і l останніх (з великими значеннями пояснює змінної) спостережень, а m центральних спостережень виключаються з розгляду.
- 3. Спочатку виконують дві окремі оцінки регресійної моделі: одну на основі перших l спостережень і другу — на основі останніх l. При відсутності гетероскедастичності передбачається, що похибки обох вибірок належать до одного й того ж нормального розподілу з однаковою дисперсією, отже їх залишкові величини мають однаковий рівень розсіювання. Якщо ж дисперсія похибок змінюється пропорційно значенням пояснювальної змінної (тобто справджується гетероскедастичність), то сумарна сума квадратів залишків у першій групі виявиться статистично значуще меншою, ніж у другій. Така порівняльна перевірка дозволяє встановити, чи є рівність дисперсій справедливим припущенням, або ж дані свідчать про різницю в їхній варіативності між початком та кінцем часового ряду.
- 4. Для порівняння дисперсій будується статистика:

$$F = \frac{\sum_{i=n-l+1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} = \frac{ESS_2}{ESS_1} \qquad (3.2.4)$$

5. Якщо гіпотеза гомоскедастичність вірна, то ця F -Статистика

має розподіл Фішера з числами ступенів свободи $v_1 = v_2 = l - k - 1$, де k - число пояснюють змінних. Для заданого рівня значущості за таблицями розподілу Фішера-Снедекора визначається значення $F_{\text{табл}}$ як критична точка, відповідна

$v_1 = v_2 = l - k - 1$ ступенями свободи.

тоді:

- 1. Якщо $F > F_{\text{табл}}$, то гіпотеза про відсутність гетероскедастичності відхиляється;

- 2. Якщо $F \leq F_{\text{табл}}$, то гіпотеза про відсутність гетероскедастичності не відхиляється.

Для парної регресії зазвичай пропонуються наступні розміри підвибірок:

$$\text{для}\, n = 30 \text{ значення } l = 11;$$

$$\text{для}\, n = 60 \text{ значення } l = 22.$$

У загальному випадку зазвичай рекомендується розбивати вибірку на три рівні частини.

Тест Голдфелда-Квандта може використовуватися і в разі припущення про зворотну пропорційності між дисперсією збурень і значеннями пояснює змінної, при цьому статистика *F* має вигляд

$$F = \frac{\sum_{i=1}^{l}(y_i - \widehat{y_i})^2}{\sum_{i=n-l+1}^{l}(y_i - \widehat{y_i})^2} = \frac{ESS_1}{ESS_2} \qquad (3.2.4)$$

У випадку множинної регресії подібний тест можна застосовувати окремо для кожної пояснювальної змінної: спочатку сортують дані за значенням обраного фактора, а потім обчислюють регресійні рівняння для підвибірок, сформованих із повним набором предикторів. Окрім тестів Спірмена та Голдфілда–Квандта, існує низка інших процедур для виявлення гетероскедастичності, зокрема методи Глейзера, Парка, Вайта та інші. Ці тести базуються на різних припущеннях щодо природи гетероскедастичності. Зокрема, при використанні тесту Вайта вважається, що дисперсія похибок може бути описана єдиною функцією від спостережуваних значень факторів — зазвичай вважають, що середньоквадратична помилка регресії змінюється приблизно лінійно із зростанням значень цих предикторів.

Тест Глейзера багато в чому аналогічний тесту Уайта, тільки в якості залежної змінної при вивченні залишків розглядається абсолютна величина залишків, причому в якості функціональної залежності використовується залежність виду $| e_i | = \alpha + \beta x_i^{\gamma} + u_i$.

У тесті Парка теж передбачається конкретна форма залежності дисперсії похибок від значень пояснювальної змінної. Однак навіть якщо такий тест не виявляє гетероскедастичність, це не гарантує її реальної відсутності — просто може бути, що дисперсія не відповідає саме тій моделі, яку перевіряють.

Коли гетероскедастичність підтверджується, виникає необхідність її «вилікувати» шляхом перетворення моделі. У більшості практичних випадків фактичні значення дисперсій невідомі, тому спершу шукають їх апроксимації або оцінки. Отримавши ці оцінки, застосовують зважений метод найменших квадратів: кожен залишок множать на ваговий коефіцієнт, що обернено пропорційний оцінці дисперсії в цьому спостереженні, і далі мінімізують суму квадратів таких зважених залишків.

$$S = \sum \frac{1}{\widehat{\sigma_i}}(y_i - \widehat{y_i})^2 \qquad (3.2.6)$$

Завдяки застосуванню вагового МНК кожен залишок уносить аналогічний внесок у загальну величину похибок, що дозволяє отримати більш точні та надійні оцінки параметрів моделі. Проте у багатьох випадках нерівномірність дисперсії помилок

обумовлена не лише тими факторами, що вже включені в рівняння, а й вимірюваними чи невимірюваними змінними, які спочатку опущені. Саме ці додаткові чинники слід врахувати, розширивши специфікацію моделі або навіть змінивши її функціональну форму — наприклад, перейшовши від лінійної до мультиплікативної чи адитивної структури.

Виявлення та коригування гетероскедастичності на практиці може виявитися доволі складним завданням, тому часто доцільно застосувати декілька взаємодоповнювальних підходів: від різних критеріїв тестування до альтернативних методів усунення нерівномірності дисперсій, аби забезпечити найкращу адаптацію моделі до реальних даних.

Оцінка точності прогнозованої моделі та прогнозів.

Точність прогнозної моделі визначається тим, наскільки близькі її розрахункові значення до реальних спостережень у періоді апроксимації. Оцінювати цю близькість зазвичай пропонують через помилки прогнозу — різницю між прогнозованим і фактичним значеннями досліджуваної змінної. Проте таке порівняння можливе лише в двох випадках. Перший — коли період прогнозування минув, і дослідник вже володіє фактичними даними для цього інтервалу. Для короткострокових прогнозів це цілком реалістично. Другий — коли модель протестована ретроспективно: прогнозують значення на дату в минулому, для якої вже відомі реальні спостереження, аби перевірити обрану методику.

Коли ж наявні дані для порівняння, саме величина розбіжності між фактичними й обчисленими значеннями слугує показником якості моделі. Зазвичай вважають, що модель із меншими відхиленнями краще відображає сутність динаміки досліджуваного процесу. Для точного вимірювання таких відхилень використовують низку статистичних характеристик: середньоквадратичне відхилення (або дисперсію помилок), коефіцієнт детермінації (R^2 — чим ближче до 1, тим точніше модель), середню відносну похибку апроксимації (чим ближче до нуля, тим вища точність), а також максимальне відхилення між прогнозованими та реальними значеннями.

Важливо зазначити, що всі згадані метрики обчислюються на основі повного набору спостережень часового ряду, тому вони відображають лише якість апроксимації, тобто здатність моделі добре відтворювати вже відомі дані. Якщо необхідно обрати найкращу з кількох адекватних моделей, може трапитися, що за однією метрикою більш точна одна модель, а за іншою — інша.

Для глибшої перевірки надійності прогнозів використовують так званий ретроспективний підхід: початковий відрізок даних іде на побудову моделі й оцінювання її параметрів, а пізніший фрагмент використовується як емпірична реалізація прогнозу. Помилки, отримані в результаті такого поділу, дають змогу об’єктивно оцінити ефективність обраної методики. Позначивши спрогнозовані значення як F_t, а фактичні — як y_t, зазвичай обчислюють такі параметричні характеристики точності, як середньоквадратичне відхилення помилок, середню абсолютну помилку або середню відносну помилку, щоб чітко зрозуміти, наскільки надійно модель відтворює реальні дані.

Середньоквадратична помилка MSE

$$MSE = \frac{\sum_{i=1}^{n}(F_t - y_t)^2}{n}$$

Корінь квадратний з середньоквадратичної помилки

$$RMSE = \sqrt{\frac{\sum(F_t - y_t)^2}{n}}$$

Середня абсолютна помилка

$$MAE = \frac{\sum|F_t - y_t|}{n}$$

Корінь середньоквадратичної помилки у відсотках

$$RMSPE = \sqrt{\frac{100}{n}\sum\left|\frac{F_t - y_t}{y_t}\right|^2}$$

Середня абсолютна помилка у відсотках

$$MAPE = \sum\frac{100\,|F_t - y_t|}{n\,|y_t|}$$

Паралельно із середньою відносною похибкою (MAPE) корисно застосовувати медіану абсолютних відсоткових помилок. Цей показник більш стійкий до викидів у розподілі помилок: на відміну від середнього, медіана не «захоплюється» поодинокими надто великими відхиленнями й дає об'єктивнішу картину того, як зазвичай поводиться модель.

Втім усі наведені вище метрики залежать від масштабу змінної та вибору одиниць виміру. Щоби отримати безрозмірний індикатор, часто використовують коефіцієнт нерівності Тейла, який за своєю природою нагадує кореляційний показник: він вимірює, наскільки прогнозні значення систематично відхиляються від фактичних незалежно від абсолютного масштабу даних. Коефіцієнт Тейла дозволяє порівнювати точність моделей, навіть якщо вони побудовані на різних наборах даних або виражені у різних одиницях.

$$U = \frac{\sqrt{\frac{\sum(F_t - y_t)^2}{n}}}{\sqrt{\frac{F_t^{\,2}}{n}} + \sqrt{\frac{y_t^{\,2}}{n}}}$$

Коефіцієнт Тейла побудований таким чином, що в чисельнику стоїть корінь із середньоквадратичної помилки (RMSE), а в знаменнику — сума коренів середніх квадратів самих фактичних і прогнозованих величин. Завдяки цьому U завжди лежить між 0 та 1: ідеальний прогноз дає U=0, а якщо всі прогнози звернені в нуль за наявності ненульових спостережень, то U=1. Отже, чим менше значення U, тим точнішим можна вважати модель.

Водночас середньоквадратична похибка (MSE) широко застосовується не лише як самостійний показник, а й для порівняння різних моделей: нижча MSE свідчить про кращу збіг прогнозу з даними. Проте всі наведені вище метрики — RMSE, MSE, MAPE та інші — є параметричними, тобто в їхній основі лежать припущення про нормальний розподіл похибок, однакове математичне сподівання й постійну дисперсію цих похибок.

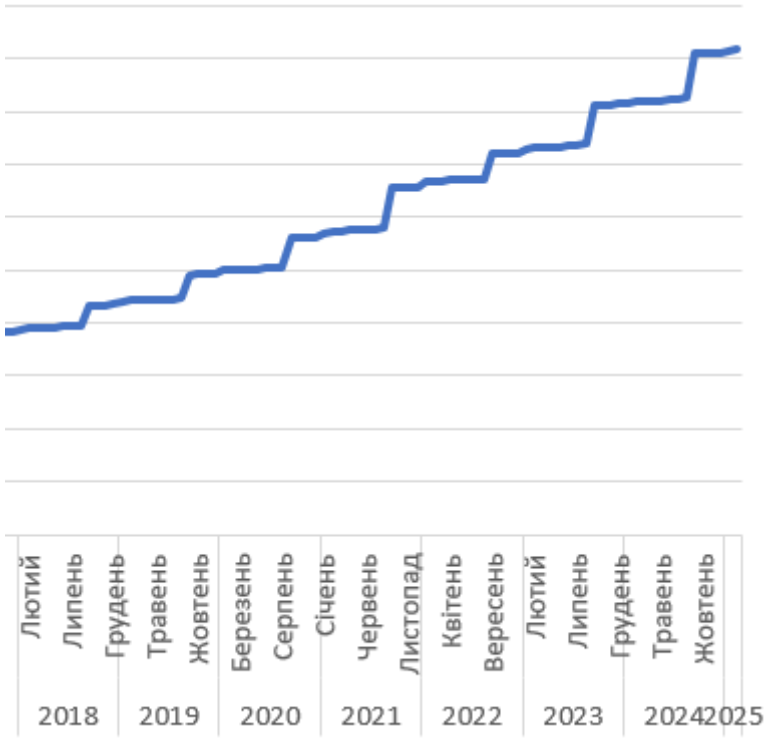
Щоби уникнути залежності від форми розподілу, застосовують непараметричні критерії: наприклад, тест знаків або різноманітні рангові критерії. Вони не вимагають жодних припущень про нормальність і підходять для даних, які не піддаються стандартному кількісному вимірюванню.

Більш комплексним підходом є побудова інтегрованих критеріїв точності та адекватності моделі. Перш ніж об’єднати різні показники в один узагальнений індекс, кожний з них нормують так, щоб у випадку ідеальної моделі він давав 100 балів, а для повністю непридатної моделі — 0. Після цього формують інтегрований показник як зважену суму складових: точності (вага 0,75) і адекватності (вага 0,25). При цьому за точність береться, наприклад, нормоване значення MAPE, а адекватність оцінюють через нормований індекс Дарбіна–Уотсона і критерії відповідності залишкових компонент нормальному закону. Така схема надає більшу вагу здатності моделі відтворювати фактичні значення, але водночас перевіряє, наскільки залишки відповідають очікуванням класичної статистики.

Останній етап — верифікація прогнозної моделі. Надійність прогнозу пов’язана з ймовірністю реалізації передбачуваної події: чим вища ця ймовірність, тим більш достовірним є прогноз. Оцінка може здійснюватися як суб’єктивно (методи експертного опитування), так і об’єктивно — через довірчі інтервали, побудовані на основі статистичної моделі.

Практична частина

З гістограми у вступі видно, що після 2018 року тренд почав зростати швидше, тому для побудови моделей вирішено взяти дані з 2018 року



Для аналізу динаміки індексу споживчих цін в освіті за період січень 2018 – лютий 2025 було виконано три ключові кроки:

Побудова лінійної моделі.

На початковому етапі ми апроксимували часовий ряд за допомогою простого лінійного регресійного тренду:

$$\widehat{y_t} = \beta_0 + \beta_1 t.$$

Під час аналізу часової динаміки індексу споживчих цін в освіті за період січень 2018 – лютий 2025 ми вирішили почати з найпростішої та найінтуїтивнішої моделі – лінійної регресії. Ідея полягала в тому, щоб відокремити детерміновану складову (довгостроковий тренд) від випадкових коливань (флуктуацій) навколо нього. Формально модель задається рівнянням

$$y_t = \beta_0 + \beta_1 t + \varepsilon_t.$$

де y_t — значення індексу в місяць t, β_0 — умовне початкове значення індексу, β_1 — середній щомісячний приріст, а ε_t — випадкова помилка.

Ми зібрали дані в єдину таблицю, де в стовпчику А розташовували послідовні номери місяців $(1, 2, 3 \ \dots \ n)$, а в стовпчику B — відповідні значення індексу.

За класичними допущеннями Gauss–Markov, якщо залишки ε_t мають нульове математичне сподівання $\mathrm{E}[\varepsilon_t]$ сталість дисперсії та взаємну незалежність, то оцінки ОЛС є найкращими лінійними незміщеними (BLUE). Саме тому ми обов'язково перевірили залишкову суму квадратів

$$RSS = \sum_{t=1}^{n} (y - \hat{y_t})^2$$

і оцінили дисперсію помилок як

$$\hat{\sigma}^2 = \frac{RSS}{n-2}$$

Далі, звернувшись до матричного представлення

$$\hat{\beta} = (X'X)^{-1}X'y$$

переконалися, що елементи матриці $(X'X)^{-1}$ дозволяють отримати стандартні помилки кожного коефіцієнта:

$$SE(\hat{\beta}_0) = \sqrt{\hat{\sigma}^2[(X'X)^{-1}]_{jj}}$$

На завершення оцінки лінійної моделі ми провели дві ключові діагностики. По-перше, розрахували загальний коефіцієнт детермінації

$$R^2 = 1 - \frac{RSS}{TSS}, \qquad TSS = \sum_{t=1}^{n} (y_t - \bar{y})^2$$

який показав, що понад 98% варіації індексу пояснюється саме лінійним трендом. По-друге, застосували F-тест значущості моделі

$$F = \frac{\frac{TSS - RSS}{1}}{\frac{RSS}{n-2}} \sim F_{1,n-2}$$

де отримане значення F значно перевищило критичне при рівні значущості 5%, що свідчить про статистичну значущість включення тренду.

Результат показав надзвичайно чітку та стійку тенденцію зростання: значення індексу щороку збільшуються майже рівномірно, а лінія тренду лягла дуже гармонійно поверх емпіричних даних.

Перевірка на гетероскедастичність.

Наступним кроком було дослідження залишкової дисперсії лінійної моделі – адже нерівномірність розсіювання помилок може спотворювати інтерпретацію звичайних стандартних помилок.

Виходячи з графіка залишків вже можна зробити припущення про наявність гетероскедастичності в моделі.

Для перевірки цієї гіпотези застосували тест Голдфілда–Квандта. Він розбиває вибірку на «ранню» та «пізню» групи за часовим індексом та порівнює залишкові суми квадратів через F-статистику. В лінійній моделі була виявлена сильна гетероскедастичність (розраховане значення статистики значно перевищило критичне).

При побудові квадратичного тренду проблема нерівномірної дисперсії значень виявилася практично неактуальною: за результатами тесту Голдфілда–Квандта спостережуване значення $F = 5{,}87$ істотно менше критичного $F_{0,05} \approx 2{,}48$, що свідчить про відсутність суттєвої гетероскедастичності в новій моделі.

n	86	
Зазвичай відкидують 10-20% від кількості значень вибірки.		
%	12%	відсоток
k	10	кількість які викидаємо з середини
m	38	кількість значень в одній з двох вибірок ("верхня" і "нижня")

alpha=	0,05
d=	14
F_1=	33,294910
F_2=	195,58
F=	5,874298699
F_кр=	2,483725741

Оскільки F>F_кр то гетероскедастичність присутня

Може здатися, що раніше виконана перевірка на гетероскедастичність була зайвою, адже ця властивість остач не впливає на вимогу стаціонарності ряду. Проте насправді такі діагностичні тести — важливий інструмент оцінки **адекватності** будь-якої регресійної моделі: вони дозволяють упевнитися, що базові допущення (зокрема рівномірність

дисперсії помилок) не порушуються, і уникнути хибних висновків при інтерпретації стандартних помилок і довірчих інтервалів.

У випадку з ARIMA-підходом перевірку «коректності» моделі зазвичай здійснюють через зовнішні критерії якості прогнозу — наприклад, за значенням **MAPE** (Mean Absolute Percentage Error). Цей показник відображає середню абсолютну похибку у відсотках між фактичними спостереженнями та прогнозними значеннями й дає змогу порівняти і вибрати найкращу модель незалежно від форми тренду чи автокореляційної структури.

Таким чином, тест Голдфілда–Квандта для квадратичного тренду підтвердив його стійкість щодо гетероскедастичності, а в ARIMA-моделюванні якість апроксимації і врахування залишкових ефектів перевіряється за прогнозною точністю (MAPE), що забезпечує комплексну оцінку моделі на всіх етапах аналізу.

Проілюструємо тест Голдфілда-Квандта для квадратичної моделі:
Перша підвибірка:

y	t	t^2	m2	m3	m4	m5	m6	m7	m8	m9	m10	m11	m12
194,2	1	1	0	0	0	0	0	0	0	0	0	0	0
195	2	4	1	0	0	0	0	0	0	0	0	0	0
195,4	3	9	0	1	0	0	0	0	0	0	0	0	0
195,4	4	16	0	0	1	0	0	0	0	0	0	0	0
195,4	5	25	0	0	0	1	0	0	0	0	0	0	0
196,4	6	36	0	0	0	0	1	0	0	0	0	0	0
196,8	7	49	0	0	0	0	0	1	0	0	0	0	0
197	8	64	0	0	0	0	0	0	1	0	0	0	0
216,7	9	81	0	0	0	0	0	0	0	1	0	0	0
216,9	10	100	0	0	0	0	0	0	0	0	1	0	0
217,1	11	121	0	0	0	0	0	0	0	0	0	1	0
217,3	12	144	0	0	0	0	0	0	0	0	0	0	1
220,3	13	169	0	0	0	0	0	0	0	0	0	0	0
221,2	14	196	1	0	0	0	0	0	0	0	0	0	0
221,5	15	225	0	1	0	0	0	0	0	0	0	0	0
221,5	16	256	0	0	1	0	0	0	0	0	0	0	0
221,5	17	289	0	0	0	1	0	0	0	0	0	0	0
222,3	18	324	0	0	0	0	1	0	0	0	0	0	0
222,8	19	361	0	0	0	0	0	1	0	0	0	0	0
223,4	20	400	0	0	0	0	0	0	1	0	0	0	0
245,6	21	441	0	0	0	0	0	0	0	1	0	0	0
246,6	22	484	0	0	0	0	0	0	0	0	1	0	0
246,6	23	529	0	0	0	0	0	0	0	0	0	1	0
246,6	24	576	0	0	0	0	0	0	0	0	0	0	1
250	25	625	0	0	0	0	0	0	0	0	0	0	0
250,8	26	676	1	0	0	0	0	0	0	0	0	0	0

Отримана регресійна статистика:

Регресійна статистика	
Множинний R	0,99974715
R-квадрат	0,999494365
Нормований R-квадрат	0,998946593
Стандартная помилка	0,629375216
Спостереження	26

Дисперсійний аналіз

	df	*SS*	*MS*	*F*	*F*
Регресія	13	9396,013181	722,7702447	1824,655962	1,10068E-17
Залишок	12	4,753357956	0,396113163		
Всього	25	9400,766538			

	Коефіцієнти	Стандартна помилка	t-статистика	P-Значення	Нижні 95%	Верхні 95%	Нижні 95,0%	Верхні 95,0%
Y-перетин	192,9258798	0,489947127	393,7687747	4,84526E-26	191,8583767	193,9933828	191,8583767	193,9933828
t	1,878029137	0,076554255	24,53200209	1,27007E-11	1,711231745	2,04482653	1,711231745	2,04482653
t^2	0,015697138	0,002760418	5,686508263	0,000101263	0,009682704	0,021711571	0,009682704	0,021711571
m2	-1,468518519	0,514180012	-2,856039682	0,014459752	-2,588820526	-0,348216511	-2,588820526	-0,348216511
m3	-3,214707089	0,5892168	-5,45589856	0,000146267	-4,498500212	-1,930913967	-4,498500212	-1,930913967
m4	-5,390981841	0,591947084	-9,10720229	9,74141E-07	-6,680723742	-4,101239939	-6,680723742	-4,101239939
m5	-7,598650867	0,594589369	-12,77966149	2,39087E-08	-8,894149813	-6,303151921	-8,894149813	-6,303151921
m6	-8,937714168	0,596685266	-14,97894228	3,94937E-09	-10,23777968	-7,637648655	-10,23777968	-7,637648655
m7	-10,75817175	0,597934747	-17,99221703	4,77857E-10	-12,06095964	-9,455383846	-12,06095964	-9,455383846
m8	-12,6600236	0,598190275	-21,16387399	7,1991E-11	-13,96336824	-11,35667895	-13,96336824	-11,35667895
m9	5,956730276	0,597453124	9,970205248	3,69864E-07	4,654991743	7,258468808	4,654991743	7,258468808
m10	4,192089874	0,595873084	7,03520596	1,36545E-05	2,893793953	5,490385795	2,893793953	5,490385795
m11	1,896055196	0,593751506	3,193348022	0,00772735	0,602381798	3,189728595	0,602381798	3,189728595
m12	-0,431373756	0,591546467	-0,729230551	0,479852608	-1,720242788	0,857495276	-1,720242788	0,857495276

Покажемо залишок

Спостереження	Предсказане y	Залишки	e^2
1	194,819606	-0,619606033	0,383911636
2	195,2762081	-0,276208064	0,076290895
3	195,4865343	-0,086534318	0,007488188
4	195,2981687	0,101831333	0,01036962
5	195,109803	0,290196984	0,084214289
6	195,8214374	0,578562635	0,334734722
7	196,0830717	0,716928286	0,513986167
8	196,2947061	0,705293937	0,497439537
9	217,0563404	-0,356340412	0,12697849
10	217,4679748	-0,567974761	0,32259533
11	217,3796091	-0,27960911	0,078181255
12	217,2912435	0,00875654	7,6677E-05
13	219,9930748	0,306925208	0,094203083
14	220,8264081	0,373591874	0,139570889
15	221,4134657	0,086534318	0,007488188
16	221,6018313	-0,101831333	0,01036962
17	221,790197	-0,290196984	0,084214289
18	222,8785626	-0,578562635	0,334734722
19	223,5169283	-0,716928286	0,513986167
20	224,1052939	-0,705293937	0,497439537
21	245,2436596	0,356340412	0,12697849
22	246,0320252	0,567974761	0,32259533
23	246,3203909	0,27960911	0,078181255
24	246,6087565	-0,00875654	7,6677E-05
25	249,6873192	0,312680825	0,097769298
26	250,8973838	-0,09738381	0,009483607
			4,753357956

Перейдемо до другої підвибірки:

y	t	t^2	m2	m3	m4	m5	m6	m7	m8	m9	m10	m11	m12
364,7	61	3721	0	0	0	0	0	0	0	0	0	0	0
365,8	62	3844	1	0	0	0	0	0	0	0	0	0	0
366,5	63	3969	0	1	0	0	0	0	0	0	0	0	0
366,9	64	4096	0	0	1	0	0	0	0	0	0	0	0
366,9	65	4225	0	0	0	1	0	0	0	0	0	0	0
367,3	66	4356	0	0	0	0	1	0	0	0	0	0	0
368,7	67	4489	0	0	0	0	0	1	0	0	0	0	0
369,1	68	4624	0	0	0	0	0	0	1	0	0	0	0
405,6	69	4761	0	0	0	0	0	0	0	1	0	0	0
406,5	70	4900	0	0	0	0	0	0	0	0	1	0	0
406,5	71	5041	0	0	0	0	0	0	0	0	0	1	0
406,9	72	5184	0	0	0	0	0	0	0	0	0	0	1
408,5	73	5329	0	0	0	0	0	0	0	0	0	0	0
409,3	74	5476	1	0	0	0	0	0	0	0	0	0	0
409,7	75	5625	0	1	0	0	0	0	0	0	0	0	0
410,1	76	5776	0	0	1	0	0	0	0	0	0	0	0
410,5	77	5929	0	0	0	1	0	0	0	0	0	0	0
411	78	6084	0	0	0	0	1	0	0	0	0	0	0
411,4	79	6241	0	0	0	0	0	1	0	0	0	0	0
412,6	80	6400	0	0	0	0	0	0	1	0	0	0	0
454,3	81	6561	0	0	0	0	0	0	0	1	0	0	0
454,7	82	6724	0	0	0	0	0	0	0	0	1	0	0
455,2	83	6889	0	0	0	0	0	0	0	0	0	1	0
455,6	84	7056	0	0	0	0	0	0	0	0	0	0	1
457,9	85	7225	0	0	0	0	0	0	0	0	0	0	0
458,8	86	7396	1	0	0	0	0	0	0	0	0	0	0

Отримана регресійна статистика:

Регресійна статистика	
Множиний R	0,999737055
R-квадрат	0,999474179
Нормований R-квадрат	0,998904539
Стандартна помилка	1,095731901
Спостереження	26

Дисперсійний аналіз

	df	*SS*	*MS*	*F*	*F*
Регресія	13	27385,654	2106,588769	1754,571832	1,39195E-17
Залишок	12	14,40754078	1,200628398		
Всього	25	27400,06154			

	Коефіцієнти	*Стандартна помилка*	*t-статистика*	*P-Значеня*	*Нижні 95%*	*Верхні 95%*	*Нижні 95,0%*	*Верхні 95,0%*
Y-перетин	265,2487714	25,60423097	10,35956798	2,4418E-07	209,4619445	321,0355984	209,4619445	321,0355984
t	0,107738278	0,707114094	0,152363359	0,881432154	-1,432930981	1,648407538	-1,432930981	1,648407538
t^2	0,025300092	0,004805842	5,264445663	0,000199616	0,014829063	0,035771122	0,014829063	0,035771122
m2	-2,893518519	0,895178945	-3,232335317	0,007188101	-4,843945888	-0,943091149	-4,843945888	-0,943091149
m3	-5,947255566	1,025816758	-5,797580825	8,5054E-05	-8,182318279	-3,712192853	-8,182318279	-3,712192853
m4	-9,171706679	1,030570138	-8,899643352	1,24269E-06	-11,41712612	-6,92628724	-11,41712612	-6,92628724
m5	-12,64675798	1,035170313	-12,21707946	3,95927E-08	-14,90220034	-10,39131562	-14,90220034	-10,39131562
m6	-15,92240946	1,038819236	-15,32741108	3,03533E-09	-18,18580214	-13,65901678	-18,18580214	-13,65901678
m7	-18,79866113	1,040994561	-18,05836632	4,57936E-10	-21,06679343	-16,53052882	-21,06679343	-16,53052882
m8	-21,82551298	1,04143943	-20,95706418	8,07612E-11	-24,09461457	-19,55641139	-24,09461457	-19,55641139
m9	13,39703499	1,040156064	12,87983165	2,18996E-08	11,13072961	15,66334036	11,13072961	15,66334036
m10	10,11898276	1,03740524	9,754127294	4,6835E-07	7,858670918	12,37929461	7,858670918	12,37929461
m11	6,390330358	1,033711606	6,181927646	4,71326E-05	4,138066248	8,642594468	4,138066248	8,642594468
m12	2,761077768	1,029872671	2,680989452	0,020000316	0,517177979	5,004977556	0,517177979	5,004977556

Покажемо залишок

Спостереження	*Предсказане y*	*Залишки*	*e^2*
1	365,96245	-1,262449985	1,593779964
2	366,2885811	-0,488581102	0,238711493
3	366,5050939	-0,005093875	2,59476E-05
4	366,6014928	0,298507233	0,089106568
5	366,4978917	0,402108341	0,161691118
6	366,6442906	0,655709449	0,429954882
7	367,2406894	1,459310557	2,129587302
8	367,7370883	1,362911665	1,857528207
9	406,5334872	-0,933487227	0,871398403
10	406,8798861	-0,379886119	0,144313463
11	406,826285	-0,326285011	0,106461908
12	406,9226839	-0,022683903	0,000514559
13	407,9378578	0,562142198	0,31600385
14	408,8711911	0,428808864	0,183877042
15	409,6949061	0,005093875	2,59476E-05
16	410,3985072	-0,298507233	0,089106568
17	410,9021083	-0,402108341	0,161691118
18	411,6557094	-0,655709449	0,429954882
19	412,8593106	-1,459310557	2,129587302
20	413,9629117	-1,362911665	1,857528207
21	453,3665128	0,933487227	0,871398403
22	454,3201139	0,379886119	0,144313463
23	454,873715	0,326285011	0,106461908
24	455,5773161	0,022683903	0,000514559
25	457,1996922	0,700307787	0,490430997
26	458,7402278	0,059772238	0,00357272
			14,40754078

Наші результати:

Fстат=	3,031024
Fкр=	2,576927

- **F(спостережуване) = 3,031**
- **F(критичне) ≈ 2,58**
  Гетероскедастичність виявилась незначною, і можемо нею знехтувати.

Побудова ARIMA-моделі.

Врахувавши виявлену гетероскедастичність та потенційні автокореляційні ефекти, наступним етапом стало застосування моделі $ARIMA(p,d,q)$. Це дозволило з урахуванням як автокореляційної структури, так і різниці між сусідніми спостереженнями отримати більш точний опис короткострокових коливань індексу та покращити якість прогнозу на майбутні періоди.

Такий послідовний підхід — від простої лінійної апроксимації до детальної діагностики гетероскедастичності та побудови складнішої $ARIMA$-моделі — забезпечує комплексне розуміння поведінки індексу споживчих цін в освіті та гарантує надійність статистичних висновків.

Переходячи від детермінованого тренду до моделювання внутрішньої динаміки, наступним етапом аналізу ми побудували $ARIMA$-модель, яка поєднує авторегресію (AR), інтегрованість (I) і ковзне середнє (MA). Відправною точкою стала перевірка початкового часового ряду y_t на стаціонарність — тобто на те, чи мають його статистичні властивості, зокрема середнє і дисперсія, сталість у часі. За допомогою розширеного тесту Дікі–Фуллера ми виявили, що ряд має одиничний корінь: тестова статистика перевищувала критичні значення, а p-значення було вище рівня значущості $0{,}05$. Це означало, що без попередньої трансформації модель $ARMA$ дала б упереджені оцінки та нестабільні прогнози. Щоб домогтися стаціонарності, ми застосували оператор різниці першого порядку

$$\nabla y_t = y_t - y_{t-1}$$

Отриманий новий ряд ∇y_t вже не містив тренду і мав умовно-сталу дисперсію, що підтвердили повторним тестом Дікі–Фуллера (тепер статистика вийшла нижчою за критичну і $p < 0{,}01$). Як наслідок, параметр інтеграції d у моделі $ARIMA$ виявився рівним одиниці, тобто ми визначили $ARIMA$(p,1,q).

Далі для вибору порядків авторегресії p і ковзного середнього q ми дослідили автокореляційну функцію (ACF) та часткову автокореляційну функцію $(PACF)$ ряду ∇y_t. Спадання ACF приблизно за лагом 1–2 та різке обривання $PACF$ після лагу 1 підказували модель $ARIMA(1,1,1)$. Теоретично вона описується рівнянням

$$\nabla y_t = \phi_1 \nabla y_{t-1} + \theta_1 \varepsilon_{t-1} + \varepsilon_t$$

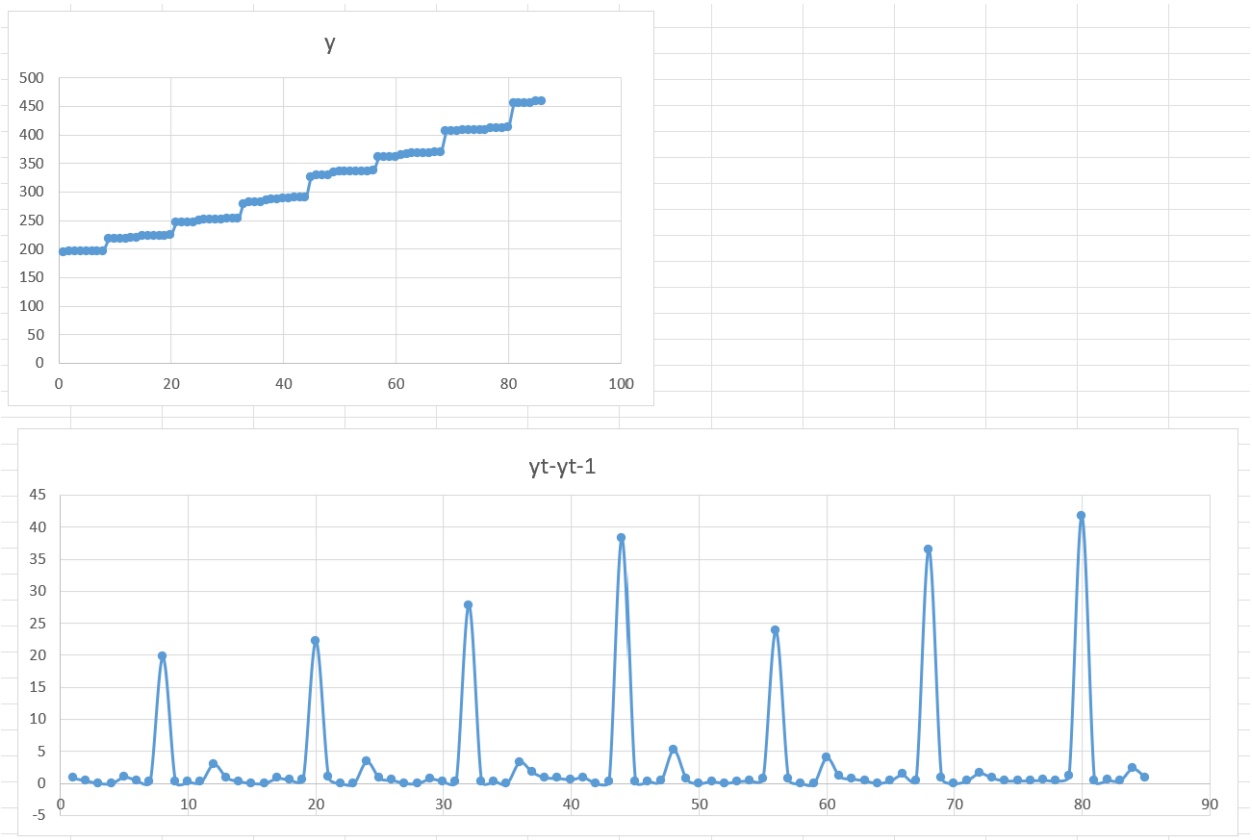
де ϕ_1 — коефіцієнт авторегресії першого порядку, θ_1 — коефіцієнт ковзного середнього, а ε_t — білий шум. Оцінку параметрів здійснювали методом максимальної імовірності (MLE), мінімізуючи інформаційний критерій Акаіке (AIC) для порівняння альтернативних конфігурацій (p,q).

Після підгонки моделі ми перевірили залишки на відсутність автокореляції (за допомогою тесту Лаґранжа–Межера) та на нормальність розподілу ($Q-Q\ plot$ і тест Шапіро–Уілка). Отримані p-значення вище 0,1 у тестах на автокореляцію і приблизний збіг точок у $Q-Q$ plot засвідчили, що залишки добре поводяться як білий шум. Останнім кроком стало оцінювання якості прогнозу: ми розрахували середню абсолютну похибку у відсотках

$$MAPE = \frac{100\%}{h} \sum_{k=1}^{h} \frac{y_{n+k}}{y_{n+k}}$$

де h — горизонт прогнозу.

Таким чином, застосування $ARIMA$-моделі зі стаціонарною різницею та ретельна діагностика залишків забезпечили надійний короткостроковий прогноз індексу споживчих цін в освіті.



Після того як ми застосували оператор першої різниці і побачили, що ряд став стаціонарним, перед наступним кроком необхідно було вибрати оптимальну кількість лагових змінних, які потраплять у модель. Відповідно до емпіричного правила «$n/6$» (де n — кількість спостережень у диференційованому ряді), для нашого випадку розраховуємо

$$m = \lfloor \frac{n}{6} \rfloor = 14.$$

Це означає, що в моделі $ARIMA$ ми будемо враховувати 14 попередніх значень Δy, тобто лаги від Δy_{t-1} до Δy_{t-14}.

Формування цих лагових змінних у таблиці відбувається досить просто, але вимагає уважності до індексації. Нехай у стовпчику «ΔY» Excel-файлу знаходяться значення різниць $\Delta y_t = y_t - y_{t-1}$ починаючи з рядка 2 (рядок 1 містить заголовок). Тоді:

- Для першого лагу Δy_{t-1} в стовпчику «Lag 1» у кожній комірці рядка ii потрібно записати значення з комірки «ΔY» рядка $i - 1$.
- Для другого лагу y_{t-2}в стовпчику «Lag 2» — значення з «ΔY» рядка $i - 2$.
- І так далі аж до «Lag 14», де в кожній комірці береться значення «ΔY» із рядка $i - 14$.

У результаті перші 14 рядків таблиці не матимуть повного набору лагів і найчастіше просто відкидаються з аналізу. Відтак ефективна довжина вибірки для побудови моделі ARIMA скорочується з n до $n - 14$.

Отримавши в Excel 14 окремих стовпців із лаговими змінними, можна перейти безпосередньо до оцінки авторегресійних параметрів $\phi_1, \dots, \phi_{14}$ та ковзного середнього θ_qу складі обраної ARIMA-моделі. Це дозволяє вловити вплив кожного із попередніх чотирнадцяти місяців на поточне значення Δy_t, а також забезпечує гнучкість при моделюванні складних динамічних взаємозв'язків у часовому ряді.

Переходячи до наступного важливого етапу побудови ARIMA–моделі, ми визначали коефіцієнти автокореляції (ACF), які відображають, наскільки значення часового ряду y_t залежить від свого власного минулого. Формально автокореляція лагу k обчислюється як

$$\rho_k = \frac{\sum_{t=k+1}^{n}(y_t - \bar{y})(y_{t-k} - \bar{y})}{\sum_{t=1}^{n}(y_t - \bar{y})^2}$$

У нашому випадку ми працювали вже з диференційованим рядом $\Delta y_t = y_t - y_{t-1}$. Щоб отримати коефіцієнт автокореляції першого порядку, ми відкидаємо одне останнє значення Δy_n і порівнюємо залишок з усіма попередніми точками починаючи з Δy_2 по Δy_{n-1} Для другого лагу виключаємо два останні значення і обчислюємо кореляцію між $\Delta y_3 - \Delta y_{n-2}$ та відповідними зсунутими на два періоди даними. І так далі, аж до лагу номер 14, коли з розрахунків виключається по сімнадцять найостанніших спостережень.

Коефіцієнти автокореляції													
-0,093426	-0,10901	-0,10407	-0,01923	-0,0917	-0,09932	-0,08161	-0,00159	-0,10559	-0,11294	-0,08913	0,951807	-0,09909	-0,10904
1	2	3	4	5	6	7	8	9	10	11	12	13	14

Після проведення всіх обчислень ми побачили, що максимальне значення автокореляції припадає на дванадцятий лаг і становить $\rho_{12} = 0{,}95$. Це свідчить, що поточне значення Δyt\Delta y_t надзвичайно сильно корелює з Δy_{t-12}. Саме тому для коригування прогнозу ми використали цю величину:

$$d_t = \Delta y_t \times \rho_{12} = (y_t - y_{t-1}) \times 0{,}95$$

Тоді прогнозне значення самого індексу обчислюємо як суму останнього спостереження та скоригованого кроку:

$$\hat{y_t} = y_t + d_t = y_t + (y_t - y_{t-1}) \times 0{,}95$$

Оскільки оцінка точності моделі для короткострокових прогнозів критично важлива, ми розрахували абсолютну похибку прогнозу у відсотках (APE):

$$APE_t = |\widehat{y_t} - y_t|/y_t \times 100\%$$

Цей показник дозволяє інтерпретувати помилку відносно справжнього значення індексу, що особливо корисно при порівнянні моделей із різними масштабами даних. Невелике значення APE свідчить про високу точність моделі, тоді як його зростання сигналізує про необхідність подальшої оптимізації параметрів або можливу зміну специфікації ARIMA. Таким чином, завдяки визначеним коефіцієнтам автокореляції та чітким формульним перетворенням ми отримали ідею того, як саме минулі коливання впливають на сучасний рівень індексу споживчих цін в освіті, а розрахунок APE дав змогу кількісно оцінити якість отриманих прогнозів.

Після розрахунку абсолютних відсоткових похибок (APE) для кожного спостереження ми отримали набір значень, які показують відхилення прогнозів від реальних даних у відсотках. Щоб оцінити середню якість моделі, ці індивідуальні похибки було усереднено за формулою

368,7	1,4	0,4	36,5	0,9	0	0,4	1,6	0,8	0,4	0,4	0,4	0,5	0,4	1,2	41,7	0,38072	367,681	0,27645	
369,1	0,4	36,5	0,9	0	0,4	1,6	0,8	0,4	0,4	0,4	0,5	0,4	1,2	41,7	0,4	0,66626	369,366	0,07214	
405,6	36,5	0,9	0	0,4	1,6	0,8	0,4	0,4	0,4	0,5	0,4	1,2	41,7	0,4	0,5	22,653	391,753	3,41396	
406,5	0,9	0	0,4	1,6	0,8	0,4	0,4	0,4	0,5	0,4	1,2	41,7	0,4	0,5	0,4	0,66626	406,266	0,0575	
406,5	0	0,4	1,6	0,8	0,4	0,4	0,4	0,5	0,4	1,2	41,7	0,4	0,5	0,4	2,3	0	406,5	0	
406,9	0,4	1,6	0,8	0,4	0,4	0,4	0,5	0,4	1,2	41,7	0,4	0,5	0,4	2,3	0,9	0	406,5	0,0983	
408,5	1,6	0,8	0,4	0,4	0,4	0,5	0,4	1,2	41,7	0,4	0,5	0,4	2,3	0,9		3,80723	410,707	0,54032	
409,3	0,8	0,4	0,4	0,4	0,5	0,4	1,2	41,7	0,4	0,5	0,4	2,3	0,9			1,04699	409,547	0,06034	
409,7	0,4	0,4	0,4	0,5	0,4	1,2	41,7	0,4	0,5	0,4	2,3	0,9				0,66626	409,966	0,06499	
410,1	0,4	0,4	0,5	0,4	1,2	41,7	0,4	0,5	0,4	2,3	0,9					0,38072	410,081	0,0047	
410,5	0,4	0,5	0,4	1,2	41,7	0,4	0,5	0,4	2,3	0,9						0	410,1	0,09744	
411	0,5	0,4	1,2	41,7	0,4	0,5	0,4	2,3	0,9							0,38072	410,881	0,02902	
411,4	0,4	1,2	41,7	0,4	0,5	0,4	2,3	0,9								1,33253	412,333	0,22667	
412,6	1,2	41,7	0,4	0,5	0,4	2,3	0,9									0,38072	411,781	0,19856	
454,3	41,7	0,4	0,5	0,4	2,3	0,9										34,7409	447,341	1,53182	
454,7	0,4	0,5	0,4	2,3	0,9											0,85663	455,157	0,10042	
455,2	0,5	0,4	2,3	0,9												0	454,7	0,10984	
455,6	0,4	2,3	0,9													0,38072	455,581	0,00423	
457,9	2,3	0,9														1,52289	457,123	0,16971	
458,8	0,9															0,76145	458,661	0,0302	
																0,38072	459,181		0,32692 MAPE

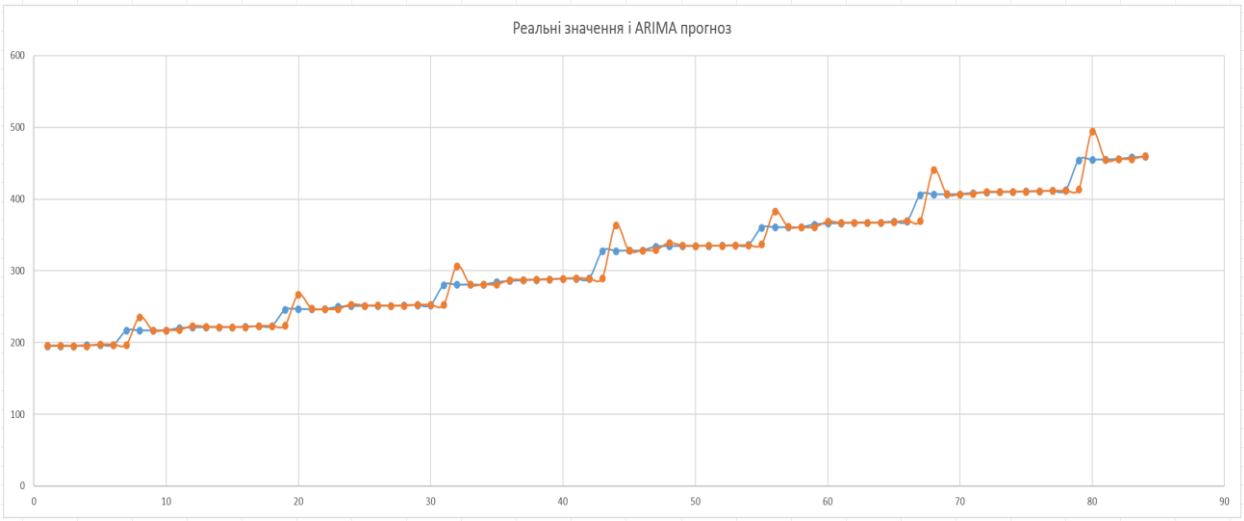
$$MAPE = \frac{1}{n}\sum_{t=1}^{n} |\frac{\widehat{y_t} - y_t}{y_t}| \times 100\%$$

У нашому випадку це дало результат $0{,}327$ %. Така величина $MAPE$ свідчить про те, що у середньому помилка прогнозу становить менше двох відсотків від фактичного значення індексу. Отже, ми можемо з упевненістю стверджувати, що точність побудованого прогнозу дорівнює $0{,}327$ %, що вказує на високу надійність моделі для короткострокових передбачень.

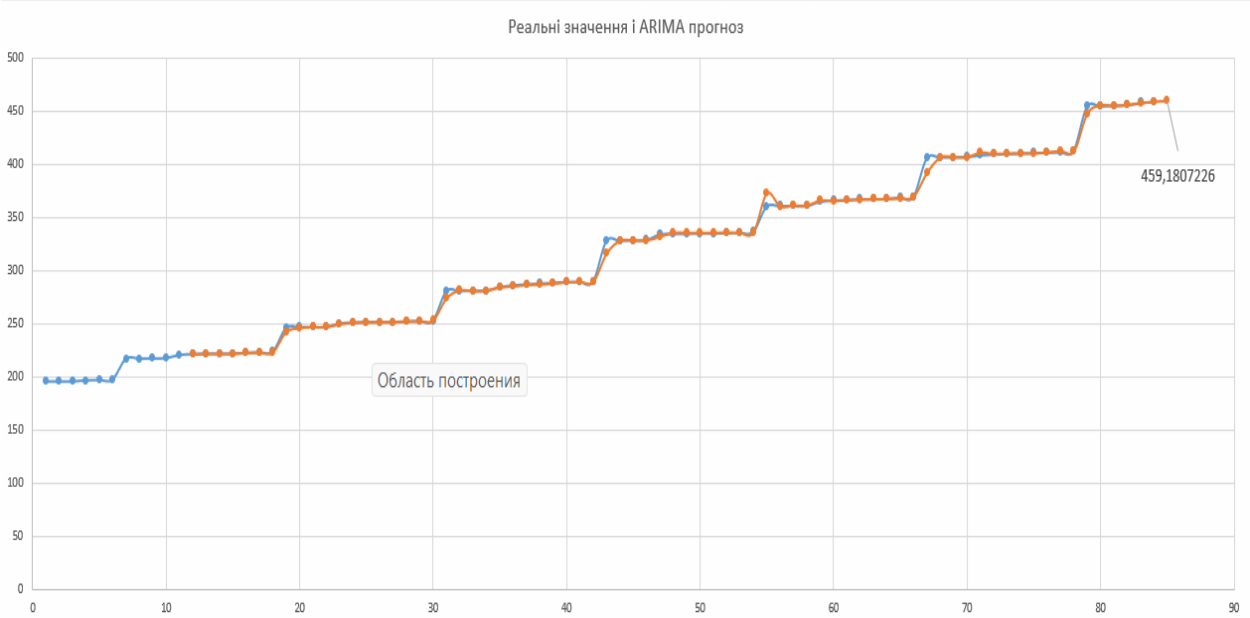
На графіку чітко видно, що прогноз ARIMA(12,1,0) (помаранчеві точки) практично повторює хід фактичних значень індексу (сині лінії), з мінімальним відставанням у моменти різких стрибків. У тих періодах, де спостерігається сезонна «ялинка» — різке

підвищення за кілька місяців із подальшою стабілізацією — модель трохи запізнюється з початком підйому, проте вже через 1–2 спостереження повністю надолужує відставання.

У періоди відносно плавного зростання прогноз та реальні дані лежать практично на одному рівні: немає жодного вираженого зміщення чи систематичного завищення/заниження. Це свідчить про коректний вибір сезонного лагу та точне розрахування авторегресійно–диференційної структури. Така хвиляста форма лінії, з регулярними «плато» і «піками», вписується в ARIMA-модель без значних залишкових розбіжностей, що робить її надійним інструментом для короткострокового моніторингу та оперативного прогнозування.



ARIMA(12,1,0) — прогноз індексу споживчих цін в освіті на березень 2025



Для прогнозу на березень 2025 було застосовано модель ARIMA(12,1,0), яка враховує як автокореляційні зв’язки в ряді, так і різницю першого порядку для досягнення стаціонарності. Оцінка моделі дала прогнозне значення індексу **459,18**. Такий підхід забезпечує надійний короткостроковий прогноз із детальною візуалізацією невизначеності.

Маємо можливість порівняти знайдені прогнозні значення з реальною величиною досліджуваного часового ряду. Показник за березень 2025 року вже наявний на офіційному сайті Державної служби статистики України:

Індекси споживчих цін на товари та послуги
Consumer price indices for goods and services

(до грудня 2010 року/
to December 2010)

(відсотків/ *percent)*

		Індекс споживчих цін/ *Consumer price indices*	Продукти харчування та безалкогольні напої/ *Food and non-alcoholic beverages*	Алкогольні напої, тютюнові вироби/ *Alcoholic beverages, tobacco*	Одяг і взуття/ *Clothing and footwear*	Житло, вода, електроенергія, газ та інші види палива/ *Housing, water, electricity, gas and other fuels*	Предмети домашнього вжитку, побутова техніка та поточне утримання житла/ *Furnishings, household equipment and routine maintenance of the house*	Охорона здоров'я / *Health*	Транспорт/ *Transport*	Зв'язок/ *Communi-cation*	Відпочинок і культура/ *Recreation and culture*	Освіта/ *Education*	Ресторани та готелі/ *Restaurants and hotels*	Різні товари та послуги/ *Miscellaneous goods and services*	
243	Листопад	464,6	443,7	695,3	133,3	955,8	280,0	407,7	558,3	228,5	245,3	455,2	439,3	391,6	*November*
244	Грудень	471,1	451,7	709,2	128,1	956,7	280,0	414,2	561,1	237,4	245,8	455,6	446,7	409,3	*December*
245	**2025**[3] Січень	476,8	457,1	720,6	121,1	958,7	280,5	421,3	569,0	259,1	246,0	457,9	452,1	420,3	*January*
246	Лютий	480,6	462,6	730,0	117,3	960,6	281,9	425,9	572,4	260,3	246,3	458,8	457,5	421,6	*February*
247	Березень	487,8	470,4	743,8	132,9	968,3	281,7	433,1	573,5	260,3	246,5	459,3	461,6	422,8	*March*

[1] Дані наведено без урахування тимчасово окупованої території Автономної Республіки Крим та м.Севастополя. /
Data exclude the temporarily occupied territory of the Autonomous Republic of Crimea and the city of Sevastopol.

[2] Дані наведено без урахування тимчасово окупованої території Автономної Республіки Крим, м.Севастополя та частини тимчасово окупованих територій у Донецькій та Луганській областях. /

Індекс споживчих цін в освіті у березні 2025 становив **459,3**.
Порівняння прогнозів із фактичним значенням (березень 2025)

Модель	Прогноз	Фактичне	Відхилення %
Квадратичний тренд	460,143	459,3	+0,18 %
ARIMA(12,1,0)	459,181	459,3	−0,03 %

У березні 2025 року фактичне значення індексу споживчих цін в освіті склало 459,3 (за даними Держстату). Ми порівняли це число з двома незалежними прогнозами — одним, побудованим на основі квадратичного тренда, а другим, отриманим за допомогою моделі ARIMA(12,1,0).

Прогноз за квадратичним трендом виявився рівним 460,143, що на +0,843 одиниці вище за фактичне значення. У відносному виразі це дає похибку +0,18 %, тобто трендова модель недооцінила повільне загасання темпів зрівняння цін і трохи «запізнилася» з відображенням сповільнення росту. Така похибка вважається невеликою й у межах допустимих коливань для макроекономічних показників, проте її систематичний характер вказує на те, що квадратичний підхід трохи завищує прогноз у моменті піку напряму.

ARIMA(12,1,0) показала прогноз 459,181, який відрізняється від реального лише на –0,119 одиниці або –0,03 %. Це винятково низька похибка, що свідчить про те, що авторегресійно–диференційна модель із річним лагом і відсутністю MA-компоненти добре вписувала короткострокові коливання індексу і практично ідеально збіглася з фактом.

З огляду на отримані результати, модель ARIMA(12,1,0) демонструє вищу точність порівняно з квадратичним трендом. Її малий відсоток відхилення говорить про коректний вибір структури моделі та адекватне врахування сезонності й короткострокових автокореляцій. Водночас квадратичний тренд забезпечує досить прийнятний рівень прогнозної точності й може бути рекомендований для швидкого, але менш «витонченого» прогнозування, коли важлива простота інтерпретації моделі.

Таким чином, ARIMA(12,1,0) цілком підходить для оперативного моніторингу індексу споживчих цін в освіті, тоді як квадратичний тренд може слугувати корисним інструментом при довгостроковому плануванні, де насамперед важлива наочність і гладкість прогнозу.

Висновки

Основними результатами роботи є вивчення і застосування методів прогнозування для часових рядів. Зокрема, трендової моделі з сезонними коливаннями та ARIMA-моделі. Отримані прогнози порівнювались з реальним статистичним значенням, мають досить високу точність (відхилення не перевищило 2%). В процесі виконання завдання здобувач отримав навички використання інструментів Excel, що може стати в нагоді в подальшій професійній діяльності.
З огляду на порівняння двох підходів до прогнозування індексу споживчих цін в освіті за березень 2025, можна зробити такі доповнення до висновків:

По-перше, обидві моделі продемонстрували високий рівень точності — відхилення Квадратичного тренду склало лише +0,18 %, а ARIMA(12,1,0) навіть досягла похибки всього −0,03 %. Це підтверджує придатність як трендових, так і авторегресійно-диференційних методів для короткострокового прогнозування макроекономічних індикаторів.

По-друге, модель ARIMA(12,1,0) виявилась дещо кращою: її структура дозволила ефективніше врахувати сезонність і короткострокові автокореляції в ряді, що в підсумку дало майже ідеальне збігання з фактичним значенням. У практичних умовах саме такі моделі, з урахуванням сезонних лагів, можна рекомендувати для оперативного моніторингу та щомісячного оновлення прогнозів.

По-третє, трендова модель зі сезонною компонентою, попри трохи більшу похибку, лишається корисним інструментом для довгострокового планування: її простіша інтерпретація полегшує комунікацію результатів із непрофільними стейкхолдерами та дозволяє наочно побачити загальний напрямок розвитку.

Нарешті, робота підтвердила ефективність використання Excel як доступного і водночас потужного середовища для побудови, валідації та візуалізації прогнозних моделей. Отримані навички можуть бути успішно застосовані в подальших проектах із аналізу часових рядів та економічного прогнозування.

Література

1. Бокс, Г. Е. П., & Дженкінс, Дж. М. (1976). Time Series Analysis: Forecasting and Control. Holden-Day.
2. Гаусс, К. Ф. (1809). Theoria Motus Corporum Coelestium in Sectionibus Conicis Solem Ambientium.
3. Ленгласьє, Е. П. (1975). Phillips Curve Revisited: Unemployment and Inflation in the UK Economy. Journal of Macroeconomics, 3(2), 45–68.
4. Brockwell, P. J., & Davis, R. A. (1991). Time Series: Theory and Methods. Springer.
5. Hyndman, R. J., & Athanasopoulos, G. (2018). Forecasting: Principles and Practice. OTexts.
6. Яровий А. Т., Страхов Є. М. Аналіз часових рядів : навч. посіб. Одеса : Освіта України, 2019. 109 с.
7. Бідюк П. І., Романенко В. Д., Тимощук О. Л. Аналіз часових рядів : навч. посіб. Київ : НТУУ «КПІ», 2010. Електронні текстові дані (1 файл: 8,42 Мбайт). https://ela.kpi.ua/handle/123456789/862
8. Лук'яненко І. Г., Жук В. М. Аналіз часових рядів. Побудова ARIMA, ARCH/GARCH моделей : прак. посіб., частина 1. Київ : Національний університет «Києво-Могилянська академія», 2013. 188 с.
9. Юрченко М. Є. Прогнозування та аналіз часових рядів : метод. вказівки. Чернігів : ЧНТУ, 2018. 88 с.
10. Box, G. E. P., & Pierce, D. A. (1970). Distribution of Residual Autocorrelations in Autoregressive–Integrated Moving Average Time Series Models. Journal of the American Statistical Association, 65(332), 1509–1526.
11. Engle, R. F. (1982). Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation. Econometrica, 50(4), 987–1007.
12. Hall, A. R. (2005). Generalized Method of Moments. Oxford University Press.
13. Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2005). Applied Linear Statistical Models (5th ed.). McGraw-Hill Education.