**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ**

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ**

**«ХАРКІВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ»**

Кафедра комп’ютерної математики і аналізу даних

**ЗВІТ**

про виконання Лабораторної роботи №4 за темою

«Аналіз часового ряду та виділення детермінованої складової»

з дисципліни «Аналіз даних і часових рядів»

Група КН-122а

Виконавець Жарський Н.Д.

Викладач Гардер С.Є.

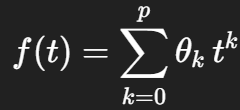
Харків 2025

**Мета роботи**

1. Навчитися виділяти та оцінювати детерміновану (трендову) складову часового ряду за допомогою поліноміальної інтерполяції МНК
2. Перевірити залишки після детрендінгу двома критеріями серій:
   1. за медіаною
   2. за висхідними/низхідними серіями
3. Провести графічну діагностику залишків (ACF, гістограма, QQ-плот)

**Теоретичні відомості**

1. **МНК-оцінка полінома**  
    Для ряду  будемо шукати



де θ знаходяться як



1. **Критерії випадковості залишків**

* **Серії за медіаною**: розбиваємо відхилення від медіани на «+»/«–» і підраховуємо загальну кількість серій R та довжину найдовшої
* **Висхідні/низхідні серії**: аналогічно, але знакова послідовність будується на основі різниць

Для обох критеріїв спостережувані Z-статистики порівнюються з *±*

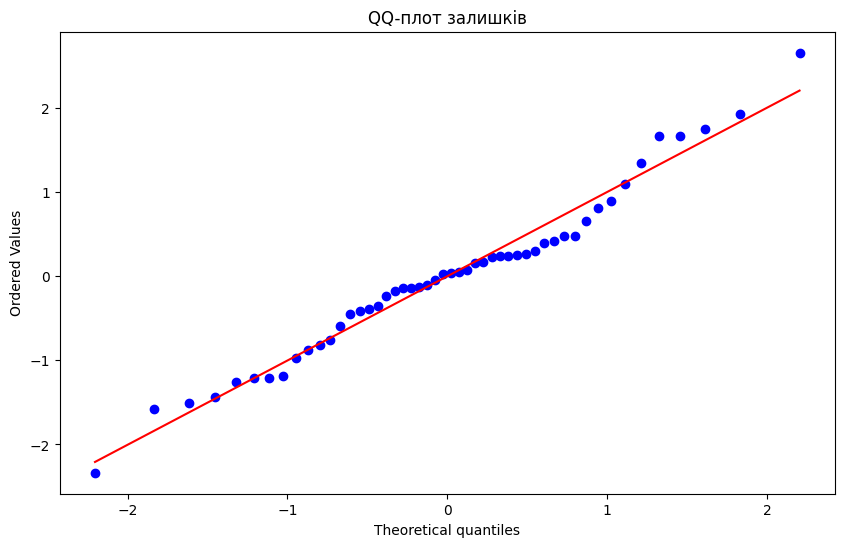
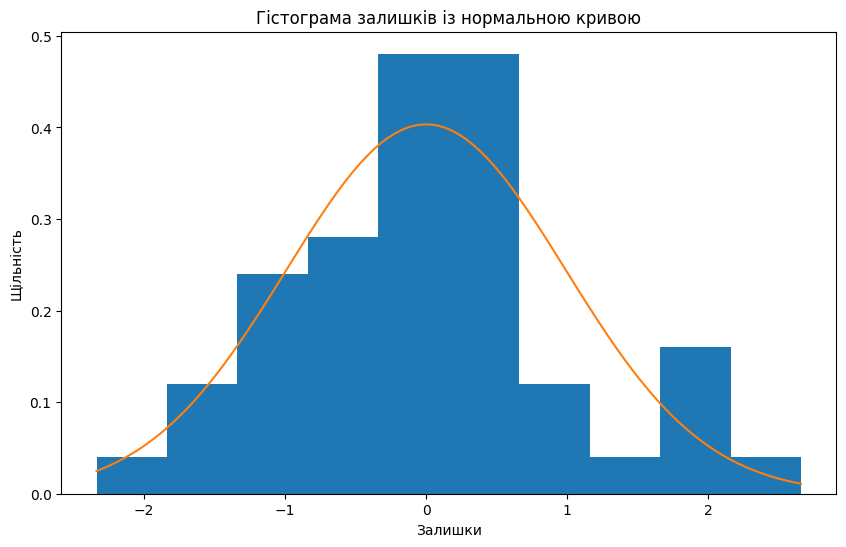
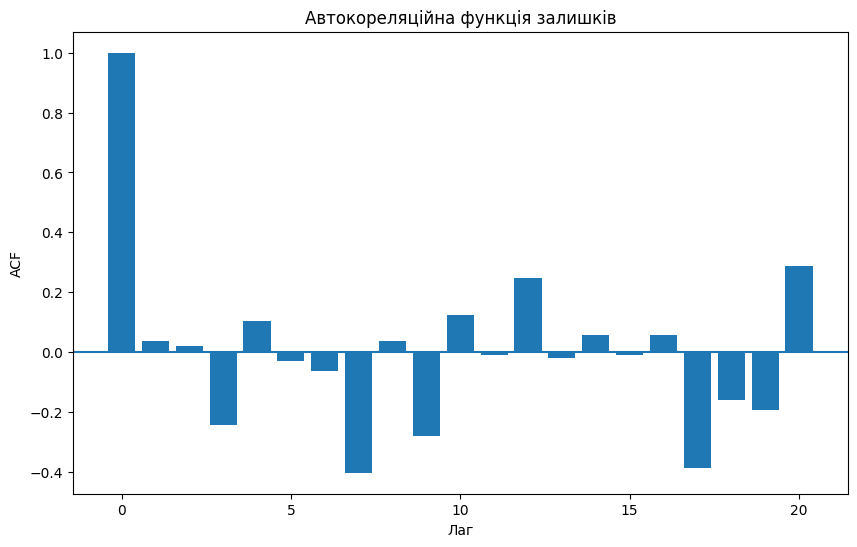
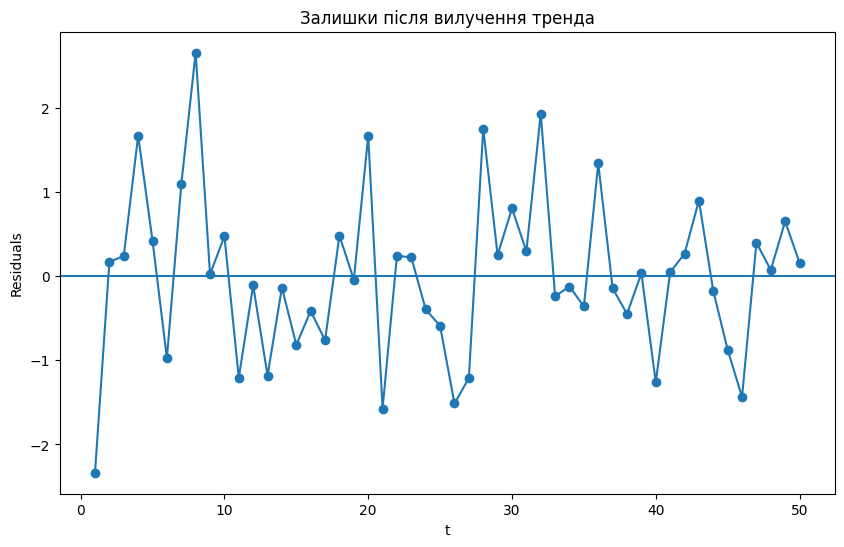
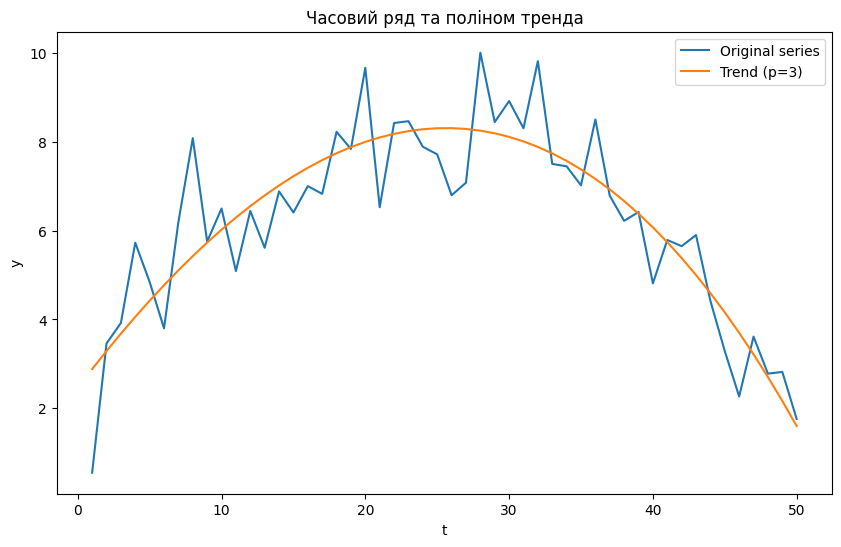
1. **Графічні методи діагностики залишків**
   * ACF (лаги 1…20)
   * Гістограма із накладеною нормальною кривою
   * QQ-плот

**Опис реалізації**

1. **Зчитування даних** із файлу ts.csv (три стовпці: t, X1 (пол.2+шум), X2 (пол.3+шум))
2. **Автоматичний підбір степеня** p∈{0,1,2,3} за критерієм автокореляції лагу 1: вибираємо найменший p, для якого ∣ρ1 ∣<1.96/
3. **Оцінка полінома тренда** МНК
4. **Виділення залишків** −f(t)
5. **Статистичні тести** на серії (медіана та напрямні серії)
6. **Графічна діагностика**:
   * рис. 1: часовий ряд + тренд;
   * рис. 2: залишки;
   * рис. 3: ACF залишків;
   * рис. 4: гістограма залишків із нормальною кривою;
   * рис. 5: QQ-плот залишків

**Результати**

* **Обраний степінь полінома**: p=3
* **Коефіцієнти тренда**: 3-й коефіцієнт насправді це 10^-17, тому він тут як 0
* **Критерій рядів за медіаною**:  
   R=22,Z=−4.021 → H0 відхиляється (занадто багато рядів)
* **Критерій висхідних/низхідних рядів**:  
   R=36,Z=1.266 → H0 не відхиляється
* **Графічний аналіз залишків**:
  + АCF показує невисоку автокореляцію для більшості лагів (рис. 3)
  + Гістограма (рис. 4) та QQ-плот (рис. 5) демонструють приблизну нормальність, хоча є відхилення в обох хвостах



**Висновки**

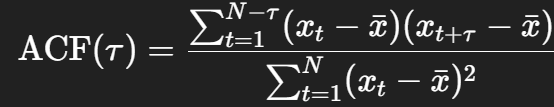
1. Поліном третього ступеня достатньо добре описує тренд, але залишається помітна детермінована складова (критерій за медіаною відхиляє H0 )
2. За критерієм висхідно/низхідних серій та графічними методами залишки можна вважати близькими до випадкових, але існує легка асиметрія
3. Для покращення детрендінгу можна розглянути:
   1. поліном вищого ступеня
   2. нелінійні (сплайнові) методи
   3. сезонне розкладання
4. **Тренд і поліном** (Рис. 1): поліном 3-го ступеня адекватно описує детерміновану складову — крива досить наближає головну форму вихідного ряду
5. **Залишки** (Рис. 2): точки залишків коливаються навколо нуля без вираженого тренда, але видно декілька послідовних “підйомів” та “спадів”, що підтверджує відхилення в тесті за медіаною
6. **ACF залишків** (Рис. 3): автокореляція на більшості лагів знаходиться близько до нуля, хоча на деяких (наприклад, лаг 1, лаг 7, лаг 20) спостерігається слабка кореляція. Загалом це свідчить про майже випадковий характер залишків
7. **Гістограма із нормальною кривою** (Рис. 4): розподіл залишків приблизно симетричний, пікова щільність близька до нормальної, проте є невеликі відхилення в обох хвостах
8. **QQ-плот залишків** (Рис. 5): більшість точок лежить поруч із теоретичною лінією нормального розподілу, але два-три відхилення у верхньому та нижньому хвостах свідчать про легку девіацію від суворої нормальності

Отже, за комбінованим аналізом тренда, статистичними критеріями та графічними діаграмами залишки можна вважати близькими до випадкових із майже нормальним розподілом, хоча часткові автокореляційні залишки й тест за медіаною вказують на не до кінця усунуту детерміновану складову. Для повного “очищення” ряду я б зробив подальше випробування вищого ступеня полінома або інших методів детрендінгу

1. Які види тимчасових рядів ви знаєте? Наведіть приклади.

* Трендові ряди (постійна або плавна зміна рівня:, напр. річна динаміка ВВП)
* Сезонні ряди (періодичні коливання: температура повітря по місяцях)
* Циклічні ряди (коливання більш довготривалі, ніж сезонні: економічні цикли)
* Стохастичні (випадкові) ряди (шум без чіткої структури: рандомні відхилення на фондовому ринку)

1. **У чому полягають характерні відмінності часових рядів від просторових вибірок?**
   * У часових рядах спостереження впорядковані за часом, між сусідніми значеннями часто існує кореляція
   * У просторових вибірках спостереження незалежні, розташовані у просторі без часової послідовності
2. Що таке автокореляція залишків?  
    Автокореляція залишків — це кореляція значення залишку із власним значенням на попередніх лагах . Висока автокореляція вказує, що модель тренда не вловила всю детерміновану складову
3. Як визначити ступінь полінома тренда?  
    Методом послідовних різниць і перевірки автокореляції лагу 1: почергово віднімаємо першу, другу … різницю, доки ACF лагу 1 не стане статистично незначущою (тобто |ρ1 ∣<1.96/N ) – тоді порядок полінома p дорівнює кількості різниць
4. Що таке автокореляційна функція?  
    Функція, що відображає кореляцію між значеннями ряду, віддаленими на лаг τ:

* 

1. Який вид має АКФ для стаціонарного та нестаціонарного рядів?
   * Стаціонарний ряд: ACF швидко спадає до нуля з ростом лаґу
   * Нестаціонарний ряд (з трендом): ACF повільно спадає або зберігає високі значення навіть на великих лаґах
2. Що означає властивість ергодичності ряду?  
    Ергодичність означає, що часові середні та дисперсії одного реалізованого ряду рівні відповідним математичним (ймовірнісним) моментам популяції: просторові та часові оцінки дають однаковий результат при N→∞