



# Лекція 6

**Особливості дослідження часових рядів. Різновиди моделей часових рядів.**

# Аналіз часового ряду (time series)

Нагадаємо приклади часових рядів з попередньої лекції:

- **Моментальний часовий ряд**

Дата надання позички	01.10.	05.10.	12.10.	23.10.	03.11.	07.11.
Розмір наданої позички, тис. гр. од.	3747	3710	3839	3783	3747	3710

- **Інтервалний часовий ряд**

Місяць	Січень	Лютий	Березень	Квітень	Травень
Валовий внутрішній продукт, млн грн	6578	7016	7353	7353	7941

- **Часовий ряд, утворений із середніх значень**

Місяць	Січень	Лютий	Березень
Середня зарплата загалом, грн/міс.	152,2	153,7	165,8

# Аналіз часового ряду (time series)

**Умови коректності застосування математичного апарату для аналізу часових рядів:**

- порівнянність;
- однорідність;
- сталість;
- достатня сукупність спостережень (репрезентативність).

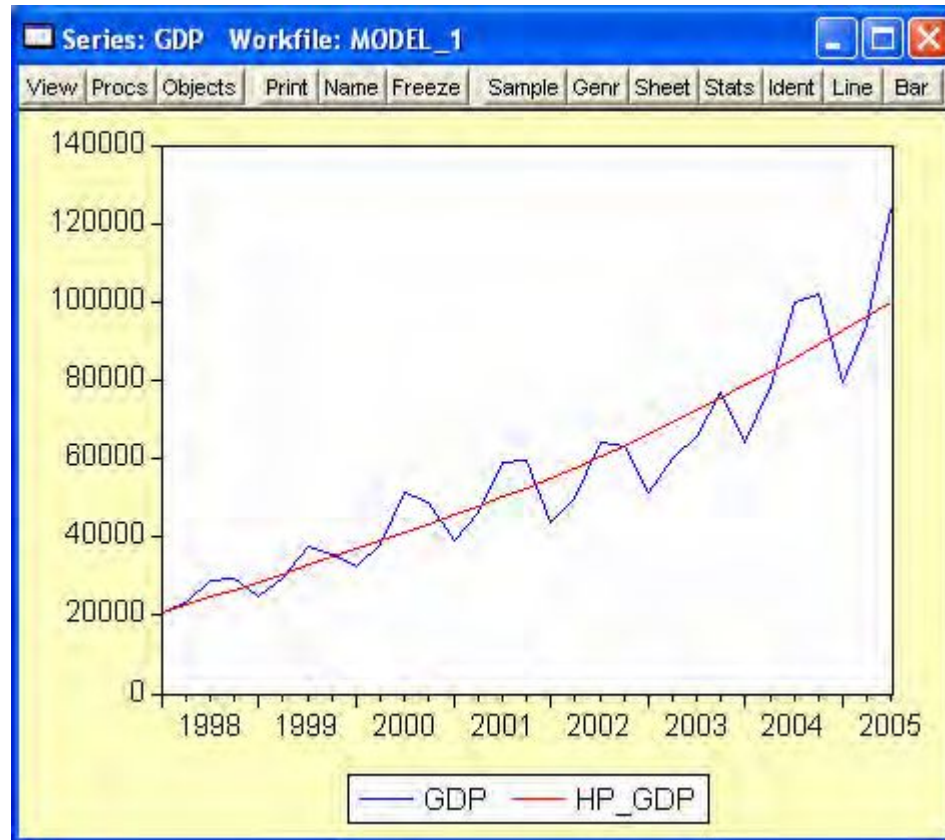
# Аналіз часового ряду (time series)

- **Мета:** побудова математичної моделі ряду, за допомогою якої можна пояснити поведінку ряду і здійснити прогноз на майбутні періоди.
- Основні **задачі** дослідження часових рядів:
  - Статистичний аналіз (визначення природи й закону розподілу ряду, вивчення статистичних параметрів ЧР);
  - Прогнозування на підставі адекватних моделей.
- **Природа ряду:**
  1. **Стаціонарні** (білий шум, випадкове блукання, марківський процес)
  2. **Нестационарні** (типів TS і DS; тренд-сезонні; нелінійні)

# Аналіз часового ряду (time series)

- **Мета:** побудова математичної моделі ряду, за допомогою якої можна пояснити поведінку ряду і здійснити прогноз на майбутні періоди.
- Основні етапи:
  - (1) Побудова і вивчення **графіка часового ряду** стосовно наявності тренду, сезонності, циклічності та аномальних даних;
  - (2) Визначення основних числових характеристик – **описової статистики** (середнє, медіана, дисперсія, стандартне відхилення, коефіцієнти асиметрії й ексцесу);
  - (3) побудова **гістограми** для перевірки динамічного ряду на відповідність закону нормального розподілу;

# Графічний аналіз ЧР



Графік 1 – Приклад ЧР із зростаючим трендом (виділений червоним) та сезонною складовою (у вигляді сезонних коливань)

# Виявлення сезонності

- У тому випадку, коли коливання показника не мають чіткої тенденції, індекси сезонності визначаються за формулою:

$$I_{\text{сез}} = \frac{\bar{y}_i}{\bar{y}_{\text{заг}}} \times 100\%,$$

- де  $\bar{y}_i$  — середнє значення показника за  $i$ -й період року (зазвичай береться період не менше 3х років – якщо це місячні дані, то розраховується середнє за 3 роки по кожному місяцю)
- $\bar{y}_{\text{заг}}$  — загальне середнє значення за всі роки.

Значні відхиленні (> 10%) від усередненого показника (100%) та їх повторення у певні періоди часу вказують на наявність сезонності.

Зазвичай, сезонну компоненту розраховують і виключають із ряду (тобто ряд коректують на сезонність), потім виявляють тенденцію, або трендову компоненту.

# Виявлення циклічних (регулярних) коливань

Для пошуку циклічних закономірностей у динаміці певних показників використовується **спектральний аналіз** (аналіз Фур'є, перетворення Фур'є, гармонійний аналіз; Spectral (Fourier) analysis), що передбачає перетворення динамічного ряду на послідовність синусоїдальних і косинусоїдальних функцій різних частот.

*Основними характеристиками при спектральному аналізі є частота ( $f$ , довжина хвилі функції, що виражається кількістю циклів за одиницю часу) та період ( $T=1/f$ , тривалість одного повного циклу).*



# Методи згладжування часового ряду

## 5. Мультиплікативна модель Холта-Вінтерса

Дана модель використовує індекс сезонності:

$$S_{t+p} = (a_t + b_t p) c_{t+p},$$

де

$$a_t = \alpha \left( \frac{y_t}{c_{t-s}} \right) + (1 - \alpha)(a_{t-1} + b_{t-1}),$$

$$b_t = \beta (a_t - a_{t-1}) + (1 - \beta) b_{t-1},$$

$$c_t = \gamma \left( \frac{y_t}{a_t} \right) + (1 - \gamma) c_{t-s}, \quad 0 \leq \alpha, \beta, \gamma \leq 1, \quad t = \overline{2s+1, T},$$

де  $s$  - кількість циклів сезонності. Цей метод вимагає щонайменше 2 повних цикли сезонності по минулих спостереженнях для побудови розрахунків.

Прогноз на період  $T + p$  будується наступним чином:

$$\hat{y}_{T+p} = (y_T + p b_T) c_{T-s+p}, \quad p = 1, 2, \dots, s,$$

$$\hat{y}_{T+p} = (y_T + p b_T) c_{T-2s+p}, \quad p = s+1, s+2, \dots, 2s.$$

# Виявлення аномальних рівнів (при нормальному розподілі даних) метод Ірвіна ґрунтується на порівнянні сусідніх значень модифікований ряду

Послідовний розрахунок за трьома  
спостереженнями величин:

$$\hat{\sigma}_y = \sqrt{\frac{\sum_{t=1}^n (y_{t-1} - \bar{y}_t)^2 + (y_{t+1} - \bar{y}_t)^2}{2}}$$

$$\bar{y}_t = \frac{(y_{t-1} + y_{t+1})}{2}, t = 2, 3, \dots, n-1,$$

# Метод Ірвіна (продовження)

2. Обчислити величину:

$$\lambda_t = \frac{|y_t - y_{t-1}|}{\hat{\sigma}_y} \quad t=2, 3, \dots, n.$$

3. Розраховані ковзні значення  $\lambda_t$  порівнюють із критичним значенням  $\lambda$  для  $n=3$  ( $\lambda_3=2,3$ ).

Якщо вони не перевищують критичне, то відповідні рівні вважаються нормальними.

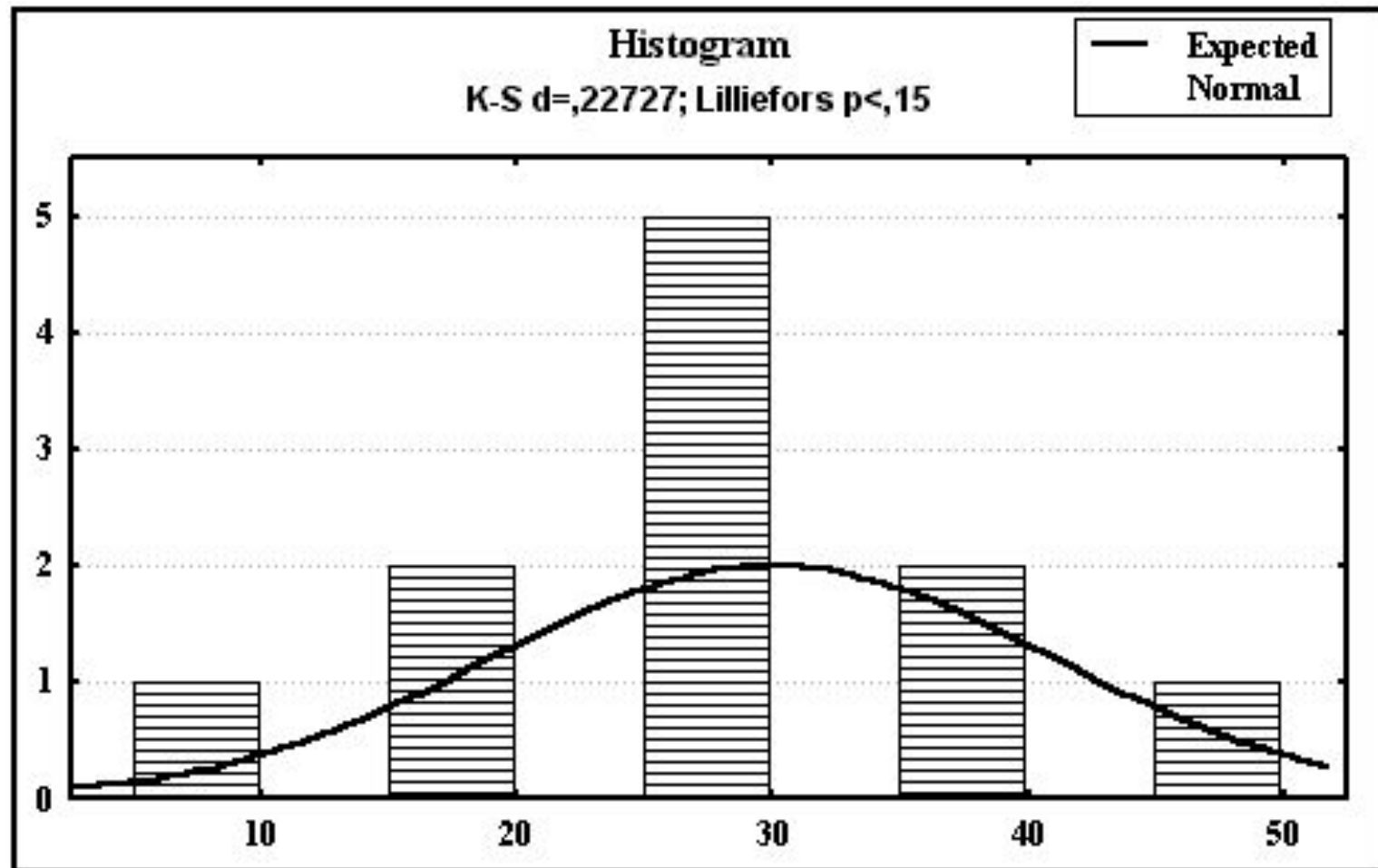
# Перевірка динамічного ряду на відповідність закону нормального розподілу

Точна форма нормального розподілу – “дзвоноподібна крива” – визначається лише двома параметрами: середнім і стандартним відхиленням. Для нормального розподілу характерним є те, що 68% всіх значень знаходяться в межах  $\pm 1$  стандартне відхилення (standard deviation) від середнього, а діапазон  $\pm 2$  стандартних відхилення включає 95% значень. Про відхилення від нормального розподілу свідчать коефіцієнти асиметрії (skewness) та ексцесу (kurtosis).

# Перевірка динамічного ряду на відповідність закону нормального розподілу

Найпростішим способом перевірки динамічного ряду на відповідність нормальному розподілу є побудова *гістограми* (histograms) – діаграми, яка показує, скільки значень попадає в кожний з інтервалів, на які можна розбити весь діапазон зміни певного показника, тобто частоту розподілу значень по інтервалах. На гістограму зазвичай накладається крива нормального розподілу, що дозволяє “на око” оцінити різні характеристики емпіричного розподілу.

# Перевірка динамічного ряду на відповідність закону нормального розподілу



# Перевірка динамічного ряду на відповідність закону нормального розподілу

Точнішу інформацію про відповідність динамічного ряду нормальному розподілу можна отримати за допомогою критеріїв нормальності, наприклад, **критерію Пірсона** ( ), критерію Колмогорова-Смірнова (K-S test for normality) чи критерію Ліллефорса (Lilliefors test for normality).

## Обрахунок стандартних статистичних параметрів, що характеризують динамічний ряд

Для стаціонарних динамічних рядів обраховуються *стандартне відхилення* (standard deviation, застаріла назва – середньоквадратичне відхилення) та *стандартна помилка* (standard error). Стандартне відхилення розраховується за формулою:

$$SD = \sqrt{\sigma^2} = \sqrt{\frac{SS}{n-1}} = \sqrt{\frac{\sum (y_t - \bar{y})^2}{n-1}}$$

де  $\sigma^2$  – дисперсія, SS – сума квадратів (sum of squares), n – величина динамічного ряду.

Стандартна помилка розраховується за формулою:

$$SE = \sqrt{\frac{\sigma^2}{n}}$$



# Обрахунок стандартних статистичних параметрів, що характеризують динамічний ряд

Для нестационарних динамічних рядів основним статистичним параметром є відхилення реальних значень від тренду – *середньоквадратична помилка* (інколи – стандартне відхилення), що розраховується за формулою:

$$S_y = \sqrt{\frac{\sum (y_t - \hat{y}_t)^2}{n}}$$

де  $y_t$  – реальні значення,  $\hat{y}_t$  – трендові значення,  $n$  – кількість неспівпадінь реальних значень із трендовими

# Етапи аналізу ЧР на стаціонарність

## **(3) Перевірка часового ряду на стаціонарність:**

3.1. Побудова автокореляційної (корелограми) і часткової автокореляційної функції для рівнів часових рядів, їх перших та других різниць.

3.2. Тест на наявність одиничних корнів – розширений тест Дікі-Фуллера (augmented Dickey – Fuller test, ADF, або unit root test).

## **(4) Приведення до стаціонарного виду шляхом:**

4.1. Виділення тренда – логарифмування – позначається літерою *I*; знаходження перших або других різниць ряду – позначається літерою *d* (“difference”);

4.2. Сезонного згладжування –*sa* (“seasonal ajustment”);

4.3. Усунення аномальних значень, нормалізація ряду.

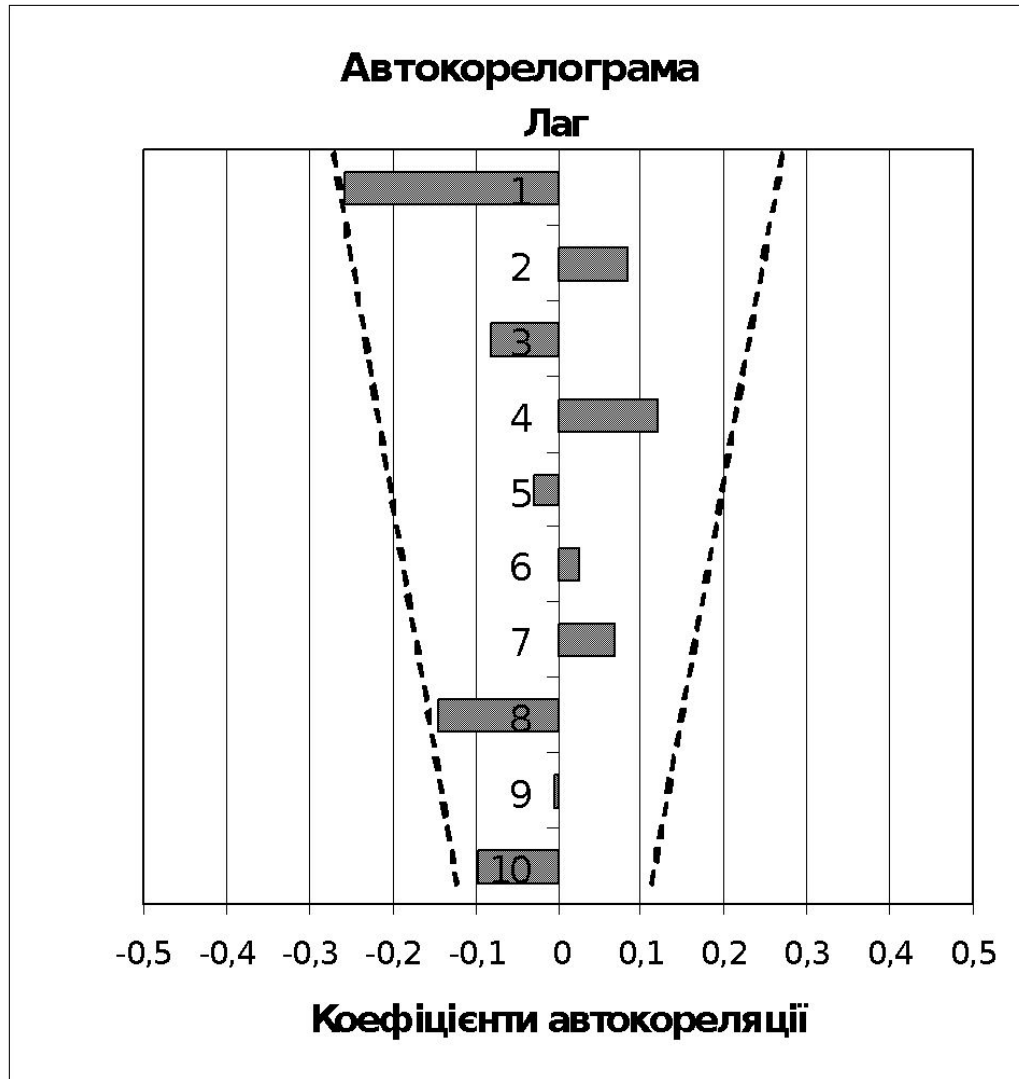
# Дослідження стаціонарності часового ряду в EViews

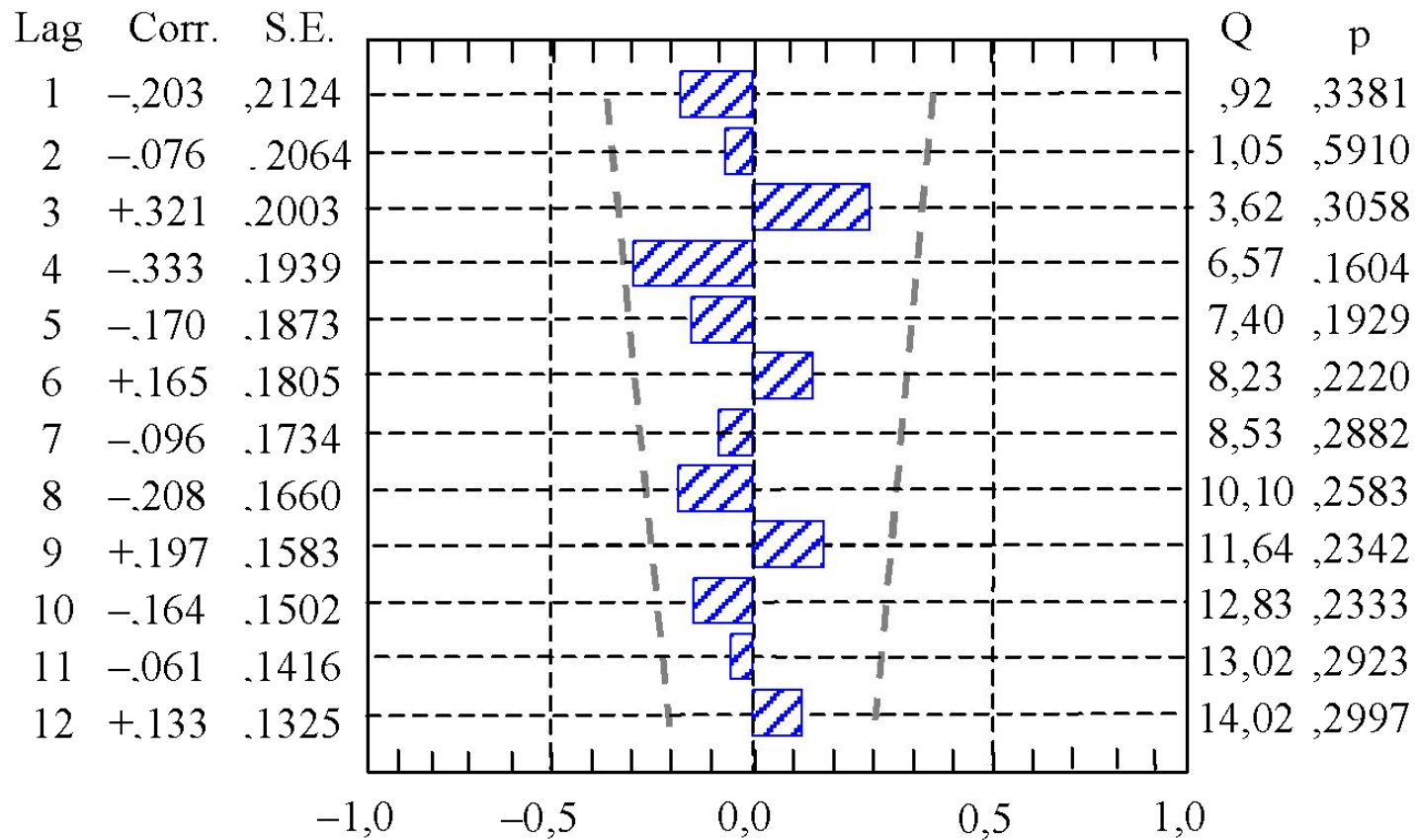
3.1. Побудова автокореляційної (корелограми) і часткової автокореляційної функції з різними часовими лагами для рівнів часових рядів, їх перших та других різниць .

$$r_a = \frac{\sum y_t y_{t-k} - n\bar{y}_t^2}{\sum y_t^2 - n\bar{y}_t^2}$$

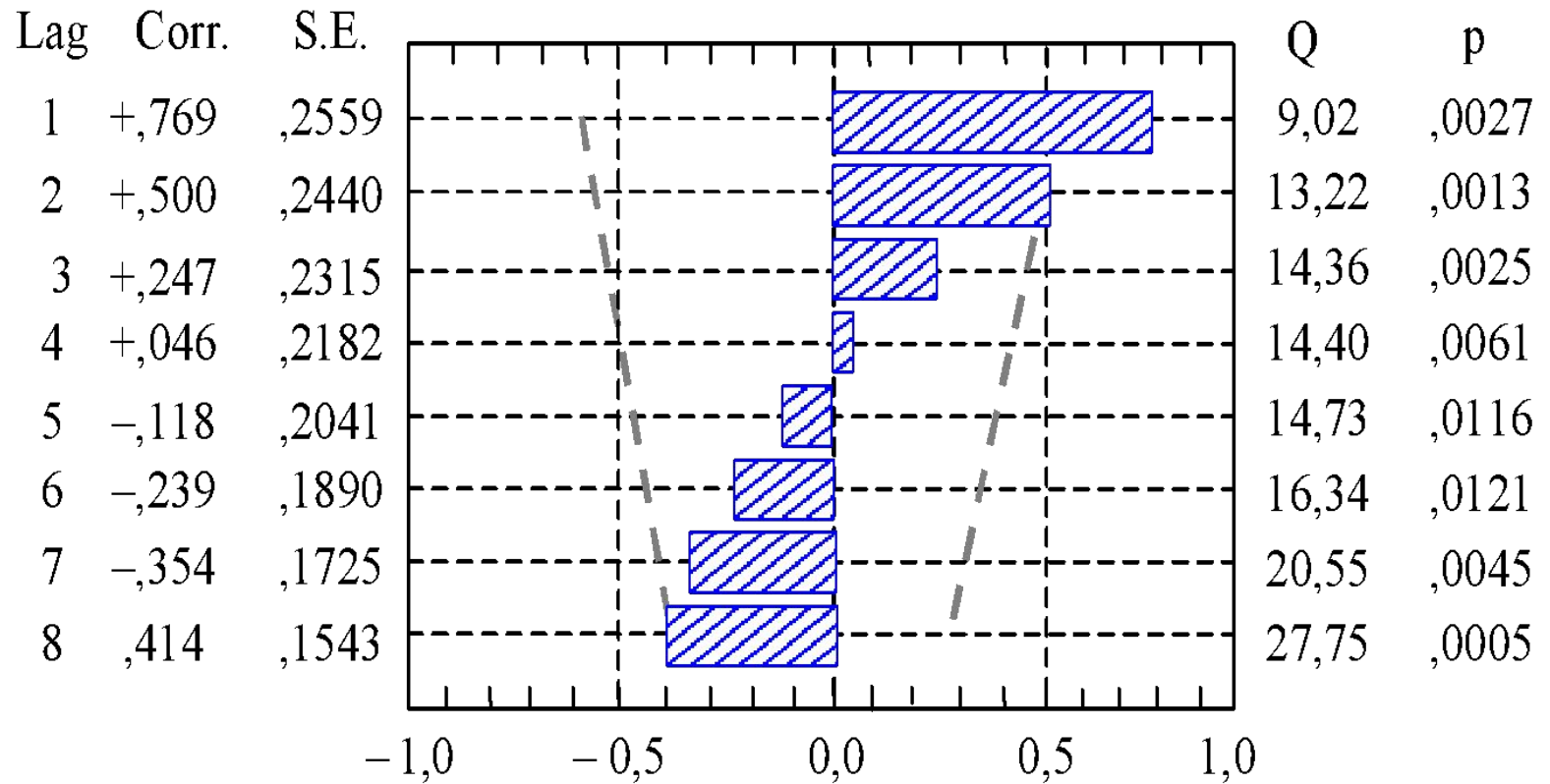
Автокореляційна функція показує ступінь тісноти зв'язку між  $t$  спостереженнями часового ряду. Часткова автокореляційна функція розраховується за формулою часткового коефіцієнту кореляції. Корелограма і графік часткової автокореляційної функції мають швидко спадати зі зростанням  $t$  у разі стаціонарного ряду. Перевірка гіпотези про те, що автокореляція відсутня до  $k$ -го лагу включно, здійснюється за допомогою статистики Льюнга-Бокса та її  $p$ -значення.

# Автокорелограма

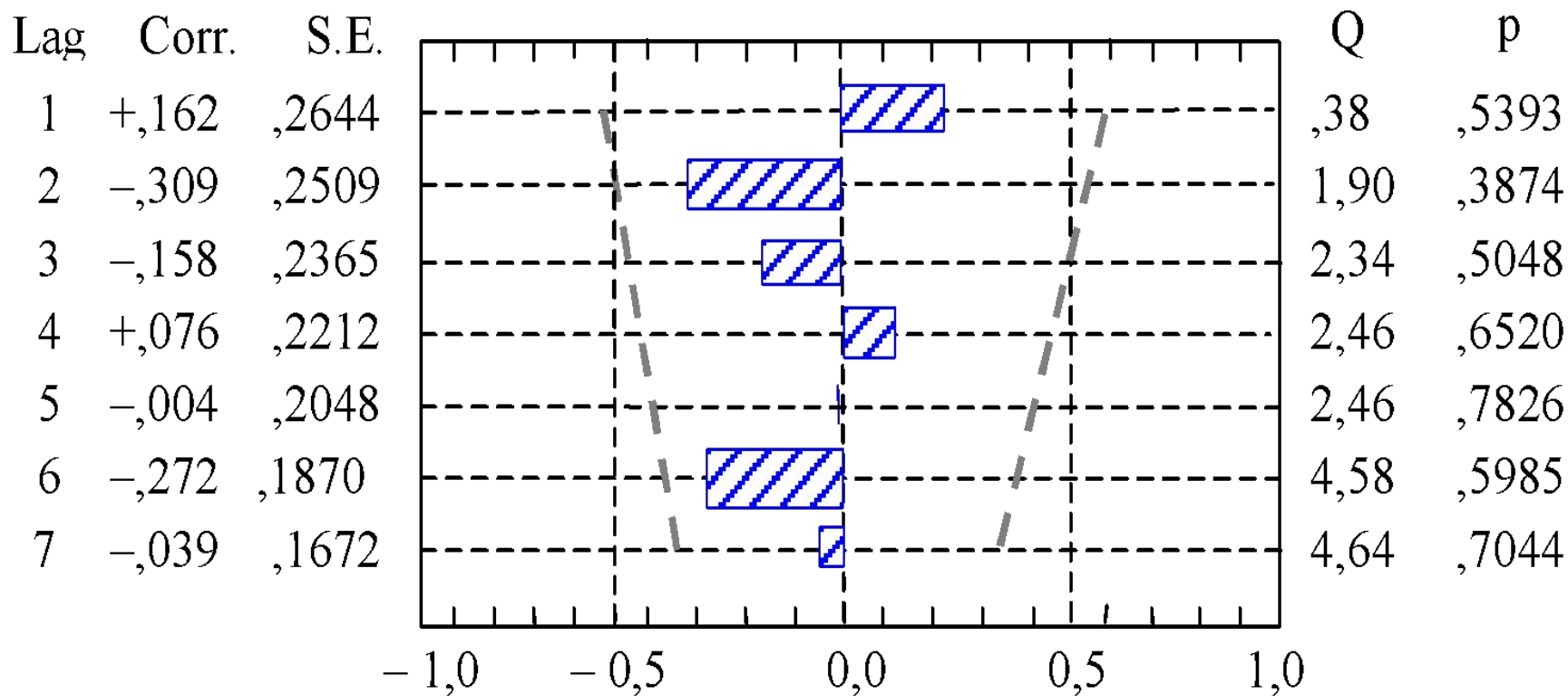




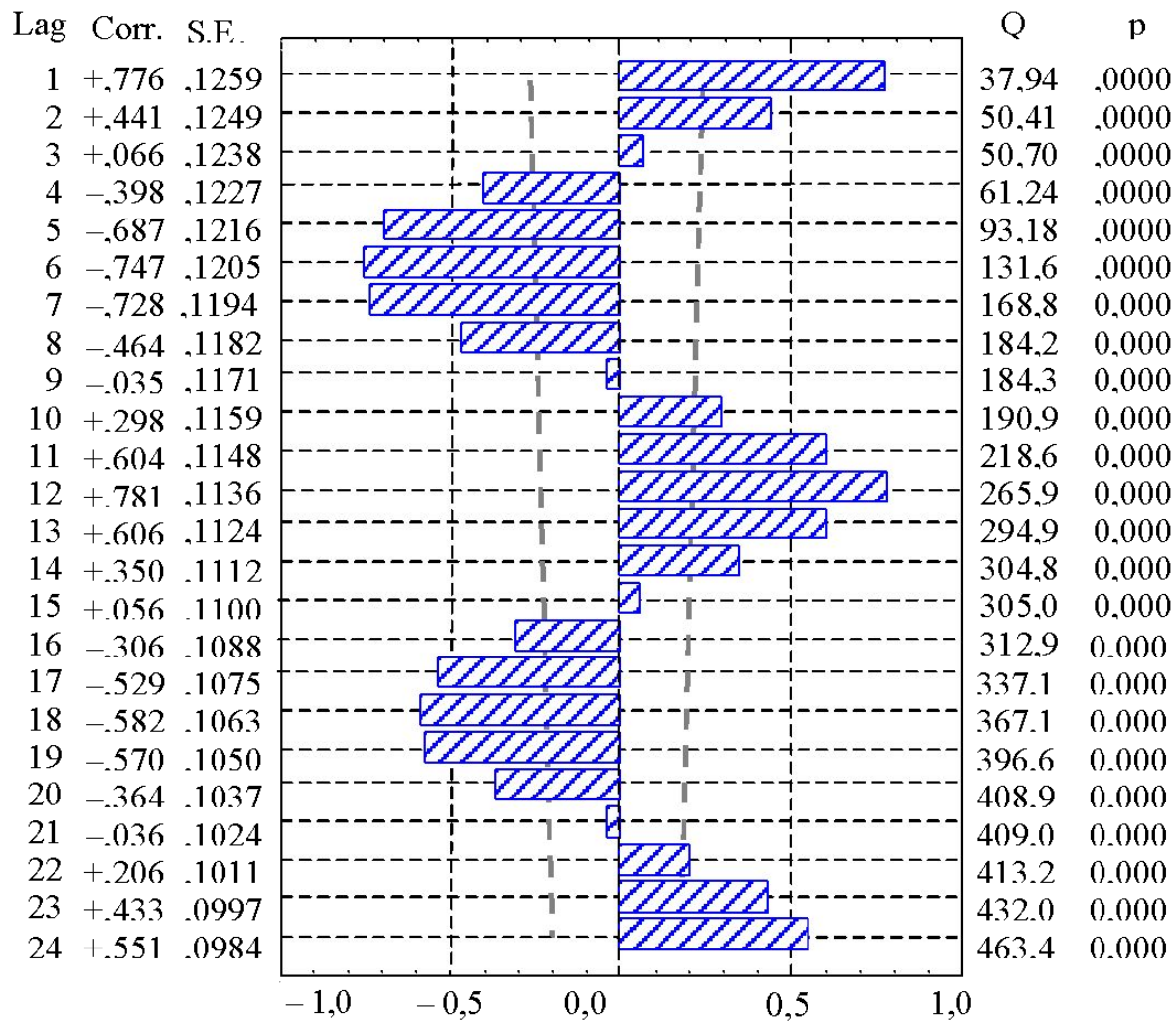
Корелограма для ряду, що не має тренду  
(стаціонарного ряду)



Корелограма для випадку лінійно-адитивного тренду (нульові різниці)

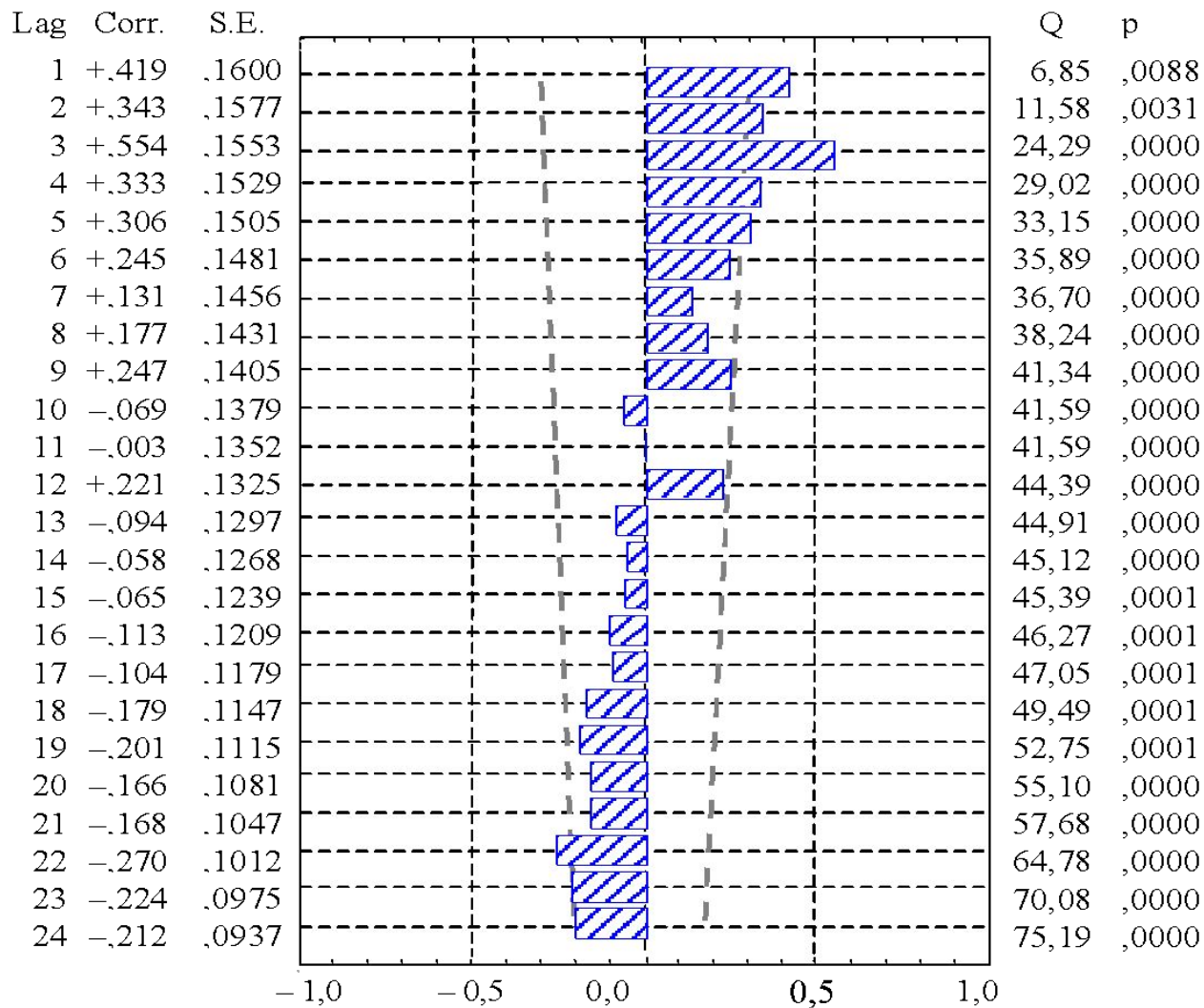


Корелограма для випадку лінійно-адитивного тренду (перші різниці)

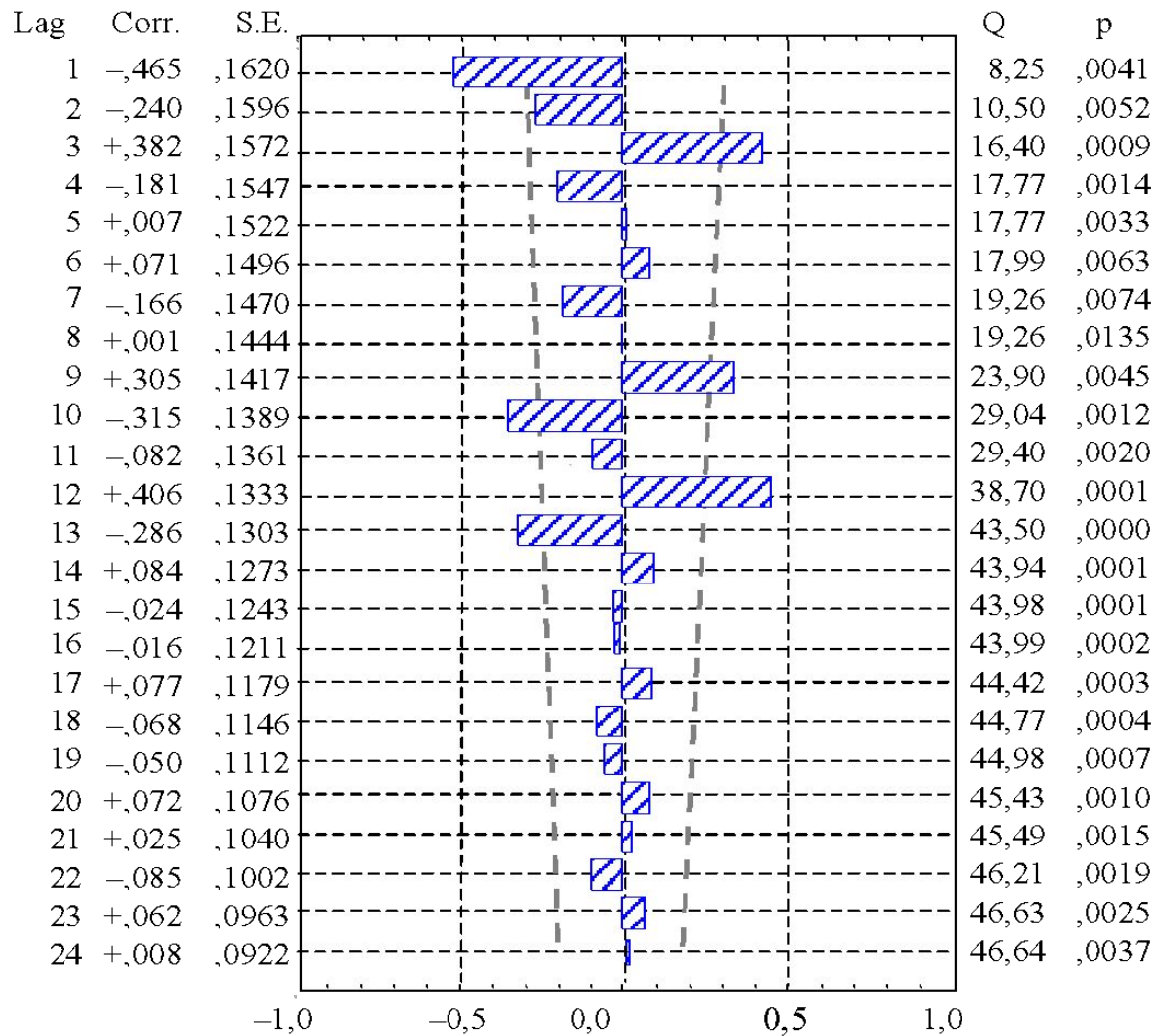


Корелограма ряду із сезонним коливанням  
(нульові різниці)





Корелограма ряду з лінійним трендом та сезонно-адитивною складовою (нульові різниці)



Корелограма ряду з лінійним трендом та сезонно-адитивною складовою (перші різниці)

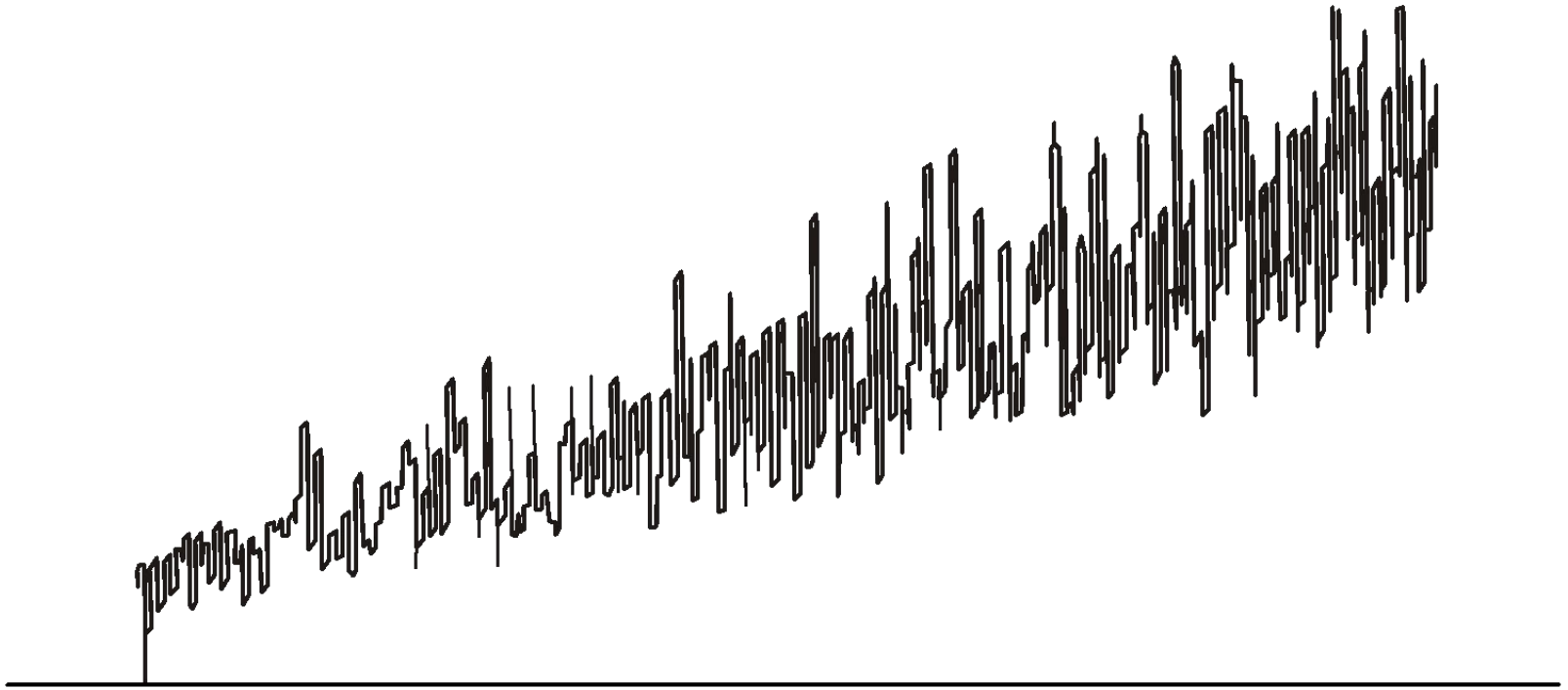


Ряд з лінійно-адитивним трендом

Зміни реального ВВП за 1996—2000 роки, % зміни  
(індекс грудня 1995 року = 100)



Сезонний ряд



Ряд з лінійно-мультиплікативним трендом



Динаміка попиту з лінійним трендом і сезонно-адитивною складовою



Динаміка ВВП із сезонно-мультиплікативним трендом

# Дослідження стаціонарності часового ряду в EViews

3.13.2. Тест на наявність одиничних корнів – розширений тест Дікі-Фуллера (augmented Dickey – Fuller test, ADF, або unit root test).

Часовий ряд має одиничний корінь, або порядок інтеграції один ( $y(k) \sim I(1)$ ), якщо його перші різниці ( $\Delta y(k) = y(k) - y(k-1)$ ) утворюють стаціонарний ряд ( $\Delta y(k) \sim I(0)$ ).

ADF-тест дозволяє тестувати гіпотезу про наявність одиничних коренів у моделях, де кількість лагів більше одного.

ADF-статистики розраховується спочатку для ряду, потім для перших, других і т.д. різниць. Умова стаціонарності виконується, якщо значення ADF-статистики не перевищує відповідне критичне значення на встановленому рівні значущості. При цьому ряд,  $k$ -і різниці якого стаціонарні, називають інтегрованим рядом  $k$ -го порядку і позначають  $I(k)$ . Стаціонарний ряд позначають  $I(0)$