

# ***Автокореляція***



# **Зміст**

- 1. Природа автокореляції. Основні поняття та означення.**
- 2. Тестування автокореляції. Критерій Дарбіна-Уотсона.**
- 3. Приклад тестування автокореляції.**

# 1. Природа автокореляції та її наслідки

Розглянемо класичну лінійну багатofакторну модель

$$y = a_0 + a_1x_1 + a_2x_2 + \dots + a_mx_m + u$$

або в матричному вигляді

$$Y = Xa + u$$

де  $y$  – вектор-стовпець залежної змінної розмірності  $n \times 1$ ;

$X$  – матриця незалежних змінних розміром  $n \times (m+1)$ ;

$a$  – вектор-стовпець невідомих параметрів розмірності  $(m+1) \times 1$

$u$  – вектор-стовпець випадкових помилок розмірності  $n \times 1$

$$\text{cov}(u_i, u_j) = 0, \quad i \neq j$$

Відсутність залежності між залишками

$$u_i \quad \text{і} \quad u_{i \pm k}$$

буде гарантувати відсутність зв'язку і між випадковими величинами

$$u_i \quad \text{і} \quad u_{i \pm 1}$$

тобто між сусідніми відхиленнями.  
Таким чином, **коваріація**

$$\text{cov}(u_i, u_{i \pm 1}) = 0, \quad i = 2, 3, \dots, n$$

У випадку, коли

$$\text{cov}(u_i, u_{i\pm 1}) \neq 0, \quad i = 2, 3, \dots, n$$

залежність, а значить і кореляція між сусідніми відхиленнями, буде існувати.

Ця кореляція називається **автокореляцією** (послідовною кореляцією), і є показником наявності зв'язку між упорядкованими в часі випадковими величинами.

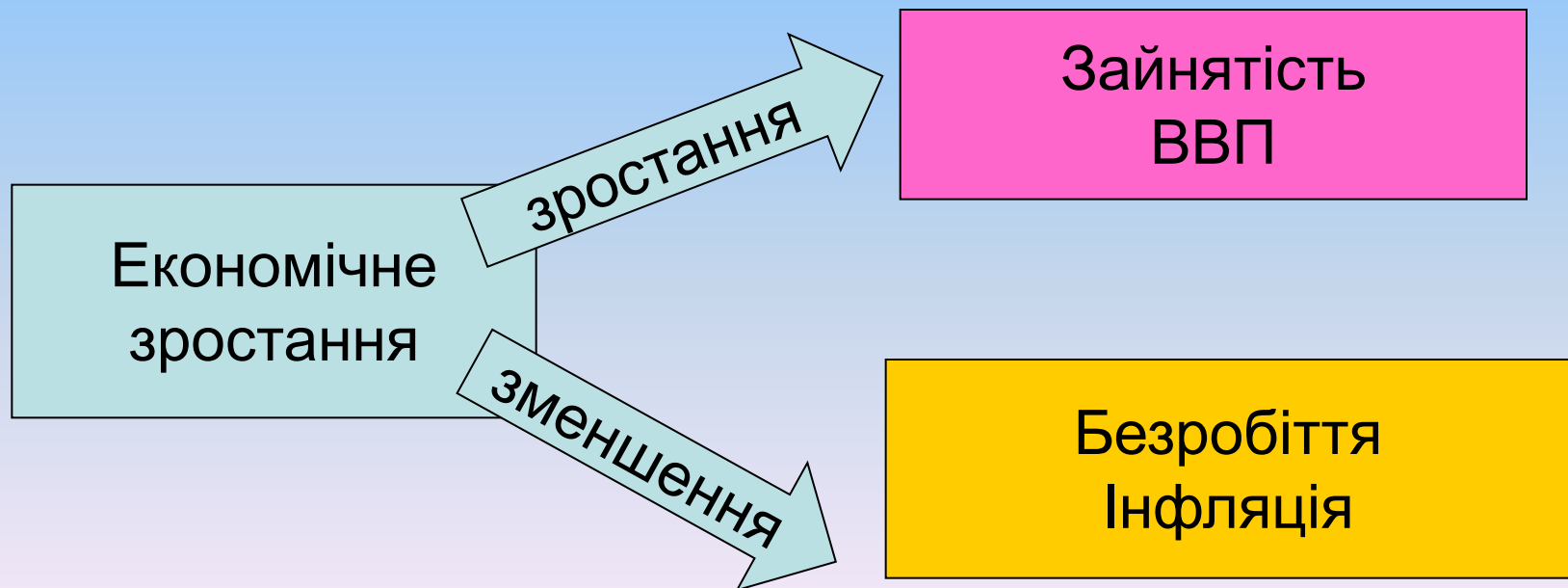
Головними причинами автокореляції можуть бути: *помилка специфікації, інерційність* в зміні економічних показників, ефект павутиння.

### *1. Помилки специфікації.*

В цьому випадку в моделі можуть бути не враховані важливі пояснювальні змінні, або може бути неправильно вибрана залежність між регресандом  $Y$  та регресорами  $X_i$ , що, як правило, викликає відхилення реальних значень  $Y = y_i$  від функції регресії.

## 2. Інерційність.

Багатьом економічним показникам, наприклад, інфляції, безробіттю, валовому продукту ВВП і т. ін., притаманна певна циклічність, яка пов'язана із хвилеподібним явищем ділової активності.



### 3. Статистична обробка інформації.

При обробці статистичної інформації за певний період часу використовують *усереднені дані*, одержані на інтервалах часу, а це призводить до згладжування коливань, які можуть існувати для кожного інтервалу, що може бути однією з причин появи автокореляції.



## НАСЛІДКИ АВТОКОРЕЛЯЦІЇ

1. Оцінки параметрів моделі можуть бути незміщеними, але неефективними, тобто вибіркові дисперсії вектора оцінок ***a*** можуть бути невиправдано великими.

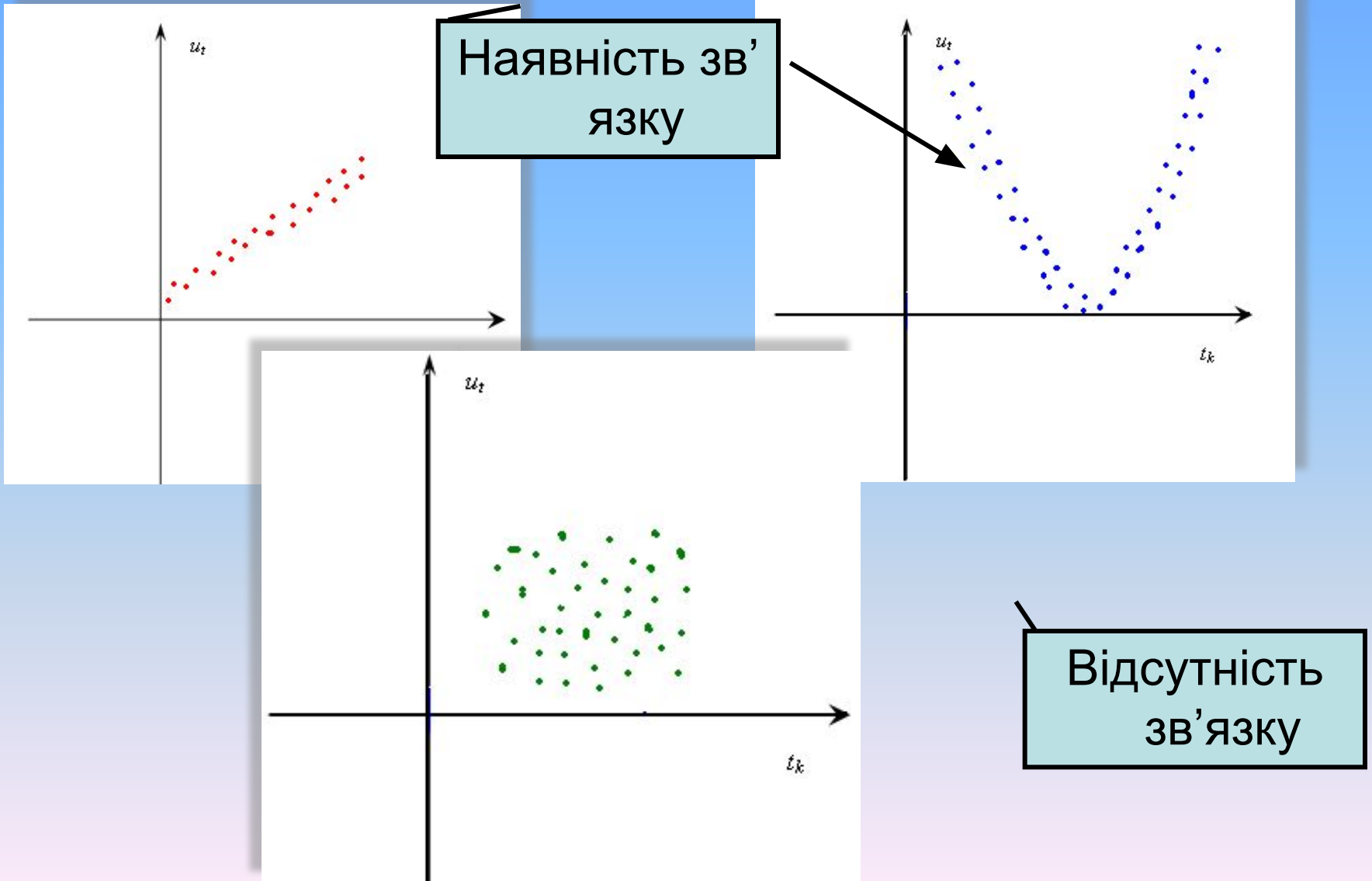
2. Статистичні критерії ***t*** і ***F***-статистик, які отримані для класичної лінійної моделі, не можуть бути використані для дисперсійного аналізу, бо їх розрахунок не враховує наявності коваріації залишків.

3. Неефективність оцінок параметрів економетричної моделі, як правило, призводить до **неефективних прогнозів**, тобто прогнозні значення матимуть велику вибіркову дисперсію.

***Висновки.*** За наявності автокореляції поширеним методом оцінювання невідомих параметрів є узагальнений метод найменших квадратів. Отримані за допомогою УМНК оцінки є незміщеними та ефективними.

## 2. Тестування наявності автокореляції

*Графічний метод*



Тестування наявності автокореляції, як правило, здійснюється за ***d*-тестом Дарбіна — Уотсона**.

**Інші тести**: критерій фон Неймана, нециклічний коефіцієнт автокореляції, циклічний коефіцієнт автокореляції.

# Критерій Дарбіна — Уотсона

**Крок 1.** Розраховується значення  $d$  - статистики за формулою

$$DW = d = \frac{\sum_{t=2}^n (u_t - u_{t-1})^2}{\sum_{t=1}^n u_t^2}$$

**Зауваження.** Доведено, що значення  $d$  -статистики Дарбіна — Уотсона перебуває в межах

$$0 \leq DW \leq 4$$

**Крок 2.** Задаємо рівень значущості  $\alpha$ . За таблицею Дарбіна — Уотсона при заданому рівні значущості  $\alpha$ , кількості факторів  $m$  і кількості спостережень  $n$  знаходимо два значення

$$DW_1 \text{ і } DW_2$$

Якщо  $0 < DW < DW_1$  наявна **додатна** автокореляція.

Якщо  $DW_1 \leq DW \leq DW_2$  або  $4 - DW_2 \leq DW \leq 4 - DW_1$

ми **не можемо** зробити висновки ані про наявність, ані про відсутність автокореляції

$DW$  (потрапляє в зону невизначеності).

Якщо

$$4 - DW_1 < DW < 4$$

маємо **від'ємну** автокореляцію.

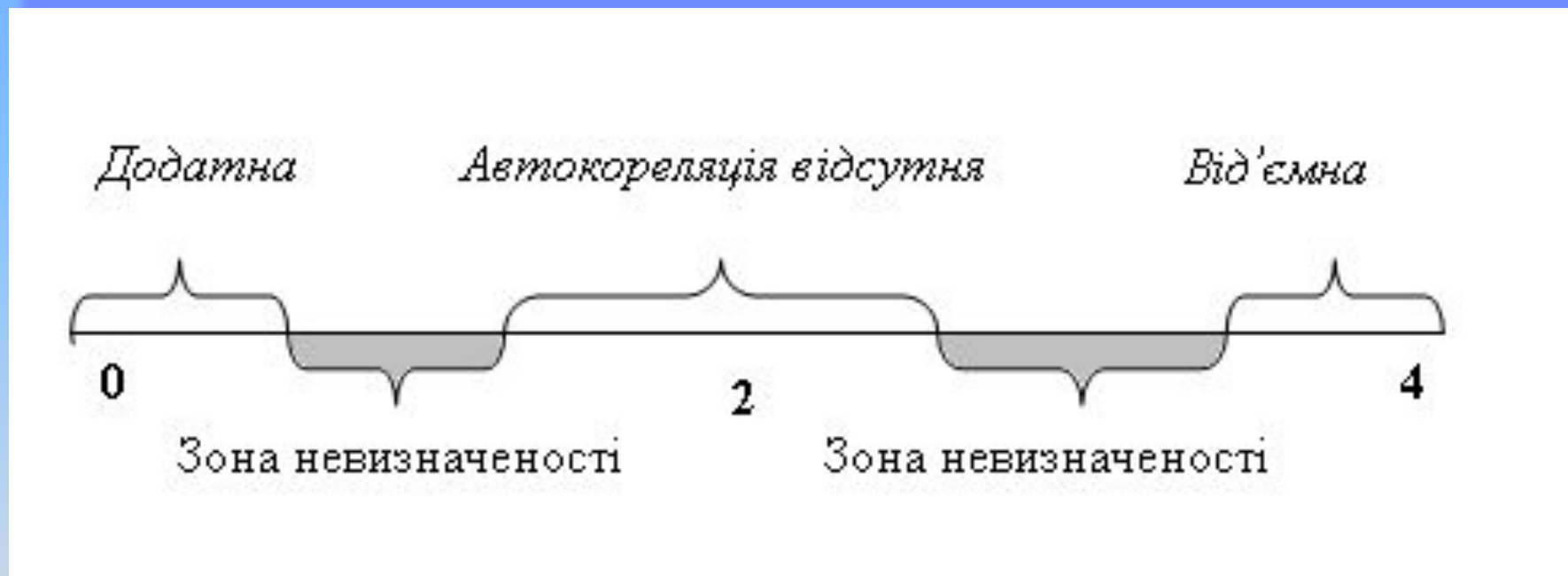
Якщо

$$DW_2 < DW < 4 - DW_2$$

автокореляція **відсутня**

## Графічне зображення

Зони автокореляційного зв'язку за критерієм Дарбіна-Утсона



$DW_1$

$DW_2$

$4 - DW_2$

$4 - DW_1$



# Критерій фон Неймана

Розраховується

$$Q = Q_{\text{факт}} = \frac{\sum_{t=2}^n (u_t - u_{t-1})^2}{\sum_{t=1}^n u_t^2} \cdot \frac{n}{n-1}$$

Звідси

$$Q = DW \cdot \frac{n}{n-1}$$

Отже, при

$$n \rightarrow \infty \quad Q = DW$$

Фактичне значення критерію фон Неймана порівнюється з табличним при вибраному рівні значущості  $\alpha$  і заданій кількості спостережень:

$$Q_{\text{табл}} = Q_{(\alpha, n)}$$

Якщо

$$Q_{\text{факт}} < Q_{\text{табл}}$$

то існує *додатна* автокореляція.

## Приклад оцінювання параметрів моделі з автокорельованими залишками

На основі двох взаємопов'язаних часових рядів про роздрібний товарообіг і доходи населення побудувати модель, що характеризує залежність роздрібного товарообігу від доходу.

Специфікація моделі:

$y$ – роздрібний товарообіг,  
 $x$ – дохід.



товарообіг

дохід

$$y = 2,3136 + 0,8683 * x$$

	1	2	3	4	5	6	7	8
	y	x	yr	u <sub>t</sub>	u <sub>t</sub> <sup>2</sup>	u <sub>t-1</sub>	(u <sub>t</sub> -u <sub>t-1</sub> ) <sup>2</sup>	u <sub>t</sub> u <sub>t-1</sub>
1	25,5	27,6	26,27868	-0,77868	0,606343			
2	26,4	27,4	26,10502	0,29498	0,087013	-0,77868	1,1527458	-0,2297
3	27,9	28,7	27,23381	0,66619	0,443809	0,29498	0,13779686	0,196513
4	28,1	29,5	27,92845	0,17155	0,029429	0,66619	0,24466873	0,114285
5	28,8	30,9	29,14407	-0,34407	0,118384	0,17155	0,26586398	-0,05903
6	29,3	31,4	29,57822	-0,27822	0,077406	-0,34407	0,00433622	0,095727
7	29,8	31,8	29,92554	-0,12554	0,01576	-0,27822	0,02331118	0,034928
8	30,7	32,2	30,27286	0,42714	0,182449	-0,12554	0,30545518	-0,05362
9	31,5	33,6	31,48848	0,01152	0,000133	0,42714	0,17273998	0,004921
10	32,4	34,7	32,44361	-0,04361	0,001892	0,01152	0,00383932	-0,0005
<b>SUM case 1-10</b>		<b>307,8</b>	<b>290,3987</b>	<b>0,00176</b>	<b>1,562628</b>	<b>0,04487</b>	<b>2,30995726</b>	<b>0,103527</b>

$$DW = d = \frac{\sum_{t=2}^{10} (u_t - u_{t-1})^2}{\sum_{t=1}^{10} u_t^2} = \frac{2,3099}{1,5626} = 1,4782$$

Задаємо  $\alpha=0,05$  і при  $n=10$  і  $m=1$  знайдемо за таблицею  $d$ -статистику Дарбіна — Уотсона критичні значення критерію:

$$DW_{\text{нижня}} = 0,879$$
$$DW_{\text{верхня}} = 1,320$$

Оскільки  $DW_2 < DW < 4 - DW_2$   
то автокореляція  
*відсутня.*

# Параметризація моделі з автокорельованими залишками

Параметри моделі з автокорельованими залишками можна оцінити на основі чотирьох методів:

1. Ейткена (УМНК);
2. Перетворення вихідної інформації;
3. Кочрена — Оркатта;
4. Дарбіна.

## Узагальнений метод найменших квадратів (метод Ейткена)

Оператор оцінювання УМНК можна записати так:

$$\hat{a} = (X^T \Omega^{-1} X)^{-1} X^T \Omega^{-1} Y$$

де  $\Omega^{-1}$ -матриця, обернена до дисперсійно-коваріаційної матриці залишків  $\Omega$ .



$$\Omega^{-1} = \frac{1}{1-\rho^2} \cdot \begin{pmatrix} 1 & -\rho & 0 & 0 & 0 & 0 \\ -\rho & 1+\rho^2 & -\rho & 0 & 0 & 0 \\ 0 & -\rho & 1+\rho^2 & -\rho & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

На практиці для розрахунку  $\rho$  використовується співвідношення

$$\rho \approx r \approx \frac{\sum_{t=2}^n u_t u_{t-1}}{\sum_{t=1}^n u_t^2}$$

або

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