УДК 330.101.52(075.8)

Андрієнко В. М.
кандидат економічних наук,
доцент кафедри
економічної кібернетики та інформаційних технологій
Одеського національного політехнічного університету
Семенов А. С.
кандидат фізико-математичних наук,
доцент кафедри
економічної кібернетики та інформаційних технологій
Одеського національного політехнічного університету

Andriyenko V. M.
Candidate of Economic Sciences,
Senior Lecturer at Department
of Economic Cybernetics and Information Technologies
Odessa National Polytechnic University
Semenov A. S.
Candidate of Physical and Mathematical Sciences,
Senior Lecturer at Department of
Economic Cybernetics and Information Technologies
Odessa National Polytechnic University

МЕТОДИКА СТАТИСТИЧНОГО АНАЛІЗУ ЕКОНОМІЧНИХ ЧАСОВИХ РЯДІВ

Анотація. Стаття присвячена проблемі комплексного статистичного аналізу емпіричних даних складних економічних систем. Запропоновано методику, що уточнює особливості застосування методів і полягає в логічній взаємопов'язаній послідовності основних етапів аналізу. Такий підхід може виявити характерні особливості емпіричних даних і дати адекватне уявлення про досліджувану систему. Наведено тестовий розрахунок на реальних даних. На основі результатів аналізу визначено, ідентифіковано математичну модель динаміки процесу.

Ключові слова: часовий ряд, стаціонарний ряд, тренд, періодограма, корелограма, сезонна складова, циклічна складова.

Вступ та постановка проблеми. Статистичний опис розвитку економічних процесів у часі здійснюється за допомогою часових рядів. Статистичний аналіз інформації, представленої у вигляді часових рядів, є необхідною складовою частиною сучасних економічних досліджень. Основні праці з методології аналізу часових рядів належать Т. Андерсону, В.Н. Афанасьєву, В.В. Витязеву, В.В. Іванову, Ю.Л. Лукашину, М. Четиркіну, Дж. Боксу (G. Box). Д. Бриллінджеру (D. Brillinger), Т. Боллерслеву (T. Bollerslev), Д. Дікі (D. Dickey), В. Фуллеру (W. Fuller), М. Хатанака (M. Hatanaka), В. Ендерсу (W. Enders).

Набір методів аналізу досить широкий. Деякі методи є більш-менш універсальними, інші – більш спеціалізованими та вимагають подальшого розроблення й апробування.

Велика різноманітність наявних методів, недостатня обізнаність фахівців про особливості використання тих чи інших методів, складність застосовуваного математичного апарату створюють для аналітиків труднощі та навіть призводять до формулювання неправильних висновків. Дослідженню часових рядів приділяється багато уваги в роботах вітчизняних та зарубіжних науковців.

Аналіз останніх досліджень і публікацій. Нині наукові роботи з аналізу часових рядів присвячені здебільшого розгляду теоретичного аспекту [1; 2; 3] або вивченню окремих властивостей рядів і не передбачають всебічного аналізу [4–7]. Відсутні спеціальні роботи, що відображають опис окремих деталей методів.

Метою роботи є розвиток наявної методології статистичного аналізу часових рядів, вдосконалення методик статистичного аналізу.

Результати дослідження. Часові ряди відрізняються від простих статистичних вибірок, адже мають такі ознаки:

– послідовні в часі показники часових рядів є взаємозалежними, особливо це відноситься до близько розташованих спостережень;

– залежно від моменту спостереження показники часового ряду мають різну інформативність, інформаційна цінність спостережень зменшується за ступенем їх віддалення від поточного моменту часу;

– зі збільшенням кількості показників часового ряду точність статистичних характеристик не збільшується пропорційно числу спостережень, а за появи нових закономірностей розвитку вона може навіть зменшуватися.

Більшість методів аналізу належить до стаціонарних (в широкому або вузькому сенсі) процесів. Це пов'язане з тим, що для стаціонарних випадкових процесів доведені теореми, що дають змогу отримувати коректні оцінки параметрів відповідних розподілів за даними спостережень, тобто по деякому часовому ряду.

Основна мета аналізу часового ряду полягає у створенні прогнозу його значень на майбутні періоди, а головним завданням аналізу часового ряду є визначення типу ряду. Нестаціонарний часовий ряд може мати трендову, сезонну, циклічну та випадкову складові. Перші три компоненти є регулярними. Якщо їх виключити, то вийде стаціонарний ряд. Випадкова складова присутня завжди. Однак часовий ряд може мати більш складну структуру, наприклад мати стохастичний тренд або хаотичну складову чи бути фракталом. В такому разі його неможливо привести до стаціонарного виду. Його подальший аналіз проводиться за допомогою спеціальних методів, а саме фрактального та хаос-динамічного аналізу. Для нестаціонарного ряду важливо зрозуміти, під впливом яких компонент формується його значення, і привести його, якщо це можливо, до стаціонарного типу.

Під трендом розуміють зміну, яка визначає загальний напрям розвитку, основну тенденцію часового ряду. Це систематична складова довгострокової дії. Поряд із тривалими тенденціями у часових рядах економічних процесів часто мають місце регулярні коливання, тобто періодичні складові рядів динаміки. Якщо період коливань не перевищує одного року, то їх називають сезонними. За більшого періоду коливання вважають, що в часових рядах має місце циклічна складова. Графічне зображення основних складових (компонент) наведено на рис. 1 [8].

Тому під час аналізування часових рядів застосовується комплекс різних методів. Періодичну компоненту можна визначити за допомогою спектрального аналізу. Статистичну залежність або незалежність значень часового ряду й регулярні складові виявляють на основі кореляційного та спектрального аналізу.

Наведемо загальну схему статистичного аналізу.

1) Графічне представлення вихідного ряду в часовій області.

Зазвичай візуальне вивчення графіка вихідного ряду дає змогу виявити або присутність в даних постійного доданка, або низькочастотний компонент (тренд). Для вивчення високочастотних компонент ряду обидві ці складові потрібно виключити з даних. Постійний доданок виключається центруванням ряду, а центрований ряд виходить з вихідного X_t відніманням середнього значення $\overline{X}_t$, а саме виконують дію $X_t - \overline{X}_t$, де $\overline{X}_t$ – середнє значення ряду. При цьому середнє значення ряду обчислюється за формулою $\overline{X}_t = \frac{1}{T}\sum_{t=1}^{T} X_t$, де T – число членів вихідного ряду.

2) Перевірка гіпотези про наявність тренду.

На практиці для перевірки гіпотези про наявність тренду використовуються параметричні тести на сталість математичного очікування та дисперсії за критеріями Стьюдента і Фішера [1]. Ці тести можна застосовувати тільки в разі нормального розподілу даних. Якщо ця умова не виконується або тренд відсутній, то слід перейти до пункту 4. Якщо виявлено тренд, то слід перейти до наступного пункту.

3) Моделювання тренду.

Для виключення тренду необхідно задати його модель. Якщо природа тренду має теоретичне пояснення, то моделювання тренду проводиться на основі цієї теорії. Найчастіше природа тренду невідома. Тоді як формальну модель використовують апроксимацію тренду за допомогою лінійної комбінації якихось поліномів. При цьому до складу моделі входить також вільний член. Параметри моделі тренду визначаються за методом найменших квадратів.

4) Обчислення періодограми.

5) Графічне представлення періодограми.

Періодограма дає змогу виявити високочастотні компоненти ряду, які зазвичай не є суттєвими, тому що викликані випадковими чинниками, а також зображують різкі відхилення. Якщо такі є, то проводиться згладжування ряду. Але якщо періодограма на низьких частотах звертається в нескінченність, то це свідчить про наявність тренду. Якщо ряд містить періодичні складові, то вони виявляться на графіку періодограми у вигляді максимумів (піків). Однак будь-яка виявлена математичними методами періодична складова повинна бути обґрунтована зі змістовної точки зору. Тому що формальне застосування методів внаслідок неправильно вибраних періодів, неправильно проведеного згладжування, нарешті, за наявності великих шумів може привести до виявлення неправдивих періодичних складових. Сезонну й регулярні компоненти потрібно виключити з ряду, після чого треба перейти до пункту 6.

6) Обчислення корелограми.

7) Графічне зображення корелограми.

Якщо корелограма експоненціальним чином убуває або має вигляд затухаючої синусоїди, то досліджуваний ряд є стаціонарним. Тоді слід перейти до пункту 8. В іншому разі ряд нестаціонарний, тому потрібно визначити природу нестаціонарності ряду. Якщо корелограма повільно убуває, а періодограма на низьких частотах звертається в нескінченність, то ряд володіє тривалою пам'яттю. Такий ряд не можна привести до стаціонарного виду, а далі його потрібно досліджувати спеціальними методами. Слід перейти до пункту 9.

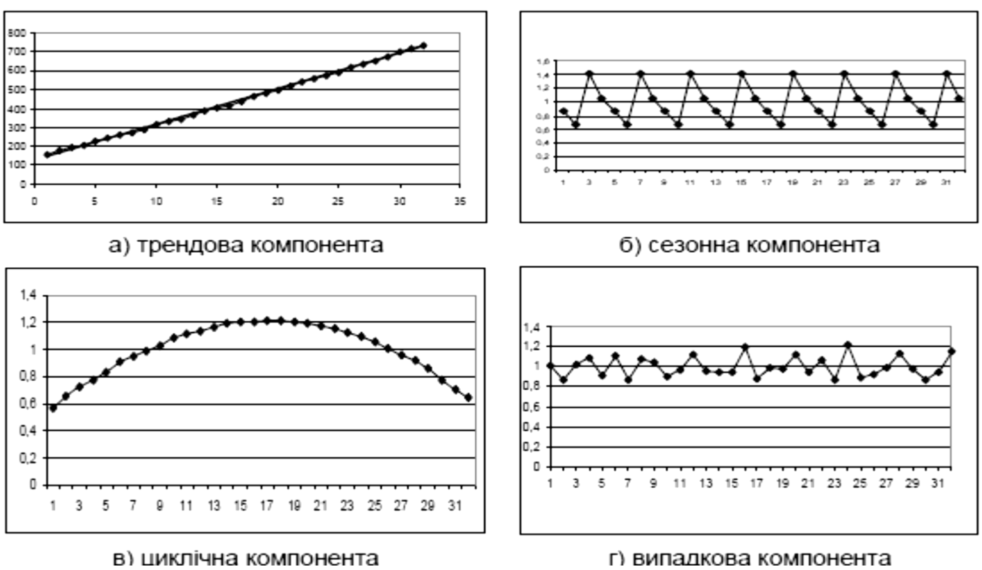

Рис. 1. Приклад основних компонент часового ряду

8) Тестування на стаціонарність

Тестування проводять за методом Дікі-Фуллера (*DF*-тест), методика якого полягає в такому [9; 10].

Часовий ряд має одиничний корінь, порядок інтеграції один, якщо його перші різниці утворюють стаціонарний ряд. Ця умова записується так: $X_t \sim I(1)$, якщо $\Delta X_t = X_t - X_{t-1}$, є стаціонарним $\Delta X_t \sim I(0)$. На практиці перевіряють значення коефіцієнта *a* в авторегресійному рівнянні першого порядку *AR* (1):

$$X_t = aX_{t-1} + \varepsilon_t\,, \quad (1)$$

де X_t – часовий ряд;

ε_t – помилка.

Якщо $a = 1$, то процес має одиничний корінь, тоді ряд X_t є не стаціонарним, є інтегрованим часовим рядом першого порядку: *I(1)*. Якщо $|a| < 1$, то ряд стаціонарний – *I(0)*. Для фінансово-економічних процесів значення $|a| > 1$ не характерне, тому що в цьому разі процес є «вибуховим». Виникнення таких процесів малоймовірне, тому що фінансово-економічне середовище є досить інерційним, що не дає змогу приймати нескінченно великі значення за малі проміжки часу.

Наведене авторегресійне рівняння (1) можна переписати так:

$$\Delta X_t = bX_{t-1} + \varepsilon_t\,, \quad (2)$$

де $b = a - 1$;

Δ – оператор різниці першого порядку $\Delta X_t = X_t - X_{t-1}$.

Тому перевірка гіпотези про одиничний корінь тут означає перевірку нульової гіпотези про рівність нулю коефіцієнта *b*. Оскільки випадок «вибухових» процесів виключається, то тест є одностороннім, тобто альтернативною гіпотезою є гіпотеза про те, що коефіцієнт *b* менше нуля. Статистика тесту (*DF*-статистика) – це звичайна *t*-статистика для перевірки значущості коефіцієнтів лінійної регресії. Однак розподіл цієї статистики відрізняється від класичного розподілу *t*-статистики, тобто розподілу Стьюдента або асимптотично нормального розподілу. Розподіл *DF*-статистики виражається через винерівський процес і називається розподілом Дікі-Фуллера. Гіпотеза про рівність нулю коефіцієнта *b* відхиляється, якщо $t_{факт} < t_{критич}$, $t_{факт}$ – це обчислена за емпіричними даними *t*-статистика Стьюдента, а $t_{критич}$ – це критичне значення статистики Дікі-Фуллера. Існує три версії тесту (тестових регресcій):

1) без константи й тренду:

$$\Delta X_t = bX_{t-1} + \varepsilon_t$$

2) з константою, але без тренду:

$$\Delta X_t = b_0 + bX_{t-1} + \varepsilon_t$$

3) з константою та лінійним трендом:

$$\Delta X_t = b_0 + b_1 t + bX_{t-1} + \varepsilon_t$$

Для кожної з трьох тестових регресій існують свої критичні значення *DF*-статистики. Якщо значення статистики лежить лівіше критичного значення (критичні значення є негативними) за цього рівня значущості, то нульова гіпотеза при одиничному корені відхиляється, а процес визнається стаціонарним (в сенсі цього тесту). В іншому разі гіпотеза не відкидається, а процес може містити одиничні коріння, тобто бути нестаціонарним (інтегрованим) часовим рядом. Нижче наведена таблиця критичних значень для різних рівнів значимості.

Тест Дікі-Фуллера, як і багато інших тестів, перевіряє наявність лише одного одиничного кореня. Однак процес може мати теоретично декілька одиничних коренів. Тоді тест може бути некоректним. Оскільки зазвичай передбачається, що більше трьох одиничних коренів навряд чи можуть зустрічатися в реальних економічних часових рядах, теоретично обґрунтованим є тестування перш за все другої різниці ряду. Якщо гіпотеза одиничного кореня для цього ряду відкидається, то тестується одиничний корінь в перших різницях. Якщо на цьому етапі гіпотеза не відкидається, то вихідний ряд має два одиничні кореня. Якщо відкидається, то перевіряється одиничний корінь в самому часовому ряді, як описано вище. На практиці часто все роблять у зворотній послідовності, що є не зовсім коректним.

Якщо в тестові регресії додати лаги перших різниць часового ряду, то розподіл *DF*-статистики (отже, критичні значення) не зміниться. Такий тест називають розширеним тестом Дікі-Фуллера (*ADF*). Необхідність включення лагів перших різниць пов’язана з тим, що процес може бути вищого порядку, ніж перший. Розглянемо на прикладі моделі *AR(2)*:

$$X_t = a_1 X_{t-1} + a_2 X_{t-2} + \varepsilon_t \quad (3)$$

Цю модель можна представити так:

$$\Delta X_t = (a_1 + a_2 - 1)X_{t-1} - a_2 \Delta X_{t-1} + \varepsilon_t\,. \quad (4)$$

Якщо часовий ряд має один одиничний корінь, то перші різниці за визначенням стаціонарні. Для перевірки наявності одиничних коренів у цій моделі слід провести стандартний *DF*-тест для коефіцієнта при X_{t-1}, причому в тестову регресію повинен бути доданий лаг першої різниці залежної змінної.

Крім того, модель може бути *ARMA*-процесом, тому слід перевірити наявність одиничного кореня для декількох лагів. Однак необхідно врахувати, що збільшення числа лагів приводить до зниження потужності тесту. Зазвичай обмежуються трьома-чотирма лагами.

Якщо перші різниці стаціонарні, то слід перейти до наступного пункту, в іншому разі варто перевірити на стаціонарність другі різниці. Якщо другі різниці стаціонарні, то слід перейти до наступного пункту, в іншому разі ряд має специфічні особливості. Для подальшого дослідження потрібні інші методи аналізу, а саме фрактальний та динамічний. Тоді слід перейти до пункту 9.

9) Кінець алгоритму.

Для реалізації описаних вище методів існують різні комп’ютерні програми, наприклад *Statistika, EViews, SPSS, AtteStat*. Тест Дікі-Фуллера включений в усі сучасні економетричні пакети.

Таблиця 1

Критичні значення статистики Дікі-Фуллера при рівнях значущості в 1%, 2,5%, 5%

Розмір вибірки	AR-модель	AR-модель з константою	AR-модель з константою та трендом
	1%; 2,5%; 5%	1%; 2,5%; 5%	1%; 2,5%; 5%
25	-2,66; -2,26; -1,95	-3,75; -3,75; -3,00	-4,38; -3,95; -3,6
50	-2,62; -2,25; -1,95	-3,58; -3,58; -2,93	-4,15; -3,80; -3,5
100	-2,60; -2,24; -1,95	-3,51; -3,17; -2,89	-4,04; -3,69; -3,45
>100	-2,58; -2,23; -1,95	-3,43; -3,12; -2,86	-3,96; -3,66; -3,41

Ряди, які після аналізу можна вважати стаціонарними, використовуються для створення деякої прогнозної моделі для досліджуваних економічних процесів.

Наведемо ілюстрацію виконання кроків алгоритму для даних табл. 2. В табл. 2 наведено дані котирувань акцій компанії «Камаз» за 132 торгові дні на момент закриття торгів за 2007 р. Дані отримано з архівів сайту http://stocks.investfunds.ru (по 11 даних у стовпці).

Вихідні дані поміщаємо у стовпець **A** на аркуші *Excel*, обчислюємо середнє значення $M(X_t) = 55{,}24$. Потім центруємо вихідний ряд, віднімаючи з нього середнє значення, і поміщаємо у стовпець **B** (рис. 2). Графік центрованого ряду зображено на рис. 3.

З рисунка видно, що в окремі інтервали часу спостерігалась тенденція, але потім відбувалося коригування, через що ряд змінював напрямок. Можна зробити висновок про те, що ряд, швидше за все, не є стаціонарним, через що потрібне його перетворення.

Досліджуємо розподіл ряду за допомогою пакета *AtteStat* [11]. У списку модулів вибираємо модуль «Проверка нормальности» (рис. 4) і заповнюємо поля у вікні, що з'явилося:

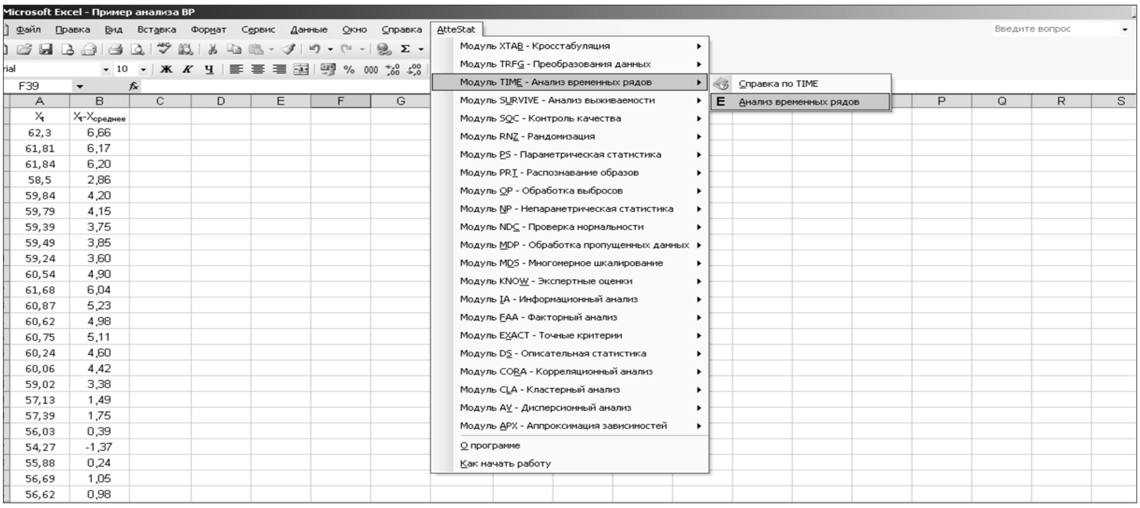

Рис. 2. Список модулів програми AtteStat

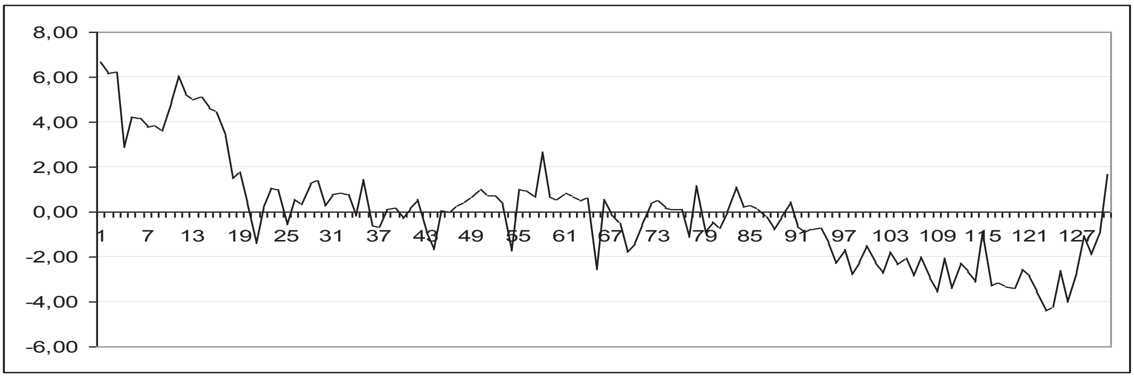

Рис. 3. Графік центрованого ряду

Таблиця 2

Дані котирувань акцій компанії «Камаз»

	1	2	3	4	5	6	7	8	9	10	11	12
1	62,3	60,87	56,69	55,45	55,71	56,57	55,53	56,82	55,44	54,15	52,23	52,09
2	61,81	60,62	56,62	57,09	55,64	56,31	55,12	54,74	56,07	53,36	53,37	51,27
3	61,84	60,75	55,07	55,05	55,89	58,28	53,85	55,2	54,95	52,89	52,98	51,4
4	58,5	60,24	56,17	54,95	56,1	56,28	54,25	54,92	54,72	53,88	52,55	53,03
5	59,84	60,06	55,96	55,77	56,29	56,21	55,1	55,61	54,88	53,28	54,73	51,62
6	59,79	59,02	56,93	55,78	56,66	56,46	56,04	56,74	54,9	53,6	52,34	52,96
7	59,39	57,13	57,05	55,34	56,38	56,31	56,13	55,84	54,22	52,79	52,49	54,58
8	59,49	57,39	55,9	55,79	56,37	56,13	55,81	55,94	53,34	53,66	52,3	53,73
9	59,24	56,03	56,4	56,2	55,98	56,26	55,74	55,73	53,98	52,67	52,24	54,81
10	60,54	54,27	56,46	54,77	53,91	53,08	55,77	55,39	52,84	52,06	53,09	57,32
11	61,68	55,88	56,39	53,97	56,62	56,19	54,54	54,84	53,39	53,59	52,73	57,34

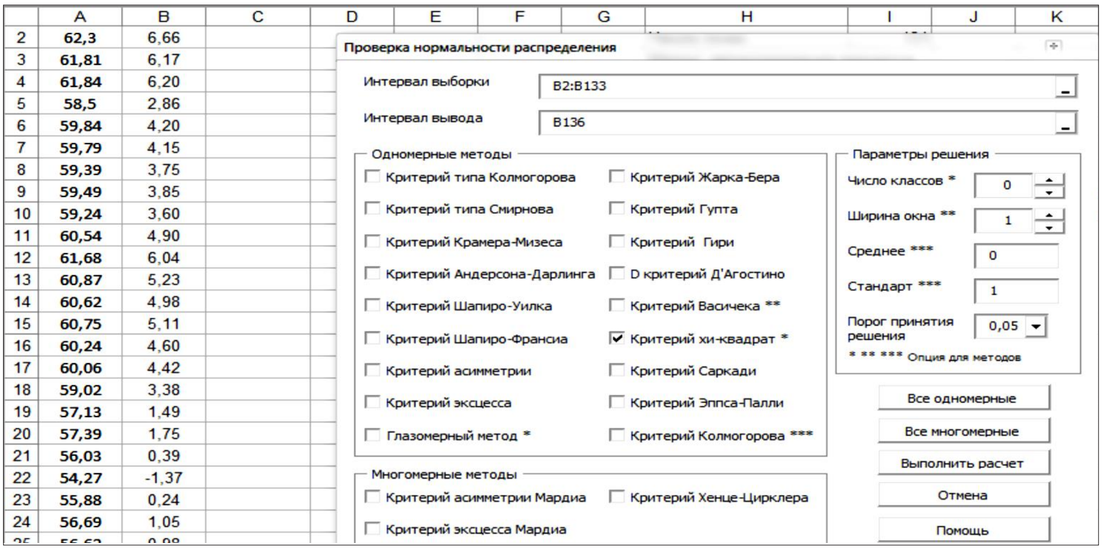

Рис. 4. Вікно модуля «Проверка нормальности»

	A	B	C	D	E
1	X_t	X_t-$X_{среднее}$			
2	62,3	6,66	Проверка нормальности распределения		
3	61,81	6,17	Выдача обычно включает:		
4	61,84	6,20	Статистика, Р-значение двустороннее, вывод		
5	58,5	2,86	Выбранное пороговое значение		
6	59,84	4,20	0,05		
7	59,79	4,15	Численность выборки		
8	59,39	3,75	131,00		
9	59,49	3,85	Критерий хи-квадрат Фишера		
10	59,24	3,60	Классы	8,00	
11	60,54	4,90	-3,68	11,00	8,98
12	61,68	6,04	-2,30	24,00	19,22
13	60,87	5,23	-0,92	24,00	28,87
14	60,62	4,98	0,46	47,00	30,43
15	60,75	5,11	1,84	7,00	22,51
16	60,24	4,60	3,22	6,00	11,69
17	60,06	4,42	4,60	8,00	4,26
18	59,02	3,38	5,97	4,00	1,09
19	57,13	1,49	Статистика критерия, р-значение		
20	57,39	1,75	36,03	0,00	Гипотеза о нормальности отклоняется

Рис. 5. Результат перевірки гіпотези про нормальність розподілу

– «Интервал виборки» – B2: B133 (зазначаємо діапазон, в якому розташований центрований ряд);

– «Интервал вывода» – B136 (зазначаємо будь-яку вільну комірку);

– «Одномерные методы» (зазначаємо «Критерий хи-квадрат»);

– «Параметры решения» (зазнчаємо «Порог принятия решения», а саме 0,05);

– натискаємо кнопку «Выполнить расчет».

Результат наведено на рис. 5. Гіпотеза про нормальний розподіл відхиляється, отже, не можна застосовувати параметричні методи.

Обчислимо тепер корелограму та періодограму. Вибираємо модуль «Анализ временных рядов» (рис. 6). У діалоговому вікні аналізу слід зазначити завдання, які необхідно виконати, діапазони вхідних та вихідних даних (рис. 6), натиснути кнопку «Выполнить расчет». Графіки корелограми і періодограми наведено на рис. 7, 8.

Графік коррелограми спадає експоненціальним чином на перших тридцяти лагах, отже, можна вважати, що досліджуваний ряд є стаціонарним типу *AR (p)*. За межі довірчого інтервалу виходять багато значень автокореляцій, але істотно відмінними від нуля є значення перших двох автокореляцій $\overline{p}_1 = 0,86, \overline{p}_2 = 0,81$ Тому параметр процесу *p* можна прийняти таким, що дорівнює двом, тобто $p = 2$.

Періодограма практично на всіх частотах має близькі до нуля значення. Висока інтенсивність виявляється на низьких частотах, що вказує на наявність шумів у цьому діапазоні. Такі шуми є суттєвими, їх не слід віддаляти методом згладжування ряду.

Протестуємо стаціонарність ряду за методом Дікі-Фуллера. Рівняння тесту Дікі-Фуллера для вихідного ряду є регресією перших різниць $\Delta X_t = X_t - X_{t-1}$ на X_{t-1} Обчислюємо перші різниці та поміщаємо у стовпець **C** (рис. 9). Скористаємось надбудовою Excel «Анализ данных». Для цього підемо таким шляхом: «Главное меню» → «Данные» → «Анализ данных» → «Регрессия». У вікні програми заповнюємо опції:

– «Входной интервал *Y*» – діапазон комірок, в яких розташовані перші різниці ΔX_t ;

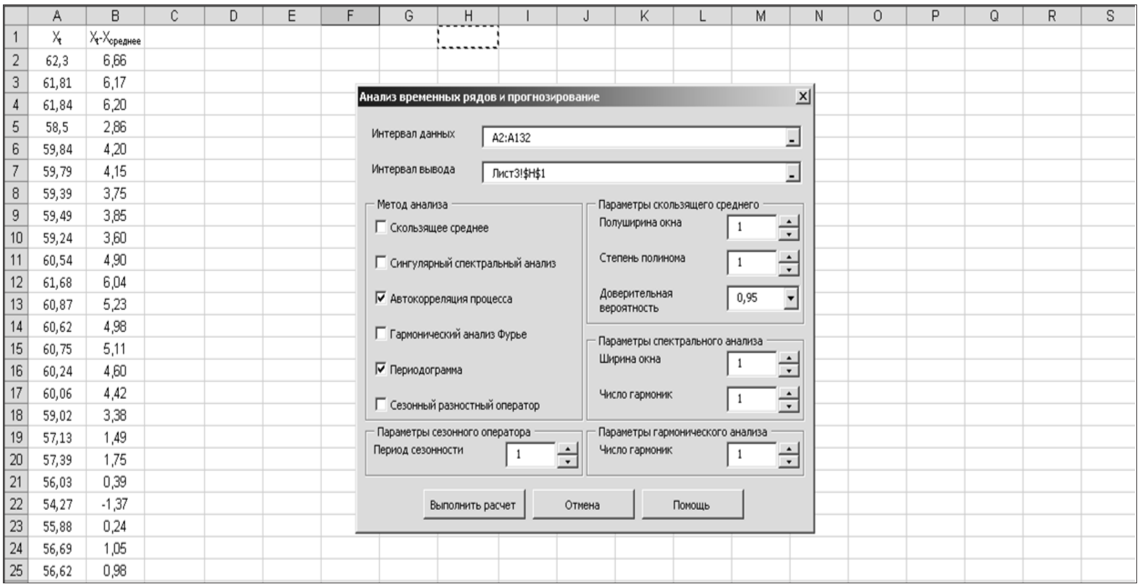

Рис. 6. Діалогове вікно аналізу часових рядів

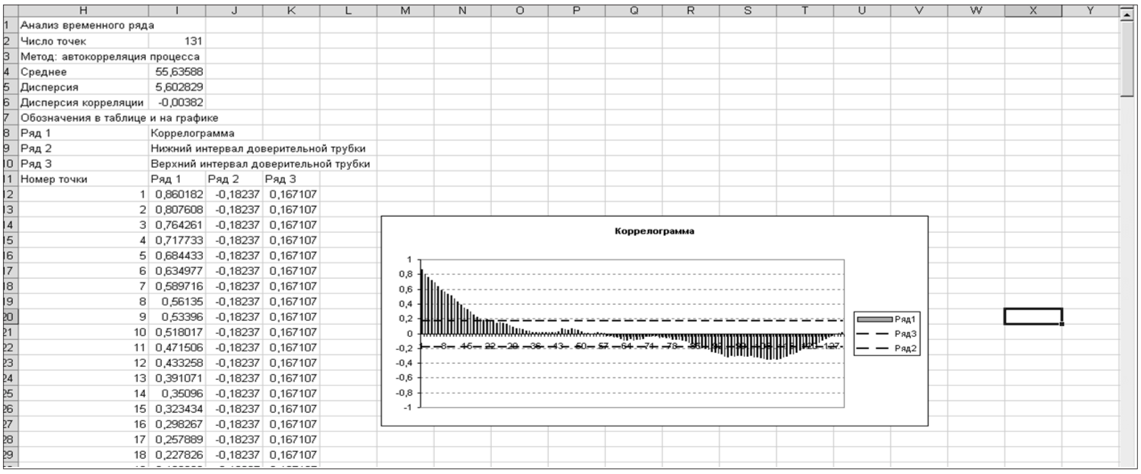

Рис. 7. Корелограма

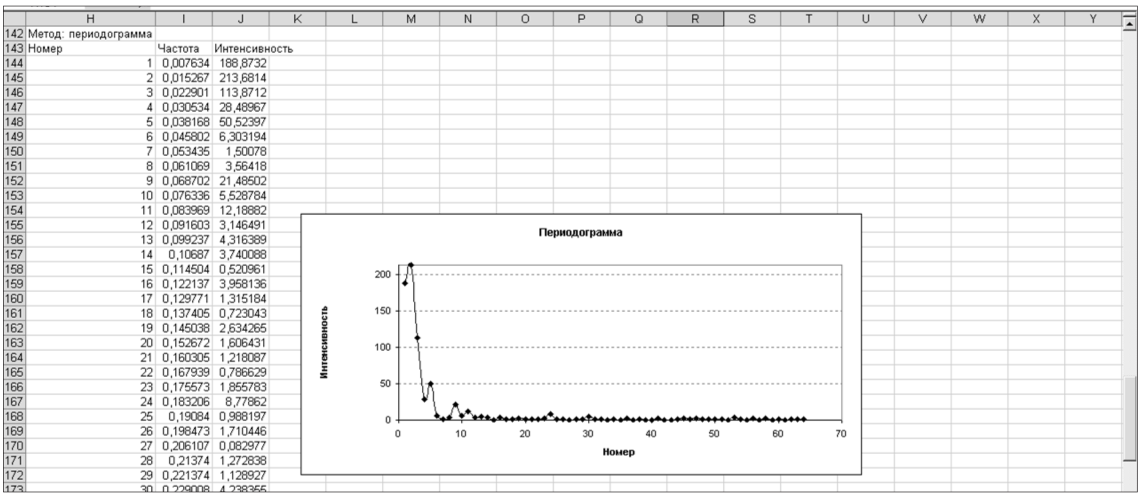

Рис. 8. Періодограма

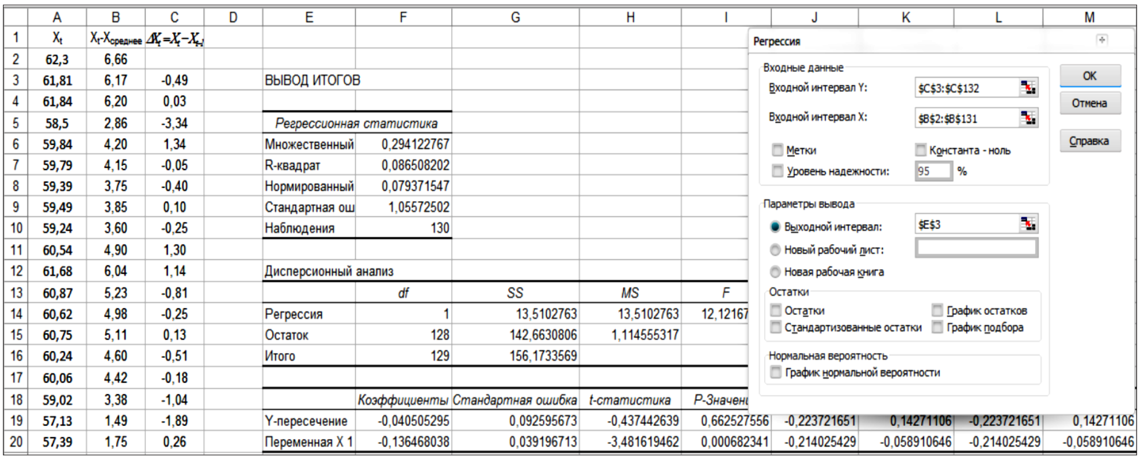

Рис. 9. Обчислення перших різниць, оцінювання регресії тесту Дікі-Фуллера

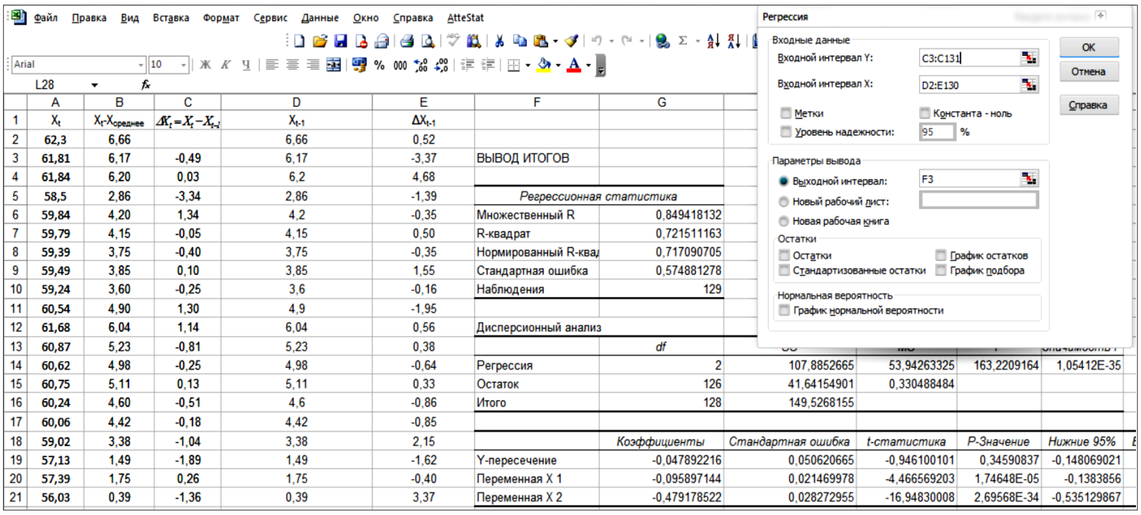

Рис. 10. Обчислення перших різниць та оцінка регресії тесту Дікі-Фуллера

– «Входной интервал *X*» – діапазон комірок, в яких розташовані дані ряду X_{t-1}, тобто дані вихідного ряду зі зміщенням на один лаг назад по відношенню до ΔX_t;

– «Уровень надежности» – 95%;

– «Выходной интервал» – будь-яка вільна комірка.

Після цього натискаємо ОК, отримуємо результатну таблицю (рис. 9).

З рисунка видно (останній рядок таблиці), що коефіцієнт при X_{t-1} з негативним знаком (коефіцієнт при x_1) і дорівнює -0,136 а значення *t*-статистики таке: $t_{факт}$ = -3,481. Це значення менше критичного значення *t*-статистики на рівні у 5% (табл. 1). Отже, можна стверджувати, що вихідний ряд з ймовірністю у 0,95 є стаціонарним.

Протестуємо тепер на стаціонарність авторегресійний процес *AR(2)*. Обчисляємо другі різниці та поміщаємо у стовпець **E**, а ряд X_{t-1} поміщаємо у стовпець **D** (рис. 10).

З рисунка видно (останній рядок таблиці), що коефіцієнт при X_{t-1} з негативним знаком (коефіцієнт при x_1) і дорівнює -0,095 а значення *t*-статистики таке: $t_{факт}$ = -4,466. Це значення менше критичного значення *t*-статистики на рівні у 5% (табл. 1). Отже, можна стверджувати, що авторегресійний ряд *AR(2)* з ймовірністю у 0,95 є стаціонарним.

Побудуємо авторегресійну модель *AR (2)*. Для обчислення коефіцієнтів моделі складемо систему [12]:

$$\begin{cases} \rho_1 = \phi_1 + \phi_2 \rho_1 \\ \rho_2 = \rho_1 \phi_1 + \phi_2 \end{cases}. \tag{5}$$

Підставивши значення $\rho_1 = 0,86$, $\rho_2 = 0,81$, отримаємо:

$$\begin{cases} 0,86 = \phi_1 + 0,86\phi_2 \\ 0,81 = 0,86\phi_1 + \phi_2 \end{cases}. \tag{6}$$

Вирішуючи цю систему, отримаємо $\phi_1 = 0,43$, $\phi_2 = 0,5$. Таким чином, модель приймає такий вигляд:

$$X_t = 0,43X_{t-1} + 0,5X_{t-2} \tag{7}$$

Перевіримо адекватність моделі за критеріями *AIC* і *BIC*. Інформаційні критерії *AIC* і *BIC* обчислюються за такими формулами:

$$AIC(p,q) = \ln \bar{\sigma}^2 + 2\frac{p+q}{T}, \tag{8}$$

$$BIC(p,q) = \ln \bar{\sigma}^2 + \ln T \frac{(p+q)}{T}, \tag{9}$$

де T – число спостережень, $\bar{\sigma}^2$ – залишкова дисперсія, $\bar{\sigma}^2 = \frac{e_i^2}{T-p-q}$, початковий $e_i = X_t - X_{прогн}$, X_t – початковий часовий ряд, $X_{прогн}$ – ряд, отриманий з використанням моделі *AR(2)*. Обчислимо за форму-

	A	B	C	D	E	F	G
1	X_t	X_t-$X_{среднее}$	$X_{прогн}$	e_i	e_i^2		
2	62,3	6,66	-				
3	61,81	6,17	-			σ^2 =	**1,0100**
4	61,84	6,20	5,99	0,22	0,05		
5	58,5	2,86	5,75	-2,89	8,36	**AIC(2,0)=**	**0,0408**
6	59,84	4,20	4,33	-0,13	0,02		
7	59,79	4,15	3,24	0,91	0,84	**BIC(2,0)=**	**0,0849**
8	59,39	3,75	3,89	-0,13	0,02		
9	59,49	3,85	3,69	0,16	0,03		
10	59,24	3,60	3,53	0,07	0,00		
11	60,54	4,90	3,48	1,43	2,04		
12	61,68	6,04	3,91	2,13	4,55		
13	60,87	5,23	5,05	0,18	0,03		
14	60,62	4,98	5,27	-0,29	0,08		
15	60,75	5,11	4,76	0,35	0,13		
16	60,24	4,60	4,69	-0,09	0,01		

Рис. 11. Обчислення критеріїв адекватності моделі

Рис. 12. Графіки вихідних і прогнозних даних

лою (7) прогнозні значення $X_{прогн}$, потім $e_i = X_t - X_{прогн}$ і e_i^2 (рис. 11).

З таблиці та рисунка видно, що статистики *AIC* і *BIC* практично дорівнюють нулю. Отже, модель адекватна результатам спостережень і може застосовуватися для прогнозування динаміки ціни акцій. На рис. 12 наведено графіки фактичної та прогнозованої ціни.

Прогнозований графік відображає тенденцію вихідного ряду на всьому часовому діапазоні, фактичні та прогнозні значення різняться незначно, оцінка середньоквадратичного відхилення така: $\bar{\sigma} = 1$.

Висновки. Наведений приклад свідчить про те, що в комплексі статистичні методи аналізу стаціонарних рядів є могутнім інструментом всебічного дослідження часових рядів, отримання точних, науково обґрунтованих результатів, але потребують адаптації під час використання їх у нестаціонарному випадку.

Список використаних джерел:

1. Березька К.М., Маслій В.В. Методологічні аспекти застосування моделі нечітких часових рядів для прогнозування податкових надходжень. Актуальні проблеми економіки. 2011. № 1. С. 227–235.
2. Андрієнко В.М. Теоретичні та методологічні аспекти моделювання статистичних рядів даних. Развитие науки в XXI веке: сборник статей VII Международной конференции. Ч. 2. Харьков. С. 68–76.
3. Андриенко В.М., Арсирий Е.А. Комплексная методология анализа, моделирования и прогнозирования временных рядов. Современный научный вестник. Серия «Математика». 2010. № 13 (95). С. 71–92.
4. Баклан І.В., Степанкова Г.А. Імовірнісні моделі для аналізу та прогнозування часових рядів. Искусственный интеллект. 2008. № 3. С. 505–515.
5. Баклан І.В., Степанкова Г.А. Про деякі нові особливості використання прихованих марковських моделей для аналізу та прогнозування часових рядів. Искусственный интеллект. 2010. № 4. С. 337–341.
6. Дзюба С.А. Технология прогнозирования временных рядов с сезонной составляющей. Менеджмент в России и за рубежом. 2010. № 3. С. 18–27.
7. Андриенко В.М., Арсирий Е.А. Интеллектуальный анализ временных рядов со стохастическим трендом. Восточно-европейский журнал передовых технологий. 2011. № 4/4 (52). С. 4–8.
8. Гур'янова Л.С., Клебанова Т.С., Сергієнко О.А., Прокопович С.В. Економетрика: навчальний посібник для студентів напряму підготовки «Економічна кібернетика» всіх форм навчання. Харків: ХНЕУ ім. С. Кузнеця, 2015. 384 с.
9. Dickey D.A., Fuller W.A. Distribution of the Estimators for Autoregressive Time Series with a Unit Root. Journal of the American Statistical Association. 1979. № 74. P. 427–431.
10. Тест Дики-Фулера // Википедия. URL: https://ru.wikipedia.org/wiki.
11. Гайдышев И.П. Моделирование стохастических систем: һуководство пользователя программы Attestat. 2013. URL: http://attestatsoft.narod.ru.
12. Андрієнко В.А., Андрієнко В.М. Аналіз фондових ринків. Одеса: Астропринт, 2011. 292 с.

МЕТОДИКА СТАТИСТИЧЕСКОГО АНАЛИЗА ЭКОНОМИЧЕСКИХ ВРЕМЕННЫХ РЯДОВ

Аннотация. Статья посвящена проблеме комплексного статистического анализа эмпирических данных сложных экономических систем. Предложена методика, которая уточняет особенности применения методов и заключается в логической взаимосвязанной последовательности основных этапов анализа. Такой подход может выявить характерные особенности эмпирических данных и дать адекватное представление об исследуемой системе. Приведен тестовый расчет на реальных данных. На основе результатов анализа определена, идентифицирована математическая модель динамики процесса.

Ключевые слова: временной ряд, стационарный ряд, тренд, периодограмма, коррелограмма, сезонная составляющая, циклическая составляющая.

METHOD OF STATISTICAL ANALYSIS OF ECONOMIC TIME RANKS

Summary. The article is devoted to the problem of complex statistical analysis of empirical data of complex economic systems. A method is proposed that specifies the specifics of the application of methods, and consists of a logically interrelated sequence of the main stages of analysis. Such an approach can reveal the characteristic features of empirical data and give an adequate idea of the system under study. The test calculation is based on real data. Based on the results of the analysis, a mathematical model of the dynamics of the process is identified, identified.

Key words: time series, stationary series, trend, periodogram, correlogram, seasonal component, cyclic component.