

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
імені ТАРАСА ШЕВЧЕНКА

АНТОНІНА ЄРІНА
ОЛЬГА МАЗУРЕНКО

СТАТИСТИЧНИЙ АНАЛІЗ ЧАСОВИХ РЯДІВ

НАВЧАЛЬНИЙ ПОСІБНИК

КИЇВ 2022

УДК 311.17(075.8)
Є71

Рецензенти:

Сергій Герасименко, докт. екон. наук, професор
(Національна академія статистики, обліку та аудиту)

Аліна Ходжаян, докт. екон. наук, професор
(Київський національний університет імені Тараса Шевченка)

*Затверджено вченою радою економічного факультету
Київського національного університету імені Тараса Шевченка
Протокол № 13 від 21 червня 2022 р.*

Єріна А., Мазуренко О.

Є71 **Статистичний аналіз часових рядів: навчальний посібник.** Київ. КНУШ, 2022. 164 с.
ISBN 978-966-

У навчальному посібнику розглядаються статистичні методи аналізу, моделювання та прогнозування перебігу складних соціально-економічних явищ і процесів, що у сучасному світі є невіддільним інструментом забезпечення обґрунтованості рішень на всіх рівнях управління — від невеликої фірми до національної економіки в цілому. Всеосяжність суспільного розвитку і специфіка динамічних процесів у різних сферах людської діяльності актуалізує набуття студентами навичок і вмінь аналізу часових рядів, прогнозування та верифікації прогнозів. Аналітичні можливості та межі застосування конкретних методів дослідження закономірностей динаміки ілюструються на прикладах, різних за соціально-економічним змістом та інформаційною базою з використанням сучасного програмного забезпечення.

Для студентів, магістрантів, аспірантів, науковців і практиків – для усіх, хто прагне опанувати сучасні методи статистичного аналізу закономірностей динаміки та прогнозування соціально-економічних процесів.

УДК 311.17(075.8)

*Розповсюджувати та тиражувати
без офіційного дозволу КНУШ забороняється*

ISBN 978-966-

© Єріна А., Мазуренко О., 2022
© КНУШ, 2022

ЗМІСТ

Передмова	5
Розділ 1. Методологічні засади статистичного аналізу часових рядів	7
1.1. Сутність і види часових рядів	7
1.2. Характеристики інтенсивності динаміки	12
1.3. Вимоги до інформації в часових рядах	20
<i>Питання і задачі для самопідготовки</i>	23
<i>Рекомендована література</i>	27
Розділ 2. Компонентний склад часових рядів	28
2.1. Детерміновані компоненти часового ряду	28
2.2. Тестування наявності тренду в структурі часового ряду ...	36
2.3. Аналіз властивостей випадкової компоненти	40
<i>Питання і задачі для самопідготовки</i>	44
<i>Рекомендована література</i>	46
Розділ 3. Аналіз і прогнозування тенденцій розвитку	47
3.1. Згладжування часових рядів	47
3.2. Трендові моделі на основі кривих зростання	50
3.3. Екстраполяція тренду як метод прогнозування	58
<i>Питання і задачі для самопідготовки</i>	64
<i>Рекомендована література</i>	68
Розділ 4. Аналіз і прогнозування закономірностей циклічного характеру	69
4.1. Аналіз циклічних закономірностей розвитку на основі S-кривих	69
4.2. Варіанти аналізу циклів оновлення	76
<i>Питання і задачі для самопідготовки</i>	82
<i>Рекомендована література</i>	84

Розділ 5. Аналіз сезонної компоненти	85
5.1. Амплітуда сезонних коливань	85
5.2. Сезонна декомпозиція часових рядів	92
5.3. Моделювання сезонної хвилі	96
<i>Питання і задачі для самопідготовки</i>	101
<i>Рекомендована література</i>	104
Розділ 6. Адаптивні методи прогнозування	105
6.1. Метод експоненційного згладжування	105
6.2. Модель ARIMA: ідентифікація та прогностні властивості	114
6.3. Застосування моделі ARIMA в аналізі інтервенцій	122
<i>Питання і задачі для самопідготовки</i>	127
<i>Рекомендована література</i>	132
Розділ 7. Множинна регресія на часових рядах	133
7.1. Особливості вимірювання кореляції часових рядів	133
7.2. Нелінійна регресія.	140
7.3. Регресія на панельних даних.	144
<i>Питання і задачі для самопідготовки</i>	148
<i>Рекомендована література</i>	151
Додатки	153

ПЕРЕДМОВА

*Для ученого цифри і таблиці статистичні
суть нині тим, чим для давніх астрологів були зорі небесні...*
Іван Франко

Статистичні методи аналізу та прогнозування динамічних процесів на основі часових рядів активно застосовують у сучасному світі в усіх сферах практичної діяльності і є невіддільним інструментом забезпечення обґрунтованості управлінських рішень на всіх рівнях управління — від невеликої фірми до національної економіки в цілому.

Оволодіння багатим арсеналом методів статистичного аналізу часових рядів з використанням комп'ютерних технологій, набуття навичок і вмінь обробки статистичної інформації, поданої часовими рядами, є важливою складовою професійної підготовки бакалаврів освітньої програми «Економічна аналітика та статистика». Саме цій меті підпорядковано курс «Статистичний аналіз часових рядів». Відповідно до програми названого курсу в навчальному посібнику розглядаються: методологічні принципи статистичного аналізу часових рядів, вимоги до інформації в часових рядах і їх компонентний склад, моделювання та прогнозування динамічних процесів, різнобічні моделі динаміки (трендові, сезонного ритму, повного циклу), критерії точності і надійності прогнозів, адаптивні методи короткострокового прогнозування, багатofакторні динамічні моделі.

У навчальному посібнику наголошується на використанні комп'ютерних технологій задля поглибленого аналізу, моделювання і прогнозування закономірностей розвитку та верифікації прогнозів. Унікальні можливості аналізу часових рядів і прогнозування надає інтегрована система обробки даних *Statistica*, модулі якої, передусім *Time series / Forecasting*, містять групи логічно пов'язаних між собою аналітичних процедур і методів аналізу динамічних процесів, а також потужну графічну систему для візуалізації даних.

Логічна структура аналізу часових рядів ілюструється на умовних прикладах соціально-економічного змісту. Для кожного типу моделей розглядаються принципи формування інформаційної бази, вибору процедур аналізу, інтерпретації результатів. Акцентуючи увагу студентів на параметрах моделей, таблиці з результатами аналізу і графіки подаються у стандартному вигляді англійською мовою. Специфікація моделей і змістовна інтерпретація параметрів розкривається в коментарях до таблиць і графіків.

Для систематизації і кращого засвоєння навчального матеріалу, а також набуття практичних навичок аналізу часових рядів до кожного розділу додаються питання і задачі для самопідготовки.

Вирішальний внесок в написання розділів 1–3 належить Ользі Мазуренко, розділів 4–7 — Антоніні Єріній.

МЕТОДОЛОГІЧНІ ЗАСАДИ СТАТИСТИЧНОГО АНАЛІЗУ ЧАСОВИХ РЯДІВ

- 1.1. Сутність і види часових рядів
- 1.2. Вимірювання інтенсивності динаміки
- 1.3. Вимоги до інформації в часових рядах

1.1. Сутність і види часових рядів

Соціально-економічні явища динамічні — постійно змінюються їх обсяги і структура, еволюціонують властиві їм характеристики. Аналіз закономірностей соціально-економічного розвитку — одне з важливих завдань статистики. Інформаційною базою аналізу динамічних закономірностей слугують часові ряди (англ. *Time series*).

Часовий ряд — це послідовність упорядкованих у хронологічному порядку значень статистичного показника, що характеризує зміну того чи іншого соціально-економічного явища в часі. Числа послідовності $y_1, y_2, y_3, \dots, y_n$ називають *рівнями часового ряду*. Підрядковий індекс $t = 1, 2, 3, \dots, n$ вказує на порядковий номер того інтервалу (моменту) часу, до якого належить значення показника, а кількість рівнів n називають довжиною часового ряду. За ознакою часу ряди поділяють на моментні та інтервальні. Рівень моментного ряду фіксує стан явища на певний момент часу, наприклад, золотовалютні резерви країни на початок року, кількість безробітних на початок кварталу тощо. Рівень інтервального ряду — це агрегований результат процесу, який залежить від тривалості часового інтервалу, наприклад, виробництво електроенергії за квартал, вилов риби за сезон тощо. Часовий ряд одного показника називають *одновимірним*, двох і більше, — *багатовимірним*.

За статистичною природою показників y_t вирізняють часові ряди абсолютних, відносних та середніх величин. Так, рівні *інтервального ряду* абсолютних величин — це сумарні результати процесу за відповідні проміжки часу: чим довший інтервал, тим більшим буде рівень ряду. Натомість рівень моментного ряду не залежить від тривалості періоду між сусідніми датами (момента-

ми), а сума рівнів ряду позбавлена економічного змісту. Часові ряди середніх і відносних величин називають *похідними*.

У часових рядах важливу інформацію несуть не лише значення окремих рівнів ряду, а також їхня послідовність. Саме характер послідовних змін значень y_t відбиває особливості руху процесу за певний період. Під впливом безлічі факторів довгострокової і короткострокової дії в одних рядах рівні протягом тривалого часу зростають або зменшуються з різною інтенсивністю, в інших зростання і зменшення y_t чергуються з певною періодичністю, наприклад, цикли економічної кон'юнктури. З року в рік більш-менш регулярно повторюються сезонні піднесення і спади (використання виробничих потужностей і робочої сили, попит на ринку споживчих товарів тощо). Окрім закономірних коливань рівнів, часовим рядам властиві також випадкові коливання, пов'язані з масовим процесом. Аналіз часового ряду дозволяє виявити закономірності динаміки та екстраполювати їх на майбутнє.

В табл. 1.1 представлені часові ряди абсолютних і відносних величин, на основі яких можна простежити закономірності динаміки глобальних демографічних показників за 1990—2025 рр.: збільшення чисельності і зростання середнього віку жителів планети; поступове затухання темпів нарощення чисельності населення, ущільнення територіального розселення і посилення процесу урбанізації.

Таблиця 1.1

НАСЕЛЕННЯ СВІТУ ЗА ДАНИМИ ООН

Рік	Чисельність населення, млн чол	Середньорічний приріст населення, %	Середній вік населення, років	Щільність населення, чол/кв км	Рівень урбанізації, % міського населення
1990	5 309 668	1,82	24	41	43,0
1995	5 735 123	1,55	25	44	44,8
2000	6 126 622	1,33	26	47	46,6
2005	6 519 636	1,25	27	50	49,1
2010	6 929 725	1,23	29	53	51,5
2015	7 349 472	1,18	30	57	53,8
2020	7 758 157	1,09	31	60	53,9
2025 (прогноз)	8 141 661	0,97	32	63	57,8

Рівні усіх наведених рядів з плином часу змінюються. Задля узагальнення притаманних часовому ряду властивостей, визначення типових характеристик розвитку і вирішення інших завдань аналізу використовують усереднені рівні. Зауважимо, що середня величина буде типовою характеристикою лише за умови однорідності часового ряду, коли причинний комплекс формування закономірностей розвитку більш-менш стабільний.

Середні рівні використовують, насамперед, для узагальнення коливальних рядів. Наприклад, аналізуючи динаміку сільськогосподарського виробництва, оперують не річними, а більш сталими середньорічними показниками за певні періоди. За допомогою середніх рівнів ряду динаміки забезпечують також порівнянність чисельника і знаменника різних за ознакою часу показників. Приміром, необхідно сформувати часовий ряд споживання фруктів у розрахунку на одну особу. Для цього за кожний рік необхідно обсяг споживання фруктів поділити на кількість населення. Очевидно, що кількість населення на той чи інший момент часу (моментний показник) непорівнянна з річним обсягом споживання фруктів (інтервальний показник). Щоб забезпечити порівнянність цих показників, слід обчислити середньорічну кількість населення.

Порядок розрахунку середнього рівня часового ряду залежить від статистичної структури показника. В інтервальному ряду абсолютних величин, рівні якого динамічно адитивні, використовують середню арифметичну просту:

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i,$$

де n — довжина ряду.

У моментному ряду, за припущення про рівномірну зміну показника між датами, середню розраховують як півсуму значень на початок і кінець періоду:

$$\bar{y} = \frac{y_1 + y_n}{2}.$$

Якщо кількість моментів $n > 2$ і між суміжними датами однакові інтервали, розрахунок виконують за формулою *середньої хронологічної*:

$$\bar{y} = \frac{y_1 + y_n + \sum_{i=2}^{n-1} y_i}{n-1}.$$

У моментних рядах із різними інтервалами між датами застосовують середню арифметичну зважену:

$$\bar{y} = \frac{1}{\sum D_t} \sum_{i=1}^m y_i D_t,$$

де D_t — інтервал часу між датами, m — кількість інтервалів.

Оскільки сума рівнів моментного ряду не має самостійного економічного змісту, то визначальною властивістю цієї середньої є сума добутоків кількості одиниць сукупності на тривалість періоду їх наявності чи функціонування. Наприклад, упродовж квітня змінювався денний запас сировини на складі фірми: з 1-го по 14-е запас був на рівні 40 т, з 15-го по 24-е — 37 т, з 25-го по 30-е — 45 т. Отже, загальний обсяг запасів сировини за місяць становив: $40 \times 14 + 37 \times 10 + 45 \times 6 = 1200$ тонно-днів, звідси середній денний запас $\bar{y} = 1200 : 30 = 40$ т.

Вирізняють часові ряди стаціонарні й нестаціонарні. Ряди, в яких рівні коливаються навколо певного середнього рівня, причому ні середня, ні дисперсія не виявляють істотних змін, називають *стаціонарними*. В разі, коли хоча б одна статистична характеристика ряду (середня чи дисперсія) змінюється, називають *нестационарними*. Приклади стаціонарного і нестаціонарного ряду ілюструють рис. 1.1 та 1.2.

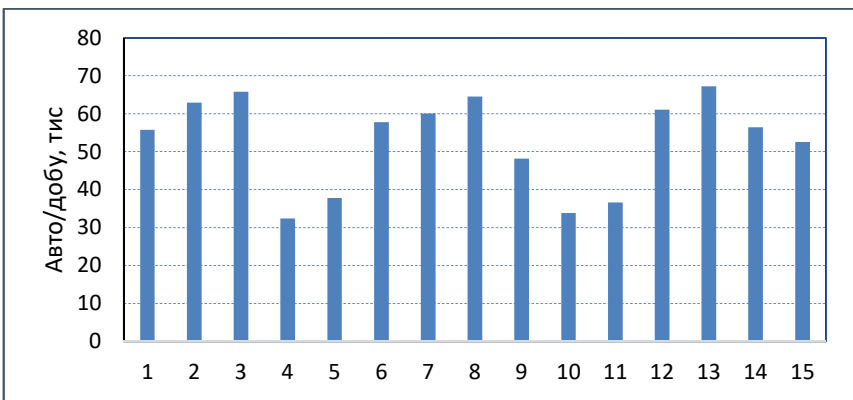


Рис. 1.1. Додова інтенсивність транспортного потоку по магістралі

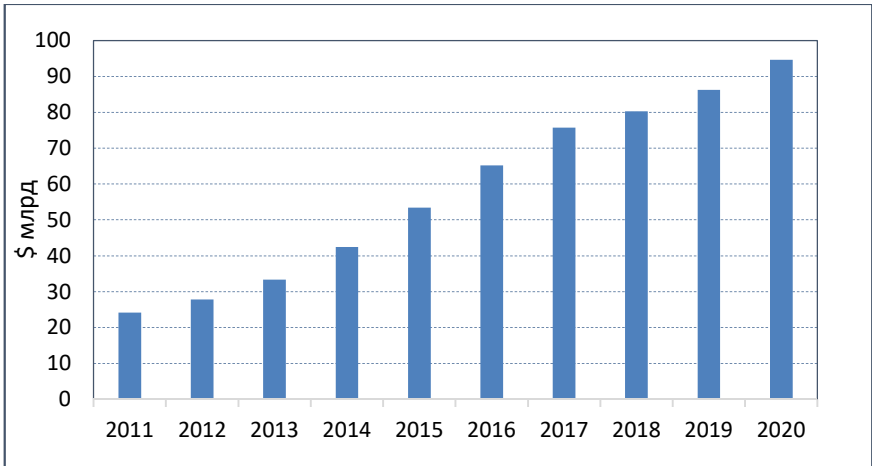


Рис. 1.2. Динаміка експорту послуг повітряного транспорту

В економічній практиці переважають нестационарні процеси. Для більшості з них характерна систематична зміна рівнів з нерегулярними коливаннями, коли піки і спади чергуються з різною інтенсивністю.

Характерною особливістю часового ряду є залежність послідовних рівнів. Значення показника y_t певною мірою залежить від минулих значень i , своєю чергою, впливає на майбутні. За незмінності комплексу умов формування динамічного процесу правомірно очікувати у майбутньому такі ж властивості і такий характер його розвитку, які були виявлені у минулому, а отже, можливим є статистичне прогнозування динамічних процесів.

В аналізі динаміки соціально-економічних процесів постають завдання:

- оцінити інтенсивність і сталість змін;
- виявити й описати характер змін за певний період часу, протягом якого процес еволюціонує, змінюється, прогресує;
- передбачити подальший рух процесу за межами ряду.

Поглиблений аналіз часових рядів соціально-економічних показників пов'язаний із більш складними узагальненнями, з визначенням основної тенденції, вимірюванням коливань і взаємозв'язків.

1.2. Вимірювання інтенсивності динаміки

На основі аналізу часових рядів можна простежити розвиток явища, показати його основні тенденції. Вибір відповідних прийомів та способів аналізу залежить від завдань дослідження та визначається характером вихідних даних.

Задля числової характеристики закономірностей і особливостей розвитку динамічного процесу на окремих етапах статистика:

- вимірює абсолютну і відносну швидкість зростання або зниження рівня за окремі проміжки часу;
- визначає узагальнюючі характеристики рівня та швидкості його зміни за той чи інший період;
- виявляє і чисельно характеризує основні тенденції розвитку об'єктів дослідження на окремих етапах;
- здійснює прогнозування розвитку об'єктів у майбутньому (екстраполяція) або відновлення пропущених даних (інтерполяція);
- оцінює точність прогнозу, здійснює його верифікацію;
- виявляє фактори, які спричиняють закономірні зміни і розвиток об'єктів дослідження.

Напрямок та інтенсивність змін описують низкою абсолютних і відносних характеристик динаміки, серед яких: абсолютні прирости, темпи зростання і темпи приросту (зменшення), абсолютне значення 1 % приросту інші. Розрахунок характеристик динаміки ґрунтується на порівнянні рівнів ряду. Базою порівняння для поточного рівня y_t може бути попередній рівень ряду y_{t-1} або будь-який віддалений у часі рівень. Якщо належить порівняти кілька послідовних рівнів ряду, база порівняння для них може бути постійною чи змінною. За постійну базу вибирається або початковий рівень ряду, або рівень, який вважають вихідним для розвитку явища, що вивчається. Характеристики динаміки, обчислені відносно постійної бази, називають *базисними*. Якщо кожний рівень ряду y_t порівнюється з попереднім y_{t-1} , характеристики динаміки називають *ланцюговими*. Схематично варіанти порівняння ілюструє рис. 1.3.

Зміст і взаємозв'язок характеристик інтенсивності динамічних процесів наведено в табл. 1.2.

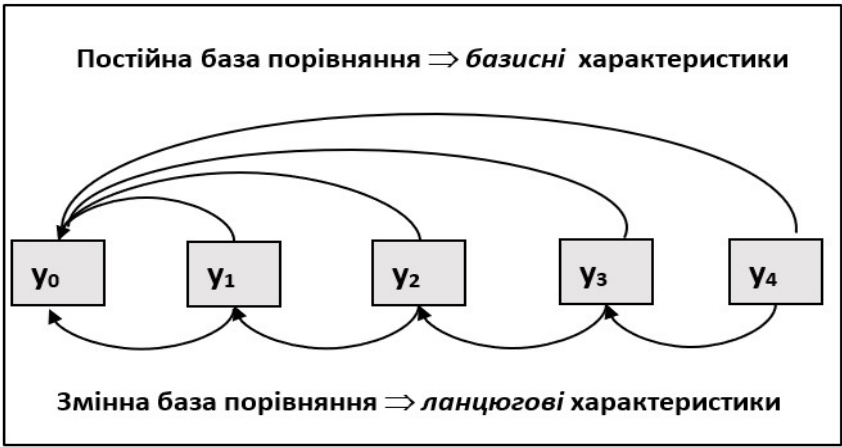


Рис. 1.3. Схеми порівняння при обчисленні характеристик інтенсивності динаміки

Таблиця 1.2

ЗМІСТ І ВЗАЄМВ'ЯЗОК ХАРАКТЕРИСТИК ДИНАМІКИ

Базисний	Зміст характеристики динаміки	Ланцюговий
$\Delta_t = y_t - y_0$	Абсолютний приріст характеризує абсолютний розмір збільшення (чи зменшення) рівня ряду	$\Delta_t = y_t - y_{t-1}$
$k_t = \frac{y_t}{y_0}$	Темп зростання (індекс) показує, у скільки разів y_t більший (менший) від бази порівняння	$k_t = \frac{y_t}{y_{t-1}}$
$T_t = \frac{\Delta_t}{y_0} 100 = 100(k_t - 1)$	Темп приросту (у %) показує, на скільки процентів y_t більший (менший) від бази порівняння	$T_t = \frac{\Delta_t}{y_{t-1}} 100 = 100(k_t - 1)$
$A_t = \frac{\Delta_t}{T_t} = \frac{y_0}{100}$, (для усіх t однакові)	Абсолютне значення 1 % приросту залежить від бази порівняння	$A_t = \frac{\Delta_t}{T_t} = \frac{y_t - y_{t-1}}{100 \left(\frac{y_t - y_{t-1}}{y_{t-1}} \right)} = \frac{y_{t-1}}{100}$

Абсолютний приріст Δ_t визначають як різницю між двома рівнями ряду, він характеризує збільшення (чи зменшення) рівня ряду y_t за певний часовий інтервал в абсолютному вимірі. Знаки «+», «-» свідчать про напрям зміни рівня ряду.

Темп зростання, далі *Індекс* k_t — це кратне відношення двох рівнів. Він показує, у скільки разів рівень y_t більший (менший) від рівня, взятого за базу порівняння: при збільшенні рівня ряду $k_t > 1$, при зменшенні — $k_t < 1$. Темпи зростання виражаються коефіцієнтами або процентами.

Ланцюгові абсолютні прирости Δ_t і темпи зростання k_t відображають відповідно абсолютну й відносну швидкість динаміки. Вони взаємозв'язані. Якщо подати $y_t = y_{t-1} + \Delta_t$, то

$$k_t = \frac{\bar{y}_{t-1} + \Delta_t}{\bar{y}_{t-1}} = 1 + \frac{\Delta_t}{\bar{y}_{t-1}}.$$

Величину $\frac{\Delta_t}{y_{t-1}}$ називають відносним прискоренням або *темпом приросту* і позначають символом T_t . Ця величина функціонально пов'язана з темпом зростання, але на відміну від останнього завжди виражається процентами:

$$T_t = 100 (k_t - 1).$$

Темп приросту показує, на скільки процентів рівень y_t більший (менший) від бази порівняння, взятої за 100 %. Залежно від напряму зміни рівня ряду темп приросту, як і абсолютний приріст, може бути додатним або від'ємним числом. Якщо, скажімо, $k_t = 102\%$, то $T_t = 102 - 100 = 2\%$; при $k_t = 97\%$ темп приросту $T_t = 97 - 100 = -3\%$.

Очевидно, що при стабільній абсолютній швидкості темпи приросту будуть зменшуватися, а при стабільних темпах приросту абсолютна швидкість, навпаки, буде зростати. Стабільні темпи зростання можливі за умови прискорення абсолютної швидкості.

Співвідношення абсолютного приросту і темпу приросту визначає вагомість кожного процента приросту — *абсолютне значення 1 % приросту* ($A_t, \%$), як видно, це відношення становить 0,01 рівня, взятого за базу порівняння. Розрахунок абсолютного значення 1 % приросту має економічний зміст лише для ланцюгових темпів приросту, для базисних темпів приросту значення A_t % однакові.

Очевидно, що ланцюгові й базисні характеристики динаміки взаємопов'язані:

а) сума ланцюгових абсолютних приростів дорівнює кінцевому базисному:

$$\sum_1^n \Delta_t = \sum_1^n (y_t - y_{t-1}) = y_n - y_0.$$

б) добуток ланцюгових темпів зростання дорівнює кінцевому базисному:

$$k_1 \cdot k_2 \dots k_n = \prod_1^n k_t = K_n = \frac{y_n}{y_0}.$$

Щодо темпів приросту, то вони не мають таких властивостей, як абсолютні прирости чи темпи зростання. Ланцюгові й базисні темпи приросту співвідносяться через темпи зростання.

Середня абсолютна та відносна швидкість динаміки

Із часом змінюються не лише рівні динамічних рядів, а й обчислені на їх основі абсолютні прирости й темпи зростання. Задля узагальнення притаманних динамічному ряду властивостей, визначення типових характеристик розвитку і вирішення інших завдань аналізу використовують середні величини — середні абсолютні прирости і середні темпи приросту (зменшення).

Середній абсолютний приріст (абсолютна швидкість динаміки) обчислюється діленням загального приросту за весь період на довжину цього періоду у відповідних квантах часу (рік, квартал, місяць тощо):

$$\bar{\Delta} = \frac{\bar{y}_n - y_0}{n} = \frac{\sum_{t=1}^n \Delta_t}{n}.$$

Наприклад, за 2018 р. автомобільним транспортом перевезено 2072 тис. т вантажів, за 2021 р. — 2126 тис. т. Середньорічний абсолютний приріст обсягів перевезення вантажів за 2018 — 2021рр. становить

$$\bar{\Delta} = (2126 - 2072) : 3 = 18 \text{ тис. т.}$$

При обчисленні середнього темпу зростання враховують правило складних процентів, за якими нагромаджується приріст на приріст, тобто змінюється відносна швидкість динаміки. Тому *середній темп зростання* обчислюють за формулою середньої геометричної з ланцюгових темпів зростання:

$$\bar{k} = \sqrt[n]{k_1 \cdot k_2 \dots k_n} = \sqrt[n]{\prod_{t=1}^n k_t},$$

де n — кількість темпів зростання за однакові інтервали часу.

Наприклад, за останні три роки невпинно зростали тарифи на автоперевезення. Темпи зростання становили: 2019 р. — 1,03; 2020 р. — 1,08; 2021 р. — 1,05. Середньорічний темп зростання

$$\bar{k} = \sqrt[3]{1,03 \cdot 1,08 \cdot 1,05} = 1,053 \text{ або } 105,3 \text{ \%}.$$

З огляду на взаємозв'язок ланцюгових і базисних темпів зростання, формулу середньої геометричної можна записати так:

$$\bar{k} = \sqrt[n]{K_n} = \sqrt[n]{\frac{y_n}{y_0}} \text{ або } \bar{k} = \sqrt[n-1]{\frac{y_n}{y_1}}.$$

Скажімо, вартість споживчого кошика за три роки зросла на 12,5 %. Середньорічний темп зростання становить

$$\bar{k} = \sqrt[3]{1 + 0,125} = \sqrt[3]{1,125} = 1,04.$$

Тобто щороку споживчий кошик дорожчав у середньому на 4 %.

Розрахунок середнього темпу зростання можна виконувати на основі логарифмів:

$$\lg \bar{k}_t = \frac{1}{n} \sum \lg k_n \text{ або } \lg \bar{k}_t = \frac{1}{n} (\lg y_n - \lg y_0).$$

Наприклад, прямі іноземні інвестиції в галузь у 2018 р. становили 172,0, у 2021 р. — 213,7 млн. дол. США. Десяткові логарифми цих величин 2,2355 і 2,3304 відповідно. Звідси $\lg \bar{k} = \frac{1}{3} (2,3304 - 2,2355) = 0,0316$. Антилогарифм для 0,0316 дорівнює 1,0754, отже, середньорічний темп зростання становить

$\bar{k} = 1,0754$, тобто за означений період прямі іноземні інвестиції у галузь щорічно зростали у середньому на 7,54 %.

Коли середній темп зростання розраховують на основі темпів зростання по періодах різної довжини, використовують середню геометричну зважену:

$$\bar{k} = \sqrt[t]{k^{t1} \cdot k^{t2} \dots k^{tn}}.$$

Приміром, середній темп зростання податкових надходжень до бюджету за останні три роки становив 1,07, за два попередні роки — 1,05. Звідси за п'ять років середньорічний темп зростання податкових надходжень був на рівні 106,2 %:

$$\bar{k} = \sqrt[5]{1.07^3 \cdot 1.05^2} = 1,062.$$

Таким чином, середній темп зростання можна обчислити на основі:

- ланцюгових темпів зростання k_i ;
- кінцевого (за весь період) темпу зростання K_n ;
- кінцевого y_n і базисного y_0 рівнів ряду

У дод. X наведено фрагмент таблиці обчислення середніх темпів зростання за 2-5 років.

Середні темпи приросту визначають на основі середніх темпів зростання

$$\bar{T} = 100(\bar{k} - 1).$$

При інтерпретації середньої абсолютної чи відносної швидкості динаміки необхідно вказувати часовий інтервал, до якого належать середні, та часову одиницю вимірювання (рік, квартал, місяць, доба тощо).

Порівняння інтенсивності динамічних процесів

На різних етапах розвитку зміна рівнів ряду може відбуватися з різною швидкістю — швидше чи повільніше. Порівнюючи однойменні характеристики динаміки, що склалися на відповідних етапах, можна оцінити *прискорення* чи *уповільнення* розвитку. Якщо тривалість етапів однакова, порівнюються базисні або середні характеристики динаміки, якщо різна, — лише середні характеристики динаміки. Так, різниця абсолютних приростів

$\delta = \Delta_2 - \Delta_1$ вказує на абсолютне прискорення ($\delta > 0$) або уповільнення ($\delta < 0$) динаміки.

Для позитивних абсолютних приростів можна визначити відносне прискорення (уповільнення) шляхом співвідношенням відповідних абсолютних приростів $k_{\nabla} = \Delta_t : \Delta_{t-1}$. Значення $k_{\nabla} > 1$ свідчить про прискорення, $k_{\nabla} < 1$ — про уповільнення. Якщо інтервали часу не однакові, порівнюють середні абсолютні прирости відповідних інтервалів.

При порівнянні інтенсивності динаміки двох паралельних рядів: одного показника за різними об'єктами (групами, регіонами, країнами) або різних показників одного об'єкта співвідносяться темпи зростання (базисні або середні). Такі співвідношення називають *коефіцієнтами випередження*:

$$k_{\text{вип}} = k' : k''$$

де k' , k'' — відповідно темпи зростання порівнюваних динамічних рядів.

Для наочності та зручності тлумачення цього коефіцієнта дільником k' є більший за значенням темп зростання. Наприклад, за звітний рік залишки коштів на банківських рахунках фізичних осіб зросли: в національній валюті на 37 %, в іноземній валюті — на 58 %. Коефіцієнт випередження зростання залишків коштів в іноземній валюті відносно залишків коштів у національній валюті становить $k_{\text{вип}} = 1,58 : 1,37 = 1,16$.

Порівняння інтенсивності динаміки двох взаємопов'язаних показників, з яких один розглядається як причина x , а другий — як наслідок y , здійснюють на основі темпів приросту. Співвідношення темпів приросту називають емпіричним *коефіцієнтом еластичності*:

$$\gamma = \frac{T_y}{T_x}$$

Цей коефіцієнт показує, на скільки процентів змінюється y зі зміною x на 1 %. Наприклад, ціна кави на світовому ринку зросла на 5 %, а попит зменшився на 2 %. Цінова еластичність попиту на

цей товар $\gamma = \frac{-2}{+5} = -0,4$, тобто зі зростанням цін на 1 % попит на товар зменшується на 0,4 %.

Оцінювання структурних зрушень

Внутрішня структура будь-якої статистичної сукупності динамічна. Скажімо, в межах однієї компанії змінюється структура активів і капіталу, вікова й професійна структура персоналу, структура витрат і зобов'язань, асортимент і якість продукції, що виробляється, тощо. Статистичними характеристиками структури є частки, тож аналіз структурних зрушень ґрунтується на порівнянні часток.

Зміна часток окремих складових сукупності свідчить про структурні зрушення. Інтенсивність структурних зрушень оцінюється за допомогою лінійного \bar{l}_d або квадратичного σ_d коефіцієнтів структурних зрушень, розрахунок їх базується на відхиленнях часток:

$$\bar{l}_d = \frac{\sum_1^m |d_{j1} - d_{j0}|}{m}; \quad \sigma_d = \sqrt{\frac{\sum_1^m (d_{j1} - d_{j0})^2}{m}},$$

лінійний квадратичний

де d_{j0} та d_{j1} — частки відповідно базисного та поточного періодів, m — кількість складових сукупності.

Так, за даними табл. 1.3 структура ресурсів комерційного банку змінилася: зросла частка власного капіталу, зменшилися частки залученого і позикового капіталу. Лінійний коефіцієнт структурних зрушень становить $\bar{l}_d = \frac{12}{3} = 4$, тобто структура банківських ресурсів зрушилася в середньому на 4 процентних пункти. Завдяки своїм математичним властивостям квадратичний коефіцієнт структурних зрушень дещо більший — $\sigma_d = \sqrt{\frac{56}{3}} = 4,3$ пункти.

Таблиця 1.3

ДО РОЗРАХУНКУ КОЕФІЦІЄНТА СТРУКТУРНИХ ЗРУШЕНЬ

Джерела формування ресурсів	Базисний період, d_0	Поточний період, d_1	Відхилення часток, $d_1 - d_0$	Модулі відхилень, $ d_1 - d_0 $	Квадрати відхилень, $(d_1 - d_0)$
Власні	8	14	6	6	36
Залучені	59	55	-4	4	16
Позикові	33	31	-2	2	4
Разом	100	100	0	12	56

1.3. Вимоги до інформації в часових рядах

Важливою умовою обґрунтованості і правильності висновків за результатами аналізу часових рядів є забезпеченість *порівнянності даних*, що характеризують динаміку досліджуваного об'єкта за певний період. Порівнянність означає, що рівні часових рядів повинні мати однакові одиниці вимірювання, однакову періодичність обліку, однаковий ступінь агрегування, обчислюватися за тією самою методикою. Причини непорівнянності даних різні, в економіці та соціології найбільш поширені такі:

- зміни в методології обліку та розрахунку показника, зокрема використання різних одиниць вимірювання;
- зміна цін в розрахунках вартісних показників;
- зміна кола охоплених дослідженням об'єктів;
- зміна адміністративно-територіальних меж об'єкта дослідження;
- структурні зміни об'єкта дослідження, наприклад, укрупнення компанії шляхом злиття або поглинання;
- різні критичні моменти реєстрації подій чи тривалість періодів, до яких належать зареєстровані дані тощо.

У більшості випадків вдається усунути непорівнянність шляхом перерахунку значень досліджуваних показників, застосовують також формальні процедури зімкнення часових рядів.

Порівнянність даних — найважливіша умова обґрунтованості і правильності висновків за результатами аналізу часових рядів. Забезпечується вона на етапах збирання та обробки даних. У разі, коли дані непорівнянні, їх приводять до порівнянного виду прямим перерахунком за допомогою порівняльних цін, курсів валют, коефіцієнтів переведення, індексів тощо. Використовують також стандартизовану структуру. Іноді удаються до спеціального прийому зімкнення уривистих часових рядів — «статистичних ключів». Суть цього прийому полягає в об'єднанні двох і більше рядів в один зімкнений ряд. Припустимо, внаслідок реорганізації будівельної компанії у квітні місяці непорівнянними виявилися помісячні обсяги виконаних робіт у I-му півріччі. Подолати уривистість ряду можна двома способами:

- Перший спосіб передбачає перерахунок абсолютних показників. Коефіцієнтом перерахунку є співвідношення рівнів того періоду, у якому відбулися зміни: ($k = Y_3 : y_3$). Помноживши рівні

першого ряду на цей коефіцієнт, отримаємо єдиний зімкнений (порівнянний) часовий ряд абсолютних величин за весь період.

- Другий спосіб відносних рівнів, коли за базу порівняння для кожного ряду беруть один і той самий період, протягом якого відбулися зміни (у наведеному прикладі це третій рівень). Два ряди відносних величин динаміки об'єднуються в один.

Схеми зімкнення уривистих часових рядів представлені в табл. 1.4.

Таблиця 1.4

СХЕМИ ЗІМКНЕННЯ УРИВИСТИХ ЧАСОВИХ РЯДІВ

t	y _t до змін	Y _t після змін	Зімкнений часовий ряд	
			абсолютних величин	відносних величин
1	y ₁		$Y_1 = k y_1$	$y_1 : y_3$
2	y ₂		$Y_2 = k y_2$	$y_2 : y_3$
3	y ₃	Y ₃	Y ₃	1,0
4		Y ₄	Y ₄	$Y_4 : Y_3$
5		Y ₅	Y ₅	$Y_5 : Y_3$

Другою передумовою є *однорідність даних*, тобто відсутність аномальних, нетипових для даного часового ряду спостережень, так званих *викидів*. Формально аномальність виявляється як несподіваний стрибок чи спад з подальшим поступовим відновленням попереднього рівня. Часовий ряд може містити декілька аномальних значень. Зазвичай їх можна виявити візуально, за допомогою графічного представлення часових рядів, що унаочнює рис. 1.4.

На етапі попереднього аналізу часових рядів необхідно встановити тип викидів. Найчастіше це наслідки технічних помилок під час збирання, запису та передачі інформації. Такого роду викиди підлягають усуненню, оскільки застосування статистичних методів до «засміченої» інформації призводить до зміщення оцінок і спотворює результати аналізу. Серед способів усунення аномальних значень найпростішим є заміна їх істинними чи збереженими за певною методикою значеннями.

З іншого боку аномальні рівні в часових рядах можуть відображати об'єктивний, реальний розвиток процесу і свідчити про відхилення від загальної тенденції розвитку, наприклад, через

«стрибок» курсу національної валюти чи різного роду форс-мажорні обставини. Аномальні значення такого типу не підлягають усуненню, їх необхідно вивчати.

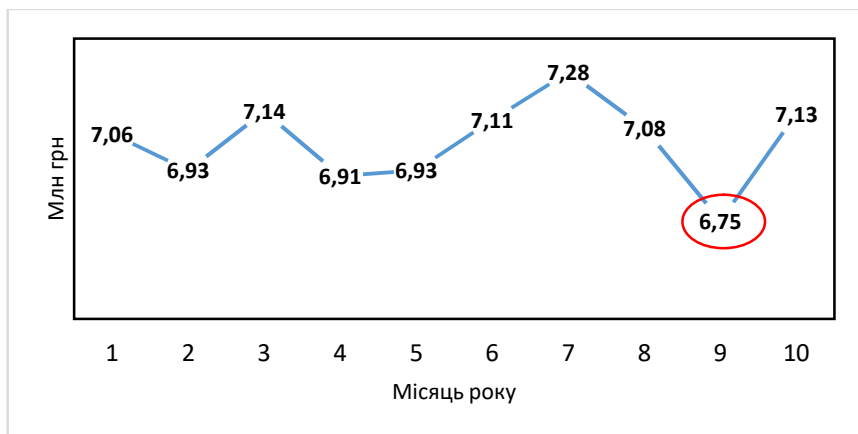


Рис. 1.4. Помісячна динаміка дебіторської заборгованості компанії

Для виявлення та заміни аномальних значень існує низка аналітичних методів. Розглянемо один з них — критерій Ірвіна, згідно з яким аномальним вважається рівень y_t , який відрізняється від попереднього y_{t-1} на величину λ_t :

$$\lambda_t = \frac{|y_t - y_{t-1}|}{\sigma_y},$$

де λ_t — статистика критерія Ірвіна, $t = 2, 3 \dots n$,
 σ_y — стандартне відхилення:

$$\sigma_y = \sqrt{\frac{\sum (y_t - \bar{y})^2}{n-1}}; \quad \bar{y} = \frac{\sum y_t}{n}.$$

Порядок тестування гіпотези про відсутність викидів у часовому ряду здійснюють за стандартною процедурою: розраховані значення λ_2, λ_3 і т. д. порівнюються з критичними $\lambda_{1-\alpha}$. Коли $\lambda_t > \lambda_{1-\alpha}$, гіпотеза відхиляється і рівень ряду y_t визнається аномальним. В додатку II наведені критичні значення λ_t для рівня істотності $\alpha = 0,05$ і довжини часового ряду n . Як свідчать дані, чим довший ряд, тим менше значення $\lambda_{1-\alpha}$.

Розглянемо застосування критерія Ірвіна на прикладі помісячної динаміки дебіторської заборгованості компанії (рис. 1.5). Середньомісячний рівень заборгованості $\bar{y} = 7,03$ млн грн., стандартне відхилення $\sigma_y = 0,192$, значення λ_t для кожного місяця представлені в табл. 1.4. Порогове значення критерія Ірвіна для $n = 10$ становить $\lambda_{1-\alpha} = 1,5$.

Таблиця 1.5

**ЧАСОВИЙ РЯД ДЕБІТОРСЬКОЇ ЗАБОРГОВАНОСТІ
І ВІДПОВІДНІ РІВНЯМ РЯДУ ЗНАЧЕННЯ КРИТЕРІЯ ІРВІНА**

t	1	2	3	4	5	6	7	8	9	10
y_t млн грн	7,06	6,93	7,14	6,91	6,93	7,11	7,28	7,08	6,75	7,13
λ_t	—	0,68	1,09	1,19	0,10	0,94	0,88	1,04	1,71	1,97

Порогове значення критерія перевищують два рівні ряду: $y_9 = 6,75$ і наступний $y_{10} = 7,13$. Проте на рис. 1.4 рівень y_{10} не видається аномальним. Насправді амплітуда зміни наступного за аномальним рівня ряду y_{t+1} може бути просто фоновим відображенням зміни саме аномального рівня.

Третьою передумовою порівняльного аналізу часових рядів є *повнота даних*, тобто достатня кількість спостережень. Ця передумова зумовлена тим, що закономірність може виявитися лише за наявності мінімально допустимого обсягу спостережень. На допустиму довжину часових рядів накладає обмеження також використання певного математичного апарату. Наприклад, щодо періодичних коливань бажано мати інформацію щонайменше за три повних періоди коливання, при використанні регресійного аналізу довжина часового ряду має кратно перевищувати кількість параметрів моделі, нелінійні динамічні процеси, яким властива довготривала пам'ять, потребують довшого ряду тощо.

Отже, для успішного вивчення динамічних процесів важливо забезпечити порівнянність даних, в разі наявності аномальних значень необхідно їх усунути чи замінити, довжину часового ряду формувати з урахуванням конкретних цілей дослідження. Відповідність наявної інформації вищезазначеним вимогам перевіряється на етапі попереднього аналізу часових рядів.

Питання і задачі для самопідготовки

1. Вкажіть, які з наведених показників подаються у формі динамічного ряду — моментного чи інтервального:

— кількість суб'єктів підприємницької діяльності — платників єдиного податку — на початок кожного року;

— сума податкових надходжень від платників єдиного податку за кожний рік;

— заборгованість комерційних банків акціонерам і бюджету станом на 1 квітня 2022 р.;

— дефіцит державного бюджету за 2021 рік.

2. На початок року в країні було зареєстровано 216 товарних і товарно-сировинних бірж, на кінець року — 232. За рік було укладено угод на суму 3024 млрд. грн. Визначте обсяг угод, укладених у середньому однією біржею за рік.

3. Згідно з даними звітів транспортної компанії залишки оборотних коштів в запасах товарно-матеріальних цінностей становили, млн грн.: на 1.01 — 200; 1.04 — 221; 1.07 — 209; 1.10 — 212; на 1.01 нового року — 220. Визначте середньоквартальний залишок оборотних коштів.

4. Зафіксована зміна товарних запасів компанії протягом вересня, млн. грн. Запас станом на: 01.09 — 39; 07.09 — 38; 15.09 — 44; 25.09 — 37.

Визначте середньоденний товарний запас.

5. Емісія цінних паперів суб'єктами фондового ринку за перший квартал поточного року становила, млрд. грн:

Емітенти цінних паперів	На початок року	Абсолютний приріст випуску акцій за		
		січень	лютий	березень
Органи місцевого самоврядування	21,5	5,9	7,1	4,2
Юридичні особи	9,4	1,7	1,2	1,8

Визначте по кожному емітенту цінних паперів:

— абсолютний приріст емісії акцій за квартал і обсяг емісії на початок квітня;

— темп приросту емісії акцій за квартал і абсолютне значення 1 % приросту.

6. Видатки на оборону у Зведеному бюджеті країни становили:

Рік	2019	2020	2021
Видатки, млрд грн	3,5	5,2	6,3

Визначте:

— абсолютний приріст і темп приросту видатків на оборону за 2020-2021 рр.;

— середньорічний абсолютний приріст;

— середньорічний темп приросту.

7. Динаміка кількості промислових підприємств, що впроваджували інновації, характеризується даними:

Інновації	2019 рік	Щорічна зміна, %	
		2020	2021
Освоювання виробництва нових видів продукції	1250	-30	-20
Впровадження нових технологічних процесів	360	+30	+10

Визначте по кожному напрямку інновацій:

— темп приросту (зменшення) кількості підприємств, що впроваджували інновації, за 2020-2021рр;

— абсолютний приріст (зменшення) кількості підприємств, що впроваджували інновації, за два роки і абсолютне значення 1 % змін.

8. Динаміку надходжень страхових платежів наведені в таблиці ланцюгові характеристики. Використовуючи їх взаємозв'язки, за кожний рік обчисліть доходи страхових компаній за цим видом страхування, абсолютні прирости, темпи приросту і абсолютне значення 1 % приросту.

Рік, № з/п	Доходи, млн. грн.	Ланцюгові характеристики динаміки			
		Абсолютний приріст, млн. грн	Темп зростання	Темп приросту, %	Абсолютне значення 1 % приросту, млн. грн.
1	21	×	×	×	×
2		4			
3			1,26		
4				40	
5					
6				48	0,65

9. Динаміка вилову риби підприємствами рибної промисловості характеризується даними:

Рік	2010	2015	2021
Виллов риби, тис. т	304	250	144

Визначте характеристики динаміки вилову риби:

— абсолютні прирости (зменшення) за 2011—2015 рр. і за 2016—2021 рр.

— абсолютне і відносне уповільнення динаміки.

10. За поточний період відбулися зміни в структурі ресурсів банківської установи:

Джерела формування банківських ресурсів	% до підсумку	
	базисний період	поточний період
Власні	8	14
Залучені	59	55
Позикові	33	31
Разом	100	100

Оцініть інтенсивність структурних зрушень, зробіть висновки.

11. Динаміка вкладів населення в банках регіону характеризується даними, млн. грн.:

Вклади	На кінець 2019 р.	Абсолютний приріст за рік	
		2020	2021
У національній валюті	70	24	18
В іноземній валюті	54	22	34

За допомогою коефіцієнта випередження визначте, у якій валюті населення здійснювало заощадження більш активно.

12. За даними обстежень домогосподарств у минулому році їхні доходи (з урахуванням трансфертів) зросли на 25 %. Прирости витрат становили, %:

— на придбання товарів довготривалого використання — 27;

— на збільшення фінансових активів — 18.

Для кожного виду витрат визначте коефіцієнт еластичності від доходу, поясніть їх зміст. Зробіть висновки.

13. Що означає порівнянність даних в аналізі часових рядів? Як її забезпечити?

14. Маємо дані про випуск продукції холдинг-компанією, організаційна структура якої у 2019 р. була реорганізована. Застосовуючи процедуру зімкнення ряду за допомогою «статистичних ключів», проаналізуйте зміну обсягу випуску продукції у 2021 р. порівняно з 2017 р.

Випуск продукції, млн грн.	2017	2018	2019	2020	2021
До реорганізації	180	207	225		
Після реорганізації			315	340	378

15. Поясніть зміст терміну «аномальність даних». Як перевірити наявність аномальних даних у часовому ряду?

16. Динаміка попиту фінансової компанії на кредитні ресурси (y_t , млн євро) характеризується даними:

Тиждень, t	1	2	3	4	5	6	7	8	9
y_t , млн євро	47	39	54	56	75	62	68	54	65

Динаміку попиту на кредитні ресурси зобразить графічно. За допомогою тесту Ірвіна перевірте наявність у часовому ряду аномальних рівнів, зробіть висновки.

Рекомендована література

1. Єріна А.М., Пальян З.О. Статистика: підручник. Київ, 2010. С. 256. — 290.

2. Статистика: структурно-логічні схеми та задачі: навч. посіб. / Єріна А.М. та ін. Київ. 2007. С. 233—264.

3. Introduction to the Fundamentals of Time Series Data and Analysis. URL:

<https://www.aptech.com> > blog (Last accessed: 30.05.2022).

- 2.1. Детерміновані компоненти часового ряду
- 2.2. Тестування наявності тренду в структурі часового ряду
- 2.3. Аналіз властивостей випадкової компоненти

2.1. Детерміновані компоненти часового ряду

Часові ряди характеризують фізичні, технічні, соціально-економічні процеси розвитку, яким властиві дві взаємопов'язані риси: динамічність та інерційність. Динамічність проявляється зміною рівнів і варіації показників, що характеризують процес, інерційність — сталістю механізму формування процесу, напрямку та інтенсивності динаміки протягом певного часу. Поєднуючи ці риси, часовий ряд у будь-який момент t містить залишки минулого, основи сучасного і зародки майбутнього.

Саме діалектична єдність мінливості й сталості, динамічності й інерційності формує закономірність розвитку. Під впливом безлічі факторів довгострокової і короткострокової дії в одних рядах рівні протягом тривалого часу зростають або зменшуються з різною інтенсивністю, в інших зростання і зменшення рівнів чергуються з певною періодичністю (наприклад, одинадцятирічні цикли градових опадів, зумовлені циклами сонячної активності). З року в рік більш-менш регулярно повторюються сезонні піднесення і спади (використання виробничих потужностей і робочої сили, попит на ринку споживчих товарів тощо). Окрім закономірних коливань рівнів, часовим рядам притаманні також випадкові коливання, пов'язані з масовим процесом.

У практиці статистичного аналізу закономірностей динаміки кожний рівень часового ряду y_t розглядають як результат комбінації різних складових, а часовий ряд в цілому за структурно-утворюючими елементами умовно поділяють на чотири *компоненти*:

- довгострокову, детерміновану часом еволюцію D_t ;
- циклічні коливання різної природи C_k ;
- сезонні коливання S_k , які відображають внутрішньорічну динаміку і більш-менш регулярно повторюються рік у рік;
- нерегулярні, випадкові флуктуації e_t (шум).

Більшість динамічних процесів у межах періоду з відносно стабільними умовами розвитку виявляють певну закономірність зміни рівнів ряду — загальну тенденцію (*trend*). Тенденція визначає напрямок розвитку, довгострокову еволюцію, зумовлену впливом множини постійно діючих, специфічних для кожного процесу факторів і умов розвитку. Одним рядам притаманна тенденція до зростання, іншим — до зниження рівнів. Зростання чи зниження рівнів часового ряду відбувається по-різному: рівномірно, прискорено чи уповільнено.

Поряд із довготривалими тенденціями у часових рядах часто виникають періодичні, більш-менш регулярні коливання:

— циклічні коливання, коли в процесі розвитку відбувається постійна і послідовна повторюваність певних стадій (фаз), наприклад, цикли ділової активності, демографічні, інвестиційні та інші цикли;

— сезонні коливання, пов'язані зі зміною пори року і мають виражену річну періодичність; найбільш виразно проявляються у сільському господарстві, на транспорті, у побутовому обслуговуванні, у торгівлі і споживанні тощо.

Регулярні, періодичні коливання можна подати схематично у вигляді синусоїди: протягом певного часу рівні ряду зростають, досягають певного максимуму, потім спадають до певного мінімуму, знову зростають і т. д. Неодмінними атрибутами циклів є наявність двох крайніх точок: найвищої точки піднесення — *пік* (вершина) і найнижчої точки падіння — *дно*. Відтинок часу від початку і до закінчення циклу (від одного піку до іншого, або від одного дна до іншого) визначає тривалість циклу. За тривалістю цикли поділяють на три види:

- малі, або короткі цикли Кітчина, тривалістю до 5 років;
- середні, виробничі або бізнес-цикли Жугляра, тривалістю 8—10 років; до середніх відносять і цикли С. Кузнеця, або так звані будівельні цикли;
- довгострокові цикли, або довгі хвилі М. Кондратьєва тривалістю близько 50 років.

Очевидно, триваліші цикли вбирають у себе коротші. Скажімо, середньострокові цикли Жугляра включають декілька

циклів Кітчина. Взаємозв'язок циклів різної тривалості ілюструє рис. 2.1.

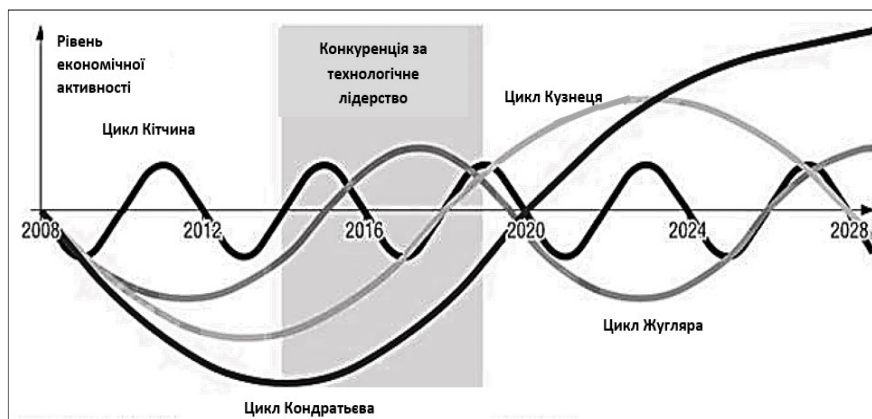


Рис. 2.1. Інтерференція циклів згідно з теорією довгих хвиль М. Кондратьєва

Циклічність є об'єктивною закономірністю економічного розвитку. Це складне й багатогранне явище. Простежується чіткий взаємозв'язок між амплітудою коливання економічних циклів та різними факторами кон'юнктури. Так, наприклад, відхилення від рівноваги першого порядку (попит і пропозиція) приводить до коротких хвиль, рівноваги другого порядку (переливання капіталу в нове обладнання, машини, модернізацію виробництва) досягають коливаннями середніх хвиль, рівновага третього порядку, яка стосується глобальних змін щодо виробничих технологій, сировинної бази, енергетичних питань, спричиняє довгі хвилі. Саме в цей період відбуваються науково-технічні винаходи, відкриття, зміни технологій, які мають найсуттєвіший вплив на соціально-економічне життя суспільства. Кожна наступна фаза довгого циклу є результатом кумулятивних процесів попередньої.

Слід зауважити, що, окрім циклічних, мають місце і нециклічні коливання економіки. Останні спричиняються, зокрема, диспропорціями між окремими сферами і галузями економіки, а тому отримали назву *структурних криз*. Серед структурних криз енергетичні, продовольчі, сировинні, фінансові (валютні), екологічні та інші. Особливістю сучасного етапу розвитку світової економіки є посилення впливу на неї саме нециклічних криз.

На відміну від циклічної компоненти, яка описує тривалі періоди відносного піднесення і спаду з різною амплітудою коливань, *сезонна компонента* відображає притаманну природним явищам і людській діяльності повторюваність протягом року з регулярною періодичністю. В аналізі часових рядів сезонність сама по собі може бути предметом дослідження, а може розглядатись як джерело короткострокових коливань часового ряду, які підлягають фільтрації.

Існують також короткострокові цикли в межах місяця чи навіть доби (інтенсивність транспортних потоків, споживання електроенергії тощо).

Тренд, циклічну та сезонну компоненти називають систематичними, детермінованими чи регулярними, їх визначають за певним правилом як функцію часу t . Оскільки часові ряди економічних показників часто виявляються надто «короткими» для ідентифікації циклічної компоненти формальними методами, в аналізі таких рядів тренд і циклічну компоненту часто об'єднують в одну тренд-циклічну компоненту D_t (рис. 2.2). Така умовна конструкція дозволяє залежно від мети дослідження вивчати тренд-циклічну, регулярну складову часового ряду, фільтруючи сезонні коливання, або вивчати сезонні коливання, елімінуючи тренд-циклічну компоненту. Виявляючи та прогнозуючи кожен компоненту ряду окремо, завдяки певній композиції отриманих результатів, можна спрогнозувати майбутні значення часового ряду.

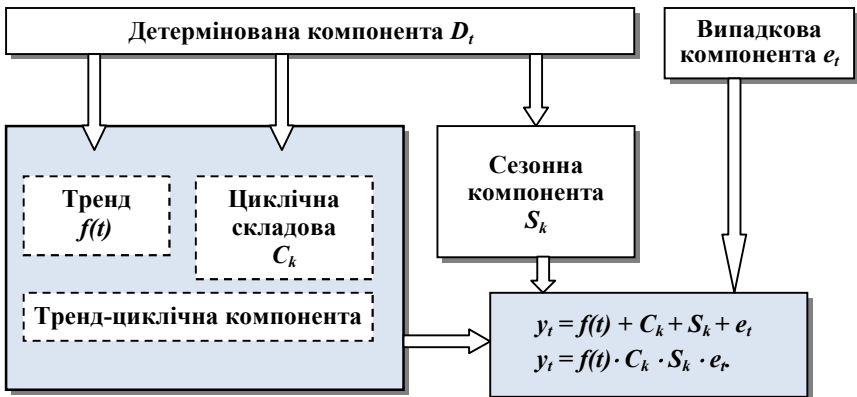


Рис. 2.2. Компоненти часового ряду

Відхилення між фактичними рівнями часового ряду і детермінованою компонентою $e_t = y_t - D_t$ представляють нерегулярну,

випадкову компоненту. Випадкова компонента є обов'язковою складовою будь-якого часового ряду, що пояснюється складним переплетенням параметрів економічної системи, впливом на них великої кількості взаємопов'язаних факторів, які часом діють у різних напрямках.

Залежно від способу взаємодії окремих компонент ряду між собою застосовують адитивну, мультиплікативну або змішану модель часового ряду. В *адитивній* моделі вплив тренд-циклічної компоненти, сезонних коливань та випадкових величин підсумовується:

$$y_t = D_t + S_k + e_t.$$

У *мультиплікативній* моделі зв'язок окремих компонент часового ряду виражається добутком:

$$y_t = D_t \cdot S_k \cdot e_t.$$

У *змішаній*, адитивно-мультиплікативній, моделі одні складові ряду підсумовуються, інші помножуються:

$$y_t = D_t \cdot S_k + e_t.$$

За характером поєднання перелічених компонент вирізняють щонайменше чотири типи часових рядів:

- до першого типу належать ряди з трендом без періодичних коливань;
- часові ряди другого типу, окрім тренду, мають сезонні коливання, амплітуда яких із часом змінюється;
- до третього типу належать часові ряди без тренду і періодичних коливань;
- часові ряди четвертого типу найскладніші, вони містять тренд, циклічні та сезонні коливання.

Випадкові флуктуації накладаються на тренд і періодичні коливання в часових рядах будь-якого типу.

Отже, найскладніший (за комбінацією компонент) часовий ряд можна подати як суму тренд-циклічної, сезонної та випадкової компонент. Тренд-цикли як і сезонна компонента за своєю природою можуть бути адитивними чи мультиплікативними. Для адитивного тренд-циклу характерні стабільні абсолютні прирости, наприклад, обсяги продажу певного товару зростають щороку на 10 млн грн; для мультиплікативного — стабільні темпи приросту, наприклад, обсяги продажу товару щороку зростають на 10 %. Для адитивної сезонності характерне стабільне збільшення амплітуди коливань на певну величину, для мультиплікативної

— збільшення амплітуди на певний процент. В адитивному випадку ряд матиме постійні сезонні коливання незалежно від зміни рівнів ряду, у мультиплікативному, навпаки, сезонні коливання змінюються залежно від рівнів ряду. Відхилення фактичних рівнів часового ряду від тренд-циклічної компоненти розглядається як стаціонарний випадковий процес.

Загальне уявлення про динаміку досліджуваного процесу і компонентний склад часового ряду можна отримати на основі візуального аналізу графіків часового ряду. Насамперед це стосується таких компонентів ряду, як тренд і сезонні коливання. Суть декомпозиції часового ряду на складові розглянемо на прикладі помісячного виїзного потоку туристів з регіону за шість років (табл. 2.1).

Таблиця 2.1

ВИЇЗНИЙ ПОТІК ТУРИСТІВ З РЕГІОНУ ЗА 2015—2020 рр., тис. осіб

Рік Місяць	2015	2016	2017	2018	2019	2020
Січень	5567	5457	5556	5759	6327	7227
Лютий	5219	4963	5266	5848	5946	6957
Березень	5886	5656	6663	7263	6738	9059
Квітень	6632	6576	6719	6970	6102	8935
Травень	7092	6835	7201	7253	8358	8735
Червень	7502	8188	8633	7059	9239	9955
Липень	7616	7758	8148	7196	9111	9431
Серпень	7594	7718	8164	8513	9154	9371
Вересень	7110	8278	8854	9208	9573	10 823
Жовтень	6264	6425	6911	7238	8123	8259
Листопад	5183	5396	5820	6193	7097	7730
Грудень	5289	6187	6409	7283	8107	8515

Аналіз часового ряду починається з виділення трендової компоненти. Її присутність неважко помітити, проаналізувавши графік часового ряду. Як правило, для економічних даних дуже типовим є повільне зростання чи падіння протягом тривалого періоду часу. На рис. 2.3 чітко простежується тренд до зростання і сезонні коливання (щороку більшість туристичних поїздок припадає на літні місяці і другий помітно менший пік наприкінці року).

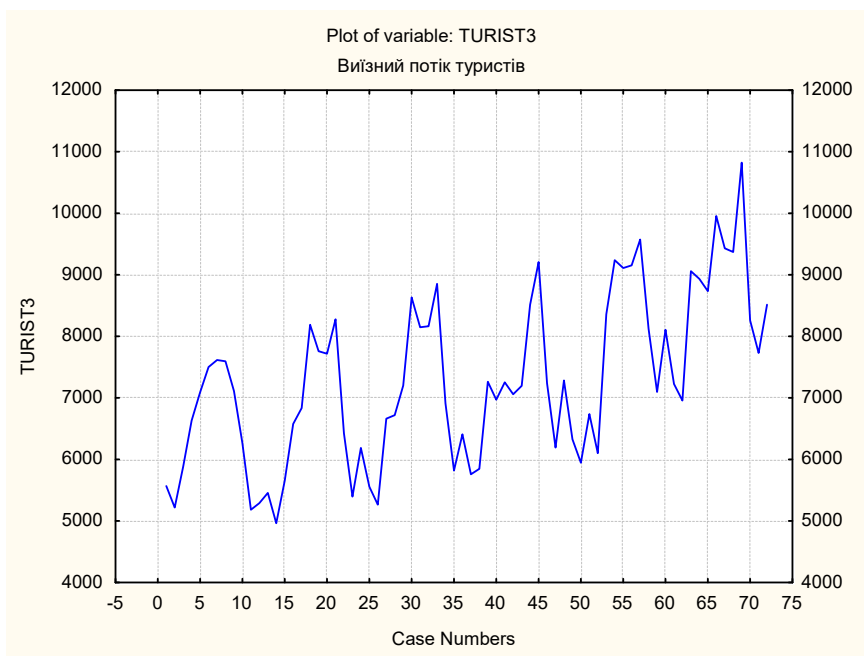


Рис. 2.3. Помісячна динаміка кількості виізних туристів

Опишемо тенденцію ряду за допомогою ковзної середньої з інтервалом згладжування $m = 12$. Ковзна середня фільтрує сезонні коливання і врівноважує випадкові, завдяки чому характер динаміки проявляється чіткіше, що видно на рис. 2.4. Траєкторія зміни ковзної середньої вказує на три етапи розвитку виізного туризму: поступове збільшення потоку на початку досліджуваного періоду змінюється певною стабілізацією і останнім часом демонструє новий етап зростання.

Нерегулярні коливання ряду формуються під впливом різних факторів короткострокової дії і розглядаються як стаціонарний випадковий процес, до якого можна застосувати відповідні методи аналізу. Як видно на рис. 2.5, нерегулярна компонента часового ряду виізного потоку туристів максимально наближається до стаціонарного процесу. Якщо випадкова компонента не є стаціонарною, здійснюють певні перетворення (трансформації) ряду, аби привести її до стаціонарного виду.

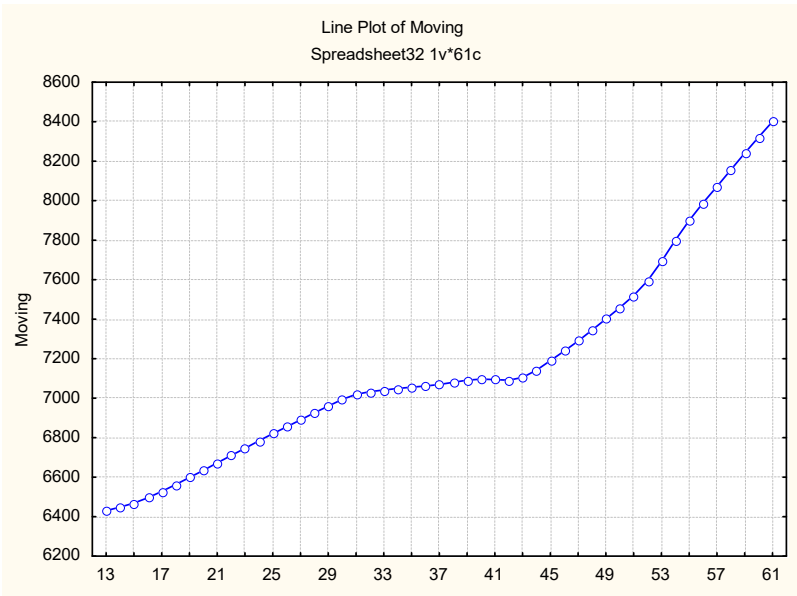


Рис. 2.4. Тренд-циклічна компонента часового ряду

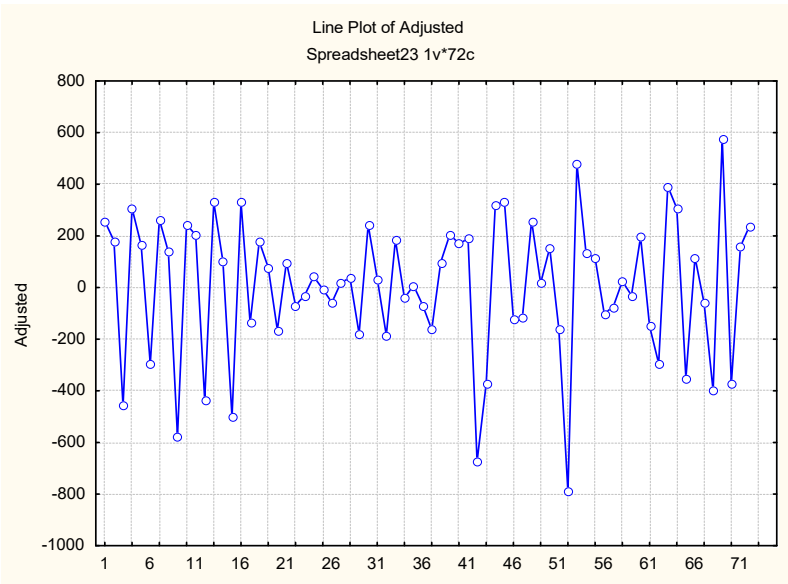


Рис. 2.5. Нерегулярні коливання часового ряду

Таким чином, на стадії проведення графічного аналізу можна дослідити компонентний склад часових рядів, а також зробити перші кроки до вибору моделі динаміки і подальшого прогнозування.

В процесі аналізу соціально-економічних процесів постають завдання:

- виявити й описати характер змін показника за певний період часу, протягом якого явище еволюціонує, змінюється, прогресує;

- оцінити інтенсивність і сталість змін;

- передбачити подальший рух процесу за межами ряду.

Згладжування та декомпозиція часових рядів слугують методологічною базою прогнозування динаміки соціально-економічних процесів. Послідовність етапів дослідження динамічних процесів така:

- візуалізація процесу і стислий опис характерних особливостей ряду, визначення структурних компонент часового ряду;

- виокремлення й аналіз детермінованої складової ряду;

- вибір статистичної моделі, яка б адекватно описала тренд-циклічну компоненту ряду;

- фільтрація сезонних (циклічних) коливань ряду;

- аналіз випадкової складової, перевірка адекватності моделі.

У контексті системного підходу до аналізу і прогнозування взаємопов'язаних часових рядів виникає необхідність оцінювання їх кореляції і крос-кореляції.

2.2. Тестування наявності тренду в структурі часового ряду

Часовим рядам, що відображають розвиток соціально-економічних процесів, властива, як правило, тенденція до зростання або спадання. Проте не завжди тенденція простежується досить чітко, тож візуальний аналіз часового ряду не може виявити її наявність. Постає необхідність з'ясувати, чи існує тенденція у такому часовому ряду взагалі. Основні підходи до вирішення даної задачі засновані на статистичному тестуванні гіпотези про випадковість ряду. Серед них метод перевірки різниці середніх, критерій квадратів послідовних різниць (критерій Аббе), критерій Фостера-Стюарта, критерій серій на базі медіани, критерій інверсій, метод поворотних точок, кореляція рангів та ін. Не зважаючи на те, що названі методи засновані на різних підходах, всі вони виконують функцію тестування гіпотези про випадковість часового ряду. Розглянемо методику перевірки гіпоте-

зи такого типу на прикладі критерія Фостера-Стюарта. Саме цей критерій, на відміну від інших, надає надійні результати щодо наявності в часовому ряду тренду як середніх, так і дисперсії.

В основі реалізації методу Фостера-Стюарта лежить принцип порівняння поточного рівня ряду з попередніми у прямому і зворотному напрямках (round-trip). У процесі порівняння для всіх $t = 2 \div n$ визначають дві величини: одна з них є індикатором того, що поточне значення y_t перевищує всі попередні рівні ряду, позначимо її u_t , друга, навпаки, вказує, що всі попередні рівні перевищують поточне значення, позначимо її l_t :

$$u_t = \begin{cases} 1, \text{ якщо } y_t > y_{t-1} > \dots > y_1, \\ 0, \text{ для інших;} \end{cases}$$

$$l_t = \begin{cases} 1, \text{ якщо } y_t < y_{t-1} < \dots < y_1, \\ 0, \text{ для інших;} \end{cases}$$

Порядок надання значень цим величинам показано в табл. 2.2 на прикладі динаміки урожайності соняшника. Сумарні їх значення формують дві статистики критерія Фостера-Стюарта: різниця призначена для виявлення тренду в середніх

$$D = \sum_2^n u_t - \sum_2^n l_t$$

сума для виявлення тренду в дисперсіях

$$S = \sum_2^n u_t + \sum_2^n l_t.$$

Статистика S набуває значень на відрізку $[0; n - 1]$, статистика D — на відрізку $[-(n - 1); n - 1]$.

Таблиця 2.2

ДО РОЗРАХУНКУ СТАТИСТИК КРИТЕРІЯ ФОСТЕРА-СТЮАРТА

Порядковий номер року	y_t , ц з 1 га	u_t	l_t	Порядковий номер року	y_t , ц з 1 га	u_t	l_t
1	12,3	0	0	9	19,1	0	0
2	16,3	1	0	10	13,6	0	0
3	10,7	0	1	11	17,8	0	0
4	17,8	1	0	12	18,3	0	0
5	16,4	0	0	13	20,6	0	0
6	18,7	1	0	14	16,4	0	0
7	17,3	0	0	15	21,5	1	0
8	20,9	1	0	Разом для $n = 15$	X	5	1

За даними таблиці означені статистики становлять:

$$D = 5 - 1 = 4; S = 5 + 1 = 6.$$

Перевірка гіпотез про випадковість часового ряду, тобто про відсутність тренду середньої і стабільність дисперсії заснована на двох статистиках критерія t -Стьюдента:

$t_D = \frac{D}{\sigma_d}$	$t_S = \frac{S-\mu}{\sigma_S}$,
$\mu = \sqrt{2 \sum_2^n \frac{1}{t}}; \sigma_d = \sqrt{\mu}$	$\sigma_S = \sqrt{\mu - 4 \sum_2^n \frac{1}{t^2}}$

Для зручності розрахунків значення величин, σ_D і σ_S табульовані (дод. III). Згідно додатку для $n = 15$ маємо:

$$\mu = 4,636; \sigma_D = 2,153; \sigma_S = 1,521.$$

Звідси значення статистик t -критерія:

$$t_D = 4 : 2,153 = 1,85;$$

$$t_S = (6 - 4,636) : 1,521 = 0,896.$$

Критичне значення t -критерія з рівнем істотності $\alpha = 0,10$ і числом ступенів свободи ($\nu = n - 1 = 14$) становить $t_{1-0,10}(14) = 1,76$ (див. дод. I).

Коли розраховане значення t -критерія менше критичного, нульова гіпотеза про випадковість часового ряду приймається, в іншому випадку відхиляється і з імовірністю $(1-\alpha)$ визнається наявність тренду.

При $|t_D| > t_{1-\alpha}$ йдеться про наявність тренду середньої, при $|t_S| > t_{1-\alpha}$, — про наявність тренду дисперсії. У нашому прикладі t_D перевищує критичне значення ($1,85 > 1,76$), а t_S менше критичного ($0,896 < 1,76$). Відтак, гіпотеза про відсутність тренду середніх з імовірністю 0,90 відхиляється, а гіпотеза про відсутність тренду дисперсії визнається доведеною. Іншими словами: критерій Фостера-Стюарта підтверджує наявність позитивного тренду середніх і відсутність тренду дисперсії.

За наявності тренду і сезонних (циклічних) коливань між послідовними рівнями часового ряду може виникнути ситуація, коли значення y_t певною мірою залежить від попередніх значень i , своєю чергою, впливає на наступні значення. Залежність між рівнями ряду, зсувеними на p часових одиниць, називають **автокореляцією** (*autocorrelations*), а часовий зсув між ними — **лагом**.

Для вимірювання автокореляції використовують однойменний коефіцієнт r_p , значення якого залежить від величини лагу p і не виходить за межі ± 1 . Лаг p вказує на порядок коефіцієнта автокореляції. Так, коефіцієнт першого порядку r_1 вимірює кореляцію поточних рівнів часового ряду y_t з попередніми рівнями y_{t-1} , другого порядку r_2 — кореляцію y_t з рівнями, зсуненими на 2 часові одиниці y_{t-2} і т. д. Чим довший часовий ряд, тим вищим може бути порядок коефіцієнтів автокореляції. Проте збільшення порядку лагу не безмежне, максимальна його величина співвідноситься з довжиною часового ряду n в пропорції $[n : 4]$.

У практиці тестування стаціонарності часових рядів застосовують переважно циклічний коефіцієнт автокореляції першого порядку r_1 , який обчислюється по аналогії з лінійним коефіцієнтом кореляції:

$$r_1 = \frac{\overline{(y_t y_{t+1})} - \bar{y}_t \bar{y}_{t+1}}{\sigma_{y_t} \sigma_{y_{t+1}}}, \text{ або } r_1 = \frac{\overline{(y_t y_{t-1})} - \bar{y}_t \bar{y}_{t-1}}{\sigma_{y_t} \sigma_{y_{t-1}}},$$

де σ_{y_t} — стандартне відхилення.

Аналогічно можна визначити коефіцієнти автокореляції другого і більш високих порядків.

Характеризуючи ступінь залежності двох послідовних рівнів часового ряду, коефіцієнт автокореляції r_1 є мірою неперервності ряду, і за його значенням можна судити про наявність лінійного (або близького до лінійного) тренду. В разі, коли часовому ряду властивий чітко виражений тренд, $r_1 \rightarrow 1$, коли ж рівні ряду незалежні, $r_1 \rightarrow 0$. Для часових рядів, яким властивий нелінійний тренд (експонента, поліном 2-ого ступеню тощо), коефіцієнт автокореляції також може прямувати до нуля. Водночас слід зауважити, що значення r_p не дає підстав робити висновок щодо характеру тренду (зростаючий чи спадний). Більшості часових рядів економічних даних властива позитивна автокореляція, однак при цьому вони можуть мати спадний тренд. Відносно високі значення коефіцієнта автокореляції при $p > 2$ свідчать про наявність у часовому ряду циклічних коливань з періодичністю p часових одиниць.

Послідовність коефіцієнтів автокореляції першого, другого і вищих порядків r_p формує автокореляційну функцію, яку зображають графічно у вигляді *корелограми* з абсцисою p і ординатою r_p . Корелограма ілюструє, як часто і з яким запізненням зміна значення y_t впливає на наступні рівні ряду. За швидкістю згасання корелограми можна зробити висновок про характер динаміки.

Для кожного окремого лагу корелограма містить значення коефіцієнта r_p , його стандартну похибку ($S.E.$) і 95 %-й довірчий інтервал, що дає можливість перевірити істотність коефіцієнтів автокореляції. Точками позначено дві вертикальні прямі, що визначають 95 %-ні довірчі межі коефіцієнта r_p . Якщо жоден r_p не виходить за межі довірчого інтервалу, часовий ряд вважається стаціонарним.

На рис. 2.6 наведено корелограму часового ряду середніх цін 1 кв. м житла на вторинному ринку. Найбільше значення коефіцієнта автокореляції першого порядку ($r_1 = 0,848$) і характер згасання функції підтверджують лінійний характер динаміки показника.

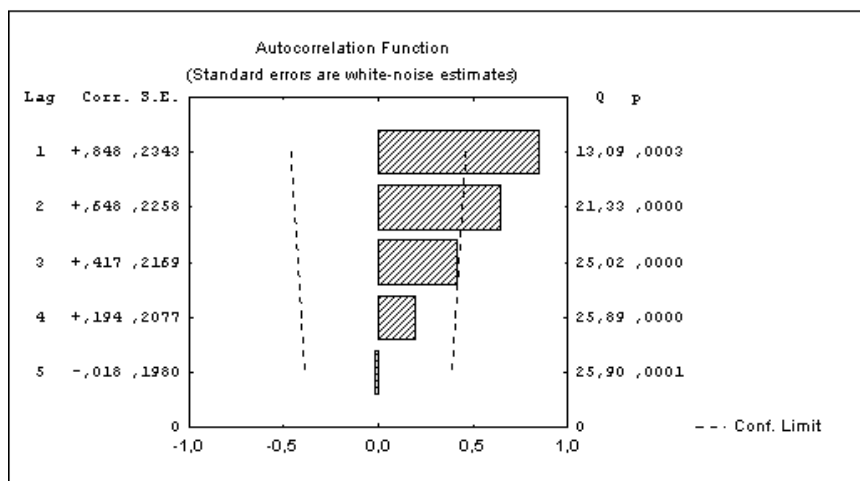


Рис. 2.6. Автокореляція середніх цін 1 кв. м житла на вторинному ринку

В разі, коли жоден з коефіцієнтів автокореляції не є значимим, можна зробити припущення відносно структури цього ряду: або ряд не містить тренду і циклічних коливань, або містить сильний нелінійний тренд, ідентифікація якого передбачає додатковий аналіз.

2.3. Аналіз властивостей випадкової компоненти

Випадкова, нерегулярна компонента часового ряду e_t — це залишкова послідовність після фільтрації з часового ряду детермінованих компонент — тренду, циклічних чи сезонних коливань:

$$e_t = y_t - (D_t + S_k).$$

Залишкову послідовність e_t формують:

- випадкові коливання, зумовлені відносно слабким і різноспрямованим впливом факторів, які не піддаються обліку чи реєстрації, а також похибки реєстрації та обчислень;
- випадкові, непередбачені збурення (інтервенцій), зумовлені раптовими сильними зовнішніми впливами.

Вплив випадкових факторів на формування залишкової послідовності визначає стохастичну її природу, а розподіл залишків належить до класу стаціонарних часових рядів, тобто має нульову середню та постійну дисперсію.

Отже, властивості випадкової компоненти:

- незалежність значень залишкової послідовності;
- відповідність розподілу залишкової послідовності нормальному закону;
- нульове значення математичного сподівання розподілу.

Розглянемо найбільш поширені методи перевірки цих властивостей випадкової компоненти.

Для оцінювання незалежності значень залишкової послідовності застосовують коефіцієнт автокореляції залишків першого порядку, тобто корелюються ряди залишкових величин e_t з лагом $p = 1$:

e_1	e_2	e_3	...	e_{n-1}	e_n
e_2	e_3	e_4	...	e_n	e_1

Припускаючи, що $\sum e_t = \sum e_{t+1} = 0$, формула розрахунку циклічного коефіцієнта автокореляції спрощується: $r_1 = \frac{\sum_1^n e_t e_{t+1}}{\sum_1^n e_t^2}$.

Розрахунок коефіцієнта автокореляції залишкових величин наведено в табл. 2.3.

За даними таблиці: $r_1 = \frac{-27,62}{141,92} = -0,195$, фактичне значення коефіцієнта автокореляції порівнюють з критичним для заданого рівня істотності α (дод. VI). Для часового ряду довжиною $n = 10$ критичне значення коефіцієнта автокореляції при $\alpha = 0,05$ становить 0,564. Оскільки фактичне значення r_1 менше критичного, можна стверджувати, що автокореляція залишкових величин відсутня.

Таблиця 2.3

**ДО РОЗРАХУНКУ КОЕФІЦІЄНТА
АВТОКОРЕЛЯЦІЇ ЗАЛИШКОВОЇ ПОСЛІДОВНОСТІ**

№ з.п.	e_t	e_{t+1}	$e_t e_{t+1}$	e_t^2	$(e_t - e_{t+1})^2$
1	0,2	-1,8	-0,36	0,04	4
2	-1,8	-4	7,2	3,24	4,84
3	-4	6,6	-26,4	16	112,36
4	6,6	4,3	28,38	43,56	5,29
5	4,3	-1,9	-8,17	18,49	38,44
6	-1,9	-5,2	9,88	3,61	10,89
7	-5,2	3,5	-18,2	27,04	75,69
8	3,5	-3,7	-12,95	12,25	51,84
9	-3,7	2	-7,4	13,69	32,49
10	2	0,2	0,4	4	3,24
Разом	0	0	-27,62	141,92	339,08

У програмних засобах перевірку наявності автокореляції здійснюють переважно за тестом Дарбіна–Уотсона, статистика якого DW набуває значень в межах $0 < DW < 4$ і функціонально пов'язана з коефіцієнтом автокореляції залишкових величин r_1 :

$$DW = \frac{\sum_2^n (e_t - e_{t+1})^2}{\sum_1^n e_t^2} = 2(1 - r_1)$$

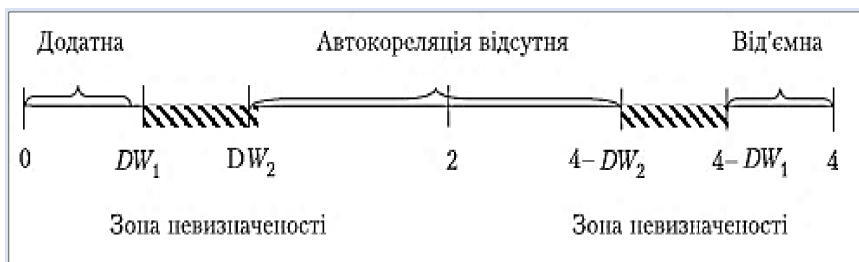
Для наведеної в табл. 2.3 залишкової послідовності:

$$DW = 2[1 - (-0,195)] = 2,39.$$

На відміну від інших тестів, за результатами перевірки гіпотези про відсутність автокореляції за критерієм DW можливі три висновки:

- автокореляція відсутня;
- гіпотеза про відсутність автокореляції відхиляється;
- висновок залишається невизначеним.

На рис. 2.7 схематично зображена шкала перевірки наявності автокореляції залишків на основі порівняння фактичного значення DW з двома критичними значеннями: верхнім DW_U і нижнім DW_L .

Рис. 2.7. Шкала тестування автокореляції за критерієм DW

За відсутності автокореляції між суміжними членами часового ряду значення DW становить приблизно 2, за високої додатної автокореляції DW наближається до 0, за високої від'ємної автокореляції — до 4. У дод. VII наведено критичні значення DW для додатної автокореляції при $\alpha = 0,05$. Перевірку наявності від'ємної автокореляції здійснюють на основі значень $(4 - DW)$, наприклад, за даними табл. 2.3 $DW = 4 - 2,39 = 1,6$. Критичні значення DW залежать від кількості членів ряду n та кількості параметрів детермінованої компоненти m . Так, для залишкової послідовності табл. 2.3 ($n = 10, m = 1$) критичні значення DW становлять: $DW_L = 0,88, DW_U = 1,32$, фактичне значення $DW = 1,61$ потрапляє в зону допустимих значень нульової гіпотези, тобто з імовірністю 0,95 можна стверджувати, що в залишковій послідовності автокореляція відсутня.

Перевірку відповідності розподілу залишкової послідовності нормальному закону розподілу можна здійснити, зокрема, за допомогою нормованого розмаху — R/S -критерію, який визначається відношенням розмаху варіації залишкових величин до стандартного відхилення s_e :

$$\frac{R}{S} = \frac{e_{\max} - e_{\min}}{s_e}, \text{ де } s_e = \sqrt{\frac{\sum e_t^2}{n - m}}.$$

У залишковій послідовності табл. 2.3 $s_e = \sqrt{\frac{141,92}{10 - 1}} = 3,97$.

Звідси

$$\frac{R}{S} = \frac{6,6 - (-5,2)}{3,97} = 2,97.$$

Розраховану статистику R/S -критерія порівнюють з критичними значеннями (дод. III). В разі, коли значення R/S не виходить за

критичні межі, гіпотеза про нормальність розподілу з імовірністю 0,95 визнається доведеною, в іншому випадку гіпотеза відхиляється.

Перевіримо нормальність розподілу залишкової послідовності (табл. 2.3). Межі критичних значень R/S для $n = 10$ і рівня істотності $\alpha = 0,05$ становлять, відповідно, 2,69 і 3,69. Оскільки $R/S = 2,97$ потрапляє в інтервал між критичними значеннями, можна зробити висновок, що розподіл залишкової послідовності відповідає нормальному закону. В такому разі тестування гіпотези про рівність нулю математичного сподівання розподілу здійснюють, як правило, на основі t -критерія Стьюдента. Якщо розрахункове значення t менше критичного для заданого рівня істотності α та числа ступенів свободи $(n-1)$, тобто $t < t_{1-\alpha}(n-1)$, гіпотеза про рівність нулю математичного сподівання залишкової послідовності приймається з імовірністю $(1-\alpha)$, в іншому випадку гіпотеза відхиляється. Наведена в табл. 2.3 залишкова послідовність задовольняє усім властивостям випадкової компоненти часового ряду.

Питання і задачі для самопідготовки

1. Назвіть основні структуроутворюючі елементи (компоненти) часового ряду.
2. Дайте визначення і наведіть приклад стаціонарного часового ряду.
3. Яка компонента часового ряду і в якій формі відбиває перевагу закономірності над випадковістю?
4. Як називають загальний напрямок розвитку процесу, поданого часовим рядом?
5. Розкрийте поняття сезонної та циклічної компонент часового ряду.
6. Поясніть способи взаємодії окремих компонент часового ряду в адитивних, мультиплікативних та змішаних моделях.
7. Як перевірити часовий ряд на стаціонарність? Тестується гіпотеза про відсутність чи наявність тренду?
8. Наведіть приклади нестаціонарних часових рядів. Якими методами їх можна звести до стаціонарних?
9. Яку функцію в аналізі часових рядів виконує коефіцієнт автокореляції?
10. У часовому ряду заборгованості за комунальні послуги ($n = 12$ кварталів) коефіцієнт автокореляції становить $r_1 = -0,539$. Чи підтверджує такий рівень автокореляції наявність тренду?

11. Яку інформацію щодо структури часового ряду містить корелограма?

12. Який висновок можна зробити на основі автокореляційної функції часового ряду довжиною $n = 25$?

Лag, p	1	2	3	4	5	6	7
r_p	0,65	0,47	0,18	0,06	-0,04	-0,09	0,05

13. Поясніть особливості перевірки гіпотези про відсутність автокореляції в часовому ряду за критерієм Дарбіна-Уотсона.

14. Чому випадкову компоненту розглядають як обов'язкову складову будь-якого часового ряду?

15. Які завдання вирішує статистичний аналіз залишкової послідовності?

16. Обгрунтуйте послідовність етапів дослідження динамічних процесів.

17. Динаміка правопорушень в країні за статтею «Торгівля людьми» характеризується даними, випадків (y_t):

t	y_t	t	y_t	t	y_t	t	y_t
1	109	6	115	11	117	16	110
2	107	7	120	12	124	17	116
3	108	8	119	13	126	18	118
4	109	9	112	14	119	19	124
5	118	10	111	15	114	20	121

Перевірте гіпотезу про наявність у часовому ряду тренду середньої і дисперсії з використанням критерія Фостера-Стюарта. Висновки зробіть з імовірністю 0,95.

18. Динаміку використання вторинної сировини (брухту і відходів чорних металів) характеризують річні дані:

№ з/п	1	2	3	4	5	6	7	8	9
Тис. т	22,5	19,4	21,1	21,4	20,4	19,1	22,1	23,3	24,6

Обчисліть коефіцієнт автокореляції першого порядку, з імовірністю 0,95 перевірте його істотність. Чи є підстави для висновку про наявність у часовому ряду лінійного тренду?

19. Використовуючи коефіцієнт автокореляції і критерій Дарбіна-Уотсона, перевірте наявність автокореляції в залишковій послідовності моделі витрат компанії на модернізацію діючого устаткування, млн грн:

t	e_t	t	e_t	t	e_t	t	e_t
1	-15,4	6	-14,3	11	11,9	16	-30,7
2	3,8	7	-30,5	12	21,4	17	-17,6
3	13,9	8	-19,2	13	15,8	18	15,0
4	29,4	9	-3,6	14	-14,5	19	18,6
5	33,6	10	4,5	15	-9,8	20	-12,3

20. За даними задачі №19 перевірте відповідність залишкової послідовності нормальному закону розподілу з використанням R/S-критерія.

Рекомендована література

1. Єріна А.М., Єрін Д.Л. Статистичне моделювання та прогнозування : підручник. Київ, КНЕУ. 2014. С. 129—148.
2. Методи і моделі економічного прогнозування : навч. посіб. / Кічор В.П. та ін. Львів, 2019. С. 61—102.
3. Time Series Analysis: The Basics. URL: <https://www.abs.gov.au> › home (Last accessed: 30.05.2022).

- 3.1. Згладжування часових рядів
- 3.2. Трендові моделі на основі кривих зростання
- 3.3. Екстраполяція тренду як метод прогнозування

3.1. Згладжування часових рядів

Через постійний перерозподіл впливу факторів, які формують динаміку процесу, змінюється інтенсивність динаміки, частота й амплітуда коливань. Як наслідок, часові ряди навіть за наявності еволюції досліджуваного процесу не виявляють чітко вираженої тенденції розвитку і до них застосовують більш еластичні спеціальні статистичні прийоми і способи *згладжування (Smoothing)*. В процесі згладжування відбувається фільтрування (поглинання) коливань часового ряду, відповідно, оператори, за допомогою яких виконується процедура згладжування, називають фільтрами. На практиці застосовують переважно лінійні фільтри, найчастіше *ковзну середню* з інтервалом згладжування $m < n$.

Суть згладжування методом ковзних середніх полягає в заміні рівнів первинного ряду y_t в межах вибраного часового інтервалу (вікна N) середньою арифметичною або в разі наявності викидів — медіаною. При розрахунку ковзних середніх кожний наступний інтервал утворюється на основі попереднього заміною одного рівня:

y_1	y_2	...	y_m
y_2	y_3	...	y_{m+1}
y_3	y_4	...	y_{m+2}

Інтервали зміщуються по шкалі часу, формуючи $(n - m + 1)$ ковзних середніх. У середніх взаємозрівноважуються коливання рівнів первинного ряду, зумовлені впливом випадкових причин,

внаслідок чого варіація середніх порівняно з варіацією рівнів первинного ряду значно зменшується, а властива ряду тенденція розвитку виявляється чіткіше. Що ширший інтервал згладжування m , тим більш плавним буде згладжений ряд, водночас він виявляється коротшим за первинний ряд на $(m - 1)$ рівнів, що потребує уважного ставлення до вибору ширини інтервалу m .

Оскільки середня припадає на середину інтервалу, то доцільно формувати інтервали з непарного числа рівнів первинного ряду ($m = 3, 5, 7$ і т. д.). Коли необхідно згладити дрібні безладні коливання, інтервал згладжування беруть якомога більшим; коли потрібно зберегти дрібніші коливання, інтервал згладжування зменшують.

В разі, коли первинному ряду динаміки притаманна певна періодичність коливань, то інтервал згладжування має бути рівним або кратним періоду коливань. Якщо інтервал згладжування — парне число, здійснюють додаткову процедуру центрування, суть якої полягає у послідовному усередненні кожної пари значень ковзних середніх \bar{y}_i . Така ситуація виникає у разі згладжування поквартальних ($m = 4$), помісячних ($m = 12$) чи погодинних ($m = 24$) даних. Слід зауважити, що незважену ковзну середню застосовують лише для рядів, яким властива лінійна тенденція.

Якщо в межах інтервалу згладжування припускають нелінійну тенденцію, застосовують зважену ковзну середню, тобто в межах інтервалу кожному рівню надають певну вагу. Способи формування вагової функції різні. В одних випадках ваги відповідають

членам розкладання бінома $(\frac{1}{2} + \frac{1}{2})^{2p}$, наприклад, в пакеті *Statistica*, модуль *Time Series / Forecasting*, 3-точкове згладжування ковзною середньою здійснюють за опцією 4253H *Filter* з вагами $a_r = 0,25, 0,5, 0,25$; в інших випадках до даних інтервалу згладжування добирається певний поліном. Так, для поліномів другого та третього порядків числова послідовність ваг при інтервалі згладжування $m = 5$ має вигляд: $\{-3; 12; 17; 12; -3\}$, при $m = 7 - \{-2; 3; 6; 7; 6; 3; -2\}$. Як бачимо з приведених формул, ваги симетричні відносно центрального рівня і сума їх з урахуванням спільного множника, винесеного за дужки, дорівнює 1. В системі ваг присутні від'ємні значення, завдяки чому згладжена крива значною мірою зберігає різні вигини тренду, що ілюструє рис. 3.1.

Основна перевага ковзної середньої — наочність та простота тлумачення тенденції. Проте не слід забувати, що ряд ковзних

середніх коротший за первинний ряд, тож втрачається інформація про крайні члени ряду. І що ширший інтервал згладжування, то відчутніші втрати, особливо нової інформації. Окрім того, маючи спільну основу розрахунку, ковзні середні виявляються залежними, що в разі згладжування значних коливань навіть за відсутності циклів у первинному ряду може вказувати на циклічність процесу (ефект Слуцького).

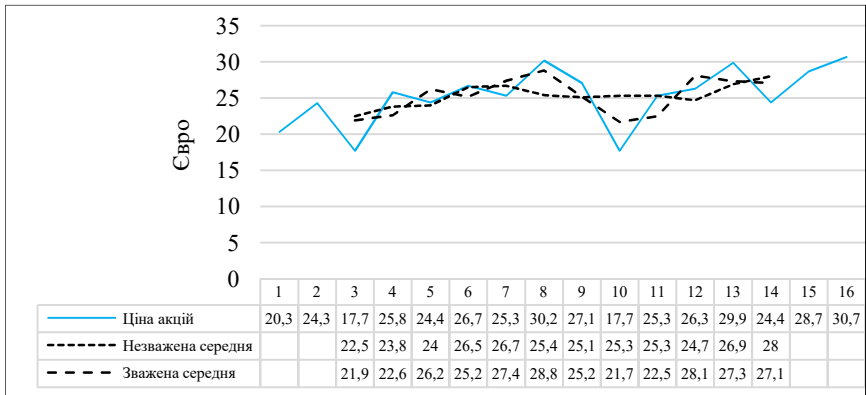


Рис. 3.1. Згладжування ціни акцій компанії ABC методом зваженої і незваженої ковзних середніх ($m = 5$)

У симетричних фільтрах стара і нова інформація рівновагомі, а під час прогнозування більш важлива нова інформація. У такому разі використовують асиметричні фільтри. Найпростіший з них — ковзна середня, яка замінює не центральний, а останній член інтервалу — так звана *адаптивна ковзна середня*.

При використанні зважених асиметричних фільтрів вагова функція формується з урахуванням ступеня новизни інформації. Такою є середня з експоненційно розподіленими вагами:

$$Y_t = \sum_{r=0}^t a(1-a)^r y_{t-r},$$

де Y_t — експоненційна середня, тобто згладжене значення рівня часового ряду на момент t ;

$a(1-a)^r$ — вага $(t-r)$ -го рівня;

a — параметр згладжування, який визначає вагу t -го рівня, значення його коливаються в межах від 0 до 1.

Чим віддаленіший від t -го моменту рівень ряду, тим менша його відносна вага і вклад у тенденцію. Так, при $a = 0,2$ ваги ста-

новлять: для t -го моменту $0,2$, для $(t - 1)$ -го моменту — $0,2(1 - 0,2) = 0,16$; для $(t - 2)$ -го моменту — $0,2(1 - 0,2)^2 = 0,128$ і т. д. Надаючи більшу вагу новій інформації, експоненційна середня адаптується до нових умов, що робить її досить ефективним і надійним методом *короткострокового прогнозування* (див. розд. 6).

3.2. Трендові моделі на основі кривих зростання

Згладжування ілюструє властивий часовому ряду тренд, але не спроможне кількісно його виміряти. Між тим в аналізі динамічних процесів важливо не лише виявити наявність і напрямок тренду, але й оцінити закономірність розвитку аби на її підставі спрогнозувати можливі майбутні зміни. Оцінювання тенденцій різноманітних процесів навколишнього світу виконують трендові моделі, які аналітично задаються у формі математичних функцій змінної часу: $Y_t = f(t)$, де $t = 1, 2, \dots, n$.

Певна річ, час не є чинником конкретного динамічного процесу, змінна часу t акумулює комплекс постійно діючих умов та причин, які визначають цей процес. Описати властивий часовому ряду тренд аналітичним методом означає надати досліджуваному процесу одноманітного характеру розвитку у вигляді плавної кривої, яка пов'язує єдиним «законом руху» всі послідовні рівні часового ряду в межах певного часового інтервалу. Процедура описування наявного в часових рядах тренду охоплює два етапи:

— обґрунтування (вибір) типу функції, яка б адекватно описувала характер тренду;

— оцінювання параметрів функції.

Вибір типу функції ґрунтується на теоретичному аналізі суті явища, яке вивчається, і характеру його динаміки. Зазвичай перевага надається функціям, параметри яких мають чіткий економічний зміст. Трендові моделі здебільшого апроксимують зміни показника в часі та підходять для моделювання будь-яких динамічних рядів. Вони формують клас так званих *кривих зростання*. Ці криві поділяють на три класи залежно від того, який характер динаміки вони описують. До першого класу відносять функції, що описують процеси з монотонним характером розвитку — це поліноміальні криві та проста експоненційна функція. Другий клас представляє функція, еластична до процесів з межею наси-

чення, — модифікована експонента. До третього класу відносять S-подібні криві, які описують процеси з межею насичення і точками перегину, — криву Гомперца та логістичну криву.

В дослідженнях соціально-економічних процесів найчастіше застосовують поліноми першого, другого, третього порядку. Поліноми більш високих ступенів, особливо в коротких часових рядах, використовувати недоцільно, оскільки такі функції можуть відтворювати властиві ряду випадкові відхилення, що суперечить сенсу виявлення тенденції.

Поліноміальні функції різного порядку мають вигляд:

$$Y_t = a + \sum_1^n b_i t^i.$$

де a, b — параметри функцій;

$i = 1, 2, \dots, p$ — порядок поліному.

Параметри поліномів визначають методом найменших квадратів (МНК) шляхом розв'язання системи нормальних рівнянь. Функції нелінійного типу приводяться до лінійного виду переважно шляхом логарифмування.

Найпоширеніші види трендових моделей, реалізованих у програмному комплексі MS Excel, наведено на рис. 3.2.

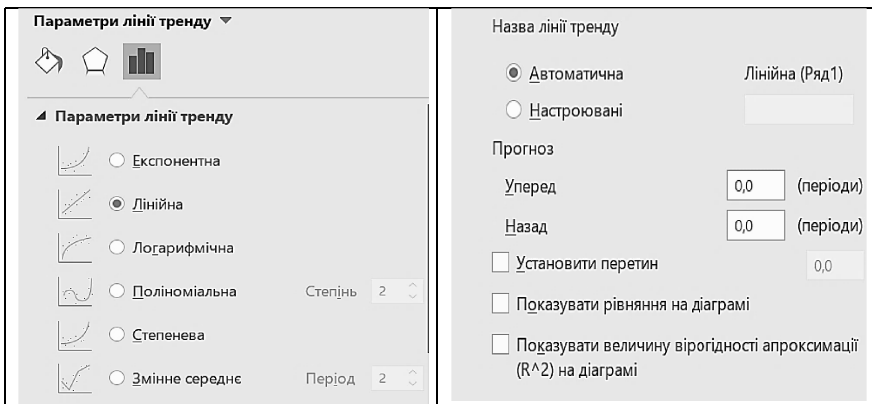


Рис. 3.2. Налаштування параметрів лінії тренду у програмному комплексі MS Excel

Суттєвою підмогою при виборі функцій є аналіз ланцюгових характеристик інтенсивності динаміки. Коли ланцюгові абсолютні прирости відносно стабільні, не мають чіткої тенденції до

зростання чи зменшення, вирівнювання ряду виконують на основі лінійної функції. Так, лінійний тренд $Y_t = a + bt$ описує процеси, які рівномірно змінюються в часі і мають стабільні абсолютні прирости ординат. Якщо ж відносно стабільними є ланцюгові темпи приросту, то такому характеру динаміки найбільш адекватна експонента $Y_t = ab^t$. У зазначених функціях: t — порядковий номер періоду (дати), a — значення показника при $t = 0$. Параметр b характеризує швидкість динаміки: середню абсолютну в лінійній функції і середню відносну — в експоненті. На рис. 3.3 представлена динаміка обсягів даних, що обробляються і зберігаються в сучасному світі. Бурхливе зростання обсягів створеної інформації описується експонентою, а доступні ресурси по зберіганню даних — поліномом першого ступеня.



Рис. 3.3. Динаміка обсягів даних, що обробляється і зберігається в сучасному світі, zettabyte (ZB)

Джерело: Data Age 2025

Коли характеристики швидкості динаміки змінюються (зростають чи зменшуються), для описування тенденції використовують інші функції. Скажімо, поліном 2-го ступеня (парабола) $Y_t = a + bt + ct^2$ описує процеси, характерною особливістю яких є рівноприскорене зростання або рівноприскорене зменшення ординат, b — початкова абсолютна швидкість, яка з плином часу змінюється на константу $2c$. При цьому для однієї гілки параболи прирощення додатні, для другої — від'ємні. Параметри b і c визначають форму параболи. Залежно від значень цих параметрів крива відображає різну еволюцію розвитку процесу (табл. 3.1).

Таблиця 3.1

АНАЛІТИЧНІ МОЖЛИВОСТІ ПАРАБОЛИ

Параметри		Парабола описує
b	c	
$b > 0$	$c > 0$	прискорене зростання
$b < -2Tc$	$c > 0$	уповільнений спад
$-2Tc < b < 0$	$c > 0$	парабола має мінімум
$b < 0$	$c < 0$	прискорений спад
$b > -2Tc$	$c < 0$	уповільнене зростання (вершина після t)
$0 < b < -2Tc$	$c < 0$	парабола має максимум

Завдяки таким універсальним властивостям параболу використовують для описування широкого спектра економічних процесів, серед них дві форми параболі визнаються класичними: коли $c > 0$, гілки параболі спрямовані вгору — парабола має *мінімум*, коли $c < 0$, гілки параболі спрямовані вниз — парабола має *максимум*. Екстремум функції досягається у точці $t = -b/2c$.

Наприклад, динаміка захворювань під час епідемії грипу описується параболою $Y_t = 264 + 45t - 1,5t^2$. Звідси екстремум функції $t = -45 / 2 \cdot (-1,5) = 15$ днів. Максимум захворювань буде зафіксовано через 15 днів від початку відліку часу і становитиме $Y_{\max} = 264 + 45 \cdot 15 - 1,5 \cdot 15^2 \cong 601$ особа.

У полінома 3-го ступеня $Y_t = a + bt + ct^2 + dt^3$ знак прирощення ординати може змінюватися один чи два рази. Саме таку динаміку виявили глобальні мілітарні витрати за 2009—2021 рр. (рис. 3.4).

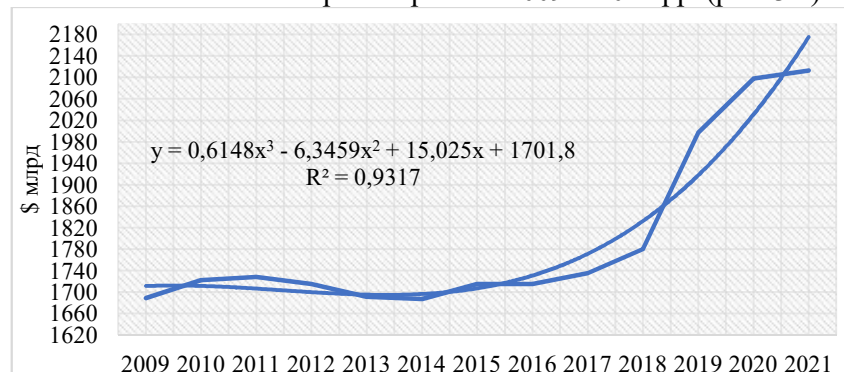


Рис. 3.4. Динаміка глобальних мілітарних витрат за 2009—2021 рр., млрд доларів США у постійних цінах

Джерело: Trends in world military expenditure, 2021

Початок ХХІ ст. ознаменувався збільшенням світових витрат на оборону, і така динаміка була характерна для усіх ключових геополітичних гравців. Після нетривалої стабілізації нові збройні конфлікти загострили ситуацію з безпекою на Близькому Сході та в Східній Європі, спричинили ще більш потужний виток нарощення військових витрат.

Для моделювання процесів, характерною властивістю динаміки яких є відносно стабільні темпи приросту, найбільш адекватна експонента. Експонентою описують так звану динаміку популяцій, тобто процеси зміни в часі основних біологічних показників (чисельності, біомаси, структури). Чисельність популяції зростає спочатку повільно, а потім стрімко, згідно експоненціальному закону, тобто крива зростання популяції приймає J-подібний вид. Будь-яка популяція теоретично здатна до необмеженого росту чисельності, якщо її не лімітують фактори зовнішнього середовища.

Алгебраїчно експонента може набувати різних еквівалентних форм, які приводяться до лінійного вигляду заміною y_t десятковими або натуральними логарифмами (табл. 3.2).

Таблиця 3.2

ЕКВІВАЛЕНТНІ ФОРМИ ЕКСПОНЕНТИ

Форма експоненти	Трансформації параметрів	Лінеаризована форма експоненти
$Y_t = ab^t$	Основна форма (показникова) $b > 0$	$\lg Y = \lg a + t \lg b$ $\ln Y = \ln a + t \ln b$ $= \ln a + \lambda t$
$Y_t = ae^{\lambda t}$	b замінюється на e^λ , де $\lambda = \ln b$; $e = 2,718$ — основа натурального логарифма; $\ln e = 1$	$\ln Y = \ln a + \lambda t \ln e =$ $= \ln a + \lambda t$
$Y_t = e^{a+\lambda t}$	a замінюється на e^a , b замінюється на e^λ , звідси $a = \ln a$ і $\lambda = \ln b$	$\ln Y = \ln e^a + \ln e^{\lambda t} =$ $= \ln a + \lambda t$
$Y_t = a(1-r)^t$	b замінюється на $(1-r)$, де $r = b - 1$	$\ln Y = \ln a + t \ln (1 - r)$

Основна форма експоненти — показникова $Y_t = ab^t$, де b — стабільна відносна швидкість зміни ординати, тобто темп зростання: при $b > 1$ ордината монотонно (зі стабільним темпом) зростає, при $b < 1$, навпаки, монотонно зменшується. Абсолютний приріст пропорційний досягнутому рівню.

В експонентах за формою $Y_t = ae^{\lambda t}$ і $Y_t = e^{a+\lambda t}$ параметр $\lambda = \ln b$ за змістом є стабільним темпом приросту. При $\lambda > 0$ експоненційний

тренд показує тенденцію прискореного розвитку, наприклад, зростання захворюваності населення в період пандемії або темп гіперінфляції.

Форму експоненти $Y_t = a(1 - r)^t$ найчастіше використовують у фінансових розрахунках. Припустимо, капітал фірми C_0 перебуває в банку протягом t років з річним процентом r . Через t років капітал становитиме $C_t = C_0(1 + r)^t$. Очевидно, на основі цієї функції можна оцінити майбутні витрати чи прибутки C_t у вартості поточного року C_0 . Наприклад, необхідно порівняти кошти, які витрачаються на різні інвестиційні проекти у поточному періоді C_0 і майбутні очікувані прибутки відповідних проектів C_t :

$$C_0 = \frac{C_t}{(1+r)^t}.$$

Переходячи до форми $Y_t = ae^{\lambda t}$, можна отримати поширене у фінансових розрахунках так зване *правило 70*, на основі якого визначають термін подвоєння капіталу:

$$C_t = 2C_0 = C_0 e^{\lambda t}, \text{ звідси } \Rightarrow t = \frac{\ln 2}{\lambda} = \frac{0,6931}{\lambda} = \frac{0,70}{\lambda}.$$

Степенева форма тренду $Y = at^b$ при $a > 0$, що є типовим для економічних процесів, залежно від знаку параметра b описує неупинне зростання процесу чи його спад. Коли $b = 1$, степенева форма тренду перетворюється на лінійний, коли $b = 2$ — на параболічний.

Оцінювання параметрів трендових рівнянь найчастіше здійснюють *методом найменших квадратів* (МНК), основною умовою якого є мінімізація суми квадратів відхилень фактичних значень y_t від теоретичних Y_t , визначених за трендовим рівнянням

$$\sum_1^n (y_t - Y_t)^2 = \min.$$

Параметри многочленів визначають безпосередньо розв'язуванням систем $p + 1$ нормальних рівнянь. За допомогою МНК можна визначити параметри розглянутих нелінійних функцій, але попередньо треба усунути нелінійність, тобто перетворити їх на лінійні (лінеаризувати). Найчастіше для цього застосовують логарифмічне перетворення даних. Саме в такий спосіб зводяться до лінійного виду експоненти будь-якої форми, степенева і логарифмічна функції, що надає змогу визначити параметри цих функцій методом найменших квадратів.

Побудову трендової моделі МНК розглянемо за даними динаміки інвестицій в нові потужності відновлюваної енергетики в регіоні (табл. 3.3). Для часового ряду характерний тренд до зрос-

тання інвестицій відносно стабільними темпами, що дає підстави описати тренд експонентою

$$Y_t = ae^{\lambda t}.$$

Таблиця 3.3

ДИНАМІКА ІНВЕСТИЦІЙ В РОЗВИТОК ВІДНОВЛЮВАНОЇ ЕНЕРГЕТИКИ

Рік	Інвестиції, млрд дол.	Ланцюгові		Рік	Інвестиції, млрд дол.	Ланцюгові	
		абсолютні прирости	темпи зростання			абсолютні прирости	темпи зростання
2011	22,8	—		2016	37,3	2,1	1,060
2012	21,7	-1,1	0,952	2017	42,2	4,9	1,131
2013	24,3	2,6	1,119	2018	48,4	6,2	1,147
2014	27,8	3,5	1,144	2019	52,8	4,4	1,091
2015	35,2	7,4	1,266	2020	57,5	4,7	1,089

Розрахунки виконані у програмному комплексі MS Excel, криву зростання у формі експоненти наведено на рис. 3.5. Експонента набуває вигляду $Y_t = 18,366e^{0,117t}$, тобто інвестиції в розвиток відновлюваної енергетики щороку зростали в середньому на 11,7%. Про адекватність моделі реальному процесу свідчить коефіцієнт детермінації $R^2 = 0,9752$.

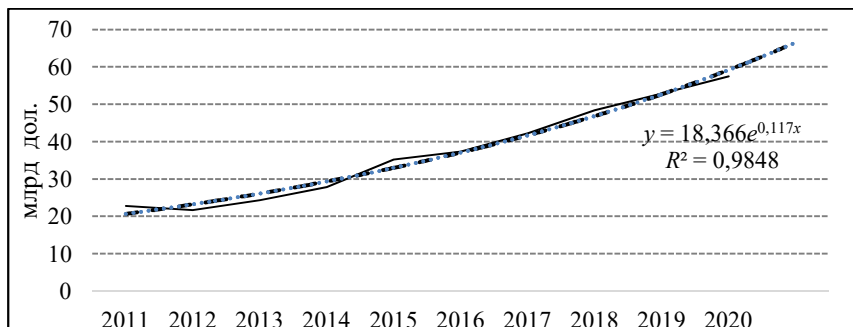


Рис. 3.5. Динаміка інвестицій в розвиток відновлюваної енергетики

Важливим етапом моделювання динаміки є оцінювання адекватності моделі реальному процесу. Основою перевірки гіпотези про адекватність моделі є аналіз залишків, тобто відхилень між первинними рівнями ряду і вирівняними за трендом $e_t = y_t - Y_t$. За умови, що модель вибрана правильно, випадкова складова є стаціонарним процесом з математичним сподіванням $M(e) = 0$ і дисперсією

$$s_e^2 = \frac{\sum (y_t - f_t)^2}{n - m},$$

де n — довжина ряду; m — кількість параметрів функції (для лінійної функції та експоненти $m = 2$).

Для оцінювання адекватності трендового рівняння реальному процесу, тобто наближення теоретичних рівнів Y_t до фактичних y_t , використовують також стандартну похибку $s_e = \sqrt{s_e^2}$, коефіцієнт детермінації R^2 та та F -критерій для перевірки істотності R^2 :

$$R^2 = 1 - \frac{\sum (y_t - Y_t)^2}{\sum (y_t - \bar{y})^2}; F = \frac{R^2}{1 - R^2} \cdot \frac{m - 1}{n - m};$$

Усі ці характеристики трендової моделі наводяться в інформаційному полі

Regression Summary for Dependent Variable модуля *Multiply Regression Results* пакета *Statistica* (рис. 3.6):

Multiple Regression

Dependent: Інвестиції, ln Multiple R = ,9875 F = 314,8365

$R^2 = ,9752$ df = 1,8

No. of cases: 10 adjusted $R^2 = ,9721$ p = ,0000

F(1,8) = 314,84 Standard error of estimate: ,05972

Рис. 3.6. Характеристики адекватності експоненти реальному процесу інвестицій в розвиток відновлюваної енергетики

Перевірку істотності коефіцієнта детермінації здійснюють традиційно: фактичне його значення порівнюють з критичним (див. додаток IX). Для часового ряду інвестицій в розвиток відновлюваної енергетики фактичне значення $R^2 = 0,975$ значно перевищує критичне $R_{1-0,05}^2(1; 8) = 0,399$, що з імовірністю 0,95 підтверджує адекватність експоненти реальному процесу.

Іншою умовою адекватності тренду є відсутність автокореляції залишкових величин. Для перевірки відсутності суттєвої авто-

кореляції у залишковій послідовності застосовують різні критерії, серед них найбільш поширений критерій Дарбіна-Уотсона (див. підр. 2.3).

Клас моделей, що описують тенденції розвитку, досить широкий. Вибір типу моделі у конкретному дослідженні ґрунтується передусім на теоретичному аналізі специфіки процесу і його внутрішньої структури. На основі такого аналізу в загальних рисах визначають характер динаміки (рівномірний, рівноприскорений, з насиченням тощо) й окреслюється коло функцій, які спроможні апроксимувати цей процес.

Зміна рівнів часового ряду виявляє динамічність досліджуваних процесів. Разом із динамічністю цим процесам притаманна інерційність: зберігається механізм формування явищ і характер розвитку (темпи, напрям, коливання). Діалектична єдність мінливості й сталості, динамічності та інерційності формує характер динаміки і уможливорює статистичне прогнозування досліджуваних процесів.

3.3. Екстраполяція тренду як метод прогнозування

Прогнозування динамічних процесів на основі трендових моделей ґрунтується на ідеї екстраполяції. Суть екстраполяції полягає у перенесенні на майбутнє виявлених у минулому та сьогоденні закономірностей розвитку об'єкта прогнозування. Такий підхід виходить з інерційності соціально-економічних явищ і процесів: з одного боку, протягом відносно тривалого часу зберігається характер розвитку об'єкта прогнозування (напрямок, темпи, коливання), з другого, — більш-менш сталою залишається його внутрішня структура, взаємозв'язки і співвідношення з іншими об'єктами.

Прогнозним результатом екстраполяції є очікувані в майбутньому значення характеристик процесу. Очевидно, що майбутнє неможливо спостерігати, а очікуваний результат — виміряти, його можна лише передбачити за певних умов, скажімо, «...якщо тенденція не зміниться, то...» або «...якщо станеться подія А, то ...» тощо. Якщо умови зміняться, то автоматично зміниться і результат прогнозування. Отже, статистичний прогноз, визначений за схемою «...якщо, то...», завжди умовний. Визнаючи можливий стан об'єкта (явища, процесу) у майбутньому, статистичний прогноз базується на інерційності системи, тобто на

припущенні збереження у прогнозованому періоді закономірностей і тенденцій, що склалися в «передісторії». Такий прогноз ніби абстрагується від можливих радикальних змін ситуації, його завдання — з'ясувати, чого найімовірніше можна очікувати в майбутньому за умови збереження існуючих тенденцій.

Іншою особливістю статистичного прогнозу є визначеність його в часі. Часовий горизонт прогнозу — це граничний термін, на котрий здійснюється розвідка майбутнього і з певною імовірністю визначається прогноз. Часовий горизонт прогнозу називають також *періодом попередження*.

Формально операцію екстраполяції можна подати як визначення функції

$$Y_{t+v} = f(Y_t, v),$$

де Y_{t+v} — прогнозне значення на період попередження v ;

Y_t — база прогнозування, найчастіше це останній, визначений за трендом рівень ряду.

v — період попередження.

Отже, прогнозний, очікуваний рівень Y_{t+v} залежить від бази та часового горизонту прогнозування. За тривалістю періоду попередження v прогнози класифікують на: короткострокові (до 1 року), середньострокові (до 5 років) та довгострокові (від 5 до 20 років і більше). Період попередження, своєю чергою, залежить від специфіки об'єкта прогнозування, інтенсивності динаміки, сталості виявлених закономірностей і тенденцій.

Так, припускаючи, що умови, за яких зростали обсяги інвестицій в розвиток джерел відновлюваної енергії, найближчим часом не зміняться, визначимо прогноз на 2022 рік. Базою прогнозування є теоретичний рівень 2020 р. ($Y_t = 56,0$), період упередження $v = 2$. Очікувані у 2022 р. обсяги інвестицій сягнуть 59,8 млрд дол.:

$$Y_{t+v} = 56,0 + 1,9 \cdot 2 = 59,8$$
$$\text{або } Y_{t+v} = 50,3 + 1,9 \cdot (3 + 2) = 59,8.$$

Метод екстраполяції дає точковий прогноз показника. Очевидно, що «влучення в точку» малоімовірне, адже тренду властива невизначеність, передусім через похибки параметрів. Джерелом цих похибок є обмежена сукупність спостережень y_t , кожне з яких містить випадкову компоненту e_t . Випадкова компонента буде присутня і за межами динамічного ряду, а отже її необхідно врахувати. На практиці, як правило, визначають довірчі межі

прогнозного рівня з певною ймовірністю $(1-\alpha)$. Ширина довірчого інтервалу залежить від варіації рівнів динамічного ряду навколо тренду та ймовірності висновку

$$Y_{t+v} \pm t_{1-\alpha} s_p,$$

де $t_{1-\alpha}$ — коефіцієнт довіри за розподілом Стьюдента;

s_p — похибка прогнозу.

Похибка прогнозу s_p залежить від залишкової дисперсії s_e^2 , довжини динамічного ряду (передісторії) n та періоду попередження v :

$$s_p = s_e \sqrt{\frac{n+1}{n} + \frac{3(n+2v-1)^2}{n(n^2-1)}}.$$

Отже, чим довший період передісторії, тим похибка менша, а збільшення горизонту прогнозу, навпаки, веде до зростання похибки прогнозу.

Визначимо довірчі межі очікуваного у 2022 р. обсягу інвестицій в розвиток відновлюваної енергії $Y_{t+v} = 59,8$ млрд дол. Стандартна похибка $s_e = 0,81$, коригувальний коефіцієнт дорівнює

$\sqrt{\frac{7+1}{7} + \frac{3(7+2 \cdot 2-1)^2}{7(7^2-1)}} = 1,426$, звідси похибка прогнозу $s_p = 0,81 \cdot 1,426 = 1,155$ або 1,9 % відносно очікуваного рівня. Критичне значення двобічного t -критерію для $\alpha = 0,05$ і $df = (n-2) = 7-2 = 5$ становить 2,57 (див. додаток I). Таким чином, $t_{sp} = 2,57 \cdot 1,155 \approx 3,0$, а довірчі межі прогнозу — $59,8 \pm 3,0$. З ймовірністю 0,95 слід очікувати, що у 2022 р. обсяги інвестицій в розвиток відновлюваної енергетики сягнуть щонайменше $59,8 - 3,0 = 56,8$ млрд дол.

Нагадаємо, що принципова можливість екстраполяції ґрунтується на припущенні, що умови, які визначали тенденцію у минулому, не зазнають істотних змін у майбутньому. Якщо умови зміняться, то автоматично зміниться і результат прогнозування.

У процесі обчислення інтервального прогнозу для лінійного тренду (експоненти) можна скористатися таблицями Є. Четиркіна [4]. У додатку V наведено значення $Z^* = t_{0,90} z$ для $n = 7 \dots 20$ та $v = 1 \dots 4$. Як бачимо, вони прямо пропорційні до періоду попередження v й обернено пропорційні до довжини динамічного ряду n . Тобто, збільшення горизонту прогнозу виявляється постійним розширенням довірчого інтервалу і посиленням невизначеності прогнозованого процесу. Оптимальний горизонт прогнозу визначається окремо для кожного часового ряду з урахуванням його особливос-

тей. Як правило, вона не перевищує 1/3 довжини часового ряду річних даних, а для квартальних і місячних — не більше двох років.

Прогнозування розглядають як один з інструментів мінімізації невизначеності в процесі прийняття управлінських рішень, відтак методика прогнозування передбачає систему наукових доведень, використання методів і прийомів з різним ступенем формалізації, узгодженість окремих висновків і оцінок щодо майбутнього розвитку процесу. Практична робота зі складання статистичного прогнозу полягає в тому, щоби на основі наявної інформації про стан та закономірності розвитку досліджуваного об'єкта (процесу) у минулому (передісторії), використовуючи необхідний інструментарій, здобути інформацію про його стан у майбутньому. Послідовність і зміст операцій розробки статистичних прогнозів унаочнює рис. 3.7.



Рис. 3.7. Послідовність і зміст операцій розробки статистичних прогнозів соціально-економічних явищ і процесів

Важливим етапом статистичного прогнозування є *верифікація* (від лат. *verus* — істинний) прогнозу, тобто визначення ступеня його відповідності дійсному стану об'єкта в майбутньому. При верифікації прогнозів вирізняють такі поняття, як точність і надійність (достовірність) прогнозу. На етапі верифікації використовують різні критерії, способи і процедури, які дозволяють усебічно оцінювати якість прогнозу.

Критерії точності і надійності прогнозів

Точність прогнозу виявляється розбіжностями між прогнозними і фактичними даними (реалізацією прогнозу) на часовому інтервалі періоду попередження. Про точність прогнозу прийнято судити за величиною похибки прогнозу, тобто точність прогнозу — це поняття прямо протилежне до похибки прогнозу. Говорячи про високу точність прогнозу, оперують малою похибкою, й навпаки, якщо похибка велика, то точність прогнозу невисока.

Надійність (достовірність) прогнозу визначається імовірністю реалізації відповідної прогностичної оцінки у заданий часовий інтервал. Чим вона вище, тим вище і надійність. Імовірність реалізації прогнозу, який ґрунтується на статистичній моделі, пов'язують з довірчим інтервалом.

Коли прогноз подають у вигляді точкової оцінки, мірою точності прогнозу використовують середня абсолютна похибка (*MAD*): в одиницях вимірювання показника, розрахунок їх ґрунтується на відхиленнях між фактичними y_t і прогнозними Y_t даними:

– середня абсолютна похибка (*MAD*):

$$MAD = \frac{\sum |y_t - Y_t|}{v};$$

– середня квадратична, або стандартна, похибка (*RMSE*):

$$RMSE = \sqrt{\frac{\sum (y_t - Y_t)^2}{v}},$$

де v — період попередження.

Критичних меж коливання середніх абсолютних похибок прогнозу немає, прогноз тим кращий, чим менші значення *MAD* і *RMSE*. Оскільки значення цих похибок залежать від масштабу

вимірювання показника, задля порівняння точності прогнозів різної менних процесів або прогнозів, визначених за різними моделями, використовують відносні похибки, виражені процентами. Серед них похибка апроксимації *MAPE* (*Mean Absolute Percentage Error*):

$$MAPE = \frac{100}{v} \sum \left| \frac{y_t - Y_t}{y_t} \right|.$$

Для високоточних прогнозів $MAPE < 10\%$, в інтервалі значень від 10 до 20 %, точність прогнозу вважають задовільною. Отже, коли стосовно конкретного показника значення *MAPE* відповідає певному критерію точності, скажімо, не перевищує 5 чи 10 %, то прогнозна модель вважається прийнятною й рекомендується для практичного використання.

Мірою якості прогнозів визнають також коефіцієнт розбіжності (невідповідності) Г. Тейла:

$$U = \frac{\sqrt{\sum (y_t - Y_t)^2}}{\sqrt{\sum y_t^2}}.$$

Особливість коефіцієнта Тейла полягає в тому, що його можна використати для порівняння якості прогнозів, визначених на основі різних методів і моделей. Так, $U = 0$, коли всі прогнозні й фактичні значення збігаються (випадок ідеального прогнозування), $U = 1$, коли точність прогнозу за трендом така сама, як і «найвна» екстраполяція на основі характеристик інтенсивності динаміки, $U > 1$, коли трендова модель дає гірший результат, верхньої межі коефіцієнт U не має.

Імовірність реалізації прогнозу, який ґрунтується на статистичній моделі, пов'язують з довірчим інтервалом. Більшість наявних методів верифікації прогнозів оперують статистичними процедурами, які зводяться до побудови довірчих меж прогнозу, себто до побудови інтервальних прогнозів. Відношення кількості випадків, коли фактична реалізація потрапляє у довірчий інтервал прогнозу, до загальної кількості прогнозів також розглядають як міру якості прогнозу:

$$u = \frac{p}{p + q},$$

де p і q — кількість прогнозів, підтверджених і непідтверджених фактичними даними.

Якщо всі прогнози підтверджуються, $q = 0$ і $u = 1$, коли не підтверджуються, то p і u дорівнюють нулю.

Розглянувши сукупність критеріїв, способів та процедур оцінювання точності та надійності прогнозу, можна дійти висновку, що на етапі верифікації більшою мірою здійснюється оцінка методу прогнозування, за допомогою якого отримано результат, ніж оцінка якості самого результату. Досі не знайдено ефективного підходу до оцінки якості прогнозу до його реалізації.

Питання і задачі для самопідготовки

1. Яку роль відіграє згладжування в аналізі детермінованої компоненти часового ряду?

2. Поясніть логіку згладжування часового ряду методом ковзних середніх.

3. Скільки тричленних ковзних середніх можна обчислити в часовому ряду довжиною $n=15$?

4. Поясніть відмінності і сфери застосування зваженою і незваженою ковзних середніх.

5. У чому сутність адаптивної ковзної середньої.

6. Поясніть сферу застосування і види *кривих зростання*.

7. Яку функцію слід застосувати для вирівнювання часових рядів з відносно стабільними: а) абсолютними приростами? б) темпами приросту?

8. Поквартальне нарощення валютних коштів, розміщених на рахунках у банківських установах регіону (млн дол) описує функція $Y_t = 48,2 * 1,012^t$. Поясніть зміст параметрів.

9. Поясніть зміст параметрів параболи $Y_t = a + bt + ct^2$, окресліть сферу її використання.

10. Окресліть принципову можливість екстраполяції виявленого тренду і характерні риси статистичних прогнозів,

11. Від чого залежить часовий горизонт прогнозування?

12. Розкрийте поняття точності статистичних прогнозів. Як співвідносяться точність і надійність прогнозу?

13. Яку функцію в оцінюванні точності статистичних прогнозів виконує похибка апроксимації?

14. Динаміка кількості перевезених пасажирів авіа транспортом характеризується даними (тис.)

Рік, № з/п	Перевезено пасажирів	Рік, № з/п	Перевезено пасажирів	Рік, № з/п	Перевезено пасажирів
1	1915	10	3229	19	8107
2	1724	11	3813	20	6473
3	1485	12	4351	21	6303
4	1164	13	4929	22	8278
5	1087	14	6181	23	10556
6	1164	15	5131	24	12529
7	1290	16	6107	25	13706
8	1768	17	7505	26	8107
9	2375	18	8106	27	4798

Засобами MS Excel зобразіть часовий ряд графічно, здійсніть згладжування ряду методом ковзної середньої (період згладжування — 3, 4 і 5 років). Зробіть висновки.

15. Динаміка курсу акцій компанії Tesla характеризується даними, дол.:

День торгів	Курсова ціна	День торгів	Курсова ціна
1	654	7	649
2	656	8	663
3	651	9	666
4	637	10	690
5	603	11	686
6	627	12	676

Виконайте згладжування часового ряду засобами MS Excel:

а) методом 4 -членної простої ковзної середньої;

б) методом 4 -членної зваженої ковзної середньої з вагами (0,5; 0,3; 0,15; 0,05)

За кожним методом визначте прогнозну ціну акції на 13 день торгів. Зробіть висновки.

16. Динаміка кількості передплатників міжнародного стрімінгового сервісу Netflix (млн осіб) описується лінійним трендом $Y = 108,1 + 7,87t$, де $t = 1, 2, \dots, 7$, зі стандартною похибкою 0,275. Визначте прогнозну кількість підписників на період попередження $v = 1$ і довірчі межі прогнозу з імовірністю 0,90.

17. Розвиток ринку агроробототехніки (млн дол) за останнє десятиліття описується функцією $Y_t = 20,1e^{0,04t}$. Поясніть зміст параметрів моделі і визначте прогнозний обсяг ринку на наступний рік.

18. Динаміка кількості поштових відправлень (посилки) характеризується даними, тис.

Рік № з/п	Кількість відправлень	Рік	Кількість відправлень	Рік № з/п	Кількість відправлень
1	32	8	205	15	248
2	38	9	209	16	277
3	30	10	191	17	329
4	35	11	239	18	299
5	101	12	230	19	381
6	138	13	234	20	388
7	159	14	234	21	411

Здійсніть згладжування часового ряду методом зваженої ковзної середньої, скористайтесь процедурами модуля *Time Series / Forecasting* системи Statistica, опція 4253H *Filter* 3-точкове згладжування. Зобразіть часовий ряд графічно, зробіть висновки

19. Складено два варіанти прогнозу єдиного податку для фізичних осіб-підприємців у регіоні (млн грн.) у першому кварталі нового року. Використовуючи стандартну похибку, оцініть точність прогнозу за кожним варіантом, зробіть висновок.

Прогнозний період	Прогнозний рівень визначений на основі моделі		Фактичний рівень
	поліном першого ступеня	експоненційне згладжування	
Січень	315	302	312
Лютий	360	314	340
Березень	355	346	362

20. За даними поквартальної динаміки надходження заявок на охорону винаходів до патентних відомств складено прогнози на період попередження $v = 1, 2, 3$. Фактичне надходження заявок відхиляється від прогнозованого. Використовуючи похибку апроксимації, порівняйте точність прогнозів, зробіть висновки.

Прогнозний період	За прогнозом		Фактично	
	патенти	товарні знаки	патенти	товарні знаки
$p + 1$	604	4115	548	3902
$p + 2$	612	4178	536	3876
$p + 3$	618	4242	524	4115

21. Страхові виплати у ризикових видах страхування характеризуються даними:

№ року з/п	1	2	3	4	5	6	7	8	9
Страхові платежі, млн грн	137,6	141,3	144,8	145,7	150,4	149,5	150,8	155,1	156,4

Опишіть тенденцію ряду поліномом першого ступеня, поясніть зміст параметрів. Визначте прогнозні рівні надходження страхових платежів на період попередження $v = 1, 2, 3$, стандартну похибку і довірчі межі прогнозу з імовірністю 0,95. Оцініть автокореляцію залишкових величин з лагом $p = 1$, зробіть висновки про адекватність тренду реальному процесу виплат у ризикових видах страхування.

22. Динаміка експорту соняшникової олії характеризується даними, тис. т:

№ року з/п	1	2	3	4	5	6	7	8	9
Експорт олії	86	92	97	106	115	128	142	154	170

Опишіть тенденцію ряду експонентою, поясніть зміст параметрів. Визначте прогнозні рівні експорту олії на період попередження $v = 1, 2, 3$ і стандартну похибку прогнозу.

23. Доходи компанії Google Alphabet від реклами (млрд дол США) становили:

2002	2003	2004	2005	2006	2007	2008	2009	2010
0,4	1,5	3,2	6,1	10,6	16,6	21,8	23,8	29,3
2011	2012	2013	2014	2015	2016	2017	2018	2019
37,9	50,18	55,51	65,67	74,54	89,96	110,38	136,22	147,9

Опишіть тенденцію ряду експонентою, поясніть зміст параметрів. Припускаючи, що виявлена тенденція збережеться, визначте очікувані доходи компанії у наступні 2 роки, розрахуйте стандартну похибку прогнозу.

24. За помісячними даними Державної податкової служби податковий борг і надані розстрочки становили:

Місяць	I	II	III	IV	V	VI	VII	VIII
Сума податкового боргу, млрд грн	104,0	101,8	101,6	102,2	101,9	101,6	107,1	106,8
Розстрочки, кількість	3	6	107	184	206	199	195	219

Опишіть тенденції наведених часових рядів поліномом 2-го ступеня, поясніть зміст параметрів, для кожного ряду обчисліть екстремуми, зробіть висновки.

25. За даними трьох фірм про обсяги експорту продукції (млн дол.) для кожної з них обґрунтуйте характер динаміки і оберіть найбільш адекватний тип кривої зростання. Припускаючи, що характер динаміки не зміниться, обчисліть прогнозні обсяги експорту продукції на 2022 рік. Оцініть точність прогнозів.

Фірма	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
A	66,7	63,4	58,5	56,9	53,1	51,3	50,3	48,3	46,8	46,6
B	67,1	63,8	58,9	57,3	63,5	66,7	60,7	48,7	47,2	47,0
C	67,3	63,9	59,0	57,4	53,6	51,8	50,9	48,8	47,4	47,1

Рекомендована література

1. Галушак М. П., Галушак О. Я., Кужда Т. І. Прогнозування соціально-економічних процесів : навч. посіб. Тернопіль, 2021. С. 74—94, 145—151.

2. Єріна А.М., Єрін Д.Л. Статистичне моделювання та прогнозування : підручник. Київ, КНЕУ. 2014. С. 129—164.

3. Методи і моделі економічного прогнозування : навч. посіб. / В.П. Кічор та ін. Львів, 2019. С 61—102.

4. Четыркин Е.М. Статистические методы прогнозирования. Москва, 1977. 199 с.

АНАЛІЗ І ПРОГНОЗУВАННЯ ЗАКОНОМІРНОСТЕЙ ЦИКЛІЧНОГО ХАРАКТЕРУ

- 4.1. Аналіз циклічних закономірностей розвитку на основі S-кривих
- 4.2. Варіанти аналізу циклів оновлення

4.1. Аналіз циклічних закономірностей розвитку на основі S-кривих

Поліноми та проста експонента забезпечують задовільні результати екстраполяції в межах еволюційного етапу розвитку. Коли ж еволюційні етапи чергуються з революційними, аналіз і прогнозування закономірностей розвитку потребують моделей іншого типу. Наприклад, на товарному ринку попит на новий товар попервах незначний, потім після визнання споживачами, він стрімко зростає, але по мірі насичення ринку темпи зростання уповільнюються, згасають, попит стабілізується на певному рівні. Аналогічні фази розвитку мають процеси нововведень та винаходів, ефективність використання ресурсів тощо.

Для динамічних процесів такого типу характерний ефект насичення, коли з плином часу темпи зростання (зниження) обсягів уповільнюються і, наближаючись до певної межі, стабілізуються. У практиці статистичного моделювання обмеженої динаміки використовують клас кривих, що мають горизонтальну асимптоту $K \neq 0$. Серед них модифікована експонента, крива Гомпертца, логістична крива (Перла-Ріда).

Модифікована експонента (рис. 4.1) на відміну від простої містить додаткову складову — асимптоту функції K , яка визначає межу розвитку процесу:

$$Y_t = K + ab^t$$

де a — різниця між граничним значенням K і досягнутим рівнем; коли $a < 0$, асимптота знаходиться вище кривої, коли $a > 0$ — нижче кривої;

b характеризує співвідношення послідовних приростів ординати; за умови рівномірного розподілу ординат по осі часу ці співвідношення будуть сталими:

$$b = \frac{Y_t - Y_{t-1}}{Y_{t-1} - Y_{t-2}} = \text{const.}$$

При вирішенні економічних завдань найчастіше доводиться мати справу з кривою, у якій $a < 0$, $b < 1$. У цьому випадку зростання рівнів відбувається з уповільненням і наближається до певної межі (рис. 4.1).

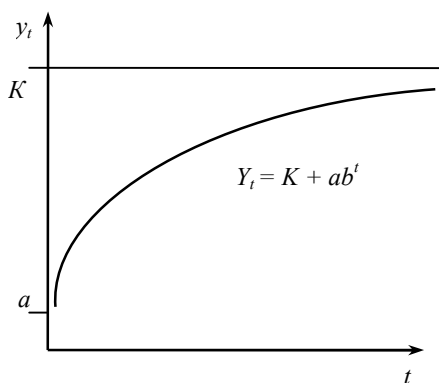


Рис. 4.1. Модифікована експонента при $a < 0$ і $b < 1$

Як свідчить практика статистичного моделювання процесів з обмеженою динамікою, модифіковану експоненту використовують переважно у дослідженнях процесів, на які діють певні обмежувальні, стримуючі фактори внутрішнього і зовнішнього середовища, і вплив цих факторів зростає зі збільшенням досягнутого рівня Y_t . Для економічного процесу такими обмеженнями може бути брак ресурсів чи виробничих потужностей, ємність ринку, поява нових конкурентів, правові рамки тощо.

У разі, коли обмежувальні фактори впливають лише після певного моменту, до якого процес розвивався за експоненціальним законом, такий процес апроксимується S-подібною кривою з точкою перегину, у якій прискорене зростання змінюється уповільненням. Схематично S-подібний процес можна поділити на дві приблизно однакові частини AC і CB (рис. 4.2). До точки C прирости ординати нижчі за пряму AB, що пояснюється опором середовища, наприклад, споживачі не готові сприйняти нововве-

дення. Ближче до точки C приріст прискорюються, досягає максимуму, і S -крива перетинає пряму AB . Наближаючись до асимптоти, прирости уповільнюються і поступово згасають.

По суті S -подібні криві описують два послідовні процеси: один з прискоренням розвитку, другий — з уповільненням і переходом до насичення. Асимптота K визначає межу розвитку процесу, точка перегину C — кульмінаційний момент розвитку. Здатність S -подібних моделей відображати діалектику взаємодії еволюційного й революційного розвитку надзвичайно важлива в дослідженнях процесів, яким властива послідовна зміна стадій повного циклу розвитку: зародження, поширення (дифузія), насичення.

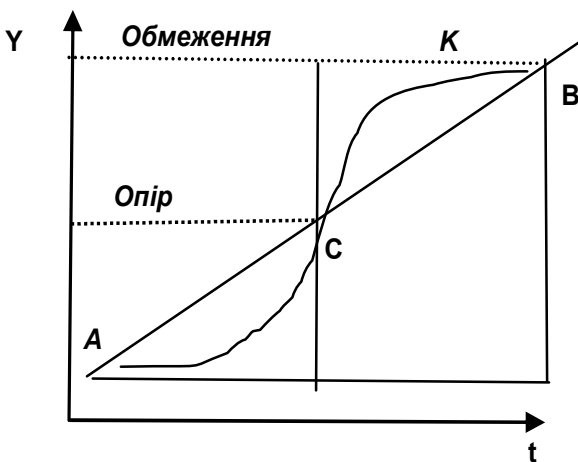


Рис. 4.2. Графік S -подібної кривої

Серед S -подібних кривих найбільш поширені крива Гомпертца і логістична крива. Криву Гомпертца застосовують переважно в демографічній і страховій статистиці, логістичну криву — в статистиці торгівлі, в маркетингу, при вивченні ринкових та інноваційних процесів.

Алгебраїчно крива Гомпертца має вигляд:

$$Y_t = K a^{b^t},$$

де K — асимптота функції, a , b — параметри, причому $b < 1$.

Крива Гомпертца описує динамічні процеси, у яких відношення послідовних приростів ординат в логарифмах залишаються-

ся незмінними за рівномірного розподілу ординат на часовій осі t і зводиться до модифікованої експоненти логарифмуванням $\lg Y_t = \lg K + b' \lg a$, тож характерними для неї є постійні відношення приростів ординат у логарифмах. Точка перегину становить $t_p = \ln(-\frac{1}{\ln a}) / \ln b$ зі значенням функції $Y_t = K / e$.

Логістичну функцію найчастіше записують у вигляді модифікованої експоненти заміною Y_t на $1/Y_t$. У такому разі

$$Y_t = \frac{K}{1 + be^{-at}}.$$

де e — основа натурального логарифма, a, b — параметри кривої; K — граничне значення функції. Замінивши натуральний логарифм на десятковий, отримаємо інший вид логістичної функції:

$$Y_t = \frac{K}{1 + 10^{a+bt}}.$$

Коли показник процесу — частка, яка змінюється в межах від 0 до 1, то формула логістичної функції спрощується:

$$Y_t = \frac{1}{e^{a+bt} + 1} \quad \text{або} \quad \frac{1}{Y_t} = 1 + e^{a+bt},$$

де e — основа натурального логарифма; a, b — параметри кривої.

В логістичній кривій точкою перегину є $t_p = \frac{1}{\ln b} \ln \frac{K}{a}$ зі значенням функції $Y_t = 1/2K$.

Отже, обидві S-подібні криві — Гомпертца і логістична — описують повні цикли розвитку динамічних процесів. Відмінність між ними в тому, що у кривій Гомпертца сталими є відношення перших різниць логарифмів до рівновіддалених одна від одної ординат, а в логістичній кривій стали відношення перших різниць обернених значень ординат. При цьому логістична крива центрально симетрична відносно точки перегину, а крива Гомпертца — асиметрична. Для обох функцій базовою кривою слугує модифікована експонента.

Для оцінювання параметрів S-кривих запропоновано низку процедур. Серед них спрощені методи (метод трьох сум і метод трьох точок), звичайний метод найменших квадратів (OLS), нелі-

нійний метод найменших квадратів (NLS), оцінки на основі максимальної правдоподібності (MLE) та інші.

Розглянемо спрощений метод оцінювання параметрів S-кривих методом трьох сум. Суть його полягає в тому, що часовий ряд поділяють на три періоди однакової довжини m , для кожного періоду визначають суми: $\Sigma_1 y$, $\Sigma_2 y$, $\Sigma_3 y$, які власне є базою розрахунку параметрів функції. В разі, коли довжина ряду $n = 3m + 1$, то $(m + 1)$ включають до другого періоду, а перший і третій періоди містять по m елементів. Коли $n = 3m + 2$, тоді m елементів включають до другого періоду, а $(m + 1)$ — до першого і третього.

За правилом трьох сум спершу визначається параметр b , а потім два інші — a і K . Оскільки логістична крива і крива Гомпертца за допомогою певних перетворень приводяться до модифікованої експоненти, розглянемо порядок розрахунку параметрів модифікованої експоненти. Формули розрахунку мають такий вигляд:

$$b = \sqrt{\frac{\Sigma_3 y - \Sigma_2 y}{\Sigma_2 y - \Sigma_1 y}};$$

$$a = (\Sigma_2 y - \Sigma_1 y) \frac{b - 1}{(b^m - 1)^2};$$

$$K = \frac{1}{m} \left[\frac{\Sigma_1 y \times \Sigma_3 y - (\Sigma_2 y)^2}{\Sigma_1 y + \Sigma_3 y - 2 \Sigma_2 y} \right].$$

В табл. 4.1 наведена динаміка експорту продукції фабрикою хутряних виробів за 12 років ($m = 4$).

Таблиця 4.1

ДИНАМІКА ЕКСПОРТУ ХУТРЯНИХ ВИРОБІВ, млн євро

Рік	Млн євро	Рік	Млн євро	Рік	Млн євро
2009	22,4	2013	35,8	2017	54,5
2010	24,8	2014	40,4	2018	55,8
2011	26,0	2015	47,6	2019	56,2
2012	28,5	2016	56,2	2020	55,7

За даними таблиці: $\Sigma_1 y = 101,7$; $\Sigma_2 y = 180,0$; $\Sigma_3 y = 222,2$. Звідси:

$$b = \sqrt[4]{\frac{222,2 - 180,0}{180,0 - 101,7}} = \sqrt[4]{\frac{42,2}{78,3}} = \sqrt[4]{0,53895} = 0,857;$$

$$a = (180,0 - 101,7) \frac{0,857 - 1}{(0,53895 - 1)^2} = -52,7;$$

$$K = \frac{1}{4} \left[\frac{101,7 \times 222,2 - 180,0^2}{101,7 + 222,2 - 2 \times 180,0} \right] = 67,9.$$

Модель має вигляд $Y_t = 67,9 - 52,7 \times 0,857^t$, тобто, високі темпи нарощення експорту хутряних виробів з часом уповільнюються і, наближаючись до рівня насиченості (67,9 млн євро), стабілізуються.

Як свідчить практика моделювання повних циклів, найкращі передбачення і найбільш ефективні оцінки параметрів (з точки зору стандартної похибки) дає NLS (The nonlinear least squares). В системі *Statistica* оцінювання можна здійснити в модулі *Nonlinear Estimation*. Вид моделі користувач визначає самостійно за процедурою *User-specified regression, custom loss function*.

Практика застосування S-кривих для вивчення динамічних процесів з урахуванням стадій розвитку досить поширена. Сучасна модель економічного зростання актуалізує дослідження інноваційних процесів, розвиток яких зазвичай рухається за логістичною кривою. Описуючи загальну форму траєкторії інноваційного циклу, S-подібна крива формалізує логіку процесу і фіксує складну динаміку його розвитку. За допомогою логістичної кривої можуть бути представлені процеси як дифузії, так і заміщення технології, перехід від одного покоління технології до іншого. Об'єднуючи різні сегменти інноваційного процесу в одну загальну тенденцію, логіста забезпечує можливість розв'язання найбільш складних завдань прогнозування: передбачення можливих змін у розвитку процесу, визначення ймовірності їх успішного здійснення у визначений термін і можливості використання цієї інформації при обґрунтуванні управлінських рішень. Рисунок 4.3. ілюструє процес заміщення технології та перехід до виробництва нового товару учасниками інноваційного процесу в контексті реалізації ними окремих інноваційних проєктів.

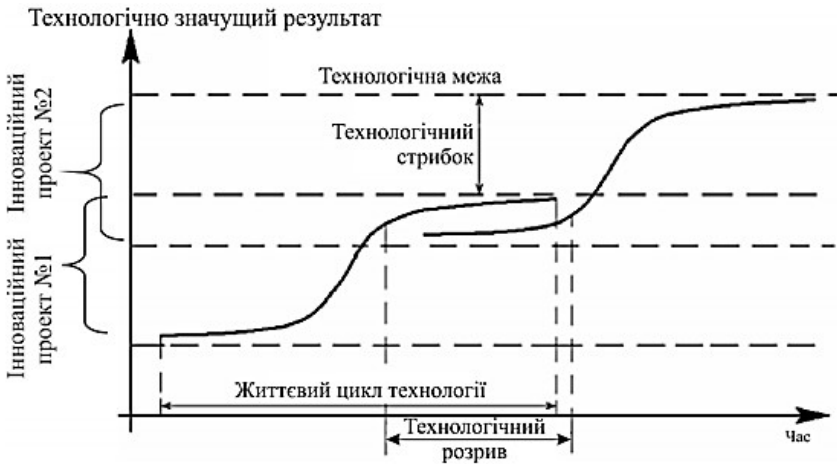


Рис. 4.3. Процес заміщення технологій

Джерело: [5]

В контексті інтенсивного розвитку інноваційного циклу найбільш важливим є другий сегмент, саме він визначає ступінь конкурентоспроможності і внутрішній потенціал обраної інновації. Управління інноваційним процесом стосовно першого сегменту спрямовано на зменшення його тривалості, стосовно третього — на пошук нових технологій. Таким чином, визначення меж між сегментами логістичної кривої є одним з першорядних завдань формування інноваційної стратегії тих суб'єктів господарювання, які спроможні генерувати ідеї і здійснювати процес їх комерціалізації на усіх стадіях інноваційного циклу.

Отже, клас моделей, які спроможні описати складні динамічні процеси, досить широкий. Вибір типу моделі у конкретному дослідженні ґрунтується передусім на теоретичному аналізі специфіки процесу, його внутрішньої структури, взаємозв'язків з іншими процесами. На основі такого аналізу в загальних рисах визначається характер динаміки (рівномірний, рівноприскорений, з насиченням тощо) й окреслюється коло функцій, які спроможні апроксимувати цей процес.

Серйозною підмогою у виборі конкретної моделі слугують формальні методи, скажімо, аналіз послідовних різниць. Сталість різниць p -ого порядку розглядається як симптом того, що процес описується поліномом p -го порядку. Якщо приблизно однакові різниці 1-го порядку $\Delta'_t = y_t - y_{t-1}$, використовують лінійний тренд,

якщо однакові різниці 2-го порядку $\Delta''_t = \Delta_t - \Delta_{t-1}$, — параболу і т. д. Певні складнощі можуть виникнути при виборі експоненти, адже S-подібна крива до точки перегину описує експоненціальний тренд, а сама точка перегину може бути за межами часового ряду. В разі, коли межа насичення теоретично можлива і процес у майбутньому може згасати або існують певні обмеження для його подальшого розвитку (правові, матеріальних ресурсів, виробничих потужностей тощо), перевага надається S-подібним кривим. У табл. 4.3 наведено основні апріорні тести, за якими визначається тип моделі для конкретного динамічного процесу.

Таблиця 4.3

АПРІОРНІ ТЕСТИ ВИБОРУ ФУНКЦІОНАЛЬНОГО ВИДУ ТРЕНДОВОЇ МОДЕЛІ

Характеристика	Властивості характеристик	Вид трендової моделі
Δ'_t	Майже однакові	Поліном 1-го ступеня
Δ''_t	Змінюються лінійно	Поліном 2-го ступеня
Δ'_t / y_{t-1}	Майже однакові	Проста експонента
$\lg \Delta'_t$	Змінюються лінійно	Модифікована експонента
$\lg[\Delta'_t / y_{t-1}]$	Змінюються лінійно	Крива Гомпертца
$\lg[\Delta'_t / (y_{t-1})^2]$	Змінюються лінійно	Логістична крива

Коли первинному часовому ряду властива значна варіація рівнів y_t , перед аналізом послідовних різниць проводять згладжування ряду методом ковзних середніх. У разі спадної тенденції послідовні різниці розраховують, починаючи з кінця.

Очевидно, що клас трендових моделей не вичерпується розглянутими функціями. Можна експериментувати з перетвореннями часового ряду, застосовуючи кусочно-лінійні сплайни з точками розриву, на різних ділянках ряду одночасно підганяти різні функції і т. ін.

4.2. Варіанти аналізу циклів оновлення

Проблема оновлення безпосередньо пов'язана з дослідженням процесів простого і розширеного відтворення властивостей динамічних і складних за структурою систем. Оновлюваною (англ. *renewable*) називають сукупність, склад якої постійно оновлюється

ся — одні елементи вибувають, інші з'являються. У кожному одиницю часу t оновлювана сукупність об'єднує множину елементів (одиниць), які надійшли в різний час і мають різний вік. Процес оновлення пов'язаний із тривалістю перебування окремих елементів (одиниць) у межах сукупності. Вік (чи термін служби) окремих елементів сукупності — величина випадкова, оскільки на її значення впливає безліч причин, а процес вибуття має стохастичний, ймовірнісний характер. Дослідивши закономірність розподілу імовірностей вибуття одиниць сукупності залежно від віку, можна визначити кількість одиниць, що доживають до певного віку, ймовірність досягнення того чи іншого віку, середню тривалість майбутнього перебування в даному віці та інші характеристики. Методами статистики визначаються ймовірності вибуття елементів різного віку, і на цій основі складаються таблиці вибуття.

Класичним прикладом оновлюваної сукупності є людська популяція. Смертність і народжуваність — це процеси заміни старших поколінь, які поступово вимирають, новими. Процеси оновлення поколінь описує матрична модель *відтворення населення*. Історично саме в демографічній, а згодом в актуарній (страховій) статистиці набув розвитку метод прогнозування на основі матричної моделі відтворення населення, яка ґрунтується на даних демографічних таблиць смертності або дожиття.

На відміну від екстраполяції, матрична модель відтворення населення, описуючи процес оновлення поколінь, встановлює зв'язок між чисельністю і віковою структурою населення, з одного боку, та режимом народжуваності і порядком вимирання, — з другого. До класу матричних моделей, що враховують вікову інтенсивність демографічних процесів, належить дискретна модель вікових пересувань. Довжину періоду, на який проводиться одне пересування, називають *кроком пересувань*. Найчастіше цей крок становить один рік.

Перспективні розрахунки чисельності та складу населення методом вікових пересувань ґрунтуються на матеріалах переписів населення та вікових коефіцієнтах смертності й народжуваності. Розрахунки виконують окремо для чоловіків і окремо для жінок через різний рівень їхньої смертності. Стартують з даних останнього перепису населення, тобто рік проведення перепису є *базовим, нульовим*. Кожного наступного року особи, враховані під час перепису, стануть на рік старші, а їхня чисельність зменшиться на кількість тих, хто не доживе до наступного року. Імовірності

дожиття та вимирання одночасно живущих різних поколінь містяться в таблицях смертності (дожиття), складених методом умовного покоління (базова чисельність $l_x = 100\ 000$). Зміст і взаємозв'язок коефіцієнтів смертності і дожиття з іншими показниками демографічних таблиць наведено в табл. 4.4.

Таблиця 4.4

ОСНОВНІ ПОКАЗНИКИ ТАБЛИЦІ СМЕРТНОСТІ АБО ДОЖИТТЯ

Вік населення, років	Кількість осіб, які доживають до віку x	Імовірність померти в інтервалі віку від x до $(x + 1)$	Імовірність дожити в інтервалі віку від x до $(x + 1)$	Кількість осіб, які помирають в інтервалі віку від x до $(x + 1)$
x	l_x	q_x	p_x	d_x
0	$l_0 = 100\ 000$	$q_x = d_x : l_x$	$p_x = l_{x+1} : l_x$	$d_x = l_x - l_{x+1}$
1	$l_{x+1} = l_x - d_x$			
...				
100				
Взаємозв'язок показників	$l_{x+1} = l_x \times p_x$	$q_x = 1 - p_x$	$p_x = 1 - q_x$	$d_x = l_x \times q_x$

Перспективна чисельність населення кожної вікової групи S_{x+1} визначається за формулою¹

$$S_{x+1} = S_x P_x,$$

де S_x — чисельність населення x -вікової групи,

P_x — коефіцієнт дожиття до наступного $(x + 1)$ віку, на відміну від імовірності дожиття p_x , визначається не числами тих, хто доживає до віку $(x + 1)$, а середніми числами осіб, які живуть в інтервалі віку від x до $(x + 1)$:

$$P_x = \frac{L_{x+1}}{L_x},$$

¹ В основі методів демографічних таблиць лежить фундаментальна передумова замкнутості (закритості) покоління, відповідно до якої міграція населення не враховується.

де L_x розраховують у кожній x -віковій групі як середню арифметичну просту: $L_x = \frac{l_x + l_{x+1}}{2}$.

Скажімо, якщо на момент останнього перепису вихідна чисельність населення x -вікової групи становила S_x^0 , то через рік, у віці $(x + 1)$, буде становити $S_{x+1} = S_x^0 P_x$, де P_x — коефіцієнт дожиття в інтервалі віку від x до $(x + 1)$. Процедура пересування застосовують до всіх вікових груп реального населення, крім нульової. Чисельність нульової групи визначається на основі прогнозних рівнів народжуваності і смертності немовлят.

Схему розрахунку кількості тих, хто житиме у кожному наступному році прогнозного періоду, серед тих, хто врахований переписом населення, наведено у табл. 4.5. Описана процедура ітеративно повторюється стільки разів, скільки років охоплює прогнозний період. Прогнозну чисельність усього населення за окремими віковими групами визначають як суму відповідної чисельності чоловіків і жінок

Таблиця 4.5

СХЕМА ПЕРЕСУВАННЯ ВІКОВИХ ГРУП

Вік, x	Коефіцієнт дожиття, P_x	Вихідна чисельність населення, S_x^0	Прогнозна чисельність населення		
			S_x^1	S_x^2	S_x^3
0	P_0	S_0^0			
1	P_1	S_1^0	$S_0^0 P_0$		
2	P_2	S_2^0	$S_1^0 P_1$	$S_1^1 P_1$	
3	P_3	S_3^0	$S_2^0 P_2$	$S_2^1 P_2$	$S_2^2 P_2$
4	P_4	S_4^0	$S_3^0 P_3$	$S_3^1 P_3$	$S_3^2 P_3$
5	P_5	S_5^0	$S_4^0 P_4$	$S_4^1 P_4$	$S_4^2 P_4$

Методологічно демографічний прогноз виступає у вигляді оцінки майбутньої чисельності та статеві-вікової структури населення за даними про динаміку демографічних параметрів у минулому з урахуванням прийнятих гіпотез щодо їхньої динаміки у

майбутньому. Тобто в процесі прогнозування минулі тенденції переносяться на майбутнє. Обґрунтованість і точність прогнозу залежить від точності оцінок стартового населення і точності передбачуваних у майбутньому параметрів народжуваності і смертності.

Аналогічні процеси оновлення відбуваються в сукупностях матеріальних предметів, скажімо введення в експлуатацію та виведення з експлуатації основних засобів, товарів тривалого користування тощо. Методами статистики можна визначити термін служби об'єктів певного типу, можна встановити «порядок вибуття», тобто з'ясувати, скільки об'єктів, уведених в експлуатацію одночасно, необхідно буде вивести з експлуатації через одну, дві, три і більше встановлених одиниць часу (день, тиждень, місяць, рік). Визначаючи термін служби засобів, припускають, що в процесі їх експлуатації буде забезпечено всі умови для нормального функціонування (технічний огляд, ремонт тощо).

Статистичними методами, за аналогією з таблицями смертності населення, складають таблиці вибуття більш-менш однорідних за терміном служби об'єктів. При цьому коефіцієнти дожиття і коефіцієнти вибуття об'єктів розглядають як імовірності настання певних подій. Коефіцієнт вибуття p_t визначається відношенням кількості об'єктів, що вибули з експлуатації в інтервалі часу від t до $(t + 1)$, тобто $(N_t - N_{t+1})$, до кількості початково введених об'єктів N_0 :

$$p_t = \frac{N_t - N_{t+1}}{N_0}.$$

За умови простого відтворення, щоб урівноважити виведення об'єктів з експлуатації, необхідно в період t ввести в експлуатацію $N_0(t)$ об'єктів:

$$N_0(t) = N_0(t-1)p_1 + N_0(t-2)p_2 + \dots + N_0(t-w)p_w,$$

де w — верхня межа терміну експлуатації, за цей період процес оновлення відбувається у повному обсязі.

Це основне рівняння оновлення, в якому $N_0(t)$ визначається як зважена сума об'єктів, уведених в експлуатацію в попередні періоди. Вагами виступають відповідні коефіцієнти вибуття p_t , сума ваг дорівнює одиниці. На основі рівняння оновлення можна пос-

лідовно визначити кількість об'єктів, що підлягають заміні в періоди $t = 1, 2, \dots, (w-1)$:

$$N_0(1) = N_0(0) p_1$$

$$N_0(2) = N_0(1) p_1 + N_0(0) p_2$$

.....

$$N_0(w-1) = N_0(w-2) p_1 + N_0(w-3) p_2 + \dots + N_0(0) p_{w-1}$$

У табл. 4.6 наведено приклад застосування рекурентних рівнянь оновлення для розрахунку кількості об'єктів, що підлягають заміні. Припустимо, що початково введено в експлуатацію 1000 приладів, максимальний термін служби яких три роки ($w = 3$), а коефіцієнти вибуття становлять: $p_1 = 0,2$; $p_2 = 0,3$; $p_3 = 0,5$. У другій графі таблиці наведено кількість нових приладів, які належить ввести в експлуатацію в окремі роки.

Таблиця 4.6

РОЗРАХУНОК КІЛЬКОСТІ ПОНОВЛЮВАНИХ ПРИЛАДІВ

Термін служби, t	$N_0(t)$	Кількість вибулих об'єктів, які необхідно поновити		
		$p_1 = 0,2$	$p_2 = 0,3$	$p_3 = 0,5$
A	1	2	3	4
1	1000	100	200	500
2	100	20	30	50
3	220	44	66	110
4	574	115	172	287
5	231	46	69	116
6	328	66	98	164
7	422	84	127	211

Графи 2—4 показують, скільки з цих об'єктів необхідно буде замінити після першого року, після двох і після трьох років експлуатації. Наприклад, через чотири роки служби заміні підлягає 574 прилади, з яких 44 було введено за рік до того ($220 \times 0,2$), 30 — за два роки ($100 \times 0,3$) і 500 — за три роки ($1000 \times 0,5$). У таблиці кількість приладів, які щороку підлягають заміні $N_0(t)$, можна визначити підсумовуванням чисел по косих лініях.

Як бачимо, динаміка оновлень зазнає коливань. Це так звані цикли оновлення, вони зумовлені концентрацією введення в експлуатацію нових об'єктів. Оскільки цикли оновлення тягнуть за собою коливання попиту і ринкової кон'юнктури, у системі управління важливо передбачити шляхи згладжування цих циклів. За умови розширеного відтворення методика моделювання оновлюваних сукупностей дещо ускладнюється.

Питання і задачі для самопідготовки

1. Наведіть приклади циклічних закономірностей розвитку соціально-економічних процесів.
2. Назвіть складові елементи життєвого циклу галузі, корпорації, товару.
3. У чому полягає особливість статистичного моделювання повного циклу розвитку з ефектом насичення?
4. Окресліть сферу використання і поясніть зміст параметрів модифікованої експоненти.
5. Динаміку витрат на будівництво автомобільних доріг (млн євро) описує модифікована експонента $Y_t = 4899 - 4571 * 0,97208^t$. Поясніть зміст параметрів.
6. Розкрийте сутність і окресліть сферу використання S -подібних кривих.
7. Доведіть, що модифікована експонента є базовою для логістичної кривої і кривої Гомпертца.
8. Динаміку страхових платежів, що передаються в перестраховування нерезидентам, описує крива Гомпертца $Y_t = 4644,5 * (0,0961435^{0,93176^t})$. Визначте функцію Y_t в точці перегину, поясніть зміст цієї величини.
9. Динаміка витрат на рекламу рухається за логістичною кривою

$$Y_t = 1/(0,00193 + 0,00187 * (0,8734)^t).$$

Визначте функцію Y_t в точці перегину, поясніть зміст цієї величини.

10. Яка статистична характеристика динаміки лежить в основі апріорних тестів вибору функціонального виду кривих зростання?
11. Наведіть приклади оновлюваних сукупностей.
12. Які завдання покладаються на модель оновлення в демографічному прогнозуванні?
13. Доведіть взаємозв'язок коефіцієнтів смертності і дожиття з іншими показниками демографічних таблиць.
14. Поясніть логіку основного рівняння оновлення. Використовуючи його, визначте, скільки об'єктів необхідно ввести в експлуатацію через 3 роки в умовах простого відтворення, якщо термін служби 3 роки, а ймовірність вибуття становить: $P_1 = 0,1$; $P_2 = 0,3$; $P_3 = 0,6$.

15. Динаміка витрат на ведення мисливського господарства характеризується даними:

№ року з/п	1	2	3	4	5	6	7	8	9
Млн грн	8,2	9,4	13,8	21,4	35,2	38,6	39,4	40,5	41,0

Використовуючи правило трьох сум, опишіть динаміку витрат на ведення мисливського господарства модифікованою експонентою, поясніть зміст параметрів.

16. Динаміка чисельності населення мегаполісу характеризується даними:

Рік	1955	1960	1965	1970	1975	1980	1985
Млн осіб	1,22	1,34	1,56	1,83	2,42	3,29	3,94
Рік	1990	1995	2000	2005	2010	2015	2020
Млн осіб	4,56	4,85	5,12	5,27	5,32	5,34	5,35

Використовуючи процедури модуля *Nonlinear Estimation* системи *Statistica*, опишіть динаміку чисельності населення модифікованою експонентою, перевірте адекватність моделі, поясніть зміст параметрів.

17. Потижнева динаміка захворюваності населення в регіоні на початку пандемії характеризувалася даними:

Тиждень	1	2	3	4	5	6
Тис. осіб	8,2	9,4	13,8	21,4	35,2	50,6
Тиждень	7	8	9	10	11	12
Тис. осіб	58,8	62,7	68,8	70,4	75,5	77,9

Використовуючи процедури модуля *Nonlinear Estimation* системи *Statistica*, опишіть динаміку кількості захворювань логістичною кривою, перевірте адекватність моделі, поясніть зміст параметрів.

18. Використовуючи основне рівняння оновлення, визначте, скільки об'єктів необхідно ввести в експлуатацію через 1, 2, 3, 4, 5 років в умовах простого відтворення, якщо термін їхньої служби 4 роки, а ймовірність вибуття становить: $P_1 = 0,1$; $P_2 = 0,2$; $P_3 = 0,3$; $P_4 = 0,4$.

Рекомендована література

1. Єріна А.М., Єрін Д.Л. Статистичне моделювання та прогнозування : підручник. Київ, КНЕУ. 2014. С. 129—148.
2. Лук'яненко І., Краснікова Л. Економетрика: підручник. Київ. 1998. С.138—158.
3. Методи і моделі економічного прогнозування : навч. посіб. / Кічор В.П. та ін.; Львів, 2019. С. 106—130.
4. Сардак С.Е. Життєвий цикл соціально-економічних систем / *Маркетинг і менеджмент інновацій*. 2016. Вип.1. С. 157—168.
5. Стояновський А. Р., Сидоров Ю. І. Визначення оптимального терміну реалізації інноваційного проекту в межах інноваційної структури / *Науковий вісник НЛТУ України*. 2011. Вип. 21.10. С. 322—330

- 5.1. Амплітуда сезонних коливань
- 5.2. Сезонна декомпозиція часових рядів
- 5.3. Моделювання сезонної хвилі

5.1. Амплітуда сезонних коливань

Сезонні коливання відображають внутрішньорічну динаміку і більш-менш регулярно повторюються з року в рік. Наприклад, виробництво й переробка сільськогосподарської продукції, споживання енергоносіїв, завантаження транспорту, діяльність рекреаційних закладів. Вони спричиняють нерівномірне використання протягом року виробничих потужностей і робочої сили, коливання попиту і пропозиції на ринку споживчих товарів тощо. Такі процеси потребують регулювання, а отже, і статистичного вивчення.

Можна окреслити щонайменше три напрями статистичного аналізу сезонних коливань:

- визначити внутрішньорічну амплітуду сезонних коливань, максимальний і мінімальний рівні ряду, аби прийняти певні рішення, скажімо, про те, скільки готелів закрити після завершення сезону;
- скоригувати прогностні значення показників на сезонність;
- виключити сезонну компоненту із часового ряду заради вивчення інших компонент, передусім, тренду.

Сезонні коливання виявляються й аналізуються на основі рядів помісячних або поквартальних даних. Кожний рівень ряду y_t належить до певного сезонного циклу s , довжина якого становить 12 місяців або 4 квартали. Для вимірювання конфігурації сезонного ритму застосовують *індекси сезонності*.

У часових рядах, які не виявляють чіткої тенденції розвитку, індекси сезонності I_t визначають відношенням місячних (квартальних) рівнів ряду y_t до середнього рівня за сезонний цикл \bar{y} і виражаються процентами:

$$I_t = 100 \frac{y_t}{\bar{y}}$$

Іноді замість середньої арифметичної використовують медіану, яка не залежить від крайніх значень індексів.

За наявності тенденції розвитку попередньо проводять згладжування чи вирівнювання часового ряду, а в індексі сезонності замість середньої \bar{y} використовують теоретичні рівні Y_t , визначені для кожного місяця (кварталу) року за функцією тренду $Y_t = f(t)$ або методом ковзних середніх:

$$I_t = 100 \frac{y_t}{Y_t}$$

Сукупність індексів в межах сезонного циклу утворює «сезонну хвилю». Якщо сезонні коливання з року в рік змінюються, виявити сталу сезонну хвилю можна усередненням помісячних чи поквартальних індексів сезонності за кілька років (не менше трьох):

$$\bar{I}_c = \frac{1}{n} \sum_1^n I_t,$$

де n — кількість років.

Порядок обчислення сезонної хвилі розглянемо на прикладі помісячної динаміки яйценосності курей на одній з птахофабрик агрохолдингу (табл. 5.1). Середня несучість курей за рік становить 234 штуки, середньомісячна — $234 : 12 = 19,5$, діапазон варіації від 12,1 у січні до 27,0 у червні. Індеси сезонності обчислені відношенням середньомісячної несучості курей до середньомісячної річної, наприклад, для січня $100 (12,1 : 19,5) = 62,2\%$. Сума індексів сезонного циклу становить 1200, середній індекс — 100 %.

Таблиця 5.1

ПОМІСЯЧНІ ІНДЕКСИ СЕЗОННОСТІ НЕСУЧОСТІ КУРЕЙ

Місяць	Середньомісячна несучість	Індекс сезонності	Місяць	Середньомісячна несучість	Індекс сезонності
Січень	12,1	62,2	Липень	26,7	136,9
Лютий	16,7	85,6	Серпень	24,7	126,7
Березень	18,3	93,8	Вересень	20,3	104,1
Квітень	19,3	99,1	Жовтень	15,4	78,8
Травень	25,8	132,6	Листопад	15,0	76,9
Червень	27,0	138,5	Грудень	12,7	65,1

Характер сезонної хвилі схематично ілюструє рис. 5.1: пік несутості курей припадає на червень (138,5 %), мінімум — на січень (62,2 %).

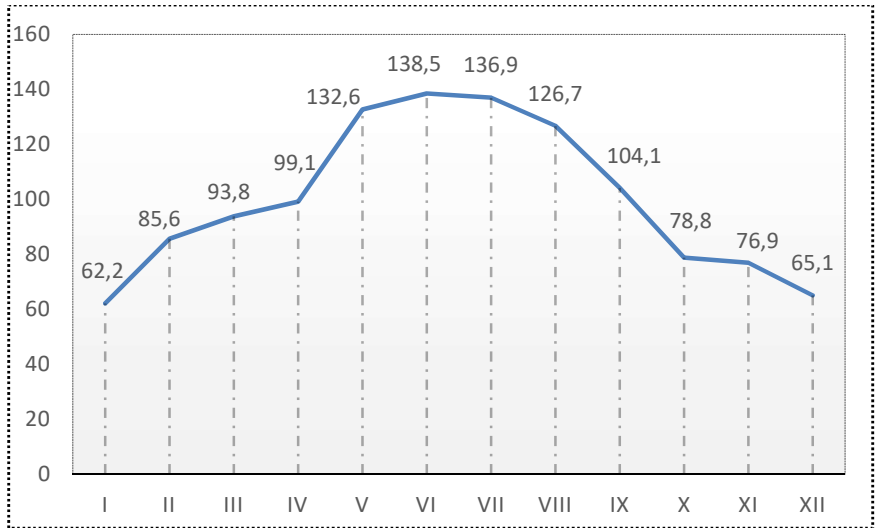


Рис. 5.1. Сезонна хвиля яйценосності курей

Для вимірювання сили сезонних коливань використовують:

— амплітуду коливань за сезонний цикл:

$$R_t = (I_{\max} - I_{\min}) = \frac{y_{\max} - y_{\min}}{\bar{y}} 100;$$

— середнє лінійне \bar{l}_t або середнє квадратичне σ_t відхилення індексів сезонності:

$$\bar{l}_t = \frac{1}{12} \sum_1^{12} |I_t - 100| \quad \sigma_t = \sqrt{\frac{1}{12} \sum_1^{12} (I_t - 100)^2}$$

У досліджуваному часовому ряду (див. табл. 5.1) амплітуда сезонних коливань становить

$$R_t = 100 (27,0 - 12,1) : 19,5 = (138,5 - 62,2) = 76,3 \text{ п.п.}$$

— середнє квадратичне відхилення:

$$\sigma_t = \sqrt{\frac{8512,83}{12}} = 26,6 \text{ п.п.}$$

Розрахунок індексів сезонності за наявності тенденції подано у табл. 5.2 на прикладі поквартальної динаміки продажу безалкогольних напоїв за три роки (млн дкл.). Тенденцію ряду описує рівняння $Y_t = 38,8 + 1,445 t$, де $t = 1, 2, 3, \dots, 12$.

Таблиця 5.2

**ДО РОЗРАХУНКУ СЕЗОННИХ КОЛИВАНЬ
З УРАХУВАННЯМ ТЕНДЕНЦІЇ РОЗВИТКУ**

Рік, квартал	Млн дкл., y_t	Теоретичні рівні $Y_t = 38,8 + 1,445 t$	Індекс сезонності $I_t = 100 \frac{y_t}{Y_t}$	Відхилення $(I_t - 100)$	$(I_t - 100)^2$
1 рік — Q1	24,4	40,3	60,6	-39,4	1552,36
Q2	52,6	41,7	126,2	26,2	686,44
Q3	60,4	43,1	140,1	40,1	1608,01
Q4	34,0	44,6	76,3	-23,7	561,69
II рік — Q1	32,7	46,0	71,1	-28,3	800,89
Q2	56,2	47,5	118,4	18,4	338,56
Q3	67,3	48,9	137,7	37,7	1421,29
Q4	36,2	50,4	71,9	-28,1	789,61
III рік — Q1	37,8	51,8	73,0	-27,0	729,0
Q2	65,3	53,3	122,5	22,5	506,25
Q3	73,1	54,7	133,7	33,7	1135,69
Q4	38,4	56,1	68,4	-31,6	998,56
Разом	578,4	578,4	1200	X	11128,35

На основі річних поквартальних індексів визначають середньорічні індекси сезонності для кожного кварталу \hat{I}_t :

для першого кварталу $(60,6 + 71,1 + 73,0) : 3 = 68,2 \%$;

для другого — 122,4; для третього — 137,2; для четвертого — 72,2 %.

Максимальний обсяг продажу безалкогольних напоїв припадає на третій квартал, мінімум — на перший квартал. Амплітуда коливань становить $(137,2 - 68,2) = 69,0$ п. п., середнє квадратичне відхилення $\sigma_t = \sqrt{\frac{11128,35}{12}} = 30,35$ п. п.

Прогнозування сезонних процесів ґрунтується на декомпозиції часового ряду. Припускають, що у майбутньому збережеться тенденція і такий самий характер сезонних коливань, як у попередні роки. За таких умов прогноз на будь-який місяць (квартал), визначений методом екстраполяції тренду, коригується індексом сезонності:

$$\hat{Y}_{t+v} = \hat{I}_t \cdot Y_{t+v},$$

де v — період попередження (місяців, кварталів).

У нашому прикладі очікуваний обсяг продажу безалкогольних напоїв у першому кварталі наступного року становить $Y_{t+1} = 38,8 + 1,445 \cdot 13 = 57,6$ млн. дкл., а скоригований на сезонність $\hat{Y}_{t+1} = 57,6 \cdot 0,682 = 39,3$ млн. дкл. Аналогічно визначені і скориговані на сезонність прогнози на 2—4 квартали наведено у табл. 5.3.

Таблиця 5.3

ДО ВИЗНАЧЕННЯ СКОРИГОВАНИХ НА СЕЗОННІСТЬ ПРОГНОЗІВ

Квартал наступного року	Q1	Q2	Q3	Q4
Період попередження, v	1	2	3	4
$(t + v)$	13	14	15	16
Екстраполяція тренду $Y_{t+v} = 38,8 + 1,445(t+v)$	57,6	59,0	60,5	61,9
Середній індекс (коефіцієнт) сезонності для кварталу, \hat{I}_t	0,682	1,224	1,372	0,722
Прогноз, скоригований на сезонність* $\hat{Y}_{t+v} = Y_{t+v} \cdot \hat{I}_t$	39,3	72,25	83,0	44,7

Випадкова компонента динамічного ряду виявляється відхиленнями фактичних рівнів від теоретичних. Абсолютною мірою випадкових коливань є стандартне відхилення s_e , яке обчислюють на основі залишкової дисперсії $s_e = \sqrt{s_e^2}$ (табл. 5.4).

Таблиця 5.4

ДО РОЗРАХУНКУ ЗАЛИШКОВОЇ ДИСПЕРСІЇ

Рік, квартал	Млн. дкл., y_t	Теоретичні рівні, скориговані на сезонність Y_t	$y_t - \widehat{Y}_t$	$(y_t - \widehat{Y}_t)^2$
I рік — Q1	24,4	27,5	-3,1	9,5
Q2	52,6	51,0	1,6	2,4
...
III рік — Q4	38,4	38,3	0,1	0,01
Разом	578,4	578,4	0	32,3

За даними таблиці залишкова дисперсія становить

$$s_e^2 = \frac{1}{n-m} \sum_1^n (y_t - Y_t)^2 = \frac{32,3}{12-2} = 3,23,$$

звідки стандартне відхилення $s_e = \sqrt{3,23} = 1,79$ млн. дкл.

Разом із абсолютною мірою випадкових коливань використовують коефіцієнт коливань (аналог коефіцієнта варіації)

$$V_e = 100 \frac{s_e}{\bar{y}},$$

де \bar{y} — середній рівень часового ряду.

Різницю $100 - V_e$ розглядають як оцінку сталості динаміки. У нашому прикладі $V_e = 100 \frac{1,79}{48,2} = 3,72\%$.

За наявності в часовому ряду сезонних коливань і тенденції закономірність динаміки можна описати у вигляді адитивної моделі, яка включає фактор часу t і змінні бінарного типу u_k , які набувають значень нуль або одиниця. Змінну u_k називають *dummy*-змінною і розглядають її як умовний код, що вказує на належність чи неналежність t -го рівня ряду до відповідного k -го сезону року (кварталу чи місяцю). Кількість *dummy*-змінних в рівнянні має бути на одиницю менше кількості періодів, тобто для кварталних даних — три змінні, для місячних — 11. Загальний вигляд адитивної моделі сезонності за наявності тренду:

$$Y_t = a_0 + \sum_1^{m-1} b_k u_k + ct,$$

де Y_t — теоретичний рівень t -го рівня часового ряду з урахуванням сезонних коливань;

b_k — ефект k -го сезону порівняно з базовим;

c — ефект неідентифікованих факторів, які формують тенденцію ряду.

Так, в разі поквартальної сезонності, один з кварталів визнається базовим, тож до моделі включаються три *dummy*-змінні. Припустимо, що базовим є I квартал, тоді означення для 2-4 кварталів будуть такі:

Квартал	u_2	u_3	u_4
I	0	0	0
II	1	0	0
III	0	1	0
IV	0	0	1

Оскільки *dummy*-змінні набувають значень 1 і 0, то практично визначається відхилення рівня ряду відповідного кварталу під впливом сезонності порівняно з I кварталом.

Як приклад, наведемо адитивну модель сезонності кількості безробітних, квартальні дані за три роки:

$$Y_t = 12,417 - 3,688 u_2 - 5,010 u_3 - 2,031 u_4 - 0,344 t.$$

Коефіцієнт детермінації становить 0,984, усі параметри моделі істотні. За змістом параметри характеризують:

$a_0 = 12,417$ — це рівень I кварталу з нульовими значеннями *dummy*-змінних;

$c = -0,344$ вказує на тенденцію зниження рівнів ряду при елімінуванні сезонності, тобто щоквартально чисельність безробітних знижувалася в середньому на 344 людини, незалежно від впливу сезонності;

параметри при змінних u_2 , u_3 і u_4 фіксують не рівні сезонності відповідного кварталу, а відхилення сезонності в II, III і IV кварталах порівняно з I кварталом, незалежно від наявної тенденції.

В дослідженнях сезонності за місячними даними в адитивну модель треба включати 11 *dummy*-змінних. В такому разі, аби забезпечити надійність оцінок параметрів, довжина первинного часового ряду помісячної інформації має бути щонайменше 7—8 років.

5.2. Сезонна декомпозиція часових рядів

Суть сезонної декомпозиції полягає в тому, щоб виявити та відокремити вплив сезонних факторів від інших, розкласти часовий ряд на сезонну, тренд-циклічну і нерегулярну компоненти. Загальний алгоритм сезонної декомпозиції складається з п'яти послідовних етапів:

- 1) визначається тренд у формі функції часу або ковзної середньої;
- 2) оцінюється сезонний фактор у межах сезонного циклу, довжина якого становить 12 місяців або 4 квартали;
- 3) здійснюється фільтрація сезонної компоненти;
- 4) визначається тренд-циклічна компонента;
- 5) визначається нерегулярна, випадкова компонента.

Обчислювальні процедури на кожному етапі залежать від типу моделі: адитивна, мультиплікативна чи змішана.

Зміст етапів сезонної декомпозиції в змішаній (адитивно-мультиплікативній) моделі унаочнює рис. 5.2.



Рис. 5.2. Етапи сезонної декомпозиції часового ряду (адитивно-мультиплікативна модель)

Розглянемо методику сезонної декомпозиції часового ряду за даними виїзного потоку туристів, виконану в модулі *Time Series / Forecasting* системи Statistica. Фрагмент таблиці результатів сезонної декомпозиції наведено в табл. 5.5.

Таблиця 5.5

СЕЗОННА ДЕКОМПОЗИЦІЯ ПОТОКУ ВИЇЗНИХ ТУРИСТІВ

Case	Seasonal Decomposition: Multipl. Season (12) TURIST3 : Виїзний потік туристів						
	TURIST3	Moving Averages	Ratios	Seasonal Factors	Adjusted Series	Smoothed Trend-c.	Irreg. Compon.
1	5567			83,46	6670,3	6454,7	1,033
2	5219			78,80	6622,9	6477,7	1,022
3	5886			95,87	6139,9	6523,7	0,941
4	6632			98,36	6742,8	6585,8	1,024
5	7092			102,90	6891,8	6636,1	1,038
6	7502			118,00	6357,8	6617,7	0,961
7	7616	6412,8	118,78	113,53	67,08,2	6515,3	1,029
8	7594	6403,7	118,59	116,06	6543,4	6385,4	1,025
9	7110	6382,3	111,40	122,18	5819,2	6271,2	0,928
10	6264	6363,2	98,44	97,48	6425,7	6212,1	1,034
11	5183	6358,5	81,51	81,79	6337,2	6196,9	1,023
12	5289	6337,1	83,46	91,57	5775,8	6200,3	0,932
13	5457	6394,3	85,34	83,46	6538,5	6222,3	1,051
14	4963	6406,1	77,47	78,80	6298,0	6248,1	1,008
15	5656	6416,4	88,15	95,87	5899,9	6316,5	0,934
16	6576	65,13,8	100,96	98,36	6685,9	6486,5	1,031

Сезонний цикл — 12 місяців, тип моделі — мультиплікативна:

- первинний ряд вирівнюється ковзною середньою Y_t ; оскільки вікно згладжування $m = 12$, відсутні значення середніх за січень-червень 2015 р. і за липень-грудень 2020 р.;

- розраховуються індекси сезонності співвідношенням первинного ряду і ковзних середніх

$$I_t = 100 \frac{y_t}{Y_t};$$

- сезонний фактор S_t визначають усередненням індексів сезонності відповідного місяця діленням на 5 років:

$$S_t = (88,15 + \dots) / 5 = 83,46$$

- фільтрацію сезонного фактора здійснюють коригуванням первинного ряду на сезонний фактор

$$Y_t^K = \frac{y_t}{S_t};$$

- тренд-циклічну компоненту визначають шляхом експоненційного згладжування скоригованого на сезонність ряду з параметром адаптації $a = 0,3$;

- випадкова складова. У мультиплікативній моделі випадкову складову визначають відношенням скоригованого на сезонність ряду до тренд-циклічної компоненти, наприклад, для 1-го рівня ряду:

$$6670,3 : 6454,7 = 1,0334.$$

До часового ряду, який не виявляє чітко вираженої тенденції до розвитку, більш еластичною виявляється ковзна середня, інтервал згладжування якої дорівнює сезонному циклу (4 або 12). Коригування ковзної середньої на сезонність здійснюється так само, як коригування лінійного тренду.

За результатами розрахунку можна перекоонатися, що в мультиплікативній моделі первинний ряд — це добуток тренд-циклічної, сезонної й нерегулярної компонент. Так, для 7-го рівня ряду:

$$6515,3 \cdot 1,1353 \cdot 1,0296 = 7616,0$$

Тренд-циклічну компоненту і ряд із сезонною поправкою унаочнює рис. 5.3. Видно, що скоригований ряд не містить сезонних коливань, проте зазнає незначних коливань навколо тренд-циклічної компоненти. Тобто модель адекватно описує тренд і сезонність часового ряду.

Припускаючи, що в майбутньому збережеться тенденція і такий самий характер коливань, прогноз на будь-який місяць (квартал), визначений методом екстраполяції тренд-циклічної компоненти, коригується індексом сезонності:

$$Y^*_{t+v} = I_t \cdot Y_{t+v},$$

де v — період попередження.

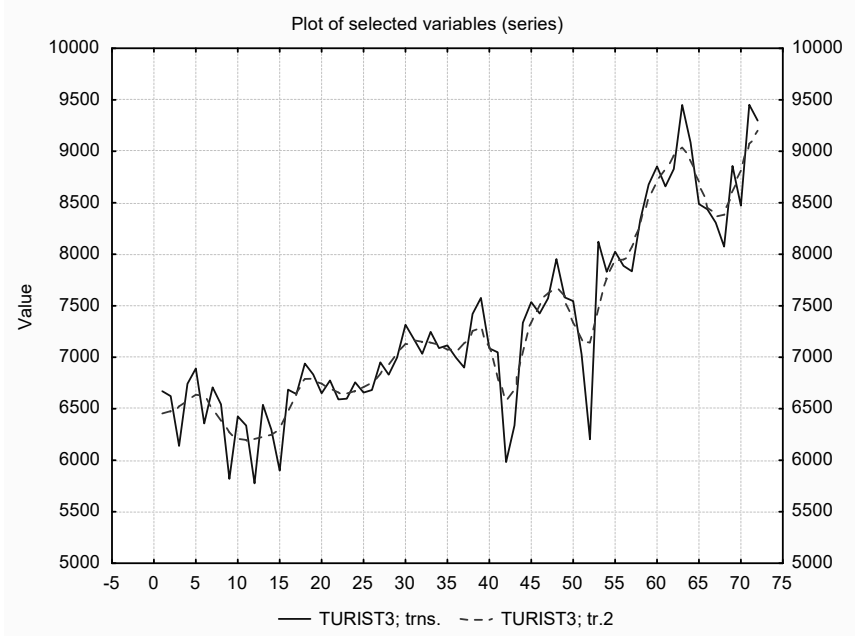


Рис. 5.3. Тренд-циклічна компонента і ряд, скоригований на сезонність

Композиція цих прогнозів залежить від типу моделі:

- в адитивній моделі — це підсумок прогнозів тренд-циклічної компоненти і сезонного фактора :

$$Y^*_{t+v} = TCt + St + et;$$

- у мультиплікативній моделі — добуток прогнозу тренд-циклічної компоненти і сезонного фактора:

$$Y^*_{t+v} = TC_t \cdot S_t \cdot e_t.$$

- у адитивно-мультиплікативній моделі:

$$Y^*_{t+v} = TC_t \cdot S_t + e_t.$$

За результатами сезонної декомпозиції часового ряду виїзного потоку туристів (мультиплікативна модель) ретроспективно за процедурою *Census I* визначимо прогноз на квітень 2016 р. (останній рядок табл. 5.5), виокремлено три компоненти часового ряду: тренд-циклічну, сезонну й випадкову. Значення тренд-циклічної компоненти на цей місяць становило 6576,0 тис. осіб, усереднений індекс сезонності (сезонний фактор) — 1,0096. Звідси прогноз виїзного потоку туристів

$$Y^*_{t+v} = 6486,5 \cdot 1,0096 = 6548,8.$$

Фактичне значення 6576,0 тис. осіб, тож похибка прогнозу становить $(6576,0 - 6548,8) = -27,2$ тис. осіб або $-0,4\%$. Це свідчить про високі прогностичні властивості моделей, які ґрунтуються на сезонній декомпозиції часових рядів.

5.3. Моделювання сезонної хвилі

Періодичним називають часовий ряд, коливання рівнів якого навколо середнього значення повторюються з певною періодичністю, тобто $Y_t = Y_{t+p}$, де p — період циклу, тобто інтервал часу, необхідний для того, щоб цикл коливань (флуктуації ряду) почав повторюватися. Період циклу вимірюється тривалістю часу за цикл, а кількість циклів за одиницю часу — частотою $f = 1/p$. Наприклад, періодична хвиля сонячної активності за річними даними має частоту 0,0909, тож на повний цикл хвилі необхідно 11 років ($1 : 0,0909$). Тобто частота і період (тривалість циклу) — взаємно обернені величини.

Окрім періоду й частоти коливань, характеристиками періодичного часового ряду є фаза та амплітуда коливань. Фаза φ — це відстань між початком відліку часу з точкою $t = 0$ і найближчим піком піднесення, амплітуда A — відстань від середнього значення ряду до піку піднесення чи до дна спаду. Схематично періодичний часовий ряд подано на рис. 5.4. Якщо припустити, що рівні ряду — помісячні дані, то період циклу дорівнює 10 місяців.

Періодичні часові ряди можуть коливатися навколо певного зростаючого чи спадного середнього рівня (ряд з трендом середнього). Так само амплітуда із часом може зростати чи зменшуватися (ряд з трендом дисперсії). Ряди з трендом середньої чи дисперсії необхідно звести до стаціонарного виду. Стаціонарний періодичний ряд можна записати у вигляді гармонійної моделі: Y

$$Y_t = \mu + A \cos 2\pi f(t - \varphi).$$

Скажімо, на рис. 5.4 при $t = \varphi$ маємо $Y_t = \mu + A$, оскільки $\cos 0 = 1$.

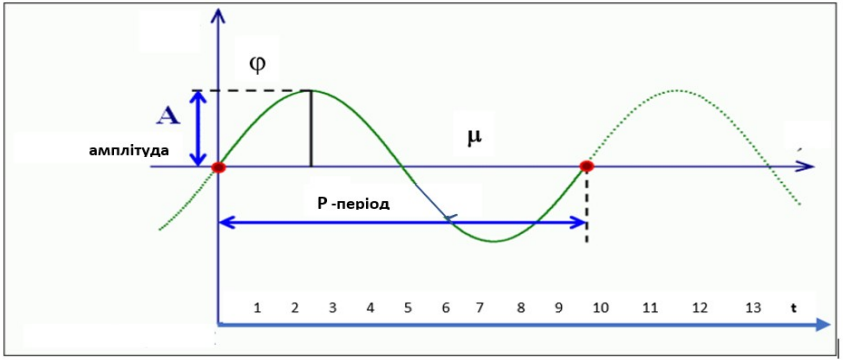


Рис. 5.4. Фрагмент періодичного часового ряду

Гармонійну модель частіше записують з використанням тригонометричних функцій — синусів і косинусів — без явного згадування про фазу:

$$Y = \mu + \sum_{k=1}^m (b_k \cos \omega_k t + d_k \sin \omega_k t),$$

де ω — кутова частота, яка вимірюється в радіанах за одиницю часу і дорівнює $\omega = 2\pi f$;

b_k, d_k — коефіцієнти k -ї гармоніки;

k — визначає гармоніку ряду (найчастіше $1 \leq k \leq 4$);

t — часова ознака.

Послідовні значення змінної часу t виражаються радіанною мірою і визначаються в інтервалі від 0 до n з приростом $2\pi/n$, де n — кількість членів часового ряду. Гармонійна функція розкладає часовий ряд на правильні періодичні хвилі, кожна з яких зображується синусоїдою чи косинусоїдою. Пошуки такої строгої періодичності в економічних рядах виявилися не досить успішними, оскільки будь-які зовнішні впливи влітаються в подальший рух економічного процесу і стають його частиною, порушуючи правильність чергування максимумів і мінімумів. Водночас не можна заперечити можливість виокремлення гармонійних компонент у рядах, яким властивий сезонний цикл тривалістю 12 місяців. Послідовність значень часової ознаки t в сезонних рядах записують так:

ωt , радіани	0	$\pi/6$	$\pi/3$	$\pi/2$	$2\pi/3$	$5\pi/6$	π	$7\pi/6$	$4\pi/3$	$3\pi/2$	$5\pi/3$	$11\pi/6$
Y_t	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Y_7	Y_8	Y_9	Y_{10}	Y_{11}	Y_{12}

У модель гармонійного аналізу можна включити кілька гармонік з різною довжиною хвиль, скажімо: 12; 6; 4; 3; 2,4 і 2 місяці. Апроксимація динаміки рядом Фур'є полягає у виборі таких гармонік, які б сумарно відображали періодичні коливання фактичних рівнів часового ряду. Адекватність гармонійної функції реальному процесу залежить від того, наскільки сталими є частота й амплітуда коливань.

Параметри моделі визначаються методом найменших квадратів. Завдяки властивостям ортогональності функцій синуса і косинуса система нормальних рівнянь зводиться до тотожностей:

$$\begin{aligned}\Sigma y_t &= \mu n \\ 2 \Sigma y_t \cos \omega t &= n b_k ; \\ 2 \Sigma y_t \sin \omega t &= n d_k,\end{aligned}$$

звідси для $n = 12$:

$$a = \Sigma y_t : 12, b_k = (\Sigma y_t \cos \omega t) : 6; d = (\Sigma y_t \sin \omega t) : 6.$$

Отже, μ — це не що інше, як середньомісячний рівень ряду; b_k , d_k — коефіцієнти k -ї гармоніки, які визначають амплітуду коливань

$$A_k = \sqrt{b_k^2 + d_k^2}.$$

Очевидно, чим більша амплітуда коливань, тим вагоміший внесок k -ї гармоніки в загальну дисперсію процесу. Оцінкою такого внеску слугує дисперсійне відношення

$$R_k^2 = \frac{\delta_k^2}{\sigma_y^2},$$

де $\delta_k^2 = 0,5A^2 = 0,5(b_k^2 + d_k^2)$ — дисперсія k -ї гармоніки (крім останньої),

$$\sigma_y^2 = \frac{\Sigma_1^n (y_t - \bar{y})^2}{n} — \text{загальна дисперсія процесу.}$$

Розглянемо методику побудови гармонійної моделі за даними помісячної динаміки виробництва молока в сільськогосподарському підприємстві (табл. 5.6). Для даного часового ряду логічним є період коливань 12 місяців, тож визначимо параметри для першої гармоніки ($k = 1$):

$$Y_1 = a + b_1 \cos \omega t + d_1 \sin \omega t.$$

Для розрахунку коефіцієнтів гармонійної моделі використаємо сумарні величини табл. 5.6. Значення коефіцієнтів моделі становили

$$\begin{aligned} \Sigma y_t &= 122,4; & \mu &= 122,4 : 12 = 10,233333; \\ \Sigma y_t \cos \omega t &= -28,4318; & b &= -28,4318 : 6 = -4,73863; \\ \Sigma y_t \sin \omega t &= 10,3522; & d &= 10,3522 : 6 = 1,725367 \\ \text{Дисперсія ряду} & & \sigma_y^2 &= 13,55 \end{aligned}$$

Таблиця 5.6

ДО РОЗРАХУНКУ СЕЗОННОЇ ХВИЛІ ВИРОБНИЦТВА МОЛОКА

Місяць	t	$\cos \omega t$	$\sin \omega t$	Виробництво, тон, y_t	$y \cos \omega t$	$y \sin \omega t$	Y_t
Січень	0	1	0	6,8	6,8	0	5,4
Лютий	$\pi/6$	0,866	0,5	7,0	6,062	3,5	7,0
Березень	$\pi/3$	0,5	0,866	9,4	4,7	8,140	9,3
Квітень	$\pi/2$	0	1	10,5	0	10,5	11,9
Травень	$2\pi/3$	-0,5	0,866	13,1	-6,55	11,345	14,1
Червень	$5\pi/6$	-0,866	0,5	16,2	-14,029	8,1	15,2
Липень	π	-1	0	16,1	-16,1	0	15,0
Серпень	$7\pi/6$	-0,866	-0,5	14,0	-12,124	-7,0	13,4
Вересень	$4\pi/3$	-0,5	-0,866	10,2	-5,1	-8,833	11,1
Жовтень	$3\pi/2$	0	-1	7,6	0	-7,6	8,5
Листопад	$5\pi/3$	0,5	-0,866	5,6	2,8	-4,850	6,3
Грудень	$11\pi/6$	0,866	-0,5	5,9	5,109	-2,95	5,2
Разом	\times	\times	\times	122,4	-28,432	10,352	122,4

Порівняння розрахованих за моделлю теоретичних Y_t і фактичних y_t рівнів ряду свідчить про високі апроксимуючі властивості гармонійної моделі першого порядку. Амплітуда і дисперсія першої гармоніки становлять:

$$A_k = \sqrt{(-4,739)^2 + 1,725^2} = \sqrt{25,4336} = 5,04,$$

$$\delta_k^2 = 0,5A^2 = 0,5(b_k^2 + d_k^2) = 0,5 \cdot 25,4336 = 12,72.$$

Звідси дисперсійне відношення дорівнює:

$$R_k^2 = \frac{\delta_k^2}{\sigma_y^2} = \frac{12,72}{13,55} = 0,938,$$

тобто перша гармоніка пояснює 93,8 % сезонних коливань часового ряду.

Аналогічно визначаються гармоніки вищих порядків; значення їх послідовно приєднуються до значень першої гармоніки (доданки розміщуються за зменшеними періодами). Так, ряд Фур'є з двома гармоніками ($k = 2$) має вигляд

$$Y_t = a + b_1 \cos \omega t + d_1 \sin \omega t + b_2 \cos \omega_2 t + d_2 \sin \omega_2 t.$$

Розрахунок гармонійної функції можна здійснити за алгоритмом *Spectral (Fourier) analysis* модуля *Time Series / Forecasting*. Програма видає підсумкову таблицю з частотами, періодами, коефіцієнтами при косинусах і синусах, значеннями періодограм, оцінками спектральної площини і вагами, які використовуються при розрахунку оцінок спектральної площини. В табл. 5.7 наведено основні параметри гармонійної моделі виробництва молока. Найбільші значення коефіцієнтів (косинус-коефіцієнт $b = -4,856$; синус-коефіцієнт $d = 1,282$) розміщені проти частоти 0,083, якій відповідає період 12 місяців. Істотним виявився також косинус-коефіцієнт для частоти 0,167 (тривалість періоду — 6 місяців).

Розглянутий приклад засвідчує, що моделювання сезонних коливань на основі гармонік ряду Фур'є можна розглядати як один з інструментів обґрунтування управлінських заходів з регулювання нерівномірності соціально-економічних процесів.

Таблиця 5.7

РЕЗУЛЬТАТИ ГАРМОНІЙНОГО АНАЛІЗУ ВИРОБНИЦТВА МОЛОКА

Spectral analysis: Production (Spreadsheet1)							
No. of cases: 12							
	Frequency	Period	Cosine Coeffs	Sine Coeffs	Periodogram	Density	Hamming Weights
0	0,000		-0,000	-0,000	0,000	73,537	0,036
1	0,083	12,00	-4,858	1,282	151,434	74,801	0,241
2	0,167	6,00	1,081	-0,235	7,344	39,935	0,446
3	0,250	4,00	-0,286	-0,052	0,506	7,612	0,241
4	0,333	3,00	-0,069	-0,357	0,794	0,853	0,036
5	0,417	2,40	0,136	-0,241	0,459	0,451	
6	0,500	2,00	-0,119	-0,000	0,085	0,316	

Питання і задачі для самопідготовки

1. Сформулюйте цілі аналізу сезонної компоненти часового ряду.
2. Як визначаються індекси сезонності в рядах за наявності і відсутності тенденції?
3. Поясніть сутність сезонної декомпозиції часового ряду.
4. Назвіть етапи сезонної декомпозиції часового ряду.
5. Поясніть відмінності між адитивним і мультиплікативним типами сезонної компоненти.
6. Розкрийте особливість сезонної декомпозиції в змішаній (адитивно-мультиплікативній) моделі.
7. У чому сутність моделі гармонійного аналізу?
8. Розкрийте поняття і методику розрахунку амплітуди гармонічних коливань?
9. Природно-кліматичні умови, які склалися в регіоні у минулому році, характеризуються даними:

Показник	Температура за Фаренгейтом (°F)	Опади, мм
Середньомісячний рівень	53,5	3,5
Максимальний	75	4,3
Мінімальний	33	3,0

Порівняйте середньомісячні коливання температури і опадів. Зробіть висновки.

10. Динаміка екотуризму в регіоні характеризується даними, тис. туристів:

Квартал року	1-й рік	2-й рік	3-й рік
I	34	33	35
II	50	57	45
III	102	105	100
IV	36	43	38

Зобразіть сезонну хвилю екотуризму графічно, обчисліть індекси сезонності та амплітуду сезонних коливань екотуризму, зробіть висновки.

11. Тенденція поквартального експорту лохини (тис.т) за останні два роки описується лінійним трендом:

$$Y = 86,0 + 0,13t, \text{ де } t = 1, 2, \dots, 8.$$

Визначте прогнозні обсяги експорту лохини на кожний квартал наступного року, скоригувавши їх на сезонність: для 1-го кварталу — 70 %; для 2-го — 68 %; для 3-го — 135 %; для 4-го — 127 %.

12. Продаж плодів, ягід, горіхів через мережу торговельних підприємств характеризується даними, тис. т.

Рік	Квартал			
	1	2	3	4
1	8,3	6,6	9,8	7,8
2	9,2	7,5	10,4	8,4
3	9,7	8,6	10,8	9,0
4	10,0	9,1	11,3	10,2
5	10,8	10,2	11,7	10,6

За кварталними даними обчисліть:

- індекси сезонності, поквартальний сезонний фактор;
- прогнозні рівні продажу плодів, ягід, горіхів на кожний квартал наступного року;
- скоригуйте їх на сезонний фактор.

13. Динаміка реалізації населенню скрапленого газу характеризується даними, т

Квартал року	1-й рік	2-й рік	3-й рік
I	24	28	35
II	22	36	28
III	16	20	24
IV	30	31	38

За кварталними даними :

- опишіть тенденцію реалізації скрапленого газу лінійним трендом і на його основі очікувані поквартальні обсяги реалізації у наступному році;
- обчисліть індекси сезонності для кожного кварталу і на їх основі амплітуду сезонних коливань;
- припускаючи, що виявлена тенденція і характер сезонних коливань у найближчій перспективі не зміняться, скоригуйте прогнозні обсяги реалізації на сезонний фактор.

14. Динаміка доходів турагенства характеризується даними, млн. дол.:

Місяць	1-й рік	2-й рік	3-й рік
Січень	68	68	70
Лютий	67	70	74
Березень	91	94	93
Квітень	102	107	102
Травень	120	131	120
Червень	152	162	148
Липень	155	161	149
Серпень	131	140	126
Вересень	101	104	93
Жовтень	76	76	79
Листопад	53	57	57
Грудень	54	59	61

Здійсніть сезонну декомпозицію часового ряду за алгоритмом *Seasonal decomposition (Census I)* модуля *Time Series / Forecasting* — системи Statistica:

1) опишіть тенденцію часового ряду, вибір функціонального виду трендової моделі обґрунтуйте;

2) оцініть сезонний фактор у межах сезонного циклу;

3) здійсніть фільтрацію сезонної компоненти ряду;

4) визначте тренд-циклічну компоненту ряду;

5) проаналізуйте випадкову компоненту.

За результатами сезонної декомпозиції часового ряду визначте прогнозні рівні доходів турагенства на кожний місяць наступного року.

15. Сезонність реалізації антигрипозної вакцини описується моделлю гармонійного аналізу (перша гармоніка)

$$Y_t = 166,3 + 3,1 \cos t + 0,60 \sin t.$$

Визначте амплітуду коливань та оцініть адекватність моделі реальному процесу, якщо загальна варіація помісячних даних характеризується дисперсією $\sigma^2 = 5,86$.

16. Параметри моделей гармонійного аналізу дорожньо-транспортних пригод становили:

Рік	a	b	d
2018	21	8,7	1,15
2021	26	8,1	2,13

Для кожного року:

— визначте амплітуду сезонних коливань, зробіть висновки;
 — оцініть адекватність моделі реальному процесу, якщо загальна варіація помісячних даних відповідно становила: $\sigma_1^2 = 86$, $\sigma_2^2 = 69$

17. Обсяги відпущеної теплової енергії в регіоні характеризуються даними:

Місяць	Тис. Гкал	Місяць	Тис. Гкал	Місяць	Тис. Гкал
1	69	5	14	9	11
2	62	6	8	10	29
3	60	7	8	11	58
4	31	8	9	12	72

Опишіть динаміку обсягів відпущеної теплової енергії за алгоритмом *Spectral (Fourier) analysis* модуля *Time Series / Forecasting* пакету Statistica. Оцініть адекватність моделі.

18. Концентрація шкідливих речовин в атмосферному повітрі міст України, (мг/куб.м) у 2018—2020 рр характеризується даними:

Рік	Квартал			
	I	II	III	IV
2018	0,282	0,365	0,267	0,241
2019	0,284	0,381	0,275	0,247
2020	0,292	0,388	0,269	0,245

Здійсніть графічний аналіз часового ряду, зробіть висновки щодо наявності тренду і сезонної компоненти. Використовуючи модуль *Multiple Regression* системи Statistica опишіть тенденцію забруднення повітря, в матрицю ознакової множини введіть *dummy*-змінні, поясніть зміст параметрів моделі, оцініть адекватність моделі; зробіть висновки.

Рекомендована література

1. Галушак М. П., Галушак О. Я., Кужда Т. І. Прогнозування соціально-економічних процесів: навч. посіб. Тернопіль. 2021. С. 132—144.
2. Єріна А.М., Єрін Д.Л. Статистичне моделювання та прогнозування : підручник. Київ, КНЕУ. 2014. С. 172—190.
3. Методи і моделі економічного прогнозування : навч. посіб. / Кічор В.П. та ін.; Львів, 2019. С.138—157.

- 6.1. Метод експоненційного згладжування
- 6.2. Модель ARIMA: ідентифікація та прогностні властивості
- 6.3. Застосування моделі ARIMA в аналізі інтервенцій

6.1. Метод експоненційного згладжування

Важливу роль у підвищенні адекватності прогнозів відіграють адаптивні методи прогнозування. Суть адаптації полягає у пристосуванні прогностної моделі до мінливої динаміки досліджуваного процесу шляхом послідовного використання поточних (нових чи оновлених) даних. Біля витоків адаптивного напрямку моделювання часових рядів лежить модель експоненційного згладжування, узагальнення якої привело до появи сімейства досить гнучких, самоналагоджувальних моделей, здатних враховувати інформаційну цінність різних членів часового ряду, що дозволяє виявляти зміни в тенденції, а також будь-які коливання, в яких простежується закономірність. Серед них метод авторегресії, гармонійних ваг, сплайн-функції та інші.

Експоненційне згладжування (*Exponential Smoothing*) — один із простих і прагматично зрозумілих методів моделювання динамічних процесів. В основу його закладена ідея постійного перегляду прогностних значень з урахуванням надходження нових даних. Як зазначено в підрозд. 3.1, суть експоненційного згладжування полягає в заміні первинних рівнів часового ряду адаптивними середніми Y_t , розрахованими з використанням спеціальної системи ваг $a(1-a)^r$:

$$Y_t = \sum_{r=0}^t a(1-a)^r y_{t-r}.$$

Послідовно розклавши формулу експоненційної середньої за елементами суми, дістанемо

$$Y_t = ay_t + a(1-a)y_{t-1} + a(1-a)^2y_{t-2} + \dots + a(1-a)^t y_0$$

або

$$Y_t = ay_t + (1 - a)[ay_{t-1} + a(1 - a)y_{t-2} + \dots + a(1 - a)^{t-1}y_0].$$

Друга складова останньої формули — не що інше, як експоненційна середня для $(t - 1)$ -ої одиниці часу. Отже, експоненційну середню можна подати як лінійну комбінацію поточного і попередніх рівнів ряду:

$$Y_t = ay_t + (1 - a)Y_{t-1}.$$

Для розрахунку експоненційної середньої Y_t необхідно визначити початкову величину Y_0 і параметр згладжування (адаптації) — a . В якості початкової величини можна використати середній рівень за минулий (до часового ряду) період або за відсутності таких даних — перший рівень ряду, тобто $Y_0 = y_1$.

Вибір оптимального значення параметра адаптації a — більш складне завдання експоненційного згладжування. Необхідно врахувати, що зі збільшенням значення a зростає вага поточних спостережень, тобто задля підвищення швидкості реакції на мінливу динаміку процесу необхідно збільшити значення a , але водночас зменшуються фільтраційні можливості експоненційної середньої. Точного методу для вибору оптимальної величини параметра a немає. В окремих випадках пропонують визначити його, виходячи з довжини інтервалу згладжування m :

$$a = \frac{2}{m+1} \quad \text{або} \quad 1 - a = \frac{m-1}{m+1},$$

де m — кількість останніх членів часового ряду, на яких ґрунтується прогноз.

Зважаючи на залежність суми вагових коефіцієнтів $\sum a_r$ від параметра a на певному часовому інтервалі m , можна для наперед заданої суми $\sum a_r$ орієнтовно визначити параметр a :

$$a = 1 - \sqrt[m]{1 - \sum_1^m a_r}.$$

Наприклад, для інтервалу згладжування $m = 10$ сума ваг $\sum a_r = 0,90$, звідси $a = 1 - \sqrt[10]{1 - 0,9} \approx 0,2$. Тобто при $a = 0,2$ приблизно 90 % ваг експоненційної середньої припадає на десять останніх членів часового ряду. На практиці найчастіше використовують значення a в інтервалі від 0,1 до 0,3.

Надаючи більшу вагу новій інформації, експоненційна середня швидше реагує на зміну тренду, легко адаптується до нових

умов, а тому розглядається як дієвий інструмент короткострокового прогнозування.

У процесі прогнозування вдаються до подвійного і потрійного згладжування, тобто згладжений ряд послідовно піддається новому згладжуванню. Експоненційна середня другого порядку Y_t^* визначається за такою самою рекурентною формулою на основі згладженого ряду Y_t :

$$Y_t^* = aY_t + (1 - a)Y_{t-1}.$$

Якщо припустити наявність лінійного тренду, прогнозний рівень Y_{t+1} можна розрахувати за формулою

$$Y_{t+1} = \frac{(2 - a)Y_t - Y_t^*}{1 - a}.$$

Для конкретних значень параметра a ця формула набуває вигляду:

- для $a = 0,1$ $Y_{t+1} = 2,111 Y_t - 1,111 Y_t^*$
- для $a = 0,2$ $Y_{t+1} = 2,250 Y_t - 1,250 Y_t^*$
- для $a = 0,3$ $Y_{t+1} = 2,429 Y_t - 1,429 Y_t^*$

Довірчі межі прогнозного рівня визначаються традиційно

$$Y_{t+1} \pm t_{1-\alpha} \sigma_y \sqrt{1 + \frac{a}{2 - a}},$$

де $\sigma_y = \sqrt{\frac{\sum (y_t - \bar{y})^2}{n - 1}}$ — стандартне відхилення первинного часового ряду;

t — статистика критерію Стьюдента для ймовірності $(1 - \alpha)$.

Очевидно, що за умови значної варіації рівнів часового ряду довірчі межі будуть досить широкими.

Як ілюстрацію методики застосування експоненційного згладжування для прогнозування розглянемо біржові котирування цін на фуражний ячмінь (дол./т) на торгах протягом тижня (табл. 6.1). Коефіцієнт адаптації $a = 0,2$. Початковий рівень згладжування $Y_1 = 223$, розраховані за рекурентною формулою експоненційні середні першого порядку Y_t , становлять:

$$Y_2 = 229 \cdot 0,2 + 223 \cdot 0,8 = 224,20,$$

$$Y_3 = 224 \cdot 0,2 + 224,20 \cdot 0,8 = 224,16 \text{ і т. д.}$$

Аналогічно визначаються експоненційні середні другого порядку Y_t^* :

$$Y_2^* = 224,20 \cdot 0,2 + 223 \cdot 0,8 = 223,21,$$

$$Y_3^* = 224,16 \cdot 0,2 + 223,21 \cdot 0,8 = 223,42 \text{ і т. д.}$$

Таблиця 6.1

ДО РОЗРАХУНКУ ЕКСПОНЕНЦІЙНИХ СЕРЕДНИХ

День тижня	Біржова ціна, дол./т, y_t	Y_t	Y_t^*
Понеділок	223	223,00	223,00
Вівторок	229	224,20	223,21
Середа	224	224,16	223,42
Четвер	226	224,53	223,64
П'ятниця	232	226,02	224,12
Прогноз ціни на понеділок	228,40	×	×

На понеділок прогноуються біржові котирування цін на фуражний ячмінь у розмірі: $2,250 \cdot 226,02 - 1,250 \cdot 224,12 = 228,40$ дол./т.

Базову модель експоненційного згладжування можна застосувати при моделюванні рядів, яким властива сезонна компонента. Алгоритми прогнозування на основі експоненційного згладжування реалізовані в пакетах прикладних програм. Зокрема, у системі *Statistica* експоненційне згладжування здійснюють за алгоритмом *Exponential Smoothing & Forecasting*. У діалоговому вікні *Seasonal and Non-Seasonal Exponential Smoothing* (рис. 6.1) необхідно визначитися щодо наявності та виду тренду, сезонної компоненти і параметра згладжування:

- коли часовий ряд не містить сезонної компоненти, установки щодо наявності та типу тренду необхідно здійснити в полі *None*, за відсутності тренду — *No trend*; лінійний тренд — *Linear trend*, експонента — *Exponential*, загасаючий тренд — *Damped trend*;

- для часових рядів помісячних чи поквартальних даних необхідно вказати лаг (lag) сезонної компоненти *Seasonal component* — 12 чи 4 та спосіб її взаємодії з іншими компонентами часового ряду: адитивний — у полі *Additive*, мультиплікативний — у полі *Multiplicative*;

– у полі параметрів згладжування необхідно вказати параметр $Alpha$, інші параметри ($Delta$, $Gamma$, Phi) — для спеціальних моделей.

Праворуч в опції *Forecast* вказують період попередження прогнозу.

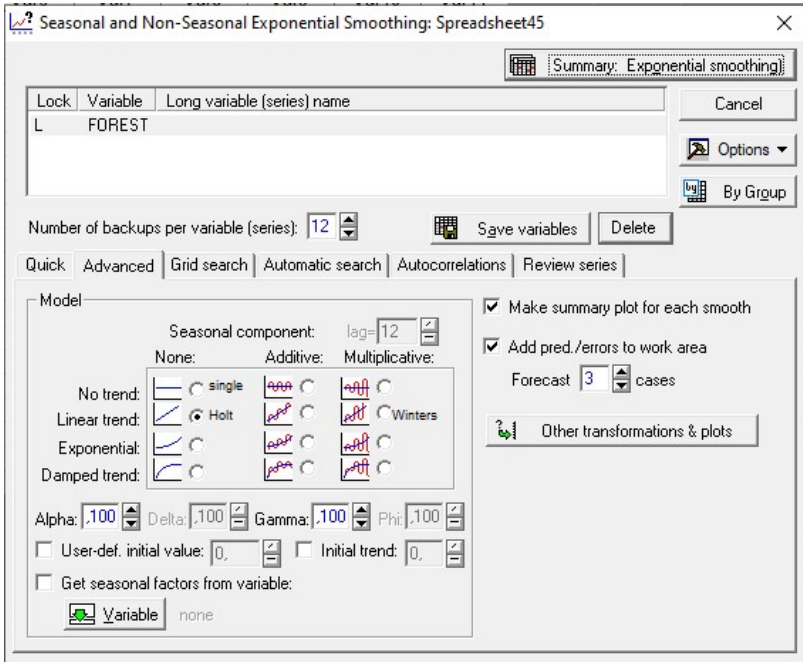


Рис. 6.1. Діалогове вікно експоненційного згладжування

Як приклад, здійснимо згладжування часового ряду природного поновлення лісів в регіоні, тис га. Установки: *Linear trend*, параметр $Alpha = 0,100$, період попередження *Forecast cases* — 3. За командою на виконання *Summary: Exponential Smoothing* система автоматично визначить оптимальні параметри моделі та прогнозні значення. Візуально експоненційне згладжування часового ряду унаочнює рис. 6.2, а результати прогнозування подаються у таблицях 6.2. та 6.3.

Табл. 6.2 містить фрагмент згладженого часового ряду природного поновлення лісів. Залишки, тобто відхилення між теоретичними і первинними рівнями ряду, дають підстави судити про адекватність моделі реальному процесу.

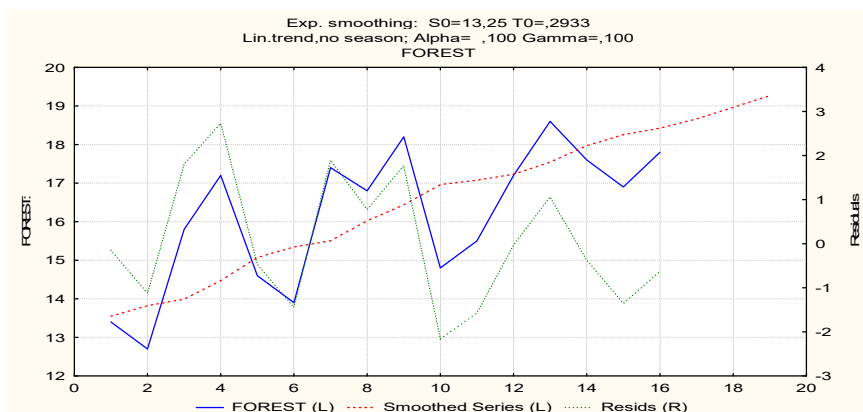


Рис. 6.2. Експоненційне згладжування площі природного поновлення лісу в регіоні, тис.га

Таблиця 6.2

**ПРОГНОЗНІ ЗНАЧЕННЯ ПРИРОДНОГО ПОНОВЛЕННЯ ЛІСУ
В РЕГІОНІ, тис. га**

Case	Exp. Smoothing: SO=13,25 TO=,2933 Lin. trend, no season; Alpha =, 100 FOREST		
	FOREST	Smoothed Series	Resids
1	13,4	13,5	-0,15
2	12,7	13,8	-1,12
3	15,8	14,0	1,81
4	17,2	14,5	2,73
5	14,6	15,1	-0,47
6	13,9	15,5	-1,44
7	17,4	15,3	1,89
8	16,8	16,0	0,78
9	18,2	16,4	-2,16
10	14,8	17,0	-1,57
11	15,5	17,1	-0,03
12	17,2	17,2	1,06
13	18,6	17,5	-0,37
14	17,6	18,0	-1,35
15	16,9	18,3	-0,63
16	17,8	18,4	
17		18,7	
18		19,0	
19		19,3	

Високі апроксимуючі властивості моделі експоненційного згладжування часового ряду природного поновлення лісу підтверджують оцінки похибок прогнозу (табл. 6.3). Серед наведених: середня абсолютна похибка (MAE) — 1,209 тис. га, похибка апроксимації (MAPE) — 7,5 %; від'ємне значення середньої процентної похибки (MPE = -0,246) свідчить про те, що модель дає послідовно завищені прогнози, що підтверджують прогнозні значення в табл. 6.2.

Таблиця 6.3

ОЦІНКИ ПОХИБОК ПРОГНОЗУ ПРИРОДНОГО ПОНОВЛЕННЯ ЛІСУ

Summary of error	Exp. Smoothing: SO=13,25 TO=,2933 Lin. trend, no season; Alpha =, 100 Gamma =, 100 FOREST
	Error
Mean of error	0,045
Mean absolute error	1,209
Sums of squares	32,253
Mean squares	2,016
Mean percentage error	-0,246
Mean abc. perc. error	7,523

Зауважимо, що метод експоненційного згладжування працює на довгих часових рядах, тож на коротких рядах, особливо за наявності тренду, метод не встигає відобразити всі зміни, що певною мірою впливає на точність прогнозу.

Експоненційне згладжування в технічному аналізі ринків

Завдяки своїм прогностичним властивостям експоненційне згладжування часових рядів застосовують у технічному аналізі фінансових ринків (валютному, фондовому). Технічний аналіз — це метод прогнозування зміни ринкової ціни у майбутньому на основі аналізу цінової динаміки у минулому. На будь-якому інтервалі часового ряду цін переважає певний напрямок руху — тренд, сила якого залежить від фази життєвого циклу динамічного процесу. Вирізняють три типи тренду:

- висхідний (bullish), коли кожна наступна ціна вища за попередню;
- спадний (bearish), коли кожна наступна ціна менша за попередню;
- боковий (flat, sideward) — період консолідації цін.

Технічний аналіз, абстрагуючись від причинного механізму формування кон'юнктури ринку, відстежує траєкторію цінової динаміки, виявляє панівний тренд і прогнозує ймовірні його зміни. Найбільш складним у технічному аналізі є передбачення моменту розвороту тренду. Для цього використовують трендові лінії, що окреслюють трендовий канал: на висхідному ринку знизу — лінія підтримки (support line), зверху — лінія опору (resistance line); на спадному ринку навпаки, зверху — лінія підтримки, знизу — лінія опору. Лінія підтримки — це лінія зафіксованих за певний проміжок часу мінімальних цін, лінія опору — лінія зафіксованих максимальних цін. Перетин часовим рядом лінії опору чи лінії підтримки сигналізує про можливу зміну тренду. Рис. 6.3 ілюструє динаміку розвороту курсових котирувань EUR/USD. Чітко видно розворот тренду 4 жовтня, спадний тренд змінює свій напрямок, перетинає трендову лінію опору, і остання перетворюється на лінію підтримки.



Рис. 6.3. Розворот тренду курсових котирувань EUR/USD

У моніторингу цін на фондових і валютних біржах традиційно використовують дві експоненційні середні цін закриття з різними періодами згладжування: швидку (12-ти денну) і повільну (26-денну) з параметрами адаптації 0,15 і 0,075, відповідно. Лінії цих середніх розглядаються як трендові лінії підтримки й опору. Значне відхилення між ними свідчить про силу тренду, а їх перетин дає упереджуваний сигнал про можливі його зміни. Коли швидка середня перетинає повільну середню зверху, це сигнал про появу нового спадного тренду, коли знизу — про появу нового висхідного тренду.

Завдяки сучасним засобам телекомунікацій та комп'ютерним мережам інформація щодо біржових котирувань доступна в режимі реального часу. З огляду на основне правило ринку «торгуй за трендом», в аналізі інформації використовують низку статистичних індикаторів, які здійснюють тестування кон'юнктури ринку, сигналізують про зміни біржового тренду, виявляють тренд, за яким торгівля мінімально ризикова (рис. 6.4).



Рис. 6.4. Торговля по конвергенції/дивергенції MACD

До більшості торговельних платформ для фінансових і товарних ринків включений індикатор MACD (Moving Average Convergence/Divergence — сходження / розбіжність ковзних середніх). MACD-аналіз включає три послідовні операції:

- 1) визначається *оперативна MACD-лінія* — лінія відхилень між швидкою (12 днів) і повільною (26 днів) експоненційними середніми; вона допомагає визначити напрямок ринкового тренду — висхідний чи спадний;

2) оперативна лінія згладжується за допомогою 9-ти денної ковзної середньої, згладжена лінія виконує роль *сигнальної*;

3) *відхилення між оперативною і сигнальною лініями* застосовують при визначенні потенційних точок розвороту тренду.

На сигналах трендових індикаторів біржових цін базується правило buy/sell: купувати, коли лінія MACD перетинає сигнальну лінію знизу догори, продавати, коли лінія MACD перетинає сигнальну лінію зверху донизу. Проте, слід пам'ятати, що система моніторингу — це лише інструмент прийняття рішень. Інструмент, яким трейдер фінансового ринку має досконало володіти, аби зменшити ризик збитків і примножити потенціал прибутків.

6.2. Модель *ARIMA*: ідентифікація та прогнозні властивості

ARIMA (AutoRegressive Integrated Moving-Average) — це клас лінійних статистичних моделей, які ефективно описують як стаціонарні, так і нестаціонарні часові ряди. Формується *ARIMA* об'єднанням моделей авторегресії порядку p і ковзної середньої залишків порядку q . Для моделі *ARIMA* важливо, аби часовий ряд був стаціонарним, тобто середня була сталою, а дисперсія та автокореляція не змінювалися в часі. Тому нестаціонарні ряди приводяться до стаціонарного виду за допомогою оператора кінцевих різниць порядку d : для фільтрації лінійного тренду використовують різниці першого порядку ($d = 1$), для фільтрації параболічного тренду — різниці другого порядку ($d = 2$) і т. д. Процедуру послідовного взяття різниць застосовують допоки ряд не стане стаціонарним.

Модель авторегресії — $AR(p)$ описує залежність поточного рівня часового ряду y_t від попередніх значень із часовим лагом p . Наприклад, процес авторегресії першого порядку ($p = 1$) описує ступінь залежності двох послідовних членів ряду:

$$y_t = \phi_1 y_{t-1} + e_t,$$

де y_t, y_{t-1} — поточне і попереднє значення показника;

ϕ_1 — коефіцієнт авторегресії;

e_t — випадковий шум (залишки).

Для оцінки ступеня залежності рівнів ряду використовують коефіцієнти автокореляції r_p , які функціонально пов'язані з коефіцієнтами авторегресії; в $AR(1)$ -процесі $a_1 = r_1$.

Модель ковзної середньої — $MA(q)$ описує лінійну комбінацію нормально розподілених випадкових величин з нульовим середнім:

$$y_t = \mu + e_t - \theta e_{t-1},$$

де μ — константа (довгострокове середнє значення процесу);

e_t — випадковий шум поточного періоду;

θe_{t-1} — попередній випадковий шум.

Загальний вигляд змішаної моделі $ARIMA(p, q)$ -процесу:

$$y_t = c + a_1 y_{t-1} + e_t - \theta e_{t-1}.$$

Підбираючи значення коефіцієнтів a і θ , можна сформулювати модель, яка адекватно опише будь-який часовий ряд з циклічними й нерегулярними коливаннями. На рис. 6.5 схематично подана логічна структура моделі $ARIMA$.

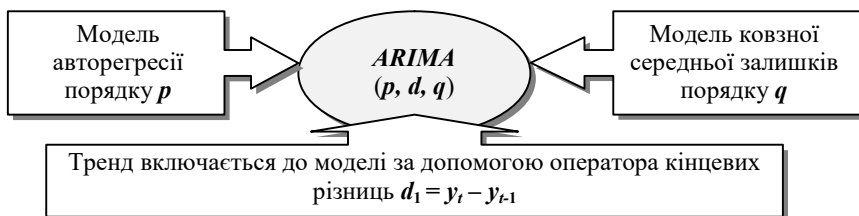


Рис. 6.5. Логічна структура моделі $ARIMA$

Модель $ARIMA$ порядку (p, d, q) досить гнучка й описує широкий спектр процесів. Найпростіші види моделей $ARIMA$ для несезонних процесів:

$(p, 0, 0)$ — авторегресійна функція;

$(0, 0, q)$ — ковзна середня;

$(p, 0, q)$ — комбінована модель авторегресії і ковзної середньої;

$(0, 1, q)$ — експоненційна середня;

$(p, 1, q)$ — нестационарний процес з лінійним трендом.

На практиці більшість часових рядів можна апроксимувати однією з названих моделей з достатнім ступенем точності. За наявності сезонних коливань у моделі враховується їхня періодичність з лагом s , що дорівнює сезонному циклу (для квартальних даних $s = 4$, для помісячних $s = 12$), і аналогічного змісту сезонними параметрами (ps, ds, qs) . Отже, повну сезонну модель можна подати як $ARIMA(p, d, q)(ps, ds, qs)$. Наприклад, $ARIMA(0,1,1)(0,1,1)_s$ містить нуль регулярних параметрів авторегресії ($p = 0$) і один параметр ковзної середньої ($q = 1$). Ці параметри

визначаються після фільтрування тренду за допомогою оператора різниць ($d = 1$) і оператора сезонних різниць ($ds = 1$).

Оскільки вид моделі *ARIMA*, адекватність її реальному процесу і прогнозні властивості залежать від параметрів p , d , q , то ключовим моментом моделювання є визначення цих параметрів — *ідентифікація моделі*. Вирізняють два аспекти ідентифікації:

- аналіз стаціонарності й визначення порядку оператора переходу до кінцевих різниць d ;
- визначення порядку змішаної моделі, яка описує стаціонарний ряд (параметрів p , q).

Основним інструментом ідентифікації моделі стаціонарного часового ряду є автокореляційні функції. Обґрунтування порядку параметрів моделі зводиться до візуального аналізу автокорелограм і ґрунтується на принципі економії, за яким $(p + q) \leq 2$. Сезонний лаг також визначають на етапі ідентифікації моделі. Порядок визначення параметрів моделі *ARIMA* за видом автокореляційних функцій подано в табл. 6.4.

Таблиця 6.4

**ВИЗНАЧЕННЯ ПОРЯДКУ ПАРАМЕТРІВ МОДЕЛІ *ARIMA*
ЗА ВИДОМ АВТОКОРЕЛЯЦІЙНОЇ ФУНКЦІЇ**

Кількість параметрів моделі <i>ARIMA</i>	Автокореляційна функція, <i>ACF</i>	Частинна автокореляційна функція, <i>PACF</i>
<i>Один параметр авторегресії</i>	Експоненціально згасає	Найбільше значення коефіцієнта з лагом 1, немає кореляції на інших лагах
<i>Два параметри авторегресії</i>	Має форму синусоїди або експоненціально згасає	Найбільше значення коефіцієнтів з лагами 1, 2, немає кореляції на інших лагах
<i>Один параметр ковзної середньої</i>	Має найбільше значення коефіцієнта на лагу 1, відсутня кореляція на інших лагах	Згасає монотонно або змінює знак
<i>Два параметри ковзної середньої</i>	Має найбільші значення коефіцієнтів на лагах 1, 2, відсутня кореляція на інших лагах	Має форму синусоїди або експоненціально згасає
<i>Один параметр авторегресії і один параметр ковзної середньої</i>	Експоненціально згасає після лагу 1	Експоненціально згасає після лагу 1

Практичне використання моделі *ARIMA* пов'язують з іменами Г. Бокса і Г. Дженкінса [1]. Запропоновані ними процедури формування моделі, алгоритми розрахунку параметрів і визначення прогнозів передбачають використання програмних засобів. В системі *Statistica* *ARIMA*-модель реалізується в модулі *Time Series /Forecasting*, процедура *ARIMA & Autocorrelation Functions*.

Як приклад ідентифікації *ARIMA*-моделі розглянемо часовий ряд вартості офісної нерухомості (дол./м²). Перший висновок щодо стаціонарності (нестаціонарності) ряду можна зробити, аналізуючи автокореляційну функцію часового ряду (рис. 6.6). Поступове зменшення коефіцієнтів автокореляції залежно від лагу свідчить про нестаціонарний характер первинного ряду. Для фільтрації тренду використаємо оператор різниць першого порядку. Після заміни первинного ряду першими різницями автокореляційна функція перетвореного ряду зі зростанням лагу експоненційно згасає (рис. 6.7), що дає підстави розглядати ряд перших різниць як стаціонарний.

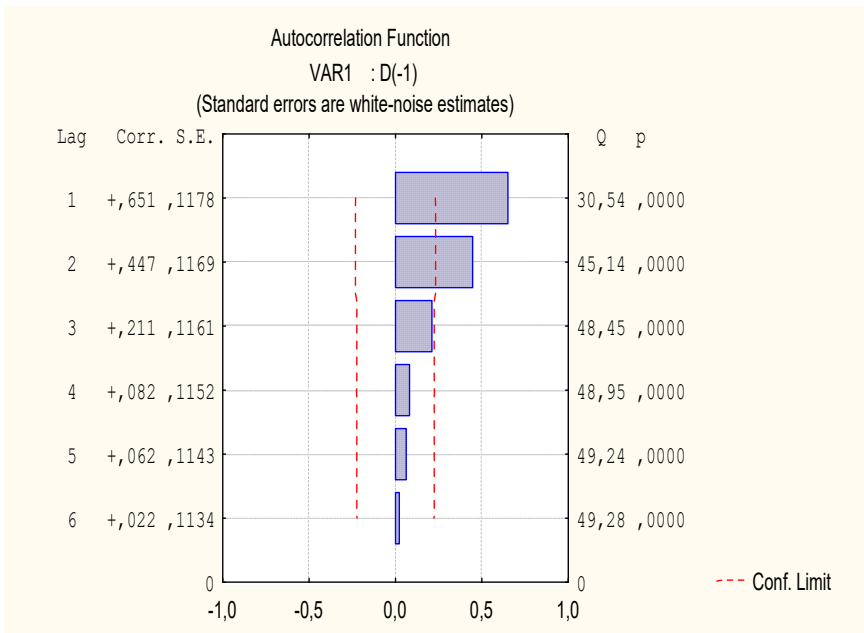


Рис. 6.6. Автокореляційна функція часового ряду вартості офісної нерухомості

На лагу $p = 1$ коефіцієнти автокореляцій однакові (у нашому прикладі 0,651). Найбільше значення частинної автокореляційної функції з одиничним лагом вказує на порядок авторегресії $p = 1$. Отже, часовий ряд вартості офісної нерухомості можна описати моделлю $ARIMA(1, 1, 0)$.

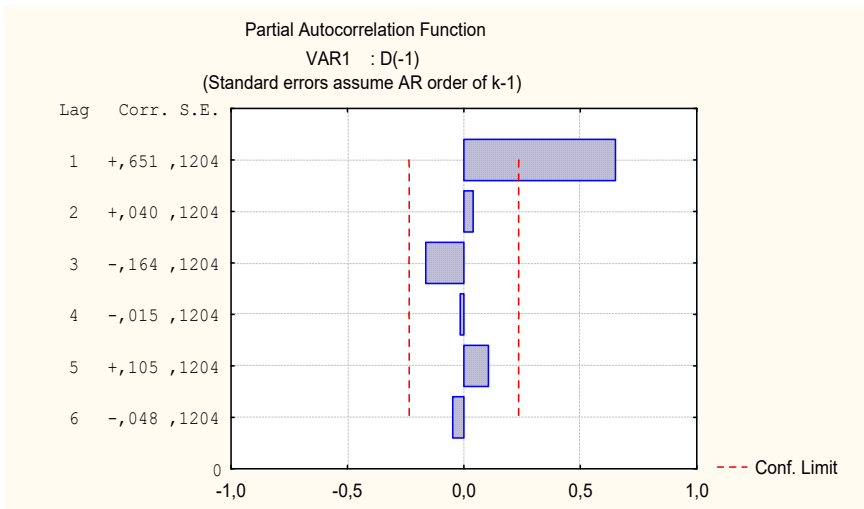


Рис. 6.7. Частинна автокореляційна функція часового ряду вартості офісної нерухомості

Наступний крок після ідентифікації моделі — оцінювання параметрів моделі. Він ґрунтується на процедурі мінімізації функції втрат. Оцінювання проводиться за перетвореними даними (після застосування оператора різниць). Адекватна реальному процесу модель $ARIMA(p, d, q)$ є базою визначення прогнозних значень ряду та їхніх довірчих меж.

Для ідентифікації $ARIMA$ -моделі у діалоговому вікні *Single Series ARIMA* передбачена спеціальна група опцій *ARIMA model parameters* (рис. 6.8), які фіксують порядок параметрів моделі (*Estimate constant, Seasonal lag*) та опції трансформації ряду:

p — *Autoregressive* — параметр авторегресії (регулярний);

P — *Seasonal* — сезонний параметр авторегресії;

q — *Moving average* — параметр ковзної середньої (регулярний).

Q — *Seasonal* — сезонний параметр ковзної середньої.

Рис. 6.8. Діалогове вікно процедури *Single Series ARIMA*

Оскільки модель часового ряду вартості офісної нерухомості ідентифікована як $ARIMA(1,1,0)$, у діалоговому вікні *Single Series ARIMA* необхідно вказати: $p = 1$, *Difference* (1. Lag — 1, № of passes — 1). Параметри моделі наведено в табл. 6.5. Значення параметра $p(1)$, його стандартна похибка (0,057), t -критерій (13,902), а також p -level < 1 % свідчать про адекватність моделі.

Таблиця 6.5

**ОЦІНКА ПАРАМЕТРІВ МОДЕЛІ $ARIMA(1, 1, 0)$ ДЛЯ ЧАСОВОГО РЯДУ
ВАРТОСТІ ОФІСНОЇ НЕРУХОМОСТІ, дол./м²**

	Input: APART-1 : Transformations: D(1) Model:(1,1,0) MS Residual=0,00208					
Paramet.	Param.	Asympt. Std.Err.	Asympt. t(120)	p	Lower 95% Conf	Upper 95% Conf
p(1)	0,788	0,057	13,903	0,000	0,675	0,902

Розраховані на основі моделі $ARIMA(1,1,0)$ прогнози значення вартості офісної нерухомості на період попередження $v = 1, 2, 3$ наведено в табл. 6.6. Як свідчать дані, зі збільшенням періоду попередження стандартна похибка зростає, відповідно збільшується ширина довірчого інтервалу.

Таблиця 6.6

ПРОГНОЗ ВАРТОСТІ ОФІСНОЇ НЕРУХОМОСТІ

Forecasts; Model:(1,1,0) Input: APART-1: Start of origin: 1 End of origin: 60				
CaseNo.	Forecast	Lower 90,0000%	Upper 90,0000%	Std.Err.
61	2239,000	2163,000	2315,00	45,000
62	2229,000	2074,000	2384,00	93,000
63	2220,000	1981,000	2459,00	144,000

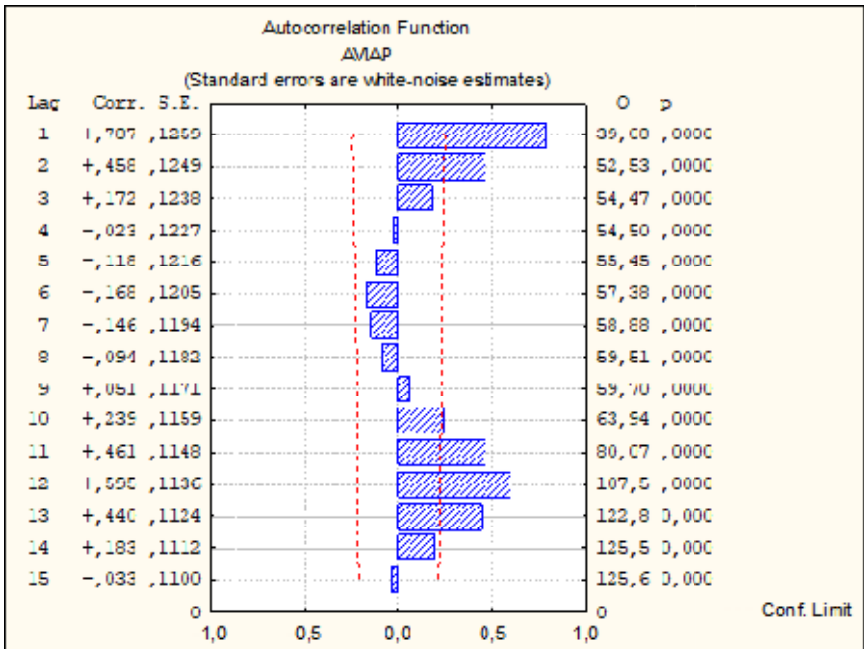


Рис. 6.9. Автокореляційна функція часового ряду перевезення авіапасажирів

Порядок ідентифікації *ARIMA*- моделі на основі корелограм поширюється на моделювання сезонних процесів. Єдина відмінність полягає у тому, що в сезонних рядах коефіцієнти автокореляції будуть істотними також на лагах, кратних сезонному циклу, що ілюструє автокореляційна функція часового ряду кількості перевезених авіакомпанією пасажирів (рис. 6.9). Сезонність — 12 місяців. Аби наблизити такий ряд до стаціонарного, необхідно фільтрувати не лише тренд, а й сезонну компоненту, для цього в діалоговому вікні (рис. 6.8, рамка *Difference*) вказати другий, сезонний *Lag2*.

Модель часового ряду перевезення авіапасажирів набуває вигляду $ARIMA(0,1,1)(0,1,1)_s$. Наведені в табл. 6.7 значення параметрів моделі, *t*-критерія і *p-level* свідчать про адекватність моделі.

Таблиця 6.7

ОЦІНКА ПАРАМЕТРІВ МОДЕЛІ $ARIMA(0,1,1)(0,1,1)_s$
ДЛЯ ЧАСОВОГО РЯДУ ПЕРЕВЕЗЕННЯ АВІАПАСАЖИРІВ

Paramet.	Input: AVIAP (Spreadsheet1) Transformations: D(1),D(1) Model:(0,1,1)(0,1,1) Seasonal lag: 12 MS Residual= 9,9864					
	Param.	Asympt. Std.Err.	Asympt. t(56)	p	Lower 95% Conf	Upper 95% Conf
q(1)	0,979	0,025	38,841	0,000	0,928	1,029
Qs(1)	-0,422	0,107	-3,928	0,000	-0,637	-0,207

За опцією *Forecast cases* програма визначає прогнози для одного повного сезонного циклу, починаючи з останнього спостереження.

Оцінювання адекватності *ARIMA*-моделі здійснюють на основі аналізу залишків (*Residuals*), тобто відхилень між фактичними і визначеними в рамках моделі рівнями ряду $e_t = (y_t - Y_t)$. В процесі аналізу залишків перевіряється дві передумови адекватності моделі:

- залишки нормально розподілені;
- залишки незалежні, тобто відсутня автокореляція.

У модулі *Time Series/Forecasting* дотримання цих передумов перевіряють візуалізацією гістограм залишків, автокореляційних функцій, інших графіків (вкладки *Distribution of residuals* та *Autocorrelations*). Як свідчить подана на рис. 6.10 автокорелограма, залишки практично некорельовані, що також підтверджує адекватність *ARIMA*-моделі реальному процесу.

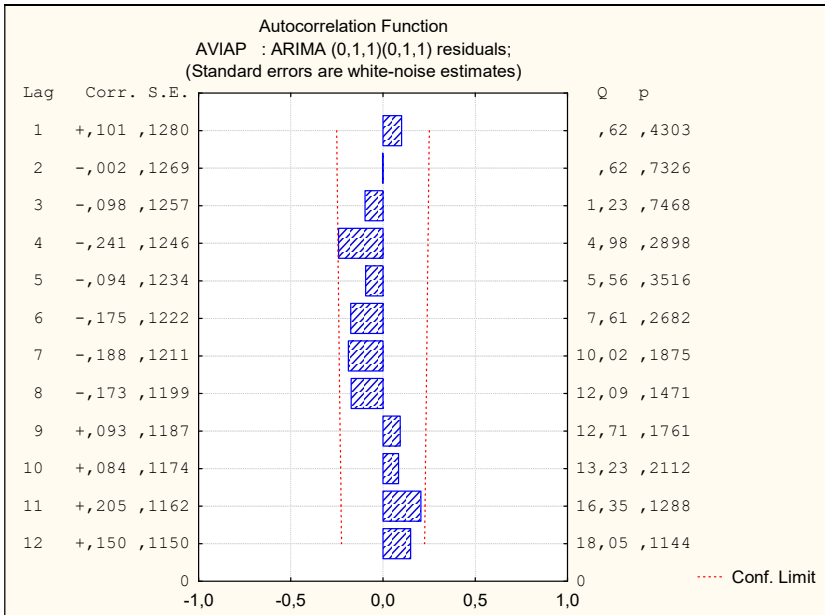


Рис. 6.10. Автокореляція залишків *ARIMA*-моделі кількості перевезених авіапасажирів

Отже, об'єднуючи різні компоненти часового ряду, *ARIMA*-модель адаптується до специфіки будь-якого динамічного процесу і забезпечує належну точність прогнозу.

6.3. Застосування моделі *ARIMA* в аналізі інтервенцій

Багато динамічних процесів під впливом раптових зовнішніх збурень (інтервенцій) стрімко змінюються і набувають стрибкоподібного виду, що значно ускладнює прогнозування їх у середньостроковій перспективі. Скажімо, зміна форми власності, зміна законодавства, фінансова криза тощо. В процесі аналізу часових рядів такого типу важливо виявити, як саме інтервенція впливає на характер динаміки і оцінити її ефект. Наприклад, чи вплинули зміни до кримінального кодексу на інтенсивність злочинів, чи сприяли зміни в податковій системі додатковим надходженням до бюджету і т. ін. Методологічною основою аналізу перерваного часового ряду є статистичне порівняння тенденцій ряду до і після інтервенції.

Коли часовому ряду властивий лінійний або приведений до лінійного виду тренд, найпростіший варіант оцінювання ефекту впливу інтервенцій на тенденцію розвитку — це сегментована регресія. Нехай точка зламу тенденції t_p відома, часовий ряд у цій точці поділяється на два сегменти [до і після]. Для ідентифікації сегментів у трендову модель лінійно вводяться дві *dummy*-змінні u_t . Значення першої змінної u_1 в інтервалі до точки зламу зростають від 1 до t_p , у другому — фіксуються на рівні t_p . Другій змінній u_2 до точки зламу приписуються значення 0, після цієї точки $u_t = 1, 2, \dots, n$, де n — останнє значення ряду [3, с. 148-149]. Часовий ряд такого типу наведено в табл. 6.8, точка зламу тенденції — п'ятий рівень t_5 . Трендова модель має вигляд:

$$Y_t = a_0 + a_1 u_1 + a_2 u_2.$$

Параметри трендової моделі a_1 і a_2 характеризують середню абсолютну швидкість динаміки, відповідно, до і після зламу тенденції.

Таблиця 6.8

**МАТРИЦЯ DUMMY-ЗМІННИХ
ДЛЯ УРАХУВАННЯ ЗЛАМУ ТЕНДЕНЦІЇ**

№ з/п	1	2	3	4	5	6	7	8	9
y_t	12	13	16	19	20	21	22	23	24
u_1	1	2	3	4	5	5	5	5	5
u_2	0	0	0	0	0	1	2	3	4

За даними таблиці $Y_t = 9,476 + 2,162 u_1 + 0,905 u_2$. Отже, до точки зламу тенденції t_5 абсолютний приріст рівнів ряду становив у середньому 2,162, після зламу — 0,907, тобто інтервенція призвела до зменшення приросту функції на 1,255. Параметр $a_0 = 9,476$ є тим значенням, яке відсікає на осі ординат перша лінія тренду. Зауважимо, що через значну варіацію чи складність моделювання автокореляційної структури не всі часові ряди піддаються сегментованій регресії.

Однією з альтернатив аналізу інтервенцій є модель *Interrupted ARIMA (Intervention Analysis)* — перервана *ARIMA*, яка, враховуючи основні тенденції, автокореляцію та сезонність, здатна гнучко моделювати різні типи впливів.

Діалогове вікно процедури *Interrupted ARIMA* наведено на рис. 6.11.

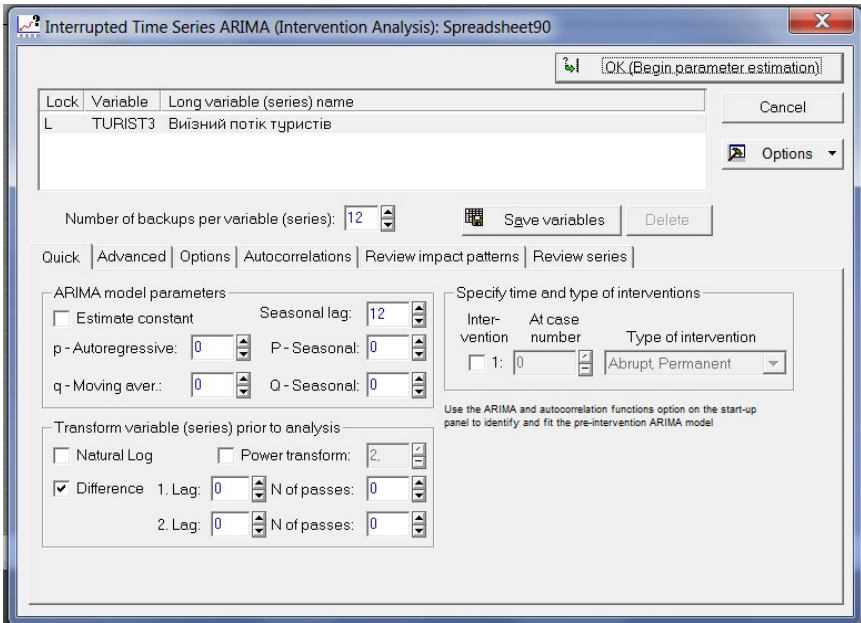


Рис. 6.11. Діалогове вікно процедури *Interrupted Time Series ARIMA (Intervention Analysis)*

На відміну від діалогового вікна одновимірної *ARIMA*, праворуч розміщено блок опцій, призначених для специфікації інтервенцій, — *Specify time and type of interventions* — Вказати час і тип інтервенції:

- *Intervention* — номер інтервенції (до 6 різних інтервенцій);
- *At case number* — номер спостереження, в якому інтервенція почалася (момент інтервенції);
- *Type of intervention* — тип інтервенції.

Часовий ряд з інтервенціями ілюструє рис. 6.12: кількість інтервенцій — 3, початок їх в точках 49, 61 і 73.

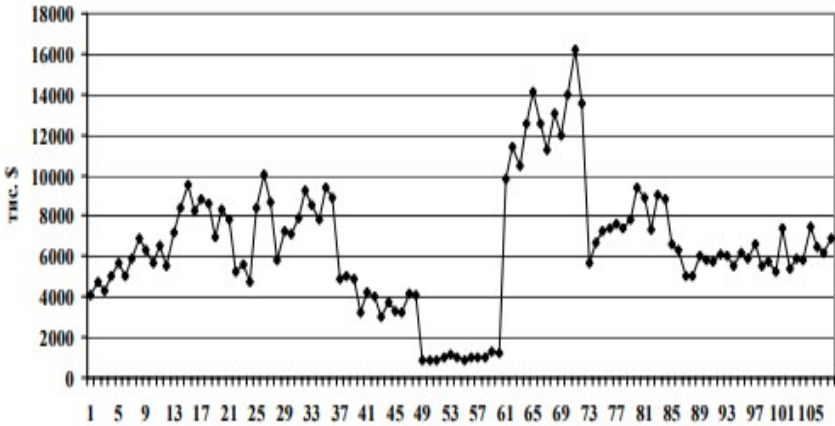


Рис. 6.12. Часовий ряд з інтервенціями в точках 49, 61 і 73.

Процедурою передбачено три типи інтервенції:

- *Abrupt, Permanent* — стрибкоподібний стійкий, коли часовий ряд стрімко обривається до певного рівня і на цьому рівні залишається;
- *Gradual, Permanent* — поступово стійкий, спричиняє поступове стійке збільшення чи зменшення значень ряду;
- *Abrupt, temporary* — стрибкоподібний тимчасовий, коли миттєва інтервенція спричиняє стрімкий стрибок ряду, який поступово зменшується, і характер зміни часового ряду відновлюється.

За опцією *Review types of impact patterns* можна переглянути типи інтервенції, вибрати серед них найбільш прийнятний і перейти до ідентифікації моделі та визначення параметрів аналогічно як в одновимірній *ARIMA*.

Як приклад розглянемо часовий ряд втрачених записів від кібербезпечності за всіма причинами (рис. 6.13). На тлі висхідної динаміки кількості втрачених чи скомпрометованих записів втрати в період [10—12] були безпрецедентно високими. Такий тип інтервенції можна ідентифікувати як *Abrupt, temporary*

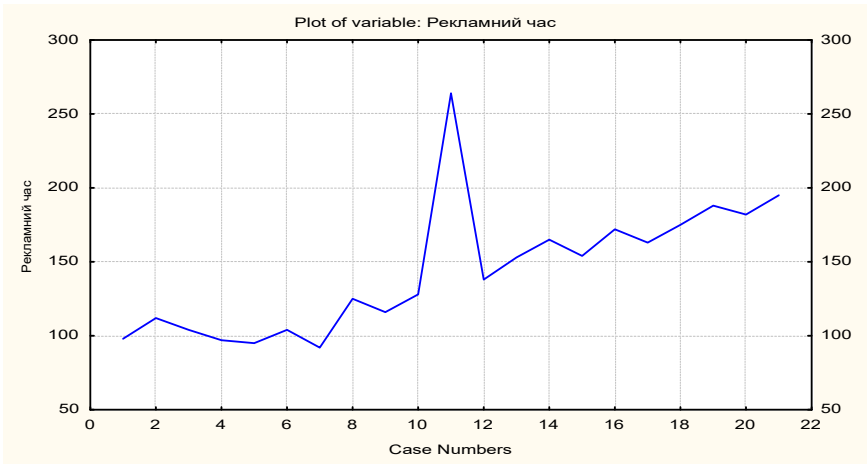


Рис. 6.13. Динаміка кількості втрачених записів від кібербезлочинів

Відповідно до графіка динаміки втрачених записів обираємо тип інтервенції — *Abrupt, temporary* — стрибкоподібний тимчасовий. Фільтрацію тренду здійснює оператор різниць першого порядку ($d = 1$), модель набуває вигляду $ARIMA(1,1,0)$. Модель поряд з параметром авторегресії $p(1)$ містить параметр ω — *Omega (1)*, який вимірює вплив зовнішнього фактора в період інтервенції (табл. 6.10). Вплив цього фактора істотний.

Таблиця 6.10

**ПЕРЕРВАНА ARIMA ДИНАМІКИ ВТРАЧЕНИХ ЗАПИСІВ
ВІД КІБЕРБЕЗЛОЧИНІВ**

Paramet.	Input: КІБЕРЗЛОЧИНИ Transformation: D (1) (Interrupted ARIMA) Model: (1,1,0) Residual = 985,94							
	Param	Asympt Std.Err	Asympt t(18)	p	Lower 95 % Conf	Upper 95 % Conf	Interv Cfse No.	Interv Type
p(1)	-0,369	0,276	-1,336	0,1982	-0,984	0,211		
Omega (1)	99,012	36,919	2,681	0,0152	21,448	176,578	11	Abr/Perm

Параметр ω для раптових постійних утручань можна інтерпретувати як кількість незворотних змін, що відбулися в момент утручання. При інтервенції типу стрибкоподібної, стійкої — *Abrupt, Permanent* параметр ω розглядається як показник стійкості змін ча-

сового ряду. В разі поступово стійкої інтервенції — *Gradual, Permanent* вплив зовнішнього фактора описується двома параметрами *omega* ω і *дельта* δ . Якщо значення цих параметрів не виходять за межі $[0;1]$, динамічний процес вважається стабільним.

Отже, модель типу *Interrupted ARIMA* є гнучким інструментом оцінювання впливу різного типу великомасштабних утручань, враховуючи основні тенденції, автокореляцію і сезонність.

Стосовно вимог до обсягу інформаційної бази моделювання *ARIMA*. Не існує чітких вказівок, скільки часових точок потрібно для застосування моделі. Часто цитоване значення «як мінімум 50 моментів часу» не має емпіричної основи і формально не перевірялося. Насправді універсальний підхід — це спрощення. Чим мінливіші та зашумлені дані, тим більше спостережень потрібно, щоб відрізнити основні закономірності від шуму. У нескладних випадках *ARIMA* може задовільно працювати з короткими часовими рядами, якщо є достатня кількість часових точок для оцінки всіх параметрів. За наявності сезонності має бути достатньо моментів часу для виявлення сезонних ефектів та врахування сезонних відмінностей.

Вправи та питання для самопідготовки

1. Поясніть сутність і сферу використання адаптивних методів прогнозування.
2. Розкрийте особливості метода експоненційного згладжування.
3. Як вибрати оптимальне значення параметра згладжування?
4. Чи можна застосувати експоненційне згладжування для прогнозування часових рядів, що містять сезонну компоненту?
5. Окресліть сферу використання методу експоненційного згладжування.
6. Поясніть методіку використання експоненційних середніх у технічному аналізі фінансових ринків.
7. На основі яких трендових індикаторів біржових цін формується рішення щодо купівлі/продажу?
8. Поясніть логічну структуру моделі *ARIMA*.
9. У чому полягає сутність авторегресії?
10. Як враховуються в моделі *ARIMA* сезонні коливання?
11. Назвіть основні інструменти ідентифікації моделі *ARIMA*.
12. Поясніть структуру моделей: а) *ARIMA* (0, 1, 1); б) *ARIMA* (1, 1, 0); в) *ARIMA* (0,1,1)(0,1,1)₄.
13. У чому полягає сутність сегментованої регресії, коли її використовують?
14. Після реконструкції підприємства тенденція зростання обсягів виробництва змінилася. Точка зламу тенденції (до і після) відома. Чи

можна застосувати для аналізу послідовності такого типу тренду *dummy*-змінні? Скільки *dummy*-змінних необхідно ввести ?

15. Поясніть необхідність і методику оцінювання ефектів інтервенції на основі часового ряду.

16. Які типи інтервенцій реалізовані в модулі *Time Series / Forecasting* процедурами *Interrupted ARIMA*?

17. Протягом тижня ф'ючерси на золото і срібло на ринку Форекс становили, доларів за 1 тр. унцію:

День	понеділок	вівторок	середа	четвер	п'ятниця	понеділок
Срібло	25,31	25,15	25,72	26,11	25,94	25,37
Золото	1847	1865	1846	1857	1848	1853

Методом експоненційного згладжування визначте прогностні курси срібла і золота на понеділок наступного тижня. Параметр згладжування a обґрунтуйте. Порівняйте прогностні курси з фактичними (остання графа таблиці), зробіть висновки.

18. Індекс споживчих цін в країні характеризується даними, %:

№ з/п	1	2	3	4	5	6
%	99,4	108,2	112,3	110,3	111,6	118,5
№ з/п	7	8	9	10	11	12
%	123,3	109,1	99,8	101,5	100,7	100,4

Методом експоненційного згладжування визначте прогностні значення індексу споживчих цін. Параметр згладжування a обґрунтуйте, зробіть висновки.

19. Щоденна динаміка курсу акцій компанії Amazon на початку року характеризується даними:

№ з/п	Доларів	№ з/п	Доларів	№ з/п	Доларів
1	116,9	13	166,2	25	155,9
2	117,8	14	156,9	26	156,8
3	128,7	15	146,6	27	153,1
4	141,6	16	144,5	28	158,6
5	129,1	17	154,1	29	166,6
6	132,2	18	158,5	30	166,9
7	132,3	19	177,9	31	171,2
8	136,2	20	173,0	32	164,8
9	156,8	21	169,5	33	166,4
10	145,1	22	153,0	34	168,5
11	148,6	23	155,5	35	170,9
12	116,9	24	147,1	36	171,1

Здійсніть експоненційне згладжування часового ряду біржових цін з 7-и і 15-денним періодами згладжування і параметрами адаптації 0,15 і 0,08 відповідно. Чи можна розглядати лінії цих середніх як трендові лінії підтримки й опору? Чи надають лінії цих середніх упереджуваний сигнал про можливі зміни тренду? За результатами аналізу зробіть висновки.

20. Динаміка витрат країни на обслуговування державного боргу характеризується даними, млрд дол.

№ з/п	Млрд дол.	№ з/п	Млрд дол.
1	4,96	9	7,45
2	5,68	10	7,78
3	7,01	11	8,56
4	8,05	12	8,94
5	8,12	13	9,75
6	8,03	14	10,52
7	7,77	15	11,77
8	7,83	16	13,59

Опишіть динаміку витрат країни на обслуговування державного боргу авторегресійною функцією. Обґрунтуйте порядок авторегресії, зробіть висновки

21. Обсяги продажів Toyota Motor Corporation на автомобільному ринку США становили, тис.

Рік, № з/п	Обсяги продажів, тис	Рік, № з/п	Обсяги продажів, тис.
1	182,2	11	214,9
2	222,8	12	183,9
3	192,3	13	222,2
4	215,3	14	202,4
5	209,6	15	209,2
6	208,8	16	248,3
7	223,1	17	169,7
8	203,1	18	188,8
9	191,1	19	207,9
10	190,4	20	207,4

За наведеними даними:

- перевірте часовий ряд на стаціонарність на основі автокореляційних функцій (*ACF* та *PACF*), зробіть висновок;
- застосуйте до часового ряду модель *ARIMA*, обґрунтуйте структуру моделі;
- здійсніть аналіз залишків, зробіть висновок щодо адекватності моделі;
- визначте прогноз на період попередження $v = 6$.

22. Середні закупівельні ціни на зернові та зернобобові культури на внутрішньому ринку становили, грн за тону:

Місяць	1-й рік	2-й рік	3-й рік	4-й рік	5-й рік
1	3279	3768	4036	4355	3279
2	3538	3799	4152	4365	3538
3	3582	3867	4217	4344	3582
4	3598	3913	4313	4334	3598
5	3628	3911	4342	4290	3628
6	3605	3901	4352	4276	3605
7	3471	3870	4319	4198	3471
8	3383	3842	4401	4149	3383
9	3376	3828	4438	4087	3376
10	3381	3821	4394	3990	3381
11	3382	3780	4358	3914	3382
12	3414	3772	4315	3868	3414

За наведеними даними:

- на основі автокореляційних функцій (*ACF*) та (*PACF*) перевірте часовий ряд на стаціонарність і наявність сезонних коливань, зробіть висновок;
- застосуйте до часового ряду модель *ARIMA*, обґрунтуйте структуру моделі;
- здійсніть аналіз залишків, зробіть висновок щодо адекватності моделі;
- визначте прогноз на період попередження $v = 6$.

23. Помісячна динаміка проблемних кредитів у сукупних валових витратах банків характеризується даними, %.

№ з/п	Проблемні кредити, %	№ з/п	Проблемні кредити, %
1	58,3	15	8,9
2	57,1	16	10,7
3	58,2	17	13,2
4	59,9	18	12,3
5	60,1	19	11,6
6	59,9	20	12,4
7	58,3	21	12,8
8	50,1	22	13,2
9	48,1	23	13,5
10	7,2	24	13,4
11	9,1	25	13,4
12	8,6	26	13,6
13	7,9	27	13,5
14	9,8	28	13,8

Візуальний аналіз часового ряду свідчить про наявність інтервенції. Зобразіть характер динаміки проблемних кредитів графічно і на основі графіка оберіть тип інтервенції. Оцініть ефект впливу інтервенції на тенденцію проблемних кредитів, застосувавши до часового ряду модель *Interrupted ARIMA* модуля *Time Series / Forecasting*. Зробіть висновки.

24. Динаміка прямих іноземних інвестицій в економіку країни характеризується даними, млрд дол :

Порядковий номер року	Інвестиції, млрд дол.	Порядковий номер року	Інвестиції, млрд дол.
1	1,148	14	0,515
2	1,162	15	0,673
3	1,138	16	0,728
4	1,244	17	0,942
5	1,267	18	1,174
6	1,227	19	1,232
7	1,215	20	1,257
8	1,207	21	1,304
9	0,292	22	1,345
10	0,312	23	1,282
11	0,377	24	1,314
12	0,413	25	1,345
13	0,407	26	1,282

Зобразіть характер динаміки прямих іноземних інвестицій графічно, доведіть наявність і обґрунтуйте тип інтервенції. Оцініть ефект впливу інтервенції на тенденцію часового ряду, застосувавши модель *Interrupted ARIMA* модуля *Time Series / Forecasting*. Поясніть зміст параметрів моделі, зробіть висновки.

25. Поквартальна динаміка добування водних біоресурсів характеризується даними, тис. т:

Квартал, № з/п	Тис т	Квартал, № з/п	Тис т	Квартал, № з/п	Тис т
1	230,5	9	226,4	17	240,7
2	240,8	10	221,1	18	242,2
3	232,8	11	88,5	19	232,9
4	218,2	12	92,6	20	253,1
5	209,5	13	226,4	21	246,4
6	225,8	14	235,2	22	240,2
7	230,2	15	226,1	23	249,8
8	218,6	16	237,3	24	260,3

Зобразіть характер динаміки добування водних біоресурсів графічно, доведіть наявність і обґрунтуйте тип інтервенції. Оцініть ефект впливу інтервенції на тенденцію часового ряду, застосувавши модель *Interrupted ARIMA* модуля *Time Series / Forecasting*. Поясніть зміст параметрів моделі, зробіть висновки.

Рекомендована література

1. Бокс Дж., Дженкінс Г. Анализ временных рядов, прогноз и управление / пер. с англ. Москва. 1974. С. 39—132; 330—361.
2. Галушак М. П., Галушак О. Я., Кужда Т. І. Прогнозування соціально-економічних процесів: навч. посіб. Тернопіль. 2021. С. 95—103.
3. Дрейпер Н., Смит Г. Прикладной регрессионный анализ / пер. с англ. Москва. 1973. С. 95—266.
4. Єріна А.М., Єрін Д.Л. Статистичне моделювання та прогнозування: підручник. Київ, КНЕУ. 2014. С. 129—148.
5. Методи і моделі економічного прогнозування : навч. посіб. / Кічор В.П. та ін. Львів, 2019. С. 161—193.

- 7.1. Особливості вимірювання кореляції часових рядів
- 7.2. Нелінійна регресія
- 7.3. Регресія на панельних даних

7.1. Особливості вимірювання кореляції часових рядів

Усі явища навколишнього світу взаємопов'язані і взаємозумовлені. Для соціально-економічних явищ характерні переважно кореляційні зв'язки, які через складність взаємодії визначальних факторів і вплив випадкових причин проявляються не у кожному окремому випадкові, а лише в середньому. Кореляційний зв'язок аналітично описує функція (рівняння) регресії. Залежно від кількості факторів x_i , включених до рівняння регресії, вирізняють парну (однофакторну) $Y = f(x)$ і множинну (багатофакторну) регресію $Y = f(x_1, x_2 \dots x_m)$; за аналітичною формою функція регресії може бути лінійною або нелінійною.

Коли інформаційна база регресійної моделі представлена часовими рядами, виникають певні методологічні труднощі, спричинені залежністю рівнів ряду, їхньою автокореляцією. Наявність останньої порушує одну з передумов регресійного аналізу — незалежність спостережень — і призводить до викривлення результатів.

Автокореляція свідчить про наявність тренду. Тому одна з умов застосування методів регресійного аналізу для вимірювання зв'язків між часовими рядами — усунення трендів з обох рядів.

Для усунення тренду застосовують два підходи: різницеви перетворень і відхилень від тренду. Різницеві перетворення означають заміну первинних рівнів пов'язаних між собою часових рядів y_t, x_t , послідовними різницями (ланцюговими абсолютними приростами). Такий підхід усуває тенденцію у формі полінома k -го порядку: різниці першого порядку усувають лінійний тренд, другого порядку — параболічний тренд і т. д. Так, в разі лінійно-

го тренду первинні рівні рядів y_t , x_t замінюють першими різницями: $\Delta y = y_t - y_{t-1}$ та $\Delta x = x_t - x_{t-1}$. Однофакторна регресія набуває вигляду:

$$\Delta y = a + b\Delta x_1,$$

де b — коефіцієнт регресії; показує, як у середньому змінюється y зі зміною x_t на одиницю власної шкали вимірювання і за інших рівних умов;

a — вільний член рівняння регресії.

Наприклад, чітко виражену лінійну тенденцію до зростання виявляють наведені в табл. 7.1 взаємопов'язані часові ряди доходів населення x і заощаджень y (в млрд грн). Якщо до первинних даних застосувати МНК, дістанемо рівняння регресії $Y = -10,47 + 0,70x$. Коефіцієнт детермінації наближається до одиниці ($R^2 = 0,995$), що зумовлено наявністю тренду в часових рядах.

Таблиця 7.1

ДО РОЗРАХУНКУ РЕГРЕСІЇ ЗА ДАНИМИ ЧАСОВИХ РЯДІВ,
ЩО ВИЯВЛЯЮТЬ ЛІНІЙНУ ТЕНДЕНЦІЮ

Рік, № з/п	x	y	Перші різниці	
			Δy	Δx
1	24	7	—	—
2	29	10	5	3
3	32	12	3	2
4	36	14	4	2
5	40	17	4	3
6	45	21	5	4
7	48	23	3	2
8	51	25	3	2
9	55	29	4	4
$Y = f(x)$	$Y_t = -10,47 + 0,70x$ ($t_x = 36,49$)		$\Delta y = -0,231 + 0,77\Delta x$ ($t_{\Delta x} = 2,57$)	
R^2	0,995		0,524	
Критерій DW	1,15		1,40	

Значення критерія Дарбіна-Уотсона $DW = 1,15$ не потрапляє в зону допустимих значень гіпотези про відсутність автокореляції залишків $[1,32 \div 2,68]$, тож задля забезпечення адекватності моделі необхідно усунути трендову компоненту в обох часових рядах. Для усунення тренду застосуємо метод різницевих перетворень. Рівняння регресії на перших різницях набуває вигляду:

$$\Delta y = -0,231 + 0,77\Delta x,$$

тобто зі зростанням доходу на 1 млрд грн заощадження зростають у середньому на 0,77 млрд. Коефіцієнт детермінації становить 0,524. Значення критерія Дарбіна-Уотсона $DW = 1,40$ потрапляє в зону допустимих значень гіпотези про відсутність автокореляції залишків $[1,32 \div 2,68]$, що свідчить про адекватність регресійної моделі.

Коли тренд нелінійний, для його усунення застосовують інший спосіб — *відхилення від тренду*, суть якого — в заміні первинних рівнів рядів y_t , x_t відхиленнями $dy = y_t - f(t)$; $dx = x_t - f(t)$. При цьому тренди пов'язаних між собою часових рядів можуть бути різними. У лінійній регресії *на* відхиленнях від тренду $dy = a + bdx$ параметр b показує, як у середньому змінюється y зі зміною x на одиницю за інших рівних умов. Коли в обох часових рядах тренди лінійні, параметр $a = 0$, а параметр b стає коефіцієнтом пропорційності, показуючи, у скільки разів відхилення в ряду y більше (менше) відхилень у ряду x .

Усуненню автокореляції сприяє також включення до моделі змінної часу t , тобто $Y_t = f(x, t)$. Зміст параметрів моделі такого типу розглянемо на прикладі взаємозв'язку динаміки імпорту бавовняної нитки y_t і цін на цей товар на світовому ринку x_t . За даними табл. 7.2 обсяги імпорту бавовняної нитки (млн втулок) щомісяця зменшувалися, що зумовлено як зміною цін, так і зміною інших, неідентифікованих факторів, які формують тенденцію ряду. Саме вплив останніх акумулює змінна часу t . Зв'язок між цими показниками можна подати лінійною функцією

$$Y_t = a + bx + ct.$$

де b вимірює ефект впливу факторної ознаки x на результат y за відсутності тенденції;

c — середній приріст (зменшення) результату y під впливом неідентифікованих факторів, які рівномірно змінюються в часі, за незмінного рівня факторної ознаки x .

Таблиця 7.2

ПОМІСЯЧНА ДИНАМІКА ІМПОРТУ БАВОВНЯНОЇ НИТКИ

№ з/п	1	2	3	4	5	6	7	8	9	10	11
y_t	1749	1702	1769	1600	1431	1325	1302	1341	1232	1180	1162
x_t	33,7	34,4	47,3	62,1	73,2	85,9	82,7	76,8	80,2	80,6	84,1

У табл. 7.3 наведено параметри динамічної регресійної моделі імпорту бавовняної нитки. Модель має вигляд

$$Y_t = 1998,11 - 4,586x - 42,244t,$$

тобто підвищення світових цін на текстильне волокно через неврожай бавовни на один долар спричинило зменшення імпорту бавовняної нитки в середньому на 4,58 млн втулок. За рахунок інших факторів, передусім, зміни технології і задіяних внутрішніх ресурсів, імпорт бавовняної нитки щомісяця скорочувався в середньому на 42,24 млн втулок.

Про адекватність моделі свідчить сукупний коефіцієнт детермінації $R^2 = 0,951$ і значення t -критерія для кожного фактора.

Таблиця 7.3

ПАРАМЕТРИ РЕГРЕСІЙНОЇ МОДЕЛІ ІМПОРТУ БАВОВНЯНОЇ НИТКИ

N=11	Regression Summary for Dependent Variable: var3 import R=,975 R²=,951 Adjusted R²=,939 F(2,8)=78,276 p<,000 Std.Error of estimate: 56,649					
	Beta	Std.Err. of Beta	B	Std.Err. of B	t(8)	p-level
Intercept			1998,114	75,004	26,640	0,000
Var1 (x)	-0,3977	0,1575	-4,586	1,816	-2,525	0,036
Var2 (t)	-0,6097	0,1575	-42,244	10,915	-3,870	0,005

Для перевірки гіпотези про відсутність автокореляції в залишках застосуємо критерій Дарбіна-Уотсона. Фактичне значення $DW = 1,74$ потрапляє в зону допустимих значень гіпотези про відсутність автокореляції $[1,60 \div 2,40]$. Незначна серіальна кореляція (*Serial. Corr.* = 0,071) підтверджує незалежність спостережень і стійкість значень коефіцієнтів регресії.

Розглянуті приклади стосуються кореляції двох часових рядів. Принцип моделювання за даними трьох і більше рядів не змінюється. Загальний вигляд лінійної регресії на часових рядах має вигляд:

$$Y_t = a_0 + \sum_1^m b_i x_i + ct,$$

де b_i — чистий ефект впливу i -го фактора на y_t ;

c — ефект неідентифікованих факторів, які формують тренд часового ряду y_t .

У динамічній моделі можна відобразити не лише тренд, а й більш складні компоненти ряду, скажімо, періодичні чи сезонні коливання.

Розглянуті методи регресійного аналізу на часових рядах не завжди забезпечують адекватність моделей. Регресія на відхиленнях від тренду часом має низький коефіцієнт детермінації, регресія із введеним до моделі фактором часу t зводиться до трендової моделі (коефіцієнти регресії виявляються неістотними, а залишки корельованими).

Моделі з лаговими змінними

Характерною рисою механізму формування варіації та динаміки соціально-економічних показників є запізнення впливу факторів, коли причина і наслідок розірвані в часі. Наприклад, інвестиції в іригацію і введення в дію зрошувальних земель, витрати на R&D і продуктивність праці, доходи населення і попит на товари тривалого користування, збільшення грошової маси і рівень споживчих цін. Часові лаги зумовлені тривалістю виробничого циклу, інерційністю процесів, наявністю зворотного зв'язку тощо. Для оцінки ефекту впливу i -го фактора на результат y з певним запізненням p до моделі включають *лагові змінні* x_{ip} . Динамічні моделі з лаговими змінними, кожна з яких має певну кількість лагів, поділяють на два класи:

1) моделі з розподіленими лагами (*distributed lags*, *DL*-модель), в яких ознакова множина містить не лише поточні, а й лагові значення факторів; у таких моделях вплив фактора x розподіляється на кілька часових періодів:

$$Y_t = a_0 + b_0 x_t + b_1 x_{t-1} + b_2 x_{t-2} + \dots + b_k x_{t-k} + e_t.$$

Загальний вигляд *DL*-моделі з m факторами:

$$Y = a_0 + \sum_{i=1}^m \sum_{p=0}^k b_{i,p} x_{i,p},$$

де $b_{i,p}$ — коефіцієнт регресії, характеризує ефект впливу i -го фактора з певним запізненням (лаг $p = 0, 1, \dots, k$); m — кількість включених до моделі факторів;

2) змішані авторегресійні моделі (*autoregressive distributed, ADL-модель lags*), в яких ознакова множина містить і лагові змінні факторних ознак, і змінні авторегресії результативної ознаки y_{t-r} :

$$Y_t = a_0 + b_0x_t + b_1x_{t-1} + b_2x_{t-2} + c_1y_{t-1} + c_2y_{t-2} + e_t.$$

Загальний вигляд *ADL*-моделі з одним фактором:

$$Y_t = a_0 + \sum_{r=1}^q a_r y_{t-r} + \sum_{p=0}^k b_p x_{t-p} + e_t.$$

В обох моделях коефіцієнт b_0 показує вплив фактора x на змінну показника-функції у поточному періоді, його називають короткостроковим мультиплікатором. Сума коефіцієнтів регресійної моделі ($b_0 + b_1 + b_2 + \dots + b_k$) дістала назву довгострокового, загального мультиплікатора; будь-яка інша сума коефіцієнтів є проміжним мультиплікатором. Співвідношення коефіцієнтів при лагових змінних b_i і загального мультиплікатора $\sum b_i$ характеризують пропорції довгострокового впливу x на y за певний період. Нехай модель з розподіленими лагами має вигляд:

$$Y_t = \text{const} + 0,42x_t + 0,26x_{t-1} + 0,12x_{t-2}.$$

Довгостроковий мультиплікатор становить $0,42 + 0,26 + 0,12 = 0,8$. Поділивши коефіцієнти b_i на $0,8$, отримаємо частки впливу за кожний період: $0,525$, $0,325$ і $0,15$. Тобто $52,5\%$ впливу x на y відчувається одразу, $32,5\%$ — з лагом $p = 1$ і решта з $p = 2$.

Теоретично модель з розподіленими лагами можна узагальнити на будь-яку кількість лагових змінних (нескінченна модель), проте практична реалізація такої моделі наражається на непереборні труднощі, зумовлені обмеженістю динамічних рядів і складністю їхньої внутрішньої структури. Основною проблемою оцінювання моделі з розподіленими лагами є мультиколінеарність, оскільки суміжні рівні більшості часових рядів високорельовані.

Апріорі складно визначити, скільки лагових змінних необхідно включити до моделі. Переважно включаються такі лагові змінні, для яких лаги обґрунтовані теоретично і перевірені емпірично. Інструментом визначення лагів слугує крос-кореляційна функція, яка описує взаємні кореляції між рядами x_i та y , зсуненими відносно один одного на лаг p .

У табл. 7.4 наведено коефіцієнти крос-кореляції попиту на авто y і двох факторів: середньодушового доходу x_1 і цін x_2 . Як бачимо, зі збільшенням лагу крос-кореляційна функція згасає.

Таблиця 7.4

КОЕФІЦІЄНТИ КРОС-КОРЕЛЯЦІЇ ПОПИТУ
НА АВТО З ДОХОДАМИ І ЦІНАМИ

Лag	x_1	x_2
0	0,823	0,612
1	0,646	0,441
2	0,416	0,187
3	0,098	0,098

Для фактора x_1 істотними виявилися лаги $p = 0; 1; 2$; для фактора x_2 — лаги $p = 0; 1$. Модель набуває вигляду:

$$Y = a_0 + b_{10}x_1 + b_{11}x_{1,t-1} + b_{12}x_{1,t-2} + b_{20}x_2 + b_{21}x_{2,t-1} + ct,$$

де параметри $b_{1,p}$ і $b_{2,p}$ характеризують ефекти впливу факторів з відповідними лагами, параметр c — вплив неідентифікованих факторів (наявність товару на ринку, мода, смаки тощо). Коли коефіцієнт регресії $b_{i,p}$ істотний, значить i -й фактор впливає на y з певним запізненням (лагом p).

Модель з визначеною довжиною лагів зводиться до множинної регресії, а її параметри оцінюють традиційним МНК. Для оцінювання моделей з нескінченною довжиною лагів застосовують спеціальні методи, серед яких методи Алмона і Койка.

У системі Statistica аналіз розподілених лагів реалізований у модулі *Time Series / Forecasting* комплексом процедур *Distributed Lags Analysis*. У стартовому вікні цієї процедури необхідно вказати *Method* оцінювання коефіцієнтів регресії і довжину лагу. В табл. 7.5 наведені оцінки коефіцієнтів регресії, значення t -критерію і p -level для залежності середньомісячних обсягів продажу компанії від витрат на рекламу.

Таблиця 7.5

КОЕФІЦІЄНТИ РЕГРЕСІЙНОЇ МОДЕЛІ З РОЗПОДІЛЕНИМИ ЛАГАМИ

Lag	Polyn. Distr. Lags; Regression Coefficients)			
	Indep: advertising costs Dep: sales volumes Lag: 3 R= ,9994 R-square= ,9987 N: 15			
	Regressn Coeff.	Standard Error	t(11)	p
0	6,2439	1,3150	4,7483	0,0006
1	4,1961	1,3356	3,1419	0,0094
2	1,7422	1,2791	1,3621	0,2004
3	0,6095	1,2652	0,4817	0,6394

Як свідчать дані, зі зростанням величини лагу значення коефіцієнтів регресії убувають. За t -критерієм істотними з імовірністю 0,95 виявилися коефіцієнти регресії з лагом 0 і 1, тобто лівова частка ефекту збільшення витрат на рекламу проявляється одразу і дещо менша — протягом першого місяця, далі ефект зростання витрат на рекламу не істотний.

7.2. Нелінійна регресія

Вимірювання кореляції часових рядів часто здійснюють на базі відносних величин — індексів. Це пояснюється більшою їх сталістю в часі порівняно з абсолютними величинами. Окрім того, з'являється можливість виключити мультиколінеарність та автокореляцію залишків. Описуються такі взаємозв'язки степеневою функцією:

$$Y = Ax_1^{b_1} x_2^{b_2} x_3^{b_3} \dots x_m^{b_m},$$

де b_i — коефіцієнт еластичності, який показує, на скільки % у середньому зміниться y зі зміною x_i на 1 % за умови незмінності інших факторів. Тобто коефіцієнт еластичності — це відносний ефект впливу i -го фактора на y .

При формуванні інформаційної бази нелінійних моделей у вигляді степеневих функцій абсолютні значення всіх взаємопов'язаних показників, окрім фактора часу t , замінюються власними логарифмами. Отже, степенева функція лінійна в логарифмах, а тому її параметри можна визначати методом найменших квадратів. На практиці степеневі функції використовують в аналізі різних економічних процесів: виробничі функції, функції споживання, попиту на товари різних категорій та ін.

Серед степеневих функцій найвідомішою є виробнича функція Кобба–Дугласа, яка описує співвідношення між факторами і результатом виробництва на будь-якому рівні економічної діяльності (компанія, галузь, регіон, економіка в цілому):

$$Q = A K^\alpha L^\beta,$$

де Q — результат виробництва; K — основний капітал; L — витрати праці (або кількість працівників).

Параметри α і β — це коефіцієнти еластичності: α характеризує відносний приріст результату на одиницю приросту капіталу при $L = \text{const}$, а β — відносний приріст результату на одиницю приросту витрат праці при $K = \text{const}$. Капітал і витрати праці розглядаються як фактори екстенсивного розвитку (залучення нових

ресурсів). При трудомісткому виробництві $\alpha > \beta$, при капіталомісткому, навпаки, $\beta > \alpha$.

У виробничій функції закладено умову, за якою $(\alpha + \beta) = 1$, тобто результат зростає у такій же пропорції, як і фактори. Коли $(\alpha + \beta) < 1$, має місце спадна ефективність факторів, тобто одночасне зростання кожного з них на 1 % забезпечує зростання результату Q менше 1 %. У разі, коли $(\alpha + \beta) > 1$, відзначається зростаюча ефективність використання факторів. Проте техніко-технологічний прогрес і підвищення продуктивності праці не дають утвердитися постійній рівності $(\alpha + \beta) = 1$. Вплив на результат виробництва техніко-технологічних факторів ураховується в моделі змінною часу $e^{\lambda t}$. Модель набуває вигляду:

$$Q = A K^{\alpha} L^{\beta} e^{\lambda t},$$

де λ характеризує темп приросту функції під впливом неуречевлених факторів (впровадження у виробництво новітніх технологій і техніки, інтелектуалізація праці тощо). Цей варіант моделі називають виробничою функцією Тінбергена. Застосувавши до неї логарифмічне диференціювання, дістанемо модель, яка описує взаємозв'язок темпів приросту:

$$q = \alpha k + \beta l + 100\lambda,$$

де q, k, l — темпи приросту, результату, капіталу і витрат праці, відповідно.

На основі наведеного взаємозв'язку темпів приросту можна визначити внесок екстенсивних та інтенсивних факторів у розвиток процесів відтворення:

$$d_{\text{екс}} = (\alpha k + \beta l) / q, \quad d_{\text{інт}} = \lambda / q.$$

У табл. 7.6 наведено параметри макроекономічних функцій трьох індустріально розвинутих країн за період з 1950-го по 1977 р.: Q — валовий національний продукт; K — основний капітал; L — трудові ресурси.

Таблиця 7.6

ПАРАМЕТРИ ВИРОБНИЧИХ ФУНКЦІЙ

Країна	α	β	λ
США	0,447	0,553	0,0134
Великобританія	0,506	0,494	0,0053
Японія	0,397	0,603	0,0466

Джерело: [4, с. 41].

За значеннями параметрів функції можна зробити висновки про особливості економічного поступу у повоєнні роки кожної з цих країн. Так, досягнення науково-технологічного розвитку найінтенсивніше впроваджувалися в економіку Японії: параметр λ вищий порівняно зі США в 3,5 рази, порівняно з Великобританією — у 9 разів. Водночас японська економіка характеризується найнижчою капіталомісткістю ($\alpha = 0,397$) і відносно високим рівнем ефективності використання трудових ресурсів ($\beta = 0,603$). Для американської економіки характерна збалансованість співвідношення еластичностей по капіталу і труду. Середньорічні темпи приросту в США в повоєнні роки становили, %: валового національного продукту — 3,38, основного капіталу — 2,79, трудових ресурсів — 1,46. Звідси внесок факторів у формування динаміки ВНП:

$$\text{екстенсивних } d_{\text{екс}} = \frac{0,447 \cdot 2,79 + 0,553 \cdot 1,46}{3,38} = 0,625;$$

$$\text{інтенсивних } d_{\text{інт}} = \frac{1,34}{3,38} = 0,375.$$

На практиці використовують різні модифікації виробничої функції. Наприклад, розділивши обидві її частини на L , дістанемо функцію продуктивності праці:

$$W = AF^\alpha e^{\lambda t},$$

де W — продуктивність праці; F — капіталоозброєність праці.

На темпах приросту взаємозв'язок цих показників записується так:

$$w = \alpha f + \lambda = \alpha (k - l) + \lambda.$$

Внесок екстенсивних та інтенсивних факторів у динаміку продуктивності праці визначається аналогічно:

$$d_{\text{екс}} = \alpha (k - l) / w$$

$$d_{\text{інт}} = \lambda / w.$$

До темпів приросту застосовують класичну регресію. Як приклад розглянемо зв'язок між темпами приросту капіталоозброєності (Var1) і продуктивності праці (Var2) в одній з галузей промисловості (табл.7.7).

Таблиця 7.7

**ТЕМПИ ПРИРОСТУ КАПІТАЛООЗБРОЄНОСТІ
І ПРОДУКТИВНОСТІ ПРАЦІ**

Показник	Ланцюгові темпи приросту t – го року, %						
	$t=1$	$t=2$	$t=3$	$t=4$	$t=5$	$t=6$	$t=7$
Var1	1,4	2,5	1,7	0,9	1,0	2,6	2,8
Var2	2,6	2,7	2,8	1,5	1,8	2,4	3,2

Результати розрахунків за стандартною процедурою модуля *Multiple Regression* наведено в табл. 7.8. Параметри моделі становлять: $\alpha = 0,581$;

$\lambda = 1,358$, тобто приріст капіталоозброєності праці на 1 % спричиняє зростання продуктивності праці на 0,581 %. За рахунок інших факторів щорічний приріст продуктивності праці становить у середньому 1,358 %. Середньорічні темпи приросту капіталоозброєності праці — 1,8 %, продуктивності праці — 2,4 %. Отже, внесок екстенсивних факторів у динаміку продуктивності праці становить $(1,8 \cdot 0,581) : 2,4 = 0,436$; інтенсивних — $1,358 : 2,4 = 0,566$.

Таблиця 7.8

ПАРАМЕТРИ РЕГРЕСІЙНОЇ МОДЕЛІ НА ТЕМПАХ ПРИРОСТУ

Regression Summary for Dependent Variable: Var2 (---.sta) R=,7762 R ² =,6024 Adjusted R ² =,5229 F(1,5)=7,577 p<,04019 Std.Error of estimate:,408					
	BETA	B	St. Err. of B	t(5)	p-level
Intercept		1,358	0,418	3,245	0,0223
Var1	0,776	0,581	0,211	2,753	0,0402

Степеневою функцією описують також взаємозв'язок між попитом C , середньодушовим доходом населення D і цінами на товар P :

$$C = A D^{\alpha} P^{\beta} e^{\lambda t},$$

де α і β — коефіцієнти еластичності попиту від доходу та цін, відповідно.

Тренд попиту, зумовлений звичками, модою тощо, представлений у моделі змінною часу $e^{\lambda t}$.

7.3. Регресія на панельних даних

Через обмеженість часових рядів соціально-економічних явищ неможливо врахувати в моделі всі особливості розвитку процесу. Аби розширити інформаційну базу моделі, практикують об'єднання просторових і динамічних рядів. Можливі різні варіанти формування такої змішаної статично-динамічної інформації. Розглянемо найбільш поширений варіант регресії на панельних даних (*panel data model*).

Панельні дані є результатом спостереження об'єктів однієї сукупності в послідовні періоди часу. Об'єктами спостереження можуть бути індивідууми, домашні господарства, компанії, регіони, країни. Масиви панельних даних поєднують у собі як просторові вибірки, так і дані часових рядів, тобто для кожного моменту часу ми маємо просторову вибірку (*cross-sectional date*) і для кожного об'єкта з вибіркової сукупності — часовий ряд (*time-series date*). В соціології та психології такі дані називають лонгitudними.

Завдяки тому, що панельні дані містять одночасно *cross-sectional date* та *time-series date*, враховують індивідуальні відмінності між об'єктами і властиві їм тренди, виникає можливість формувати більш гнучкі моделі порівняно з класичними. Вирізняють моделі: наскрізної регресії, регресії з детермінованими індивідуальними ефектами і з випадковим індивідуальним ефектом.

В моделі наскрізної регресії припускають, що всі об'єкти демонструють однакову поведінку в усі моменти спостереження. Просторові й динамічні ряди об'єднуються в один масив інформації, одиницею якого є об'єкто-період. Для 10 об'єктів за 5 років маємо $10 \cdot 5 = 50$ об'єкто-періодів. Отже, об'єкто-періоди лише розширяють інформаційну базу моделі, параметри моделі не залежать ані від об'єктів, ані від моментів спостереження, їх визначають методом найменших квадратів. Проте часто об'єднання просторово-динамічних рядів, розширяючи інформаційну базу моделі, водночас наділяє її особливими властивостями, які необхідно врахувати в моделі.

Головна особливість статично-динамічної інформації — залежність спостережень. Залежними виявляються не лише рівні часових рядів, але й ряди в цілому (і просторові, і часові), оскільки належність рівня y_t до того чи іншого ряду фіксована. Так, залежність між часовими рядами — це результат просторової варіа-

ції, яка через інерційність процесів зберігається певний час. Залежність просторових рядів відображає синхронність динаміки показників за окремими об'єктами, зумовлену спільними умовами розвитку. Ігнорування цих особливостей інформаційної бази моделювання призводить до помилкових висновків.

Коли об'єкти різняться, скажімо, якістю управління, надійністю ділового партнерства тощо, але індивідуальні особливості стабільні протягом періоду спостереження, до такої інформаційної бази застосовують регресійну модель з детермінованими індивідуальними ефектами (*fixed effect model*). Ознакова множина такої моделі доповнюється *dummy*-змінними окремих об'єктів u_j , коефіцієнти регресії при цих змінних характеризують індивідуальні ефекти.

На панельних даних за тривалий період часу можна виявити тенденції розвитку як окремих об'єктів, так і сукупності в цілому. Властивий усім об'єктам тренд враховується в моделі змінною часу t . Проте через нерівномірність розвитку окремих об'єктів сукупності поряд зі спільним трендом істотними можуть виявитись індивідуальні тренди. Для їх фільтрації можна використати змінні динамічної взаємодії: для факторів — $x_i t$, для об'єктів — $u_j t$. З урахуванням зазначених особливостей регресійну модель для сукупності об'єкто-періодів можна записати так:

$$Y = a_0 + \sum_{i=1}^m b_i x_i + \sum_{i=1}^m c_i x_i t + \sum_{j=1}^{n-1} a_j u_j + \sum_{j=1}^{n-1} d_j u_j t + f t,$$

Параметри моделі вимірюють:

b_i — чистий, елімінований від взаємозв'язків у рамках моделі, ефект впливу фактора x_i ;

c_i — зміну ефектів впливу b_i у часі;

a_j — різницю між значеннями функції на j -му об'єкті порівняно і об'єктом, взятим за базу порівняння;

d_j — зміну цих відмінностей у часі;

f — спільний для всіх об'єктів тренд — вплив неідентифікованих у моделі факторів;

a_0 — вільний член рівняння. Для кожного j -го об'єкта вільний член рівняння дорівнює сумі ($a_0 + a_j$); на відміну від a_0 , сума має економічний зміст — вимірює вплив факторів, які визначають специфіку цього об'єкта.

Отже, регресійна модель на панельних даних містить дві групи параметрів. Одна група — це оцінки ефектів впливу факторів і зміну їх у часі, друга описує особливості сукупності, специфіку

розвитку окремих об'єктів. Уникнути перевантаження моделі і зберегти максимум інформації для оцінки параметрів можна, скориставшись алгоритмом покрокового регресійного аналізу.

Як приклад розглянемо параметри моделі продуктивності праці в агрогосподарствах (тис. грн на одного працівника), які спеціалізуються на вирощуванні винограду і фруктів та мають власні переробні цехи. Інформаційний масив сформовано за даними 18 господарств за 5 років. До ознакового простору моделі включені фактори: x_1 — економічна оцінка сільськогосподарських угідь, бали; x_2 — частка садів і виноградників у загальній площі сільськогосподарських угідь; x_3 — частка праці механізаторів у загальній кількості відпрацьованих людино-днів. Для оцінки тенденцій ефектів впливу кожного з цих факторів включені змінні динамічної взаємодії x_{it} . Два нетипові господарства представлені в моделі *dumtmy*-змінними u_j , а індивідуальні їх тренди — змінними динамічної взаємодії $u_j t$.

Параметри моделі визначалися за процедурами модуля *Multiply Regression*. Істотними виявилися ефекти впливу всіх факторів b_1, b_2, b_3 , параметр при змінній динамічної взаємодії другого фактора c_{2t} , параметри при структурних змінних обох господарств a_1, a_2 , параметр при змінній динамічної взаємодії другого господарства d_{2t} . Значення параметрів наведено у табл. 7.9.

Таблиця 7.9

ПАРАМЕТРИ РЕГРЕСІЙНОЇ МОДЕЛІ НА ПАНЕЛЬНИХ ДАНИХ

Параметр моделі	b_1	b_2	b_3	c_{2t}	a_1	a_2	d_{2t}
Значення параметра	39,86	15,63	20,46	1,17	-42,65	56,78	-3,52

Коефіцієнти регресії b_i інтерпретуються традиційно як чисті ефекти впливу включених до моделі факторів. При цьому, як свідчить параметр c_{2t} , ефект впливу спеціалізації (частки садів і виноградників у загальній площі сільськогосподарських угідь) на продуктивність праці щороку зростає в середньому на 1,17 тис. грн. Істотність параметрів a_1 та a_2 підтверджують нетиповість господарств, представлених у моделі *dumtmy*-змінними. За рахунок специфічних умов функціонування цих господарств рівень продуктивності праці на першому з них нижчий за середній на 42,65 тис. грн, на другому, навпаки, на 56,78 тис. грн вищий. Останній параметр має тенденцію до зменшення щороку в середньому на 3,52 тис. грн.

Прогнозні можливості регресійних моделей реалізуються процедурою *Predict dependent var.* модуля *Multiply Regression*.

Можливі інші варіанти застосування змішаної статично-динамічної інформації, зокрема, так звана *динамізація просторових моделей*. Суть цього методу полягає в тому, що для кожного t -го року визначається статична модель $Y_t = f(x_1, x_2, x_3)$. Коефіцієнти регресії статичних моделей утворюють динамічні ряди. Якщо ефект впливу i -го фактора змінюється в часі, то така зміна виявиться трендом ряду b_i . Методом екстраполяції тренду можна визначити очікуваний ефект впливу на період попередження v . Водночас визначається прогнозний рівень самого фактора $x_{i,t+v}$. Поєднання цих прогнозів дає прогноз функції Y_{t+v} :

$$Y_{t+v} = a_{0,t+v} + b_{1,t+v}x_{1,t+v} + b_{2,t+v}x_{2,t+v} + b_{3,t+v}x_{3,t+v}.$$

За відсутності тренду коефіцієнта регресії в прогнозній моделі використовують середнє його значення. У табл. 7.10 наведено фрагменти динамічних рядів параметрів регресійної моделі продуктивності праці в цементній промисловості, т на одного робітника. Фактори: x_1 — електроозброєність праці, кВт-год/люд.-год; x_2 — продуктивність цементних печей, т/год; x_3 — коефіцієнт використання календарного часу роботи цементних печей. Тривалість періоду — 5 років.

Таблиця 7.10

ЧАСОВІ РЯДИ КОЕФІЦІЄНТІВ РЕГРЕСІЇ

Рік	b_1	b_2	b_3
1	11,8	11,3	18,5
2	11,5	11,9	19,1
3	11,3	12,2	17,7
4	10,6	13,4	18,2
5	9,9	13,7	18,6

Як свідчать часові ряди коефіцієнтів регресії, в даній галузі промисловості відбувається перерозподіл ефектів впливу факторів на продуктивність праці: зменшується вплив електроозброєності праці x_1 , зростає вплив продуктивності устаткування x_2 і майже незмінним залишається вплив використання календарного часу устаткування x_3 .

Прогнозування ефектів впливу факторів та їхніх рівнів можна здійснити у будь-який спосіб, обґрунтувавши функціональний вид прогнозної моделі.

Питання і задачі для самопідготовки

1. У чому полягають особливості кореляції часових рядів?
2. Як привести нестационарний часовий ряд до стаціонарного?
3. Поясніть сутність методу різницевих перетворень часових рядів.
4. Розкрийте особливості і сферу використання регресійної моделі з урахуванням фактору часу tt .
5. Яку функцію в регресійному аналізі виконує змінна динамічної взаємодії xt ?
6. Попит на сталь за помісячними даними описується моделлю:

$$Y = 1,5 - 1,27 x_1 + 6,22 x_1 t + 4,65 x_2 - 0,03 t,$$

де y — млн т проданої сталі, x_1 — ціна сталі, центів за фунт, x_2 — індекс промислового виробництва. Поясніть зміст параметрів моделі.

7. Поясніть особливості регресійної моделі з розподіленими лагами (DL)?
8. Зазначте спільні риси і відмінності між моделями з розподіленими лагами (DL) і змішаною авторегресійною моделлю (ADL).
9. Розкрийте сутність коротко- і довгострокового мультиплікаторів DL-моделі. Визначте частку впливу фактора x на результат у поточному періоді на основі моделі $Y_t = \text{const} + 6,24x_t + 4,20x_{t-1} + 1,74x_{t-2} + 0,61x_{t-3}$.
10. Яку функцію в аналізі взаємопов'язаних часових рядів виконують коефіцієнти крос-кореляції?
11. Розкрийте сутність і окресліть сферу використання степеневих функцій.
12. Залежність інфляції (динаміки споживчих цін) від темпів нарощення грошової маси (агрегат М0) описується функцією: $Y = 0,8x^{0,265}$. Поясніть зміст параметрів моделі, визначте очікуваний рівень інфляції, якщо грошову масу збільшити на 10 %.
13. Поясніть зміст параметрів виробничих функцій Кобба-Дугласа і Тінбергена.
14. Як оцінити внесок екстенсивних та інтенсивних факторів у темпи економічного зростання?
15. Розкрийте особливості структури регресійної моделі на панельних даних. Наведіть приклади.
16. Динаміка урожайності і собівартості виробництва картоплі в регіоні характеризується даними:

Рік, № з/п	Собівартість грн /ц, y_t	Урожайність, ц з га, x_t	Рік, № з/п	Собівартість грн /ц, y_t	Урожайність ц з га, x_t
1	118	108	8	172	148
2	154	101	9	240	102
3	130	106	10	224	130
4	139	124	11	275	98
5	151	103	12	247	139
6	196	106	13	214	173
7	162	149	14	217	168

З використанням засобів MS Excel оцініть ефекти впливу урожайності картоплі на собівартість:

- на основі рівняння регресії з урахуванням фактора часу;
- застосовуючи спосіб різницевих перетворень часових рядів;
- застосовуючи спосіб відхилення від трендів.

На основі отриманих характеристик щільності та істотності зв'язку зробіть висновки щодо адекватності кожної моделі.

17. Вплив інаестицій (x) на динаміку основних засобів і необоротних активів (y) описує рівняння регресії

$$Y_t = 0,82 + 0,76_{t-1} + 1,54_{t-2} + 0,76_{t-3} + 0,44_{t-4} + 0,26_{t-5},$$

де t — роки.

Поясніть зміст параметрів моделі, визначте короткостроковий і довгостроковий мультиплікатори, зробіть висновки.

18. Динаміка оборонного бюджету країни за даними 2005–2021 рр. описується моделлю: $Y = -23,46 + 0,87 y_{t-1} + 0,0215 x_{t-1}$, де y — витрати на оборону, x — валовий національний продукт. Визначте прогнозний рівень витрат на оборону на 2022 рік, якщо у 2021 р. валовий національний продукт становив 2310 млрд дол., витрати на оборону — 83,5 млрд дол. Очікуваний обсяг валового національного продукту на 2022 рік — 2395 млрд дол.

19. Динаміка продуктивності праці y_t і витрат на виконання наукових досліджень і розробок x_t в агрохімічній компанії характеризується даними:

t	y_t , тис дол.	x_t , млн дол	t	y_t , тис дол.	x_t , млн дол
1	17,9	1,52	10	16,1	1,14
2	16,8	1,23	11	17,4	1,58
3	18,4	1,78	12	17,3	1,45
4	18,3	1,45	13	16,8	1,24
5	17,3	1,48	14	16,1	1,35
6	16,3	1,12	15	17,5	1,68
7	15,1	1,34	16	18,8	1,65
8	15,6	1,38	17	18,9	1,56
9	16,2	1,45	18	19,2	1,5

З використанням процедури *Distributed lags* модуля *Time Series / Forecasting* системи *Statistica*:

а) оцініть ефекти впливу витрат на виконання наукових досліджень і розробок на продуктивність праці з лагом 1, 2, 3; визначте ко-

ротко- і довгостроковий мультиплікатори, надайте їм економічну інтерпретацію;

б) здійсніть аналіз залишкових величин, зробіть висновок щодо адекватності моделі.

20. Розвиток промислового сектора регіону характеризується даними, Q — обсяг виробництва, млн дол.; K — середньорічна вартість основного капіталу, млн дол.; L — середньорічна чисельність зайнятих, тис. чол.:

Рік, № з/п	Q	K	L	Рік, № з/п	Q	K	L
1	70,1	44,4	33,3	8	106,1	58,7	56,7
2	78,6	47,3	38,5	9	109,2	61,5	57,9
3	81,7	48,2	40,4	10	112,8	63,8	60,5
4	88,7	49,8	45,8	11	117,1	64,2	63
5	90,8	52,2	46,2	12	119,1	65	65,1
6	96,4	54,3	50	13	123,2	65,4	66,6
7	99,7	54,4	52,1	14	124	66	67,2

Для вимірювання взаємозв'язку між результатом і факторами виробництва застосуйте степеневу функцію Кобба-Дугласа. За результатами аналізу:

- надайте економічну інтерпретацію параметрам функції;
- зробіть висновки щодо адекватності моделі;
- на основі степеневої функції визначте точковий прогноз обсягів виробництва на наступний рік за умови, що вартість основного капіталу зросте на 2,5 %, а чисельність зайнятих — на 2,0 %.

21. На основі даних зад. 20 сформулюйте інформаційну базу моделі продуктивності праці, оцініть еластичність капіталоозброєності на продуктивність праці, зробіть висновки. Доведіть взаємозв'язок між функціями виробництва і продуктивності праці.

22. За наведеними даними про темпи приросту та еластичність капіталоозброєності праці в різних галузях промисловості визначте темпи приросту продуктивності праці, а також вплив на динаміку продуктивності праці екстенсивних та інтенсивних факторів. Зробіть висновки.

Галузь промисловості	Еластичність капіталоозброєності праці	Темп приросту капіталоозброєності	λ , %
<i>A</i>	0,35	2,2	1,2
<i>B</i>	0,48	3,4	2,3
<i>C</i>	0,27	1,9	1,7

23. За наведеними в таблиці даними визначте темпи приросту ВВП кожної країни, оцініть вплив на динаміку ВВП екстенсивних та інтенсивних факторів. Зробіть висновки.

Країна	Витрати капіталу		Витрати праці		λ , %
	Темп приросту	Еластичність	Темп приросту	Еластичність	
<i>A</i>	2,5	0,46	2,2	0,53	1,4
<i>B</i>	4,3	0,35	1,7	0,65	2,6

24. За трирічними поквартальними даними 8 підприємств цементної промисловості побудовано модель продуктивності праці (виробництво цементу на одного робітника, т). Ознакова множина моделі: x_1 — коефіцієнт забезпеченості основними напівфабрикатами (клінкером) власного виробництва; x_2 — коефіцієнт використання календарного фонду робочого часу цементних печей; x_3 — погодинна продуктивність цементних млинів, т. Істотними виявилися параметри при структурних змінних: u_1 — цементно-шиферного комбінату; u_2 — цементно-гірничого комбінату. Коефіцієнт детермінації — 0,925. Параметри моделі становлять:

Ознака	Коефіцієнт регресії		
	факторної ознаки	структурної змінної	змінної динамічної взаємодії
x_1	-0,52	—	—
x_2	0,85	—	0,04
x_3	1,04	—	—
u_1	—	67,12	-4,32
u_2	—	-81,43	3,06

Поясніть зміст параметрів моделі. Зробіть специфікацію моделі для цементно-шиферного і цементно-гірничого комбінатів.

Рекомендована література

1. Єріна А.М. Єрін Д.Л. Статистичне моделювання та прогнозування: підручник. Київ. КНЕУ. 2014. С. 264—286
2. Лук'яненко І., Краснікова Л. Економетрика: підручник. Київ, 1998. С. 310 — 325.
3. Методи і моделі економічного прогнозування: навч. посібн. / Кічор В.П. та ін. Львів. 2019. С. 201—247.
4. Теория и практика статистического моделирования экономики / Под ред Е.М. Четыркина и А.Класа. Москва. 1986. 272 с.

ДОДАТКИ

ДОДАТОК I

КРИТИЧНІ ТОЧКИ РОЗПОДІЛУ СТЬЮДЕНТА ДЛЯ $\alpha = 0,05$
| T | - ДВОБІЧНИЙ КРИТЕРІЙ ; T — ОДНОБІЧНИЙ КРИТЕРІЙ

k	$ t $	t	k	$ t $	t
5	2,57	2,01	18	2,10	1,73
6	2,45	1,94	20	2,09	1,73
7	2,37	1,89	22	2,07	1,72
8	2,31	1,86	24	2,06	1,71
9	2,26	1,83	26	2,06	1,71
10	2,23	1,81	28	2,05	1,70
11	2,20	1,80	30	2,05	1,70
12	2,18	1,78	40	2,02	1,68
13	2,16	1,77	50	2,01	1,68
14	2,14	1,76	60	2,00	1,68
15	2,13	1,75	100	1,98	1,66
16	2,12	1,75	∞	1,96	1,64

ДОДАТОК II

КРИТИЧНІ ЗНАЧЕННЯ СТАТИСТИКИ ТЕСТУ ІРВІНА λ_n ДЛЯ $\alpha = 0,05$

n	2	3	10	20	30	50	100
λ_n	2.8	2,2	1.5	1.3	1.2	1.1	1,0

ДОДАТОК III

КРИТИЧНІ ЗНАЧЕННЯ СТАТИСТИКИ ТЕСТУ R/S ДЛЯ $\alpha = 0,05$

Кількість спостережень	Межі R/S- критерія	
	Нижня	Верхня
10	2,67	3,69
11	2,74	3,80
12	2,80	3,91
13	2,86	4,00
14	2,92	4,09
15	2,97	4,17
16	3,01	4,24
17	3,06	4,31
18	3,10	4,37
19	3,14	4,43
20	3,18	4,49
25	3,34	4,71
30	3,47	4,89

ДОДАТОК IV

КРИТИЧНІ ЗНАЧЕННЯ СТАТИСТИКИ
ТЕСТУ ФОСТЕРА — СТЮАРТА ДЛЯ $\alpha = 0,05$

n	μ	σ_1	σ_2
10	3,858	1,288	1,964
15	4,636	1,521	2,153
20	5,195	1,677	2,279
25	5,632	1,791	2,373
30	5,990	1,882	2,447
35	6,294	1,956	2,509
40	6,557	2,019	2,561
45	6,790	2,072	2,606
50	6,998	2,121	2,645
55	7,187	2,163	2,681
60	7,360	2,201	2,713
65	7,519	2,236	2,742
70	7,666	2,268	2,769
75	7,803	2,297	2,793
80	7,931	2,324	2,816
85	8,051	2,349	2,837
90	8,165	2,373	2,857
95	8,273	2,395	2,876
100	8,375	2,416	2,894

ДОДАТОК V

ЗНАЧЕННЯ $Z^*=T_{0,90}Z$ ДЛЯ ОЦІНКИ ДОВІРЧИХ МЕЖ ПРОГНОЗУ
ПРИ $\alpha = 0,10$ (ЛІНІЙНИЙ ТРЕНД)

n	Період попередження v			
	1	2	3	4
7	2,638	2,875	3,140	3,427
8	2,463	2,639	2,836	3,050
9	2,342	2,479	2,631	2,796
10	2,252	2,361	2,483	2,614
11	2,183	2,272	2,371	2,478
12	2,127	2,202	2,383	2,372
13	2,084	2,146	2,215	2,290
14	2,046	2,100	2,159	2,222
15	2,015	2,062	2,113	2,168
16	1,988	2,029	2,073	2,121
17	1,965	2,002	2,040	2,083
18	1,945	1,977	2,012	2,049
19	1,928	1,957	1,988	2,021
20	1,912	1,937	1,965	1,995

ДОДАТОК VI

**КРИТИЧНІ ЗНАЧЕННЯ
ЦИКЛІЧНОГО КОЕФІЦІЄНТА АВТОКОРЕЛЯЦІЇ
($\alpha=0,05$)**

<i>n</i>	Додатні значення	Від'ємні значення
5	0,253	-0,753
6	0,345	-0,708
7	0,370	-0,674
8	0,371	-0,625
9	0,366	-0,593
10	0,360	-0,564
11	0,353	-0,539
12	0,348	-0,516
13	0,341	-0,497
14	0,335	-0,479
15	0,328	-0,462
20	0,299	-0,399
25	0,276	-0,356
30	0,257	-0,356
35	0,242	-0,300
40	0,229	-0,279
50	0,208	-0,248
60	0,191	-0,225

ДОДАТОК VII

КРИТИЧНІ ЗНАЧЕННЯ КРИТЕРІЮ ДАРБІНА-УОТСОНА

n	$m = 1$		$m = 2$		$m = 3$		$m = 4$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
6	0,61	1,40						
7	0,70	1,36	0,47	1,90				
8	0,76	1,33	0,56	1,78	0,37	2,29		
9	0,82	1,32	0,63	1,70	0,46	2,13	0,30	2,59
10	0,88	1,32	0,70	1,64	0,53	2,02	0,38	2,41
11	0,93	1,32	0,76	1,60	0,60	1,93	0,44	2,28
12	0,97	1,33	0,81	1,58	0,66	1,86	0,51	2,18
13	1,01	1,34	0,86	1,56	0,72	1,82	0,57	2,09
14	1,05	1,35	0,91	1,55	0,77	1,78	0,63	2,03
15	1,08	1,36	0,95	1,54	0,81	1,75	0,69	1,98
16	1,11	1,37	0,98	1,54	0,86	1,73	0,73	1,94
18	1,16	1,39	1,05	1,54	0,93	1,70	0,82	1,87
20	1,20	1,41	1,10	1,54	1,00	1,68	0,89	1,83
22	1,24	1,43	1,15	1,54	1,05	1,66	0,96	1,80
25	1,29	1,45	1,21	1,55	1,12	1,65	1,04	1,77
28	1,33	1,48	1,26	1,56	1,18	1,65	1,10	1,75
32	1,37	1,50	1,31	1,57	1,24	1,65	1,18	1,73
35	1,40	1,52	1,34	1,58	1,28	1,65	1,22	1,73
40	1,44	1,54	1,39	1,60	1,34	1,66	1,29	1,72
45	1,48	1,57	1,43	1,62	1,38	1,67	1,34	1,72
50	1,50	1,59	1,46	1,63	1,42	1,67	1,38	1,72
60	1,55	1,62	1,51	1,65	1,48	1,69	1,44	1,73
70	1,58	1,64	1,55	1,67	1,53	1,70	1,49	1,74
80	1,61	1,66	1,59	1,69	1,56	1,72	1,53	1,74
90	1,64	1,68	1,61	1,70	1,59	1,73	1,57	1,75
100	1,65	1,69	1,63	1,72	1,61	1,74	1,59	1,76
150	1,72	1,75	1,71	1,76	1,69	1,77	1,68	1,79
200	1,76	1,78	1,75	1,79	1,74	1,80	1,73	1,81

ДОДАТОК VIII

ЗНАЧЕННЯ СИНУСІВ І КОСИНУСІВ ДЛЯ 1-4 ГАРМОНІК

Місяць	Радіанна міра	Гармоніка							
		Перша		Друга		Третя		Четверта	
		sin	cos	sin	cos	sin	cos	sin	cos
1	$\pi/6$	0,5	0,866	0,866	0,5	1	0	0,866	-0,5
2	$\pi/3$	0,866	0,5	0,866	-0,5	0	-1	-0,866	-0,5
3	$\pi/2$	1	0	0	-1	-1	0	0	1
4	$2\pi/3$	0,866	-0,5	-0,866	-0,5	0	1	0,866	-0,5
5	$5\pi/6$	0,5	-0,866	-0,866	0,5	1	0	-0,866	-0,5
6	π	0	-1	0	1	0	-1	0	1
7	$7\pi/6$	-0,5	-0,866	0,866	0,5	-1	0	0,866	-0,5
8	$4\pi/3$	-0,866	-0,5	0,866	-0,5	0	1	-0,866	-0,5
9	$3\pi/2$	-1	0	0	-1	1	0	0	1
10	$5\pi/3$	-0,866	0,5	-0,866	-0,5	0	-1	0,866	-0,5
11	$11\pi/6$	-0,5	0,866	-0,866	0,5	-1	0	-0,866	-0,5
12	0	0	1	0	1	0	1	0	1

ДОДАТОК ІХ

КРИТИЧНІ ЗНАЧЕННЯ КОЕФІЦІЄНТА ДЕТЕРМІНАЦІЇ R^2
 ДЛЯ РІВНЯ ІСТОТНОСТІ $\alpha = 0,05$ І ЧИСЛА СТУПЕНІВ СВОБОДИ

$$k_1 = m - 1; k_2 = n - m$$

k_2	k_1				
	1	2	3	4	5
5	0,569	699	764	806	835
6	500	632	704	751	785
7	444	575	651	702	739
8	399	527	604	657	697
9	362	488	563	618	659
10	332	451	527	582	624
12	283	394	466	521	564
14	247	348	417	471	514
16	219	312	378	429	477
18	197	283	345	394	435
20	179	259	318	364	404
24	151	221	273	316	353
28	130	193	240	279	314
32	115	171	214	250	282
36	102	153	192	226	256
40	093	139	176	207	234
50	075	113	143	170	194
60	063	095	121	144	165
80	047	072	093	110	127
100	038	058	075	090	103
120	032	049	063	075	087
200	019	030	038	046	053

ДОДАТОК X

ТАБЛИЦЯ СЕРЕДНІХ ТЕМПІВ ЗРОСТАННЯ, %

Середньорічний темп зростання	Темпи зростання			
	за 2 роки	за 3 роки	за 4 роки	за 5 роки
100,1	100,2	100,3	100,4	100,5
100,2	100,4	100,6	100,8	101,0
100,3	100,6	100,9	101,21	101,51
100,4	100,8	101,2	101,61	102,02
100,5	101,0	101,51	102,02	102,53
100,6	101,2	101,81	102,42	103,04
100,7	101,4	102,11	102,83	103,55
100,8	101,61	102,42	103,24	104,06
100,9	101,81	102,72	103,65	104,58
101,0	102,01	103,03	104,06	105,1
101,1	102,21	103,34	104,47	105,62
101,2	102,41	103,64	104,89	106,15
101,3	102,62	103,95	105,3	106,67
101,4	102,82	104,26	105,72	107,2
101,5	103,02	104,57	106,14	107,73
101,6	103,23	104,88	106,56	108,26
101,7	103,43	105,19	106,98	108,79
101,8	103,63	105,5	107,4	109,33
101,9	103,84	105,81	107,82	109,87
102,0	104,04	106,12	108,24	110,41
102,1	104,24	106,43	108,67	110,95
102,2	104,45	106,75	109,09	111,49
102,3	104,65	107,06	109,52	112,04
102,4	104,86	107,37	109,95	112,59
102,5	105,06	107,69	110,38	113,14
102,6	105,27	108,0	110,81	113,69

Продовження дод. X

Середньорічний темп зростання	Темпи зростання			
	за 2 роки	за 3 роки	за 4 роки	за 5 роки
102,7	105,47	108,32	111,25	114,25
102,8	105,68	108,64	111,68	114,81
102,9	105,88	108,95	112,11	115,37
103,0	106,09	109,27	112,55	115,93
103,1	106,3	109,59	112,99	116,49
103,2	106,5	109,91	113,43	117,06
103,3	106,71	110,23	113,87	117,63
103,4	106,92	110,55	114,31	118,2
103,5	107,12	110,87	114,75	118,77
103,6	107,33	111,19	115,2	119,34
103,7	107,54	111,52	115,64	119,92
103,8	107,74	111,84	116,09	120,5
103,9	107,95	112,16	116,54	121,08
104,0	108,16	112,49	116,99	121,67
104,1	108,37	112,81	117,44	122,25
104,2	108,58	113,14	117,89	122,84
104,3	108,78	113,46	118,34	123,43
104,4	108,99	113,79	118,8	124,02
104,5	109,2	114,12	119,25	124,62
104,6	109,41	114,44	119,71	125,22
104,7	109,62	114,77	120,17	125,82
104,8	109,83	115,1	120,63	126,42
104,9	110,04	115,43	121,09	127,02
105,0	110,25	115,76	121,55	127,63
105,1	110,46	116,09	122,01	128,24
105,2	110,67	116,43	122,48	128,85
105,3	110,88	116,76	122,95	129,46

Продовження дод. X

Середньорічний темп зростання	Темпи зростання			
	за 2 роки	за 3 роки	за 4 роки	за 5 роки
105,4	111,09	117,09	123,41	130,08
105,5	111,3	117,42	123,88	130,7
105,6	111,51	117,76	124,35	131,42
105,7	111,72	118,09	124,82	131,94
105,8	111,94	118,43	125,3	132,56
105,9	112,15	118,76	125,77	133,19
106,0	112,36	119,1	126,25	133,82
106,1	112,57	119,44	126,72	134,45
106,2	112,78	119,78	127,2	135,09
106,3	113,0	120,12	127,68	135,73
106,4	113,21	120,46	128,16	136,37
106,5	113,42	120,79	128,65	137,01
106,6	113,64	121,14	129,13	137,65
106,7	113,85	121,48	129,62	138,3
106,8	114,06	121,82	130,1	138,95
106,9	114,28	122,16	130,59	139,6
107,0	114,49	122,5	131,08	140,26
107,1	114,7	122,85	131,57	140,91
107,2	114,92	123,19	132,06	141,57
107,3	115,13	123,54	132,56	142,23
107,4	115,35	123,88	133,05	142,9
107,5	115,56	124,23	133,55	143,56
107,6	115,78	124,58	134,04	144,23
107,7	115,99	124,92	134,54	144,9
107,8	116,21	125,27	135,04	145,58
107,9	116,42	125,62	135,55	146,25
108,0	116,64	125,97	136,05	146,93

Закінчення дод. X

Середньорічний темп зростання	Темпи зростання			
	за 2 роки	за 3 роки	за 4 роки	за 5 роки
108,1	116,86	126,32	136,55	147,61
108,2	117,07	126,67	137,06	148,3
108,3	117,29	127,02	137,57	148,98
108,4	117,51	127,38	138,08	149,67
108,5	117,72	127,73	138,59	150,37
108,6	117,94	128,08	139,1	151,06
108,7	118,16	128,44	139,61	151,76
108,8	118,37	128,79	140,12	152,46
108,9	118,59	129,15	140,64	153,16
109,0	118,81	129,5	141,16	153,86
109,1	119,03	129,86	141,68	154,57
109,2	119,25	130,22	142,2	155,28
109,3	119,46	130,58	142,72	155,99
109,4	119,68	130,93	143,24	156,71
109,5	119,9	131,29	143,77	157,42
109,6	120,12	131,65	144,29	158,14
109,7	120,34	132,01	144,82	158,87
109,8	120,56	132,38	145,35	159,59
109,9	120,78	132,74	145,88	160,32
110,0	121,0	133,1	146,41	161,05

Навчальне видання

АНТОНІНА ЄРІНА
ОЛЬГА МАЗУРЕНКО

СТАТИСТИЧНИЙ АНАЛІЗ ЧАСОВИХ РЯДІВ

НАВЧАЛЬНИЙ ПОСІБНИК

Видано в авторській редакції
Художник обкладинки *Т. Зябліцева*
Верстка *О. Федосенко*

Підп. до друку 02.09.22. Формат 60×84/16.
Папір офсет. № 1. Гарнітура Тип Таймс. Друк офсет.
Ум.-друк. арк. 7,74. Обл.-вид. арк. 10,84