

ДЕРЖАВНИЙ ВИЩИЙ НАВЧАЛЬНИЙ ЗАКЛАД
«ЗАПОРІЗЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ»
МІНІСТЕРСТВА ОСВІТИ І НАУКИ, МОЛОДІ ТА СПОРТУ УКРАЇНИ

В.В. Леонт'єва, Н.О. Кондрат'єва

МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ ВИРОБНИЧИХ ПРОЦЕСІВ

Навчально-методичний посібник
для студентів денного відділення математичного факультету
напряму підготовки «Прикладна математика»
освітньо-кваліфікаційного рівня «Бакалавр»

Затверджено
вченою радою ЗНУ
Протокол №1 від 27.09.2011 р.

Запоріжжя
2011

УДК: 519.866 : 330.4 (075.8)

ББК: У212.311 я73

Л 478

Леонт'єва В.В., Кондрат'єва Н.О. Математичне моделювання виробничих процесів: навчально-методичний посібник для студентів денного відділення математичного факультету напряму підготовки «Прикладна математика» освітньо-кваліфікаційного рівня «Бакалавр». – Запоріжжя: ЗНУ, 2011. – 120 с.

У навчально-методичному посібнику викладено методологічні, методичні матеріали та інструментарій, які використовуються при вивченні курсу «Математичне моделювання виробничих процесів». Використання посібника у навчальному процесі дозволяє здобути фундаментальні теоретичні знання та сформувані практичні навички оволодіння методами аналізу, моделювання та прогнозування виробничих процесів, а також використовувати адекватні математичні моделі у сфері виробництва та підприємництва.

Навчально-методичний посібник містить теоретичний матеріал, методичні рекомендації та приклади розв'язку типових задач за кожною темою, корисні при виконанні лабораторних робіт. Також до посібника включені варіанти індивідуальних завдань до лабораторних робіт, список рекомендованої літератури за курсом, словник корисних термінів та додатки. Крім того, за кожною темою наводиться перелік питань до самоконтролю, який дозволяє студенту перевірити рівень отриманих при вивченні теми знань. Зміст посібника відповідає навчальним програмам з курсу «Математичне моделювання виробничих процесів».

Рекомендується для студентів математичного факультету денного відділення галузі знань «Фізико-математичні науки», напряму підготовки «Прикладна математика».

Рецензент *д.т.н., професор Гоменюк С.І.*

Відповідальний за випуск *д.т.н., професор Грищак В.З.*

ЗМІСТ

ВСТУП.....	6
1. ТИПОВА НАВЧАЛЬНА ПРОГРАМА ДИСЦИПЛІНИ «МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ ВИРОБНИЧИХ ПРОЦЕСІВ»	9
1.1 Зміст курсу	9
1.2 Мета та завдання навчального курсу	10
2. ТЕОРЕТИЧНА ЧАСТИНА. МЕТОДИЧНІ ПОРАДИ ТА ІЛЮСТРАТИВНИЙ МАТЕРІАЛ (ЗАДАЧІ, ПРИКЛАДИ ТА ЇХ РОЗВ'ЯЗАННЯ) ДО ВИВЧЕННЯ ТЕМ КУРСУ	11
2.1 Часові ряди та їх попередній аналіз.....	11
2.1.1 Основні поняття та визначення	11
2.1.2 Попередній аналіз часових рядів. Коригування рівнів часового ряду	12
2.1.3 Попередній аналіз часових рядів. Виявлення тенденції у часовому ряді.....	14
2.1.3.1 Критерій серій	15
2.1.3.2 Критерій «висхідних і спадних» серій.....	16
2.1.3.3 Критерій, заснований на порівнянні середніх рівня ряду. 18	
2.1.3.4 Метод Форстера-Стюарта	21
2.1.3.5 Критерій квадратів послідовних різниць (критерій Аббе)24	
2.1.4 Питання для самоконтролю	26
2.1.5 Рекомендована література.....	26
2.2 Методи екстраполяції в прогнозуванні виробничих процесів	26
2.2.1 Сутність екстраполяції	26
2.2.2 Класифікація методів екстраполяції	27
2.2.3 Прості методи екстраполяції.....	27
2.2.3.1 Екстраполяція середнього	28
2.2.3.2 Екстраполяція на основі аналітичних показників рядів динаміки	30
2.2.3.3 Екстраполяція тренду	37
2.2.3.4 Екстраполяція на основі індексу сезонності	43
2.2.4 Питання для самоконтролю	46
2.2.5 Рекомендована література.....	46
2.3 Адаптивні методи прогнозування.....	46
2.3.1 Сутність адаптивних методів прогнозування	46
2.3.2 Метод ковзної середньої та його застосування	48
2.3.2.1 Сутність методу.....	48
2.3.2.2 Застосування методу ковзної середньої до побудови прогнозу за допомогою адитивної та мультиплікативної моделей часового ряду.....	49

2.3.3	Метод експоненціального згладжування.....	61
2.3.3.1	Сутність методу.....	61
2.3.3.2	Застосування методу експоненціального згладжування до побудови прогнозів	68
2.3.4	Метод гармонійних ваг	74
2.3.4.1	Сутність методу.....	74
2.3.4.2	Застосування методу гармонійних ваг до побудови прогнозів	77
2.3.5	Питання для самоконтролю	80
2.3.6	Рекомендована література.....	81
2.4	Методи аналізу якості побудованих моделей та прогнозів виробничих процесів	81
2.4.1	Методи аналізу якості побудованих моделей.....	81
2.4.1.1	Перевірка рівності нулю математичного сподівання рівнів ряду залишків	81
2.4.1.2	Перевірка умови випадковості виникнення окремих відхилень від тренду	82
2.4.1.3	Наявність (відсутність) автокореляції у відхиленнях від моделі зросту	82
2.4.1.4	Відповідність ряду залишків нормальному закону розподілу	84
2.4.2	Методи аналізу якості побудованих прогнозів.....	89
2.4.2.1	Абсолютні показники	89
2.4.2.2	Порівняльні показники	91
2.4.2.3	Якісні показники	93
2.4.3	Питання для самоконтролю	95
2.4.4	Рекомендована література.....	95
3.	ПРАКТИЧНА ЧАСТИНА. МЕТОДИЧНІ ПОРАДИ ДО ВИКОНАННЯ ЛАБОРАТОРНИХ РОБІТ ЗА КУРСОМ.....	96
3.1	Значення лабораторної роботи при вивченні курсу «Математичне моделювання виробничих процесів»	96
3.2	Порядок визначення варіанту лабораторної роботи.....	96
3.3	Рекомендації до виконання та вимоги до захисту лабораторних робіт..	96
3.4	Критерії оцінювання знань студентів.....	97
3.5	Завдання до лабораторних робіт за темами дисципліни «Математичне моделювання виробничих процесів»	97
3.5.1	Лабораторна робота 1. Попередній аналіз часових рядів	97
3.5.1.1	Методичні поради до вивчення теми	97
3.5.1.2	Завдання на лабораторну роботу	98
3.5.2	Лабораторна робота 2. Екстраполяція в прогнозуванні виробничих процесів	99
3.5.2.1	Методичні поради до вивчення теми	99
3.5.2.2	Завдання на лабораторну роботу	99

3.5.3	Лабораторна робота 3. Побудова прогнозу за допомогою адитивної та мультиплікативної моделей часового ряду	103
3.5.3.1	Методичні поради до вивчення теми	103
3.5.3.2	Завдання на лабораторну роботу	103
3.5.4	Лабораторна робота 4. Прогнозування за допомогою методу експоненціального згладжування.....	104
3.5.4.1	Методичні поради до вивчення теми	104
3.5.4.2	Завдання на лабораторну роботу	105
3.5.5	Лабораторна робота 5. Побудова прогнозу за допомогою методу гармонійних ваг	105
3.5.5.1	Методичні поради до вивчення теми	105
3.5.5.2	Завдання на лабораторну роботу	106
3.5.6	Лабораторна робота 6. Аналіз якості побудованих моделей та прогнозів	106
3.5.6.1	Методичні поради до вивчення теми	106
3.5.6.2	Завдання на лабораторну роботу	107
	СЛОВНИК КОРИСНИХ ТЕРМІНІВ.....	108
	РЕКОМЕНДОВАНА ЛІТЕРАТУРА	111
	Додаток А Статистико-математичні таблиці	112
	Додаток Б Таблиця логарифмів	115
	Додаток В Метод найменших квадратів (МНК).....	116

ВСТУП

Виробництво і наука в Україні зараз переживають складні часи, тому проблема відновлення виробництва та збільшення ролі науки в його розвитку стає визначною.

Оскільки відновлення і розвиток виробництва повинні ґрунтуватися на освоєнні нових методів керівництва та впровадженні новітніх технологій й устаткування, які дозволяють забезпечити випуск конкурентно-здатної продукції за якісними та економічними показниками, то розробники передових технологій й устаткування повинні знати та широко використовувати ефективні методи наукових досліджень, що надасть можливість скоротити тривалість та вартість рішення поставлених задач. До таких ефективних методів слід віднести математизацію досліджень, яка (разом з тими областями знань, до яких належать ці дослідження) припускає в першу чергу побудову математичної моделі досліджуваного процесу, яка достатньо точно й адекватно його описує. Головною перевагою такого способу дослідження є те, що за наявності математичної моделі для отримання рішення поставлених конкретних задач виникає можливість подальше дослідження процесу замінити аналізом його моделі. При цьому розрізняють так звані теоретичні (евристичні) моделі, які описують на мікро- та макрорівнях механізм фізико-хімічних та механічних процесів, що відбуваються в об'єкті, та дозволяють здійснити широке прогнозування результату процесу для умов, ще не реалізованих на практиці, а також моделі емпіричні (статистичні), які є результатом обробки експериментальних даних, отриманих на досліджуваному об'єкті (чи його фізичній моделі) в певному діапазоні зміни умов процесу та найчастіше будуються у вигляді полінома того або іншого степеня. Потрібно зауважити, що адекватність статистичної моделі досліджуваному об'єкту доводиться лише кількісно за експериментальними даними. Такі моделі широко застосовують на практиці, коли вирішуються питання вдосконалення і оптимізації конкретного виробництва або об'єкту, проте сфера їх застосування обмежена вивченим діапазоном зміни умов та допускає дуже обережну екстраполяцію, зазвичай з експериментальною перевіркою отриманих рішень. Отримати ці моделі можна за допомогою, наприклад, кореляційно-регресійного аналізу, методів планування експерименту, що дозволяють отримати ефективну математичну модель при малих витратах праці і часу.

Як відомо, сучасне господарське життя засноване на величезній множині різноманітних виробничих процесів, що відіграють важливу роль в сучасній економіці. Оскільки під **виробничим процесом** розуміється процес створення доданої вартості шляхом цілеспрямованого перетворення одного набору товарів (послуг) в інший набір товарів (послуг), тобто процес, який полягає в тому, що за допомогою комбінування або перетворення одних продуктів (послуг) створюються нові, то цілком визначено можна сказати, що виробничі процеси протікають в системах, орієнтованих як на виробництво товарів, так і на обслуговування (виробництво послуг). Буквально усі види господарської

діяльності пов'язані з процесами виробництва. Так, наприклад, до класу виробничих процесів відносять такі види господарської діяльності, як нафтовидобуток, біржова торгівля, транспортування вантажів, навчання студентів у вищих навчальних закладах, надання аудиторських послуг і т. ін.

Різноманіття вироблюваних продуктів, існуючих технологій та методів виробничої організації й керування обумовлюють величезну різноманітність можливих виробничих систем і складних соціально-технічних процесів, які протікають в них. При цьому математичне моделювання виробничих процесів дозволяє виділити найважливіші чинники виробництва та оптимально їх розподілити при існуючих обмеженнях на виробничу діяльність, вказати шляхи зниження виробничих витрат, порівняти різні сценарії зростання та запропонувати найкращі стратегії розвитку виробничої системи.

При дослідженні багатьох природних та виробничих процесів виникає задача аналізу в динаміці подій та їх послідовностей, які не піддаються класичним математичним методам аналізу. Річ у тому, що багато чинників таких процесів можна розглядати як випадкові, а такого роду об'єктами займаються спеціальні області математики – теорія ймовірності, математична статистика, та, зокрема, теорія випадкових процесів і часових рядів. Основними завданнями в таких дослідженнях є детальне вивчення цих процесів, виділення їх істотних характеристик, яке може призвести до можливості прогнозування розвитку цих процесів в майбутньому. Також представляє інтерес виділення внутрішніх закономірностей, яким підпорядкований розвиток цих процесів. Для такого роду досліджень розроблено немало різних методів, проте не усі з них дійсно ефективні, а застосування інших вимагає досить глибокої теоретичної підготовки.

Отже, із математичним моделюванням виробничих процесів тісно пов'язане прогнозування основних виробничих показників. При цьому, як відомо, поняття «прогноз», «прогнозування» стосуються майбутніх, тобто невідомих подій. Інтерес до майбутнього випливає з практичної потреби сьогодення. Існують явища, майбутнє яких ми не знаємо, але вони мають важливе значення для рішень, прийнятих нами в сьогоденні. Отже, виникає необхідність прогнозувати ці явища.

Прогнозування сьогодні входить практично в усі галузі науки, господарської і політичної діяльності. Але основним є соціально-економічне й виробниче прогнозування, оскільки розвиток суспільства, економіки та виробництва визначає темпи росту інших елементів суспільної системи. Разом з тим виробництво є однією з найбільш важливих і складних сфер для прогнозу. Робота підприємств в умовах ринкових відносин все більше потребує в науково-обґрунтованому прогнозуванні їхньої діяльності. Пов'язано це насамперед з тим, що процес вироблення й ухвалення рішення містить у собі визначення безлічі альтернативних рішень, їхню оцінку і вибір єдиного, оптимального рішення. І результати прогнозування при цьому мають сенс як «матеріалу для міркування», на основі якого можна істотно підвищити наукову обґрунтованість рішень, які приймаються.

У зв'язку із викладеним вище, вивчення дисципліни «Математичне моделювання виробничих процесів» спрямовано на оволодіння студентами необхідним інструментарієм моделювання та прогнозування виробничих процесів, оскільки вміння правильно вибирати і застосовувати методи прогнозування безсумнівно впливає на якість рішень, прийнятих у виробничо-господарській діяльності.

Метою вивчення даного курсу є здобуття фундаментальних теоретичних знань і формування практичних навичок оволодіння методами моделювання та прогнозування виробничих процесів, а також вивчення і набуття досвіду застосування сучасних комп'ютерних технологічних засобів у процесі розробки прогнозів. При цьому в центрі уваги перебувають методи і моделі аналізу тенденцій та причинно-наслідкових зв'язків, які є необхідною умовою прогнозування окремих показників виробництва і комплексного економічного розвитку країни.

Навчальними планами для студентів математичного факультету денного відділення (галузь знань «Фізико-математичні науки», напрям підготовки «Прикладна математика») передбачено виконання лабораторних робіт з дисципліни «Математичне моделювання виробничих процесів» за розділами: «Аналіз динаміки виробничих процесів. Прості методи прогнозування виробничих процесів», «Адаптивні методи прогнозування виробничих процесів. Оцінка побудованих моделей та прогнозів». Лабораторна робота є важливою формою вивчення курсу «Математичне моделювання виробничих процесів» і виступає засобом набуття й поглиблення знань та практичних навичок студентів. Виконання лабораторних робіт забезпечує студентам систематичне вивчення курсу на основі опрацювання рекомендованих підручників і навчальних посібників, самостійний пошук необхідних літературних джерел та матеріалів, їх аналіз та узагальнення. Метою виконання лабораторних робіт є вивчення та закріплення навчального матеріалу з тем «Часові ряди та їх попередній аналіз», «Методи екстраполяції в прогнозуванні виробничих процесів», «Адаптивні методи прогнозування», «Методи аналізу якості побудованих моделей та прогнозів виробничих процесів», отриманого на лекціях, лабораторних заняттях та при самостійному вивченні матеріалу.

У даному навчально-методичному посібнику наводяться теоретичні відомості з дисципліни «Математичне моделювання виробничих процесів», розглядаються приклади їх практичного застосування шляхом розв'язання типових задач за кожною темою, надаються методичні поради, рекомендації до виконання та варіанти лабораторних робіт, які повинні виконати студенти в процесі вивчення даної дисципліни. Крім того, даний посібник містить словник корисних термінів, список рекомендованої літератури (основної та додаткової) за курсом «Математичне моделювання виробничих процесів» та додатки, використання яких значно спрощує сприйняття теоретичної інформації та вирішення конкретних задач практики.

1. ТИПОВА НАВЧАЛЬНА ПРОГРАМА ДИСЦИПЛІНИ «МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ ВИРОБНИЧИХ ПРОЦЕСІВ»

1.1 Зміст курсу

Модуль I. Аналіз динаміки виробничих процесів. Прості методи прогнозування виробничих процесів.

Тема 1. Часові ряди та їх попередній аналіз.

Надаються вихідні поняття, сутність, характерні риси та види часових рядів. Дається поняття про декомпозицію часового ряду та основні компоненти часових рядів. Наводиться характеристика аномальних спостережень та умов їх виникнення, розкриваються методи виявлення та усунення аномальних спостережень у часових рядах. Розглядаються основні методи перевірки гіпотези про існування тренду у часових рядах. Надається поняття про ідентифікацію моделі часового ряду та про методи аналізу динаміки часових рядів.

Тема 2. Методи екстраполяції в прогнозуванні виробничих процесів.

Описуються основні випадки, у яких можуть застосовуватися методи екстраполяції. Наводяться сутність та основні відмінності екстраполяції та інтерполяції. Надаються основні методи простої та складної прогновної екстраполяції. Розглядається прогнозування за одномірним динамічним рядом методами екстраполяції середнього та тренду. Наводяться основні аналітичні показники динаміки часового ряду та розглядається їх застосування у прогнозуванні. Розкривається метод екстраполяції часових рядів на основі індексу сезонності.

Модуль II. Адаптивні методи прогнозування виробничих процесів. Оцінка побудованих моделей та прогнозів.

Тема 3. Адаптивні методи прогнозування.

Розглядається різниця між адаптивними методами прогнозування та методами простої прогновної екстраполяції. Розкривається сутність адаптивних методів. Надаються наступні адаптивні методи прогнозування: метод ковзної середньої, метод експоненціального згладжування часових рядів, метод гармонійних ваг. Розглядаються основні умови застосування адаптивних методів прогнозування.

Тема 4. Методи аналізу якості побудованих моделей та прогнозів виробничих процесів.

Наводяться основні методи аналізу якості побудованих моделей прогнозування: перевірка рівності нулю математичного сподівання рівнів ряду залишків; перевірка умови випадковості виникнення окремих відхилень від тренду; перевірка наявності (відсутності) автокореляції у відхиленнях від моделі зросту; перевірка відповідності ряду залишків нормальному закону розподілу. Надаються абсолютні, порівняльні та якісні показники точності прогнозів. Розглядаються сутність та принципи побудови діаграми «прогноз-реалізація».

1.2 Мета та завдання навчального курсу

Метою вивчення курсу «Математичне моделювання виробничих процесів» є здобуття фундаментальних теоретичних знань і формування практичних навичок оволодіння методами моделювання та прогнозування виробничих процесів, а також вивчення і набуття досвіду застосування сучасних комп'ютерних технологічних засобів у процесі розробки прогнозів. При цьому в центрі уваги перебувають методи і моделі аналізу тенденцій та причинно-наслідкових зв'язків, які є необхідною умовою прогнозування окремих показників виробництва і комплексного економічного розвитку країни.

Для досягнення поставленої мети необхідно вирішити *наступні завдання*:

1. Дати систематичні знання стосовно постановки та розв'язання задач математичного моделювання виробничих процесів.
2. Ознайомити з основними етапами та методами аналізу впорядкованих у часі послідовностей спостережень основних показників виробничих процесів.
3. Розкрити основні методи прогнозування виробничих процесів.
4. Розглянути основні критерії оцінки адекватності, якості й точності математичних моделей виробничих процесів та отриманих за ними прогнозів.
5. Висвітлити особливості практичних аспектів математичного моделювання виробничих процесів.

У результаті вивчення дисципліни студент повинен:

– *знати*:

1. Предмет та об'єкт математичного моделювання виробничих процесів.
2. Методи виявлення та усунення аномальних спостережень у впорядкованих у часі послідовностях спостережень основних показників виробничих процесів.
3. Основні критерії та методи виявлення тенденції у часових рядах, що описують динаміку основних показників виробничих процесів.
4. Основні методи прогнозування виробничих процесів.
5. Основні критерії оцінки адекватності, якості й точності математичних моделей та отриманих за ними розв'язків.

– *вміти*:

1. Формулювати математичну постановку задач математичного моделювання та прогнозування виробничих процесів.
2. Обирати методи стосовно задачі, що розв'язується.
3. Розв'язувати задачі математичного моделювання виробничих процесів.
4. Визначати найімовірніший напрямок розвитку об'єкту моделювання та прогнозування на основі побудованої математичної моделі.
5. Оцінювати адекватність, якість й точність математичної моделі.
6. Проводити аналіз отриманих розв'язків та робити відповідні висновки.

2. ТЕОРЕТИЧНА ЧАСТИНА. МЕТОДИЧНІ ПОРАДИ ТА ІЛЮСТРАТИВНИЙ МАТЕРІАЛ (ЗАДАЧІ, ПРИКЛАДИ ТА ЇХ РОЗВ'ЯЗАННЯ) ДО ВИВЧЕННЯ ТЕМ КУРСУ

2.1 Часові ряди та їх попередній аналіз

2.1.1 Основні поняття та визначення

При аналізі багатьох виробничих показників найчастіше використовують щорічні, щоквартальні, щомісячні, щоденні дані. Наприклад, це можуть бути річні дані за ВВП, ВВП, обсягом чистого експорту, інфляції і т.д., місячні дані за обсягом продажу продукції, щоденні обсяги випуску будь-якої фірми. Для раціонального аналізу необхідно систематизувати моменти одержання відповідних статистичних даних.

У цьому випадку слід упорядкувати дані за часом їх одержання й побудувати так звані **часові ряди**, які є інформаційною базою для аналізу виробничих, соціальних, економічних та ін. процесів та під якими розуміються ряди динаміки, впорядковані за часом, або сукупності спостережень вимірюваної величини в різні моменти часу.

Сукупність спостережень $y(t_1), y(t_2), \dots, y(t_n)$ аналізованої величини $Y(t)$, здійснених у послідовні моменти часу t_1, t_2, \dots, t_n , називається **часовим рядом**.

Окремі спостереження часового ряду називають **рівнями ряду**.

Під **довжиною** часового ряду розуміють час, що минув від першого до останнього моменту спостереження. Часто довжиною ряду називають кількість рівнів n , які утворюють часовий ряд.

Для аналізу часового ряду $y(t_1), y(t_2), \dots, y(t_n)$ порядок у послідовності t_1, t_2, \dots, t_n є суттєвим, тобто час виступає одним із визначальних чинників. Це відрізняє часовий ряд від звичайної випадкової вибірки. Принциповою відмінністю часового ряду від простих статистичних сукупностей є наступні:

- по-перше, рівні часового ряду не є незалежними, тобто, якщо майбутні значення змінної можна визначити, то вони є функцією від минулих значень цієї змінної;

- по-друге, рівні часового ряду неоднаково розподілені. Закон розподілу ймовірностей цих випадкових величин і, зокрема, їхні математичні сподівання та дисперсії, можуть залежати від часу.

Отже, не можна поширювати властивості та правила статистичного аналізу випадкових вибірових спостережень на часові ряди. Порушення умови незалежності між спостереженнями призводить до негативних наслідків застосування цих методів.

В подальшому для зручності часовий ряд будемо визначати як послідовність $y_1, y_2, \dots, y_t, \dots, y_n$, де y_t – значення рівня ряду у момент часу t , ($t = \overline{1, n}$).

Перш ніж будувати прогноз за заданим часовим рядом потрібно провести попередній аналіз часового ряду, який полягає у виявленні й усуненні аномальних значень рівнів ряду, а також у визначенні наявності тенденції у вихідному часовому ряді.

2.1.2 Попередній аналіз часових рядів. Коригування рівнів часового ряду

На даному етапі попереднього аналізу часового ряду проводиться перевірка усіх рівнів ряду на нормальність, виявлення та усунення аномальних значень рівнів ряду.

Аномальним рівнем називається окреме значення рівня часового ряду, яке не відповідає потенційним можливостям досліджуваної економічної системи і яке, залишаючись у якості рівня ряду, впливає на значення основних характеристик часового ряду, у тому числі й на відповідну трендову модель.

Формально аномальність рівнів ряду виявляється як несподіваний стрибок (або спад) із подальшим поступовим встановленням попереднього рівня. Аномальність призводить до зміщення оцінок i , отже, до спотворення результатів аналізу. Отже, перш ніж проводити подальший аналіз часового ряду необхідно виявити та, якщо це є можливим, усунути аномальні значення ряду.

Для виявлення аномальних рівнів часових рядів використовуються методи, розраховані для статистичних сукупностей, одним з яких є **метод Ірвіна**, який ґрунтується на порівнянні сусідніх значень ряду та розрахунку характеристики λ_t й припускає використання наступної формули:

$$\lambda_t = \frac{|y_t - y_{t-1}|}{\sigma_y}, \quad t = 2, 3, \dots, n, \quad (2.1.1)$$

де y_t, y_{t-1} – значення рівня ряду відповідно у момент часу t та $t-1$;

$$\sigma_y = \sqrt{\frac{\sum_{t=1}^n (y_t - \bar{y})^2}{n}} - \text{середньоквадратичне відхилення}; \quad \bar{y} = \frac{1}{n} \cdot \sum_{t=1}^n y_t - \text{середній}$$

рівень ряду; n – число рівнів ряду.

Розрахункові значення $\lambda_2, \lambda_3, \dots, \lambda_n$ порівнюються з табличними значеннями критерію Ірвіна λ_α , наведеними у табл. 2.1.1:

- якщо $\lambda_t < \lambda_\alpha$, то значення рівня ряду y_t вважається нормальним;
- якщо $\lambda_t \geq \lambda_\alpha$, то значення рівня ряду y_t вважається аномальним.

Таблиця 2.1.1

Значення критерію Ірвіна λ_α при рівні значущості $\alpha = 0,05$

n	2	3	10	20	30	50	100
λ_α	2,8	2,3	1,5	1,3	1,2	1,1	1,0

Після виявлення аномальних рівнів ряду обов'язковим є визначення причин їх виникнення. Якщо точно встановлене, що вони викликані помилками першого роду, то вони усуваються декількома способами:

- заміною аномальних рівнів простою середньою арифметичною двох сусідніх з ним рівнів ряду;
- заміною аномальних рівнів відповідними значеннями за кривою, яка апроксимує даний часовий ряд;
- коригуванням аномальних рівнів за наступною схемою:

а) розраховується нове значення рівня ряду:

$$y_t^{(1)} = \lambda_\alpha \sigma_y + y_{t-1}, \quad (t = \overline{1, n}),$$

де $y_t^{(1)}$ – нове (скориговане) значення рівня ряду, у якому верхній індекс вказує номер ітерації коригування;

б) виявлене аномальне значення y_t у вихідному ряді замінюється на нове значення $y_t^{(1)}$;

в) визначаються характеристики $\bar{y}^{(1)}$ і $\sigma_y^{(1)}$ нового ряду;

г) проводиться перевірка рівнів отриманого ряду на аномальність:

1) якщо скориговане значення $y_t^{(1)}$ в отриманому ряді буде нормальним, то воно займає місце відповідного y_t в вихідному ряді та процес коригування вважається завершеним;

2) якщо ж в отриманому ряді будуть виявлені аномальні значення рівнів, то рекомендується повторити процедуру коригування виявлених аномальних рівнів часового ряду, тобто розрахувати нові значення $y_t^{(2)} = \lambda_\alpha \sigma_y^{(1)} + y_{t-1}^{(1)}$ й перевірити отриманий ряд на аномальність рівнів.

Отже, процес коригування носить ітераційний характер. Ітераційний процес завершується, коли в часовому ряді не буде жодного аномального спостереження.

Наведемо приклад виявлення та усунення аномальних спостережень у заданому часовому ряді.

Приклад 2.1.1. Нехай заданий часовий ряд:

$$y_t = 1,0; 1,1; 1,3; 3,0; 1,4 .$$

Визначимо основні характеристики ряду: $\bar{y} = 1,56$, $\sigma_y = 0,7338$.

Обчислимо λ -статистики згідно (2.1.1):

$$\lambda_2 = \frac{|y_2 - y_1|}{\sigma_y} = \frac{|1,1 - 1,0|}{0,7338} = \frac{0,1}{0,7338} = 0,136 < \lambda_\alpha = 2 ;$$

$$\lambda_3 = \frac{|y_3 - y_2|}{\sigma_y} = \frac{|1,3 - 1,1|}{0,7338} = \frac{0,2}{0,7338} = 0,2725 < \lambda_\alpha = 2 ;$$

$$\left. \begin{aligned} \lambda_4 &= \frac{|y_4 - y_3|}{\sigma_y} = \frac{|3,0 - 1,3|}{0,7338} = \frac{1,7}{0,7338} = 2,317 \geq \lambda_\alpha = 2; \\ \lambda_5 &= \frac{|y_5 - y_4|}{\sigma_y} = \frac{|1,4 - 3,0|}{0,7338} = \frac{1,6}{0,7338} = 2,180 \geq \lambda_\alpha = 2. \end{aligned} \right\}$$

Нове значення ряду:

$$y_t^{(1)} = \lambda_\alpha \sigma_y + y_{t-1} = 2 \cdot 0,7338 + 1,3 = 2,7676.$$

Отже, новий часовий ряд має вигляд

$$y_i = 1,0; 1,1; 1,3; 2,7676; 1,4.$$

Знову обчислюємо \bar{y} , σ_y , λ_t для отриманого ряду та перевіряємо усі його рівні на аномальність. Якщо $\lambda_t \geq 2$, то процес повторюється доти, доки рівень ряду не стане нормальним.

2.1.3 Попередній аналіз часових рядів. Виявлення тенденції у часовому ряді

Наступним кроком попереднього аналізу часового ряду є перевірка наявності тренду у вихідному часовому ряді (без аномальних спостережень).

Під **трендом** (від англ. trend — напрям, тенденція) розуміють тривалу тенденцію розвитку часового ряду, тобто певний напрямок розвитку, тривалу еволюцію (тенденцію до росту, стабільності або до зниження рівнів явища).

В окремих випадках присутність тренду у часовому ряді можна помітити, проаналізувавши графік часового ряду. Як правило, для економічних даних дуже типовим є повільне зростання чи падіння протягом тривалого періоду часу. Наявність **тренду** у виробничих часових рядах можна пояснити демографічними, технологічними змінами, а також змінами в структурі виробництва, попиту, тощо. Дія таких факторів є постійною, тому дослідники мають змогу описувати такі зміни за допомогою апроксимуючих кривих.

В загальному випадку, коли не є можливим візуальне визначення присутності тренду у часовому ряді, для виявлення факту наявності або відсутності тренду, тобто для перевірки гіпотези про його існування $H_0: E y(t) = a = const$, використовують **наступні критерії**:

- 1) критерій серій;
- 2) критерій «висхідних і спадних» серій;
- 3) критерій, заснований на порівнянні середніх рівня ряду;
- 4) метод Форстера-Стюарта;
- 5) критерій квадратів послідовних різниць (критерій Аббе).

Розглянемо кожний з них докладно.

Нехай є часовий ряд $y_1, y_2, \dots, y_t, \dots, y_n$. Прийmemo наступні гіпотези:
 H_0 : тренда немає; H_1 : тренд є.

2.1.3.1 Критерій серій

Розглянемо етапи застосування критерію:

1) Упорядкуємо члени ряду за зростанням:

$$y_1 < y_2 < \dots < y_t < \dots < y_n.$$

2) Визначимо медіану ряду:

$$y_{med} = \begin{cases} \frac{y_{n+1}}{2}, & \text{якщо } n - \text{непарне,} \\ \frac{1}{2} \left(\frac{y_n}{2} + \frac{y_{n+1}}{2} \right), & \text{якщо } n - \text{парне.} \end{cases} \quad (2.1.2)$$

3) Утворюємо послідовність «плюсів» і «мінусів», відповідних до вихідного ряду, не впорядкованому за зростанням, за правилом:

$$\begin{cases} \text{якщо } y_t > y_{med}, & \text{то } y_t \text{ відповідає "+"}, \\ \text{якщо } y_t < y_{med}, & \text{то } y_t \text{ відповідає "-"} \end{cases} \quad (2.1.3)$$

4) Підраховуємо загальне число серій ν і довжину найдовшої серії τ .

Під **серією** розуміється послідовність «плюсів» («мінусів»), які ідуть підряд.

Наприклад,

$$\underbrace{-, -}_{\text{серія}}, \underbrace{+, +}_{\text{серія}}, -, \underbrace{+, +, +, +}_{\text{серія}}, - \Rightarrow \nu = 3, \text{ і } \tau = 4.$$

5) Перевіримо виконання наступних нерівностей:

$$\nu > \left[\frac{1}{2} \left(n + 2 - 1,96 \sqrt{n-1} \right) \right], \quad (2.1.4)$$

$$\tau < [1,43 \ln(n+1)]. \quad (2.1.5)$$

Квадратні дужки в правих частинах нерівностей (2.1.4), (2.1.5) означають цілу частину числа. Нагадаємо, що **ціла частина числа** $A - [A]$ – це ціле число, яке є найближчим до A та не перевищує його.

Перевірка гіпотези H_0 ґрунтується на тому, що при умові випадковості ряду (за відсутністю тренду) довжина найдовшої серії τ не повинна бути надто великою, а загальне число серій ν – надто малим. Тому для того, щоб не була відкинута гіпотеза H_0 про випадковість вихідного ряду (про відсутність тренду) повинні виконуватися нерівності (для 5% рівня значущості) (2.1.4), (2.1.5).

Якщо хоча б одна з зазначених нерівностей (2.1.4), (2.1.5) виявиться порушеною, то гіпотеза H_0 (гіпотеза про те, що тренду немає) відкидається з імовірністю помилки $0,05 \leq \alpha \leq 0,0975$ (тобто це говорить про те, що тренд є).

Розглянемо на прикладі перевірку гіпотези H_0 за допомогою критерію серій.

Приклад 2.1.2. Нехай є часовий ряд:

$$y_t = 2, 4, 7, 5, 3, 9, 8, 6, 1, 5.$$

Упорядкуємо члени даного ряду за зростанням:

$$y_t = 1, 2, 3, 4, 5, 5, 6, 7, 8, 9.$$

Визначимо медіану ряду (2.1.2). Оскільки у нашому випадку $n = 10$, тобто воно є парним, то:

$$y_{med} = \frac{1}{2} \left(y_{\frac{n}{2}} + y_{\frac{n}{2}+1} \right) = \frac{1}{2} (y_5 + y_6) = \frac{1}{2} (5 + 5) = 5.$$

Утворюємо послідовність «плюсів» і «мінусів», відповідних до вихідного ряду, за правилом (2.1.3). Для нашого прикладу одержуємо наступну послідовність «плюсів» і «мінусів» (з вихідного ряду, невпорядкованого за зростанням):

$$-, -, +, +, -, +, +, +, -, +.$$

Знайдемо загальне число серій ν та довжину найдовшої серії τ . Для цього розглянемо отриману нами послідовність «плюсів» і «мінусів»:

$$\underbrace{-, -}_{\text{серія}}, \underbrace{+, +}_{\text{серія}}, -, \underbrace{+, +, +}_{\text{серія}}, -, +.$$

Маємо: $\nu = 3$, і $\tau = 3$.

Перевіримо виконання нерівностей (2.1.4), (2.1.5):

$$\left\{ \begin{array}{l} 3 > \left[\frac{1}{2} (10 + 2 - 1,96 \sqrt{10-1}) \right] = [6,12] = 6, \Rightarrow \begin{array}{l} 3 < 6 \\ \text{(нерівність} \\ \text{не виконується)} \end{array} \\ 3 < [1,43 \ln(10+1)] = [3,43] = 3. \Rightarrow \begin{array}{l} 3 \leq 3 \\ \text{(нерівність} \\ \text{виконується)} \end{array} \end{array} \right.$$

Оскільки не виконується перша з нерівностей, то гіпотеза H_0 відкидається, а це означає, що в заданому ряді є тренд.

Висновок: тренд є.

2.1.3.2 Критерій «висхідних і спадних» серій

У даному критерії ми не впорядковуємо наш вихідний ряд за зростанням, а розглядаємо його таким, яким він заданий, тобто розглядаємо ряд

$$y_1, y_2, \dots, y_t, \dots, y_n.$$

Послідовність застосування критерію:

1) утворення послідовності «плюсів» і «мінусів», відповідних до вихідного ряду, за правилом:

2.1.3.3 Критерій, заснований на порівнянні середніх рівня ряду

Сутність цього критерію полягає в наступному:

1) Вихідний часовий ряд розбивають на дві приблизно однакові за кількістю рівнів частини:

$$y_t = \underbrace{y_1, y_2, \dots, y_k}_{n_1}, \underbrace{y_{k+1}, \dots, y_{n-1}, y_n}_{n_2}$$

$$y^1 = y_1, y_2, \dots, y_{k-1}, y_k \quad \text{і} \quad y^2 = y_{k+1}, y_{k+2}, \dots, y_{n-1}, y_n$$

При цьому кожна частина розглядається як самостійна вибірка сукупність, що має нормальний розподіл.

2) Знаходимо середні рівні ряду:

$$\bar{y}_1 = \frac{1}{n_1} \cdot \sum_{i=1}^k y_i; \quad \bar{y}_2 = \frac{1}{n_2} \cdot \sum_{i=k+1}^n y_i.$$

Якщо часовий ряд має тенденцію (тренд), то середні, обчислені для кожної сукупності, \bar{y}_1 і \bar{y}_2 повинні суттєво різнитися між собою, тобто $\bar{y}_1 \neq \bar{y}_2$.

Якщо ж розбіжність буде незначущою (несуттєвою), тобто $\bar{y}_1 \approx \bar{y}_2$, то часовий ряд не має тенденції (тренду).

Таким чином, перевірка наявності тренду в досліджуваному ряді зводиться до перевірки гіпотези про рівність середніх двох нормально розподілених сукупностей.

3) Обчислюємо дисперсії:

$$S_1^2 = \frac{1}{n_1} \cdot \sum_{i=1}^k (y_i - \bar{y}_1)^2; \quad S_2^2 = \frac{1}{n_2} \cdot \sum_{i=k+1}^n (y_i - \bar{y}_2)^2$$

Якщо часовий ряд має тренд, то дисперсії $S_1^2 \neq S_2^2$ (суттєво різняться між собою).

Якщо ж $S_1^2 \approx S_2^2$ (розбіжність між ними є незначущою), то часовий ряд не має тенденції (тренду).

4) Перевіряємо гіпотезу про рівність дисперсій при рівні значущості $\alpha = 0,05$: $H_0: \sigma_1^2 = \sigma_2^2$. Альтернативною гіпотезою в цьому випадку є $H_1: \sigma_1^2 \neq \sigma_2^2$.

Для цього обчислюємо F – критерій Фішера за формулою

$$F_{розр} = \begin{cases} \frac{S_1^2}{S_2^2}, & \text{якщо } S_1^2 > S_2^2 \\ \frac{S_2^2}{S_1^2}, & \text{якщо } S_1^2 < S_2^2 \end{cases} \quad (2.1.9)$$

та за таблицею А.1 (Додаток А) знаходимо критичне значення критерію Фішера $F_{кр} = (\alpha; n_2 - 1, n_1 - 1)$ при рівні значущості $\alpha = 0,05$.

Порівняймо розрахункове значення F – критерію $F_{розр}$ з табличним $F_{кр}$:

- якщо $F_{розр} < F_{кр}$, то приймається гіпотеза H_0 : $\sigma_1^2 = \sigma_2^2$;
- якщо $F_{розр} \geq F_{кр}$, то приймається гіпотеза H_1 : $\sigma_1^2 \neq \sigma_2^2$.

5) Перевіряється основна гіпотеза:

$$H_0: \bar{y}_1 = \bar{y}_2; \quad H_1: \bar{y}_1 \neq \bar{y}_2.$$

Обчислюємо розрахункове значення критерію Стьюдента у вигляді

$$T_{розр} = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}} \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}} \quad (2.1.10)$$

та за таблицею А.3 (Додаток А) знаходимо критичне значення критерію Стьюдента $t_{кр} = (\alpha; n - 2)$ при рівні значущості $\alpha = 0,05$.

Порівняймо розрахункове значення критерію Стьюдента $T_{розр}$ з табличним $t_{кр}$:

- якщо $|T_{розр}| < t_{кр}$, то приймається гіпотеза H_0 : $\bar{y}_1 = \bar{y}_2$;
- якщо $|T_{розр}| \geq t_{кр}$, то приймається гіпотеза H_1 : $\bar{y}_1 \neq \bar{y}_2$.

6) Формулюємо висновок:

$$\left. \begin{array}{l} \text{– якщо} \\ \bar{y}_1 \neq \bar{y}_2 \\ S_1^2 \neq S_2^2 \quad (\sigma_1^2 \neq \sigma_2^2) \\ F_{расч} \geq F_{кр} \\ |T_{расч}| \geq t_{кр} \end{array} \right\} \Rightarrow \text{тренд існує ;}$$

$$\left. \begin{array}{l} \text{– якщо} \\ \bar{y}_1 = \bar{y}_2 \\ S_1^2 = S_2^2 \quad (\sigma_1^2 = \sigma_2^2) \\ F_{розр} < F_{кр} \\ |T_{розр}| < t_{кр} \end{array} \right\} \Rightarrow \text{тренд не існує.}$$

Розглянемо на прикладі перевірку гіпотези H_0 за допомогою критерію, заснованого на порівнянні середніх рівня ряду.

Приклад 2.1.4. Нехай є дані по врожайності зернових в одній з областей, ц із 1 га:

Таблиця 2.1.3

Врожайність зернових в одній з областей, ц з 1 га

Рік	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Врожайність зернових	14,1	9,3	19,4	19,7	5,4	24,2	13,8	24,5	14,7	16,6	5,6	16,2	25,3	11,9	18,5
	Ряд 1 $n_1 = 7$							Ряд 2 $n_2 = 8$							
	Вихідний часовий ряд $n = n_1 + n_2 = 15$														

Поділимо вихідний ряд на дві частини. В результаті отримаємо 2 нових ряди з кількістю рівнів n_1 та n_2 відповідно.

Знаходимо середні рівні ряду за кожною з отриманих частин вихідного ряду:

$$\bar{y}_1 = \frac{1}{n_1} \cdot \sum_{i=1}^k y_i = 15,13; \quad \bar{y}_2 = \frac{1}{n_2} \cdot \sum_{i=k+1}^n y_i = 16,66.$$

Оскільки розбіжність між середніми є несуттєвою (незначущою), тобто $\bar{y}_1 \approx \bar{y}_2$, то, отже, часовий ряд не має тенденції (тренду).

Обчислюємо дисперсії:

$$S_1^2 = \frac{1}{n_1} \cdot \sum_{i=1}^k (y_i - \bar{y}_1)^2 = \frac{1}{7} \cdot \sum_{i=1}^7 (y_i - 15,13)^2 = 36,08;$$

$$S_2^2 = \frac{1}{n_2} \cdot \sum_{i=k+1}^n (y_i - \bar{y}_2)^2 = \frac{1}{8} \cdot \sum_{i=8}^{15} (y_i - 16,66)^2 = 36,15.$$

Оскільки $S_1^2 \approx S_2^2$ (розбіжність між ними є незначущою), то часовий ряд не має тенденції (тренду).

Перевіряємо гіпотезу про рівність дисперсій ($H_0: \sigma_1^2 = \sigma_2^2$; $H_1: \sigma_1^2 \neq \sigma_2^2$) при рівні значущості $\alpha = 0,05$. Знаходимо розрахункове та критичне значення критерію Фішера:

$$\left. \begin{aligned} F_{розр} &= \frac{S_2^2}{S_1^2} = \frac{36,15}{36,08} = 1,002 \\ F_{кр} &= (0,05; 7, 6) = 4,21 \end{aligned} \right\} \Rightarrow F_{розр} < F_{кр} \Rightarrow \begin{aligned} &\text{приймається} \\ &\text{гіпотеза} \\ &H_0: \sigma_1^2 = \sigma_2^2. \end{aligned}$$

За даними спостережень дисперсії генеральних сукупностей дорівнюють $\sigma_1^2 = \sigma_2^2$, різниця між виправленими вибірковими дисперсіями (S_1^2 і S_2^2) є незначущою (розбіжність між ними є випадковою).

Тоді можна перевірити основну гіпотезу: $H_0: \bar{y}_1 = \bar{y}_2$; $H_1: \bar{y}_1 \neq \bar{y}_2$. Знаходимо розрахункове та критичне значення критерію Стьюдента:

$$\left. \begin{array}{l} T_{розр} \approx -0,49 \\ t_{кр} = (0,05; 13) = 2,16 \end{array} \right\} \Rightarrow |T_{розр}| < t_{кр} \Rightarrow \begin{array}{l} \text{приймається} \\ \text{гіпотеза} \\ H_0: \bar{y}_1 = \bar{y}_2. \end{array}$$

Розбіжність між обчисленими середніми є незначущою.

Звідси висновок: оскільки

$$\left\{ \begin{array}{l} \bar{y}_1 = \bar{y}_2, \\ S_1^2 = S_2^2 \quad (\sigma_1^2 = \sigma_2^2), \\ F_{розр} < F_{кр}, \\ |T_{розр}| < t_{кр}, \end{array} \right.$$

то, отже, тренд врожайності зернових відсутній.

2.1.3.4 Метод Форстера-Стюарта

Даний метод має більше можливостей і дає більш надійні результати, ніж попередні критерії.

Крім тренду самого ряду (тренду в середньому), він дозволяє встановити існування тренду дисперсії часового ряду:

- якщо тренду дисперсії немає, то розкид рівнів ряду постійний;
- якщо дисперсія збільшується, то ряд «розхитується».

Реалізація даного методу складається з **4 етапів**:

1 етап: Порівнюється кожний рівень вихідного часового ряду (починаючи з 2^{го} рівня) з усіма попередніми.

При цьому визначаються 2 числові послідовності u_t й l_t ($t = 2, 3, \dots, n$):

- а) якщо будь-який рівень ряду перевищує за своєю величиною кожний з попередніх рівнів, то величині u_t привласнюється значення «1», в інших випадках вона дорівнює «0», тобто:

$$u_t = \begin{cases} 1, & \text{якщо } y_t > y_{t-1}, y_{t-2}, \dots, y_1, \\ 0, & \text{в інших випадках;} \end{cases}$$

- б) якщо рівень ряду менший за всі попередні рівні, то отже величина l_t дорівнює «1», в інших випадках вона дорівнює «0», тобто:

$$l_t = \begin{cases} 1, & \text{якщо } y_t < y_{t-1}, y_{t-2}, \dots, y_1, \\ 0, & \text{в інших випадках.} \end{cases}$$

2 етап: Розраховуються величини s і d :

$$s = \sum_{t=2}^n (u_t + l_t); \quad d = \sum_{t=2}^n (u_t - l_t).$$

Величини s і d є асимптотично нормальними й мають незалежні розподіли. Вони суттєво залежать від порядку розташування рівнів у часі.

- Величина s , яка характеризує зміну часового ряду, здобуває значення

$$0 \leq s \leq (n-1)$$

(всі рівні ряду однакові)
(ряд є монотоним)

- Величина d характеризує зміну дисперсії рівнів часового ряду й змінюється у межах

$$-(n-1) \leq d \leq (n-1)$$

(ряд поступово спадає (згасає))
(ряд поступово зростає (розхитується))

3 етап: За допомогою величини s можна перевірити, чи існує тенденція зміни в дисперсії, а величина d дозволяє виявити тенденцію в середній. З цією метою на даному етапі перевіряються 2 гіпотези щодо того, чи можна вважати випадковими:

- а) відхилення величини s від μ (де μ – математичне сподівання величини s для ряду), у якому рівні розподілені випадково та чи суттєво s відрізняється від μ ;
- б) відхилення величини d від 0 і чи суттєво вони відрізняються один від одного.

Ці гіпотези перевіряються за допомогою розрахункових значень t -критерію Стьюдента для середньої й дисперсії відповідно:

$$T_s = \frac{|s - \mu|}{\sigma_1}; \quad T_d = \frac{|d - 0|}{\sigma_2};$$

де μ – математичне сподівання величини s для ряду, у якому рівні розташовані випадково; σ_1 – середньоквадратичне відхилення (помилка) для величини s : $\sigma_1 = \sqrt{2 \ln n - 3,4253}$; σ_2 – середньоквадратичне відхилення (помилка) для величини d : $\sigma_2 = \sqrt{2 \ln n - 0,8456}$.

Значення μ, σ_1, σ_2 для різних n наведені у таблиці А.2 (див. Додаток А). Величини T_s й T_d мають розподіл Стьюдента з $k = n - 1$ ступенями свободи.

4 етап: Розрахункові значення T_s й T_d порівнюються з табличними значеннями t -критерію Стьюдента з $k = n - 1$ ступенями свободи й при заданому рівні значущості α :

- якщо $T_d \geq t_{кр}(\alpha, k)$, то гіпотеза про відсутність тренду в середній відхиляється;

2.1.3.5 Критерій квадратів послідовних різниць (критерій Аббе)

Якщо є підстави вважати, що випадковий розкид спостережень y_t щодо своїх середніх значень відповідає нормальному закону розподілу ймовірностей, то для з'ясування питання про можливий систематичний зсув середнього в процесі вибіркового обстеження, доцільно скористатися критерієм Аббе, який в цьому випадку є найбільш потужним.

Нехай y_1, \dots, y_n – взаємно незалежні нормально розподілені випадкові величини з однаковими, але невідомими дисперсіями.

Критерій Аббе призначений для перевірки гіпотези про рівність середніх (про відсутність тенденції)

$$H_0 : My_1 = My_2 = \dots = My_n$$

проти альтернативи (про наявність тенденції)

$$H_1 : |My_{i+1} - My_i| > 0 \quad (\text{для всіх } i = 1, 2, \dots, n-1).$$

Статистика критерію Аббе (у сучасній формі) являє собою відношення (для $4 \leq n \leq 60$ взаємно незалежних нормально розподілених вибірок)

$$q_{emn} = \frac{1}{2} \cdot \frac{\sum_{i=1}^{n-1} (y_{i+1} - y_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2},$$

де $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ – середнє арифметичне (математичне сподівання) з вибірки;

n – кількість значень у вибірці (кількість членів рівнів ряду).

Якщо альтернативна гіпотеза H_1 (про наявність тренду) є вірною, то знаменник q_{emn} є більшим за чисельник. Тому значення цієї статистики будуть меншими за тих, які спостерігаються, коли є справедливою основна гіпотеза про рівність середніх H_0 .

Згідно із критерієм Аббе:

- якщо $q_{emn} < q_n(p)$, то приймається гіпотеза H_1 (тренд є);
- якщо $q_{emn} \geq q_n(p)$, то приймається гіпотеза H_0 (тренду немає).

Критичне (табличне) значення $q_n(p)$ визначається з таблиці А.4 для q – критерію Аббе, де p – заданий рівень значущості (див. Додаток А).

При великих n відношення q розподілене приблизно, як

$$1 + \frac{\xi}{\sqrt{n + \frac{1}{2}(1 + \xi^2)}},$$

де ξ – нормальна випадкова величина з параметрами $(0, 1)$.

Якщо $n > 60$, то для обчислення $q_n(p)$ рекомендується скористатися наближеною формулою

$$q_n(p) \approx 1 + \frac{\Psi(P)}{\sqrt{n + \frac{1}{2} \{1 + [\Psi(P)]^2\}}},$$

де $\Psi(P)$ – є P -квантиль нормального розподілу.

У якості таких величин, для яких застосуємо критерій Аббе, можуть виступати вибіркові частки або відсотки, середні арифметичні й інші статистики вибірових розподілів, якщо вони близькі до нормального.

Приклад 2.1.6: У таблиці 2.1.5 надана динаміка відсотку студентів 4 курсу, які склали іспити на «відмінно» у зимові сесії протягом 10 років роботи одного з факультетів університету. Потрібно встановити, чи є тенденція до підвищення успішності.

Таблиця 2.1.5

Динаміка відсотку відмінників 4 курсу за 10 років роботи факультету

№ з/п	Навчальний рік, t	y_t , %
1	2001-2002	10,8
2	2002-2003	16,4
3	2003-2004	17,4
4	2004-2005	22,0
5	2005-2006	23,0
6	2006-2007	21,5
7	2007-2008	26,1
8	2008-2009	17,2
9	2009-2010	27,5
10	2010-2011	33,0

Прийmemo наступні гіпотези:

H_0 – про відсутність тенденції (тренду) (тобто про рівність відсотків);

H_1 – про наявність тренду у часовому ряді y_t .

Розв'язання

Усередняємо відсотки y_t , тобто знаходимо

$$\bar{y}_t = \frac{1}{n} \sum_{t=1}^n y_t = 21,49 \approx 21,5.$$

Обчислюємо значення q_{emn} :

$$q_{emn} = \frac{1}{2} \cdot \frac{293,48}{367,8} = 0,39897 \approx 0,4.$$

Порівнюємо значення q_{emn} з табличним значенням $q_n(p)$ при заданому рівні значущості 0,05, тобто із критичним значенням $q_{10}(0,05) = 0,5311$, яке знаходимо з таблиці А.4 (Додаток А).

Оскільки $q_{emn} < q_n(p)$, гіпотеза про рівність відсотка «відмінників» відхиляється та приймається альтернативна гіпотеза про наявність тенденції у вихідному часовому ряді. Отже, висновок: тренд існує.

2.1.4 Питання для самоконтролю

1. Що називають часовим рядом? Відмінність його від динамічного ряду.
2. Який ряд називають стаціонарним?
3. Які спостереження у часовому ряді називають аномальними?
4. Розкрийте сутність методу Ірвіна.
5. Які існують методи коригування аномальних рівнів ряду?
6. Що називається трендом?
7. За якими критеріями перевіряється наявність тенденції у часовому ряді?
8. Наведіть основні етапи використання критерію, заснованого на порівнянні середніх рівня ряду.
9. Розкрийте сутність методу Форстера-Стюарта.
10. За яких умов використовується критерій послідовних різниць?

2.1.5 Рекомендована література



Для самостійного вивчення теми рекомендується література: основна – [2-5, 7, 8, 10, 11].

2.2 Методи екстраполяції в прогнозуванні виробничих процесів

2.2.1 Сутність екстраполяції

У теорії і практиці в процесі прогнозування досить часто використовують методологію екстраполяції.

Під **екстраполяцією** розуміється метод, за яким прогнозовані показники розраховуються як продовження динамічного ряду на майбутнє за виявленою закономірністю розвитку у минулому та теперішньому.

Тобто суть методу екстраполяції полягає у перенесенні минулих і сьогоденних закономірностей, взаємозв'язків на майбутні періоди. Він дає можливість показати, яких результатів можна досягти в майбутньому, якщо рухатися з тією ж швидкістю і з тим же прискоренням.

Необхідним елементом при цьому є побудова та аналіз так званого **часового ряду**, який класифікує значення показників у часі у розрізі окремих періодів та описує динаміку їх розвитку. Метод дозволяє знайти рівень ряду за його межами, у майбутньому.

При чому, за умови використання цього методу, необхідною умовою для отримання якісних результатів прогнозування є стабільність основних факторів і тенденцій минулих років, яка зберігається на період прогнозу. Отже, важливою передумовою застосування даного методу є наявність **тренду**, тобто тенденції часового ряду, яка характеризує основну закономірність руху в часі. Тому, перш ніж застосувати метод екстраполяції до прогнозування економічних показників необхідно перевірити наявність тренду у часовому ряді (див. підрозділ 2.1).

Якщо ж розвиток показників (рівнів) часового ряду у попередніх періодах характеризується значною нестабільністю і суттєвим коливанням показників, то їх екстраполяція на майбутні періоди буде неможливою, а отже, недоцільним є використання відповідних методів.

Крім того, з метою забезпечення реальності прогнозу слід пам'ятати загальне емпіричне правило, суть якого полягає в тому, що термін прогнозування не повинен перевищувати третьої частини довжини бази прогнозу, тобто наявного часового ряду. Так наприклад, для прогнозу на 5 років необхідно мати дані хоча би за 15 років. Відповідно чим більша база прогнозу і менший період прогнозу, тим надійнішими будуть його результати.

Екстраполяцію в загальному випадку можна представити у вигляді певного значення функції

$$y_{t+l}^* = f(y_i, l, a_j),$$

де y_{t+l}^* – прогнозоване значення ряду динаміки; l – час (період) упередження; y_i – рівень ряду, прийнятий за базу екстраполяції; a_j – параметр рівняння тренду.

Задача прогнозу полягає у визначенні виду екстраполюючих функцій на основі вихідних емпіричних даних.

2.2.2 Класифікація методів екстраполяції

Залежно від особливостей зміни рівнів у ряді динаміки прийнято виділяти наступні **методи екстраполяції**:

- **прості методи**, які засновані на припущенні щодо сталості в майбутньому абсолютних значень рівнів, середнього рівня ряду, середнього абсолютного приросту, середнього темпу росту;
- **складні методи** – засновані на виявленні основної тенденції, тобто застосуванні статистичних формул, що описують тренд.

2.2.3 Прості методи екстраполяції

Прості методи екстраполяції використовуються в керуванні виробництвом, оскільки мають ряд переваг:

- досить простий апарат дослідження;
- швидкість виконання розрахунків;
- наявність невеликого масиву інформації.

До таких методів належать:

- екстраполяція на основі середнього рівня ряду (екстраполяція середнього);
- екстраполяція на основі аналітичних показників рядів динаміки;
- екстраполяція на основі індексу сезонності.

Розглянемо основні методи простої прогнозовної екстраполяції та їх застосування в економічному прогнозуванні.

2.2.3.1 Екстраполяція середнього

Даний метод ґрунтується на принципі, за яким прогнозований рівень приймається рівним середньому значенню рівнів ряду в минулому, тобто

$$y_{t+l}^* = \bar{y},$$

де \bar{y} – середній рівень ряду.

У цьому випадку екстраполяція дає прогностичну *точкову оцінку* з середньою помилкою прогнозу, яка визначається за формулою

$$\mu = \pm \sqrt{\frac{\sigma^2}{n}},$$

де $\sigma = \frac{\sum (y - \bar{y})^2}{n}$ – дисперсія.

Точний збіг цих оцінок з фактичними даними – явище малоімовірне. Отже, прогноз повинен бути представлений у вигляді інтервалу значень.

Довірчий інтервал прогнозу для середньої при невеликій кількості спостережень визначається у вигляді

$$y_{t+l}^* = \bar{y} \pm t_\alpha S \sqrt{1 + \frac{1}{n}}, \quad (2.2.1)$$

де t_α – табличне значення t -критерію Стьюдента з $n-1$ ступенями свободи та

рівнем імовірності p ; $S = \sqrt{\frac{\sum (y - \bar{y})^2}{n-1}}$ – середньоквадратичне відхилення для вибірки.

Недолік розглянутого підходу полягає в тому, що довірчий інтервал тут не пов'язаний з періодом упередження l .

Приклад 2.2.1. Скласти прогноз продажу хлібобулочних виробів у магазині на 11-й, 12-й та 13-й дні його роботи, якщо відомі обсяги продажу (у тоннах) хлібобулочних виробів у магазині за минулі 10 днів роботи (табл. 2.2.1).

Таблиця 2.2.1

Обсяги продажу хлібобулочних виробів у магазині за минулі 10 днів роботи, т

1-й день	2-й день	3-й день	4-й день	5-й день	6-й день	7-й день	8-й день	9-й день	10-й день
2,5	2,8	2,0	2,4	2,3	2,9	2,7	2,2	2,3	2,8

Розв'язання

Враховуючи певну стабільність споживання товару, складемо прогноз продажу на основі його середнього обсягу реалізації за день, тобто за формулою середньої

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i = \frac{y_1 + y_2 + \dots + y_n}{n},$$

де y_1, y_2, \dots, y_n – обсяг продажу хлібобулочних виробів за кожний день роботи, т; n – число днів періоду.

Тоді маємо:

$$\bar{y} = \frac{24,9}{10} = 2,49 \text{ (т або близько 2,5 т у день)}.$$

Таким чином, екстраполяція продажу хлібобулочних виробів у магазині показала, що на 11-й день роботи, так само як і на 12-й і 13-й дні роботи, продаж товару може скласти 2,5 т.

Розрахуємо середню помилку прогнозу:

$$\mu = \pm \sqrt{\frac{\sigma^2}{n}} = \pm \sqrt{\frac{(0,081)^2}{10}} = \pm 0,0256 \text{ т} = \pm 25,6 \text{ кг}.$$

де μ – середня помилка; σ – дисперсія.

Отже, прогноз щоденного продажу становить 2,5 т і при незмінності формуючих попит даних може мати відхилення (25,6 кг). Зрозуміло, що такий прогноз буде слушний, якщо в найближчому майбутньому не закритється на ремонт сусідня булочна або не буде введений в експлуатацію недалеко від магазину новий багатоповерховий будинок, тобто якщо буде спостерігатися стабільність основних факторів і тенденцій.

У цьому випадку екстраполяція дає прогностичну точкову оцінку. Точний збіг цих оцінок з фактичними даними – явище малоімовірне. Отже, прогноз повинен бути поданий у вигляді інтервалу значень.

Знайдемо довірчий інтервал прогнозу за (2.2.1) у вигляді

$$\bar{y} - t_\alpha S \sqrt{\frac{n+1}{n}} \leq y_{n+1}^* \leq \bar{y} + t_\alpha S \sqrt{\frac{n+1}{n}}. \quad (2.2.2)$$

Для цього розрахуємо середньоквадратичне відхилення

$$S = \sqrt{\frac{\sum (y - \bar{y})^2}{n-1}} = \sqrt{\frac{0,81}{9}} = \sqrt{0,09} = 0,3$$

та за допомогою табл. А.3 (Додаток А) знайдемо табличне значення t -критерію Стьюдента

$$t_\alpha = t(\alpha; n-1) = t(0,05; 9) = 2,26.$$

Тоді довірчий інтервал прогнозу (2.2.2) прийме вигляд

$$2,49 - 2,26 \cdot 0,3 \sqrt{1,1} \leq y_{11}^* \leq 2,49 + 2,26 \cdot 0,3 \sqrt{1,1},$$

звідки

$$1,78 \leq y_{11}^* \leq 3,2.$$

Отже, ми одержали довірчий інтервал, у якому перебуває значення нашого прогнозу на 11-й день. Оскільки довірчий інтервал тут не пов'язаний з періодом упередження l , то на 12-й і 13-й дні роботи він зберігається таким самим.

2.2.3.2 Екстраполяція на основі аналітичних показників рядів динаміки

Для рядів динаміки соціально-економічних показників характерним є відображення динамічності економіки. Збільшення виробничих ресурсів, підвищення технічного рівня, продуктивності праці, поліпшення соціальних умов виробництва ведуть за собою нарощення виробництва, економічне зростання. Водночас ряд інших факторів (наприклад, інфляція) гальмують це зростання, призводять до зниження виробництва й економічного спаду.

На основі рівнів ряду динаміки визначають систему показників, які дають всебічну характеристику напряму та інтенсивності змін явища у часі. Позначимо: y_0 – початковий рівень часового ряду; y_n – кінцевий рівень часового ряду; y_i – умовно прийнятий (*i*-й) рівень часового ряду; n – кількість рівнів часового ряду.

Показники аналізу ряду динаміки можуть розраховуватися на постійній і змінній базисах порівняння. При цьому прийнято називати порівнюваний рівень **звітним**, а рівень, з яким проводиться порівняння, – **базисним**.

Залежно від того, яка база (рівень, з яким проводиться порівняння) взята для порівняння, розрізняють характеристики базисні й ланцюгові. Якщо база порівняння постійна (показники відносяться до начального рівня ряду), то характеристики динаміки називають **базисними**. Якщо база порівняння змінюється (показники відносяться до попереднього рівня), то характеристики динаміки будуть називатися **ланцюговими**.

Наведемо основні аналітичні показники динамічного ряду, які використовуються в прогнозуванні (табл. 2.2.2).

Таблиця 2.2.2

Основні аналітичні показники динамічного ряду

Характеристики	Розрахункові формули
1	2
<p>1. Абсолютний приріст – показує, наскільки даний рівень ряду перевищує рівень ряду, прийнятий за базу порівняння.</p> <p>Абсолютний приріст може бути додатнім (динаміка зростання), від’ємним (зменшення) або нульовим (без змін).</p>	<p>ланцюговий: $\Delta y_i = y_i - y_{i-1}$</p> <p>базисний: $\Delta y_i^0 = y_i - y_0$</p> <p>зв’язок між ними: $\sum \Delta y_i = \Delta y_i^0$</p>
<p>2. Середній абсолютний приріст (середня швидкість росту) – характеризує приріст за одиницю часу</p>	<p>ланцюговий: $\bar{\Delta y} = \frac{1}{n} \sum_{i=1}^n \Delta y_i$,</p> <p>де n – кількість ланцюгових темпів росту</p> <p>базисний: $\bar{\Delta y}_0 = \frac{y_n - y_0}{n - 1} = \frac{1}{n - 1} \sum_{i=1}^{n-1} \Delta y_i$,</p> <p>де n – кількість періодів</p>

1	2
<p>3. Коефіцієнт росту – показує, у скільки раз збільшився або зменшився рівень ряду щодо базисного</p> <p>Коефіцієнти росту, виражені у відсотках, прийнято називати темпами росту.</p>	<p>ланцюговий: $K_i = \frac{y_i}{y_{i-1}}$</p> <p>базисний: $K_i^0 = \frac{y_i}{y_0}$</p> <p>за весь період: $K_n = \frac{y_n}{y_0}$</p> <p>зв'язок між ними: $\prod K_i = K_n$</p> <p>При цьому:</p> <ul style="list-style-type: none"> - за $K_{\text{росту}} > 1$ рівень підвищується - за $K_{\text{росту}} < 1$ рівень зменшується - за $K_{\text{росту}} = 1$ - не змінюється
<p>4. Коефіцієнт приросту – показує, наскільки рівень ряду більший за той, з яким проводиться порівняння.</p> <p>При цьому додатне значення коефіцієнту приросту свідчить про зростання рівня, а від'ємне – про зниження (падіння).</p> <p>Коефіцієнти приросту, виражені у відсотках, прийнято називати темпами приросту.</p>	<p>ланцюговий: $k_i = \frac{\Delta y_i}{y_{i-1}}$</p> <p>базисний: $k_i^0 = \frac{\Delta y_i^0}{y_0}$</p> <p>за весь період: $k_n = \frac{\Delta y_n}{y_0}$</p> <p>зв'язок з коефіцієнтом росту:</p> $k_{\text{приросту}} = K_{\text{росту}} - 1 \text{ (або 100\%)}$ <p>(ланцюговий, базисний, за весь період)</p>
<p>5. Середній коефіцієнт росту – характеризує середню швидкість зміни показника, що вивчається, за певний період</p>	<p>ланцюговий: $\bar{K} = \sqrt[n]{K_1 \cdot K_2 \cdot \dots \cdot K_n}$, де n – кількість ланцюгових темпів росту</p> <p>базисний: $\bar{K}_0 = \sqrt[n-1]{\frac{y_n}{y_0}}$,</p> <p>де n – кількість періодів</p>
<p>6. Середній коефіцієнт приросту</p>	$\bar{k}_{\text{приросту}} = \bar{K}_{\text{росту}} - 1 \text{ (або 100\%)}$
<p>7. Абсолютний розмір 1% приросту – показує, що означає підвищити або знизити рівень певного суспільного явища на один відсоток</p>	<p>ланцюговий: $\Delta y_{1\%} = \frac{y_i - y_{i-1}}{\frac{y_i - y_{i-1}}{y_{i-1}} \cdot 100} = \frac{y_{i-1}}{100}$</p> <p>за весь період: $\bar{\Delta} y_{1\%} = \frac{\bar{\Delta} y}{\bar{k}_{\text{приросту}}}$</p>

1	2
8. Пункти росту – це різниця двох суміжних базових темпів росту (або приросту) у відсотках або коефіцієнтах; використовуються, коли проводиться порівняння достатньо віддалених у часі показників	у відсотках: $P_{\text{росту}} = K_{\text{росту}_i} - K_{\text{росту}_{i-1}}$ у коефіцієнтах: $P_{\text{росту}} = k_{\text{приросту}_i} - k_{\text{приросту}_{i-1}}$
9. Середній рівень ряду для інтервального ряду з рівними інтервалами – розраховується як середня арифметична проста	$\bar{y} = \frac{1}{n} \cdot \sum_{i=1}^n y_i$
10. Середній рівень ряду для інтервального ряду з нерівними інтервалами – розраховується як середня арифметична зважена	$\bar{y} = \frac{\sum y_i \cdot t}{\sum t},$ де t – число періодів часу, протягом яких рівень не змінюється
11. Середній рівень ряду для моментного ряду з рівними інтервалами – розраховується як середня хронологічна проста	$y_{xp} = \frac{\frac{1}{2}y_1 + y_2 + y_3 + \dots + y_{n-1} + \frac{1}{2}y_n}{n-1}$
12. Середній рівень ряду для моментного ряду з нерівними інтервалами – розраховується як середня хронологічна зважена	$y_{xp.зв.} = \frac{(y_1 + y_2)t_1 + (y_2 + y_3)t_2 + \dots + (y_{n-1} + y_n)t_{n-1}}{2\sum t_{n-1}}$

Екстраполяція на основі аналітичних показників рядів динаміки спирається на використання в прогнозуванні основних аналітичних показників часового ряду (табл. 2.2.2) та використовується при недостатній інформації про передісторію розвитку явища (немає досить довгого ряду або інформація задана тільки двома значеннями: на початок і кінець періоду).

Виділимо найбільш вживані методи екстраполяції на основі аналітичних показників рядів динаміки:

1. Метод середнього абсолютного приросту, за яким для побудови прогнозу на будь-який період необхідно визначити середній абсолютний приріст і послідовно додати його до останнього рівня ряду стільки раз, на скільки періодів екстраполюється ряд. Для побудови прогнозу за даним методом використовується наступна формула:

$$\hat{y}_{n+T} = y_n + \bar{\Delta}y_0 \cdot T,$$

де \hat{y} – прогнозне значення показника; $\bar{\Delta}y_0 = \frac{y_n - y_0}{n-1}$ – середній абсолютний приріст; T – величина горизонту прогнозу (час упередження), $T = 1; 2; 3; \dots$

Застосування в екстраполяції середнього абсолютного приросту припускає, що розвиток явища відбувається за арифметичною прогресією та використовується за умови, що загальна тенденція є лінійною.

2. Метод середнього темпу росту.

Екстраполяція за цим методом може здійснюватися, якщо тенденція ряду динаміки буде представлена показовою (експоненціальною) функцією. Для обрахунку перспективних показників використовується наступна формула:

$$\hat{y}_{n+T} = y_n (\bar{K}_0)^T,$$

де $\bar{K}_0 = \sqrt[n-1]{\frac{y_n}{y_0}}$ – середній коефіцієнт росту.

Зазначимо, що істотним недоліком показників середнього абсолютного приросту й середнього коефіцієнта росту є те, що значення їх повністю залежать від крайніх рівнів динамічного ряду. Проміжні значення, які багато в чому визначають тенденцію зміни показників, по суті не беруть участі в розрахунках. Отже, методи, які використовують ці показники, не є досить точними і виникає потреба в усуненні зазначеного недоліку, яке може бути здійснено за допомогою аналітичного вирівнювання рядів динаміки.

Крім того, перед тим, як застосовувати методи екстраполяції на основі аналітичних показників рядів динаміки, необхідно перевірити, яким саме методом – на основі середнього абсолютного приросту або середнього коефіцієнту росту – краще описується тенденція змін динамічного ряду. Тобто необхідно оцінити надійність застосовуваного методу, за яким в подальшому можна отримати якісний прогноз на майбутній період. Для такої оцінки використовується так званий **метод «прогноз екс-пост»**.

Суть його полягає в наступному:

1) вихідний часовий ряд розбивають на 2 частини:

$$y_t = \underbrace{y_1, y_2, \dots, y_k}_{n_1}, \underbrace{y_{k+1}, \dots, y_{n-1}, y_n}_{n_2}$$

$$\swarrow \qquad \searrow$$

$$y^1 = y_1, y_2, \dots, y_{k-1}, y_k \quad \text{і} \quad y^2 = y_{k+1}, y_{k+2}, \dots, y_{n-1}, y_n$$

$$(t_1 = \overline{1, k}) \qquad \qquad \qquad (t_2 = \overline{k+1, n})$$

2) за даними першої частини, умовно прийнятої за «передісторію», будується рівняння (модель), на базі якої складається прогноз для другої частини (другого періоду), результати якого потім порівнюються з фактичними даними;

3) та модель, за якою отримані найменші відхилення між прогнозованими та фактичними даними, визначається найкращою, тобто такою, яка забезпечить побудову найбільш точних прогнозів. За обраною таким чином моделлю здійснюється побудова прогнозу на наступний $(n+T)$ -й період.

Метод «прогноз екс-пост» застосовується у випадку, коли точно невідомо, за яким з методів (моделей) будувати прогноз. У протилежному випадку, коли відомо, що обраний метод (модель) адекватно описує досліджуване явище (процес), то використання методу «прогноз екс-пост» є недоцільним. Отже, за своєю сутністю та призначенням зазначений метод виступає як метод перевірки адекватності обраних для прогнозування моделей.

Розглянемо на прикладі практичне застосування методів екстраполяції на основі аналітичних показників рядів динаміки та методу «прогноз екс-пост».

Приклад 2.2.2. Нехай надані дані про середнє споживання кондитерських виробів на одну людину за регіоном у рік (див. табл. 2.2.3) (дані за дві п'ятирічки і базисний рік). Необхідно побудувати прогноз споживання кондитерських виробів на наступну п'ятирічку.

Таблиця 2.2.3

Середньорічне споживання кондитерських виробів за регіоном

Номер року, t	1 (баз)	2	3	4	5	6	7	8	9	10	11
Споживання кондитерських виробів на 1 людину у рік, кг	10,7	11,5	12,2	13,4	15,0	15,0	15,9	17,2	18,1	19,8	21,2
	Ряд 1 $n_1 = 6$					Ряд 2 $n_2 = 5$					
	Весь часовий ряд $n = n_1 + n_2 = 11$										

Перш ніж застосувати для побудови прогнозу споживання кондитерських виробів на наступну п'ятирічку методи екстраполяції на основі аналітичних показників рядів динаміки, перевіримо, яким методом краще описується поведінка даних динамічного ряду. Скористаємось методом «прогноз екс-пост».

Розіб'ємо вихідний часовий ряд на 2 приблизно однакові за кількістю рівнів ряду частини:

$$y_t = \underbrace{y_1, y_2, \dots, y_6}_{t_1=1,6; n_1=6}, \underbrace{y_7, \dots, y_{11}}_{t_2=7,11; n_2=5}.$$

Використовуючи дані перших шести років – базисний рік ($y_0 = y_1 = 15$) і роки першої п'ятирічки y_2, \dots, y_6 , розрахуємо:

- середній абсолютний приріст:

$$\bar{\Delta}y_0 = \frac{y_6 - y_0}{6 - 1} = \frac{15,0 - 10,7}{5} = 0,9 \text{ кг};$$

- середньорічний коефіцієнт росту:

$$\bar{K}_0 = \sqrt[6-1]{\frac{y_6}{y_0}} = \sqrt[5]{\frac{15,0}{10,7}} = 1,07.$$

Складемо прогноз на основі середнього абсолютного приросту на період $t_2 = \overline{7,11}$ за формулою $\hat{y}_{k+T} = y_k + \bar{\Delta}y_0 \cdot T$, де $k = 6$, $T = \overline{1,5}$:

$$\hat{y}_7 = 15 + 0,9 \cdot 1 = 15,9 \text{ кг};$$

$$\hat{y}_8 = 15 + 0,9 \cdot 2 = 16,8 \text{ кг};$$

$$\hat{y}_9 = 15 + 0,9 \cdot 3 = 17,7 \text{ кг};$$

$$\hat{y}_{10} = 15 + 0,9 \cdot 4 = 18,6 \text{ кг};$$

$$\hat{y}_{11} = 15 + 0,9 \cdot 5 = 19,5 \text{ кг}.$$

Результати прогнозу зведемо в таблицю й порівняємо з фактичними даними (див. табл. 2.2.4)

Таблиця 2.2.4

**Оцінка якості прогнозу, складеного
на основі середнього абсолютного приросту**

№	Фактичне значення, кг	Прогнозне значення, кг	Відхилення	
			Абсолютне (стовп. 2 - стовп. 3), кг	Відносне (стовп. 4 / стовп. 2) 100%
1	2	3	4	5
1	15,9	15,9	0	0
2	17,2	16,8	0,4	2,3
3	18,1	17,7	0,4	2,2
4	19,8	18,6	1,2	6,1
5	21,2	19,5	1,7	8,0
Середнє значення		—	0,7	3,7

Складемо прогноз на основі середньорічного коефіцієнта росту на період $t_2 = \overline{7,11}$ за формулою $\hat{y}_{k+T} = y_k (\bar{K}_0)^T$, де $k = 6$, $T = \overline{1,5}$:

$$\hat{y}_7 = 15 \cdot 1,07^1 = 16 \text{ кг};$$

$$\hat{y}_8 = 15 \cdot 1,07^2 = 17,2 \text{ кг};$$

$$\hat{y}_9 = 15 \cdot 1,07^3 = 18,4 \text{ кг};$$

$$\hat{y}_{10} = 15 \cdot 1,07^4 = 19,7 \text{ кг};$$

$$\hat{y}_{11} = 15 \cdot 1,07^5 = 21 \text{ кг}.$$

Результати прогнозу зведемо в таблицю, порівняємо з фактичними даними й оцінимо якість прогнозу (див. табл. 2.2.5).

Порівнюючи результати прогнозів, наведених у табл. 2.2.4 і табл. 2.2.5, можна зробити наступний висновок: використання середньорічного коефіцієнта росту забезпечує більш високу точність прогнозу. Про це свідчать відхилення за всіма роках та в цілому (середнє значення відхилень) за п'ятирічку.

Отже, для нашого прикладу найточнішим методом побудови прогнозів є метод на основі середньорічного коефіцієнту росту. За ним і будемо складати прогноз на наступну п'ятирічку.

Таблиця 2.2.5

**Оцінка якості прогнозу, складеного
на основі середньорічного коефіцієнта росту**

№	Фактичне значення, кг	Прогнозне значення, кг	Відхилення	
			Абсолютне (стовп. 2 - стовп. 3), кг	Відносне (стовп. 4 / стовп. 2) 100%
1	2	3	4	5
1	15,9	16,0	- 0,1	- 0,6
2	17,2	17,2	0	0
3	18,1	18,4	- 0,3	- 1,7
4	19,8	19,7	0,1	0,5
5	21,2	21,0	0,2	0,9
Середнє значення		-	0,1	0,7

Для складання прогнозу за межі представлених даних, тобто на перспективу, розрахуємо середньорічний коефіцієнт росту на основі другої п'ятирічки з використанням базисного періоду (яким у цьому випадку є $y_6 = 15$):

$$\bar{K}_0 = n-k \sqrt[n-k]{\frac{y_n}{y_k}} = 11-6 \sqrt[5]{\frac{y_{11}}{y_6}} = 11-6 \sqrt[5]{\frac{21,2}{15,0}} = \sqrt[5]{1,413} = 1,071.$$

Прогноз споживання кондитерських виробів на наступну п'ятирічку будемо складати за формулою:

$$\hat{y}_{n+T} = y_n (\bar{K}_0)^T.$$

Отже, маємо прогноз споживання кондитерських виробів на наступні 5 років, складений з урахуванням збереження тенденцій, які склалися в «передісторії»:

$$\hat{y}_{n+1} = \hat{y}_{12} = 21,2 \cdot 1,071^1 = 22,7 \text{ кг};$$

$$\hat{y}_{n+2} = \hat{y}_{13} = 21,2 \cdot 1,071^2 = 24,3 \text{ кг};$$

$$\hat{y}_{n+3} = \hat{y}_{14} = 21,2 \cdot 1,071^3 = 26 \text{ кг};$$

$$\hat{y}_{n+4} = \hat{y}_{15} = 21,2 \cdot 1,071^4 = 28 \text{ кг};$$

$$\hat{y}_{n+5} = \hat{y}_{16} = 21,2 \cdot 1,071^5 = 29,9 \text{ кг}.$$

Розглянемо ще один приклад застосування методу екстраполяції на основі середнього коефіцієнту росту.

Приклад 2.2.3. Скласти прогноз обсягу платних послуг у фірмі на 2012 і 2013 рр. Відомо, що в 2005 р. обсяг платних послуг у фірмі становив 207 грошових одиниць (грош. од.), а в 2011 р. – 228 грош. од. у порівнянних цінах. Умови роботи фірми, які склалися, суттєво не змінюється.

Розв'язання

Для визначення прогнозу методом екстраполяції за приростом обсягу послуг, який склався, треба визначити його середньорічний зріст за минулі роки й екстраполювати на майбутні періоди. Розрахуємо середній коефіцієнт росту за формулою

$$k = n \sqrt[n]{\frac{y_n}{y_1}},$$

де n – число періодів динамічного ряду, ($n = 6$); y_n – обсяг платних послуг у звітному році, грош. од.; y_1 – обсяг платних послуг у базисному році, грош. од.

Середній коефіцієнт росту попиту на платні послуги фірми складе:

$$k = 6 \sqrt[6]{\frac{y_6}{y_1}} = 6 \sqrt[6]{\frac{228}{207}} = 1,0162, \text{ або } 101,62 \ \%.$$

Відповідно прогноз обсягу послуг складе:

- на 2012 р.: $228 \times 1,0162 = 231,70$ (грош. од.);
- на 2013 р.: $231,70 \times 1,0162 = 235,46$ (грош. од.).

Висновок. Прогноз обсягу у фірмі складе: на 2012 р. – 231,70 грош. од., на 2013 р. – 235,46 грош. од.

2.2.3.3 Екстраполяція тренду

Під **екстраполяцією тренду** розуміють продовження виявленої в процесі аналізу тенденції за межі побудованого на основі емпіричних даних ряду динаміки.

Передумовою використання цього методу прогнозування є сталість чинників, що формують виявлений тренд, а принциповим моментом – виявлення тренду, характерного для досліджуваного ряду динаміки.

Для такого виду екстраполяції характерне знаходження плавної лінії, яка відображує закономірності розвитку в часі. Тренд звичайно застосовується як основна складова прогнозованого часового ряду, на яку накладаються інші складові, наприклад сезонні коливання.

Екстраполяція на основі тренду включає:

- збір інформації емпіричного ряду показника за минулі періоди;
- вибір оптимального виду функції, яка описує зазначений ряд з урахуванням його згладжування й вирівнювання;
- визначення параметрів обраної екстраполяційної функції;
- розрахунок прогнозу на майбутнє за обраною функцією.

Вибір оптимального виду функції, яка описує часовий ряд, здійснюється за допомогою побудови так званих **моделей кривих росту**, в основу яких покладено принцип отримання за допомогою методу найменших квадратів (МНК) оцінки детермінованої компоненти часового ряду, яка характеризує основну тенденцію.

Плавну криву (гладку функцію), яка апроксимує часовий ряд, прийнято називати **кривою росту**.

Підбір такої кривої є **аналітичним (не механічним) вирівнюванням** часового ряду, тобто вирівнюванням ряду за допомогою аналітичних формул.

При вивченні загальної тенденції методом аналітичного вирівнювання виходять із того, що зміни рівнів ряду динаміки можуть бути з тим або іншим ступенем точності наближені виражені певними математичними функціями. Вид апроксимуючої функції визначається характером динаміки розвитку конкретного явища. Логічний аналіз при виборі виду функції (рівняння) може бути заснований на розрахованих показниках динаміки, а саме:

- якщо відносно стабільні абсолютні прирости (перші різниці рівнів приблизно рівні), то згладжування може бути виконане за прямою;
- якщо абсолютні прирости рівномірно збільшуються (другі різниці рівнів приблизно рівні), можна прийняти параболу другого порядку;
- при прискорено зростаючих або спадаючих абсолютних приростах – параболу третього порядку;
- при відносно стабільних темпах росту – показову функцію.

Найпростішим шляхом вибору кривої росту (апроксимуючої функції) є візуальний підхід, який спирається на графічне зображення часового ряду. Необхідно підібрати таку криву росту, форма якої відповідає фактичному розвитку процесу. Якщо на графіку вихідного ряду тенденція розвитку недостатньо чітко проглядається, то можна провести деякі стандартні перетворювання ряду (наприклад, згладжування), а потім підібрати функцію, яка відповідає графіку отриманого в такий спосіб нового ряду.

Отже, метою аналітичного вирівнювання є визначення аналітичної або графічної залежності. На практиці за вихідним часовим рядом задають вигляд та знаходять параметри апроксимуючої функції, а потім аналізують поведінку відхилень від тенденції. При цьому апроксимуючу функцію вибирають в такий спосіб, щоб вона давала змістовне пояснення досліджуваного процесу. Найчастіше при вирівнюванні використовуються наступні залежності: лінійна, параболічна й експоненціальна.

Параметри апроксимуючої функції визначається за допомогою МНК так, щоб графік цієї функції розташовувався на мінімальному віддаленні від точок вихідних даних часового ряду. Докладніше про застосування МНК до визначення невідомих параметрів рівняння можна ознайомитися у матеріалах, наведених у Додатку В.

Для визначення прогнозного значення показника часового ряду за методом екстраполяції тренду необхідно підставити до отриманої моделі кривої росту відповідне значення часового параметру.

Отриманий у такий спосіб прогноз називають **точковим**, тому що для кожного моменту часу визначається тільки одне значення прогнозованого показника.

На практиці для доповнення точкового прогнозу бажане визначити границі можливої зміни прогнозованого показника, тобто обчислити інтервальний прогноз.

Розбіжність між фактичними даними та точковим прогнозом, отриманим шляхом екстраполяції тенденції за кривими росту, може бути викликана:

- а) суб'єктивною помилковістю вибору виду кривої;
- б) похибкою оцінювання параметрів кривих;
- в) похибкою, пов'язаною з відхиленням окремих спостережень від тренду, який характеризує деякий середній рівень ряду на кожний момент часу.

Похибка, пов'язана з джерелами б) та в), може бути відображена у вигляді довірчого інтервалу прогнозу.

Довірчий інтервал, який враховує невизначеність, пов'язану з положенням тренду, та можливість відхилення від цього тренду, визначається у вигляді

$$\hat{y}_{n+l} \pm t_{\alpha} S_p, \quad (2.2.3)$$

де n – довжина часового ряду; l – період упередження; \hat{y}_{n+l} – точковий прогноз на момент $(n+l)$; t_{α} – значення t -статистики Стьюдента; S_p – середня квадратична помилка прогнозу.

Припустимо, що тренд характеризується прямою $\hat{y}_t = a_0 + a_1 t$.

Оскільки оцінки параметрів визначаються за вибірковою сукупністю, представленою часовим рядом, то вони містять похибку. При цьому похибка параметру a_0 призводить до вертикального зрушення прямої, похибка параметру a_1 – до зміни кута нахилу прямої щодо осі абсцис. З урахуванням розкиду конкретних реалізацій відносно ліній тренду, дисперсію S_p^2 можна представити у вигляді

$$S_p^2 = \frac{S_y^2}{n} + S_y^2 \frac{(t_1 - \bar{t})^2}{\sum_{t=1}^n (t - \bar{t})^2} + S_y^2, \quad (2.2.4)$$

де S_y^2 – дисперсія відхилень фактичних спостережень від розрахункових; t_1 – час упередження, для якого робиться екстраполяція; $t_1 = n+l$; t – порядковий номер рівнів ряду, $t = \overline{1, n}$; \bar{t} – порядковий номер рівня, який знаходиться в середині ряду, $\bar{t} = \frac{n+1}{2}$.

Тоді довірчий інтервал можна подати у вигляді (рис. 2.2.1)

$$\hat{y}_{n+l} \pm t_{\alpha} S_y \sqrt{\frac{n+1}{n} + \frac{(t_1 - \bar{t})^2}{\sum_{t=1}^n (t - \bar{t})^2}}. \quad (2.2.5)$$

Позначимо корінь у виразі (2.2.5) через K . Значення K залежить тільки від довжини ряду n й періоду упередження l . Тому можна скласти таблиці значень K або $K^* = t_\alpha K$. Тоді інтервальна оцінка буде мати вигляд

$$\hat{y}_{n+l} \pm S_y K^*. \quad (2.2.6)$$

Вираз, аналогічний (2.2.5), можна одержати для полінома другого ступеня:

$$\hat{y}_{n+l} \pm t_\alpha S_y \sqrt{1 + \frac{t_1^2}{\sum t^2} + \frac{\sum t^4 - (2\sum t^2)t_1^2 + nt_1^4}{n\sum t^4 - (\sum t^2)^2}} \quad (2.2.7)$$

або

$$\hat{y}_{n+l} \pm S_y K^*. \quad (2.2.8)$$

Дисперсія відхилень фактичних спостережень від розрахункових визначається у вигляді

$$S_y^2 = \frac{\sum_{t=1}^n (y_t - \hat{y}_t)^2}{n - k}, \quad (2.2.9)$$

де y_t та \hat{y}_t – відповідно фактичні та розрахункові значення рівнів ряду, n – довжина ряду, k – число оцінюваних параметрів вирівнюючої кривої.

Отже, ширина довірчого інтервалу залежить від рівня значущості, періоду упередження, середнього квадратичного відхилення від тренду й ступеня полінома.

Чим вище ступінь полінома, тем ширше довірчий інтервал при тому самому значенні S_y , оскільки дисперсія рівняння тренду обчислюється як зважена сума дисперсій відповідних параметрів рівняння.

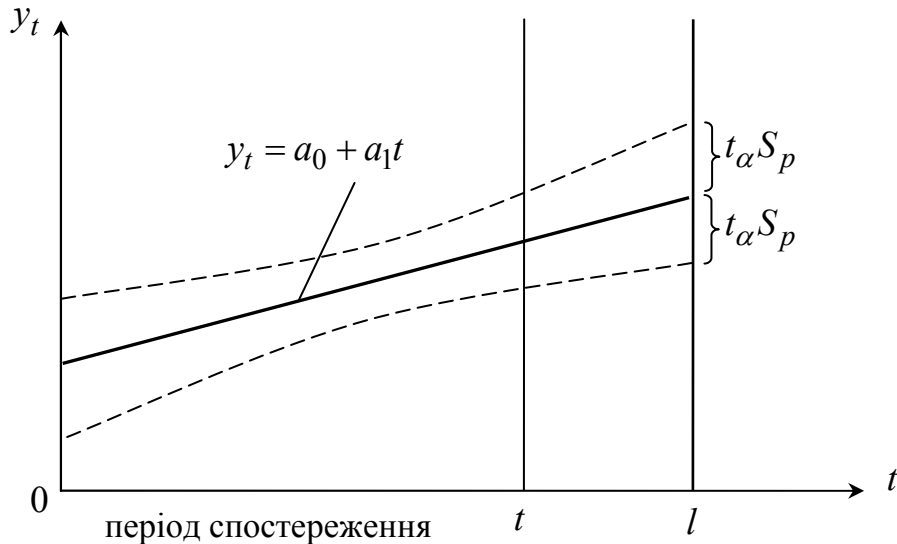


Рисунок 2.2.1 – Довірчі інтервали прогнозу для лінійного тренду

Довірчі інтервали прогнозів, отриманих з використанням рівняння експоненти, визначають аналогічним образом. Відмінність полягає в тому, що як і при обчисленні параметрів кривої, так і при обчисленні середньої квадратичної помилки використовують не самі значення рівнів часового ряду, а їх логарифми.

У табл. 2.2.6 наведені значення K^* залежно від довжини часового ряду n та періоду упередження l для прямої й параболи. Вочевидь, що при збільшенні довжини рядів (n) значення K^* зменшуються, з ростом періоду упередження l значення K^* збільшуються. При цьому вплив періоду упередження є неоднаковим для різних значень n : чим більшою є довжина ряду n , тем менший вплив чине період упередження l .

Таблиця 2.2.6

Значення K^* для оцінки довірчих інтервалів прогнозу на основі лінійного тренду й параболічного тренду при довірчій ймовірності 0,9 (7)

Довжина ряду n	Лінійний тренд			Довжина ряду n	Параболічний тренд		
	Період упередження l				Період упередження l		
	1	2	3		1	2	3
7	2,6380	2,8748	3,1399	7	3,948	5,755	8,152
8	2,4631	2,6391	2,8361	8	3,459	4,754	6,461
9	2,3422	2,4786	2,6310	9	3,144	4,124	5,408
10	2,2524	2,3614	2,4827	10	2,926	3,695	4,698
11	2,1827	2,2718	2,3706	11	2,763	3,384	4,189
12	2,1274	2,2017	2,2836	12	2,636	3,148	3,808
13	2,0837	2,1463	2,2155	13	2,536	2,965	3,516
14	2,0462	2,1000	2,1590	14	2,455	2,830	3,286
15	2,0153	2,0621	2,1131	15	2,386	2,701	3,100
16	1,9883	2,0292	2,0735	16	2,330	2,604	2,950
17	1,9654	2,0015	2,0406	17	2,280	2,521	2,823
18	1,9455	1,9776	2,0124	18	2,238	2,451	2,717
19	1,9280	1,9568	1,9877	19	2,201	2,391	2,627
20	1,9117	1,9375	1,9654	20	2,169	2,339	2,549
21	1,8975	1,9210	1,9461	21	2,139	2,293	2,481
22	1,8854	1,9066	1,9294	22	2,113	2,252	2,422
23	1,8738	1,8932	1,9140	23	2,090	2,217	2,371
24	1,8631	1,8808	1,8998	24	2,069	2,185	2,325
25	1,8538	1,8701	1,8876	25	2,049	2,156	2,284

Отже, аналітичне вирівнювання часового ряду за допомогою моделей кривих росту дозволяє не тільки визначити загальну тенденцію зміни явища на розглянутому відрізку часу, але й виконувати розрахунки для таких періодів, у відношенні яких немає вихідних даних.

Розглянемо застосування методу екстраполяції тренду на наступному прикладі.

Приклад 2.2.4. Скласти прогноз попиту на шкіряне взуття на 2011 р., використовуючи метод екстраполяції тренду, якщо відомі дані (табл. 2.2.7) про продаж населенню товару за останні 5 років (у грошових одиницях).

Таблиця 2.2.7

**Обсяг продажу населенню шкіряного взуття
в регіоні за останні 5 років, грош. од.**

роки, t	2007 р.	2008 р.	2009 р.	2010 р.	2011 р.
обсяг продажу, y	179,3	193,3	206,0	216,9	226,6

Розв'язання

За даними табл. 2.2.7 побудуємо графік продажу шкіряного взуття в регіоні за останні 5 років (рис. 2.2.2).

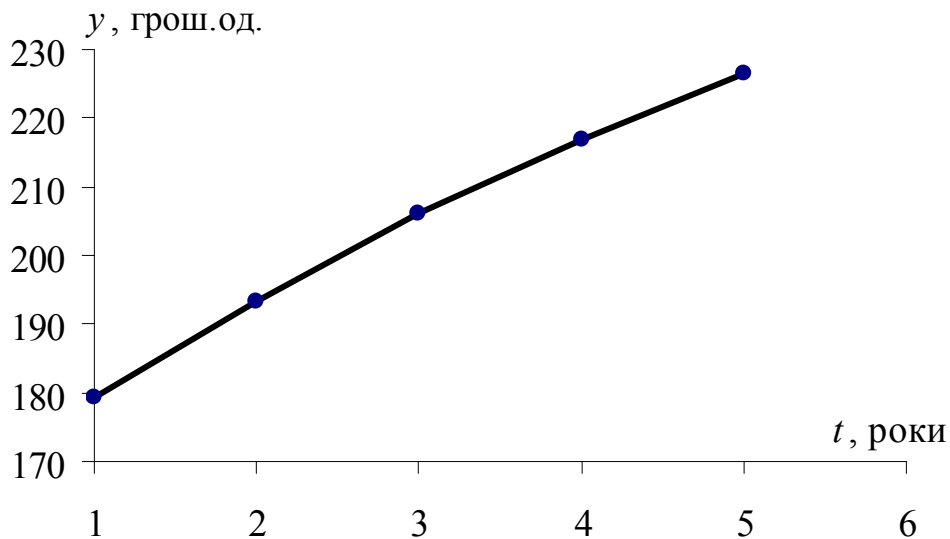


Рисунок 2.2.2 – Динаміка товарообігу шкіряного взуття в регіоні за 5 років

З рис. 2.2.2 видно, що лінія обсягу продажу за роками близька до лінійної залежності, яка виражається як

$$y = a + bt, \quad (2.2.10)$$

де y – обсяг товарообігу (попит); t – роки; a – параметр, який характеризує вплив основних факторів на продаж шкіряного взуття (ціни, грошових доходів населення, товарного забезпечення й ін.); b – параметр, що характеризує вплив допоміжних факторів на продаж шкіряного взуття (погоди, моди, сезонності, режиму роботи підприємств торгівлі, кваліфікації працівників та ін.).

Для знаходження параметрів рівняння (2.2.10) скористаємось методом найменших квадратів (Додаток В) та розв'яжемо систему рівнянь

$$\begin{cases} a \cdot n + b \sum t = \sum y, \\ a \sum t + b \sum t^2 = \sum ty. \end{cases} \quad (2.2.11)$$

Для визначення усіх коефіцієнтів системи (2.2.11) проведемо додаткові розрахунки (див. табл. 2.2.8).

Таблиця 2.2.8

Роки t	Попит y	$t \cdot y$	t^2
1-й	179,3	179,3	1
2-й	193,3	386,6	4
3-й	206,0	618,0	9
4-й	216,9	867,6	16
5-й	226,6	1133,0	25
Разом: 15	1022,1	3184,5	55

За даними табл. 2.2.8 маємо систему

$$\begin{cases} 5a + 15b = 1022,1, \\ 15a + 55b = 3184,5, \end{cases}$$

звідки отримаємо невідомі коефіцієнти

$$\begin{cases} a = 168,96; \\ b = 11,82. \end{cases} \quad (2.2.12)$$

Отже, з урахуванням (2.2.12) рівняння (2.2.10) матиме вигляд:

$$y = 168,96 + 11,82t. \quad (2.2.13)$$

За знайденим рівнянням (2.2.13) можна провести розрахунки тренду, тобто визначити вирівняні значення попиту за минулі роки й розрахувати прогноз на майбутнє. Так, для прогнозу на 7-й рік необхідно в отримане рівняння підставити черговий рік – 7:

$$y = 168,96 + 11,82 \cdot 7 = 251,7 \text{ (грош. од.)}.$$

Висновок. Попит на шкіряне взуття у 2013 р. може скласти 251,7 грош. од. (у цінах 2011 р.)

2.2.3.4 Екстраполяція на основі індексу сезонності

У процесі господарської діяльності окремі галузі промисловості, торгівля, збут зустрічаються із циклічними коливаннями, викликаними сезонним характером виробництва й споживання товарів і послуг.

Сезонні коливання – це більш-менш постійні внутрішньорічні коливання у часовому ряді, обумовлені специфічними умовами виробництва й споживання даного товару або послуги. Такі коливання показують коливання навколо тренду та пояснюється сезонним характером виробництва, споживання. Наприклад, у четвертому кварталі кожного року перед Новим роком значно зростає споживання товарів.

Для організації виробництва й реалізації продукції сезонних виробництв необхідно вивчити тенденцію сезонних коливань, які склалися, та побудувати прогноз на найближчу перспективу.

Для вивчення сезонних коливань використовуються спеціальні показники, називані **індексами сезонності**. Сукупність індексів сезонності утворює **сезонну хвилю**.

Сезонні коливання формуються під впливом не лише природно-кліматичних, але й соціально-економічних факторів. Сила і напрям дії окремих факторів формує різну конфігурацію сезонної хвилі. За своїм характером сезонна компонента може бути адитивною або мультиплікативною. Для адитивної компоненти характерні сталі коливання навколо середнього рівня чи тренду, для мультиплікативної — зростання амплітуди коливань з часом.

Індекс сезонності за кожним місяцем визначається за формулою:

$$i_c = \frac{\bar{y}_i}{\bar{y}}, \quad i = 1, 2, \dots, 12, \quad (2.2.14)$$

де \bar{y}_i – середнє значення показника за прийнятий проміжок часу (місяць):

$$\bar{y}_i = \frac{\sum y_i}{k}; \quad k \text{ – кількість років; } \bar{y} \text{ – середнє значення показника за весь період:}$$

$$\bar{y} = \frac{\sum \bar{y}_i}{n} = \frac{\sum \sum y_{ij}}{k \cdot n}; \quad n \text{ – кількість місяців } (n = 1, 2, 3, \dots, 12).$$

Отримані в такий спосіб n індексів сезонності утворюють ряд даних, називаний **сезонною хвилею**.

Для складання помісячного прогнозу на наступний рік використовується наступна залежність:

$$\hat{Q}_i = \frac{\hat{Q} \cdot i_c}{100}, \quad i = 1, 2, \dots, 12, \quad (2.2.15)$$

де \hat{Q}_i – очікуваний місячний обсяг показника; \hat{Q} – очікуваний річний обсяг показника; i_c – індекс сезонності; n – кількість періодів ($n = 12$).

Розглянемо застосування даного методу на наступному прикладі.

Приклад 2.2.5. За даними, що характеризують обсяг реалізації продукції хлібозаводами (див. табл. 2.2.9) розрахуємо індекси сезонності, побудуємо сезонну хвилю й прогноз обсягу реалізації продукції на окремі місяці наступного року.

Розв'язання

Індекси сезонності для даних табл. 2.2.9 визначаються за формулою (2.2.14) та подаються в стовпці 8 табл. 2.2.9, причому для даного прикладу \bar{y}_i – середнє значення показника за кожний місяць, стовп. 7, табл. 2.2.9); $k = 1, 2, 3, 4$.

Ряд отриманих індексів сезонності утворює сезонну хвилю, яка може бути зображена графічно.

Побудуємо прогноз обсягу реалізації продукції на окремі місяці наступного року. Індекс сезонності для складання прогнозу використовується в такий спосіб.

Обсяг реалізації хлібобулочних виробів

Рік Місяць	1-й рік	2-й рік	3-й рік	4-й рік	Разом за 4 роки (гр 2+3+4+5)	У середньому за 4 роки (гр 6/4), \bar{y}_i	Індекс сезонності $\left(\frac{\bar{y}_i}{\bar{y}}\right) 100$	Прогноз обсягу реалізації продукції на наступний рік
<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>6</i>	<i>7</i>	<i>8</i>	<i>9</i>
01	5,3	5,4	5,5	6,4	22,6	5,65	74,6	6,1
02	5,4	5,6	5,7	6,7	23,4	5,85	77,2	6,3
03	6,2	6,0	5,9	6,9	25,0	6,25	82,5	6,7
04	6,4	6,6	6,7	7,3	27,0	6,75	89,1	7,3
05	7,0	7,2	7,5	7,7	29,4	7,35	97,0	7,9
06	7,5	7,7	8,0	8,2	31,4	7,85	103,6	8,5
07	8,0	8,1	8,5	8,7	33,3	8,33	110,0	9,0
08	8,5	8,6	8,8	9,1	35,0	8,75	115,5	9,4
09	8,9	9,0	9,2	9,5	36,6	9,15	120,8	9,9
10	8,3	8,5	9,0	9,1	34,9	8,72	115,1	9,4
11	8,0	8,3	8,6	8,4	33,3	8,33	110,0	9,0
12	7,5	7,9	8,3	8,0	31,7	7,93	104,7	8,5
Разом	87,0	88,9	91,7	96,0	363,6	-	1200,0	98,0
У середньому						7,575	100,0	

Нехай на наступний рік об'єднання хлібо заводів припускає реалізувати 98 тис.т. хлібобулочних виробів. Для того, щоб сформувавши план реалізації продукції використовується залежність (2.2.15). Результати отриманих у такий спосіб розрахунків наведено в стовпці 9 таблиці 2.2.9. Таким чином, прогноз обсягу реалізації продукції на наступний рік за кожним місяцем складає відповідно:

6,1; 6,3; 6,7; 7,3; 7,9; 8,5; 9,0; 9,4; 9,9; 9,4; 9,0; 8,5.

Прогнозування сезонних процесів ґрунтується на декомпозиції динамічного ряду. Припускають, що у майбутньому збережеться тенденція і такий же характер коливань. За таких умов прогноз y_{t+l} на будь-який місяць (квартал), визначений методом екстраполяції тренду, коригується індексом сезонності:

$$y_{t+l}^* = I_t \cdot y_{t+l},$$

де l – період упередження,

I_t – індекс сезонності t -го кварталу.

Приклад 2.2.6. Нехай поквартальна динаміка обсягів імпорту пального (тис. т) за два роки ($n = 8; t_1 = 1; t_n = 8$) описується трендом $y_t = 923,7 + 33,8t$, за яким теоретичний обсяг імпорту у восьмому кварталі становить $y_8 = 923,7 + 33,8 \cdot 8 = 1194,1$ тис. т, а у 1-му кварталі наступного року ($l = 1$) передбачається $y_{t+l} = 1194,1 + 33,8 \cdot 1 = 1227,9$.

Якщо середній індекс сезонності 1-го кварталу $I_t = 1,34$, то скоригований на сезонність прогнозний рівень дорівнює $y_{t+l}^* = 1,34 \cdot 1227,9 = 1645,4$ тис. т.

2.2.4 Питання для самоконтролю

1. Що називають екстраполяцією?
2. В чому полягає відмінність між простими та складними методами екстраполяції?
3. Які методи екстраполяції відносять до простих?
4. Розкрийте сутність екстраполяції середнього. Як за таким методом побудувати прогноз на наступні 5 років?
5. В чому полягає відмінність коефіцієнту росту від коефіцієнта приросту?
6. Що розуміється під пунктами росту?
7. В чому полягає сутність методу «екс-пост»? Як за допомогою цього методу перевірити адекватність обраної прогнозної моделі?
8. Розкрийте сутність методу екстраполяції тренду.
9. Що називають сезонною хвилею?
10. Наведіть основні етапи використання методу екстраполяції на основі індексу сезонності.

2.2.5 Рекомендована література



Для самостійного вивчення теми рекомендується література:
основна – [1, 2, 4, 5, 9, 10].

2.3 Адаптивні методи прогнозування

2.3.1 Сутність адаптивних методів прогнозування

Розглянуті у підрозділі 2.2 методи простої екстраполяції часових рядів, в тому числі і методи, що використовують рівняння часового тренду, мають наступні недоліки:

1. Вони ґрунтуються на припущенні збереження тенденцій, що склалися у минулому і прогнозованому періоді, так що значення аналітичних показників динаміки і параметрів рівняння тренду визначаються кількістю періодів «передісторії». Якби ці параметри залишалися незмінними для різних відрізків часу, то перераховані методи у зв'язку з простотою їх використання були б бездоганними.

2. Оскільки для більшості економічних, соціальних, виробничих і т.п. процесів характерні як стабільні, так і стрибкоподібні зміни, то це істотним чином відображується на величині параметрів рівняння, розрахованих для різних періодів, а отже, і на результатах прогнозу.

3. При використанні в прогнозуванні рівняння тренду усі рівні однаковою мірою впливають на параметри рівняння і, таким чином, однаковою мірою визначають і рівень прогнозованих показників.

Вказані недоліки привели до необхідності удосконалення методів прогнозування, що ґрунтуються на використанні одного часового ряду. В результаті для кращого відображення особливостей зміни досліджуваного показника наприкінці періоду спостереження були розроблені так звані *адаптивні методи прогнозування*, кожен з яких має певний механізм пристосування до нових умов.

Суть цих методів полягає в тому, що здійснюється постійна адаптація результатів прогнозів до нової інформації, тобто прогнози стають чутливішими до нових даних. А це значною мірою збільшує точність прогнозів. Одночасно ці методи прогнозування передбачають різну цінність рівнів часового ряду, що є основою для побудови прогнозів.

Таким чином, адаптивні методи прогнозування дозволяють будувати самокориговані моделі, які, враховуючи результат прогнозу, зробленого на попередньому кроці, і різну інформаційну цінність членів динамічного ряду, здатні оперативно реагувати на умови, що змінюються, і на цій основі дати на найближчу перспективу точніші прогнози.

В основі адаптивних методів лежить обчислення послідовних в часі значень прогнозованого показника з урахуванням міри впливу попередніх рівнів. Із віддаленням від кінця часового ряду рівні ряду динаміки все менше впливають на результати прогнозу. Отже, результати прогнозу, головним чином, залежать від рівнів ряду, які щонайближче знаходяться на початку прогнозного періоду, тобто пізніші і віддаленіші в часі спостереження.

До найбільш поширених адаптивних методів прогнозування відносяться наступні:

- *метод ковзної середньої* – передбачає, що наступний у часі показник за своєю величиною дорівнює середній за останні місяці (роки) величині, наприклад середній за останні 3 місяці;

- *метод експоненціального згладжування* – метод, при якому прогноз визначається як сума фактичних і прогнозних даних за період, зважених за допомогою спеціальних коефіцієнтів, які визначаються статистичним шляхом;

- *метод гармонійних ваг* – (близький до методу експоненціального згладжування, використовує той же принцип) – в його основі лежить зважування ковзного показника, але замість ковзної середньої використовується ідея ковзного тренду;

Загальним для усіх методів цієї групи є надання найбільших ваг останнім спостереженням при оцінці параметрів.

Розглянемо кожний із зазначених методів в прогнозуванні виробничих процесів.

2.3.2 Метод ковзної середньої та його застосування

2.3.2.1 Сутність методу

Метод ковзної середньої є одним з найбільш широко відомих методів згладжування часових рядів.

Сутність методу полягає в наступному: згладжування за допомогою ковзною середньої засноване на тому, що в середніх величинах взаємно гасяться випадкові відхилення.

Це погашення відбувається внаслідок наступних **етапів методу**:

1 етап: Перші рівні часового ряду замінюються середньою арифметичною величиною усередині обраного інтервалу часу. Отримане значення відноситься до середини обраного періоду.

2 етап: Період зрушується на одне спостереження й розрахунки середньої повторюється, причому періоди визначення середньої беруться весь час однаковими. Таким чином, у кожному випадку середня є центрованою, тобто віднесеною до серединної точки інтервалу згладжування і являє собою рівень для цієї точки.

Особливістю методу ковзної середньої є те, що рівень показників, який знаходиться як найближче до прогнозованого періоду, впливає на значення прогнозованих показників, у порівнянні з віддаленими періодами.

Перевага методу ковзної середньої полягає в тому, що на значення прогнозованих показників впливають тією чи іншою мірою всі дані «передісторії», у той час як в інших розглянутих методах (наприклад, у методі екстраполяції на основі аналітичних показників динаміки) значення коефіцієнтів, на основі яких будується прогноз, визначаються тільки крайніми величинами часового ряду.

До методу ковзної середньої можна зробити наступні зауваження:

Зауваження 1: При згладжуванні часового ряду ковзними середніми в розрахунках беруть участь усі рівні ряду.

Зауваження 2: Чим ширше інтервал ковзання, тим більш плавним буде тренд.

Зауваження 3: Згладжений ряд коротше первісного на $(k - 1)$ спостережень, де k – величина інтервалу згладжування. При великих значеннях k коливання згладженого ряду значно знижуються. Одночасно помітно скорочується число спостережень, що створює чималі труднощі.

Зауваження 4: Вибір інтервалу згладжування залежить від мети дослідження. При цьому слід керуватися тим, у який саме період часу нас цікавить дія, а отже, і усунення впливу випадкових факторів. Наприклад, для згладжування часового ряду продуктивності праці, планування якої розраховано на 5-річний період, доцільно обрати 5-річний період згладжування.

Зауваження 5: З вибором інтервалу згладжування пов'язане питання про кількість рівнів ряду, які задіяні у розрахунках ковзної середньої, та про техніку розрахунків.

Розглянемо 2 варіанти розрахунків величини ковзної середньої:

а) якщо число членів інтервалу згладжування є непарним, то отримані значення ковзної середньої припадають на середній член інтервалу ковзання, тобто ковзна середня буде обчислюватися за формулою

$$\hat{y}_{i+m} = \frac{y_i + y_{i+1} + \dots + y_{i+2m}}{2m+1} = \frac{\sum_{k=0}^{2m} y_{i+k}}{2m+1};$$

б) при парній кількості рівнів отримані значення ковзної середньої не можна віднести до жодного рівню ряду – ковзні середні будуть розташовуватися в проміжках між рівнями, тобто

$$\hat{y}_{i+m} = \frac{\frac{1}{2}y_i + y_{i+1} + \dots + \frac{1}{2}y_{i+2m}}{2m} = \frac{\frac{1}{2}y_i + \sum_{k=0}^{2m-1} y_{i+k} + \frac{1}{2}y_{i+2m}}{2m},$$

де прийняті наступні позначення: \hat{y}_{i+m} – величина $(i+m)$ -ї ковзної середньої; y_i – i -й рівень часового ряду ($i=1, 2, \dots, n-2m$); m – задане ціле позитивне число, за допомогою якого визначається величина інтервалу згладжування:

$$1 \leq m \leq \left[\frac{n}{2} \right] = \begin{cases} \frac{n}{2}, & \text{якщо } n \text{ – парне число;} \\ \frac{n-1}{2}, & \text{якщо } n \text{ – непарне число;} \end{cases}$$

n – число рівнів часового ряду; k – змінний індекс для інтервалу згладжування ($k=0, 1, 2, \dots, 2m$).

2.3.2.2 Застосування методу ковзної середньої до побудови прогнозу за допомогою адитивної та мультиплікативної моделей часового ряду

Перш ніж перейти до розгляду практичного застосування методу ковзної середньої до прогнозування виробничих процесів наведемо основні компоненти часового ряду та надамо характеристику адитивної та мультиплікативної моделей часового ряду.

У загальному випадку часовий ряд економічного показника можна розкласти на чотири структурно утворюючі елементи (компоненти): тренд (T), сезонні коливання (S), циклічні коливання (C), випадкова компонента (U). При цьому, деякі компоненти ряду можуть бути відсутні у тих чи інших часових рядах.

Розглянемо окремо вплив кожного фактору на часовий ряд.

Тренд відображає вплив на економічний показник деяких постійних чинників, дія яких акумулюється в часі. У широкому сенсі під **трендом** розуміють будь-який упорядкований процес, що відрізняється від випадкового. Відносно функції тренду припускається, що вона є певною гладкою функцією, степінь гладкості якої заздалегідь не відомий.

Під **степенем гладкості** розуміють мінімальну степінь поліному, який найкраще згладжує компоненту тренду. На рис. 2.3.1 зображено умовний часовий ряд із тенденцією, що лінійно зростає.

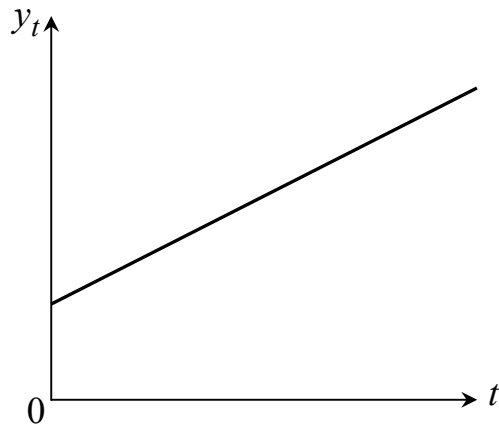


Рисунок 2.3.1

Під **сезонними коливаннями** розуміють стійкі зміни значень ряду, пов'язані з календарним значенням періоду спостережень. Такі коливання обумовлені дією деякого механізму, який періодично повторюється у певний період року, та, отже, мають періодичний або близький до нього характер упродовж одного року. Наприклад, ціни на сільгосппродукцію влітку вищі, ніж взимку; рівень безробіття в курортних містах у зимовий період зростає відносно до літнього; у четвертому кварталі кожного року перед Новим роком значно зростає споживання товарів. Сезонні коливання можуть виникати з причин, пов'язаних з діяльністю людини (свята, відпустки, релігійні традиції тощо). На рис. 2.3.2 зображено умовний часовий ряд, який містить лише сезонну компоненту.

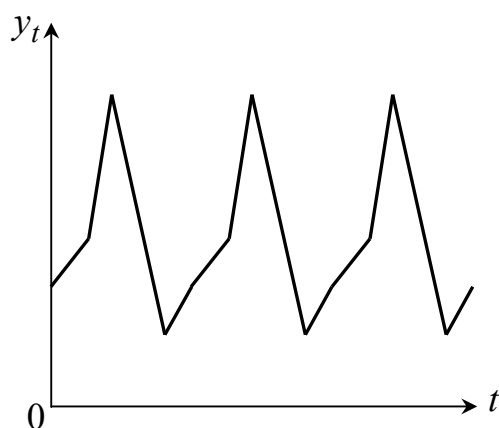


Рисунок 2.3.2

Під **циклічними коливаннями** розуміють стійкі зміни значень ряду, які відбуваються через однакові проміжки часу та не пов'язані з будь-яким календарним періодом спостережень. Такі коливання схожі на сезонні, але виявляються на триваліших інтервалах часу. Характерним прикладом служать цикли ділової активності, демографічні та астрофізичні цикли.

Наприклад, за багаторічними спостереженнями активність сонця має циклічність у 10,5-11 років, причому сплески сонячної радіації впливають на врожайність зернових культур, репродуктивну властивість тварин тощо. Отже за наявності циклічних коливань у часовому ряді динаміка економічного показника містить характерні зміни, що повторюються з однаковою циклічністю.

Тренд, сезонна і циклічна компоненти часового ряду об'єднують у **детерміновану (систематичну) компоненту** та розглядають модель ряду у вигляді $y_t = f(t) + u_t, \forall t$.

Випадкові чинники являють собою залишкові коливання ряду біля середнього значення, тобто залишкову послідовність даних, яка залишається після виділення з часового ряду компонент тренду, сезонних та циклічних коливань. Такі чинники не підлягають вимірюванню, але неминуче супроводжують будь-який економічний процес і визначають стохастичний характер його елементів. До випадкових чинників можна віднести помилки вимірювання, випадкові збурення тощо. На рис. 2.3.3 зображено умовний часовий ряд, який містить лише випадкову компоненту.

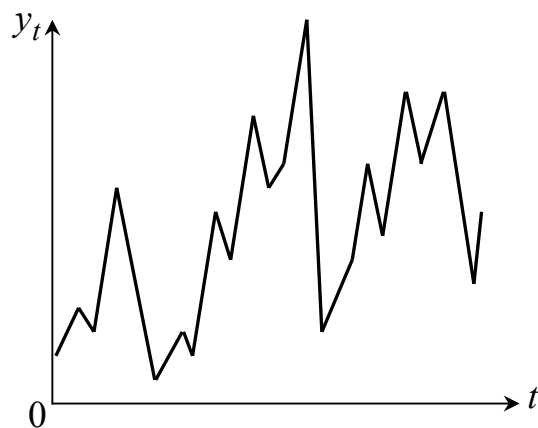


Рисунок 2.3.3

В залежності від взаємозв'язку компонент ряду можуть бути побудовані:

– **адитивна модель часового ряду** ($y_t = T + C + S + U$) – характеризується тим, що характер циклічних і сезонних флуктацій (коливань) залишається постійним, а спостережувані значення в середньому збільшуються або зменшуються за період на однакову величину;

– **мультиплікативна модель часового ряду** ($y_t = T \cdot C \cdot S \cdot U$) – характеризується тим, що характер циклічних і сезонних флуктацій (коливань) залишається постійним тільки стосовно тренду, при цьому спостережувані значення збільшуються або зменшуються за період на певний відсоток.

Побудова адитивної або мультиплікативної моделей часового ряду є найпростішим підходом до моделювання часових рядів, що містять сезонні коливання, тобто коли часовий ряд не містить циклічних коливань та, отже, розглядається модель ряду у вигляді $y_t = T + S + U$ (адитивна модель) або $y_t = T \cdot S \cdot U$ (мультиплікативна модель). У зв'язку з вище зазначеним розглянемо більш детально сезонну компоненту часового ряду.

При вивченні багатьох виробничих та соціально-економічних явищ і процесів часто виявляються певні коливання, які повторюються. Цим коливанням властиві більш-менш стійкі зміни рівнів ряду протягом досліджуваного періоду: рік у рік у певні місяці рівень явища підвищується, а в інші – знижується. Такі коливання особливо помітні в явищах сезонного характеру та є результатом впливу соціальних і природно-кліматичних причин, загальних економічних факторів, а також численних і різноманітних факторів, які часто є регульованими. Зазначені коливання прийнято називати **сезонними**.

Сезонні коливання найчастіше за все зустрічаються у добувних і переробних галузях – сільському господарстві, рибній й лісовій промисловості, а також у торгівлі, транспорті, будівництві, туризмі і т.д. Погодні зміни впливають на побутове споживання палива й електроенергії, на асортименти взуття, верхнього одягу (зимова, весняно-осіння, літня), фруктів, овочів і багатьох інших товарів. У будівництві найбільше погравлення діяльності проявляється влітку; у цей же період року спостерігається максимальний наплив туристів. Сезонність може проявлятися не тільки у місячних, але й у денних, тижневих даних. Так, кафе, ресторани, театри відчувають збільшення попиту наприкінці тижня.

Сезонність проявляється при повному або майже повному припиненні виробництва на якийсь проміжок часу, обумовлений самою природою продукту й способом його виготовлення. Дозрівання зернових, наприклад, вимагає кілька місяців, а в садівництві після посадки саджанців проходить кілька років до одержання готового продукту.

Отже, викликані різними причинами, сезонні коливання як у виробництві, так й в обігу, звичайно негативно впливають на результати виробничої діяльності через те, що викликають порушення ритмічності виробництва, обумовлюють нерівномірність використання трудових ресурсів та устаткування впродовж року і т.д. Багато галузей економіки взаємозалежні, тому проблема сезонності – загальна проблема економіки різних країн. Нерівномірність виробництва того або іншого продукту веде до нерівномірності його споживання, споживання ж, у свою чергу, впливає на виробництво.

Вплив сезонних коливань повністю усунути неможливо, але деякі підприємства намагаються його знизити, приймаючи заходи раціональної комбінації галузей, механізації трудомістких процесів і т.д. От із цієї причини сезонні коливання, відображені в рядах динаміки, необхідно вивчати й вимірювати.

Розглянемо процес моделювання часових рядів, що містять сезонні коливання, за допомогою побудови адитивної та мультиплікативної моделей часового ряду.

Вибір однієї із цих моделей ґрунтується на аналізі структури часового ряду:

- якщо амплітуда сезонних коливань приблизно постійна, то будують адитивну модель;
- якщо амплітуда сезонних коливань непостійна, тобто зростає або зменшується, то будують мультиплікативну модель.

Для цього необхідно представити часовий ряд графічно та проводити аналіз структури ряду за його виглядом або проводити аналіз за безпосередньо представленими даними в вихідному ряді. При цьому слід враховувати, що порівняння даних слід проводити за відповідні періоди, а не за минулі періоди, тобто, наприклад, дані за грудень одного року треба порівнювати з даними грудня минулих років, а не з листопадом.

Процес побудови відповідної моделі ряду включає **наступні етапи**:

- 1 етап:** Вирівнювання вихідного ряду методом ковзної середньої. Розрахунки значень сезонної компоненти S .
- 2 етап:** Усунення сезонної компоненти з вихідних рівнів ряду й одержання вирівняних даних $(T+U)$ в адитивній або $(T \cdot U)$ в мультиплікативній моделі.
- 3 етап:** Аналітичне вирівнювання рівнів $(T+U)$ або $(T \cdot U)$ й розрахунки значень T з використанням отриманого рівняння тренду.
- 4 етап:** Розрахунки отриманих за моделлю значень $(T+S)$ або $(T \cdot S)$.
- 5 етап:** Розрахунки абсолютних і/або відносних помилок. Оцінка точності моделі часового ряду.
- 6 етап:** Складання прогнозу на наступний прогнозний період.

Розглянемо процес побудови адитивної моделі на наступному прикладі.

Приклад 2.3.1. Нехай надані дані про кількість продукції (тис. шт.) (стовп. 2, табл. 2.3.1), проданою фірмою протягом останніх 20 кварталів (стовп. 1, табл. 2.3.1). Необхідно побудувати адитивну модель часового ряду та скласти поквартальний прогноз на наступний рік.

Розв'язання

Розв'язання поставленої задачі будемо проводити поетапно.

1 етап: Проведемо вирівнювання ряду методом ковзної середньої.

Для цього:

- просумуємо рівні ряду за 4 квартали послідовно (стовп. 3, табл. 2.3.1);
- розділимо отримані суми на 4 і знайдемо ковзні середні, які вже не містять сезонної компоненти (стовп. 4, табл. 2.3.1);
- знайдемо центровані ковзні середні, для чого обчислимо середні значення з двох послідовних ковзних середніх (стовп. 5, табл. 2.3.1);
- обчислимо оцінки сезонної компоненти як різниці між фактичним рівнем продаж і центрованими ковзними середніми (стовп. 6, табл. 2.3.1).

Розрахунки оцінок сезонної компоненти

	Квартал	Обсяг продаж, тис. шт.	Разом за 4 квартали	Ковзна середня за 4 квартали	Центрована ковзна середня	Оцінка сезонної компоненти
	1	2	3	4	5	6
1 рік	1	8,4				
	2	8,6	35,3	8,825		
	3	8,8	35,4	8,85	8,8375	-0,0375
	4	9,5	35,9	8,975	8,9125	0,5875
2 рік	5	8,5	36,3	9,075	9,025	-0,525
	6	9,1	36,7	9,175	9,125	-0,025
	7	9,2	37,9	9,475	9,325	-0,125
	8	9,9	38,7	9,675	9,575	0,325
3 рік	9	9,7	39,6	9,9	9,7875	-0,0875
	10	9,9	40,5	10,125	10,0125	-0,1125
	11	10,1	41,3	10,325	10,225	-0,125
	12	10,8	42,1	10,525	10,425	0,375
4 рік	13	10,5	43	10,75	10,6375	-0,1275
	14	10,7	44,4	11,1	10,925	-0,225
	15	11	45,8	11,45	11,275	-0,275
	16	12,2	47,4	11,85	11,65	0,55
5 рік	17	11,9	48,9	12,225	12,0375	-0,1375
	18	12,3	49,9	12,475	12,35	-0,05
	19	12,5				
	20	13,2				

Використовуємо отримані оцінки сезонної компоненти (стовп. 6, табл. 2.3.1) для розрахунків сезонності S у табл. 2.3.2.

Таблиця 2.3.2

Коригування значень сезонної компоненти

Показники	рік	Квартал			
		Q_1	Q_2	Q_3	Q_4
	1	—	—	-0,0375	0,5875
2	-0,525	-0,025	-0,125	0,375	
3	-0,0875	-0,1125	-0,125	0,375	
4	-0,1375	-0,225	-0,275	0,55	
5	-0,1375	-0,05	—	—	
Разом за квартал		-0,8875	-0,4125	-0,5625	1,8375
Середня оцінка сезонної компоненти для кварталу, \bar{S}		-0,2218	-0,1031	-0,1406	0,4593
Скоригована оцінка сезонної компоненти, S_i		-0,2203	-0,1015	-0,1390	0,4609

Для цього:

– знайдемо середні квартальні оцінки сезонної компоненти, використавши дані всіх кварталів (табл. 2.3.2);

– сума значень сезонної компоненти за всіма кварталами повинна дорівнювати нулю, тобто $\sum S_i = 0$. Тому значення сезонної компоненти коригуються на величину, отриману як частку від поділу суми оцінок сезонної компоненти на число сезонів;

– розрахуємо коригувальний коефіцієнт

$$k = \frac{\sum_{i=1}^4 \bar{S}_i}{4} = -0,00156;$$

– знайдемо скориговані оцінки сезонної компоненти, які визначаються шляхом віднімання коригувального коефіцієнта з середньої оцінки сезонної компоненти для кварталу (рядок 4, табл. 2.3.2), тобто у вигляді $S_i = \bar{S}_i - k$:

$$S_1 = (-0,2218) - (-0,00156) = -0,2203,$$

$$S_2 = (-0,1031) - (-0,00156) = -0,1015,$$

$$S_3 = (-0,1406) - (-0,00156) = -0,1390,$$

$$S_4 = 0,4593 - (-0,00156) = 0,4609;$$

– перевіримо виконання властивості, за якою $\sum S_i = 0$:

$$\sum_{i=1}^4 S_i = (-0,2203) + (-0,1015) + (-0,1390) + 0,4609 = 0.$$

2 етап: Усунемо сезонну компоненту з вихідних рівнів ряду й одержимо вирівняні дані $T + U = y_i - S$ (стовп. 5, табл. 2.3.3).

Зазначимо, що в останньому рядку таблиці 2.3.3 наведені суми за кожним стовпцем таблиці.

3 етап: Визначимо компоненту тренду T .

Для цього проведемо аналітичне вирівнювання ряду $(T + U)$ за допомогою лінійного тренду $T = a_0 + a_1 t$. Для визначення коефіцієнтів рівняння тренду скористаймося методом найменших квадратів (МНК). У табл. 2.3.3 отримані величини

$$\begin{aligned} \sum t &= 210, & \sum (T + U) &= 206,7994, \\ \sum t^2 &= 2870, & \sum (T + U) t &= 2324,378, \end{aligned} \quad n = 20,$$

за допомогою яких складаємо систему нормальних рівнянь

$$\begin{cases} \sum (T + U) = a_0 n + a_1 \sum t \\ \sum (T + U) t = a_0 \sum t + a_1 \sum t^2 \end{cases} \Rightarrow \begin{cases} 206,7994 = 20 a_0 + 210 a_1, \\ 2324,378 = 210 a_0 + 2870 a_1, \end{cases}$$

розв'язання якої дозволяє отримати невідомі коефіцієнти

$$\begin{cases} a_0 = 7,9244 \\ a_1 = 0,2301 \end{cases}$$

Отже, маємо лінійний тренд виду $T = 7,9244 + 0,2301 \cdot t$.

Підставляючи в отримане рівняння тренду послідовно значення $t = 1, \dots, 20$, одержимо значення тренду для кожного рівня часового ряду (стовп. 7, табл. 2.3.3).

Таблиця 2.3.3

**Розрахунки вирівняних значень T та помилок U
в адитивній моделі**

t	t^2	y_i	S_i	$T + U$	$(T+U)t$	T	$T + S$	U	U^2	$(y_i - \bar{y})^2$
1	2	3	4	5	6	7	8	9	10	11
1	1	8,4	-0,22	8,62	8,62	8,15	7,93	0,47	0,22	3,76
2	4	8,6	-0,10	8,70	17,40	8,38	8,28	0,32	0,10	3,03
3	9	8,8	-0,14	8,94	26,82	8,61	8,48	0,32	0,11	2,37
4	16	9,5	0,46	9,04	36,16	8,84	9,31	0,19	0,04	0,71
5	25	8,5	-0,22	8,72	43,60	9,07	8,85	-0,35	0,13	3,39
6	36	9,1	-0,10	9,20	55,21	9,31	9,20	-0,10	0,01	1,54
7	49	9,2	-0,14	9,34	65,37	9,54	9,40	-0,20	0,04	1,30
8	64	9,9	0,46	9,44	75,51	9,77	10,23	-0,33	0,11	0,19
9	81	9,7	-0,22	9,92	89,28	10,00	9,78	-0,08	0,01	0,41
10	100	9,9	-0,10	10,00	100,02	10,23	10,12	-0,22	0,05	0,19
11	121	10,1	-0,14	10,24	112,63	10,46	10,32	-0,22	0,05	0,06
12	144	10,8	0,46	10,34	124,07	10,69	11,15	-0,35	0,12	0,21
13	169	10,5	-0,22	10,72	139,36	10,92	10,70	-0,20	0,04	0,03
14	196	10,7	-0,10	10,80	151,22	11,15	11,04	-0,34	0,12	0,13
15	225	11	-0,14	11,14	167,09	11,38	11,24	-0,24	0,06	0,44
16	256	12,2	0,46	11,74	187,83	11,61	12,07	0,13	0,02	3,46
17	289	11,9	-0,22	12,12	206,05	11,84	11,62	0,28	0,08	2,43
18	324	12,3	-0,10	12,40	223,23	12,07	11,96	0,34	0,11	3,84
19	361	12,5	-0,14	12,64	240,14	12,30	12,16	0,34	0,12	4,67
20	400	13,2	0,46	12,74	254,78	12,53	12,99	0,21	0,05	8,18
210	2870	206,8	0,0	206,80	2324,4	206,8	206,8	-0,01	1,55	40,33

4 етап: Знайдемо значення рівнів ряду, отримані за адитивною моделлю як $(T + S)$ (стовп. 8, табл. 2.3.3).

5 етап: Розрахуємо абсолютну помилку й перевіримо якість отриманої моделі:

- знайдемо абсолютну помилку як $U = y_i - (T + S)$ (стовп. 9, табл. 2.3.3);
- перевіримо якість отриманої моделі, використовуючи дані квадратів абсолютних помилок (стовп. 10, табл. 2.3.3). Для цього знайдемо суму

квадратів абсолютних помилок $\sum U^2 = 1,55$ та суму квадратів відхилень вихідних рівнів ряду від його середнього рівня (стовп. 11, табл. 2.3.3):

$$\sum (y_i - \bar{y})^2 = 40,33,$$

де $\bar{y} = \frac{\sum y}{n} = 10,34$.

По відношенню до суми квадратів відхилень вихідних рівнів ряду від його середнього рівня $\sum (y_i - \bar{y})^2$ величина $\sum U^2$ складе 3,84 %, виходячи зі співвідношення:

$$\begin{array}{rcl} 40,33 & - & 100 \% \\ 1,55 & - & x \end{array} \Rightarrow x = \frac{1,55 \cdot 100 \%}{40,33} \approx 3,84 \%$$

Отже, адитивна модель пояснює $100 \% - 3,84 \% = 96,16 \%$ загальної варіації обсягу продажів за 20 кварталів, тобто точність моделі складає приблизно 96 %.

6 етап: Цей етап полягає в складанні прогнозу для кожного з кварталів майбутнього (в цьому випадку 6-го) року.

Додаємо останнє центроване ковзне середнє за квартал (стовп. 5, табл. 2.3.1) до його скоригованої оцінки сезонної компоненти (рядок 4, табл. 2.3.2). Вертаючись до таблиці 2.3.1 (стовп. 5) ми бачимо, що останнє центроване ковзне середнє відноситься до II кварталу 5 року, наступне – до I кварталу 5 року, потім – до IV кварталу 4 року, і, нарешті, – до III кварталу 4 року.

У результаті одержимо поквартальний прогноз на 6 рік:

$$\left\{ \begin{array}{l} Q_1: 12,0375 + (-0,2203) = 11,8172, \text{ т.е. } 11\,817 \text{ шт.} - \text{I кв.} \\ Q_2: 12,35 + (-0,1015) = 12,2485, \text{ т.е. } 12\,249 \text{ шт.} - \text{II кв.} \\ Q_3: 11,275 + (-0,1390) = 11,1360, \text{ т.е. } 11\,136 \text{ шт.} - \text{III кв.} \\ Q_4: 11,65 + 0,4609 = 12,1111, \text{ т.е. } 12\,111 \text{ шт.} - \text{IV кв.} \end{array} \right.$$

Розглянемо побудову мультиплікативної моделі на наступному прикладі.

Приклад 2.3.2. Нехай надані поквартальні дані про обсяг експорту однієї з областей України за 5 років (млн. дол.) (стовп. 2, табл. 2.3.4). Необхідно побудувати мультиплікативну модель часового ряду та скласти поквартальний прогноз на наступний рік.

Розв'язання

Розв'язання поставленої задачі будемо проводити поетапно.

Розрахунки оцінок сезонної компоненти

	Квартал	Обсяг експорту млн. дол.	Разом за 4 квартали	Ковзна середня за 4 квартали	Центрована ковзна середня	Оцінка сезонної компоненти
	1	2	3	4	5	6
1 рік	1	19,3				
	2	12,3	60,4	15,1		
	3	13,2	62,6	15,65	15,375	0,858537
	4	15,6	66,1	16,525	16,0875	0,969697
2 рік	5	21,5	70,1	17,525	17,025	1,262849
	6	15,8	74,4	18,6	18,0625	0,87474
	7	17,2	79,2	19,8	19,2	0,895833
	8	19,9	82,5	20,625	20,2125	0,984539
3 рік	9	26,3	85,6	21,4	21,0125	1,251636
	10	19,1	88,0	22,0	21,7	0,880184
	11	20,3	91,4	22,85	22,425	0,90524
	12	22,3	93,6	23,35	23,1	0,965368
4 рік	13	29,7	96,8	24,2	23,775	0,249211
	14	21,1	99,9	24,975	24,5875	0,85816
	15	23,7	102,0	25,5	25,2375	0,939079
	16	25,4	104,8	26,2	25,85	0,982592
5 рік	17	31,8	106,9	26,725	26,4625	1,201701
	18	23,9	108,9	27,225	26,975	0,886006
	19	25,8				
	20	27,4				

1 етап: Проведемо вирівнювання ряду методом ковзної середньої.

Для цього:

- просумуємо рівні ряду за 4 квартали послідовно (стовп. 3, табл. 2.3.4);
- розділимо отримані суми на 4 і знайдемо ковзні середні, які вже не містять сезонної компоненти (стовп. 4, табл. 2.3.4);
- знайдемо центровані ковзні середні, для чого обчислимо середні значення з двох послідовних ковзних середніх (стовп. 5, табл. 2.3.4);
- обчислимо оцінки сезонної компоненти як частку від поділу фактичного рівня експорту на центровані ковзні середні (стовп. 6, табл. 2.3.4).

Використовуємо отримані оцінки сезонної компоненти (стовп. 6, табл. 2.3.4) для розрахунків сезонності S у табл. 2.3.5.

Коригування значень сезонної компоненти

Показники	рік	Квартал			
		Q_1	Q_2	Q_3	Q_4
	1	—	—	0,8585	0,9696
	2	1,2628	0,8747	0,8958	0,9845
	3	1,2516	0,8801	0,9052	0,9653
	4	1,2492	0,8581	0,9390	0,9825
	5	1,2017	0,8860	—	—
Разом за квартал		4,9653	3,4989	3,5985	3,9019
Середня оцінка сезонної компоненти для кварталу, \bar{S}		1,2413	0,8747	0,8996	0,9755
Скоректована оцінка сезонної компоненти, S_i		1,2441	0,8767	0,9016	0,9776

Для цього:

– знайдемо середні квартальні оцінки сезонної компоненти, використавши дані всіх кварталів (рядок 3, табл. 2.3.5);

– сума значень сезонної компоненти за всіма кварталами повинна дорівнювати числу періодів у циклі. У нашому прикладі, цикл – 1 рік, у якому відповідно 4 квартали. Отже, $\sum S_i = 4$.

– тому остаточний варіант сезонної компоненти буде отриманий коригуванням, яке полягає у поділі на коефіцієнт k середньої оцінки сезонної компоненти для кварталу:

$$k = \frac{\sum_{i=1}^4 \bar{S}_i}{4} = 0,9977875.$$

Тоді сезонна компонента розраховується у вигляді $S_i = \frac{\bar{S}_i}{k}$:

$$S_1 = \frac{1,2413}{0,9978} = 1,2441, \quad S_2 = \frac{0,8747}{0,9978} = 0,8767,$$

$$S_3 = \frac{0,8996}{0,9978} = 0,9016, \quad S_4 = \frac{0,9755}{0,9978} = 0,9776.$$

Отримані в такий спосіб значення занесемо до таблиці 2.3.5.

Перевіримо виконання властивості, за якою $\sum S_i = 4$:

$$\sum_{i=1}^4 S_i = 1,2441 + 0,8767 + 0,9016 + 0,9776 = 4.$$

2 етап: Усунемо сезонний компонент з вихідних рівнів ряду й одержимо вирівняні дані у вигляді $T \cdot U = \frac{y_i}{S}$ (стовп. 5, табл. 2.3.6).

Таблиця 2.3.6

**Розрахунки вирівняних значень T та помилок U
у мультиплікативній моделі**

t	t^2	y_i	S_i	$T \cdot U$	$(T \cdot U)t$	T	$T \cdot S$	U	U^2	$(y_i - \bar{y})^2$
1	2	3	4	5	6	7	8	9	10	11
1	1	19,3	1,24	15,51	15,51	14,30	17,79	1,09	1,18	5,20
2	4	12,3	0,88	14,03	28,06	15,07	13,21	0,93	0,87	86,12
3	9	13,2	0,90	14,64	43,92	15,84	14,28	0,92	0,85	70,22
4	16	15,6	0,98	15,96	63,83	16,62	16,24	0,96	0,92	35,76
5	25	21,5	1,24	17,28	86,41	17,39	21,63	0,99	0,99	0,01
6	36	15,8	0,88	18,02	108,15	18,16	15,92	0,99	0,98	33,41
7	49	17,2	0,90	19,08	133,54	18,94	17,07	1,01	1,02	19,18
8	64	19,9	0,98	20,36	162,85	19,71	19,27	1,03	1,07	2,82
9	81	26,3	1,24	21,14	190,27	20,48	25,48	1,03	1,07	22,28
10	100	19,1	0,88	21,79	217,89	21,25	18,63	1,03	1,05	6,15
11	121	20,3	0,90	22,52	247,67	22,03	19,86	1,02	1,04	1,64
12	144	22,3	0,98	22,81	273,73	22,80	22,29	1,00	1,00	0,52
13	169	29,7	1,24	23,87	310,37	23,57	29,33	1,01	1,03	65,93
14	196	21,1	0,88	24,07	336,98	24,35	21,34	0,99	0,98	0,23
15	225	23,7	0,90	26,29	394,30	25,12	22,65	1,05	1,10	4,49
16	256	25,4	0,98	25,98	415,71	25,89	25,31	1,00	1,01	14,59
17	289	31,8	1,24	25,56	434,57	26,67	33,17	0,96	0,92	104,45
18	324	23,9	0,88	27,26	490,76	27,44	24,05	0,99	0,99	5,38
19	361	25,8	0,90	28,62	543,70	28,21	25,44	1,01	1,03	17,81
20	400	27,4	0,98	28,03	560,56	28,99	28,34	0,97	0,93	33,87
210	2870	431,6	20,0	432,82	5058,8	432,82	431,3	19,99	20,01	530,07

3 етап: Визначимо компоненту T .

Для цього проведемо аналітичне вирівнювання ряду $(T \cdot U)$ за допомогою лінійного тренду $T = a_0 + a_1 t$. Для визначення коефіцієнтів рівняння тренду скористаймося методом найменших квадратів (МНК). У табл. 2.3.6 отримані величини

$$\begin{aligned} \sum t &= 210, & \sum (T \cdot U) &= 432,8224, & n &= 20, \\ \sum t^2 &= 2870, & \sum (T \cdot U)t &= 5058,783, & & \end{aligned}$$

за допомогою яких складаємо систему нормальних рівнянь

$$\begin{cases} \sum (T \cdot U) = a_0 n + a_1 \sum t \\ \sum (T \cdot U)t = a_0 \sum t + a_1 \sum t^2 \end{cases} \Rightarrow \begin{cases} 432,8224 = 20 a_0 + 210 a_1, \\ 5058,783 = 210 a_0 + 2870 a_1, \end{cases}$$

розв'язання якої дозволяє отримати невідомі коефіцієнти

$$\begin{cases} a_0 = 13,5230, \\ a_1 = 0,77315. \end{cases}$$

Отже, маємо лінійний тренд у вигляді $T = 13,5230 + 0,77315 \cdot t$

Підставляючи в отримане рівняння тренду послідовно значення $t = 1, \dots, 20$, одержимо значення тренду для кожного рівня часового ряду (стовп. 7, табл. 2.3.6).

По відношенню до суми квадратів відхилень вихідних рівнів ряду від його середнього рівня $\sum (y_i - \bar{y})^2$ величина $\sum U^2$ складе 3,78 %, виходячи зі співвідношення:

$$\frac{530,072}{21,03378} - 100 \% \Rightarrow x = \frac{20,0122 \cdot 100 \%}{530,072} \approx 3,78 \%$$

Отже, побудована мультиплікативна модель пояснює $100 \% - 3,78 \% = 96,22 \%$ загальної варіації експорту, тобто точність моделі складає приблизно 96 %.

4 етап: Цей етап полягає в складанні прогнозу для кожного із кварталів майбутнього, у цьому випадку 6 року.

Для цього помножимо останнє центроване ковзне середнє за квартал (стовп. 5, табл. 2.3.4) на його скориговану оцінку сезонної компоненти (рядок 4, табл. 2.3.5).

Вертаючись до таблиці 2.3.4 (стовп. 5) ми бачимо, що останнє центроване ковзне середнє відноситься до II кварталу 5 року, наступне – до I кварталу 5 року, потім – до IV кварталу 4 року, і, нарешті, – до III кварталу 4 року.

У результаті одержимо поквартальний прогноз на 6 рік:

$$\begin{cases} Q_1: 26,4625 \cdot 1,2441 = 32,922, \text{ т.е. } \$ 32,922 \text{ млн.} - \text{I кв.} \\ Q_2: 26,975 \cdot 0,8767 = 23,649, \text{ т.е. } \$ 23,649 \text{ млн.} - \text{II кв.} \\ Q_3: 25,2375 \cdot 0,9016 = 22,754, \text{ т.е. } \$ 22,754 \text{ млн.} - \text{III кв.} \\ Q_4: 25,85 \cdot 0,9776 = 25,271, \text{ т.е. } \$ 25,271 \text{ млн.} - \text{IV кв.} \end{cases}$$

2.3.3 Метод експоненціального згладжування

2.3.3.1 Сутність методу

Експоненціальне згладжування є узагальненням методу ковзної середньої.

Сутність цього методу полягає в тому, що кожний рівень часового ряду згладжується за допомогою зваженої ковзної середньої, причому вага кожного спостереження зменшується в міру його віддалення від моменту, для якого визначається згладжене значення, тобто від кінця ряду.

Особливість методу полягає в тому, що в процедурі вирівнювання кожного спостереження використовуються тільки значення попередніх рівнів ряду динаміки, узятих з певною вагою.

Згладжене значення рівня ряду S_t на момент t визначається за формулою

$$S_t = \alpha y_t + (1 - \alpha)S_{t-1}, \quad (2.3.1)$$

де S_t, S_{t-1} – значення експоненціальної середньої відповідно в моменти часу t і $(t-1)$; y_t – значення економічного процесу в момент t ; α – вага (**параметр згладжування**) t -го значення ряду динаміки, причому $\alpha = const$ та $0 < \alpha < 1$.

З формули (2.3.1) видно, що при обчисленні експоненціальної середньої S_t використовується лише попередня експоненціальна середня S_{t-1} й останнє спостереження y_t , а всі попередні рівні ряду «забуваються».

Якщо S_{t-1} розглядати як прогноз на 1 крок уперед, то величина $(y_t - S_{t-1})$ є погрішність цього прогнозу, а новий прогноз S_t виходить у результаті коректування попереднього прогнозу з урахуванням його помилки:

$$S_t = \alpha y_t + (1 - \alpha)S_{t-1} = \alpha y_t + S_{t-1} - \alpha S_{t-1} = S_{t-1} + \alpha(y_t - S_{t-1}).$$

У цьому й полягає **суть процесу адаптації моделі** до вихідних даних.

Послідовне застосування формули (2.3.1) дає можливість обчислити експоненціальну середню через значення всіх рівнів даного ряду динаміки.

Так, величину S_{t-1} у формулі (2.3.1) можна представити у вигляді суми фактичного значення рівня y_{t-1} й згладженого значення попереднього йому спостереження S_{t-2} , узятих з відповідними вагами, тобто

$$S_{t-1} = \alpha y_{t-1} + (1 - \alpha)S_{t-2}.$$

Процес такого розкладання можна продовжити для членів ряду S_{t-2}, S_{t-3} і т.д.

У результаті одержимо вираження, у яким середнє згладжене значення є комбінацією всіх попередніх рівнів ряду, тобто

$$\begin{aligned} S_t &= \alpha y_t + (1 - \alpha)S_{t-1} = \alpha y_t + (1 - \alpha) \cdot [\alpha y_{t-1} + (1 - \alpha)S_{t-2}] = \\ &= \alpha y_t + \alpha(1 - \alpha)y_{t-1} + (1 - \alpha)^2 \cdot [\alpha y_{t-2} + (1 - \alpha)S_{t-3}] = \\ &= \alpha y_t + \alpha(1 - \alpha)y_{t-1} + \alpha(1 - \alpha)^2 y_{t-2} + \dots + \alpha(1 - \alpha)^k y_{t-k} + \dots + (1 - \alpha)^t y_0, \end{aligned} \quad (2.3.2)$$

де $t = \overline{2, k}$; y_0 – характеризує початкові умови процесу.

Склавши у формулі (2.3.2) усі члени, що містять параметр α , одержимо:

$$S_t = \alpha \sum_{j=0}^{t-1} (1 - \alpha)^j y_{t-j} + (1 - \alpha)^t y_0, \quad t = \overline{2, k}, \quad (2.3.3)$$

де j – число періодів відставання від моменту t .

Згідно з формулою (2.3.3) відносна вага кожного попереднього рівня знижується за експонентою* у міру його віддалення від моменту, для якого обчислюється згладжене значення, звідси й походить назва даного методу.

* За формулою Тейлора маємо: $(1 - \alpha) \approx e^{-\alpha}$.

Оскільки $(1-\alpha) < 1$, то при $t \rightarrow \infty$ маємо:

$$(1-\alpha)^t \rightarrow 0, \quad \alpha \sum_{j=0}^{t-1} (1-\alpha)^j \rightarrow 1$$

і, отже,

$$S_t = \alpha \sum_{j=0}^{t-1} (1-\alpha)^j y_{t-j}.$$

Тобто, експоненціальна середня – зважена сума всіх членів ряду, зі спадаючими експоненціально вагами залежно від давнини спостережень.

Формула (2.3.1) визначає також експоненціальні середні першого порядку, тобто середні, отримані безпосередньо при згладжуванні вихідних даних ряду динаміки.

У тих випадках, коли тенденція після згладжування вихідного ряду визначена недостатньо ясно, процедуру згладжування повторюють, тобто обчислюють експоненціальні середні 2-го, 3-го і т.д. порядків, користуючись наступними вираженнями:

$$S_t^{[2]} = \alpha S_t^{[1]} + (1-\alpha) S_{t-1}^{[2]}$$

$$S_t^{[3]} = \alpha S_t^{[2]} + (1-\alpha) S_{t-1}^{[3]}$$

.....

$$S_t^{[k]} = \alpha S_t^{[k-1]} + (1-\alpha) S_{t-1}^{[k]}$$

де $S_t^{[k]}$ – експоненціальна середня k -го порядку в крапці t ($k = 1, 2, \dots, n$).

Перед тем, як проводити експоненціальне згладжування, необхідно перевірити за допомогою візуального підходу (по виду графіка вихідних показників) або методу «прогноз екс-пост», яким рівнянням описується тенденція змін динамічного ряду – лінійним або квадратичним.

Крім того, при виборі виду рівняння необхідно розв'язати два питання: наскільки логічно й статистично обране рівняння відповідає досліджуваним процесам і явищам.

Під **логічною адекватністю** розуміють здатність рівняння адекватно, тобто найбільше точно відбивати природу досліджуваних явищ у широкому змісті, а відносно часового тренду – здатність відображати закономірність тенденції, що склалася.

Статистична адекватність означає відповідність рівняння окремим критеріям, які виражаються системою статистичних характеристик:

а) **коефіцієнт парної кореляції** для лінійного рівняння

$$r = \frac{n \sum yt - \sum y \sum t}{\sqrt{(n \sum t^2 - (\sum t)^2)(n \sum y^2 - (\sum y)^2)}}$$

і **парне кореляційне співвідношення** для всіх нелінійних рівнянь, які відображують тісноту зв'язку між результативним показником (функцією) і факторіальною ознакою (аргументом)

$$\eta = \sqrt{1 - \frac{\sum (y_t - \hat{y}_t)^2}{\sum (y_t - \bar{y}_t)^2}}.$$

Якщо $r, \eta \leq 0,5$ \Rightarrow зв'язок слабкий \Rightarrow адекватність низька;

Якщо $0,5 \leq r, \eta \leq 0,7$ \Rightarrow зв'язок середній \Rightarrow адекватність середня;

Якщо $r, \eta \geq 0,7$ \Rightarrow зв'язок сильний \Rightarrow адекватність висока;

б) **середня помилка апроксимації**

$$\bar{\varepsilon} = \frac{1}{n} \sum \frac{|y_t - \hat{y}_t|}{y_t} \cdot 100$$

Якщо $\bar{\varepsilon} \leq 10\%$ \Rightarrow рівняння з високим рівнем адекватності;

в) **середнє квадратичне відхилення** між фактичними й розрахунковими значеннями функції:

– абсолютне:
$$\sigma_{абс} = \sqrt{\frac{\sum (y_t - \hat{y}_t)^2}{n-1}};$$

– відносне:
$$\sigma_{отн} = \sqrt{\frac{\sum \frac{(y_t - \hat{y}_t)^2}{y_t}}{n-1}} \cdot 100;$$

г) **середнє відхилення** між фактичними й розрахунковими значеннями функції:

– абсолютне:
$$\bar{\Delta}_{абс} = \frac{1}{n} \sum |y_t - \hat{y}_t|;$$

– відносне:
$$\bar{\Delta}_{отн} = \bar{\varepsilon};$$

Чим менше значення наведених показників, тем вище якість обраного рівняння.

Розглянемо наступні **розрахункові формули методу для двох випадків:**

– якщо часовий тренд описується лінійним рівнянням;

– якщо часовий тренд описується квадратичним рівнянням.

1. Лінійне рівняння

Для лінійної моделі ($\hat{y}_t = a_0 + a_1 t$) експоненціальні середні першого й другого порядку розраховуються за формулами:

$$S_t^{[1]} = \alpha y_t + (1 - \alpha) S_{0(y)}^{[1]}$$

$$S_t^{[2]} = \alpha S_t^{[1]} + (1 - \alpha) S_{0(y)}^{[2]}.$$

Початкові умови $S_{0(y)}^{[1]}$ й $S_{0(y)}^{[2]}$ визначаються в такий спосіб:

$$S_{0(y)}^{[1]} = a_0 - \frac{1 - \alpha}{\alpha} a_1$$

$$S_{0(y)}^{[2]} = a_0 - \frac{2(1 - \alpha)}{\alpha} a_1.$$

(2.3.4)

В (2.3.4) для вираження значень коефіцієнтів a_0 і a_1 необхідно брати коефіцієнти рівняння тренду

$$\hat{y}_t = a_0 + a_1 t, \quad (2.3.5)$$

отримані методом найменших квадратів із системи рівнянь

$$\begin{cases} n a_0 + a_1 \sum t = \sum y \\ a_0 \sum t + a_1 \sum t^2 = \sum t y \end{cases}.$$

Щоб виразити коефіцієнти рівнянь тренду (2.3.5) через експоненціальні середні, використовується система рівнянь, що зв'язує оцінки коефіцієнтів $\overline{a_0}$ і $\overline{a_1}$ з названими експоненціальними середніми:

$$S_t^{[1]} = \overline{a_0} - \frac{1-\alpha}{\alpha} \overline{a_1}$$

$$S_t^{[2]} = \overline{a_0} - \frac{2(1-\alpha)}{\alpha} \overline{a_1},$$

дозволивши яку відносно $\overline{a_0}$ й $\overline{a_1}$, одержимо:

$$\overline{a_0} = 2 S_t^{[1]} - S_t^{[2]}$$

$$\overline{a_1} = \frac{\alpha}{1-\alpha} [S_t^{[1]} - S_t^{[2]}].$$

Прогноз у випадку лінійного тренду здійснюється по формулі

$$y_{t+l}^* = \overline{a_0} + \overline{a_1} l,$$

де l – час попередження прогнозу (періоди, на які складається прогноз).

Помилку прогнозу визначається за формулою

$$\sigma_{y_{t+l}^*} = \sigma_y \cdot \sqrt{\frac{\alpha}{(2-\alpha)^5} [1 + 4(1-\alpha) + 5(1-\alpha)^2 + 2\alpha(4-3\alpha)l + 2\alpha^2 l^2]},$$

де σ_y – середня квадратична помилка відхилення від лінійного тренду, яка обчислюється за формулою

$$\sigma_y = \sqrt{\frac{\sum (y_i - y_{t+l}^*)^2}{f}},$$

де f – число ступенів волі, обумовлене залежно від числа членів ряду n і числа оцінюваних параметрів кривій, що вирівнює.

2. Квадратичне рівняння

Якщо ряд динаміки описується параболою другого порядку

$$\hat{y}_t = a_0 + a_1 t + a_2 t^2, \quad (2.3.6)$$

то основні показники експоненціального згладжування розраховуються за наступними формулами:

– експоненціальні середні:

$$\begin{aligned} S_t^{[1]} &= \alpha y_t + (1-\alpha)S_{t-1}^{[1]} \\ S_t^{[2]} &= \alpha S_t^{[1]} + (1-\alpha)S_{t-1}^{[2]} \\ S_t^{[3]} &= \alpha S_t^{[2]} + (1-\alpha)S_{t-1}^{[3]} \end{aligned}$$

– початкові умови:

$$\begin{aligned} S_{0(y)}^{[1]} &= a_0 - \frac{1-\alpha}{\alpha} a_1 + \frac{(1-\alpha)(2-\alpha)}{\alpha^2} a_2 \\ S_{0(y)}^{[2]} &= a_0 - \frac{2(1-\alpha)}{\alpha} a_1 + \frac{2(1-\alpha)(3-2\alpha)}{\alpha^2} a_2 \\ S_{0(y)}^{[3]} &= a_0 - \frac{3(1-\alpha)}{\alpha} a_1 + \frac{3(1-\alpha)(4-3\alpha)}{\alpha^2} a_2 \end{aligned} \quad (2.3.7)$$

Як і для лінійного рівняння коефіцієнти a_0, a_1, a_2 рівняння (2.3.7) відповідають параметрам рівняння (2.3.6) і визначаються по методу найменших квадратів із системи рівнянь

$$\begin{cases} n a_0 + a_1 \sum t + a_2 \sum t^2 = \sum y \\ a_0 \sum t + a_1 \sum t^2 + a_2 \sum t^3 = \sum t y \\ a_0 \sum t^2 + a_1 \sum t^3 + a_2 \sum t^4 = \sum t^2 y \end{cases}$$

Для зв'язку оцінок $\overline{a_0}, \overline{a_1}, \overline{a_2}$ коефіцієнтів a_0, a_1, a_2 з експоненціальними середніми використовується система рівнянь

$$\begin{aligned} S_t^{[1]} &= \overline{a_0} - \frac{1-\alpha}{\alpha} \overline{a_1} + \frac{(1-\alpha)(2-\alpha)}{\alpha^2} \overline{a_2} \\ S_t^{[2]} &= \overline{a_0} - \frac{2(1-\alpha)}{\alpha} \overline{a_1} + \frac{2(1-\alpha)(3-2\alpha)}{\alpha^2} \overline{a_2}, \\ S_t^{[3]} &= \overline{a_0} - \frac{3(1-\alpha)}{\alpha} \overline{a_1} + \frac{3(1-\alpha)(4-3\alpha)}{\alpha^2} \overline{a_2} \end{aligned}$$

дозволивши яку відносно $\overline{a_0}, \overline{a_1}, \overline{a_2}$, одержимо наступні оцінки параметрів тренду:

$$\begin{aligned} \overline{a_0} &= 3 [S_t^{[1]} - S_t^{[2]}] + S_t^{[3]} \\ \overline{a_1} &= \frac{\alpha}{2(1-\alpha)^2} [(6-5\alpha)S_t^{[1]} - 2(5-4\alpha)S_t^{[2]} + (4-3\alpha)S_t^{[3]}] \\ \overline{a_2} &= \frac{\alpha^2}{(1-\alpha)^2} [S_t^{[1]} - 2S_t^{[2]} + S_t^{[3]}] \end{aligned}$$

Модель, за якою здійснюється прогноз:

$$y_{t+l}^* = \overline{a_0} + \overline{a_1} \cdot l + \frac{1}{2} \overline{a_2} \cdot l^2.$$

Помилка прогнозу визначається за формулою

$$\sigma_{y_{t+l}}^* = \sigma_y \sqrt{2a + 3a^2 + 3a^2 l^2},$$

де

$$\sigma_y = \sqrt{\frac{\sum (y_i - y_{t+l}^*)^2}{n-3}}.$$

Довірчі інтервали (верхня й нижня границі) як для лінійної, так і для квадратичної моделі перебувають по формулі:

$$y_{t+l}^* \pm \sigma_{y_{t+l}}^*.$$

При практичній використанні методу експоненціального згладжування виникають **деякі утруднення**. Основними з них є:

1) Вибір значення параметра згладжування.

При короткостроковій прогнозуванні бажане якнайшвидше відбити зміни ряду й у той же час якнайкраще «очистити» ряд від випадкових коливань.

Від чисельного значення параметра α залежить, наскільки швидко буде зменшуватися вага попередніх спостережень і відповідно до цього ступінь їх впливу на рівень, що згладжується.

Чим більше значення параметра α , тем менше позначається вплив попередніх рівнів і відповідно до менших виявляється вплив, що згладжує, експоненціальної середньої.

Отже, з одного боку, слід збільшувати вагу більш свіжих спостережень, підвищуючи α , з іншого боку, для згладжування випадкових відхилень величину α потрібно зменшувати.

У цей час є **кілька шляхів вибору параметра α** :

- Автор методу експоненціального згладжування англійський учений Р.Г. Браун запропонував наступну формулу розрахунків α :

$$\alpha = \frac{2}{m+1},$$

де m – число рівнів, що входять в інтервал згладжування.

Величина m , отже, і α визначаються в цьому випадку емпірично. У якості задовільного практичного компромісу він рекомендує брати α в межах від 0,1 до 0,3.

- Пошук оптимального значення параметра α може здійснюватися також шляхом перебору різних його значень.

У цьому випадку в якості оптимального обирається те значення α , при якому отримана найменша дисперсія помилки прогнозування, обчислена або при реалізації процедури згладжування всього ряду динаміки або на невикористаній у розрахунках ділянці ряду, спеціально залишеного для перевірки якості прогнозних моделей.

2) Визначення початкової умови y_0 .

Задачу вибору параметра y_0 , який визначає початкові умови, пропонується розв'язувати в такий спосіб:

- якщо є дані про розвиток явища в минулому, то в якості y_0 можна використовувати середню арифметичну всіх наявних рівнів ряду динаміки або якоїсь їхньої частини, тобто

$$\frac{1}{n} \sum_{i=1}^n y_i ;$$

- якщо ж таких відомостей немає, то в якості y_0 використовують вихідне (перше) значення рівня ряду динаміки y_1 . Вага, що приписується до цього рівня, швидко зменшується в міру віддалення від першого рівня ряду, разом із цим швидко зменшується його вплив на розмір експоненціальної середньої.

Метод експоненціального згладжування в порівнянні з іншими методами прогнозу має **переваги й недоліки**.

У числі переваг методу необхідно відзначити наступні:

- точність методу, яка збільшується зі збільшенням числа рівнів ряду;
- оптимальне значення єдиного параметра α можна знайти емпіричним шляхом;
- коефіцієнти моделі оцінюються спільно таким чином, щоб зменшити автокореляцію в залишках.

Недоліками методу виступають наступні:

- уся специфіка ряду має бути відображена в єдиному параметрі α , що обмежує клас моделей, припустимих у рамках даного методу;
- відсутній точний метод для вибору оптимальної величини параметра згладжування α ;
- метод розглядає часовий ряд ізольовано від інших явищ, і навіть якщо є додаткова інформація, вона може бути використана лише шляхом регулювання швидкості адаптації;
- точність прогнозу за цим методом падає зі збільшенням прогностного інтервалу;
- метод ефективний для короткострокових прогнозів, в інших умовах його можна використовувати для одержання наближених оцінок.

2.3.3.2 Застосування методу експоненціального згладжування до побудови прогнозів

Для ілюстрації застосування методу експоненціального згладжування для прогнозування розглянемо наступний приклад.

Приклад 2.3.3. Нехай заданий часовий ряд випуску натурального цементу (млн. т.) за період 1990-2011 рр. (табл. 2.3.7).

Таблиця 2.3.7

Обсяги випуску натурального цементу за період 1990-2011 рр., млн. т.

t	1	2	3	4	5	6	7	8	9	10	11
y_t	10,2	12,1	13,9	16,0	19,0	22,5	24,9	28,9	33,3	38,8	45,5
t	12	13	14	15	16	17	18	19	20	21	22
y_t	50,9	57,3	61,0	64,9	72,4	80,0	84,8	87,5	89,7	95,2	100,3

Перш за все, нанесемо усі вхідні дані на графік (рис. 2.3.2).

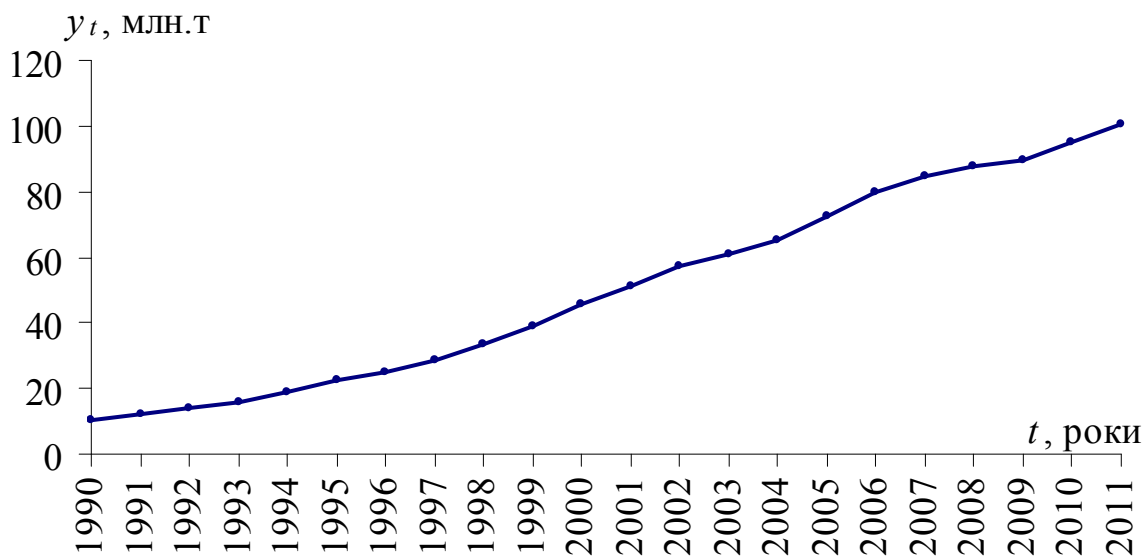


Рисунок 2.3.2 – Зміна випуску цементу за 1990-2011 рр.

Побудова прогнозу за методом експоненціального згладжування здійснюється у декілька етапів:

1. Передусім знайдемо тренд заданого ряду. З рис. 2.3.2 видно, що початковий часовий ряд найефективніше описується квадратичною функцією

$$y_t = a_0 + a_1 t + a_2 t^2. \quad (2.3.8)$$

Методом найменших квадратів визначаємо коефіцієнти рівняння (2.3.8):

$$a_0 = 3,5617, \quad a_1 = 3,0326, \quad a_2 = 0,0694.$$

Отже, тренд цього ряду достатньо добре описується параболою виду

$$y_t = 3,5617 + 3,0326 t + 0,0694 t^2. \quad (2.3.9)$$

Далі за формулою (2.3.9) розраховуються значення рівнів \hat{y}_t та відхилення фактичних значень рівнів ряду y_t від розрахункових значень \hat{y}_t (див. табл. 2.3.8).

Результати розрахунків

Роки	Фактичні значення рівнів	Розрахункові значення рівнів	Відхилення фактичних значень рівнів від розрахункових
t	y_t	\hat{y}_t	$y_t - \hat{y}_t$
1	10,2	6,7	3,5
2	12,1	9,9	2,2
3	13,9	13,3	0,6
4	16,0	16,8	-0,8
5	19,0	20,5	-1,5
6	22,5	24,3	-1,8
7	24,9	28,2	-3,3
8	28,9	32,3	-3,4
9	33,3	36,5	-3,2
10	38,8	40,8	-2,0
11	45,5	45,3	0,2
12	50,9	50,0	0,9
13	57,3	54,7	2,6
14	61,0	59,6	1,4
15	64,9	64,7	0,2
16	72,4	69,9	2,5
17	80,0	75,2	4,8
18	84,8	80,6	4,2
19	87,5	86,2	1,3
20	89,7	92,0	-2,3
21	95,2	97,9	-2,7
22	100,3	103,9	-3,6

2. Визначаються основні показники експоненціального згладжування.

В якості інтервалу згладжування для розрахунку параметра α використовувався період в 11 років (це було обумовлено тим, що до 11-го року (2001 р.) в основному було закінчене технічне переозброєння цементної промисловості), тобто $m = 11$. Звідси, використовуючи формулу Р.Г. Брауна, знаходимо, що

$$\alpha = \frac{2}{m+1} = 0,1666667.$$

Для визначення оцінок коефіцієнтів $\overline{a_0}, \overline{a_1}, \overline{a_2}$ окрім значення α необхідно знайти експоненціальні середні $S_{t(y)}^{[1]}, S_{t(y)}^{[2]}, S_{t(y)}^{[3]}$.

Для їх розрахунку скористаємося формулою $S_{t(y)}^{[k]} = \alpha S_{t(y)}^{[k-1]} + (1-\alpha)S_{t-1(y)}^{[k]}$.

Знаходимо початкові умови за формулами:

$$\begin{aligned} S_{0(y)}^{[1]} &= a_0 - \frac{1-\alpha}{\alpha} a_1 + \frac{(1-\alpha)(2-\alpha)}{\alpha^2} a_2 \\ S_{0(y)}^{[2]} &= a_0 - \frac{2(1-\alpha)}{\alpha} a_1 + \frac{2(1-\alpha)(3-2\alpha)}{\alpha^2} a_2 . \\ S_{0(y)}^{[3]} &= a_0 - \frac{3(1-\alpha)}{\alpha} a_1 + \frac{3(1-\alpha)(4-3\alpha)}{\alpha^2} a_2 \end{aligned} \quad (2.3.10)$$

При вираженні початкових умов у формулах (2.3.10) необхідно в якості коефіцієнтів a_0, a_1, a_2 брати коефіцієнти рівняння тренду (2.3.9), отримані методом найменших квадратів.

У нашому прикладі $S_{0(y)}^{[1]} = -7,783$; $S_{0(y)}^{[2]} = -15,655$; $S_{0(y)}^{[3]} = -20,057$.

Далі підраховуються безпосередньо самі експоненціальні середні за формулами:

$$\begin{aligned} S_t^{[1]} &= \alpha y_t + (1-\alpha)S_{t-1(y)}^{[1]} \\ S_t^{[2]} &= \alpha S_t^{[1]} + (1-\alpha)S_{t-1(y)}^{[2]} . \\ S_t^{[3]} &= \alpha S_t^{[2]} + (1-\alpha)S_{t-1(y)}^{[3]} \end{aligned}$$

Отримані значення заносяться до табл. 2.3.9.

3. Для побудови моделі прогнозу використовуються наступні формули для визначення експоненціальних середніх:

$$\begin{aligned} \bar{a}_0 &= a_0 - \frac{1-\alpha}{\alpha} a_1 + \frac{(1-\alpha)(2-\alpha)}{\alpha^2} a_2 ; \\ \bar{a}_1 &= a_0 - \frac{2(1-\alpha)}{\alpha} a_1 + \frac{2(1-\alpha)(3-2\alpha)}{\alpha^2} a_2 ; \\ \bar{a}_2 &= a_0 - \frac{3(1-\alpha)}{\alpha} a_1 + \frac{3(1-\alpha)(4-3\alpha)}{\alpha^2} a_2 , \end{aligned}$$

звідки визначаються оцінки коефіцієнтів :

$$\begin{aligned} \bar{a}_0 &= 3 [S_t^{[1]} - S_t^{[2]}] + S_t^{[3]} ; \\ \bar{a}_1 &= \frac{\alpha}{2(1-\alpha)^2} [(6-5\alpha)S_t^{[1]} - 2(5-4\alpha)S_t^{[2]} + (4-3\alpha)S_t^{[3]}] ; \\ \bar{a}_2 &= \frac{\alpha^2}{(1-\alpha)^2} [S_t^{[1]} - 2S_t^{[2]} + S_t^{[3]}] \end{aligned}$$

і отримані значення заносяться до табл. 2.3.9.

Далі здійснюємо прогноз на період 1990-2011 рр. за формулою

$$y_{t+l}^* = \bar{a}_0 + \bar{a}_1 \cdot l + \frac{1}{2} \bar{a}_2 \cdot l^2 \quad (\text{в даному випадку } l = 1),$$

тобто фактично ми здійснюємо прогноз на період 1990-2011 рр. за формулою

$$y_{t+l}^* = \bar{a}_0 + \bar{a}_1 + \frac{1}{2} \bar{a}_2.$$

Отримані дані заносимо до табл. 2.3.9 та знаходимо відхилення фактичних рівнів часового ряду y_t від прогнозованих y_t^* , тобто $\sigma = y_t - y_t^*$.

Таблиця 2.3.9

t	$S_{t(y)}^{[1]}$	$S_{t(y)}^{[2]}$	$S_{t(y)}^{[3]}$	\bar{a}_0	\bar{a}_1	\bar{a}_2	y_t^*	$y_t - y_t^*$
2	-4,7856	-13,8439	-19,0213	8,1535	3,4416	0,1552	11,6728	0,4272
3	-1,9713	-11,8651	-17,8286	11,8528	3,6295	0,1572	15,5608	-1,6608
4	0,6739	-9,7753	-16,4864	14,8611	3,6598	0,1495	18,5957	-2,5957
5	3,2282	-7,6080	-15,0067	17,5022	3,6111	0,1375	21,1820	-2,1820
6	5,8569	-5,3639	-13,3995	20,2627	3,5819	0,1274	23,9083	-1,4083
7	8,6307	-3,0314	-11,6715	23,3150	3,6017	0,1209	26,9772	-2,0772
8	11,3423	-0,6358	-9,8322	26,1021	3,5639	0,1113	29,7216	-0,8216
9	14,2686	1,8482	-7,8855	29,3755	3,6124	0,1075	33,0416	0,2584
10	17,4405	4,4469	-5,8301	33,1505	3,7396	0,1087	36,9495	1,8555
11	21,0004	7,2058	-3,6574	37,7262	3,9900	0,1173	41,7749	3,7251
12	25,0837	10,1855	-1,3503	43,3442	4,3918	0,1345	47,8033	3,0967
13	29,3864	13,3856	1,1057	49,1080	4,7629	0,1488	53,9453	3,3547
14	34,0387	16,8278	3,7261	55,3586	5,1680	0,1644	60,6088	0,3912
15	38,5322	20,4452	6,5126	60,7736	5,3622	0,1662	66,2189	-1,3189
16	42,9268	24,1921	9,4592	65,6633	5,4277	0,1601	71,1710	1,2290
17	47,8390	28,1333	12,5715	71,6888	5,6816	0,1658	77,4533	2,5467
18	53,1992	32,3109	15,8614	78,5262	6,0419	0,1775	84,6569	0,1431
19	58,4660	36,6701	19,3295	84,7172	6,2304	0,1782	91,0367	-3,5367
20	63,3050	41,1093	22,9595	89,5467	6,1385	0,1618	95,7661	-6,0661
21	67,7042	45,5417	26,7232	93,2105	5,8369	0,1338	99,1142	-3,9142
22	72,2868	49,9993	30,6025	97,4652	5,6717	0,1156	103,1947	-2,8947
23	76,956	54,492	34,584	101,975	5,566	0,102	107,59	

4. Складається прогноз на 2012-2015 рр.

При побудові прогнозу виробництва цементу на 2012-2015 рр. використовуються наступні величини (див. останній рядок у табл. 2.3.9) експоненціальних середніх

$$S_{t(y)}^{[1]} = 76,956, \quad S_{t(y)}^{[2]} = 54,492, \quad S_{t(y)}^{[3]} = 34,584,$$

відповідні оцінки коефіцієнтів моделі прогнозу

$$\bar{a}_0 = 101,975, \quad \bar{a}_1 = 5,566, \quad \bar{a}_2 = 0,102,$$

прогнозне значення $y_{2012}^* = 107,59$.

Результати прогнозу на 2012-2015 рр. занесемо до табл. 2.3.10.

Таблиця 2.3.10

Прогноз випуску цементу на 2012-2015 рр., млн. т.

роки	l	прогноз	помилка прогнозу	Довірчі інтервали	
				верхня межа	нижня межа
2012	1	107,59	1,80	109,39	105,79
2013	2	113,31	1,89	115,20	111,42
2014	3	119,13	2,02	121,15	117,11
2015	4	125,06	2,19	127,25	122,87

У таблиці 2.3.10:

– прогноз здійснюється за формулою

$$y_{t+l}^* = \bar{a}_0 + \bar{a}_1 \cdot l + \frac{1}{2} \bar{a}_2 \cdot l^2, \text{ де в даному випадку } l = \bar{1,4}.$$

– помилка прогнозу знаходиться за формулою

$$\sigma_{y_{t+l}^*} = \sigma_y \sqrt{2a + 3a^2 + 3a^2 l^2}, \quad l = \bar{1,4}$$

$$\text{де } \sigma_y = \sqrt{\frac{\sum_{i=2}^{n-1} (y_i - y_i^*)^2}{n-3}} = \sqrt{\frac{142,5574}{19}} = 2,739164; \quad y_i^* - \text{ визначається з}$$

табл. 2.3.9.

– довірчі інтервали знаходяться наступним чином:

$$\text{верхня межа: } y_{t+l}^* + \sigma_{y_{t+l}^*};$$

$$\text{нижня межа: } y_{t+l}^* - \sigma_{y_{t+l}^*}.$$

5. На останньому етапі графічно зображуються весь часовий ряд випуску цементу за 1990-2011 рр., прогноз на 2012-2015 рр. та довірчі інтервали для прогнозованих значень (див. рис. 2.3.3).

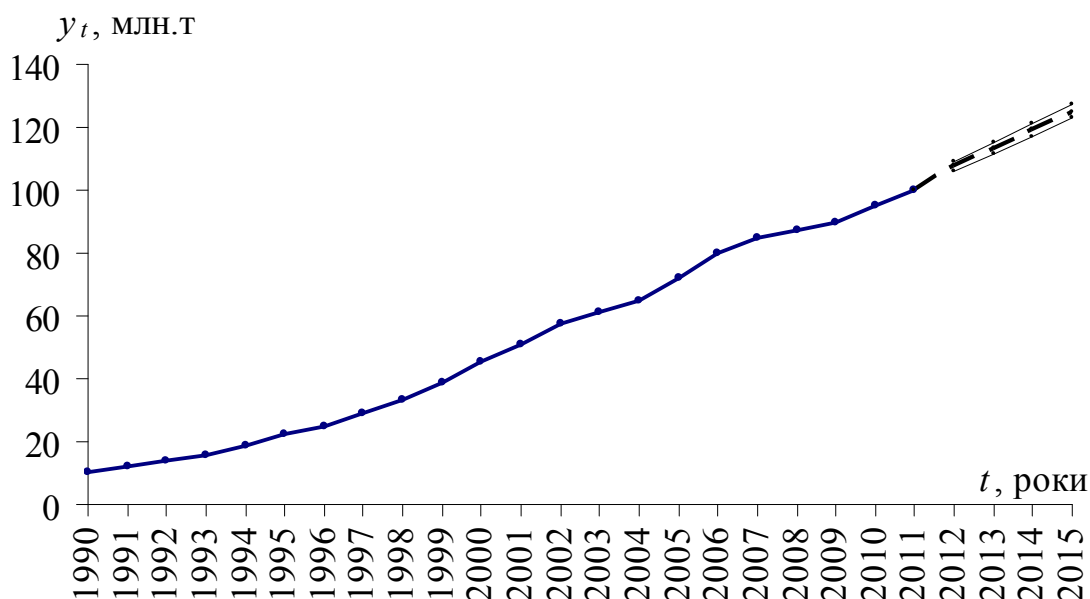


Рисунок 2.3.3 – Зміна випуску цементу за 1990-2015 рр.

2.3.4 Метод гармонійних ваг

2.3.4.1 Сутність методу

Метод гармонійних ваг, розроблений польським статистиком З. Хелвигом, є близьким до методу простого експоненціального згладжування, оскільки використовує той же принцип.

Особливість методу: в його основі лежить зважування ковзного показника, проте замість ковзної середньої використовується ковзний тренд.

Суть методу: екстраполяція проводиться за ковзним трендом, окремі точки ламаної лінії зважуються за допомогою гармонійних ваг, що дозволяє пізнішим спостереженням надавати більші ваги.

Метод гармонійних ваг базується на наступних передумовах:

- 1) період часу, за який вивчається виробничий процес, має бути досить тривалим, щоб можна було визначити його закономірності;
- 2) початковий ряд динаміки не повинен мати стрибкоподібних змін;
- 3) прогнозоване явище повинне мати інерційність, тобто для настання великої зміни в характеристиках процесу необхідно, щоб пройшов значний час;
- 4) відхилення від ковзного тренду (E_t) мають випадковий характер;
- 5) автокореляційна функція, розрахована на основі послідовних різниць, має зменшуватися зі зростом t , тобто вплив пізнішої інформації має сильніше відображатися на прогнозованій величині, ніж на ранній інформації.

Для отримання точного прогнозу методом гармонійних ваг потрібне виконання усіх вищезгаданих передумов для початкового ряду динаміки.

Прогнозування за цим методом здійснюється у декілька етапів:

1 етап: Заданий ряд динаміки перевіряється на виконання перших чотирьох передумов, на яких базується метод.

2 етап: Початковий ряд динаміки розбивається на фази k . При цьому число фаз k має бути меншим за число членів ряду n , тобто $k < n$. Зазвичай фаза дорівнює 3-5 рівням.

Для кожної фази розраховується лінійний тренд виду

$$y_{i(t)} = a_i + b_{ik} t \quad (i = 1, 2, \dots, n - k + 1),$$

для оцінки параметрів якого використовується МНК (Додаток В).

При цьому

$$\text{для } i = 1: \quad t = 1, 2, \dots, k;$$

$$\text{для } i = 2: \quad t = 2, 3, \dots, k + 1;$$

$$\text{для } i = n - k + 1: \quad t = n - k + 1, n - k + 2, \dots, n.$$

3 етап: За допомогою отриманих на 2 етапі $(n - k + 1)$ рівнянь визначаються значення ковзного тренду. З цією метою виділяються ті значення $y_{i(t)}$, для яких $t = i$, їх позначають $y_{j(t)}$. Нехай їх буде n_j .

Далі на цьому етапі знаходиться середнє значення $\bar{y}_{j(t)}$ за формулою

$$\bar{y}_{j(t)} = \frac{1}{n_j} \sum_{j=1}^{n_j} y_{j(t)}, \quad j = 1, 2, \dots, n_j. \quad (2.3.11)$$

4 етап: Перевіряється гіпотеза про те, що відхилення від ковзного тренду є стаціонарним в широкому сенсі випадковим процесом. З цією метою розраховується автокореляційна функція. Якщо від періоду до періоду її значення зменшуються, то п'ята передумова методу виконується.

5 етап: Розраховуються основні параметри методу гармонійних ваг:

а) обчислюються прирости у вигляді

$$\omega_{t+1} = f_{(t+1)} - f_{(t)} = \bar{y}_{t+1} - \bar{y}_t; \quad (2.3.12)$$

б) розраховується середня приростів за формулою

$$\omega = \sum_{t=1}^{n-1} C_{t+1}^n \cdot \omega_{t+1}, \quad t = 1, 2, \dots, n - 1, \quad (2.3.13)$$

де C_{t+1}^n – гармонійні коефіцієнти, що задовольняють умовам

$$C_{t+1}^n > 0, \quad \sum_{t=1}^{n-1} C_{t+1}^n = 1. \quad (2.3.14)$$

Вираз (2.3.13) дозволяє більш пізній інформації надавати великі ваги, оскільки прирости обернено пропорційні часу, який відділяє ранню інформацію від пізньої для моменту $t = n$.

Якщо найраніша інформація має вагу $m_2 = \frac{1}{n-1}$, то вага інформації, що відноситься до наступного моменту часу, дорівнює

$$m_3 = m_2 + \frac{1}{n-2} = \frac{1}{n-1} + \frac{1}{n-2}.$$

в) розраховуються гармонійні ваги. У загальному вигляді ряд гармонійних ваг визначають за формулою

$$m_{t+1} = m_t + \frac{1}{n-t} \quad \text{або} \quad m_{t+1} = \sum_{i=1}^t \frac{1}{n-i}, \quad t = 2, 3, \dots, n-1, \quad (2.3.15)$$

звідки $\sum_{t=1}^{n-1} m_{t+1} = n-1$.

г) визначаються гармонійні коефіцієнти C_{t+1}^n , які задовольняють умовам (2.3.14). Для їх отримання треба гармонійні ваги m_{t+1} поділити на $(n-1)$, тобто

$$C_{t+1}^n = \frac{m_{t+1}}{n-1}. \quad (2.3.16)$$

6 етап: На даному етапі проводиться побудова прогнозу. Прогнозування здійснюється так само, як і при простих методах прогнозу – шляхом збільшення до останнього значення ряду динаміки середнього приросту, тобто за формулою

$$\hat{y}_t = \bar{y}_t^* + \bar{\omega} \quad (2.3.17)$$

за початкової умови $\bar{y}_t^* = \bar{y}_{j(t)}$.

Довірчі інтервали прогнозу знаходяться із використанням нерівності Чебишева для випадкової величини ω_{t+1} :

$$P \left\{ \left| \omega_{t+1} - E(\omega) \right| > a\sigma_\omega \right\} < \frac{1}{a^2}, \quad (2.3.18)$$

де a – задане ціле додатне число; σ_ω – визначається у вигляді

$$\sigma_\omega = \rho_\omega = \sqrt{\sum_{t=1}^{n-1} C_{t+1}^n (\omega_{t+1} - \bar{\omega})^2}. \quad (2.3.19)$$

Оскільки значення змінної ω_{t+1} корельовані між собою, то можна припустити, що при прогнозі a – є величина змінна. Вона є функцією віддаленості від l , де $l = 1, 2, 3, \dots, n-1$.

Функція $a(l)$ визначається з формули

$$a(l) = a \sum_{t=1}^{l+1} C_{n-t+1}^n. \quad (2.3.20)$$

Тоді довірчі межі для прогнозних значень y_{t+1}^* наступні:

$$y_{n+l}^* - a(l)\rho_\omega \leq y_{t+l} \leq y_{n+l}^* + a(l)\rho_\omega. \quad (2.3.21)$$

Наостанок зазначимо, що метод гармонійних ваг застосовується, коли є упевненість, що тенденція в майбутньому описується плавною кривою, тобто коли у часовому ряді відсутні сезонні та циклічні коливання.

2.3.4.2 Застосування методу гармонійних ваг до побудови прогнозів

Проілюструємо застосування методу гармонійних ваг на наступному прикладі.

Приклад 2.3.4. Нехай заданий часовий ряд динаміки виробництва продукції (тис. грн.) за 2003-2011 рр. (табл. 2.3.11)

Таблиця 2.3.11

Динаміка виробництва продукції (тис. грн.) за 2003-2011 рр.

2003 р.	2004 р.	2005 р.	2006 р.	2007 р.	2008 р.	2009 р.	2010 р.	2011 р.
10,0	11,1	12,1	12,5	13,7	13,9	19,6	15,9	19,0

Прогнозування за цим методом здійснюється у декілька етапів:

1 етап: Заздалегідь ряд динаміки перевіряється на виконання перших чотирьох передумов, на яких базується метод.

2 етап: Для здійснення прогнозу початковий ряд динаміки розбивається на фази k , причому число фаз k має бути менше числа членів ряду n . Зазвичай фаза дорівнює 3-5 рівням. Далі знаходимо параметри рівнянь окремих фаз руху ковзного тренду.

У нашому прикладі приймемо $k = 3$, тоді маємо $(9 - 3 + 1) = 7$ «нових» часових рядів

- 1) 10,0 ; 11,1 ; 12,1 при $t = 1, 2, 3$
- 2) 11,1 ; 12,1 ; 12,5 при $t = 2, 3, 4$
- 3) 12,1 ; 12,5 ; 13,7 при $t = 3, 4, 5$
- 4) 12,5 ; 13,7 ; 13,9 при $t = 4, 5, 6$
- 5) 13,7 ; 13,9 ; 19,6 при $t = 5, 6, 7$
- 6) 13,9 ; 19,6 ; 15,9 при $t = 6, 7, 8$
- 7) 19,6 ; 15,9 ; 19,0 при $t = 7, 8, 9$

та $(9 - 3 + 1) = 7$ рівнянь

$$\begin{aligned} \hat{y}_{1(t)} &= 8,97 + 1,05t, & t = 1, 2, 3; & & \hat{y}_{5(t)} &= 11,4 + 0,45t, & t = 5, 6, 7; \\ \hat{y}_{2(t)} &= 9,8 + 0,7t, & t = 2, 3, 4; & & \hat{y}_{6(t)} &= 7,8 + 1,0t, & t = 6, 7, 8; \\ \hat{y}_{3(t)} &= 9,57 + 0,8t, & t = 3, 4, 5; & & \hat{y}_{7(t)} &= -1,1 + 2,2t, & t = 7, 8, 9. \\ \hat{y}_{4(t)} &= 9,87 + 0,7t, & t = 4, 5, 6; & & & & \end{aligned}$$

3 етап: За допомогою отриманих на 2 етапі рівнянь визначаємо значення ковзного тренду.

При $t = 1$ маємо одне значення $\bar{y}_{1(t)}$, яке отримуємо з рівняння

$$\begin{aligned} \hat{y}_{1(t)} &= 8,97 + 1,05 \cdot 1, \\ \text{тобто } \bar{y}_{1(t)} &= \hat{y}_{1(t)} = 10,02. \end{aligned}$$

При $t = 2$ маємо два значення $\bar{y}_{2(t)}$, які отримаємо з рівнянь:

$$\hat{y}_{2(t)} = 8,97 + 1,05 \cdot 2,$$

$$\hat{y}_{2(t)} = 9,8 + 0,70 \cdot 2,$$

тобто $\hat{y}_{2(t)} = 11,07$ та $\hat{y}_{2(t)} = 11,2$, звідки $\bar{y}_{2(t)} = \frac{11,07 + 11,2}{2} = 11,13$.

Аналогічно знаходимо усі значення $\bar{y}_{j(t)}$.

$$\bar{y}_{3(t)} = 11,99; \quad \bar{y}_{5(t)} = 13,53; \quad \bar{y}_{7(t)} = 14,55; \quad \bar{y}_{9(t)} = 18,74.$$

$$\bar{y}_{4(t)} = 12,68; \quad \bar{y}_{6(t)} = 13,99; \quad \bar{y}_{8(t)} = 16,15;$$

4 етап: Перевіряється гіпотеза про те, що відхилення від ковзного тренду є стаціонарним в широкому сенсі випадковим процесом.

З цією метою по виділених фазах розраховуються автокореляційні функції, тобто послідовність значень коефіцієнтів автокореляції r_τ , де τ – лаг ($\tau = 1, 2, \dots, l$). При розрахунку коефіцієнтів автокореляції r_τ із зростанням порядку τ число корельованих пар зменшується, а відомо, що при невеликому числі спостережень значущими виявляються лише високі коефіцієнти кореляції. Отже, найбільше значення τ має бути таким, щоб число пар спостережень виявилось достатнім для обчислення коефіцієнтів автокореляції r_τ . В практиці орієнтуються на правило $\tau \leq n/4$.

У нашому випадку, при розрахунку коефіцієнтів автокореляції між першими двома рядами по виділених фазах k , маємо наступний ряд

$$\begin{array}{cccc} \text{ряд 1} & & & \\ \underbrace{10,0 \quad 11,1 \quad 12,1 \quad 12,5}_{y_1 \quad y_2 \quad y_3 \quad y_4} & & & n = 4 \\ & & & \tau = n/4 = 1 \\ & & \underbrace{\hspace{10em}} & \\ & & \text{ряд 2} & \end{array}$$

Коефіцієнти автокореляції розраховуються за формулою парного коефіцієнта кореляції. Так, коефіцієнт автокореляції першого порядку ($\tau = 1$) є не що інше, як парний коефіцієнт кореляції між послідовними рівнями динамічного ряду, тобто зв'язок між рядами $y_1, y_2, y_3, \dots, y_{n-1}$ та $y_2, y_3, \dots, y_{n-1}, y_n$:

$$r_1 = \frac{\sum_{t=1}^{n-1} y_t y_{t+1} - \frac{1}{n-1} \left(\sum_{t=1}^{n-1} y_t \right) \left(\sum_{t=1+1}^n y_t \right)}{\sqrt{\left(\sum_{t=1}^{n-1} y_t^2 - \frac{1}{n-1} \left(\sum_{t=1}^{n-1} y_t \right)^2 \right) \left(\sum_{t=1+1}^n y_t^2 - \frac{1}{n-1} \left(\sum_{t=1+1}^n y_t \right)^2 \right)}}.$$

Аналогічно обчислюються коефіцієнти автокореляції другого порядку ($\tau = 2$), третього і так далі.

Загальна формула для розрахунку коефіцієнта автокореляції порядку τ запишеться так:

$$r_{\tau} = \frac{\sum_{t=1}^{n-\tau} y_t y_{t+\tau} - \frac{1}{n-1} \left(\sum_{t=1}^{n-\tau} y_t \right) \left(\sum_{t=1+\tau}^n y_t \right)}{\sqrt{\left(\sum_{t=1}^{n-\tau} y_t^2 - \frac{1}{n-1} \left(\sum_{t=1}^{n-\tau} y_t \right)^2 \right) \left(\sum_{t=1+\tau}^n y_t^2 - \frac{1}{n-1} \left(\sum_{t=1+\tau}^n y_t \right)^2 \right)}},$$

де τ – лаг ($\tau = 1, 2, \dots, l$), n – число рівнів ряду, t – індекс року.

У нашому прикладі маємо наступну послідовність автокореляційних функцій:

$$r_{y_{1(t)}, y_{2(t)}} = 0,976960; \quad r_{y_{3(t)}, y_{4(t)}} = 0,782467; \quad r_{y_{5(t)}, y_{6(t)}} = -0,140202;$$

$$r_{y_{2(t)}, y_{3(t)}} = 0,866025; \quad r_{y_{4(t)}, y_{5(t)}} = 0,633375; \quad r_{y_{6(t)}, y_{7(t)}} = -0,979783.$$

Значення отриманих автокореляційних функцій показують, що величина їх зменшується, тобто виконується одна з основних умов (п'ята передумова) застосування методу гармонійних ваг для прогнозування виробничих явищ.

5 етап: Розраховуються основні параметри методу гармонійних ваг.

Розраховуються прирости у вигляді

$$\omega_{t+1} = f_{(t+1)} - f_{(t)} = \bar{y}_{t+1} - \bar{y}_t,$$

тобто для нашого прикладу:

$$\omega_2 = 1,11; \quad \omega_3 = 0,86; \quad \omega_4 = 0,70; \quad \omega_5 = 0,85;$$

$$\omega_6 = 0,46; \quad \omega_7 = 0,56; \quad \omega_8 = 1,6; \quad \omega_9 = 2,59,$$

і гармонійні ваги у вигляді

$$m_{t+1} = m_t + \frac{1}{n-t} \quad (t = 2, 3, \dots, n-1),$$

тобто для нашого прикладу:

$$m_2 = 0,125; \quad m_3 = 0,268; \quad m_4 = 0,435; \quad m_5 = 0,635;$$

$$m_6 = 0,885; \quad m_7 = 1,218; \quad m_8 = 1,717; \quad m_9 = 2,717,$$

причому $\sum_{t=1}^{n-1} m_{t+1} = n-1 = 8$.

Гармонійні коефіцієнти отримаємо за формулою $C_{t+1}^n = \frac{m_{t+1}}{n-1}$:

$$C_2 = 0,0156; \quad C_3 = 0,0335; \quad C_4 = 0,0543; \quad C_5 = 0,0793;$$

$$C_6 = 0,1106; \quad C_7 = 0,1522; \quad C_8 = 0,2147; \quad C_9 = 0,3397.$$

Усі ці коефіцієнти задовольняють умовам

$$C_{t+1}^n > 0, \quad \sum_{t=1}^{n-1} C_{t+1}^n = 1 \quad (t = 1, 2, \dots, n-1).$$

Знаходимо середній абсолютний приріст у вигляді

$$\bar{\omega} = \sum_{t=1}^{n-1} C_{t+1}^n \cdot \omega_{t+1} = 1,51.$$

6 етап: Розраховуються прогнознi значення виробництва реалізованої продукції на 2012-2016 рр. (див. табл. 2.3.12) за формулою $\hat{y}_t = \bar{y}_t^* + \bar{\omega}$ за початкової умови $\bar{y}_t^* = \bar{y}_{j(t)}$.

Таблиця 2.3.12

Прогноз виробництва реалізованої продукції на 2012-2016 рр., тис. грн.

Роки	$t+l$ ($t=9$)	y_{t+l}^*	l	$a(l)$	$a(l)\rho_\omega$	Довірчі інтервали	
						$y_{t+l}^* - a(l)\rho_\omega$	$y_{t+l}^* + a(l)\rho_\omega$
2012	10	20,25	1	2,218	1,913	18,40	22,16
2013	11	21,76	2	2,826	2,437	19,32	24,20
2014	12	23,27	3	3,269	2,819	20,45	26,09
2015	13	24,78	4	3,586	3,092	21,69	27,87
2016	14	26,29	5	3,803	3,279	23,01	29,57

Довірчі інтервали прогнозу знаходяться, використовуючи нерівність Чебишева (2.3.18), у якій a приймається заданим ($a=4$),

$$\rho_\omega = \sqrt{\sum_{t=1}^{n-1} C_{t+1}^n (\omega_{t+1} - \bar{\omega})^2} = 0,862.$$

Функція $a(l)$ визначається з (2.3.20) та подається у табл. 2.3.12.

За представленими у табл. 2.3.12 результатами обчислень, знаходимо довірчі межі для передбачених значень ω_{t+1} у вигляді (2.3.21). Усі розрахунки прогнозу та довірчих меж також подаються у табл. 2.3.12.

2.3.5 Питання для самоконтролю

1. Назвіть основні компоненти часового ряду.
2. В чому полягає відмінність між циклічними та сезонними коливаннями?
3. Розкрийте сутність адаптивних методів прогнозування.
4. В чому полягає відмінність між адитивною та мультиплікативною моделями часового ряду?
5. В чому полягає сутність методу ковзної середньої? Як за цим методом будується прогноз?
6. Як обирається вигляд кривої росту?
7. В чому полягає сутність методу експоненціального згладжування?
8. Яким чином обчислюється параметр згладжування в методі експоненціального згладжування?
9. В чому полягає сутність методу гармонійних ваг?
10. Основні етапи застосування методу гармонійних ваг.

2.3.6 Рекомендована література



Для самостійного вивчення теми рекомендується література: основна – [2, 3, 5, 8, 10], додаткова – [1, 4].

2.4 Методи аналізу якості побудованих моделей та прогнозів виробничих процесів

2.4.1 Методи аналізу якості побудованих моделей

Важливим етапом моделювання виробничих процесів є перевірка адекватності побудованої моделі реальному явищу. Для цього досліджують ряд залишків $\varepsilon_t = y_t - \tilde{Y}_t$, тобто відхилення розрахункових значень від фактичних.

Для оцінки адекватності побудованих моделей досліджуються властивості залишкової компоненти, тобто розбіжності рівнів, розрахованих за моделлю, та фактичних спостережень.

Найбільш важливими властивостями залишкової компоненти є:

- незалежність рівнів ряду залишків;
- випадковість рівнів ряду залишків;
- відповідність рівнів ряду залишків нормальному закону розподілу.

2.4.1.1 Перевірка рівності нулю математичного сподівання рівнів ряду залишків

Перевірка рівності математичного сподівання рівнів ряду залишків нулю здійснюється в ході перевірки відповідної нульової гіпотези $H_0: |\bar{\varepsilon}| = 0$.

З цією метою будується t – статистика:

$$t_{расч} = \frac{|\bar{\varepsilon}|}{S_{\varepsilon}} \sqrt{n},$$

де $\bar{\varepsilon}$ – середнє арифметичне значення рівнів ряду залишків ε_t ; S_{ε} – середньоквадратичне відхилення для цієї послідовності, розраховане за формулою для малої вибірки:

$$S_{\varepsilon} = \sqrt{\frac{\sum_{t=1}^n (\varepsilon_t - \bar{\varepsilon})^2}{n-1}}.$$

На рівні значущості α гіпотеза H_0 відхиляється (модель є неадекватною), якщо $t_{расч} > t_{\alpha, \nu}$, де $t_{\alpha, \nu}$ – критерій розподілу Стьюдента з довірчою ймовірністю $(1 - \alpha)$ і ступенями свободи $\nu = n - 1$.

Якщо ж гіпотеза H_0 приймається ($t_{расч} < t_{\alpha, \nu}$), то модель вважається адекватною.

2.4.1.2 Перевірка умови випадковості виникнення окремих відхилень від тренду

Для здійснення такої перевірки використовується критерій піків, заснований на поворотних точках.

Значення випадкової змінної вважається *поворотною точкою*, якщо воно одночасно більше (менше) сусідніх з ним елементів.

Якщо залишки випадкові, то поворотна точка приходиться приблизно на кожні 1,5 спостережень. Якщо їх більше, то обурення швидко коливаються, і це не може бути пояснено тільки випадковістю. Якщо ж їх менше, то послідовні значення випадкового компонента додатно корельовані.

Критерій випадковості відхилень від тренду при рівні ймовірності 0,95 можна представити як

$$p > \left[\frac{2}{3}(n-2) - 1,96\sqrt{\frac{16n-29}{90}} \right], \quad (2.4.1)$$

де p – фактична кількість поворотних точок у випадковому ряді; 1,96 – квантиль нормального розподілу для 5% -го рівня значущості. Квадратні дужки означають, що від результату обчислення слід узяти цілу частину.

Якщо нерівність (2.4.1) не виконується, то ряд залишків не можна вважати випадковим (тобто він містить регулярну компоненту), отже, модель не є адекватною.

2.4.1.3 Наявність (відсутність) автокореляції у відхиленнях від моделі зросту

Наявність (відсутність) автокореляції у відхиленнях від моделі зросту перевіряють за допомогою *критерію Дарбіна-Уотсона*. З цією метою будується статистика Дарбіна-Уотсона (d – статистика), в основі якої лежить формула

$$d = \frac{\sum_{t=2}^n (\varepsilon_t - \varepsilon_{t-1})^2}{\sum_{t=1}^n \varepsilon_t^2}, \quad (2.4.2)$$

За відсутності автокореляції $d \approx 2$, а при повній автокореляції $d \approx 0$ або $d \approx 4$. Отже, оцінки, що отримуються за критерієм, є не точковими, а інтервальними.

Верхні (d_2) та нижні (d_1) критичні значення, які дозволяють прийняти або відхилити гіпотезу про відсутність автокореляції, залежать від кількості рівнів динамічного ряду і числа незалежних змінних моделі. Значення цих меж для рівня значущості $\alpha = 0,05$ подані у табл. 2.4.1.

Статистичні таблиці критичних рівнів при рівні значущості $\alpha = 0,05$

Кількість спостережень	Перший коефіцієнт автокореляції, $r(1)$	Модель				Межа RS -критерію	
		однопараметрична ($m = 1$)		двохпараметрична ($m = 2$)		нижня	верхня
		d_1	d_2	d_1	d_2		
10	0,360	—	—	—	—	2,67	3,69
15	0,328	1,08	1,36	0,95	1,54	2,96	4,14
20	0,300	1,20	1,41	1,10	1,54	3,18	4,49
25	0,276	1,28	1,45	1,20	1,55	3,34	4,71
30	0,257	1,35	1,49	1,28	1,57	3,47	4,89

При порівнянні розрахункового значення d – статистики (2.4.2) з табличним можуть виникнути такі ситуації:

- $d_2 < d < 2$ — ряд залишків не корельований;
- $d < d_1$ — залишки містять автокореляцію;
- $d_1 < d < d_2$ — область невизначеності, коли немає підстав ні прийняти, ні відкинути гіпотезу про існування автокореляції;
- $d > 2$ — наявність негативної кореляції.

Перед порівнянням з табличними значеннями d – критерій слід перетворити за формулою $d' = 4 - d$.

Встановивши наявність автокореляції залишків, переходять до поліпшення моделі. Якщо ж ситуація виявилася невизначеною, застосовують інші критерії.

Зокрема, можна скористатися першим коефіцієнтом автокореляції, розрахованим за формулою

$$r(1) = \frac{\sum_{t=2}^n \varepsilon_t \varepsilon_{t-1}}{\sum_{t=1}^n \varepsilon_t^2}. \quad (2.4.3)$$

Для прийняття рішення про наявність або відсутність автокореляції в досліджуваному ряді фактичне значення (узяте по модулю) коефіцієнта автокореляції $r(1)$ зіставляється з табличним (критичним) для 5%-го рівня значущості (ймовірності допустити помилку при прийнятті нульової гіпотези про незалежність рівнів ряду). Критичні значення коефіцієнта $r(1)$ наведені у табл. 2.4.1.

При цьому:

- якщо фактичне значення коефіцієнта автокореляції менше за табличне, то гіпотеза про відсутність автокореляції в ряді може бути прийнята та модель вважається адекватною;
- якщо фактичне значення більше за табличне, роблять висновок про наявність автокореляції у ряді динаміки i , отже, модель неадекватна.

2.4.1.4 Відповідність ряду залишків нормальному закону розподілу

Відповідність ряду залишків нормальному закону розподілу можна перевірити за допомогою RS -критерію:

$$RS = \frac{[\varepsilon_{\max} - \varepsilon_{\min}]}{S_{\varepsilon}}, \quad (2.4.4)$$

де ε_{\max} та ε_{\min} – відповідно максимальний та мінімальний рівні ряду залишків; S_{ε} – середньоквадратичне відхилення ряду залишків:

$$S_{\varepsilon} = \sqrt{\sum_{t=1}^n \frac{\varepsilon_t^2}{n-1}}.$$

Якщо розрахункове значення RS потрапляє між табульованими межами (з табл. 2.4.1) із заданим рівнем ймовірності, то гіпотеза про нормальний розподіл ряду залишків приймається. Модель вважається адекватною. В цьому випадку допускається будувати довірчий інтервал прогнозу.

Якщо усі пункти перевірки (усіх перерахованих властивостей залишкової компоненти ряду) дають позитивний результат, то обрана трендова модель є адекватною реальному ряду динаміки, і, отже, її можна використовувати для побудови прогнозних оцінок. Інакше (якщо хоч би одна з властивостей не виконується) – модель потрібно покращувати.

Розглянемо на наступному прикладі застосування всіх зазначених критеріїв, виконання яких необхідне для того, щоб модель прогнозу була адекватною.

Приклад 2.4.1. Нехай заданий часовий ряд (табл. 2.4.2).

Таблиця 2.4.2

Заданий часовий ряд

t	1	2	3	4	5	6	7	8	9
y_t	41	46	49	48	65	55	61	59	65

На основі даних, наведених у табл. 2.4.2, необхідно:

- 1) побудувати лінійну модель $\hat{y}_t = a_0 + a_1 t$, параметри якої оцінити МНК;
- 2) оцінити адекватність побудованої моделі на основі дослідження:
 - випадковості залишкової компоненти за критерієм піків;
 - незалежності рівнів ряду залишків за d – критерієм (в якості критичних значень слід використовувати рівні $d_1 = 1,08$ й $d_2 = 1,36$) та за першим коефіцієнтом автокореляції, критичний рівень якого $r(1) = 0,36$;
 - нормальності розподілу залишкової компоненти за RS -критерієм з критичними рівнями 2,67 – 3,69;
- 3) для оцінки точності моделі використовуйте середньоквадратичне відхилення та середню (по модулю) відносну помилку;

- 4) побудувати точковий та інтервальний прогнози на два кроки вперед (для ймовірності $P = 70\%$ використовуйте коефіцієнт, що дорівнює 1,12);
- 5) відобразити на графіку фактичні дані, результати розрахунків та прогнозування.

Розв'язання:

1) **Побудова лінійної моделі (кривої зросту) прогнозу.**

Модель будується у вигляді $\hat{y}_t = a_0 + a_1 t$, де параметри моделі оцінюються за МНК-методом: $a_0 = 40,5$, $a_1 \approx 2,77$. Отже, модель прогнозу має вигляд $\hat{y}_t = 40,5 + 2,77 t$.

За моделлю $\hat{y}_t = 40,5 + 2,77 t$ визначаємо прогнозні значення часового ряду та, порівнюючи їх з початковими даними, отримуємо ряд залишків, що підлягає оцінюванню (див. табл. 2.4.3).

Таблиця 2.4.3

Ряд залишків

t	Прогнозні значення, \hat{y}_t	Ряд залишків, $\varepsilon_t = y_t - \hat{y}_t$
1	43,27	-2,27
2	46,03	-0,03
3	48,80	0,20
4	51,57	-3,57
5	54,33	10,67
6	57,10	-2,10
7	59,87	1,13
8	62,63	-3,63
9	65,40	-0,40

2) **Оцінка якості побудованої моделі.**

Для цього досліджуємо адекватність моделі.

Модель є адекватною, якщо математичне сподівання значень залишкового ряду близьке або дорівнює нулю та якщо значення залишкового ряду випадкові, незалежні й підпорядковані нормальному закону розподілу.

При **перевірці незалежності (відсутність автокореляції)** визначається відсутність у ряді залишків систематичної складової, наприклад, за допомогою d – критерію Дарбіна-Уотсона за формулою (2.4.2):

$$d = 2,84, \quad d' = 4 - 2,84 = 1,16.$$

Оскільки d' потрапило до інтервалу $d_1 < d' < d_2$, то за цим критерієм не можна зробити висновок про виконання властивості незалежності.

Необхідно обчислити коефіцієнт автокореляції першого порядку за формулою (2.4.3):

$$r(1) = -0,44,$$

тобто фактичне значення по модулю більше за табличне. Це означає, що у ряді динаміки є автокореляція, отже, модель за цим критерієм неадекватна.

Перевірку випадковості рівнів ряду залишків проведемо на основі критерію поворотних точок за формулою (2.4.1), тобто

$$p > \left[\frac{2}{3}(n-2) - 1,96 \sqrt{\frac{16n-29}{90}} \right].$$

У нашому випадку $p = 6$ (див. рис. 2.4.1) – це значення: 0,2; -3,57; 10,67; -2,1; 1,13; -3,63.

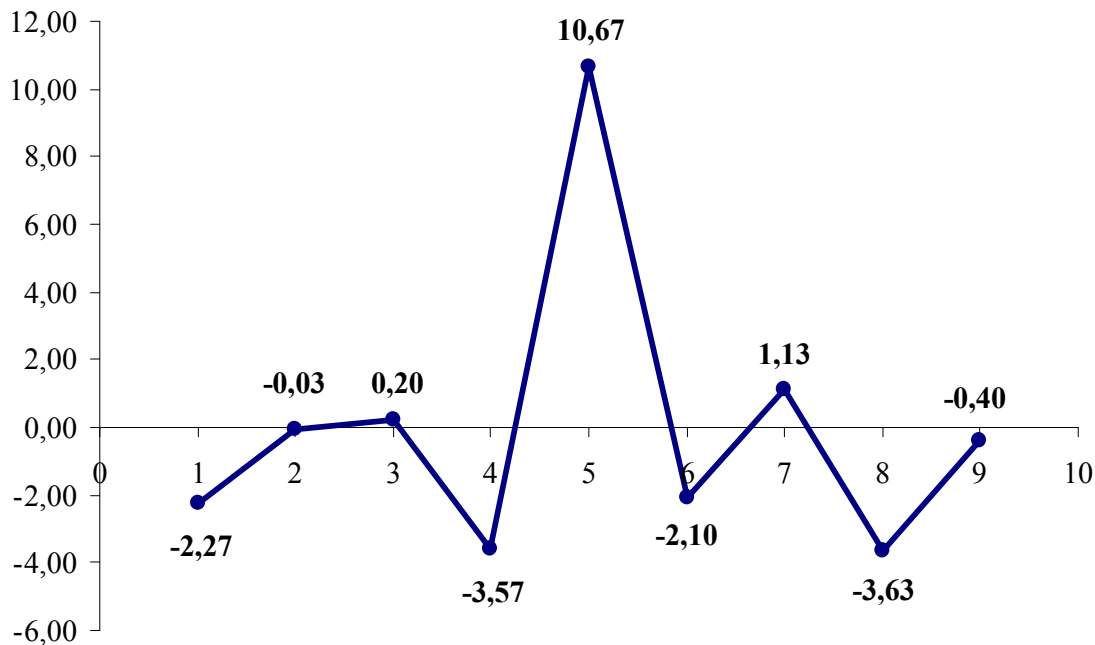


Рисунок 2.4.1 – Графік залишків

Досліджувана нерівність виконується ($6 > 2$). Отже, виконується і властивість випадковості. Модель за цим критерієм адекватна.

Відповідність ряду залишків нормальному закону розподілу визначимо за допомогою RS -критерію:

$$RS = \frac{[\varepsilon_{\max} - \varepsilon_{\min}]}{S_{\varepsilon}} = \frac{[10,67 - (-3,63)]}{4,34} = 3,23.$$

Розрахункове значення потрапляє до інтервалу $(2,67 - 3,69)$, отже, виконується властивість нормальності розподілу. Модель за цим критерієм адекватна.

Перевірка рівності нулю математичного сподівання рівнів ряду залишків.

У нашому випадку $\bar{\varepsilon} = 0$, тому гіпотеза про рівність математичного сподівання значень залишкового ряду нулю виконується.

У табл. 2.4.4 зібрані дані аналізу ряду залишків.

Аналіз ряду залишків

Властивість, що перевіряється	Використовувані статистики		Межа		Висновок
	найменування	значення	нижня	верхня	
Незалежність	<i>d-критерій Дарбіна-Уотсона</i>	$d = 2,84$ $d' = 1,16$	1,08	1,36	не можна зробити висновок за критерієм $ r(1) > 0,36$ неадекватна
	коефіцієнт автокореляції $r(1)$	-0,44			
Випадковість	Критерій піків (поворотних точок)	$p = 6$	2		$6 > 2$ адекватна
Нормальність	<i>RS-критерій</i>	3,23	2,67	3,69	адекватна
Середнє = 0?	<i>t-статистика Стьюдента</i>	0,0	2,306		$0 < 2,306$ адекватна
Висновок: модель статистично неадекватна					

3) Побудова точкового та інтервального прогнозів на 2 кроки вперед.

На основі побудованої моделі розраховуються точкові й інтервальні прогнози.

Точковий прогноз на основі часових моделей отримується підстановкою до моделі (рівняння тренду) відповідного значення чинника часу, тобто $t = n + 1, n + 2, \dots, n + k$, де k – кількість кроків, на які ми поглиблюємося в майбутнє (для нашої моделі $k = 2$).

Для нашої моделі отримаємо точковий прогноз на 2 кроки вперед:

$$\hat{y}_{10} = 40,5 + 2,77 \cdot 10 = 68,17, \quad \hat{y}_{11} = 40,5 + 2,77 \cdot 11 = 70,93.$$

При цьому точний збіг фактичних даних і прогностичних точкових оцінок, отриманих шляхом екстраполяції кривих, що характеризують тенденцію, має малу ймовірність. У зв'язку з цим доцільною є побудова інтервальних прогнозів.

Інтервальні прогнози будуються на основі точкових прогнозів. Для їх побудови будуються довірчі інтервали.

Довірчим інтервалом називається такий інтервал, відносно якого можна із задалегідь обраною ймовірністю стверджувати, що він містить значення прогнозованого показника.

Ширина інтервалу залежить від якості моделі, тобто міри близькості її до фактичних даних, числа спостережень, горизонту прогнозування та обраного користувачем рівня ймовірності.

При побудові довірчого інтервалу прогнозу розраховується величина $U(k)$, яка для лінійної моделі має вигляд

$$U(k) = S_{\hat{y}} t_{\alpha} \sqrt{1 + \frac{1}{n} + \frac{(n+k-\bar{t})^2}{\sum_{i=1}^n (t_i - \bar{t})^2}},$$

де $S_{\hat{y}}$ – стандартна помилка (середньоквадратичне відхилення від моделі)

$$S_{\hat{y}} = \sqrt{\sum_{t=1}^n \frac{\varepsilon_t^2}{n-m-1}};$$

m – кількість чинників в моделі (для лінійної моделі $m=1$); t_{α} – табличне значення t -статистики Стьюдента при заданому рівні значущості α та числі спостережень. Так, при довірчій ймовірності 70% значення рівня значущості $\alpha = 0,3$. Тоді критерій Стьюдента $t_{\alpha} = 1,12$ при $\nu = n - 2 = 7$.

Для нашої моделі

$$U(1) = 4,64 \cdot 1,12 \sqrt{1 + \frac{1}{9} + \frac{(9+1-5)^2}{60}} = 6,42; \quad U(2) = 6,79.$$

Довірчий інтервал прогнозу матиме наступні межі:

- верхня межа прогнозу : $\hat{y}_{n+k} + U(k)$;
- нижня межа прогнозу : $\hat{y}_{n+k} - U(k)$.

Для нашої моделі довірчий інтервал прогнозу :

- при $k=1$: $68,17 \pm 6,42$;
- при $k=2$: $70,93 \pm 6,79$.

Отримані результати за моделлю представимо в табл. 2.4.5.

Таблиця 2.4.5

Результати прогнозування

$n+k$	$U(k)$	Прогноз	Довірчий інтервал	
			верхня межа	нижня межа
10	$U(1) = 6,42$	68,17	74,59	61,75
11	$U(2) = 6,79$	70,93	77,73	64,14

Якщо побудована модель адекватна, то з вибраною користувачем ймовірністю можна стверджувати, що при збереженні закономірностей розвитку, що склалися, прогнозована величина потрапляє в інтервал, утворений верхньою та нижньою межею. У нашому випадку, оскільки модель не адекватна, затверджувати це не можна.

Побудуємо графік для нашого прикладу (рис. 2.4.2).

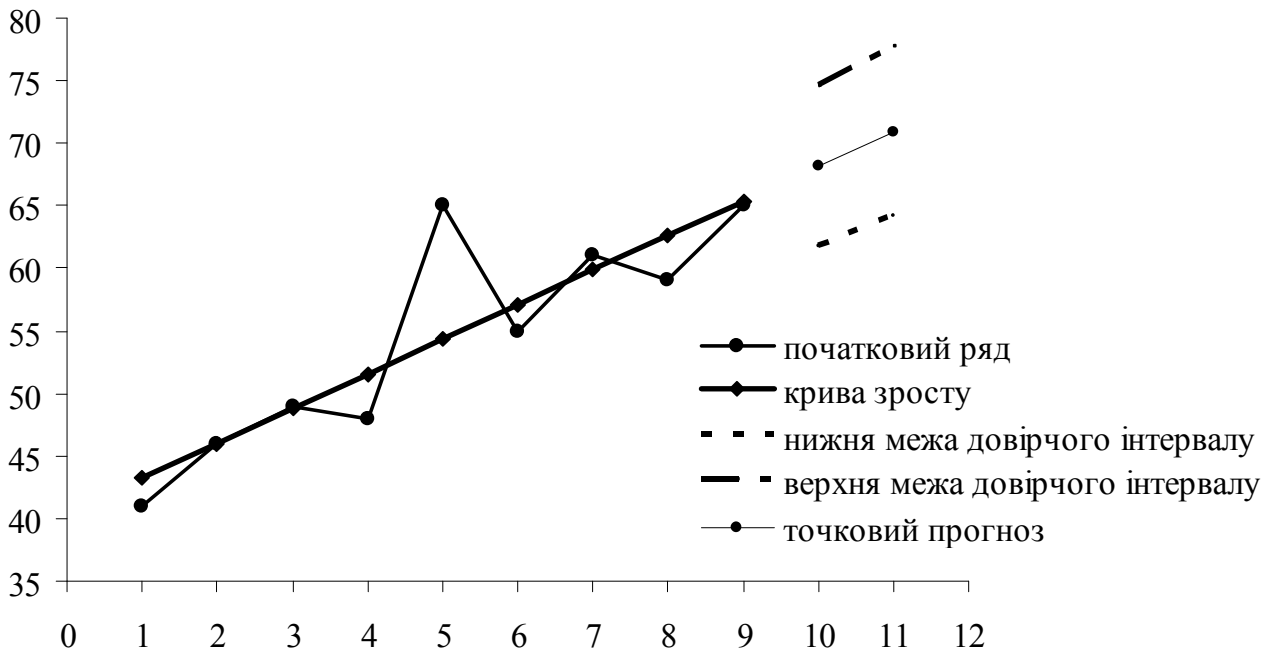


Рисунок 2.4.2 – Фактичні і прогнозні значення часового ряду

2.4.2 Методи аналізу якості побудованих прогнозів

Усі показники, використовувані для аналізу якості прогнозу, можна розділити на три групи: абсолютні, порівняльні та якісні.

2.4.2.1 Абсолютні показники

До них відносяться такі показники, які дозволяють кількісно визначити величину помилки прогнозу в одиницях виміру прогнозованого об'єкту або в %. Це середньоквадратична помилка σ_t , абсолютна помилка Δ_{np} й середня абсолютна помилка $\bar{\Delta}_{np}$, відносна помилка ε_{np} та середня відносна помилка $\bar{\varepsilon}_{np}$. Розглянемо їх обчислення.

Абсолютна помилка прогнозу може бути визначена як різниця між фактичним значенням (y_t) та прогнозом (y_t^*), тобто

$$\Delta_{np} = y_t - y_t^*, \quad (2.4.5)$$

середнє абсолютне значення помилки:

$$\bar{\Delta}_{np} = \frac{\sum_{i=1}^n |y_t - y_t^*|}{n}. \quad (2.4.6)$$

Середнє абсолютне значення $\bar{\Delta}_{np}$ завжди невід'ємне, навіть коли Δ_{np} від'ємне (тобто при $y_t^* > y_t$).

Середньоквадратична помилка прогнозу розраховується за формулою

$$\sigma_t = \sqrt{\frac{\sum_{t=1}^n (y_t^* - y_t)^2}{n}}, \quad (2.4.7)$$

де n – період упередження.

Слід зазначити, що існує зв'язок між середнім абсолютним відхиленням $\bar{\Delta}_{np}$ та стандартним відхиленням σ_t . Для великого класу статистичних розподілів значення стандартного відхилення дещо більше значення середнього абсолютного відхилення та строго пропорційно йому. Константа пропорційності для різних розподілів коливається між 1,2 та 1,3. Найчастіше на практиці береться 1,25, тому

$$\sigma_t = 1,25 \cdot \bar{\Delta}_{np}. \quad (2.4.8)$$

Недоліком даних показників є те, що значення цих характеристик істотно залежить від масштабу виміру рівнів досліджуваних явищ. Тому абсолютна помилка прогнозу Δ_{np} може бути виражена у відсотках відносно фактичних значень показника. Така помилка називається **відносною помилкою прогнозу** та визначається як

$$\varepsilon_{np} = \frac{y_t - y_t^*}{y_t} 100, \quad (2.4.9)$$

а **середня відносна помилка** розраховується наступним чином:

$$\bar{\varepsilon}_{np} = \frac{1}{n} \sum_{t=1}^n \frac{|y_t - y_t^*|}{y_t} 100. \quad (2.4.10)$$

Цей показник, як правило, використовується при порівнянні точності прогнозів різнорідних об'єктів прогнозування, оскільки характеризує відносну точність прогнозу.

Наведемо типові значення ε_{np} для середньострокових прогнозів та їх інтерпретації:

ε_{np} , %	Інтерпретація
<10	висока точність
10-20	хороша точність
20-50	задовільна точність
>50	незадовільна точність

Якщо на практиці $y_t > 0$, то ε_{np} стає нескінченною величиною. Тому при прогнозуванні в більшості випадків, для яких обчислюється ε_{np} , дані не можуть набувати нульових значень. Якщо ж $y_t = 0$, доцільно пропустити обчислення, зменшуючи при цьому і число n на одиницю.

Подібний підхід до оцінки точності прогнозу можливий тільки за умови, коли період упередження вже закінчився та є фактичні дані про прогнозований показник, а також при ретроспективному прогнозуванні. У останньому випадку наявна інформація ділиться на дві частини, одна з яких охоплює більше ранні дані, а інша – пізніші. За допомогою даних першої групи оцінюються параметри моделі прогнозу, а дані другої групи розглядаються як фактичні дані прогнозованого показника. Отримана ретроспективно помилка прогнозу в якійсь мірі характеризує точність вживаної методики прогнозування.

Середня абсолютна та середньоквадратична помилки фіксують середнє значення помилки на кожному кроці прогнозу без урахування цієї помилки. Середня помилка дозволяє визначити, який вид помилки є найбільш типовим – недооцінка або переоцінка прогнозованого показника.

Необхідно мати на увазі, що $\bar{\Delta}_{np}$ та σ_{np} дорівнюють нулю тільки тоді, коли $y_t = y_t^*$ для кожного t , тобто у разі досконалого прогнозу. Аналогічне твердження несправедливе для абсолютної помилки Δ_{np} , оскільки тут може мати місце взаємопогашення помилок. Для розрахунку цих показників можуть бути використані як абсолютні величини змінних, так і їх прирости або темпи приростів.

2.4.2.2 Порівняльні показники

Ці показники засновані на порівнянні помилки даного прогнозу з еталонними прогнозами певного виду.

Один з типів таких показників (K) може бути в загальному вигляді представлений таким чином:

$$K = \sqrt{\frac{\sum_{t=1}^n (p_t - y_t)^2}{\sum_{t=1}^n (p_t^* - y_t)^2}}, \quad (2.4.11)$$

де p_t^* – прогнозоване значення величини еталонного прогнозу. В якості еталонного прогнозу може бути обрана проста екстраполяція, постійний темп приросту і т. п.

Частковим випадком показників такого типу є коефіцієнт невідповідності, в якому $p_t^* = 0$ для усіх t :

$$KH = \sqrt{\frac{\sum_{t=1}^n (y_t^* - y_t)^2}{\sum_{t=1}^n y_t^2}}, \quad (2.4.12)$$

При цьому $KH = 0$ у разі досконалого прогнозу; $KH = 1$, коли прогноз має ту ж помилку, що і «найвна» екстраполяція незмінності; KH не має верхньої кінцевої межі.

Можна побудувати різні *модифікації коефіцієнта невідповідності*. Розглянемо деякі з них.

1. Коефіцієнт невідповідності KH_1 , обчислюваний як відношення середньоквадратичної помилки прогнозу до тієї ж помилки, яка мала б місце, якщо прийняти в якості прогнозу для кожного року середнє значення змінної за увесь період:

$$KH_1 = \sqrt{\frac{\sum_{t=1}^n (y_t^* - y_t)^2}{\sum_{t=1}^n (\bar{y} - y_t)^2}}, \quad (2.4.13)$$

$$\text{де } \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i.$$

Якщо $KH_1 > 1$, то прогноз на рівні середнього значення дав би кращий результат, чим наявний прогноз.

2. Коефіцієнт розбіжності V , що обчислюється як відношення середньоквадратичної помилки прогнозу до тієї ж помилки, яка мала б місце, якщо прийняти в якості прогнозу для кожного року екстраполююче значення за аналітичним трендом, тобто

$$V = \sqrt{\frac{\sum_{t=1}^n (y_t^* - y_t)^2}{\sum_{t=1}^n (y_t - \bar{y})^2}}, \quad (2.4.14)$$

де \bar{y}_t – екстраполююче значення досліджуваної величини на момент t .

Як і у попередньому випадку, $V > 1$ означає, що прогноз методом простої екстраполяції дає кращий результат.

До порівняльних показників слід віднести і *коефіцієнт кореляції* R між прогнозованими і фактичними значеннями змінної:

$$R = \frac{\frac{1}{n} \sum_{t=1}^n (y_t^* - \bar{y}^*)(y_t - \bar{y})}{\sqrt{\frac{1}{n} \sum_{t=1}^n (y_t^* - \bar{y}^*)^2 \cdot \frac{1}{n} \sum_{t=1}^n (y_t - \bar{y})^2}}. \quad (2.4.15)$$

Одним з недоліків використання коефіцієнта кореляції в якості вимірника точності прогнозів є те, що повна додатна кореляція не припускає досконалого прогнозу, а говорить лише про існування лінійної залежності між рядами прогнозних та фактичних величин, тобто якщо $R = 1$, існують такі константи α і β ($\beta > 0$), що $y_t^* = \alpha + \beta y_t$. При цьому для досконалого прогнозу потрібне, щоб $\alpha = 0$ і $\beta = 1$. Внаслідок цього коефіцієнт кореляції найбільш придатний для аналізу прогнозів змінних, які циклічно розвиваються.

2.4.2.3 Якісні показники

До них слід віднести показники, які дозволяють провести деякий аналіз видів помилок прогнозів, розкласти їх на які-небудь складові. Особливо такий аналіз важливий для змінних, що циклічно змінюються, коли необхідно прогнозувати не лише загальний напрям розвитку, але і поворотні точки циклу.

Одним з методів такого аналізу є діаграма «прогноз-реалізація». Суть методу полягає в побудові області точкових прогнозів в координатах, в яких по одній осі відкладається реальне значення змінної, по іншій – прогнозоване (рис. 2.4.3).

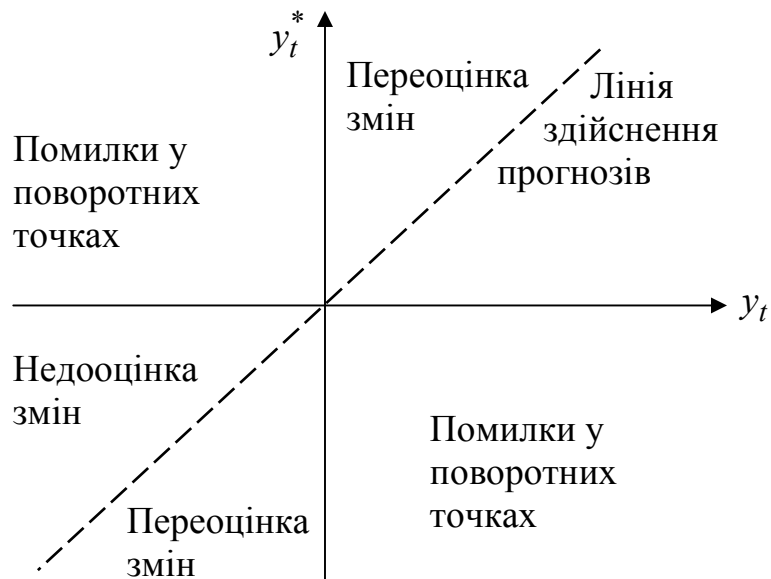


Рисунок 2.4.3 – Діаграма «прогноз-реалізація»:
 y_t – фактичні значення змінних, y_t^* – прогноз

Наприклад, для розглянутого прикладу 2.4.1 діаграма «прогноз-реалізація» буде виглядати наступним чином (рис. 2.4.4).

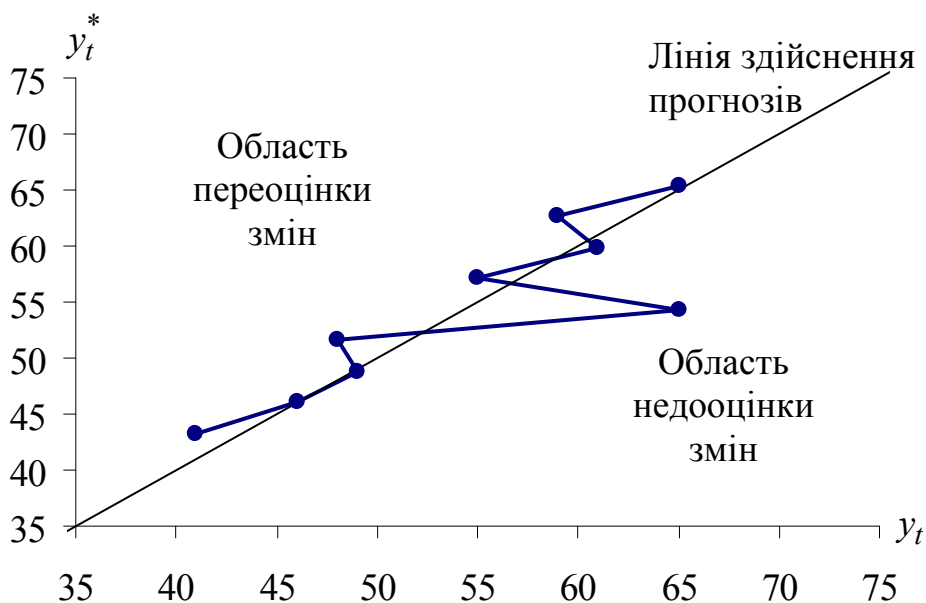


Рисунок 2.4.4 – Діаграма «прогноз-реалізація» для прикладу 2.4.1

Використання діаграми дозволяє змістовно оцінити якість різних прогнозів, розрахувати деякі коефіцієнти, що аналізують якість прогнозування поворотних точок, виділити найбільш типові помилки (недооцінки або переоцінки змін).

Для аналізу загальнішого вигляду помилок прогнозів може бути використана інша класифікація за джерелами помилок.

Г.Тейл розклав помилку прогнозу на долі невідповідності. Усі розглянуті вище показники якості прогнозу містять у своїй основі середньоквадратичну помилку. Квадрат її можна представити таким чином:

$$\frac{1}{n} \sum_{t=1}^n (y_t^* - y_t)^2 = (\bar{y}^* - \bar{y})^2 + (S_{y^*} - S_y)^2 + 2(1-R)S_{y^*}S_y, \quad (2.4.16)$$

де \bar{y}^* і \bar{y} – середні значення прогнозних і фактичних значень змінних; R – коефіцієнт кореляції між прогнозними і фактичними значеннями; S_{y^*} і S_y – дисперсії прогнозних і фактичних значень змінних:

$$S_{y^*} = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t^* - \bar{y}^*)^2}, \quad S_y = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \bar{y})^2}.$$

Таке розкладання середньоквадратичної помилки дозволяє досліджувати її природу. Поділимо праву частину рівняння на ліву, що дає нам в правій частині

$$V^\mu = \frac{(\bar{y}^* - \bar{y})^2}{\frac{1}{n} \sum_{t=1}^n (y_t^* - y_t)^2}; \quad V^S = \frac{(S_{y^*} - S_y)^2}{\frac{1}{n} \sum_{t=1}^n (y_t^* - y_t)^2}; \quad V^l = \frac{2(1-R)S_{y^*}S_y}{\frac{1}{n} \sum_{t=1}^n (y_t^* - y_t)^2}. \quad (2.4.17)$$

Подані у (2.4.17) коефіцієнти – так звані коефіцієнти (долі) невідповідності:

а) V^μ – **доля зміщення** – показує наявність помилки в оцінці центральної тенденції, тобто $V^\mu = 0$, коли середнє арифметичне значення прогнозів відрізняється від середнього арифметичного значення фактичних даних. На діаграмі «прогноз-реалізація» відсутність цієї помилки означає, що центр тяжіння точкових прогнозів лежить на лінії досконалих прогнозів.

б) V^S – **доля дисперсії** – відображує міру збігу стандартних відхилень прогнозу та фактичних значень, і $V^S = 0$ у тому випадку, коли $S_{y^*} = S_y$. Таким чином, цей показник відображає відповідність міри нестійкості прогнозних значень міри нестійкості фактичної динаміки.

в) V^l – **доля коваріації**. Коефіцієнт $V^l = 0$, коли коефіцієнт кореляції між прогнозними та фактичними значеннями дорівнює одиниці. На діаграмі «прогноз-реалізація» це відповідає випадку, коли усі точки лежать на одній прямій. Аналіз цього критерію дозволяє виділити ті випадки, коли прогноз, будучи задовільним за двома першими критеріями, має взаємну компенсацію помилок для різних спостережень.

Очевидно, що сума їх складає одиницю: $V^\mu + V^S + V^l = 1$.

Усі розглянуті вище показники точності прогнозу використовуються при перевірці точності прогнозу, отриманого у вигляді точкових оцінок. Усі вони мають велику цінність при зіставленні різних методик прогнозування та можуть бути використані тільки за наявності інформації про фактичні значення досліджуваного показника. Якщо ж практично цих даних немає, то в цьому випадку проблема точності може розглядатися в плані зіставлення апіорних якостей, властивостей, властивих альтернативним прогностичним моделям. Так, при прогнозуванні статистичними методами поняття апіорної точності прогнозу можна пов'язати з розміром довірчого інтервалу. В цьому випадку модель прогнозу вважається точнішою, якщо при тій самій довірчій ймовірності вона дає вузьчий довірчий інтервал.

Вибір показників точності прогнозів залежить від задач, які ставить перед собою дослідник при аналізі точності прогнозів. Необхідно пам'ятати, що перевірка точності одного прогнозу мало що може дати дослідникові, оскільки на формування досліджуваного явища впливає безліч різноманітних чинників. Тому повний збіг або значна розбіжність прогнозів та його реалізації може бути наслідком просто особливо сприятливих (чи несприятливих) збігів обставин. Звідси витікає, що про якість прогнозів вживаних методик і моделей можна судити лише по сукупності зіставлень прогнозів та їх реалізацій.

2.4.3 Питання для самоконтролю

1. Назвіть основні методи перевірки адекватності побудованих моделей.
2. В чому полягає перевірка рівності нулю математичного сподівання рівнів ряду залишків?
3. Розкрийте сутність умови випадковості виникнення окремих відхилень від тренду. Як проводиться перевірка такої умови?
4. Як проводиться визначення наявності (відсутності) автокореляції у відхиленнях від моделі зросту?
5. Що називається поворотними точками? Як вони обчислюються?
6. Як перевіряється відповідність ряду залишків нормальному закону розподілу?
7. В чому полягає особливість абсолютних показників точності прогнозів? Які показники до них відносяться?
8. Що показує коефіцієнт невідповідності? Які існують його модифікації?
9. Як будується діаграма «прогноз-реалізація»? Що вона відображує?
10. Що показують долі невідповідності? У яких випадках їх значення дорівнюють нулю?

2.4.4 Рекомендована література



Для самостійного вивчення теми рекомендується література: основна – [1, 4, 5, 11], додаткова – [4].

3. ПРАКТИЧНА ЧАСТИНА. МЕТОДИЧНІ ПОРАДИ ДО ВИКОНАННЯ ЛАБОРАТОРНИХ РОБІТ ЗА КУРСОМ

3.1 Значення лабораторної роботи при вивченні курсу «Математичне моделювання виробничих процесів»

Навчальними планами для студентів математичного факультету денного відділення напряму підготовки «Прикладна математика» передбачено виконання лабораторних робіт з дисципліни «Математичне моделювання виробничих процесів» за розділами: «Аналіз динаміки виробничих процесів. Прості методи прогнозування виробничих процесів», «Адаптивні методи прогнозування виробничих процесів. Оцінка побудованих моделей та прогнозів». Лабораторна робота є важливою формою вивчення курсу «Математичне моделювання виробничих процесів» і виступає засобом набуття й поглиблення знань та практичних навичок студентів.

Виконання лабораторних робіт забезпечує студентам систематичне вивчення курсу на основі опрацювання рекомендованих підручників і навчальних посібників, самостійний пошук необхідних літературних джерел та матеріалів, їх аналіз та узагальнення.

Метою виконання лабораторних робіт є вивчення та закріплення навчального матеріалу з тем «Часові ряди та їх попередній аналіз», «Методи екстраполяції в прогнозуванні виробничих процесів», «Адаптивні методи прогнозування», «Методи аналізу якості побудованих моделей та прогнозів виробничих процесів», отриманого на лекціях, лабораторних заняттях та при самостійному вивченні матеріалу.

3.2 Порядок визначення варіанту лабораторної роботи

Завдання до лабораторних робіт за курсом наведені у посібнику у даному розділі. Кожне із завдань до лабораторних робіт виконується студентами індивідуально. Варіанти лабораторних робіт подані у таблицях після завдань до лабораторних робіт. Номер варіанта у таблицях відповідає номеру студента за списком у журналі.

3.3 Рекомендації до виконання та вимоги до захисту лабораторних робіт

У лабораторній роботі студент повинен розв'язати подані задачі методами, зазначеними у завданнях, а також привести графіки, де це потрібно. Також повинен показати набуті теоретичні знання з цього курсу, розкриваючи зміст отриманих результатів.

Перед виконанням кожного завдання лабораторної роботи, рекомендується опрацювати лекційний матеріал за темою та наведені у 2 розділі даного посібника теоретичні відомості, а також ознайомитися з відповідними темами у зазначених до кожної теми літературних джерелах для самоопрацювання.

Якість лабораторної роботи оцінюється, перш за все, по тому, наскільки студентом самостійно, правильно розв'язані поставлені задачі і глибоко розглянутий зміст отриманих розв'язків. При захисті лабораторної роботи студенту можуть бути задані додаткові питання з теорії, яка стосується відповідної теми лабораторної роботи.

3.4 Критерії оцінювання знань студентів

Оцінювання знань студентів під час поточного контролю відбувається на підставі наступних критеріїв:

- правильність відповідей (правильне, чітке, достатньо глибоке викладення теоретичних понять);
- ступінь усвідомлення програмного матеріалу;
- новизна навчальної інформації; рівень використання наукових (теоретичних) знань;
- вміння користуватися засвоєними теоретичними знаннями на практиці.

Відповіді студентів оцінюються також за формою з точки зору логічності, чіткості, виразності викладу.

3.5 Завдання до лабораторних робіт за темами дисципліни «Математичне моделювання виробничих процесів»

3.5.1 Лабораторна робота 1. Попередній аналіз часових рядів

3.5.1.1 Методичні поради до вивчення теми

Виконання лабораторної роботи за даною темою надасть студентам можливість оволодіти практичними навичками проведення попереднього аналізу часових рядів у частині коригування рівнів часового ряду (виявлення та усунення аномальних значень рівнів ряду) та ознайомитися з основними критеріями та методами перевірки гіпотези про існування тренду у часовому ряді.

Для виконання лабораторної роботи передбачається наявність у студентів знань з основних питань теми «Часові ряди та їх попередній аналіз», а саме: вихідних понять, сутності, характерних рис, видів та основних компонентів часових рядів, декомпозиції часового ряду, методах виявлення та усунення аномальних значень рівнів часових рядів, основних критеріях і методах перевірки гіпотези про наявність тенденції у часових рядах. Крім того, необхідним є володіння практичними навичками застосування техніки регресійного аналізу.



Для самостійного вивчення теми рекомендується література:
основна – [2-5, 7, 8, 10, 11].

3.5.1.2 Завдання на лабораторну роботу

1. Перевірте заданий часовий ряд (див. табл. 3.5.1) на наявність у ньому аномальних спостережень та усунете їх, якщо вони мають місце.

2. За даними нового часового ряду (з усіма нормальними рівнями ряду) перевірте гіпотезу про наявність тренду за допомогою відомих 5 критеріїв й методів.

3. Зробіть загальний висновок щодо існування тренду у часовому ряді.

Таблиця 3.5.1

Варіанти завдання для виконання лабораторної роботи №1 (№ варіанта у таблиці відповідає номеру студента за списком у журналі)

	1 вар.	2 вар.	3 вар.	4 вар.	5 вар.	6 вар.	7 вар.	8 вар.	9 вар.	10 вар	11 вар	12 вар
t	y_t	y_t	y_t	y_t	y_t	y_t	y_t	y_t	y_t	y_t	y_t	y_t
1	2,8	1,0	2,6	4,4	3,2	2,3	1,4	3,2	3,3	3,7	6,0	4,6
2	3,3	1,5	3,1	4,9	3,7	2,8	1,9	3,7	3,8	4,2	6,5	5,1
3	4,0	2,2	3,8	5,6	4,4	3,5	2,6	4,4	4,5	4,9	7,2	5,8
4	3,5	1,7	3,3	5,1	3,9	3,0	2,1	3,9	4,0	4,4	6,7	5,3
5	3,1	1,3	2,9	4,7	3,5	2,6	1,7	3,5	3,6	4,0	6,3	4,9
6	3,8	2,0	4,0	5,4	4,2	3,3	2,8	4,6	4,7	5,1	7,4	6,0
7	4,5	2,7	4,7	6,1	4,9	4,0	3,5	5,3	5,4	5,8	8,1	6,7
8	4,0	2,2	4,2	5,6	4,4	3,5	3,0	4,8	4,9	5,3	7,6	6,2
9	3,5	1,7	3,7	5,1	3,9	3,0	2,5	4,3	4,4	4,7	7,1	5,7
10	4,2	2,4	4,4	5,8	4,6	3,7	3,2	5,0	5,0	5,5	7,8	6,4
11	4,9	3,1	5,1	6,5	5,3	4,4	3,9	5,7	5,8	6,2	8,5	7,1
12	4,7	2,9	4,9	6,3	5,1	4,2	3,7	5,5	5,6	6,0	8,3	6,9
13	4,4	2,6	4,6	6,0	4,8	3,9	3,4	5,2	5,3	5,7	8,0	6,6
14	3,7	1,9	3,9	5,3	4,1	3,2	2,7	4,5	4,6	5,0	7,3	5,9
15	4,6	2,8	4,8	6,2	5,0	4,1	3,6	5,4	5,5	5,9	8,2	6,8
16	5,5	3,7	5,7	7,1	5,9	5,0	4,5	6,3	6,4	6,8	9,1	7,7
17	5,0	3,2	5,2	6,6	5,4	4,5	4,0	5,8	5,9	6,3	8,6	7,2
18	4,9	3,1	5,0	6,5	5,3	4,4	3,8	5,6	5,7	6,1	8,4	7,0
19	5,2	3,4	5,4	6,8	5,6	4,7	4,2	6,0	6,1	6,5	8,8	7,4
20	5,7	3,9	5,9	7,3	6,1	5,2	4,7	6,5	6,6	7,0	9,3	7,9
21	4,9	3,1	5,1	6,5	5,3	4,4	3,9	5,7	5,8	6,2	8,5	7,1
22	4,0	2,2	4,2	5,6	4,4	3,5	3,0	4,8	4,9	5,3	7,6	6,2
23	4,6	2,8	4,8	6,2	5,0	4,1	3,6	5,4	5,5	5,9	8,2	6,8
24	5,1	3,3	5,3	6,7	5,5	4,6	4,1	5,9	6,0	6,4	8,7	7,3
25	4,7	2,9	4,9	6,3	5,1	4,2	3,7	5,5	5,6	6,0	8,3	6,9
26	3,9	2,1	4,1	5,5	4,3	3,4	2,9	4,7	4,8	5,2	7,5	6,1
27	15,5	13,7	15,7	17,1	15,9	15,0	4,5	6,3	6,4	6,8	9,1	7,7
28	13,2	11,4	13,4	14,8	13,6	12,7	14,0	14,0	12,1	12,5	14,8	13,4
29	12,0		0,9	3,4	2,2	1,3		1,5		7,0	9,3	7,9
30	1,8				2,2	1,3		11,3		16,7		

3.5.2 Лабораторна робота 2. Екстраполяція в прогнозуванні виробничих процесів

3.5.2.1 Методичні поради до вивчення теми

Виконання лабораторної роботи за даною темою надасть студентам можливість оволодіти практичними навичками застосування простих методів прогнозування екстраполяції, які використовуються в керуванні виробничими процесами в економіці, а саме екстраполяції середнього, екстраполяції за допомогою аналітичних показників динаміки часового ряду, екстраполяції тренду та екстраполяції на основі індексу сезонності. Крім того, у процесі виконання даної роботи студенти набудуть навичок застосування методу «Прогноз екс-пост» до перевірки адекватності моделей та вибору оптимальної моделі за певним критерієм.

Для виконання лабораторної роботи передбачається наявність у студентів знань з основних питань тем «Часові ряди та їх попередній аналіз» та «Методи екстраполяції в прогнозуванні виробничих процесів», а саме: вихідних понять, сутності, видів, основних компонентів та декомпозиції часових рядів; сутності, видів екстраполяції, основних методів простої прогнозування екстраполяції та умов їх застосування. Крім того, необхідним є володіння практичними навичками застосування техніки регресійного аналізу, матричного числення та методів розв'язання систем алгебраїчних рівнянь.



Для самостійного вивчення теми рекомендується література: основна – [1, 2, 4, 5, 9, 10].

3.5.2.2 Завдання на лабораторну роботу

1. На підставі заданого часового ряду (див. табл. 3.5.2):

– побудувати прогноз на $(n+4)$ -й період методом екстраполяції середнього;

– визначити (за допомогою табл. 2.2.2) аналітичні показники динаміки (представити їх у вигляді таблиці) та за допомогою екстраполяції на основі аналітичних показників рядів динаміки скласти прогноз на 6 років за моделлю, обраною за допомогою методу «прогноз екс-пост».

2. На підставі заданого часового ряду (див. табл. 3.5.3) побудувати прогноз на $(n+1)$ -й і $(n+2)$ -й період за допомогою екстраполяції за середнім коефіцієнтом росту.

3. На підставі заданого часового ряду (див. табл. 3.5.4) скласти прогноз попиту на наступний рік, використовуючи метод екстраполяції тренду.

4. За даними, які характеризують обсяг реалізації продукції (див. таблиці 3.5.5-3.5.10), розрахувати індекси сезонності, побудувати сезонну хвилю (таблично та графічно) й прогноз обсягу реалізації продукції на окремі місяці наступного року, використовуючи метод екстраполяції на основі індексу сезонності, якщо очікуваний річний обсяг показника складає 80 тис. т.

Номер варіанта у наведених нижче таблицях 3.5.2-3.5.10 відповідає номеру студента за списком у журналі.

Таблиця 3.5.2

	1 вар.	2 вар.	3 вар.	4 вар.	5 вар.	6 вар.	7 вар.	8 вар.	9 вар.	10 вар.	11 вар.	12 вар.
t	y_t	y_t	y_t	y_t	y_t	y_t	y_t	y_t	y_t	y_t	y_t	y_t
1	48,6	412,2	19,7	20	110,6	32,2	15	12,6	11,8	113,9	72,1	38
2	48,2	412,4	19,5	20,3	113,9	30,9	14,4	12,3	15,9	115,7	74,8	38,4
3	48,7	412,6	19,1	20,2	112,1	28,1	14,8	12,8	14,7	117,2	74,4	37,2
4	48,5	412	19	20,6	111,4	31,9	14,9	12,9	12,5	116,6	75,2	39,8
5	48,4	411,7	18,8	20,1	112,6	30,7	15,5	13,1	13,9	118,2	76,7	40,1
6	48,2	411,9	18,6	19,9	115,7	31,6	15,7	12,2	14,6	117,7	72,2	40,9
7	48,8	412,1	19,2	19,7	118,2	32	16	12,7	14,8	117,8	71,7	47,8
8	48,9	412,3	19,3	20,1	118,5	34,3	16,1	12,9	15,4	118,3	73,2	44,7
9	49,1	412,4	18,9	20,4	120,1	38,4	16,3	13	16,1	119	75,5	43,2
10	48,9	412,9	19	20,8	119,9	41,1	15,7	13,3	15,5	119,8	78,3	41
11	48,6	413	19,2	20,5	117,3	39,3	15,1	13,6	15,7	121,2	80	42,4
12	48,7	413,2	19,1	20,7	119,4	38,5	14,3	14	15	119,9	83,7	42,1
13	48,4	412,9	19,8	21,1	121,8	40,3	14,8	13,2	16,2	119,5	82,4	43,7

Таблиця 3.5.3

	1 вар.	2 вар.	3 вар.	4 вар.	5 вар.	6 вар.	7 вар.	8 вар.	9 вар.	10 вар.	11 вар.	12 вар.
t	y_t	y_t	y_t	y_t	y_t	y_t	y_t	y_t	y_t	y_t	y_t	y_t
2000	127,5	148,7	114,3	124,4	510,5	465,2	454,7	352,7	49,7	847,6	292,4	605,7
2009	138,3	161,4	128,9	136,8	528,3	478,5	495,1	396,0	56,1	876,2	315,9	648,2

Таблиця 3.5.4

	1 вар.	2 вар.	3 вар.	4 вар.	5 вар.	6 вар.	7 вар.	8 вар.	9 вар.	10 вар.	11 вар.	12 вар.
t	y_t	y_t	y_t	y_t	y_t	y_t	y_t	y_t	y_t	y_t	y_t	y_t
1	253,5	322,2	159,7	121	254,6	493,2	145	562,6	311,8	150,1	176,2	435,9
2	259,2	330,4	161,5	127,3	271,3	500,9	154,4	569,3	325,9	159,6	194,9	455,4
3	266,1	342,6	164,1	135,2	290,8	518,1	166,8	574,8	334,7	169,2	212,4	467,2
4	273,5	349,2	167	141,6	301,9	539,9	177,3	582,9	341,5	182,6	237,2	479,8
5	276,4	356	172,8	148	312,8	560,2	184,5	587,8	357,9	188,2	253,7	485,1
6	283,2	359,7	176,6	152,4	326,2	568,6	186,7	593,2	368,6	195,7	270,1	500,9
7	287,8	366,1	179,2	165,9	341,5	580	193	601,5	378,8	202,5	288,7	507,8
8	294,5	368,7	184,3	172,1	355,9	603,3	197,6	612,9	391,4	210,3	297,2	514,7

Таблиця 3.5.5

	1 варіант					2 варіант				
	1 рік	2 рік	3 рік	4 рік	5 рік	1 рік	2 рік	3 рік	4 рік	5 рік
січень	23,5	23,8	24,7	26	25,6	1,2	1,5	1,6	1,8	2,1
лютий	24,2	24,4	25,5	26,1	26,5	1,3	1,7	1,7	1,9	2,4
березень	25,1	25,6	25,7	26,2	26,7	1,5	1,8	1,9	2	2,5
квітень	26,5	26,2	26,3	26,8	27	1,7	1,9	2	2,1	2,6
травень	27,4	27	27,2	27,3	27,8	1,4	1,6	1,7	1,8	2,3
червень	27,2	26,7	27	27,1	27,5	1,2	1,4	1,5	1,6	2,1
липень	26,8	25,9	26,3	26,5	27	1,1	1,3	1,2	1,3	1,8
серпень	26,5	25,0	25,9	26,2	26,4	1,8	1,6	1,7	1,6	2
вересень	27	26	26,4	26,6	26,9	1,9	2	2,1	2,3	2,4
жовтень	27,4	27,3	27,1	27,5	27,6	2	2,1	2,3	2,4	2,6
листопад	27,8	28	28,1	28,6	28,7	2,1	2,3	2,5	2,7	2,9
грудень	27,9	28,7	29	29,3	29,8	2,3	2,5	2,6	2,9	3,1

Таблиця 3.5.6

	3 варіант					4 варіант				
	1 рік	2 рік	3 рік	4 рік	5 рік	1 рік	2 рік	3 рік	4 рік	5 рік
січень	16,2	16,5	16,8	17	17,5	3,1	3,6	4,7	6	5,6
лютий	16,3	16,6	16,9	17,1	17,6	4,2	4,4	5,5	6,1	6,5
березень	16,5	16,7	17,1	17,3	17,9	5,3	5,6	5,7	6,2	6,7
квітень	16,2	16,3	16,8	17	17,3	6	6,2	6,3	6,8	7
травень	16,1	16,2	16,6	16,7	17	7,4	7,3	7,2	7,3	7,8
червень	16	15,9	16,2	16,4	16,5	7,2	6,7	7	7,1	7,5
липень	16,2	16,4	16,4	16,6	16,9	6,8	6,9	6,3	6,8	7,2
серпень	16,4	16,5	16,6	16,7	17,1	6,5	6,6	5,9	6,2	6,4
вересень	16,8	16,7	16,8	17	17,2	7	7,2	6,4	6,6	7
жовтень	17	16,9	17	17,2	17,4	7,4	7,8	7,1	7,4	7,6
листопад	17,1	17,2	17,3	17,5	17,6	7,6	8,1	8,2	8,6	8,8
грудень	17,3	17,4	17,6	17,8	18	7,9	8,7	9	9,3	9,4

Таблиця 3.5.7

	5 варіант					6 варіант				
	1 рік	2 рік	3 рік	4 рік	5 рік	1 рік	2 рік	3 рік	4 рік	5 рік
січень	5	5,2	5,4	5,4	5,8	8,1	8,5	8,6	9	9,3
лютий	5,3	5,7	5,7	5,9	6,4	8,4	8,6	8,9	9,1	9,6
березень	5,5	5,8	5,9	6	6,5	8,5	8,8	9,1	9,3	9,8
квітень	5,8	5,9	6,3	6,2	6,6	8,2	8,3	8,8	9	9,1
травень	5,3	5,6	5,7	5,8	6,2	8,1	8,2	8,7	8,8	9
червень	5,2	5,3	5,4	5,5	6,1	8	7,9	8,2	8,3	8,5
липень	5,1	5,2	5,2	5,3	5,7	8,2	8,4	8,4	8,6	8,9
серпень	5,8	5,6	5,7	5,6	6	8,4	8,5	8,7	8,7	9,1
вересень	5,9	6	6,1	6,3	6,4	8,8	8,7	8,9	9	9,2
жовтень	6	6,1	6,2	6,4	6,5	9	8,9	9	9,2	9,4
листопад	6,2	6,3	6,5	6,7	6,9	9,1	9,2	9,3	9,5	9,6
грудень	6,4	6,5	6,8	6,9	7,3	9,5	9,4	9,6	9,8	10,4

Таблиця 3.5.8

	7 варіант					8 варіант				
	1 рік	2 рік	3 рік	4 рік	5 рік	1 рік	2 рік	3 рік	4 рік	5 рік
січень	1,1	1,3	1,7	2	2,2	1,1	1,2	1,4	1,5	1,7
лютий	1,2	1,4	1,9	2,1	2,4	1,4	1,5	1,7	1,6	1,8
березень	1,3	1,6	2	2,2	2,7	1,5	1,7	1,9	2	2,1
квітень	2	2,1	2,3	2,8	2,9	1,7	2	2	2,1	2,3
травень	2,4	2,3	2,5	2,9	3	1,9	2	2,1	2,3	2,5
червень	2	1,7	2	2,1	2,5	2	2,1	2,2	2,5	2,8
липень	1,8	1,5	1,3	1,8	2,2	2,2	2,3	2,5	2,7	2,9
серпень	1,5	1,3	1	1,2	2,1	2,4	2,6	2,8	3	3,1
вересень	2	1,5	1,4	1,6	2,5	1,9	2	2,1	2,3	2,4
жовтень	2,4	1,8	1,7	1,8	2,6	1,3	1,5	1,8	2	1,9
листопад	2,6	2,1	2,2	2,3	2,8	1,1	1,2	1,5	1,4	1,5
грудень	2,9	2,4	2,5	2,7	2,9	1,4	1,6	1,8	2	2,2

Таблиця 3.5.9

	9 варіант					10 варіант				
	1 рік	2 рік	3 рік	4 рік	5 рік	1 рік	2 рік	3 рік	4 рік	5 рік
січень	3	3,1	3,3	3,9	4	5,1	5,2	5,4	5,6	5,8
лютий	3,2	3,3	3,4	4	4,1	5,4	5,6	5,7	5,9	6
березень	3,4	3,5	3,7	4,1	4,2	5,5	5,8	5,9	6	6,2
квітень	3,8	3,7	3,9	4,2	4,3	5,8	5,9	6	6,2	6,4
травень	4	3,9	4	4,4	4,5	5,9	6	6,1	6,3	6,5
червень	4,1	4,2	4,3	4,5	4,6	6	6,1	6,2	6,4	6,8
липень	4,5	4,4	4,6	4,8	4,9	6,2	6,4	6,6	6,7	6,9
серпень	3,4	3,5	3,7	3,7	4,1	6,4	6,5	6,8	6,9	7,3
вересень	3,3	3,4	3,6	3,6	4	5,9	6	6,1	6,3	6,4
жовтень	3	3,2	3,1	3,2	3,4	5	5,5	5,6	6,1	6
листопад	2,8	3	3	3,1	3,3	4,8	5,3	5,4	5,6	5,8
грудень	3,5	3,6	3,8	3,9	4	5,4	5,6	5,8	6	6,5

Таблиця 3.5.10

	11 варіант					12 варіант				
	1 рік	2 рік	3 рік	4 рік	5 рік	1 рік	2 рік	3 рік	4 рік	5 рік
січень	2	2,5	2,6	2,8	3,1	3	3,5	3,7	4	4,2
лютий	2,1	2,7	2,7	2,9	3,4	3,2	3,6	3,8	4,1	4,3
березень	2,2	2,8	2,8	3	3,5	3,4	3,7	3,9	4,1	4,3
квітень	2,5	2,9	2,9	3,2	3,6	3,8	3,7	4	4,2	4,4
травень	2,4	2,7	2,7	2,8	3,3	4	3,9	4,1	4,3	4,5
червень	2,2	2,4	2,6	2,7	3,1	4,1	4,2	4,3	4,5	4,8
липень	2,1	2,3	2,5	2,5	2,8	4,3	4,4	4,6	4,8	5
серпень	2	2,3	2,3	2,4	2,4	3,4	3,5	3,9	3,7	4,1
вересень	2,9	3	3	3,1	3,2	3	3,3	3,8	3,3	4
жовтень	3	3,1	3,3	3,5	3,6	2,8	3,1	3,5	3	3,7
листопад	3,2	3,4	3,5	3,7	3,9	3,4	3,6	4	3,8	4,6
грудень	3,4	3,6	3,7	4	4,2	3,9	4	4,2	4,4	5

3.5.3 Лабораторна робота 3. Побудова прогнозу за допомогою адитивної та мультиплікативної моделей часового ряду

3.5.3.1 Методичні поради до вивчення теми

Виконання лабораторної роботи за даною темою надасть студентам можливість оволодіти практичними навичками моделювання часових рядів, які містять сезонні коливання, за допомогою адитивної та мультиплікативної моделей часового ряду, а також застосування методу ковзної середньої до вирівнювання часового ряду.

Для виконання лабораторної роботи передбачається наявність у студентів знань з основних питань тем «Часові ряди та їх попередній аналіз» та «Адаптивні методи прогнозування», а саме: вихідних понять, сутності, видів, основних компонентів та декомпозиції часових рядів; адитивної та мультиплікативної моделей часового ряду; сутності адаптивних методів екстраполяції, методу ковзної середньої та умовах його застосування. Крім того, необхідним є володіння практичними навичками застосування техніки регресійного аналізу, матричного числення та методів розв'язання систем алгебраїчних рівнянь.



Для самостійного вивчення теми рекомендується література: основна – [2, 4-6, 9], додаткова – [1-4].

3.5.3.2 Завдання на лабораторну роботу

1. Проаналізувати структуру заданого часового ряду (див. табл. 3.5.11) та вибрати найбільш прийнятну (адитивну або мультиплікативну) модель для опису часового ряду.

2. Побудувати обрану адитивну (або мультиплікативну) модель часового ряду, за якою скласти поквартальний прогноз на наступний рік, проаналізувати його й порівняти отримані результати.

3. Побудувати графіки за фактичними та прогнозними даними.

Варіанти завдання для виконання лабораторної роботи подано у табл. 3.5.11. Номер варіанта у даній таблиці відповідає номеру студента за списком у журналі.

Варіанти завдання для виконання лабораторної роботи №3

		1	2	3	4	5	6	7	8	9	10	11	12
		вар.	вар.	вар.	вар.	вар.	вар.	вар.	вар.	вар.	вар.	вар.	вар.
1 рік	1	100	72,6	32,5	181,0	19,3	18,6	307,2	411,7	94,7	47,6	17,9	80,2
	2	101	59,3	38,4	172,1	12,3	18,8	308,0	412,1	81,4	47,8	10,0	80,6
	3	102	64,8	30,9	178,3	13,2	19,0	308,4	412,2	90,6	49,0	12,2	81,2
	4	101	78,7	39,3	171,7	15,6	19,5	309,8	413,2	90,0	49,1	21,8	80,9
2 рік	5	100,4	74,0	32,9	183,7	21,5	18,7	308	411,9	95,1	48,0	18,0	80,3
	6	100,9	61,8	38,5	175,5	15,8	18,9	309	412,3	81,5	48,2	10,4	80,7
	7	102,0	67,4	34,3	180,4	17,2	19,2	309,0	412,4	90,8	49,3	12,6	81,6
	8	101,4	79,9	39,5	173,2	19,9	19,8	310,6	414	90,3	49,6	22,4	81,1
3 рік	9	100,7	78,3	33,4	184,8	26,3	19,0	308,3	412,0	95,8	48,2	18,3	80,6
	10	101,2	65,8	38,7	175,9	19,1	19,2	309,2	412,6	81,8	48,7	10,8	81,0
	11	102,8	69,5	36,1	181,8	20,3	19,6	309,8	412,9	91,0	49,5	12,9	82,2
	12	101,6	87,9	40,0	176,7	22,3	20,2	311,1	414	90,8	49,9	25,1	81,3
4 рік	13	101,1	92,2	33,6	186,2	29,7	19,3	309,2	412,4	96,2	48,8	19,0	80,9
	14	101,5	73,7	39,2	177	21,1	19,5	309,9	412,9	82,2	48,9	10,9	81,2
	15	103,3	77,9	36,7	183	23,7	19,6	310,3	413,0	91,4	50,1	13,0	82,6
	16	102	93,0	40,3	179	25,4	20,5	311,4	414	91,3	50,5	26,0	81,8
5 рік	17	102	94,3	34,0	189,0	31,8	19,5	310,0	413	98	48,9	19,5	81,4
	18	102,0	80,6	40,3	178	23,9	19,8	310,4	413	82,4	49,2	11,0	81,6
	19	104	84,1	37,1	186	25,8	20,0	311,0	414	94,8	50,6	13,7	83,1
	20	102,8	98,5	40,9	181	27,4	21,0	311,8	415	91,5	51,0	26,4	82,2

3.5.4 Лабораторна робота 4. Прогнозування за допомогою методу експоненціального згладжування

3.5.4.1 Методичні поради до вивчення теми

Виконання лабораторної роботи за даною темою надасть студентам можливість оволодіти практичними навичками застосування методу експоненціального згладжування до прогнозування часових рядів.

Для виконання лабораторної роботи передбачається наявність у студентів знань з основних питань тем «Часові ряди та їх попередній аналіз» та «Адаптивні методи прогнозування», а саме: вихідних понять, сутності, видів та основних компонентів часових рядів; сутності адаптивних методів екстраполяції, сутності, основних етапів методу експоненціального згладжування та умов його застосування. Крім того, необхідним є володіння практичними навичками застосування техніки регресійного аналізу.



Для самостійного вивчення теми рекомендується література: основна – [4, 5, 9-11].

3.5.4.2 Завдання на лабораторну роботу

1. За наведеними даними (табл. 3.5.12) скласти прогноз на наступні 5 років.
2. Побудувати графіки за фактичними та прогнозними даними.

Таблиця 3.5.12

Варіанти завдання для виконання лабораторної роботи №4 (№ варіанта у таблиці відповідає номеру студента за списком у журналі)

	1 вар.	2 вар.	3 вар.	4 вар.	5 вар.	6 вар.	7 вар.	8 вар.	9 вар.	10 вар.	11 вар.	12 вар.
t	y_t	y_t	y_t	y_t	y_t	y_t	y_t	y_t	y_t	y_t	y_t	y_t
1	24,0	3,6	97,9	14,3	140,0	57,8	82,9	39,8	54,3	18,4	16,3	7,4
2	25,9	4,2	116,2	16,9	166,1	68,6	98,4	47,2	56,2	21,8	19,4	8,8
3	27,7	4,9	133,4	19,5	190,8	78,8	113,0	54,2	58,0	25,0	22,2	10,1
4	29,8	5,6	153,6	22,4	219,7	90,7	130,1	62,4	60,1	28,8	25,6	11,6
5	32,8	6,7	182,4	26,6	260,9	107,7	154,5	74,1	63,1	34,2	30,4	13,8
6	36,3	7,9	216,0	31,5	308,9	127,6	182,9	87,8	66,6	40,5	36,0	16,3
7	38,7	8,7	239,0	34,9	341,9	141,2	202,4	97,1	69,0	44,8	39,8	18,0
8	42,7	10,1	277,4	40,5	396,8	163,9	235,0	112,7	73,0	52,0	46,2	20,9
9	47,1	11,7	319,7	46,6	457,2	188,8	270,7	129,9	77,4	59,9	53,3	24,1
10	52,6	13,6	372,5	54,3	532,7	220,0	315,4	151,3	82,9	69,8	62,1	28,1

3.5.5 Лабораторна робота 5. Побудова прогнозу за допомогою методу гармонійних ваг

3.5.5.1 Методичні поради до вивчення теми

Виконання лабораторної роботи за даною темою надасть студентам можливість оволодіти практичними навичками застосування методу гармонійних ваг до побудови прогнозу за заданим часовим рядом.

Для виконання лабораторної роботи передбачається наявність у студентів знань з основних питань тем «Часові ряди та їх попередній аналіз» та «Адаптивні методи прогнозування», а саме: вихідних понять, сутності, видів, основних компонентів часових рядів, методах виявлення та усунення аномальних значень рівнів часових рядів; сутності складної екстраполяції та адаптивних методів прогнозування; сутності, основних етапів, умов застосування та основних передумов методу гармонійних ваг.

Крім того, необхідним є володіння практичними навичками застосування техніки регресійного аналізу та методів розв'язання систем алгебраїчних рівнянь.



Для самостійного вивчення теми рекомендується література:
основна – [4, 5, 9, 11], додаткова – [1, 2, 4].

3.5.5.2 Завдання на лабораторну роботу

1. За заданим часовим рядом (див. табл. 3.5.13) перевірте виконання усіх передумов методу гармонійних ваг.

Якщо у часовому ряді наявні аномальні спостереження, потрібно їх усунути та застосовувати метод гармонійних ваг до «нового» часового ряду (з усіма нормальними рівнями ряду).

2. За даними часового ряду скласти прогноз на наступні 5 років.

3. Побудувати графіки за фактичними та прогнозними даними.

Таблиця 3.5.13

Варіанти завдання для виконання лабораторної роботи №5 (№ варіанта у таблиці відповідає номеру студента за списком у журналі)

	1 вар.	2 вар.	3 вар.	4 вар.	5 вар.	6 вар.	7 вар.	8 вар.	9 вар.	10 вар.	11 вар.	12 вар.
t	y_t	y_t	y_t	y_t	y_t	y_t	y_t	y_t	y_t	y_t	y_t	y_t
1	37,6	38,8	2,68	24,8	9,7	2,47	54,3	18,8	31,58	30,2	10,8	7,39
2	40,9	45,5	2,67	26,7	15,9	2,46	56,2	20,45	34,36	32,5	11,6	8,77
3	42,9	50,9	2,78	28,5	17,8	2,56	58,0	21,45	36,04	34,7	12,4	10,0
4	46,3	57,3	2,85	30,6	20,1	2,62	60,1	23,15	38,89	37,2	13,3	11,1
5	48,4	72,0	3,72	31,7	21,4	3,42	63,1	24,2	40,66	38,6	13,8	12,1
6	50,0	64,9	2,94	35,2	32,6	2,70	70,9	25,0	42,00	42,8	15,3	16,3
7	71,1	70,4	4,57	59,2	35,1	4,20	69,0	35,55	59,72	72,1	25,8	18,04
8	54,2	79,5	3,86	42,6	22,9	3,55	73,0	27,1	45,53	51,8	18,5	16,2
9	59,3	84,8	5,91	49,2	26,3	5,44	44,1	29,65	49,81	59,9	21,4	21,9

3.5.6 Лабораторна робота 6. Аналіз якості побудованих моделей та прогнозів

3.5.6.1 Методичні поради до вивчення теми

Виконання лабораторної роботи за даною темою надасть студентам можливість оволодіти практичними навичками проведення аналізу якості побудованих моделей та отриманих за ними прогнозів.

Для виконання лабораторної роботи передбачається наявність у студентів знань з основних питань тем «Часові ряди та їх попередній аналіз» та «Методи аналізу якості побудованих моделей та прогнозів», а саме: вихідних понять, сутності, видів, основних компонентів часових рядів; основних методів аналізу якості побудованих моделей, за допомогою яких перевіряється адекватність моделей; застосування абсолютних, порівняльних та якісних показників точності прогнозів. Крім того, необхідним є володіння практичними навичками застосування техніки регресійного аналізу.



Для самостійного вивчення теми рекомендується література:
основна – [4, 5, 9, 11].

3.5.6.2 Завдання на лабораторну роботу

1. Для заданого часового ряду (табл. 3.5.14) потрібно:
 - а) побудувати лінійну модель $\hat{y}_t = a_0 + a_1t$, параметри якої оцінити МНК;
 - б) оцінити адекватність побудованої моделі на основі дослідження:
 - випадковості залишкової компоненти за критерієм піків;
 - незалежності рівнів ряду залишків за d –критерієм (в якості критичних значень слід використовувати рівні $d_1 = 1,08$ й $d_2 = 1,36$) та за першим коефіцієнтом автокореляції, критичний рівень якого $r(1) = 0,36$;
 - нормальності розподілу залишкової компоненти за RS -критерієм з критичними рівнями 2,67 – 3,69;
 - рівність математичного сподівання рівнів ряду залишків нулю за t -критерієм Стьюдента;
 - в) побудувати точковий та інтервальний прогнози на два кроки вперед (для ймовірності $P = 70\%$ прийміть коефіцієнт, що дорівнює 1,11);
 - г) відобразити на графіці фактичні дані, результати розрахунків та прогнозування.

Критичні значення d –критерію, першого коефіцієнта автокореляції $r(1)$ та RS -критерію наведені в табл. 2.4.1.
2. За допомогою усіх абсолютних, порівняльних та якісних показників точності прогнозів оцінити адекватність та якість отриманих прогнозів. Побудувати діаграму «прогноз-реалізація».

Таблиця 3.5.14

Варіанти завдання для виконання лабораторної роботи №6
(№ варіанта у таблиці відповідає номеру студента за списком у журналі)

	1	2	3	4	5	6	7	8	9	10	11	12
	вар.	вар.	вар.	вар.	вар.	вар.	вар.	вар.	вар.	вар.	вар.	вар.
t	y_t	y_t	y_t	y_t	y_t	y_t	y_t	y_t	y_t	y_t	y_t	y_t
1	37,6	38,8	2,6	24,8	13,7	14,2	54,3	18,8	31,58	30,2	10,8	7,40
2	40,9	45,5	3,7	26,7	15,9	15,7	56,2	20,45	34,36	32,5	11,6	8,70
3	42,9	50,9	5,8	28,5	17,8	16,8	58,0	21,45	36,04	34,7	12,4	10,0
4	46,3	57,3	4,8	30,6	20,1	17,3	60,1	23,15	38,89	37,2	13,3	11,5
5	48,4	72,0	5,7	31,7	21,4	15,2	63,1	24,2	40,66	38,6	13,8	12,6
6	50,0	64,9	5,9	35,2	22,6	18,9	70,9	25,0	42,00	42,8	15,3	16,3
7	71,1	70,4	6,5	43,8	25,1	18,4	69,0	35,55	59,72	64,4	25,8	20,4
8	54,2	79,5	6,8	42,6	22,9	19,9	73,0	27,1	45,53	51,8	18,5	23,2
9	59,3	84,8	7,9	49,2	26,3	22,1	80,1	29,65	49,81	59,9	21,4	28,9
10	62,4	90,1	8,2	52,4	29,1	25,3	84,6	31,2	52,42	63,1	25,2	30,8

СЛОВНИК КОРИСНИХ ТЕРМІНІВ

Адаптація – здатність системи цілеспрямовано набувати пристосування у складних середовищах, а також сам процес такого пристосовування, при якому можуть змінюватись кількісні характеристики системи, а також її структура.

Адаптивні методи прогнозування – методи, які базуються на обчисленні послідовних в часі значень прогнозованого показника з урахуванням міри впливу попередніх рівнів.

Адитивна модель ряду динаміки – модель, у якій характер циклічних і сезонних коливань залишається постійним, а спостережувані значення в середньому збільшуються або зменшуються за період на однакову величину.

Аномальний рівень – окреме значення рівня часового ряду, яке не відповідає потенційним можливостям досліджуваної системи і яке, залишаючись в якості рівня ряду, чинить істотний вплив на значення основних характеристик часового ряду, у тому числі і на відповідну трендову модель.

Верифікація математичної моделі – перевірка моделі, зіставлення її з дійсністю, з'ясування відповідності реальним даним і змістовному уявленню про об'єкт й мету моделювання.

Верифікація (достовірність) прогнозу – це оцінка ймовірності та точності здійснюваного прогнозу для заданого довірчого інтервалу.

Виробнича система (виробництво) – економічна система, в якій організований і протікає виробничий процес.

Виробничий процес – це процес створення доданої вартості шляхом цілеспрямованого перетворення одного набору товарів в інший.

Динамічний ряд – послідовність спостережень одного показника, впорядкованих залежно від послідовно зростаючих або спадних значень іншого показника. Складовими елементами рядів є цифрові значення показника (так звані **рівні** цих рядів), і моменти або інтервали часу, до яких вони відносяться.

Динамічні моделі описують зміни модельованої системи в часі, тому час в цих моделях представлений явно – як неперервна або як дискретна величина, скінченим набором чи нескінченною послідовністю дискретних значень.

Динамічна система – еволюційна система, в широкому розумінні довільна (зокрема, фізична, економічна) система, що змінюється в часі.

Дискретна модель – математична модель, невідомі якої набувають дискретних значень (цілочислові, булеві тощо). Така модель використовується для спрощення процесу моделювання, коли події у досліджуваній (зокрема, виробничій) системі відбуваються лише в окремі дискретні моменти часу.

Екстраполяція – метод, за яким прогнозовані показники розраховуються як продовження динамічного ряду на майбутнє за виявленою закономірністю розвитку, тобто це визначення значень функції $y(t)$ в точках t , які лежать поза відрізком $[t_0, t_n]$, за її значеннями в точках $t_0 < t_1 < \dots < t_n$.

Коригування прогнозу – це процес уточнення прогнозу на підставі результатів його перевірки. Коригування складається з оцінки достовірності й точності прогнозу на підставі додаткової інформації, раніше не врахованої.

Крива зросту – плавна крива (гладка функція), що апроксимує часовий ряд.

Математична модель – абстракція реальної дійсності (світу), в якій відношення між реальними елементами, що цікавлять дослідника, замінені відношеннями між математичними категоріями. Ці відношення зазвичай подаються у формі рівнянь і/чи нерівностей, відношеннями формальної логіки між показниками (змінними), які характеризують функціонування реальної системи, що моделюється.

Метод найменших квадратів (МНК) – метод регресійного аналізу, який застосовується для оцінювання параметрів кривої зросту та полягає у визначенні параметрів моделі, які мінімізують суми квадратів відхилень спостережуваних (фактичних) значень залежної змінної від прогнозних значень, отриманих за моделлю.

Моделювання – процес побудови, вивчення та застосування моделей, тобто спосіб вивчення об'єкта (процесу, явища) за допомогою деякого його спрощеного представлення – моделі. Змістом методу моделювання є побудова моделі на підставі попереднього вивчення об'єкта і виокремлення його суттєвих характеристик, її експериментальний і/чи теоретичний аналіз, зіставлення результатів з даними про об'єкт, корегування моделі тощо. Побудова моделі передбачає наявність суб'єкта (дослідника), об'єкта моделювання й мети, задля якої воно здійснюється.

Модель – це уявний або реальний об'єкт, який у процесі свого вивчення замінює об'єкт-оригінал. Вона конструюється суб'єктом дослідження чи управління так, щоб відобразити найсуттєвіші для дослідження характеристики об'єкта (властивості, взаємозв'язки, структурні й функціональні параметри тощо). Використання моделі ґрунтується на її здатності замінювати, заміщувати об'єкт стосовно до мети дослідження та на її властивості бути при цьому доступнішою для вивчення у відповідному аспекті, ніж модельований об'єкт.

Мультиплікативна модель ряду динаміки характеризується тим, що характер циклічних і сезонних коливань залишається постійним тільки по відношенню до тренду, при цьому спостережувані значення збільшуються або зменшуються за період на певний відсоток.

Неперервна модель – математична модель, кожне невідоме якої змінюється неперервно в області свого визначення; крім того, зазвичай, також припускається, що в описі моделі використані тільки неперервні залежності її елементів від первинних даних і невідомих (елементи опису моделі — праві й ліві частини рівнянь тощо).

Період (час) упередження прогнозу – це проміжок часу, на який складається прогноз. Його тривалість залежить від часу функціонування об'єкту, інтенсивності зростання його показників, тривалості дії виявлених тенденцій і закономірностей.

Прогноз – це науково-обґрунтований опис можливих станів об'єкту в майбутньому, а також альтернативних шляхів, термінів та способів досягнення цього стану.

Прогнозна ретроспекція – це етап прогнозування, на якому досліджується історія розвитку об'єкту прогнозування та прогнозного фону з метою отримання їх систематизованого опису.

Прогнозний діагноз – це етап прогнозування, на якому досліджується систематизований опис об'єкту та прогнозного фону з метою виявлення тенденцій їх розвитку, вибору та розробки моделей й методів подальшого прогнозування.

Прогнозний фон – це сукупність зовнішніх умов, істотних для вирішення завдання прогнозу.

Прогнозування – це вид пізнавальної діяльності людини, яка спрямована на формування знань наперед про об'єкт та його розвиток на основі аналізу минулих тенденцій його розвитку та економічних закономірностей. Прогнозування – це процес розробки прогнозів. **Об'єкт прогнозування** – можливий розвиток розглядуваного об'єкту в майбутньому з урахуванням розвитку зовнішнього середовища. **Предмет прогнозування** – це кількісні і якісні зміни, які відбуваються в об'єкті під впливом сукупності або окремих чинників в межах прогнозованого періоду.

Проспекція – це етап прогнозування, де за результатами діагнозу розробляється прогноз та виробляється його коригування.

Система – комплекс взаємопов'язаних елементів разом із відношеннями (зв'язками) між цими елементами та між їхніми атрибутами, які спільно реалізують певні цілі.

Складність об'єкту прогнозування – це характеристика об'єкту, яка визначається різноманітністю його елементів, властивостей та відношень.

Соціально-економічна система – складна ймовірнісна динамічна система, що охоплює процеси виробництва, обміну, розподілу й споживання матеріальних та інших благ.

Стаціонарний часовий ряд (у вузькому сенсі) – ряд $y(t)$, якщо середнє, дисперсія і коваріації $y(t)$ не залежать від t , тобто якщо від t не залежить закон розподілу вірогідності випадкової величини $y(t)$.

Товаром узагальнено називаються усі види економічного продукту будь-якої виробничої системи. Споживані у виробничому процесі товари називаються **чинниками виробництва** (або **ресурсами**). Отримані в результаті виробничого процесу товари називаються **продуктами виробництва**.

Тренд – це довготривала тенденція зміни досліджуваного часового ряду, стійкий рух ряду у певному напрямі. Тренди можуть бути описані різними рівняннями – лінійними, логарифмічними, степеневими і т.д. Фактичний тип тренду встановлюють на основі підбору його функціональної моделі статистичними методами або згладжуванням початкового часового ряду.

Трендові моделі – моделі, які визначають загальний напрям розвитку, головну тенденцію (тренд) часових рядів (рядів динаміки).

Часовий ряд – сукупність спостережень $y(t_1), y(t_2), \dots, y(t_n)$ аналізованої величини $Y(t)$, вироблених в послідовні моменти часу t_1, t_2, \dots, t_n , тобто впорядкована в часі послідовність спостережень.

РЕКОМЕНДОВАНА ЛІТЕРАТУРА**Основна:**

1. Экономико-математическое моделирование: Учебник для студентов вузов / Под общ. ред. И.Н. Дрогобыцкого. – М.: Экзамен, 2004. – 800 с.
2. Глівенко С.В., Соклов М.О., Теліженко О.М. Економічне прогнозування: Навч. посібник. – 2-ге вид. – Суми: Університетська книга, 2001. – 207 с.
3. Горелова В.Л., Мельникова Е.Н. Основы прогнозирования систем. – М.: Высшая школа, 1986. – 286 с.
4. Дуброва Т.А. Статистические методы прогнозирования: Учебн. пособие для вузов. – М.: ЮНИТИ-ДАНА, 2003. – 206 с.
5. Єріна А.М. Статистичне моделювання та прогнозування: Навч. посібник. – К.: КНЕУ, 2001. – 170 с.
6. Макроэкономические модели планирования и прогнозирования. – М.: Статистика, 1970.- 263 с.
7. Основы экономического и социального прогнозирования / Под ред. В.П. Мосина. – М.: Высшая школа, 1985. – 199 с.
8. Прогнозирование и планирование экономики: Учебн. пособие / В.И. Борисович, Г.А. Кандаурова. – Мн.: Интерпресс сервис; экоперспектива, 2001. – 380 с.
9. Рабочая книга по прогнозированию. – М.: Мысль, 1982. – 429 с.
10. Равікович Є.І. Макроекономічне прогнозування: Навч.-метод. посібник для самост. вивч. дисц. – К.: КНЕУ, 2003. – 139 с.
11. Теория прогнозирования и принятия решений. Учебн. пособие / Под ред. С.А. Саркисяна. – М.: Высшая школа, 1977. – 352 с.

Додаткова:

1. Басовский Л.Е. Прогнозирование и планирование в условиях рынка. Учебн. пособие. – М.: Инфра-М, 2003. – 260 с.
2. Ольхин В.Я. Элементы линейной алгебры: Учебн. пособие / Международный педагогический университет. – Магадан: Изд-во МПУ, 1997. – 43 с.
3. Орлова И.В. Экономико-математическое моделирование: Практическое пособие к решению задач. – М.: Вузовский учебник, 2007. – 144 с.
4. Стохастическое моделирование и прогнозирование / Под ред. А.Г. Гранберга. – М.: Финансы и статистика, 1990. – 385 с.
5. Гришин А.Ф., Кочерова Е.В. Статистические модели: построение, оценка, анализ: Учебн. пособие. – М.: Финансы и статистика, 2005. – 416 с.

Додаток А
Статистико-математичні таблиці

Таблиця А.1

Таблиця значень F -критерію Фішера при рівні значущості $\alpha = 0,05$

$k_1 \backslash k_2$	1	2	3	4	5	6	8	12	24	∞
1	161,45	199,5	215,72	224,57	230,17	233,97	238,89	243,91	249,04	254,32
2	18,51	19,00	19,16	19,25	19,30	19,33	19,37	19,41	19,45	19,50
3	10,13	9,55	9,28	9,12	9,01	8,94	8,84	8,74	8,64	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,04	5,91	5,77	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,82	4,68	4,53	4,36
6	5,99	5,14	4,76	4,53	4,39	4,28	4,15	4,00	3,84	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,73	3,57	3,41	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,44	3,28	3,12	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,23	3,07	2,90	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,07	2,91	2,74	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	2,95	2,79	2,61	2,40
12	4,75	3,88	3,49	3,26	3,11	3,00	2,85	2,69	2,50	2,30
13	4,67	3,80	3,41	3,18	3,02	2,92	2,77	2,60	2,42	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,70	2,53	2,35	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,64	2,48	2,29	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,59	2,42	2,24	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,55	2,38	2,19	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,51	2,34	2,15	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,48	2,31	2,11	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,45	2,28	2,08	1,84
21	4,32	3,47	3,07	2,84	2,68	2,57	2,42	2,25	2,05	1,81
22	4,30	3,44	3,05	2,82	2,66	2,55	2,40	2,23	2,03	1,78
23	4,28	3,42	3,03	2,80	2,64	2,53	2,38	2,20	2,00	1,76
24	4,26	3,40	3,01	2,78	2,62	2,51	2,36	2,18	1,98	1,73
25	4,24	3,38	2,99	2,76	2,60	2,49	2,34	2,16	1,96	1,71
26	4,22	3,37	2,98	2,74	2,59	2,47	2,32	2,15	1,95	1,69
27	4,21	3,35	2,96	2,73	2,57	2,46	2,30	2,13	1,93	1,67
28	4,20	3,34	2,95	2,71	2,56	2,44	2,29	2,12	1,91	1,65
29	4,18	3,33	2,93	2,70	2,54	2,43	2,28	2,10	1,90	1,64
30	4,17	3,32	2,92	2,69	2,53	2,42	2,27	2,09	1,89	1,62
35	4,12	3,26	2,87	2,64	2,48	2,37	2,22	2,04	1,83	1,57
40	4,08	3,23	2,84	2,61	2,45	2,34	2,18	2,00	1,79	1,51
45	4,06	3,21	2,81	2,58	2,42	2,31	2,15	1,97	1,76	1,48
50	4,03	3,18	2,79	2,56	2,40	2,29	2,13	1,95	1,74	1,44
60	4,00	3,15	2,76	2,52	2,37	2,25	2,10	1,92	1,70	1,39
70	3,98	3,13	2,74	2,50	2,35	2,23	2,07	1,89	1,67	1,35
80	3,96	3,11	2,72	2,49	2,33	2,21	2,06	1,88	1,65	1,31
90	3,91	3,10	2,71	2,47	2,32	2,20	2,04	1,86	1,64	1,28
100	3,94	3,09	2,70	2,46	2,30	2,19	2,03	1,85	1,63	1,26

Продовження табл. А.1

125	3,92	3,07	2,68	2,44	2,29	2,17	2,01	1,83	1,60	1,21
150	3,90	3,06	2,66	2,43	2,27	2,16	2,00	1,82	1,59	1,18
200	3,89	3,04	2,65	2,42	2,26	2,14	1,98	1,80	1,57	1,14
300	3,87	3,03	2,64	2,41	2,25	2,13	1,97	1,79	1,55	1,10
400	3,86	3,02	2,63	2,40	2,24	2,12	1,96	1,78	1,54	1,07
500	3,86	3,01	2,62	2,39	2,23	2,11	1,96	1,77	1,54	1,06
1000	3,85	3,00	2,61	2,38	2,22	2,10	1,95	1,76	1,53	1,03
∞	3,84	2,99	2,60	2,37	2,21	2,09	1,94	1,75	1,52	1,00

Таблиця А.2

Значення середньої μ й стандартних помилок σ_1 і σ_2 (для $10 \leq n \leq 50$)

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

Таблиця А.3

Критичні значення t -критерію Стюдента при рівнях значущості $\alpha = 0,1$,
 $\alpha = 0,05$ та $\alpha = 0,01$ (двосторонній)

Число ступенів свободи	Рівень значущості α			Число ступенів свободи	Рівень значущості α		
	0,10	0,05	0,01		0,10	0,05	0,01
1	6,3138	12,706	63,657	31	1,6955	2,0395	2,7440
2	2,9200	4,3027	9,9248	32	1,6939	2,0369	2,7385
3	2,3534	3,1825	5,8409	33	1,6924	2,0345	2,7333
4	2,1318	2,7764	4,6041	34	1,6909	2,0322	2,7284
5	2,0150	2,5706	4,0321	35	1,6896	2,0301	2,7238
6	1,9432	2,4469	3,7074	36	1,6883	2,0281	2,7195
7	1,8946	2,3646	3,4995	37	1,6871	2,0262	2,7154
8	1,8595	2,3060	3,3554	38	1,6860	2,0244	2,7116
9	1,8331	2,2622	3,2498	39	1,6849	2,0227	2,7079
10	1,8125	2,2281	3,1693	40	1,6839	2,0211	2,7045
11	1,7959	2,2010	3,1058	45	1,6794	2,0141	2,6896
12	1,7823	2,1788	3,0545	50	1,6759	2,0086	2,6778
13	1,7709	2,1604	3,0123	55	1,6730	2,0040	2,6682
14	1,7613	2,1448	2,9768	60	1,6707	2,0003	2,6603
15	1,7530	2,1315	2,9467	65	1,6686	1,9971	2,6536
16	1,7459	2,1199	2,9208	70	1,6669	1,9944	2,6479
17	1,7396	2,1098	2,8982	75	1,6654	1,9921	2,6430
18	1,7341	2,1009	2,8784	80	1,6641	1,9901	2,6387

Продовження табл. А.3

19	1,7291	2,0930	2,8609	85	1,6630	1,9883	2,6349
20	1,7247	2,0860	2,8453	90	1,6620	1,9867	2,6316
21	1,7207	2,0796	2,8314	95	1,6611	1,9853	2,6286
22	1,7171	2,0739	2,8188	100	1,6602	1,9840	2,6259
23	1,7139	2,0687	2,8073	110	1,6588	1,9818	2,6213
24	1,7109	2,0639	2,7969	120	1,6577	1,9799	2,6174
25	1,7081	2,0595	2,7874	130	1,6567	1,9784	2,6142
26	1,7056	2,0555	2,7787	140	1,6558	1,9771	2,6114
27	1,7033	2,0518	2,7707	150	1,6551	1,9759	2,6090
28	1,7011	2,0484	2,7633	160	1,6544	1,9749	2,6069
29	1,6991	2,0452	2,7564	200	1,6525	1,9719	2,6006
30	1,6973	2,0423	2,7500	∞	1,6449	1,9600	2,5758

Таблиця А.4

Критерій Аббе, $q_n(p)$

$n \backslash p$	0,001	0,01	0,05	$n \backslash p$	0,001	0,01	0,05
4	0,2949	0,3128	0,3902	32	0,4963	0,6089	0,7177
5	2080	2690	4102	33	5027	6141	7216
6	1817	2808	4451	34	5090	6193	7256
7	1848	3070	4680	35	5150	6242	7292
8	2018	3314	4912	36	5208	6290	7328
9	0,2210	0,3544	0,5121	37	0,5265	0,6337	0,7363
10	2408	3759	5311	38	5319	6381	7396
11	2598	3957	5482	39	5373	6425	7429
12	2778	4140	5638	40	5425	6467	7461
13	2949	4309	5778	41	5475	6508	7491
14	0,3112	0,4466	0,5908	42	0,5524	0,6548	0,7521
15	3266	4611	6027	43	5571	6587	7550
16	3413	4746	6137	44	5616	6622	7576
17	3552	4872	6237	45	5660	6659	7603
18	3684	4989	6330	46	5701	6693	7628
19	0,3809	0,5105	0,6417	47	0,5743	0,6727	0,7653
20	3926	5203	6498	48	5781	6757	7676
21	4037	5301	6574	49	5817	6787	7698
22	4142	5393	6645	50	5853	6814	7718
23	4241	5479	6713	51	5887	6842	7739
24	0,4334	0,5562	0,6776	52	0,5922	0,6869	0,7759
25	4423	5639	6836	53	5955	6896	7779
26	4509	5713	6893	54	5989	6924	7799
27	4591	5784	6946	55	6020	6949	7817
28	4670	5850	6996	56	6051	6974	7836
29	0,4748	0,5915	0,7046	57	0,6083	0,6999	0,7853
30	4822	5975	7091	58	6114	7024	7872
31	4895	6034	7136	59	6145	7049	7891
				60	6174	7071	7906

Додаток Б

Таблиця логарифмів

Таблиця Б.1

Таблиця для обчислення логарифмів

a	$\ln a$	a	$\ln a$	a	$\ln a$	a	$\ln a$
1	0	26	3,258097	51	3,931826	76	4,330733
2	0,693147	27	3,295837	52	3,951244	77	4,343805
3	1,098612	28	3,332205	53	3,970292	78	4,356709
4	1,386294	29	3,367296	54	3,988984	79	4,369448
5	1,609438	30	3,401197	55	4,007333	80	4,382027
6	1,791759	31	3,433987	56	4,025352	81	4,394449
7	1,945910	32	3,465736	57	4,043051	82	4,406719
8	2,079442	33	3,496508	58	4,060443	83	4,418841
9	2,197225	34	3,526361	59	4,077537	84	4,430817
10	2,302585	35	3,555348	60	4,094345	85	4,442651
11	2,397895	36	3,583519	61	4,110874	86	4,454347
12	2,484907	37	3,610918	62	4,127134	87	4,465908
13	2,564949	38	3,637586	63	4,143135	88	4,477337
14	2,639057	39	3,663562	64	4,158883	89	4,488636
15	2,708050	40	3,688879	65	4,174387	90	4,499810
16	2,772589	41	3,713572	66	4,189655	91	4,510860
17	2,833213	42	3,737670	67	4,204693	92	4,521789
18	2,890372	43	3,761200	68	4,219508	93	4,532599
19	2,944439	44	3,784190	69	4,234107	94	4,543295
20	2,995732	45	3,806662	70	4,248495	95	4,553877
21	3,044522	46	3,828641	71	4,262680	96	4,564348
22	3,091042	47	3,850148	72	4,276666	97	4,574711
23	3,135494	48	3,871201	73	4,290459	98	4,584967
24	3,178054	49	3,891820	74	4,304065	99	4,595120
25	3,218876	50	3,912023	75	4,317488	100	4,605170

Додаток В

Метод найменших квадратів (МНК)

На практиці часто виникають такі задачі, коли маючи обмежену кількість експериментальних даних, треба спрогнозувати поведінку досліджуваного об'єкту в майбутньому. Для цієї мети широко використовують рівняння різного вигляду, які з той чи іншою похибкою моделюють поведінку об'єкта. Підбір таких рівнянь називають **апроксимацією експериментальних даних**. Зокрема, апроксимація усередині області одержання експериментальних даних називається **інтерполяцією**, а за межами цієї області – **екстраполяцією**.

Нехай на основі експерименту потрібно встановити функціональну залежність $\hat{y}_t = f(t_1, t_2, \dots, t_n)$, де \hat{y}_t – розрахункові значення вихідного ряду; t – час; n – число спостережень. Вигляд функції $f(t_1, t_2, \dots, t_n)$ встановлюється або із теоретичних міркувань, або на основі аналізу графіка функції. Для цього слід побудувати в прямокутній декартовій системі координат точки, відповідні експериментальним значенням. За виглядом побудованої графічної залежності можна вибрати функцію, яка найкраще апроксимує експериментальні дані. Якщо, наприклад, з графіка видно, що дані найкраще описуються лінійною залежністю, то слід апроксимувати їх за допомогою лінійної функції $y = at + b$.

Взагалі, існує необмежена кількість кривих $\hat{y}_t = f(t_1, t_2, \dots, t_n)$, які можна провести через множину спостережуваних точок. Щоб визначити, яку саме криву з них потрібно вибрати, потрібно мати в розпорядженні певний критерій, який би дозволяв вибрати з множини можливих кривих «найкращу» з точки зору даного критерію. Найбільш поширеним критерієм є критерій мінімізації суми квадратів відхилень, згідно з яким параметри шуканого рівняння підбираються таким чином, щоб графік функції розташовувався на мінімальному віддаленні від точок вихідних даних.

Для оцінювання параметрів обраного рівняння застосовується метод найменших квадратів (МНК), сутність якого полягає у визначенні параметрів моделі, які мінімізують суми квадратів відхилень фактичних значень залежної змінної y_t від прогнозних значень \hat{y}_t , отриманих за моделлю, тобто

$$S = \sum_{i=1}^n (\hat{y}_t - y_t)^2 \rightarrow \min,$$

де \hat{y}_t – розрахункові значення вихідного ряду; y_t – фактичні значення вихідного ряду; n – число спостережень.

Оскільки t_1, t_2, \dots, t_n та y_t є відомими, то сума квадратів їх відхилень знаходиться в залежності від невідомих параметрів $a_0, a_1, a_2, \dots, a_n$. Отже, S можна представити у вигляді $S = f(a_0, a_1, a_2, \dots, a_n)$.

Відомо, що в точці екстремуму диференційованої функції її перша похідна дорівнює нулю. Оскільки функція $S = f(a_0, a_1, a_2, \dots, a_n)$ є диференційованою, то можна взяти частинні похідні суми S за окремими параметрами та прирівняти їх нулю:

$$\sum t = \frac{n(n+1)}{2}; \quad \sum t^2 = \frac{n(n+1)(2n+1)}{6};$$

$$\sum t^3 = \frac{n^2(n+1)^2}{4}; \quad \sum t^4 = \frac{n(n+1)(2n+1)(3n^2+3n-1)}{30}.$$

Найбільш поширені види апроксимуючих функцій та відповідних систем нормальних рівнянь, побудованих за МНК, подані у табл. В.1.

Таблиця В.1

Системи лінійних рівнянь для оцінки параметрів поліномів за МНК

Вид залежності, рівняння апроксимуючої кривої	Система нормальних рівнянь
Пряма лінія $\hat{y}_t = a_0 + a_1 \cdot t$	$\begin{cases} n \cdot a_0 + a_1 \sum t = \sum y_t \\ a_0 \sum t + a_1 \sum t^2 = \sum y_t t \end{cases}$
Гіпербола $\hat{y}_t = a_0 + \frac{a_1}{t}$	$\begin{cases} n \cdot a_0 + a_1 \sum \frac{1}{t} = \sum y_t \\ a_0 \sum \frac{1}{t} + a_1 \sum \frac{1}{t^2} = \sum \frac{y_t}{t} \end{cases}$
Парабола 2-го порядку $\hat{y}_t = a_0 + a_1 t + a_2 t^2$	$\begin{cases} n \cdot a_0 + a_1 \sum t + a_2 \sum t^2 = \sum y_t \\ a_0 \sum t + a_1 \sum t^2 + a_2 \sum t^3 = \sum y_t t \\ a_0 \sum t^2 + a_1 \sum t^3 + a_2 \sum t^4 = \sum y_t t^2 \end{cases}$
Парабола 3-го порядку $\hat{y}_t = a_0 + a_1 t + a_2 t^2 + a_3 t^3$	$\begin{cases} n \cdot a_0 + a_1 \sum t + a_2 \sum t^2 + a_3 \sum t^3 = \sum y_t \\ a_0 \sum t + a_1 \sum t^2 + a_2 \sum t^3 + a_3 \sum t^4 = \sum y_t t \\ a_0 \sum t^2 + a_1 \sum t^3 + a_2 \sum t^4 + a_3 \sum t^5 = \sum y_t t^2 \\ a_0 \sum t^3 + a_1 \sum t^4 + a_2 \sum t^5 + a_3 \sum t^6 = \sum y_t t^3 \end{cases}$
Показникова крива $\hat{y}_t = a_0 \cdot a_1^t$	$\begin{cases} n \cdot \lg a_0 + \lg a_1 \sum t = \sum \lg y_t \\ \lg a_0 \sum t + \lg a_1 \sum t^2 = \sum (\lg y_t \cdot t) \end{cases}$
Степенева функція $\hat{y}_t = a_0 \cdot t^{a_1}$	$\begin{cases} n \cdot \lg a_0 + a_1 \sum \lg t = \sum \lg y_t \\ \lg a_0 \sum t + a_1 \sum (\lg t)^2 = \sum (\lg y_t \cdot \lg t) \end{cases}$

Використання процедури оцінки, заснованої на МНК, припускає обов'язкове задоволення наступних передумов:

1. Випадкові помилки мають нульову середню, кінцеві дисперсії й коваріації.

2. Кожне вимірювання випадкової помилки характеризується нульовими середніми, які не залежать від значень спостережуваних змінних.

3. Дисперсії кожної випадкової помилки однакові, їх величини незалежні від значень спостережуваних змінних (гомоскедастичність).

4. Відсутність автокореляції помилок, тобто значення помилок різних спостережень незалежні один від одного.

5. Нормальність. Випадкові помилки мають нормальний розподіл.

6. Значення ендогенної змінної x є вільними від помилок вимірювання та мають кінцеві середні значення й дисперсії.

МНК широко застосовується для одержання конкретних прогнозів, що пояснюється його простотою й легкістю реалізації на ЕОМ. Недолік методу полягає в тому, що модель тренду жорстко фіксується й за допомогою МНК можна одержати надійний прогноз на невеликий період упередження. Тому МНК відноситься, головним чином, до методів короткострокового прогнозування. Крім того, істотними труднощами МНК є правильний вибір виду моделі.

Вибір моделі в кожному конкретному випадку здійснюється виходячи з цілого ряду статистичних критеріїв (наприклад, дисперсії, кореляційного відношення й ін.). Слід зазначити, що названі критерії є критеріями апроксимації, а не прогнозу. Однак, беручи до уваги прийняту гіпотезу про стійкість процесу в майбутньому, можна припускати, що в цих умовах модель, найбільш удала для апроксимації, буде найкращою й для прогнозу.

У ряді випадків для вибору виду функціональної залежності використовується метод, згідно з яким певні співвідношення між змінами вхідної й вихідної величин припускають ту або іншу функціональну залежність.

Дійсно, якщо виконана умова $\frac{\Delta y}{\Delta x} = const$, то ухвалюється лінійна модель

$$y = a_0 + a_1x,$$

де a_0, a_1 – коефіцієнти, які визначаються за МНК; $\Delta y, \Delta x$ – збільшення залежної й незалежної змінної, тобто $\Delta y = y_t - y_{t-1}$; $\Delta x = x_t - x_{t-1}$.

Далі маємо:

- якщо $\frac{\Delta \ln y}{\Delta x} = const$, то приймається модель $y = a_0 x^{a_1}$;

- якщо $\frac{\Delta \ln y}{\Delta \ln x} = const$, то модель обирається у вигляді $y = a_0 a_1^x$;

- якщо $\frac{\Delta y^2}{\Delta x^2} = const$, то обирається модель $y = a_0 + a_1x + a_2x^2$;

- якщо $\frac{\Delta \left(\frac{x}{y} \right)}{\Delta x} = const$, то приймається $y = \frac{x}{a_0 + a_1x}$.

Аналогічні співвідношення можна визначити й для ряду інших залежностей.



Для поглиблення знань з застосування методу найменших квадратів до побудови моделей часових рядів рекомендується література: основна – [1, 4, 5, 11], додаткова – [3, 5].

Навчально-методичне видання
(українською мовою)

Леонтьєва Вікторія Володимирівна
Кондрат'єва Наталія Олександрівна

МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ ВИРОБНИЧИХ ПРОЦЕСІВ

Навчально-методичний посібник
для студентів денного відділення математичного факультету
напряму підготовки «Прикладна математика»
освітньо-кваліфікаційного рівня «Бакалавр»

Рецензент	<i>д.т.н., професор С.І. Гоменюк</i>
Відповідальний за випуск	<i>д.т.н., професор В.З. Грищак</i>
Коректор	<i>к.ф.-м.н. В.В. Леонтьєва</i>

Адреса редакційної колегії:
Запорізький національний університет
69600, м. Запоріжжя, вул. Гоголя, 63, каб. 102а
тел. (061) 289-12-12

Підписано до друку 03.10.2011. Формат А5
Папір офсетний 80 гр/м². Гарнітура «Times New Roman»
Друк – лазерний. Наклад 20 прим. 120 с.

Віддруковано у цифровій друкарні «Сору Арт»
м. Запоріжжя, вул. Жуковського 57.
тел. (061) 289-22-00
тел. (061) 178-69-27
<http://copyart.uaprom.net/>
e-mail: copyshopzp08@rambler.ru