



Національний університет
"Києво-Могилянська академія"

І. Г. Лук'яненко, В. М. Жук

АНАЛІЗ ЧАСОВИХ РЯДІВ
Побудова ARIMA, ARCH/GARCH моделей
з використанням пакета E.Views 6.0

Частина перша

Практичний посібник для роботи
в комп'ютерному класі

Київ
2013

ББК 65.261

Λ52

Посібник містить детальні інструкції для самостійної побудови ARIMA, ARCH/GARCH моделей за допомогою пакета E.Views 6.0. Оскільки методологія та алгоритми побудови цих моделей стандартні, посібник можна використовувати і під час роботи з іншими пакетами прикладних програм.

Важливою складовою посібника є додатковий матеріал. У розділі «Додатки» зібрана інформація щодо особливостей застосування критеріїв прогнозованої якості у випадку моделювання фінансових часових рядів, програми автоматичної реалізації процедури Хенона–Рісанена тощо. Всі етапи побудови ARIMA, ARCH/GARCH моделей та прогнозування за ними проілюстровані рисунками, графіками, таблицями. До кожної теми додано ситуативні вправи з розв'язками та завдання для самостійного виконання, а також пам'ятки зі стислим описом певних теоретичних положень або необхідних формул. Наведено список рекомендованої літератури.

Розраховано на студентів економічних спеціальностей програм бакалаврату та магістратури, а також аспірантів, викладачів, фахівців – усіх, хто прагне оволодіти навичками побудови ARIMA й ARCH/GARCH моделей та зрозуміти особливості їх застосування для прогнозу на практиці.

Рецензенти:

Ільчук В. П., доктор економічних наук, професор, завідувач кафедри фінансів Чернігівського державного технологічного університету
Камінський А. Б., доктор економічних наук, професор кафедри економічної кібернетики Київського національного університету імені Тараса Шевченка

*Ухвалено до друку Вченою радою НаУКМА
(протокол № 6 від 20 червня 2013 року)*

ISBN

© І. Г. Лук'яненко, В. М. Жук, 2013
© Д. Л. Виговський, обкладинка, 2013
© НаУКМА, 2013

ЗМІСТ

Тема 1. ARIMA моделі: практика побудови із використанням програмного пакета E.Views 6.0. Детальні інструкції побудови ARIMA моделі на прикладі реального часового ряду6

Вступ	6
1.1. У яких випадках доцільно використовувати ARIMA моделі на практиці?	6
1.2. Організація даних для побудови ARIMA моделей	7
1.2.1. Які дані можна використовувати для аналізу?.....	7
1.2.2. Як створити робочий файл у E.Views та імпортувати або копіювати в нього дані?	9
1.2.3. Як здійснити первинний статистичний аналіз даних?	15
1.3. Перевірка часового ряду на стаціонарність. Визначення порядку інтеграції	18
1.3.1. Перевірка часового ряду на стаціонарність. Перетворення нестационарного часового ряду в стаціонарний	18
1.3.2. Проблема сезонності у часових рядах	22
1.4. Ідентифікація ARIMA моделі	25
1.4.1. Як визначити значущі лаги для побудови моделі?.....	25
1.4.2. Як знайти найкращу специфікацію для ARIMA моделі (визначити оптимальний порядок AR- та MA-складових)?.....	28
1.5. Остаточне оцінювання ARIMA моделі та її перевірка на адекватність ..	32
1.5.1. Як переоцінити і перевірити модель на адекватність?.....	32
1.6. Прогнозування на основі побудованої ARIMA моделі.....	43
1.6.1. Обчислення прогнозних значень на основі побудованої моделі ...	43
1.6.2. Як побудувати інтервали довіри для прогнозу?	46

Завдання, творчі роботи, ситуаційні вправи до теми 1

Завдання для самоперевірки 1	50
Творча робота 1	51
Практичний кейс 1	54
Завдання для самоперевірки 2	66

Додатки до теми 1

Додаток 1_А. Приклад визначення оптимальних лагів AR- та MA-складових в ARIMA моделі	67
Додаток 1_Б. Сезонні перетворення на прикладі ARIMA(2,2) моделі	73
Додаток 1_В. Критерії прогновної якості	74
Додаток 1_Г. Порівняльна таблиця критеріїв та орієнтирів при ідентифікації ARIMA моделі	78

Додаток 1_Г. Програма в командах пакета E.Views 6.0 щодо вибору найкращої специфікації ARIMA за допомогою процедури Хенона–Рісанена та інструкція до її виконання	80
Додаток 1_Д. Статистичні дані значень показника реального ефективного обмінного курсу євро для країн Євросоюзу (змінна REER)	88
Додаток 1_Е. Згенеровані дані часових рядів при виконанні практичного кейса для аналізу дії шоків	90

**Тема 2. Практика побудови ARCH/GARCH моделей із використанням
програмного пакета E.Views 6.0. Інструкції побудови
ARCH/GARCH моделей на прикладі реального часового ряду.....**

Вступ	97
2.1. Для чого використовувати ARCH/GARCH моделей за наявності ARIMA методології?	98
2.2. На основі яких даних можна будувати ARCH/GARCH моделей? Особливості імпортування високочастотних даних в E.Views	99
2.3. У чому полягає сутність ARCH/GARCH моделювання.....	108
2.4. Як виявити ARCH/GARCH ефект за допомогою пакета E.Views?	109
2.5. Як побудувати та оцінити ARCH/GARCH моделі?	116
2.6. Критерії перевірки ARCH/GARCH моделей на адекватність.....	135
2.7. Приклад практичного застосування результатів ARCH/GARCH моделей.....	143

Завдання, творчі роботи, ситуаційні вправи до теми 2

Завдання для самоперевірки 1	148
Творча робота 1	149
Практичний кейс 1.....	150
Практичний кейс 2	153
Завдання для самоперевірки 2.....	160
Завдання для самоперевірки 3.....	160

Додатки до теми 2

Додаток 2_А. Дані щоденних значень індексу ПФТС (змінна PFTS)	161
Додаток 2_Б. Дані котирувань (дохідності акцій) компаній «Алчевський металургійний комбінат» (ALMK) та «Металургійний комбінат “Азовсталь”» (AZST)	162
Пам’ятки до тем 1 та 2	165
Пам’ятка 1. Загальна класифікація економетричних моделей.....	165
Пам’ятка 2. Властивості найпростіших AR(1), AR(2), MA(1), ARMA(1,1) процесів	165

Пам'ятка 3. Поняття стаціонарності часового ряду, Дікі–Фулер тест перевірки часового ряду на стаціонарність. Визначення порядку ARIMA(p,d,q) моделі	169
Пам'ятка 4. Генерація за допомогою генератора випадкових чисел найпростіших AR- та MA-процесів	172
Пам'ятка 5. Тестування наявності або відсутності GARCH ефекту в часовому ряду (залишках моделі) за допомогою BDS тесту	175
Пам'ятка 6. Робота з програмою E.Views.....	180
Пам'ятка 7. Необхідні команди для виконання завдань кейса 1 теми 1	182
Література до теми 1–2.....	184

ARIMA моделі: практика побудови із використанням програмного пакета E.Views 6.0. Детальні інструкції побудови ARIMA моделей на прикладі реального часового ряду

Вступ

Основні етапи аналізу даних і побудови ARIMA моделей на практиці з використанням пакета E.Views проілюстровані на прикладі розробки ARIMA моделі для часового ряду реального ефективного обмінного курсу євро для країн Євросоюзу (РЕОК). Фактичні дані, які використовуються для побудови моделі, наведено в Додатку 1_A (таблиця 1А, часовий ряд REER).

Особливості й тонкощі моделювання на кожному етапі прокоментовано та проілюстровано за допомогою програмного пакета E.Views 6.0.

Наведений ілюстративний матеріал можна використати як для виконання на комп'ютері домашнього завдання і творчої роботи з теми ARIMA моделювання, так і для побудови реальних ARIMA моделей на практиці.

Зауважимо, що в межах одного програмного пакета однакові операції можна здійснити кількома способами, і лише один або кілька з них описані в посібнику. Це відкриває простір для самостійної творчості, яка в опануванні будь-якої справи, в тому числі економетрики, заохочується. Нагадаємо, що моделювання не тільки наука, а й мистецтво, яке залежить від розуміння проблеми, знань, досвіду та інтуїції. Творчих вам успіхів на цьому нелегкому, але надзвичайно захопливому шляху!

1.1. У яких випадках доцільно використовувати ARIMA моделі на практиці?

При прийнятті оперативних рішень на макро- та мікрорівні часто необхідно спиратись на якісні прогностні значення певних

показників, не заглиблюючись у детальний аналіз факторів, які впливають на їх зміну. Припустимо, що ми маємо часовий ряд місячних даних щодо реального ефективного обмінного курсу євро для країн Євросоюзу (РЕОК) і нам необхідно визначити прогнозне значення цього показника на наступні два місяці, а також побудувати інтервали довіри для прогнозу.

Загалом, на зміну показника РЕОК можуть впливати різні фактори, зокрема ВВП, індекси цін, а також інші чинники економічного життя не лише країн Євросоюзу, а й їхніх країн – торгових контрагентів. Збір даних та аналіз всіх цих факторів для побудови багатофакторної регресійної моделі може бути тривалим і потребувати великих витрат ресурсів, невідповідних отриманому результату.

Саме в таких випадках доцільно використовувати методи прогнозування є методи прогнозування на основі часових рядів, зокрема побудову ARIMA моделі часового ряду.

Зауважимо, що побудова ARIMA моделей є також ефективним допоміжним інструментарієм для обчислення прогнозних значень окремих факторів, що впливають на змінну залежного показника під час прогнозування на основі багатофакторних регресійних моделей, моделей симульативних рівнянь тощо.

Крім того, на практиці можна будувати різновиди ARIMA моделей, зокрема ARMAX моделі, які одночасно з лаговими змінними досліджуваного показника можуть враховувати і додаткові екзогенні фактори в різних формах. Такі своєрідні «гібриди» лінійної багатофакторної регресії та ARIMA моделей за правильної специфікації доволі ефективні та можуть забезпечити зменшення похибки прогнозу порівняно з класичними багатофакторними регресійними моделями або навіть з чистими ARIMA моделями.

1.2. Організація даних для побудови ARIMA моделей

1.2.1. Які дані можна використовувати для аналізу?

При використанні даних для ARIMA моделювання існують два основних обмеження. Перше з них є теоретичне і стосується стаціонарності часового ряду, оскільки ARIMA моделі можуть буду-

ватися тільки для стаціонарних часових рядів. Звичайно, не всі часові ряди на практиці стаціонарні, але операції різних порядків або інших перетворень, наприклад, завдяки шляхом логарифмування, вони можуть стати придатними для ARIMA моделювання. Так, на наведеному нижче графіку (див. рис. 1.1.) відтворено динаміку квартальних значень монетарного агрегату МЗ для України з 1996 до 2006 року (дані НБУ, млн грн).

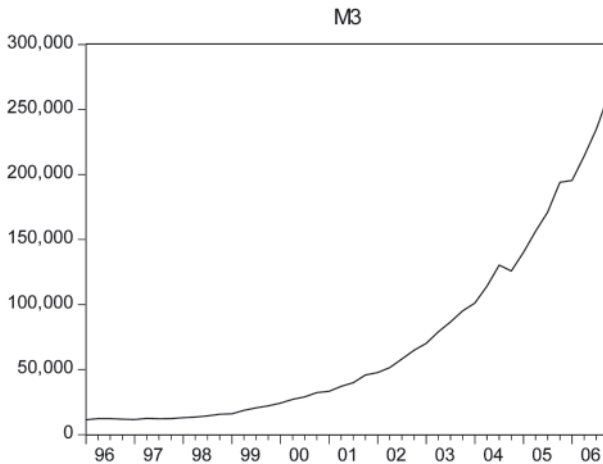


Рис. 1.1. Динаміка монетарного агрегату МЗ за 1996–2006 роки

Як можна побачити з рисунка, дані зростають подібно до експоненційної функції. Перші та другі різниці такого ряду не є стаціонарним рядом. Утім, логарифмування ряду дає змогу експоненційний тренд у даних зробити лінійним, і перші різниці ряду стають стаціонарним рядом, відповідно ARIMA модель можна будувати на логарифмах перших різниць такого ряду.

Друге обмеження технічне і більше стосується правильної організації даних. Так, для спрощення імпортування часових рядів даних у пакет E.Views їх потрібно правильно підготувати. Бажано, щоб у значень змінних десятковим розділювачем була крапка, дані наводилися в стовпчик або в рядок – без проміжних підсумків (як-от *загалом за квартал, рік* тощо). Краще використовувати файл даних у форматах .xls, .csv, .txt і перед імпор-

туванням переконатися, що він не містить кириличних символів. Зауважимо, що пакет E.Views також підтримує і деякі інші, менш поширені, формати баз даних, які в цьому посібнику не розглядаємо.

Нарешті, варто пам'ятати, що ARIMA аналіз пов'язаний суто з аналізом часових рядів, для яких необхідно знати: хронологічні межі ряду; одиницю виміру; регулярність (чи виміряні дані через однакові проміжки часу); частоту вимірів (періодичність).

1.2.2. Як створити робочий файл у E.Views та імпортувати або копіювати в нього дані?

Перед імпортом даних у пакет E.Views слід створити робочий файл, виділивши в головному меню послідовно: **File>New>Workfile**. Відкриється вікно робочого файлу (рис. 1.2)

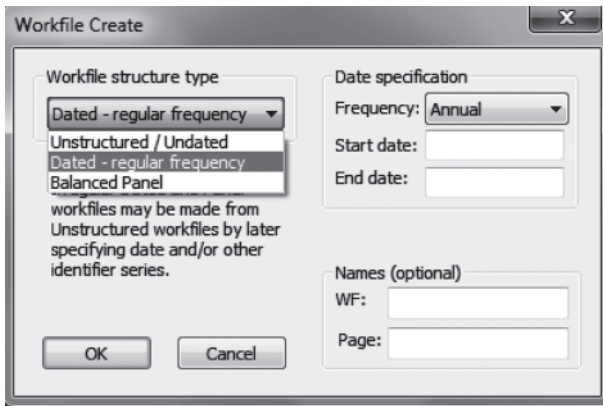


Рис. 1.2. Вікно для створення робочого файлу в E.Views

У вікні, що відкривається, потрібно вказати тип даних (вхідної інформації) для побудови моделі. У випадку одного часового ряду можливі два варіанти: структуровані (регулярні) та неструктуровані (нерегулярні) дані.

Опцію **Unstructured/undated (Неструктуровані/недатовані)** слід обрати, якщо дані для побудови моделі є, наприклад, щоденними (наявність вихідних і святкових днів порушує періодичність),

щохвилинними (характерно для біржових показників), неперіодичними. У такому разі ряд ефективно індексувати просто за номером спостереження. Звичайно за необхідності можна створити окремий ряд даних із датами у форматі **PPPP / MM / ДД**, а потім процедурою **Proc>Structure/Resize Current page** головного меню програми (зауважимо, що в кожного вікна в програмі теж є меню, яке називатимемо «меню вікна») задати необхідні дати для введених спостережень.

Опцію **Dated – regular frequency** (регулярні дані) варто обрати тоді, коли дані виміряні через однакові проміжки часу, наприклад, рік, квартал, місяць, тиждень, день тощо (див. рис. 1.3). Зауважимо, що пакет E.Views надає дві можливості організації щоденних даних із 5-денним та 7-денним робочими тижнями*.

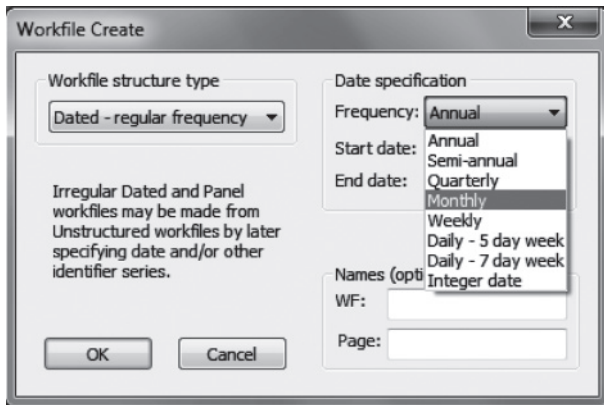


Рис. 1.3. Специфікація даних у пакеті E.Views

У створеному робочому файлі вказують кількість спостережень, яку необхідно імпортувати (цей показник може бути завищений).

У наведеному прикладі для організації даних (нагадаємо, що вхідна інформація містить щомісячні дані щодо реального ефективного курсу євро з січня 2000 до червня 2008 року у відсотках, та припускаємо, що вводити ми її будемо з таблиці програми Excel,

* Для організації українських щоденних даних краще обрати неструктурований варіант організації даних з огляду на велику кількість свят, а отже, на відсутність необхідних спостережень у визначені періоди часу.

файл REER.xls) обираємо опцію **Monthly**, оскільки дані є щомісячними (фактичні дані наведено у додатку 1_A, таблиця 1А, часовий ряд REER). Поля стосовно початку й кінця інтервалу заповнюють таким чином:

Start date: 2000m1 або (2000:1)

End date: 2008m06 або (2008:6).

Примітка: В E.Views використовуються різні розділювачі для виокремлення номера кварталу, місяця чи півріччя. Наприклад: запис 2000m2 означає: лютий 2000 року, 1999q4 – четвертий квартал 1999 року, 2003s2 – друге півріччя 2003 року. Якщо тип періодичності вже заданий, можна використовувати універсальний розділювач – двокрапку, тобто 2000:2, 1999:4, 2003:2. При застосуванні двокрапки треба бути уважним. Наприклад, для квартальних даних запис 1999:5 помилковий, а для щомісячних інтерпретуватиметься як травень відповідного року.

Після задання типу даних, а також дат початку та кінця спостережень можна здійснювати безпосереднє введення даних у робочий файл двома способами: шляхом імпортування або простого копіювання даних. Прокоментуємо кожний з них детальніше.

Введення даних у робочий файл (спосіб 1).

Імпортування

Якщо необхідні для розрахунків статистичні дані організовані в форматах .xls, .txt, .dat, .csv та подібних, їх можна безпосередньо імпортувати у робочий файл. Для цього почергово обираємо в головному меню **File>Import>Read Text-Lotus-Excel**. У списку знизу можна обрати або ASCII-файли, які можуть бути з різними розширеннями (у тому числі CSV, PRN, DAT тощо) або XLS-файли для імпорту таблиць з програми Excel. Зауважимо, що вибір ASCII-файлів нетиповий для фінансових статистичних рядів, однак розглянемо і таку можливість (на випадок роботи з базами так званих сирих (необроблених) даних, наприклад, з бухгалтерських програм чи опитувань). При виборі цієї опції відкриється вікно, зображене на рис. 1.4.

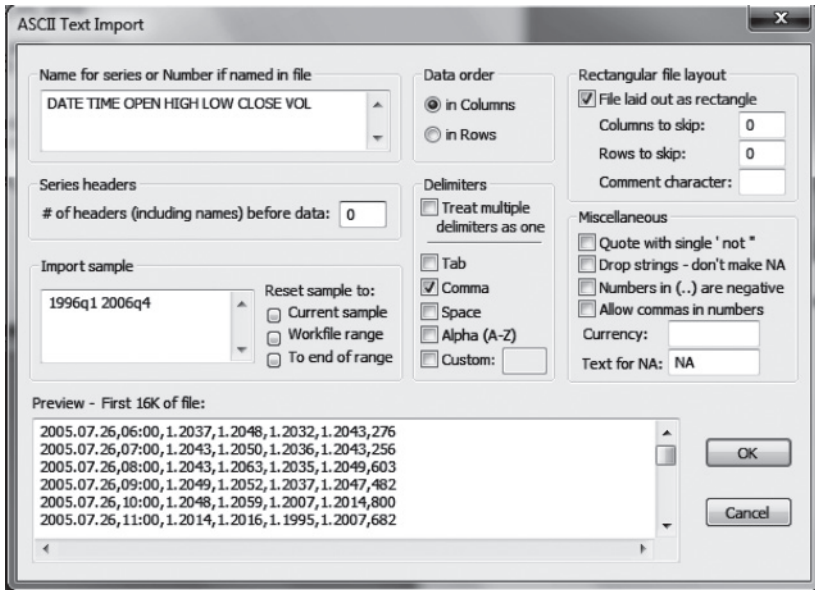


Рис. 1.4. Вікно імпортування даних у форматі ASCII

У вікні можна задати розділювач або кілька розділювачів (у нашому випадку задана кома), кількість колонок чи рядків для пропуску, кількість рядків заголовків. Якщо заголовки є у файлі, в лівому верхньому полі слід лише вказати кількість рядів даних для імпорту, і програма сама зчитає назви заголовків. Крім того, необхідно звернути увагу, чи дані організовані як рядки, чи як колонки. Якщо файл не містить назв часових рядів, їх можна вписати в те саме поле, розділивши проміжками по порядку. Також треба звернути увагу на вікно **Import sample**, в якому задають часовий інтервал для імпортованих даних. Наприклад, на рис. 1.4 зображений варіант неправильного імпорту – імпортуються погодинні дані, а відрізок вказаний для квартальних даних іншого часового проміжку. Тому потрібно бути дуже уважними, особливо при роботі з кількома робочими файлами одночасно.

Імпорт даних із таблиць з програми Excel (опція XLS) найпоширеніший на практиці. Вибір цієї опції здійснюється послідовністю команд **File>Import>Read Text-Lotus-Excel**, після чого оби-

рається безпосередньо файл (Excel) із даними, які необхідно імпортувати. Це показано на рис. 1.5.

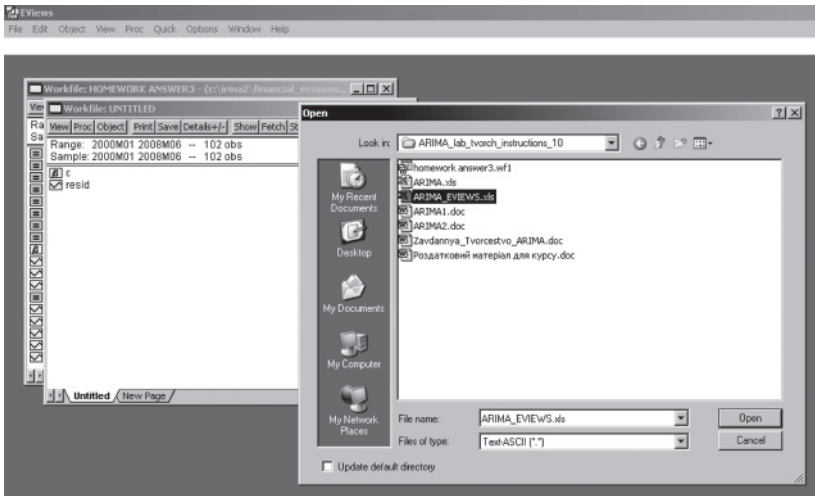


Рис. 1.5. Вікно для вибору файлу з даними, підготовленими для імпорту

Після виділення необхідного файлу та натискання клавіші «Open», відкриється вікно, як на рис. 1.6.

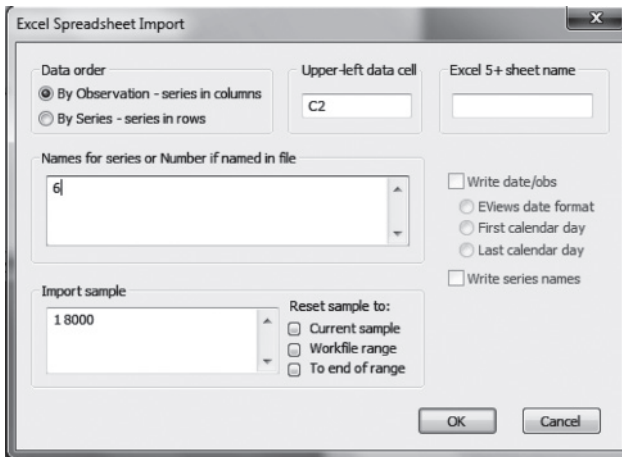


Рис. 1.6. Вікно імпортування даних у форматі XLS

У цьому вікні задають атрибути, як у розглянутому вище випадку імпорту даних у форматі ASCII. Слід звернути увагу на те, що програма не визначає автоматично верхню ліву клітинку даних у файлі, тому варто її задати. **Задавати треба клітинку, в якій починаються безпосередньо дані!** Якщо задати клітинку із заголовками, програма їх не розпізнає.

Якщо заголовки наявні у файлі, в опції **Names for series or Numbers if named in file**, слід лише вказати кількість колонок (рядів) даних для імпорту, і програма сама читає назви заголовків. Якщо файл не містить назв часових рядів, їх можна вписати в те саме поле, розділивши проміжками, по порядку. Крім того, необхідно зазначити, чи дані організовані як рядки, чи колонки в опції **Data orders**.

Переваги імпортування даних особливо відчутні, якщо обсяг файлу дуже великий. Крім того, це найкращий спосіб введення даних з екзотичних текстових форматів (такі часто трапляються в базах даних бухгалтерських програм, таких, наприклад, як поширена на Заході Peachtree).

Недоліком імпортування даних є те, що, попри велику кількість налаштувань, користувач не отримує повного контролю над тим, як читаються дані, і тому з першого разу дані можуть бути імпортовані неякісно, з великою кількістю «невизначених» (NA) спостережень. У таких випадках необхідно уважно проаналізувати причину неякісного імпорту, виявити проблему та повторити процедуру введення даних ще раз.

ПАМ'ЯТАЙТЕ: файл, який імпортується, слід перед імпортом закрити в усіх інших програмах.

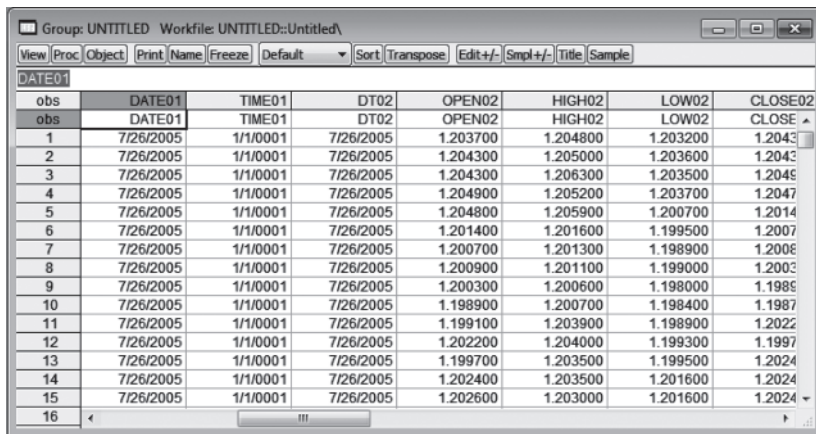
Введення даних у робочий файл (спосіб 2).

Копіювання

Копіювання даних здійснюється безпосередньо в робочий файл. Для цього слід активувати вікно готового робочого файлу і послідовно обрати в головному меню опції: **Quick>Empty Group (Edit Series)**. У таблицю, яка відкриється, потрібно вставити скопійовані дані з будь-якого табличного редактора, виконавши перед тим процедури копіювання (Ctrl+C), наприклад з Excel-файлу, та вставки (Ctrl+V) у відкриту робочу таблицю E.Views-файлу.

Звичайно перед копіюванням слід переконатися, що ряди даних розміщені в колонках, вони не містять кириличних символів тощо. На рисунку 1.7 зображено вікно з уже вставленими після копіювання даними.

ПАМ'ЯТАЙТЕ: щоб одночасно перенести скопійовані дані та заголовки, варто виділити будь-яку верхню клітинку таблиці E.Views-файлу і натиснути на клавіатурі стрілочку «вгору». Тоді вставляються дані з заголовками, оскільки рядок заголовків також стає придатним для редагування (див. рис. 1.7).



obs	DATE01	TIME01	DT02	OPEN02	HIGH02	LOW02	CLOSE02
1	7/26/2005	1/1/0001	7/26/2005	1.203700	1.204800	1.203200	1.204300
2	7/26/2005	1/1/0001	7/26/2005	1.204300	1.205000	1.203600	1.204300
3	7/26/2005	1/1/0001	7/26/2005	1.204300	1.206300	1.203500	1.204900
4	7/26/2005	1/1/0001	7/26/2005	1.204900	1.205200	1.203700	1.204700
5	7/26/2005	1/1/0001	7/26/2005	1.204800	1.205900	1.200700	1.201400
6	7/26/2005	1/1/0001	7/26/2005	1.201400	1.201600	1.199500	1.200700
7	7/26/2005	1/1/0001	7/26/2005	1.200700	1.201300	1.198900	1.200800
8	7/26/2005	1/1/0001	7/26/2005	1.200900	1.201100	1.199000	1.200300
9	7/26/2005	1/1/0001	7/26/2005	1.200300	1.200600	1.198000	1.198900
10	7/26/2005	1/1/0001	7/26/2005	1.198900	1.200700	1.198400	1.198700
11	7/26/2005	1/1/0001	7/26/2005	1.199100	1.203900	1.198900	1.202200
12	7/26/2005	1/1/0001	7/26/2005	1.202200	1.204900	1.199300	1.199700
13	7/26/2005	1/1/0001	7/26/2005	1.199700	1.203500	1.199500	1.202400
14	7/26/2005	1/1/0001	7/26/2005	1.202400	1.203500	1.201600	1.202400
15	7/26/2005	1/1/0001	7/26/2005	1.202600	1.203000	1.201600	1.202400

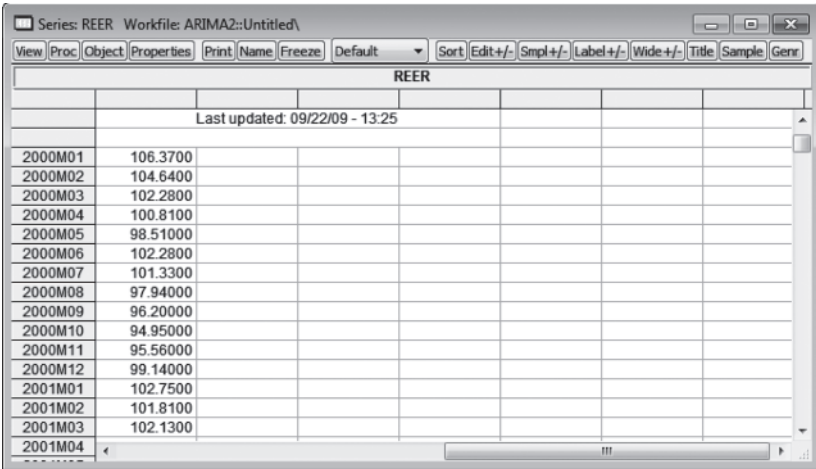
Рис. 1.7. Вікно E.Views-файлу
зі скопійованими даними та заголовками

Якщо в ряді даних є текст або інші нецифрові записи, такий ряд програма розпізнає не як ряд даних, а як об'єкт «alpha» – набір текстових рядків. У такому випадку слід виправити дані, видалити такий окремий ряд (об'єкт) (не з таблиці об'єкта «group», як на рисунку, а з робочого файлу) і скопіювати дані наново.

1.2.3. Як здійснити первинний статистичний аналіз даних?

E.Views надає широкий спектр можливостей для первинного аналізу введених даних, їх графічної візуалізації та групування. Зокрема, можна зберігати дані як окремі ряди (**Series**) або об'єдну-

вати їх у групи (**Groups**). При цьому називати можна як ряди, так і групи. Оскільки ARIMA моделі, як правило, базуються на моделюванні поведінки однієї змінної, то найбільше уваги приділимо аналізу об'єкта Series – ряду даних. Загалом, введений ряд можна відобразити в таблиці (Spreadsheet), яку для часового ряду REER (реальний ефективний обмінний курс) наведено на рис. 1.8.



REER	
Last updated: 09/22/09 - 13:25	
2000M01	106.3700
2000M02	104.6400
2000M03	102.2800
2000M04	100.8100
2000M05	98.51000
2000M06	102.2800
2000M07	101.3300
2000M08	97.94000
2000M09	96.20000
2000M10	94.95000
2000M11	95.56000
2000M12	99.14000
2001M01	102.7500
2001M02	101.8100
2001M03	102.1300
2001M04	

Рис. 1.8. Представлення часового ряду у формі таблиці

Зауважимо, що меню вікна, відображеного у вигляді таблиці часового ряду, включає достатню кількість опцій (рис. 1.8). Так, опція **Edit+/-** дає можливість редагувати (+) чи не редагувати (-) ряд даних. За замовчуванням редагувати значення ряду даних не можна. Опція **Default** (вид подачі даних) дає змогу залишати дані без змін або обирати вид їх перетворення. Відповідно потрібно обрати підопцію (**Безпосередні дані**; **Різниці**; **Різниці за рік**; **% зміна** тощо).

Більш детально опис кожної опції меню вікна можна знайти в розділі Help E.Views-пакета.

Якщо необхідно скопіювати ряд даних, то для цього необхідно послідовно обрати в меню вікна **Object>Copy object**. Новоствореному рядові можна дати назву за допомогою команди **Name**. З безіменним рядом важко працювати (за замовчуванням ряду на-

дають індифікатори SER01, SER02... за кількістю рядів у робочому файлі).

Для графічної візуалізації ряду використовують команду **View>Graph**^{*}, яка дає можливість відобразити графік часового ряду в різних формах (див. рис. 1.9).

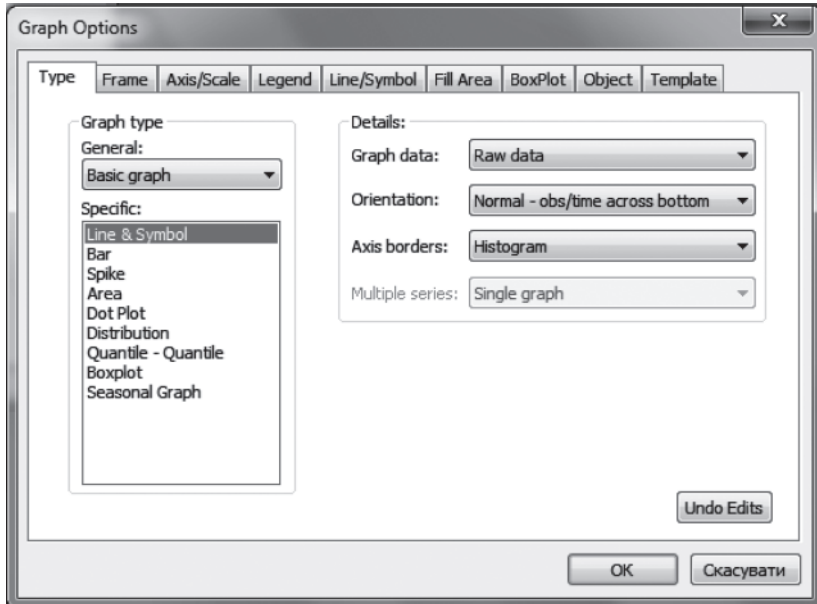


Рис. 1.9. Вікно вибору опцій для побудови графіка часового ряду

Версія E.Views 6.0 надає значно більші графічні можливості порівняно з попередніми версіями. Головне вікно вибору типів графіків представлено на рисунку 1.9. За допомогою численних параметрів цього графічного інтерфейсу можна налаштувати параметри відображення окремих складових графіка. Наприклад, оберемо простий лінійний граф на основі необроблених даних з додатковою гістограмою розподілу вздовж осі, відповідно отримаємо графік як на рис. 1.10.

^{*} Надалі для прикладу побудови ARIMA будемо використовувати ряд реального ефективного курсу (PEOK) євро для країн єврозони з січня 2000 року до червня 2008 року (щомісячні дані).

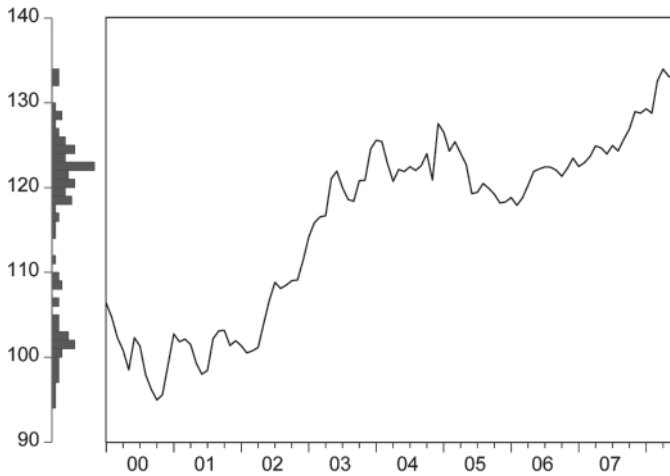


Рис. 1.10. Графік часового ряду REER євро

На графіку бачимо позитивний тренд і як довгострокову, так і короткострокову циклічність. Тож, гістограма розподілу, яка додатково представлена ліворуч, має два піки. До того ж, варто зазначити, що графік не перетинає значення 0 (це логічно для відсоткового курсу). Зроблені висновки будуть потрібні на наступному етапі – перевірка часового ряду на стаціонарність.

1.3. Перевірка часового ряду на стаціонарність. Визначення порядку інтеграції

1.3.1. Перевірка часового ряду на стаціонарність.

Перетворення нестаціонарного часового ряду в стаціонарний

Найпоширеніший тест перевірки часового ряду на стаціонарність – розширений тест Дікі–Фулера. Для його виконання для виділеного ряду необхідно поступово обрати в меню: **View>Unit root test**, відкриється вікно, зображене на рис. 1.11.

У полі «тип тесту» слід, згідно з попереднім аналізом графіка часового ряду (нагадаємо, що ряд має чітко виявлену тенденцію до

зростання), вказати: **Trend and intercept**, припускаючи в базовому рівнянні Дікі–Фулера наявність лінійного тренду та ненульового значення перетину. Крім того, в шостій версії EViews автоматично можна оптимізувати кількість лагів у базовому рівнянні Дікі–Фулера, обравши відповідну опцію в підменю: **Automatic selection**. У цьому прикладі було обрано оптимізацію кількості лагів за інформаційним критерієм Шварца.

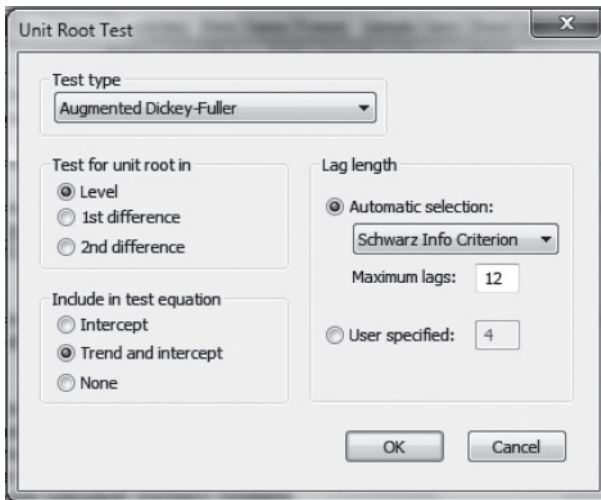


Рис. 1.11. Вікно для визначення параметрів Дікі–Фулера тесту перевірки часового ряду на стаціонарність

Зауважимо, що перший крок – це тестування стаціонарності у первісному часовому ряді. Протестуємо дані в рівнях, обравши в опції **Test for unit root in** підопцію **Levels** (рівні). Натиснувши клавішу ОК, отримаємо таке вікно, як на рис. 1.12.

Основний результат Дікі–Фулера тесту – отримання розрахункового значення та критичних значень т-статистики МакКінона. Як можна побачити на рис. 1.12, абсолютна величина розрахункового значення т-статистики МакКінона (-2.193343) менша за абсолютні величини критичного значення при 1 %, 5 % та 10 % рівня значущості. Крім того, p-value (імовірність) тесту дорівнює 0,487 (48,7 %, тобто **p-value > 10 %**). Отже, ми не можемо відкинути ну-

льову гіпотезу щодо наявності одиничного кореня в часовому ряді, оскільки мінімальна ймовірність того, що ряд має одиничний корінь і не є стаціонарним, становить 49 % ($p\text{-value} > 10\%$).

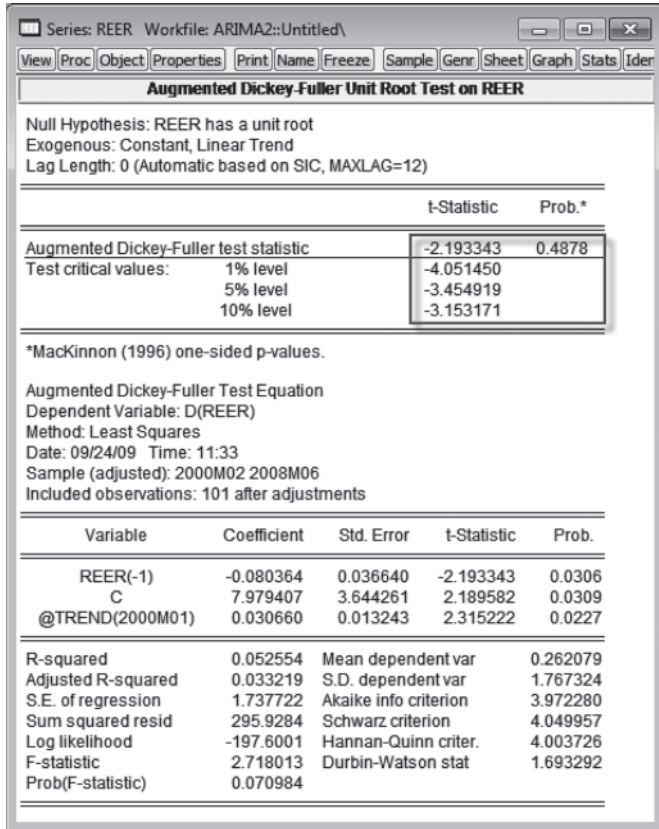


Рис. 1.12. Вихідне вікно результатів виконання Дікі–Фулера тесту перевірки часового ряду в рівнях на стаціонарність

Таким чином, ми не відкидаємо нульову гіпотезу про наявність одиничного кореня. Це означає, що ряд в рівнях нестационарний. Потрібно перевірити, чи зможемо ми операцією перших різниць перетворити ряд на стаціонарний.

Отже, варто провести аналогічний Дікі–Фулера тест, але вже для ряду в перших різницях. Для цього необхідно повторити всі

зазначені вище процедури, обравши в полі опції **Test for unit root in** підопцію **1st Difference**. Зауважимо: якщо ряд в перших різницях не має вираженого тренду, його можна не задавати в базовому рівнянні Дікі–Фулера тесту; константу бажано вводити завжди. На рис. 1.13 відображено результати розрахунків за Дікі–Фулера тестом досліджуваного часового ряду в перших різницях.

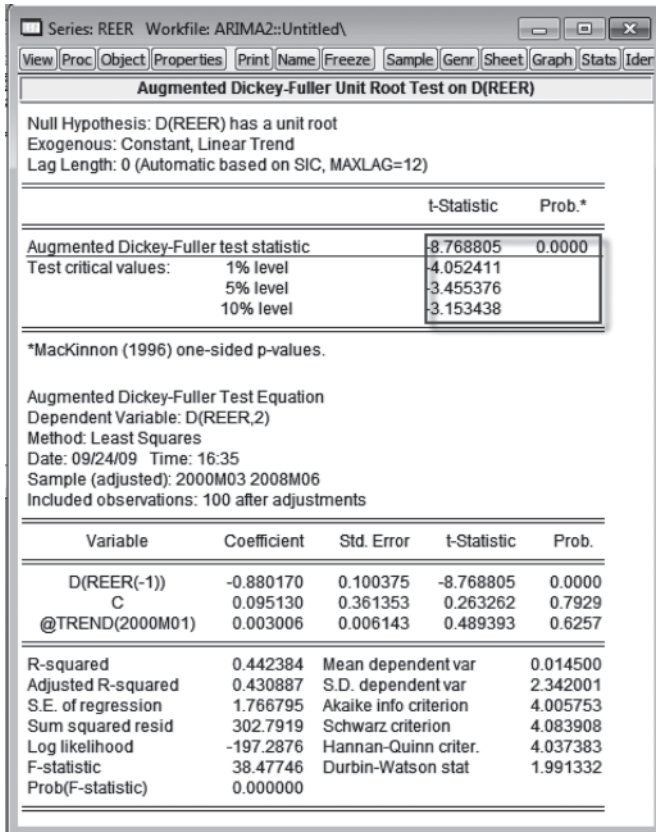


Рис. 1.13. Результати виконання Дікі–Фулера тесту перевірки часового ряду в перших різницях на стаціонарність

Результати виконання Дікі–Фулера тесту для ряду в перших різницях свідчать про стаціонарність перетвореного ряду. Як можна побачити, абсолютна величина розрахункового значення

t-статистики МакКінона (–8,76) менша за абсолютні величини критичного значення при 1 %, 5 % та 10 % рівнях значущості. Крім того, p-value (імовірність) тесту дорівнює 0,000 (0 %, тобто **p-value < 10 %**). Це означає, що ми відкидаємо нульову гіпотезу щодо наявності одиничного кореня (нестационарності) в ряді перших різниць із мінімальною імовірністю помилитися (майже в 0 % випадків зі 100 %).

Таким чином, ряд у перших різницях є стаціонарним, а ряд у рівнях **має порядок інтеграції 1**.

Отже, ARMA модель будуватимемо для ряду в перших різницях і на цьому етапі маємо **AR(?)I(1)MA(?)**.

1.3.2. Проблема сезонності у часових рядах

Якщо ряд має сезонну періодичність, тобто коливання з періодом у чотири спостереження для квартальних даних, у дванадцять – для щомісячних тощо, то можна застосувати оператор сезонних різниць. Для цього ряд перетворюють, наприклад, так: **d(REER, 1, 12)** – перші різниці і сезонне диференціювання для щомісячних даних з лагом дванадцять.

Отриманий внаслідок такої операції ряд можна записати таким чином:

$$(1-L)^1(1-L)^{12}X_t = (1-L)(X_t - X_{t-12}) = X_t - X_{t-12} - X_{t-1} + X_{t-13}.$$

Аналогічний результат можна отримати, виконавши, наприклад, команду **d(REER–REER (–12))**. Зауважимо, що для ілюстрації було обрано назву ряду REER з розглянутого прикладу.

Однак, щоб визначити, чи варто застосовувати сезонне диференціювання, необхідно проаналізувати корелограм часового ряду. Для отримання графіка корелограма для об'єкта Series (ряд REER) використаємо опції меню **View>Correlogram** і виберемо перші різниці. Відповідний корелограм наведено на рис. 1.14.

Зауважимо, що за наявності сезонності в аналізованому часовому ряді значення ACF/PACF для 12-го лагу мали би бути статистично значущими (статистично значимо відрізнятись від нуля). Утім, як можна побачити на рис. 1.14, вони є незначущими (стовп-

чки на 12-му лазі не заходять за пунктирну лінію), тобто значення ACF/PACF з лагом 12 не виходять за межі інтервалу довіри, відповідно, статистично незначно відрізняються від нуля.

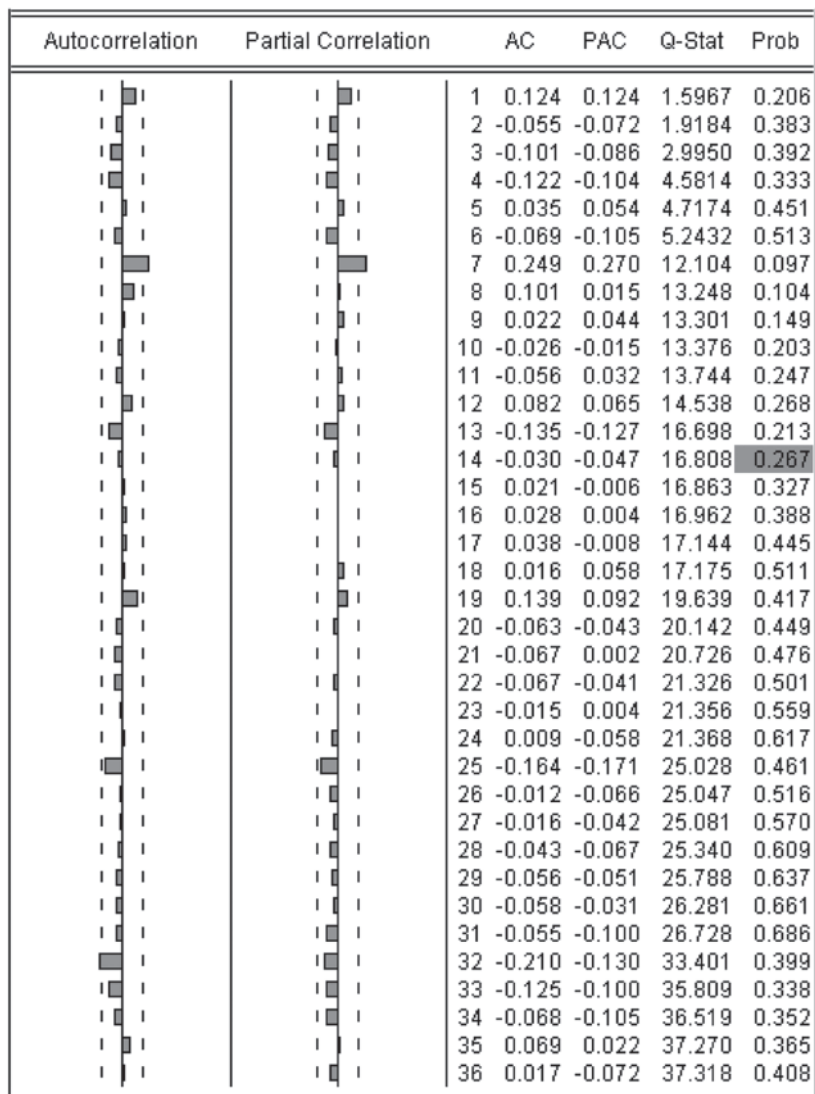


Рис. 1.14. Графік корелограма ряду REER у перших різницях

Якщо, незважаючи на цей факт, є потреба врахувати наявність сезонності в моделі, варто згенерувати новий ряд із сезонним диференціюванням, обравши в меню **Quick> Generate series**.

При цьому відкривається вікно, зображене на рис. 1.15, у якому в опції **Enter Equation** потрібно записати формулу для розрахунку нового ряду. Для нашого прикладу створюємо новий ряд REER12, який утворюється зі старого шляхом сезонного диференціювання для щомісячних даних із лагом дванадцять.

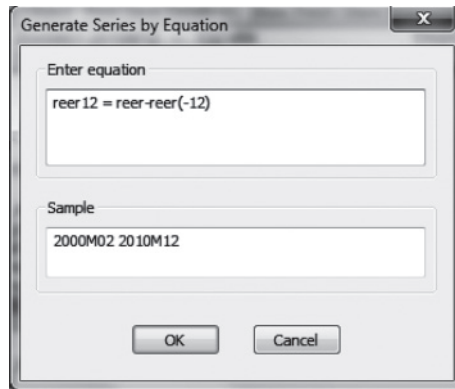


Рис. 1.15. Генерація нового ряду з сезонним диференціюванням

Проаналізуємо для новоствореного ряду REER12 корелограм у перших різницях (див. рис. 1.16).

Як можна побачити на графіку, операція сезонного диференціювання в нашому випадку спродувала статистично значущі 12 та 1-й лаги, незначущий 7-й лаг (який до сезонного диференціювання був значущим!). Тобто тенденції даних були штучно викривлені (інакше кажучи, з'явилися артефакти). Отже, для нашого прикладу сезонні перетворення не мають сенсу, оскільки сезонність чітко не проявлена.

Загалом, сезонне диференціювання слід застосовувати в таких випадках:

- 1) графіки **ACF/PACF** показують наявність статистично значущих лагів порядку 12, 24, 36 тощо для місячних даних (або 4, 8, 12... для квартальних); як наслідок, значення коефіцієнтів автокореляції повільно зменшуються до нуля;

Sample: 2000M01 2008M06
Included observations: 89

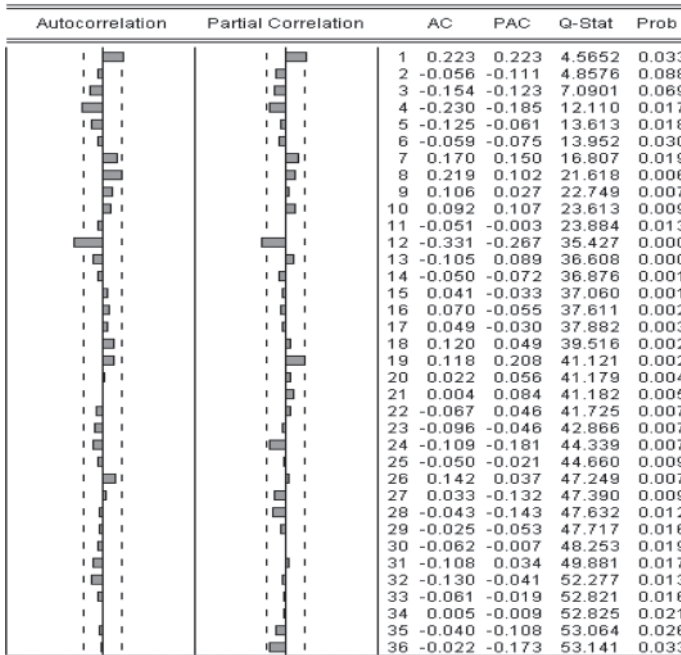


Рис. 1.16. Графік корелограма ряду з сезонним диференціюванням у перших різницях

- 2) диференціювання несуттєво впливає на значущість інших лагів, прибираючи сезонні перетворення.

1.4. Ідентифікація ARIMA моделі

1.4.1. Як визначити значущі лаги для побудови моделі?

Для попереднього визначення загального вигляду специфікації майбутньої ARIMA моделі й кількості лагів для кожної складової скористаємося графіками автокореляційної і часткової автокореляційної функції досліджуваного показника, для нашого прикладу – реального ефективного обмінного курсу в перших різницях (нагадаємо, що саме в перших різницях ряд стаціонарний). Заува-

жимо, що візуальний аналіз ACF/PACF дає змогу зробити висновок, чи можна вважати часовий ряд чистим AR- або MA-процесом, чи він є змішаним ARMA процесом. У випадку чистих процесів аналіз графіків ACF/PACF функцій уможлиблює зробити попередній висновок про максимально можливу кількість лагів. За змішаного процесу необхідно застосовувати спеціальні процедури ідентифікації, зокрема процедуру Хенона-Рісанена. Основні властивості графіків ACF- та PACF-функцій для MA-, AR- та ARMA процесів наведено в таблиці 1.1.

Таблиця 1.1. Властивості ACF- та PACF-функцій для MA-, AR- та ARMA процесів

Процеси	Поведінка ACF	Поведінка PACF
AR(p)	Нескінченно зменшується до нуля	Скінченний: дорівнює нулю після лагу p. Максимальний порядок процесу – останній ненульовий лаг в PACF (p)
MA(q)	Скінченний: дорівнює нулю після лагу q. Максимальний порядок процесу – останній ненульовий лаг в ACF(q)	Нескінченно зменшується до нуля
ARMA(p,q)	Нескінченно зменшується до нуля Максимальні порядки (p,q) процесу визначаються за спеціальними процедурами	Нескінченно зменшується до нуля Максимальні порядки (p,q) процесу визначаються за спеціальними процедурами

Отже, для попереднього визначення типу процесу та кількості лагів для включення у модель, повернемося знову до графіка автокореляційної й часткової автокореляційної функції перших різниць реального обмінного курсу (нагадаємо, що отримати його можна, задавши послідовність команд **View>Correlorgam**). Графік наведено на рис. 1.17 (див. також аналогічний рис. 1.14)

Аналіз наведеного рисунка показує, що в цьому випадку значущим є 7-й лаг, а також, можливо, 32-й. Це цілком імовірно, враховуючи «велику хвилю» майже трирічної циклічності показника (за візуальним аналізом графіка). Крім того, аналіз поведінки ACF/PACF свідчить про змішаний характер процесу. Отже, для

виявлення порядків AR- та MA-складової необхідно застосовувати спеціальні процедури.

Date: 10/31/10 Time: 19:14

Sample: 2000M01 2008M06

Included observations: 101









































































Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.124	0.124	1.5967	0.206
		2	-0.055	-0.072	1.9184	0.383
		3	-0.101	-0.086	2.9950	0.392
		4	-0.122	-0.104	4.5814	0.333
		5	0.035	0.054	4.7174	0.451
		6	-0.069	-0.105	5.2432	0.513
		7	0.249	0.270	12.104	0.097
		8	0.101	0.015	13.248	0.104
		9	0.022	0.044	13.301	0.149
		10	-0.026	-0.015	13.376	0.203
		11	-0.056	0.032	13.744	0.247
		12	0.082	0.065	14.538	0.268
		13	-0.135	-0.127	16.698	0.213
		14	-0.030	-0.047	16.808	0.267
		15	0.021	-0.006	16.863	0.327
		16	0.028	0.004	16.962	0.388
		17	0.038	-0.008	17.144	0.445
		18	0.016	0.058	17.175	0.511
		19	0.139	0.092	19.639	0.417
		20	-0.063	-0.043	20.142	0.449
		21	-0.067	0.002	20.726	0.476
		22	-0.067	-0.041	21.326	0.501
		23	-0.015	0.004	21.356	0.559
		24	0.009	-0.058	21.368	0.617
		25	-0.164	-0.171	25.028	0.461
		26	-0.012	-0.066	25.047	0.516
		27	-0.016	-0.042	25.081	0.570
		28	-0.043	-0.067	25.340	0.609
		29	-0.056	-0.051	25.788	0.637
		30	-0.058	-0.031	26.281	0.661
		31	-0.055	-0.100	26.728	0.686
		32	-0.210	-0.130	33.401	0.399
		33	-0.125	-0.100	35.809	0.338
		34	-0.068	-0.105	36.519	0.352
		35	0.069	0.022	37.270	0.365
		36	0.017	-0.072	37.318	0.408

Рис. 1.17. Графік корелограма ряду REER у перших різницях

Примітка. Якщо автокореляції виявляють сезонний характер, треба спробувати застосувати сезонну компоненту в рівнянні. Особливість введення такої компоненти в модель буде пояснено далі.

1.4.2. Як знайти найкращу специфікацію для ARIMA моделі (визначити оптимальний порядок AR- та MA-складових)?

Якщо процес не є чистим авторегресійним процесом або процесом ковзного середнього, слід визначити порядок AR- та MA-складових, наприклад, за допомогою процедури Хенона–Рісанена, яку розглянемо детальніше.

За цією процедурою спочатку звичайним методом найменших квадратів оцінюють AR-складову ARMA/ARIMA моделі. При цьому оптимальними лагами для включення в модель вважають такі, за яких досягається мінімальне значення Акайк-інформаційного критерію (AIC).

Процедуру починають з послідовного оцінювання AR моделей різного порядку, від першого до останнього, наприклад, AR(1), AR(2)... AR(k). Для кожної з моделей аналізуються значення AIC-критерію (зауважимо, що в пакеті E.Views цей критерій розраховують автоматично і подають у вихідній інформації). Оптимальним вважають таке значення лагу (наприклад, p^*), при якому досягається найменше значення Акайк критерію.

Примітка. У випадку, коли таких лагів-кандидатів кілька або є сумніви щодо остаточного вибору лагу, можна додатково застосувати таку процедуру:

- 1) першою оцінити регресію, в яку включити максимально можливу кількість лагів. Проаналізувати значення AIC-критерію для цієї моделі;
- 2) видалити з оціненої регресії лаг із найбільшим значенням p-value для відповідного коефіцієнта і порівняти отримане значення AIC критерію з попереднім значенням;
- 3) якщо AIC став меншим, повторити процедуру з наступним лагом. Якщо значення AIC критерію зросло, повернути видалену на попередньому кроці лагову змінну в рівняння і видалити змінну з лагом, для якого p-value є наступним за величиною;

4) продовжувати процедуру до моменту, коли будь-який видалений лаг збільшуватиме значення AIC.

Початкове рівняння може містити не одну лагову змінну, а декілька, зокрема це можуть бути статистично значущі в графіках автокорелограма лагові змінні.

Зауважимо, що для нашого прикладу в початкове рівняння AR-складової було введено 7-й та 32-й лаги як статистично значущі, а потім по черзі додавались 1-й, 2-й і т. д. лаги, що допомогло значно скоротити процес відбору. Крім того, використовувалася процедура, описана в примітці 1.

Для оцінки першого рівняння AR-складової (а також всіх наступних) необхідно послідовно обрати в головному меню **Quick/Estimate equation**. Після цього відкриється вікно, в якому потрібно записати специфікацію рівняння в форматі E.Views (див. рис. 1.18).

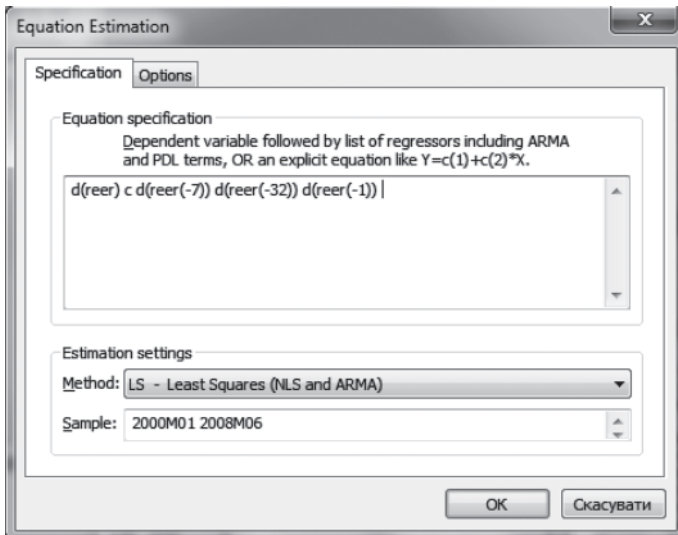


Рис. 1.18. Вікно для оцінки початкового рівняння для визначення оптимального порядку AR-складової моделі ARMA/ARIMA

Результати оцінювання цього рівняння наведено на рис. 1.19.

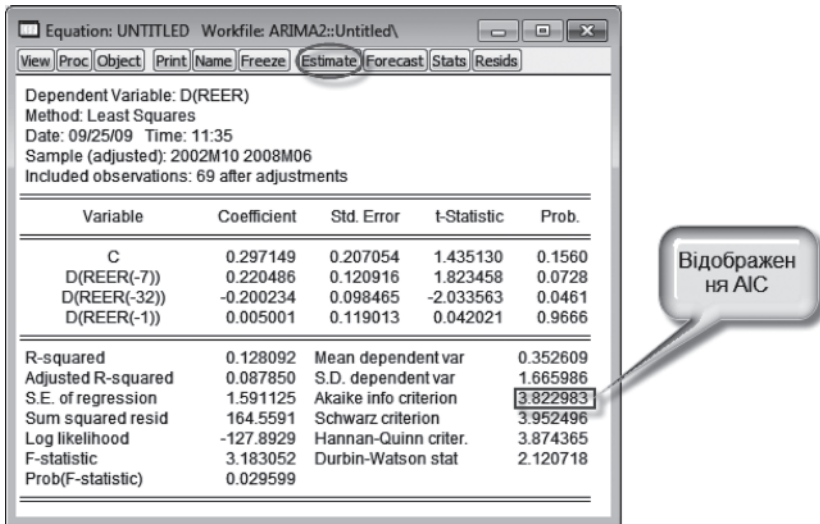


Рис. 1.19. Результати оцінювання початкового рівняння для визначення оптимального порядку AR-складової моделі ARMA/ARIMA

Як можна побачити на рис. 1.19, значення AIC критерію (інформаційного Акайк критерію) дорівнює 3,822983. Зауважимо: для оцінювання нового рівняння, не обов'язково його записувати заново, можна відредагувати старе, скориставшись підопцією меню цього вікна **Estimate**. Завдяки натисненню **Estimate** ми маємо можливість відредагувати рівняння, що оцінюється, додавши, змінивши або видаливши з нього будь-яку змінну.

Результати оцінювання послідовності моделей AR-складових із відповідними значеннями Акайк критерію наведено в додатку 1_A.

Послідовне оцінювання AR-складових показало, що мінімальне значення AIC критерію, яке дорівнювало **3,777527948**, було досягнуте для специфікації загального вигляду:

$$D(reer) = m + \alpha_6 D(reer(-6)) + \alpha_7 D(reer(-7)) + \alpha_{32} D(reer(-32)) \quad (1.1)$$

або в більш стандартному вигляді:

$$X_t = m + \alpha_6 X_{t-6} + \alpha_7 X_{t-7} + \alpha_{32} X_{t-32}, \quad (1.2)$$

де $X_t = D(reer)_t$ – перші різниці ефективного реального обмінного курсу в t -період часу.

Результати оцінювання AR-складової з мінімальним значенням Акайк критерію наведені на рис. 1.20.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.377405	0.199959	1.887412	0.0636
D(REER(-7))	0.224080	0.115516	1.939830	0.0567
D(REER(-32))	-0.200516	0.095571	-2.098081	0.0398
D(REER(-6))	-0.196756	0.113135	-1.739131	0.0867

R-squared	0.166837	Mean dependent var	0.332909
Adjusted R-squared	0.128383	S.D. dependent var	1.663986
S.E. of regression	1.555371	Akaike info criterion	3.777528
Sum squared resid	157.2466	Schwarz criterion	3.907041
Log likelihood	-126.3247	Hannan-Quinn criter.	3.828910
F-statistic	4.338651	Durbin-Watson stat	2.061120
Prob(F-statistic)	0.007544		

Рис. 1.20. Результати оцінювання AR-складової моделі ARMA/ARIMA з мінімальним значенням Акайк критерію

Після визначення оптимальної AR-складової необхідно утворити ряд залишків цієї моделі для наступного використання при визначенні оптимального порядку MA-складової моделі ARMA/ARIMA. Для цього у вікні регресії послідовною серією команд **Proc>Make residual series** створюємо ряд залишків з ім'ям **RESIDS_TMP**.

Отже, після того, як визначено оптимальний порядок (лаг) AR-складової, необхідно оптимізувати лаги MA-частини. У рівнянні з

оптимальною комбінацією лагів авторегресії слід виділити ряд залишків (для нашого прикладу – RESIDS_TMP).

Процедура пошуку оптимального порядку МА-складової моделі ARMA/ARIMA аналогічна процедурі визначення оптимального порядку AR-складової. Однак на цьому етапі до визначеної AR-складової поступово додають МА(1), МА(2)... МА(q)-складові і розраховують значення Шварц критерію. Модель, яка має найменше значення Шварц критерію, – претендент для подальшого аналізу.

Зауважимо, що фактично оцінюється послідовність моделей загального вигляду: *оптимальна AR-складова + лагові значення залишків, попередньо отримані для цієї складової*, наприклад: D(REER) C D(REER(-6)) D(REER(-7)) D(REER(-32) RESIDS_TMP (-1) (специфікація надається в форматі E.Views).

Мінімальне значення критерію Шварца, яке дорівнює 3,312357838, досягається для такої загальної специфікації моделі в форматі E.Views:

```
d(reer) c d(reer(-7)) d(reer(-6)) d(reer(-32)) resid_
tmp(-5) resid_tmp(-7) resid_tmp(-32).
```

Результати розрахунків за цією моделлю на реальних даних наведено на рис. 1.21. Специфікація інших послідовно оцінених моделей з відповідними значеннями Шварц критерію наведено в Додатку 1_A (таблиця 2А).

1.5. Остаточне оцінювання ARIMA моделі та перевірка її на адекватність

1.5.1. Як переоцінити і перевірити модель на адекватність?

Попереднє оцінювання всіх описаних вище моделей для визначення оптимального порядку AR- та МА-складових здійснювалося за допомогою методу найменших квадратів. Однак застосування цього методу оцінювання є некоректним у випадку ARIMA моделювання, оскільки він може давати зміщені оцінки невідомих параметрів (оцінки не будуть BLUE). Тому після того, як визначено

попередню специфікацію (форму) моделі, її потрібно переоцінити нелінійним методом найменших квадратів (NLS)!!!

Equation: OPTIMAL_ARMA Workfile: ARIMA2::Untitled\				
View Proc Object Print Name Freeze Estimate Forecast Stats Resids				
Dependent Variable: D(REER)				
Method: Least Squares				
Date: 09/24/09 Time: 21:22				
Sample (adjusted): 2005M06 2008M06				
Included observations: 37 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.569747	0.219863	2.591377	0.0146
D(REER(-7))	-0.491165	0.271860	-1.806682	0.0809
D(REER(-32))	-0.094682	0.272522	-0.347429	0.7307
D(REER(-6))	-0.324923	0.110209	-2.948248	0.0061
RESIDS_TMP(-7)	0.634031	0.273346	2.319522	0.0274
RESIDS_TMP(-5)	0.379368	0.162527	2.334184	0.0265
RESIDS_TMP(-32)	-0.127658	0.289361	-0.441174	0.6623
R-squared	0.424432	мінімізований критерій Шварца		
Adjusted R-squared	0.309319	Mean dependent var	0.273514	
S.E. of regression	1.000527	S.D. dependent var	1.203299	
Sum squared resid	30.03166	Akaike info criterion	3.007590	
Log likelihood	-48.64041	Schwarz criterion	3.312358	
F-statistic	3.687074	Hannan-Quinn criter.	3.115035	
Prob(F-statistic)	0.007305	Durbin-Watson stat	1.889418	

Рис. 1.21. Результати розрахунків ARIMA моделі з мінімальним значенням Шварц критерію

Програма зробить це автоматично, якщо замінити відповідні лагові значення змінних на ARMA терміни: **REER(-1)** на **AR(1)**, а **RESIDS_TMP(-1)** на **MA(1)** тощо.

Для переоцінювання моделі необхідно послідовно обрати в головному меню **Quick / Estimate Equation**, після чого у вікно, що відкриється ввести визначену специфікацію ARIMA моделі в термінах AR та MA, як показано на рис. 1.22, та натиснути **OK**.

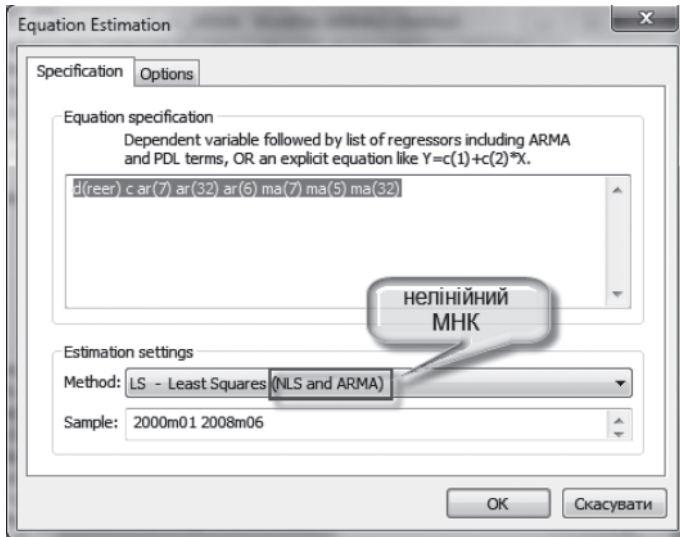


Рис. 1.22. Переоцінювання ARIMA моделі методом нелінійних найменших квадратів (NLS)

Результати відповідних розрахунків на реальній інформації наведено в таблиці 1.2.

Для удосконалення моделі можна спробувати врахувати ефект сезонності, тобто додати компонент сезонності з лагом 12, оскільки в нашому прикладі ми використовуємо місячні дані. У додатку 1_Б наведено приклад сезонного перетворення для найпростішої ARMA(2,2) моделі.

Перевіримо, наприклад, ефект від введення у специфіковану раніше модель сезонного МА-компонента 12-го порядку, відповідно оцінивши модель такого загального вигляду в форматі E.Views:

```
d(reer) c ar(7) ar(32) ar(6) ma(7) ma(5) ma(32) sma(12).
```

Нагадаємо, що ми додаємо до вже оціненого рівняння сезонний МА-компонент 12-го порядку (з урахуванням дванадцяти місяців року в щомісячних даних).

Таблиця 1.2. Результати переоцінювання ARIMA моделі методом ННМК

Dependent Variable: D(REER)				
Method: Least Squares				
Sample (adjusted): 2002M10 2008M06				
Included observations: 69 after adjustments				
Failure to improve SSR after 14 iterations				
MA Backcast: 2000M02 2002M09				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.314449	0.154533	2.034832	0.0461
AR(7)	0.257061	0.147646	1.741064	0.0866
AR(32)	-0.278988	0.094719	-2.945433	0.0045
AR(6)	-0.148600	0.107588	-1.381200	0.1722
MA(7)	-0.356392	0.168457	-2.115628	0.0384
MA(5)	0.111018	0.103226	1.075480	0.2863
MA(32)	0.532527	0.122453	4.348825	0.0001
R-squared	0.457453	Mean dependent var		0.352609
Adjusted R-squared	0.404948	S.D. dependent var		1.665986
S.E. of regression	1.285135	Akaike info criterion		3.435530
Sum squared resid	102.3974	Schwarz criterion		3.662179
Log likelihood	-111.5258	Hannan-Quinn criter.		3.525449
F-statistic	8.712629	Durbin-Watson stat		2.151351
Prob(F-statistic)	0.000001			
Inverted AR Roots	.96+.09i	.96-.09i	.92+.27i	.92-.27i
	.84-.45i	.84+.45i	.74-.62i	.74+.62i
	.62-.75i	.62+.75i	.46+.84i	.46-.84i
	.28+.91i	.28-.91i	.08+.95i	.08-.95i
	-.10-.97i	-.10+.97i	-.27+.93i	-.27-.93i
	-.44-.84i	-.44+.84i	-.61+.73i	-.61-.73i
	-.75-.60i	-.75+.60i	-.86-.46i	-.86+.46i
	-.92-.30i	-.92+.30i	-.94-.10i	-.94+.10i
Inverted MA Roots	.98+.09i	.98-.09i	.93-.28i	.93+.28i
	.85+.46i	.85-.46i	.75-.63i	.75+.63i
	.63+.77i	.63-.77i	.47+.86i	.47-.86i
	.28+.92i	.28-.92i	.08-.97i	.08+.97i
	-.11+.98i	-.11-.98i	-.28-.96i	-.28+.96i
	-.45+.87i	-.45-.87i	-.61-.75i	-.61+.75i
	-.76-.61i	-.76+.61i	-.88-.46i	-.88+.46i
	-.94-.29i	-.94+.29i	-.97-.10i	-.97+.10i

При цьому форма рівняння змінюється з

$$(1 - \rho_6 L^6 - \rho_7 L^7 - \rho_{32} L^{32}) REER_t = (1 + \theta_5 L^5 + \theta_7 L^7 + \theta_{32} L^{32}) e_t \quad (1.3)$$

на

$$(1 - \rho_6 L^6 - \rho_7 L^7 - \rho_{32} L^{32}) REER_t = (1 + \theta_5 L^5 + \theta_7 L^7 + \theta_{32} L^{32}) (1 + \omega L^{12}) e_t. \quad (1.4)$$

Якщо розкрити дужки і виконати перетворення, то модель з сезонним МА-компонентом буде мати такий фінальний вигляд:

$$REER_t = \rho_6 REER_{t-6} + \rho_7 REER_{t-7} + \rho_{32} REER_{t-32} + e_t + \theta_5 e_{t-5} + \theta_7 e_{t-7} + \theta_{32} e_{t-32} + \omega \theta_5 e_{t-5-12} + \omega \theta_7 e_{t-7-12} + \omega \theta_{32} e_{t-32-12}. \quad (1.5)$$

Оцінивши модель (1.5) з сезонним компонентом (SARIMA), отримуємо покращення значень AIC і Шварц критерію, а також коефіцієнта детермінації (див. таблицю 1.3).

Таблиця 1.3. Результати розрахунків за SARIMA моделлю (1.5)

Dependent Variable: D(REER)				
Method: Least Squares				
Sample (adjusted): 2002M10 2008M06				
Included observations: 69 after adjustments				
Failure to improve SSR after 14 iterations				
MA Backcast: 1999M02 2002M09				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.342175	0.173819	1.968575	0.0536
AR(7)	0.256469	0.145861	1.758308	0.0837
AR(32)	-0.259085	0.092480	-2.801527	0.0068
AR(6)	-0.103992	0.117264	-0.886821	0.3787
MA(7)	-0.364066	0.157180	-2.316242	0.0239
MA(5)	0.027431	0.106155	0.258410	0.7970
MA(32)	0.575102	0.110001	5.228169	0.0000
SMA(12)	0.175916	0.144354	1.218648	0.2277
R-squared	0.504178	Mean dependent var		0.352609
Adjusted R-squared	0.447280	S.D. dependent var		1.665986
S.E. of regression	1.238579	Akaike info criterion		3.374458

Sum squared resid	93.57873	Schwarz criterion		3.633485
Log likelihood	-108.4188	Hannan-Quinn criter.		3.477222
F-statistic	8.861147	Durbin-Watson stat.		2.168432
Prob(F-statistic)	0.000000			
Inverted AR Roots	.96-.09i	.96+.09i	.91+.27i	.91-.27i
	.84+.45i	.84-.45i	.74-.62i	.74+.62i
	.62+.75i	.62-.75i	.46-.84i	.46+.84i
	.28+.91i	.28-.91i	.08-.95i	.08+.95i
	-.10+.96i	-.10-.96i	-.27+.93i	-.27-.93i
	-.44+.84i	-.44-.84i	-.60-.73i	-.60+.73i
	-.75-.60i	-.75+.60i	-.86+.46i	-.86-.46i
	-.92+.29i	-.92-.29i	-.94+.10i	-.94-.10i
Inverted MA Roots	.99-.09i	.99+.09i	.94-.27i	.94+.27i
	.86-.46i	.86+.46i	.84-.22i	.84+.22i
	.75+.63i	.75-.63i	.63-.77i	.63+.77i
	.61-.61i	.61+.61i	.48+.86i	.48-.86i
	.29+.93i	.29-.93i	.22+.84i	.22-.84i
	.09+.97i	.09-.97i	-.11-.98i	-.11+.98i
	-.22-.84i	-.22+.84i	-.28+.96i	-.28-.96i
	-.45+.87i	-.45-.87i	-.61+.61i	-.61-.61i
	-.62+.75i	-.62-.75i	-.76+.61i	-.76-.61i
	-.84+.22i	-.84-.22i	-.88+.46i	-.88-.46i
	-.94-.30i	-.94+.30i	-.97-.10i	-.97+.10i

Для того, щоб модель (1.5) вважалась адекватною, достатньо виконання однієї умови – залишки оціненого фінального рівняння мають бути білим шумом. Для перевірки залишків на білий шум, можна використати корелограми і Дікі–Фулер тест.

Для цього виділимо залишки нашої регресії в окремий ряд (**Proc>Make residual series**) і відтворимо його корелограм (див. рис. 1.23).

Як можна побачити з графіка автокорелограма, всі коефіцієнти автокореляції статистично незначущі, отже попередньо можна зробити висновок про стаціонарність залишків.

Sample: 2000M01 2008M06
Included observations: 69

























































Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 -0.090	-0.090	0.5812	0.44E
		2 -0.146	-0.156	2.1448	0.342
		3 0.020	-0.009	2.1749	0.537
		4 -0.056	-0.081	2.4140	0.66C
		5 -0.007	-0.020	2.4177	0.78E
		6 -0.055	-0.082	2.6560	0.851
		7 0.011	-0.007	2.6660	0.914
		8 0.090	0.067	3.3199	0.91E
		9 -0.128	-0.119	4.6678	0.862
		10 -0.048	-0.061	4.8599	0.90C
		11 -0.023	-0.080	4.9036	0.93E
		12 0.031	0.009	4.9861	0.95E
		13 -0.015	-0.046	5.0049	0.97E
		14 -0.081	-0.092	5.5829	0.97E
		15 0.153	0.109	7.7198	0.93E
		16 0.136	0.137	9.4367	0.894
		17 -0.092	-0.014	10.231	0.894
		18 -0.074	-0.069	10.751	0.90E
		19 0.130	0.122	12.410	0.867
		20 -0.076	-0.077	12.992	0.87E
		21 -0.038	-0.009	13.143	0.904
		22 0.095	0.086	14.074	0.89E
		23 0.119	0.126	15.576	0.87E
		24 -0.030	0.015	15.671	0.90C
		25 -0.175	-0.107	19.079	0.79E
		26 0.012	0.026	19.096	0.832
		27 0.008	-0.045	19.103	0.86E
		28 -0.033	-0.012	19.233	0.891

Рис. 1.23. Корелограм залишків моделі (1.5)

Перевіримо залишки на білий шум також на основі формального тесту Дікі–Фулера, результати якого наведено на рис. 1.24.

Тест Дікі–Фулера також показує відсутність одиничного кореня, тобто залишки стаціонарні. Таким чином, можемо стверджувати, що залишки оціненої моделі (1.5) – це білий шум, відповідно, модель слід вважати адекватною.

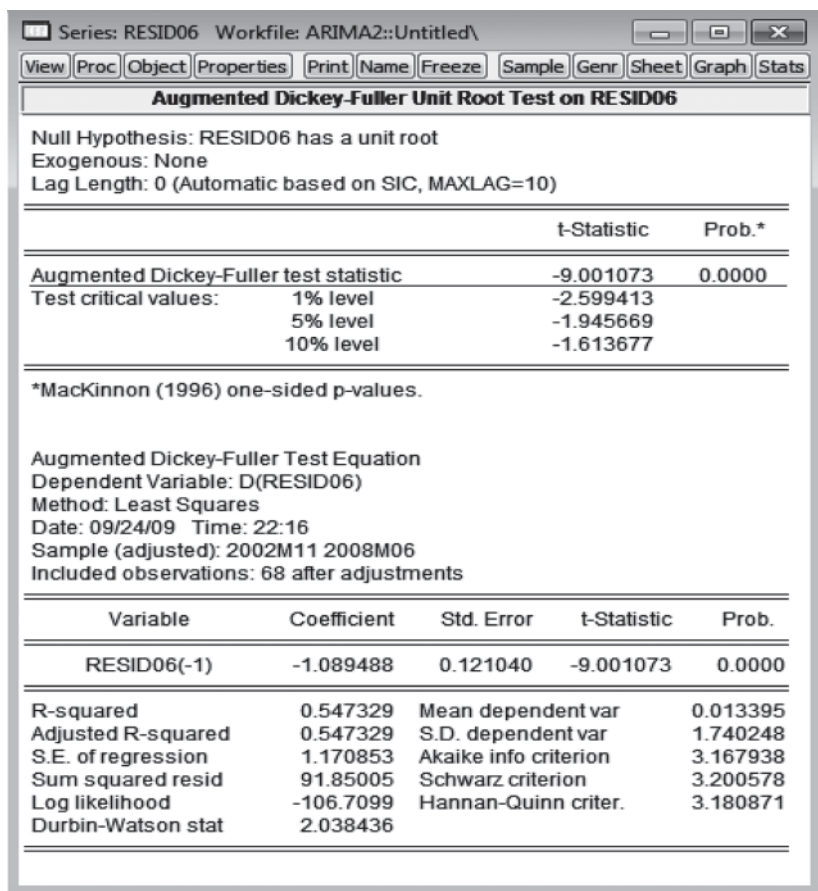


Рис. 1.24. Результати перевірки залишків моделі (1.5) на білий шум

Після перевірки моделі на адекватність можна проаналізувати її властивості і поведінку. Для цього використовують команду **View>ARMA Structure**. За нею відкривається вікно, в якому можна обрати різні опції, зокрема, аналіз моделі на стаціонарність та обертовність корелограма, а також відобразити імпульсну функцію відгуків (див. рис. 1.24).

Опція **Roots** дає можливість проаналізувати стаціонарність і обертовність ARMA моделі.

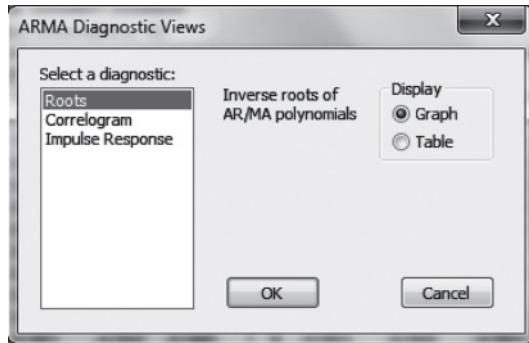


Рис. 1.24. Вікно діагностики фінальної ARIMA моделі

Інформацію можна оформити як графік або таблицю. Якщо на графіку всі корені AR розташовані в межах одиничного кола, модель стаціонарна. Якщо всі МА-корені розміщені в межах одиничного кола, то модель обертовна. Рис. 1.25 і таблиця 1.4 ілюструють властивості стаціонарності та обертовності моделі.

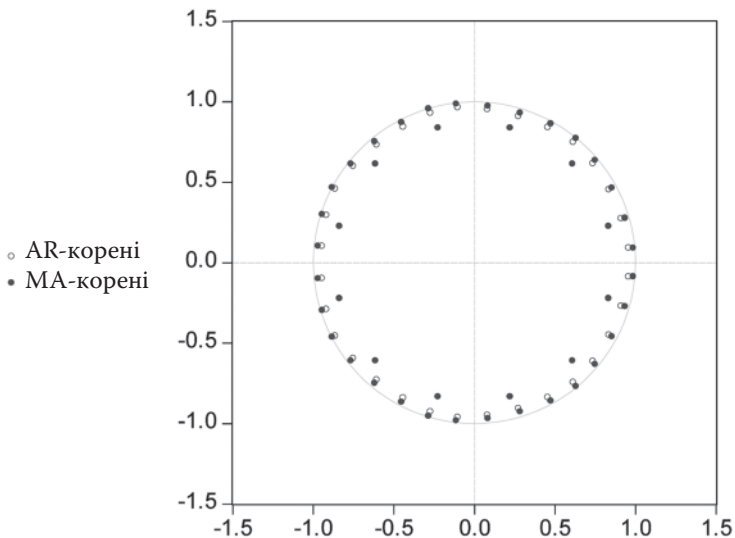


Рис. 1.25. Графік значень коренів характеристичного полінома

Таблиця 1.4. Результати діагностики моделі (1.5) на стаціонарність та обертовність

Inverse Roots of AR/MA Polynomial(s)		
Specification: D(REER) C AR(7) AR(32) AR(6) MA(7)		
MA(5) MA(32) SMA(12)		
Date: 09/24/09 Time: 22:21		
Sample: 2000M01 2008M06		
Included observations: 69		
AR Root(s)	Modulus	Cycle
-0.860976 ± 0.456851i	0.974676	2.367663
-0.102008 ± 0.963902i	0.969284	3.748398
0.615791 ± 0.745984i	0.967311	7.134201
-0.270685 ± 0.928095i	0.966763	3.387927
0.959148 ± 0.089738i	0.963337	67.35187
-0.916992 ± 0.293211i	0.962729	2.218550
0.738590 ± 0.615696i	0.961560	9.041836
-0.749217 ± 0.597130i	0.958066	2.545160
0.458897 ± 0.839808i	0.957008	5.868311
0.084100 ± 0.950923i	0.954634	4.237993
0.914812 ± 0.272507i	0.954537	21.70260
0.838807 ± 0.452176i	0.952922	12.70836
-0.440319 ± 0.843311i	0.951343	3.061994
0.276442 ± 0.908543i	0.949668	4.926338
-0.942718 ± 0.099903i	0.947997	2.069551
-0.603674 ± 0.730213i	0.947435	2.778189
No root lies outside the unit circle.		
ARMA model is stationary.		
MA Root(s)	Modulus	Cycle
0.632470 ± 0.771345i	0.997493	7.107640
-0.281988 ± 0.955330i	0.996079	3.382023
-0.880338 ± 0.464497i	0.995365	2.365583
0.988042 ± 0.088285i	0.991978	70.50489
-0.109448 ± 0.984681i	0.990745	3.736670
-0.942254 ± 0.297744i	0.988177	2.215877
0.475780 ± 0.861787i	0.984400	5.892216
0.752215 ± 0.633700i	0.983567	8.974805

Продовження табл. 1.4

$-0.764496 \pm 0.612361i$	0.979510	2.547675
$-0.450741 \pm 0.868487i$	0.978487	3.065680
$0.936912 \pm 0.274752i$	0.976367	22.02656
$0.086603 \pm 0.970398i$	0.974254	4.240275
$-0.968871 \pm 0.101857i$	0.974211	2.068982
$0.856029 \pm 0.461454i$	0.972484	12.70851
$0.286600 \pm 0.928645i$	0.971865	4.941752
$-0.616514 \pm 0.750753i$	0.971452	2.782229
$-0.611778 \pm 0.611778i$	0.865185	2.666667
$-0.835704 \pm 0.223926i$	0.865185	2.181818
$0.835704 \pm 0.223926i$	0.865185	24.00000
$0.223926 \pm 0.835704i$	0.865185	4.800000
$0.611778 \pm 0.611778i$	0.865185	8.000000
$-0.223926 \pm 0.835704i$	0.865185	3.428571
No root lies outside the unit circle.		
ARMA model is invertible.		

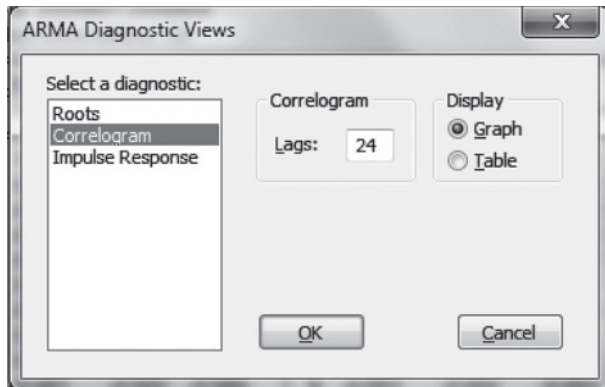


Рис. 1.26. Вікно діагностики фінальної ARIMA моделі

Опція **Корелограм** дозволяє порівняти коефіцієнти автокореляції та часткової автокореляції (**ACF/PACF**) фактичних значень часового ряду (**Actual**, сині піки) з відповідними коефіцієнтами розрахованих (теоретичних) значень часового ряду (**Theoretical**, горизонтальна лінія) (див. рис. 1.27).

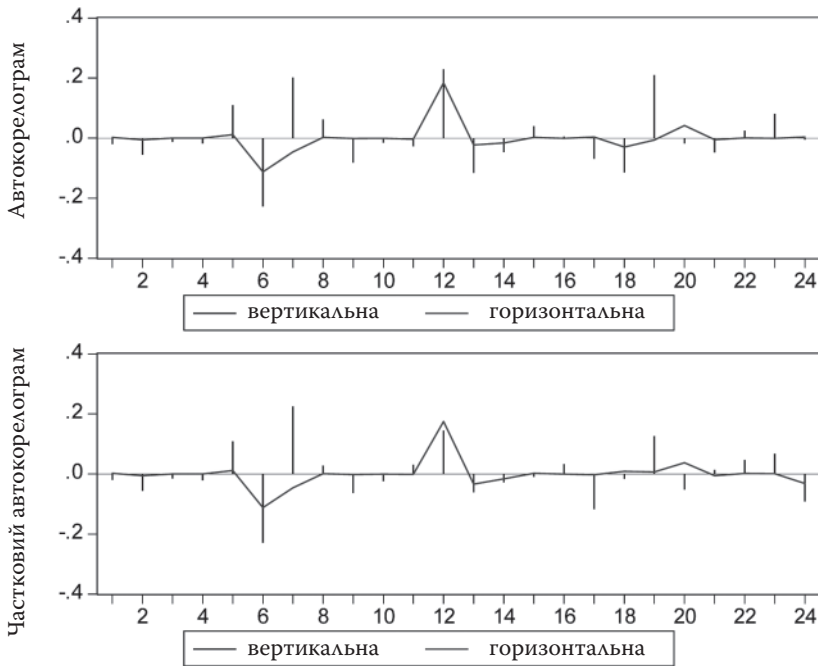


Рис. 1.27. Графіки ACF/PACF фактичних і теоретичних (розрахункових) рівнів часового ряду (DREER)

Як можна побачити з аналізу графіка, частина піків не відтворюється оціненою моделлю (нагадаємо, що R^2 становить близько 40 %).

Підопція **Impulse response** дає змогу вивчати динамічну зміну досліджуваного показника у відповідь на шок в межах оціненої моделі (див. рис. 1.28). Після перевірки моделі на адекватність та додаткової діагностики необхідно перевірити її прогностну якість й обчислювати прогностні значення ряду.

1.6. Прогнозування на основі побудованої ARIMA моделі

1.6.1. Розрахунок прогностних значень на основі побудованої моделі

Для розрахунку прогностних значень необхідно у вікні з оціненим рівнянням обрати опцію **Forecast**, яка активує вікно, відображене на рис. 1.29.

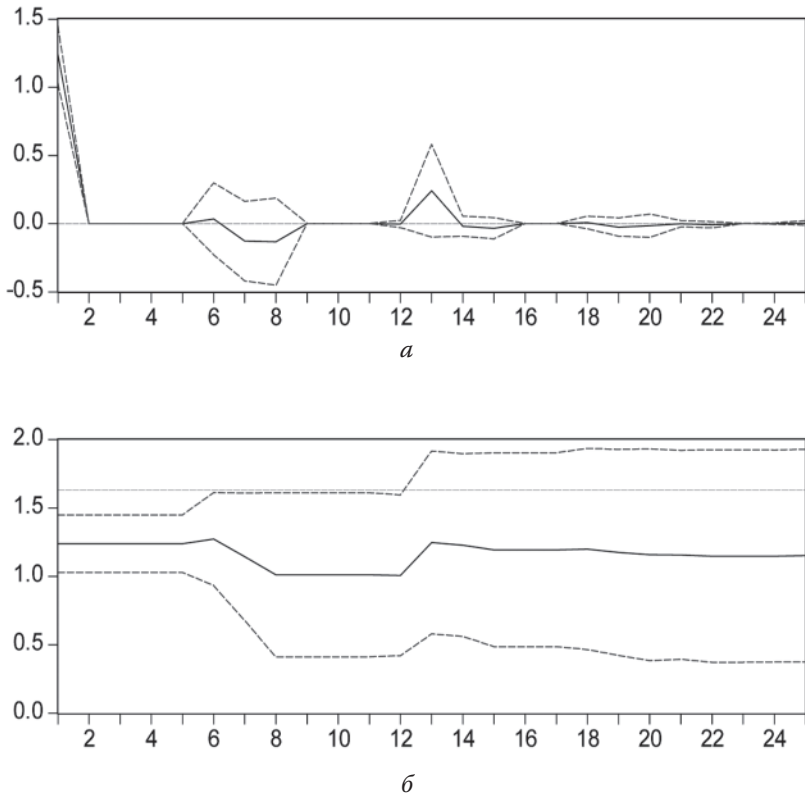


Рис. 1.28. Імпульсна ± 2 S.E. (а) та кумулятивна ± 2 S.E. (б) функції відгуків досліджуваного показника на зовнішні шоки

Зауважимо, що перед розрахунком прогнозних значень у підопції **Forecast name** необхідно задати ім'я ряду, в якому вони будуть зберігатися. За замовчуванням до імені фактичних значень часового ряду додається розширення «f» (у нашому випадку прогнозні дані будуть зберігатися в файлі з ім'ям «reerf»).

У підопції **Forecast sample** задається період для визначення прогнозу.

Рис. 1.29. Вікно для розрахунку прогнозу з відповідними опціями

Для графічного відображення потрібно виділити також **Forecast graph** у підопції **Output**. Натиснувши **OK**, отримаємо результати прогнозних розрахунків, значення основних критеріїв прогнозової якості, а також графічне відображення прогнозних значень та 95 % інтервалів довіри (див. рис. 1.30).

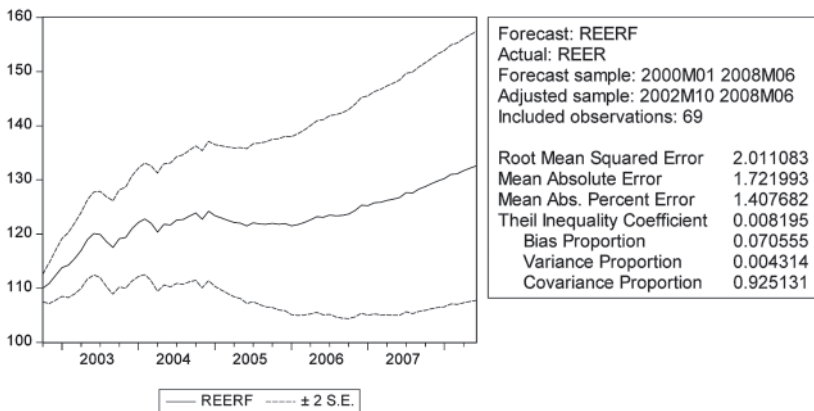


Рис. 1.30. Результати розрахунків прогнозу на основі оціненої моделі (1.5)

Зауважимо, що спочатку можна оцінити прогностну якість моделі протягом певного періоду, на якому її було оцінено. Для цього просто не слід змінювати значення вибірки у вікні **Forecast sample** (див. рис. 1.29).

Для оцінювання прогностної якості моделі необхідно проаналізувати значення розрахованих критеріїв прогностної якості (див. рис. 1.30). Одним із найпоширеніших на практиці є критерій середньої абсолютної процентної похибки (MAPE – Mean Absolute Percent Error), який показує середню абсолютну похибку прогнозу в відсотках. Для нашого прикладу значення MAPE = 1,4 %, що свідчить про високу прогностну якість моделі. Детальніше інформацію про критерії прогностної якості моделей наведено в додатку 1_В та пам'ятці 6.

1.6.2. Як побудувати інтервали довіри для прогнозу?

Для того, щоб отримати числові значення інтервалів довіри для прогностних значень, необхідно перш за все зберегти значення середнього квадратичного відхилення прогнозу (кореня з його дисперсії). Для цього у вікні **Forecast** (див. рис. 1.30) в підопції **S.E. (Optional)** слід вписати назву нового ряду, в якому будуть зберігатися значення середнього квадратичного відхилення, наприклад, **reerf_se** (і програма створить цей ряд, вписавши туди стандартні похибки прогнозу).

Щоб відобразити графік з інтервалами довіри для всього прогностного періоду, потрібно спочатку розрахувати їхні значення.

При цьому на першому етапі в робочому файлі треба виділити ряди **reer**, **reerf**, **reerf_se**, натиснувши клавішу **Ctrl**, і двічі клацнути на виділеному місці (див. рис. 1.31).

Потім слід вибрати в меню, що з'явиться, **Open Group**; натиснувши опцію **Edit+/-** у таблиці з даними, ввімкнути редагування; назву ряду **reerf_se** змінити на **reerf+2*reerf_se**. Поряд у порожній колонці замість назви написати **reerf-2*reerf_se** (програма автоматично обчислює формулу, прописану в колонці!!!) (див. рис. 1.32).

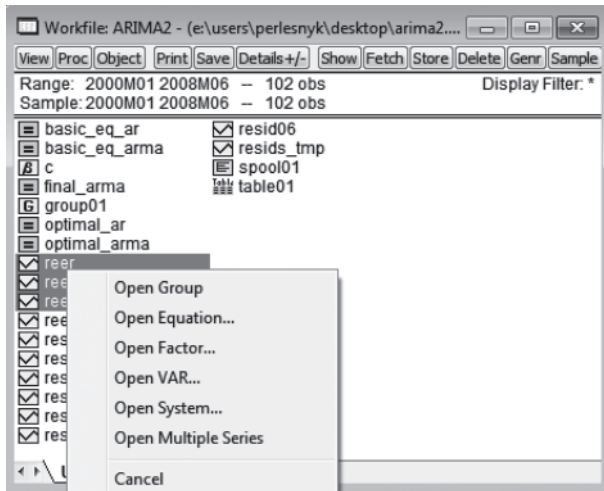


Рис. 1.31. Вікно для порівняння фактичних і розрахункових значень

obs	REER	REERF	REERF-2*REERF SE
2000M01	106.370	106.3700	NA
2000M02	104.640	104.6400	NA
2000M03	102.280	102.2800	NA
2000M04	100.810	100.8100	NA
2000M05	98.5100	98.51000	NA
2000M06	102.280	102.2800	NA
2000M07	101.330	101.3300	NA
2000M08	97.9400	97.94000	NA
2000M09	96.2000	96.20000	NA
2000M10	94.9500	94.95000	NA
2000M11	95.5600	95.56000	NA
2000M12	99.1400	99.14000	NA
2001M01	102.750	102.7500	NA
2001M02	101.810	101.8100	NA
2001M03	102.130	102.1300	NA
2001M04	101.490	101.4900	NA
2001M05			

Рис. 1.32. Введення формул для обчислення інтервалів довіри

На другому етапі в меню вікна, відображеного на рис. 1.32 послідовно обрати **View>Graph** і **Line&Symbol**. Отримаємо графік прогнозних значень з 95 % інтервалів довіри (див. рис. 1.33).

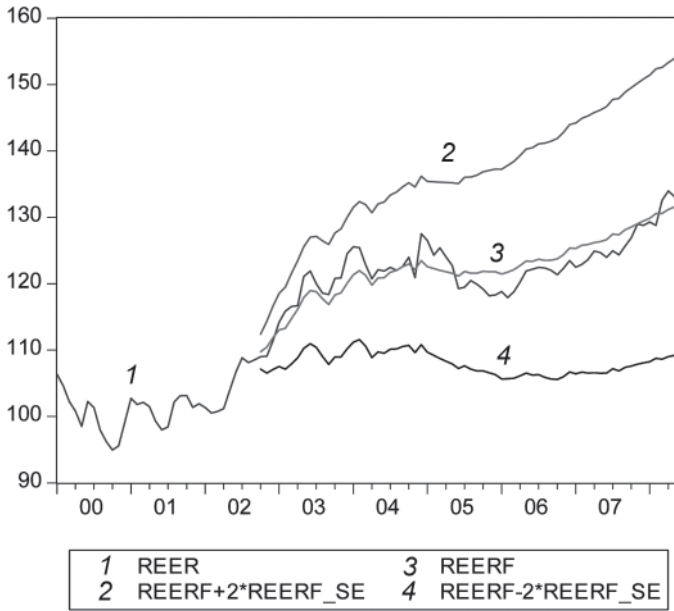


Рис. 1.33. Графічне відображення прогностичних значень та 95 % інтервалу довіри

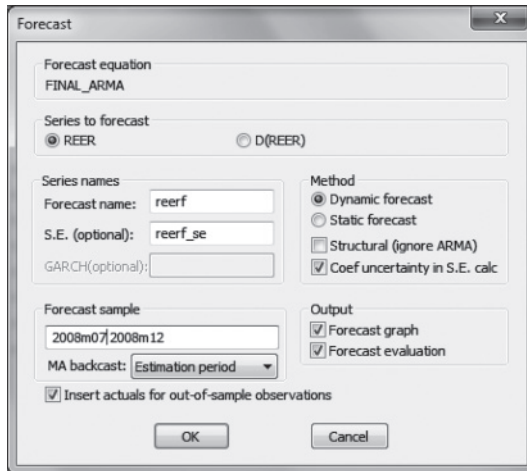


Рис. 1.34. Вікно Forecast зі зміненими підопціями для отримання прогнозу на наступні 6 місяців за фінальною ARIMA моделлю

Щоб зберегти графік або вигляд будь-якого іншого об'єкта, можна використовувати команду **Freeze** вікна об'єкту (див. рис. 1.32).

Для отримання прогнозу поза межами оціненого періоду, треба повторити аналогічні дії, попередньо розширивши у вікні **Forecast** (рис. 1.29) вибірку (**Sample size**), наприклад, на наступні 6 місяців (команда **Proc>Structure/Resize Current Page** головного меню при виділеному вікні робочого файлу), і для більшої точності динамічного прогнозу зсунувши початок прогнозного інтервалу, як показано на рис. 1.34.

Якщо не закривати групу з відображенням графіка, він автоматично оновиться (див. рис. 1.35).

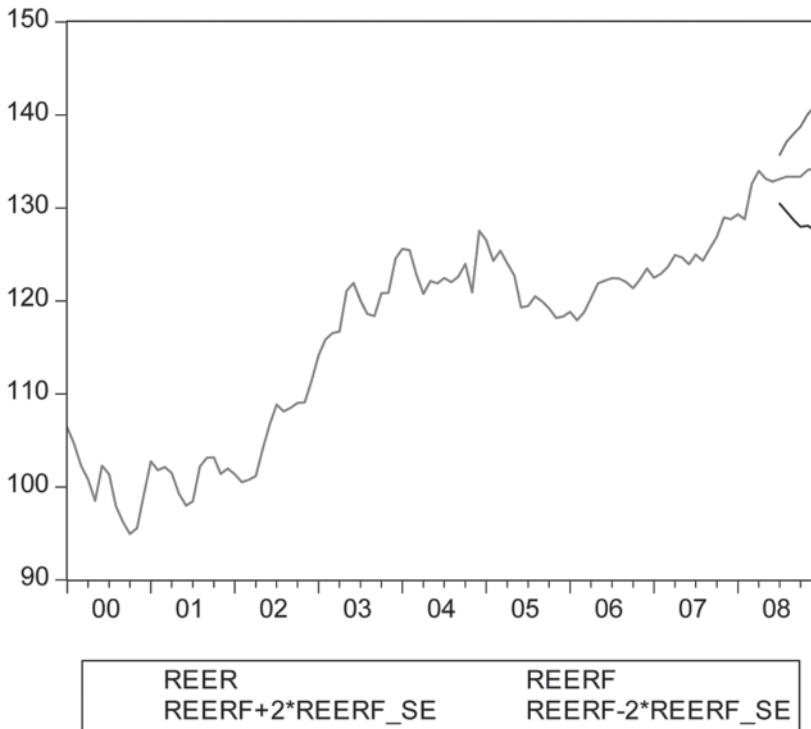


Рис. 1.35. Графік прогнозних значень та інтервали довіри для прогнозу на шість наступних періодів



ЗАВДАННЯ, ТВОРЧІ РОБОТИ, СИТУАЦІЙНІ ВПРАВИ ДО ТЕМИ 1

Завдання для самоперевірки 1. Самостійна побудова та аналіз ARIMA моделі для часового ряду показника реального ефективного обмінного курсу євро для країн Євросоюзу (РЕОК).

Використовуючи значення **показника реального ефективного обмінного курсу євро для країн Євросоюзу (РЕОК)**, наведені в додатку 1_E й таблиці 1Е (часовий ряд REER), та надані інструкції, повторіть самостійно всі етапи побудови ARIMA моделі для цього часового ряду.

Скористайтеся інструкцією:

- а) створіть робочий файл у E.Views та імпортуйте або скопіюйте в нього дані;
- б) здійсніть первинний статистичний аналіз даних;
- в) перевірте часовий ряд на стаціонарність. Визначте порядок інтеграції;
- г) специфікуйте (ідентифікуйте) ARIMA модель для досліджуваного часового ряду, визначивши кількість лагів у моделі та порядок AR- та MA-складових;
- г) перевірте побудовану ARIMA модель на адекватність;
- д) за побудованою моделлю розрахуй прогноз на шість наступних періодів. Побудуйте довірчі інтервали прогнозу та проаналізуйте прогнозу якість розробленої ARIMA моделі;
- е) зробіть висновки.

Дайте відповіді на такі запитання:

1. Які основні труднощі виникли під час виконання цього завдання?
2. Чи допоміг аналіз графіків автокорелограма та часткового автокорелограма (ACF та PACF) щодо визначення кількості лагів для первинного включення в модель?

3. Чи всі етапи виконання процедури Хенона–Рісанена були зрозумілими?
4. За якими основними процедурами можна протестувати залишки моделі на білий шум?
5. Чому в процесі попереднього аналізу вибору відповідної специфікації ARIMA моделі використовується метод найменших квадратів, а переоцінювання фінальної моделі здійснюється нелінійним методом найменших квадратів?
6. Які основні відмінності в побудові моделей для чистих процесів і змішаних процесів?
7. Що таке порядок інтеграції часового ряду (моделі), для чого його необхідно визначати?
8. У чому, на Ваш погляд, полягає проблема оцінки прогнозової якості ARIMA моделей?
9. Які основні труднощі виникли під час обчислення інтервалів довіри для прогнозу? Для якого періоду упередження Ви б порекомендували використовувати оцінену Вами ARIMA модель?
10. Які основні висновки Ви можете зробити, підсумовуючи виконане завдання?
11. Спробуйте оновити надані Вам реальні дані часового ряду показника реального ефективного обмінного курсу євро для країн Євросоюзу (РЕОК) за останні роки та здійснити нові обчислення. Чи збігаються специфікації оновленої моделі та попередньо отриманої Вами? У чому полягає різниця між ними (якщо вона є)? Обґрунтуйте відповідь.

Творча робота 1. Побудова та аналіз ARIMA моделі за допомогою пакета E.Views 6.0.

Завдання. Використовуючи надані інструкції з ARIMA моделювання, матеріали підручника І. Г. Лук'яненко, Ю. О. Городніченка «Сучасні економетричні методи в фінансах» (Літера, 2003, С. 56–69), а також інші дидактичні матеріали, виконайте творчу роботу. Для цього:

- а) знайдіть статистику (бажано місячну (квартальну)), включно до останнього місяця поточного року) фінансових або інших економічних показників (див. сайт НБУ: <http://>

bank.gov.ua; офіційний сайт Комітету Верховної Ради України з питань бюджету: <http://budget.rada.gov.ua>; офіційний сайт Державного комітету статистики України: <http://www.ukrstat.gov.ua>;

- б) побудуйте ARIMA модель для обраного часового ряду, попередньо дослідивши його на стаціонарність. Обґрунтуйте адекватність побудованої моделі;
- в) за побудованою моделлю складіть прогноз на три наступні періоди. Розрахуйте критерії прогнозовної якості та побудуйте 96 % інтервалів довіри. Проаналізуйте прогнозну якість побудованої ARIMA моделі;
- г) обґрунтуйте важливість та можливість застосування ARIMA інструментарію загалом і для прогнозування обраного вами показника. Які труднощі виникли під час виконання самостійної творчої роботи?
- г) підготуйте презентацію для захисту творчої роботи.

При виконанні творчої роботи в пакеті E.Views мають бути відображені послідовно такі етапи: підготовка та аналіз даних; аналіз часового ряду; формулювання, оцінювання й застосування моделі; висновки.

Етап перший. Підготовка та аналіз даних

Відкрийте файл(и) з даними (*.xls) для обраного Вами показника і проаналізуйте: кількість спостережень, періодичність даних (місяць, рік, квартал...), одиницю виміру.

Створіть робочий файл E.Views із визначеними параметрами розміру і періодичності. Як ім'я файлу вкажіть своє прізвище. Імпортуйте дані з файла *.xls будь-яким із доступних способів.

Очікуваний результат: створено робочий файл E.Views, у якому є об'єкти «series» з даними. Спостереження мають бути індексовані за датою спостереження.

Етап другий. Аналіз часового ряду

Виведіть графіки динаміки показників. Проаналізуйте зовнішній вигляд графіка: наявність тренду, циклу, сезонності, наявність нульових значень.

Виведіть на графікові гістограму рядів, поясніть, що означає їхня форма.

Проаналізуйте ряди на стаціонарність, використовуючи ADF тест. Опишіть нульову й альтернативну гіпотези, випадки, коли вони приймаються або відкидаються.

Визначте порядок інтеграції (пригадайте, операціями яких різниць ряд зводиться до стаціонарного).

Проаналізуйте ACF/PACF стаціонарного процесу. Визначте наявність сезонності та скоригуйте в разі необхідності.

Визначте форму моделі для опису процесу – AR, MA, ARMA, і кількість лагів для перших двох, якщо можливо.

Очікуваний результат: для кожного часового ряду визначено та обґрунтовано порядок інтеграції – форму, в якій він поводить ся як стаціонарний ряд (наприклад, перші чи другі різниці, логарифмування, коригування сезонності тощо). Визначено базові характеристики авторегресійного процесу чи процесу ковзного середнього.

Етап третій. Формулювання, оцінювання й застосування моделі

Якщо процес змішаний, виконайте процедуру Хенона–Рісанена для визначення порядків авторегресійної та МА-складової. Варто подумати, як спростити алгоритм здійснення процедури. Правильних варіантів моделі для кожного досліджуваного показника може бути кілька.

Перевірте моделі на коректність, проаналізувавши залишки, корені лагового полінома та корелограм. У разі потреби спробуйте вдосконалити модель увівши фактори сезонності. Орієнтуйтеся на мінімізацію значення AIC та Schwarz IC.

Проведіть тестування прогностної якості кожної моделі та прогнозування майбутніх значень ряду терміном на півроку. Вкажіть 95 % довірчого інтервалу для прогнозу. Побудуйте відповідний графік.

Етап четвертий. Висновки

Зробіть висновки стосовно проведеної творчої роботи.

Практичний кейс 1. Аналіз дії шоків на прикладі * часового ряду розриву ВВП для випадку простого AR(1) процесу: $Y_t = m + \alpha Y_{t-1} + u_t$.

Пояснимо, як визначено розрив ВВП і його взаємозв'язок із фактичним рівнем ВВП у цьому прикладі.

Позначимо через Y_t^f фактичний рівень ВВП в Україні в гривнях, Y_t^p – потенційний рівень ВВП. Нехай, $Y_t = \ln \left(\frac{Y_t^f}{Y_t^p} \right)$ – логарифмічне відхилення (лог-відхилення) фактичного ВВП від потенційного значення. Часовий ряд Y_t і є ВВП розривом, тобто різницею між фактичним ВВП та його потенційним значенням. У макроекономічних моделях припускається, що ВВП розрив у довгостроковому періоді дорівнює нулю, оскільки макросистема еволюціонує до свого стаціонарного стану, коли фактичний рівень ВВП досягає свого потенційного значення і тільки короткострокові збурення можуть призвести до відхилення фактичного від потенційного значення. Для розуміння того, що характеризує Y_t у цьому випадку, візьмемо похідну з часу:

$$\frac{dY_t}{dt} = \frac{d \left(\ln \left(\frac{Y_t^f}{Y_t^p} \right) \right)}{dt} = \frac{\frac{dY_t^f}{dt}}{Y_t^f} - \frac{\frac{dY_t^p}{dt}}{Y_t^p} \approx \frac{\Delta Y_t^f}{Y_t^f} - \frac{\Delta Y_t^p}{Y_t^p}. \quad (1)$$

З формули (1) випливає, що $(Y_t \times 100(\%))$ – це відхилення темпу росту фактичного ВВП від потенційного у відсотках.

Проілюструємо наведені твердження за допомогою простого прикладу. Нехай потенційний ріст ВВП в Україні за останні 9 років становить 5,7 %. Тоді падіння ВВП, наприклад у 2009 році, на 14,8 % означатиме, що $\left. \frac{dY_t}{dt} \right|_{t=2009} \times 100 = -14,8 \% - 5,7 \% = -20,5 \%$.

В умовах розглянутого кейса, припустимо, що згенеровано величину Y_p , яка дорівнює нулю на досліджуваному часовому проміжку. На основі цього часового ряду ВВП розриву можна дослідити дію різних шоків.

* Усі приклади в цьому завданні гіпотетичні.

Зауважимо, що стабільну ситуацію, коли ВВП не відхилятиметься від свого потенційного значення, може порушити якийсь шок. Наприклад, відбувся неочікувано різкий ріст цін на сталь на зовнішніх ринках, що спричинив позитивний ВВП розрив для України, або масовий відплив капіталу закордонних хедж-фондів з Української біржі (негативний ВВП розрив) тощо.

Шоки бувають очікуваними і неочікуваними. З першими маємо справу, коли населенню заздалегідь відомо про настання шоку. Наприклад, НБУ за умови повної довіри до нього оголосив про перехід до інфляційного таргетування через дев'ять місяців. Також шоки можна умовно поділити на прості, коли шок діє протягом одного періоду, а потім зникає, та складні, коли шок діє кілька періодів.

Припустимо, що розглядається період часу від першого кварталу 1996 року до четвертого кварталу 2010 року, який включає 60 спостережень. Протягом цього періоду аналізується дія трьох можливих типів шоків (випадкових величин, які позначимо відповідно u_t , w_t , ε_t) на зміну часового ряду розриву ВВП – постійного складного; наперед невідомого непостійного; неочікуваного.

Постійний складний шок, який, за припущенням, відображає реакцію ВВП на зміну погодних умов. Динаміка шоку відповідає нормальному процесу $N(0,1)$. Для аналізу дії цього шоку на ВВП необхідно згенерувати спочатку часовий ряд незалежних нормально розподілених випадкових величин з нульовим математичним сподіванням та одиничною дисперсією $N(0,1)$, що складається з 60 спостережень. Загалом, за визначенням, він є білим шумом *.

Припустимо, що аномально сильну весняно-літню посуху 2007 року можна вважати шоком, внаслідок дії якого значення Y_t у другому та третьому кварталах 2007 року знизилися на 0,03 та 0,04 відповідно. Для відображення дії цього шоку в згенерованому ряді випадкових величин u_t спостереження u_{46} та u_{47} (другий і третій квартали 2007 року) необхідно замінити відповідно на $(u_{46} - 0,03)$ та $(u_{47} - 0,04)$, що відображатиме аномально сильну весняно-

* При генерації часових рядів за допомогою команди `grand` слід переконатися, що отримані випадкові значення збігаються для всіх альтернативних сценаріїв. Для цього потрібно встановити початкове значення генератора на певному рівні, наприклад, виконавши команду `rndseed = 33`.

літню посуху 2007 року, яка охопила близько двох третин території України.

Зауважимо, що бажано визначити початкове значення бази генератора випадкових чисел, щоб отримувати кожного разу одні й ті самі випадкові числа. Це спростить порівняння результатів у процесі аналізу. Нехай початкове значення становить для цього випадку 33.

Наперед невідомий непостійний шок, який, за припущенням, відображає реакцію ВВП України на фінансову кризу в Росії – дефолт 17 серпня 1998 року. У цьому випадку такий шок буде представлений рядом w_t , значення якого дорівнюють нулю для всіх спостережень, окрім третього кварталу 1998 року. Припустимо, що ми знаємо про настання негативного неочікуваного шоку, але невідоме його очікуване значення. Отже, для третього кварталу 1998 року необхідно згенерувати негативне випадкове значення $w_{11} < 0$. При цьому оберемо початкове значення, яке, наприклад, дорівнює 7.

Шок неочікуваний. Він наперед невідомий для населення, але про нього знають, наприклад, аналітики інвестиційної компанії. Аналітики інвестбанку, вивчаючи на основі відкритих джерел результати переговорів між двома крупними компаніями, на початку 2002 року дійшли висновку, що в першому кварталі цього року буде підписано угоду про співпрацю, яка передбачатиме здійснення інвестиції в економіку України розміром 800 млн грн. На основі прогнозу аналітики вважають, що таке капіталовкладення призведе до збільшення ВВП у наступному кварталі на 100 млн грн, що означає, що лог-відхилення становитиме 0,0002.

Крім того, також гіпотетично припустимо: завдяки розслідуванню журналістів аналітики дізналися, що представники керівництва певної великої компанії планують реалізувати спеціальну схему з приховання 150 млн дол. Зокрема, на початку 2003 року планується оголосити тендер на постачання обладнання, а переможцем тендера може бути оголошена підставна компанія, яка, закупивши обладнання по ціні 250 млн дол., продасть її певній українській компанії за 400 млн дол. Очікується, що така операція призведе до падіння ВВП у першому кварталі 2003 року на 150 млн грн з відповідним лог-відхиленням $-0,0005$.

Три аналітики – Петро з компанії «Дракон», Василь з компанії «Тисячоліття» та Джон із компанії «Літак» – для оцінювання акцій, зокрема необхідної дохідності (required rate of return), використовують макроекономічну факторну модель (macroeconomic factor model), яка прогнозує дохідність на основі зміни макрпоказників, серед яких і ріст ВВП. Наприклад, вони розробляють моделі наприкінці 1995 року. Кожен із них використовує для аналізу ВВП авторегресійну модель першого порядку і припускає, що в першому кварталі фактичний ВВП дорівнює потенційному. Але Петро, Василь і Джон мають різні погляди щодо еволюції лог-відхилення ВВП. Зокрема, ріст лог-відхилення ВВП на одиницю призводить до росту чи падіння ВВП в наступному періоді, на думку Петра, на 0,7, Василя – на 1,2, Джона – на 1,0.

Треба проаналізувати, яка з трьох запропонованих аналітиками AR(1) моделей краще відтворюватиме дію описаних шоків. Варто також звернути увагу на особливості затухання ефектів від різних шоків.

Для проведення аналізу потрібно побудувати та вивести графіки таких часових рядів:

- 1) перший графік – три шоки;
 - 2) наступні три графіки – ВВП розриви для кожного шоку.
- При цьому на кожному графіку слід відобразити три ВВП розриви кожного аналітика, які реагують на один із шоків.

Треба описати шоки та лог-відхилення, особливість реакції лог-відхилення на дію шоків; побудувати функції автокореляції шоків та лог-відхилень; протестувати шоки та лог-відхилення на стаціонарність на основі графічного аналізу, функцій автокореляції та тесту Дікі–Фулера.

Яка модель, на Вашу думку, найкраще описує реакцію ВВП України на дію шоків? Чому? Вкажіть позитивні та негативні сторони моделей.

Пам'ятка. Необхідні команди для виконання завдань кейса

У E.Views більшість операцій можна виконати як через засоби графічного інтерфейсу, так і за допомогою введення текстових команд у командному вікні. Останній спосіб зручний, тому що у ви-

падку багаторазового виконання операцій можна повторно використувати команди, змінюючи деякі їхні параметри.

1. Генерування часових рядів за формулою: *series назва_змінної = формула*. Наприклад:

- а) *series y_p_1 = 0* створює змінну (або ж перезаписує значення наявної змінної *y_p_1*) зі значенням усіх спостережень 0;
- б) *series y_p_2 = nrnd* створює змінну (або ж перезаписує значення наявної змінної *y_p_2*) зі значенням усіх спостережень, що відповідають випадковим незалежним значенням нормально розподіленої випадкової величини $N(0,1)$. Відповідно *series y_p_3 = .3 + .5*nrnd* – значення $N(0.3, 0.5)$;
- в) *series y_p_4 = .7* y_p_4(-1) + u* – простий AR(1) процес із рядом збурень, збереженням у змінній *u*.

2. Виділення підвибірки: *smpl початок кінець* або *smpl @all* – для вибору всіх спостережень активного робочого файлу.

Усі команди, в тому числі *series*, діють на спостереження в межах активної підвибірки.

Приклади:

- а) *smpl 1999q2 2005q3* – з другого кварталу 1999 року до 3 кварталу 2005 року;
- б) *smpl @first +2 @last+2* – вибір другого спостереження з вибірки;
- в) *smpl @first @last* – аналог *@all* – вибір усіх спостережень.

3. Модуль числа: *abs(число)*

Наприклад: *abs(-6)* дорівнює 6.

abs(nrnd) – модуль нормально розподіленої випадкової величини.

3. Встановлення початкового значення для генератора псевдовипадкових чисел: *rndseed значення*.

rndseed 15 – встановити значення на 15.

Детальні інструкції до виконання практичного кейса 1 за допомогою пакета E.Views 6.0

1. Створюємо новий робочий файл, який міститиме 60 кварталних спостережень часового ряду ВВП розриву з ім'ям **gdp-gap**. Припустимо, що початок часового ряду – перший квартал

1996 року, а кінець – четвертий квартал 2000 року. Вікно відповідного робочого файлу наведено на рис. 1.

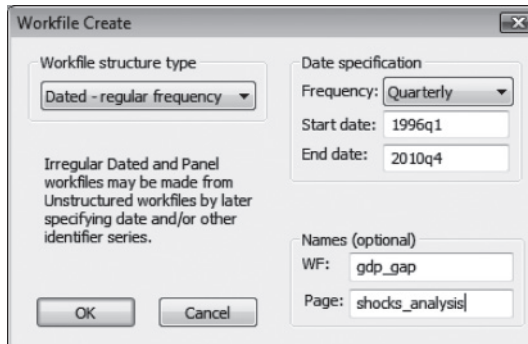


Рис. 1. Вікно для задання часового ряду ВВП розриву

2. Для аналізу дії шоків у межах AR(1) моделей, запропонованих різними аналітиками, згенеруємо спочатку відповідні часові ряди за допомогою команд пакета E.Views 6.0. У програмний рядок вводимо команди, показані в таблиці 1, та натискаємо **Enter**. Результати виконання заданих команд наведено на рис. 2. Таким чином згенеруємо дев'ять ВВП розривів, які використовуватимемо для моделювання різних шоків різними аналітиками.

Таблиця 1. Команди, що генерують початкові значення часових рядів ВВП розривів при різних типах шоків, створені різними аналітиками

Аналітик	Шок 1	Шок 2	Шок 3
Петро	series $y_{p_1} = 0$	series $y_{p_2} = 0$	series $y_{p_3} = 0$
Василь	series $y_{v_1} = 0$	series $y_{v_2} = 0$	series $y_{v_3} = 0$
Джон	series $y_{d_1} = 0$	series $y_{d_2} = 0$	series $y_{d_3} = 0$

Зауважимо, що введено такі умовні позначення:

y_{p_1} , y_{p_2} , y_{p_3} – часові ряди, що відповідно відображатимуть дію шоків 1, 2, 3 моделлю, яку запропонував аналітик Петро;

y_{v_1} , y_{v_2} , y_{v_3} – часові ряди, що відповідно відображатимуть дію шоків 1, 2, 3 моделлю, яку запропонував аналітик Василь;

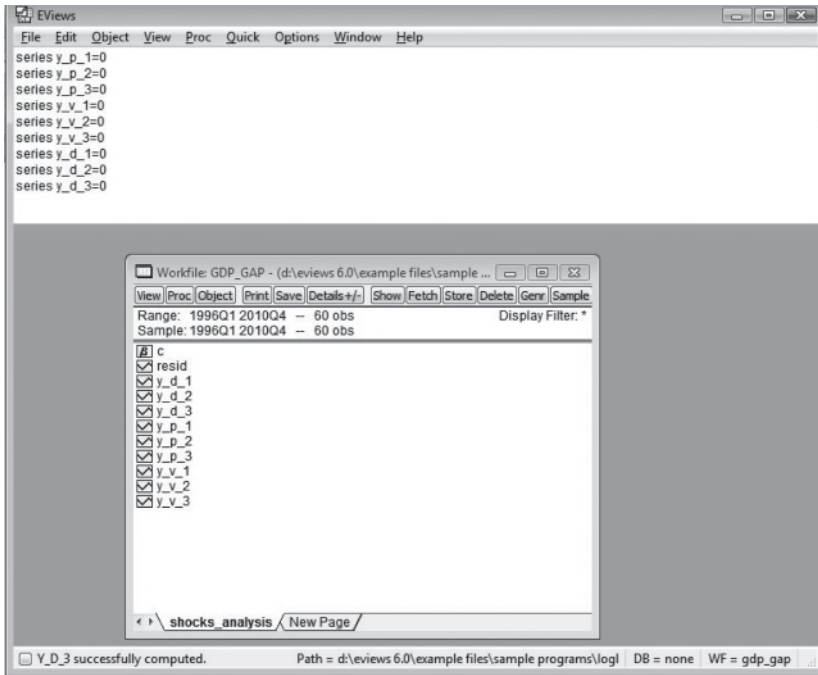


Рис. 2. Результати виконання команд зі створення часових рядів для аналізу дії трьох типів шоків трьома аналітиками

y_{d_1} , y_{d_2} , y_{d_3} – часові ряди, що відповідно відображатимуть дію шоків 1, 2, 3 моделлю, яку запропонував аналітик Джон.

3. За визначеними в кейсі умовами згенеруємо перший шок (шок 1). Для цього в програмний рядок вводимо **rndseed 33** (для задання початкового значення). Далі використовуємо команду **series u = nrnd** (для генерування випадкових величин). Згідно з умовами завдання, змінюємо два згенерованих значення для врахування посухи (дії шоку): спостереження u_{46} та u_{47} (другий і третій квартали 2007 року) замінюємо відповідно на $(u_{46} - 0,03)$ та $(u_{47} - 0,04)$. Зауважимо, що детально опис команд для генерування випадкових чисел командами пакета E.Views 6.0. наведено у пам'ятці 4 та 7.

4. За визначеними в кейсі умовами згенеруємо другий шок (шок 2) такою послідовністю команд: `series w = 0`, далі – `smpl @first+10 @first+10`, потім – `rndseed 7`, і в кінці – `series w = -abs(nrnd)`.

5. За визначеними в кейсі умовами згенеруємо третій шок (шок 3) такою послідовністю команд: `smpl @all`, далі – `series $\epsilon = 0$` . Відображаємо два шоки: для другого кварталу 2002 року $\epsilon = 0,0002$, а для першого кварталу 2003 року $\epsilon = -0,0005$.

Вікно E.Views після генерації шоків буде мати вигляд як на рис. 3.

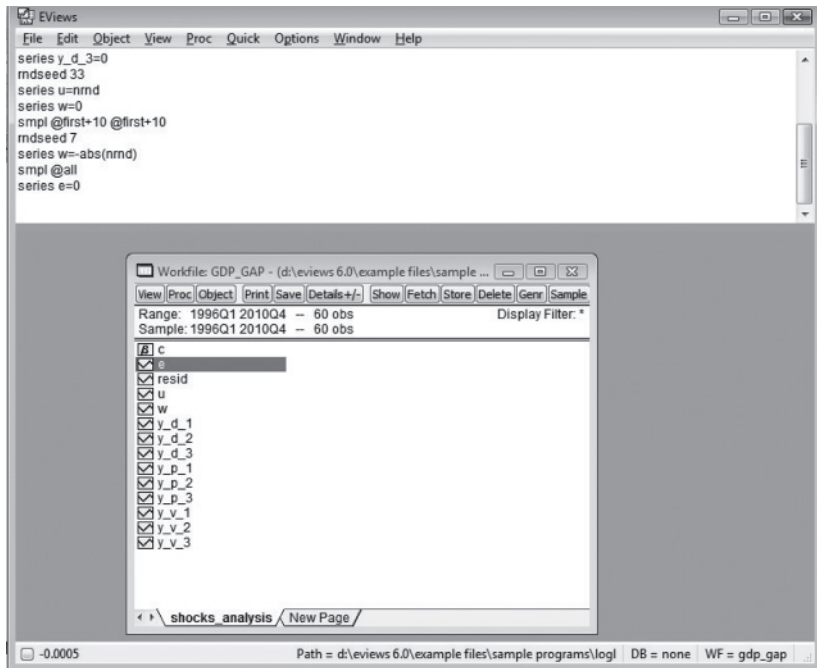


Рис. 3. Генерація часових рядів, що відображують дію трьох шоків за умовами завдання

6. Генеруємо лог-розрив, що отримується в моделях, запропонованих трьома аналітиками для трьох різних шоків. Для цього визначимо вибірку `smpl @first+1 @last`.

7. Далі згенеруємо лог-відхилення з урахуванням шоків для трьох різних моделей, згідно з командами, наведеними в таблиці 2.

Таблиця 2. Генерація відповідних моделей кожного аналітика для визначення лог-відхилень під дією різних шоків

Аналітик	Шок 1	Шок 2	Шок 3
Петро	series $y_{p_1} = .7^* y_{p_1}(-1) + u$	series $y_{p_2} = .7^* y_{p_2}(-1) + w$	series $y_{p_3} = .7^* y_{p_3}(-1) + \varepsilon$
Василь	series $y_{v_1} = -1.2^* y_{v_1}(-1) + u$	series $y_{v_2} = -1.2^* y_{v_2}(-1) + w$	series $y_{v_3} = -1.2^* y_{v_3}(-1) + \varepsilon$
Джон	series $y_{d_1} = 1.0^* y_{d_1}(-1) + u$	series $y_{d_2} = 1.0^* y_{d_2}(-1) + w$	series $y_{d_3} = 1.0^* y_{d_3}(-1) + \varepsilon$

На рис. 4 відображено вікно E.Views після генерації лог-відхилень.

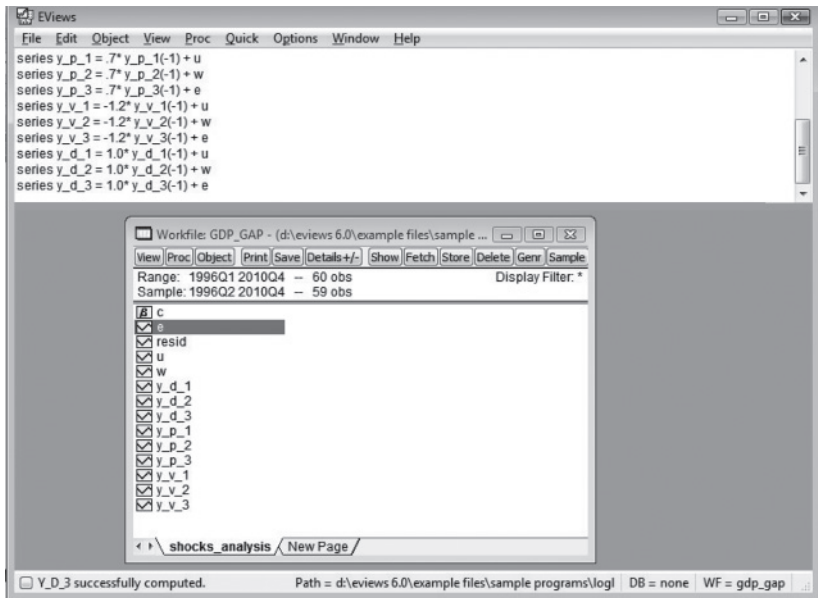


Рис. 4. Вікно E.Views після генерації лог-відхилень для кожної з трьох моделей і трьох шоків

8. Виберемо всю вибірку: `smpl @all`.

9. Зобразимо графічно отримані числові ряди: шоки та лог-відхилення. Вікно для завдання відповідних опцій наведено на рис. 5, а відповідний спільний графік для трьох шоків зображено на рис. 6. Зауважимо також, що всі згенеровані дані наведені в додатку Е до теми 1 і таблиці Е1–Е4.

Шоки (шок Е відобразимо на правій вертикальній осі):

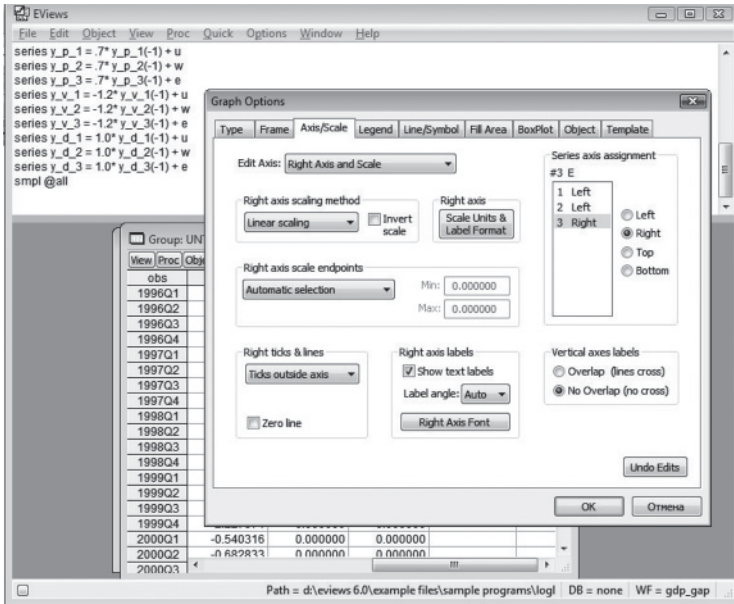


Рис. 5. Вікно для графічного відображення графіків шоків

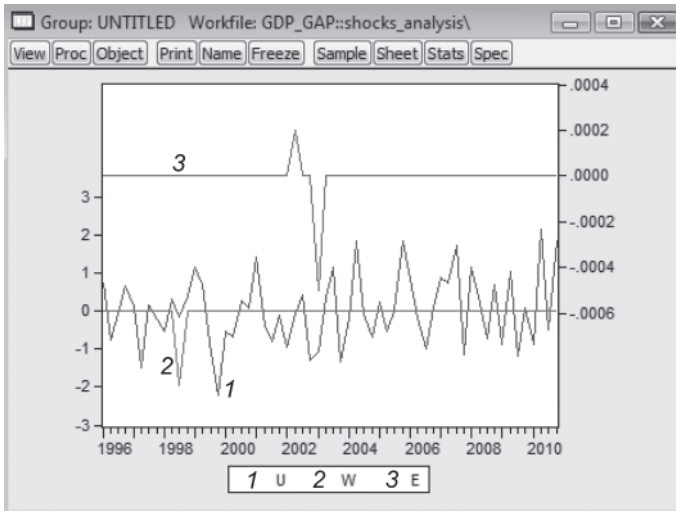


Рис. 6. Спільний графік дії трьох шоків

10. Графіки часових рядів лог-відхилень під дією різних шоків, виходячи з кожної з трьох AR(1) моделей, запропонованих трьома різними аналітиками, наведено відповідно на рис. 7–9. Зауважимо, що результати, отримані за моделлю Василя, відображено за допомогою правої осі.

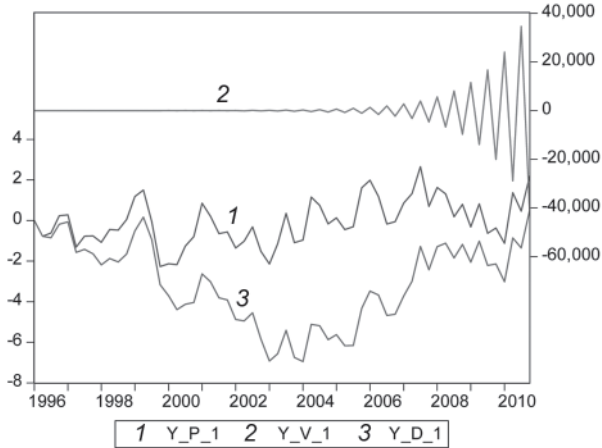


Рис. 7. Часові ряди лог-відхилень, отримані за трьома AR(1) моделями з урахуванням дії шоку 1

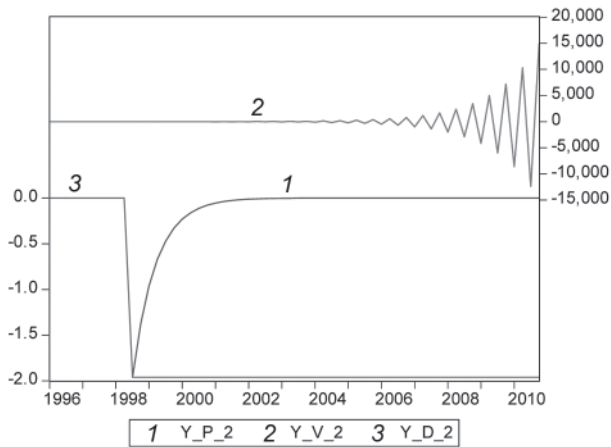


Рис. 8. Часові ряди лог-відхилень, отримані за трьома AR(1) моделями з урахуванням дії шоку 2

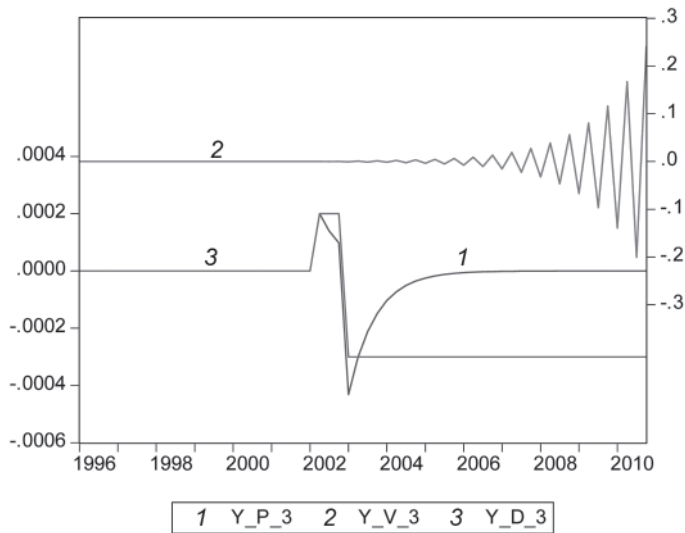


Рис. 9. Часові ряди лог-відхилень, отримані за трьома AR(1) моделями з урахуванням дії шоку 3

Як показує аналіз наведених графіків, головна різниця між шоками полягає в тому, що збурення w та ϵ тимчасове, а шок u , який описує погоду, стабільний. Тому, якщо модель опису ВВП розриву стаціонарна, то тимчасові збурення не призводитимуть до постійного відхилення фактичного ВВП від потенційного значення.

Що ж до особливостей моделювання динаміки ВВП розриву, то на графіках чітко простежується будь-який шок, чи постійний і великий за своїм значенням, чи тимчасовий і майже економічно незначущий, призводитиме до значного з часом відхилення ВВП розриву від нуля. Це зумовлено тим, що модель Василя нестаціонарна: ВВП розрив коливається біля нуля, але його дисперсія постійно зростає. Також нестаціонарною є модель Джона. З одного боку, модель Джона та стаціонарна модель Петра мають схожу реакцію на перший (погодний) шок. З іншого – нестаціонарність моделі Джона виявляється при аналізі поведінки ВВП розриву під час дії шоків 2 та 3: значення ВВП розриву відхиляється від почат-

кового рівноважного значення назавжди. Водночас модель Петра передбачає повернення ВВП до свого потенційного значення.

Реальна ситуація найкраще описується моделлю Петра, яка передбачає повернення ВВП розриву до свого попереднього стаціонарного стану. Але слід зазначити, що й інші моделі можуть бути корисними. Наприклад, модель Василя може добре описувати коливання змінних протягом бізнес-циклу, а модель Джона – вплив шоків, що призводять до постійних, а не тимчасових змін.

Завдання для самоперевірки 2. Ілюстрування дії шоків на стаціонарні та нестаціонарні часові ряди на прикладі простого AR(1) процесу: $Y_t = m + \alpha Y_{t-1} + u_t$

Часовий ряд цього процесу слід згенерувати таким чином:

1. 180 спостережень згенерувати зі 180 незалежними значеннями нормально розподілених випадкових величин з математичним сподіванням «нуль» та дисперсією 1 ($N(0,1)$). Загалом, цей ряд, який позначимо через u_t , $t = 1, \dots, 180$, – це білий шум, оскільки спостереження незалежно отримані з розподілу з однаковою дисперсією.

2. Потім спостереження u_{90} замінити на $(u_{90} + 19)$.

3. Згенерувати часовий ряд Y_t , $t = 1, \dots, 180$, усі початкові значення якого дорівнюють нулю.

4. Остаточні 179 спостережень згенерувати як AR(1) процес загального вигляду: $Y_t = m + \alpha Y_{t-1} + u_t$ для t від 2 до 180 з коефіцієнтами відповідно, $\alpha = 0,45$; $\alpha = 0,9$; $\alpha = 1$.

У результаті буде отримано три часових ряди з однаковими випадковими величинами.

Створіть спільний графік даних часових рядів і проаналізуйте дію шоку. Який з трьох AR(1) слід вважати нестаціонарним. Чому? Зробіть висновок щодо дії шоків на стаціонарні та нестаціонарні часові ряди. У чому полягають їхні особливості?

ДОДАТКИ ДО ТЕМИ 1

Додаток 1_А. Приклад визначення оптимальних лагів AR- та MA-складових в ARIMA моделях

Таблиця 1А. Специфікація послідовно оцінених AR-складових із розрахованими значеннями Акайк критерію

АІС	Специфікація AR-складової
[3.822982558]	D(reer) c d(reer(-7)) d(reer(-32)) d(reer(-1))
[3.811106399]	D(reer) c d(reer(-7)) d(reer(-32)) d(reer(-2))
[3.840074517]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-1)) d(reer(-2))
[3.822612483]	D(reer) c d(reer(-7)) d(reer(-32)) d(reer(-3))
[3.851583041]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-1)) d(reer(-3))
[3.839552166]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-2)) d(reer(-3))
[3.868530967]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-1)) d(reer(-2)) d(reer(-3))
[3.821491915]	D(reer) c d(reer(-7)) d(reer(-32)) d(reer(-4))
[3.850451385]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-1)) d(reer(-4))
[3.83813003]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-2)) d(reer(-4))
[3.867099336]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-1)) d(reer(-2)) d(reer(-4))
[3.850035131]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-3)) d(reer(-4))
[3.879007098]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-1)) d(reer(-3)) d(reer(-4))
[3.866512786]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-2)) d(reer(-3)) d(reer(-4))
[3.895492765]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-1)) d(reer(-2)) d(reer(-3)) d(reer(-4))
[3.808455027]	D(reer) c d(reer(-7)) d(reer(-32)) d(reer(-5))
[3.837337903]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-1)) d(reer(-5))
[3.825789169]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-2)) d(reer(-5))
[3.854692647]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-1)) d(reer(-2)) d(reer(-5))

AIC	Специфікація AR-складової
[3.837122261]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-3)) d(reer(-5))
[3.866028102]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-1)) d(reer(-3)) d(reer(-5))
[3.854328393]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-2)) d(reer(-3)) d(reer(-5))
[3.883255897]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-1)) d(reer(-2)) d(reer(-3)) d(reer(-5))
[3.835688144]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-4)) d(reer(-5))
[3.864572569]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-1)) d(reer(-4)) d(reer(-5))
[3.852554303]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-2)) d(reer(-4)) d(reer(-5))
[3.881459744]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-1)) d(reer(-2)) d(reer(-4)) d(reer(-5))
[3.864312514]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-3)) d(reer(-4)) d(reer(-5))
[3.893221159]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-1)) d(reer(-3)) d(reer(-4)) d(reer(-5))
[3.881033217]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-2)) d(reer(-3)) d(reer(-4)) d(reer(-5))
[3.909963831]	D(reer) c d(reer(-7)) d(reer(-32)) d(reer(-1)) d(reer(-2)) d(reer(-3)) d(reer(-4)) d(reer(-5))
[3.777527948]	D(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6))
[3.805498306]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-1)) d(reer(-6))
[3.792370538]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-2)) d(reer(-6))
[3.820370451]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-1)) d(reer(-2)) d(reer(-6))
[3.805939259]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-3)) d(reer(-6))
[3.834004068]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-1)) d(reer(-3)) d(reer(-6))
[3.820590233]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-2)) d(reer(-3)) d(reer(-6))
[3.848698107]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-1)) d(reer(-2)) d(reer(-3)) d(reer(-6))
[3.804602129]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-4)) d(reer(-6))
[3.832572602]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-1)) d(reer(-4)) d(reer(-6))
[3.818888691]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-2)) d(reer(-4)) d(reer(-6))
[3.846889324]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-1)) d(reer(-2)) d(reer(-4)) d(reer(-6))

AIC	Специфікація AR-складової
[3.832951742]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-3)) d(reer(-4)) d(reer(-6))
[3.861021738]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-1)) d(reer(-3)) d(reer(-4)) d(reer(-6))
[3.847022749]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-2)) d(reer(-3)) d(reer(-4)) d(reer(-6))
[3.875137279]	D(reer) c d(reer(-7)) d(reer(-32)) d(reer(-1)) d(reer(-2)) d(reer(-3)) d(reer(-4)) d(reer(-6))
[3.789963771]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-5)) d(reer(-6))
[3.817521848]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-1)) d(reer(-5)) d(reer(-6))
[3.805043736]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-2)) d(reer(-5)) d(reer(-6))
[3.832639634]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-1)) d(reer(-2)) d(reer(-5)) d(reer(-6))
[3.818472928]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-3)) d(reer(-5)) d(reer(-6))
[3.846133379]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-1)) d(reer(-3)) d(reer(-5)) d(reer(-6))
[3.833377889]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-2)) d(reer(-3)) d(reer(-5)) d(reer(-6))
[3.861092872]	D(reer) c d(reer(-7)) d(reer(-32)) d(reer(-1)) d(reer(-2)) d(reer(-3)) d(reer(-5)) d(reer(-6))
[3.816746578]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-4)) d(reer(-5)) d(reer(-6))
[3.844300736]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-1)) d(reer(-4)) d(reer(-5)) d(reer(-6))
[3.831239042]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-2)) d(reer(-4)) d(reer(-5)) d(reer(-6))
[3.85883139]	D(reer) c d(reer(-7)) d(reer(-32)) d(reer(-1)) d(reer(-2)) d(reer(-4)) d(reer(-5)) d(reer(-6))
[3.845196082]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-3)) d(reer(-4)) d(reer(-5)) d(reer(-6))
[3.872859222]	D(reer) c d(reer(-7)) d(reer(-32)) d(reer(-1)) d(reer(-3)) d(reer(-4)) d(reer(-5)) d(reer(-6))
[3.859490113]	D(reer) c d(reer(-7)) d(reer(-32)) d(reer(-2)) d(reer(-3)) d(reer(-4)) d(reer(-5)) d(reer(-6))
[3.887209271]	D(reer) c d(reer(-7)) d(reer(-32)) d(reer(-1)) d(reer(-2)) d(reer(-3)) d(reer(-4)) d(reer(-5)) d(reer(-6))
[3.794024216]	d(reer) c d(reer(-7)) d(reer(-32))

Таблиця 2А. Фрагмент специфікації послідовно оцінених МА-складових із розрахованими значеннями Шварц критерію

Schwarz	Специфікація МА-складової
[4.067028028]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) resids_tmp(-7) resids_tmp(-1)
[4.063237424]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) resids_tmp(-7) resids_tmp(-2)
[4.123235921]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) resids_tmp(-7) resids_tmp(-1) resids_tmp(-2)
[4.072123359]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) resids_tmp(-7) resids_tmp(-3)
[4.133464036]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) resids_tmp(-7) resids_tmp(-1) resids_tmp(-3)
[4.129703243]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) resids_tmp(-7) resids_tmp(-2) resids_tmp(-3)
[4.189330337]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) resids_tmp(-7) resids_tmp(-1) resids_tmp(-2) resids_tmp(-3)
[4.069201223]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) resids_tmp(-7) resids_tmp(-4)
[4.130841825]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) resids_tmp(-7) resids_tmp(-1) resids_tmp(-4)
[4.12579244]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) resids_tmp(-7) resids_tmp(-2) resids_tmp(-4)
[4.185929753]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) resids_tmp(-7) resids_tmp(-1) resids_tmp(-2) resids_tmp(-4)
[4.135722087]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) resids_tmp(-7) resids_tmp(-3) resids_tmp(-4)
[4.197150611]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) resids_tmp(-7) resids_tmp(-1) resids_tmp(-3) resids_tmp(-4)
[4.192091968]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) resids_tmp(-7) resids_tmp(-2) resids_tmp(-3) resids_tmp(-4)
[4.251716961]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) resids_tmp(-7) resids_tmp(-1) resids_tmp(-2) resids_tmp(-3) resids_tmp(-4)
[4.069072822]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) resids_tmp(-7) resids_tmp(-5)
[4.131109023]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) resids_tmp(-7) resids_tmp(-1) resids_tmp(-5)
[4.126533385]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) resids_tmp(-7) resids_tmp(-2) resids_tmp(-5)
[4.187185209]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) resids_tmp(-7) resids_tmp(-1) resids_tmp(-2) resids_tmp(-5)
[4.135637566]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) resids_tmp(-7) resids_tmp(-3) resids_tmp(-5)

Schwarz	Специфікація МА-складової
[4.197613008]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) residс_tmp(-7) resids_tmp(-1) residс_tmp(-3) residс_tmp(-5)
[4.193054864]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) residс_tmp(-7) resids_tmp(-2) residс_tmp(-3) residс_tmp(-5)
[4.253420723]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) residс_tmp(-7) resids_tmp(-1) residс_tmp(-2) residс_tmp(-3) residс_tmp(-5)
[4.132795623]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) residс_tmp(-7) resids_tmp(-4) residс_tmp(-5)
[4.19498559]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) residс_tmp(-7) resids_tmp(-1) residс_tmp(-4) residс_tmp(-5)
[4.189174378]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) residс_tmp(-7) resids_tmp(-2) residс_tmp(-4) residс_tmp(-5)
[4.24994827]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) residс_tmp(-7) resids_tmp(-1) residс_tmp(-2) residс_tmp(-4) residс_tmp(-5)
[4.199350501]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) residс_tmp(-7) resids_tmp(-3) residс_tmp(-4) residс_tmp(-5)
[4.261395072]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) residс_tmp(-7) resids_tmp(-1) residс_tmp(-3) residс_tmp(-4) residс_tmp(-5)
[4.255572116]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) residс_tmp(-7) resids_tmp(-2) residс_tmp(-3) residс_tmp(-4) residс_tmp(-5)
[4.315920597]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) residс_tmp(-7) resids_tmp(-1) residс_tmp(-2) residс_tmp(-3) residс_tmp(-4) resids_tmp(-5)
[4.072123503]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) residс_tmp(-7) resids_tmp(-6)
[4.133346644]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) residс_tmp(-7) resids_tmp(-1) residс_tmp(-6)
[4.129487101]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) residс_tmp(-7) resids_tmp(-2) residс_tmp(-6)
[4.188530487]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) residс_tmp(-7) resids_tmp(-1) residс_tmp(-2) residс_tmp(-6)
[4.138686385]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) residс_tmp(-7) resids_tmp(-3) residс_tmp(-6)
[4.199765269]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) residс_tmp(-7) resids_tmp(-1) residс_tmp(-3) residс_tmp(-6)
[4.195940068]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) residс_tmp(-7) resids_tmp(-2) residс_tmp(-3) residс_tmp(-6)
[4.254513715]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) residс_tmp(-7) resids_tmp(-1) residс_tmp(-2) residс_tmp(-3) residс_tmp(-6)
[4.13576706]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) residс_tmp(-7) resids_tmp(-4) residс_tmp(-6)

Schwarz	Специфікація МА-складової
[4.197245246]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) residс_тmp(-7) resids_тmp(-1) residс_тmp(-4) residс_тmp(-6)
[4.192118777]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) residс_тmp(-7) resids_тmp(-2) residс_тmp(-4) residс_тmp(-6)
[4.25139511]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) residс_тmp(-7) resids_тmp(-1) residс_тmp(-2) residс_тmp(-4) residс_тmp(-6)
[4.202288102]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) residс_тmp(-7) resids_тmp(-3) residс_тmp(-4) residс_тmp(-6)
[4.263536376]	D(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) residс_тmp(-7) resids_тmp(-1) residс_тmp(-3) residс_тmp(-4) residс_тmp(-6)
[4.258402045]	D(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) residс_тmp(-7) resids_тmp(-2) residс_тmp(-3) residс_тmp(-4) residс_тmp(-6)
[4.317057888]	D(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) residс_тmp(-7) resids_тmp(-1) residс_тmp(-2) residс_тmp(-3) residс_тmp(-4) resids_тmp(-6)
[4.135624842]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) residс_тmp(-7) resids_тmp(-5) residс_тmp(-6)
[4.197579114]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) residс_тmp(-7) resids_тmp(-1) residс_тmp(-5) residс_тmp(-6)
[4.192953744]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) residс_тmp(-7) resids_тmp(-2) residс_тmp(-5) residс_тmp(-6)
[4.252867195]	D(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) residс_тmp(-7) resids_тmp(-1) residс_тmp(-2) residс_тmp(-5) residс_тmp(-6)
[4.202189299]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) residс_тmp(-7) resids_тmp(-3) residс_тmp(-5) residс_тmp(-6)
[4.264074029]	D(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) residс_тmp(-7) resids_тmp(-1) residс_тmp(-3) residс_тmp(-5) residс_тmp(-6)
[4.259468283]	D(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) residс_тmp(-7) resids_тmp(-2) residс_тmp(-3) residс_тmp(-5) residс_тmp(-6)
[4.319013456]	D(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) residс_тmp(-7) resids_тmp(-1) residс_тmp(-2) residс_тmp(-3) residс_тmp(-5) resids_тmp(-6)
[4.199319671]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) residс_тmp(-7) resids_тmp(-4) residс_тmp(-5) residс_тmp(-6)
[4.261504191]	D(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) residс_тmp(-7) resids_тmp(-1) residс_тmp(-4) residс_тmp(-5) residс_тmp(-6)
[4.255643445]	D(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) residс_тmp(-7) resids_тmp(-2) residс_тmp(-4) residс_тmp(-5) residс_тmp(-6)
[4.315767272]	D(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) residс_тmp(-7) resids_тmp(-1) residс_тmp(-2) residс_тmp(-4) residс_тmp(-5) resids_тmp(-6)

Schwarz	Специфікація МА-складової
[4.265875489]	D(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) resids_tmp(-7) resids_tmp(-3) resids_tmp(-4) resids_tmp(-5) resids_tmp(-6)
[4.327904374]	D(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) resids_tmp(-7) resids_tmp(-1) resids_tmp(-3) resids_tmp(-4) resids_tmp(-5) resids_tmp(-6)
[4.322031205]	D(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) resids_tmp(-7) resids_tmp(-2) resids_tmp(-3) resids_tmp(-4) resids_tmp(-5) resids_tmp(-6)
[4.381635944]	D(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) resids_tmp(-7) resids_tmp(-1) resids_tmp(-2) resids_tmp(-3) resids_tmp(-4) resids_tmp(-5) resids_tmp(-6)
[3.381646632]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) resids_tmp(-7) resids_tmp(-32)
[3.398023696]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) resids_tmp(-7) resids_tmp(-1) resids_tmp(-32)
[3.477708642]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) resids_tmp(-7) resids_tmp(-2) resids_tmp(-32)
[3.488983927]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) resids_tmp(-7) resids_tmp(-1) resids_tmp(-2) resids_tmp(-32)
[3.427284316]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) resids_tmp(-7) resids_tmp(-3) resids_tmp(-32)
[3.452315786]	d(reer) c d(reer(-7)) d(reer(-32)) d(reer(-6)) resids_tmp(-7) resids_tmp(-1) resids_tmp(-3) resids_tmp(-32)

Додаток 1_Б. Сезонні перетворення на прикладі ARIMA (2,2) моделі

Сезонне перетворення ARMA (2,2) моделі

$$X_t = \rho_1 X_{t-1} + \rho_2 X_{t-2} + e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} \quad (1)$$

можна зробити, ввівши у модель AR сезонну складову або МА сезонну складову з лагом 4 для квартальних даних або з лагом 12 для місячних даних тощо.

Зазначимо, що процес перетворення ARMA(2,2) можна переписати із застосуванням лагового оператора

$$(1 - \rho_1 L - \rho_2 L^2) X_t = (1 + \theta_1 L + \theta_2 L^2) e_t. \quad (2)$$

Введення сезонного AR-компонента четвертого порядку (для квартальних даних) змінить модель (2) таким чином:

$$(1 - \rho_1 L - \rho_2 L^2) (1 - \delta L^4) X_t = (1 + \theta_1 L + \theta_2 L^2) e_t. \quad (3)$$

Введення сезонного МА-компонента четвертого порядку (для квартальних даних) надасть моделі (2) іншого вигляду:

$$(1 - \rho_1 L - \rho_2 L^2) X_t = (1 + \theta_1 L + \theta_2 L^2) (1 + \omega L^4) e_t.$$

Визначення того, який саме компонент (або компоненти) сезонності треба ввести у модель, здійснюється шляхом порівняння і відбору альтернативних моделей (модель без сезонних компонентів, модель із введенням SAR (сезонного AR) компонента, модель із введенням сезонного МА-компонента, модель з обома типами компонентів). Критеріями такого порівняння обов'язково мають бути тест залишків на білий шум, стаціонарність/обертівність моделей та інформаційні критерії Акайка і Шварца. Можна використовувати й такі критерії, як зіставлення корелограмів, коефіцієнт детермінації, критерії прогнозної якості.

Додаток 1_В. Критерії прогнозної якості

Після отримання прогнозу за ARIMA моделями постає питання про його точність. Її важко оцінити, виходячи лише з одного прогнозного значення. На практиці точність (якість) прогнозу або помилка прогнозу оцінюються, як правило, з урахуванням вибірки прогнозних даних усього періоду упередження. Відповідно, весь часовий ряд (вибірка даних) розбивається на дві частини. Одну з них використовують для оцінювання параметрів ARIMA моделей, а іншу розглядають як період упередження, для якого є фактичні дані, і застосовують для обчислення помилок прогнозу – різниці між фактичними значеннями та розрахунковими значеннями, отриманими за розробленою моделлю. Однак на практиці, за невеликої кількості спостережень, для попереднього розрахунку

прогнозної якості моделі часто використовують весь часовий ряд, без поділу на частини.

Зазвичай помилки прогнозу можуть бути як позитивними, так і негативними величинами, що робить недоцільним їх просте сумування, оскільки позитивні та негативні помилки взаємно компенсуються. Відповідно, при агрегації використовують квадрати помилок або їхні абсолютні значення.

Прийmemo такі позначення: $e_{t+s} = Y_{t+s} - \hat{Y}_{t+s}$, де Y_{t+s} – фактичне значення часового ряду в $(t+s)$ період; \hat{Y}_{t+s} – прогнозне значення часового ряду в $(t+s)$ період; e_{t+s} – помилка прогнозу в $(t + s)$ період.

Для визначення точності прогнозу існує ряд формальних критеріїв.

Розглянуто найбільш поширені критерії оцінювання якості (точності) прогнозу.

1. Середнє абсолютне відхилення (mean absolute deviation, MAD) – середня величина (за модулем) усіх помилок прогнозу:

$$MAD(e) = \frac{1}{n} \sum_{i=1}^n |e_i - \bar{e}|. \quad (1)$$

2. Середній квадрат помилки MSE (mean square error):

$$MSE(e) = \frac{1}{n} \sum_{i=1}^n e_i^2. \quad (2)$$

Оскільки цю статистику складають квадрати помилок, то очевидно, що великі помилки роблять найбільший внесок в суму квадратів.

Середній квадрат помилки є однією з найбільш поширених на практиці статистик (функцій втрат), яка використовується і для аналізу точності прогнозу. Критерій має переваги в ситуації, коли значні прогнозні помилки вважаються диспропорційно більш серйозними за незначні. Утім, це може бути і недоліком, якщо наявність значних помилок не вважається більш серйозною проблемою, ніж наявність незначних. Загалом, багато вчених рекомендують у випадку наявності нетипових значень використовувати критерії, що базуються на абсолютних величинах прогнозних помилок, а не їхніх квадратах.

Варто зауважити, що величини розглянутих критеріїв прогнозовної здатності не достатньо інформативні. Інтерпретація таких критеріїв здійснюється шляхом порівняння їхніх розрахункових значень для альтернативних моделей (з тією самою множиною залежних змінних і на тому самому прогнозному горизонті) і визначення оптимальної, найбільш точної з них.

Зазначимо, що величина MSE та аналогічних статистик залежить від обраної одиниці виміру, що, звичайно, дещо ускладнює порівняння значення цього критерію для різних економетричних моделей.

3. Корінь з середнього квадрата помилки RMSE (root mean square error):

$$RMSE(e) = \sqrt{\frac{1}{n} \sum_{i=1}^n e_i^2}. \quad (3)$$

4. Середня абсолютна помилка MAE (mean absolute error):

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i|. \quad (4)$$

Критерій MAE дає змогу визначити середнє значення помилки незалежно від знака. На відміну від критерію MSE, він не надає більшої ваги великим значенням помилки і, таким чином, відповідає лінійній (відносно пропорційних помилок) функції втрат.

5. Середня абсолютна відсоткова (процентна) помилка MAPE:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|e_i|}{Y_i} \cdot 100\% = \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_i}{\hat{Y}_i} - 1 \right| \cdot 100\%. \quad (5)$$

MAPE не залежить від масштабу. Вважається, що менше 10 % значення MAPE дає високу точність прогнозу, а отже, і якості моделі; від 10 % до 20 % – добру точність, від 20 % до 50 % – задовільну; більше 50% – незадовільну.

На жаль, цей критерій не рекомендують використовувати в ситуаціях, коли фактичні значення ряду можуть бути меншими за одиницю (за абсолютною величиною). У таких випадках цей критерій може мати занадто великі значення. Проілюструємо це твердження прикладом.

Припустимо, що було отримано прогнозне значення $\hat{Y}_i = 2,9$, тоді як фактичне значення дорівнювало $Y_i = 0,0001$. Припустимо, період упередження містить 400 спостережень. Відповідно, якщо додати цю одну складову помилки прогнозу до критерію MSE, ми отримаємо таке значення:

$$\frac{1}{400} (0,0001 - 2,9)^2 = 0,0210. \quad (6)$$

Отримане в (6) значення хоча і завелике, проте достатньо реальне в багатьох випадках для певного ряду даних. Однак, якщо ми додамо цю складову до критерію MAPE, то отримаємо:

$$\frac{1}{400} \left| \frac{0,0001 - 2,9}{0,0001} \right| = 72,4975. \quad (7)$$

Отримане в (7) значення надзвичайно велике. Це демонструє проблеми застосування MAPE критерію в випадках, коли часові ряди даних можуть приймати значення з різними знаками або менші за одиницю (за абсолютною величиною).

MAPE має переваги, якщо застосовується в процесі випадкового блукання в рівнях, де нуль-прогнозне значення критерію дорівнює одиниці. Таким чином, якщо прогнозна модель дає значення MAPE менше одиниці, вона вважається кращою за процес випадкового блукання. Цей критерій не надійний у випадку, коли часовий ряд може приймати значення (за абсолютною величиною) менші за одиницю.

6. Нерівність Тейла (U-статистика):

$$U = \sqrt{\frac{1/n \sum_{i=1}^n (\Delta Y_i - \Delta \hat{Y}_i)^2}{1/n \sum_{i=1}^n (\Delta Y_i)^2}}.$$

Критерій Тейла є досить популярним у практичних дослідженнях. Він був запропонований відомим статистиком Г. Тейлом і відомий як «коефіцієнт нерівності» (або «коефіцієнт невідповіднос-

ті»). Популярність критерію Тейла зумовлена тим, що він не залежить від масштабу і може легко використовуватися для порівняння точності прогнозів, отриманих за допомогою різних методів та моделей.

Значення критерію потребують в межах від нуля до одиниці; дорівнюють нулю – для досконалого прогнозу, та одиниці – для прогнозу «жодних змін». Відповідно, чим ближче значення цього критерію до 0, тим меншою є помилка прогнозу.

Описані вище критерії точності використовують як інформацію, що дає змогу зробити остаточний висновок про якість отриманого прогнозу або з множини моделей вибрати найкращу, або обрати оптимальний прогноз.

Оптимальним вважається той прогноз, який можна одержати за наявних обставин. Часто його називають прогнозом раціональних сподівань.

Слід зауважити, що при використанні економетричної моделі для прогнозування майже всі сучасні економетричні пакети прикладних програм видають широкий спектр значень критеріїв точності прогнозу. Завдання дослідника грамотно їх проінтерпретувати та зробити остаточний висновок про якість розробленої моделі або про якість отриманого на її основі прогнозу.

Додаток 1_Г. Порівняльна таблиця критеріїв та орієнтирів при ідентифікації ARIMA моделей

Таблиця 1Г. Порівняльна таблиця критеріїв та орієнтирів процесу ідентифікації ARIMA моделей

Орієнтир (здебільшого достатня, але не необхідна умова правильної специфікації ARIMA моделі)	Критерій (необхідна, але часто недостатня умова для правильної специфікації ARIMA моделі)
У якості орієнтирів нижче подано ті основні процедури, які дають змогу уявити особливості поведінки досліджуваного процесу. Їх можна враховувати при побудові моделі, але модель може бути адекватною (правильно специфікованою) і ні	У якості критеріїв нижче подано ті процедури, які є ознаками адекватності моделі. Їх обов'язково треба здійснити. Негативний результат такої процедури означає, що модель не можна використовувати на практиці

Орієнтир (здебільшого достатня, але не необхідна умова правильної специфікації ARIMA моделі)	Критерій (необхідна, але часто недостатня умова для правильної специфікації ARIMA моделі)
<i>Стаціонарність ряду (визначення порядку інтеграції ряду)</i>	
<p>Аналіз графіка часового ряду. Дає змогу виявити наявність і характер тренду, зміну варіативності ряду, що корисно для подальшого тестування, але не є підставою для тверджень про стаціонарність.</p> <p>Аналіз графіків ACF/PACF. Дає можливість виявити деякі випадки нестаціонарності ряду, а також наявну сезонність</p>	<p>Тест на одиничний корінь, зокрема, розширений тест Дікі-Фулера (ADF) з автоматичним підбором кількості лагів (завдяки цій рисі може оптимізуватися і до простого тесту Дікі-Фулера, з нулем лагів). Дає змогу виявити велику кількість випадків нестаціонарності. Винятками є нелінійні процеси, наприклад, GARCH-процеси. Тест їх не виявляє, вважаючи стаціонарними</p>
<i>Визначення лагів AR- та MA-складових моделі (параметрів p та q)</i>	
<p>Аналіз графіків ACF/PACF. Дає змогу визначити порядок для чистих AR- або MA-процесів, тобто максимальну кількість лагів, які специфікують чистий AR(p) та MA(q)-процеси, а також значущі лаги для змішаних ARIMA процесів для проведення процедури Хенона-Рісанена. Водночас це не означає, що незначущі лаги за ACF/PACF не можуть бути включені в модель. Також метод є чутливим до обсягу вибірки: що більша вибірка, то важче відрізнити слабкі лаги від сильних.</p> <p>Значення t-статистики та p-value відповідного коефіцієнта в моделі – це орієнтир для послідовного перебирання лагів у процедурі Хенона-Рісанена. Однак це не означає, що статистично незначущі лаги обов'язково необхідно включити з моделі</p>	<p>Аналіз поліпшення/погіршення значень інформаційних критеріїв Акайке та Шварца (AIC, Schwarz). Якщо певний лаг виключається з моделі і це призводить до зростання відповідного критерію, лаг не треба виключати, оскільки він важливий. Так само якщо певний лаг було виключено і значення критерію покращилося (зменшилося), то лаг можна вважати незначущим для моделі і його слід виключити з остаточної специфікації</p>
<i>Адекватність та правильна специфікація моделі</i>	
<p>Аналіз коефіцієнта детермінації (R^2). Що більший коефіцієнт детермінації, то модель ефективна. Утім, цілком можливо, що адекватна модель матиме невисокий коефіцієнт детермінації, наприклад, 0,4. Так само некоректні моделі можуть мати високі значення показника</p>	<p>Аналіз залишків моделі (аналіз на білий шум). Якщо залишки моделі не є білим шумом, то модель неадекватна і не може застосовуватися на практиці, незалежно від значень інших критеріїв та орієнтирів</p>

Орієнтир (здебільшого достатня, але не необхідна умова правильної специфікації ARIMA моделі)	Критерій (необхідна, але часто недостатня умова для правильної специфікації ARIMA моделі)
<p>Порівняння корелограмів (теоретичного та емпіричного). Коли модель відтворює важливі автокореляційні моменти реального ряду, можемо припустити, що така модель більше відповідає реальному процесу. Проте така відповідність не означає, що модель коректна.</p> <p>Аналіз критерію прогнозної якості, зокрема MAPE. Середня абсолютна відсоткова похибка – один із поширених критеріїв прогнозної якості моделі. Але оцінювання прогнозної якості не дарма робиться після оцінки адекватності моделі: якщо модель неадекватна, то, який би малий відсоток не складала похибка прогнозу, модель непридатна для практичного застосування</p>	<p>Аналіз моделі на стабільність (перевірка на стаціонарність AR-складової та на обертовність MA-складової оціненої моделі). Інколи після переоцінювання ARIMA моделі нелінійним методом найменших квадратів у виводі програми з'являються повідомлення: AR process non-stationary або MA roots irreversible, тобто відповідні корені лежать за межами одиничного кола. Навіть якщо це «єдиний недолік моделі», модель слід вважати непридатною до практичного застосування</p>

Додаток 1_Г. Програма в командах пакета E.Views 6.0 щодо вибору найкращої специфікації ARIMA за допомогою процедури Хенона–Рісанена та інструкція до її виконання

Інструкція до виконання програми:

1) зайдіть у **File>Open>Program** і виберіть файл програми (hannon1.prg);

2) відкрийте робочий файл (**Workfile**) з потрібними часовими рядами. Перейдіть у вікно програми (робочий файл лишиться активним, про що свідчитиме статус-рядок внизу головного вікна E.Views);

3) натисніть **Run** у вікні програми. Відкриється вікно, де потрібно ввести параметри. Є два способи використати програму:

- [ser_name] [num_lags] [t-stat minimum] [t-stat maximum] – програма здійснить процедуру оптимізації для всіх лагів із першого до [num lags] (параметри вводять без лапок чи дужок через проміжок), не перевіряючи лаги, під час першого

оцінювання всіх лагів дають t -stat мінімальний [t -stat minimum] і максимальний [t -stat maximum]. Перші не включаються в модель, другі включаються раз і назавжди. Це зменшує час виконання програми;

- [ser_name] 0 [t -stat minimum] [t -stat maximum] [«lag1 lag2 lag3....»] – робить те саме, але перевіряє лише лаги, введені в останньому параметрі.

Приклади:

reer 5 1 2 – протестує лаги з першого до п'ятого, зразу відкинувши ті, для яких значення t -stat<1, і без перевірки включити ті, для яких значення t -stat>2;

dlog(reer, 2, 12) 0 0.1 2.5 «1 2 3 4 5 6 7 32» – перевірить другі різниці логарифмів показника, сезонно диференційовані з дванадцятим лагом. Перевіряє перший – сьомий і тридцять другий лаги, які зразу дають t -stat між 0.1 і 2.5 тощо;

d(cpi) 5 0 100 – перевірить усі комбінації з п'яти лагів перших різниць показника.

Текст програми в командах пакета E.Views 6.0 щодо вибору найкращої специфікації ARIMA за допомогою процедури Хенона–Рісанена

```
subroutine smpl_ar (string %eqname, vector lags_vector)
%ar_part = %est_series_name + " c "
for !lagn = 1 to @obs(lags_vector)
    if lags_vector(!lagn)>0 then
        call gen_ar_lag(!lagn)
        %ar_part = %ar_part + %res_str_lag
    endif
next
equation {%eqname}.ls {%ar_part}
endsub

subroutine smpl_ma (string %eqname, string %arpart, vector lags_vector,
string %res_series_name)
%arma_part = %arpart
for !lagn = 1 to @obs(lags_vector)
    if lags_vector(!lagn)>0 then
        %arma_part = %arma_part + %res_series_name + "(-" + @
str(lags_vector(!lagn)) + ")" "
    endif
next
```

```

equation {%eqname}.ls {%arma_part}
endsub

subroutine gen_ar_lag(scalar !lag)
%res_str_lag = ""
if @instr(@trim(%est_series_name), "(")>0 then
  if @instr(%est_series_name, ",")>0 then
    %res_str_lag = @mid(@trim(%est_series_name), 1, @
instr(%est_series_name, ",")-1)+ "(-" + @str(lags_vector(!lag)) + ")" +
@mid(@trim(%est_series_name), @instr(%est_series_name, ",")+ " "
    else
    %res_str_lag = @mid(@trim(%est_series_name), 1, @
instr(%est_series_name, ")")-1)+ "(-" + @str(lags_vector(!lag)) + ") "
  endif
else
  %res_str_lag = %est_series_name + "(-" + @str(lags_
vector(!lag)) + ")" "
endif
endsub

subroutine final_arma_est(vector arlags, vector malags)
%part = %est_series_name + " c "
for !lagn = 1 to @obs(arlags)
  if arlags(!lagn)>0 then
    %part = %part + "ar(" + @str(arlags(!lagn)) + ") "
  endif
next
for !lagn = 1 to @obs(malags)
  if malags(!lagn)>0 then
    %part = %part + "ma(" + @str(malags(!lagn)) + ") "
  endif
next
equation final_arma.ls {%part}
endsub

subroutine concat_vectors(string %resname, vector v1, vector v2)
vector(@obs(v1)+@obs(v2)) {%resname}
for !d = 1 to @obs({%resname})
  if !d<=@obs(v1) then
    {%resname}(!d) = g_lags(!d)
  else
    {%resname}(!d) = v2(!d-@obs(v1))
  endif
next
endsub

```

```

subroutine adv_aic_optimize(string %eqname, matrix lags_matrix, vector
g_lags)
!nlgs = @obs(@columnextract(lags_matrix, 1))
vector(2^!nlgs) aics = na
for !eqn = 1 to 2^!nlgs
    vector lvector = @columnextract(lags_matrix, !eqn)
    call concat_vectors("eqvector", g_lags, lvector)
    call smpl_ar(%eqname, eqvector)
    aics(!eqn) = {%eqname}.@aic
    call tab_out ("["+@str{%eqname}.@aic+"]", 1, 1)
    call tab_out (%ar_part, 2, 0)
next
scalar best_aic = @min(aics)
scalar best_ar = 0
for !n = 1 to @obs(aics)
    if aics(!n) = best_aic then
        best_ar = !n
    endif
next
vector opt_ar_vector = @columnextract(lags_matrix, best_ar)
call concat_vectors("optimal_ar_vector", g_lags, opt_ar_vector)
call smpl_ar("optimal_ar", optimal_ar_vector)
    call tab_out ("--minimum AIC equation--", 2, 0)
    call tab_out ("["+@str(optimal_ar.@aic)+""]", 1, 1)
    call tab_out (%ar_part, 2, 0)
show optimal_ar
endsub

subroutine adv_schw_optimize(string %eqname, string %arprt, string
%residname, matrix lags_matrix, vector g_lags)
!nlgs = @obs(@columnextract(lags_matrix, 1))
vector(2^!nlgs) schw = na
for !eqn = 1 to 2^!nlgs
    vector lvector = @columnextract(lags_matrix, !eqn)
    call concat_vectors("eqvector", g_lags, lvector)
    call smpl_ma(%eqname, %arprt, eqvector, %residname)
    schw(!eqn) = {%eqname}.@schwarz
    call tab_out ("["+@str{%eqname}.@schwarz)+""]", 1, 1)
    call tab_out (%arma_part, 2, 0)
next
!best_schw = @min(schw)
!best_ma = 0
for !n = 1 to @obs(schw)
    if schw(!n) = !best_schw then

```

```

!best_ma = !n
endif

next
vector opt_arma_vector = @columnextract(lags_matrix, !best_ma)
call concat_vectors("optimal_arma_vector", g_lags, opt_arma_vector)
call smpl_ma("optimal_arma", %arprt, optimal_arma_vector, %residname)
    call tab_out ("--minimum Schwarz equation--", 2, 0)
    call tab_out ("["+@str(optimal_arma.@schwarz)+"]", 1, 1)
    call tab_out (%arma_part, 2, 0)
show optimal_arma
endsub

subroutine gen_combi(string %lcname, vector lvect)
!nlags = @obs(lvect)
matrix(!nlags, 2^!nlags) {%lcname}
matrix(!nlags, 2^!nlags) combi
for !j = 1 to 2^!nlags
    !num = !j
    for !i = 1 to !nlags
        combi(!i, !j) = 2*(!num/2-@floor(!num/2))
        {%lcname}(!i, !j) = combi(!i, !j)*lvect(!i)
        !num = (!num - combi(!i, !j))/2
    next
next
endsub

subroutine t_stat_divide(string %name, vector studs, vector slags)
!lleng = @obs(slags)
vector(!lleng) good_{%name}
vector(!lleng) {%name}
scalar good_{%name}_index = 1
scalar {%name}_index = 1
scalar max = !max_t
scalar min = !min_t
for !k = 1 to @obs(studs)
    if @abs(studs(!k))>=!max_t then
        good_{%name}(good_{%name}_index) = slags(!k)
        good_{%name}_index = good_{%name}_index+1
    endif
    if @abs(studs(!k))>=!min_t and @abs(studs(!k)) < !max_t then
        {%name}({%name}_index) = slags(!k)
        {%name}_index = {%name}_index+1
    endif
next
!k = 1

```

```

while good_{%name}(!k)>0
    if !k < !leng then
        !k = !k+1
    else
        !k = !k+1
        exitloop
    endif
wend
if !k-1=0 then !k=2 endif
vector good_{%name}_t = @subextract(good_{%name}, 1, 1, !k-1, 1)
good_{%name} = good_{%name}_t
delete good_{%name}_t
!k = 1
while {%name}(!k)>0
    if !k < !leng then
        !k = !k+1
    else
        !k = !k+1
        exitloop
    endif
wend
if !k-1=0 then !k=2 endif
vector {%name}_t = @subextract({%name}, 1, 1, !k-1, 1)
{%name} = {%name}_t
delete {%name}_t
endsub

subroutine str_parse_to_vector(string %src, string %vname)
%rest = @trim(@mid(%src, 2, @len(%src)-1))
!tok_n = 0
while @instr(%rest, " ")
    %rest = @mid(%rest, 1+@instr(%rest, " "))
    !tok_n = !tok_n+1
wend
vector(!tok_n+1) {%vname}
%rest = @trim(@mid(%src, 2, @len(%src)-1))
for !vind = 1 to !tok_n
    %tok = @trim(@mid(%rest, 1, @instr(%rest, " ")))
    %rest = @trim(@mid(%rest, 1+@instr(%rest, " ")))
    {%vname}(!vind) = @val(%tok)
next
{%vname}(!tok_n+1) = @val(@trim(@mid(%rest, 1, @len(%rest)-1)))
endsub

subroutine tab_out (string %text, scalar !pos, scalar !stay)

```

```

logtable(!log_pos, !pos) = %text
if !stay=0 then !log_pos = !log_pos + 1 endif
endsub

subroutine vect_to_str( vector vectsrc)
%tout = "["
for !n = 1 to @obs(vectsrc)
    %tout = %tout + @str(vectsrc(!n)) + " "
next
%tout = %tout + "]"
endsub

'ENTRY POINT
table logtable
!log_pos = 1
%est_series_name = %0
%num_lags = %1
!min_t = @val(%2)
!max_t = @val(%3)
if @str(%1) = 0 then
    call str_parse_to_vector(%4, "all_lags")
else
    vector({%num_lags}) all_lags
    for !n = 1 to {%num_lags}
        all_lags(!n) = !n
    next
endif
call tab_out("-----", 1, 1)
call tab_out("-----", 2, 0)
call tab_out("AR SPECIFICATION", 2, 0)
call tab_out("-----", 1, 1)
call tab_out("-----", 2, 0)
call smpl_ar("basic_eq_ar", all_lags)
vector stud_basic = basic_eq_ar.@tstats
vector stud_basic_t = @subextract(stud_basic, 2, 1)
stud_basic = stud_basic_t
delete stud_basic_t
call t_stat_divide ("lags", stud_basic, all_lags)
call tab_out("strong lags", 1, 1)
call vect_to_str(good_lags)
call tab_out(%tout, 2, 0)
call tab_out("weak lags", 1, 1)
call vect_to_str(lags)
call tab_out(%tout, 2, 0)
call gen_combi("lcombi", lags)

```

```

call tab_out ("AIC", 1, 1)
call tab_out (" AR Specification", 2, 0)
call adv_aic_optimize("temp_ar", lcombi, good_lags)
optimal_ar.makesresids resids_tmp
call smpl_ma("basic_eq_arma", %ar_part, all_lags, "resids_tmp")
vector stud_basic_ma = basic_eq_arma.@tstats
vector stud_basic_ma_t = @subextract(stud_basic_ma, @obs(stud_basic_
ma)-@obs(all_lags)+1, 1)
stud_basic_ma = stud_basic_ma_t
delete stud_basic_ma_t
call tab_out("-----", 1, 1)
call tab_out("-----", 2, 0)
call tab_out("MA SPECIFICATION", 2, 0)
call tab_out("-----", 1, 1)
call tab_out("-----", 2, 0)
call t_stat_divide ("rlags", stud_basic_ma, all_lags)
call tab_out("strong lags", 1, 1)
call vect_to_str(good_rlags)
call tab_out(%tout, 2, 0)
call tab_out("weak lags", 1, 1)
call vect_to_str(rlags)
call tab_out(%tout, 2, 0)
call gen_combi("rlcombi", rlags)
call tab_out ("Schwarz", 1, 1)
call tab_out (" MA Specification", 2, 0)
call adv_schw_optimize("temp_arma", %ar_part, "resids_tmp", rlcombi,
good_rlags)
call final_arma_est(optimal_ar_vector, optimal_arma_vector)
show final_arma
delete aics
delete all_lags
delete best_aic
delete best_ar
delete combi
delete eqvector
delete good_lags
delete good_lags_index
delete good_rlags
delete good_rlags_index
delete lags
delete lags_index
delete rlags
delete rlags_index
delete lcombi
delete lvector

```

```

delete max
delete min
delete opt_ar_vector
delete optimal_ar_vector
delete opt_arma_vector
delete optimal_arma_vector
delete rcombi
delete schw
delete stud_basic
delete stud_basic_ma
delete temp_ar
delete temp_arma
freeze logtable
delete logtable

```

Додаток 1_Д. Статистичні дані значень показника реального ефективного обмінного курсу євро для країн Євросоюзу (змінна REER)

Таблиця 1Д. Статистичні дані показника реального ефективного обмінного курсу євро для країн Євросоюзу (змінна REER) за 01.2000 – 06.2008 рр.

№	Період спостереження	Значення показника REER	№	Період спостереження	Значення показника REER
1	01.2000	106.37	52	04.2004	120.74
2	02.2000	104.64	53	05.2004	122.12
3	03.2000	102.28	54	06.2004	121.89
4	04.2000	100.81	55	07.2004	122.46
5	05.2000	98.51	56	08.2004	122.02
6	06.2000	102.28	57	09.2004	122.59
7	07.2000	101.33	58	10.2004	123.98
8	08.2000	97.94	59	11.2004	120.91
9	09.2000	96.2	60	12.2004	127.52
10	10.2000	94.95	61	01.2005	126.52
11	11.2000	95.56	62	02.2005	124.31
12	12.2000	99.14	63	03.2005	125.39
13	01.2001	102.75	64	04.2005	124.04
14	02.2001	101.81	65	05.2005	122.72
15	03.2001	102.13	66	06.2005	119.28

Продовження табл. 1Д

№	Період спостереження	Значення показника REER	№	Період спостереження	Значення показника REER
16	04.2001	101.49	67	07.2005	119.46
17	05.2001	99.28	68	08.2005	120.47
18	06.2001	97.98	69	09.2005	119.92
19	07.2001	98.42	70	10.2005	119.2
20	08.2001	102.18	71	11.2005	118.17
21	09.2001	103.11	72	12.2005	118.31
22	10.2001	103.15	73	01.2006	118.81
23	11.2001	101.38	74	02.2006	117.91
24	12.2001	101.94	75	03.2006	118.79
25	01.2002	101.35	76	04.2006	120.27
26	02.2002	100.51	77	05.2006	121.89
27	03.2002	100.76	78	06.2006	122.19
28	04.2002	101.16	79	07.2006	122.45
29	05.2002	104.02	80	08.2006	122.39
30	06.2002	106.71	81	09.2006	122.05
31	07.2002	108.84	82	10.2006	121.35
32	08.2002	108.11	83	11.2006	122.3
33	09.2002	108.51	84	12.2006	123.47
34	10.2002	109.02	85	01.2007	122.49
35	11.2002	109.09	86	02.2007	122.93
36	12.2002	111.42	87	03.2007	123.66
37	01.2003	114.19	88	04.2007	124.92
38	02.2003	115.84	89	05.2007	124.67
39	03.2003	116.55	90	06.2007	123.95
40	04.2003	116.7	91	07.2007	124.96
41	05.2003	121.08	92	08.2007	124.33
42	06.2003	121.92	93	09.2007	125.67
43	07.2003	120	94	10.2007	126.89
44	08.2003	118.59	95	11.2007	128.95
45	09.2003	118.37	96	12.2007	128.77
46	10.2003	120.8	97	01.2008	129.3
47	11.2003	120.87	98	02.2008	128.79
48	12.2003	124.55	99	03.2008	132.58
49	01.2004	125.58	100	04.2008	133.97
50	02.2004	125.43	101	05.2008	133.12
51	03.2004	122.82	102	06.2008	132.84

Додаток 1_Е. Згенеровані дані часових рядів при виконанні практичного кейса аналізу дії шоків

Таблиця 1Е. Згенеровані часові ряди випадкових величин (ϵ , u , w)

№ спостереження	Період спостереження	ϵ	u	w
1	1996Q1	0.000000	0.739635	0.000000
2	1996Q2	0.000000	-0.776305	0.000000
3	1996Q3	0.000000	-0.071131	0.000000
4	1996Q4	0.000000	0.666700	0.000000
5	1997Q1	0.000000	0.115757	0.000000
6	1997Q2	0.000000	-1.506333	0.000000
7	1997Q3	0.000000	0.149131	0.000000
8	1997Q4	0.000000	-0.217799	0.000000
9	1998Q1	0.000000	-0.542871	0.000000
10	1998Q2	0.000000	0.308336	0.000000
11	1998Q3	0.000000	-0.163249	-1.960928
12	1998Q4	0.000000	0.384684	0.000000
13	1999Q1	0.000000	1.145022	0.000000
14	1999Q2	0.000000	0.674262	0.000000
15	1999Q3	0.000000	-1.105390	0.000000
16	1999Q4	0.000000	-2.227971	0.000000
17	2000Q1	0.000000	-0.540316	0.000000
18	2000Q2	0.000000	-0.682833	0.000000
19	2000Q3	0.000000	0.265490	0.000000
20	2000Q4	0.000000	0.087869	0.000000
21	2001Q1	0.000000	1.407769	0.000000
22	2001Q2	0.000000	-0.383278	0.000000
23	2001Q3	0.000000	-0.790797	0.000000
24	2001Q4	0.000000	-0.113336	0.000000
25	2002Q1	0.000000	-0.961818	0.000000
26	2002Q2	0.000200	-0.069498	0.000000
27	2002Q3	0.000000	0.403586	0.000000
28	2002Q4	0.000000	-1.298044	0.000000
29	2003Q1	-0.000500	-1.081956	0.000000
30	2003Q2	0.000000	0.358444	0.000000
31	2003Q3	0.000000	1.155250	0.000000
32	2003Q4	0.000000	-1.338826	0.000000
33	2004Q1	0.000000	-0.197912	0.000000
34	2004Q2	0.000000	1.831121	0.000000

Продовження табл. 1Е

№ спостереження	Період спостереження	ε	и	ш
35	2004Q3	0.000000	-0.065846	0.000000
36	2004Q4	0.000000	-0.687300	0.000000
37	2005Q1	0.000000	0.243602	0.000000
38	2005Q2	0.000000	-0.540377	0.000000
39	2005Q3	0.000000	0.010836	0.000000
40	2005Q4	0.000000	1.824617	0.000000
41	2006Q1	0.000000	0.857288	0.000000
42	2006Q2	0.000000	-0.202299	0.000000
43	2006Q3	0.000000	-1.006288	0.000000
44	2006Q4	0.000000	0.067986	0.000000
45	2007Q1	0.000000	0.897404	0.000000
46	2007Q2	0.000000	0.720991	0.000000
47	2007Q3	0.000000	1.724538	0.000000
48	2007Q4	0.000000	-1.153281	0.000000
49	2008Q1	0.000000	1.136093	0.000000
50	2008Q2	0.000000	0.174710	0.000000
51	2008Q3	0.000000	-0.732159	0.000000
52	2008Q4	0.000000	0.669060	0.000000
53	2009Q1	0.000000	-0.873199	0.000000
54	2009Q2	0.000000	1.033051	0.000000
55	2009Q3	0.000000	-1.196914	0.000000
56	2009Q4	0.000000	0.071998	0.000000
57	2010Q1	0.000000	-0.869403	0.000000
58	2010Q2	0.000000	2.157254	0.000000
59	2010Q3	0.000000	-0.494915	0.000000
60	2010Q4	0.000000	1.863117	0.000000

Таблиця 2Е. Часові ряди y_{p_1} , y_{p_2} , y_{p_3} лог-відхилень ВВП, отриманих за моделлю першого аналітика

№ спостереження	Період спостереження	y_{p_1}	y_{p_2}	y_{p_3}
1	1996Q1	0.000000	0.000000	0.000000
2	1996Q2	-0.776305	0.000000	0.000000
3	1996Q3	-0.614544	0.000000	0.000000
4	1996Q4	0.236519	0.000000	0.000000
5	1997Q1	0.281320	0.000000	0.000000

Продовження табл. 2Е

№ спостереження	Період спостереження	y_{p_1}	y_{p_2}	y_{p_3}
6	1997Q2	-1.309409	0.000000	0.000000
7	1997Q3	-0.767456	0.000000	0.000000
8	1997Q4	-0.755018	0.000000	0.000000
9	1998Q1	-1.071383	0.000000	0.000000
10	1998Q2	-0.441632	0.000000	0.000000
11	1998Q3	-0.472392	-1.960928	0.000000
12	1998Q4	0.054010	-1.372650	0.000000
13	1999Q1	1.182829	-0.960855	0.000000
14	1999Q2	1.502242	-0.672598	0.000000
15	1999Q3	-0.053820	-0.470819	0.000000
16	1999Q4	-2.265645	-0.329573	0.000000
17	2000Q1	-2.126267	-0.230701	0.000000
18	2000Q2	-2.171220	-0.161491	0.000000
19	2000Q3	-1.254364	-0.113044	0.000000
20	2000Q4	-0.790186	-0.079131	0.000000
21	2001Q1	0.854639	-0.055391	0.000000
22	2001Q2	0.214969	-0.038774	0.000000
23	2001Q3	-0.640318	-0.027142	0.000000
24	2001Q4	-0.561559	-0.018999	0.000000
25	2002Q1	-1.354909	-0.013299	0.000000
26	2002Q2	-1.017935	-0.009310	0.000200
27	2002Q3	-0.308968	-0.006517	0.000140
28	2002Q4	-1.514322	-0.004562	9.80E-05
29	2003Q1	-2.141981	-0.003193	-0.000431
30	2003Q2	-1.140943	-0.002235	-0.000302
31	2003Q3	0.356590	-0.001565	-0.000211
32	2003Q4	-1.089213	-0.001095	-0.000148
33	2004Q1	-0.960361	-0.000767	-0.000104
34	2004Q2	1.158868	-0.000537	-7.25E-05
35	2004Q3	0.745362	-0.000376	-5.08E-05
36	2004Q4	-0.165546	-0.000263	-3.55E-05
37	2005Q1	0.127719	-0.000184	-2.49E-05
38	2005Q2	-0.450974	-0.000129	-1.74E-05
39	2005Q3	-0.304846	-9.02E-05	-1.22E-05
40	2005Q4	1.611225	-6.31E-05	-8.53E-06
41	2006Q1	1.985145	-4.42E-05	-5.97E-06
42	2006Q2	1.187302	-3.09E-05	-4.18E-06

Закінчення табл. 2Е

№ спостереження	Період спостереження	y_{p_1}	y_{p_2}	y_{p_3}
43	2006Q3	-0.175176	-2.17E-05	-2.93E-06
44	2006Q4	-0.054638	-1.52E-05	-2.05E-06
45	2007Q1	0.859158	-1.06E-05	-1.43E-06
46	2007Q2	1.322401	-7.43E-06	-1.00E-06
47	2007Q3	2.650219	-5.20E-06	-7.02E-07
48	2007Q4	0.701872	-3.64E-06	-4.92E-07
49	2008Q1	1.627403	-2.55E-06	-3.44E-07
50	2008Q2	1.313893	-1.78E-06	-2.41E-07
51	2008Q3	0.187565	-1.25E-06	-1.69E-07
52	2008Q4	0.800356	-8.74E-07	-1.18E-07
53	2009Q1	-0.312950	-6.12E-07	-8.26E-08
54	2009Q2	0.813987	-4.28E-07	-5.79E-08
55	2009Q3	-0.627123	-3.00E-07	-4.05E-08
56	2009Q4	-0.366988	-2.10E-07	-2.83E-08
57	2010Q1	-1.126295	-1.47E-07	-1.98E-08
58	2010Q2	1.368847	-1.03E-07	-1.39E-08
59	2010Q3	0.463278	-7.20E-08	-9.72E-09
60	2010Q4	2.187412	-5.04E-08	-6.81E-09

Таблиця 3Е. Часові ряди y_{v_1} , y_{v_2} , y_{v_3} лог-відхилень ВВП, отриманих за моделлю другого аналітика

№ спостереження	Період спостереження	y_{v_1}	y_{v_2}	y_{v_3}
1	1996Q1	0.000000	0.000000	0.000000
2	1996Q2	-0.776305	0.000000	0.000000
3	1996Q3	0.860434	0.000000	0.000000
4	1996Q4	-0.365821	0.000000	0.000000
5	1997Q1	0.554742	0.000000	0.000000
6	1997Q2	-2.172024	0.000000	0.000000
7	1997Q3	2.755560	0.000000	0.000000
8	1997Q4	-3.524470	0.000000	0.000000
9	1998Q1	3.686493	0.000000	0.000000
10	1998Q2	-4.115456	0.000000	0.000000
11	1998Q3	4.775298	-1.960928	0.000000
12	1998Q4	-5.345673	2.353114	0.000000
13	1999Q1	7.559829	-2.823737	0.000000
14	1999Q2	-8.397533	3.388484	0.000000

Продовження табл. 3Е

№ спостереження	Період спостереження	y_v_1	y_v_2	y_v_3
15	1999Q3	8.971649	-4.066181	0.000000
16	1999Q4	-12.99395	4.879417	0.000000
17	2000Q1	15.05242	-5.855301	0.000000
18	2000Q2	-18.74574	7.026361	0.000000
19	2000Q3	22.76038	-8.431633	0.000000
20	2000Q4	-27.22459	10.11796	0.000000
21	2001Q1	34.07727	-12.14155	0.000000
22	2001Q2	-41.27601	14.56986	0.000000
23	2001Q3	48.74041	-17.48383	0.000000
24	2001Q4	-58.60183	20.98060	0.000000
25	2002Q1	69.36038	-25.17672	0.000000
26	2002Q2	-83.30195	30.21206	0.000200
27	2002Q3	100.3659	-36.25448	-0.000240
28	2002Q4	-121.7372	43.50537	0.000288
29	2003Q1	145.0026	-52.20645	-0.000846
30	2003Q2	-173.6447	62.64774	0.001015
31	2003Q3	209.5289	-75.17728	-0.001218
32	2003Q4	-252.7735	90.21274	0.001461
33	2004Q1	303.1303	-108.2553	-0.001753
34	2004Q2	-361.9253	129.9063	0.002104
35	2004Q3	434.2445	-155.8876	-0.002525
36	2004Q4	-521.7807	187.0651	0.003030
37	2005Q1	626.3804	-224.4782	-0.003636
38	2005Q2	-752.1968	269.3738	0.004363
39	2005Q3	902.6470	-323.2486	-0.005236
40	2005Q4	-1081.352	387.8983	0.006283
41	2006Q1	1298.479	-465.4779	-0.007539
42	2006Q2	-1558.378	558.5735	0.009047
43	2006Q3	1869.047	-670.2882	-0.010857
44	2006Q4	-2242.788	804.3459	0.013028
45	2007Q1	2692.243	-965.2150	-0.015634
46	2007Q2	-3229.971	1158.258	0.018761
47	2007Q3	3877.690	-1389.910	-0.022513
48	2007Q4	-4654.381	1667.892	0.027015
49	2008Q1	5586.393	-2001.470	-0.032418
50	2008Q2	-6703.497	2401.764	0.038902
51	2008Q3	8043.465	-2882.117	-0.046682
52	2008Q4	-9651.489	3458.540	0.056019

Закінчення табл. 3Е

№ спостереження	Період спостереження	y_{v_1}	y_{v_2}	y_{v_3}
53	2009Q1	11580.91	-4150.248	-0.067223
54	2009Q2	-13896.06	4980.298	0.080667
55	2009Q3	16674.08	-5976.357	-0.096800
56	2009Q4	-20008.82	7171.629	0.116161
57	2010Q1	24009.72	-8605.954	-0.139393
58	2010Q2	-28809.50	10327.15	0.167271
59	2010Q3	34570.91	-12392.57	-0.200725
60	2010Q4	-41483.23	14871.09	0.240870

Таблиця 4Е. Часові ряди y_{d_1} , y_{d_2} , y_{d_3} лог-відхилень ВВП, отриманих за моделлю третього аналітика

№ спостереження	Період спостереження	y_{d_1}	y_{d_2}	y_{d_3}
1	1996Q1	0.000000	0.000000	0.000000
2	1996Q2	-0.776305	0.000000	0.000000
3	1996Q3	-0.847436	0.000000	0.000000
4	1996Q4	-0.180736	0.000000	0.000000
5	1997Q1	-0.064979	0.000000	0.000000
6	1997Q2	-1.571312	0.000000	0.000000
7	1997Q3	-1.422181	0.000000	0.000000
8	1997Q4	-1.639980	0.000000	0.000000
9	1998Q1	-2.182851	0.000000	0.000000
10	1998Q2	-1.874515	0.000000	0.000000
11	1998Q3	-2.037764	-1.960928	0.000000
12	1998Q4	-1.653080	-1.960928	0.000000
13	1999Q1	-0.508058	-1.960928	0.000000
14	1999Q2	0.166204	-1.960928	0.000000
15	1999Q3	-0.939186	-1.960928	0.000000
16	1999Q4	-3.167157	-1.960928	0.000000
17	2000Q1	-3.707472	-1.960928	0.000000
18	2000Q2	-4.390305	-1.960928	0.000000
19	2000Q3	-4.124815	-1.960928	0.000000
20	2000Q4	-4.036946	-1.960928	0.000000
21	2001Q1	-2.629177	-1.960928	0.000000
22	2001Q2	-3.012456	-1.960928	0.000000
23	2001Q3	-3.803252	-1.960928	0.000000

Продовження табл. 4Е

№ спостереження	Період спостереження	y_{d_1}	Y_{d_2}	y_{d_3}
24	2001Q4	-3.916588	-1.960928	0.000000
25	2002Q1	-4.878406	-1.960928	0.000000
26	2002Q2	-4.947904	-1.960928	0.000200
27	2002Q3	-4.544318	-1.960928	0.000200
28	2002Q4	-5.842362	-1.960928	0.000200
29	2003Q1	-6.924318	-1.960928	-0.000300
30	2003Q2	-6.565874	-1.960928	-0.000300
31	2003Q3	-5.410624	-1.960928	-0.000300
32	2003Q4	-6.749450	-1.960928	-0.000300
33	2004Q1	-6.947362	-1.960928	-0.000300
34	2004Q2	-5.116241	-1.960928	-0.000300
35	2004Q3	-5.182087	-1.960928	-0.000300
36	2004Q4	-5.869387	-1.960928	-0.000300
37	2005Q1	-5.625785	-1.960928	-0.000300
38	2005Q2	-6.166162	-1.960928	-0.000300
39	2005Q3	-6.155327	-1.960928	-0.000300
40	2005Q4	-4.330710	-1.960928	-0.000300
41	2006Q1	-3.473422	-1.960928	-0.000300
42	2006Q2	-3.675721	-1.960928	-0.000300
43	2006Q3	-4.682009	-1.960928	-0.000300
44	2006Q4	-4.614024	-1.960928	-0.000300
45	2007Q1	-3.716619	-1.960928	-0.000300
46	2007Q2	-2.995629	-1.960928	-0.000300
47	2007Q3	-1.271090	-1.960928	-0.000300
48	2007Q4	-2.424372	-1.960928	-0.000300
49	2008Q1	-1.288279	-1.960928	-0.000300
50	2008Q2	-1.113569	-1.960928	-0.000300
51	2008Q3	-1.845728	-1.960928	-0.000300
52	2008Q4	-1.176668	-1.960928	-0.000300
53	2009Q1	-2.049866	-1.960928	-0.000300
54	2009Q2	-1.016815	-1.960928	-0.000300
55	2009Q3	-2.213729	-1.960928	-0.000300
56	2009Q4	-2.141731	-1.960928	-0.000300
57	2010Q1	-3.011134	-1.960928	-0.000300
58	2010Q2	-0.853881	-1.960928	-0.000300
59	2010Q3	-1.348796	-1.960928	-0.000300
60	2010Q4	0.514322	-1.960928	-0.000300

Практика побудови ARCH/GARCH моделі із використанням програмного пакета E.Views 6.0. Інструкції побудови ARCH/GARCH моделі на прикладі реального часового ряду

Вступ

Більшість класичних економетричних моделей є лінійними за параметрами, тобто такими, в яких зміна досліджуваного показника залежить пропорційно від зміни певного фактора або лагового значення самого показника (простіше, параметр помножується на змінну моделі). Властивості лінійних моделей детально досліджені, зрозумілі та можуть бути легко проінтерпретовані. Це сприяло поширенню цього класу моделей на практиці. Зауважимо, що деякі нелінійні моделі шляхом логарифмування або за допомогою інших операцій трансформування можуть бути зведені до лінійних. Відповідно, для них правильні всі основні положення класичного лінійного аналізу.

Варто врахувати також, що багато фінансових взаємозв'язків внутрішньо нелінійні. Це пов'язано з особливостями фінансових даних, наприклад, з тенденцією фінансових часових рядів (наприклад, доходів по активам) мати розподіл з нетиповими (більш піднятими) критичними зонами («хвостами») та зміщеннями щодо середнього (leptokurtosis). Крім того, досить часто фінансові часові ряди змінюються кластерами або пулами (volatility clustering or volatility pooling), коли більшим відхиленням часового ряду передують високі, а незначним – несуттєві відхилення (незалежно від знака). Можливий також ефект левереджу: тенденція до збільшення волатильності не при зростанні, а при падінні певних фінансово-економічних показників (асиметричність інформації).

Загалом моделювання та прогнозування волатильності (volatility) на фондовому ринку стало пріоритетним напрямом як теоретичних, так і прикладних досліджень останніх років. Волатильність, яка вимірюється стандартним відхиленням або дисперсією

дохідності (наприклад, цінних паперів) часто використовується як грубе наближення до виміру загального ризику фінансових інструментів (цінних паперів). Значна кількість моделей виміру ринкового ризику (value-at-risk) потребує оцінювання або прогнозування волатильності (volatility, розкиду) як вхідного параметру. Зазначимо, що параметр волатильності (volatility) цін на фондовому ринку безпосередньо використовується в формулі Блека–Шольса (Black–Scholes) визначення ціни опціонного контракту.

Отже, важливим у фінансових дослідженнях є моделювання волатильності (volatility), яка вимірюється в основному середньоквадратичним відхиленням або дисперсією та належить до нелінійних процесів. Як правило, точне визначення певного типу нелінійності можливе лише в контексті певної моделі. Однак достатньо узагальнено економічний часовий ряд (процес) можна вважати нелінійним, якщо тривалі та короткі шоки мають різний вплив на зміну цього ряду, тобто зміна ряду як відповідь на шок непропорційна щодо його розміру. Додатково можна припустити, що ряд нелінійний, якщо його реакція на шоки залежить від того, негативні чи позитивні ці шоки.

У 1997 році Кембелом Ло та МакКінлі (Campbell Lo, MacKinlay) було сформовано загальну концепцію щодо класу нелінійних процесів, які описують нелінійну залежність поточного значення певного показника від поточного та лагових значень випадкових величин: Найпопулярнішими серед нелінійних моделей є ARCH, GARCH, що їх використовують для моделювання та прогнозування волатильності (volatility) фінансових рядів, та моделі перемикання (switching models) для опису різної поведінки часових рядів у різні проміжки часу.

2.1. Для чого використовувати ARCH/GARCH моделі за наявності ARIMA методології?

Необхідність аналізувати дані фінансових ринків, які генерують часові ряди з періодичністю в одну хвилину, п'ять хвилин, годину, день, дала поштовх для розвитку «високочастотної економетрики». На думку провідних західних експертів, у цій частині еко-

нометрика менше дотична до економіки і більше – до обробки сигналів, наприклад, звуку.

Крім того, природа фінансових процесів часто така:

- розподіл середніх значень і дисперсії в часі не здійснюється за нормальним законом. У багатьох фінансових рядах функція щільності, порівняно з нормальним законом, відрізняється більш піднятою критичною зоною (наявністю так званих товстих хвостів). Інакше кажучи, більшість фінансових часових рядів має такі розподіли значень, у яких екстремальні значення показника трапляються частіше, ніж це передбачено нормальним законом;
- у процесах відбувається кластеризація волатильності: спонтанні сильні шоки показника не затухають, а тривають ще певний час. Наприклад, паніка на ринку акцій може зумовити різкі коливання ціни акцій упродовж певного часу, тобто в цей час дисперсія значень може бути в рази більшою.

Водночас ARIMA моделі, незважаючи на потужний потенціал відображення складної динаміки, належать до класу лінійних моделей зі сталою дисперсією, тому не можуть коректно відображати внутрішню нелінійність, притаманну багатьом описаним вище фінансовим процесам. На відміну від них, ARCH/GARCH моделі належать до класу нелінійних моделей з умовною дисперсією, яка змінюється в часі, що дає змогу, крім середнього значення досліджуваного показника, одночасно моделювати динаміку його дисперсії, а відтак коректно описувати такі явища, як кластеризація волатильності, асиметричність інформації тощо.

2.2. На основі яких даних можна будувати ARCH/GARCH моделі? Особливості імпортування високочастотних даних у E.Views?

Наявність ARCH/GARCH ефекту в часовому ряді варто перевіряти, коли йдеться про дані, періодичність більша за щомісячну (тобто тиждень, доба, година, хвилина). Зауважимо, що ARCH/GARCH ефект рідко спостерігається в часових рядах із місячною, квартальною та річною періодичністю.

Прикладами фінансових часових рядів з високою частотністю можуть слугувати ціни та індекси фондових та товарних бірж. У практичних дослідженнях їх часто аналізують у формі *геометричних доходів* (*geometric returns*), тобто *логарифма темпу зростання ціни активу* (або темпу зростання значення біржового індексу) y_t :

$$r_t = \log\left(\frac{y_t}{y_{t-1}}\right) = \log(y_t) - \log(y_{t-1}), \quad (2.1)$$

де r_t – геометричний дохід від активу. В E.Views ця трансформація здійснюється оператором різниці логарифмів $\text{dlog}(y)$. Якщо спробувати побудувати ARIMA на основі таких даних, постане низка проблем: ACF/PACF не матиме яскраво виражених сильних лагів, буде вказувати на білий шум або міститиме слабкі лаги без регулярної структури, які майже не затухають. Коефіцієнт детермінації R^2 моделі буде незначним, оскільки лінійна модель із припущенням сталої дисперсії нездатна відтворити динаміку коливань.

Для ілюстрації виявлення ARCH/GARCH ефекту в часовому ряду розглянемо ряд щоденних значень індексу ПФТС за період із 10.03.97 до 10.17.09 (для економії місця тільки перші 108 спостережень цього ряду наведено в додатку 2_A до теми 2, таблиця 1A, файл pfts.xls). Створимо робочий файл у E.Views. Спочатку – для неструктурованих спостережень (у цьому прикладі маємо 2928 спостережень). У вікні робочого файлу обираємо опцію **Unstructured/undated** для характеристики типу вхідних даних та задаємо кількість спостережень: **2928** в опції **Observations**. Крім того, за бажанням, можна задати ім'я ряду даних в опції **Names**. Для розглядуваного прикладу зазначимо ім'я: **PFTS** (див. рис. 2.1).

Оскільки дані є нерегулярними, для того, щоб пакет E.Views розпізнавав всі спостереження як співвіднесені з конкретними моментами часу, потрібно, крім ряду значень показника, імпортувати і ряд відповідних їм дат, що мають бути введені в форматі, який E.Views розпізнає як дату. Наприклад, якщо початковий ряд дат записаний у стандартному форматі Excel: **ДД.ММ.РРРР**, його слід трансформувати у формат **РРРР/ММ/ДД** (коректно розпізнається E.Views *).

* У самому робочому файлі програма замінить формат дати на свій, ММ/ДД/РРРР.

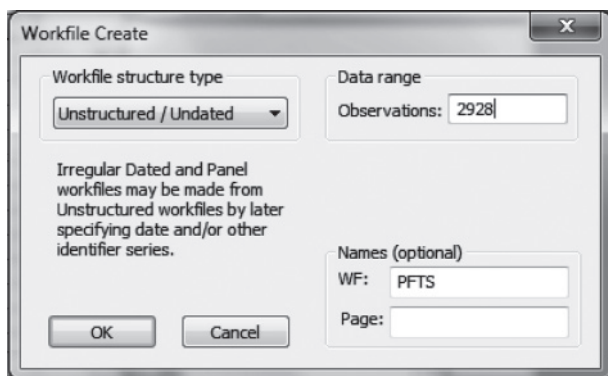


Рис. 2.1. Вікно E.Views для початкового визначення формату даних

Зазначену трансформацію в Excel можна зробити послідовністю таких дій:

- 1) виділити всі клітинки з датами;
- 2) у контекстному меню клітинок обрати «Формат клітинок» (див. рис. 2.2):

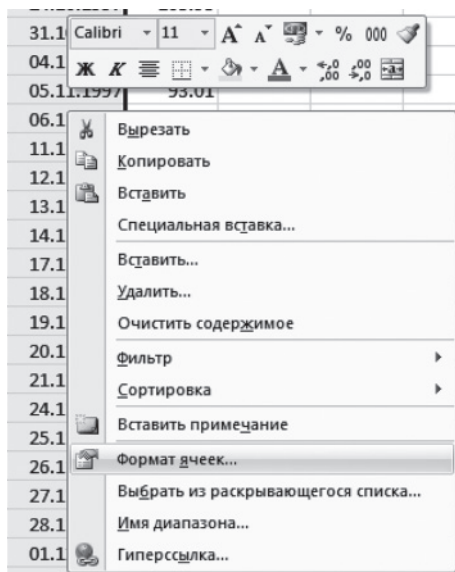


Рис. 2.2. Вікно Excel для форматування інформації

- 3) у вікні, що відкривається, обрати «Всі формати» і ввести формат дати, як показано на рисунку 2.3.

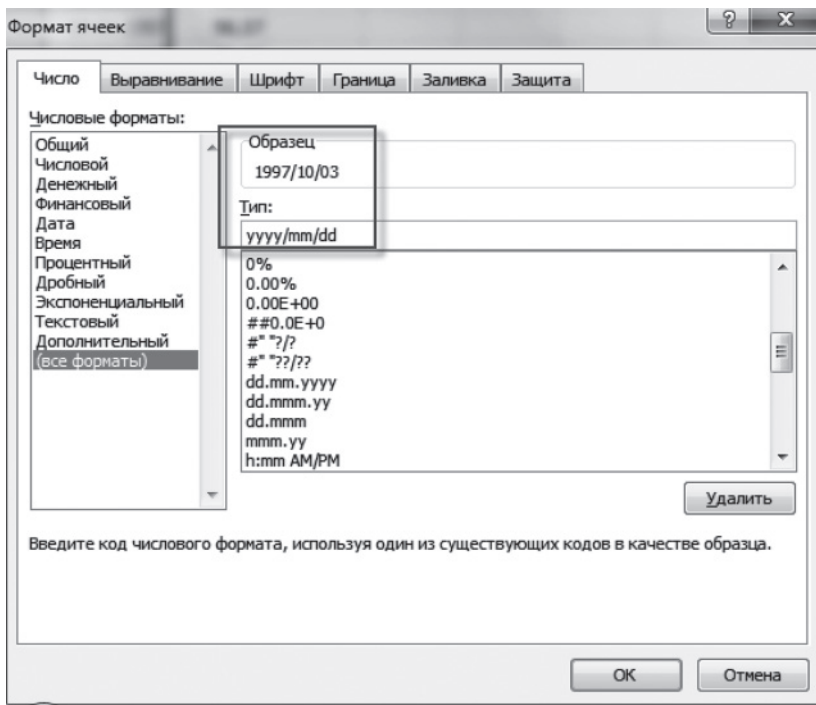
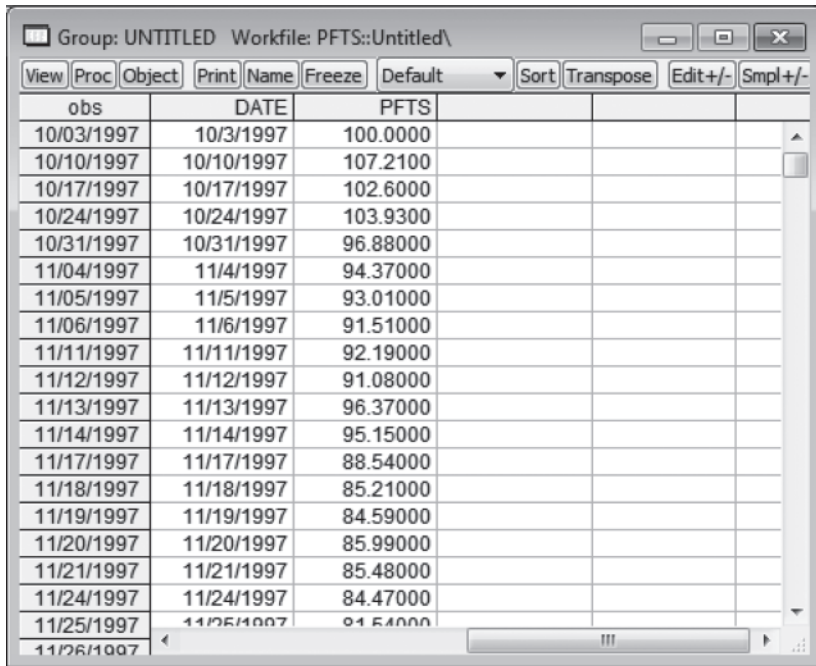


Рис. 2.3. Вікно Excel для вибору типу формату інформації

Якщо йдеться про російську версію Excel, формат дати може бути такий: ГГГГ/ММ/ДД. Щоб переконатися, що дата розпізнана правильно (спочатку рік, потім місяць, потім день), потрібно подивитися на приклад, який висвічується в правому верхньому куті відкритого вікна (див. рис. 2.3).

Після цього слід відкрити у E.Views нову порожню групу (**Quick>Empty Group...**), в яку можна просто скопіювати обидва ряди – ряд дат і ряд даних. Ознака того, що дати правильно розпізнані програмою, – вони будуть переформатовані як ММ/ДД/РРРР. Якщо при послідовності виконання наведених вище операцій була допущена помилка, наприклад, ряд дат розпізнаний як

об'єкт ALPHA, слід видалити цей об'єкт і провести процедуру імпорту дат знову. На рис. 2.4, відображено результати коректного введення інформації як самих значень ряду індексу ПФТС, так і ряду відповідних дат.



obs	DATE	PFTS
10/03/1997	10/3/1997	100.0000
10/10/1997	10/10/1997	107.2100
10/17/1997	10/17/1997	102.6000
10/24/1997	10/24/1997	103.9300
10/31/1997	10/31/1997	96.88000
11/04/1997	11/4/1997	94.37000
11/05/1997	11/5/1997	93.01000
11/06/1997	11/6/1997	91.51000
11/11/1997	11/11/1997	92.19000
11/12/1997	11/12/1997	91.08000
11/13/1997	11/13/1997	96.37000
11/14/1997	11/14/1997	95.15000
11/17/1997	11/17/1997	88.54000
11/18/1997	11/18/1997	85.21000
11/19/1997	11/19/1997	84.59000
11/20/1997	11/20/1997	85.99000
11/21/1997	11/21/1997	85.48000
11/24/1997	11/24/1997	84.47000
11/25/1997	11/25/1997	81.51000
11/26/1997		

Рис. 2.4. Введена в робочий файл E.Vews інформація щодо значень ряду індексу ПФТС та ряду відповідних дат

Після введення необхідних даних потрібно послідовно виконати команди: **Proc/ «Structure/Resize...»**. При цьому слід вказати в полі **Workfile structure type** значення **Dated-specified by date series**, а в полі **Date series** – ім'я імпортованого ряду дат (у нашому випадку – **date**) (див. рис. 2.5).

Послідовність зазначених вище дій дає змогу привести у відповідність спостереження і нерегулярні моменти часу, в які ці спостереження були здійснені.

Якщо випадково ряд дат виявиться регулярним, програма автоматично зробить файл регулярно датованим. Тобто описаний

підхід імпортування універсальний. Його можна застосовувати як до регулярних, так і до нерегулярних часових рядів.

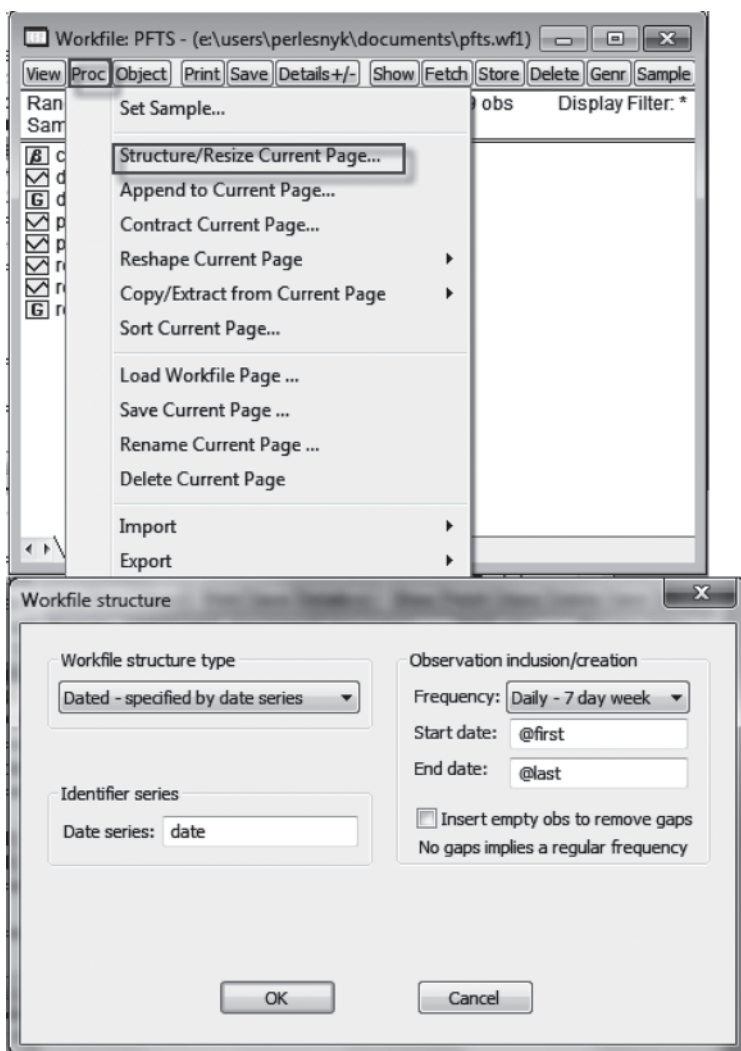


Рис. 2.5. Вікно E.Views для приведення у відповідність спостереження в нерегулярні моменти часу з конкретними датами

Імпортовану інформацію, необхідно перевірити на наявність «відсутніх» («NA» (not available)) спостережень, оскільки в такому випадку при оцінюванні ARCH/GARCH моделі з'являтиметься повідомлення про помилку. Для перевірки великої вибірки на відсутні спостереження дані можна відсортувати в порядку зростання або зменшення.

Для цього необхідно відкрити часовий ряд (для нашого прикладу, pfts, об'єкт series) і натиснути кнопку **Sort**. У вікні, що з'явиться, обрати ряд значень (**PFTS**) замість ряду дат (**Observation Order**) та порядок сортування. У нашому прикладі обрано порядок зростання: (**Ascending**) (див. рис. 2.6).

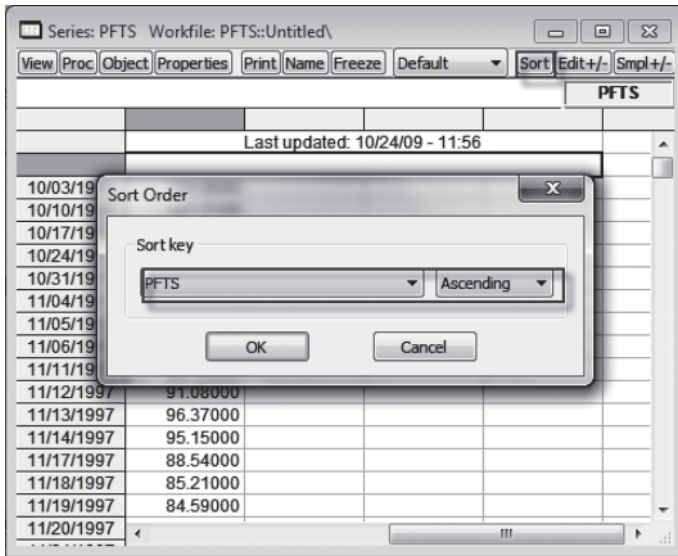


Рис. 2.6. Вікно для сортування інформації

Отримавши відсортований ряд даних, можна побачити, що відсутніх даних («NA») немає, натомість є **нульове значення (за 11/30/1999)** – неможливе. Таке нульове значення бажано замінити, наприклад, середнім арифметичним попереднього та наступного значень (для цього ряд треба знову відсортувати за порядком спостережень і замінити нульове значення на визна-

чене середнє сусідніх значень, яке в нашому прикладі дорівнює +36,28). Зауважимо, що запропонована операція заміни суттєво не вплине на результати розрахунків, особливо при великій вибірці.

Для візуального аналізу поведінки досліджуваного показника бажано побудувати його графік (приклад наведено на рис. 2.7).

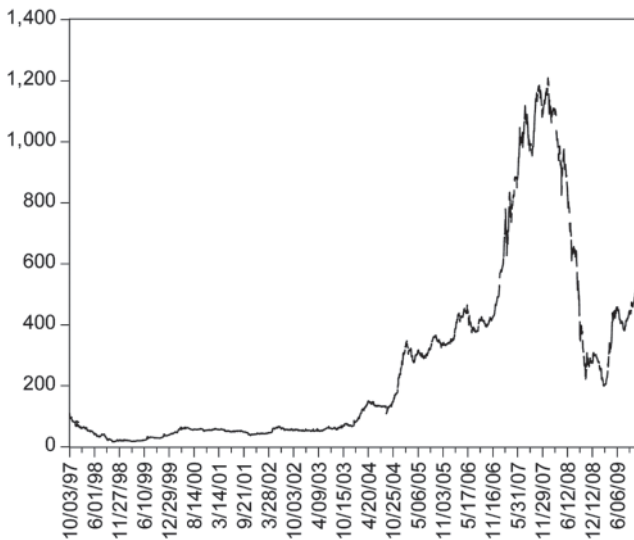


Рис. 2.7. Графік динаміки індексу ПФТС у відсотках

Перетворимо первісний ряд індексу ПФТС у ряд перших різниць логарифмів його значень ($d\log(pfts)$). Графік перетвореного ряду зображено на рис. 2.8. Візуальний аналіз показує наявність кількох зон кластеризації волатильності (зміну волатильності пучками (кластерами)).

Рисунок 2.8 відображає одну з типових ситуацій наявності ARCH/GARCH ефекту в часовому ряді, а саме кластеризацію волатильності, що свідчить про гетероскедастичність умовної дисперсії. Для прикладу на рисункові 2.9 відображено згенерований ряд випадкових нормально розподілених величин зі сталою дисперсією.

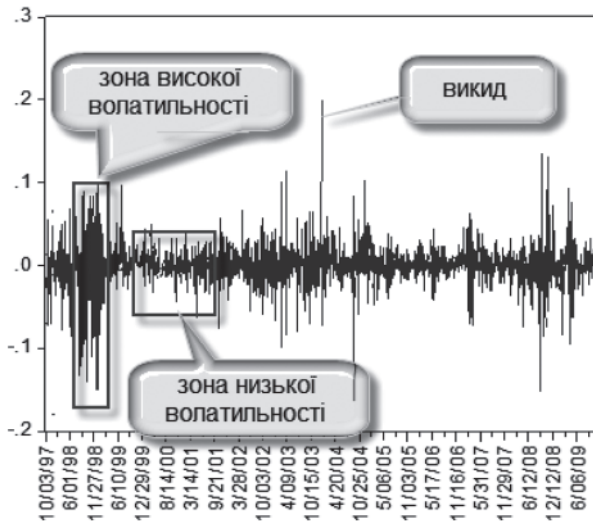


Рис. 2.8. Ряд перших різниць логарифмів значень індексу ПФТС ($D(\log(PFTS))$)

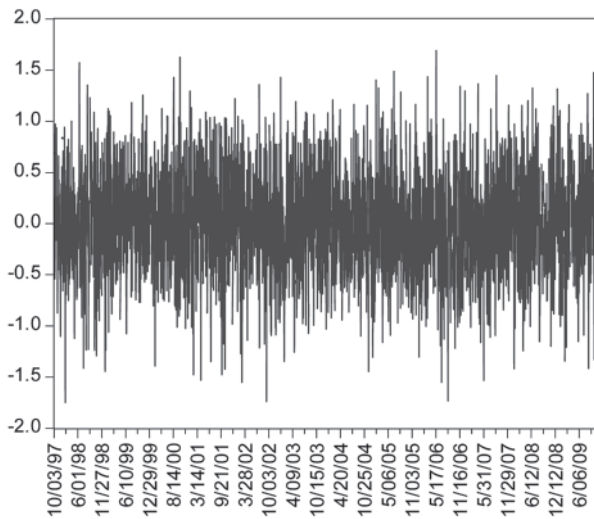


Рис. 2.9. Графік ряду згенерованих нормально розподілених випадкових величин

За наявності ARCH/GARCH ефекту в часовому ряді необхідно змодельовати вимір волатильності показника (наприклад, дисперсію), а не лише його середнє значення, тому застосовувати ARIMA методологію в цьому випадку некоректно.

2.3. У чому полягає сутність ARCH/GARCH моделювання?

Загалом, ARCH/GARCH методологію можна охарактеризувати як методологію моделювання дисперсії досліджуваного показника. Оскільки дисперсія – момент другого порядку, модель дисперсії нелінійна, отже, її не можна оцінювати методами, розробленими для лінійних моделей, зокрема за допомогою ARIMA моделей.

Абревіатура ARCH перекладається як авторегресійна умовна гетероскедастичність, тобто зміна в часі дисперсії випадкових величин (збурень), які формалізовано можна представити таким чином:

$$a_{t+1} = y_{t+1} - \mu_y(y_1, y_2, y_3 \dots y_t), \quad (2.2)$$

де a_{t+1} – збурення (випадкова величина) в $(t+1)$ період часу, y_{t+1} – значення показника в $(t+1)$ період часу; $\mu_y(y_1, y_2, y_3 \dots y_t)$ – середнє значення показника, оцінене на основі даних, які передують моменту $(t + 1)$.

Зауважимо, що абревіатура GARCH перекладається як узагальнена авторегресійна модель з умовною гетероскедастичністю. У GARCH моделях на відміну від ARCH моделей у рівнянні дисперсії враховано, крім лагових змінних випадкових величин, ще й лагові змінні умовної дисперсії.

Оскільки ARCH/GARCH вважаються моделями дисперсії, середнє значення показника можна моделювати на основі різних підходів, зокрема, простим середнім, лінійною регресійною моделлю, ARIMA, SARIMA моделями тощо. Власне загальну ARCH/GARCH модель ми можемо представити як послідовність фільтрів (див. рис. 2.10).

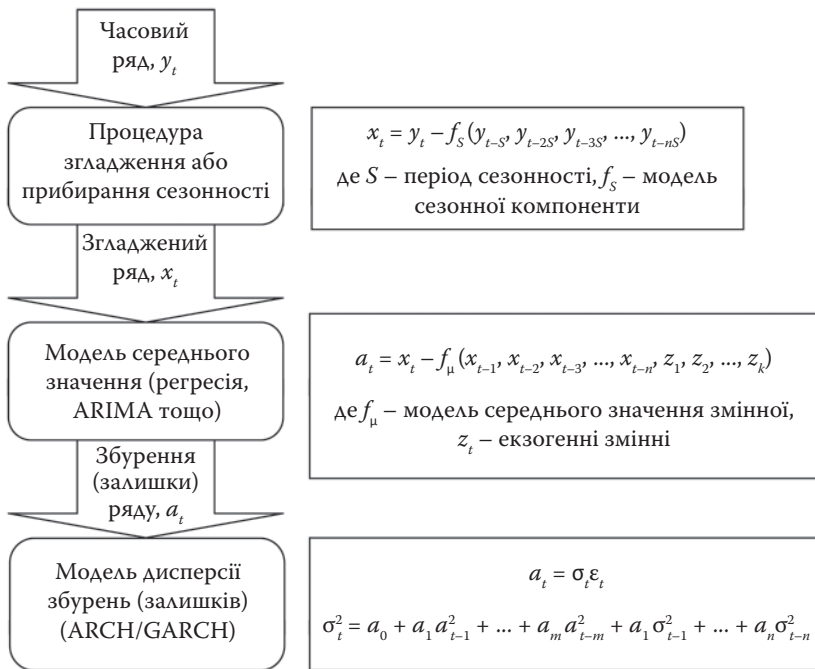


Рис. 2.10. Узагальнений алгоритм послідовної реалізації ARCH/GARCH моделі

Зауважимо, що процедури згладжування або десеզонування часового ряду не обов'язкові та виконуються за необхідності. Загалом, ARCH/GARCH модель можна вважати «доповненням» як до одно-, багатofакторної лінійної регресії, так і до ARIMA моделей тощо. Зауважимо, однак, що оцінювання ACRH/GARCH моделей неможливо здійснити методом найменших квадратів (МНК). Для їх оцінювання розроблено спеціальні процедури, зокрема метод максимальної вірогідності тощо.

2.4. Як виявити ARCH/GARCH ефект за допомогою пакета E.Views?

Для тестування наявності ARCH/GARCH ефекту в часовому ряді необхідно протестувати залишки на білий шум. Зауважимо,

що залишки ряду можна отримати на основі оцінювання моделі середнього досліджуваного ряду (модель, у якій показник є функцією лише від константи (середнього) та випадкової величини) з наступним розрахунком залишків. Залишки регресії, які мають бути білим шумом (мати математичне сподівання нуль та стали безумовну дисперсію), можуть містити структури умовної гетероскедастичності, що свідчить про наявність ARCH/GARCH ефекту. А отже, і про необхідність застосовувати для такого ряду або ARCH, або GARCH. Зауважимо, що останні найбільш розповсюджені в прикладних дослідженнях останніх років. Для виявлення таких ефектів у пакеті E.Views існує спеціальний тест – ARCH-тест.

Щоб проілюструвати використання цього тесту на практиці, скористаємося рядом перших різниць логарифмів індексу ПФТС. За загальною схемою спочатку проаналізуємо корелограм ряду, зображеного на рис. 2.11.

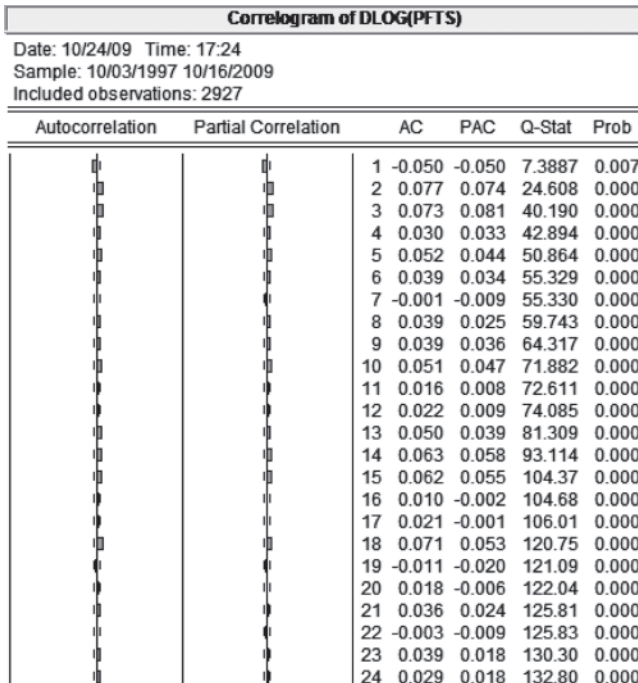


Рис. 2.11. Корелограм ряду різниць логарифмів індексу ПФТС

Візуальний аналіз графіка показує відсутність статистично значущих коефіцієнтів автокореляції цього ряду. Крім того, їхні значення майже не затухають зі збільшенням кількості лагів. Це дає змогу зробити попередній висновок про можливу наявність ARCH/GARCH ефекту. Для точніших висновків, потрібно протестувати залишки цього ряду, отримані, наприклад, з простої регресійної моделі за допомогою стандартних тестів на білий шум. Зауважимо, що залишки для тестування на білий шум можна отримати із більш складних моделей математичного сподівання (середнього) аналізованого показника. Наприклад, для отримання залишків побудуємо ARIMA модель на перших різницях логарифмів значень цього індексу ПФТС. Зокрема оцінимо * ARMA модель такого загального вигляду (зазначимо, що етапи побудови та оцінювання наведеної ARIMA моделі опускаємо, оскільки вони стандартні й розглядалися раніше):

```
dlog(pfts) c ar(1) ar(2) ar(3) ar(5) ar(10) ar(4) ar(6)
ar(8) ar(9) ma(1) ma(5) ma(8) (2.3).
```

Результати розрахунків за моделлю (2.3) наведено в таблиці 2.1.

Таблиця 2.1. Результати розрахунків DLOG(PFTS) на основі моделі (2.3)

Dependent Variable: DLOG(PFTS)				
Method: Least Squares				
Sample (adjusted): 11/14/1997 10/16/2009				
Included observations: 2917 after adjustments				
Convergence achieved after 12 iterations				
MA Backcast: 10/24/1997 11/13/1997				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.000898	0.000905	0.992755	0.3209
AR(1)	-0.411276	0.063008	-6.527324	0.0000
AR(2)	0.023816	0.017608	1.352565	0.1763
AR(3)	0.071114	0.019509	3.645205	0.0003

* Часто заради спрощення оцінювання будують ARCH із рівнянням середнього, представленим константою. Тобто оцінюється лінійна регресія, наприклад: $dlog(pfts) \sim c$, і залишки цієї регресії тестуються на ARCH ефект. У цьому посібнику розглядається складніший випадок.

Продовження табл. 2.1

AR(5)	0.471102	0.067336	6.996242	0.0000
AR(10)	0.032071	0.019159	1.673951	0.0942
AR(4)	0.038912	0.019404	2.005360	0.0450
AR(6)	0.075318	0.017219	4.374152	0.0000
AR(8)	0.426442	0.036653	11.63453	0.0000
AR(9)	0.056714	0.021191	2.676353	0.0075
MA(1)	0.345800	0.061415	5.630511	0.0000
MA(5)	-0.435158	0.068138	-6.386444	0.0000
MA(8)	-0.443204	0.029757	-14.89396	0.0000
R-squared	0.038085	Mean dependent var		0.000654
Adjusted R-squared	0.034110	S.D. dependent var		0.022785
S.E. of regression	0.022393	Akaike info criterion		-4.755721
Sum squared resid	1.456155	Schwarz criterion		-4.729078
Log likelihood	6949.219	Hannan-Quinn criter.		-4.746124
F-statistic	9.581477	Durbin-Watson stat		1.994186
Prob(F-statistic)	0.000000			
Inverted AR Roots	.97	.53-.60i	.53+.60i	.05-.89i
	.05+.89i	-.07+.26i	-.07-.26i	-.78+.63i
	-.78-.63i	-.86		
Inverted MA Roots	.93	.52+.63i	.52-.63i	.04+.91i
	.04-.91i	-.77-.63i	-.77+.63i	-.86

Після оцінювання моделі отримаємо ряд залишків (зауважимо, що їх можна зберегти, задавши певне бажане ім'я). Для відображення графіка залишків натиснемо кнопку **resids** (див. рис. 2.12).

Навіть візуально можна побачити наявність нетипової для білого шуму поведінки залишків. Перевіримо, чи мають залишки ARCH структуру за допомогою формальних тестів, наприклад ARCH LM тесту. Для цього в E.Views необхідно для виділеного ряду залишків виконати послідовність таких команд меню: **View>Residual tests>Heteroscedasticity tests...** У вікні, що відкривається, задати кількість лагів (див. рис. 2.13).

Зауважимо, що для виконання тесту з виявлення ARCH ефекту (наявності умовної гетероскедастичності) у залишках достатньо задати один лаг. Результати виконання тесту наведено в таблиці 2.3.

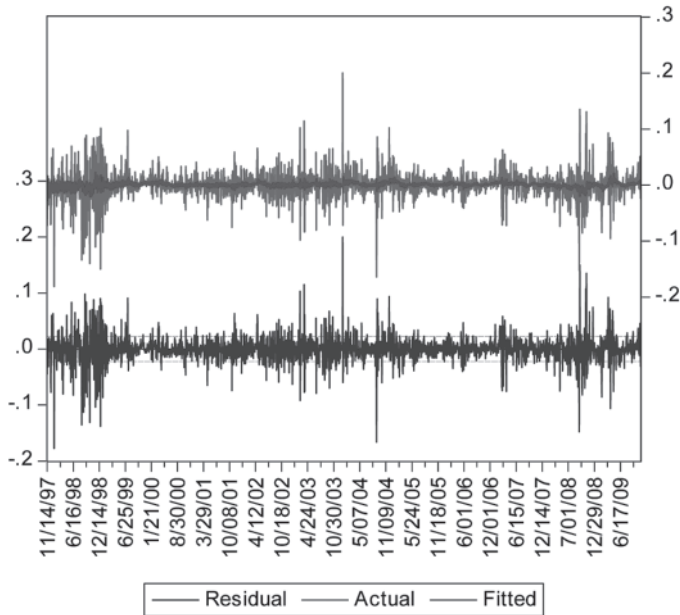


Рис. 2.12. Графік ряду залишків, отриманих після оцінювання моделі (2.3)

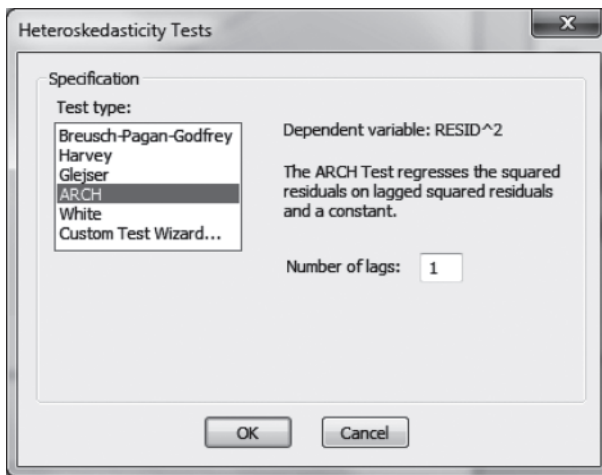


Рис. 2.13. Вікно EViews ARCH тесту

Таблиця 2.3. Результати виконання ARCH тесту в E.Views

Heteroskedasticity Test: ARCH				
F-statistic	63.09070	Prob. F(1,2914)	0.0000	
Obs*R-squared	61.79606	Prob. Chi-Square(1)	0.0000	
Test Equation:				
Dependent Variable: RESID^2				
Method: Least Squares				
Date: 10/24/09 Time: 17:41				
Sample (adjusted): 11/17/1997 10/16/2009				
Included observations: 2916 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic Prob.	
C	0.000427	3.36E-05	12.69553	0.0000
RESID^2(-1)	0.145575	0.018328	7.942966	0.0000
R-squared	0.021192	Mean dependent var	0.000499	
Adjusted R-squared	0.020856	S.D. dependent var	0.001765	
S.E. of regression	0.001746	Akaike info criterion	-9.862009	
Sum squared resid	0.008886	Schwarz criterion	-9.857909	
Log likelihood	14380.81	Hannan-Quinn criter.	-9.860532	
F-statistic	63.09070	Durbin-Watson stat	2.019658	
Prob(F-statistic)	0.000000			

Імовірність (p-value) розрахованої статистики nR^2 (Obs×R-squared) дорівнює нулю, що свідчить про можливість відкинути нуль-гіпотезу (немає ARCH ефекту) на користь альтернативної (ARCH ефект наявний). Тест, власне, показує, що квадрат залишків (збурень) – «авторегресійний» процес із лагом 1. Додатково можемо проаналізувати корелограм ряду залишків, який наведено на рис. 2.14.

Корелограм залишків схожий на корелограм білого шуму. Але чи це дійсно так? Проаналізуємо корелограм квадратів залишків, який відображено на рис. 2.15.

Як можна побачити на наведеному рис. 2.15, квадрати залишків поведуться як «авторегресійний» процес 3–4 порядку. Єдина відмінність: квадрати залишків – нелінійна функція, тому моделювати їх лінійно неоректно. Саме тому такі процеси мають окрему назву – ARCH/GARCH процеси.

Correlogram of RESID02							
Date: 10/24/09 Time: 17:44							
Sample: 10/03/1997 10/16/2009							
Included observations: 2917							
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob		
		1	0.003	0.003	0.0246	0.875	
		2	0.016	0.016	0.7467	0.688	
		3	0.013	0.013	1.2475	0.742	
		4	0.007	0.007	1.3944	0.845	
		5	-0.001	-0.001	1.3976	0.925	
		6	-0.007	-0.008	1.5519	0.956	
		7	-0.027	-0.027	3.7045	0.813	
		8	0.006	0.006	3.7955	0.875	
		9	-0.010	-0.009	4.0636	0.907	
		10	-0.019	-0.019	5.1518	0.881	
		11	-0.016	-0.016	5.9288	0.878	
		12	-0.002	-0.001	5.9427	0.919	
		13	0.015	0.015	6.5609	0.923	
		14	0.033	0.033	9.8059	0.776	
		15	0.029	0.029	12.306	0.656	
		16	-0.013	-0.016	12.839	0.684	
		17	-0.008	-0.011	13.027	0.734	
		18	0.022	0.020	14.389	0.703	
		19	-0.030	-0.030	16.999	0.590	
		20	-0.011	-0.010	17.340	0.631	
		21	0.008	0.010	17.516	0.679	
		22	-0.025	-0.024	19.348	0.624	
		23	0.004	0.004	19.389	0.678	
		24	0.010	0.013	19.660	0.716	
		25	-0.019	-0.016	20.691	0.710	
		26	0.011	0.009	21.054	0.739	
		27	0.019	0.018	22.082	0.733	
		28	0.021	0.019	23.413	0.712	
		29	0.020	0.016	24.583	0.700	
		30	0.014	0.013	25.160	0.717	
		31	-0.017	-0.018	25.994	0.722	
		32	0.037	0.034	29.948	0.571	
		33	-0.029	-0.028	32.447	0.494	
		34	-0.009	-0.006	32.675	0.533	

Рис. 2.14. Корелограм залишків моделі (2.3)

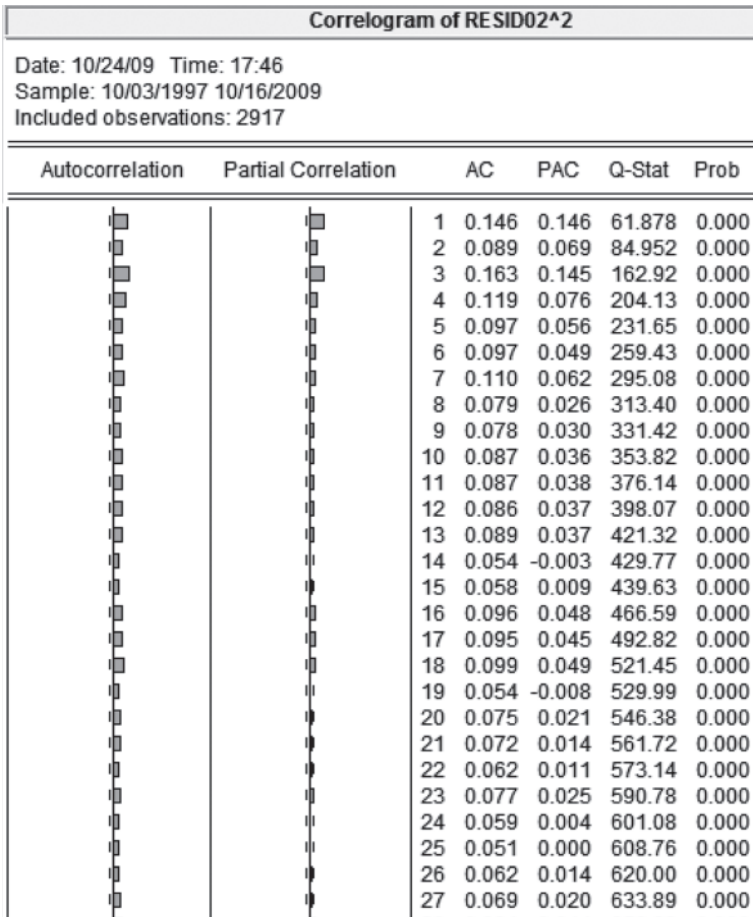


Рис. 2.15. Корелограм квадратів залишків моделі (2.3)

2.5. Як побудувати та оцінити ARCH/GARCH модель?

Виявлення ARCH/GARCH ефекту в часовому ряді перших різниць логарифмів індексів ПФТС свідчить про необхідність моделювання його динаміки на основі ARCH/GARCH методології, тобто за допомогою двох моделей: середньої (яку попередньо було змодельовано як ARIMA процес) та нелінійної моделі дисперсії.

Тож перетворимо попередньо оцінену лінійну модель середнього в ARCH/GARCH модель. Для цього слід у відкритому вікні оцінювання регресії (кнопка **Estimate об'єкту Equation**) вибрати метод оцінювання **ARCH** (див. рис. 2.16).

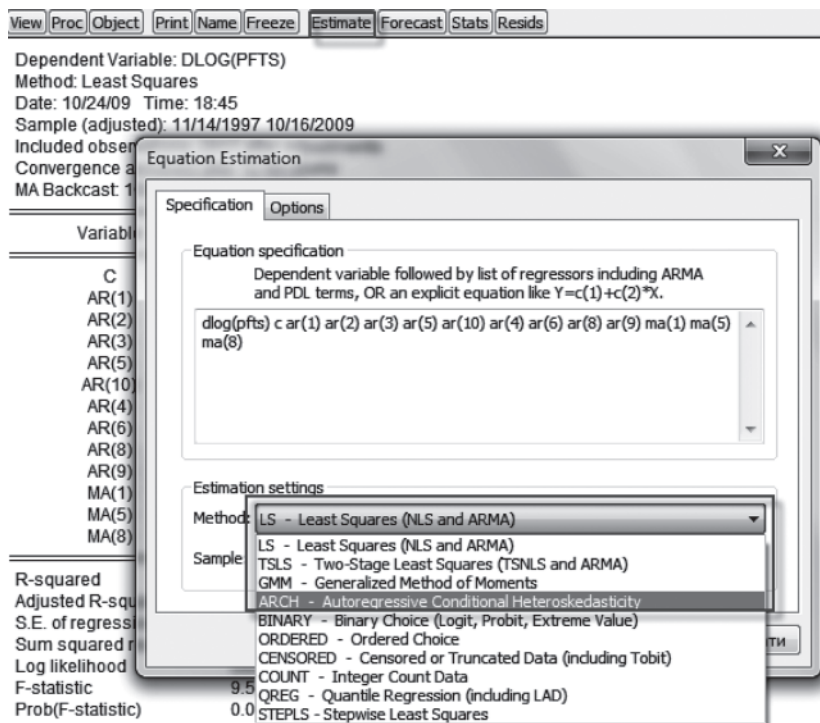


Рис. 2.16. Вікно E.Views для переоцінювання лінійної ARIMA моделі (2.3) з урахуванням гетероскедастичності дисперсії залишків (ефекту нелінійності дисперсії)

При цьому вікно оцінки регресії матиме вигляд, як на рис. 2.17.

У вікні, що відкривається, потрібно задати порядок ARCH та GARCH складових. На рис. 2.17 обраний порядок обох складових дорівнює 1. На практиці часто застосовують GARCH(1,1) моделі. Зауважимо: якби ми обрали лише ARCH складову, то мали б ARCH(1) модель, у якій дисперсія залежала тільки від лагів квадратів випадкової величини (збурення). Якщо повернутися до гра-

фіка корелограма квадратів залишків (див. рис. 2.15), то можна побачити, що статистично значущі лаги частково автокореляційної функції (PACF) наявні до 4-го періоду. Ці лаги можна співвіднести з порядком GARCH. Отже, на цьому етапі специфікуємо GARCH(4,4) модель.

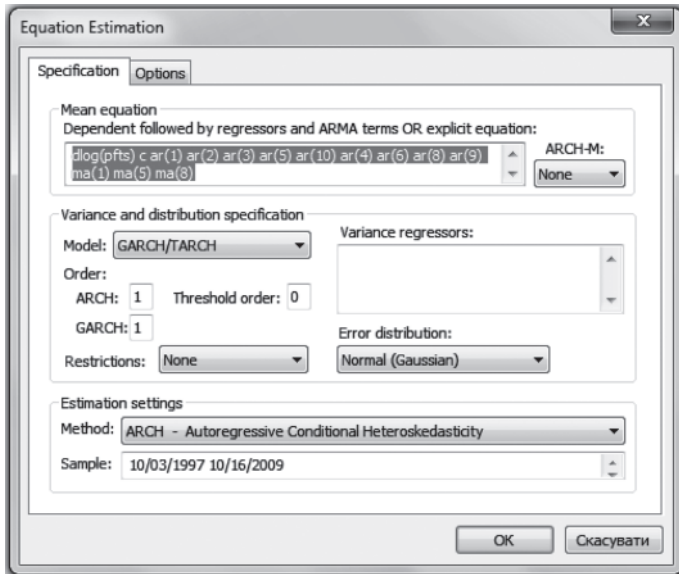


Рис. 2.17. Вікно E.Views для визначення порядку складових GARCH моделі та завдання моделі середнього

Модель 1. Припущення: дисперсія (варіація) слідує процесу GARCH(4, 4)

Будуємо GARCH(4, 4) (у дужках традиційно зазначений порядок ARCH і GARCH складових). Відповідне вікно специфікації наведено на рис. 2.18.

Процес оцінювання GARCH моделей високих порядків триває довго. За ним можна стежити в лівому нижньому куті вікна E.Views. (Порада: не потрібно під час оцінювання натискати кнопки Esc і F1, інакше оцінювання доведеться повторити заново). Варто зауважити: чим складніша модель середнього (Mean equation) та чим більша кількість ARCH/GARCH лагів, тим довше і повільніше триває

оцінювання. Результати розрахунків за GARCH(4,4) моделлю перших різниць логарифмів індексу ПФТС наведено в таблиці 2.4.

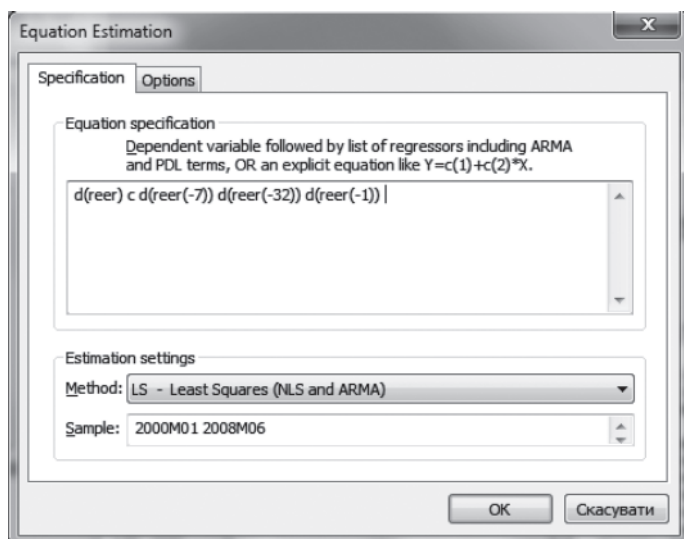


Рис. 2.18. Вікно E.Views для визначення порядку складових GARCH моделі

Таблиця 2.4. Результати розрахунків GARCH(4,4) моделі ряду $D(\log(PFTS))$

Dependent Variable: DLOG(PFTS)				
Method: ML – ARCH (Marquardt) - Normal distribution				
Sample (adjusted): 11/14/1997 10/16/2009				
Included observations: 2917 after adjustments				
Convergence achieved after 28 iterations				
MA Backcast: 10/24/1997 11/13/1997				
Presample variance: backcast (parameter = 0.7)				
GARCH = C(14) + C(15)*RESID(-1)^2 + C(16)*RESID(-2)^2 + C(17)*RESID(-3)^2 + C(18)*RESID(-4)^2 + C(19)*GARCH(-1) + C(20)*GARCH(-2) + C(21)*GARCH(-3) + C(22)*GARCH(-4)				
Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0.001475	0.000645	2.286375	0.0222

Продовження табл. 2.4

AR(1)	-0.160431	0.021605	-7.425663	0.0000
AR(2)	0.070884	0.021184	3.346101	0.0008
AR(3)	0.062338	0.011110	5.611142	0.0000
AR(5)	0.096660	0.012161	7.948383	0.0000
AR(10)	-0.024765	0.018810	-1.316545	0.1880
AR(4)	0.059017	0.011193	5.272594	0.0000
AR(6)	0.021222	0.009297	2.282711	0.0224
AR(8)	0.778746	0.009962	78.16902	0.0000
AR(9)	-0.057329	0.020596	-2.783501	0.0054
MA(1)	0.209357	0.003567	58.68677	0.0000
MA(5)	-0.053247	0.002453	-21.70762	0.0000
MA(8)	-0.807583	0.003411	-236.7849	0.0000
Variance Equation				
C	1.17E-06	4.75E-07	2.454862	0.0141
RESID(-1)^2	0.308488	0.019785	15.59208	0.0000
RESID(-2)^2	0.158253	0.166094	0.952792	0.3407
RESID(-3)^2	-0.277354	0.021232	-13.06301	0.0000
RESID(-4)^2	-0.141994	0.140531	-1.010405	0.3123
GARCH(-1)	0.265444	0.530458	0.500404	0.6168
GARCH(-2)	0.947959	0.425275	2.229047	0.0258
GARCH(-3)	-0.147083	0.283728	-0.518394	0.6042
GARCH(-4)	-0.109543	0.177623	-0.616716	0.5374
R-squared	0.014813	Mean dependent var		0.000654
Adjusted R-squared	0.007666	S.D. dependent var		0.022785
S.E. of regression	0.022697	Akaike info criterion		-5.142396
Sum squared resid	1.491385	Schwarz criterion		-5.097308
Log likelihood	7522.185	Hannan-Quinn criter.		-5.126155
F-statistic	2.072740	Durbin-Watson stat		2.217895
Prob(F-statistic)	0.002840			
Inverted AR Roots	.98	.64-.67i	.64+.67i	.22
	-.02-.97i	-.02+.97i	-.15	-.72+.68i
	-.72-.68i	-.99		
Inverted MA Roots	.96	.66-.69i	.66+.69i	-.02+.97i
	-.02-.97i	-.72-.69i	-.72+.69i	-.99

У таблиці 2.4 **Variance Equation** наведено результати оцінювання **ARCH/GARCH частини загальної моделі**. Лаги вигляду $\text{RESID}(-1)^2$ – це ARCH-лаги (лаги квадратів залишків). Їхня кількість відповідає порядку ARCH-процесу. Лаги типу $\text{GARCH}(-1)$ – GARCH-лаги (лаги оціненої дисперсії залишків), їхня кількість відповідає порядку GARCH-складової.

Зауважимо, що інтерпретації виведених значень p-value для кожного оціненого параметра й у лінійній регресії аналогічні.

Для отримання прогнозу та візуального аналізу якості відтворення моделлю динаміки середнього значення і дисперсії, виконаємо команду **Forecast**, натиснувши відповідну кнопку меню. Відкриється вікно, зображене на рис. 2.19.

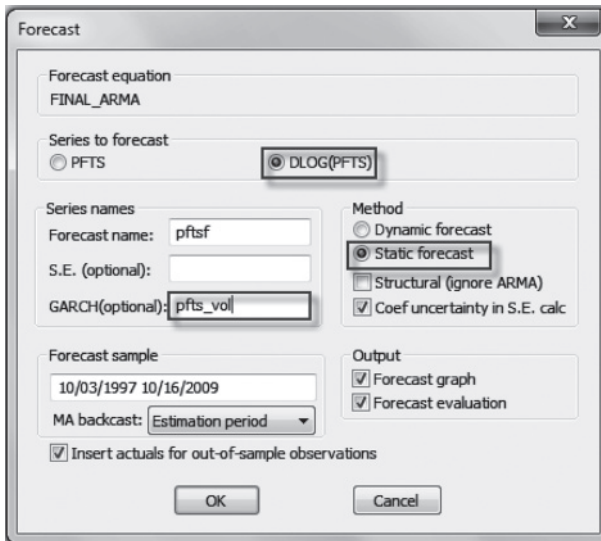


Рис. 2.19. Вікно E.Views для задавання параметрів прогнозу

У вікні, що відкриється, необхідно специфікувати, буде прогноз здійснюватися для часового ряду в рівнях, перших різницях чи інших перетвореннях. Слід використати опцію **Series to forecast**, задати ім'я прогнозних значень середнього. Послідовно застосовуємо опцію **Series names/Forecast names** та ім'я (ідентифікатор) прогнозного значення дисперсії, опцію **Series names/GARCH(optional)**.

Також потрібно задати статичну або динамічну процедуру прогнозування (**Static forecast/Dynamic forecast**) та вибірку (опція **Forecast sample**).

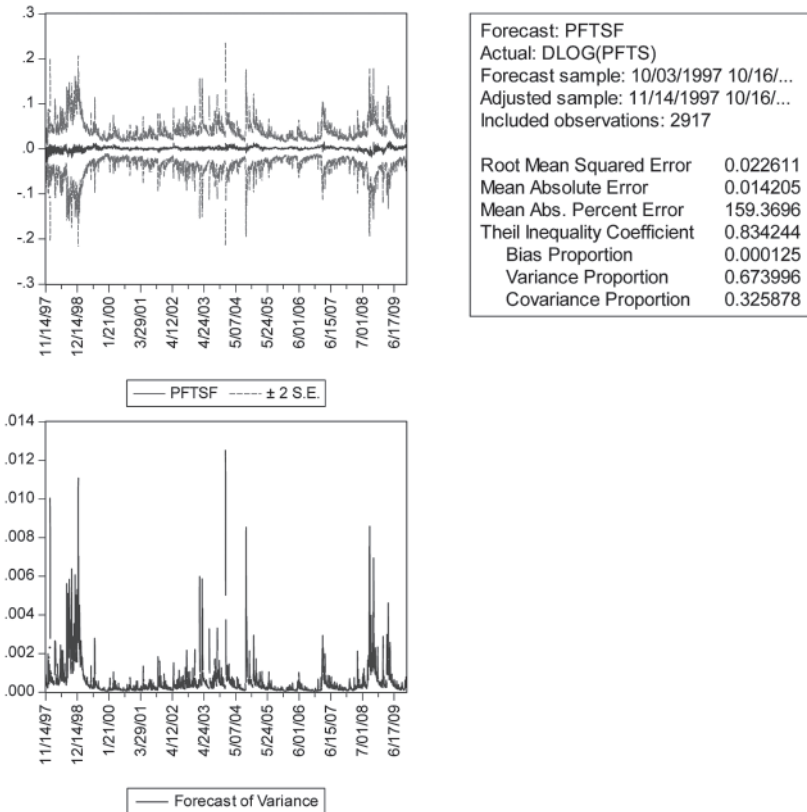


Рис. 2.20. Результати розрахунків прогнозу за рядом перших різниць логарифмів індексу ПФТС

В умовах нашого прикладу обираємо прогноз для ряду в перших різницях логарифмів індексу ПФТС, задавши опцію **DLOG(PFTS)**, тобто фактично прогнозуємо ряд у геометричних доходах. Таким чином матимемо похибки моделі, ніби під збільшувальним склом. Вибираємо опцію **Static forecast** для того, щоб прогноз спирався на реальні історичні дані попередніх періодів,

а не на попередні прогнози (Dynamic forecast). Статичний прогноз має горизонт в один період, але для ілюстративного прикладу цього достатньо. Зауважимо, що на практиці частіше використовуються функція динамічного прогнозування.

Також задамо назву ряду для оціненої дисперсії, наприклад, **pfts_vol**. Натиснувши ОК, отримаємо графічне відображення результатів прогнозування (рис. 2.20).

Якщо трансформація ряду продукує значення, які статистично несуттєво відрізняються від нуля, або інші проблеми в даних, то використовувати MAPE критерій для оцінювання якості прогнозу не варто. Краще орієнтуватися на інші критерії прогновної якості, зокрема, MAE, Theil критерій тощо (детальніше про особливості застосування критеріїв прогновної якості до моделей часових рядів див. додаток 1_В, до теми 1: ARIMA моделювання). Для усунення цієї проблеми можна також перейти до прогнозу в рівнях.

Для побудови прогнозу в рівнях при прогнозуванні замість **DLOG(PFTS)** слід обрати **PFTS**. Результати розрахунків наведено на рис. 2.21.

Зауважимо, що помилка прогнозу в цьому випадку незначна, тобто прогноз у рівнях доволі точний. Однак попередні розрахунки та відображення графіків геометричних доходів (перших різниць логарифмів) дають можливість наочніше уявити зміну дисперсії пулами.

Щоб графічно відобразити, наскільки точно оцінена дисперсія повторює реальну волатильність, необхідно створити два додаткових часових ряди. Для цього послідовно обираємо **Quick>Empty Group** і створюємо два ряди: **RESID02^2**, тобто ряд квадратів залишків ARIMA-рівняння (2.3) як оцінювання реальної дисперсії та **PFTS_VOL** – ряд, у якому на етапі прогнозування було збережено значення оціненої дисперсії. Після цього виводимо графік зміни цих двох показників (створеної групи) командою (**Line graph**). Графічне відображення оцінювання реальної та розрахованої дисперсій наведено на рис. 2.22.

Для того, щоб зберегти утворену групу, необхідно задати їй ім'я (зберігаємо її під ім'ям COMP для нашого прикладу) командою **Name**.

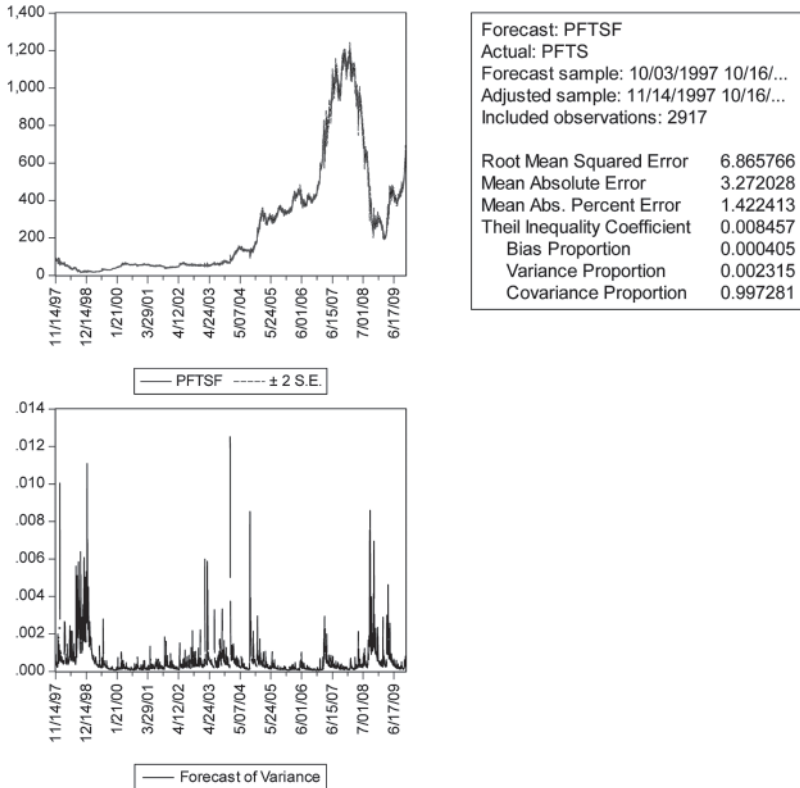


Рис. 2.21. Результати розрахунків прогнозу в рівнях за рядом індексу ПФТС

Далі проаналізуємо деякі підходи щодо можливого поліпшення отриманих результатів моделювання за ARCH/GARCH підходом для нашого прикладу.

Модель 2. Припущення: розподіл похибок – t -розподіл Стьюдента

За умови, що похибки моделі розподілені не за нормальним законом, а наприклад, за t -розподілом Стьюдента (зауважимо, що можна протестувати й інші закони розподілу за аналогічною схемою), треба переоцінити модель, обравши відповідну опцію (Student's t), як зазначено на рис. 2.23.

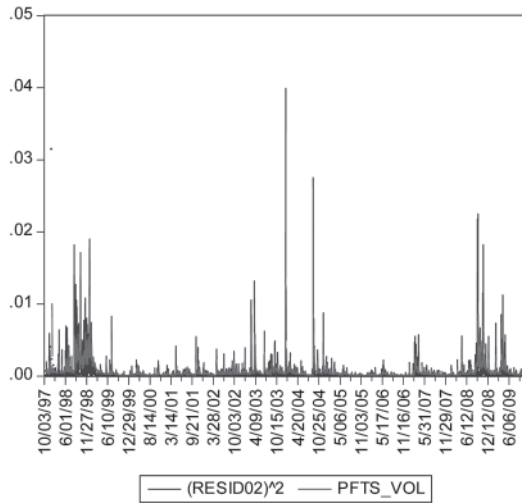


Рис. 2.22. Графік значень оцінювання реальної та прогнозової дисперсії ряду ПФТС

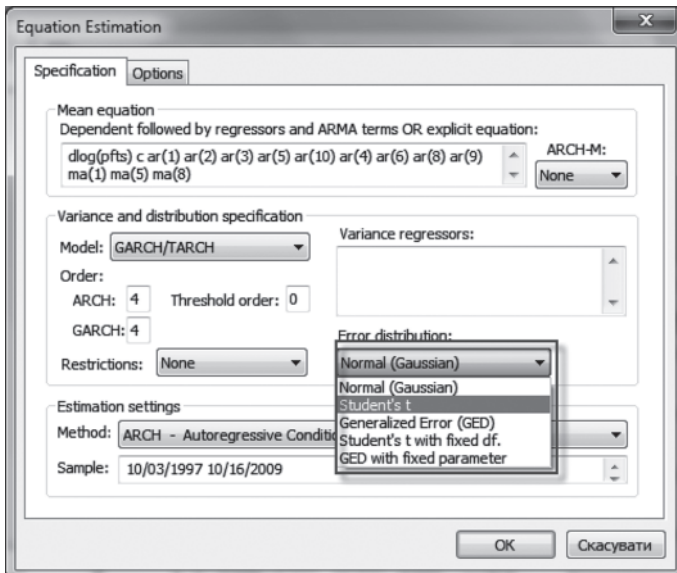


Рис. 2.23. Переоцінювання моделі GARCH (4,4) за умови, що залишки мають t -розподіл Стюдента

Натиснувши ОК, отримаємо результати переоцінювання моделі за новим припущенням (таблиця 2.5).

Таблиця 2.5. Результати переоцінювання моделі GARCH (4,4) за умови, що залишки мають не нормальний закон розподілу, а *t*-розподіл Стьюдента

Dependent Variable: DLOG(PFTS)				
Method: ML - ARCH (Marquardt) - Student's t distribution				
Sample (adjusted): 11/14/1997 10/16/2009				
Included observations: 2917 after adjustments				
Convergence achieved after 397 iterations				
MA Backcast: 10/24/1997 11/13/1997				
Presample variance: backcast (parameter = 0.7)				
GARCH = C(14) + C(15)*RESID(-1)^2 + C(16)*RESID(-2)^2 + C(17)*RESID(-3)^2 + C(18)*RESID(-4)^2 + C(19)*GARCH(-1) + C(20)*GARCH(-2) + C(21)*GARCH(-3) + C(22)*GARCH(-4)				
Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0.001060	0.000389	2.723951	0.0065
AR(1)	0.274582	0.288904	0.950426	0.3419
AR(2)	0.045824	0.016128	2.841333	0.0045
AR(3)	0.048669	0.021378	2.276624	0.0228
AR(5)	0.458084	0.146279	3.131586	0.0017
AR(10)	-0.000157	0.017023	-0.009198	0.9927
AR(4)	0.043844	0.024753	1.771240	0.0765
AR(6)	0.001045	0.022433	0.046561	0.9629
AR(8)	-0.003929	0.125636	-0.031270	0.9751
AR(9)	-0.020727	0.014793	-1.401130	0.1612
MA(1)	-0.282293	0.287948	-0.980361	0.3269
MA(5)	-0.431511	0.136790	-3.154556	0.0016
MA(8)	-0.020605	0.120884	-0.170451	0.8647
Variance Equation				
C	1.28E-06	5.22E-07	2.447743	0.0144
RESID(-1)^2	0.527776	0.078593	6.715300	0.0000

Продовження табл. 2.5

RESID(-2)^2	-1.068233	0.162166	-6.587297	0.0000
RESID(-3)^2	0.910833	0.134870	6.753389	0.0000
RESID(-4)^2	-0.343643	0.048891	-7.028822	0.0000
GARCH(-1)	2.209388	0.087833	25.15437	0.0000
GARCH(-2)	-1.961152	0.197801	-9.914783	0.0000
GARCH(-3)	0.852342	0.188806	4.514380	0.0000
GARCH(-4)	-0.122598	0.069951	-1.752621	0.0797
T-DIST. DOF	3.232113	0.223772	14.44375	0.0000
R-squared	0.022091	Mean dependent var		0.000654
Adjusted R-squared	0.014657	S.D. dependent var		0.022785
S.E. of regression	0.022617	Akaike info criterion		-5.356139
Sum squared resid	1.480367	Schwarz criterion		-5.309001
Log likelihood	7834.928	Hannan-Quinn criter.		-5.339159
F-statistic	2.971581	Durbin-Watson stat		2.118238
Prob(F-statistic)	0.000004			
Inverted AR Roots	.95	.47	.31-.79i	.31+.79i
	-.01	-.01-.46i	-.01+.46i	-.45
	-.65+.51i	-.65-.51i		
Inverted MA Roots	.92	.31+.79i	.31-.79i	.18-.32i
	.18+.32i	-.36	-.63-.51i	-.63+.51i

За результатами обчислень бачимо, що отримали кращі значення p-value для рівняння дисперсії, а також кращі значення усіх інформаційних критеріїв, F-статистики тощо. Коефіцієнт детермінації також дещо збільшився, хоча все ще залишається незначним. На рис. 2.24–2.25 відображено результати прогнозування за переоціненою моделлю.

Як можна побачити з рис. 2.25, змодельована дисперсія переоціненої моделі більш реалістично «відтворює» оцінку реальної дисперсії порівняно з попереднім випадком.

Розглянемо можливість поліпшення результатів моделювання за GARCH(4,4) моделлю, припустивши можливість наявності ефекту асиметричності.

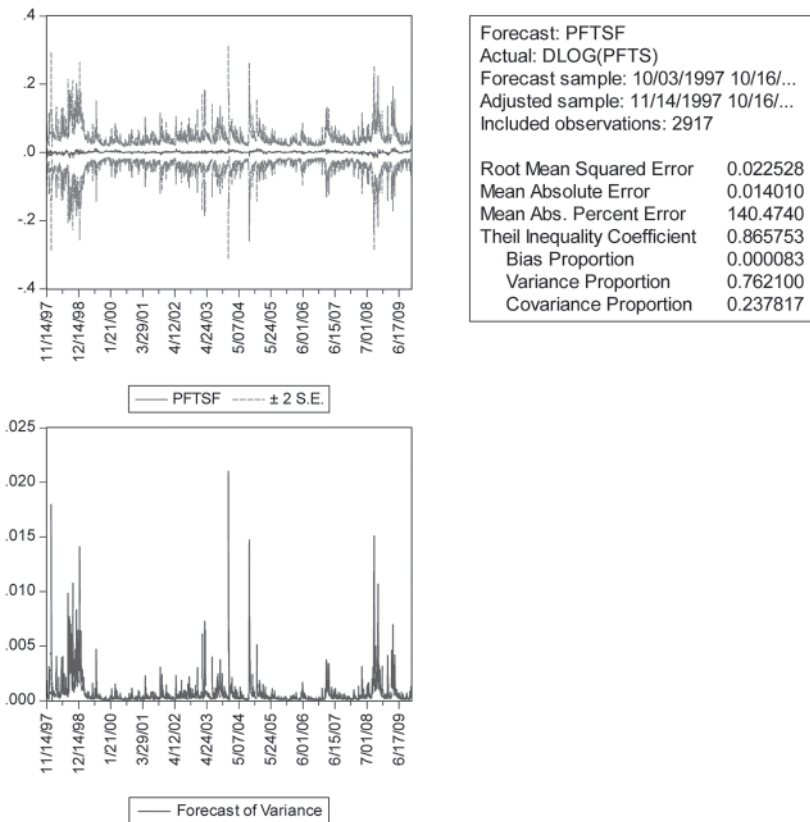


Рис. 2.24. Результати розрахунків прогнозу за рядом перших різниць логарифмів індексу ПФТС за моделлю з припущенням, що випадкові величини відповідають t -розподілу Стюдента

Модель 3. Припущення: ринок по-різному поводить у разі «хороших новин» і «поганих новин» (можлива наявність ефекту асиметричності)

Розглянемо можливість покращення результатів моделювання за GARCH(4,4) моделлю, припустивши наявність ефекту асиметричності в часовому ряді. У цьому випадку переоцінимо GARCH(4,4) модель у вигляді TARCH (Threshold ARCH) моделі, яка за побудовою враховує можливу асиметричність інформації.

Для цього потрібно повернутись у вікно **Equation Estimation** та змінити опцію **Threshold order**, задавши в ній одиницю. Відповідне вікно E.Views наведено на рис. 2.26.

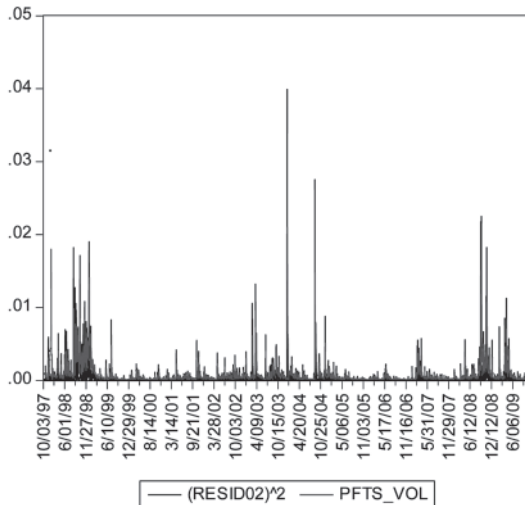


Рис. 2.25. Графік значень оцінювання реальної та прогнозової дисперсії ряду ПФТС за моделлю з припущенням, що випадкові величини відповідають t -розподілу Стюдента

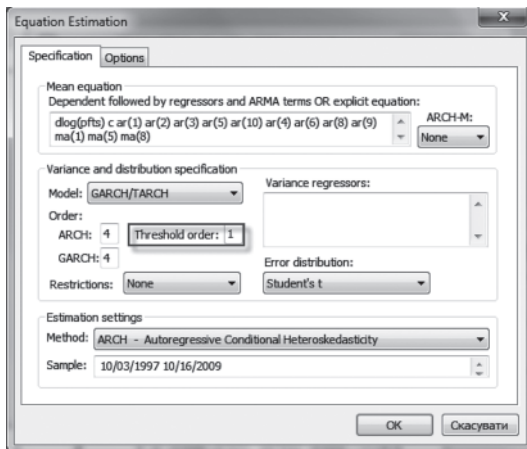


Рис. 2.26. Вікно E.Views для специфікації TGARCH моделі

Результати розрахунків за оціненою TGARCH моделлю наведені в таблиці 2.6.

Таблиця 2.6. Результати розрахунків за переоціненою TGARCH моделлю

Dependent Variable: DLOG(PFTS)				
Method: ML - ARCH (Marquardt) - Normal distribution				
Included observations: 2917 after adjustments				
Convergence achieved after 18 iterations				
MA Backcast: 10/24/1997 11/13/1997				
Presample variance: backcast (parameter = 0.7)				
GARCH = C(14) + C(15)*RESID(-1)^2 + C(16)*RESID(-1)^2*(RESID(-1)<0)				
+ C(17)*RESID(-2)^2 + C(18)*RESID(-2)^2*(RESID(-2)<0) + C(19)				
*RESID(-3)^2 + C(20)*RESID(-3)^2*(RESID(-3)<0) + C(21)*RESID(-4)^2 + C(22)*GARCH(-1) + C(23)*GARCH(-2) + C(24)*GARCH(-3) +				
C(25)*GARCH(-4)				
Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0.001400	0.000625	2.240807	0.0250
AR(1)	-0.160500	0.029224	-5.492138	0.0000
AR(2)	0.065941	0.021638	3.047455	0.0023
AR(3)	0.059753	0.012031	4.966530	0.0000
AR(5)	0.116574	0.018455	6.316752	0.0000
AR(10)	-0.026650	0.019366	-1.376083	0.1688
AR(4)	0.052995	0.012033	4.404308	0.0000
AR(6)	0.016485	0.010297	1.600892	0.1094
AR(8)	0.771280	0.013092	58.91240	0.0000
AR(9)	-0.045335	0.021205	-2.137924	0.0325
MA(1)	0.190967	0.018126	10.53573	0.0000
MA(5)	-0.082875	0.015450	-5.364211	0.0000
MA(8)	-0.798187	0.008443	-94.53522	0.0000
Variance Equation				
C	1.59E-06	3.77E-07	4.222046	0.0000
RESID(-1)^2	0.218873	0.031338	6.984208	0.0000
RESID(-1)^2*(RESID(-1)<0)	0.160040	0.041000	3.903405	0.0001
RESID(-2)^2	0.150612	0.046783	3.219398	0.0013
RESID(-2)^2*(RESID(-2)<0)	-0.412798	0.053858	-7.664592	0.0000
RESID(-3)^2	-0.199972	0.039024	-5.124356	0.0000

Продовження табл. 2.6

RESID(-3)^2*(RESID(-3)<0)	0.253351	0.035132	7.211339	0.0000
RESID(-4)^2	-0.119788	0.022040	-5.435042	0.0000
GARCH(-1)	0.939391	0.076094	12.34509	0.0000
GARCH(-2)	-0.148080	0.089332	-1.657646	0.0974
GARCH(-3)	0.356422	0.059459	5.994457	0.0000
GARCH(-4)	-0.193654	0.030454	-6.358899	0.0000
R-squared	0.016860	Mean dependent var	0.000654	
Adjusted R-squared	0.008701	S.D. dependent var	0.022785	
S.E. of regression	0.022685	Akaike info criterion	-5.148190	
Sum squared resid	1.488286	Schwarz criterion	-5.096953	
Log likelihood	7533.635	Hannan-Quinn criter.	-5.129734	
F-statistic	2.066496	Durbin-Watson stat	2.187344	
Prob(F-statistic)	0.001706			
Inverted AR Roots	.98	.64+.67i	.64+.67i	.22
	-.02-.97i	-.02+.97i	-.16	-.73+.68i
	-.73-.68i	-.99		
Inverted MA Roots	.96	.65+.69i	.65-.69i	-.01+.97i
	-.01-.97i	-.72-.69i	-.72+.69i	-.99

Як можна побачити в таблиці 2.6, змінні **RESID(-1)^2*(RESID(-1)<0)** лагові. Вони дають змогу врахувати ефект асиметричності. Зауважимо, що всі вони значущі, але не всі позитивні, що частково може вказувати на можливість ефекту асиметричності. Однак маємо погіршення інформаційних критеріїв порівняно з попередньою моделлю, а також низки інших параметрів – *F*-статистики, *R*-squared тощо. Графік значень оцінювання реальної та розрахованої дисперсії наведено на рис. 2.27.

Модель 4. Припущення: волатильність ринку має істотний вплив на встановлення середнього значення дохідності

Щоб перевірити припущення про вплив ризику на дохідність ринку, слід побудувати та оцінити ARCH-М модель. Вплив ризику в цьому класі моделей обчислюється за допомогою введення стандартного відхилення, дисперсії чи логарифму дисперсії в рівняння середнього (**Mean Equation**). У E.Views для налаштування цієї моделі потрібно повернутись у вікно **Equation Estimation** та обрати

опцію ARCH-M, додатково задавши, яке саме значення вводиться у рівняння середнього значення: дисперсія, середньоквадратичне відхилення, логарифм дисперсії тощо. Відповідне вікно E.Views наведено на рис. 2.28.

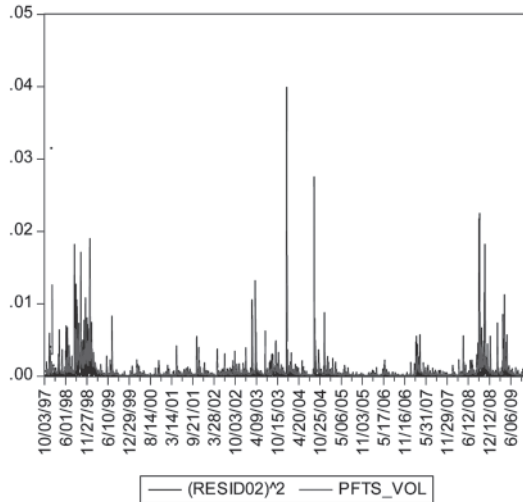


Рис. 2.27. Графік значень оцінювання реальної та прогнозової дисперсії ряду ПФТС за TGARCH моделлю

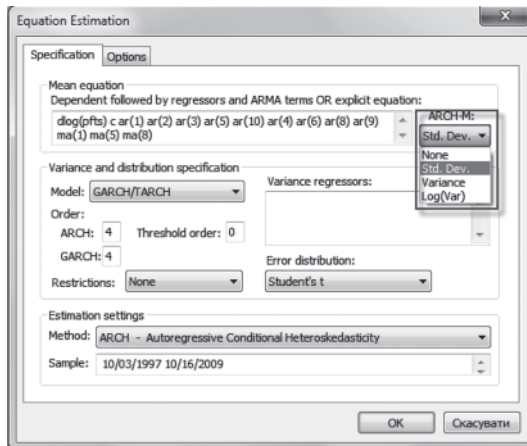


Рис. 2.28. Вікно E.Views для специфікації ARCH-M моделі

Детальний аналіз розрахунків за цією моделлю не наводимо, відмітимо лише, що суттєвого покращення результатів у цьому випадку не спостерігається.

Модель 5. Припущення: на дисперсію впливають додаткові екзогенні чинники

Зауважимо, що в рівняння дисперсії можна ввести додаткові змінні. Припустимо, що дисперсія залежить від других різниць логарифмів індексу ПФТС, обрахованих для попереднього періоду (гравці на ринку беруть до уваги «прискорення» ціни як фактор оцінки ризику). В E.Views для задання додаткових факторів у рівняння дисперсії потрібно повернутися у вікно **Equation Estimation** та ввести ім'я (ідентифікатори) відповідних змінних в опцію **Variance regressors**. Відповідне вікно E.Views наведено на рис. 2.29.

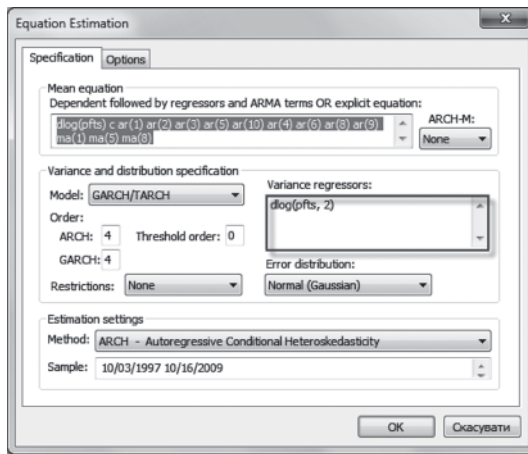


Рис. 2.29. Вікно E.Views для введення додаткових змінних у рівняння дисперсії

Детальний аналіз розрахунків за цією моделлю ми не вказуємо. Зауважимо лише, що значення інформаційних критеріїв у моделі з врахуванням додаткових змінних у рівнянні дисперсії погіршилися порівняно з іншою оціненою моделлю.

Модель 6. Припущення: наявний не квадратичний, а експоненційний ефект асиметричності

Для аналізу припущення про можливість асиметричного експоненційного ефекту слід оцінити експоненційну модель умовної гетероскедастичності EGARCH. Вона має низку теоретичних переваг, зокрема не потребує додаткових обмежень на додатність дисперсії (σ^2). Для того, щоб оцінити таку модель, потрібно повернутись у вікно **Equation Estimation**, задати в опції **Model: EGARCH**. Також змінити порядок асиметрії з 1 на 3. Крім того, для EGARCH, як правило, застосовується припущення про розподіл залишків за GED (узагальненим розподілом похибок), що необхідно позначити в опції **Error distribution**. Відповідне вікно див. на рис. 2.30.

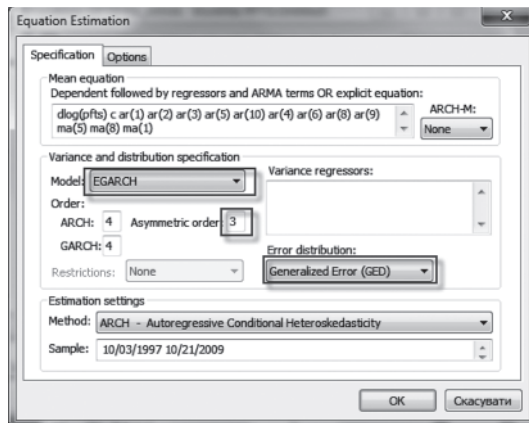


Рис. 2.30. Вікно E.Views для специфікації EGARCH моделі

Рівняння дисперсії для EGARCH моделі має такий вигляд:

$$\begin{aligned} \text{LOG}(\text{GARCH}) = & C(14) + C(15) \times \text{ABS}(\text{RESID}(-1) / @\text{SQRT}(\text{GARCH}(-1))) + \\ & C(16) \times \text{ABS}(\text{RESID}(-2) / @\text{SQRT}(\text{GARCH}(-2))) + C(17) \times \text{ABS}(\text{RESID}(-3) / \\ & @\text{SQRT}(\text{GARCH}(-3))) + C(18) \times \text{ABS}(\text{RESID}(-4) / @\text{SQRT}(\text{GARCH}(-4))) \\ & + C(19) \times \text{RESID}(-1) / @\text{SQRT}(\text{GARCH}(-1)) + C(20) \times \text{RESID}(-2) / \\ & @\text{SQRT}(\text{GARCH}(-2)) + C(21) \times \text{RESID}(-3) / @\text{SQRT}(\text{GARCH}(-3)) + \\ & C(22) \times \text{LOG}(\text{GARCH}(-1)) + C(23) \times \text{LOG}(\text{GARCH}(-2)) + \\ & C(24) \times \text{LOG}(\text{GARCH}(-3)) + C(25) \times \text{LOG}(\text{GARCH}(-4)). \end{aligned} \quad (2.4)$$

У рівнянні (2.4) доданки типу $C(n) \times \text{ABS}(\text{RESID}(-k) / \sqrt{\text{GARCH}(-k)}) = c_n \frac{|a_{t-k}|}{\sqrt{\sigma_{t-k}^2}}$ слід трактувати як ARCH-складові,

доданки типу $C(n) \times \text{LOG}(\text{GARCH}(-k)) = c_n \log(\sigma_{t-k}^2)$ – як GARCH-складові, доданки типу $C(n) \times \text{RESID}(-k) / \sqrt{\text{GARCH}(-k)} = c_n \frac{a_{t-k}}{\sqrt{\sigma_{t-k}^2}}$, які відрізняються від ARCH-складових відсутністю

модуля в чисельнику, – як складові асиметрії. Їхня кількість і максимальний лаг відповідають порядку асиметрії (три в нашому випадку).

Зауважимо, що за жодним з додаткових припущень не спостерігалося суттєвого покращення результатів порівняно з моделлю 2 (GARCH(4,4) модель з припущенням, що випадкові величини відповідають не нормальному, а t -розподілу Стьюдента).

2.6. Критерії перевірки ARCH/GARCH моделі на адекватність

Для того, щоб перевірити, чи певна ARCH/GARCH модель є коректною для застосування на практиці, необхідно з'ясувати: залишки моделі мають ARCH структуру або є білим шумом. Таку перевірку здійснюють за допомогою ARCH-LM тесту послідовності таких команд: **View>Residual Tests> ARCH LM Test**. На рис. 2.31 відображене вікно тесту з результатами для п'яти лагів (ARCH-LM тест із п'ятьма лагами).

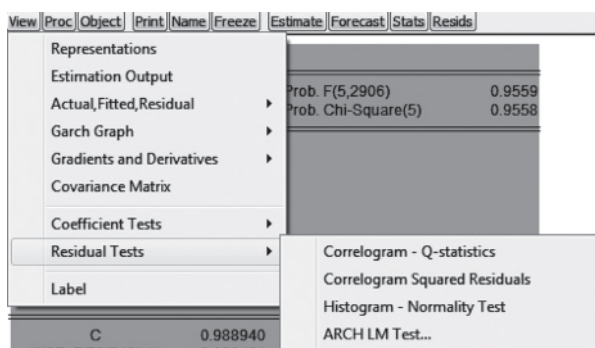


Рис. 2.31. Вікно E.Views для задання ARCH LM Test

Результати розрахунків за цим тестом наведено в таблиці 2.7.

Таблиця 2.7. Результати тестування залишків ARCH/GARCH моделі на білий шум за ARCH LM тестом

Heteroskedasticity Test: ARCH				
F-statistic	0.215733	Prob. F(5,2906)		0.9559
Obs*R-squared	1.080492	Prob. Chi-Square(5)		0.9558
Test Equation:				
Dependent Variable: WGT_RESID^2				
Method: Least Squares				
Date: 10/24/09 Time: 21:03				
Sample (adjusted): 11/21/1997 10/16/2009				
Included observations: 2912 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.988940	0.098023	10.08889	0.0000
WGT_RESID^2(-1)	-0.009421	0.018549	-0.507918	0.6115
WGT_RESID^2(-2)	-0.002417	0.018549	-0.130325	0.8963
WGT_RESID^2(-3)	-0.007844	0.018549	-0.422892	0.6724
WGT_RESID^2(-4)	-0.009507	0.018546	-0.512592	0.6083
WGT_RESID^2(-5)	-0.011545	0.018550	-0.622401	0.5337
R-squared	0.000371	Mean dependent var		0.950237
Adjusted R-squared	-0.001349	S.D. dependent var		4.827204
S.E. of regression	4.830459	Akaike info criterion		5.989818
Sum squared resid	67806.66	Schwarz criterion		6.002133
Log likelihood	-8715.175	Hannan-Quinn criter.		5.994254
F-statistic	0.215733	Durbin-Watson stat		2.000072
Prob(F-statistic)	0.955944			

Як можна побачити з результатів обчислень, отримане p-value розрахованої статистики nR^2 (Obs×R-squared) становить 0,9558. Це більше за 10 %, отже, можна зробити висновок: залишки – білий шум, тобто застосування GARCH моделі було коректним, оскільки дозволило зняти ARCH/GARCH ефект у квадратах залишків.

Для побудови та виведення графіка квадратів дисперсії та його корелограма, виділимо ряд стандартизованих залишків регресії послідовністю команд **Proc>Make residual series**. У вікні, що від-

кривається (див. рис. 2.32), обираємо **Standardized** (для стандартизованих значень залишків).

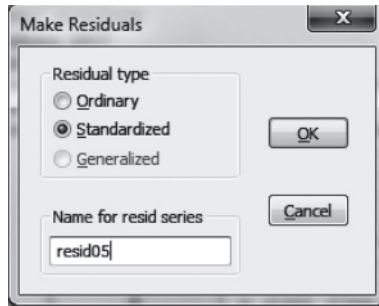


Рис. 2.32. Вікно E.Views для задання стандартизованих залишків

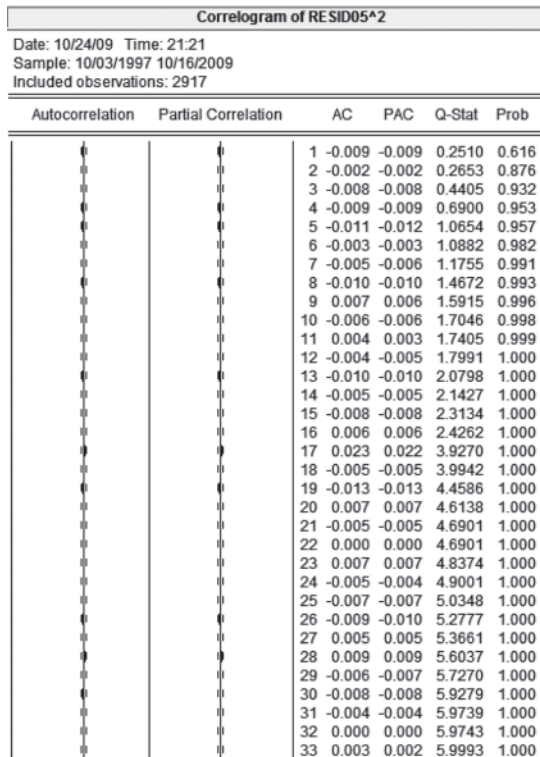


Рис. 2.33. Корелограм квадратів залишків

Щоб отримати ряд квадратів цих залишків, виконуємо команду **Quick>Empty Group**. Одну з колонок групи називаємо **resid05^2**. Відповідний ряд буде автоматично обрахований. Вибираємо корелограм у рівнях (**View>Correlorgam(1)**), відповідний графік наведено на рис. 2.33.

Як можна побачити на рис. 2.33, корелограм квадратів стандартизованих залишків має вигляд білого шуму.

Також можна проаналізувати графік самих залишків, наведених на рис. 2.34.

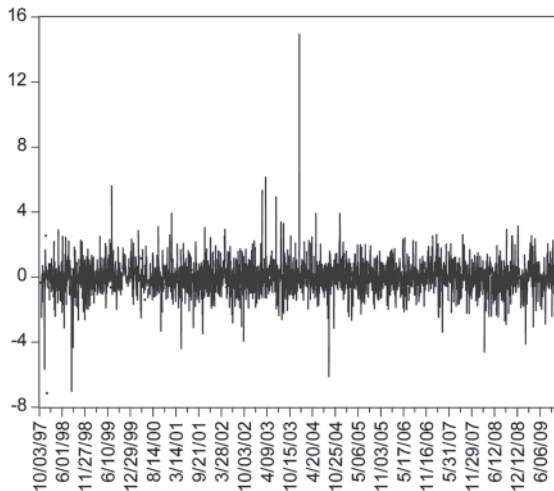


Рис. 2.34. Графік стандартизованих залишків

Як можна побачити, розкид значень досить рівномірний, немає явної кластеризації волатильності. Порівнюємо розподіл залишків із нормальним і t -розподілом відповідно. Для цього обираємо **View>Graph...** Крім того, обираємо **Quantile-Quantile** як тип графіка. Відповідне вікно зображено на рис. 2.35.

У вікні, зображеному на рис. 2.35, також обираємо **Theoretical** для того, щоб розподіл обчислених залишків можна було порівнювати з теоретично розрахованим (нормальним законом розподілу або t -розподілом Стьюдента). Потім обираємо **Options**, відкривається додаткове вікно, відображене на рис. 2.36.

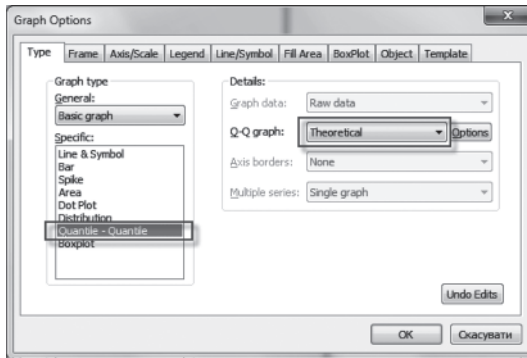


Рис. 2.35. Вікно E.Views для графічного порівняння різних типів розподілів залишків

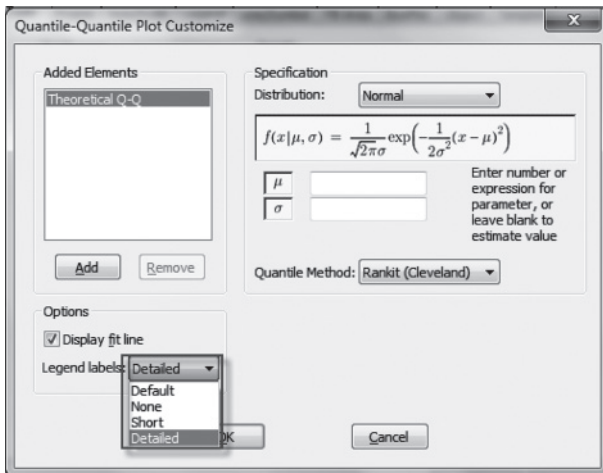


Рис. 2.36. Вікно E.Views для задання законів розподілу

Вікно, зображене на рис. 2.36, достатньо інформативне, з формулами щодо функцій щільності зазначених законів розподілу. У цьому вікні спочатку оберемо **Normal**, потім опцію **Legend Labels** змінимо на **Detailed**. Якщо ми лишаємо порожніми поля для μ і σ – параметрів розподілу, – то система сама розрахує найімовірніші їхні значення. На рис. 2.37 зображено результати побудови розподілу залишків.

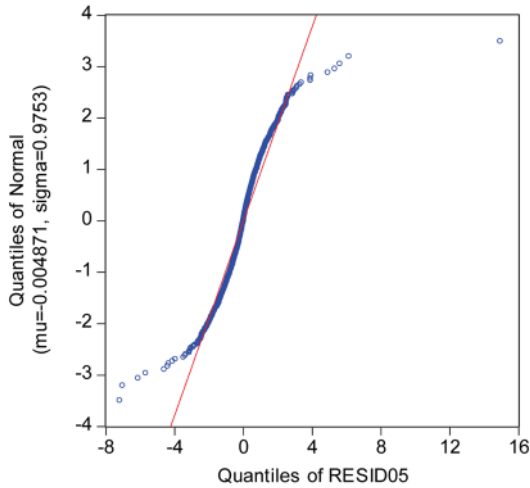


Рис. 2.37. Графік розподілу залишків порівняно з нормальним законом

Як можна побачити на графіку (див. рис. 2.37), розподіл залишків близький до нормального з середнім значенням 0,005 і стандартним відхиленням 0,98. Утім, «хвости», що не лежать на червоній лінії, мають досить багато спостережень. Спробуємо змінити теоретичний розподіл на t -розподіл Стюдента, щоб порівняти, наскільки добре розподіл залишків може бути апроксимовано t -розподілом Стюдента. Відповідне вікно наведено на рис. 2.38.

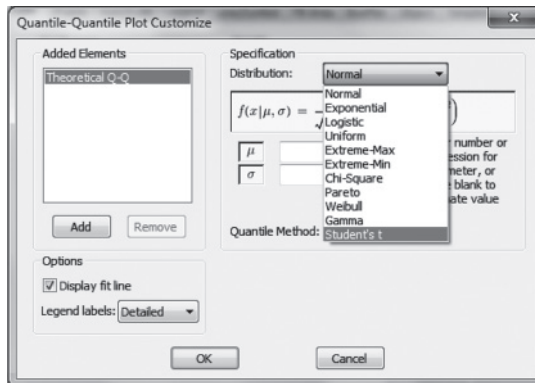


Рис. 2.38. Вікно E.Views для зміни теоретичного закону розподілу

Результати порівняння теоретичного закону t -розподілу Стюдента з законом, за яким обчислюють залишки, наведені на рис. 2.39.

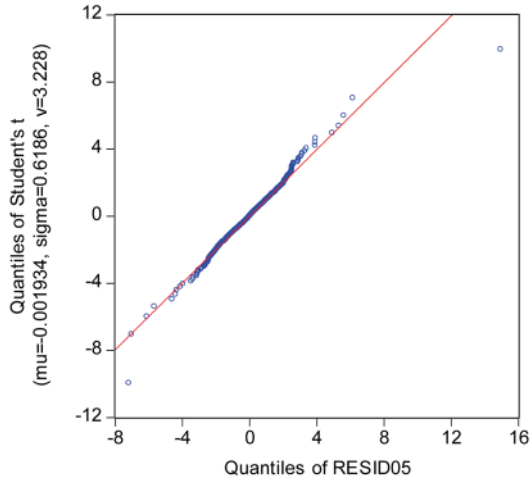


Рис. 2.39. Графік порівняння розподілу розрахованих залишків з теоретичним t -розподілом Стюдента

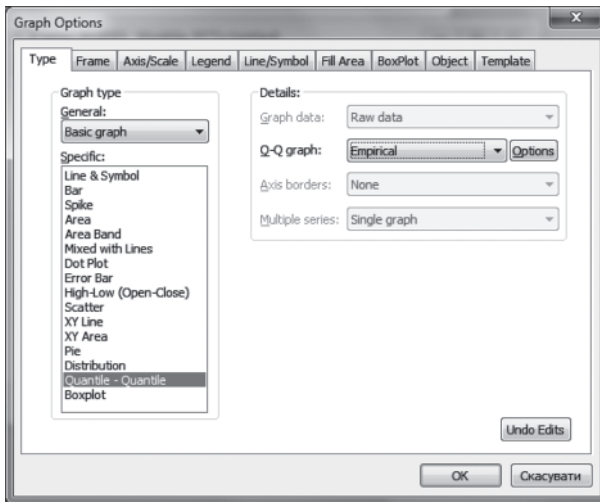


Рис. 2.40. Вікно E.Views для побудови графіків квадратів залишків

Як можна побачити на графіку (див. рис. 2.39), лише за винятком кількох спостережень, емпіричний розподіл залишків збігається з t -розподілом за параметрами: середнє значення 0,002, стандартне відхилення 0,62 і кількість ступенів свободи 3,228.

За допомогою такого потужного інструмента, як графік квантилей, можемо також порівняти, наскільки точно відтворено розподіл дисперсії. Для цього виділимо раніше створену групу **COMP** з рядами RESID02^2 і PFTS_VOL та побудуємо для них графік Quantile-Quantile. Відповідне вікно зображено на рис. 2.40.

У полі Q-Q graph вікна, зображеного на рис. 2.40, обираємо **Empirical**, щоб порівняти розподіли двох колонок групи. Результати порівняння наведено на рис. 2.41.

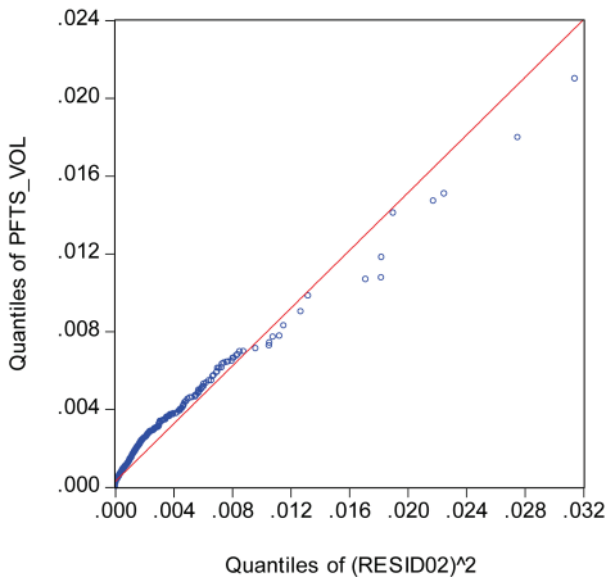


Рис. 2.41. Порівняння законів розподілу RESID02^2 та PFTS_VOL

Як можна побачити на графіку, тільки незначна кількість викидів, найбільших значень дисперсії, перебуває далеко від червоної лінії. Відповідно, можна зробити висновок, що GARCH модель досить якісно відтворила розподіл дисперсії.

2.7. Приклад практичного застосування результатів ARCH/GARCH моделей

Можливість використання результатів ARCH/GARCH моделювання на практиці розглянемо на прикладі гравця, який торгує цінними паперами, прив'язаними до індексу ПФТС. Для того, щоб отримати прибуток, він, наприклад, має розробити стратегію і відповісти на такі запитання:

- коли купити цінні папери так, щоб ціна їх в наступному періоді зросла?
- Наскільки ризикованою буде така операція?
- У яких межах ціна знаходитиметься наступного дня в 60 випадках зі 100 (шанс програти в 20 % – це схильність нашого гравця до ризику програшу)?

Припустимо, гравець знає методологію GARCH моделювання. Щодня він робить прогноз на наступний день за розглянутими GARCH моделями. Уявимо, що, наприклад, наприкінці дня 16 жовтня 2009 року, він зробив би такі дії:

- збільшив свій робочий файл на одне спостереження:

Proc>Structure/Resize current page (див. рис. 2.42).

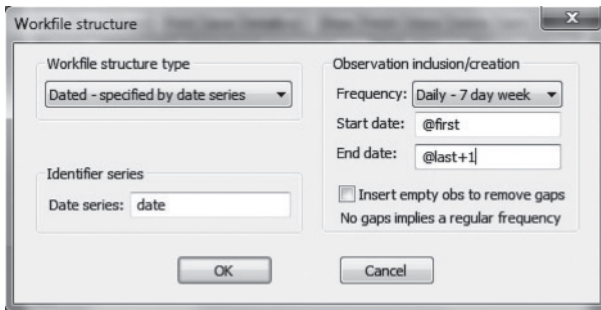


Рис. 2.42. Вікно E.Views для зміни вибірки

Наступною датою має бути понеділок, 19 жовтня. Коригуємо відповідно ряд date.

Припустимо, що далі гравець провів би процедуру прогнозування за розробленою моделлю так, щоб отримати прогноз у додане спостереження, зберігши стандартні похибки. Він врахував би, що

статичний прогноз не дасть значень стандартної похибки, тому обрав би опцію динамічного прогнозу, але скоротив діапазон прогнозу до одного дня (19 жовтня). Відповідне вікно зображене на рис. 2.43.

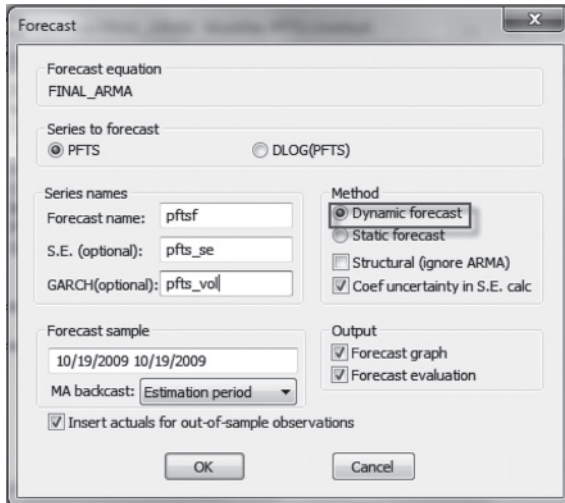


Рис. 2.43. Вікно E.Views для прогнозування

Після здійснених розрахунків він отримав би результати прогнозу, наведені на рис. 2.44.

Уявимо, що далі гравець відкрив би групу з таких рядів даних:

```
PFTSF
PFTSF-0.97*PFTS_SE
PFTSF+0.97*PFTS_SE
```

Останні два ряди відображають 60 % довірчий інтервал з урахуванням змодельованої дисперсії показника. Як було проілюстровано раніше, залишки моделюються t -розподілом з приблизно трьома ступенями вільності, отже, відповідні множники для довірчого інтервалу наш гравець знайшов у таблиці t -розподілу Стьюдента. Таким чином, гравець став би на 80 % упевнений, що індекс (а отже, і ціна паперів) буде 653,40–675,68. Значення показника зростає і коливатиметься в межах середнього значення 653,54. Значення індексів та обчислені інтервали довіри наведені на рис. 2.45.

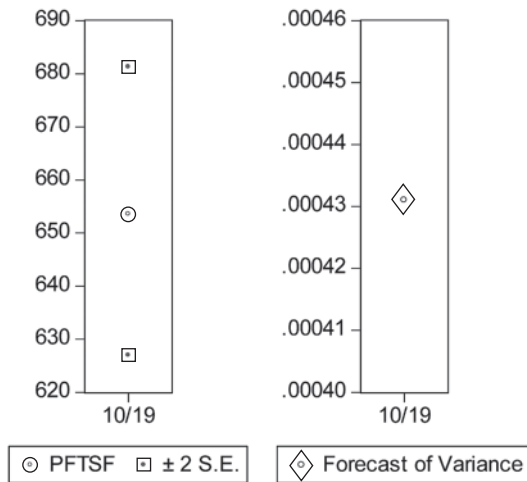


Рис. 2.44. Результати прогнозних значень, отримані гравцем за описаною процедурою

obs	PFTSF	PFTSF-0.97...	PFTSF+0.97...
10/15/2009	646.0500	NA	NA
10/16/2009	649.8500	NA	NA
10/19/2009	653.5450	640.3678	666.7221

Рис. 2.45. Результати обчислень гравця щодо зміни індексу ПФТС

Крім того, з прогнозованого ряду `pfts_vol` гравець би виявив, що ризик, який вимірюється дисперсією, імовірно, не буде таким високим, як у п'ятницю. Відповідні результати наведено на рис. 2.46.

PFTS_VOL			
10/15/2009	0.000387		
10/16/2009	0.000931		
10/19/2009	0.000431		

Рис. 2.46. Результати розрахунків значень дисперсії

Отже, якби гравець 16 жовтня купив актив, імовірність отримати програш для нього становила б (за t -розподілом) $P(y_{t+1} < y_t) = F((649,85 - 653,54)/13,58) = 0,4 = 40\%$, де y_t , y_{t+1} відповідно поточна і прогнозована ціни активу, $F(\cdot)$ – функція розподілу t . Відтак гравець поки що утримується від дій.

Перевіримо точність такого прогнозу, адже нам відомі дані за 17 жовтня: 648,99.

Гравець зазнав би втрат, якби купив актив. Він може на основі нових даних переоцінити модель і продовжити свої розрахунки. Тому додає нове спостереження в ряд *pfts* і переоцінює модель. Результати обчислень за статичним прогнозом наведено на рис. 2.47, а результати розрахунків дисперсії – на рис. 2.48.

Статичний прогноз показує падіння середнього значення.

obs	PFTSF	PFTSF-0.97...	PFTSF+0.97...
10/16/2009	651.0939	630.0041	672.1837
10/19/2009	653.5317	639.5490	667.5144
10/20/2009	651.5171	NA	NA

Рис. 2.47. Результати розрахунку гравцем значень та інтервалів довіри для індексу ПФТС

obs	RESID01*2				
10/16/2009	0.004063				
10/19/2009	0.111749				
10/20/2009	NA				

Рис. 2.48. Результати розрахунку гравцем значень дисперсії

Бачимо, що оцінка дисперсії була надто заниженою. Цілком можливо, на ринок вплинули невраховані екзогенні фактори. Зауважимо, що запропонована модель в цілому неточно відображає динаміку середнього та дисперсії.

Перевіримо, чи є історичні моменти в часі, коли можна було б зробити точний прогноз, а отже, і заробити, використовуючи раніше досліджену модель. Для цього здійснимо статичний прогноз для минулих (історичних) даних часового ряду і знайдемо спостереження, для якого нижня межа довірчого інтервалу була вищою за попереднє історичне значення (тобто коли модель радила купити актив) і коли наступне історичне значення було більшим за попереднє. Для горизонту прогнозу на один день таких випадків немає. Для того, щоб подивитися горизонт прогнозу на два дні, створимо групу з такими рядами:

$$((PFTSF(2) - 0.97 * PFTS_SE(2)) > PFTS) * (PFTS(2) - PFTS) * ((PFTS(2) - PFTS) > 0) \\ ((PFTSF(2) + 0.97 * PFTS_SE(2)) < PFTS) * (PFTS - PFTS(2)) * ((PFTS(2) - PFTS) < 0).$$

Перший ряд показує можливий виграш при підвищених цінах, другий – при знижених цінах. Побудува відповідних рядів доводить: модель дає змогу заробляти на зниженні або підвищенні ціни за два дні, якщо відомий достовірний прогноз на один день (див. результати, наведено на рис. 2.49).

obs	PFTSF	PFTSF-0.97...	PFTSF+0.97...	((PFTSF(2)-...	((PFTSF(2)+...
9/11/2009	471.3668	464.8369	477.8968	0.000000	0.000000
9/14/2009	473.4360	468.1890	478.6831	12.33000	0.000000
9/15/2009	473.6090	468.5492	478.6689	10.69000	0.000000
9/16/2009	479.5788	473.2908	485.8668	0.000000	0.000000
9/17/2009	486.0850	479.6050	492.5649	0.000000	0.000000
9/18/2009	490.5716	484.7272	496.4159	0.000000	0.000000
9/21/2009	492.6677	487.2497	498.0856	14.39000	0.000000
9/22/2009	494.2970	488.8667	499.7272	12.32000	0.000000
9/23/2009	503.6958	495.7341	511.6576	8.110000	0.000000
9/24/2009	508.7592	502.5412	514.9772	0.000000	0.000000
9/25/2009	516.4807	509.4638	523.4975	32.85000	0.000000
9/28/2009	517.3513	511.3157	523.3869	26.85000	0.000000
9/29/2009	528.4733	519.6356	537.3109	0.000000	0.000000
9/30/2009	550.1726	533.7794	566.5657	17.05000	0.000000
10/01/2009	556.5392	547.2208	565.8575	0.000000	0.000000
10/02/2009	576.6611	561.3597	591.9625	26.57000	0.000000
10/05/2009	574.6975	562.7429	586.6522	12.56000	0.000000

Рис. 2.49. Результати розрахунків прогнозів
із періодом упередження в два дні

Аналіз рисунку показує, що за досліджуваний період, використовуючи результати обчислень за моделлю, можна було б заробити 173,72 відсотки. Звичайно, можна також окремо порахувати, скільки б ми втратили, враховуючи рекомендації за результатами моделювання.

Зауважимо, що було розглянуто лише ілюстративний приклад, а не реальну фінальну модель. Для практичного використання реальні моделі слід будувати ретельно, отже, результати мають бути точнішими не лише за дисперсією, а й за значеннями математичного сподівання досліджуваного показника.



ЗАВДАННЯ, ТВОРЧІ РОБОТИ, СИТУАЦІЙНІ ВПРАВИ ДО ТЕМИ 2

Завдання для самоперевірки 1. Самостійна побудова та аналіз GARCH моделі для часового ряду показника індексу ПФТС (змінні PFTS).

Самостійно доповніть часовий ряд наданих Вам значень індексу ПФТС включно до теперішнього часу (див. Додаток 2_А до теми 2, таблиця 1А). Використовуючи значення **індексу ПФТС та надані інструкції**, повторіть самостійно всі етапи побудови GARCH моделі для цього часового ряду.

Інструкції для виконання завдання:

- а) створіть робочий файл в E.Views та імпортуйте або скопіюйте в нього дані;
- б) здійсніть первинний статистичний аналіз даних;
- в) перевірте часовий ряд на наявність GARCH ефекту (див. практичний кейс 1);
- г) специфікуйте (ідентифікуйте) GARCH модель для досліджуваного часового ряду;
- г) перевірте побудовану GARCH модель на адекватність;
- д) за побудованою моделлю розрахуйте прогноз на один-два наступних періоди. Побудуйте довірчі інтервали прогнозу та проаналізуйте прогнозу якість розробленої GARCH моделі;
- е) зробіть висновки.

Дайте відповіді на такі запитання:

1. Які основні труднощі виникли при виконанні Вами цього завдання?
2. Які основні відмінності, на Вашу думку, є між ARIMA та GARCH методологією?
3. У чому, на Ваш погляд, полягає проблема оцінювання прогнозної якості при переході від лінійних до нелінійних моделей, зокрема GARCH моделей?

4. Які основні труднощі виникли при обчисленні інтервалів довіри для прогнозу? Для якого періоду упередження Ви порекомендували б використовувати оцінену Вами GARCH модель?
5. Які основні висновки Ви можете зробити, підсумовуючи виконане вами завдання?
6. Проведіть розрахунки на обмеженому часовому ряді показника індексу ПФТС, який подано в Додатку 2_А, таблиця 1А. Порівняйте результати розрахунків. Яка різниця між ними спостерігається? Прокоментуйте.

Творча робота 1. Побудова та аналіз GARCH моделі за допомогою пакета E.Views 6.0.

Завдання. Використовуючи надані інструкції з GARCH моделювання, а також інший додатковий матеріал, виконайте творчу роботу, для цього:

- 1) знайдіть статистику (бажано щоденну, щотижневу, місячну) включно до поточного періоду часу для досліджуваного вами показника;
- 2) побудуйте GARCH моделі для обраного часового ряду, попередньо дослідивши наявність GARCH ефекту. Обґрунтуйте адекватність побудованої моделі;
- 3) за побудованою моделлю складіть прогноз на три наступні періоди. Розрахуйте критерії прогнозовної якості та побудуйте 96-відсоткові інтервали довіри. Проаналізуйте прогнозовану якість побудованої GARCH моделі;
- 4) обґрунтуйте важливість та можливість застосування GARCH інструментарію загалом та для прогнозування обраного Вами показника. Які основні труднощі виникли при виконанні самостійної творчої роботи?
- 5) підготуйте презентацію для захисту творчої роботи.

При виконанні творчої роботи в пакеті E.Views мають бути відображені послідовно такі етапи: підготовка та аналіз даних; формулювання, оцінювання і застосування моделі.

Етап перший. Підготовка та аналіз даних

Відкрийте файл(и) з даними (*.xls) для обраного показника і проаналізуйте кількість спостережень, періодичність даних (день, тиждень, місяць тощо), одиницю виміру.

Створіть робочий файл E.Views із параметрами розміру і періодичності, визначеними вище. Як ім'я файла вкажіть своє прізвище. Імпортуйте дані з файла *.xls будь-яким із доступних способів.

Очікуваний результат: створено робочий файл E.Views, у якому є об'єкти «series» із даними. Спостереження мають бути індексовані за датою спостереження.

Етап другий. Аналіз часового ряду

Виведіть графік динаміки досліджуваного показника. Проаналізуйте графік: наявність волатильності кластерами, асиметричність, нелінійність, наявність циклу і сезонності, наявність нульових значень тощо.

Протестуйте ряд на наявність GARCH ефекту.

Визначте специфікацію моделі для опису процесу.

Очікуваний результат: визначено базові характеристики часового ряду та можлива специфікація GARCH моделі або її різновидів.

Етап третій. Формулювання, оцінювання і застосування моделі

Оцініть обрану специфікацію моделі та перевірте її коректність, проаналізувавши залишки.

Проведіть тестування прогнозної якості кожної моделі та прогнозування майбутніх значень ряду терміном на два-три періоди. Вкажіть 95 % довірчий інтервал для прогнозу. Побудувати відповідний графік.

Етап четвертий. Висновки

Зробіть висновки щодо проведеної творчої роботи.

Практичний кейс 1. Тестування наявності або відсутності ARCH/GARCH ефекту в часовому ряді (залишках моделі) за допомогою BDS тесту та інструкції для застосування BDS тесту в пакеті Eviews для тестування наявності або відсутності ARCH/GARCH ефекту в часовому ряді (залишках моделі).

Вступ

BDS тест застосовується до певних змінних часового ряду або залишків певної економетричної моделі і перевіряє, чи є дослі-

джувані змінні (залишки) незалежними та однаково розподіленими величинами. У більшості випадків тест застосовується саме для перевірки залишків економетричних моделей. Дослідження показали, що тест має високу потужність (test power) для аналізу нелінійних зв'язків. Наприклад, можна перевіряти залишки лінійної ARIMA моделі на нелінійність. Слід зазначити, що для тестування потрібно використовувати «стандартизовані» залишки, тобто кінцеві, як правило – з нормальним стандартизованим розподілом. Різницю між ними проілюструємо за допомогою найпростіших AR (1) та ARCH(1) моделей, наведених у таблиці 1.

Таблиця 1. Стандартизовані та нестандартизовані залишки для AR(1) та ARCH(1) моделей для часового ряду Y

Модель	AR(1)	ARCH (1)
Задання розрахунку моделі в форматі пакета E.Views	Y C AR(1)	Y C ARCH(1)
Загальний вигляд економетричної моделі	$Y_t = a + u_t,$ $u_t = \rho u_{t-1} + \varepsilon_t,$ $\varepsilon_t \sim iid(0,1)$	$Y_t = x'_t \delta + u_t,$ $u_t = \varepsilon_t (\alpha_0 + \alpha_1 u_{t-1}^2)^{1/2},$ $\varepsilon_t \sim iid(0,1)$
«Стандартизовані» залишки	$\hat{\varepsilon}_t$	$\hat{\varepsilon}_t$
«Нестандартизовані» залишки	\hat{u}_t	\hat{u}_t

Сам тест ґрунтується на розрахунку відстані між спостереженнями і порівнянні цієї відстані з параметром epsilon.

Виконання BDS тесту в пакеті E.Views

Для ілюстрації застосування BDS тесту в пакеті E.Views розглянемо щохвилинні котирування «Алчевського металургійного комбінату» та оцінимо ARMA(9,9) модель дохідності акцій (дані наведені в Додатку 2_Б, таблиця 1Б, назва ряду даних спостережень – ALMK). Збережемо залишки як RESID_ARIMA_ALMK та протестуємо, чи вони незалежно та однаково розподілені, а поведінка дохідності акцій лінійна та може бути описана моделлю ARIMA(9,9).

Для цього відкриваємо вікно залишків і в меню **View** вибираємо **BDS Independence test...** З'являється вікно BDS тесту, наведено на рис. 1.

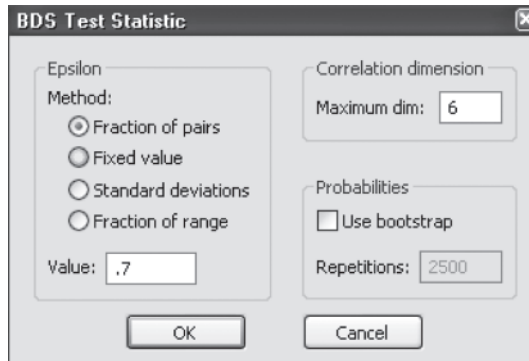


Рис. 1. Робоче вікно BDS тесту

У вікні в опції **Epsilon** можна обрати чотири варіанти розрахунку змінної ϵ , яка є вхідним параметром для оцінювання статистики тесту:

- 1) підопція **Fraction of pairs** добирає значення ϵ таким чином, що певна частка відстаней менша за ϵ ;
- 2) підопція **Fixed value** – значення ϵ фіксоване;
- 3) підопція **Standard deviations** – значення ϵ розраховується на основі стандартного відхилення значень змінної;
- 4) підопція **Fraction of range** – значення ϵ розраховується як частка від різниці між максимальним та мінімальним значеннями змінної.

Зауважимо, що найбільш інваріантним до різних типів розподілів змінної є варіант **fraction of pairs**.

У підопції **Value** потрібно вказати значення, яке використовуватиметься для розрахунку ϵ . Значення, що приймається за замовчуванням – 0,7. Воно підходить для невеликих вимірів, які задаються в опції **Correlation dimension – Maximum dim**. При збільшенні вимірності варто збільшити також **ϵ** для підвищення сили (power) тесту. Для прикладу виберемо вимір 6. Зазначимо, що при малому розмірі вибірки можна обрати бутстреп-метод для генерування розподілу статистики. Чим більше повторень (Repetitions), тим краще.

Результати виконання BDS тесту за обраними опціями наведено на рисунку 2.

Series: RESID_ARIMA_ALMK Workfile: DATA_METAL_FUND::Dat... View Proc Object Properties Print Name Freeze Sample Genr Sheet Graph Stats Ident

BDS Test for RESID_ARIMA_ALMK
Date: 11/07/11 Time: 11:05
Sample: 1 262
Included observations: 262

Dimension	BDS Statistic	Std. Error	z-Statistic	Prob.
2	0.013771	0.005028	2.738853	0.0062
3	0.026756	0.007987	3.350076	0.0008
4	0.032436	0.009505	3.412414	0.0006
5	0.037130	0.009902	3.749925	0.0002
6	0.035110	0.009543	3.679007	0.0002

Raw epsilon 0.004695
Pairs within epsilon 47907.00 V-Statistic 0.703263
Triples within epsilon 9514137. V-Statistic 0.535116

Dimension	C(m,n)	c(m,n)	C(1,n-(m-1))	c(1,n-(m-1))	c(1,n-(m-1))*k
2	17011.00	0.505227	23604.00	0.701040	0.491456
3	12347.00	0.369549	23383.00	0.699859	0.342793
4	9165.000	0.276446	23301.00	0.702832	0.244010
5	6837.000	0.207837	23099.00	0.702183	0.170707
6	5088.000	0.155882	22948.00	0.703064	0.120773

Рис. 2. Вікно з відображеними результатами виконання BDS тесту

Зауважимо, що значення ймовірності BDS статистики (p-value) для різних вимірів виділені червоним квадратом. Нульова гіпотеза: залишки (змінна) – незалежно та однаково розподілена величина. Отже, ми не приймаємо гіпотезу про лінійність дохідності акцій «Алчевського металургійного комбінату» для всіх вимірів, оскільки розраховані значення p-value менші за 0,1 (10 %).

Слід зазначити, що немає певного строгого правила для вибору epsilon та кількості вимірів. Правило, можливо, існує для певних типів моделей, але не для всіх загалом. Тому за необхідності й для глибшого розуміння тесту звертайтеся до пов'язаних із ним робіт, наприклад, Hsieh та LeBaron (1988), Hsieh (1991), Brock та ін. (1991), Brock, Hsieh і LeBaron (1991) та ін.

Практичний кейс 2. Побудова моделей ARIMA та ARCH/GARCH для портфельного аналізу дохідності акцій і визначення кращої моделі на основі BDS тесту (Програмування в командах пакета E.Views)

Великі приватні інвестиційні фонди Близького Сходу, поверхнево досліджуючи ринок України, вирішили, що одна з найприваб-

лівіших галузей – металургійна промисловість. Зокрема, на їхню думку, це підтверджується ростом капіталізації компаній – лідерів ринку – порівняно з середньоринковими показниками, очікуваним ростом попиту на їхню продукцію, наявністю значного потенціалу для зниження витрат і відповідного підвищення ефективності та впровадження кращих технологій. У зв'язку з цим близькосхідні інвестори зацікавилися формуванням інвестиційного портфеля металургійних компаній, куди буде зроблено портфельну інвестицію розміром 10 млн дол. На початковому етапі інвестори хочуть, щоб фонд формувався з двох великих компаній ринку, представлених на фондовій біржі України. На початковому етапі портфель має бути мінімально ризикованим.

Інвестор звернувся до представництва інвестиційної компанії Balclays Capital у Києві з проханням сформуванню такий портфель та здійснювати управління ним протягом шести місяців. Також інвестори просять мінімізувати витрати на управління портфелем шляхом автоматизації з використання технічного аналізу та відмовитись на цьому етапі від фундаментального аналізу.

Керівник дослідницького департаменту Balclays Capital поставив аналітику департаменту наступне завдання: протягом однієї години написати програму в E.Views для портфельного менеджера, яка в автоматичному режимі:

- 1) зчитуватиме щохвилини щохвилинні дані з файлу в форматі Excel (data_metal_fund.xls). (Дані наведені в додатку 2_Б до теми 2, таблиця 1Б, назва рядів спостережень відповідно – ALMK та AZST). На початковому етапі в базі є 263 спостереження котирувань (цін) на акції і щохвилини додаватиметься одне спостереження, яке надходитиме з автоматизованої торговельної системи української біржі. У першому рядковій міститься тікер акції, дані починаються з першого стовпчика. Інвестор обрав компанію «Алчевський металургійний комбінат» (ALMK) та компанію «Металургійний комбінат «Азовсталь»» (AZST);
- 2) будуватиме моделі ARIMA(9,9) та ARCH/GARCH(9,9) для кожної акції;
- 3) на основі BDS тесту визначатиме кращу модель (ARIMA, ARCH чи GARCH);

- 4) на основі кращої моделі для кожної акції прогнозуватиме її дохідність наступної хвилини;
- 5) визначатиме оптимальну кількість акцій, яка мінімізуватиме ризик портфеля з використанням отриманого прогнозу;
- 6) виводитиме на екран рекомендацію портфельному менеджеру для кожної акції: купувати, продавати акції чи не змінювати позиції.

Примітка. Нехай портфель P складається з двох акцій A та B , а w – вага кожної акції в портфелі. Тоді дохідність акції становитиме:

$$E(r_p) = w_A E(r_A) + w_B E(r_B),$$

а ризикованість:

$$\sigma_p^2 = (w_A \sigma_A + w_B \sigma_B)^2 = w_A^2 \sigma_A^2 + w_B^2 \sigma_B^2 + 2w_A w_B \rho_{AB} \sigma_A \sigma_B.$$

Розв'язок задачі мінімізації:

$$\min_{w_A} \sigma_p^2 = (w_A \sigma_A + w_B \sigma_B)^2 = w_A^2 \sigma_A^2 + w_B^2 \sigma_B^2 + 2w_A w_B \rho_{AB} \sigma_A \sigma_B$$

з умовою, що $w_A + w_B = 1$

$$\in w_A = \frac{\sigma_B^2 - \rho_{AB} \sigma_A \sigma_B}{\sigma_A^2 + \sigma_B^2 - 2\rho_{AB} \sigma_A \sigma_B} \text{ та } w_B = 1 - w_A.$$

Зауважимо, що BDS тест має кілька опцій, описаних у E.Views 6.0 Command Ref.pdf на сторінці 357. Теорія й посилання на першоджерело міститься в Eviews 6 User Guide I.pdf на сторінці 328.

У цій задачі використовуйте параметри тесту за замовчуванням.

За необхідності використовуйте інші припущення, які не суперечать умові завдання.

Розв'язок:

Програма та її опис подані нижче. Кожної хвилини Ви можете вводити нові значення цін на акції, відкриваючи, зберігаючи зміни та закриваючи Excel-файл.

!o=0		Визначаємо кількість хвилин, які минули
!repeat=2		Введіть кількість хвилин, які повинна працювати програма
while !o<=!repeat		Перевіряємо кількість хвилин, які працює програма
close robot_file		Закриваємо файл перед кожною наступною хвилиною
wfopen "D:\Path\data_metal_fund.xls"		Відкриваємо дані
!wealth = 10000000		Вводимо розмір інвестиційного портфеля
genr almk_r=almk/almk(-1)-1		Генеруємо дохідність акцій
genr azst_r=azst/azst(-1)-1		Генеруємо дохідність акцій
group g almk azst		Створюємо групу з акцій
for !i=1 to g.@count	Оцінювання регресій	Запускаємо цикл від одиниці до кількості акцій
%iname = g.@seriesname(!i)		Присвоюємо змінній назву акції
equation eq_garch_{{%iname}}.arch(9,9) {{%iname}}_r c		Оцінюємо ARCH/GARCH
series resid_garch_{{%iname}}=resid		Генеруємо залишки від регресії, які тестуватимемо BDS тестом на лінійність
equation eq_arima_{{%iname}}.ls {{%iname}}_r c ar(1) ar(2) ar(3) ar(4) ar(5) ar(6) ar(7) ar(8) ar(9) ma(1) ma(2) ma(3) ma(4) ma(5) ma(6) ma(7) ma(8) ma(9)		Оцінюємо ARIMA
series resid_arima_{{%iname}}=resid		Генеруємо залишки від регресії, які тестуватимемо BDS тестом на лінійність
Next		Закінчуємо цикл
pagestruct(end=@last+1)		Розширюємо розмір файлу (range) на одне спостереження, щоб у майбутньому записати прогноз цін на акції
smpl @first @last-1		Звужуємо вибірку до попередніх розмірів до розширення розміру файлу
for !i=1 to g.@count		Запускаємо цикл від одиниці до кількості акцій
%iname = g.@seriesname(!i)		Присвоюємо змінній назву акції
freeze(bds_garch_{{%iname}}) resid_garch_{{%iname}}.bdstest		Запускаємо BDS тест для залишків ARCH/GARCH та копіюємо дані в таблицю

freeze(bds_arma_{%iname}) resid_arma_{%iname}.bdstest	На основі BDS тесту визначаємо, яка модель краща та прогнозуємо ціну	Запускаємо BDS тест для залишків ARIMA та копіюємо дані в таблицю
if bds_garch_{%iname}(12,5)>bds_arma_{%iname}(12,5) then		Нульова гіпотеза BDS тесту: залишки нормально і незалежно розподілені. Чим більше р-значення, тим «краще»
smpl @last @last		Обмежуємо вибірку на останньому спостереженні файла, куди впишемо прогноз
eq_garch_{%iname}.forecast {%iname}_r_f		Прогнозуємо
smpl @first @last-1		Повертаємо розмір вибірки до попереднього значення
Else		
smpl @last @last		Обмежуємо вибірку на останньому спостереженні файла, в який впишемо прогноз
eq_arma_{%iname}.forecast {%iname}_r_f		Прогнозуємо
smpl @first @last-1		Повертаємо розмір вибірки до попереднього значення
Endif		
Next		Закінчуємо цикл
smpl @first @last		Максимально розширюємо вибірку, включаючи також прогнозне значення
!w_almk=2*(@vars(azst_r)-2*@covs(azst_r, almk_r))/(2*@vars(azst_r)+2*@vars(almk_r)-4*@covs(azst_r, almk_r))	Розрахунок кількості акцій в оптимальному портфелі	Розраховуємо ваги кожної акції
!w_azst=1-!w_almk		Розраховуємо ваги кожної акції
smpl @last @last		Обмежуємо вибірку на останньому спостереженні файла
for !i=1 to g.@count		Запускаємо цикл від одиниці до кількості акцій
%iname = g.@seriesname(!i)		Присвоюємо змінній назву акції
{%iname}={%iname}_r_f+1)*{%iname}(-1)		Розраховуємо прогнозну кількість акцій, яка повинна бути в оптимальному портфелі
Next		
!quotes_n=g.@count		Визначаємо кількість акцій
matrix (3, !quotes_n) quotes		Створюємо матрицю
for !i=1 to g.@count		Запускаємо цикл від одиниці до кількості акцій

%iname = g.@seriesname(!i)	Присвоюємо змінній назву акції
quotes(1,i) = @max({%iname})	Записуємо в матрицю прогнозну ціну акції
quotes(2,i) = !w_{%iname}*!wealth/ quotes(1,i)	Записуємо в матрицю оптимальну кількість акцій у портфелі
Next	Закінчуємо цикл
smpl @all	Максимально розширюємо вибірку
if !o=0 then	Перевіряємо, чи минула хоча б одна хвилина
table(3+!quotes_n,5) message_to_po	Створюємо таблицю
setline (message_to_po, 2)	Креслимо лінію
message_to_po (1,2) = "DEAR PORTFOLIO MANAGER,"	Записуємо звернення
for !i=1 to g.@count	Запускаємо цикл від одиниці до кількості акцій
%iname = g.@seriesname(!i)	Присвоюємо змінній назву акції
if !w_{%iname}>0 then	Перевіряємо, чи вага акції позитивна
message_to_po (!i+2,1) = "BUY "+@str(@round(quotes(2,!i)))+ STOCKS OF "+%iname	Якщо позитивна, то виводимо кількість, яку треба купити
else if !w_{%iname}<0 then	Перевіряємо, чи вага акції негативна (коротка позиція)
message_to_po (!i+2,1) = "SELL "+ @str(@abs(@round(quotes(2,!i))))+ STOCKS OF "+%iname	Якщо негативна, то виводимо кількість, яку треба продати по короткій позиції
Else	
message_to_po (!i+2,1) = "DO INCLUDE INTO PORTFOLIO STOCKS OF "+%iname	В іншому випадку не включаємо акцію в портфель
Endif	
Endif	
Next	Наступне значення циклу
delete e* a* resid_* bds_* g	Видаляємо все, крім матриці quote, де записані кількість акцій у портфелі й таблиці-повідомлення
wfsave "D:\Path\robot_file"	Зберігаємо файл
Else	Якщо минула хоча б одна хвилина, то...
pageload "D:\Path\robot_file"	Відкриваємо сторінку в попередньо збереженому файлі, яка матиме назву data_metal_fund1, тобто поточний файл матиме дві сторінки data_metal_fund та data_metal_fund1
matrix (3, !quotes_n) q_for_copy	Створюємо нову матрицю на сторінці data_metal_fund1

q_for_copy=quotes	Цій новій матриці присвоюємо значення портфеля з попередньої хвилини
copy(c=a) data_metal_fund1\q_for_copy data_metal_fund\	Копіюємо матрицю з попередніми значеннями на сторінку data_metal_fund, де поточні значення портфеля зберігаються у матриці quote
pageselect data_metal_fund	Активуємо сторінку data_metal_fund
table(3+!quotes_n,5) message_to_po	Створюємо таблицю
setline (message_to_po, 2)	Креслимо лінію
message_to_po (1,2) = "DEAR PORTFOLIO MANAGER,"	Записуємо звернення
for !i=1 to g.@count	Запускаємо цикл від одиниці до кількості акцій
%iname = g.@seriesname(!i)	Присвоюємо змінній назву акції
if q_for_copy(2,!i)-quotes(2,!i)>0 then	Перевіряємо, чи потрібно докуповувати акції
message_to_po (!i+2,1) = "BUY "+@str(@round(q_for_copy(2,!i)-quotes(2,!i)))+ " STOCKS OF "+%iname	Якщо треба докуповувати, то виводимо кількість, яку треба купити
else if q_for_copy(2,!i)-quotes(2,!i)<0 then	Перевіряємо, чи потрібно продавати акції
message_to_po (!i+2,1) = "SELL "+@str(@abs(@round(q_for_copy(2,!i)-quotes(2,!i)))+ " STOCKS OF "+%iname	Якщо потрібно продавати, то виводимо кількість, яку треба продати
Else	
message_to_po (!i+2,1) = "KEEP IN THE PORTFOLIO PREVIOUS NUMBER OF STOCKS OF "+%iname	В іншому випадку не змінюємо позиції
Endif	
Endif	
Next	
delete e* a* resid_* bds_* g q_for_copy	Видаляємо все, крім матриці quote, де записана кількість акцій у портфелі та таблиці-повідомлення
pagedelete data_metal_fund1	Видаляємо сторінку
wfsave "D:\Path\robot_file"	Зберігаємо файл

Виведення повідомлення на екран всі наступні рази

Endif		
show message_to_po		Виводимо повідомлення на екран
%time1=@time	Чекаємо нової хвилини	Визначаємо поточний час
!minute1=@val(@right(%time1,2))		Вирізаємо хвилини
!minute2=!minute1		Генеруємо новий скаляр
while !minute1=!minute2		Перевіряємо, чи пройшла хвилина
%time2=@time		Визначаємо поточний час
!minute2=@val(@right(%time2,2))		Вирізаємо хвилини
Wend		Переходимо назад до перевірки хвилин
!o=!o+1		Визначаємо, скільки хвилин пройшло
Wend		Переходимо майже на початок програми

Завдання для самоперевірки 2. Тестування наявності або відсутності GARCH ефекту в часовому ряді (залишках моделі) за допомогою BDS тесту.

Для виконання завдання використовуйте обраний часовий ряд.

- 1.1. Використовуючи інструктивні матеріали тестування наявності або відсутності GARCH ефекту в часовому ряді (залишках моделі) за допомогою BDS тесту перевірте наявність GARCH ефекту в обраному вами часовому ряді (залишках моделі) за допомогою BDS тесту.
- 1.2. Запишіть нульову та альтернативну гіпотези BDS тесту та правило прийняття рішення.
- 1.3. Опишіть детально процедуру тестування за цим тестом, а також кожну опцію, що з'являється у вікні BDS тесту.
- 1.4. Зробіть висновок щодо наявності або відсутності GARCH ефекту в обраному часовому ряді (або моделі).
- 1.5. У чому полягали основні складнощі при виконанні BDS тесту

Завдання для самоперевірки 3. Яким є правило прийняття рішення в BDS тесті? Наведіть коротке пояснення. Як Ви розумієте термін (Correlation dimension – Maximum dim) та його роль в прийнятті рішення щодо наявності або відсутності GARCH ефекту в часовому ряді (залишках моделі)?

ДОДАТКИ ДО ТЕМИ 2

Додаток 2_А. Дані щоденних значень індексу ПФТС (змінна PFTS) за період з 10.03.97 до 04.30.98

Дата	Індекс ПФТС (PFTS)	Дата	Індекс ПФТС (PFTS)	Дата	Індекс ПФТС (PFTS)
1997/10/03	100	1997/12/19	73.29	1998/03/12	65.28
1997/10/10	107.21	1997/12/22	73.2	1998/03/13	64.46
1997/10/17	102.6	1997/12/23	73.35	1998/03/16	64.76
1997/10/24	103.93	1997/12/24	74.64	1998/03/17	64.06
1997/10/31	96.88	1997/12/26	79.65	1998/03/18	63.54
1997/11/04	94.37	1997/12/29	79.74	1998/03/19	62.82
1997/11/05	93.01	1998/01/12	80.05	1998/03/20	61.86
1997/11/06	91.51	1998/01/13	80.12	1998/03/23	60.37
1997/11/11	92.19	1998/01/14	81.44	1998/03/24	58.76
1997/11/12	91.08	1998/01/15	80.34	1998/03/25	59.58
1997/11/13	96.37	1998/01/30	67.06	1998/03/26	58.84
1997/11/14	95.15	1998/02/02	67.88	1998/03/27	61.36
1997/11/17	88.54	1998/02/03	66.59	1998/03/30	62.96
1997/11/18	85.21	1998/02/04	66.59	1998/03/31	59.17
1997/11/19	84.59	1998/02/06	65.8	1998/04/01	62.2
1997/11/20	85.99	1998/02/09	65.75	1998/04/02	61.45
1997/11/21	85.48	1998/02/10	65.29	1998/04/03	61.51
1997/11/24	84.47	1998/02/11	65.33	1998/04/06	56.91
1997/11/25	81.54	1998/02/12	65.68	1998/04/07	54.59
1997/11/26	82.45	1998/02/13	65.71	1998/04/08	55.74
1997/11/27	82.12	1998/02/16	65.69	1998/04/09	53.58
1997/11/28	81.93	1998/02/17	63.27	1998/04/10	52.72
1997/12/01	81.97	1998/02/20	61.48	1998/04/13	52.51
1997/12/02	81.74	1998/02/23	61.56	1998/04/14	52.81
1997/12/03	81.02	1998/02/24	61.21	1998/04/15	52.85

Продовження додатка 2_А

Дата	Індекс ПФТС (PFTS)	Дата	Індекс ПФТС (PFTS)	Дата	Індекс ПФТС (PFTS)
1997/12/04	80.74	1998/02/25	60.8	1998/04/16	53.2
1997/12/05	80.95	1998/02/26	60	1998/04/17	52.5
1997/12/08	80.85	1998/02/27	61.03	1998/04/20	52.63
1997/12/09	80.34	1998/03/02	62.18	1998/04/21	52.74
1997/12/10	74.88	1998/03/03	62.4	1998/04/22	52.75
1997/12/11	69.86	1998/03/04	62.68	1998/04/23	52.72
1997/12/12	70.57	1998/03/05	63.13	1998/04/24	51.15
1997/12/15	68.1	1998/03/06	61.9	1998/04/27	50.38
1997/12/16	67.28	1998/03/09	63.74	1998/04/28	49.65
1997/12/17	70.64	1998/03/10	64.25	1998/04/29	49.89
1997/12/18	73.34	1998/03/11	65.31	1998/04/30	52.66

**Додаток 2_Б. Дані котирувань (дохідності акцій)
компаній «Алчевський металургійний комбінат» (ALMK)
та «Металургійний комбінат “Азовсталь”» (AZST)**

№	ALMK	AZST	№	ALMK	AZST	№	ALMK	AZST
1	0.0915	1.51	90	0.0936	1.45	178	0.0936	1.5
2	0.0924	1.464	91	0.0941	1.479	179	0.0939	1.5
3	0.093	1.475	92	0.0941	1.456	180	0.094	1.496
4	0.093	1.5	93	0.0941	1.456	181	0.0938	1.505
5	0.0931	1.48	94	0.094	1.463	182	0.0941	1.5
6	0.0937	1.47	95	0.094	1.458	183	0.0941	1.505
7	0.0932	1.465	96	0.0936	1.456	184	0.0941	1.496
8	0.0932	1.478	97	0.0939	1.447	185	0.0942	1.496
9	0.0925	1.47	98	0.0939	1.441	186	0.0947	1.5
10	0.0932	1.48	99	0.0932	1.435	187	0.0949	1.502
11	0.0928	1.484	100	0.093	1.445	188	0.095	1.495
12	0.0933	1.48	101	0.0931	1.443	189	0.095	1.505
13	0.0933	1.47	102	0.0932	1.432	190	0.0948	1.5
14	0.0934	1.47	103	0.0933	1.443	191	0.0952	1.509
15	0.0938	1.47	104	0.0936	1.444	192	0.0942	1.504
16	0.0938	1.47	105	0.0935	1.454	193	0.0944	1.519

Продовження додатка 2_Б

№	ALMK	AZST	№	ALMK	AZST	№	ALMK	AZST
17	0.0942	1.469	106	0.0935	1.458	194	0.0944	1.524
18	0.0941	1.477	107	0.0933	1.45	195	0.0938	1.525
19	0.094	1.48	108	0.093	1.457	196	0.0941	1.53
20	0.0942	1.483	109	0.0929	1.458	197	0.0935	1.53
21	0.0936	1.488	110	0.0924	1.46	198	0.0943	1.549
22	0.0933	1.477	111	0.0925	1.464	199	0.0938	1.547
23	0.0932	1.476	112	0.0928	1.464	200	0.094	1.546
24	0.093	1.48	113	0.093	1.464	201	0.0938	1.54
25	0.093	1.48	114	0.0931	1.467	202	0.0934	1.525
26	0.0932	1.472	115	0.0931	1.46	203	0.0933	1.52
27	0.0935	1.472	116	0.0927	1.455	204	0.0935	1.524
28	0.0936	1.466	117	0.0925	1.462	205	0.0931	1.525
29	0.0937	1.466	118	0.0923	1.456	206	0.0938	1.53
30	0.0935	1.466	119	0.0927	1.46	207	0.0937	1.53
32	0.0937	1.469	120	0.0929	1.468	208	0.0937	1.531
33	0.0939	1.451	121	0.093	1.464	209	0.0938	1.531
34	0.094	1.455	122	0.0928	1.455	210	0.0939	1.539
35	0.0941	1.45	123	0.0923	1.455	211	0.094	1.54
36	0.0945	1.45	124	0.0924	1.46	212	0.0942	1.543
37	0.094	1.447	125	0.0927	1.464	213	0.0943	1.543
38	0.0941	1.459	126	0.0924	1.468	214	0.0944	1.543
39	0.094	1.468	127	0.0924	1.468	215	0.094	1.543
40	0.0943	1.468	128	0.092	1.474	216	0.0941	1.53
41	0.0943	1.47	129	0.0917	1.479	217	0.0939	1.53
42	0.0943	1.47	130	0.0916	1.479	218	0.0936	1.538
43	0.0931	1.47	131	0.0919	1.479	219	0.0938	1.538
44	0.0931	1.46	132	0.0917	1.48	220	0.0935	1.53
45	0.0933	1.456	133	0.0917	1.48	221	0.0935	1.539
46	0.0935	1.455	134	0.0915	1.48	222	0.0935	1.544
47	0.0938	1.451	135	0.0921	1.48	223	0.0937	1.53
48	0.0938	1.448	136	0.0922	1.484	224	0.0965	1.53
49	0.0945	1.44	137	0.0919	1.484	225	0.0965	1.515
50	0.0938	1.433	138	0.0913	1.484	226	0.0955	1.52
51	0.0943	1.443	139	0.0913	1.485	227	0.0956	1.514
52	0.0937	1.43	140	0.0914	1.485	228	0.0957	1.502
53	0.0941	1.437	141	0.0916	1.42	229	0.0956	1.515

Закінчення додатка 2_Б

№	ALMK	AZST	№	ALMK	AZST	№	ALMK	AZST
54	0.0934	1.43	142	0.0916	1.42	230	0.0962	1.51
55	0.0932	1.428	143	0.0914	1.429	231	0.0966	1.514
56	0.0931	1.426	144	0.0914	1.44	232	0.0965	1.496
57	0.0939	1.433	145	0.0919	1.451	233	0.096	1.503
58	0.0939	1.433	146	0.092	1.46	234	0.096	1.491
59	0.0931	1.43	147	0.092	1.459	235	0.096	1.503
60	0.0931	1.44	148	0.0919	1.464	236	0.0954	1.504
61	0.0933	1.44	149	0.092	1.464	237	0.0961	1.504
62	0.0933	1.444	150	0.0921	1.46	238	0.096	1.51
63	0.0932	1.445	151	0.0921	1.474	239	0.0955	1.501
64	0.0932	1.437	152	0.0919	1.463	240	0.0957	1.508
65	0.093	1.456	153	0.0918	1.462	241	0.096	1.51
66	0.0932	1.447	154	0.092	1.474	242	0.0961	1.51
67	0.0936	1.45	155	0.0922	1.47	243	0.0955	1.51
68	0.0936	1.454	156	0.0923	1.465	244	0.0955	1.515
69	0.0935	1.446	157	0.0927	1.46	245	0.0954	1.512
70	0.093	1.447	158	0.0928	1.461	246	0.0955	1.505
71	0.0931	1.45	159	0.0925	1.47	247	0.0955	1.512
72	0.0931	1.446	160	0.0924	1.472	248	0.0955	1.512
73	0.093	1.45	161	0.0923	1.475	249	0.0955	1.505
74	0.0932	1.45	162	0.0926	1.484	250	0.0957	1.52
75	0.093	1.455	163	0.0926	1.484	251	0.0959	1.52
76	0.0933	1.455	164	0.0926	1.489	252	0.0959	1.514
77	0.0931	1.451	165	0.0925	1.491	253	0.0959	1.51
78	0.0929	1.45	166	0.0919	1.486	254	0.0958	1.5
79	0.0928	1.46	167	0.0924	1.49	255	0.0956	1.5
80	0.093	1.46	168	0.0931	1.499	256	0.0958	1.502
81	0.093	1.461	169	0.0927	1.5	257	0.0957	1.502
82	0.093	1.461	170	0.0932	1.504	258	0.0956	1.505
83	0.0931	1.453	171	0.0935	1.514	259	0.095	1.505
84	0.093	1.448	172	0.0938	1.509	260	0.0952	1.505
85	0.0931	1.444	173	0.094	1.51	261	0.0948	1.5
86	0.0931	1.453	174	0.094	1.51	262	0.095	1.5
87	0.0932	1.453	175	0.0937	1.505	263	0.095	1.499
88	0.0932	1.453	176	0.094	1.5			
89	0.093	1.453	177	0.094	1.5			

ПАМ'ЯТКИ ДО ТЕМ 1 ТА 2

Пам'ятка 1. Загальна класифікація економетричних моделей

Моделі Змінні	Статичні	Динамічні
Тільки залежна	–	ARIMA (p, d, q), ARCH, GARCH (стаціонарні дані)
Залежна та принаймні слабко екзогенні незалежні	Класична регресійна модель (при ендогенності оцінена за допомогою інструментальних змінних)	Стаціонарні дані – модель із лаговими значеннями залежної та незалежних змінних. Нестационарні, але коінтегровані дані – ECM
	Моделі з панельними даними (для кожної одиниці спостереження існує часовий ряд даних, наприклад: показники 10-ти підприємств за 5 років)	
Усі змінні ендогенні	Система одночасних рівнянь	Стаціонарні дані – VAR Нестационарні, але коінтегровані дані – VECM

Пам'ятка 2. Властивості найпростіших AR(1), AR(2), MA(1), ARMA(1,1) процесів

1. Авторегресійна модель першого порядку AR(1) та її властивості

Загальний вигляд AR(1) моделі: $AR(1): Y_t = m + \alpha Y_{t-1} + \varepsilon_t$, де Y_t, Y_{t-1} – поточне та лагове значення досліджуваного показника; m, α – невідомі параметри моделі (відповідно перетин (константа) та нахил), ε_t – випадкова величина (збурення).

1.1. Математичне сподівання AR(1) дорівнює: $E(Y_t) = \frac{m}{1-\alpha} = \mu$.
P. S. ($|\alpha| < 1$).

1.2. Дисперсія AR(1) дорівнює: $\sigma_Y^2 = E(Y_t - \mu)^2 = \frac{\sigma_\varepsilon^2}{1-\alpha^2}$.

1.3. Коефіцієнти автоковаріації AR(1):

$$\gamma_k = \alpha \gamma_{k-1} = \alpha^k \gamma_0, k=1, 2, \dots \text{ де } E(X_t X_{t-s}) = \gamma_s.$$

1.4. Коефіцієнти автокореляції AR(1): $\rho_k = \alpha \rho_{k-1} = \alpha^k, k=1, 2, \dots$, де

$$\rho_k = \frac{E(X_t X_{t-k})}{\sqrt{VAR(X_t)VAR(X_{t-k})}} = \frac{\gamma_k}{\gamma_0}. \quad (1)$$

Зауважимо, що послідовність коефіцієнтів автокореляції (1) утворює автокореляційну функцію часового ряду (autocorrelation function of the series (ACF)), яку для скорочення позначають ACF (часового ряду).

Графічне відображення автокореляційної функції називають корелограмом (correlogram) часового ряду.

Вигляд корелограма AR(1) моделі (процесу) залежить від знака параметра α .

1.5. Будь-яку AR(1) модель із перетином (константою) загального вигляду AR(1): $Y_t = m + \alpha Y_{t-1} + \varepsilon_t$ можна звести до моделі без константи: AR(1): $X_t = \alpha X_{t-1} + \varepsilon_t$ шляхом відповідної заміни змінних: $X_t = Y_t - \mu$, μ – математичне сподівання AR(1) процесу.

1.6. Зауважимо, що авторегресійні моделі різних типів визначені тільки для стаціонарних часових рядів (детальніше див. пам'ятку 3).

Умова стаціонарності AR(1) процесу: $|\alpha| < 1$

2. Авторегресійна модель другого порядку AR(2) та її властивості

Загальний вигляд AR(2) моделі: AR(2): $Y_t = m + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \varepsilon_t$.

2.1. Математичне сподівання AR(2) дорівнює: $E(Y_t) = \frac{m}{1 - \alpha_1 - \alpha_2} = \mu$.

2.2. Дисперсія AR(2) дорівнює: $\gamma_0 = \frac{(1 - \alpha_2)\sigma_\varepsilon^2}{(1 + \alpha_2)(1 - \alpha_1 - \alpha_2)(1 + \alpha_1 - \alpha_2)}$

Умова стаціонарності AR(2): $\alpha_2 + \alpha_1 < 1$; $\alpha_2 - \alpha_1 < 1$; $|\alpha_2| < 1$ – одночасно

3. Коефіцієнти автокореляції AR(2): $\rho_k = \alpha_1 \rho_{k-1} + \alpha_2 \rho_{k-2}$,
 $k = 3, 4, \dots$ з початковими умовами: $\rho_2 = \frac{\alpha_1^2}{1 - \alpha_2} + \alpha_2$, $\rho_1 = \frac{\alpha_1}{1 - \alpha_2}$.

4. Лагів поліном: $A(L) = 1 - \alpha_1 L - \alpha_2 L^2 = (1 - \lambda_1 L)(1 - \lambda_2 L)$. З його врахуванням

$AR(2): X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \varepsilon_t$ можна переписати у вигляді: $A(L)X_t = \varepsilon_t$.

Умова стаціонарності (тотожне представлення): $|\lambda_1| < 1; |\lambda_2| < 1$ (або корені характеристичного поліному $\lambda^2 - \alpha_1 \lambda - \alpha_2 = 0$ лежать всередині одиничного кола).

Р. С. Альтернативним твердженням є те, що корені полінома лагового оператора $A(Z) = 1 - \alpha_1 Z - \alpha_2 Z^2$ мають лежати поза межами одиничного кола

$$(A(Z) = 1 - \alpha_1 Z - \alpha_2 Z^2 = (1 - \lambda_1 Z)(1 - \lambda_2 Z) = 0; Z_j = \frac{1}{\lambda_j}, j = 1, 2).$$

5. Модель ковзного середнього першого порядку МА(1) та її властивості

Загальний вигляд МА(1) моделі: $X_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1}$, де X_t – точне значення досліджуваного показника, μ , θ_1 – невідомі параметри моделі (μ – константа (математичне сподівання моделі), θ_1 – параметр ковзного середнього), ε_t , ε_{t-1} – залишки моделі (точні та з лагом в один період), білий шум.

5.1. Математичне сподівання МА(1) дорівнює:

$$E(X_t) = E(\mu + \varepsilon_t + \theta_1 \varepsilon_{t-1}) = E(\mu) + E(\varepsilon_t) + \theta_1 E(\varepsilon_{t-1}) = \mu.$$

5.2. Дисперсія МА(1). Приберемо з запису константу μ , яка не впливає на дисперсію, щоб спростити перетворення. Тоді дисперсія становитиме:

$$\begin{aligned} \text{var}(X_t) &= \text{var}(\varepsilon_t + \theta_1 \varepsilon_{t-1}) = \text{var}(\varepsilon_t) + \text{var}(\theta_1 \varepsilon_{t-1}) + 2\text{cov}(\varepsilon_t, \theta_1 \varepsilon_{t-1}) = \\ &= \sigma^2 + \theta_1^2 \sigma^2 + 0 = (1 + \theta_1^2) \sigma^2. \end{aligned}$$

5.3. Коефіцієнти автоковаріації МА(1):

Автоковаріація першого порядку:

$$\begin{aligned} \gamma_1 &= E(x_t x_{t-1}) = E((\varepsilon_t + \theta \varepsilon_{t-1})(\varepsilon_{t-1} + \theta \varepsilon_{t-2})) \\ &= E(\varepsilon_t \varepsilon_{t-1}) + \theta E(\varepsilon_{t-1}^2) + \theta(\varepsilon_t \varepsilon_{t-2}) + \theta^2(\varepsilon_{t-1} \varepsilon_{t-2}) \\ &= 0 + \theta E(\varepsilon_{t-1}^2) + 0 + 0 \\ &= \theta \sigma^2. \end{aligned}$$

Автоковаріація наступних порядків:

$$\begin{aligned}\gamma_k &= E(x_t x_{t-k}) = E((\varepsilon_t + \theta \varepsilon_{t-1})(\varepsilon_{t-k} + \theta \varepsilon_{t-k-1})) \\ &= E(\varepsilon_t \varepsilon_{t-k}) + \theta E(\varepsilon_{t-1} \varepsilon_{t-k}) + \theta(\varepsilon_t \varepsilon_{t-k-1}) + \theta^2(\varepsilon_{t-1} \varepsilon_{t-k-1}) \\ &= 0 + 0 + 0 + 0 \\ &= 0.\end{aligned}$$

5.4. Коефіцієнти автокореляції MA(1):

$$\rho_1 = \frac{\gamma_1}{\text{var}(y_t)} = \frac{\theta \sigma^2}{(1 + \theta) \sigma^2} = \frac{\theta}{(1 + \theta)}, \quad \rho_k = 0,$$

ρ_1 має такий самий знак як і θ , $-1 < \theta < 1$, $-\frac{1}{2} < \rho_1 < \frac{1}{2}$.

5.5. Обернений AR(∞) процес:

Процес MA(1) можна перетворити на авторегресійний процес нескінченного порядку:

$$y_t = -\sum_{i=1}^{\infty} (-\theta)^i y_{t-i} + \varepsilon_t.$$

Цей ряд збігається (тобто обернений процес існує) за умови, якщо $|\theta| < 1$.

6. Модель авторегресії та ковзного середнього першого порядку ARMA(1,1) та її властивості

6.1. Загальний вигляд процесу ARMA(1,1):

$$z_t - \phi z_{t-1} = \mu + \varepsilon_t - \theta \varepsilon_{t-1},$$

де z_t , z_{t-1} – поточне і лагове значення досліджуваної змінної, μ – невідома константа (перетин), ϕ – невідомий параметр AR(1) складової процесу, θ – невідомий параметр MA(1) складової процесу, ε_t , ε_{t-1} – поточні і лагові значення залишків (білого шуму).

6.2. Математичне сподівання ARMA(1,1):

$$E(z_t) = E(\phi z_{t-1} + \mu + \varepsilon_t - \theta \varepsilon_{t-1}) = \frac{\mu}{1 - \phi}.$$

6.3. Дисперсія ARMA(1,1):

$$\text{var}(z_t) = \gamma_0 = \frac{\sigma^2 (1 + \theta^2 - 2\phi\theta)}{1 - \phi^2}.$$

6.4. Автоковаріація ARMA(1,1):

$$\begin{aligned}\gamma_1 &= \phi = \gamma_0 - \theta\sigma \\ \gamma_k &= \phi\gamma_{k-1}, \quad k > 1.\end{aligned}$$

6.5. Автокореляція ARMA(1,1):

$$\begin{aligned}\rho_1 &= \frac{(1 - \phi\theta)(\phi - \theta)}{1 + \theta^2 - 2\phi\theta}, \\ \rho_k &= \phi\rho_{k-1}, \quad k > 1.\end{aligned}$$

Пам'ятка 3. Поняття стаціонарності часового ряду, Дікі-Фулер тест перевірки часового ряду на стаціонарність. Визначення порядку ARIMA(p,d,q) моделі

1. Стаціонарність. Ряд $\{Y_t\}$ стаціонарний, якщо його математичне сподівання, дисперсія та попарні коефіцієнти кореляції між Y_t та Y_{t+s} не залежать від часу. При цьому попарні коефіцієнти кореляції залежать лише від лагу між ними (s).

Приклад. Розглянемо найпростішу AR(1) модель:

$$AR(1): Y_t = m + \alpha Y_{t-1} + \varepsilon_t. \quad (1)$$

Застосувавши рекурсію, її можна переписати як нескінченну суму випадкових величин:

$$AR(1): Y_t = m^* + \alpha^t Y_0 + \sum_{i=0}^{\infty} \alpha^i \varepsilon_{t-i}. \quad (2)$$

Для стаціонарних рядів $|\alpha| < 1$, для нестаціонарних (інтегрованих) рядів – $\alpha = 1$ (одиничний корінь, unit root).

Коли $\alpha = 1$, формула (2) набуває вигляду

$$Y_t = m^* + Y_0 + \sum_{i=0}^{\infty} \varepsilon_{t-i}, \quad (3)$$

і всі випадкові величини (помилки в оціненій моделі) відіграють однакову роль у поясненні поточних значень Y . Отже, Y має нескінченну дисперсію.

2. Поняття інтеграції часового ряду. Якщо перші, другі або різниці вищих порядків ряду даних стаціонарні ($I(0)$), ряд даних називається відповідно ($I(1)$), ($I(2)$)... Тобто ряд інтегрований порядку 1 або 2, або тощо. Таким чином, порядок інтеграції часового ряду – це число, яке показує кількість операцій перших різниць, яку потрібно застосувати, щоб звести нестационарний часовий ряд до стаціонарного.

Зауважимо, що ряди фінансових та макроекономічних показників, як правило, мають порядок інтеграції не вище 2.

3. Тести на стаціонарність. Розширений тест Дікі–Фулера (Augmented Dickey – Fuller (ADF) test)

Базова регресія Дікі–Фулера тесту в загальному вигляді:

$$\Delta Y_t = a_0 + a_1 t + (\theta - 1)Y_{t-1} + \sum_{i=1}^k c_i \Delta Y_{t-i} + \varepsilon_t, \quad (4)$$

$$\text{або } \Delta Y_t = a_0 + a_1 t + bY_{t-1} + \sum_{i=1}^k c_i \Delta Y_{t-i} + \varepsilon_t, \quad (5)$$

де $\Delta Y_t = Y_t - Y_{t-1}$; $b = (\theta - 1)$; a_0, a_1, b, c_i – невідомі параметри, $i = 1, 2, \dots, k$; t – часовий тренд.

Зауважимо: якщо всі $c_i = 0, i = 1, \dots, k$, то ми маємо простий Дікі–Фулера тест, у протилежному випадку Дікі–Фулера тест розширений. За правилом пальця кількість лагів, що включається в базову модель тесту, не повинна перевищувати 10 % спостережень ($k < 10\%$). У сучасних пакетах кількість лагів може визначатися автоматично.

Щодо включення перетину та тренду базову модель Дікі–Фулера тесту потрібно дотримуватися такого простого правила: якщо візуально з графіка часового ряду не можна зробити висновок про наявність тренду, то необхідно включати лише перетин, навіть якщо ряд коливається навколо нуля (в останньому випадку, на наступних етапах застосування тесту перетин виключається з базової моделі). Якщо візуально ряд має тренд, то в моделі необхідно залишити і перетин, і тренд!!!

Нульова та альтернативні гіпотези Дікі–Фулера тесту

На основі базової моделі тесту Дікі–Фулера перевіряють такі гіпотези:

$H_0 : b = 0$, або часовий ряд нестационарний: $Y_t \sim I(d), d > 0$;

$H_1 : b < 0$, або часовий ряд стаціонарний: $Y_t \sim I(d), d = 0$.

Зауважимо, тестування нульової гіпотези, що $(b = 0)$ аналогічне тому, що $\theta = 1$ в моделі (4).

Правило прийняття рішення

Підкинути нульову гіпотезу $H_0 : b = 0$, якщо $b < 0$ та розраховане значення τ – статистики більше за критичне значення статистики МакКіннона для тестування одиничного кореня при заданому рівні значущості, а

$\left| \tau \text{ stat} \right| = \left| \frac{b}{S_e(b)} \right| \geq$ – критичне значення статистики МакКіннона. У такому випадку ряд стаціонарний.

Якщо ми не можемо відкинути нульову гіпотезу, тобто $\left| \tau \text{ stat} \right| = \left| \frac{b}{S_e(b)} \right| <$ критичного значення статистики МакКіннона, то ряд нестационарний. У такому випадку потрібно визначити порядок інтеграції часового ряду і відповідними операціями перших різниць звести його до стаціонарного.

Як уже зазначали, при виконанні тесту потрібно обирати чи включати перетин, тренд та кількість лагів (від цього залежать критичні значення).

P. S. Аналогічно виконується тест Філіпс–Перона (Phillips–Perron).

Усі подібні тести мають малу потужність, тому ми порівняно рідко відхиляємо нульову гіпотезу про нестационарність. Однак краще припустити стаціонарність там, де її немає, ніж вважати інтегровані дані стаціонарними.

Тест з нульовою гіпотезою «стаціонарність» – KPSS

4. Попередня специфікація ARIMA моделей та визначення порядку чистих AR- та МА-процесів за графіком автокореляційної (ACF) та частково автокореляційної функції (PACF).

Побудова й аналіз графіків автокореляційної функції та функції часткової автокореляції дають змогу попередньо визначити, чи часовий ряд краще представити як змішаний ARIMA (ARMA) процес, чи як чисті AR-, МА-процеси. Крім того, для чистих про-

цесів аналіз корелограма та часткового автокорелограма дає можливість визначити порядок (p) для AR-процесів та (q) для MA-процесів, тобто кількість лагів, що потрібно включати у відповідні моделі. У таблиці 1 стисло схарактеризовано поведінку ACF та PACF для різних процесів авторегресії.

Таблиця 1. Властивості ACF та PACF процесів авторегресії

Процеси	Поведінка ACF	Поведінка PACF
AR(p)	Нескінченно зменшується до нуля	Скінченний: дорівнює нулю після лагу p
MA(q)	Скінченний: дорівнює нулю після лагу q	Нескінченно зменшується до нуля
ARMA(p, q)	Нескінченно зменшується до нуля	Нескінченно зменшується до нуля
Білий шум	Коефіцієнти близькі до нуля (коливаються в межах довірчого інтервалу біля нульового значення)	Коефіцієнти близькі до нуля (коливаються в межах довірчого інтервалу біля нульового значення)

Формальні додаткові підходи щодо оцінювання достатньої кількості лагів, включених у модель:

- 1) оцінити модель з лагом на один більше, ніж виходить з корелограма та перевірити значущість коефіцієнтів при останніх лагах;
- 2) мінімальне значення критерії Акайк і Шварц;
- 3) якщо модель правильно специфікована, її помилки мають бути білим шумом (перевірити корелограм помилок).

P. S. Слід пам'ятати, що нескінченний авторегресійний процес можна представити у вигляді MA(1) процесу і навпаки. Тому, якщо багато лагів Y значущі, можливо, варто замість них включити до моделі кілька (1–2) минулих помилок.

Пам'ятка 4. Генерація за допомогою генератора випадкових чисел найпростіших AR- та MA-процесів

1. Генератор випадкових нормально розподілених чисел (Normal random number generator).

Команда **nrnd** використовується для генерації випадкових нормально розподілених величин із математичним сподіванням нуль та дисперсією.

2. Генерація найпростішого AR(1) процесу за допомогою генератора випадкових чисел.

Найпростіший AR(1) процес без перетину загального вигляду $Y_t = \alpha Y_{t-1} + u_t$ можна згенерувати за допомогою такої команди:

```
AR(1) process: series y = .6*y(-1) + .5*nrnd.
```

Ця команда генерує часовий ряд, який є AR(1) процесом із початковим значенням нуль, із параметром $\alpha = 0,6$. Збурення $u = (0,5 \times \text{nrnd})$ – нормально розподілені випадкові величини з математичним сподіванням нуль та середнім квадратичним відхиленням 0,5.

3. Генерація найпростішого MA(1) процесу за допомогою генератора випадкових чисел.

Найпростіший MA(1) процес загального вигляду $U_t = \mu + \beta_0 U_{t-1} + \varepsilon_t$ можна згенерувати за допомогою, наприклад, такої команди:

```
MA(1) - процес: series U = 10 + @sqr(3)*nrnd.
```

Ця команда генерує часовий ряд, який є MA(1) процесом. Збурення – нормально розподілені випадкові величини з математичним сподіванням (середнім) 10 та дисперсією 3.

4. Команда: series x = nrnd^2+nrnd^2+nrnd^2

генерує часовий ряд x, який є сумою трьох незалежних стандартних нормально розподілених випадкових величин та має χ^2 – розподіл із трьома ступенями свободи (вільності).

Зауважимо, сума трьох часових рядів, що отримується наведеною вище командою, не аналогічна часовому ряду, який отримується командою

```
series x=3*nrnd^2,
```

оскільки остання команда включає генерацію однієї випадкової величини.

5. Команда: series x = @qchisq(rnd,3)

забезпечує альтернативний метод симуляції випадкових величин із χ^2 -розподілом.

6. Генератор рівномірно розподілених випадкових величин (uniform random number generator).

Цей генератор генерує випадкові величини, рівномірно розподілені на проміжку (0,1). Його може бути включено або в вираз для визначення часового ряду, або в певне рівняння, яке розв'язується. Наприклад, команда

```
series u=5+(12-5)*rnd
```

генерує часовий ряд *u* рівномірно розподілених величин на проміжку (5, 12).

7. Генератор рівномірно розподілених цілих випадкових величин (uniform random integer generator).

Команда `rndint` заповнює часові ряди, вектори, матриці випадковими рівномірно розподіленими цілими величинами від нуля до певного заданого максимального значення. Команда `rndint` ігнорує поточну вибірку та заповнює весь об'єкт згенерованими випадковими цілими величинами.

Загальний формат команди:

```
rndint(ім'я об'єкту, n).
```

Задається ім'я об'єкта (часового ряду, вектора, матриці) для заповнення випадковими рівномірно розподіленими цілими величинами з максимальним значенням *n*, яке має бути позитивним цілим числом.

Наприклад, для часового ряду `index` (series `index`) команда

```
rndint(index,10)
```

заповнює часовий ряд `index` випадковими цілими величинами від 0 до 10. Як і `genr`, команда `rndint` ігнорує поточні значення та заповнює часовий ряд згенерованими випадковими величинами.

Інший приклад для симетричної матриці `var` порядку 3x3 (`sym(3) var3`) – команда

```
rndint(var3,5)
```

заповнює всю матрицю `VAR3` випадковими цілими величинами від 0 до 5.

8. Тип генератора випадкових чисел (seed the random number generator).

Команда `rndseed` використовується у випадку симуляції послідовності випадкових величин, що повторюється, або при виборі генератора, що використовується.

E.Views підтримує кілька типів генераторів випадкових чисел. Зауважимо, що багато нерівномірно розподілених випадкових величин утворюються шляхом трансформації величин, отриманих з рівномірно розподілених згенерованих чисел.

Загальний формат команди: `rndseed(опції)`.

Можливі такі опції:

- a) `type=arg`. Тип випадкового генератора Кнутца (Knuth's) («`type=kn`»);
- б) Комбінований множинний рекурсивний генератор Лекуйєра (L'Ecuyer's) («`type=le`»);
- в) Матсумото та Нішимура, Мерсенна число (Mersenne Twister) («`type=mt`»).

За замовчуванням використовується перший тип генератора (Knuth lagged Fibonacci generator).

Пам'ятка 5. Тестування наявності або відсутності GARCH ефекту в часовому ряді (залишках моделі) за допомогою BDS тесту.

BDS тест дає можливість виявити наявність нелінійності, зокрема GARCH ефекту, або безпосередньо в часовому ряді досліджуваного показника, або протестувати залишки оціненої ARIMA або GARCH моделі на відсутність/наявність у залишках моделі GARCH ефекту (гетероскедастичності дисперсії залишків).

Нульова гіпотеза BDS тесту – відсутність GARCH ефекту, коли аналізовані дані відтворюють чистий шум (абсолютно випадковий процес). Наприклад: залишки моделі – білий шум.

Альтернативна – наявність GARCH ефекту: досліджувані дані можна розглядати як нелінійний процес. Наприклад, залишки моделі не є білим шумом.

Нульова гіпотеза BDS тесту: дані відтворюють чистий шум (абсолютно випадковий процес), альтернативна гіпотеза визначає нелінійні процеси.

Оскільки в загальному вигляді BDS тест використовується для тестування нелінійності, то формулювання наведених вище гіпотез цього тесту можна інтерпретувати також іншим чином. Наприклад, нульова гіпотеза BDS тесту: дані відтворюють чистий шум (абсолютно випадковий процес).

Альтернативна гіпотеза – досліджуваний процес є нелінійним процесом, детермінованим хаосом тощо.

BDS тест може застосовуватися до первинних даних часового ряду або до залишків певної моделі, наприклад, GARCH моделі, які при правильній специфікації мають бути білим шумом за BDS тестом. Розрахована BDS статистика тесту має бути статистично значущою за адекватності моделі.

Для виконання BDS тесту в E.Views (версія 4.0–6.0) необхідно виділити або досліджуваний (модельований) часовий ряд, або ряд залишків попередньо оціненої моделі. Потім слід послідовно обрати: **View/BDS/Independent test**. Вікно BDS тесту, що відкриється, зображене на рис. 1.

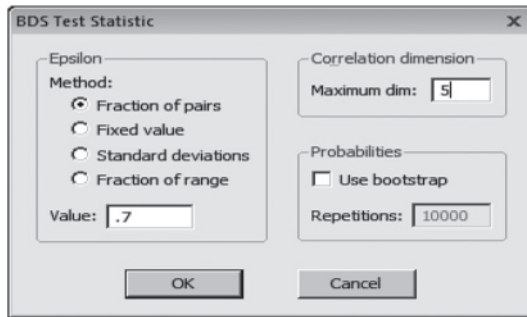


Рис. 1. Вікно BDS тесту

У вікні BDS тесту потрібно задати певні параметри, а саме:

Опція **Dimension** визначає кількість спостережень, які вводяться в тесті. Опція **Correlation dimension – Maximum dim** використовується, щоб специфікувати максимальну кількість вимірів, для яких наявний GARCH ефект. Вони вводяться у вікні, що відкрилося, і визначають порядок, для якого розраховується значення p-value (ймовірності) BDS статистики. Зауважимо, змінивши значення кількості вимірів для певного значення Epsilon Value,

ми отримаємо лише збільшення числа вимірів, котрі будуть відображені у вікні, проте p-value залишатиметься без змін. Проілюструємо цей факт відповідними змінами, наприклад, задамо послідовно такі значення: Epsilon Value = 0,7 та Epsilon Value = 0,9. Результати виконання BDS тесту для певного ряду при заданих значеннях відповідно відображені на рис. 2 та 3.

BDS Test for RESID01

Date: 11/20/11 Time: 20:28

Sample: 2006M01 2011M09

Included observations: 69

Dimension	BDS Statistic	Std. Error	z-Statistic	Prob.	
2	-0.011161	0.010590	-1.053938	0.2919	
3	-0.025394	0.017100	-1.484990	0.1375	
4	-0.025333	0.020691	-1.224389	0.2208	
5	-0.030154	0.021915	-1.375934	0.1688	
Raw epsilon		1076.834			
Pairs within epsilon		2070.000	V-Statistic	0.709877	
Triples within epsilon		85420.00	V-Statistic	0.542473	
Dimension	C(m,n)	c(m,n)	C(1,n-(m-1))	c(1,n-(m-1))	c(1,n-(m-1))*k
2	659.0000	0.478229	964.0000	0.699565	0.489391
3	415.0000	0.312971	924.0000	0.696833	0.338365
4	261.0000	0.204706	883.0000	0.692549	0.230039
5	151.0000	0.123265	842.0000	0.687347	0.153419

BDS Test for RESID01

Date: 11/20/11 Time: 20:33

Sample: 2006M01 2011M09

Included observations: 69

Dimension	BDS Statistic	Std. Error	z-Statistic	Prob.	
2	-0.011161	0.010590	-1.053938	0.2919	
3	-0.025394	0.017100	-1.484990	0.1375	
4	-0.025333	0.020691	-1.224389	0.2208	
5	-0.030154	0.021915	-1.375934	0.1688	
6	-0.041802	0.021481	-1.946046	0.0516	
7	-0.029497	0.020010	-1.474121	0.1404	
8	-0.035982	0.017980	-2.001151	0.0454	
9	-0.030548	0.015730	-1.942025	0.0521	
10	-0.022392	0.013480	-1.661194	0.0967	
11	-0.015126	0.011363	-1.331210	0.1831	
12	-0.014349	0.009452	-1.518138	0.1290	
13	-0.008751	0.007776	-1.125493	0.2604	
14	-0.005436	0.006338	-0.857723	0.3910	
15	-0.003912	0.005125	-0.763274	0.4453	
Raw epsilon		1076.834			
Pairs within epsilon		2070.000	V-Statistic	0.709877	
Triples within epsilon		85420.00	V-Statistic	0.542473	
Dimension	C(m,n)	c(m,n)	C(1,n-(m-1))	c(1,n-(m-1))	c(1,n-(m-1))*k
2	659.0000	0.478229	964.0000	0.699565	0.489391
3	415.0000	0.312971	924.0000	0.696833	0.338365
4	261.0000	0.204706	883.0000	0.692549	0.230039
5	151.0000	0.123265	842.0000	0.687347	0.153419
6	86.00000	0.073129	820.0000	0.697279	0.114932
7	52.00000	0.046099	780.0000	0.691489	0.075596
8	27.00000	0.024977	762.0000	0.704903	0.060958
9	12.00000	0.011594	728.0000	0.703382	0.042142
10	5.000000	0.005051	691.0000	0.697980	0.027443
11	2.000000	0.002114	654.0000	0.691332	0.017241
12	0.000000	0.000000	634.0000	0.702104	0.014349
13	0.000000	0.000000	598.0000	0.694541	0.008751
14	0.000000	0.000000	565.0000	0.689024	0.005436
15	0.000000	0.000000	539.0000	0.691026	0.003912

Рис. 2. Результати виконання BDS тесту для Epsilon Value = 0,7

BDS Test for RESID01					
Date: 11/20/11 Time: 21:36					
Sample: 2006M01 2011M09					
Included observations: 69					
Dimension	BDS Statistic	Std. Error	z-Statistic	Prob.	
2	-0.003498	0.003612	-0.968475	0.3328	
3	-0.018185	0.007382	-2.463281	0.0138	
4	-0.016759	0.011296	-1.483593	0.1379	
5	-0.020473	0.015126	-1.353458	0.1759	
Raw epsilon		1723.391			
Pairs within epsilon		2636.000	V-Statistic	0.903978	
Triples within epsilon		130746.0	V-Statistic	0.830323	
Dimension	C(m,n)	c(m,n)	C(1,n-(m-1))	c(1,n-(m-1))	c(1,n-(m-1))^k
2	1111.000	0.806241	1240.000	0.899855	0.809739
3	944.0000	0.711916	1194.000	0.900452	0.730100
4	805.0000	0.631373	1144.000	0.897255	0.648132
5	674.0000	0.550204	1095.000	0.893878	0.570677

BDS Test for RESID01					
Date: 11/20/11 Time: 21:37					
Sample: 2006M01 2011M09					
Included observations: 69					
Dimension	BDS Statistic	Std. Error	z-Statistic	Prob.	
2	-0.003498	0.003612	-0.968475	0.3328	
3	-0.018185	0.007382	-2.463281	0.0138	
4	-0.016759	0.011296	-1.483593	0.1379	
5	-0.020473	0.015126	-1.353458	0.1759	
6	-0.044660	0.018739	-2.383247	0.0172	
7	-0.038008	0.022059	-1.723022	0.0849	
8	-0.062644	0.025045	-2.501244	0.0124	
9	-0.064934	0.027680	-2.345825	0.0190	
10	-0.051821	0.029963	-1.729479	0.0837	
11	-0.038256	0.031902	-1.199179	0.2305	
12	-0.056970	0.033512	-1.700026	0.0891	
13	-0.048960	0.034812	-1.406415	0.1596	
14	-0.042876	0.035826	-1.196794	0.2314	
15	-0.042539	0.036576	-1.163034	0.2448	
Raw epsilon		1723.391			
Pairs within epsilon		2636.000	V-Statistic	0.903978	
Triples within epsilon		130746.0	V-Statistic	0.830323	
Dimension	C(m,n)	c(m,n)	C(1,n-(m-1))	c(1,n-(m-1))	c(1,n-(m-1))^k
2	1111.000	0.806241	1240.000	0.899855	0.809739
3	944.0000	0.711916	1194.000	0.900452	0.730100
4	805.0000	0.631373	1144.000	0.897255	0.648132
5	674.0000	0.550204	1095.000	0.893878	0.570677
6	564.0000	0.479592	1056.000	0.897959	0.524251
7	474.0000	0.420213	1009.000	0.894504	0.458220
8	398.0000	0.368178	973.0000	0.900093	0.430821
9	328.0000	0.316908	930.0000	0.898551	0.381842
10	275.0000	0.277778	886.0000	0.894949	0.329599
11	230.0000	0.243129	843.0000	0.891121	0.281385
12	190.0000	0.210410	809.0000	0.895903	0.267380
13	156.0000	0.181185	769.0000	0.893148	0.230145
14	129.0000	0.157317	731.0000	0.891463	0.200193
15	105.0000	0.134615	695.0000	0.891026	0.177155

Рис. 3. Результати виконання BDS тесту для Epsilon Value = 0,9

Як можна побачити з результатів виконання тесту для певного заданого значення Epsilon Value зміна кількості вимірів не впливає

на значення ймовірності (p -value) BDS статистики, яке свідчить про прийняття або неприйняття нульової гіпотези. Нагадаємо, що нульова гіпотеза означає, що залишки діагностованої моделі (або рівні досліджуваного часового ряду) незалежні та однаково розподілені величини, які не мають нелінійного ефекту.

При цьому ми не приймаємо гіпотезу про відсутність нелінійного ефекту для всіх вимірів із p -значенням, меншим за 0,1.

Тобто, якщо значення p -value менше за 0,1 припускається, що модель (або досліджуваний часовий ряд) матиме GARCH ефект. Значення вимірів, задане в опції Correlation-dimension max, допомагає лише більш детально відстежити, для яких вимірів значення p -value дає змогу не прийняти нуль-гіпотезу.

Підсумовуючи, можна зазначити, що математично BDS тест заснований на кореляційному інтегралі. За нульової гіпотези випадкового блукання його статистика має стандартний нормальний розподіл. За допомогою BDS тесту можна також перевірити гіпотезу лінійності. Для цього до досліджуваного часового ряду можна застосувати лінійний фільтр, тобто замість самих значень використати оцінені залишки моделі авторегресії AR (p), де кількість лагів p визначається з інформаційною умовою AIC.

BDS тест має два параметри: розмірність вкладення m і параметр масштабу. На практиці зазвичай використовуються значення m від 2 до 5 і від 0,5 до 2, де m – середньоквадратичне відхилення вихідних даних. Якщо всі значення за межами інтервалу $(-1,96, 1,96)$, то гіпотези про чистий білий шум і лінійності відкидаються.

Таким чином, нульова гіпотеза BDS тесту: залишки випадкові та однаково розподілені (лінійність залишків). Для цього обчислюється кореляційний інтеграл – кількість близьких попарних зв'язків розділена на загальну кількість зв'язків.

Базова статистика (BDS статистика) цього тесту (різниця кореляційного інтегралу і добутку попарних зв'язків) має нормальний розподіл із середнім 0. Нульова гіпотеза не приймається (відкидається), якщо значення статистики відрізняється від нуля більше (більше за абсолютною величиною), ніж критичне значення, або якщо p -value для цього значення є меншим за 0,1 (10 %).

Correlation dimension є показником кількості векторів попарних відстаней – кількості вимірів. Для кожного виміру розрахову-

ється окрема статистика. Значущості бодай однієї статистики достатньо для відкидання нульової гіпотези.

Ми відкидаємо нульову гіпотезу, якщо при значенні p -value BDS тесту є меншим за 10 % (p -value $< 0,1$).

Пам'ятка 6. Робота з програмою E.Views

У програмі **E.Views** дані зберігаються та обробляються в *робочих файлах*, що мають розширення **.wfl**.

У робочому файлі зберігаються *об'єкти* – ряди даних (**Series**), функції, рівняння тощо. Повний список об'єктів можна побачити в меню **Objects/New Object**.

Меню програми **E.Views** міститься у верхньому рядку вікна програми (воно доступне навіть до того, як створено або відкрито робочий файл). Більшість команд головного меню дублюються в меню робочого файлу (верхній рядок робочого файлу). Крім того, своє меню має кожен об'єкт робочого файлу (ряд даних, група, рівняння, система та ін.).

Під головним меню програми міститься білий прямокутник – командний рядок, через який можна вводити команди до програми. Зазвичай він нам не потрібен, оскільки відповідні команди можна обрати з меню. Однак у **HELP E.Views** можна знайти опис команд для цього рядка з прикладами введення, формули, за якими програма обчислює той чи інший показник, а також деякі відомості з економетрики.

HELP цієї програми дружній до користувача.

Нижче наведено деякі найуживаніші команди.

Створення робочого файлу

File/New/Workfile. Потім у діалоговому вікні потрібно обрати діапазон даних – річні, місячні, тижневі або такі, що не є часовими рядами (**Undated or Irregular**). Для часових рядів потрібно вказати початкову і кінцеву дати, для недатованих даних – кількість спостережень.

Відкриється файл, де зберігатимуться два (поки що порожні) ряди даних – константа та помилки (**Residuals**). Після оцінювання

регресії ці ряди даних набудуть своїх значень. Файл можна зберегти командою **Save**.

До цього файла потрібно *записати ряди даних*. Це можна зробити кількома способами:

- 1) через меню **Objects / New Object / Series**. Відкриється вікно, схоже на таблицю Excel. Натиснувши кнопку **Edit** на цьому вікні, можна вводити дані вручну або копіювати з Excel;
- 2) імпорт даних із таблиць Excel: **Procs / Import / Read Text-Lotus-Excel**. У діалоговому вікні потрібно обрати клітинку вихідного файлу, з якої почнеться імпорт даних, а також кількість рядів, які потрібно імпортувати, або одразу дати їм назви, якщо в первісному файлі вони назв не мають;
- 3) крім того, з наявних рядів даних можна створювати нові за допомогою команди **Genr**. Наприклад, якщо у вас є ряди X та Y, можна створити ряд Z, що дорівнюватиме їх добуткові, увівши у діалогове вікно команди **Genr** рівняння $Z = X \times Y$.

Двічі клацнувши мишею на ряді даних, можна відкрити його в своєму вікні для редагування (якщо дані недоступні для редагування, потрібно натиснути кнопку **Edit+/-**). Це вікно також має свої команди меню. Наприклад, через кнопку **View** можна представити ряд даних як графіки різного виду або виконати деякі статистичні тести.

Через команду **Group** (обрати, утримуючи клавішу CTRL, кілька рядів даних, натиснути праву кнопку миші й обрати з контекстного меню команду **Open / As Group**) можна об'єднати кілька рядів у групу й проробити з нею те саме, що й з окремим рядом.

Через команду **Quick / Estimate Equation** або **Object / New Object / Equation** можна *оцінити регресійну модель*. Для цього потрібно в діалоговому вікні, що відкриється, ввести регресійне рівняння у вигляді

Залежна_змінна Константа Незалежна_змінна_1 Незалежна_змінна_2 ...

Або записати рівняння у вигляді:

**Залежна_змінна=C(1)+C(2) × Незалежна_змінна_1+C(3)
× Незалежна_змінна_2 ...**

Для оцінювання ARIMA моделі рівняння вводиться у вигляді

$$Y_t \sim C \text{ Trend AR}(1) \text{ AR}(2) \dots \text{MA}(1) \text{ MA}(2) \dots$$

(при цьому константа та тренд необов'язкові).

Рівняння можна змінити, натиснувши кнопку **Estimate** на вікні *рівняння*. Крім того, можна виконати деякі тести, а також побачити графік справжнього, оціненого ряду значень залежної змінної та їх різниці (**Actual, Fitted, Residual**) через меню **View**. Цей графік дає нам перше уявлення про те, наскільки добре наше рівняння регресії оцінює зв'язок між рядами даних.

Рівняння можна зберегти через команду **Name**, щоб використовувати його надалі.

Оцінити систему рівнянь можна за допомогою команди **Object / New Object / System** з меню робочого файла, а оцінити систему динамічних рівнянь можна за допомогою команди **Quick / Estimate VAR** головного меню програми.

Пам'ятка 7. Необхідні команди для виконання завдань кейса 1 теми 1

Зауважимо, що в E.Views більшість операцій можна виконати як через засоби графічного інтерфейсу, так і за допомогою введення текстових команд у командну область вікна. Останній спосіб зручний тим, що у випадку багаторазового виконання операцій можна повторно використовувати команди, змінюючи деякі їхні параметри.

1. Генерування часових рядів за формулою: *series назва_змінної* = *формула*. Наприклад:

series y_p_1 = 0 створює змінну (або ж перезаписує значення вже наявної змінної) *y_p_1* зі значенням усіх спостережень 0;

series y_p_2 = *nrnd* створює змінну (або ж перезаписує значення вже наявної змінної) *y_p_2* зі значенням всіх спостережень, що відповідають випадковим незалежним значенням нормально розподіленої випадкової величини $N(0,1)$. Відповідно *series y_p_3* = $.3 + .5 \times \text{nrnd}$ – значення $N(0.3, 0.5)$;

series y_p_4 = $.7 \times y_p_4(-1) + u$ – простий AR(1) процес із рядом збурень, збереженим у змінній *u*.

2. Виділення підвибірки: `smpl початок кінець` або `smpl @all` – для вибору всіх спостережень активного робочого файлу.

Усі команди, в тому числі `series`, діють на спостереження в межах активної підвибірки.

Приклади:

`smpl 1999q2 2005q3` – з другого кварталу 1999 року до 3 кварталу 2005 року;

`smpl @first +2 @last+2` – вибір другого спостереження з вибірки;

`smpl @first @last` – аналог `@all` – вибір всіх спостережень.

3. Модуль числа: `abs(число)`

Наприклад:

`abs(-6)` дорівнює 6;

`abs(nrnd)` – модуль нормально розподіленої випадкової величини.

4. Встановлення початкового значення для генератора псевдовипадкових чисел: `rndseed значення`.

`rndseed 15` – встановити значення на 15.

ЛІТЕРАТУРА ДО ТЕМИ 1–2

1. Єріна А. М. Економічна статистика : практикум / А. М. Єріна, О. К. Мазуренко, З. О. Пальян. – К. : ТОВ «УВПК:ЕксОб», 2002. – 232 с.
2. Єріна А. М. Організація вибірових спостережень : навчальний посібник / А. М. Єріна. – К. : КНЕУ, 2004. – 127 с.
3. Лук'яненко І. Г. Системне моделювання показників бюджетної системи України : Принципи та інструменти : монографія. – К. : Видавничий дім «Києво-Могилянська академія», 2004. – 584 с.
4. Лук'яненко І. Г. Економетрика. Теорія та практика / І. Г. Лук'яненко, Л. І. Краснікова. — К. : Знання, 1998 — 493 с.
5. Лук'яненко І. Г. Сучасні економетричні методи в фінансах / І. Г. Лук'яненко, Ю. О. Гордніченко. – К. : Літера, 2003. – 348 с.
6. Лук'яненко І. Г. Динамічні макроеконометричні моделі. Новий концептуальний підхід. – К. : ВД «КМ Академія», 2003. – 50 с.
7. Лук'яненко І. Г. Попит на гроші. Ілюстрація класичного підходу / І. Г. Лук'яненко // Наукові записки НаУКМА. Серія: Економіка. – 2000. – Т. 18. – С. 38–42.
8. Лук'яненко І. Г. Проблеми «хибної» регресії в економічних дослідженнях / І. Г. Лук'яненко // Наукові записки НаУКМА. Спеціальний випуск. – К. : ВД «КМ Академія», 2002. – Т. 20. – Ч. II. – С. 422–424.
9. Лук'яненко І. Г. Проблеми діагностики класичних економетричних моделей / І. Г. Лук'яненко // Економічна кібернетика. Міжнародний науковий журнал. – 2004. – № 3–4 (27–28). – С. 100–107.
10. Лук'яненко І. Г. Застосування формальних тестів для обґрунтування остаточної специфікації економетричної моделі / І. Г. Лук'яненко // Вісник ЖДТУ. Серія: Економічні науки. – 2004. – № 4 (30). – С. 346–351.
11. Лук'яненко І. Г. Моделювання впливу зміни монетарних режимів на фінансово-економічний розвиток країн з перехідною економікою / І. Г. Лук'яненко, В. М. Жук // Економіка і прогнозування. – 2011. – № 2 (квітень–червень). – С. 130–152.
12. Макроекономічне моделювання та короткострокове прогнозування. – Харків : Форт, 2000. – 336 с.
13. Михайличенко С. Ю. Реальний ефективний обмінний курс гривні : економічний зміст, динаміка, моделі, застосування / С. Ю. Михайли-

- ченко, І. Г. Лук'яненко, М. С. Михайличенко. – К. : Видавничий дім «КМ Академія», 2003. – 197 с.
14. Наконечний С. І. Економетрія : підручник / С. І. Наконечний, Т. О. Терещенко, Т. П. Романюк. – К. : КНЕУ, 2000. – 296 с.
15. Райская Н. Временные лаги в динамике инфляции // Вопросы экономики. – 1996. – №8. – С. 34–41.
16. Султан К. Методологічні аспекти розробки та практичного застосування макроеконометричної моделі України / К. Султан, І. Лук'яненко, Ю. Городніченко. – К. : Видавничий дім «КМ Academia», 2000. – 204 с.
17. Черняк О. І. Динамічна економетрика / О. І. Черняк, А. В. Ставицький. – К. : КВЦІ, 2000. – 120 с.
18. Anderson T. W. Testing Multiple Equation System for Common Nonlinear Components / T. W. Anderson, F. Vahid // Journal of Econometrics. – 1998. – № 84. – P. 1–36.
19. Baillie R. T. Long Memory Processes and Fractional Integration in Econometrics / R. T. Baillie // Journal of Econometrics. – 1996. – № 73. – P. 5–60.
20. Balke N. S. Threshold Cointegration / N. S. Balke, T. Fomby // International Economic Review. – 1997. – № 38. – P. 627–646.
21. Blanchard Oliver J. What is Left of the Multiplier Accelerator / Oliver J. Blanchard // The American Economic Review. – May 1981. – Vol. 71. – Issue 2. – P. 150–154.
22. Bollerslev T. Generalized Autoregressive Conditional Heteroscedasticity / T. Bollerslev // Journal of Econometrics. – 1986. – Vol. 31. – P. 307–327.
23. Caner M. Threshold Autoregression with a Unit Root / M. Caner, B. Hansen // Econometrica. – 2001. – Vol. 69. – P. 1555–1596.
24. Carrasco M. Misspecified Structural Change. Threshold and Markov-switching Models / M. Carrasco // Journal of Econometrics. – 2002. – Vol. 109. – P. 239–273
25. Charemza W. W. New Direction in Econometric Practice / W. W. Charemza, D. F. Deadman. – Brookfield : Edward Eglar, 1992. – 343 p.
26. Davidson R. Problems with the Estimation of Moving Average Processes / R. Davidson // Journal of Econometrics. – 1981. – № 16. – P. 295–310.
27. Davidson R. Several Tests for Model specification in the Presence of Alternative Hypothesis / R. Davidson, J. MacKinnon // Econometrica. – 1981. – Vol. 49. – P. 781–793.
28. Dickey D. A. Unit Roots in Time Series Models : Tests and Implications / D. A. Dickey, W. R. Bell, R. B. Miller // American Statistician. – 1986. – № 40. – P. 12–26.
29. Dickey D. A. Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root / D. A. Dickey // Econometrica. – 1981. – № 49. – P. 1057–1072.

30. Edelberg W. Understanding the Effects of a Shock to Government Purchases / Wendy Edelberg, Martin Einchenbaum, Jonas D. M. Fisher // NBER Working Paper. – 1998. – № 6737. – P. 1011–1028.
31. Enders W. Applied Econometrics Time Series : 3rd ed. / W. Enders. – John Wiley and Sons, Inc., – 2010. – 517 p.
32. Favero Carlo A. Applied Macroeconometrics / A. Favero. Carlo – Oxford : Oxford University Press, 2001. – 282 p.
33. Frances P. H. Spurious Deterministic Seasonality / P. Frances, S. Hylleberg, H. Lee // Economics Letters. – 1995. – № 48. – P. 249–256.
34. Frances P. H. Time Series Models for Business and Economic Forecasting / Philip Hans Frances. – Cambridge University Press, 1999. – 280 p.
35. Graham E. Economic Forecasting / E. Graham, A. Timmermann // Journal of Economic Literature. – 2008. – № 46. – P. 504–508.
36. Granger C. W. J. Spurious Regressions in Econometrics / C. W. J. Granger, P. Newbold // Journal of Econometrics, 1974, № 2. – P. 111–120.
37. Granger C. W. J. Modelling Nonlinear Economic Relationships / C. W. J. Granger, T. Terasvirta. – Oxford : Oxford University Press, 1993. – 245p.
38. Greene W. Econometric analysis / William Greene. – 4th ed. – New Jersey : Upper Saddle River, Prentice-Hall, Inc., 2000. – 1004 p.
39. Hamilton J. D. Time Series Analysis / James D. Hamilton. – New Jersey : Princeton University Press, 1994. – 799 p.
40. Hannan E. J. The Estimation of the Order of an ARMA Process / E. J. Hannan // Annals of Statistics. – 1980. – № 8. – P. 1071–1081.
41. Hansen L. P. Seasonality and Approximation Errors in Rational Expectations Models / L. P. Hansen, T. J. Sargent // Journal of Econometrics. – 1993. – № 55. – P. 21–56.
42. Hendry D. F. Dynamic Econometrics / D. F. Hendry. – Oxford : Oxford University Press, 1995. – 456 p.
43. Johnston J. Econometric Methods / J. Johnston, J. Dinardo. – McGraw-Hill, 1997. – 531p.
44. Pagan A. R. The Econometrics of Financial Markets / A. R. Pagan // Journal of Empirical Finance. – 1996. – № 3. – P. 15–102.
45. Poskitt D. S. Diagnostic Tests for Multiple Time Series Models / D. S. Poskitt, A. R. Tremayne // Annals of Statistics. – 1982. – № 10. – P. 114–120.
46. Verbeek M. A Guide to Modern Econometrics / Marno Verbeek. – Chichester : John Wiley and Sons, Ltd., 2000. – 386 p.

Навчальне видання

І. Г. Лук'яненко, В. М. Жук

**Аналіз часових рядів
Побудова ARIMA, ARCH / GARCH моделей
з використанням пакета E.Views 6.0**

Частина перша

Практичний посібник для роботи в комп'ютерному класі

**Редактор О. Харитоненко
Художнє оформлення
Коректор Л. Пізнюк
Комп'ютерна верстка М. Кулікова**

Підписано до друку _____ Формат 60 × 84/16.
Гарнітура Warnock Pro. Ум. друк. арк. 10,92.
Наклад 300 прим.

Надруковано у ТОВ «Аграр Медіа Груп»
04080, м. Київ, вул. Новокостянтинівська, 4а
Тел.: (044) 361-53-06
Свідцтво про внесення суб'єкта видавничої справи до
Державного реєстру видавців ДК № 3651 від 22.12.2009

Лук'яненко І. Г., Жук В. М. Аналіз часових рядів. Частина перша : Побудова ARIMA, ARCH/GARCH моделей з використанням пакета E.Views 6.0. Практичний посібник для роботи в комп'ютерному класі / І. Г. Лук'яненко, В. М. Жук. – К. : НаУКМА ; Аграр Медіа Груп, 2013. – 187 с.

ISBN

Посібник містить детальні інструкції для самостійної побудови ARIMA, ARCH/GARCH моделей за допомогою пакета E.Views 6.0. Оскільки методологія та алгоритми побудови цих моделей стандартні, посібник можна використовувати і під час роботи з іншими пакетами прикладних програм.

Важливою складовою посібника є додатковий матеріал. У розділі «Додатки» зібрана інформація щодо особливостей застосування критеріїв прогнозної якості у випадку моделювання фінансових часових рядів, програми автоматичної реалізації процедури Хенона–Рісанена тощо. Всі етапи побудови ARIMA, ARCH/GARCH моделей та прогнозування за ними проілюстровані рисунками, графіками, таблицями. До кожної теми додано ситуативні вправи з розв'язками та завдання для самостійного виконання, а також пам'ятки зі стислим описом певних теоретичних положень або необхідних формул. Наведено список рекомендованої літератури.

Розраховано на студентів економічних спеціальностей програм бакалаврату та магістратури, а також аспірантів, викладачів, фахівців – усіх, хто прагне оволодіти навичками побудови ARIMA й ARCH/GARCH моделей та зрозуміти особливості їх застосування для прогнозу на практиці.

ББК 65.261