Шифр: «Prognoz»

Студентська наукова робота

з галузевої науки

«Кількісні методи в економіці»

на тему

„МОДЕЛЮВАННЯ ТА ПРОГНОЗУВАННЯ НЕСТАЦІОНАРНИХ ФІНАНСОВИХ ЧАСОВИХ РЯДІВ”

2012

ЗМІСТ

ВСТУП

Розвиток фінансового ринку створює ситуацію, коли, з одного боку, все більше число активів залучається до сфери ринкових відносин, а з іншого боку, методи, розроблені в рамках теорії фінансового ринку, починають активно застосовуватися до об'єктів реального інвестування (хеджування, вбудовані опціони та ін.) [1, 2]. Ситуація, що склалася, в цілому, робить позитивний вплив, як на розвиток самої фінансової теорії, так і на розвиток ринкових механізмів.

Незважаючи на це, в даний час зростає незадоволеність результатами практичної діяльності на фінансових ринках. Триває пошук причин, що викликають почуття незадоволеності, і рецептів успішного інвестування. Серед причин лідирує неадекватність умов гіпотези ефективного ринку, а серед рецептів - опціонні стратегії хеджування [3].

Основною причиною успіхів і невдач на ринку є вміння або відсутність такого приймати інвестиційні рішення на основі аналізу сформованого уявлення про майбутнє, що дозволяє оцінити розміри реального доходу та ризику.

Таким чином *актуальним* є розвиток методів моделювання та прогнозування фінансових часових рядів. Особливість такого аналізу обумовлена складною структурою ринку. Фінансові ринки постійно зазнають потрясінь, в результаті чого виникає проблема виявлення факторів, які б сигналізували про можливість кризи. Тоді можна було б змінювати напрямок руху тренду або навіть створювати кризу. Також поява тих або інших зовнішніх факторів (передвиборча кампанія, прийняття нового закону, зміна керуючої ланки компаній тощо) може проявлятися через невизначений інтервал часу в майбутньому [4]. Як наслідок, фінансові часові ряди можуть характеризуватися наявністю різких стрибків, автокореляцією та гетероскедастичностю помилок, порушенням стаціонарності [5, 6].

З врахуванням цього в даній роботі пропонуються наступні етапи дослідження. Спочатку проводиться попередня обробка вихідних даних, розбиття часових рядів на класи та віднесення до одного з класів. Більшість фінансових часових рядів є нестаціонарними, тому використовується наступна класифікація досліджуваних рядів: TS та DS класи [7]. До TS класу належать ряди, які є стаціонарними відносно детермінованого тренду, тому для них прийнято виділяти трендову складову [8]. До DS класу входять ряди з присутнім стохастичним трендом (можливо, разом з детермінованим трендом). Такі ряди зводяться до стаціонарних шляхом *k*-кратного диференціювання [9]. Ці два класи нестаціонарності вимагають різних методів моделювання, особливо у випадках використання побудованих моделей з метою прогнозування. Важливою характеристикою для прогнозування рядів TS класу є те, що вплив минулих відхилень затухає з часом, а в рядах DS класу кожне відхилення впливає з однаковою силою на всі наступні значення часового ряду [10]. У створенні та розвитку теорії, що враховує нестаціонарність часових рядів, брали участь багато вчених, серед яких Дікі Д., Фуллер В., Гренджер К., Хансен Б., Йохансен С., Юселіус К., Перрон П., Філліпс П., Сімс К., Сток Дж., Ватсон М.

Наступним кроком аналізується ефективність методів для кожного з класів. При цьому використані методи повинні враховувати можливу складність таких рядів, яка обумовлена особливістю механізмів, що їх формують. При моделюванні часових рядів ми виходили з того, що в їх динаміці спостерігаються ефекти неперервної та дискретної зміни, причому ні моменти часу, коли відбуваються стрибки в значеннях часового ряду, ні величини цих стрибків невідомі. Тому пропонується використовувати дискретно-неперервну модель, для якої параметри неперервної моделі і дискретні ефекти оцінюються ітераційним шляхом.

Далі вирішується питання вибору прогнозного значення. Зазначимо, що до складу описаної нижче дискретно-неперервної моделі включені фіктивні змінні, які формуються в процесі ітерацій. Тому при прогнозуванні використовується множинна логіт-модель, яка дозволяє будувати найімовірнісний варіант прогнозу.

І в кінці оцінюється адекватність модельних прогнозних значень та можливості їх використання при прийнятті рішень.

Апробація запропонованих алгоритмів проведена на фінансових часових рядах курсу акцій українських компаній на Варшавській фондовій біржі, яка характеризується найбільш динамічним розвитком серед бірж центральної Європи.

Об'єкт дослідження - реальні фінансові активи, що виступають в якості об'єктів інвестування.

Предмет дослідження - математичний апарат моделювання та прогнозування фінансових часових рядів.

Метою даної роботи є побудова економетричних моделей, що адекватно описують нестаціонарні ряди і дозволяють будувати високоточні короткострокові прогнози для майбутніх значень часового ряду.

Структура роботи наступна. Перший розділ присвячений особливостям дослідження нестаціонарності фінансового часового ряду. У другому розділі розглянутий алгоритм побудови дискретно-неперевної моделі в залежності від нестаціонарності часового ряду та особливості використання отриманої моделі у прогнозуванні. У третьому розділі всі моделі та методи апробуються на статистичних даних цін акцій українських компаній, перевіряється адекватність отриманих прогнозних значень при практичних дослідженнях.

РОЗДІЛ 1. НЕСТАЦІОНАРНІ ЧАСОВІ РЯДИ

Як будь-який аналіз, економетричний аналіз починається з оцінки вхідних статистичних даних. Тобто, для того щоб аналізувати часовий ряд, спочатку необхідно провести попередній аналіз даних.

Так як в часових рядах є взаємозалежність даних, то це виключає можливість застосування бази статистичного аналізу в звичайній формі. Члени часового ряду не розподілені однорідно і не є статистично незалежними [7].

Для аналізу економічних часових рядів розрізняють групи факторів, що можуть впливати на формування значень ряду. В загальному випадку їх може бути безліч – календарні ефекти, випадкові флуктуації, викиди і т.д. Так як ми розглядаємо економетричний аналіз часових рядів, тобто нас цікавить саме економічна природа таких факторів, то будемо виділяти такі основні 3 групи:

- Довгострокові, які формують загальну тенденцію в зміні аналізованної ознаки. Зазвичай ця тенденція описується за допомогою невипадкової функції $f_{\text{тр}}(t)$, яка називається функцією тренда або просто трендом. Під трендом розуміють зміну, визначальний загальний напрям розвитку, основну тенденцію часового ряду. Це систематична складова довгострокової дії .
- Короткострокові, які можуть відображати як циклічні та сезонні зміни, так і ефекти, що зникають або затухають з часом (в тому числі і структурні зсуви).
- Випадкові (нерегулярні), які не піддаються обліку і реєстрації. Їх дія на формування значень часового ряду саме і зумовлює стохастичну природу елементів $x(t)$, а отже, і необхідну інтерпретацію $x(1)$, $x(2)$, …, $x(N)$ як спостережень, які проводилися над випадковими величинами $\xi(1)$, $\xi(2)$, …, $\xi(N)$ відповідно. Будемо позначати результат дії

випадкових факторів за допомогою випадкових величин $\varepsilon(t)$ («залишків», «помилок») [4].

Щоб побудувати математичну модель, яка б адекватно описувала вхідні дані і провести подальший аналіз часових рядів, необхідно правильно провести їх класифікацію. Часові ряди поділяють на стаціонарні та нестаціонарні.

Статистичний процес називається строго стаціонарним, якщо взаємний розподіл ймовірностей l спостережень інваріантний по відношенню до загального зсуву часового аргументу, тобто спільна щільність розподілу випадкових величин $x(t_1)$, $x(t_2)$, …, $x(t_l)$ така ж, як для величин $x(t_{1+k})$, $x(t_{2+k})$, …, $x(t_{l+k})$, при будь-яких цілих значеннях k [8].

Вимога стаціонарності, яка визначається цими умовами, є достатньо жорсткою. На практиці при вивченні часових рядів розрізняють строгу стаціонарність та слабку стаціонарність або стаціонарність другого порядку. В другому випадку процес має стале середнє значення μ=E(x_t) для всіх t, яке визначає рівень, відносно якого він флуктує (коливається); сталу дисперсію σ^2=E(x_t-μ)^2 для всіх t і сталу автоковаріацію $\gamma(q) = \mathrm{cov}(x(t), x(t+q)) = E[(x(t)-\mu)(x(t+q)-\mu)]$ для всіх t, тобто коваріація між $x(t), x(t+q)$ залежить лише від величини зсуву q і не залежить від t [11].

В усіх інших випадках часовий ряд є нестаціонарним. Всі відомі нам методи побудови математичної моделі вимагають, щоб часовий ряд був стаціонарним. Тому нам також необхідно розглянути методи зведення нестаціонарних часових рядів до стаціонарних.

Бокс і Дженкінс запропонували свій підхід до визначення стаціонарності часових рядів та зведення нестаціонарних часових рядів до стаціонарних [10]. Такий підхід Бокс і Дженкінс запропонували 25 – 30 років тому, і на той час він був новітній і широко використовувався, оскільки дозволяв досить зручно і компактно описувати більшість фінансових та макроекономічних процесів. Проте в такому підході є багато незручностей і неточностей. Так не зовсім коректним є встановлення стаціонарності чи нестаціонарності даного часового ряду за виглядом його графіків функцій автокореляції та часткової автокореляції.

Розглянемо деякий інший спосіб зведення нестаціонарних часових рядів до стаціонарних. Нехай часовий ряд має лінійний тренд: $y_t = \alpha + \beta t + x_t + \varepsilon_t$, де x_t - часовий ряд, для якого характерна стаціонарність. Тоді спочатку треба провести детрендування ряду, щоб побудувати математичну модель. Наша модель матиме вигляд $y_t = \mu + \nu t + u_t$, u_t - залишки ряду. Використання корелограми отриманих залишків дозволить ідентифікувати ряд і знайти порядок моделі. Після цього проводиться оцінювання моделі з включеним до неї трендом.

Однак і такий підхід не є універсальним. По-перше, він може бути використаний лише в разі наявності в моделі тренду, а по-друге, навіть за його наявності детрендизація не завжди приводить нестаціонарний ряд до стаціонарного.

Розглянемо ще один приклад – так зване «випадкове блукання зі зсувом», тобто процес вигляду $y_t = \alpha + y_{t-1} + \varepsilon_t$, $t = 1, 2, \ldots, T$, $y(0)=y_0$, $\alpha \neq 0$ [7]. Ряд y_t має і детермінований, і стохастичний тренд. Детрендизація даного ряду призводить до ряду, який не є стаціонарним. Щоб звести ряд такого типу до стаціонарного, треба перейти до різниць $\Delta y_t = y_t - y_{t-1}$, звідки отримаємо стаціонарний ряд.

Можемо зробити висновок про те, що часові ряди необхідно розділити на класи з різними властивостями. На практиці їх розбивають на TS і DS ряди [9].

Часовий ряд y_t називається стаціонарним відносно детермінованого тренду $T(t)$, якщо ряд $y_t - T(t)$ є стаціонарним. Якщо ряд y_t стаціонарний відносно деякого детермінованого тренду, то кажуть, що ряд належить класу рядів, стаціонарних відносно детермінованого тренду, і він є TS (time stationary) рядом [8].

Часовий ряд y_t називається інтегрованим порядку k, k=1, 2, …, якщо:

- Ряд y_t не є стаціонарним або стаціонарним відносно детермінованого тренда, тобто не є TS-рядом
- Ряд $\Delta^k y_t$, який отримано в результаті k-кратного диференціювання ряда y_t є стаціонарним рядом
- ряд $\Delta^{k-1} y_t$, який отримано в результаті (k-1)-кратного диференціювання ряда y_t, не є TS-рядом.

Різницево стаціонарним, або DS (difference stationary) рядом називається інтегрований ряд порядку k=1, 2, … .

Для вирішення питання про віднесення ряду до TS або DS класу існують процедури, які вирішують поставлену задачу в класі ARMA моделей. В цьому випадку перевірка нульової гіпотези щодо належності досліджуваного часового ряду DS класу може бути зведена до перевірки того, що авторегресійний поліном містить хоча б один одиничний корінь [10]. Але кожна з цих процедур не є універсальною і має різні недоліки. Наприклад, отримується досить низька потужність, що веде до частого підтвердження нульової гіпотези, хоча насправді вона не виконується. Іншим недоліком може бути невиконання теоретичних передумов, на яких базується дана процедура. Це може привести до частого відхилення нульової гіпотези, хоча вона справджується в дійсності [7]. Тому при віднесенні часового ряду до одного з класів використовують одразу декілька процедур.

Найчастіше на практиці використовується розширений тест Дікі-Фулера [9], який оцінює методом найменших квадратів наступну модель:

$$\Delta x_t = \alpha + \rho x_{t-1} + \beta t + \sum_{i=1}^{p} \delta_i \Delta x_{t-i} + u_t,\ t = 1, 2, \ldots, N, \tag{1}$$

де u_t – незалежні, нормально розподілені залишки з нульовим математичним сподіванням; N – об'єм вибірки.

Нульова гіпотеза про те, що ряд належить DS класу, еквівалентна тому, що $\rho = 0$. Альтернативна гіпотеза – ряд належить TS класу та $\rho < 0$.

При практичному використанні тесту Дікі-Фуллера важливо правильно специфікувати регресійну модель (1). Для рядів можливі випадки відсутності тренду, вільного члена чи вибору різної довжини лагу p. Неправильна специфікація тесту може привести до неадекватних результатів [5].

Для вибору оптимальної величини лагу пропонуються різні підходи [12]. У даній роботі використовується інформаційний критерій Шварца:

$$BIC(p) = \log(\hat{\sigma}_u^2) + \frac{(p+1)\cdot\log(N)}{N}, \tag{2}$$

де $\hat{\sigma}_u^2$ – оцінка дисперсії залишків регресійної моделі (1).

При виборі оптимального p перевага надається моделям з найменшим значенням BIC та високим значенням коефіцієнта детермінації R^2.

Якщо встановлена оптимальна кількість лагів, то існують процедури для остаточної правильної специфікації моделі (1). Найбільш відомі з них описані в [13]. На практиці ефективним є метод Доладо Дж., Дженкінса Т., Сосвіля – Ріверо С. [14]. Цей метод ґрунтується на тому, що включення в модель (1) додаткових незначущих складових приводить до завищення розрахованого значення тесту. При зменшенні кількості регресорів в моделі (1) критичне значення збільшується (при заданих рівні значущості та числу ступенів вільності). Вказаний метод враховує ці особливості і полягає в наступному.

1. Оцінюємо рівняння вигляду:

$$\Delta x_t = \alpha + \beta t + (\rho_1 - 1)x_{t-1} + \sum_{i=1}^{\rho} \delta_i \Delta x_{i-1} + \varepsilon_t, \tag{3}$$

тобто включаємо в модель і вільний член, і тренд. При цьому нульова гіпотеза має вигляд H_0: DS => $\rho_1 = 1$. Якщо $\rho_1 = 1$, то ряд фактично відноситься до нового класу – в ньому наявні два типи тренду. Для спрощення покладемо $p_1 = 0$, тоді отримуємо процес типу

$$x_t = \alpha + \beta t + x_{t-1} + \varepsilon_t. \tag{4}$$

Будемо називати лінійний тренд детерміновани трендом, а випадкове блукання із зсувом – стохастичним трендом. Загальний розв'язок рівняння матиме вигляд:

$$x_t = x_0 + \sum(\alpha + \beta t + \varepsilon_t) = x_0 + \alpha t + \beta \frac{t(t+1)}{2} + \sum \varepsilon_t \, .$$

Порівняння величин t-відношення для коефіцієнту при $x_{t\text{-}1}$ з критичним значенням для тесту Дікі-Фуллера призводить до двох випадків. Висновки щодо прийняття або відхилення нульової гіпотези залежать від того, чи потрапить розрахункове значення в критичну область. Якщо розрахункове значеня не перевищує критичне, то згідно тесту Дікі-Фуллера ми повинні прийняти нульову гіпотезу. Але якщо ми провели неправильну специфікацію моделі, то лінійний тренд βt може бути включений до неї помилково. Тобто тепер необхідно відповісти на питання чи вірно ми включили тренд у модель. Для цього припустимо, що нульова гіпотеза виконується і тоді ми отримуємо модель вигляду

$$\Delta x_t = \alpha + \beta t + \sum_{i=1}^{p} \delta_i \Delta x_{i-1} + \varepsilon_t \quad . \tag{5}$$

2. Оцінимо отриману модель. Оскільки це рівняння не може мати одиничних коренів, то можемо застосувати стандартну *t*-статистику. Перевіряємо значущість МНК-оцінки $\hat{\beta}$ коефіцієнта β: якщо $\hat{\beta}$ є незначущою, то $\beta=0$. Робимо висновок про те, що тренд був включений до моделі помилково і переходимо до наступного кроку. Якщо ж $\hat{\beta}$ є значущою, то це підтверджує висновок отриманий на першому кроці і даний ряд належить класу DS.

3. Якщо МНК-оцінка $\hat{\beta}$ є незначущою, то необхідно оцінити регресію

$$\Delta x_t = \alpha + (\rho_1 - 1)x_{t-1} + \sum_{i=1}^{p} \delta_i \Delta x_{i-1} + \varepsilon_t \qquad (6)$$

за допомогою тесту Дікі-Фуллера. Якщо знову приймається нульова гіпотеза, то необхідно перевірити правильність включення вільного члена α до моделі. Будуємо регресію

$$\Delta x_t = \alpha + \sum_{i=1}^{p} \delta_i \Delta x_{i-1} + \varepsilon_t \qquad (7)$$

і за допомогою *t*-статистики і розподілу Стьюдента перевіряємо значущість оцінки $\hat{\alpha}$ коефіцієнта α. Якщо $\hat{\alpha}$ виявляється незначущою, то ми повинні виключити α з моделі, оцінити за допомогою тесту Дікі-Фуллера регресію

$$\Delta x_t = (\rho_1 - 1)x_{t-1} + \sum_{i=1}^{p} \delta_i \Delta x_{i-1} + \varepsilon_t \qquad (8)$$

і зробити кінцевий висновок належності ряду до DS (якщо нульова гіпотеза приймається) або TS (якщо нульова гіпотеза відхиляється) класу.

Передумовами тесту Дікі-Фуллера є неавтокорельованість, гомоскедастичність та нормальний розподіл залишків u_t регресійної моделі (1). У випадку невиконання однієї з цих умов на практиці використовується тест Філліпса-Перрона, який також базується на оцінюванні моделі (1), але стандартні помилки розглядає в формі Н'юї-Веста [4]. Отже, на відміну від критерію Дікі-Фуллера, випадкова складова u_t з нульовим математичним сподіванням може бути автокорельована, гетероскедастична та не обов'язково нормально розподілена. Тим самим, критерій Філліпса-Перрона розглядає більш широкий клас часових рядів.

Зазначимо, якщо досліджуваний часовий ряд буде належати DS класу, то описана методика імітації та прогнозування буде застосовуватися для диференційованого часового ряду. Для цього розширений тест Дікі-Фуллера використовується для регресійної моделі

$$\Delta^{(n)} x_t = \alpha + \rho \Delta^{(n-1)} x_{t-1} + \beta t + \sum_{i=1}^{p} \delta_i \Delta^{(n)} x_{t-1} + u_t , \tag{9}$$

де $\Delta^{(n)} x_t = \Delta^{(n-1)} x_t - \Delta^{(n-1)} x_{t-1}$ – кінцеві різниці *n*-го порядку. Значення *n* збільшується до тих пір, доки не буде встановлена стаціонарність відповідного ряду.

РОЗДІЛ 2. ДИСКРЕТНО-НЕПЕРЕРВНА МОДЕЛЬ

Припустимо, що нами отримано деякий стаціонарний часовий ряд. Далі розглянемо метод прогнозування цього ряду. Вважаємо, що в досліджуваних часових рядах спостерігаються неперервні та дискретні процеси. Тоді відповідну неперервно – дискретну регресійну модель для даного ряду можна записати наступним чином:

$$y_t = f(t, \mathbf{a}) + \mathbf{d}'\mathbf{z} + \varepsilon_t, \tag{10}$$

де y_t – значення ряду в момент часу t, $f(t,\ \mathbf{a})$ – неперервна складова, $\mathbf{d}'\mathbf{z}$ – дискретна складова, $\mathbf{a} = (a_0,\ a_1,\ ...,\ a_l)'$, $\mathbf{d} = (d_1,\ d_2,\ ...,\ d_m)'$ – вектори оцінювальних параметрів, $\mathbf{z} = (z_{t1},\ z_{t2},\ ...,\ z_{tm})'$ – вектор змінних дискретної складової; ε_t – випадкове збурення.

В якості неперервної складової будемо використовувати авторегресійні моделі, які залежать від класу нестаціонарного ряду. Якщо ряд належить TS класу, то до авторегресійної моделі включаємо детермінований тренд, якщо DS класу, то використовуємо кінцеві різниці, які є стаціонарними.

Недостатня складність у виборі неперервної складової моделі для якісного опису даного ряду може бути компенсована ітераційним шляхом. В огляді літератури з даної проблеми вказується на використання двозначних фіктивних змінних на кожній ітерації [15]. Але аналіз реальних часових рядів показує, що при такому підході на практиці не виконуються передумови лінійної регресійної моделі (гомоскедастичність, відсутність автокореляції та нормальний розподіл залишків). Тому у даній роботі пропонується використовувати тризначну фіктивну змінну.

Спочатку оцінюється неперервна складова регресійної моделі (10) і обчислюються залишки $\varepsilon_t^{(1)} = y_t - f(t, \hat{\mathbf{a}})$. На першому кроці дискретна змінна формується наступним чином:

$$z_{t1} = \begin{cases} +T^{(1)}, \ \varepsilon_t^{(1)} > \xi, \\ 0, -\xi \le \varepsilon_t^{(1)} \le \xi, \\ -S^{(1)}, \ \varepsilon_t^{(1)} < -\xi. \end{cases} \quad (11)$$

Тут ξ - точність наближення (у даній роботі значення ξ вибиралось рівним $0{,}01 \cdot (y_t^{\max} - y_t^{\min})$), $S^{(1)}$ і $T^{(1)}$ – кількість точок з даного проміжку, в яких виконуються нерівності $\varepsilon_t^{(1)} > \xi$ і $\varepsilon_t^{(1)} < -\xi$ відповідно (тоді середнє значення фіктивної змінної автоматично дорівнює нулю). Далі оцінюється регресійна модель $\varepsilon_t^{(1)} = d_1 z_{1t} + \varepsilon_t^{(2)}$.

Аналогічно на k-кроці (k=2, 3, …, m) обчислюються залишки $\varepsilon_t^{(k)} = \varepsilon_t^{(k-1)} - \hat{d}_{k-1} z_{(k-1)t}$ і

$$z_{tk} = \begin{cases} +T^{(k)}, \quad \varepsilon_t^{(k)} > \xi, \\ 0, \ -\xi \le \varepsilon_t^{(k)} \le \xi, \\ -S^{(k)}, \quad \varepsilon_t^{(k)} < -\xi. \end{cases} \quad (12)$$

Далі оцінюється регресійна модель $\varepsilon_t^{(k)} = d_k z_{kt} + \varepsilon_t^{(k+1)}$.

Практичні дослідження показують, що на 3-4-й ітерації фіктивна змінна приймає значення 0, що означає закінчення процесу ітерацій із заданою точністю.

Розглянемо тепер особливості використання дискретно-неперевної моделі для побудови короткострокових прогнозів. Метою прогнозування мають бути оцінки з наступними властивостями:

- прогнозні оцінки повинні мати можливість передбачати поведінку деякого часового ряду, а не лише відображати основні характеристики статистики історичних даних;

- техніка оцінки прогнозного значення не збільшує похибку, яка існує в статистичних даних;

- прогнозна оцінка має відповідати фізичному змісту процесу.

Описана процедура вказує на те, що за допомогою вектора **z** фіктивних змінних можна досягти необхідну точність апроксимації. Однак таке введення фіктивних змінних приводить до неоднозначності при визначенні прогнозного значення регресанда. Для прогнозної точки t^* регресанд може приймати 3^m значень, оскільки кожна з фіктивних змінних (12) може приймати одне із трьох значень, а кількість ітерацій дорівнює m. Для вибору оптимального значення фіктивної змінної в точці t^* пропонується використовувати множинну логіт – модель [16]:

$$\mathrm{P}(z_i^* = \hat{z}_i^*(j)) = \frac{e^{t^* \mathbf{b}_j}}{\sum_{j=1}^{3} e^{t^* \mathbf{b}_j}}, \qquad j = 1,\ 2,\ 3;\ i = 1,\ 2, \ldots, m, \tag{13}$$

де z_i^* – значення фіктивної змінної z_{ti} в прогнозній точці t^*, $\hat{z}_i^*(j)$ – оцінка прогнозного значення для варіанта j, $\mathbf{b} = (b_1, b_2, b_3)'$ – вектор оцінювальних параметрів.

Регресійна модель, яка відповідає (13), є істотно нелінійною відносно $\mathbf{b}$, тому для її оцінювання можна використовувати чисельний метод Ньютона–Рафсона розв’язання системи рівнянь правдоподібності, записаних у логарифмічній формі [16].

РОЗДІЛ 3. ПРАКТИЧНА ЧАСТИНА

На сьогоднішній день за класифікацією FTSE (Financial Times Stock Exchange, провайдер всесвітньо відомих фондових індексів), усі фінансові ринки за рівнем розвитку розподіляються наступним чином: розвинені; ті, що розвиваються; приграничні. Згідно з оцінкою FTSE, український фондовий ринок поки що не задовольняє вимогам навіть найнижчого рівня, відстаючи в розвитку від таких країн, як Ботсвана, Кот-д'Івуар, В'єтнам, Маврикій та Нігерія [17].

До основних проблем фондового ринку України можна віднести його недостатню розвиненість, слабку законодавчу базу, низький рівень ліквідності цінних паперів, недостатні технічну оснащеність та інформованість громадян про фондовий ринок і його інструменти [18, 19].

Тому апробація запропонованого методу проводилася для цін акцій українських акцій на Варшавській фондовій біржі, що входять до індексу WIGUkraine.

Варшавська Фондова Біржа (ВФБ) є ринком акцій у центральній Європі, який характеризується найбільш динамічним розвитком. Успіхи ВФБ є результатом дії цілого ряду факторів: потенціалу польського господарства, сприятливих законодавчих актів, сили та фінансових ресурсів інвесторів, інформаційної інфраструктури, а також праці і професійності працівників та установ польського ринку капіталу.

У першій половині 2011 року на Варшавській Фондовій Біржі мали лістинг 416 компаній, у тому числі 36 закордонних компаній з Голандії, Австрії, Угорщини, Німеччини, України, Словаччини, Республіки Чехії, США, Іспанії, Естонії. Загальна ринкова капіталізація лише національних компаній на кінець першої половині 2011 року перевищила 143 млрд євро. Разом з закордонними компаніями капіталізація ВФБ становила понад 206 млрд євро.

До виходу на Варшавську біржу заохочують прості процедури, низькі витрати, сприятливі законодавчі акти та доступність капіталу.

WIG Ukraine. Це перший і поки єдиний фондовий індекс поза Україною, у кошику якого є виключно компанії з нашої країни.

До складу індексу WIG Ukraine входять компанії, акції яких котируються на Головному ринку Варшавської фондової біржі, з офісом в Україні, або ж компанії, в яких українське підприємство є головною частиною холдингової структури з офісом поза Україною.

Частка акцій компаній у вільному обігу з WIG Ukraine повинна становити не менше 10%, а також компанії повинні належати до одного з наступних сегментів на Головному ринку WSE 250PLUS, 50PLUS або 5PLUS.

Базовою датою індексу є 31 грудня 2010 року, а його базова вартість становить 1 000 пунктів.

На 6.10.2011 до складу індексного кошику входять наступні компанії: Kernel Holding S.A. (KER, 42,267%), Astarta Holding N.V. (AST, 21,613%), Coal Energy S.A. (CLE, 9,422%), Milkiland N.V. (MLK, 5,784%), Agroton Public Limited (AGT, 4,937%), KSG Agro S.A. (KSG, 3,852%), Industrial Milk Company S.A. (IMC, 3,238%), Sadovaya Group S.A. (SGR, 3,234%), Ovostar Union N.V. (OVO, 2,967%), Westa ISIC S.A. (WES, 2,687%) (в дужках вказані котировки акцій на Варшавській фондовій біржі та доля в індексному кошику на 6.10.2011).

При апробації алгоритму використовувалась статистична інформація за період з 18.05.2011 до 6.10.2011 (об'єм вибірки N=100, якщо компанія торгувала на біржі до 18.05.2011); або з дня виходу компанії на біржу до 6.10.2011 (для CLE, OVO та WES; значення N для них дорівнює 56, 71 та 80 відповідно) .

Розглянемо детально моделювання цін та доходностей акцій на прикладі Kernel Holding S.A. (KER) та Westa ISIC S.A. (WES).

Діаграма розсіювання цін акції KER показана на рис. 1.

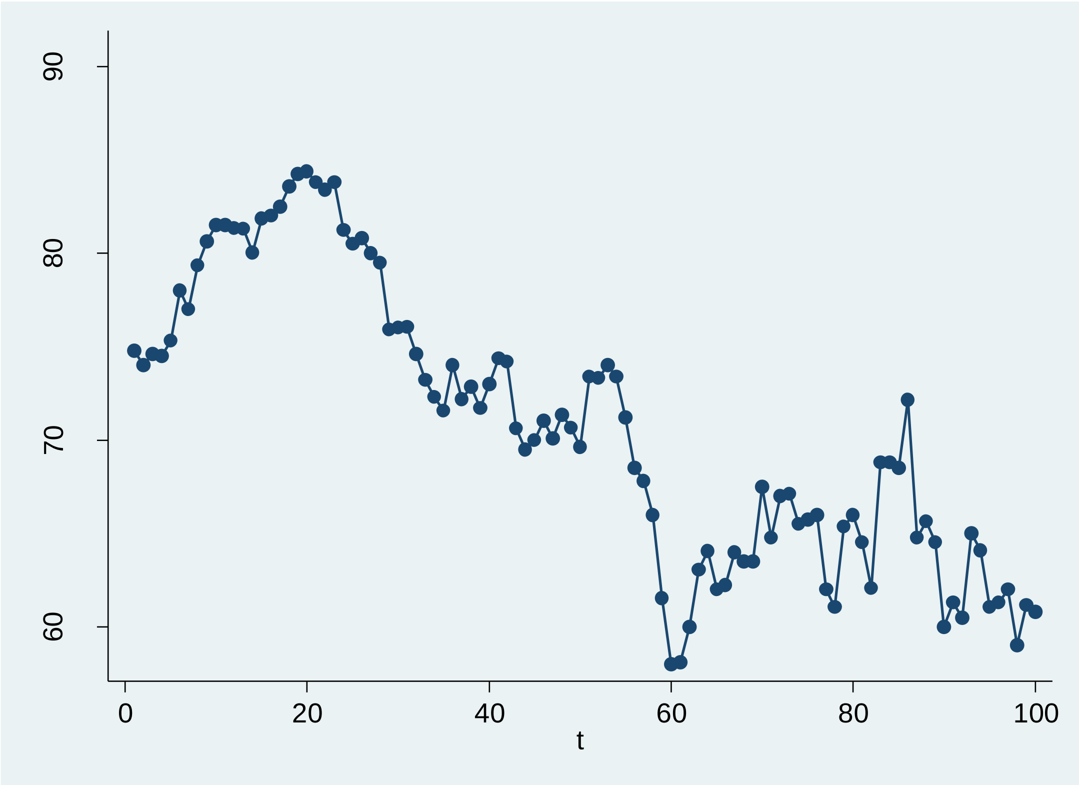


Рис.1. Статистичні дані цін акції KER

Проаналізуємо до якого класу належить даний часовий ряд: TS чи DS. Перевірка моделі (1) за розширеним тестом Дікі-Фуллера з використанням методу Доладо Дж., Дженкінса Т., Сосвіля – Ріверо С. та інформаційного критерію Шварца вказує на значущість оцінки коефіцієнта ρ (рівень значущості – 5%) при $\alpha \neq 0, \beta \neq 0$ і $\delta_i = 0\,(i = \overline{1, p})$. Це означає, що досліджуваний часовий ряд відноситься до TS класу. Зазначимо, що за допомогою інформаційного критерію Шварца встановлено: оптимальною є авторегресійна модель з одним лагом. Необхідно також перевірити залишки u_t регресійної моделі (1) на гетероскедастичність, нормальність та автокореляцію. Діаграма розсіювання залишків дає підстави для підозри щодо присутності гетероскедастичності. Використовуючи тест Бройша-

Паґана [20], підтверджуємо гіпотезу про гетероскедастичність залишків (Р-значення критерію дорівнює 12,27). Отже, не виконуються передумови розширеного тесту Дікі-Фуллера і тому використовуємо тест Філліпса-Перрона. Таким чином встановлюємо, що статистичні дані є стаціонарним часовим рядом відносно лінійного детермінованого тренду.

Розрахунки формують неперервну складову моделі (10) у вигляді

$$\hat{x}_t = \underset{(4,3220)}{13,2919} - \underset{(0,0135)}{0,0393}t + \underset{(0,0524)}{0,8388}x_{t-1}, \quad R^2 = 92,94. \tag{14}$$

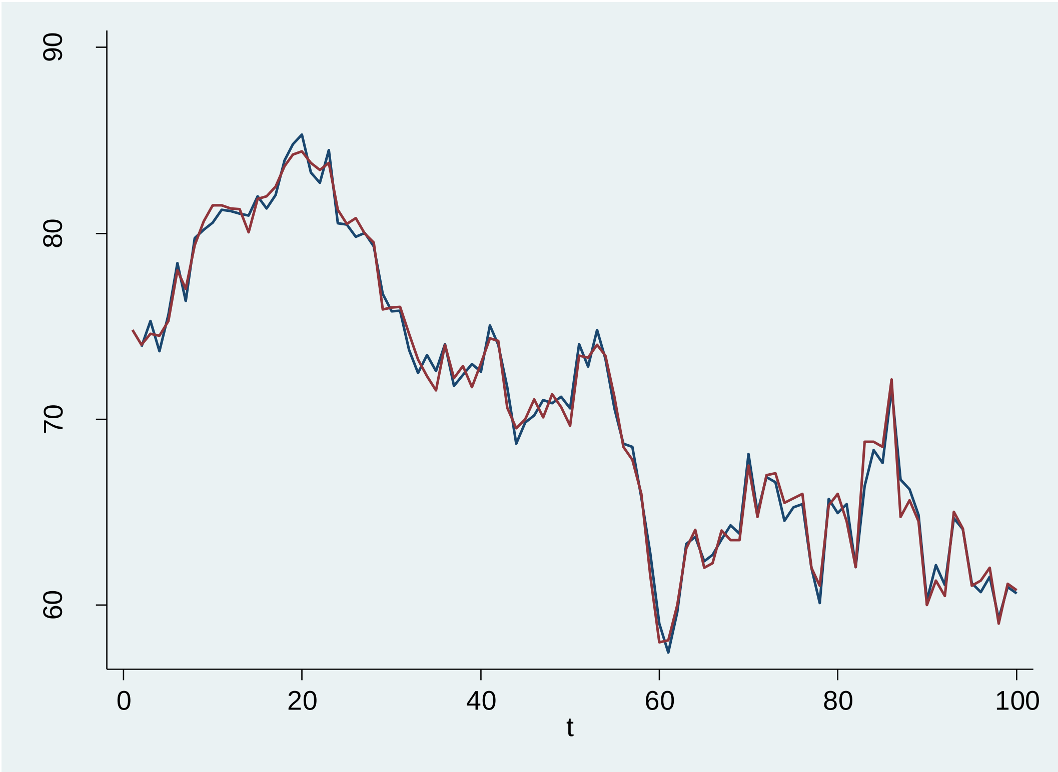
Тут під значенням МНК-оцінок вказані відповідні значення стандартних помилок цих оцінок. Перевірка значущості оцінок за критерієм Ст’юдента при рівні значущості 5% та відповідному числу ступенів вільності показує, що цьому критерію задовольняють всі оцінки.

Після нарощування за вказаним алгоритмом дискретної складової після двох ітерацій приходимо до наступної дискретно-неперервної моделі:

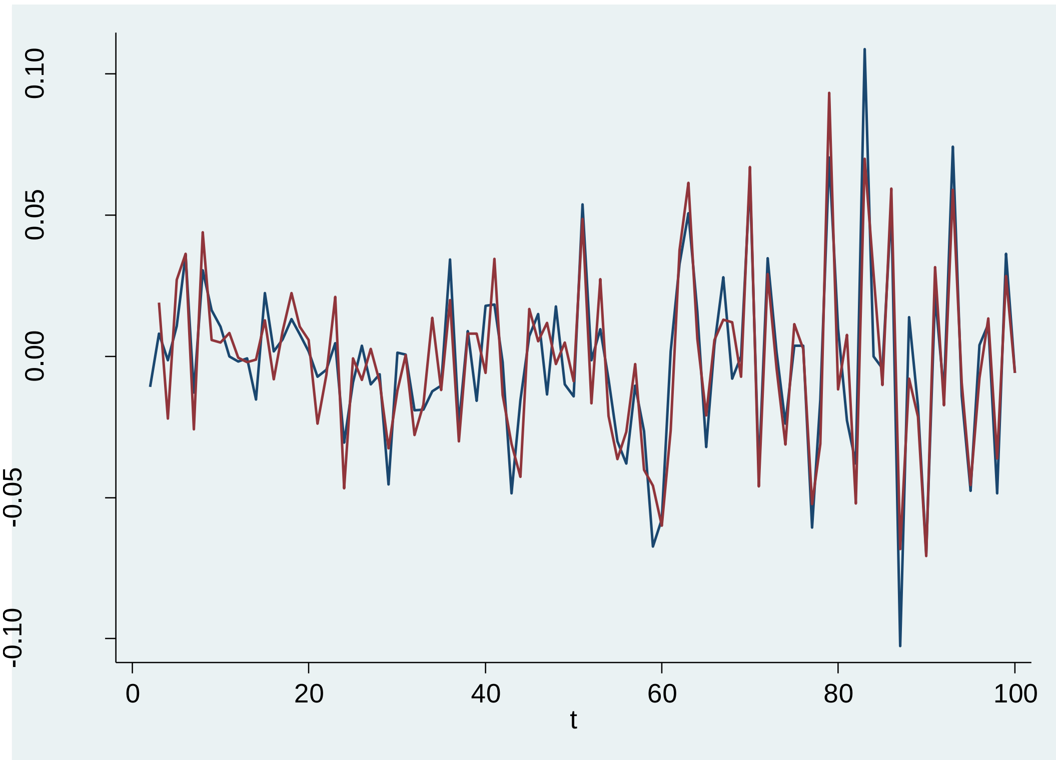
$$\underset{(c.n.)}{\hat{x}_t} = \underset{(4,3220)}{13,2919} - \underset{(0,0135)}{0,0393}t + \underset{(0,0524)}{0,8388}x_{t-1} + \underset{(0,0025)}{0,0397}z_1 + \underset{(0,0079)}{0,0908}z_2, \; R^2 = 99,15\%. \tag{15}$$

Аналіз залишків не виявляє автокорельованості (*P*-значення статистики Дарбіна-Уотсона дорівнює 2,08), значущих відхилень від нормального розподілу (*P*- значення критерію Жака-Бера дорівнює 0,27), гетероскедастичності (*P*-значення критерію Бройша-Паґана дорівнює 0,45), тому обґрунтованим є використання результатів *t*- і *F*-критеріїв [21].

На рис. 2 показані модельні та реальні значення цін акції та їх доходностей на досліджуваному проміжку часу.



(а)



(б)

Рис. 2 . Модельні та реальні значення цін акції KER та їх доходностей

Аналіз отриманих результатів вказує на високу якість апроксимації і тому дана модель може бути використана для прогнозування в точці $t^*=N+1$, що відповідає 7.10.2011. Для цього спочатку обчислюємо теоретичні значення ціни в точках досліджуваного періоду за формулою (15). Потім знаходяться значення фіктивних змінних в прогнозній точці t^* за допомогою множинної логіт-моделі (13). Після цього прогнозне значення розраховується за формулою (15) при $t = t^*$. У даному випадку отримуємо прогнозне значення 62,26. Зазначимо, що реальне значення на 7.10.2011 становить 62,75, тому відносна похибка прогнозу складає 0,79%.

Розглянемо тепер моделювання цін акцій компанії WES. Відповідна діаграма розсіювання показана на рис. 3.

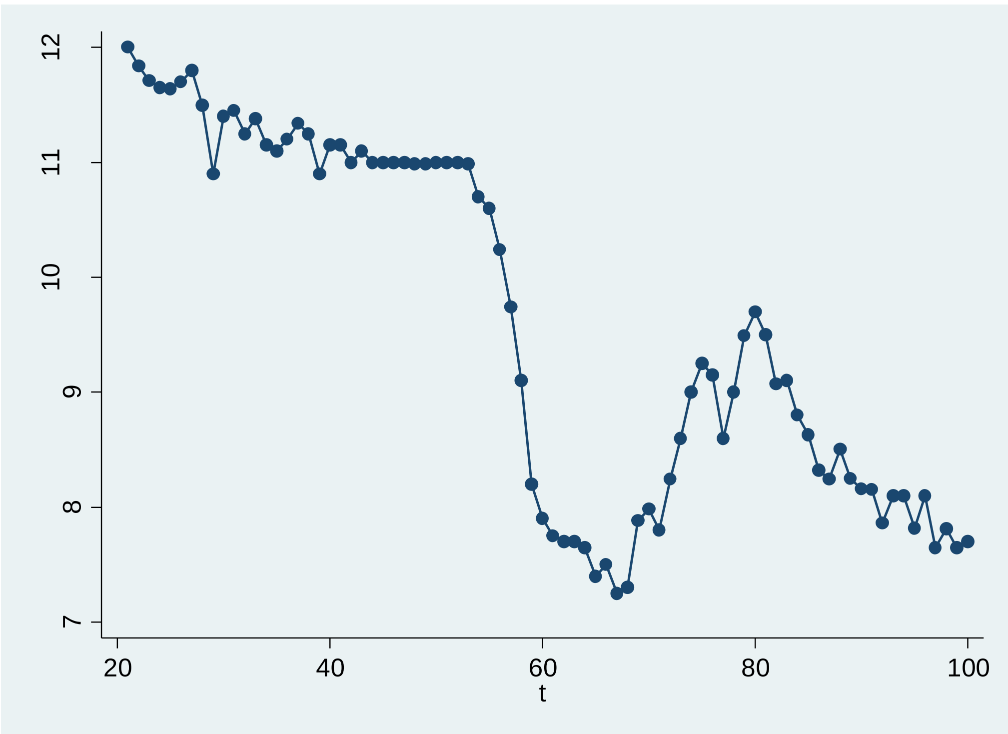


Рис. 3. Статистичні дані цін акції WES

Аналіз діаграми розсіювання вказує на належність даного часового ряду до DS класу, тому диференціюємо часовий ряд. Результати диференціювання показані на рис. 4.

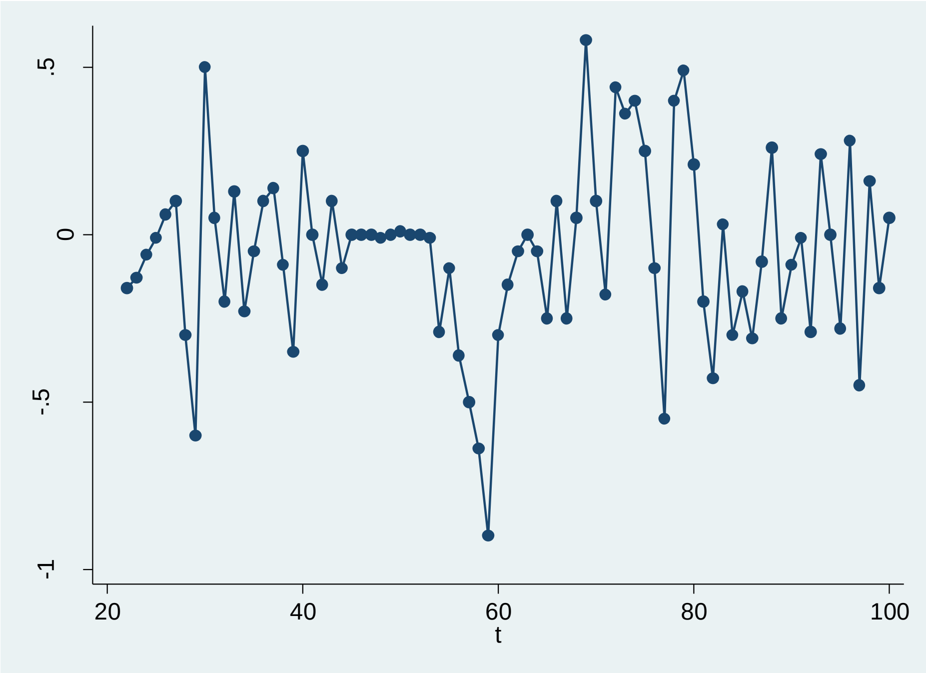


Рис. 4. Статистичні дані перших різниць цін акції WES

Використання описаної вище процедури встановлення стаціонарності часового ряду вказує на те, що часовий ряд, який відповідає рис. 4, є стаціонарним вже при $n = 2$.

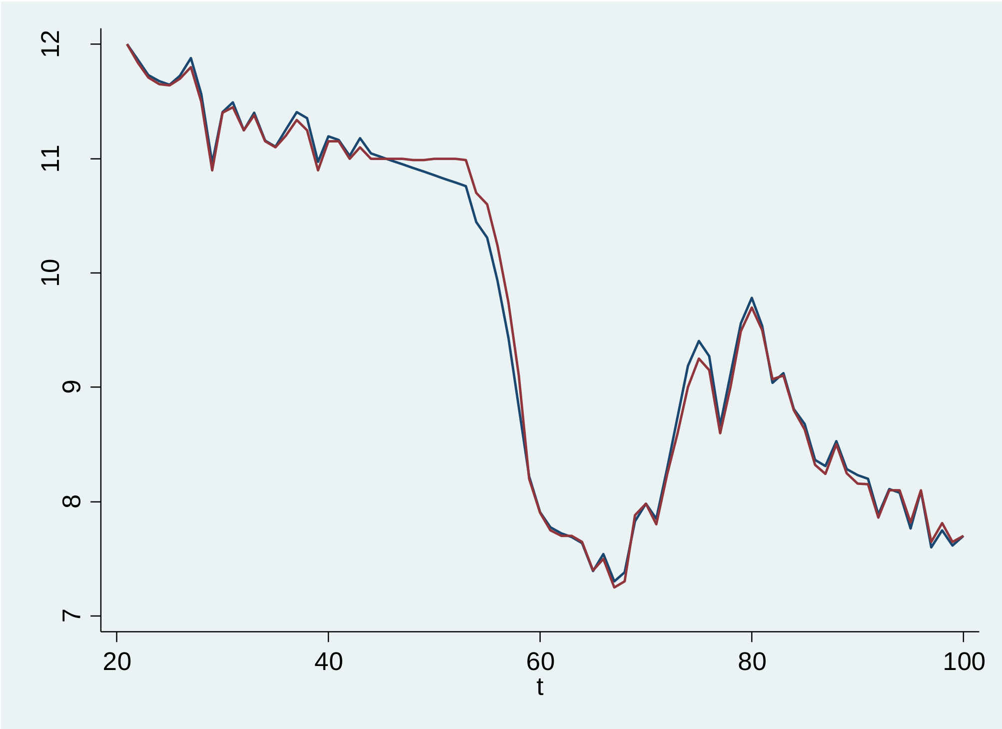
Відповідна неперервна складова моделі наступна:

$$\Delta\hat{x}_t = -\underset{(0,0415)}{0{,}0416} + \underset{(0,0922)}{0{,}2063}\,\Delta x_{t-1}. \qquad (16)$$

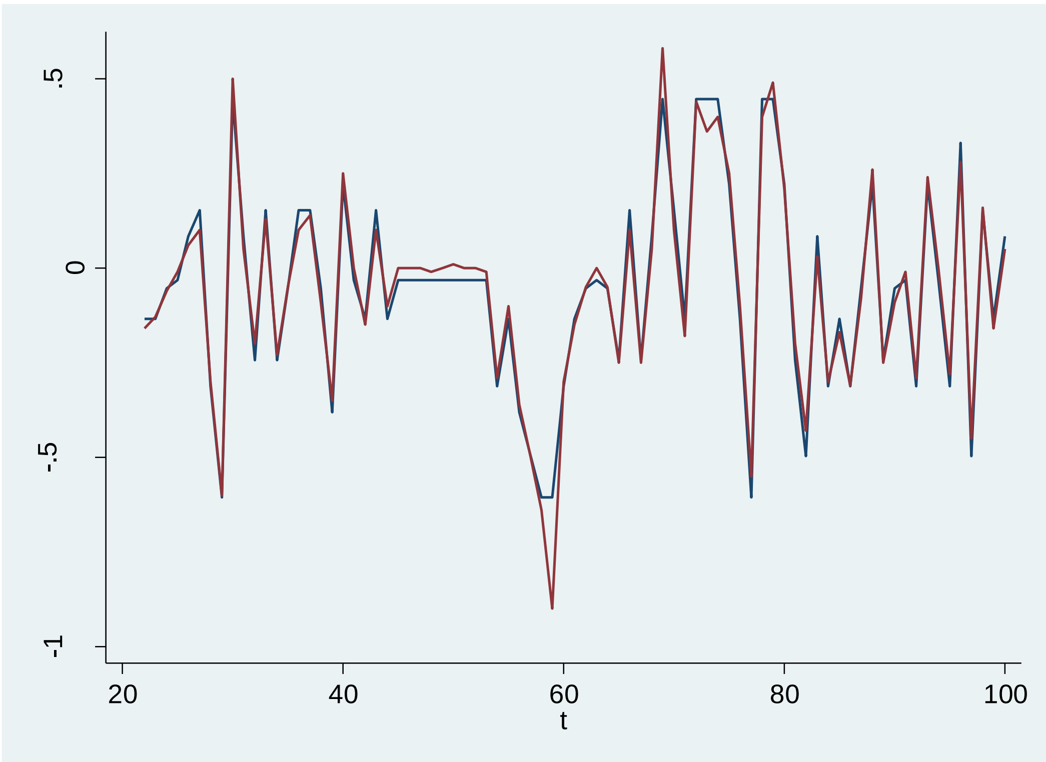
Після нарощування дискретної складової за вказаним алгоритмом після трьох ітерацій дискретно-неперервна модель приймає вигляд:

$$\underset{(c.n.)}{\Delta\hat{x}_t} = -\underset{(0,0415)}{0{,}0416} + \underset{(0,0922)}{0{,}2063}\,\Delta x_{t-1} + \underset{(0,0007)}{0{,}0099}\,z_1 + \underset{(0,0007)}{0{,}0102}\,z_2 + \underset{(0,0061)}{0{,}0345}\,z_2. \qquad (17)$$

Аналіз залишків цієї моделі не виявляє автокорельованості (*P*-значення статистики Дарбіна-Уотсона дорівнює 2,11), значущих відхилень від нормального розподілу (*P*- значення критерію Жака-Бера дорівнює 1,56), гетероскедастичності (*P*-значення критерію Бройша-Паґана дорівнює 6,30), тому обґрунтованим є використання результатів *t*- і *F*-критеріїв.



(а)



(б)

Рис. 5. Модельні та реальні значення цін акції WES (*а*) та їх різниць (*б*)

На рис. 5 показані модельні та реальні значення цін акції WES та їх різниць на досліджуваному проміжку часу.

Аналіз рис. 4 вказує на досить високі імітаційні властивості моделі (17). Використовуючи відповідну множинну логіт – модель і розраховуючи теоретичні (на досліджуваному періоді) та прогнозне значення доходностей (приростів), отримаємо прогнозне значення ціни, яке дорівнює 7,65. Реальне значення становить 7,69, тому відносна похибка прогнозу складає 0,52%.

Табл. 1. Прогнозні та реальні значення цін українських компаній

	Реальні значення	Прогнозні значення	Відносна похибка прогнозу, %	Клас часового ряду	Кількість ітерацій
KER	62,75	62,26	0,79	TS	2
AST	69,75	70,13	0,54	TS	2
MLK	20,85	21,24	1,84	DS	3
AGT	25,4	25,08	1,28	DS	3
CLE	24,25	24,24	0,04	TS	2
KSG	23,35	24,04	2,87	TS	2
SGR	9,15	9,19	0,44	TS	2
IMC	9,00	8,99	0,11	TS	2
OVO	58,60	58,92	0,54	DS	3
WES	7,69	7,65	0,52	DS	3

За вказаною методикою було проведено моделювання цін акцій інших зазначених компаній. В таблиці 1 наведені реальні та прогнозні значення цін українських компаній.

ВИСНОВКИ

У даній роботі розглянутий підхід, який дозволяє моделювати фінансові часові ряди з високими імітаційними та прогнозними властивостями. Враховується нестаціонарність вихідного часового ряду і пропонується процедура зведення дослідження до аналізу стаціонарного часового ряду. Особливістю даної роботи є побудова неперервно-дискретних моделей з використанням тризначних фіктивних змінних, що дозволяють отримувати високі апроксимаційні властивості за 2-3 ітерації. Апробація запропонованої методики на прикладі цін та доходностей акцій українських компаній на Варшавській фондовій біржі показала можливість її використання на практиці.

ЛІТЕРАТУРА

1. Гитман Л.Дж., Джонк М.Д. Основы инвестирования. – М: Дело, 1997. – 991 с.

2. Sharpe W.F. Portfolio Theory and Capital Markets. – N.Y.: McGrawffill, 1970. – 1027 c.

3. Гибсон Р. Формирование инвестиционного портфеля: управление финансовыми рисками. – М.: Альпина Бизнес Букс, 2008. – 276 с.

4. Айвазян С.А., Мхитарян B.C. Прикладная статистика и основы эконометрики. — М.: Юнити, 1998.

5. Gurierous C., Monfort A. Time Series and Dynamic Models. – Cambridge University Press, 1997. – 668 p.

6. Roberto R. Identification Through Heteroskedasticity // The Review of Economics and Statistic. – 2003. – Vol. 85, No. 4. – P. 777–792.

7. Hamilton J. Time Series Analysis. – Princeton University Press, 1994. – 799 p.

8. Носко В. П. Введение в регрессионный анализ временных рядов. –Москва, 2002. – 273с.

9. Clements M. P., Hendry D. F. Forecasting with difference-stationary and trend-stationary models // Econometrics Journal. – 2001. – No. 4. – P. 1–19.

10. Канторович Г. Г. Анализ временных рядов // Экономический журнал ВШЭ. – 2002. –№ 3. – С. 379 – 401.

11. http://www.bseu.by/russian/faculty5/stat/docs/4/EconometricsBook3.pdf.

12. Rossen A. On the predictive content of nonlinear transformations of lagged autoregression residuals and time series observations // Hamburg Institute of International Economics, Research Paper 113. – 2001. – P. 1–24 .

13. Patterson K. An Introduction to Applied Econometrics: A Time Series Approach. –Palgrave, 2000. – 797 p.

14. Dolado J.J. Cointegration and Unit Roots / T. Jenkinson, S. Sosvilla-Rivero // Journal of Economic Survey. – 1990. – Vol. 4. – P. 249–273.

15. Мартынова М.А. Инвестиционные решения в пространстве риск-устойчивых стратегий. – Воронеж: Изд-во Воронеж. гос. ун-та, 2009. – 141с.

16. Давнис В. В. Прогнозные модели экспертных предпочтений. – Воронеж: Изд-во Воронеж. гос. ун-та, 2005. – 248с.

17. Москвин С. Фондовий ринок України: місце зустрічі можна змінити // Дзеркало тижня. – 2010. – № 14. – С. 78–79.

18. Назарчук М. І. Аналіз стану та перспективи розвитку фондового ринку України// Фінанси України. –2007. – №12. – С. 83–95.

19. Міньков В. І. Деякі особливості розвитку фондового ринку України // Фінанси України. – 2005. – № 12. – С. 104–114.

20. Gujarati D. Basic Econometrics. –McGraw-Hill/Irwin, 2002.–1002 p.

21. Назаренко О.М. Основи економетрики. – К.: Центр навчальної літератури, 2005. –392 с.