

УДК 004.4:519.85

БІДЮК П. І.¹, ДЕМКІВСЬКИЙ Є. О.², ДЕМКІВСЬКА Т. І.³

¹Інститут прикладного системного аналізу НТУУ «КПІ» ім. Ігоря Сікорського

²Київський національний університет імені Тараса Шевченка

³Київський національний університет технологій та дизайну

МЕТОДИКА ПОБУДОВИ МОДЕЛЕЙ ГЕТЕРОСКЕДАСТИЧНИХ ПРОЦЕСІВ

Мета. Метою дослідження є вдосконалення методики побудови гетероскедастичних моделей часових рядів, яка ґрунтується на використанні кореляційного аналізу та множини статистичних характеристик моделей.

Методика ґрунтується на принципах системного аналізу у застосуванні до оцінювання структури і параметрів моделей за допомогою статистичних даних. Спочатку статистичні дані попередньо обробляються: фільтруються, нормуються, заповнюються пропуски і, за необхідності, виміри розмножуються. Далі на основі кореляційного аналізу виконується оцінювання структури моделі і реалізується оцінювання її параметрів за допомогою коректно вибраного методу. Оцінювання адекватності моделі та якості прогнозів виконується за вибраними статистичними критеріями.

Результати. Удосконалено методику побудови моделей гетероскедастичних процесів, яка забезпечує отримання адекватних моделей динаміки дисперсії за умови наявності інформативних даних. Наведено приклад побудови адекватної моделі динаміки дисперсії гетероскедастичного процесу. Виконані обчислювальні експерименти, спрямовані на моделювання динаміки умовної дисперсії. Проілюстровано ефективність запропонованої методики моделювання процесів, нестационарних стосовно дисперсії. Отримані характеристики побудованих моделей УАРУТ свідчать про можливість їх використання для прогнозування умовної дисперсії на практиці.

Наукова новизна. Запропоновано системний підхід до моделювання нелінійних нестационарних процесів з використанням статистичних даних та принципів системного аналізу. Побудовано нову модель досліджуваного фінансового процесу, яка забезпечує отримання високоякісного короткострокового прогнозу волатильності.

Практична значимість. Розроблена методика моделювання гетероскедастичних процесів може бути застосована для аналізу процесів такого типу у різних галузях діяльності для створення адекватних моделей та оцінювання прогнозів. Програмна реалізація методики забезпечує побудову високоякісних моделей.

Ключові слова: гетероскедастичний процес, волатильність, побудова моделі, критерії адекватності, автокореляційна функція, часткова автокореляційна функція.

Вступ. В даний час, в умовах зростання кількості фінансових ринків та збільшення кількості їх учасників, є актуальною побудова моделей, які адекватно описують фінансові та фінансово-економічні процеси з метою прийняття оптимальних рішень для врахування і мінімізації можливих ризиків та отримання максимальних прибутків.

Такі процеси можуть бути представлені у формі часових рядів. Тому виникає необхідність удосконалювати існуючі методики моделювання таких процесів. Дана робота присвячена дослідженню нестационарних процесів, представлених у формі часових рядів. В роботі розглядається нестационарність у формі гетероскедастичності, тобто дисперсія досліджуваного процесу змінюється у часі на інтервалі дослідження. Моделі, що описують динаміку дисперсії, дають можливість виконувати її короткострокове прогнозування, яке знаходить застосування у системах торгівлі на біржі, у менеджменті фінансових ризиків і т.

ін. Для дослідження використовується інтелектуальний аналіз даних (статистичний аналіз), шляхом оцінювання структури і параметрів математичних.

Постановка завдання. Завданнями даного дослідження є такі: вдосконалення методики побудови моделей гетероскедастичних процесів у формі часових рядів, яка ґрунтується на використанні кореляційного аналізу даних, множини статистичних характеристик моделі та застосуванні принципів системного аналізу; ілюстрація застосування методики моделювання нестационарних гетероскедастичних процесів до аналізу фактичних фінансових даних.

Методика ґрунтується на принципах системного аналізу даних, що застосовуються до оцінювання структури і параметрів моделей за допомогою статистичних даних. Для моделювання гетероскедастичного процесу необхідно визначити структуру моделі та оцінити її параметри, побудувати декілька моделей-кандидатів та визначити кращу модель, яка описує процес, використавши множину статистичних характеристик моделі.

Результати дослідження. Для моделювання екологічних, економічних та технологічних процесів використовується методика Дженкінса і Бокса [1]. Дана методика, яка ґрунтується на аналізі автокореляційної та часткової автокореляційної функції і застосовується для побудови моделей стаціонарних процесів (моделей авторегресії та моделей ковзного середнього) та нестационарних процесів (моделей авторегресії та інтегрованого ковзного середнього).

Інший підхід до моделювання процесів, які описані часовими рядами, запропоновано в [2]. Цей підхід побудований на використанні інтелектуального аналізу даних (кореляційного аналізу) для оцінювання структури і параметрів регресійних моделей часових рядів. Для оцінювання параметрів моделей пропонується використовувати метод найменших квадратів (МНК) і метод максимальної правдоподібності (ММП) та, інші нелінійні методи і універсальний метод Монте-Карло для марковських ланцюгів [3].

Гетероскедастичні процеси виникають при дослідженні деяких технічних, екологічних та фінансово-економічних процесів, особливо в умовах нестійкої економіки. Ці процеси відносять до слабо стаціонарних процесів. До відомого підходу моделювання таких процесів належить метод групового урахування аргументів (МГУА) [4]. Питання побудови моделей нестационарних процесів розглядається в [5 - 9]. При моделюванні гетероскедастичних процесів спочатку виконують перевірку на наявність гетероскедастичності, яка потребує додаткових обчислень. Опис тестів на гетероскедастичність наведено в [10 - 12].

Відомо, що деякі технічні, екологічні та фінансово-економічні процеси мають досить мінливий характер, тобто їх параметри постійно зростають або спадають. Волатильність – це ступінь мінливості параметра у часі. Мірою волатильності є дисперсія або середньоквадратичне відхилення.

Якщо дисперсія процесу змінюється (зменшується або збільшується) в часі, то такий процес називають гетероскедастичним, якщо дисперсія не змінюється, то такий процес називається гомоскедастичним.

$$\text{var}(z_i) = \delta_z^2 = \text{const}$$

Для моделювання гетероскедастичного процесу необхідно знайти закон, за яким змінюється дисперсія.

Розглянемо гетероскедастичний процес:



Рис.1. Графік гетероскедастичного процесу

Математичний опис динаміки дисперсії формально ґрунтується на рівняннях авторегресії та авторегресії з ковзним середнім. Подальше ускладнення таких моделей приводить до експоненційної узагальненої моделі авторегресії з умовною гетероскедастичністю. Для того щоб забезпечити скінченність значень умовної дисперсії, корені характеристичного рівняння моделі (для її авторегресійної складової) повинні знаходитися всередині кола одиничного радіуса.

Для побудови моделі авторегресійної умовно гетероскедастичної моделі (АРУГ) пропонується така послідовність кроків:

1. Побудувати авторегресійну модель низького порядку, наприклад, АР(1) з метою виділення випадкової складової досліджуваного процесу.
2. Побудувати ряд залишків моделі $\hat{\varepsilon}(k)$
3. Побудувати автокореляційну (АКФ) та часткову автокореляційну функцію (ЧАКФ) для ряду залишків моделі
4. Згенерувати новий ряд із квадратів залишків.
5. Побудувати корелограму для ряду, сформованого із квадратів залишків.

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.404	0.404	16.655	0.000
		2	0.025	-0.165	16.720	0.000
		3	-0.029	0.032	16.809	0.001
		4	-0.022	-0.020	16.859	0.002
		5	-0.019	-0.008	16.896	0.005
		6	-0.022	-0.015	16.949	0.009
		7	-0.005	0.011	16.951	0.018
		8	0.000	-0.006	16.951	0.031
		9	-0.020	-0.024	16.997	0.049
		10	-0.038	-0.023	17.156	0.071

Рис 2. Автокореляційна та часткова автокореляційна функція для побудови моделі АРУГ фінансового процесу

5. Аналізуємо ЧАКФ ряду, сформованого із квадратів залишків. Оскільки найбільшу значимість має лаг з номером 1, то ми можемо включити його в модель:

$$\varepsilon^2(k) = \alpha_0 + \alpha_1 \varepsilon^2(k-1) + v(k),$$

де $v(k)$ – випадковий процес, утворений залишками наведеної моделі.

За допомогою МНК обчислюємо оцінки коефіцієнтів рівняння першого порядку для дисперсії залишків.

Отримали таку модель умовної дисперсії:

$$\varepsilon^2(k) = 3.371626 + 0.404188 * \varepsilon^2(k-1)$$

Наводимо статистичні характеристики даної моделі:

R-squared	0.163358	Mean dependent var	5.661678
Adjusted R-squared	0.154643	S.D. dependent var	25.72049
S.E. of regression	23.64828	Akaike info criterion	9.184655
Sum squared resid	53687.14	Schwarz criterion	9.237409
Log likelihood	-448.0481	Durbin-Watson stat	1.866692

Рис. 3. Статистичні характеристики моделі моделі АРУГ (1)

Наступною моделлю, яку можна побудувати згідно із значеннями розрахованої часткової автокореляційної функції, є АРУГ(2):

$$\varepsilon^2(k) = \alpha_0 + \alpha_1 \varepsilon^2(k-1) + \alpha_2 \varepsilon^2(k-2)$$

За допомогою МНК обчислюємо коефіцієнти і отримуємо таку модель:

$$\varepsilon^2(k) = 3.973860 + 0.470585 * \varepsilon^2(k-1) - 0.165058 * \varepsilon^2(k-2).$$

R-squared	0.185958	Mean dependent var	5.719935
Adjusted R-squared	0.168638	S.D. dependent var	25.84761
S.E. of regression	23.56761	Akaike info criterion	9.188063
Sum squared resid	52210.63	Schwarz criterion	9.267693
Log likelihood	-442.6211	Durbin-Watson stat	1.989608

Рис. 4. Статистичні характеристики моделі моделі АРУГ (2)

Оскільки отримана модель адекватна за коефіцієнтом Дарбіна Ватсона 1,98, але показує неадекватність за коефіцієнтом детермінації 0,185, для підвищення адекватності гетероскедастичного процесу розширюємо модель, включивши до її складу авторегресійну складову стосовно умовної дисперсії $h(k)$, яка обчислюється за наявними статистичними даними. Припускаємо, що дисперсія збурюється процесом у вигляді білого шуму.

Якщо умовна гетероскедастичність в процесі наявна, то корелограма повинна вказувати на цей факт. Алгоритм аналізу корелограми квадратів залишків можна записати таким чином:

1. Побудувати модель АРКС (або АР) для послідовності $\{y(k)\}$ і побудувати додатковий ряд із квадратів похибок $\hat{\varepsilon}^2(k)$. Обчислити вибірку (на основі значень елементів ряду) дисперсію залишків

$$\hat{\sigma}^2 = \sum_{k=1}^N \frac{\hat{\varepsilon}^2(k)}{N-1},$$

де N – кількість елементів ряду похибок.

2. Визначити вибірку АКФ для ряду квадратів залишків

$$\rho(s) = \frac{\sum_{k=s+1}^N [\hat{\varepsilon}^2(k) - \hat{\sigma}^2][\hat{\varepsilon}^2(k-s) - \hat{\sigma}^2]}{\sum_{k=1}^N [\hat{\varepsilon}^2(k) - \hat{\sigma}^2]^2}.$$

3. Побудова ряду умовної дисперсії для цього ряду даних виконується так:

$$h(k) = \frac{1}{w-1} \sum_{i=k-\frac{w-1}{2}}^{k+\frac{w-1}{2}} [y(i) - \bar{y}]^2,$$

де $k = 2, 3, 4, \dots, w$ – розмір рухомого вікна, за яким розраховується умовна дисперсія.

Для отриманого ряду значень умовної дисперсії автокореляційна та часткова автокореляційна функція мають такий вигляд.

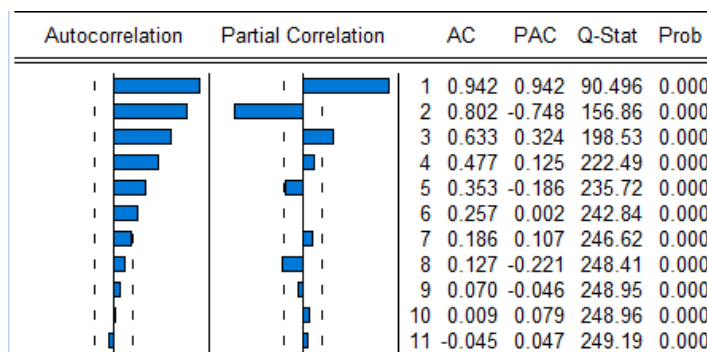


Рис. 5. Автокореляційна та часткова автокореляційна функція для побудови моделі УАРУГ

З аналізу ЧАКФ видно, що в модель УАРУГ можна включити лаги 1, 2, 3. Серед множини побудованих моделей найкращою є модель УАРУГ(3,2):

$$HH = C(1) + C(2) * HH(-1) + C(3) * HH(-2) + C(4) * HH(-3) + C(5) * RESID_2(-1) + C(6) * RESID_2(-2)$$

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	1.960486	1.444592	1.357121	0.1781
C(2)	2.078621	0.099025	20.99095	0.0000
C(3)	-1.604740	0.180503	-8.890393	0.0000
C(4)	0.471638	0.101540	4.644846	0.0000
C(5)	-0.062267	0.057662	-1.079866	0.2831
C(6)	0.046292	0.057669	0.802723	0.4242
R-squared	0.965701	Mean dependent var		36.16667
Adjusted R-squared	0.963795	S.D. dependent var		62.88463
S.E. of regression	11.96538	Akaike info criterion		7.862373
Sum squared resid	12885.32	Schwarz criterion		8.022645
Log likelihood	-371.3939	Durbin-Watson stat		1.983704

Рис. 6. Статистичні характеристики моделі моделі УАРУГ(3,2)

В результаті виконаних обчислень побудовано таку модель:

$$h(k) = 1.960486 + 2.078621 * h(k-1) - 0.1604740 * h(k-2) + 0.471638 * h(k-3) \\ + 0.062267 * \varepsilon^2(k-1) + 0.046292 * \varepsilon^2(k-2)$$

З аналізу статистичних характеристик отриманої моделі видно, що дана модель краще описує гетероскедастичний процес, оскільки значення характеристики Дарбіна Ватсона наближається до 2, що означає відсутність автокореляції залишків моделі, а коефіцієнт детермінації, який показує ступінь адекватності моделі, дорівнює 0,96, що є прийнятним значенням для прогнозуючої моделі.

Висновки. Гетероскедастичні процеси широко розповсюджені у фінансах та інших галузях досліджень. Необхідність побудови моделей гетероскедастичних процесів зумовлена тим, що у багатьох випадках необхідно знаходити прогнози оцінок умовної дисперсії, наприклад, для виконання торговельних операцій, для оцінювання фінансових ризиків та для використання у діагностичних системах. Запропонована удосконалена методика моделювання динаміки дисперсії гетероскедастичних процесів. Виконані обчислювальні експерименти, спрямовані на моделювання динаміки умовної дисперсії, ілюструють ефективність запропонованої методики моделювання процесів, нестационарних стосовно дисперсії. Отримані характеристики побудованих моделей УАРУГ свідчать про можливість їх використання для прогнозування умовної дисперсії на практиці.

Література

1. Бокс Дж. Анализ временных рядов. Прогноз и управление: Вып. 1 / Дж. Бокс, Г. Дженкинс. – М.: Мир, 1974. – 406 с.
2. Бідюк П.І., Романенко В.Д., Тимошук О.Л. Аналіз часових рядів: Навчальний посібник. – Київ.: НТУУ КІП, 2013. – 599 с.
3. Tsay R.S. Analysis of financial time series. – Hoboken, New Jersey: John Wiley & Sons Inc., 2010. – 715 p.
4. Ивахненко А.Г. Индуктивный метод самоорганизации моделей сложных систем. – Киев: Наукова думка, 1982. – 296 с.
5. Себер Дж. Прикладной регрессионный анализ. – М.: Мир, 1982. – 450 с.
6. Бард Й. Нелинейное оценивание параметров. – М.: Статистика, 1979. – 349
7. Дрейпер Н., Смит Г. Прикладной регрессионный анализ (т. 2). – М.: Финансы и статистика, 1986. – 366 с.
8. Лукашин Ю.П. Адаптивные методы краткосрочного прогнозирования временных рядов. М.: Финансы и статистика, 2003, 416 с.
9. Enders W. Applied Econometrics Time Series : 3nd ed. / W. Enders. – Jon Wiley and Sons, Inc., – 2010. – 517 p.
10. Enders W. Applied econometric time

References

1. Dzh. Boks, G. Dzhenkins, (1974). Analiz vremennykh ryadov. Prognoz i upravlenie [Time Series Analysis. Forecast and Management] Monografiya M.: Mir [in Russian]
2. Bidiuk P.I., Romanenko V.D., Tymoshchuk O.L., (2013) Analiz chasovykh riadiv [Time Series Analysis] Kyiv.: NTUU KPI [in Ukrainian]
3. Tsay R.S., (2010). Analysis of financial time series. – Hoboken, New Jersey: John Wiley & Sons Inc., 715 [in English]
4. Ivakhnenko, A.G., (1982). Induktivnyy metod samoorganizatsii modeley slozhnykh sistem [Inductive method of self-organization of complex systems models] Kyiv: Nauk. Dumka [in Russian]
5. Seber. Dzh., (1982). Prikladnoy regressionnyy analiz [Applied Regression Analysis] M.: Mir [in Russian]
6. Bard, Y., (1979). Nelineynoe otsenivanie parametrov [Non-linear parameter estimation] M.: Statistika [in Russian]
7. Dreyper, N., Smit, G., Prikladnoy regressionnyy analiz (t.2) [Applied Regression Analysis] M.: Finansy i statistika [in Russian]
8. Lukashin Yu.P., (2003), Adaptivnye metody kratkosrochnogo prognozirovaniya vremennykh ryadov. [Adaptive methods of short-term forecasting of time series], . M.: Finansy i statistika, 416 [in Russian]
9. Enders W, (2010), Applied Econometrics Time Series 3rd ed. /W. Enders. – Jon Wiley and Sons, Inc., 517 [in English]
10. Enders W., (1995), Applied econometric time series. –

series. – New York: John Wiley & Sons, Inc., 1995. – 434 p.

11. Johnston J., DiNardo J. Econometric methods. – New York: McGraw-Hill, Inc., 1997. – 530 c.

12. Демківський О.Б., Красницький С.М., Пилипенко Ю.М., Слізков А.М. – Статистична обробка результатів вимірювань та експериментальних даних в текстильній промисловості: Навчальний посібник. – Київ.: КНУТД, 2012. – 105 с.

New York: John Wiley & Sons, Inc., 434 [in English]

11. Johnston J., DiNardo J., (1997), Econometric methods. – New York: McGraw-Hill, Inc., – 530 [in English]

12. Demkivskyi, O.B., Krasnytskyi, S.M., Pylypenko, Yu.M., Slizkov A.M., (2012)– *Statystychna obrobka rezultativ vymyriuvan ta eksperymentalnykh danykh v tekstylnii promyslovosti* [Statistical processing of measurement results and experimental data in the textile industry] Kyiv. Kyiv National University of Technologies and Design [in Ukrainian]

BIDYUK PETRO

pbidyuke_00@ukr.net

ID: 6602445011

ORCID: (<http://orcid.org/>) 0000-0002-7421-3565

ResearcherID:

(<http://www.researcherid.com>) S-1161-2016

Institute for applied system analysis National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute»

DEMKIVSKIY YEVGEN

demkivsky@gmail.com

ORCID: (<http://orcid.org/>) 0000-0002-9383-7466

Taras Shevchenko National University of Kyiv

DEMKIVSKA TETIANA

demkivskiy@gmail.com

ORCID: (<http://orcid.org/>) 0000-0002-2176-163X
Kyiv National University of Technologies & Design

МЕТОДИКА ПОСТРОЕНИЯ МОДЕЛЕЙ ГЕТЕРОСКЕДАСТИЧЕСКИХ ПРОЦЕССОВ

БИДЮК П. И.¹, ДЕМКОВСКИЙ С. О.², ДЕМКОВСКАЯ Т. И.³

¹Институт прикладного системного анализа НТУУ «КПИ» им. Игоря Сикорского

²Киевский национальный университет имени Тараса Шевченка

³Киевский национальный университет технологий и дизайна

Цель. Целью исследования является совершенствование методики построения гетероскедастических моделей временных рядов, основанной на использовании корреляционного анализа данных и множества статистических характеристик модели.

Методика основывается на принципах системного анализа данных в применении к оценке структуры и параметров моделей с помощью статистических данных. Для моделирования гетероскедастических моделей временных рядов необходимо решить следующие задачи: воспользоваться принципами системного анализа для формулирования методики моделирования, обеспечить надлежащую подготовку данных для моделирования, оценить структуру и параметры моделей, а также сформировать критериальную базу для корректного выбора лучших моделей из множества оцененных кандидатов.

Результаты. Усовершенствована методика построения моделей гетероскедастических процессов, которая обеспечивает получение адекватных моделей динамики дисперсии при наличии информативных данных. Приведен пример построения адекватной модели динамики дисперсии гетероскедастического процесса. Выполнены вычислительные эксперименты, направленные на моделирование динамики условной дисперсии. Проиллюстрировано эффективность предложенной методики моделирования процессов, нестационарных относительно дисперсии. Полученные характеристики построенных моделей УАРУГ свидетельствуют о возможности их использования для прогнозирования условной дисперсии на практике.

Научная новизна. Показана возможность применения системного подхода к построению моделей гетероскедастических временных рядов и приведена последовательность выполнения операций при построении моделей указанного типа. Построена новая модель исследуемого

финансового процесса, обеспечивает получение высококачественного краткосрочного прогноза волатильности.

Практическая значимость. Предложенная методика обеспечивает построение моделей приемлемой адекватности при условии полноты и информативности статистических данных, представленных временными рядами.

Ключевые слова: временной ряд, гетероскедастический процесс, волатильность, построение модели, критерии адекватности, автокорреляционная функция, частичная автокорреляционная функция.

METHOD OF DEVELOPMENT OF HETEROSKEDASTIC PROCESSES MODELS BIDIUK P. I.¹, DEMKIVSKYI Y. O.², DEMKIVSKA T. I.³

¹Institute for applied system analysis National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute»

²Taras Shevchenko National University of Kyiv

³Kyiv National University of Technologies and Design

Purpose. The aim of the study is to improve the method of construction of heteroscedastic models of time series, which is based on the use of correlation analysis of data and multitude of statistical characteristics of the model.

Methodology is based on the principles of systematic data analysis in application to the estimation of structure and parameters of models using statistical data. For modelling of heteroscedastic time series models need to solve the following tasks: to apply system analysis principles to formulate the modeling methodology, to ensure proper modeling data preparation, to evaluate the structure and parameters of the models, and to form a criterion base for the correct selection of the best candidate models from the set of candidates evaluated.

Findings. The technique of constructing models of heteroskedastic processes is improved, which ensures obtaining adequate models of the dynamics of dispersion with the presence of informative data. An example of constructing an adequate model of the dispersion dynamics of the heteroscedastic process is given. Computational experiments aimed at modeling the conditional dispersion dynamics are performed. The efficiency of the proposed technique for modeling non-stationary variance processes is illustrated. The obtained characteristics of the constructed GARCH - models indicate that they can be used to predict conditional variance in practice.

Originality. The possibility of applying a systematic approach to the construction of heteroskedastic time series models is shown, and the sequence of operations is performed when constructing models of this type. A new model of the studied financial process has been constructed to provide high quality short-term volatility forecast.

Practical value. The proposed methodology ensures the construction of models of acceptable adequacy, provided the completeness and informativeness of the statistics presented in time series.

Keywords: time series, heteroscedastic process, volatility, model construction, adequacy criteria, autocorrelation function, partial autocorrelation function.