УДК 519.688+517.9
JEL Classification: F59
DOI: 10.15587/2312-8372.2019.182109

ВИБІР ПОРЯДКУ РЕГРЕСІЙНОЇ МОДЕЛІ ПРИ ПРОГНОЗУВАННІ ВИПАДКОВИХ НЕСТАЦІОНАРНИХ ЕКОНОМІЧНИХ ПРОЦЕСІВ

Кубів С. І.

ВЫБОР ПОРЯДКА РЕГРЕССИОННОЙ МОДЕЛИ ПРИ ПРОГНОЗИРОВАНИИ СЛУЧАЙНЫХ НЕСТАЦИОНАРНЫХ ЭКОНОМИЧЕСКИХ ПРОЦЕССОВ

Кубив С. И.

CHOICE OF THE ORDER OF THE REGRESSION MODEL FOR FORECASTING OF RANDOM NON-STATIONARY ECONOMIC PROCESSES

Kubiv S.

Об'єктом дослідження є гетероскедастичні процеси, які впливають на виробництво товарів військового призначення країн-експортерів. На сьогоднішній день збройні конфлікти є найбільш значущим фактором, який впливає на обсяги виробництва та експорту озброєння, оскільки передбачає наявність у сторін необхідної кількості озброєння та є в певному сенсі стохастичним процесом. Робота присвячена прогнозуванню стохастичних впливів на виробничі процеси товарів військового призначення країн-експортерів. В якості прикладу розглянуто економічну систему зі стохастичними впливами та проблемами вузьких місць у виробничих підрозділах. Модель процесу виходу продукції представлено у виді випадкового процесу з повільною нестаціонарністю (гетероскедастичного процесу). В ході дослідження використовувалися методи прогнозування нестаціонарних випадкових процесів. Досліджено задачу вибору та обґрунтування математичної моделі прогнозу гетероскедастичного процесу, що розглядається. Доведено, що найбільш спроможним методом короткострокового прогнозу є метод наближення Паде. Показано, що метод Паде, по суті, є методом апроксимації аналітичними (дрібно-раціональними) функціями, тому його можна інтерпретувати як метод побудови моделі авторегресії та ковзного середнього (АРКС). Розглянуті модифікації моделі АРКС, такі як модель авторегресії та інтегрованого ковзного середнього або авторегресії та фрактального інтегрованого ковзного середнього. Розроблено модифікований метод вибору порядку авторегресійної моделі за інформаційним критерієм Акаіке та за байєсівським інформаційним критерієм. Проаналізовано модельну задачу та приклади експериментальних залежностей. Запропоновано ефективну методику вибору порядку регресійних моделей, що застосовуються *при практичному прогнозуванні стохастичних процесів, яка заснована на канонічних розкладах випадкової функції. Для розбиття функції розподілу на нееквідистантні інтервали з постійними інтенсивностями потоку використовується економічний рекурентний алгоритм. Результати розрахунків можуть бути використані для оптимального вибору порядку регресійної моделі, якою апроксимується реальний процес виробництва у вигляді часового ряду з випадковими зовнішніми впливами.*

Ключові слова: *гетероскедастичність, дискретні часові ряди, модель авторегресії, стохастична система, апроксимація Паде, регресійна модель, порядок моделі, виробнича система.*

Объектом исследования являются гетероскедастические процессы, которые влияют на производство товаров военного назначения стран-экспортеров. На сегодняшний день вооруженные конфликты являются наиболее значимым фактором, влияющим на объемы производства и экспорта вооружения, поскольку предполагает наличие у сторон необходимого количества вооружения и есть в некотором смысле стохастическим процессом. Работа посвящена прогнозированию стохастических воздействий на производственные процессы товаров военного назначения стран-экспортеров. В качестве примера рассмотрено экономическую систему со стохастическими воздействиями и проблемами узких мест в производственных подразделениях. Модель процесса выхода продукции представлено в виде случайного процессу с медленной нестационарностью (гетероскедастического процесса). В ходе исследования использовались методы прогнозирования нестационарных случайных процессов. Исследована задача выбора и обоснования математической модели прогноза рассматриваемого гетероскедастического процесса. Доказано, что наиболее способным методом краткосрочного прогноза является метод приближения Паде. Показано, что метод Паде, по сути, является методом аппроксимации аналитическими (мелко-рациональными) функциями, поэтому его можно интерпретировать как метод построения модели авторегрессии и скользящего среднего (АРСС). Рассмотрены модификации модели АРСС, такие как модель авторегрессии и интегрированного скользящего среднего или авторегрессии и фрактального интегрированного скользящего среднего. Разработан модифицированный метод выбора порядка авторегрессионной модели по информационному критерию Акаике и по байесовскому информационному критерию. Проанализированы модельная задача и примеры экспериментальных зависимостей. Предложено эффективную методику выбора порядка регрессионных моделей, применяемых при практическом прогнозировании стохастических процессов, которая основана на канонических раскладах случайной функции. Для разбиения функции распределения на неэквидистантные интервалы с постоянными интенсивностями потока используется экономический рекуррентный алгоритм. Результаты расчетов могут быть использованы для оптимального выбора порядка регрессионной модели, которой аппроксимируется реальный процесс производства в виде временного ряда со случайными внешними воздействиями.

Ключевые слова: *гетероскедастичность, дискретные временные ряды, модель авторегрессии, стохастическая система, аппроксимация Паде, регрессионная модель, порядок модели.*

1. Вступ

Одним з актуальних завдань в загальній проблемі планування офсетної політики є прогнозування впливів від реалізації офсетних договорів на субпідрядників країн-експортерів. Прогнозування здійснюється переважно на основі використання математичних моделей та методів й вимірювання ризику.

Представлена робота присвячена вельми актуальній та специфічній області фінансово-економічної діяльності – прогнозуванню стохастичних впливів на виробничі процеси товарів військового призначення країн-експортерів.

Проблеми аналізу часових рядів в економіці та виробництві в останні роки привертають значну увагу. Будуючи економетричні моделі, ризик-менеджери найчастіше використовують стандартний показник – дохідність активів. У той же час, сфера економетрики зазнає різних нових можливостей, особливо у галузі короткострокового прогнозування, стохастичної варіабельності, та наявності потужних спеціалізованих програмних застосунків.

Аналіз часових рядів стосується теорії та практики оцінки виробничих можливостей у часі. У певному сенсі – це емпірична дисципліна, але як і в інших наукових галузях, основою для отримання висновків та прийняття рішень є теорія. Однак є ключова особливість, яка відрізняє аналіз часових рядів в економіці від інших різновидів аналізу часових рядів. І теоретичні передпосилання, емпіричні часові ряди містять помітний елемент невизначеності.

При дослідженні часових рядів характеристик виробничої системи, як правило, отримують різні конкуруючі моделі, особливо в виробничих умовах з стохастичністю випуску продукції, пов'язаних з проблемами виникнення вузьких місць. Отже, вибір найкращої моделі, яка описує виробничу систему, стає складним і критичним, оскільки деякі моделі, які найбільш точно відповідають спостережуваним даними, можуть неправильно прогнозувати майбутні значення через складність та неоднозначність моделі. У цій роботі автор прагне продемонструвати процедуру вибору моделі у виробничій системі зі стохастичностями з використанням налаштування коефіцієнту детермінації, байєсівського інформаційного критерію та інформаційного критерію Акаіке. Отримані результати оцінювання об’єму виробництва служать у якості початкових даних для розрахунку функцій автокореляції та множинної кореляції та вибору порядку відповідних моделей авторегресії (АР-моделей), АРКС та АР-моделей з інтегрованим ковзним середнім (АРІКС). Параметри моделі оцінюються, використовуються для прогнозів та зіставляються з вихідними і перетвореними даними для отримання суми квадратів помилок (Sum of Square Errors – SSE). Оцінка адекватності моделі зазвичай проводиться за байєсівським інформаційним критерієм (BIC) та за інформаційним критерієм Акаіке (AIC). Серед конкуруючих моделей модель АРІКС найкраще пояснює дисперсію наборів даних і має найнижчі значення БІК й ІКА. Тому найчастіше обирається в якості моделі, яка представляє досліджувану виробничу систему [1].

Також встановлено, що скоригований коефіцієнт детермінації в поєднанні з критеріями БІК й ІКА є адекватним інструментом для вибору моделі при дослідженні часових рядів, особливо за наявності стохастичних впливів [2].

У практичному аспекті дуже привабливою рисою моделей типу АРКС є їх прийнятна точність при прогнозуванні та відсутності необмеженої розбіжності екстраполюючої функції (екстраполянти). Ця особливість обумовлена наявністю асимптотичних властивостей екстраполянти як такої, що є дрібно-раціональною за визначенням, у той час як ніякий поліном не має ні горизонтальної, ані будь-якої іншої асимптоти [3].

Важливість розрахунку кількісних мір точності прогнозування добре висвітлена в літературі [4, 5]. Але конкретні рекомендації по кількісному розрізненню якісних прогнозів від невдалих, як правило, відсутні. Для цього зазвичай використовують стандартні класичні міри похибок прогнозування, такі як:

– середнє абсолютне відхилення (mean absolute deviation – MAD);

– середньоквадратична похибка (mean Square Error – MSE) або середня (у відсотках) абсолютна похибка (mean Absolute Percentage Error – MAPE).

Для таких мір менші значення вказують на кращі моделі прогнозування. Однак ці міри не в усіх випадках можуть забезпечити належним чином точність моделей прогнозу при практичному застосуванні. Це призводить до того, що користувачі не можуть зрозуміти наслідки прогнозів для їх діяльності [5]. У цій роботі пропонується простий та практичний показник точності прогнозування моделі – процентна похибка прогнозу (Percent Forecast Error – PFE).

Дослідники в тематичній області протягом десятиліть пропонували різні методи вибору моделі, наприклад, у [6] стверджується, що продуктивність моделі є функцією її передбачуваної здатності, і її вибір є виключно необхідним, оскільки він керує вибором якості обраної моделі. Поліпшення продуктивності моделі, отримане за допомогою вибору моделі, забезпечує надійний прогноз майбутнього системи [7]. У [8] відмічається, що окрім тестування адекватності моделі, мета вибору моделі включає пошук хорошого алгоритму прогнозу, який описує систему, а критерій AIC є основним методом вибору моделі [9]. У [9] запропоновано регуляризований інформаційний критерій (RIC) за мірою Kullback-Leibler, який є розширенням AIC та BIC, а потім використовується для вибору моделі. У [10] використані різні методи для вибору моделі, включаючи тестування гіпотез, діагностичні тести, методи відповідності, байєсівські підходи та методи оцінки прогнозу. У [11] використаний алгоритм вибору моделі SURE-Autometrtcs; стверджувалося, що метод працює добре. У [12] запропоновано тести гіпотези та критерії відбору з використанням остаточної помилки прогнозування (FPE) для вибору моделі.

Таким чином, вище представлені різні методи процедур вибору моделі, але їх можливо застосовувати в досить вузьких, конкретних ситуаціях, відмінних від галузі, яка тут досліджується. Це також свідчить, що не існує єдиного методу вибору моделі, а деякі процедури, рекомендовані в літературі, є досить складними, трудомісткими, і доволі абстрактними, що звужує сферу їх практичного застосування. *Отже, об'єктом дослідження є* гетероскедастичні процеси, які впливають на виробництво товарів військового призначення країн-експортерів. *Метою даної роботи* є розробка ефективної методики вибору порядку регресійних моделей, що застосовуються при практичному прогнозуванні процесів.

2. Методика проведення досліджень

Ключові змінні, необхідні для вибору моделі з використанням узгодженого коефіцієнту детермінації R^2 – це число параметрів моделі та сума квадратів похибок. Результати вимірювання об'єму випуску продукції, отримані від виробничої організації, служать вхідними даними для функції автокореляції та обчислення функції часткової автокореляції. Значення параметрів використовуються для прогнозування та порівняння їх з вихідними та перетвореними даними для отримання суми квадратів похибок.

Розглянемо найпростіше рівняння лінійної регресії:

$$y = ax + b,$$

де y – залежна змінна; x – незалежна змінна; a,b – коефіцієнти оцінки за методом найменших квадратів.

Адекватність моделі оцінюється за допомогою скоригованого коефіцієнту детермінації R^2, який дає уявлення про те, скільки точок даних потрапляє в лінію регресії для вивчення взаємозв'язку в наборі даних. Іншими словами, це певна доля дисперсії σ_y^2 залежної змінної y. Ця доля з'являється у сумарній дисперсії σ_{res} внаслідок впливу незалежної змінної x:

$$\sigma_{res}^2 = \sigma_{x|y}^2 + \sigma_y^2,$$

де $\sigma_{x|y}^2$ – умовна (по факторам впливу змінної x) дисперсія залежної змінної, тобто дисперсія випадкової похибки моделі; σ_y^2 – власне дисперсія залежної змінної y.

Скоригований коефіцієнт детермінації R^2 показує долю варіації змінної, яка пояснюється впливом незалежної змінної на залежну змінну.

Рівняння для скоригованого коефіцієнту детермінації дається виразом:

$$\overline{R}^2 = 1 - \frac{\sum_{k=1}^{n} \frac{\left(y_k - \overline{y}_k\right)^2}{(n-k+1)}}{\sum_{k=1}^{n} \frac{\left(y_k - \overline{y}_k\right)^2}{(n-k+1)}}, \quad \overline{y}_k = \frac{1}{n}\sum_{k=1}^{n} y_k,$$

де k – число параметрів; n – число незалежних змінних;

$SSE = \sum_{k=1}^{n} \left(y_k - \overline{y}_k \right)^2$ – сума квадратів залишків регресії;

$y_k, \overline{y}_k$ – фактичні та розрахункові значення пояснювальної змінної.

Відповідні рівняння для критеріїв Акаіке та для байєсівського критерію є:

$$AIC = m\,ln\left(SSE \right) + 2k;$$
$$BIC = m\,ln\left(SSE \right) + k\,ln\,m,$$

де m – число спостережень у серії.

Покажемо, що принаймні один фактор (випадковість) впливає на точність і може пояснити відмінності у відносній ефективності різних методів. Більш того, передбачається, що точність методу прогнозування залежить від декількох факторів, і що ці фактори можна виділити і кількісно визначити, а також виміряти їх вплив. Можливо, коли в ряді даних присутня висока випадковість, більш досконалі методи, такі як АРКС, можуть переповнювати модель цими даними. Це надбавка може виникнути, коли середня квадратична помилка мінімізується, тобто коли досягається стаціонарність шляхом або диференціації за сезон, або коли обрана модель АРКС (*p, q*). Наприклад, відсутність випадкових подій у залишках не завжди означає кращі результати прогнозування.

Нарешті, слід зазначити різницю між узгодженням моделі та прогнозуванням, а також типом функції втрат. Наприклад, коли має місце необмежена стандартизована функція квадратичних втрат, методи сезонного аналізу дають ті ж самі результати, як і методи з використанням даних, що коригуються за сезон. Однак для прогнозування це не так. Найкращі методи варіюються залежно від прийнятої функції втрат та кількості випадкових величин, присутніх у серії.

Особа, яка приймає рішення, використовуючи ці серії та застосовуючи єдиний метод прогнозування, отримала б дуже різні результати залежно від того, яка функція втрат мінімізувалася б та чи хотіла б вона мінімізувати похибки в підгонці моделі або на фазі прогнозування. Однак у цілому це можна зробити і за допомогою більш простих методів. Тут важливу роль відіграє згладжування експериментальних даних шляхом введення деякої фінітної вагової функції. Розглянемо задачу оптимізації такої функції згладжування.

3. Результати дослідження та обговорення

Один з найбільш універсальних методів побудованих аналітичних моделей випадкової функції – канонічні розкладення В. С. Пугачова [13]. Розглянемо методику канонічного розкладання для вибірки фіксованого об'єму.

В реальній ситуації, в системі збору та обробки, апріорні дані про статистичні характеристики процесу, як правило, або відсутні повністю, або є тільки частково, самого загального характеру. Тому при побудові алгоритму розкладання для вибірки змінного (наростаючого) об'єму необхідно одночасно оцінювати необхідні характеристики, враховуючи знову одержувані дані. До таких характеристик відносяться математичне сподівання, дисперсія, кореляційна функція процесу, щільності розподілу коефіцієнтів розкладання V_ν .

З урахуванням властивості безперервності випадкових величин V_ν , яка випливає з безперервності обвідної випадкової функції $X(t)$, можна застосувати непараметричну оцінку Парзена [14] виду:

$$F_N(V) = \frac{1}{N d_f} \sum_{k=1}^{N} g(u_k),$$

де d_f – деяка константа, яка називається коефіцієнтом розмитості;

$g(u_k)$ – вагова згладжуюча функція або функція ядра;

$u_k = \frac{(V - V_k)}{d_f}$, V_k – k-а реалізація випадкової величини V.

Маючи набір реалізацій V_k, $k = \overline{1,N}$ і задавши будь-яким чином d_f і $g(u)$, можна однозначно визначити щільність ймовірності випадкової величини V .

Практична методика вибору згладжуючого ядра і коефіцієнтів розмитості запропонована в [15]. Показано, що за умови симетричності ядра $g(u) = g(-u)$ його структура виду:

$$g(u) = \begin{cases} a - bu^2, & |u| \le c, \\ 0, & |u| > c, \end{cases}$$

є оптимальною за критерієм мінімуму інтегральної середньоквадратичної помилки (СКП) апроксимації. Тут a, b, c – деякі константи, які обираються, виходячи з особливостей розв'язуваної задачі. Відомо, що вибір функції згладжує ядра, що спадає до країв, практично завжди дає результат, найбільш близький до оптимального. У той же час для даної задачі відхилення форми ядра від наведеної вище не дуже критично. Наприклад, при використанні найпростішої – прямокутної функції ядра СКП апроксимації зростає всього на 6 %. При цьому значно знижується трудомісткість обчислень, які необхідно виконувати в реальному масштабі часу.

Наведемо деякі міркування щодо вибору коефіцієнта d_f – параметра, яким визначається інтервал ненульових значень ядра. Якщо вибрати його занадто великим, оцінка буде занадто згладженою, нечутливою до швидких відхилень випадкової величини. Якщо ж значення d_f занадто мале, оцінка буде згладжена недостатньо, буде «зашумленою». Оптимальним між цими крайніми позиціями для прямокутної функції ядра буде вибір:

$$d_f = 0{,}5 \sup_k |V_k - V_{k-1}|, \; V_k \ge V_{k-1}, \; k = \overline{2,N},$$

як половина найбільшої відстані між двома сусідніми членами випадкової послідовності. При цьому коефіцієнт d_f , очевидно, залежить від параметрів вибірки, ніж гарантується відсутність розривів області визначення оцінки, тобто відсутність зашумлення, і мінімальне «загладжування» оцінки не в середньому по всій множині вибірок, а для кожної конкретної вибірки.

Таким чином, при використанні прямокутної функції згладжування і наведеного вище правила обчислення коефіцієнтів вираз для оцінки має вигляд:

$$F_N(V) = \frac{1}{N}\sum_{k=1}^{N} g_u(V),$$

де $g_u(V) = \begin{cases} 1/2d_f\,, & V_k - d_f \le V \le V_k + d_f\,, \\ 0, & \left| V_k - V \right| > d_f\,, \end{cases} \quad k = \overline{1,N}\,.$

При практичній реалізації розглянутої методики на комп’ютері з обмеженою швидкодією і об'ємом пам'яті облік тривалості післядії реальних процесів є досить важливим завданням. Ефективною оцінкою тривалості післядії є нормований коефіцієнт кореляції процесу, виражений через координатні функції канонічного розкладання:

$$r_{\nu x}(\nu,k) = \frac{\sigma_\nu \varphi_\nu(k)}{\sigma_x(k)},\ \nu = \overline{1,N},\ k = \overline{1,N}.$$

В [13] показано, що функцію деякого спеціального виду від випадкової величини $r_{\nu x}(\nu,k)$ можна вважати розподіленою по гаусівському закону з відповідними математичним очікуванням і дисперсією. Там же визначено область прийняття гіпотези про рівність коефіцієнта кореляції нулю. Таким чином, використовуючи перевірку даної гіпотези проти простої альтернативи про нерівність коефіцієнта кореляції нулю паралельно із загальним алгоритмом обробки даних, можна на кожному етапі обчислення параметрів розкладання визначати необхідний об'єм інформації, що зберігається.

4. Висновки

У представленій роботі запропонована методика вибору моделі у виробничій системі зі стохастичностями. Методика заснована на канонічних розкладах випадкової функції. Для розбиття функції розподілу на нееквідистантні інтервали з постійними інтенсивностями потоку використовується економічний рекурентний алгоритм, що легко реалізується на комп'ютері.

Результати розрахунків можуть бути використані для оптимального вибору порядку регресійної моделі, якою апроксимується реальний процес виробництва – часовий ряд з випадковими зовнішніми впливами.

References

1. McQuarrie, A., Tsai, C.-L. (1998). *Regression and Time Series Model Selection*. World Scientific Publishing Co. Pte. Ltd, 455.

2. Tsay, R. S. (2010). *Analysis of Financial Time Series*. Hoboken: John Wiley & Sons, Inc., 677. doi: http://doi.org/10.1002/9780470644560

3. Kendall, M. (1976). *Time-series*. London: Charles Griffin, 197.

4. Lewis, C. D. (1982). *Industrial and business forecasting methods: a practical guide to exponential smoothing and curve fitting*. London, Boston: Butterworth Scientific, 143.

5. Klimberg, R., Ratick, S. (2018). Development of a Practical and Effective Forecasting Performance Measure. *Advances in Business and Management Forecasting. Vol. 12*. Emerald Publishing Ltd, 103–118. doi: http://doi.org/10.1108/s1477-407020170000012007

6. Hastie, T., Tibshirani, R., Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference and Prediction*. Springer Series in Statistics. New York: Springer-Verlag, 764.

7. Ferreira, E. (2015). Model Selection in Time Series Machine Learning Applications. *Academic dissertation of Technology and Natural Sciences of the University of Oulu*. Linnanmaa, 87.

8. Rao, C. R. (Ed.) (2013). *Handbook of Statistics. Vol. 30.* Time Series Analysis: Methods and Applications. Kidlington, Oxford: Elsevier, The Boulevard, 777.

9. Hurvich, C. M., Tsai, C.-L. (1989). Regression and time series model selection in small samples. *Biometrika, 76 (2),* 297–307. doi: http://doi.org/10.1093/biomet/76.2.297

10. Hannan, E. J., Krishanaiah, P. R., Rao, M. M. (1985). Various Model Selection Techniques in Time Series Analysis. *Handbook of Statistics. Vol. 5.* Elsevier B.V., 179–187. doi: http://doi.org/10.1016/s0169-7161(85)05008-8

11. Norhayati, Y. (2016). *SURE-Autometrtcs Algorithm for Model Selection in Multiple Equations*. Utara, 94.

12. Liebscher, E. (2012). A Universal Selection Method in Linear Regression Models. *Open Journal of Statistics, 2 (2),* 153–162. doi: http://doi.org/10.4236/ojs.2012.22017

13. Pugachev, V. S. (1962). *Teoriia sluchainykh funkcii.* Moscow: Fizmatgiz, 784.

14. Parzen, E. (1962). On Estimation of a Probability Density Function and Mode. *The Annals of Mathematical Statistics, 33 (3),* 1065–1076. doi: http://doi.org/10.1214/aoms/1177704472

15. Epanechnikov, V. A. (1969). Neparametricheskaia ocenka mnogomernoi plotnosti veroiatnosti. *Teoriia veroiatnostei i ee primeneniia, 1,* 156–161.

The object of research is heteroskedastic processes that affect the production of military goods of exporting countries. Today, armed conflicts are the most significant factor affecting the volume of production and export of weapons, since it assumes that the parties have the necessary quantity of weapons and is, in a sense, a stochastic process. The work is devoted to forecasting stochastic effects on the production processes of military goods of exporting countries. As an example, an *economic system with stochastic effects and bottleneck problems in production units is considered. The model of the output process is presented as a random process with slow non-stationarity (heteroscedastic process). The methods for predicting non-stationary random processes are used. The problem of choosing and substantiating a mathematical model for predicting a heteroskedastic process is investigated, and considered. It is proved that the most capable short-term forecasting method is the Padé approximation method. It is shown that the Padé method, in fact, is a method of approximation by analytical (finely rational) functions, therefore it can be interpreted as a method of constructing a model of autoregression and moving average (ARIMA). Modifications of the ARIMA model, such as a model of autoregression and integrated moving average or autoregression and fractal integrated moving average, are considered. A modified method is developed for choosing the order of the autoregressive model according to the Akaike information criterion and beyond the Bayesian information criterion. The model problems and examples of experimental dependencies are analyzed. An effective technique is proposed for choosing the order of regression models used in the practical forecasting of stochastic processes, based on the canonical layouts of a random function. To partition the distribution function into non-equidistant intervals with constant flow intensities, an economic recurrence algorithm is used. The calculation results can be used to optimally select the order of the regression model, which approximates the real production process in the form of a time series with random external influences.*

Keywords: *heteroskedasticity, discrete time series, autoregressive model, stochastic system, Padé approximation, regression model, model order, production system.*