

УДК 004.94: 629.039.58

І.С. СКІТЕР*, М.В. САВЕЛЬЄВ**,**

МОДЕЛЮВАННЯ АНОМАЛІЙ ЩІЛЬНОСТІ ПОТОКУ НЕЙТРОНІВ В АВТОМАТИЗОВАНІЙ СИСТЕМІ КОНТРОЛЮ ЯДЕРНОЇ БЕЗПЕКИ ОБ'ЄКТА «УКРИТТЯ»

*Інститут проблем безпеки АЕС НАН України, м. Чорнобиль, Україна

**Інститут проблем математичних машин і систем НАН України, м. Київ, Україна

Анотація. Аналіз динаміки щільності потоку нейтронів (ЩПН) від паливовмісних мас (ПВМ) в об'єкті «Укриття» (ОУ) показує наявність значень, які перевищують середні значення за різні періоди спостережень. Ідентифікація таких значень за критерієм «аномалія/не аномалія» дозволить виключити з масиву спостережень неінформативні події. Або, у разі підтвердження аномалії, сформувані ефективні дії для прийняття рішень щодо усунення причин таких подій. Для вирішення проблеми виявлення аномальних вимірювань у даний час використовується теорія статистичних рішень. Вона заснована на використанні параметричних методів. Їх використання вимагає апріорної інформації про характер розподілу вимірюваного явища і його параметри. Для ефективного вирішення проблеми виявлення й усунення аномальних вимірювань необхідно знати статистичні характеристики нормальних та аномальних компонентів. У цій роботі пропонуються статистичні критерії для оцінки аномалії у часових рядах ЩПН, які мають різні підходи до формування інтервалів спостережень, потужності та надійності виявлення аномалій. У залежності від виду закону розподілу масиву спостережень запропонований набір критеріїв, які найбільш доцільно використовувати при перевірці аномальності рівнів ряду для експоненціального закону розподілу, розподілу Пуассона та Вейбула. Проведена оцінка потужностей критеріїв у залежності від об'єму вибірок. Встановлена точність визначення аномалій за критеріями при відомих значеннях середнього та дисперсії у досліджуваній вибірці. Рекомендоване використання критерію Граббса для дослідження на аномальність рівнів вибірок з $n > 700$, для вибірок з $n < 50$ доцільним є використання критеріїв Діксона та Смоляка-Титаренка. Використання оптимальних критеріїв у залежності від характеристик досліджуваних вибірок дасть змогу підвищити математичну значущість отримуваних результатів і, як наслідок, підвищити якість управлінських рішень та ядерну безпеку на ОУ в цілому.

Ключові слова: об'єкт «Укриття», щільність потоку нейтронів, аналіз динаміки, параметричні методи, статистичні критерії, виявлення аномалій.

Abstract. Analysis of the dynamics of neutron flux density (NFD) from fuel-containing masses (FCM) in the «Shelter» Object shows the presence of values that exceed the average values for different observation periods. Identification of such values by the criterion of «anomaly / non-anomaly» will allow excluding uninformative events from the array of observations. Or, in the case of the anomaly confirmation, it will allow forming effective actions for decision-making in order to eliminate the consequences of such events. To solve the problem of detecting anomalous measurements, now there is utilized the theory of statistical solutions which is based on the use of parametric methods. The utilization of these methods requires a priori information about the nature of the distribution of the measured process and its parameters. In order to find an effective solution to the problem of detecting and eliminating anomalous measurements, it is necessary to know the statistical characteristics of normal and anomalous components. This paper proposes statistical criteria for estimating anomalies in time series of NFD which have different approaches to the formation of observation intervals, power and reliability of anomaly detection. Depending on the type of distribution of the array of observations, a set of criteria is proposed. These criteria are most expedient to use when checking the anomaly of series levels for the exponential distribution, the Poisson and Weibull distribution. The capacity of the criteria has been evaluated depending on the sample size. The article defines the accuracy of determination of anomalies by criteria with the known values of the mean and dispersion in the studied sample. As the result, it is recommended

to use the Grubbs test to study the anomaly of the sample levels with $n > 700$, and the Dixon and Smolyak-Titareno criteria for the samples with $n < 50$. The utilization of optimal criteria depending on the characteristics of the studied samples will increase the mathematical significance of the obtained results and, as a result, will improve the quality of management decisions and nuclear safety on the "Shelter" Object as a whole.

Keywords: «Shelter» Object, neutron flux density, dynamics analysis, parametric methods, statistical criteria, detection of anomalies.

DOI: 10.34121/1028-9763-2021-4-70-77

1. Вступ

Інтегрована автоматизована система управління об'єктом «Укриття» (ОУ) контролює щільність потоку нейтронів (ЩПН) від ПВМ, які знаходяться в середині ОУ. Вимірювання динаміки потоку нейтронів проводиться системою контролю ядерної безпеки (СКЯБ) за допомогою мережі датчиків, розташованих у внутрішніх приміщеннях об'єкта «Укриття» та шахті зруйнованого реактора 4-го енергоблока ЧАЕС. Необхідність контролю поведінки потоку нейтронів пов'язана з тим, що після побудови та введення в експлуатацію нового безпечного конфайнменту (НБК) з'явилась нова єдина система, в якій через зміну волого-температурного режиму змінилися процеси, що раніше відбувалися в об'єкті «Укриття» [1]. Такі зміни потенційно можуть змінити динаміку потоків нейтронів, наприклад, призвести до появи аномальних значень ЩПН, різко відмінних від середніх показників за різні періоди спостережень.

Актуальною є задача виявлення та ідентифікація таких значень за критерієм «аномальне значення / неаномальне значення». Аномальність значення ЩПН при цьому означає перевищення допустимого рівня, визначеного на основі середнього значення та квадратичного відхилення. Аналіз масиву за вказаним критерієм дозволить класифікувати аномальні події та визначати причини, які їх спричинили. Крім того, це дозволить виключити з масиву спостережень неінформативні події при побудові прогнозних моделей станів ПВМ. Або, у разі підтвердження аномалії, сформулювати ефективні дії для прийняття рішень щодо усунення причин таких подій. Реалізація поставленої задачі на першому етапі може бути проведена шляхом визначення переліку статистичних критеріїв оцінки аномальності значень показників, оцінки їх статистичних властивостей, визначення меж їх використання.

Метою статті є моделювання, аналіз та визначення потужності статистичних критеріїв, які використовуються для оцінки аномалій та серій аномальних значень при експоненціальних законах розподілу динамічних рядів даних потоку нейтронів в об'єкті «Укриття»; визначення та відбір максимально ефективних критеріїв оцінки аномальності для визначених об'єктів спостережень.

2. Особливості моделювання ЩПН для виявлення аномальних подій в ОУ

Результати моніторингу нейтронної активності, який проводиться при контролі можливих ядерно небезпечних скупчень, представляють собою динамічний ряд і є послідовністю y_1, y_2, \dots, y_n результатів вимірювань ЩПН, отриманих у рівновіддалені проміжки часу t_1, t_2, \dots, t_n . Постає задача контролю динамічного ряду спостережень на предмет виявлення різко відмінних за статистичними параметрами показників та їх якісна класифікація.

Математична модель результатів вимірювань може бути представленою в загальному вигляді як

$$Y(t_k) = [Tr(t) + S(t_k)] + e(t_k) + A(t_k), \quad (1)$$

де $[Tr(t) + S(t_k)]$ – корисна складова, $Tr(t)$ – загальна тенденція (тренд), $S(t_k)$ – сезонна компонента, $e(t_k)$ – стохастична складова, $A(t_k)$ – аномальна складова.

Як правило, для стаціонарних процесів значення адитивної стохастичної складової $e(t_k)$ є некорельованими, мають нульове математичне сподівання і представляють собою реалізацію ергодичного випадкового процесу.

Аномальними вимірами $A(t_k)$ будемо вважати значення послідовності результатів вимірювань, які різко відрізняються за амплітудою і статистичними властивостями на фоні основної групи значень реалізації.

Для вирішення завдання виявлення аномальних вимірювань на сьогодні використовуються різні підходи щодо аналізу динамічних рядів, представлених у роботах [2–4], тощо. Слід зазначити, широке застосування має також теорія статистичних рішень із використанням параметричних методів на основі апріорних відомостей як про характер розподілу вимірюваного процесу, так і про його параметри (математичне сподівання, дисперсію, кореляційну функцію та ін.) [5].

Більшість статистичних критеріїв визначення аномальних даних застосовується для умови нормального розподілу вихідних даних у вибірці. Для експоненціального закону розподілу ЩПН, розподілів типу розподілу Пуассона чи Вейбула доцільне використання таких критеріїв, як критерій Смоляка-Титаренка [6], критерій Бродського-Бицаня-Власенка [7]. Вказані критерії направлені на оцінку аномальності окремих вимірювань ЩПН.

Для викидів, які присутні у вибірці з деякою частотою (неодиничних викидів) доцільно використовувати критерій Кімбера для декількох викидів [8], критерій серії викидів для розподілу Вейбула. Для експоненціального розподілу зі щільністю ймовірностей

$f(x) = \frac{1}{\lambda} e^{-\frac{x}{\lambda}}$ логічно враховувати максимальні значення показників у вибірці.

Розрахункове значення критерію Смоляка-Титаренка для k -го вимірювання у вибірці об'ємом n визначається за (2):

$$C_k = \frac{x_{k(n)}}{\bar{x}(n)}, \quad (2)$$

де $x_{k(n)}$ – максимальне (чи «підозріле» на аномальність) значення показника у вибірці, $\bar{x}(n)$ – середнє значення у досліджуваній вибірці.

Значення $x_{k(n)}$ вважають аномальним, якщо виконується умова $C_k > C_{kp}(\alpha)$, де $C_{kp}(\alpha)$ – критичне значення критерію для заданого рівня значущості.

Для експоненціального розподілу у формі $f(x) = \lambda e^{-\lambda x}$, де λ – параметр розподілу, статистика критерію Бродського-Бицаня-Власенка для перевірки на аномальність x_k (максимального чи «підозрілого») значення з вибірки при невідомому λ визначається як

$$z_k = \frac{x_k - x_{k-1}}{x_k - x_1}. \quad (3)$$

Для перевірки аномальності показника x_k розраховують ймовірність

$$P(z \geq z_k) = (n-1)! (1-z_k)^{n-1} \prod_{j=1}^{n-2} \frac{1}{1+j(1-z_k)}. \quad (4)$$

Якщо розрахована ймовірність більше прийнятого рівня значущості α , гіпотеза про аномальність показника відхиляється.

Перевірка вибірки, в якій присутні декілька аномальних значень, проводиться за допомогою статистики Кімбера:

$$S_j = \frac{x_{n-j+1}}{\sum_{i=1}^{n-j+1} x_i}, \quad j = 1, 2, \dots, n-1. \quad (5)$$

Перевірка на аномальність серії викидів проводиться за таким алгоритмом:

1. Якщо для $S_j(\alpha)$ – критичного значення статистики – для $i = 1, 2, \dots, k$ виконується умова $S_j < S_j(\alpha)$, то з заданим рівнем значущості α гіпотеза про те, що k максимальних чи «підозрілих» значень є аномальними, відхиляється.

2. Якщо виконується умова $S_i > S_i(\alpha)$ для $i = k, k-1, \dots, j-1$ і $S_j > S_j(\alpha)$, то з заданим рівнем значущості α гіпотеза про те, що j максимальних чи «підозрілих» значень є аномальними, приймається.

3. Якщо $S_k < S_k(\alpha)$, то з заданим рівнем значущості α гіпотеза про те, що k максимальних чи «підозрілих» значень є аномальними, приймається.

Для перевірки на аномальність показників у вибірці, яка має розподіл Пуассона чи Вейбула, використовують методи трансформації статистик [9]:

– трансформована статистика критерію Граббса [10]:

$$G = \frac{s_{n-k+1,n}^2}{s^2}, \quad (6)$$

$$\text{де } s_{n-k+1,n}^2 = \sum_{i=1}^{n-k} (x_i - \overline{x_{n-k+1,n}})^2, \quad \overline{x_{n-k+1,n}} = \frac{1}{n-k} \sum_{i=1}^{n-k} x_i, \quad s^2 = \sum_{i=1}^n (x_i - \bar{x})^2, \quad \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i.$$

При виконанні умови $G < G(\alpha)$ приймається гіпотеза про те, що у досліджуваній вибірці існує k аномальних значень;

– трансформована статистика критерію Діксона [11]:

$$R_k = \frac{x_n - x_{n-k}}{x_n - x_k}. \quad (7)$$

При виконанні умови $R_k < R_k(\alpha)$ приймається гіпотеза про те, що у досліджуваній вибірці існує k аномальних значень.

Перевагами статистичних методів аналізу аномальних значень у динамічних рядах є їх універсальність, надійність, алгоритмічність, адаптивність до різних об'єктів дослідження та структури масивів спостережень.

3. Нерозв'язані задачі

Реєстрація аномальних подій СКЯБ може свідчити як про відхилення від умов безпечної експлуатації ОУ, так і про метрологічні дефекти СКЯБ – зареєстровані хибні значення вимірювань ЩПН. Необхідність експертного аналізу кожної аномальної події СКЯБ призводить до додаткових затрат інформаційних ресурсів та часу, що знижує ефективність і своєчасність прийняття рішень. Таким чином, є актуальною задача створення автоматизованого методу аналізу ЩПН. До теперішнього часу моделювання динаміки потоку нейтронів із метою дослідження аномальності рівнів ряду та перевірка аномальності статистичними чи іншими методами не проводились.

4. Експериментальна оцінка потужності критеріїв

Реальні динамічні ряди спостережень за нейтронними потоками в ОУ представляють собою масиви даних об'ємом декілька десятків тисяч позицій, результати вимірювань фіксуються в архіві довготривалого зберігання з періодичністю одна хвилина. Для побудови імітаційної моделі та аналізу обрані масиви спостережень за період 01.08.2019 – 31.08.2021.

Тому дослідження потужності приведених статистичних критеріїв умовно можна розділити на два етапи. На першому етапі проводиться оцінка потужностей критеріїв на випадкових вибірках визначеної довжини з заданим періодом дискретизації за запропонованим нижче алгоритмом. Для аналізу були обрані такі об'єми досліджуваних вибірок: $n_1 = 730$ (періодичність – 1 день), $n_2 = 96$ (періодичність – 1 тиждень), $n_3 = 73$ (періодичність – 1 декада), $n_4 = 24$ (періодичність – 1 місяць). Вибір саме таких значень пов'язаний із реальними експериментальними даними спостережень і відповідає накопиченим значенням ЩПН за день, неділю, декаду та місяць. На другому етапі проводиться апробація отриманих результатів на реальних динамічних рядах спостережень.

Алгоритм реалізації першого етапу такий:

1. За допомогою генератора випадкових чисел проводиться генерація масивів даних об'ємами $n_1 = 730$, $n_2 = 96$, $n_3 = 73$, $n_4 = 24$ за експоненціальним законом, що відповідає закону розподілу для реальних експериментальних даних. На основі реальних спостережень за потоком нейтронів за період 01.08.2019 – 31.08.2021 при генерації масивів задані параметри середнього \bar{y}_i та дисперсії σ_i^2 , максимально наближені до реальних значень середнього та дисперсії ЩПН на досліджуваному об'єкті.

2. Проводяться ранжування за зростанням кожної згенерованої вибірки і заміна останніх $0,1 n_i$ максимальних значень досліджуваної вибірки на випадкові величини з діапазону $\bar{y}_i \pm k\sigma_i$, де σ_i – середнє квадратичне відхилення i -ї вибірки, $k = [1, 0 \div 10, 0]$.

3. За допомогою статистичних критеріїв (2)–(7) для кожної генерації масивів проводиться перевірка на аномальність рівнів ряду за критерієм перевищення значення класичної міри для закону розподілу випадкових величин – $\bar{y}_i + 3\sigma_i$.

4. Для значень $k\sigma_i$ була розрахована доля правильно визначених аномальних значень за кожним із критеріїв, яка чисельно може трактуватися як потужність відповідного критерію.

5. За результатами розрахунків проведений аналіз потужності критеріїв у залежності від величини вхідної вибірки.

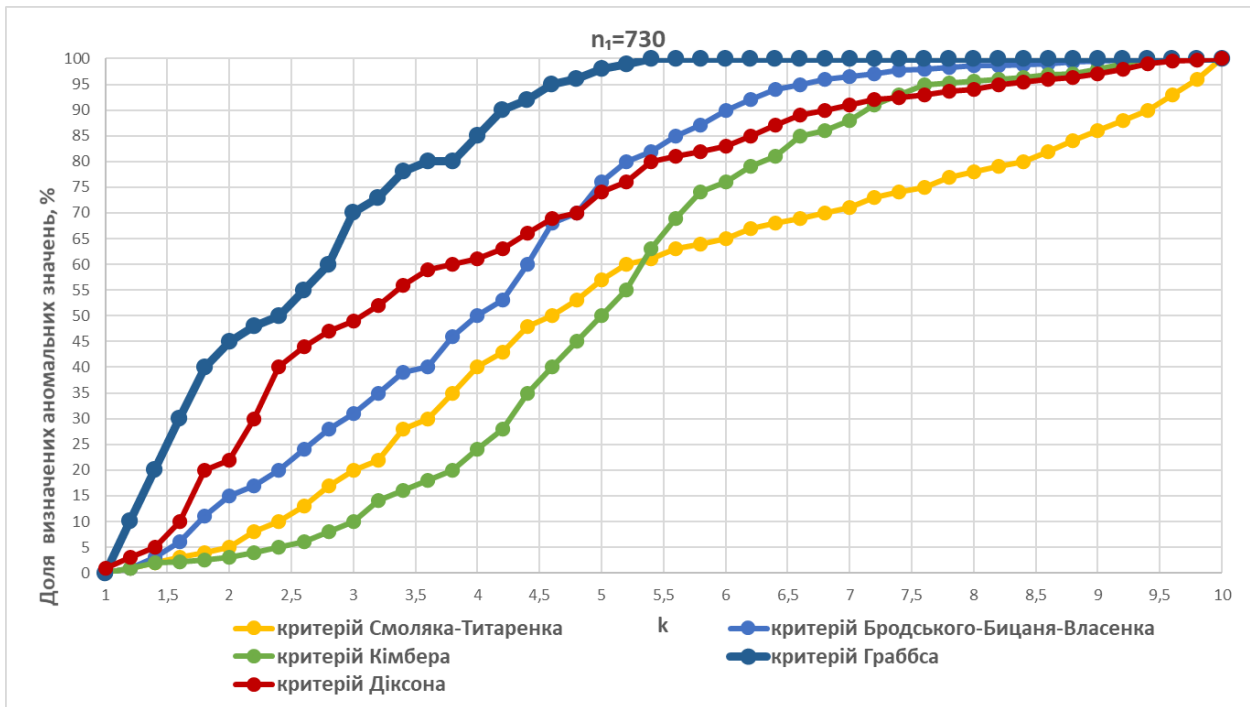
Результати імітаційного моделювання та аналізу потужностей на основі згенерованих масивів даних приведені на рис. 1.

На рисунках представлена залежність доли визначених аномальних значень від величини $\bar{y}_i + k\sigma_i$. Аналіз отриманих результатів показує, що потужність критеріїв виявлення аномальних значень у вибірці змінюється в залежності від об'єму вибірки.

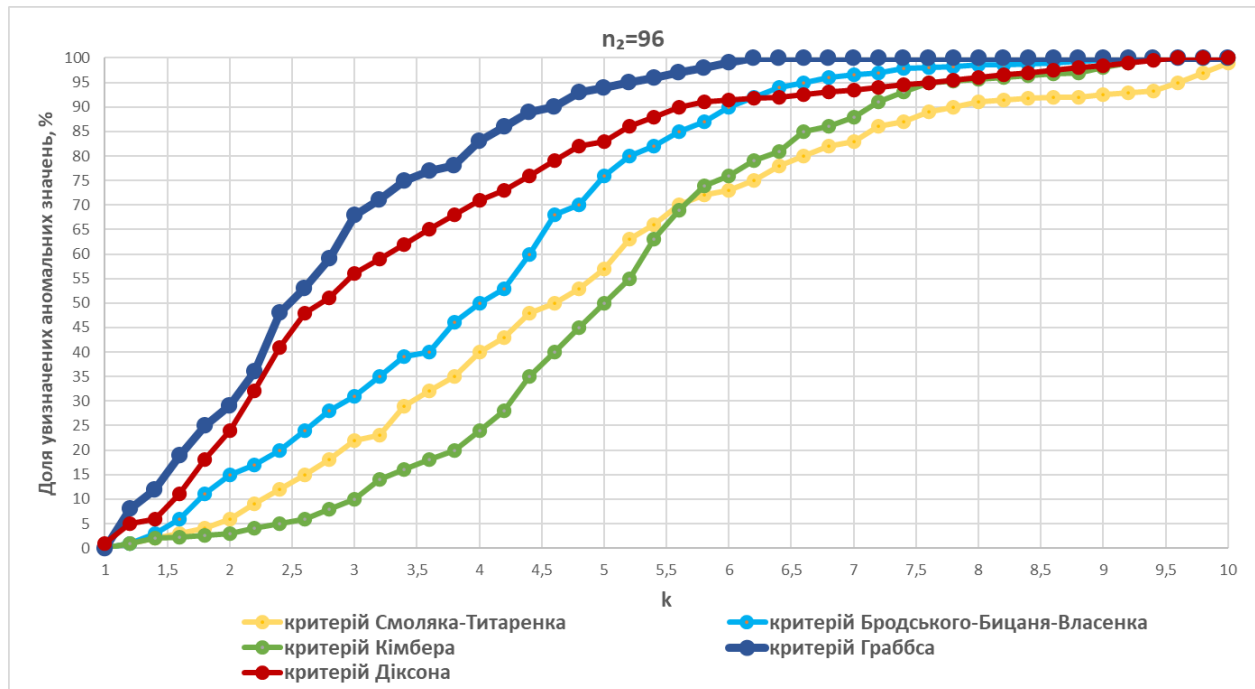
Так, для великих вибірок (наприклад, для $n_1 = 730$) максимальну потужність має критерій Граббса. При цьому критерій Бродського-Біцаня-Власенка має меншу потужність щодо визначення аномальних значень, але це компенсується малою кількістю помилок першого роду. Використання інших критеріїв для «великих» вибірок є недоцільним через їх низьку потужність.

Однак, із збільшенням рівня дискретизації динамічних рядів та зменшенням величини досліджуваних вибірок спостерігається зростання потужності критерію Діксона. Так при значеннях $\bar{y}_i + 3\sigma_i$ відсоток визначення аномальних значень за критерієм змінюється

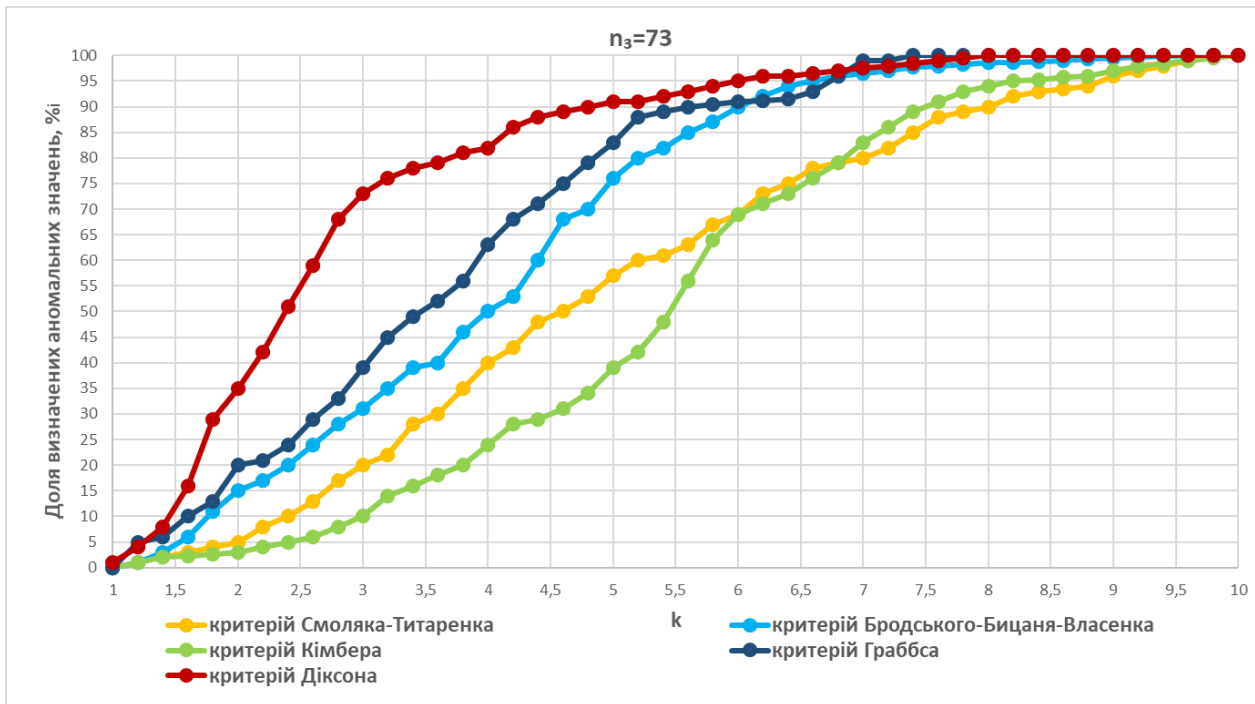
від 50% (для $n_1 = 730$) до 80% (для $n_4 = 24$), що вказує на доцільність використання саме цього критерію для визначення аномальних значень у малих вибірках.



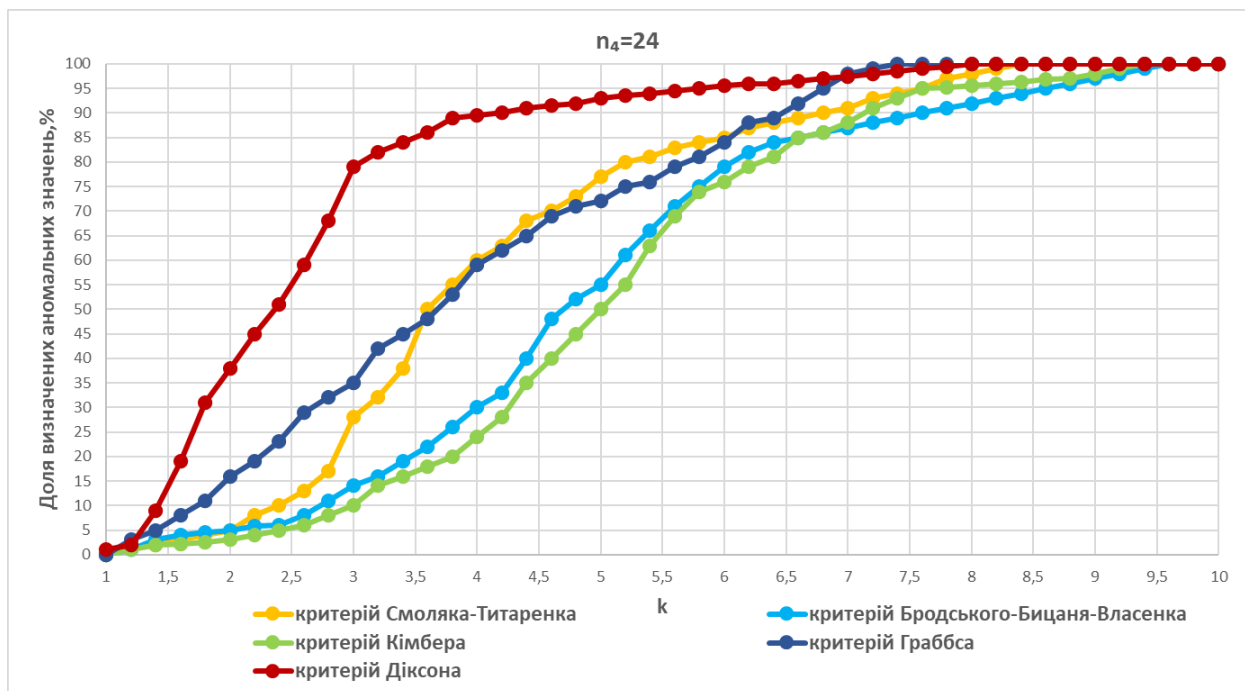
a)



б)



в)



г)

Рисунок 1 – Потужності критеріїв визначення аномалій у ряду спостережень за ЩПН для різного об'єму вибірки: а) $n_1 = 730$; б) $n_2 = 96$; в) $n_3 = 73$; г) $n_4 = 24$

5. Висновки

Таким чином, проведене дослідження дає змогу констатувати, що при розв'язку задачі ідентифікації аномальних значень у масивах даних ЩПН у системі контролю ядерної безпеки ПВМ на першому етапі необхідно визначити набір оптимальних з точки зору меж викори-

стання їх критеріїв. Мотивоване використання критеріїв у залежності від об'ємів досліджуваних вибірок різко, практично до нуля, зменшує похибки першого і другого роду. Отримані результати імітаційного моделювання дали змогу встановити, що оптимальність вибору критеріїв, яка базується на їх потужності, залежить від об'єму масиву, який підлягає оцінці.

Крім того, дискретизація динамічних рядів приводить до необхідності використання критеріїв, адекватних саме для заданого рівня дискретизації. Так для «великих» масивів даних оптимальним буде використання критерію Граббса, точність якого становить більше 70% для значень $\bar{y}_i + 3\sigma_i$ і значно зростає при незначному збільшенні показника k .

Критерій Діксона перевищує потужність критерія Граббса при невеликих вибірках ($n_3 = 73$, $n_4 = 24$), і його точність становить від 75% до 82%. Крім того, критерій є простим у розрахунках.

Також для невеликих вибірок доцільним є використання критерію Смоляка-Титаренка, який має достатню потужність і може використовуватися для розподілів типу Пуассона чи Вейбула.

Отримані результати є базою для аналізу реальних експериментальних даних ЩПН із метою класифікації подій, формування управлінських рішень. Крім того, проведені дослідження є основою для проектування та створення підсистеми аналізу аномальності в автоматизованій системі контролю ЩПН та автоматизованій системі підтримки прийняття рішень щодо забезпечення та підвищення рівня ядерної безпеки комплексу ОУ-НБК.

СПИСОК ДЖЕРЕЛ

1. Об'єкт «Укриття» в умовах нового безпечного конфайнмента: монографія / за ред. А.В. Носовського. Чорнобиль: ПБ АЕС НАН України, 2021. 344 с.
2. Lytvynov V., Stoianov N., Skiter I., Lytvyn S. Traffic Abnormalities Identification Based on the Stationary Parameters Estimation and Wavelet Function Detailization. *Mathematical Modeling and Simulation of Systems. MODS 2019. Advances in Intelligent Systems and Computing*. Springer, Cham, 2019. Vol. 1019. P. 83–95.
3. Скітер І.С. Використання EWMA-статистики для виявлення аномальної поведінки комп'ютерної мережі. *Створення та модернізація озброєння і військової техніки в сучасних умовах*: зб. тез доп. 16 наук.-техн. конф. Чернігів: ДНВЦ ЗС України, 2016. С. 18–21.
4. NCSS [Statistical software]. (n.d.). NCSS 2021. Data Analysis & Graphics: Time Series and Forecasting Methods in NCSS. Retrieved 3 February, 2021. URL: <https://www.ncss.com/software/ncss/time-series-and-forecasting-in-ncss>.
5. Долгих А.О., Байбуз О.Г. Анализ методов, моделей та программных засобів прогнозування часових рядів. *Відкриті інформаційні та комп'ютерні інтегровані технології*: зб. наук. праць Національного аерокосмічного університету ім. Жуковського «ХАІ». Харків, 2018. Т. 79. С. 74–87.
6. Смоляк С.А., Титаренко Б.П. Устойчивые методы оценивания. М.: Статистика, 1980. 208 с.
7. Бродский Я.С., Быцань Н.Н., Власенко В.М. Об исключении экстремальных значений. *Заводские лаборатории*. 1975. № 7. С. 847–849.
8. Kimber A.C., Stevens H.J. The null distribution of a test for two upper outliers in an exponential sample. *Applied Statistics*. 1987. Vol. 30. P. 153–157.
9. Fung Karen Yuen, Paul S.R. Comparisons of outlier detection procedures in Weibull or extreme-value distributions. *Communications in Statistics – Simulation and Computation*. 1985. Vol. 14, N 4. P. 895–917.
10. Grubbs F.E. Simple criteria for testing Rejection of Outlying observations. *Proc. of the American Mathematical Society*. 1950. Vol. 21. P. 27–58.
11. Dixon W.J. Rations involving extreme Values. *Proc. of the American Mathematical Society*. 1956. Vol. 22. P. 68–76.

Стаття надійшла до редакції 05.11.2021