

**НАУКОВІ ТА ПРАКТИЧНІ ПРОБЛЕМИ ВИРОБНИЦТВА
ПРИЛАДІВ ТА СИСТЕМ**

DOI: 10.20535/1970.67(1).2024.306724

УДК 519.24 (075.8)

**ЗАСТОСУВАННЯ КОНТРОЛЬНИХ КАРТ ІНДИВІДУАЛЬНИХ
СПОСТЕРЕЖЕНЬ ДЛЯ ХІ-КВАДРАТ РОЗПОДІЛЕНИХ ДАНИХ***Козир О. В., Шестак О. А.**Національний технічний університет України**«Київський політехнічний інститут імені Ігоря Сікорського», Київ, Україна**E-mail: oleg.kozyr@aer.kpi.ua*

Контрольні карти індивідуальних значень застосовуються у випадках, коли досліджувані процеси змінюються занадто повільно, у випадку коштовності отримання стандартної вибірки для побудови карт Шухарта або у випадку вивчення статистичних характеристик технологічного процесу. Контрольні карти застосовуються для контролю технологічних процесів у випадку нормального розподілу генеральної сукупності можливих значень контрольованого параметру. Інколи допускається квазі-нормальний розподіл досліджуваного параметру. На практиці, досить часто доводиться мати справу з технологічними процесами, вибірка значень контрольованого параметру якого розподілена не за нормальним законом. У таких випадках використання контрольних карт індивідуальних значень може привести до помилок I та II роду. Це обумовлено тим, що контрольні межі карт розраховані для середнього та середньоквадратичного відхилення нормального розподілу. Контрольні карти індивідуальних значень найбільш чутливі до відхилення закону розподілу вибірки від нормального, оскільки для побудови карт використовують індивідуальні значення, на які не розповсюджується центральна гранична теорема, як у випадку контрольних карт середніх значень. Тому надзвичайно важливим є розробка методів застосування розроблених контрольних карт індивідуальних значень для випадкових вибірок, генеральна сукупність яких розподілена не за нормальним законом.

У роботі проведено дослідження застосування методу статистичного контролю випадкового процесу розподіленого не за нормальним законом на основі використання контрольних карт індивідуальних значень, перетворенням вибірки значень досліджуваного параметру у вибірку розподілену за нормальним або квазі-нормальним законом, тобто виконання нормалізації з використанням методу Вох-Сох нормалізації та подальшою побудовою контрольних карт на основі отриманої вибірки. Описано математичний апарат алгоритму застосування карт для вибірок з не-нормальним розподілом. Виконано чисельне моделювання застосування контрольних карт індивідуальних значень для вибірки процесу, розподіленого за хі-квадрат розподілом, який має значну асиметричність розподілу для малих значень коефіцієнту форми, та для вибірки, яка отримана внаслідок нормалізації хі-квадрат.

Результати проведеного дослідження показали, що застосування контрольних карт індивідуальних значень для статистичного контролю для вибірок із значно асиметричним законом розподілу значень призводить до збільшення появи помилок I роду, чого майже немає у випадку застосування методу нормалізації початкової вибірки. Моделювання розладнання процесу, шляхом зміни певного індивідуального значення, також продемонструвало появу хибних сигналів розладнання процесу.

Ключові слова: контрольні карти індивідуальних значень; негауссівський розподіл; не-нормальний розподіл; хі-квадрат розподіл; Вох-Сох нормалізація.

Вступ

Контрольні карти індивідуальних спостережень (ККІС) призначені для контролю одного числового параметра процесу у тих випадках, коли неможливо отримати вибірку більше одного значення. Вони найчастіше застосовуються для контролю повільно змінних технологічних процесів, а також у випадках вивчення статистичних характеристик технологічного процесу. Крім того, стандар-

тний розмір вибірки для побудови карт Шухарта становить 20-30 підвибірок розміром 4-6 значень або $400 / (k - 1)$, де k – розмір підгрупи, отриманих випадковим чином. Якщо підвибірка складає одне значення, то загальна вибірка повинна складати 300 підвибірок [1]. Також і для таких випадків були розроблені ККІС.

Контрольні карти застосовуються для контролю технологічних процесів у випадку нормаль-

ного розподілу генеральної сукупності можливих значень контрольованого параметру. Інколи допускається квазі-нормальний розподіл досліджуваного параметру. На практиці, досить часто доводиться мати справу з технологічними процесами, вибірка значень контрольованого параметру якого розподілена не за нормальним законом. У таких випадках використання ККІС значень може привести до помилок I та II роду. Це обумовлено тим, що контрольні межі карт розраховані для середнього та середньоквадратичного відхилення нормально розподілу. ККІС найбільш чутливі до відхилення закону розподілу вибірки від нормального, оскільки для побудови карт використовують індивідуальні значення, на які не розповсюджується центральна гранична теорема, як у випадку контрольних карт середніх значень. Тому надзвичайно важливим є розробка методів застосування розроблених ККІС для випадкових вибірок, генеральна сукупність яких розподілена не за нормальним законом.

Огляд існуючих методів побудови ККІС для не нормально розподілених вибірок

Існують дослідження впливу не-нормальності закону розподілу досліджуваного процесу на побудову контрольних карт Шухарта або контрольних карт вибірових середніх значень. Публікацій, які висвітлюють використання контрольних карт індивідуальних значень для не-нормально розподілених значень вибірки випадкового процесу не так багато.

Одна з найперших робіт, у яких досліджувалось використання контрольних карт Шухарта для вибірок з ненормально розподіленими законами генеральних сукупностей, це робота [2]. Було встановлено, що для значних вибірок вплив асиметрії (крім значної) на коефіцієнти побудови контрольних меж карт Шухарта незначний, а самі коефіцієнти зберігають стабільність через дію центральної граничної теореми на карту середніх значень. Проте для значної асиметрії розподілу зростає кількість випадків виникнення помилок I роду.

У роботі [3] показано, що при значному відхиленні від нормальності закону розподілу вибірки, це призводить до значних помилок у визначенні розладнаності процесу за допомогою контрольних карт середніх значень. Щоб мінімізувати вплив цих помилок у [4] запропоновано застосовувати евристичний метод на основі зважених дисперсій для побудови контрольних меж для контрольних карт Шухарта у випадку асиметричних розподілів вибірок. Крім того, у [5] досліджувались ймовірності сигналів про розлад процесу та помилок I роду, як внаслідок впливу не-нормальності розподілу вибірок на контрольні карти середніх значень. У випадку сумішей нормальних розподілів, потрібно змінювати коефіцієнти побудови контрольних карт, через значний вплив не-

нормальності на досліджувані ймовірності. Виключення становить подвійний експоненційний розподіл.

Важливим є вплив величини підвибірки розподіленої за не-нормальним законом, та відхилення самих законів від нормального на визначення контрольних меж карт середніх значень [6]. Додатково у роботі [7] досліджено негативний вплив не-нормальності закону розподілу на контрольні межі контрольних карт та запропоновано використовувати підхід який базується на квантилях розподілу для побудови цих меж.

У роботі [8] досліджено вплив відхилення від нормального розподілу на параметри карти Шухарта для середніх значень. Проведені у роботі [8] дослідження показали наявність можливих небезпек та пасток при використанні нелінійних перетворень вибірок із не-нормальним розподілом для отримання нормально розподіленої вибірки. Щоб позбавитись вказаних недоліків, у роботі [8] запропоновано використовувати для не розладнаних процесів апроксимацію закону розподілу для визначення контрольних границь або непараметричних контрольних карт.

Іншим підходом до подолання впливу не-нормальності запропоновано у роботі [9]. Запропоновано виконувати перетворення вибірки використовуючи апроксимацію не-нормального закону розподілу за допомогою розподілу Джонсона, як метод нормалізації досліджуваного процесу та визначення ступеня відхилення від нормальності. Також, у роботі [10] запропоновано виконувати нормалізацію вибірки із не-нормальним законом розподілу за допомогою Vox-Cox перетворення для використання як звичайних карт, так і робастних.

Побудова ККІС

ККІС визначають варіацію досліджуваного процесу за індивідуальними спостереженнями, зібраними у вектор

$$x = (x_1, x_2, \dots, x_n), \quad (1)$$

де n – значення величини вибірки індивідуальних спостережень.

Побудова ККІС відбувається на основі ковзних розмахів, тобто абсолютних різниць між сусідніми значеннями

$$r = |x_i - x_{i-1}|. \quad (2)$$

Спочатку будується карта ковзних розмахів для оцінки стабільності варіації процесу. Дана карта будується так, як і карта розмахів, крім того, що розраховується не середній розмах за вибіркою, а середнє ковзних розмахів

$$\bar{r} = \frac{\sum_{i=1}^{n-1} r_i}{n-1}. \quad (3)$$

Контрольні межі розраховуються на основі середнього значення ковзних розмахів

$$UCL(\bar{R}) = \bar{R} \cdot D_4, \quad LCL(\bar{R}) = \bar{R} \cdot D_3. \quad (4)$$

Коефіцієнти D_3 та D_4 беруться із таблиці коефіцієнтів карт розмахів [11]. Дані коефіцієнти залежать від розміру підвибірки k . Оскільки ми розглядаємо індивідуальні значення, то вибираються коефіцієнти для найменшої кількості $k = 2$, вказані в (5). Відповідно, для ККІС нижня границя ковзних розмахів не розраховується [11]

$$D_4 = 3,267, \quad D_3 = 0. \quad (5)$$

Визначивши, що процес стабільний, побудувавши контрольну карту ковзних розмахів, і на ній значення не виходять за розраховану верхню контрольну межу, можна будувати ККІС. Контрольні межі ККІС розраховуються також на основі середнього значення ковзних розмахів

$$UCL(x) = \bar{x} + 3 \frac{\bar{R}}{d_2} \quad (6)$$

$$LCL(x) = \bar{x} - 3 \frac{\bar{R}}{d_2},$$

де \bar{x} - середнє значення вибірки (1), d_2 - коефіцієнт [11] для визначення оцінки середньоквадратичного відхилення за середнім значенням ковзних розмахів

$$\hat{\sigma} = \frac{\bar{R}}{d_2}. \quad (7)$$

Оскільки вибірка містить одне спостереження, то для визначення коефіцієнту d_2 вибирають найменше значення для вибірки, $k = 2$

$$d_2 = 1,128. \quad (8)$$

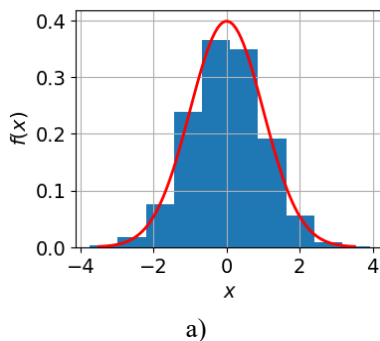
Враховуючи (7), рівняння контрольних границь (6) можна переписати відповідно до канонічного вигляду для контрольних карт Шухарта

$$UCL(x) = \bar{x} + 3\hat{\sigma} \quad (9)$$

$$LCL(x) = \bar{x} - 3\hat{\sigma}.$$

Побудова ККІС для не-нормального розподілу вибірки

Усі вище вказані рівняння розрахунку параметрів ККІС виведені із припущення про нормальність закону розподілу досліджуваного параметру процесу. У цьому випадку можна використовувати



контрольні карти для моніторингу стабільності досліджуваного процесу.

Проте, часто значення параметру процесу має не-нормальний закон розподілу значень вибірки. У багатьох випадках, закони розподілу реальних параметрів процесу мають значну асиметрію або мають розподіл генеральної сукупності логнормальний або навіть χ^2 .

У даній роботі порівнюється використання ККІС для оцінки стабільності процесу для двох законів розподілу: нормального (гауссівського) та χ^2 . Вибірки двох випадкових величин із вказаних процесів отримують для значень середнього μ , середньоквадратичного σ та величини вибірки n генеральних сукупностей. У випадку χ^2 розподілу з df ступенями свободи, для отримання значно асиметричного розподілу.

$$\mu = 0, \quad \sigma = 1, \quad n = 1000, \quad df = 2.5. \quad (10)$$

Модельовання випадкових процесів, статистична обробка даних, побудова графіків та усі числові розрахунки виконувались із використанням мови програмування *python*, та бібліотек, написаних на основі цієї мови: *numpy* – бібліотека числових розрахунків [12], *scipy* – бібліотека основних математичних алгоритмів для наукових розрахунків [13], бібліотека побудови 2 та 3-мірних графіків.

Гістограми отриманих вибірок x для двох розподілів разом з графіками теоретичних законів розподілу (густин ймовірності розподілу $f(x)$) зображено на рис. 1.

ККІС, як і у загальному контрольні карти, відображають вибірки індивідуальних спостережень у відповідності до часових значень або у загальному випадку відповідності до номеру індивідуального спостереження (рис. 2).

Перевіримо нормальність отриманих розподілів двома способами: за рахунок використання графічного підходу та статистичного показника. Графічний метод полягає у побудові нормального імовірнісного графіку, де по осі абсцис відкладаються значення вибірки x , а по осі ординат – z -оцінки або стандартні оцінки, тобто кількість стандартних відхилень від середнього для теоретичного нормального розподілу.

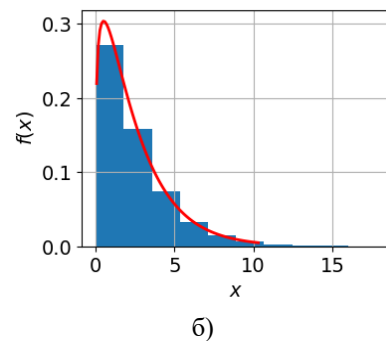


Рис. 1. Гістограми змодельованих вибірок. а) нормальний закон розподілу; б) χ^2 розподіл

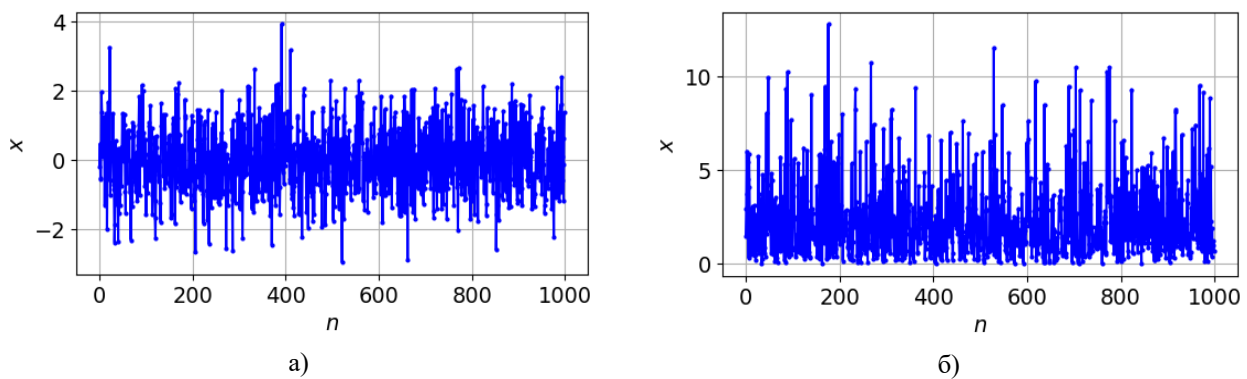


Рис. 2. Графіки змодельованих вибірок залежно від номеру індивідуального спостереження. а) нормальний розподіл; б) χ^2 розподіл

Даний графік використовується для оцінки відхилення розподілу генеральної сукупності за отриманою вибіркою від нормального закону. Для вибірки за нормальним законом розподілу точковий графік розміщується вздовж прямої лінії.

Як видно з рис. 3, вибірка, яка має нормальний розподіл (рис. 3, а) розміщена вздовж прямої

лінії і можна впевнено сказати, що вибірка отримана із нормальної генеральної сукупності. У випадку рис. 3, б, присутнє явне відхилення від прямої лінії, що явно свідчить про не-нормальність тестової вибірки.

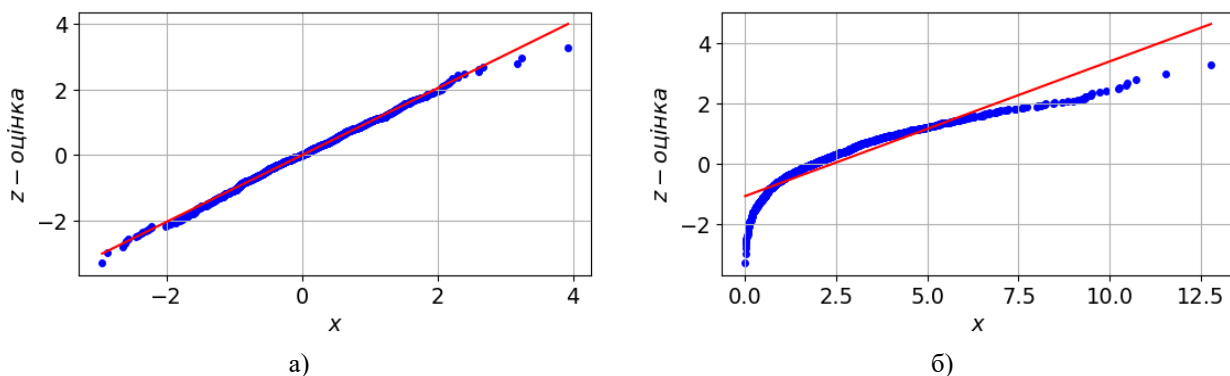


Рис. 3. Нормальні імовірнісні графіки. а) нормальний розподіл; б) χ^2 розподіл

Іншою оцінкою нормальності розподілу вибірки є статистичний показник, а саме критерій узгодженості Андерсона-Дарлінга. Даний критерій використовують для перевірки гіпотези, що досліджувана вибірка отримана із нормальної генеральної сукупності. У результаті цього тесту розраховують значення критерію Андерсона-Дарлінга ad та його модифікованого критерію ad^* для кожної вибірки. На основі цих критеріїв розраховують p -значення нульової гіпотези. Якщо значення $p \geq 0,05$, то приймається нульова гіпотеза – досліджувана вибірка отримана із нормальної генеральної сукупності. У випадку $p < 0,05$ – можна зробити висновок, що закон розподілу досліджуваної вибірки не-нормальний.

У табл. 1 наведено значення розглянутих статистик для досліджуваних двох вибірок, нормаль-

но розподіленої та з χ^2 розподілом. Отримані значення p -критерію для двох досліджуваних вибірок підтверджують, яка вибірка розподілена за нормальним законом, а яка ні.

Таблиця 1. Порівняння статистик початкових та нормалізованої вибірок

Розподіл вибірки	ad	ad^*	p
нормальний	0,705	0,705	0,066
χ^2	34,946	34,972	$5 \cdot 10^{-77}$
нормалізований	0,552	0,552	0,155

Дослідивши вибірки на нормальність їх генеральної сукупності, будемо ККІС. Відповідно до формул (2) – (9) для двох вибірок будемо карти ковзних розмахів (рис. 4, б, г) та відповідні ККІС (рис. 4, а, в).

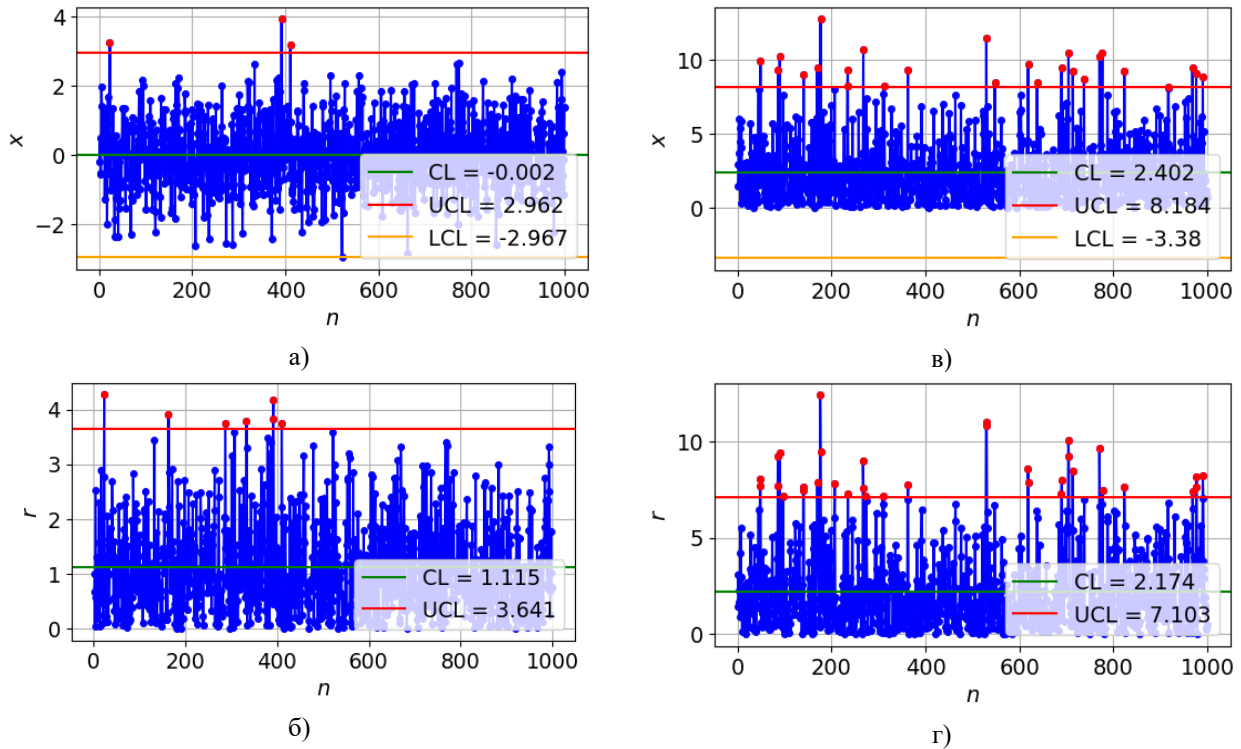


Рис. 4. ККІС для двох вибірок: а) ККІС для нормального розподілу; б) карта ковзних розмахів для випадку а); в) ККІС для χ^2 розподілу; г) карта ковзних розмахів для випадку в)

У випадку нормального розподілу вибірки, на графіках ККІС індивідуальні значення та ковзні розмахи знаходяться у контрольних межах, за винятком незначної кількості випадків виходу за них (помилка I роду), що можна пояснити тим, що межі карт містять $\pm 3\sigma$, тобто 99,7 % можливих значень. Відповідно для вибірки із $n = 1000$ значень (як у нашому дослідженні) кількість помилок I роду складає приблизно 3 значення на вибірку (рис. 4, а).

ККІС для χ^2 розподілу містить значну кількість хибних сигналів про розлад процесу (помилка I роду). У випадку процесу із малою кількістю

можливих спостережень, це є надзвичайно важливою помилкою.

Щоб позбавитись розглянутого недоліку у випадку не-нормального розподілу вибірки, перетворимо її для отримання нормально розподіленої вибірки або вибірки з квазі-нормальним законом розподілу, тобто виконаємо нормалізацію використовуючи метод Вох-Сох нормалізації [15]. У результаті отримаємо нову вибірку, гістограма розподілу якої зображена на рис. 5, а. Оскільки нова вибірка не може бути інтерпретована у одиницях досліджуваного параметру процесу, виконаємо її приведення до стандартної форми (рис. 5, б).

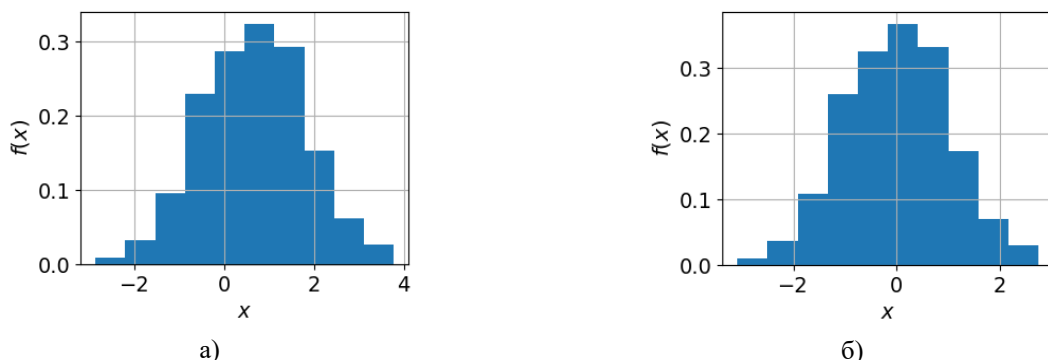


Рис. 5. Гістограми нормалізованої вибірки: а) нормалізована вибірка; б) стандартизована нормалізована вибірка

Виконаємо тестування нормальності розподілу трансформованої вибірки. З нормального імовірнісного графіку (рис. 6) видно, що індивідуальні значення нормалізованої вибірки розподілені

вздовж прямої теоретичного закону нормального розподілу, що свідчить про нормальність розподілу досліджуваної вибірки.

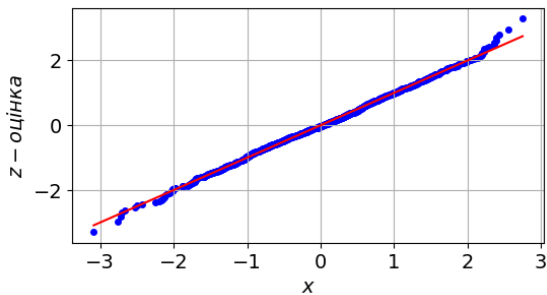
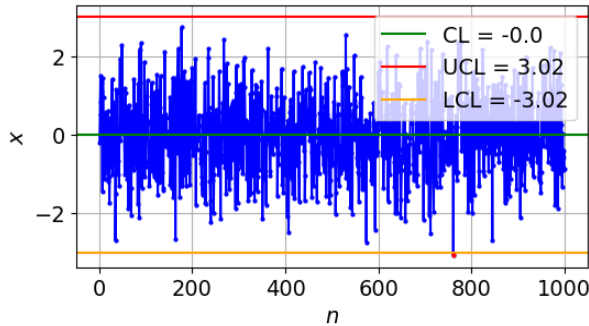


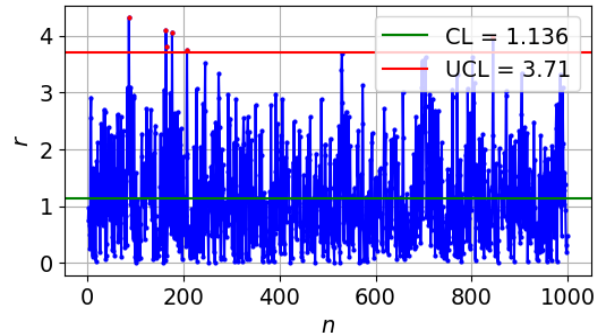
Рис. 6. Нормально ймовірнісний графік нормалізованої вибірки

Значення p -критерію з табл. 1 для нормалізованої вибірки більше межового, що також вказує на нормальність вибірки.

Побудовані графіки ККІС (рис. 7) підтверджують, що процес стабільний та відповідає вибірці, яка нормально розподілена (рис. 4, а, б).



а)



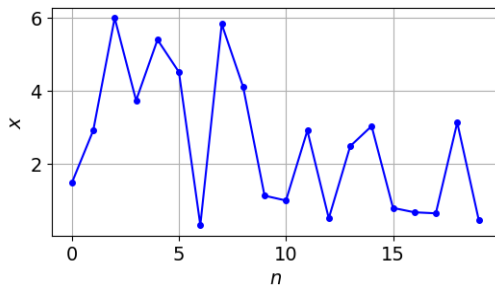
б)

Рис. 7. ККІС нормалізованої вибірки: а) індивідуальних спостережень; б) ковзних розмахів

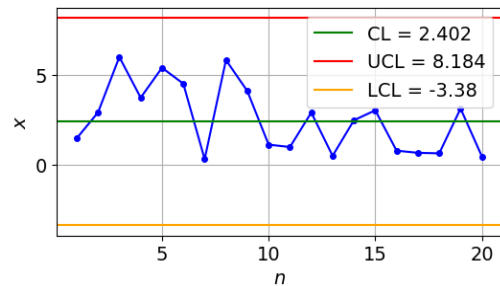
Дослідимо ККІС для розладнаного процесу

Змодельємо розладнаний процес розподілений за χ^2 розподілом та дослідимо застосування ККІС для виявлення розладнаності процесу. Створимо вибірку із 20 значень (рис. 8, а) та побудуємо

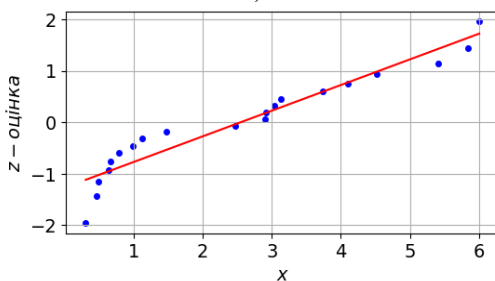
її нормально ймовірнісний графік (рис. 8, б), який вказує на ненормальність розподілу. Побудуємо ККІС створеної вибірки (рис. 8, в, г), з яких можна зробити висновок, що процес статистично керований (стабільний).



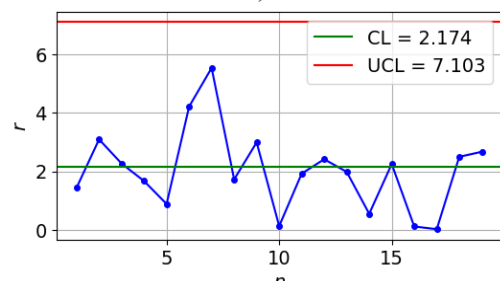
а)



в)



б)



г)

Рис. 8. Графіки малої тестової вибірки розподіленої за χ^2 розподілом. а) графік індивідуальних значень; б) нормально ймовірнісний графік; в) ККІС малої вибірки; г) графік ковзних розмахів

Збільшимо одне значення вибірки з номером $n = 9$ для імітації розладу процесу. Створимо дві нові вибірки з різними значеннями збільшення: вибірку із збільшенням в 10 разів та в 20 разів відповідно (рис. 9).

У випадку 10 разового збільшення одного значення (рис. 9, а), ККІС для χ^2 розподіленої вибірки

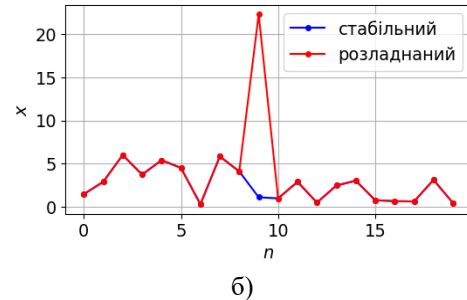
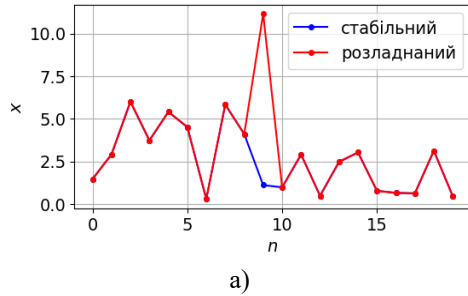


Рис. 9. Графіки вибірок змодельованих можливих розладів з χ^2 розподілом, шляхом збільшення вибіркового значення з $n = 9$. а) значення збільшено в 10 разів; б) значення збільшено у 20 разів

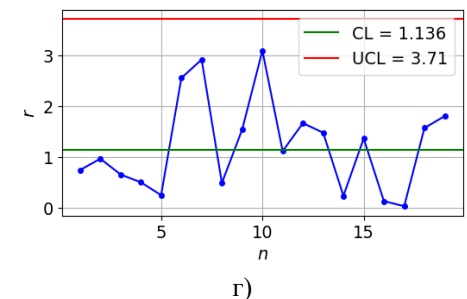
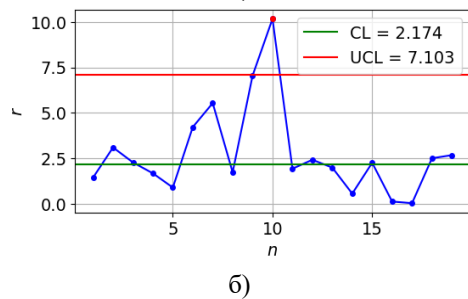
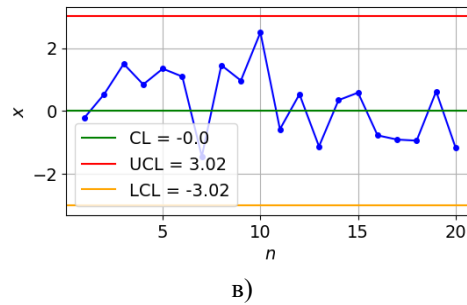
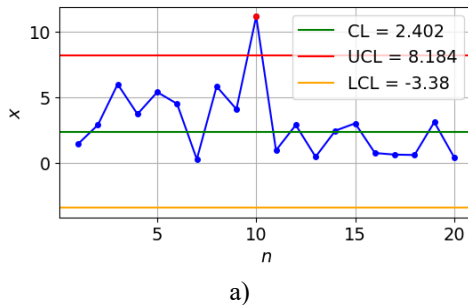


Рис. 10. ККІС вибірки з χ^2 розподілом із збільшенням одного значення в 10 разів. а) та б) початкова вибірка; в) та г) нормалізована

У випадку 20 разового збільшення одного значення (рис. 9. б), ККІС для χ^2 розподіленої вибірки (рис. 11, а, б) сигналізують про значний, порівняно з 10 разовим збільшенням, вихід процесу з ладу, а ККІС цієї ж вибірки після нормалізації також фіксує розлад процесу в тій самій точці (рис. 11, в, г). Можна зробити висновок, що нормалізація суттєво асиметричного процесу дозволяє позбавитись від помилок I роду та визначити справжній розлад процесу.

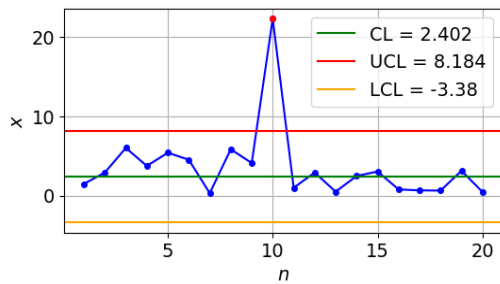
Висновки

Результати проведеного дослідження показали, що застосування контрольних карт індивідуальних значень для вибірок із значно асиметричним законом розподілу значень, таких, наприклад, які

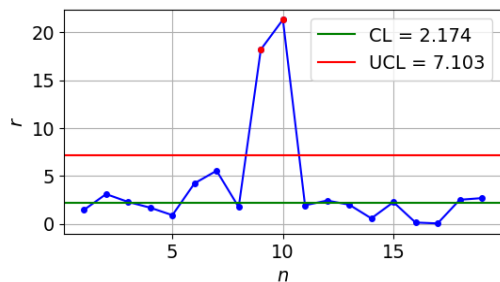
(рис. 10, а, б) сигналізують розлад процесу, а ККІС цієї ж вибірки після нормалізації розладання не фіксує (рис. 10, в, г). Проте фіксується не випадковість зміни процесу. У випадку початкової вибірки з χ^2 розподілом ККІС сигналізує помилку I роду.

мають χ^2 розподіл, у результаті контролю ККІС, призводить до збільшення появи помилок I роду, чого майже немає у випадку застосування методу нормалізації початкової вибірки із використанням Вох-Сох нормалізації. Моделювання розладання процесу внаслідок зміни певного індивідуального значення, також продемонструвало появу хибних сигналів розладання процесу, що не сигналізується на нормалізованій вибірці, проте сигналізується невідповідною зміною процесу. Запропонований алгоритм використання ККІС із послідовним застосуванням методів дослідження нормальності вибірки, коли процес досліджується, із подальшим застосуванням знайдених коефіцієнтів нормалізованої вибірки для нових значень, які отримані у процесі контролю, дозволяють застосовувати

ККІС для контролю ненормально розподіленого

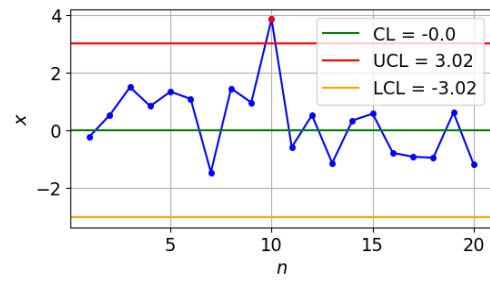


а)

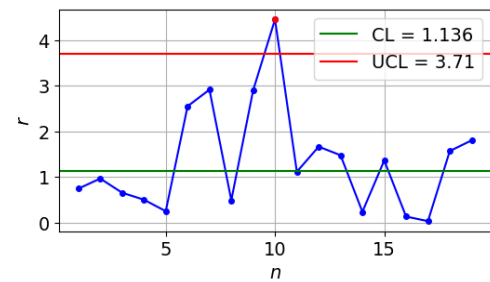


б)

процесу.



в)



г)

Рис. 11. ККІС вибірки з χ^2 розподілом із збільшенням одного значення в 20 разів: а) та б) початкова вибірка; в) та г) нормалізована

Література

- [1] Charles P. Quesenberry, "The Effect of Sample Size on Estimated Limits for \bar{X} and X Control Charts", *Journal of Quality Technology*, 25:4, 237-247, 1993. DOI: 10.1080/00224065.1993.11979470.
- [2] Burr, I. W., "The Effect of Non-Normality on Constants for \bar{X} and R Charts," *Industrial Quality Control*, vol. 23, no. 11, May 1967, pp. 563-568.
- [3] S. A. Yourstone and W. J. Zimmer, "Non-normality and the design of control charts for averages", *Decision Sciences*, 23 (5):1099-113, 1992. DOI:10.1111/j.1540-5915.1992.tb00437.x.
- [4] D.-S. Bai, and I. Choi. "X and R control charts for skewed populations", *Journal of Quality Technology*, 27 (2):120-31, 1995. DOI:10.1080/00224065.1995.11979575.
- [5] Balakrishnan, N., and S. Kocherlakota. "Effects of Nonnormality on \bar{X} Charts: Single Assignable Cause Model." *Sankhyā: The Indian Journal of Statistics, Series B (1960-2002)*, vol. 48, no. 3, 1986, pp. 439-44. JSTOR, <http://www.jstor.org/stable/25052468>.
- [6] E. G. Schilling, & P. R. Nelson, "The effect of non-normality on the control limits of X charts", *Journal of Quality Technology*, 8(4), 183-188, 1976. DOI: 10.1080/00224065.1976.11980743.
- [7] Osama H. Arif, Ali A. Al-Shomrani, Ahmed Ibrahim Shawky & Muhammad Aslam, "Control chart for log-logistic using quantile approach", *Journal of Statistics and Management Systems*, 23:8, 1571-1585, 2020. DOI: 10.1080/09720510.2020.1816691.
- [8] Marzieh Khakifirooz, V. G. Tercero-Gómez & W. H. Woodall, "The role of the normal distribution in statistical process monitoring", *Quality Engineering*, 33:3, 497-510, 2021. DOI: 10.1080/08982112.2021.1909731
- [9] M. Aichouni, A. Al-Ghonamy, L. Bachioua, "Control Charts for Non-Normal Data: Illustrative Example from the Construction Industry Business", *Mathematical and Computational Methods in Science and Engineering*, pp. 71-76, 2014. ISBN: 978-960-474-372-8 <https://api.semanticscholar.org/CorpusID:18505575>.
- [10] Figueiredo, F.; Gomes, M.I. Box-Cox transformations and robust control charts in SPC. *Adv. Math. Comput. Tools Metrol. VII 2006*, 72, 35-46. DOI:10.1142/9789812774187_0004.
- [11] ISO. ISO 7870-2:2013 Control charts – Part 2: Shewhart control charts. Geneva: International organization for standardization (ISO), 2013.
- [12] Harris, C.R., Millman, K.J., van der Walt, S.J. et al. "Array programming with NumPy", *Nature*, 585, 357-362, 2020. DOI: 10.1038/s41586-020-2649-2.
- [13] Virtanen, P., Gommers, R., Oliphant, T.E. et al. "SciPy 1.0: fundamental algorithms for scientific computing in Python", *Nat Methods*, 17, 261-272, 2020. DOI: 10.1038/s41592-019-0686-2
- [14] J. D. Hunter, "Matplotlib: A 2D Graphics Environment", *Computing in Science & Engineering*, vol. 9, no. 3, pp. 90-95, 2007.

[15] G. E. P. Box, and D. R. Cox, "An analysis of transformations", *J. Royal Statist. Soc. B*, 26 (2), pp. 211-252, 1964.

UDC 519.24 (075.8)

Oleh Kozyr, Oleksandra Shestak

National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine

INDIVIDUAL CONTROL CHARTS FOR NON-NORMAL DATA

Individual control charts are a suitable option to consider when the processes being studied change slowly. If obtaining standard samples for Shewhart control charts is expensive or in the case of studying statistical characteristics of the technological process, they are used. Control charts are used to control technological processes in the case of a normal distribution of the possible values of the controlled parameter. Sometimes a quasi-normal distribution of the studied parameter is allowed. In practice, it is often necessary to deal with technological processes in which samples are not distributed according to a normal distribution. In such cases, using individual control charts can lead to type I and II errors. This is because the control limits of the charts are calculated for the mean and standard deviation of the normal distribution. Control charts of individual values are most sensitive to the deviation of the distribution from the normal one since individual values are used to construct the limits, to which the central limit theorem does not apply, as in the case of control charts of mean values. Therefore, it is extremely important to develop methods of applying the individual control charts for random samples, the general population of which is not normally distributed.

The study examines the application of individual control charts for a random process that is non-normally distributed. This is achieved by using the normalization method to transform samples of the studied parameter into samples that are distributed normally or quasi-normally. The normalization is performed using the Box-Cox method with further construction of control charts based on the transformed samples. The mathematical apparatus of the individual control chart application for samples with non-normal distribution is described. Numerical modeling of the application of individual control charts for the chi-square distributed samples, which have a significant distribution asymmetry for small values of the shape factor, and for the samples obtained by normalizing the chi-square one, was performed.

According to the research findings, the utilization of individual control charts for samples exhibiting a significantly asymmetric distribution results in a higher likelihood of type I errors being detected. However, when the method of sample normalization is employed, this issue is almost entirely eliminated. Modeling process disruption, by changing a certain individual value, also demonstrated the emergence of false signals of process disruption.

Keywords: individual control charts, non-Gaussian distribution, non-normal distribution, chi-square distribution, Box-Cox normalization.

*Надійшла до редакції
18 грудня 2023 року*

*Рецензовано
12 лютого 2024 року*



© 2024 Copyright for this paper by its authors.
Use permitted under Creative Commons License Attribution 4.0 International
(CC BY 4.0).