

ВПОРЯДКОВАНІ ВИБІРКИ НЕКОРЕЛЬОВАНИХ ДАНИХ ЯК ОДНОЕЛЕМЕНТНІ БАГАТОКАНАЛЬНІ КОНВЕРТОРИ

Р.О. Мазманян*, докт.техн.наук

Інститут електродинаміки НАН України,
пр. Перемоги, 56, Київ, 03057, Україна,

e-mail: mazmanian@ied.org.ua

Достатня близькість медіан упорядкованих вибірок випадкових даних до нормального розподілу використовується в комп'ютерних системах управління, моніторингу та діагностування електроенергетичного обладнання. Однак залишається невідомим, чи мають подібну схожість функції щільності розподілу (probability density function, pdf) інших елементів вибірок з непарним числом елементів. Запропоновано концепцію багатоканального конвертора функції щільності ймовірності (pdf-конвертор). У статті представлено результати перевірки параметричними та статистичними методами гіпотез для каналів pdf-конвертора – елементів (статистик) упорядкованих вибірок заданих розмірів з обмежених за величиною випадкових чисел з рівномірним розподілом. Застосування різних методів досліджень за однакових умов показали достатню близькість результатів тестування за трьома статистичними критеріями. Це дозволило прийняти обґрунтований висновок щодо переваги у застосуванні запропонованого методу використання критерію χ^2 -квадрат для оцінки згоди з нормальним розподілом аналітичних pdf каналів конвертора. Також було запропоновано використання «вікна статистичної близькості» для виділення тих каналів конвертора, які не мають істотних відмінностей з нормальним розподілом. На додаток представлено емпіричну формулу, яка визначає залежність розміру вікна статистичної близькості від розміру вибірки. Результати досліджень узагальнені в створеній статистичній моделі багатоканального конвертора некорельованих даних. Бібл. 27, рис. 7.

Ключові слова: випадкові дані, рівномірний розподіл, упорядковані вибірки, функція щільності ймовірності (pdf), pdf-конвертор, оцінка щільності, критерій χ^2 -квадрат.

Вступ. Підвищена небезпека обслуговування і експлуатації електроенергетичного обладнання, наявність різнохарактерних електромагнітних полів зі змінюваним у широкому діапазоні рівнем може викликати неприпустимі спотворення вимірювальної інформації щодо його поточного стану. Перехідні процеси, викликані змінами режимів роботи як спостережуваного, так і суміжного обладнання, можуть супроводжуватися одиничними або груповими імпульсними складовими, такими, що порушують природну інерційність процесів, що спостерігаються.

Подібні порушення можуть призвести до невірної оцінки, наприклад, середніх значень контрольованих параметрів у відносно малих часових інтервалах або неприпустимих спотворень зображень полів діагностичних параметрів. Ці та інші обставини вимагають застосування спеціальних алгоритмічних процедур і програмних засобів для попередньої обробки вимірювальної інформації в комп'ютерних системах управління, контролю параметрів, моніторингу і діагностування стану електроенергетичного обладнання.

Упорядкування і вибір медіани зі списку номерів елементів упорядкованих вибірок [1, 2] ефективно використовуються задля альтернативної оцінки середнього значення різних даних. Модифікації цієї простої процедури знаходять застосування в згладжуванні [3, 4], нормалізації [5] та обробці даних статистичних досліджень [6, 7], усуненні імпульсних завад і придушенні шуму в цифрових зображеннях [8, 9]. У ряді випадків для забезпечення порівняння контрольованих параметрів у широкому діапазоні змін потрібна нормалізація даних, тобто зведення їх до єдиної номінальної шкали. Тому дослідження селективних властивостей впорядкованих вибірок в перетвореннях випадкових сигналів переважно були орієнтовані на одномірні та двовимірні медіанні фільтри [10, 11, 12].

В той же час існує багато практичних прикладів, коли неможливо ігнорувати значні за величиною випадкові складові інформації, вплив яких може зростати в процесі її обробки. Обробка даних впорядкованими вибірками також може бути піддана їхньому впливу. Тому дослідження і формалізація перетворювальних властивостей впорядкованих вибірок випадкових даних мають практичне значення для підвищення достовірності інформації, що обробляється обчислювальними методами.

Вибірка і впорядкування даних виконуються наступним чином. З деякої множини $\{N\}$ (генеральної сукупності розміром N) некорельованих, обмежених за величиною дійсних чисел і із загальною функцією розподілу, відбираються [13] прості випадкові вибірки H_1^s з непарним числом елементів $s = 2d + 1$, де d – порядок вибірок. Упорядкування вибірок за значенням елементів здійснюється процедурою сортування [14] Sort. Багаторазове виконання відбору і сортування перетворює множину $\{N\}$ випадкових величин в s множин випадкових величин $X_n^{(s)}$ з розмірами $n = \left\lfloor \frac{N}{s} \right\rfloor$ і перетвореними розподілами щільності ймовірності.

Очевидно, що процедура упорядкування викликає суттєві зміни щільності ймовірності (*pdf*) вхідної випадкової величини, а розміри вибірок і номер елемента в вибірках визначають нелінійний характер цих змін. Концептуально процедури створення вибірок випадкових даних, їхнє упорядкування та дані елементів упорядкованих вибірок утворюють окрему реалізацію s -канального конвертора щільності розподілу вхідний випадкової величини (*pdf*-конвертора) (рис. 1).

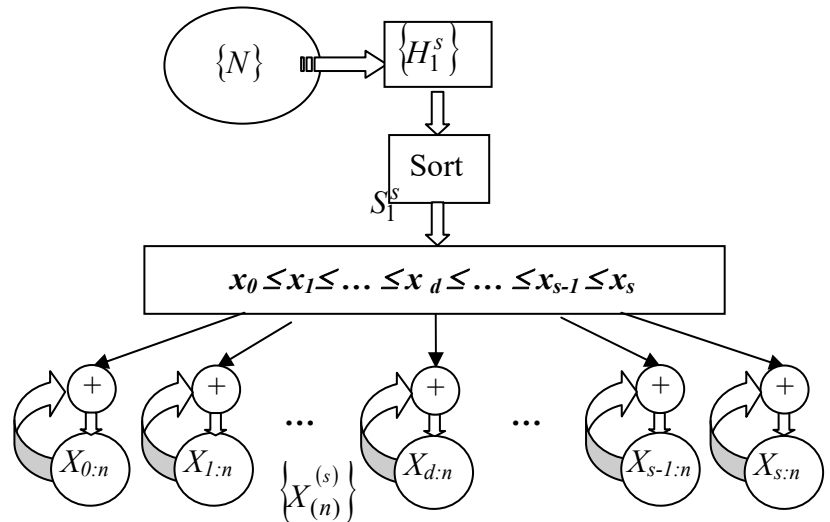


Рис. 1

Основні характеристики елементів упорядкованих вибірок з множини випадкових даних.

Основні характеристики елементів (статистик) упорядкованих вибірок з непарним числом елементів $s = 2d + 1$, де d – порядок вибірок для обмеженої за значенням випадкової величини $-X_m \leq x \leq X_m$ з рівномірним розподілом щільності визначаються наступними аналітичними виразами [15], у яких $k = 0 \dots s$ – номер каналу конвертора:

- щільність ймовірності

$$p_{d,k}(x) = \frac{1}{2X_m} \cdot \frac{(2d+1)!}{k!(2d-k)!} \cdot \left(\frac{X_m+x}{2X_m}\right)^k \cdot \left(\frac{X_m-x}{2X_m}\right)^{2d-k}; \quad (1)$$

- середні значення елементів вибірок

$$\mu_{d,k} = \frac{(2 \cdot d + 1)!}{2^{2 \cdot d + 1}} \cdot X_m \cdot \sum_{j=0}^k \sum_{m=0}^{2 \cdot d - k} \frac{(-1)^m - (-1)^j}{j! \cdot m! \cdot (k-j)! \cdot (2 \cdot d - k - m)! \cdot (m+j+2)}, \quad (2)$$

або у спрощеному вигляді

$$M_{d,k} = \frac{X_m}{d+1} \cdot (k+1) - X_m; \quad (3)$$

- дисперсія елементів упорядкованих вибірок

$$D_{d,k} = \alpha_{d,k}^{(2)} - \mu_{d,k}^2 = \left[X_m^2 \cdot \frac{(2 \cdot d + 1)!}{2^{2 \cdot d + 1}} \cdot \sum_{j=0}^k \sum_{m=0}^{2 \cdot d - k} \frac{(-1)^m + (-1)^j}{j! \cdot m! \cdot (k-j)! \cdot (2 \cdot d - k - m)! \cdot (m+j+3)} \right] - \left[X_m^2 \cdot \frac{(2 \cdot d + 1)!}{2^{2 \cdot d + 1}} \cdot \sum_{j=0}^k \sum_{m=0}^{2 \cdot d - k} \frac{(-1)^m - (-1)^j}{j! \cdot m! \cdot (k-j)! \cdot (2 \cdot d - k - m)! \cdot (m+j+2)} \right]^2; \quad (4)$$

- середньоквадратичні відхилення елементів

$$\sigma_{d,k} = \sqrt{D_{d,k}}. \quad (5)$$

Постановка задачі. Оцінки функцій щільності розподілу експериментальних даних, отриманих в результаті теоретичних досліджень, використовуються для їхньої заміни єдиною функцією або сімейством функцій, властивості яких добре відомі. У багатьох випадках особливий інтерес представляє близькість до нормального закону розподілу. Для медіан упорядкованих вибірок з випадкової некорельованої послідовності подібне оцінювання виконувалося окремо в двох аспектах [16, 17]: параметричне порівняння як оцінка похибки апроксимації функцій щільності ймовірності (1) функціями Гауса та статистична близькість до нормального закону розподілу елементів багатоканального конвертора. Строгий математичний опис перетворювальних властивостей всіх елементів упорядкованих вибірок з випадкових даних може бути використаний для порівняльного аналізу, дослідження існуючих і побудови нових методів обробки даних.

Мета досліджень полягає у визначенні перетворювальних властивостей впорядкованих вибірок (*pdf*-конверторів) з множини обмежених за величиною випадкових даних з рівномірним розподілом і отриманні оцінок параметричної та статистичної згоди або відмінності з нормальним розподілом елементів багатоканального конвертора, створенні статистичної моделі перетворення.

Оцінка параметричної близькості елементів до функції Гауса. Розглянемо можливість наближення незміщеними функціями Гауса для кожного з $0 \leq k \leq 2d$ каналів конвертора. Приймається гіпотеза H_0 , згідно з якою існує близькість досліджуваних функцій до відповідних функцій Гауса та альтернативна гіпотеза H_1 , яка її заперечує.

Відповідно до гіпотези H_0 , апроксимуючі функції Гауса [18] побудовані з використанням середньоквадратичних відхилень (5) і середніх значень $\mu_{d,k}$ (2) функції (1), що апроксимується, для порядку $d = 1 \dots 25$ упорядкованих вибірок з випадкової змінної $-X_m \leq x \leq X_m$ та її прийнятого граничного значення $X_m = 0.5$

$$G_{d,k}(x) = \frac{1}{\sigma_{d,k} \sqrt{2\pi}} \exp\left(-\frac{(x - \mu_{d,k})^2}{2\sigma_{d,k}^2}\right). \quad (6)$$

Підходящою оцінкою наближення щільності ймовірності кожного з елементів упорядкованих вибірок функцією Гауса можуть служити наведені до діапазону змін випадкової величини інтегральні значення відхилень. Для кожної з функцій каналів k впорядкованої вибірки заданого порядку d відхилення визначаються залежностями (1), (6)

$$E_{d,k} = \frac{1}{2X_m} \cdot \int_{-X}^X |p_{d,k}(x) - G_{d,k}(x)| dx. \quad (7)$$

На рис. 2, *a* показано незміщені середньоквадратичні відхилення елементів упорядкованих

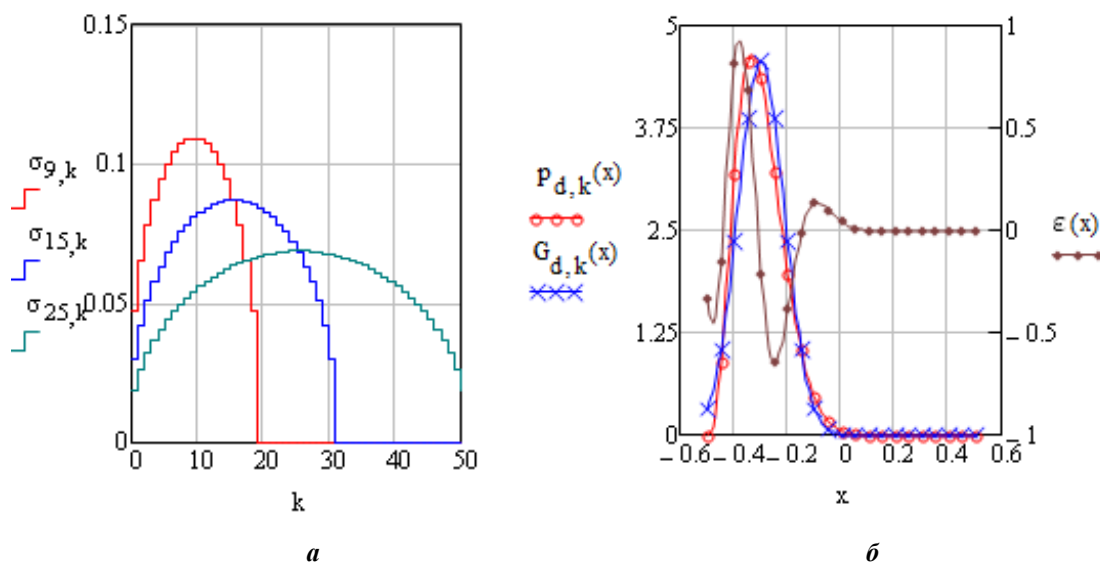


Рис. 2

вибірок для $d = 9, 15, 25$, на рис 2, *б* – функція щільності та апроксимуюча її функція Гауса для каналу

$k = 5$ вибірки з порядком $d = 15$. На цьому ж рисунку представлена різниця двох детермінованих функцій ε .

Сімейство інтегральних відхилень елементів показано на рис. 3. Тут можна відзначити очікувану центральну симетрію та мінімальні значення відхилень для медіан вибірок щодо умовно прийнятого граничного значення, рівного 0.1. Це свідчить про наявність двох груп елементів. Центральна щодо власних медіан група елементів не перевищує граничне значення і допускає згоду за умовою з гіпотезою H_0 . Друга, периферійна група, заперечує її на користь гіпотези H_1 .

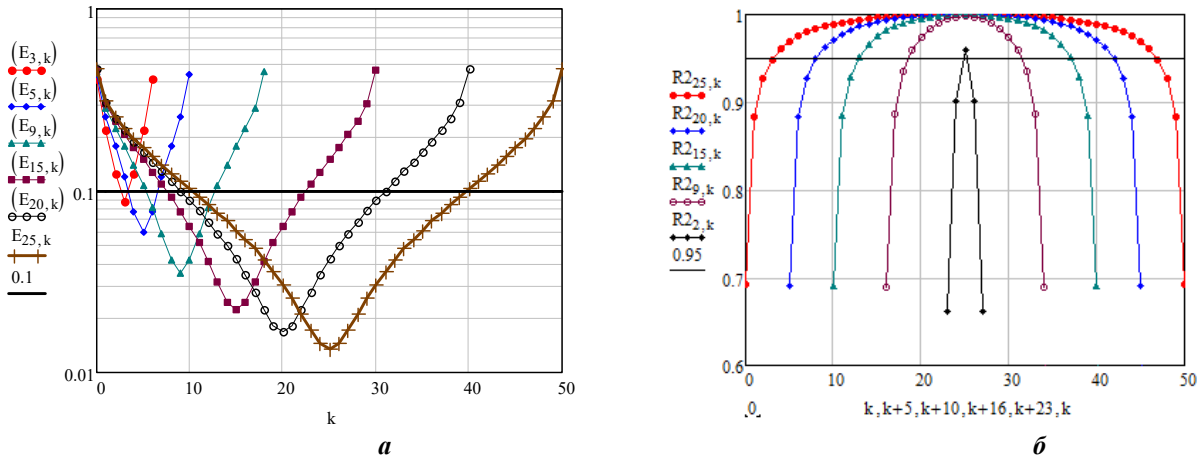


Рис. 3

Коефіцієнти детермінації R^2 [20] елементів вибірок як альтернативні оцінки якості апроксимації використано для аналітичних функцій (1), (6)

$$R^2_{d,k} = 1 - \frac{SSres_{d,k}}{SStot_{d,k}}, \quad (8)$$

де $SSres_{d,k} = \int_{-X_m}^{X_m} [p_{d,k}(x) - G_{d,k}(x)]^2 dx$ – сума квадратів залишків; $SStot_{d,k} = \int_{-X_m}^{X_m} [p_{d,k}(x) - \bar{p}_{d,k}]^2 dx$

– загальна сума квадратів; $\bar{p}_{d,k} = \int_{-X_m}^{X_m} p_{d,k}(x) dx = 1$ – середнє значення елементів вибірок.

На рис. 3, б наведено значення коефіцієнтів детермінації $R^2_{d,k}$ каналів вибірок з порядком $d = 2, 9, 15, 20, 25$ і умовний поріг 0.95, що істотно перевищує рекомендовану величину під час оцінки якості апроксимації [20]. І в цьому випадку можна відзначити відсутність строгих правил завдання порога та значно меншу параметричну чутливість критерія R^2 у порівнянні з інтегральними відхиленнями E (рис. 3, а). Відсутність обґрунтованих критеріїв параметричного порівняння не дозволяє робити вичерпні висновки щодо якості апроксимації, але при цьому придатні задля зіставлення результатів різних підходів в оцінках.

Критерії статистичної близькості каналів до нормального закону розподілу. Розглянемо критерії статистичної близькості каналів вибірок до нормального закону розподілу через міри скошеності, тобто асиметрії, та міри гостровершинності – ексцесу (куртозиса) їхніх функцій розподілу [19, 20, 21]. Для абсолютно неперервних розподілів елементів упорядкованих вибірок асиметрія та ексцес визначаються їхніми n -ми центральними моментами, місце розташування яких задається зсувами $M_{d,k}$ з (3)

$$\mu_{d,k} = \frac{1}{2X_m} \cdot \frac{(2d+1)!}{k!(2d-k)!} \int_{-X_m}^{X_m} [x - M_{d,k}]^n \left[\left(\frac{X_m + x}{2X_m} \right)^k \cdot \left(\frac{X_m - x}{2X_m} \right)^{2-d-k} \right] dx. \quad (9)$$

Коефіцієнт асиметрії визначається відношенням третього центрального моменту $\mu_{3d,k}$ до третього ступеня $\sigma_{d,k}^3$ стандартного відхилення елементів вибірок (13)

$$\gamma_{1d,k} = \mu_{3d,k} \sigma_{d,k}^{-3}. \quad (10)$$

Аналогічно розраховувалися коефіцієнти ексцесу з моменту четвертого порядку

$$\gamma_{2d,k} = \mu_{4d,k} \sigma_{d,k}^{-4}. \quad (11)$$

Зсув $M_{e_{d,k}} \approx 3$ коефіцієнтів для різних порядків вибірок визначався значеннями медіан коефіцієнта ексцесу $\gamma_{2d,d}$.

Розглянемо спосіб оцінки статистичної близькості елементів упорядкованих вибірок до нормального закону розподілу за критерієм подібності Пірсона [25]. Критерій, призначений для оцінки розходжень між вибіркової і аналітичними функціями щільності, може бути використаний і для зіставлення двох аналітичних функцій. При цьому в статистиці χ^2 [26] природним чином будуть збережені порівнювані характеристики як спостережуваної щільності розподілу – елементів вибірки $p_{d,k}(x)$, так і очікуваної – нормального закону розподілу Гауса $G_{d,k}(x)$, які визначаються формулами (1), (6).

Статистика вибірових значень критерію згоди виражається формулою

$$\chi^2 = n \cdot \sum \frac{(p_{d,k}(x_i) - G_{d,k}(x_i))^2}{G_{d,k}(x_i)}, \quad (12)$$

де n – число дискретних значень функцій розподілу щільності ймовірності.

Оцінка близькості щільності розподілу елементів до функції Гауса. Оцінка статистичної близькості щільності розподілу до нормального закону розподілу буде проводитися трьома методами – аналізом коефіцієнтів асиметрії та ексцесу і використанням критерію згоди *xi-квадрат*.

Метою перевірки є визначення сукупності елементів вибірок, для яких гіпотеза H_0 , що перевіряється, узгоджується з їхніми функціями щільності розподілу і не відхиляється. Розглянемо показники близькості, оціночні або критичні рівні та результати їхнього застосування до кожного з цих критеріїв.

Оцінки асиметрії A і ексцесу E засновані на їхній рівності нулю для нормального розподілу. Значимість відхилень від середніх оцінюється дисперсіями асиметрій $|D(A)|$ і ексцесу $D(E)$ елементів. Критерій згоди в цьому випадку формулюється так [22]: розподіл елементів, що спостерігаються, близький до нормального, якщо їхня асиметрія і ексцес задовольняють нерівностям

$$A \leq 3\sqrt{D(A)}, \quad E \leq 5\sqrt{D(E)}. \quad (13)$$

В іншому випадку гіпотезу нормальності відхиляють.

Дисперсія коефіцієнтів асиметрії елементів $\gamma_{1d,k}$ визначалася за формулою

$$D_{A_{d,k}} = \int_{-X}^X x^2 |\gamma_{1d,k}| dx. \quad (14)$$

Оціночні рівні елементів визначаються дисперсією зміщених значень коефіцієнта ексцесу $\gamma_{2d,d}$

$$D_{E_{d,k}} = \int_{-X}^X x^2 (\gamma_{2d,k} - M_{e_{d,k}}) dx. \quad (15)$$

На рис. 4, *a* наведено реалізації правил перевірки гіпотези H_0 по асиметрії для елементів вибірок порядку $d = 5$ і $d = 9$. Перевірку гіпотези H_0 по ексцесам елементів для тих же значень порядку показано на рис. 4, *б*.

Отримані результати дають змогу прийняти статистично обґрунтований висновок про існування групи центральних елементів, що не заперечують очікуваних за гіпотезою H_0 результатів. У той час як периферійні лівостороння і правостороння щодо медіани групи елементів їй суперечать.

Зіставлення числа елементів, згодних з гіпотезою H_0 в результаті статистичної перевірки за критеріями асиметрії (крива $[S^{(A)}]_d$), ексцесу (крива $[S^{(E)}]_d$) і розмірів вибірок (крива s_d), показано на рис. 5, *a*.

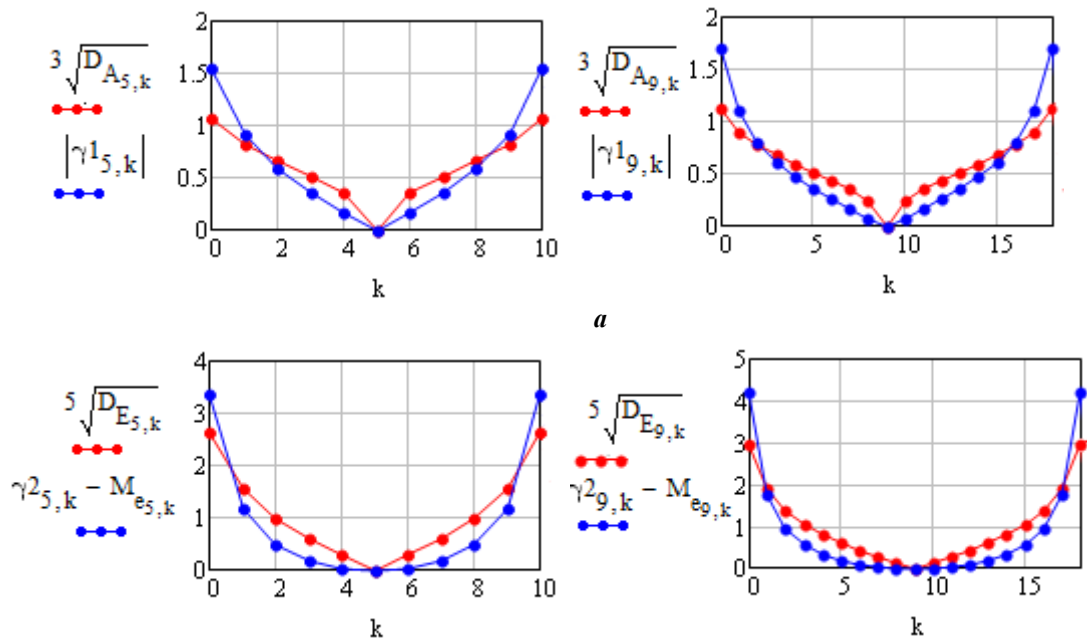


Рис. 4

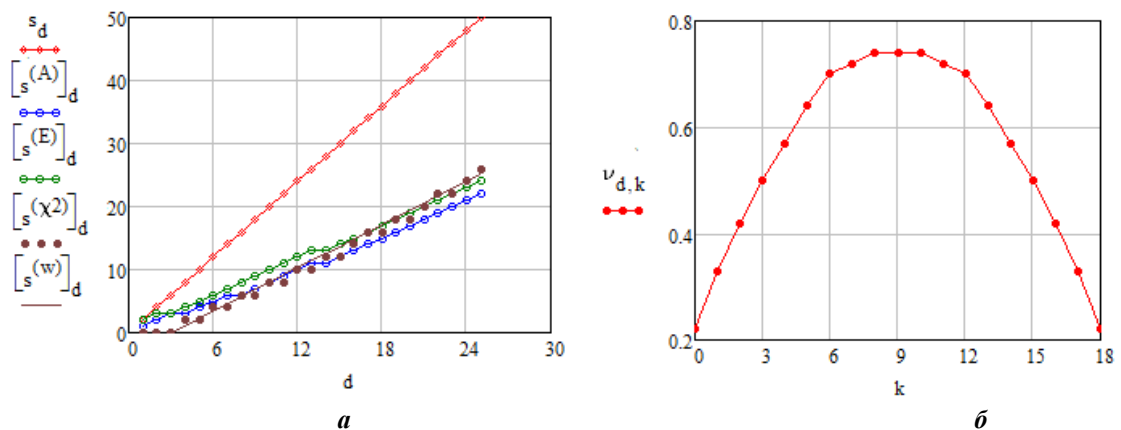


Рис. 5

Статистична перевірка випадкових величин $\{X_{(n)}^{(s)}\}$ на виході багатоканального конвертора з використанням критерію χ^2 -квадрат [23] вимагає формування двох векторів значень аналітично заданих досліджуваної та очікуваної щільності розподілу ймовірностей.

Розроблений метод використання критерію χ^2 -квадрат ґрунтується на наступній послідовності перетворень кумулятивної статистики критерію Пірсона з функції від значень порівнюваних частот подій в залежність від значень порівнюваних щільності ймовірності.

$$\chi^2 = \sum_{i=1}^N \frac{(O_i - E_i)^2}{E_i} = \sum_i \frac{N^2 \left(\frac{O_i - E_i}{N}\right)^2}{E_i} = \sum_i \frac{\left[N \cdot \left(\frac{O_i - E_i}{N}\right)\right]^2}{\frac{E_i}{N}} = N \sum_i \frac{(p_{d,k}(x_i) - G_{d,k}(x_i))^2}{G_{d,k}(x_i)}, \quad (16)$$

де O – кількість спостережень; E – очікуване число подій; N – загальна кількість подій.

При цьому в статистиці χ^2 [24, 25] природним чином будуть збережені порівнювані характеристики, як «спостережуваної» щільності розподілу елементів вибірки, так і очікуваної – нормального закону розподілу Гауса, які визначаються формулами (1), (6).

До визначення критичних значень. Значуще число із загального числа подій N і розмір інтервалів дискретизації функцій (1), (6) з незалежної випадкової величини $-X_m \leq x \leq X_m$ визначаються за допомогою, наприклад, одного з відомих методів розрахунку [26, 27].

Значимість рівновіддалених дискретних значень щільності ймовірності можна оцінити і за перевищенням встановленого порога. Для порядку вибірки $d = 9$ і порога, рівного 0.01, відносні числа значущих відліків $V_{d,k}$ мають вигляд, показаний на рис. 5, б.

Подальше рішення задачі перевірки статистичних гіпотез проводиться у звичайній послідовності: обчислюються статистики χ^2 -квадрат (18) для s випадкових величин $\{X_{(n)}^{(s)}\}$, задається рівень значущості α , приймається односторонній тип критичної області.

Граничні значення критичної області $t_{d,k}$ з квантилем розподілу χ^2 порядку 0.95 визначаються з числа ступенів свободи для кожного елемента окремо, тобто $df_{d,k} = n_{d,k} - 1 - 2$, $n_{d,k} = V_{d,k} \cdot r$, де r – загальне число відліків.

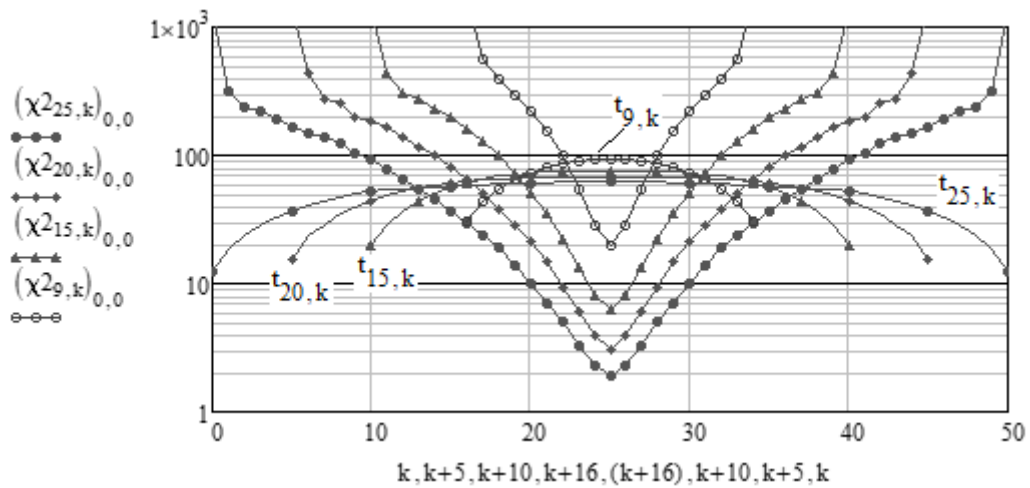
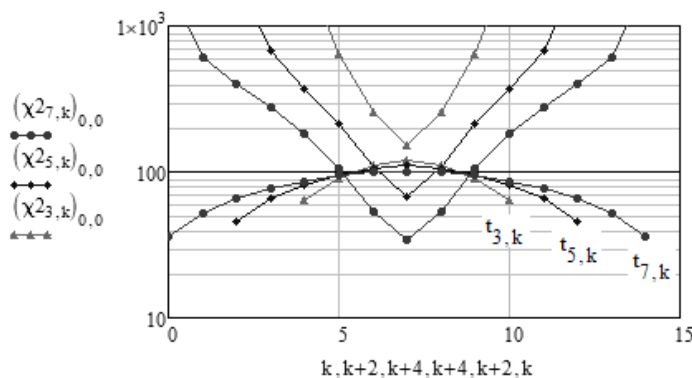


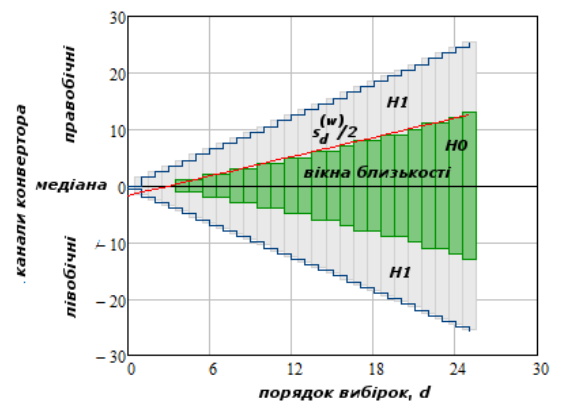
Рис. 6

Заключним етапом оцінки є застосування вирішальних правил. На рис. 6 наведено центровані щодо власних медіан статистики χ^2 -квадрат і межі правобічних критичних областей для елементів упорядкованих вибірок порядком $9 \leq d \leq 25$.

Статистики χ^2 для значень вибірок порядком $3 \leq d \leq 7$ окремо показано на рис. 7, а.



а



б

Рис. 7

Елементи функції щільності розподілу, розташовані нижче власних значень $t_{d,k}$ (рис. 6, 7, а), узгоджуються гіпотезою H_0 , що перевіряється. Ця гіпотеза суперечить функціям щільності розподілу елементів, розташованих вище критичних значень, для яких вона відхиляється на користь альтернативної гіпотези H_1 на рівні значущості α .

Отримані результати виділяють в вибірках групи елементів, для яких справедливо прийняття нульової гіпотези щодо статистичної близькості до очікуваного нормального закону розподілу (крива

$\left[s^{\chi^2} \right]_d$). Число елементів в групах залежить від порядку вибірок і має деякий розкид, викликаний різницею методів і критеріїв оцінки (рис. 5, а). Можна розглядати об'єднання цих груп в межах «вікон статистичної близькості» до нормального закону розподілу. Це призводить до необхідності прийняття результатів статистичної перевірки за критерієм згоди *хі-квадрат* для всіх вибірок, як більш суворої оцінки для малих порядків і більш рівномірною по всьому ряду розглянутих значень. Вікна статистичної близькості представлені разом з межами вибірок на рис 7, б. Розміри вікон (крива $\left[s^{(w)} \right]_d$) можна визначити лінійною залежністю від порядку вибірки наступною рівністю:

$$s_d^{(w)} = 0.571 \cdot (d - 3). \quad (17)$$

Результати досліджень узагальнюються статистичною моделлю багатоканального одноелементного конвертора щільності ймовірності, яка описується набором функцій Гауса (6), параметри яких визначаються середніми значеннями (2) або (3), дисперсіями (4) і середньоквадратичними відхиленнями (5) елементів вибірок, що виділяються вікном близькості. Крім цих елементів, до окремих груп можна віднести два елементи (8) з найбільшим і найменшим номерами в вибірках з довільним значенням порядку d . Щільність їхньої ймовірності дорівнює

$$p_d(x) = \frac{(2d+1)!}{(2X)^{2 \cdot d+1}} \cdot \begin{cases} (X-x)^{2d} & \text{для } k=0 \\ (X+x)^{2d} & \text{для } k=2d \end{cases}. \quad (18)$$

Висновки.

1. Пропоноване концепцію подання упорядкованих вибірок з некорельованої випадкової множини даних як багатоканальних одноелементних конверторів функції щільності ймовірності.

2. Визначення перетворювальних властивостей впорядкованих вибірок виконано шляхом оцінок близькості детермінованих функцій «спостережуваних» щільності ймовірності елементів (каналів конвертора) і апроксимуючих їх «очікуваних» функцій Гауса за допомогою параметричних і статистичних методів.

3. Параметричні оцінки апроксимації здійснюються за допомогою наведеного інтегрального відхилення і коефіцієнту детермінації досліджуваної та апроксимуючої функцій. Для умовних граничних значень параметричних оцінок показано наявність груп елементів для вибірок різних порядків, умовно згодних з гіпотезою H_0 .

3. Статистичну згоду двох сімейств функцій щільності ймовірності оцінено за допомогою трьох різних критеріїв – через міри скошеності, гостровершинності їхніх функцій розподілу та за допомогою критерію подібності Пірсона. Розроблений метод адаптації критерію *хі-квадрат* для оцінки сімейств аналітичних функцій щільності ймовірності зі змінними параметрами.

4. Результати досліджень узагальнені в створеної статистичної моделі багатоканального конвертора некорельованих даних. Модель включає групу статистично близькі до нормального закону розподілу випадкових величин каналів *pdf*-конверторів, які запропоновано виділяти «вікном статистичної близькості». Щільності розподілу ще двох груп каналів являють собою параболи другого і більш високих порядків.

5. Запропонована методика досліджень може бути використана для визначення перетворювальних властивостей багатоканальних одноелементних конверторів – упорядкованих вибірок випадкових величин з іншими абсолютно безперервними функціями щільності розподілу ймовірності.

Роботу виконано за рахунок коштів бюджетної програми "Розвиток наукових засад методів та засобів вимірювання експлуатаційних параметрів обладнання електростанцій» (0116U007285), КПКВК 6541030.

1. Luo D., Wan X., Liu J., Tong T. Optimally estimating the sample mean from the sample size, median, mid-range, and/or mid-quartile range. *Statistical Methods in Medical Research*. 2016. No 27(6). Pp. 1785-1805. DOI: <https://doi.org/10.1177/0962280216669183>.

2. Gervini D. Robust functional estimation using the median and spherical principal components. *Biometrika*. 2008. No 95(3). Pp. 587-600. DOI: <https://doi.org/10.1093/biomet/asn031>.

3. Qasim Nasir Husain, Mohd Bakri Adam, Mahendran Shitan, Anwar Fitrianto. Extension of Tukey's Smoothing Techniques. *Indian Journal of Science and Technology*. 2016. No 9(28). Pp.1-6. DOI: <https://doi.org/10.17485/ijst/2016/v9i28/97354>.

4. Jansen M. Multiscale Poisson data smoothing. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. 2006. No 68(1). Pp. 27-48. DOI: <https://doi.org/10.1111/j.1567-9868.2005.00531.x>.

5. Kappal S. Data Normalization Using Median & Median Absolute Deviation (MMAD) based Z-Score for Robust Predictions vs. Min-Max Normalization. *London Journal of Research in Science: Natural and Formal*. 2019. Vol. 19. Pp. 39-44. DOI:

<https://doi.org/10.13150/RG.2.2.32799.82088>

6. Cheng X-B, Wang F-K. The performance of EWMA median and CUSUM median control charts for a normal process with measurement errors. *Qual Reliab Engng Int*. 2018. Vol. 34(2). Pp. 203-213. DOI: <https://doi.org/10.1002/qre.2248>
7. Wan X., Wang W., Liu J., Tong T. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. *BMC Med Res Methodol*. 2014. Vol. 14(135). DOI: <https://doi.org/10.1186/1471-2288-14-135>.
8. Eng H-L., Ma K-K. Noise Adaptive Soft-Switching Median Filter. *IEEE transactions on image processing*. 2001. Vol. 10(2). Pp. 242-251. DOI: https://doi.org/10.1007/978-3-642-22720-2_19.
9. Villar S.A., Torcida S., Acosta G.G. Median Filtering: A New Insight. *Journal of Mathematical Imaging and Vision*. 2017. Vol. 58. No 1. Pp. 130-146. DOI: <https://doi.org/10.1007/s10851-016-0694-0>
10. Rajan P. K. Two-dimensional digital signal processing II: Transforms and median filters. *Proceedings of the IEEE*. 1982. Vol. 70(7). Pp. 780-781. DOI: <https://doi.org/10.1109/proc.1982.12396>
11. Zhu B., Jiao J., Steinhardt J. When does the Tukey Median work? URL: <https://arxiv.org/abs/2001.07805> (accessed at 20.12.2020)
12. Gabbou M., Coyle E.J., Gallagher N.C., Jr. An overview of median and stack filtering. *Circuits Systems Signal process*. 1992. Vol. 11(1). Pp. 7-45. DOI: <https://doi.org/10.1007/BF01189220>.
13. Taherdoos H. Sampling Methods in Research Methodology. How to Choose a Sampling Technique for Research. *International Journal of Academic Research in Management (IJARM)*. 2016. Vol 5(2). Pp. 18-27. DOI: <https://doi.org/10.2139/ssrn.3205035>.
14. Estivill-Castro V., Wood D. A survey of adaptive sorting algorithms. *ACM Computing Surveys*. 1992. Vol 24. Issue 4. Pp. 441-476. DOI: <https://doi.org/10.1145/146370.146381>.
15. Mazmanian R.O. Characteristics of ordered samples of a random uncorrelated signal. *Tekhnichna Elektrodynamika*. 2004. No 6. Pp. 60-64. (Rus).
16. Mazmanian R.O. Ordered samples in uncorrelated signal conversion. *Tekhnichna Elektrodynamika*. 2018. No 3. Pp. 82-89. DOI: <https://doi.org/10.15407/techned2018.03.082> (Rus).
17. Feller W. An introduction to probability theory and its applications. Vol. 1. New York: Wiley, 1968. 528 p.
18. Ozer D.J. Correlation and the coefficient of determination. *Psychological Bulletin*. 1985. Vol 97(2). Pp. 307-315. DOI: <https://doi.org/10.1037/0033-2909.97.2.307>.
19. Keya Rani Das, A.H.M. Rahmatullah Imon. A Brief Review of Tests for Normality. *American Journal of Theoretical and Applied Statistics*. 2016. Vol 5(1). Pp. 5-12. DOI: <https://doi.org/10.11648/j.ajtas.20160501.12>.
20. Jones T. A. Skewness and kurtosis as criteria of normality in observed frequency distributions. *Journal of Sedimentary Research*. 1969. Vol 39(4). Pp. 1622-1627. DOI: <https://doi.org/10.1306/74d71ec9-2b21-11d7-8648000102c1865d>.
21. DeCarlo L. T. On the meaning and use of kurtosis. *Psychological Methods*. 1997. Vol 2(3). Pp. 292-307. DOI: <https://doi.org/10.1037/1082-989X.2.3.292>.
22. Pustynnik E.I. Statistical methods of analysis and processing of observations. Moskva: Nauka, 1968. 288 p. (Rus).
23. Montgomery D.C., Runger G.C. Applied statistics and probability for engineers. Wiley PLUS Products., 2017. 479 p.
24. Massey A., Miller S.J. Tests of Hypotheses Using Statistics. Providence: Brown University, 2006. 32 p.
25. Greenwood P., Nikulin M.S. A Guide to Chi-squared Testing. New York: Wiley, 1996. DOI: <https://doi.org/10.1007/s10958-014-1835-x>.
26. Guenther W.C. Power and Sample Size for Approximate Chi-Square Tests. *The American Statistician*. 1977. Vol. 31(2). Pp. 83-85. DOI: <https://doi.org/10.2307/2683047>.
27. Nisen J.A., Schwertman N.C. A simple method of computing the sample size for Chi-square test for the equality of multinomial distributions. *Computational Statistics and Data Analysis*. 2008. Vol. 52. Pp. 4903-4908. DOI: <https://doi.org/10.1016/j.csda.2008.04.007>.

NONCORRELATED DATA ORDERED SAMPLES AS A SINGLE-ELEMENT MULTICHANNEL CONVERTER.

R.O. Mazmanian

Institute of Electrodynamics National Academy of Sciences of Ukraine,

56, Peremohy ave, Kyiv, 03057, Ukraine,

e-mail: mazmanian@ied.org.ua

The sufficient closeness of the medians of the ordered samples of random data to the normal distribution is used in computer systems for control, monitoring and diagnosing electric power equipment. However, it remains what other probability density function (pdf) of elements (sample statistics) have such similarities. This paper presents various methods for statistical testing hypotheses for pdf-converter channels as statistics of given sizes odd numbered and ordered samples of bounded and uniformly distributed random numbers. The use of various different criteria and the results of estimates studied under the same conditions showed a sufficient conformity of the results of tests for three statistical criteria. It made possible to draw a reasonable conclusion about the preferable use of the adapted chi-square test for assessing the congruence of analytical pdf channels of the converter with normal distribution. We also suggested using the "statistical closeness window" to define those channels of the converter that do not significantly differ from the normal distribution. In addition, we presented an empirical formula determining the dependence of the size of the window of the statistical closeness window on the sample size. The results of the research are summarized in a statistical model of a multichannel uncorrelated data converter. References 27, figures 7.

Keywords: random data, uniformly distribution, ordered samples, probability density function (pdf), pdf-converter, density estimation, chi-square criteria.

Надійшла 25.01.2021
Остаточний варіант 15.03.2021