

**БАГАТОКАНАЛЬНЕ КОНВЕРТУВАННЯ ВИПАДКОВИХ ДАНИХ
ПАРНИМИ ЕЛЕМЕНТАМИ ВПОРЯДКОВАНИХ ВИБІРОК**

Р.О. Мазманян*, докт. техн. наук
Інститут електродинаміки НАН України,
пр. Перемоги, 56, Київ, 03057, Україна,
e-mail: mazmanian@ied.org.ua

Концепцію багатоканальної паралельної конвертації функції розподілу щільності ймовірності, pdf, випадкових даних раніше було застосовано в дослідженні властивостей одноелементних pdf-конверторів. У розвиток концепції, в статті досліджуються перетворювальні властивості каналів spdf-конверторів, утворених сумою пар елементів упорядкованих вибірок (впорядкованих статистик). Характеристики результату конверсії, як залежності від розміру вибірок і зміщення каналів відносно її медіани, було отримано для даних з рівномірним розподілом щільності. З аналізу виключено області взаємної залежності функцій щільності елементів, що додаються, котрі далі, разом з апроксимуючими їх функціями, було нормалізовано. Незважаючи на очевидні структурні відмінності, завданням даного дослідження, як і раніше, було визначення згоди перетворених даних з деякими стандартними функціями щільності розподілу ймовірностей, зокрема, з нормальним законом розподілу. Як і раніше, оцінки близькості каналів spdf-конвертора отримано за допомогою критерію χ^2 -квадрат. Результати досліджень використано задля визначення розміру і розташування вікон статистичної близькості, побудови статистичних моделей каналів конвертора. Бібл. 20, рис. 14.

Ключові слова: випадкові дані, рівномірний розподіл, впорядковані вибірки, функція щільності ймовірності, probability density function (pdf), сума парних упорядкованих статистик, spdf-конвертор, оцінка щільності ймовірності.

Вступ. Сучасні комп'ютерно-інтегровані технології в електроенергетиці є розподіленими системами збору і передачі технологічної інформації по каналах зв'язку від локальних систем збору і попередньої обробки даних щодо параметрів об'єктів, що входять до складу енергосистеми, обробки цієї інформації задля автоматичного управління, контролю якості електричної енергії, аварійного захисту і діагностування електроенергетичного обладнання. Рішення цих завдань пов'язане з дослідженнями процесів генерації, перетворення і розподілу електричної енергії, зі створенням інформаційного, математичного та програмного забезпечення для об'єктів і комплексів електроенергетики. Оцінка якості електричної енергії в мережі і стану обладнання за наявності численних збурюючих впливів, що мають випадковий або детермінований характер, забезпечується за допомогою різноманітних поєднань спеціальних апаратних і алгоритмічних засобів. Робастні оцінки середніх значень контрольованих параметрів за допомогою медіанних фільтрів забезпечують підвищення достовірності інформації [1].

Разом з тим, ранжирування впорядкованими вибірками числових даних, наприклад, відхиленнями від робастної середньої напруги в мережі і тривалості зареєстрованих відхилень, несуть важливу інформацію про зміни параметра в умовах дії випадкових або циклічних збурень. Їхній аналіз дозволить виявити приховані закономірності в причинах і проявах порушень нормального режиму роботи окремих об'єктів або енергетичного комплексу в цілому.

Отримання оцінок контрольованого середовища за допомогою упорядкованих вибірок даних може бути ефективно використано в мультиагентних системах для енергетики [2, 3], що відрізняються здатністю реагувати на зміну середовища, цілеспрямованою поведінкою, можливістю комунікації з іншими інтелектуальними агентами. Програмно реалізовані агенти систем здійснюють функції моніторингу та діагностики, систем захисту, розподілених систем управління, а також моделювання і симуляції процесів, що дає змогу обмінюватися інформацією про конкретне обладнання [4].

Дослідження упорядкованих вибірок з множин випадкових дійсних чисел охоплюють широке

коло проблем, пов'язаних з вивченням їхніх фундаментальних властивостей і різноманітним застосуванням в обробці інформації [5, 6, 7]. Помітне місце в цих роботах займають перетворення абсолютно неперервних випадкових величин [8], які, крім практичної спрямованості, надають можливість вивчення та формалізації властивостей цього нелінійного методу обробки даних.

Неперервність випадкових величин і функцій їхньої щільності ймовірності разом з дискретними параметрами методу, які є, по суті, дискретними аргументами безперервних функцій, забезпечили вичерпне аналітичне уявлення досліджуваних властивостей впорядкованих вибірок.

У цьому сенсі, розгляд упорядкованих вибірок як систем з одним входом і багатьма виходами, або, що те ж саме, як багатоканальних *pdf*-конверторів [9], надає можливість створення цілісного опису процесів перетворення, загальних методів їхнього вивчення і формування порівнянних результатів для різних процедур і об'єктів досліджень, пов'язаних з ними.

Раніше з аналітичних функцій щільності ймовірності, математичного сподівання і дисперсії [10] було визначено умовну параметричну близькість до функції Гауса і статистична близькість до нормального закону щільності ймовірності деяких з $k = 0 \dots 2d$ елементів упорядкованих вибірок з порядком d і розміром $s = 2d + 1$ з множини випадкових, обмежених значенням X_{\max} величин [9]. Було показано, що окремі поодинокі елементи багатоканального конвертора можуть входити в групи вікон параметричної та/або статистичної близькості, розміри яких не збігаються з їхніми граничними номерами.

У цій роботі розглядаються властивості сум пар елементів, які однаково зміщені у виборах з порядком d і розміром s в протилежних щодо медіани напрямках на величину $m = 0 \dots d$.

Мета роботи полягає в дослідженні перетворювальних властивостей багатоканальних конверторів, утворених сумами парних упорядкованих статистик вибірок з множини некорельованих даних з рівномірним розподілом, створенні статистичної моделі конвертації.

Методика досліджень. 1. Спочатку *аналітичними методами досліджень* визначається результуюча функція щільності ймовірності суми *spdf* (*sum probability density function*) двох випадкових величин – довільних симетричних відносно медіани пар елементів вибірок із застосуванням інтеграла згортки [11]. Далі формулюються нульова і альтернативна її гіпотези про статистичну згоду випадкових величин – сумарних виходів пар елементів конвертора – з нормальним або іншими законами розподілу неперервних випадкових величин. Обґрунтовуються і формалізуються способи нормалізації *spdf*, тобто їхнє приведення до масштабів *pdf* елементів, що додаються, з виключенням з аналізу залежних значень в цих випадкових величинах. Нормалізовані *spdf* будуть представлені аналітичними залежностями від нормалізованих сум пар випадкових величин.

2. Зіставлення *spdf* з нормальним законом розподілу починається з подання порівнюваних функцій, як звичайної процедури наближення, апроксимації, визначення умов заміни однієї детермінованою аналітичної функції однієї або декількома іншими функціями [12, 13]. Апроксимуюча функція Гауса (*pdfG*) будується по середньоквадратичному відхиленню (СКВ) сумарної випадкової величини і далі нормалізується. Для зіставлення використовується похибка апроксимації у вигляді інтегрального показника відхилень між функціями.

3. Аналіз параметричної, в детермінованому сенсі, похибки апроксимації, визначення в новій сукупності з числом елементів, рівним порядку початкових упорядкованих вибірок d , елементів з особливими властивостями, встановлення їхнього зв'язку з величиною зміщення.

4. Отримання оцінок видимої близькості до розподілу Лапласа (*pdfL*) щільності суми елементів *spdf* з максимальним зсувом d для вибірок будь-якого порядку. І в цьому випадку виконується нормалізація розглянутих каналів парних елементів, формулюються власні основна гіпотеза $H_0^{(L)}$ і альтернативна їй гіпотеза $H_1^{(L)}$.

5. Перевірки згоди сумарних випадкових величин до нормального закону і розподілу Лапласа *чисельними методами досліджень* виконуються окремо. Для цього формуються дві пари векторів дискретних значень *pdfG* і *pdfL* з *spdf* [10], обчислюються статистики χ^2 -квадрат [14], визначаються критичні значення кожного каналу парних елементів, за якими приймаються або відхиляються сформульовані гіпотези.

6. Побудова вікон статистичної близькості і статистичних моделей багатоканальних конверторів парних елементів довільного порядку d з отриманих оцінок згоди для всіх каналів $0 \dots d$ *spdf*-конвертора.

Функції щільності ймовірності суми парних елементів упорядкованих вибірок. Розглянемо зв'язок параметрів суми двох елементів вибірок *spdf* і параметрів *pdf* доданків цієї суми

(рис. 1). Абсолютні номери розглянутих пар елементів для позитивного і негативного зсуву m щодо медіани рівні $k_m^{(p)} = d + m$, $k_m^{(n)} = d - m$, відповідно.

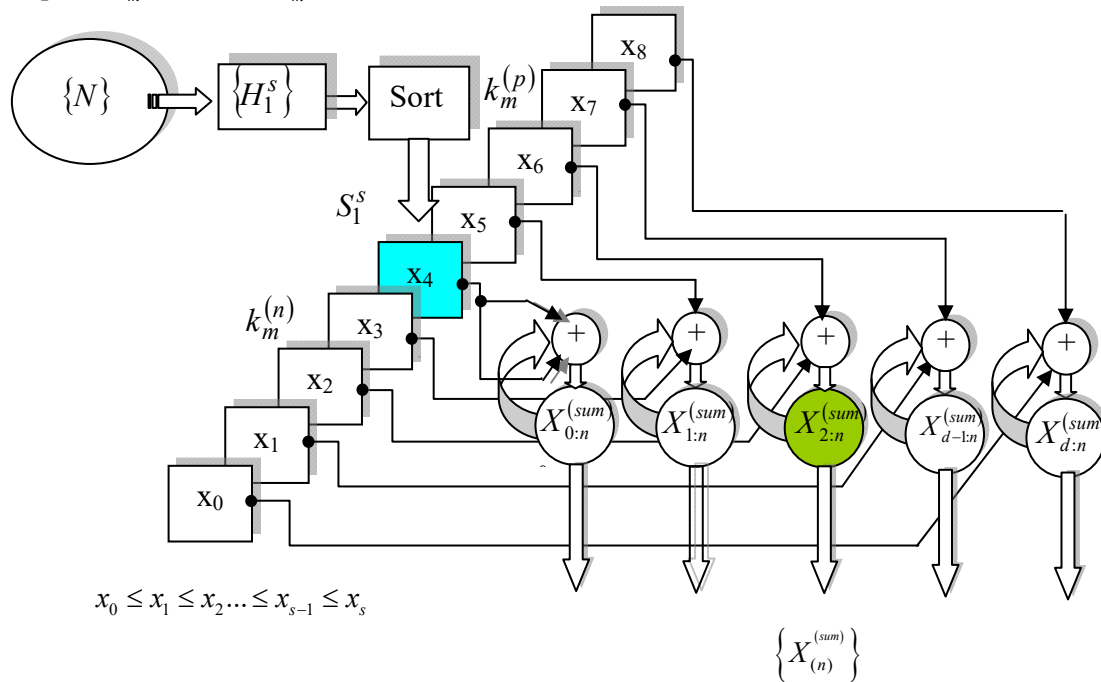


Рис. 1

Розділимо випадкові величини елементів, що додаються, зі зміною позначення на y елемента з негативним зміщенням. Функція щільності ймовірності визначається згортокою щільності ймовірності доданків елементів вибірки [11], тобто

$$p_{d,m}^{(con)}(z) = p_S(x^{(p)}, d, k_m^{(p)}) * p_S(x^{(n)}, d, k_m^{(n)}) = \int_a^b p_S(x, d, k_m^{(p)}) \cdot p_S(z - x, d, k_m^{(n)}) dx, \quad (1)$$

де щільність ймовірності елемента визначається формулою [10]

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

Очевидно, що межі змін суми двох випадкових величин подвоюються $z = x + y$, тобто $-2X_{\max} \leq z \leq 2X_{\max}$. Підстановка меж інтеграла згортки і щільності ймовірності елементів перетворює вираз (1) до вигляду

$$p_{d,m}^{(con)}(z) = S(d, m) \int_a^b (X_{\max} + x)^{k_m^{(p)}} \cdot (X_{\max} - x)^{2d-k_m^{(p)}} \cdot [X_{\max} + (z-x)]^{k_m^{(n)}} \cdot [X_{\max} - (z-x)]^{2d-k_m^{(n)}} dx, \quad (3)$$

де $a = -X_{\max}$, $b = z + X_{\max}$ для $z < 0$; $a = z - X_{\max}$, $b = X_{\max}$ для $z \geq 0$, (4)

$$S_{d,m} = \frac{1}{(2X_{\max})^{2(d+1)}} \cdot \frac{[(2d+1)!]^2}{(2d-k_m^{(n)})!(2d-k_m^{(p)})!k_m^{(n)}!k_m^{(p)}!} \quad (5)$$

– коефіцієнт, який визначається дискретними значеннями порядку вибірки d і зміщенням елементів m , що додаються.

З подвоєнням діапазону змін випадкової величини z зменшуються значення функції розподілу суми елементів $p_{d,m}^{(con)}(z)$, що підтримує рівність одиниці граничного значення функції розподілу випадкової величини z .

Для вибірок розміром $d = 9$ на рис. 2 показана функція щільності ймовірності (2) елемента $p_{d,d}(x)$ з номером $k = d$, тобто

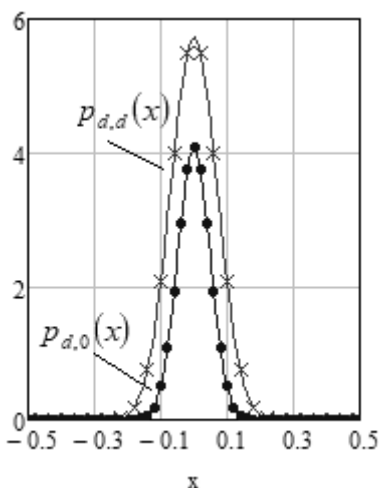


Рис. 2

медіани, і $spdf\ p_{d,0}^{(con)}(z)$ пари зі зміщенням $m=0$, яка підсумовується. Це особливий випадок, коли підсумовуються два доданки, два елементи, функції щільності яких рівні між собою. На цьому і наступних рисунках обидві функції будуть представлені в шкалі випадкової величини x , тобто для $-X_{max} \leq z \leq X_{max}$.

Про перевірку гіпотез розподілу з відомими параметрами. Зазвичай нульова H_0 і альтернативна H_1 статистичні гіпотези формуються як припущення, призначені для перевірки відповідності невідомих параметрів розподілу вибірових даних якого-небудь передбачуваного закону розподілу. Як правила прийняття або відхилення статистичних гіпотез для оцінки близькості двох різних аналітично заданих розподілів із загальними параметрами, як і раніше [9], використовується критерій χ^2 -квадрат.

Нормалізація функції щільності розподілу суми пар елементів. Необхідність співставлення функцій щільності суми і окремих елементів вимагає нормалізації [15] результатів підсумовування, тобто їхнього зведення до початкового масштабу, до загальних шкал для випадкових величин і досліджуваних pdf і $spdf$.

Нормалізацію за двома параметрами необхідно виконувати із застосуванням двох множників, які визначаються окремо для інтеграла згортки і для її незалежної змінної z . Коректуючі множники надають одночасний нормуючий вплив на результат, але розділені по точці прикладання. Збільшення вдвічі діапазону змін суми компенсується множником $\xi_{d,m}$ для суми випадкових значень елементів вибірки z . Зменшення значень щільності ймовірності суми компенсується множником $\psi_{d,m}$, що застосовується до значень функції щільності ймовірності.

Коректне застосування інтегралу згортки вимагає виключення залежних значень [16] в елементах зі зміщенням m , що додаються. Їхнє число залежить від порядку вибірки і зміщення. Тому статистична корекція результатів узгоджується з метою нормалізації і форматом множників, що нормалізують.

Включення фактора статистичної корекції в множники $\xi_{d,m}$ та $\psi_{d,m}$, що нормалізують, засноване на аналізі ймовірності виникнення залежних/незалежних пар подій в особливих випадках і врахування загального коефіцієнта нормалізації λ . Він дорівнює добутку нормалізуючих множників і має постійну величину, яка визначається відносним числом подій суми для будь-яких значень порядку вибірки і зміщення пари, тобто

$$\lambda = \xi_{d,m} \cdot \psi_{d,m} = 2. \quad (6)$$

Перший особливий випадок відноситься до статистик впорядкованих вибірок зі зміщенням $m=0$, що додаються, і однаковими їхніми номерами $k=k_0^{(p)}=k_0^{(n)}=d$ (рис. 3, а). Вони повністю взаємно залежні. Число незалежних значень цієї пари дорівнює половині від їхнього загального числа. Коректуючий вплив, що дорівнює 2, забезпечується за його рівного розподілу між двома множниками, тобто за $\xi_{d,0}=\sqrt{2}$ і $\psi_{d,0}=\sqrt{2}$.

Другий випадок пов'язаний зі зміщенням пари елементів, рівним порядку вибірки d , тобто

$$k_p = k_d^p = 2d, \quad k_n = k_d^n = 0 \quad (\text{рис. 3, б}).$$

Тут число незалежних подій $n_d^{(ind)}$ дорівнюватиме сумі подій двох елементів, тобто двом.

Корекція по незалежній змінній означає перехід від z до x зі зменшенням меж в 2 рази. Це очевидний факт, пов'язаний з самою величиною, що коректується, тому що в цій парі елементів знаходяться взаємно незалежні випадкові величини. Тут вони залишаться і після виконання цієї складової корекції для приведення значень сумарної випадкової величини

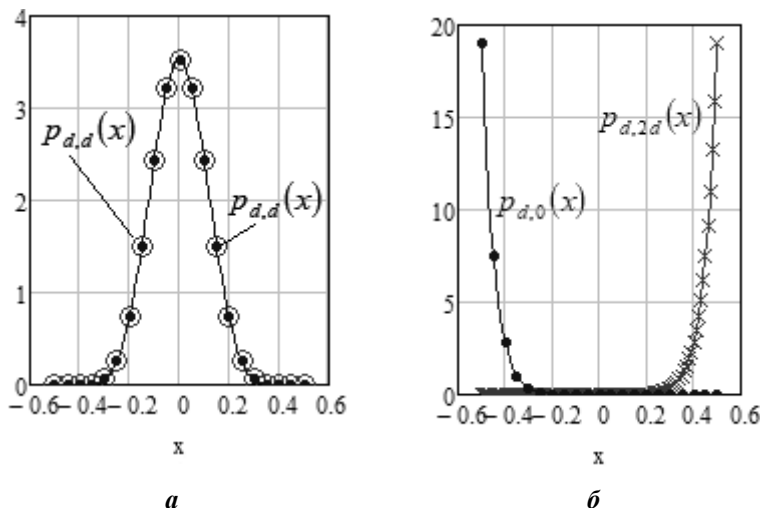


Рис. 3

до меж змін однієї випадкової величини.

Коректувальний множник $\sqrt{2} \leq \xi_{d,m} \leq 2$ для номерів пар елементів $0 \leq m \leq d$ використовується для зменшення діапазону, тобто в знаменнику випадкової величини z .

Коректувальний множник $\psi_{d,m}$ використовується для збільшення значень функцій щільності, тобто в чисельнику коефіцієнта $S(d, m)$ (5) інтеграла згортки (3). Значення множника визначаються з рівності (6)

$$\psi_{d,m} = \lambda / \xi_{d,m}. \quad (7)$$

Таким чином, у разі зміни зсуву пари доданків елементів від 0 до d число незалежних подій $n^{(ind)}$ буде змінюватися в діапазоні значень від 1 до 2, тобто від їхньої повної взаємної залежності до їхньої повної незалежності.

Лінійні наближення змін коректувальних множників в залежності від порядку вибірки визначаються наступними виразами:

$$\xi_{d,m}^{lin} = \xi_{d,0} + \Delta\xi(d) \cdot m, \quad \psi_{d,m}^{lin} = \psi_{d,0} + \Delta\psi(d) \cdot m, \quad (8)$$

де $\Delta\xi(d) = (\xi_{d,d} - \xi_{d,0})/d$, $\Delta\psi(d) = (\psi_{d,d} - \psi_{d,0})/d$ – кутові коефіцієнти за напрямками корекції.

На рис. 4, а представлено графіки зміни коректувальних множників для $d = 9$.

Чисельна оцінка залежності/незалежності випадкових величин будь-якої пари елементів упорядкованих вибірок визначається часткою площі взаємного перекриття або вільної від перекриття в загальній площі їхніх функцій щільності розподілу. Для симетричних відносно медіани пар елементів зі зміщенням $m = 3$, наприклад, (рис. 4, б), ця площа визначає частку залежних/незалежних

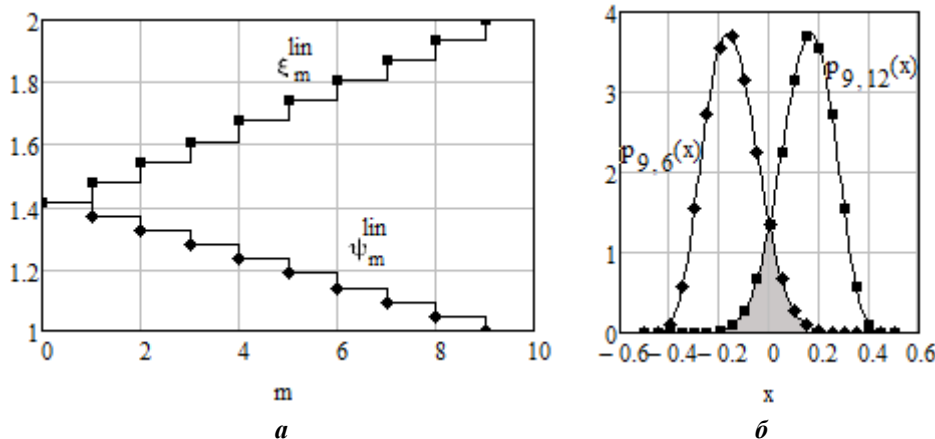


Рис. 4

подій в результаті складання. Її слід враховувати під час коректування результату.

Очевидно, що перекриття симетричних пар для правого елемента $k_{d,m}^{(p)} = d + m$ утворено площею, обмеженою негативною піввіссю абсцис і його функцією розподілу щільності ймовірності в цій області, а для лівого елемента $k_{d,m}^{(n)} = d - m$ –

позитивною піввіссю і функцією розподілу лівого елемента. Причому, в силу симетрії площі функції щільності ймовірності елементів щодо осей ординат рівні між собою. Тому сумарне число незалежних подій пари елементів визначається площею, обмеженою межами $-X_{\max} \dots 0$ і $0 \dots X_{\max}$ для лівого і правого елемента, відповідно (рис. 4, б), або подвоєному значенню числа одного з них.

Частка (ймовірність) незалежних подій для лівого елемента визначається значенням функції розподілу $P(x)$ в точці $x=0$, а для правого $1 - P(x)$. Загальна частка (ймовірність) незалежних подій для суми дорівнює подвоєному значенню ймовірності одного з елементів.

Ймовірність появи $p_{d,m}^{(ind)}(k_{d,m}^{(p)})$ незалежних подій правого від медіани елемента визначається площею її функції щільності ймовірності в діапазоні від нуля до X_{\max} (рис. 4, б)

$$p_{d,k}^{(ind)}(x) = \frac{1}{2X_{\max}} \cdot \frac{(2d+1)!}{k!(2d-k)!} \int_0^{X_{\max}} (X_{\max} + x)^{k^{(p)}} \cdot (X_{\max} - x)^{2d-k^{(p)}}. \quad (9)$$

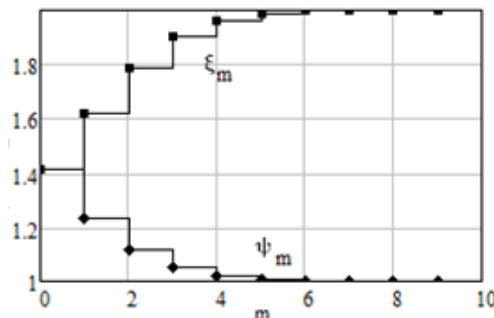
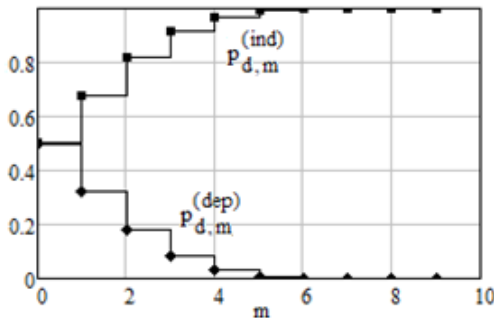
Залежні події в цьому випадку будуть мати ймовірність, яка визначається рівністю $p_{d,m}^{(dep)}(k_{d,m}^{(p)}) = 1 - p_{d,m}^{(ind)}(k_{d,m}^{(p)})$.

Зміну ймовірностей подій в залежності від зміщення показано на рис. 5, а.

Нелінійну корекцію функції щільності ймовірності (10) отримано шляхом зміщень і

масштабування подвоєної ймовірності появи незалежних подій $p_{d,m}^{(ind)}(k_{d,m}^{(p)})$

$$\xi_{d,m} = 2 \cdot (p_{ind}(k_m^{(p)}) - 0.5) \cdot (2 - \sqrt{2}) + \sqrt{2}. \quad (10)$$



Коректувальний множник по незалежній змінній визначається з рівності (7). Множники за двома напрямками графічно показано на рис. 5, б.

а

б

Рис. 5

Корекція сумарної випадкової величини $u = z \cdot (\xi_{d,m})^{-1}$, введення коректувального *spdf* множника $\psi_{d,m}$ з (7), заміна позначення в підінтегральному виразі x на y та перетворює (3) до нормалізованого вигляду

$$p_{d,m}^{(norm)}(u) = \psi_{d,m} \cdot p(u, d, m) = \psi_{d,m} \cdot S_{d,m} \times \int_a^b (X_{\max} + x)^{k_m^{(p)}} \cdot (X_{\max} - x)^{2d - k_m^{(p)}} \cdot [X_{\max} + (z - x)]^{k_m^{(n)}} \cdot [X_{\max} - (z - x)]^{2d - k_m^{(n)}} dx, \quad (11)$$

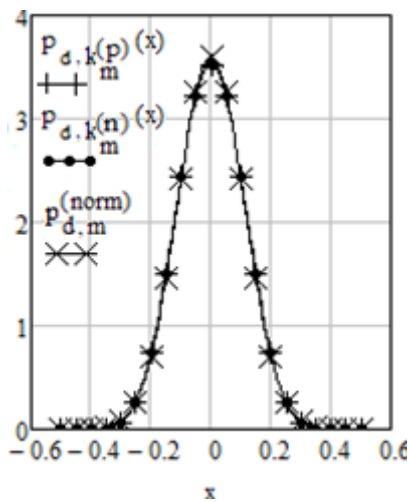


Рис. 6

де межі інтегрування визначаються з (4).

Повернення до початкової випадкової величини дає змогу завершити нормалізацію формулою

$$p_{d,m}^{(norm)}(x) = p_{d,m}^{(norm)}(u). \quad (12)$$

Функції щільності ймовірності двох одноелементних медіан разом з нормалізованою *spdf* пари елементів з нульовим зміщенням, тобто $k_{d,0}^{(p)} = k_{d,0}^{(n)} = d$, показано на рис. 6.

Нижче наведено факторіальний вираз для формули нормалізованої щільності ймовірності суми двох симетрично розташованих елементів упорядкованих вибірок, отриманий за допомогою біноміального розкладання співмножників підінтегрального виразу в рівнянні (11)

$$p_{d,m}^{(con)}(x) = \psi_{d,m} \cdot S_{d,m} \times \sum_{jp=0}^{k_m^{(p)}} \sum_{ip=0}^{2d - k_m^{(p)}} \sum_{jn=0}^{k_m^{(n)}} \sum_{in=0}^{2d - k_m^{(n)}} \frac{(-1)^{in+ip} \cdot X_{\max}^{4d - jn - jp - in - ip} \cdot (2d - k_m^{(n)})! \cdot (2d - k_m^{(p)})! \cdot k_m^{(n)}! \cdot k_m^{(p)}!}{jn! \cdot jp! \cdot in! \cdot ip! \cdot (k_m^{(n)} - jn)! \cdot (k_m^{(p)} - jp)! \cdot (2d - k_m^{(n)} - in)! \cdot (2d - k_m^{(p)} - ip)!} \times \sum_{i=0}^{jn-in} \frac{(-1)^i \cdot (jn + in)! \cdot \left(\frac{2x}{\xi_{d,m}}\right)^{jn-in-i}}{i! \cdot (jn + in - i)! \cdot (jp + ip + i + 1)!} \times \begin{cases} \left(\frac{2x}{\xi_{d,m}} + X_{\max}\right)^{jp+ip+i+1} - (-X_{\max})^{jp+ip+i+1} & x < 0 \\ X_{\max}^{jp+ip+i+1} - \left(\frac{2x}{\xi_{d,m}} - X_{\max}\right)^{jp+ip+i+1} & x \geq 0 \end{cases} \quad (13)$$

Апроксимація функцій і оцінка параметричної близькості. Апроксимація функцій може розглядатися як заміна досліджуваної функції іншої, більш простою або більш зручною для використання. Такий детерміністський підхід потребує застосування чисельних оцінок і критичних значень відмінностей, які дозволять судити про *параметричну* близькість або віддаленість заміни в

рамках прийнятих двох гіпотез. Якщо для оцінок якості наближення можуть використовуватися різні показники точності/похибки апроксимації, то вибір їхніх критичних значень не має обґрунтованих правил прийняття і відхилення гіпотез, що перевіряються.

Тому параметрична перевірка буде носити умовний характер, пов'язаний з порогом прийняття рішень. Така процедура може бути корисною, наприклад, для зіставлення параметричних висновків з результатами статистичної перевірки.

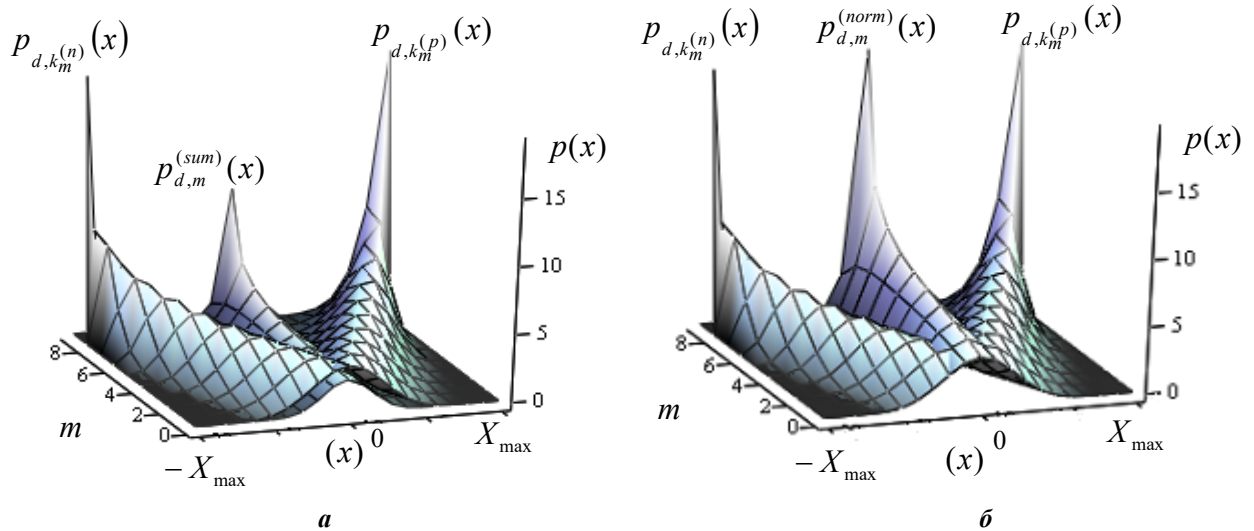


Рис. 7

Загальний вигляд функцій щільності ймовірності лівосторонніх і правобічних відносно медіани каналів *pdf*-конвертора і каналів парних сум *spdf* для вибірок з порядком $d = 9$ наведено на рис. 7, а.

Нормалізовані *spdf* разом з каналами елементів, що додаються, показані на рис. 7, б.

Побудова апроксимуючої функції. Апроксимуюча функція Гауса в першому наближенні будується за двома параметрами досліджуваної випадкової величини, тобто впорядкованої статистики розглянутого каналу, – її дисперсії і середнього значення.

Середні симетричної відносно медіани пари елементів впорядкованої вибірки [10] мають рівні, але протилежні за величиною значення

$$M_{d,k} = \frac{X_m}{d+1} \cdot (k+1) - X_m. \quad (14)$$

Тому сума довільної симетричної пари має нульовий зсув.

Дисперсія елементів упорядкованих вибірок визначається виразом [10]

$$D_{d,k} = \alpha_{d,k}^{(2)} - \mu_{d,k}^2 = \left(X_{\max}^2 \cdot \frac{(2 \cdot d + 1)!}{2^{2 \cdot d + 1}} \cdot \sum_{j=0}^k \sum_{m=0}^{2 \cdot d - k} \frac{(-1)^m + (-1)^j}{j! \cdot m! \cdot (k - j)! \cdot (2 \cdot d - k - m)! \cdot (m + j + 3)} \right) - \left(X_{\max}^2 \cdot \frac{(2 \cdot d + 1)!}{2^{2 \cdot d + 1}} \cdot \sum_{j=0}^k \sum_{m=0}^{2 \cdot d - k} \frac{(-1)^m - (-1)^j}{j! \cdot m! \cdot (k - j)! \cdot (2 \cdot d - k - m)! \cdot (m + j + 2)} \right)^2. \quad (15)$$

Тут дисперсія суми обчислюється з початкової функції щільності ймовірності суми елементів (3) без коректувального множника, тому що під час підсумовування вдвічі збільшується діапазон змін випадкової величини з одночасним зменшенням вдвічі значень функції щільності суми

$$D_{d,m}^{(sum)} = S_{d,m} \int_{-2X}^{2X} z^2 \int_a^b (X_{\max} + x)^{k_m^{(p)}} \cdot (X_{\max} - x)^{2 \cdot d - k_m^{(p)}} \cdot [X_{\max} + (z - x)]^{k_m^{(n)}} \cdot [X_{\max} - (z - x)]^{2 \cdot d - k_m^{(n)}} dx. \quad (16)$$

Межі інтегрування $[a, b]$ визначаються з (4). Дисперсія суми двох елементів впорядкованої вибірки дорівнює сумі дисперсій доданків елементів [17]. На рис. 8 показано дисперсії каналів одноелементного конвертора $D_{9,k}$ і конвертора парних елементів $D_{9,m}^{(sum)}$ для номерів елементів і їхніх зсувів, відповідно.

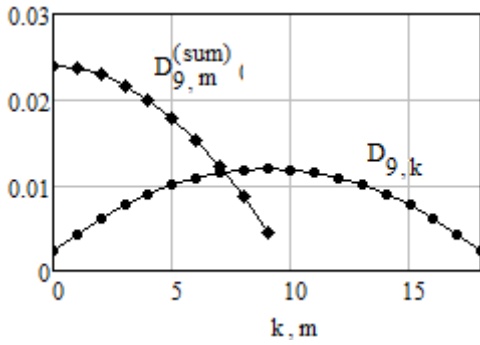


Рис. 8

Вибіркові середньоквадратичні відхилення елементів пов'язані з дисперсіями залежністю

$$\sigma_{d,m} = \sqrt{D_{d,m}}. \quad (17)$$

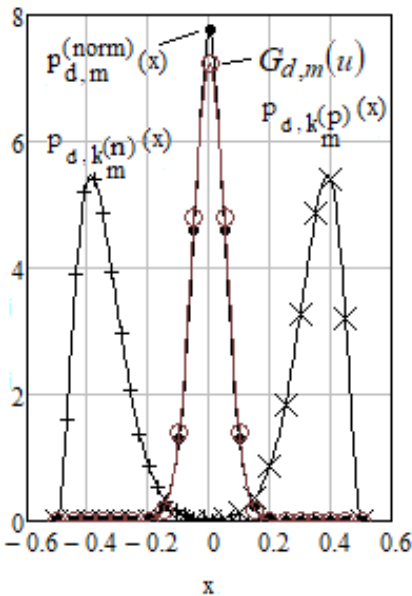
Нормалізація функції Гауса. Апроксимуюча функція Гауса визначається середньоквадратичним відхиленням пар елементів, що додаються, з введенням коректувальних множників в інтервалі змін випадкової величини.

Корекція функції Гауса здійснюється множником, як і за нормалізації *pdf*, яка апроксимується. А корекція аргументу цієї функції здійснюється його подвоєнням для підгонки під подвійний інтервал, а потім в знаменник вводиться корекція подвійного інтервалу

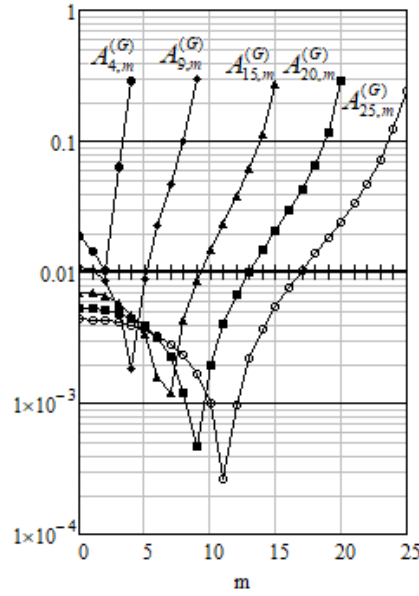
$$G_{d,m}(u) = \frac{\Psi_{d,m}}{\sigma_{d,m}^{(s)} \cdot \sqrt{2\pi}} \cdot \exp\left[-0,5(u)^2 (\sigma_{d,m}^{(s)})^{-2}\right], \quad (18)$$

де $u(x) = \frac{2}{\xi_{d,m}} \cdot ([-X_{\max} + x])$.

Процедура забезпечує рівність одиниці площі апроксимуючої функції для всіх припустимих значень його дискретних аргументів.



а



б

Рис. 9

Слід зазначити, що для суми елементів зі зміщеннями, близькими до порядку вибірки, спостерігається деяке заниження максимуму *pdf* за збереження близькості двох функцій на всьому діапазоні змін нормованої по аргументу суми випадкових величин елементів. Це показано на рис. 9, а для елементів, що додаються, зі зміщенням $m = 7$ в вибірках з порядком $d = 8$.

Наведені до діапазону змін інтегральні значення відхилень *pdf*, що апроксимується, від функції Гауса (рис. 9, б), визначаються

залежностями (11), (12) і (18)

$$A_{d,m}^{(G)} = \frac{1}{2X_{\max}} \cdot \int_{-X}^X |p_{d,m}^{(norm)}(x) - G_{d,m}(x)| dx. \quad (19)$$

Графічне подання похибок апроксимації елементів упорядкованих вибірок, що додаються, як виходів багатоканального конвертора, дозволяє судити про цілком певну їхню впорядкованість, схожу з впорядкованістю одноелементних конверторів [10]. Йдеться про наявність виходу, що займає середнє положення в ряду виходів *pdf*-конвертора, і має мінімальні відмінності від апроксимуючої функції Гауса. В цьому випадку можна стверджувати про відтворення в дещо зміненому форматі структури початкової впорядкованої вибірки з випадкових величин як одноелементного *pdf*-конвертора.

Номер цього елемента в ряду виходів конвертора парних елементів графічно визначається з рівності дисперсії суми двох елементів впорядкованої вибірки суми дисперсій елементів, що додаються [17]. Для випадку парного порядку вихідних вибірок (рис. 10, а) номер медіанного каналу *pdf*-конвертора (рис. 1) дорівнює $d / 2$

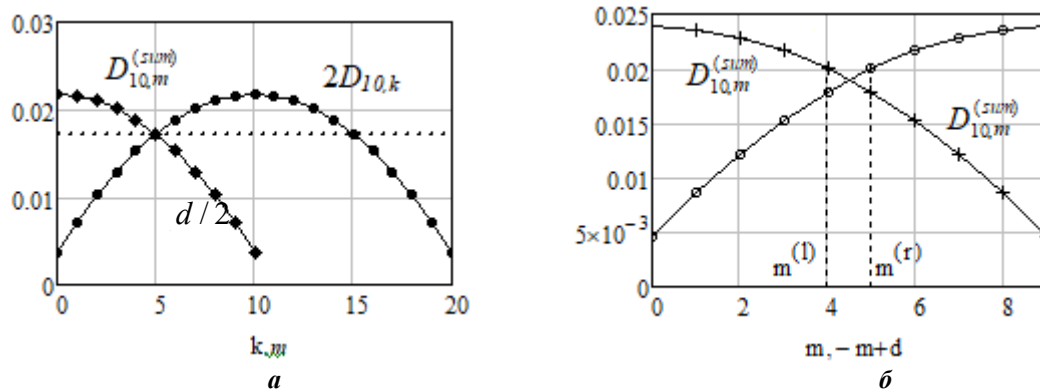


Рис. 10

Непарний порядок вихідних вибірок (рис. 10, б) формує конвертори другого типу, які відрізняються наявністю двох серединних каналів з номерами, рівними меншій $m^{(l)}$ і більшій $m^{(r)}$ цілої частині зсуву, рівного $d/2$.

Асиметрія дисперсії щодо медіанного каналу або двох серединних каналів призводить і до асиметрії інтегральних відхилень $A^{(G)}$ від функції Гауса $spdf$, що апроксимується, (рис. 9, б)

На другому етапі нормалізації апроксимуючої функції Гауса визначаються масштабуючі коефіцієнти, що доповнюють її

$$\rho_{d,m}^{(G)} = p_{d,m}^{(\max)} (x) (G_{d,m}(0))^{-1}, \quad (20)$$

де максимальні значення функцій щільності ймовірності (11), (12) рівні

$$p_{d,m}^{(\max)} = \frac{\psi_{d,m}(2d+1)!}{(2X_{\max}^{d+1} \cdot (d+m)! \cdot (d-m)!)^2} \cdot \int_{-X_{\max}}^{X_{\max}} (X_{\max} + x)^{d+m} \cdot (X_{\max} - x)^{d-m} dx, \quad (21)$$

а максимальні значення апроксимуючої функції визначаються з її формули (17)

$$G_{d,m} = \frac{1}{\sqrt{2\pi} \cdot \sigma_{d,m}^{(s)}}. \quad (22)$$

Зведення максимального значення функції Гауса до максимального значення $spdf$, що апроксимується, здійснюється коректувальним СКВ множником

$$\sigma_{d,m}^{(G)} = \rho_{d,m}^{(G)} \cdot \sigma_{d,m}^{(s)}. \quad (23)$$

Нормалізована функція Гауса

$$G_{d,m}(x) = \frac{\psi_{d,m}}{\sigma_{d,m}^{(G)} \cdot \sqrt{2\pi}} \cdot \exp\left[-0,5(2x/\xi_{d,m})^2 (\sigma_{d,m}^{(G)})^{-2}\right] \quad (24)$$

далі буде використана для статистичної оцінки каналів $spdf$ -конвертора.

Особливий випадок перетворень парними елементами. Серед усіх можливих пар елементів вибірок можна виділити пару, зміщення яких дорівнюють порядку вибірки, що розглядається. Для

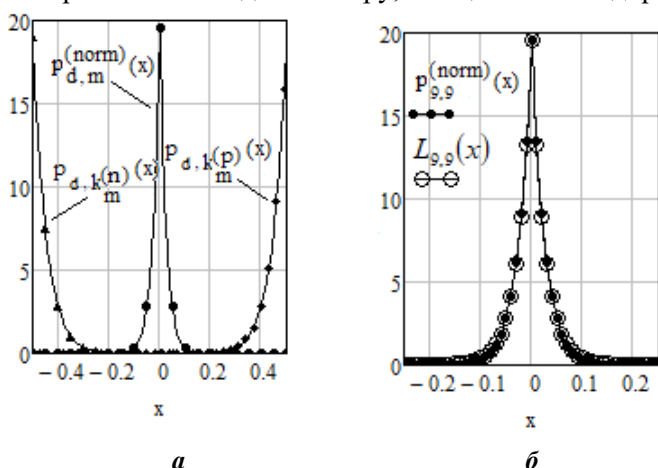


Рис. 11

крайніх лівосторонніх і правобічних елементів сума формується з монотонно спадної і монотонно зростаючої функцій щільності ймовірності, відповідно. (рис. 11, а). Всі інші утворені функціями щільності ймовірності пар елементів, що мають монотонно зростаючі і монотонно спадні області (рис. 9, а).

Сумарна щільність розподілу елементів для цих номерів має найбільшу відміну від апроксимуючої функції Гауса. Її форма має помітну схожість з функцією розподілу Лапласа.

Для перевірки припущення формулюються основна $H0^{(L)}$ і альтернативна їй $H1^{(L)}$

гіпотези про близькість або віддаленість *spdf* периферійних каналів з апроксимуючої функцією Лапласа, представленої у вигляді залежності від випадкової величини x з параметром α [18].

Нормалізація функції Лапласа. Корекція функції розподілу Лапласа не вимагає застосування коректувальних множників ξ і ψ , тому що елементи з максимальним зміщенням не містять залежних випадкових величин.

Використаємо як параметр α функції щільності ймовірності Лапласа (*pdfL*) значення СКВ $\sigma_{d,m}^{(s)}$ функції, що апроксимується, тобто, $\alpha = 2(\sigma_{d,m}^{(s)})^{-1}$

$$L_{d,m}(x) = \frac{\alpha}{2} \cdot \exp[-\alpha \cdot |x|] = (\sigma_{d,m}^{(s)})^{-1} \cdot \exp[-|2x|(\sigma_{d,m}^{(s)})^{-1}]. \quad (25)$$

Множник $\rho_{d,m}^{(L)}$ зведення *spdf* пари елементів з функцією щільності ймовірності Лапласа визначається з відношення їхніх значень (25) і (13) для незалежної випадкової величини $x = 0$

$$\rho_{d,m}^{(L)} = \frac{L_{d,m}(0)}{p_{d,m}^{(con)}(0)} = \sigma_{d,m}^{(s)} \cdot p_{d,m}^{\max}. \quad (26)$$

Після нормалізації *pdfL* (25) прийме вигляд

$$L_{d,m}(x) = \frac{\rho_{d,m}^{(L)}}{\sigma_{d,m}^{(s)}} \cdot \exp\left[-\frac{|2x| \cdot \rho_{d,m}^{(L)}}{\sigma_{d,m}^{(s)}}\right] = p_{d,m}^{\max} \cdot \exp[-|2x| \cdot p_{d,m}^{\max}]. \quad (27)$$

Результат нормалізації функції Лапласа представлений на рис. 11, б.

Середньоквадратична похибка апроксимації і в цьому разі визначається як різниця даних функцій. Наведені до діапазону змін інтегральні значення відхилень *spdf* від функції Лапласа (рис. 11, б), визначаються різницею залежностей (11), (12) і (27)

$$A_{d,m}^L = \frac{1}{2X} \cdot \int_{-X}^X |p_{d,m}^{(norm)}(x) - L_{d,m}(x)| dx. \quad (28)$$

Графічне подання похибки апроксимації функцією Лапласа для елементів із зміщенням m вибірок до $d = 25$ і умовним критичним значенням, рівним 0.01, наведено на рис. 12.

Перевірка статистичних гіпотез. Для аналізу властивостей впорядкованих вибірок з множини некорельованих випадкових величин було сформульовано два припущення щодо виду розподілу щільності ймовірності *spdf*-конвертора. В цьому випадку висунуті нульові гіпотези розглядаються незалежно один від одного, і кожна з них має власну альтернативну гіпотезу. Як і раніше [9], перевірка справедливості гіпотез для кожної з випадкових величин $\{X_{(n)}^{(sum)}\}$ на виході

багатоканального конвертора буде виконуватися за допомогою критерію згоди Пірсона χ^2 -квадрат, який для

безперервних функцій щільності ймовірності виходів конвертора $p_{d,m}^{(norm)}$ і розподілів Гауса $G_{d,m}(x)$ і Лапласа $L_{d,m}(x)$, запишеться у такий спосіб:

$$(\chi_{d,m}^{(G)})^2 = n \cdot \sum_{i=0}^n \frac{(p_{d,m}^{(norm)}(x_i) - G_{d,m}(x_i))^2}{G_{d,m}(x_i)}, \quad (29)$$

$$(\chi_{d,m}^{(L)})^2 = n \cdot \sum_{i=0}^n \frac{(p_{d,m}^{(norm)}(x_i) - L_{d,m}(x_i))^2}{L_{d,m}(x_i)}, \quad (30)$$

де n – число дискретних значень *spdf*, *pdfG* і *pdfL* (11), (12), (18), (27) від випадкової величини

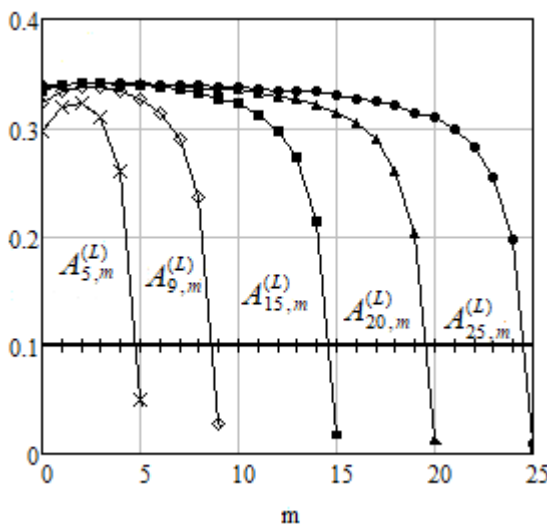


Рис. 12

$$-X_m \leq x_i \leq X_m \text{ [19, 20].}$$

Перевірки обох пар статистичних гіпотез проводяться в наступній послідовності: обчислюються роздільні статистики χ^2 -квдрат (29), (30) для n випадкових величин $\{X_{(n)}^{(s)}\}$, задаються загальні рівень значущості α , тип критичних областей і число ступенів свободи по числу інтервалів n , тобто $df = n - 1 - 2$, визначається єдина межа критичної області, рівна квантилю розподілу χ^2 порядку 0.95. Перевищення критеріями $(\chi_{d,m}^{(G)})^2$ або $(\chi_{d,m}^{(L)})^2$ узагальнюючої міри розбіжності між порівнюваними функціями відхиляє нульові гіпотези на користь прийняття висновку щодо відсутності згоди з очікуваними $pdfG$ і $pdfL$.

Результати перевірки гіпотез про узгодженість елементів окремих упорядкованих вибірок з

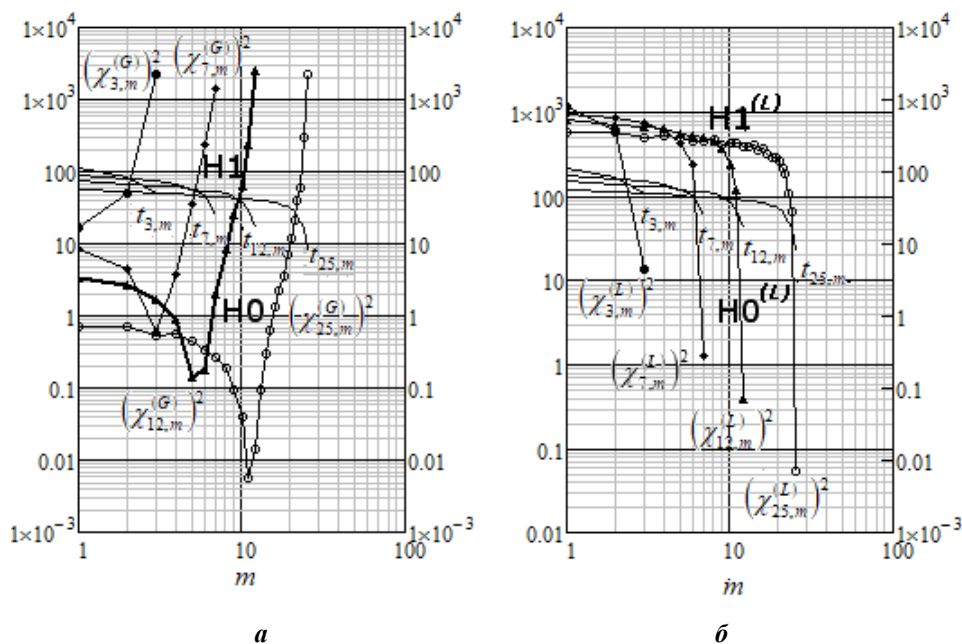


Рис. 13

законами розподілу Гауса і Лапласа показано на рис. 13, а, б, відповідно.

Пара серединних елементів дванадцятиканального $spdf$ -конвертора викликає сплюснення мінімальних значень критерію близькості до нормального закону розподілу $(\chi_{12,m}^{(G)})^2$ (рис. 13, а).

Вікна статистичної близькості до нормального закону розподілу $s_d^{(w)}$ разом з розміром конвертора парних елементів s_d для $d = 1 \dots 25$, які поєднують групи елементів, розподіл

щільності ймовірності яких згідний із нормальним законом, показано на рис. 14.

Лінійне наближення для визначення розмірів вікон близькості до нормального закону розподілу визначається формулою

$$s_d^{(G)} = 0.88 \cdot d. \quad (31)$$

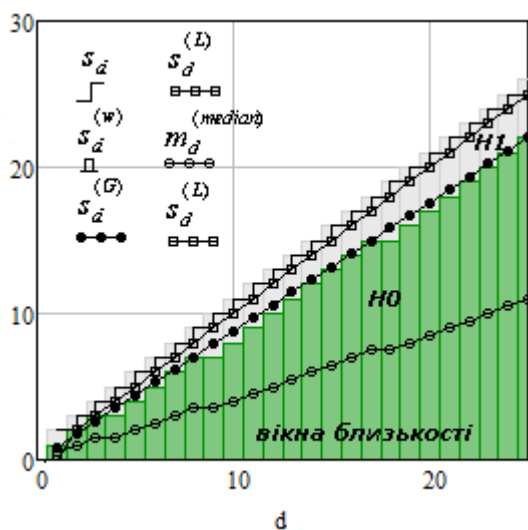


Рис. 14

Номери медіан конверторів $m_d^{(median)}$ також показано на цьому рисунку.

Вікна близькості до закону розподілу Лапласа $s_d^{(L)}$ мають одиничний розмір для всіх значень порядку вибірок.

Узагальненням результатів досліджень може служити статистична модель багатоканального конвертора парних елементів, що представляє собою набір функцій Гауса (24) і Лапласа (27), які статистично близькі до функцій щільності розподілу каналів. Параметри каналів моделі, які виділяються вікнами близькості, визначаються середніми значеннями (14), дисперсіями (16) і середньоквадратичними відхиленнями (17).

Порівняння вікон близькості каналів одноелементних pdf -конверторів [9] і $spdf$ -конверторів (31) заданих порядків показує суттєве збільшення (в 1.5

рази, приблизно) частки згодних з нормальним розподілом в конверторах парних елементів.

Висновки

1. Розроблено і випробувано методику аналітичних і чисельних досліджень перетворювальних властивостей багатоканальних конверторів пар елементів упорядкованих вибірок з множини випадкових некорельованих даних з рівномірним розподілом.

2. Визначено аналітичні залежності імовірнісних параметрів – безперервних функцій щільності ймовірності, середніх значень і моментів суми пар елементів від порядку впорядкованої вибірки і їхні зміщення елементів m щодо медіани.

3. Виведено вирази нормуючих і коректувальних множників для досліджуваних функцій щільності розподілу ймовірності і апроксимуючих їхніх функцій розподілу Гауса і Лапласа, запропоновано і реалізовано процедури нормалізації – приведення до єдиних шкал, і виключення впливу залежних випадкових величин на імовірнісні параметри;

4. Показано існування власних медіан $spdf$ -конверторів, які містять один в разі непарного або два – для парного порядку вибірок, серединних елементів в ряду його виходів, утворених сумою парних елементів.

5. Виконано статистичну перевірку гіпотез щодо близькості до нормального розподілу і закону розподілу Лапласа всіх пар елементів вибірок довільного порядку з використанням критерію згоди χ^2 -квадрат.

6. Узагальнення результатів в статистичній моделі багатоканального конвертора некорельованих даних засновано на вікнах статистичної близькості до кожного з розглянутих розподілів, що визначають розміри і положення вікон щодо каналів $spdf$ -конвертора. Параметри каналів моделі визначаються параметрами апроксимуючих функцій щільності ймовірності.

7. Запропонована методика досліджень може бути використана для визначення перетворювальних властивостей багатоканальних $spdf$ -конверторів 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. B. Horling and V. Lesser, A Survey of Multi-Agent Organizational Paradigms. *The Knowledge Engineering Review*. 2004. Vol. 19. No 4. Pp. 281–316. DOI: <https://doi.org/10.1017/S0269888905000317>.

3. Moradi, M. H., Razini, S., & Mahdi Hosseinian, S. State of art of multiagent systems in power engineering: A review. *Renewable and Sustainable Energy Reviews*. 2016. No 58. Pp. 814–824. DOI: <https://doi.org/10.1016/j.rser.2015.12.339>.

4. McArthur S. D. J. et al. Multi-Agent Systems for Power Engineering Applications—Part II: Technologies, Standards, and Tools for Building Multi-agent Systems. *IEEE Transactions on Power Systems*. 2007. Vol. 22. No 4. Pp. 1753-1759. DOI: <https://doi.org/10.1109/TPWRS.2007.908472>.

5. Shahbaz M. Q., Ahsanullah M., Hanif Shahbaz S. and Al-Zahrani B. Ordered Random Variables: Theory and Applications. Atlantis Press and Springer, 2016. 294 p. DOI: <https://doi.org/10.2991/978-94-6239-225-0>.

6. Dytso A., Cardone M., M. S. VeeduV. Poor, On estimation under noisy order statistics. Proceedings of the *International Symposium on Information Theory*, Paris, France, 2019. Pp. 36-40. URL: <https://arxiv.org/abs/1901.06294> (accessed at 20.12.2020)

7. Liu Z. and Zhu J. Signal Detection From Unlabeled Ordered Samples. *IEEE Communications Letters*. 2018. Vol. 22. No 12. Pp. 2431-2434. DOI: <https://doi.org/10.1109/LCOMM.2018.2875454>.

8. David H.A., Nagaraja H.N. Order Statistics. John Wiley & Sons, Inc., 2003. DOI: <https://doi.org/10.1002/0471722162>.

9. Mazmanian R.O. Noncorrelated data ordered samples as a single-element multichannel converter. *Tekhnichna Electrodynamika*. 2021. No 3. Pp. 74 – 82. (Ukr). DOI: https://doi.org/10.15407/techned2021_03_074

10. Mazmanian R.O. Ordered samples of random uncorrelated signal characteristics. *Tekhnichna Electrodynamika*. 2004. No 6. Pp. 60 – 64. (Rus).

11. Hogg R., McKean J., Craig A. Introduction to Mathematical Statistics. Boston: Pearson, 2019. 746 p.

12. Barron A.R., Sheu Ch.-H. Approximation of density functions by sequences of exponential families. *The Annals of Statistics*. 1991. Vol. 19. No 3. Pp. 1347-1369. DOI: <https://doi.org/10.1214/aos/1176348252>

13. Approximating the Probability Density Function of a Transformation of Random Variables. *Methodology and Computing in Applied Probability*. 2019. No 21. Pp. 633–645. DOI: <https://doi.org/10.1007/s11009-018-9629-0>
14. Montgomery, Douglas C., George C. Runger. Applied statistics and probability for engineers. John Wiley & Sons, Inc., 2002. 792 p.
15. Kappal S. Data Normalization using Median & Median Absolute Deviation (MMAD) based Z-Score for Robust Predictions vs. Min – Max Normalization. *Journal of Research in Science: Natural and Formal*. 2019. Vol. 19. No 4. Pp. 39 – 44.
16. Garthwirt J.L, Rubin H. The behavior of robust estimators on dependent data. *The Annals of Statistics*. 1975. Vol. 3. No 5. Pp. 1070-1100.
17. Grinstead C. M., Snell J.L. Introduction to probability. Providence, RI, American Mathematical Society, 1997. 510 p.
18. Kozubowski T.J., Podgorski K. Laplace probability distributions and related stochastic processes. Chapter 3. In book: *Probability: Interpretation, Theory and Applications*. New York: Nova Science Publishers, 2011. Pp. 105-145.
19. Guenther W.C. Power and Sample Size for Approximate Chi-Square Tests. *The American Statistician*. 1977. Vol. 31. No 2. Pp. 83-85.
20. 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. No 52. Pp. 4903–4908. DOI: <https://doi.org/10.1016/j.csda.2008.04.007>.

MULTICHANNEL CONVERSION OF RANDOM DATA WITH THE PAIR-ELEMENTS OF ORDERED SAMPLES

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 concept of multichannel parallel converting of probability density function (pdf) of random data was previously used for single-element pdf-converters. In development of this concept, here we investigate converting properties of spdf-converters channels formed by the sum of the pairs of ordered sample elements (order statistics). The characteristics of the conversion results as dependencies on the size of the samples and the displacement of the channels relative to the median of the samples were obtained for data with a uniform distribution density. Also where excluded the areas of mutual dependence of the density functions of the summed elements, which further where normalized together with approximating them functions. Despite the apparent structural differences, the goal of this study still was to determine the closeness of the converted data with some standard functions of the probability distribution density, in particular, with the normal distribution law. As before, the estimates of the closeness of the spdf-converter channels were obtained using the chi-square criteria. The results of the research were used to determine the size and location of the statistical closeness windows, and to construct spdf-converters statistical model. References 20, Figures 14.

Keywords: random data, uniformly distribution, ordered samples, probability density function (pdf), sum of the pairs of order statistics, spdf-converter, density estimation.

Надійшла 22.04.2021
Остаточний варіант 07.06.2021