

АВТОМАТИЗАЦІЯ ТА ІНТЕЛЕКТУАЛІЗАЦІЯ ПРИЛАДОБУДУВАННЯ

УДК 681.2.088

МЕТОД ОЦІНЮВАННЯ РОЗШИРЕНОЇ НЕВИЗНАЧЕНОСТІ РЕЗУЛЬТАТІВ ВИМІРЮВАНЬ ЗА ДОПОМОГОЮ ПЕРЕТВОРЕННЯ ДЖОНСОНА

¹⁾Самойліченко О.В., ¹⁾Єременко В.С., ¹⁾Мокійчук В.М., ²⁾Павленко Ж.О., ¹⁾Національний авіаційний університет, м. Київ, Україна; ²⁾Національний технічний університет України «Київський політехнічний інститут», м. Київ, Україна

Запропоновано та досліджено метод оцінювання розширеної невизначеності результату вимірювання. Метод базується на процедурі нормалізації даних за допомогою сімейства розподілів Джонсона. Досліджено точність визначення розширеної невизначеності запропонованим методом у порівнянні із методом заміни результуючого закону розподілу невизначеності гауссівським розподілом. Показана перевага методу оцінювання розширеної невизначеності за допомогою перетворення Джонсона

Вступ

Оцінювання сумарної невизначеності результату вимірювання за її складовими з подальшим знаходженням розширеної невизначеності регламентовано в [1]. Загальновідомими є правила підсумовування складових невизначеності та вибір коефіцієнту охопту, що характеризує інтервал розсіювання значень, який приписуються вимірюваній величині для відомих законів розподілу сумарної невизначеності. Далі будемо розглядати лише складові невизначеності, оцінені за типом А. Проте у практиці стикаються з проблемою вибору коефіцієнту охопту для випадку, коли реальний закон розподілу сумарної невизначеності не відповідає жодному з відомих законів розподілу, що може бути обґрунтовано особливостями процедури або об'єкту вимірювання. Така ситуація виникає, наприклад, для випадку, коли кількість складових сумарної невизначеності менше п'яти, і не виконуються умови центральної граничної теореми [2]. Неправильно оцінене в такому випадку значення коефіцієнту охопту K призводить до завищення чи заниження розширеної невизначеності.

Відомі наступні методи оцінювання розширеної невизначеності U для випадку негауссівського закону розподілу сумарної невизначеності: аналітичні, з використанням неперервної згортки функцій розподілу; чисельні, за допомогою моделювання методом Монте-Карло і дискретної згортки; використання сумарного ексцесу і класифікації розподілів; методи, що ґрунтуються на заміні реального розподілу нормальним та апроксимації розподілів поліномами, коригування результатів інтервального аналізу за допомогою нечітких множин [3, 4]. Неперервна згортка застосовується для випадку, коли закон розподілу результатів може бути заданий аналітично та потребує використання складного математичного апарату. Чисельні методи вимагають наявності апріорної інформації про закони розподілу складових невизначеності, наявності якісних

генераторів випадкових чисел та ін. Найбільш загальноновживаними і простими до застосування є апроксимаційні методи з використанням граничної теореми, коли результуючий закон сумарної невизначеності замінюється гаусівським, при цьому коефіцієнт охопту обирається рівним відповідному квантилю гаусівського закону розподілу. У групу цих методів входить також апроксимація реального закону розподілу сумарної невизначеності степеневими поліномами. Ці методи мають досить просту реалізацію, проте не забезпечують достатньої точності.

Постановка задачі

Аналіз можливих методів оцінювання розширеної невизначеності дозволив запропонувати досить зручний метод, що ґрунтується на апроксимації сумарного закону функцією, для якої відомий аналітичний вираз, та подальшому оцінюванні відповідних квантилів x_α та $x_{1-\alpha}$, за якими розширена невизначеність U приймається рівною інтервалу $[x_\alpha \dots x_{1-\alpha}]$.

Задача апроксимації реального розподілу може бути вирішена внаслідок опису його відомим розподілом. В [5] розглянуто три можливих підходи до описання розподілів. Перший підхід, запропонований Брусом, Грамом, Шарльє та Еджвортом, заснований на представленні емпіричної щільності розподілу у вигляді ряду від похідних щільності гаусівських розподілів. Другий підхід, запропонований Еджвортом та іншими авторами полягає у відшуванні такої функції від вихідних даних, розподіл якої, хоча б наближено, можна було б виразити через відомий розподіл. Третій підхід, запропонований Пірсоном, заснований на знаходженні сімейств кривих, за допомогою яких можливо задовільно представити емпіричні розподіли.

Існують наступні сімейства розподілів: криві Пірсона, зміна чотирьох основних параметрів якого призводить до великої кількості родин розподілів, сімейства кривих Зігеля-Тьюкі, які мають два параметри розподілу, а також розподіли Джонсона, що мають чотири параметри [6]. Використання кривих Пірсона дозволить досить точно апроксимувати розподіли, але при цьому виникає проблема методів підбору розподілів Пірсона, оскільки отримання кожного сімейства розподілів пов'язано з рішенням різних систем рівнянь. Апроксимація кривими Зігеля-Тьюкі легша у реалізації, оскільки не потребує складного математичного апарату, проте невелика кількість параметрів розподілу не дозволяє досить повно описати всі емпіричні розподіли. Отже, необхідно знайти компроміс між точністю апроксимації та простотою реалізації. Вказаним умовам відповідає система функцій, запропонована Джонсоном, яка дозволяє апроксимувати більш широкий клас розподілів, оскільки має чотири параметри. При цьому процедура вибору відповідної кривої Джонсона є досить простою.

Для отриманих даних обирають відповідну ним криву Джонсона, далі перетворюють початкову випадкову величину X (задану в деякому інтервалі) так, що результат перетворення можна розглядати як нормовану випадкову величину.

ну, розподілену за гауссівським законом, використовуючи яку визначають значення квантилів необхідного рівня.

Дане перетворення можливе за наступних умов: функція щільності розподілу випадкової величини є унімодальною; функція $g(x)$ є монотонною на заданому інтервалі; область значень функції $g(x)$ є у діапазоні від $-\infty$ до ∞ . Перевага даного підходу полягає в тому, що значення емпіричної функції розподілу випадкової величини X обчислюються як значення функції нормального розподілу. У загальному випадку перетворення має вигляд:

$$Z = \gamma + \eta \tau(X, \varepsilon, \lambda); \quad \eta > 0, \quad -\infty < \gamma < \infty, \quad \lambda > 0, \quad -\infty < \varepsilon < \infty,$$

де $\tau(X, \varepsilon, \lambda)$ – довільна функція; γ , η , ε і λ – параметри розподілу (форми, центру та масштабу); Z – нормована випадкова величина, розподілена за нормальним законом.

В [3] запропоновано три різні форми або сімейства функцій $\tau(X, \varepsilon, \lambda)$ – S_B , S_U (запропоновані Джонсоном [6], або S_L (логарифмічно нормальний розподіл). Отже, оцінки квантилів за нормалізованими даними і будуть шуканими квантилями x_α та $x_{1-\alpha}$, що дозволить вирішити поставлену задачу знаходження розширеної невизначеності.

Проведено дослідження процедури нормалізації для наступних складних випадків, які можуть мати місце на практиці, коли сумарна невизначеність складається з трьох складових у наступних комбінаціях: із гауссівським, рівномірним і трикутним (домінуючий рівномірний) законами розподілу, одна складова з арксинусоїдальним та дві з гауссівським (домінуючий арксинусоїдальний) законами розподілу, а також усі складові з рівномірним законом розподілу, одна з яких домінуюча. Параметри отриманого розподілу: сумарне математичне сподівання – $m_{\text{сум}} = m_1 + m_2 + m_3$, сумарне середнє квадратичне відхилення (с.к.в.) – $\sigma_{\text{сум}} = \sqrt{\sigma_1^2 + \sigma_2^2 + \sigma_3^2}$. Співвідношення с.к.в. визначалося за критерієм нехтовної малості похибки, а саме $\sigma_2/\sigma_1=1$, $\sigma_3/\sigma_1=3$. Зазначимо, що випадок, коли домінуючим є арксинусоїдальний закон розподілу призводить до двомодальності результуючого закону розподілу і застосування найпоширеніших методик розрахунку розширеної невизначеності призводить значних похибок [4]. Як було зазначено, процедура нормалізації передбачає унімодальність функції щільності ймовірності, але для вирішення поставлених задач це виявилось не обов'язковим. Достатньою є точна апроксимація лише крайніх областей, у яких знаходяться шукані квантили x_α , а це досягається правильним підбором типу кривої Джонсона.

Також була поставлена задача порівняння точності визначення квантилів за допомогою процедури нормалізації з точністю оцінювання квантилів, що отримані заміною результуючого закону гауссівським, оскільки останній спосіб є найбільш простим і розповсюдженим при оцінюванні розширеної невизначеності на практиці.

Дослідження проводилося із застосуванням методів імітаційного моделювання за наступною методикою:

1. За допомогою комп'ютерного моделювання було отримано три масиви випадкових чисел, із заданими законами розподілу, що імітували складові сумарної невизначеності.

2. Знаходився сумарний закон розподілу випадкової величини, яка визначалася як арифметична сума отриманих у п. 1 послідовностей.

3. Для отриманої реалізації обраховувалися оцінки b_1 та b_2 (де b_1 – квадрат нормованого показника асиметрії, а b_2 – нормований показник гостровершинності [6], на значеннями яких приймалося рішення про тип кривої Джонсона).

Для вирішення поставлених задач на підставі розрахованих оцінок (b_1 , b_2) було використано розподіл типу S_B Джонсона. Розподіл S_B Джонсона має два параметра форми γ та η : один параметр, що характеризує центр розподілу ε , один параметр масштабу λ . Випадкова величина, що має розподіл S_B Джонсона, теоретично обмежена границями ε і $\varepsilon + \lambda$. В загальному, ця умова приводить до наступних випадків:

- а) обидва крайніх значення ε і $\varepsilon + \lambda$ відомі;
- б) відомо тільки одне крайнє значення;
- в) жодне крайнє значення не відомо.

У межах проведеного дослідження розглядався третій випадок.

4. Чотири емпіричних квантиля x_i , $i=\overline{1,4}$, прирівнювалися відповідним квантилям нормованого нормального розподілу z_i , $i=\overline{1,4}$ та отримували систему чотирьох нелінійних рівнянь:

$$z_i = \hat{\gamma} + \hat{\eta} \cdot \ln \left(\frac{x_i - \hat{\varepsilon}}{\hat{\varepsilon} + \hat{\lambda} - x_i} \right) \quad i = \overline{1,4}, \quad (1)$$

яка розв'язувалась чисельними методами.

5. За отриманими значеннями параметрів розподілу γ , η , ε і λ проводилась процедура нормалізації даних. При цьому вирішують обернене до (1) рівняння:

$$x_\alpha = \frac{\exp(z_\alpha - \gamma)(\lambda + \varepsilon)}{1 + \exp\left(\frac{z_\alpha - \gamma}{\eta}\right)},$$

де x_α – шукані оцінки квантилів, α – обраний рівень довірчої ймовірності.

Отже, за оцінку розширеної невизначеності U приймалося значення x_α (або інтервал $[x_\alpha \dots x_{1-\alpha}]$). Точність запропонованого методу значною мірою залежить від точності використаного методу розв'язку системи нелінійних рівнянь і точності апроксимації на кінцях розподілу, де знаходяться шукані квантилі.

6. Для оцінювання точності результатів запропонованого методу обраховувалися умовно істинні значення квантилів необхідного рівня значимості по сукупності, що наближалася до генеральної (об'ємом $N=100000$ значень), а також значення квантилів, отримані шляхом заміни результуючого закону гауссівським.

При дослідженні процедури нормалізації представлена вище методика по-

вторювалася $N=10000$ разів, що дозволило оцінити варіацію оцінок шуканих квантилів.

Результати досліджень

Використовуючи процедуру нормалізації було апроксимовані розподіли вихідних даних кривою S_B Джонсона, результати апроксимації суми гауссівського, трикутного і рівномірного (домінуючого) законів представлено на рис. 1. Апроксимація суми двох рівномірних та аркисинусоїдального законів розподілу представлено на рис. 2.

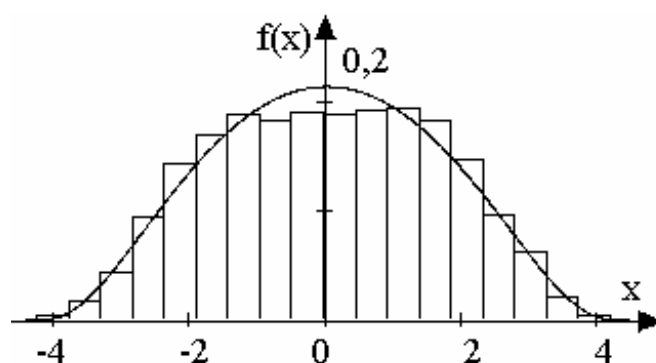


Рис. 1. Апроксимація суми гауссівського, рівномірного і трикутного законів розподілу кривою S_B Джонсона

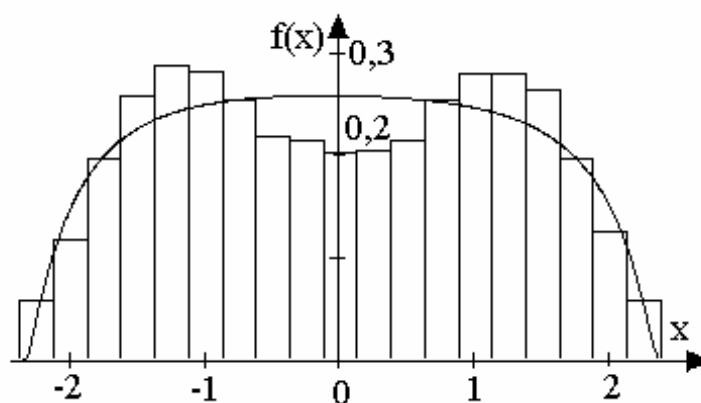


Рис. 2. Апроксимація суми гауссівського, рівномірного і аркисинусоїдального законів розподілу кривою S_B Джонсона

На рис. 2, 3 представлені гістограми розподілу оцінок квантилів рівня 0,025 та 0,05 відповідно, для випадку суми гауссівського, трикутного і рівномірного законів розподілу, $f1(x)$, $f3(x)$ – розподіли оцінок, отриманих при заміні результуючого закону гауссівським, $f2(x)$, $f4(x)$ – розподіли оцінок, отриманих за допомогою процедури нормалізації.

Пунктирною лінією позначені умовно істинні значення квантилю відповідного рівня.

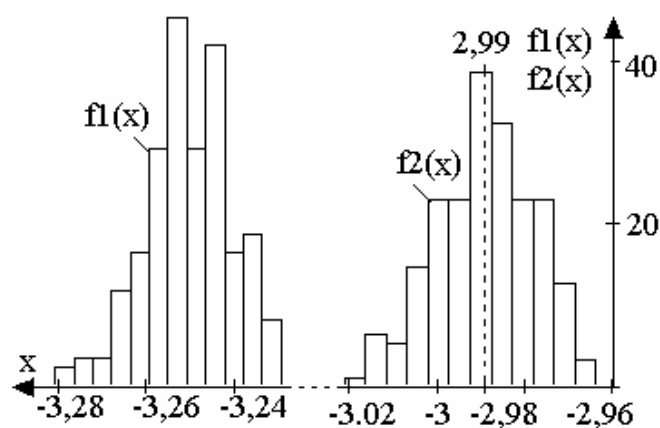


Рис. 3. Розподіли оцінок квантилів рівня 0,025

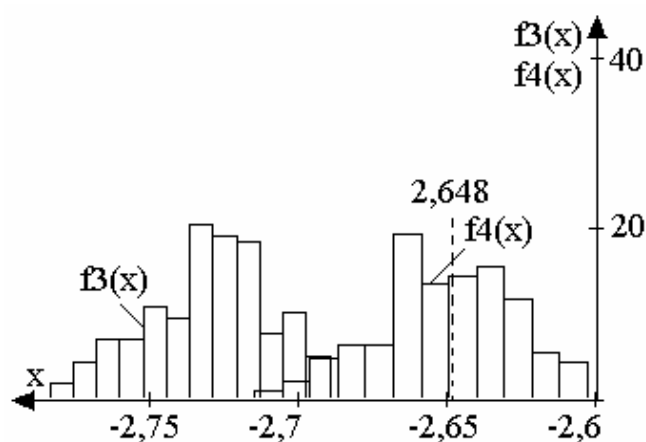


Рис. 4. Розподіли оцінок квантилів рівня 0,05

За отриманими оцінками квантилів були розраховані значення їх с.к.в. (табл.1).

Таблиця 1. Значення середніх квадратичних відхилень оцінок квантилів

Досліджувані закон розподілу	Процедура обробки	Значення с.к.в. для рівня квантиля			
		0,025	0,05	0,95	0,975
Сума 3-х рівномірних	Нормалізація	0,012	0,024	0,022	0,012
	Заміна гауссівським	0,012	0,022	0,021	0,012
Сума гауссівського, рівномірного і трикутного	Нормалізація	0,057	0,048	0,049	0,050
	Заміна гауссівським	0,049	0,045	0,043	0,061
Сума двох рівномірних та арксинусоїдального	Нормалізація	0,007	0,006	0,006	0,007
	Заміна гауссівським	0,007	0,006	0,006	0,009

З таблиці видно, що варіація обрахованих статистик залежить від рівня

квантиля, але не залежить від методу отримання оцінки. Аналіз рисунків показує, що розподіл оцінки квантиля (середнє значення), обрахованої за допомогою заміни результуючого розподілу гауссівським має зміщення за віссю x , тим самим завищуючи реальні межі розширеної невизначеності результату вимірювання, а середнє значення оцінки квантиля, обрахованої за допомогою процедури нормалізації збігається з умовно істинним значенням.

З приведених результатів можна зробити висновок, що збіжність отриманих оцінок є однаковою, проте різною є правильність оцінювання, вона значно вища для методу, що заснований на процедурі нормалізації.

Оцінкою правильності в даному випадку будемо вважати відхилення середнє арифметичне значення оцінки квантилю від умовно істинного значення квантилю. В табл. 2. представлені розраховані оцінки квантилів, умовно істинні значення квантилів, а також відносні похибки $\delta 1$ %, обраховані при використанні методу нормалізації та $\delta 2$ % – обраховані при заміні результуючого закону гауссівським законом розподілу. Значення $\delta 1$ та $\delta 2$ розраховувались за наступними формулами:

$$\delta 1 = \frac{|U_{\text{іст}} - U_{\text{норм}}|}{U_{\text{іст}}} \cdot 100\%, \quad \delta 2 = \frac{|U_{\text{іст}} - U_{\text{гаус}}|}{U_{\text{іст}}} \cdot 100\%,$$

де $U_{\text{іст}}$ – умовно істинне значення для сумарного закону; $U_{\text{норм}}$ – оцінки квантилів, отриманих при використанні запропонованого методу; $U_{\text{гаус}}$ – оцінки квантилів, отриманих при заміні розподілу гауссівським законом.

Таблиця 2. Значення похибок оцінок квантилів

Рівень квантиля	Нормалізація	Заміна гауссівським	$\delta 1$, %	$\delta 2$, %	Умовно істинні значення квантилю
Сума гауссівського, рівномірного і трикутного законів розподілу					
0,025	-2,99	-3,25	$<10^{-3}$	8,69	-2,99
0,05	-2,65	-2,73	$<10^{-3}$	3,02	-2,65
0,95	2,65	2,73	$<10^{-3}$	3,02	2,65
0,975	2,99	3,25	$<10^{-3}$	8,69	2,99
Сума трьох рівномірних законів розподілу					
0,025	-6,01	-6,50	0,04	8,14	-6,01
0,05	-5,32	-5,46	0,07	2,66	-5,32
0,95	5,32	5,46	0,07	2,66	5,32
0,975	6,01	6,50	0,04	8,14	6,01
Сума двох рівномірних та арксинусоїдального законів розподілу					
0,025	-2,06	-2,40	$<10^{-3}$	16,5	-2,06
0,05	-1,87	-2,02	$<10^{-3}$	8,02	-1,87
0,95	1,87	2,02	$<10^{-3}$	8,02	1,87
0,975	2,06	2,55	$<10^{-3}$	23,55	2,06

Іншим підходом до порівняння процедури нормалізації та заміни резуль-

туючого закону гауссівським є порівняння відсотку значень у отриманій реалізації, що знаходиться в межах оцінених квантилів. Використовуючи обраховані значення квантилів обраховувався реальний відсоток значень, що знаходився між ними, він порівнювався з необхідним відсотком $S_{95}=95\%$, $S_{90}=90\%$ (між квантилям рівня 0,025 та 0,975 повинно знаходитися 95% всіх значень, між квантилям рівня 0,05 та 0,95 – 90%), і обрахувалася відносна похибка

$$\delta 1 = \frac{|S_{95} - S_{\text{норм}}|}{S_{95}} \cdot 100\%, \quad \delta 2 = \frac{|S_{95} - S_{\text{гаус}}|}{S_{95}} \cdot 100\%.$$

В якості відсотку значень, що знаходяться між оціненими квантилям брався усереднений показник по N результатам обрахунку відсотку значень. Результати представлені в таблиці 3.

Таблиця 3. Відсоток значень (S) у вибірці, що знаходиться в межах інтервалу з заданим рівнем ймовірності P

	$P, \%$	Нормалізація	Заміна гауссівським	$\delta 1, \%$	$\delta 2, \%$
	Сума гауссівського, рівномірного і трикутного законів розподілу				
$S, \%$	90	89,99	91,37	0,01	1,52
	95	95,01	97,37	0,01	2,49
	Сума трьох рівномірних законів розподілу				
$S, \%$	90	90,01	91,20	0,01	1,33
	95	95,01	97,30	0,01	2,45
	Сума двох рівномірних та арксинусоїдального законів розподілу				
$S, \%$	90	89,95	93,96	0,01	4,40
	95	95,01	99,38	0,01	4,61

З цієї таблиці видно, що значення відносної похибки при використанні процедури нормалізації значно менше, ніж при заміні гауссівським законом. Відносна похибка обрахунку квантилів, а також відсотку реальних значень, що знаходяться в межах 90% та 95% не перевищує 0,01%, що дає підстави стверджувати про можливість використання запропонованого методу оцінки розширеної невизначеності, що ґрунтується на процедурі нормалізації, оскільки точність його є на два порядки вищою. Окрім того, для випадку суми трьох законів з домінуючим арксинусоїдальним законом процедура нормалізації також показує прийнятні результати, хоча функція щільності розподілу ймовірності сумарного закону двомодальна і усунення значних похибок (до 20%) при оцінюванні розширеної невизначеності для такого випадку потребує спеціальних підходів та алгоритмів. Процедуру нормалізації в свою чергу можна використовувати без змін, при цьому отримуючи значення відносної похибки оцінки квантилів менше 0,01%.

Висновки

1. Запропонований метод оцінювання розширеної невизначеності результату вимірювання використовуючи процедуру нормалізації для випадку, коли складові сумарної невизначеності оцінюються за типом А не потребує знаходження коефіцієнту охопту K .

2. Порівняння оцінок квантилів, а відповідно і розширеної невизначеності, показало перевагу метода нормалізації, який забезпечив відносну похибку оцінювання менше 0,01%.

3. Процедура нормалізації має порівняно просту реалізацію, єдиною математичною складністю якої є застосування чисельних методів розв'язку системи нелінійних рівнянь, які в свою чергу мають широке представлення в різноманітних програмних пакетах. Точність процедури нормалізації в значній мірі залежить від точності обрахунку параметрів розподілу і практично не залежить від закону розподілу сумарної невизначеності і рівня шуканого квантиля.

Запропонований підхід рекомендовано використовувати при опрацюванні результатів раундів міжлабораторних порівнянь, при атестації стандартних зразків та в інших випадках, коли можна отримати достатньо велику кількість результатів.

Література

1. Guide if the Expression of Uncertainty in Measurment. ISO. Geneva, 1993.
2. Вентцель Е.С. Теория вероятностей. М.: Гос. изд. физ.-мат. литературы, 1958. – 464 с.
3. Циделко В.Д. Исследование методов оценивания расширенной неопределенности / В.Д. Циделко, Н.А. Яремчук, М.В. Галевская // Системи обробки інформації. – 2008. – №4(71). – С.49-54.
4. Карпенко А.Н. Объединение погрешностей результатов измерений с использованием компенсации результатов интервального анализа / А.Н. Карпенко // Інформаційні технології та комп'ютерна інженерія. – 2008. – №3(13). – С. 70-75.
5. Кендалл М., Стьюарт А. Теория распределений. М.: Наука, 1966.- 587 с.
6. Хан Г. Статистические модели в инженерных задачах / Г.Хан, С.Шапиро. – М.: Мир, 1969. – 395 с.

Самойличенко О.В., Ерѐменко В.С., Мокийчук В.М., Павленко Ж.А. Метод оценивания рас-ширенной неопределенности результатов измерений с помощью преобразования Джонсона Предложен и исследован метод оценивания расширенной неопределенности результата измерения. Метод базируется на процедуре нормализации данных с помощью семейства распределений Джонсона. Исследована точность определения расши-ренной неопределенности предложенным методом в сравнении с методом замены результирующего закона распределе-ния неопределенности гауссовским распреде-лением. Показано преимущество метода оце-нивания расширенной неопределенности при помощи преобразования Джонсона.	Somoylichenko O.V., Yeremenko V.S., Mokiychuk V.M., Pavlenko J.A. Method Of Measuring Result's Expanded Uncertainty Estimation Based On Jonson Transformation Method of expanded uncertainty estimation was introduced and investigated. The methods is based on procedure of data normalization due to Jonson transformation. The expanded uncertainty evaluation precision of investigated method versus method of substitution to Gauss distribution was explore. The advantage of method of measuring result's expanded uncertainty estimation based on Jonson transformation was shown.
--	---

Надійшла до редакції
8 травня 2009 року