УДК 519.24

Ю.В. Куц, С.В. Шенгур, Л.М. Щербак

Національний авіаційний університет, Київ, Україна

ХАРАКТЕРИСТИКА КУТОВИХ ВИМІРЮВАНЬ ПРИ СТАТИСТИКАХ МАЛОГО ОБСЯГУ

Основними формами представлення похибки результатів вимірювань є розширена невизначеність та довірчий інтервал. В статті досліджується модель побудови довірчого інтервалу в задачі аналізу кутових даних обмеженого обсягу з використанням так званого статистичного методу "The Bootstrap method". Наведені алгоритм побудови довірчого інтервалу та результати його комп'ютерного моделювання для кутових вимірювань. Представлені результати порівняльного аналізу визначення характеристик кутових вимірювань для двох методів опрацювання вибірок малого обсягу: відомого класичного та запропонованого.

Ключові слова: *розширена невизначеність, Bootstrap, довірчий інтервал, кутові дані, статистична обробка, вибірки малого обсягу.*

Вступ

Згідно з діючими міжнародними стандартами результати вимірювань, в тому числі кутових і фазових, повинні подаватись в сукупності з показниками точності у формі невизначеності [1, 2]. Формування інтервальних оцінок статистичних характеристик кутових вимірювань за результатами вибірок значного обсягу розглядалось в [3, 4]. Запропонована в цих роботах методика передбачає визначення границь довірчих інтервалів результатів вимірювань як добуток стандартної невизначеності на коефіцієнт розширення, який залежить від закону розподілу ймовірності досліджуваного випадкового кута чи фазового зсуву сигналів.

В цей же час, обсяг вибірки $N1$ кутових даних в ряді випадків обмежений умовами проведення експериментів і не перевищує значень $N1 \in [10,\ 20]$. Це, по-перше, не дозволяє перевірити гіпотезу про закон розподілу випадкового кута і, відповідно, не дозволяє статистично обґрунтувати коефіцієнт розширення, по-друге, суттєво обмежує точність оцінки стандартної невизначеності. Проблема малих вибірок полягає у порівняно невисокій точності отримання інтервальних оцінок та вимагає пошуку нових методик опрацювання результатів вимірювання.

Останнім часом набуває поширення метод статистичної обробки малих вибірок, вперше описаний у 1979 році американським математиком Бредлі Ефроном (Bradley Efron). Цей метод дістав назву " The Bootstrap method" [5]. Якщо взяти за основу структуру побудови алгоритму опрацювання даних вимірювань цього методу, то синонімами назви такого методу можуть бути методи: "початкового завантажування", або метод "розкрутки" чи метод "перестановок".

У математичній статистиці під "розкруткою" розуміють новітній метод отримання статистичних висновків, що виділився з більш широкого класу методів повторних вибірок і вимагає значного обсягу обчислень [5]. Зручність використання методу "розкрутки" полягає у простоті його застосування для отримання значень розширених невизначеностей та довірчих інтервалів, оцінювання параметрів розподілу випадкових кутів, таких як процентні квантилі, довірчі ймовірності та коефіцієнти кореляції в умовах значної апріорної невизначеності.

У статті розглянуто застосування методу „розкрутки" для оцінки розширеної невизначеності результатів кутових спостережень в умовах вибірки малого обсягу.

Постановка задачі

Проведена серія з $N1 = 10...20$ незалежних вимірювань кута $\psi(\omega) \in [0, 2\pi)$, $\omega \in \Omega$ (ω – елементарна подія з простору подій Ω), з апріорно невідомою одновершинною щільністю розподілу ймовірності $p(\theta)$, $\theta \in [0, 2\pi)$. Вимірювання проводиться за незмінних умов. Під час вимірювання відбувається рівномірне квантування значень кутів з кроком $2\pi/m$. Даними вимірювань є вибірка $\Theta = \left\{\theta_{n1}, n1 = \overline{1, N1}\right\}$, $\theta_{n1} \in [0, 2\pi)$, обсягу $N1$, яка розглядається як реалізація випадкового вектора $\psi(\omega) = \left(\psi_1(\omega), ...\psi_{n1}, ...\psi_{N1}(\omega)\right)$ з $N1$ незалежними та однаково розподіленими випадковими кутами $\psi_{n1}(\omega) \in [0, 2\pi)$.

Визначити, використовуючи метод „розкрутки", розширену невизначеність, результат вимірювання і, загалом, довірчий інтервал для середнього значення кута та подати результат вимірювання у встановленій формі [1, 2] за умови, що систематична складова похибки вимірювання відсутня.

Розв’язок

Спочатку розглянемо сутність методу „розкрутки”. Вона полягає у:

- формуванні з вхідної вибірки Θ обмеженого обсягу N1 ряду K „розкручених” допоміжних вибірок

$$\Theta^k = \left(\theta^k_{n1},\ n1=\overline{1,N1}\right),\ k=\overline{1,K}$$

такого ж обсягу K, які формуються шляхом випадкового вибору з поверненням значень

$$\theta_{n2} \in [0,2\pi),\ \ n2=\overline{1,N2}$$

з вхідної вибірки Θ (значення N2 може вибиратись рівним декільком сотням); випадкова вибірка кутів може бути здійснена, наприклад, шляхом формування псевдовипадкової послідовності

$$n1'_1,...n1_{n1},...n1'_{N1}$$

з множини псевдовипадкових чисел $u_1,...,u_{N1}$ з рівномірним розподілом в інтервалі значень [0,1] як

$$n1_{n1} = [u_{n1}N1+1]\,, \tag{1}$$

де квадратними дужками позначена операція виділення цілої частини числа;

- визначенні для всіх множин Θ'_k, вибіркових кругових середніх кутів

$$\overline{\theta^k} = \left\{\mathrm{arctg}\frac{S^k}{Ck} + \frac{\pi}{2}\left\{2-\left(\mathrm{sign}S^k\right)\left[1+\left(\mathrm{sign}C^k\right)\right]\right\}\right\},$$

$$\overline{\theta^k} \in [0,2\pi)\,, \tag{2}$$

де sign – знакова функція, а

$$C^k = \frac{1}{N1}\sum_{n1=1}^{N1}\cos\theta^k_{n1},\quad S^k = \frac{1}{N1}\sum_{n1=1}^{N1}\sin\theta^k_{n1}\,; \tag{3}$$

- визначенні вибіркового серединного кута $\bar{\theta}$ з вхідної вибірки $\Theta = \left\{\theta_{n1}, n1=\overline{1,N1}\right\}$ згідно з виразами (2), (3);

- визначенні відхилень вибіркових кругових середніх кутів (2) від вибіркового серединного кута $\bar{\theta}$ вхідної вибірки

$$\Delta\theta^k = \left|\overline{\theta^k} - \bar{\theta}\right|; \tag{4}$$

- побудові варіаційного ряду з елементів отриманої множини $\left(\Delta\theta^k,\ k=\overline{1,K}\right)$

$$\Delta\theta^{(1)} \le \Delta\theta^{(2)} \le ...\Delta\theta^{(k)} \le ... \le \Delta\theta^{(K)}; \tag{5}$$

- оцінці розширеної невизначеності для обраної довірчої ймовірності $P_{дов} = 1-\alpha$ за відповідним квантилем

$$U_1 = \Delta\theta^{\left([0,5K\alpha+0,5]+1\right)},\ U_2 = \Delta\theta^{\left(K-[0,5K\alpha+0,5]\right)}. \tag{6}$$

Результат подається інтервальною оцінкою як

$$\left(\bar{\theta}+U_1\left(\bmod 2\pi\right),\ \bar{\theta}+U_2\left(\bmod 2\pi\right)\right),\ P_{дов}\,. \tag{7}$$

Структурну схему застосування методу „розкрутки” для опрацювання даних кутових вимірювань наведено на рис. 1.

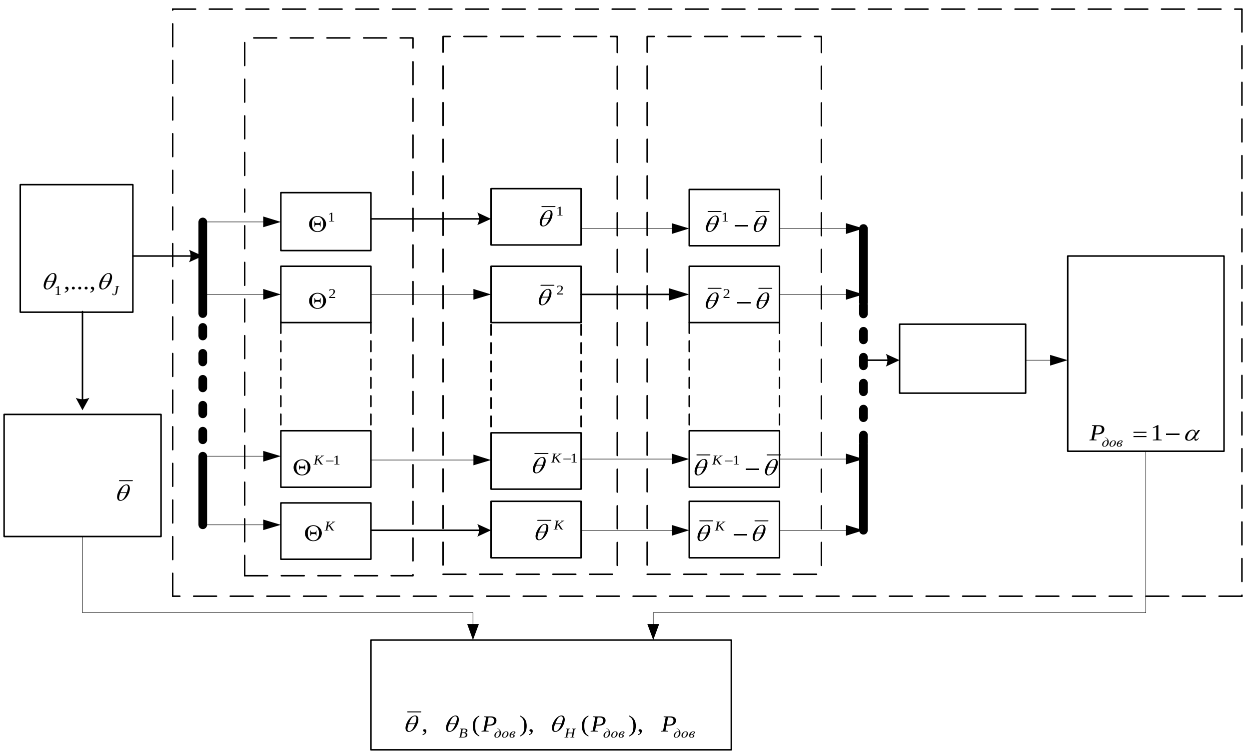

Рис. 1. Зображення послідовності арифметико-логічних операцій опрацювання за методом „розкрутки”

Результат моделювання

Перевірку ефективності методу „розкрутки” для обчислення розширеної невизначеності вибірки кутових даних обмеженого обсягу виконано комп’ютерним моделюванням методу Монте-Карло. На першому етапі сформовано генеральну вибірку обсягом 10000 значень випадкового кута з гаусівським намотаним розподілом імовірності. З цією метою формувалась вибірка відповідного обсягу гаусівської випадкової величини $\xi(\omega)$ з математичним сподіванням, що дорівнює π (тобто генеральне середнє $\overline{\theta}_{\Gamma} = \pi$), та дисперсією $\sigma^2 = 1$. До гаусівської випадкової величини застосовувалось нелінійне перетворення виду $\psi(\omega) = \xi(\omega)(\operatorname{mod} 2\pi)$, що забезпечило трансформацію вхідної випадкової величини у випадкову величину з намотаним гаусівським розподілом і такими ж параметрами – $\sigma^2 = 1$, $\overline{\theta}_{\Gamma} = \pi$.

З цієї генеральної вибірки шляхом випадкового вибору значень формувались вибірки малого обсягу $N1 = 10$. Вибірка здійснювалась шляхом формування псевдовипадкових номерів вибірки у відповідності до (1).

Приклад однієї з отриманих вибірок представлено на рис. 2. На цьому рисунку напрям відліку кутів обраний за годинниковою стрілкою, а вибіркові значення кутів позначені точками на колі одиничного радіусу.

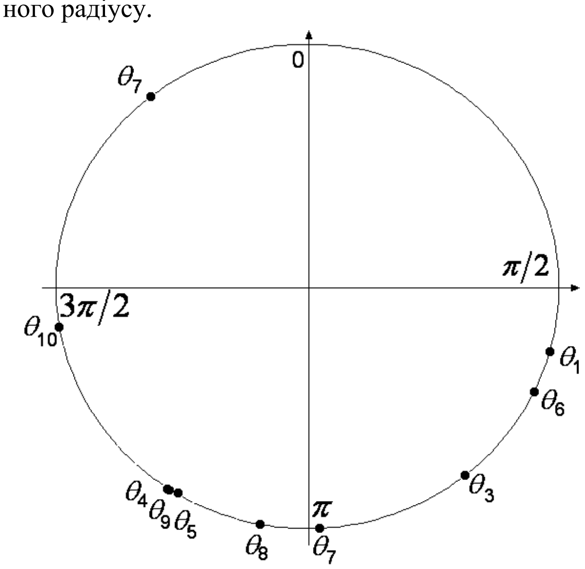

Рис. 2. Графічне зображення вибірки кутів обсягу $N1 = 10$

Для сформованих в такий спосіб малих вибірок кутів визначався довірчий інтервал за методом „розкрутки” для довірчої ймовірності $P_{\text{дов}} = 0{,}95$. За цієї умови було обрано $K = 200$.

На другому етапі знаходилась оцінка розширеної невизначеності та середнього кута для малої вибірки згідно з наведеною вище методикою. Значення розширеної невизначеності оцінювалось як різниця квантилів варіаційного ряду з номерами

$$l_1 = \left[\frac{1}{2} \cdot K \cdot \alpha + \frac{1}{2}\right] + 1 \text{ та } l_2 = K - l_1 .$$

Для $K = 200$ отримано $l_1 = 6$, $l_2 = 195$. Відповідно до цього, розширена невизначеність оцінювалась як $U = \theta_{l_2} - \theta_{l_1}$.

На рис. 3 зображено довірчий інтервал, отриманий методом „розкрутки”, для вибірки кутів, зображеної на рис. 2.

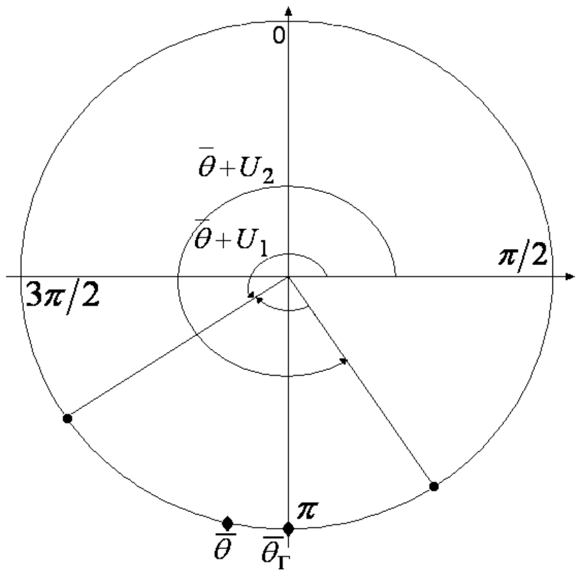

Рис. 3. Довірчий інтервал, отриманий методом „розкрутки”

Оцінка ефективності методу „розкрутки” здійснювалась шляхом порівняння отриманих в такий спосіб довірчих інтервалів (розширених невизначеностей) з відповідними оцінками, які отримувались за класичною методикою, яка передбачає визначення емпіричного середнього квадратичного відхилення для кутів малої вибірки, визначення коефіцієнтів Стьюдента з урахуванням гіпотези про близький до гаусівського розподілу кругового серединного кута і формування інтервалу як

$$\overline{\theta} \pm t_{N1,P_{\text{дов}}} \hat{\sigma} ,$$

де $t_{N1,P_{\text{дов}}}$ – коефіцієнт Стьюдента, $\hat{\sigma}$ – оцінка середньоквадратичного відхилення для кутів малої вибірки. Середньоквадратичне відхилення визначалось як [3, 5]

$$\sigma = \sqrt{-2 \ln r} , \tag{8}$$

де

$$r = \sqrt{\left(\frac{1}{N1}\sum_{n1=1}^{N1} \cos\theta_{n1}\right)^2 + \left(\frac{1}{N1}\sum_{n1=1}^{N1} \sin\theta_{n1}\right)^2} . \tag{9}$$

Отримані в серії з $M = 1000$ випробувань результати підтвердили відповідність методу „розкрутки" класичному методу визначення довірчих інтервалів.

Відповідність експериментального значення довірчої ймовірності заданому перевірялось шляхом оцінки частоти попадання генерального середнього $\bar{\theta}_{\Gamma}$ у визначений довірчий інтервал при проведенні серії з M вимірювань.

За оцінку довірчої імовірності було взято попадання генерального середнього у довірчий інтервал, тобто частоту появи події.

$$\bar{\theta}_{\Gamma} \in \left(\bar{\theta}_m + U_{m1}\left(\mathrm{mod}\,2\pi\right), \bar{\theta}_m + U_{m2}\left(\mathrm{mod}\,2\pi\right)\right), \quad (10)$$

де $\bar{\theta}_m$, U_m – значення відповідно вибіркового середнього кута та розширеної невизначеності, отримані в m-му експерименті, $m = \overline{1, M}$.

Якщо з M випробовувань випадає L експериментів, для яких виконується умова (10), довірча імовірність оцінюється як

$$P_{\text{дов}} \sim L/M .$$

В проведених експериментах отримано $L/M = 0,94$, що вказує на збереження довірчої ймовірності під час оцінки розширеної невизначеності методом „розкрутки".

Таким чином, проведені дослідження засвідчили ефективність застосування методу „розкрутки" для формування інтервальних оцінок результатів кутових спостережень за вибірками малого обсягу.

Висновки

Застосування методу „розкрутки" для оцінки розширеної невизначеності результатів кутових вимірювань для вибірок малого обсягу в умовах апріорної невизначеності розподілу випадкових кутів є досить ефективним, що підтверджується результатами моделювання.

Наведена методика статистичної обробки результатів кутових вимірювань може бути використана під час обробки експериментальних даних вимірювання кутових величин чи фазових зсувів сигналів в умовах дії завад та обмеженого обсягу даних вимірювань.

Список літератури

1. *Захаров И.П. Теория неопределенности в измерениях: учеб. пособие / И.П. Захаров, В.Д. Кукуш. – Х.: Консум, 2002. – 256 с.*
2. *Циделко В.Д. Невизначеність вимірювання. Обробка даних і подання результатів вимірювань: монографія / В.Д. Циделко, Н.А. Яремчик. – К.: Видавництво «Політехніка»", 2002. – 176 с.*
3. *Куц Ю.В. Статистична фазометрія / Ю.В. Куц, Л.М. Щербак. – Тернопіль: Тернопільський державний університет ім. І. Пулюя, 2009. – 383 с.*
4. *Куц Ю.В. Використання невизначеності для обробки даних і подання результатів кутових вимірювань / Ю.В. Куц, О.В. Монченко // Тез. докладов II научн.-техн. семинара "Неопределенность измерения: нормативные, научные, методические и производственные аспекты". – Х.: Харківський національний університет радіоелектроніки, 2005. – С. 34-36.*
5. *Fisher N.I. Statistical analysis of circular data / N.I. Fisher. Cambridge: Cambridge University Press, 2000. – 277 p.*

Надійшла до редколегії 28.04.2010

Рецензент: канд. техн. наук, проф. В.С. Єременко, Національний авіаційний університет, Київ.

ХАРАКТЕРИСТИКА УГЛОВЫХ ИЗМЕРЕНИЙ ПО СТАТИСТИКАМ МАЛОГО ОБЪЕМА

Ю.В. Куц, Л.Н. Щербак, С.В. Шенгур

Основными формами представления погрешности результатов измерений являются расширенная неопределенность и доверительный интервал. В статье исследуется модель построения доверительного интервала в задаче анализа угловых данных ограниченного объема с использованием так называемого статистического метода "The Bootstrap method". Приведены алгоритм построения доверительного интервала и результаты его компьютерного моделирования для угловых измерений. Представлены результаты сравнительного анализа определения характеристик угловых измерений для двух методов обработки выборок малого объема: известного классического и предложенного.

Ключевые слова: *расширенная неопределенность, Bootstrap, доверительный интервал, угловые данные, статистическая обработка, выборки малого объема.*

CHARACTERISTIC OF SMALL SAMPLES CIRCULAR MEASUREMENT

Y.V. Kuts, L.M. Scherbak, S.V. Shengur

Expanded uncertainty and confidence interval are the main kind of measurement results error representations. An article develops the confidence interval modeling in circular data analyzing task using method which is called "The Bootstrap method".The confidence interval obtaining algorithm and computer simulation results set in. Got results were analyzed in comparison with obtained by classical method, conformity estimates set in.

Keywords: *expanded uncertainty, Bootstrap, confidence interval, circular data, statistical data processing, small samples.*