УДК 006.91:519.2

С.В. Шенгур

Національний авіаційний університет, Київ

ОПРАЦЮВАННЯ ВИБІРОК ВИПАДКОВИХ КУТІВ З АПРІОРНО НЕВІДОМОГО РОЗПОДІЛУ

Запропонована методика оцінювання вибіркових кругових характеристик та їх невизначеності за результатами кутових спостережень в умовах апріорно невідомого закону їх розподілу. Наведено результати експериментальних досліджень оцінювання вибіркових кругових середнього, моди та медіани.

Ключові слова: *випадковий кут, вибіркові кругові характеристики, невизначеність, бутстреп.*

Вступ

Постановка проблеми. Результати вимірювання, у відповідності до вимог національних та міжнародних стандартів [1 – 2], повинні супроводжуватись оцінками точності. Основними вибірковими круговими характеристиками є середнє, мода та медіана. Як оцінки точності використовують розширену невизначеність U або довірчий інтервал $\Delta_{ä}$.

Вибіркове кругове середнє значення $\overline{\theta}$ – характеристика вибірки випадкових кутових даних, що визначає область найбільш імовірного значення випадкового кута (ВК) [3]. Вона визначається як:

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

де
$$C = \frac{1}{n}\sum_{j=1}^{n}\mathbf{cos}\,\theta_j, \quad S = \frac{1}{n}\sum_{j=1}^{n}\sin\theta_j. \quad (2)$$

Вибіркова кругова медіана – медіанному куту відповідає така точка кола, діаметр через яку ділить значення статистики навпіл.

Вибіркова мода – куту відповідає точка кола, в околі якої спостерігається максимальна концентрація значень статистики.

Довірчим інтервалом результатів вимірювання випадкового кута є довжина найменшої дуги кола одиничного радіуса, що включає найбільш достовірні значення цієї статистики [3].

Традиційно розподіл результату вимірювання випадкової величини, ґрунтуючись на результатах попередніх вимірювань, приймають близьким до нормального [3]. За наявності достатньої статистики виконують перевірку на нормальність за одним з відомих критеріїв. Проте традиційний метод має ряд обмежень у випадках, коли:

а) вимірювана величина розподілена за асиметричним законом розподілу щільності ймовірності;

б) вимірювана величина не відповідає описаним законам розподілу;

в) має місце недостатня апріорна інформація про ймовірнісну модель випадкових кутових величин.

Аналіз досліджень і публікацій. За умов невідомого розподілу ймовірності вимірюваних випадкових кутових значень і наявності статистик обмежених об'ємів, оцінку розширеної невизначеності для кругового середнього напряму U_{mean} і довірчої ймовірності $P_{äî â}$ можна визначити на основі використання аналога нерівності Чебишева на колі [3, 4]:

$$P\left\{\left|\mathbf{sin}\frac{\psi(\omega) - \overline{\theta}}{2}\right| > \varepsilon\right\} < \frac{\upsilon}{2\varepsilon^2},\ 0 < \varepsilon < 1,\ 0 < \frac{\upsilon}{2\varepsilon^2} < 1, \quad (3)$$

де υ – кругова дисперсія випадкового кута $\psi(\omega)$, $0 \le \upsilon \le 1$; $\overline{\theta}$ – кругове середнє значення кута, яке обраховується з тригонометричного моменту першого порядку відносно нульового напряму тобто $\overline{\theta} = \arg(\dot{\tau}_1)$. Нерівність (1) виконується при $|\dot{\tau}_1| > 0$. Її можна інтерпретувати таким чином: ймовірність події $\left|\mathbf{sin}\frac{\psi(\omega) - \overline{\theta}}{2}\right| > \varepsilon$ не перевищує величину $\frac{\upsilon}{2\varepsilon^2}$ за довільної функції розподілу випадкового кута $\psi(\omega)$.

Для формування U_{mean} зручнішою є інша форма цієї нерівності:

$$P\left\{\left|\left(\psi(\omega) - \overline{\theta} + \pi\right)\mathbf{mod}2\pi - \pi\right| > 2\mathbf{arcsin}\left(\varepsilon_1\sqrt{\frac{\upsilon}{2}}\right)\right\} < \frac{1}{\varepsilon_1^2},\ 0 < \varepsilon_1 < \sqrt{\frac{2}{\upsilon}}. \quad (4)$$

Із формули (4) згідно з (3) маємо

$$U_{mean}(\upsilon) = 2\mathbf{arcsin}\left(\varepsilon_1\sqrt{\frac{\upsilon}{2}}\right). \quad (5)$$

Довірчий інтервал для кругового середнього напряму формується як:

$$\overline{\theta} \pm 2\arcsin\left(\varepsilon\sqrt{V/2}\right). \quad (6)$$

Обчислення довірчого інтервалу для кругового середнього напряму на основі описаного методу не потребує апріорної інформації про ймовірнісну модель випадкової кутової величини, що дає йому перевагу над традиційним методом. Проте, ряд проведених експериментальних досліджень показав стабільно завищені результати обчислення значень довірчого інтервалу, що є його недоліком.

Останні три десятиліття набуває поширення бутстреп-метод опрацювання малих вибірок [5], основною перевагою якого є можливість статистичного опрацювання результатів вимірювання випадкових величин за невідомого закону їх розподілу та формування довірчого інтервалу, що не залежить від центру розподілу та коефіцієнту покриття.

Загальна задача обробки формулюється наступним чином. Нехай за результатами вимірювання одержано вибірку випадкових величин $X_n = \left\{x_1,...,x_j,...,x_n\right\}$, $x \in (-\infty, \infty)$. Тоді з вхідної вибірки формують ряд обсягу B (B>>n) допоміжних вибірок такого ж обсягу n шляхом випадкового вибору з поверненням значень. Для кожної допоміжної вибірки оцінюють вибіркове середнє та одержують масив його незалежних значень обсягу B, який дозволяє побудувати функцію розподілу середнього значення вхідної вибірки X. За функцією розподілу, обчисливши її параметри, оцінюють математичне сподівання, його довірчий інтервал та стандартне відхилення для X.

Мета роботи – розробити та експериментально дослідити методику оцінювання вибіркових кругових характеристик та їх точності за вибірками випадкових кутів, що належать до апріорно невідомого закону розподілу, із застосуванням бутстреп-методу.

Вирішення задачі

Сутність методу бутстреп для задач оцінювання вибіркових кругових середнього, моди, медіани та їх невизначеності полягає у виконанні наступних етапів:

1. Із вихідної вибірки випадкових кутів $\Theta = \left\{\theta_1,...\theta_j,...\theta_n\right\}$ формується нова вибірка $\Theta^{(b)} = \left\{\theta'_1,...\theta'_j,...\theta'_n\right\}$ шляхом перестановки з поверненням зі збереження обсягу вибірки (рис. 1).

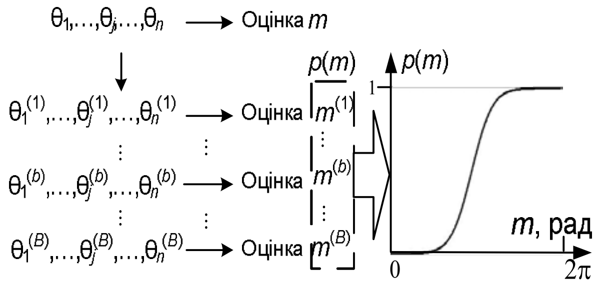

Рис. 1. Зображення бутстреп-процедури

2. Для одержаної вибірки $\Theta^{(b)}$ оцінюється кругове середнє, мода або медіана $m^{(b)}$ (від англ. mean, mode, median).

3. Пункти 1-2 повторюються задану кількість разів B (B не менше 1000), в результаті чого формується так званий «бутстреп»-розподіл $m^{(1)},...,m^{(b)},...,m^{(B)}$ (рис. 2).

4. Середнє значення вибірки $m^{(1)},...,m^{(b)},...,m^{(B)}$ приймається за оцінку ймовірнісної характеристики m для вибірки випадкових кутів Θ, а її симетричні процентилі $m^{(l+1)}$ та $m^{(B-l)}$ – як границі довірчого інтервалу цієї характеристики із заданим рівнем довіри. Значення l розраховують як:

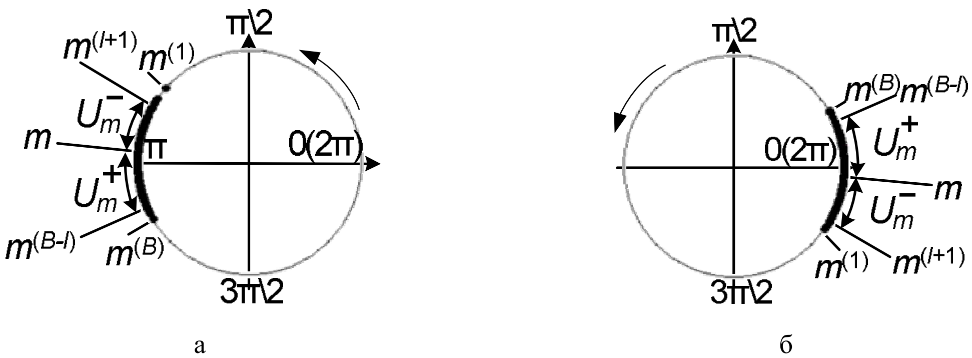

Рис. 2. Відображення “бутстреп”-розподілу на одиничному колі: а – точка розриву шкали не потрапляє в діапазон вибірки; б – точка розриву шкали потрапляє в діапазон вибірки

$$l = \left[\frac{1}{2}B\left(P_{äîâ} - 1\right) + \frac{1}{2}\right], \tag{7}$$

де [] – позначення операції виділення цілої частини числа.

Значення складових розширеної невизначеності U_m^- та U_m^+ [6] для досліджуваної кругової характеристики відповідають довжинам дуг між граничними значеннями вибірок $m^{(l+1)},..,m$ та $m,...,m^{(B-l)}$, які обчислюють за формулою вибіркового кругового розмаху W (довжина найменшої дуги кола одиничного радіусу, яка містить усі значення вибірки та визначається із варіаційного ряду (упорядкованої по величині послідовності вибіркових значень $\theta_1 \le ... \le \theta_n$)):

$$T_j = \theta_{j+1} - \theta_j,\ j = 1,...,n-1,\ T_n = 2\pi - \theta_n + \theta_1,$$

$$W = 2\pi - \mathbf{max}\left\{T_1,...,T_n\right\}. \qquad (8)$$

5. Формують результат вимірювання як

$$\left(\left(m - U_m^-\right)\mathrm{mod}\,2\pi;\ \left(m + U_m^+\right)\mathrm{mod}\,2\pi\right),\ P_{\text{ãî â}}. \qquad (9)$$

Наведений алгоритм має назву непараметричний бутстреп метод. У роботі [7] наведені загальні відомості про цей метод, але особливості його практичної реалізації та результати дослідження його ефективності відсутні.

Цей різновид "розкрутки" даних, отриманих у результаті спостереження, доцільно використовувати у випадку обмеженої апріорної інформації про математичну модель випадкової величини – за умови невідомого закону розподілу ймовірності кутової величини та у випадку малого обсягу вибірки випадкових кутів. Даний спосіб реалізується шляхом генерування ряду псевдовипадкових номерів $i_1,...,i_n$ з використанням рівномірного розподілу $U[0,1]$:

$$\begin{aligned} i_1 &= \left[nu_1 + 1\right]^+, \\ &\vdots \\ i_n &= \left[nu_J + 1\right]^+, \end{aligned} \qquad (10)$$

де u_j – значення випадкової величини з розподілом U, а квадратними дужками позначена операція виділення цілої частини числа, та формування вибірки:

$$\Theta^{(1)} = \left\{\theta_{i1},...,\theta_{ij},...,\theta_{in}\right\} = \left\{\theta_1',...,\theta_j',...,\theta_n'\right\}. \qquad (11)$$

Параметричний метод бутстреп застосовують до даних вимірювання, розподілених за відомим законом розподілу ймовірності.

Тоді для вхідної вибірки $\Theta = \left\{\theta_1,...,\theta_j,...,\theta_n\right\}$ обчислюють оцінки параметрів закону розподілу ймовірності відповідних випадкових кутів, та генерують нову вибірку $\Theta^{(1)} = \left\{\theta_1',...,\theta_j',...,\theta_n'\right\}$ як реалізацію за цим законом.

Методика оцінювання вибіркового кругового середнього та його розширеної невизначеності для вибірки випадкових кутових величин $\Theta = \left\{\theta_1,...,\theta_j,...,\theta_n\right\}$ полягає у наступному.

1. Для отриманої у результаті спостереження вибірки випадкових кутових величин $\Theta = \left\{\theta_1,...,\theta_j,...,\theta_n\right\}$ (рис. 3) обчислюється 2×1 вектор $\mathbf{Z_0} = z$, 2×2 матриця $\mathbf{U_0} = \mathbf{u}$ та 2×2 матриця $\mathbf{V_0} = \mathbf{v}$ відповідно до виразів [7]:

$$\mathbf{z} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}, \qquad (12)$$

де $z_1 = \sum_{j=1}^{n} x_j / n$, $z_2 = \sum_{j=1}^{n} y_j / n$, $x_j = \cos\theta_j$, $y_j = \sin\theta_j$, $j = 1,...,n$.

$$\mathbf{u} = \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix}, \qquad (13)$$

де $u_{11} = \sum_{j=1}^{n} (x_j - z_1)^2 / n$, $u_{22} = \sum_{j=1}^{n} (y_j - z_2)^2 / n$,

$$u_{12} = u_{21} = \sum_{j=1}^{n} (x_j - z_1)(y_j - z_2) / n.$$

$$\mathbf{v} = \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix}, \qquad (14)$$

де $v_{11} = (\beta^2 t_1 + t_2) / (1 + \beta^2)$,

$$v_{22} = (t_1 + \beta^2 t_2) / (1 + \beta^2),$$

$$v_{12} = v_{21} = \beta(t_1 - t_2) / (1 + \beta^2)^{\frac{1}{2}},$$

$$t_1 = (\beta^2 u_{11} + 2\beta u_{12} + u_{22})^{\frac{1}{2}} / (1 + \beta^2)^{\frac{1}{2}},$$

$$t_2 = (u_{11} - 2\beta u_{12} + \beta^2 u_{22})^{\frac{1}{2}} / (1 + \beta^2)^{\frac{1}{2}},$$

$$\beta = (u_{11} - u_{22}) / (2u_{12}) - \left[(u_{11} - u_{22})^2 / (4u_{12}^2) + 1\right]^{\frac{1}{2}}.$$

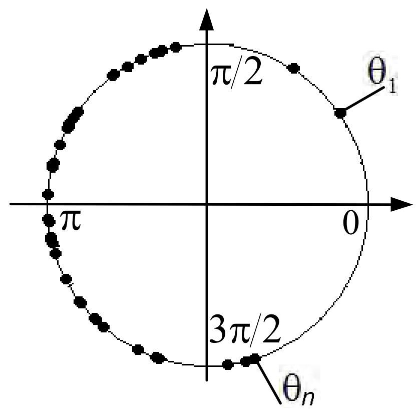

Рис. 3. Графічне відображення вибірки випадкових кутів на одиничному колі

2. На другому етапі формується ряд нових вибірок, "розкручених" із вхідної вибірки $\Theta = \left\{\theta_1,...,\theta_j,...,\theta_n\right\}$ випадкових кутів. В залежності від наявної апріорної інформації для цього обирають непараметричний або параметричний різновид бутстреп-методу.

У випадку, коли розподіл випадкової кутової величини невідомий, проте відомо, що він симетричний, доцільним є додаткове симетрування даних, яке полягає у наступному:

– генерують новий ряд псевдовипадкових номерів $u_1',...,u_n'$ з рівномірного розподілу $U[0,1]$;

– для кожного j, $j=1,...,n$ при $u'_j<0,5$ значення θ_{ij} з (11) замінюють на $2\overline{\theta}-\theta_{ij}$.

3. Для сформованої вибірки $\Theta^{(1)}=\{\theta'_1,...,\theta'_j,...,\theta'_n\}$ обчислюють вектор $\mathbf{Z_1}=z$ з (12) та 2×2 матрицю $\mathbf{W_1}=\mathbf{w}$:

$$\mathbf{w}=\begin{pmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{pmatrix}, \qquad (15)$$

де $w_{11}=(\beta^2 t_1+t_2)/(1+\beta^2)$,

$$w_{22}=(t_1+\beta^2 t_2)/(1+\beta^2),$$

$$w_{12}=w_{21}=\beta(t_1-t_2)/(1+\beta^2)^{\frac{1}{2}},$$

$$t_1=(1+\beta^2)^{\frac{1}{2}}/(\beta^2 u_{11}+2\beta u_{12}+u_{22})^{\frac{1}{2}},$$

$$t_2=(1+\beta^2)^{\frac{1}{2}}/(u_{11}-2\beta u_{12}+\beta^2 u_{22})^{\frac{1}{2}},$$

$$\beta=(u_{11}-u_{22})/(2u_{12})-\left[(u_{11}-u_{22})^2/(4u_{12}^2)+1\right]^{\frac{1}{2}}.$$

4. Для вибірки $\Theta^{(1)}=\{\theta'_1,...\theta'_j,...\theta'_n\}$ обчислюють значення оцінки кругового середнього напряму $\overline{\theta}^{(1)}$. Для цього обчислюють матрицю

$$\begin{pmatrix} c_B \\ s_B \end{pmatrix}=\mathbf{z_0}+\gamma_0 w_{\mathbf{B}}(z_{\mathbf{B}}-z_{\mathbf{0}}), \qquad (16)$$

де величини $\mathbf{z_0}$ та γ_0 обчислені для вхідної вибірки Θ, а $\mathbf{z_B}$ та $\mathbf{w_B}$ – для нової сформованої вибірки $\Theta^{(b)}$ (наприклад, для першої вибірки $\Theta^{(1)}$ $\mathbf{z_B}=\mathbf{z_1}$ та $\mathbf{w_B}=\mathbf{w_1}$).

Елементи C_B та S_B матриці (17) отримують з перетворення

$$\begin{pmatrix} C_B \\ S_B \end{pmatrix}=(c_B^2+s_B^2)^{-\frac{1}{2}}\begin{pmatrix} c_B \\ s_B \end{pmatrix} \qquad (18)$$

та використовують як синусну та косинусну складові виразу (2) для обчислення кругового середнього напряму $\overline{\theta}^{(1)}$.

5. Пункти 2 – 4 повторюють задану кількість разів B. Число B доцільно обирати не менше 1000. Обчислені оцінки кругових середніх значень для усіх вибірок $\Theta^{(b)}=\{\theta'_1,...,\theta'_j,...,\theta'_n\}$, $b=\overline{1,B}$, формують ряд $\overline{\theta}^{(1)},...,\overline{\theta}^{(b)},...,\overline{\theta}^{(B)}$, що будемо називати "бутстреп"-розподілом (рис. 2).

6. Для сформованої вибірки $\overline{\theta}^{(1)},...,\overline{\theta}^{(b)},...,\overline{\theta}^{(B)}$ оцінюють кругове середнє значення за формулою (1). Обчислене значення відповідає вибірковому круговому середньому напрямку $\overline{\theta}$ вхідної вибірки $\Theta=\{\theta_1,...,\theta_j,...,\theta_n\}$.

7. З вибірки $\overline{\theta}^{(1)},...,\overline{\theta}^{(b)},...,\overline{\theta}^{(B)}$ формують довірчий інтервал для кругового середнього напрямку вибірки випадкових кутів $\Theta=\{\theta_1,...,\theta_j,...,\theta_n\}$. Використовуючи формулу (8) для обчислення кругового розмаху статистики різниць, визначають крайні елементи вибірки $\overline{\theta}^{(1)},...,\overline{\theta}^{(b)},...,\overline{\theta}^{(B)}$, тобто елементи, що утворюють дугу найменшого розміру, яка містить всі значення вибірки, відповідно $\overline{\theta}^{(1)}$ і $\overline{\theta}^{(B)}$, та, задавшись значенням довірчої ймовірності $P_{äî â}$, розраховують номери елементів за (7).

На рис. 2 проілюстровано можливі варіанти розташування вибірки $\overline{\theta}^{(1)},...,\overline{\theta}^{(b)},...,\overline{\theta}^{(B)}$ на одиничному колі: а – вибірка розташована так, що точка розриву шкали $0\,(2\pi)$ не потрапляє в діапазон вибірки; б – точка розриву шкали $0\,(2\pi)$ потрапляє в діапазон вибірки.

Для випадку, представленого на рис. 2, а, границями довірчого інтервалу для кругового середнього напряму є значення кутів $\overline{\theta}^{(l+1)}$ та $\overline{\theta}^{(B-l)}$, представленої у вигляді варіаційного ряду $\overline{\theta}^{(1)}\le...\le\overline{\theta}^{(B)}$ вибірки $\overline{\theta}^{(1)},...,\overline{\theta}^{(b)},...,\overline{\theta}^{(B)}$.

Граничні елементи $\overline{\theta}^{(1)}$ і $\overline{\theta}^{(B)}$ представленої у вигляді варіаційного ряду вибірки $\overline{\theta}^{(1)},...,\overline{\theta}^{(b)},...,\overline{\theta}^{(B)}$, в діапазон якої потрапляє точка розриву шкали $0\,(2\pi)$, розташовані близько і не охоплюють значення вибірки дугою найменшої довжини.

Для визначення елементів, що відповідають границям довірчого інтервалу для кругового середнього напрямку, необхідно визначити порядкові номери, яким відповідають крайні елементи вибірки, та визначити елементи, що знаходяться на l+1 позицій відповідно у напрямі наростання значень кутів шкали від елементу $\overline{\theta}^{(1)}$ та у напрямі їх зменшення від елементу $\overline{\theta}^{(B)}$. При цьому необхідно враховувати можливість потрапляння точки розриву шкали $0\,(2\pi)$ в інтервали $\left[\overline{\theta}^{(1)},\overline{\theta}^{(1+(l+1))}\right]$ та $\left[\overline{\theta}^{(B-(l+1))},\overline{\theta}^{(B)}\right]$.

8. Значення складових розширеної невизначеності кругового середнього U_{mean}^{-} та U_{mean}^{+} обчислюють як довжини дуг, що охоплюють вибірки $\overline{\theta}^{(l+1)},...,\overline{\theta}^{(b)},...,\overline{\theta}$ та $\overline{\theta},...,\overline{\theta}^{(b)},...,\overline{\theta}^{(B-l)}$. Значення довжин дуг обчислюють за формулою вибіркового кругового розмаху (8).

9. Формують результат вимірювання як:

$$\left(\left(\bar{\theta}-U_{mean}^{-}\right)\bmod 2\pi;\ \left(\bar{\theta}+U_{mean}^{+}\right)\bmod 2\pi\right),\ P_{äîâ}.$$

Загальновживаний принцип обчислення оцінки моди [4] визначений на групованих даних. Проте для вибірок обсягом $n <\sim 40$ значень має місце недостатній обсяг статистики для вибору правильного способу групування даних вимірювання. З метою зменшення методичної похибки під час розбиття даних спостереження обмеженого обсягу на інтервали запропоновано коригувати результати групування із застосуванням наступної методики.

1. Для досліджуваної вибірки ВК обсягом n обчислюють значення вибіркового кругового розмаху W за (8).

2. Визначають початковий крок розбиття $\Delta\theta_1$ найменшої довжини дуги W, що містить дану вибірку як $\Delta\theta_1 = \dfrac{W}{n}$.

3. Збільшують крок розбиття $\Delta\theta_2 = \dfrac{W}{n-1}$; $\Delta\theta_3 = \dfrac{W}{n-2}$; $\Delta\theta_4 = \dfrac{W}{n-3}$ до тих пір, поки не буде отримано нерівномірний розподіл даних по інтервалам.

При обчисленні медіани та її довірчого інтервалу доцільно враховувати поправку у випадку, коли відомо, що вибірка ВК належить до симетричного розподілу [7].

Для цього кожну "розкручену" вибірку $\Theta^{(b)} = \left\{\theta'_1,...,\theta'_j,...,\theta'_n\right\}$ ВК складають з додатково згенерованою вибіркою, що належить до намотаного нормального розподілу з параметрами $\mu = \bar{\theta}^{(b)}$ (середнє значення) та $\sigma = 1/\sqrt{n}$ (СКВ).

Результати експериментальних досліджень

Обчислювальний експеримент з оцінювання вибіркових кругових середнього, моди, медіани та їх невизначеності проведено за наступною методикою.

1. Формування симетричної (асиметричної) вибірки обсягом N=100000 значень випадкових кутів з діапазону $[0, 2\pi)$.

2. Формування з вибірки п.1 вибірок обсягу $n = 10 \div 100$ значень.

3. Обчислення усереднених значень вибіркового кругового середнього, моди та медіани за результатами 1000 експериментів для кожного обсягу вибірки ВК з використанням традиційного методу та методу бутстреп.

4. Обчислення значень границь довірчих інтервалів для кругового середнього напряму з використанням методів бутстреп, Чебишева та традиційного методу для гауссівських розподілів.

5. Проведення порівняльного аналізу методів (рис. 4).

Для кожного обсягу вибірки з діапазону $n = 10 \div 100$ значень експеримент повторено 1000 разів для перевірки точності (правильності та прецизійності) оцінок, як основних показників якості результату вимірювання за ДСТУ ГОСТ ИСО 5725-1:2005.

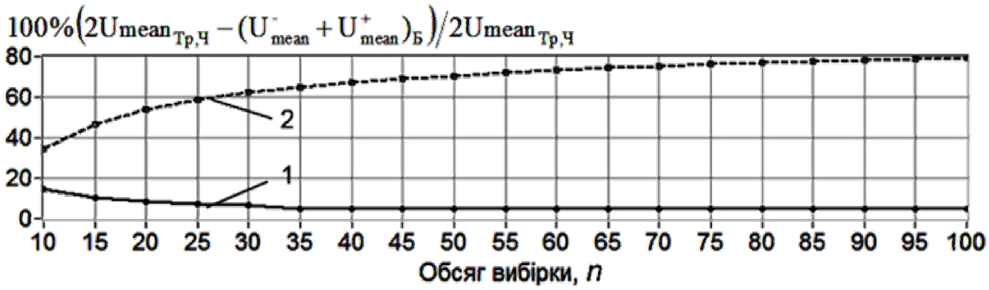

Рис. 4. Графічне відображення зменшення довжини довірчого інтервалу, обчисленого бутстреп-методом відносно традиційного методу (крива 1) та методу Чебишева (крива 2)

Правильність оцінювалась як відхилення усередненого значення одержаної оцінки від істинного $\hat{\theta}_{з\tilde{n}\grave{o}}$ (умовно істинного – у випадку складених законів розподілу генеральної сукупності) значення оцінки кругової характеристики, що процентному співвідношенні має вигляд:

$$\delta_{Á} = \frac{\hat{\theta}_{з\tilde{n}\grave{o}} - \hat{\theta}_{Á}}{\hat{\theta}_{з\tilde{n}\grave{o}}} \cdot 100\% \text{ та } \delta_{Á} = \frac{\hat{\theta}_{з\tilde{n}\grave{o}} - \hat{\theta}_{Á}}{\hat{\theta}_{з\tilde{n}\grave{o}}} \cdot 100\%. \quad (19)$$

Прецизійність оцінювалась порівнянням усереднених значень стандартних відхилень $\hat{\sigma}_{Á}$ та $\hat{\sigma}_{\grave{O}\eth}$ оцінок вибіркових кругових характеристик.

Для тестування методики як симетричний обраний гауссівський намотаний розподіл ($p(\theta) = 1/2\pi\left(1 + 2\sum\limits_{i=1}^{\infty}\rho^{i^2}\mathbf{cos}\,i(\theta-\mu)\right)$, $0 \le \theta < 2\pi$, $0 \le \mu < 2\pi$, $0 \le \rho \le 1$, де μ – круговий середній напрям, ρ – довжина результуючого вектора [8]) з параметрами $\bar{\theta} = \pi, \rho = 0{,}88\ (\sigma = 0{,}5)$, а як асиметричний – комбінація трьох розподілів Лагневіна-

Фішера ($p(\theta) = C\sin\theta\exp\{k\cos(\theta)\}$, $0 \le \theta < \pi$, де C та $k > 0$ – параметри масштабу) на колі з параметрами: I – k=7,5, c=0,3, n=40000; II – k=5,5, c=0,3, n=20000; III – k=2, c=0,3, n=40000 [9].

Результати досліджень оцінок середнього, моди та медіани зведені в табл. 1.

Таблиця 1

Результати досліджень

	n=10…20	n=20…100
Середне	$\delta_{\text{Тр}} = \delta_{\text{Б}} < 0{,}5\%$	
Середне	$\hat{\sigma}_{\text{Òð}} = \hat{\sigma}_{\text{Á}}$	
Медіана	$\delta_{\text{Тр}} = \delta_{\text{Б}} < 1\%$	
Медіана	$\dfrac{\hat{\sigma}_{\text{Òð}} - \hat{\sigma}_{\text{Á}}}{\hat{\sigma}_{\text{Á}}} \cdot 100,\% = 7...10$	$\dfrac{\hat{\sigma}_{\text{Òð}} - \hat{\sigma}_{\text{Á}}}{\hat{\sigma}_{\text{Á}}} \cdot 100,\% = 5...7$
Мода		$\delta_{\text{Тр}} = \delta_{\text{Б}} < 0{,}5\%$
Мода		$\dfrac{\hat{\sigma}_{\text{Òð}} - \hat{\sigma}_{\text{Á}}}{\hat{\sigma}_{\text{Á}}} \cdot 100,\%$ = 20-32 (Асим. розп.) = 35-41 (Сим. розп.)

Для оцінки ефективності методу бутстреп, порівняно з традиційним та методом використання аналога нерівності Чебишева на колі, розраховане процентне співвідношення усереднених за результатами 1000 експериментів довжин довірчих інтервалів, обчислених за вибірками ВК обсягом від 10 до 100 значень. Отримані результати виведено на графік, наведений на рис. 4, де у процентному співвідношенні виражено зменшення довжини довірчого інтервалу, обчисленого бутстреп-методом відносно традиційного методу (крива 1) та «методу Чебишева» (крива 2).

Висновки

Розроблено методику оцінювання вибіркових кругових середнього, моди, медіани та їх невизначеності на основі застосування бутстреп-методу. Проведено експериментальне дослідження для вибірок випадкових кутів обсягом 10…100 значень, що показало покращення прецизійності оцінювання вибіркової кругової моди на 20 – 40%, медіани – на 5 – 10%. Зменшення довжини довірчого інтервалу, обчисленого методом бутстреп відносно традиційного, знаходиться у межах від 7% до 18%; відносно методу Чебишева – 35 – 80%.

Список літератури

1. ДСТУ 2681-94 Метрологія. Терміни та визначення: Чин. з 1995-01-01. – К.: Держстандарт України, 1994. – 68 с.

2. Guide to the expression of uncertainty in measurement (GUM): First edition. – ISO, 1993. – 101 p.

3. Куц Ю.В. Статистична фазометрія / Ю.В. Куц, Л.М. Щербак. – В.: Тернопіль, 2009. – 383 с.

4. Mardia K.V. Statistics of directional data / K.V. Mardia, P.E. Jupp. – London: Academic Press Inc., 1972. – 415 p.

5. Efron B. An introduction to the bootstrap / B. Efron, R. Tibshirani. – Boca Raton, FL: Chapman&Hall/CRC, 1993. – 436 p.

6. Куц Ю.В. Подання результату кутових вимірювань в концепції невизначеності / Ю.В. Куц, С.В. Шенгур // Системи обробки інформації. – Х.: ХУПС, 2013. – Вип. 3(110). – С. 97-100.

7. Fisher N.I. Statistical analysis of circular data. / N.I. Fisher. – Cambridge: Cambridge University Press, 2000. – 277 p.

8. Куц Ю.В. Віртуальний прилад для генерування вибірок випадкових кутів / Ю.В. Куц, С.В. Шенгур // Електроніка та системи управління. – 2011. – Вип. 1 (27). – С. 35-41.

9. Jammalamadaka S. Rao. Topics in circular statistics / S. Rao Jammalamadaka, A. SenGupta. – Singapore: World Scientific Publishing Co. Pte. Ltd., 2001 – 322 p.

Надійшла до редколегії 27.08.2013

Рецензент: д-р техн. наук, проф. Л.М. Щербак, Національний авіаційний університет, Київ.

ОБРАБОТКА ВЫБОРОК СЛУЧАЙНЫХ УГЛОВ ИЗ АПРИОРНО НЕИЗВЕСТНОГО РАСПРЕДЕЛЕНИЯ

С.В. Шенгур

Предложена методика оценивания выборочных круговых характеристик и их неопределенности по результатам угловых наблюдений в условиях априорно неизвестного закона их распределения. Приведены результаты экспериментальных исследований оценивания выборочных круговых среднего, моды и медианы.

***Ключевые слова**: случайный угол, выборочные круговые характеристики, неопределенность, бутстреп.*

PRIORY UNKNOWN DISTRIBUTION RANDOM CIRCLE SAMPLE PROCESSING

S.V. Shengur

The technique of sample circular characteristics and their uncertainty estimation from priori unknown distribution is proposed. Experimental researches results of sample circular mean, mode and median estimation are shown.

***Keywords:** Random circle, sample circular characteristics, uncertainty, bootstrap.*