

сениця»-SSA і Бокса-Дженкінса. Запропоновано модель авто-регресії — спектрально інтегрованого ковзного середнього, що реалізує трендовий підхід, який полягає в моделюванні процесу як відхилення фактичних значень відносно трендової складової, в ролі якої виступає лінійна рекурентна формула методу «Гусениця»-SSA.

Ключові слова: прогнозування часових рядів, структурна ідентифікація, метод Бокса-Дженкінса, метод «Гусениця»-SSA.

Щелкалін Віталій Николаевич, ассистент, кафедра прикладной математики, Харьковский национальный университет радиоэлектроники, Украина, e-mail: vitalii.shchelkalin@gmail.com.

Щелкалін Віталій Миколайович, ассистент, кафедра прикладной математики, Харьковский национальный университет радиоэлектроники, Украина.

Shchelkalin Vitalii, Kharkiv National University of Radioelectronics, Ukraine, e-mail: vitalii.shchelkalin@gmail.com

УДК 006.91:519.2

Шенгур С. В.

СТАТИСТИЧЕСКАЯ ОБРАБОТКА СВЕРХМАЛЫХ ВЫБОРОК СЛУЧАЙНЫХ УГЛОВ

Статья посвящена статистическому анализу выборок случайных углов очень малого объема. Основной целью работы является разработка и экспериментальное исследование новой методики статистической обработки угловых величин, полученных из различных источников и сфер деятельности. В основу положен метод бутстреп. Разработанная методика позволяет повысить точность оценивания выборочного кругового среднего и его доверительного интервала. Приведены результаты экспериментальных исследований.

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

1. Введение

Результаты измерений в соответствии с национальными и международными стандартами [1, 2] должны сопровождаться показателями точности. В качестве таких показателей чаще всего используют расширенную неопределенность U либо доверительный интервал Δ_d .

Распределение результата измерения случайной величины, основываясь на результатах предварительных измерений, как правило, принимают приближенным к гауссовскому [3]. При наличии достаточной статистики выполняют проверку на гауссовость по одному из известных критериев — Колмогорова, хи-квадрат.

Однако такие критерии предусматривают построение по данным наблюдения гистограммы, и, следовательно, имеют ограничения при применении к выборкам менее 35 значений. И даже при таком объеме данных имеют нестабильный результат. На практике не редко встречаются случаи, когда выборка, полученная по результатам наблюдений, не превышает 10 значений. Это обусловлено, прежде всего, высокой стоимостью проведения эксперимента либо невозможностью его повторного проведения. Общепринятого критерия сверхмалой выборки нет. Будем использовать этот термин при $n \leq 9$ [4]. Обработка сверхмалых выборок требует специфического подхода, за исключением ситуаций, когда закон распределения случайного угла известен.

Традиционный метод расчета и представления результата измерений случайных угловых данных предусматривает получение по результатам измерений выборки углов определенного объема, определение выборочных круговых среднего и стандартного отклонения, оценки неопределен-

ности измерений. Оценку неопределенности рассчитывают как симметричный относительно выборочного кругового среднего интервал значений углов, величина которого формируется как удвоенное произведение выборочного кругового стандартного отклонения и фактор покрытия для заданного уровня доверия и известного закона распределения случайной угловой величины. Применение традиционного метода к сверхмалым выборкам, закон распределения которых априорно неизвестен, необоснованно и может привести к грубым погрешностям.

2. Анализ литературных данных и постановка проблемы

В работе [5] приведена и экспериментально апробирована методика оценивания основных выборочных круговых характеристик — среднего, моды, медианы, по выборкам объемом до 100 значений.

Цель работы — разработать и экспериментально исследовать методику оценивания точечных и интервальных характеристик случайных угловых величин по выборкам сверхмалого объема, принадлежащим к априорно неизвестному закону распределения.

3. Результаты исследований

Задачу оценивания результатов измерения случайных углов по выборкам сверхмалого объема предложено решать на основе бутстреп метода [4–6]. Бутстреп метод статистической обработки данных известен с 70-х годов XX века. Их преимущество состоит в увеличении статистики без увеличения входных данных.

Методика оценивания доверительного интервала для выборочного кругового среднего (ВКС) значения [7–10] выборки $SU \Theta = \{\theta_1, \dots, \theta_j, \dots, \theta_n\}$ объемом $n = 3 \dots 9$ предполагает применение к исследуемой выборке технологии так называемого двойного бутстреп (double bootstrap) [4]. Такая технология позволяет получить улучшенные по сравнению с [5] результаты, учитывая взаимозависимость вероятностного наполнения входной выборки и ряда полученных из нее выборок. Сформированный результат измерения имеет вид:

$$((\bar{\theta} - U_{\text{mean}}^-) \bmod 2\pi; (\bar{\theta} + U_{\text{mean}}^+) \bmod 2\pi), P_{\text{дов}}, \quad (1)$$

где $\bar{\theta}$ — рассчитанная оценка кругового среднего для входной выборки, $U_{\text{mean}}^-, U_{\text{mean}}^+$ — составляющие расширенной неопределенности кругового среднего, рассчитанные как длины дуг для $100(P_{\text{дов}})\%$ доверительного интервала.

Для проведения компьютерных экспериментов по оцениванию выборочного кругового среднего направления, а так же его показателей точности по сверхмалым выборкам случайных углов, применена методика, описанная в [5]. Из совокупности объемом не менее 10000 значений сформированы выборки объемом от 3 до 9 значений; для каждой выборки рассчитано значение выборочного кругового среднего и его неопределенности (доверительного интервала) методами: традиционным, с использованием аналога неравенства Чебышева на круге [11] бутстреп. В качестве генеральной использованы симметричная и асимметричная совокупности. Для каждого объема выборки из диапазона $n = 3 \dots 9$ значений эксперимент повторен 1000 раз для проверки точности (правильности и прецизионности) оценок, как основных показателей качества результата измерения по ДСТУ ГОСТ ИСО 5725-1:2005. Рассчитанная оценка сравнивалась с (условно) истинной, заданной для генеральной совокупности.

Результаты моделирования показали, что правильность и прецизионность оценок ВКС для двух исследуемых методов (традиционного и бутстреп) в пределах погрешности статистического эксперимента одинаковы как для симметричной, так и для асимметричной совокупностей на всех значениях объема выборки из диапазона $n = 3 \dots 9$.

Результаты исследования границ доверительного интервала для ВКС отображены на рис. 1.

Анализируя графики, можно сделать вывод о том, что бутстреп метод учитывает асимметричность генеральной совокупности, к которой принадлежит выборка СУ и обеспечивает наименьший доверительный интервал вокруг среднего направления сравнительно с другими методами. На выборках объемом 3...7 значений метод Чебышева обеспечил лучший результат сравнительно с традиционным. Учитывая эту особенность, можно рекомендовать к использованию метод Чебышева для обработки сверхмалых выборок СУ объемом до 7 значений в случаях, когда сформулированы условия ограничения затрат продуктивности вычислительных ресурсов. График процентного соотношения длин доверительного интервала, рассчитанного методом бутстреп сравнительно с традиционным:

$$100\% (2U_{\text{meanTr}} - (U_{\text{mean}}^- + U_{\text{mean}}^+)_{\text{Б}}) / 2U_{\text{meanTr}} \quad (\text{кривая 1})$$

и методом Чебышева:

$$100\% (2U_{\text{meanЧ}} - (U_{\text{mean}}^- + U_{\text{mean}}^+)_{\text{Б}}) / 2U_{\text{meanЧ}} \quad (\text{кривая 2})$$

приведен на рис. 2.

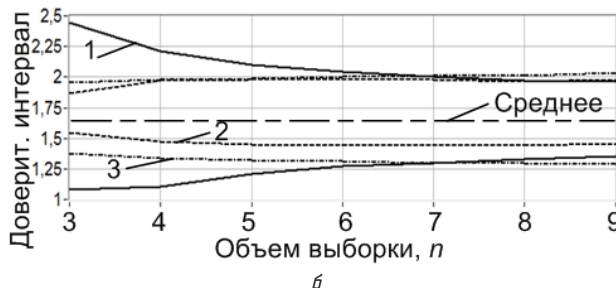
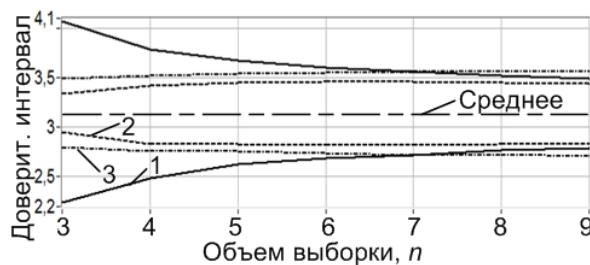


Рис. 1. Графики исследования границ доверительного интервала для ВКС в зависимости от объема выборки СУ; генеральная совокупность принадлежит к: а — симметричному распределению; б — асимметричному распределению; 1 — традиционный метод, 2 — бутстреп, 3 — метод Чебышева

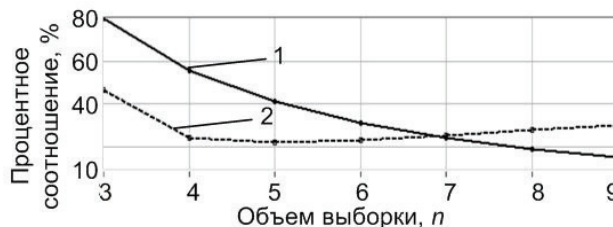


Рис. 2. График исследования процентного соотношения длин доверительного интервала для ВКС в зависимости от объема выборки СУ, рассчитанного методом бутстреп: 1 — сравнительно с традиционным методом; 2 — сравнительно с методом Чебышева

Из графика видно, что использование метода бутстреп для оценивания доверительного интервала для кругового среднего направления обеспечивает уменьшение его значения на 12–80 % сравнительно с традиционным методом и на 20–47 % сравнительно с методом Чебышева в диапазоне значений объема выборки $n = 3 \dots 9$.

4. Выводы

В работе предложена методика обработки сверхмалых выборок случайных углов. Показаны результаты экспериментальных исследований. Достигнуто повышение точности интервального оценивания выборочного кругового среднего, полученного по выборкам сверхмалого объема (3–9 значений). Повышение точности, полученное предложенным методом, выражено длиной доверительного интервала и составляет: относительно метода традиционного метода 12–80 %, относительно метода Чебышева — 20–47 %.

Литература

1. ДСТУ 2681-94. Метрологія. Терміни та визначення [Текст]. – Чин. з 1995-01-01. – К.: Держстандарт України, 1994. – 68 с.
2. Guide to the expression of uncertainty in measurement (GUM) [Text]. – First edition. – ISO, 1993. – 101 p.
3. Куц, Ю. В. Статистична фазометрія [Текст] / Ю. В. Куц, Л. М. Щербак. – В.: Тернопіль, 2009. – 383 с.
4. Fisher, N. I. Statistical analysis of circular data [Text] / N. I. Fisher. – Cambridge: Cambridge University Press, 2000. – 277 p.
5. Шенгур, С. В. Опрацювання вибірок випадкових кутів з апіорно невідомого розподілу [Текст] / С. В. Шенгур // Системи обробки інформації. – 2013. – Вип. 8. – С. 130–135.
6. Efron, B. An introduction to the bootstrap [Text] / B. Efron, R. Tibshirani. – Boca Raton, FL: Chapman&Hall/CRC, 1993. – 436 p.
7. Куц, Ю. В. Подання результату кутових вимірювань в концепції невизначеності [Текст] / Ю. В. Куц, С. В. Шенгур // Системи обробки інформації. – 2013. – Вип. 3. – С. 97–100.
8. Kuts, Y. V. Circular data point and interval evaluation [Text] / Y. V. Kuts, S. V. Shengur, A. V. Dergunov // «Aviation in the XXI-st century 2012»: proc. VI World congress. – K.: NAU, 2012. – P. 1.6.5–1.6.7.
9. Kuts, Y. V. Circular measurement data modeling and statistical processing in LabView [Text] / Y. V. Kuts, S. V. Shengur, L. M. Shcherbak // «MRRS-2011»: proceedings III International symposium: – K.: NAU, 2011. – P. 317–320.
10. Jammalamadaka, S. Rao. Topics in circular statistics [Text] / S. Rao Jammalamadaka, A. SenGupta. – Singapore: World Scientific Publishing Co. Pte. Ltd., 2001 – 322 p.
11. Mardia, K. V. Statistics of directional data [Text] / K. V. Mardia, P. E. Jupp. – London: Academic Press Inc., 1972. – 415 p.

СТАТИСТИЧНЕ ОПРАЦЮВАННЯ НАДМАЛИХ ВИБІРОК ВИПАДКОВИХ КУТІВ

Стаття присвячена статистичному аналізу вибірок випадкових кутів дуже малого обсягу. Основною метою роботи є розробка та експериментальне дослідження нової методики статистичного опрацювання кутових величин, одержаних із різних джерел, а також сфер діяльності. За основу покладений метод бутстреп. Розроблена методика дозволяє підвищити точність оцінювання вибіркового кругового середнього значення та його довірчого інтервалу. Наведені результати експериментальних досліджень.

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

Шенгур Светлана Виталиевна, кандидат технических наук, доцент, кафедра информационно-измерительных систем, Национальный авиационный университет, Киев, Украина, e-mail: shengur.s@gmail.com.

Шенгур Світлана Віталіївна, кандидат технічних наук, доцент, кафедра інформаційно-вимірювальних систем, Національний авіаційний університет, Київ, Україна.

Shengur Svitlana, National Aviation University, Kyiv, Ukraine, e-mail: shengur.s@gmail.com