

ДОСЛІДЖЕННЯ ОДНОРІДНОСТІ ПСЕВДОВИПАДКОВИХ ВИБІРОК КРИТЕРІЯМИ АНДЕРСОНА ТА ЙОГО АНАЛОГАМИ

Клименко О. Д.

Дніпровський національний університет імені Олеся Гончара, аспірант, Україна

Анотація. У роботі детально розглянуто критерії статистичної однорідності вибірок експериментальних вимірювань у випадках, коли функції розподілу ймовірностей апріорно невідомі. Зокрема, проаналізовано критерій Андерсона, який широко застосовується в практиці для оцінювання однорідності даних на основі рангових розподілів. Запропоновано комбінований критерій Буша-Вінда, який базується на логістичних випадкових величинах і за своїми властивостями є аналогом відомих статистик Ван-дер-Вардена та Клотца. Наведено відповідні математичні моделі, алгоритми для обчислень, а також критичні значення, які застосовуються для перевірки статистичних гіпотез щодо однорідності вибірок. Проведені експериментальні дослідження з використанням різної кількості вимірювань (довжина вибірки) та з різними видами закону розподілу ймовірності (логістичний, експоненціальний). Отримані результати дослідження підтвердили високу ефективність використаних критеріїв у прикладних задачах аналізу даних.

Ключові слова: статистика, критерій, алгоритм, математична модель, розподіл ймовірності, однорідність.

Дослідження статистичної однорідності вибірок є важливим завданням у математичній статистиці, оскільки воно дозволяє оцінювати достовірність експериментальних вимірювань та визначати, чи належать вони до однієї генеральної сукупності [1-6]. Розглянемо підходи щодо перевірки статистичної однорідності вибірок: критерій Андерсона та комбінований критерій Буша-Вінда. Кожен із цих критеріїв має свої особливості, сфери застосування та ефективність залежно від специфіки вибірок, що аналізуються. Критерій Андерсона широко використовується для оцінки статистичної однорідності вибірок, отриманих у ході експериментальних досліджень, коли функція розподілу ймовірностей невідома. Основна ідея цього критерію полягає в аналізі рангових характеристик вибірок та порівнянні їх у рамках єдиної сукупності даних [4]. Методологія цього підходу базується на порівнянні упорядкованих значень вибірок та їхніх об'єднаних рангових характеристик. Завдяки цьому критерію можна робити висновки про рівномірний розподіл

даних у вибірках та їхню взаємну відповідність. Важливо зазначити, що ефективність цього критерію значно зростає при достатньо великій кількості спостережень у вибірках. Запропонований налог критерію Андерсона ґрунтується на порівнянні власних рангів двох вибірок [1, 4]. Цей критерій особливо ефективний у випадках, коли потрібно оцінити взаємозв'язок між даними, отриманими одним і тим самим вимірювальним приладом. Завдяки використанню рангових характеристик можна уникнути впливу викидів та локальних аномалій у вибірках, що робить цей метод більш стійким до нерівномірного розподілу даних. Даний критерій має високу інформативність при великих обсягах вибірок, оскільки випадкові відхилення у розподілі даних стають менш значущими. Запропоновано комбінований критерій Буша-Вінда базується на використанні логістичних випадкових величин та є аналогом статистик Ван-дер-Вардена і Клотца. Він дозволяє оцінювати вибірки з невідомими статистичними закономірностями, що робить його універсальним засобом для аналізу експериментальних вимірювань. Однією з ключових переваг цього критерію є його здатність враховувати як відхилення у середніх значеннях вибірок, так і варіації у їхній дисперсії. Такий підхід до обробки випадкових величин дозволяє отримати більш точні результати при аналізі вибірок, які можуть мати суттєві відмінності у своїх параметрах. Критерій Буша-Вінда використовується у випадках, коли стандартні методи перевірки статистичної однорідності дають неоднозначні результати або не враховують всіх можливих відхилень у вибірках.

Висновки. Проведені дослідження статистичної однорідності вибірок за допомогою критеріїв: Андерсона та комбінованого критерію Буша-Вінда. Проведені дослідження показали, що кожен з методів має свої переваги та обмеження залежно від кількості даних (довжини вибірок) та виду розподілу ймовірності (логістичний, експоненціальний, тощо). Критерій Андерсона демонструє високу точність при роботі з великими вибірками та рівномірним розподілом даних. Комбінований критерій Буша-Вінда забезпечує найбільш універсальний підхід, дозволяючи аналізувати вибірки з різними статистичними параметрами розподілу. У результаті проведеного дослідження

підтверджено, що кожен із розглянутих критеріїв може бути ефективно застосований у залежності від конкретного завдання. Для перевірки вибірок з невідомими розподілами найкраще підходить комбінований критерій Буша-Вінда, оскільки він враховує широкий спектр можливих варіацій у вибірках. Отримані результати можуть бути корисними для проведення статистичного аналізу у різних наукових і прикладних дослідженнях, включаючи контроль якості продукції, оцінку надійності вимірювальних приладів, аналіз економічних даних та інші сфери, де необхідно визначати статистичну однорідність вибірок. Подальші дослідження можуть бути спрямовані на розширення спектра критеріїв, що застосовуються, та розробку нових підходів до оцінки статистичної однорідності вибірок з урахуванням складніших математичних моделей і методів машинного навчання.

ЛІТЕРАТУРА

1. Malaichuk, V., Klymenko, S., & Astakhov, D. (2023). Computer Processing of Measurements in Problems of Observation of the Condition of Technical Objects. Journal of Rocket-Space Technology, 30(4), 99-106. <https://doi.org/10.15421/452213>
2. GailF. FahoomeTwenty Nonparametric Statistics And Their Large Sample Approximations / Journal of Modern Applied Statistical Methods. No2, Wayne State University, 2002. p.248–268.
3. Hajek J. Nonparametric Statistics. Holden-Day, San Francisco, 1969. – P. 346.
4. Kobzar, A. I. (2006). Applied Mathematical Statistics: For Engineers and Researchers. M.: Fizmatlit.– P. 816
5. Таблиці функцій та критичних точок розподілів. Розділи: Теорія ймовірностей. Математична статистика. Математичні методи в психології. / уклад. М. М. Горонескуль. Харків: УЦЗУ, 2009. 90с.
6. Турчин В.М. Теорія ймовірностей і математична статистика: навч. пос. Дніпропетровськ: IMA, 2014. 556 с.

RESEARCH ON THE HOMOGENEITY OF PSEUDO-RANDOM SAMPLES USING ANDERSON'S CRITERION AND ITS ANALOGS

Klymenko O. D.

Abstract. *The paper examines in detail the criteria for statistical homogeneity of samples of experimental measurements in cases where the probability distribution functions are a priori unknown. In particular, the Anderson criterion, which is widely used in practice to assess the homogeneity of data based on rank distributions, is analyzed. The Spearman criterion is also considered, which allows comparing two*

samples obtained using the same measuring device and identifying statistical differences between them. In addition, the combined Busch-Wind criterion is proposed, which is based on logistic random variables and is analogous to the well-known Van der Waerden and Klotz statistics in its properties. The corresponding mathematical models, algorithms for calculations, and critical values used to test statistical hypotheses regarding the homogeneity of samples are presented. Experimental studies have been conducted that included logistic and exponential samples, which confirmed the high efficiency of the criteria used in applied data analysis problems.

Keywords: *statistics, criterion, algorithm, mathematical model, probability distribution, homogeneity.*

REFERENCES

1. Malaichuk, V., Klymenko, S., & Astakhov, D. (2023). Computer Processing of Measurements in Problems of Observation of the Condition of Technical Objects. *Journal of Rocket-Space Technology*, 30(4), 99-106. <https://doi.org/10.15421/452213>
2. Fahoome, G. F. (2002). Twenty Nonparametric Statistics and Their Large Sample Approximations. *Journal of Modern Applied Statistical Methods*, 2, 248–268.
3. Hajek, J. (1969). *Nonparametric Statistics*. San Francisco: Holden-Day.
4. Kobzar, A. I. (2006). *Applied Mathematical Statistics: For Engineers and Researchers*. M.: Fizmatlit. – P. 816.
5. Horoneskul, M. M. (Ed.). (2009). *Tables of Functions and Critical Points of Distributions: Probability Theory, Mathematical Statistics, Mathematical Methods in Psychology*. Kharkiv: UCZU.
6. Turchyn, V. M. (2014). *Probability Theory and Mathematical Statistics: A Textbook*. Dnipro: IMA.