

О.Д. Клименко

ПЕРЕВІРКА ОДНОРІДНОСТІ ПСЕВДОВИПАДКОВИХ ВИБІРОК КРИТЕРІЯМИ АНДЕРСОНА ТА БУША-ВІНДА

Анотація. 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. У цій статті розглянуто два підходи до перевірки статистичної однорідності вибірок: двох вибіркової критерій Андерсона та комбінований критерій Буша-Вінда. Кожен із цих критеріїв має свої особливості, сфери застосування та ефективність залежно від розміру вибірок та розподілу ймовірності, що аналізуються. Мета дослідження: оцінка ефективності двох підходів до перевірки статистичної однорідності вибірок: двох вибіркового критерію Андерсона та комбінованого критерію Буша-Вінда, а також визначення впливу розміру вибірок та статистичних параметрів псевдовипадкових величин, визначені недоліки та переваги використання цих критеріїв. Методи дослідження: у дослідженні розглядаються критерії перевірки однорідності псевдовипадкових величин: двох вибіркової критерій Андерсона та комбінований критерій Буша-Вінда, за допомогою яких визначається однорідність псевдовипадкових вибірок в умовах невідомих статистичних параметрів (розподілу ймовірності, математичного очікування, дисперсії. Наукова новизна дослідження: визначено, що, якщо вибірки у дослідженні, $n \leq 100$, то критерій Андерсона є більш надійним для аналізу псевдовипадкових величин із невідомими статистичними параметрами, так як він демонструє вищу стійкість до аномалій, у порівнянні з комбінованим критерієм Буша-Вінда. Визначено, що при дослідженні малих вибірок ($n=20 \div 40$) обидва критерія показують високу ймовірність прийняття правильного рішення при аналізі псевдовипадкових величин на однорідність з різними статистичними параметрами. Практична новизна дослідження: для аналізу випадкових величин з невідомими статистичними параметрами для коротких вибірок запропоновано використовувати двох вибіркової критерій Андерсона та комбінований критерій Буша-Вінда, але для підвищення надійності результатів при дослідженні, в умовах забруднення, доцільно комбінувати статистичні критерії з методами очищення даних або використовувати робастні аналоги. Сфери застосування: дослідження може бути використане в галузях, де необхідно оцінювати статистичну однорідність вибірок, зокрема в експериментальних науках, соціології, медицині та інженерії, для оцінки якості вимірювань та виявлення можливих відхилень у даних.

Ключові слова: математична статистика, метод, вибірка, розподіл ймовірності, критерій Андерсона, критерій Буша-Вінда.

Вступ. 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. У цій статті розглянуто два підходи до перевірки статистичної однорідності вибірок: двохвибірковий критерій Андерсона та комбінований критерій Буша-Вінда. Кожен із цих критеріїв має свої особливості, сфери застосування та ефективність залежно від специфіки вибірок, що аналізуються. Двох вибірковий критерій Андерсона широко використовується для оцінки статистичної однорідності вибірок, отриманих у ході експериментальних досліджень з невідомими статистичними параметрами. Методологія цього підходу базується на порівнянні упорядкованих значень вибірок та їх 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. 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. Це 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 [1-7]. Проведемо дослідження щодо перевірки однорідності вибірок випадкових величин за двома критеріями.

Теоретичні основи критерію Андерсона. Критерій Андерсона широко застосовується під час перевірки статистичної однорідності вибірок експериментальних вимірів з невідомими функціями розподілу ймовірності. Цей критерій, що формується за рангами впорядкованих вибірок $X_1^*(k)$ та $X_2^*(k)$, а також їх об'єднаної вибірки $X_{12}^*(i)$, де $i = 1, 2, \dots, 2n$, $X_{12}^*(i) = X_1^*(k) + X_1^*(k - n)$, $X_{21}^*(i) = X_2^*(k) + X_2^*(k - n)$. Ранги визначаються за формулами

$$R(X_1^*(k)) = \sum_{k=1}^{2n} (X_1^*(k) - X_{12}^*(k)), \quad R(X_2^*(k)) = \sum_{k=1}^{2n} (X_2^*(k) - X_{21}^*(k)).$$

де $\text{sgn}(x - a)$ – функція одиничного стрибка, що дорівнює одиниці, якщо $x \geq a$, та нулю, якщо $x < a$.

Критерій Андерсона описується формулою

$$V_1(n) = \frac{1}{2n^2} \left[\sum_{k=1}^n (R(X_1^*(k)) - k)^2 + (R(X_2^*(k)) - k)^2 \right] - \frac{4n^2 - 1}{12n},$$

Це випадкові величини та їх статистичні закономірності при $n \geq 40$ схожі на розподіл Смирнова-Крамера-фон-Мізеса [1]. Якщо вибірки короткі і статистично однорідні, то з ймовірністю 0,95 критерій Андерсона $V_1(n) \leq V_{0.1}(n)$ менше критичного значення

$$V_{0.1}(n) = \left(0,4614 - \frac{0,4}{n} - \frac{0,6}{n^2} \right) \left(1 + \frac{1}{n} \right).$$

Якщо виконується ця нерівність, то приймається рішення про однорідність досліджуваних вибірок.

Теоретичні основи комбінованого критерію Буша-Вінда. Якщо дві вибірки $X_1(k)$ і $X_2(k)$ можуть змінюватися математичні очікування і дисперсії, зрушення і масштаби, то передбачається перевіряти їх статистичну однорідність, використовуючи комбінований критерій Буша-Вінда. Класична математична статистика цей критерій апроксимує стандартних нормальних випадкових величин [2-4]. Замінімо їх логістичними випадковими величинами – функцією $F(x)$ та зворотною функцією $U(x)$.

$$F(x) = \frac{1}{1 + \exp\left(-\frac{\pi}{\sqrt{3}}x\right)},$$

$$U(x) = \frac{\sqrt{3}}{\pi} \left(\ln(F(x)) - \ln(1 - F(x)) \right).$$

Критерій Буша-Вінда формується на статистиці, еквівалентній статистиці Вандер-Вардена та Клотца [1, 2]. Досліджуються дві вибірки $X_1(k)$ і $X_2(k)$ з невідомими статистичними закономірностями. Формується об'єднана вибірка $X_{12}(k)$ за формулою

$$X_{12}(k) = X(i/2) + X(i - n/2), i = 1, 2, \dots, 2n,$$

де $X(i/2) = X_1(k)$, $X_1(i) = 0$, $i > 0$; $X(i/2) = X_2(i - n)$, $i > n$, $X_2(i) = 0$, $i < n$.

Визначимо ранги $X_{12}(k)$

$$R(X_{12}(k)) = \sum_{i=1}^n \operatorname{sgn}(X_{12}(k) - X_{12}(i)).$$

Використовуючи ранги, перетворимо вибірку $X_{12}(i)$ на впорядковану вибірку

$$X_{12}^*(i) = X_{12}(k) \sum_{k=1}^{2n} \left(\operatorname{sgn}(i - R(X_{12}(k))) + \operatorname{sgn}(i - 1 - R(X_{12}(k))) \right),$$

Ранги впорядкованої вибірки $R(X_{12}^*(i))$ відомі, це їх номери $i = 1, 2, \dots, 2n$. Для розуміння розглянемо дві впорядковані вибірки

$$X_{12}^*(i/1) = 2, 1; 3, 2; 4, 8; 5, 7; 8, 1; 9, 2; 12, 4.$$

$$X_{12}^*(i/2) = 0,8; 0,9; 2,6; 4,3; 7,3; 8; 9,4; X_2(i) = 12,7.$$

Сформуємо за ними одну сумарну впорядковану вибірку $X_{12}^*(i/1)$ та її ранги (табл.1).

Таблиця 1

$X_{12}^*(i)$	0,8	0,9	2,1	2,6	3,2	4,3	4,8	5,7	7,3	8,1	9,2	9,4	12,4	12,7
$R(X_{12}^*(i))$	1	2	3	4	5	6	7	8	9	10	11	12	13	14

У тому числі ранги вибірок $X_{12}^*(i/1) = X_1^*$, $X_{12}^*(i/2) = X_2^*$. Їх ранги $R(X_{12}^*(i/1))$ та $R(X_{12}^*(i/2))$ представлені в таблиці 2.

Таблиця 2

$R(X_{12}^*(i/1))$	3	5	7	8	10	11	13
$R(X_{12}^*(i/2))$	1	2	4	6	9	12	14

Ранги $R(X_{12}^*(i/1))$ використовуються для оцінки критерію Буша-Вінда. За ними спочатку визначаються показники, еквівалентні критеріям Ван-дер-Вардена $S(X_{12}^*(i/1))$ і Клотца $T(X_{12}^*(i/1))$ логістичним функціям

$$U\left(\frac{k}{2n+1}\right) = \frac{\sqrt{3}}{\pi} \left[\ln\left(\frac{k}{2n+1}\right) - \ln\left(1 - \frac{k}{2n+1}\right) \right],$$

$$U\left(\frac{R(X_1^*(k))}{2n+1}\right) = \frac{\sqrt{3}}{\pi} \left[\ln\left(\frac{R(X_1^*(k))}{2n+1}\right) - \ln\left(1 - \frac{R(X_1^*(k))}{2n+1}\right) \right].$$

Аналоги критеріїв Ван-дер-Вардена та Клотца визначаються за вибіркою $X_{12}^*(k/1) = X_1^*(k), k = 1, 2, \dots, n$

$$S(X_{12}^*) = \frac{\sqrt{\frac{2n-1}{n^2}} \sum_{k=1}^n U\left(\frac{R(X_{12}^*(i/1))}{n+1}\right)}{\sqrt{\frac{1}{2n} \sum_{k=1}^{2n} U^2\left(\frac{k}{2n}\right)}},$$

$$T(X_{12}^*) = \frac{\sqrt{\frac{2n-1}{n^2}} \sum_{k=1}^n U^2\left(\frac{R(X_{12}^*(k/1))}{n+1}\right) - \frac{1}{2} \sum_{k=1}^n U^2\left(\frac{k}{2n}\right)}{\sqrt{\frac{1}{2n} \sum_{k=1}^{2n} U^2\left(\frac{k}{2n}\right)}}.$$

Це випадкові величини вибірок $S(X_{12}^*)$ та $T(X_{12}^*)$.

Критерії Буша-Вінда визначається за цими показниками $S(X_{12}^*) = S_{12}^*$ та $T(X_{12}^*) = T_{12}^*$. Класична математична статистика розглядає їх як стандартні нормальні випадкові величини та описується їх апроксимаціями. Замінімо такий опис статистики стандартною логістичною функцією розподілу ймовірності

$$F(S_{12}^*) = \frac{1}{1 + \exp\left(-\frac{\pi}{\sqrt{3}} S_{12}^*\right)}, \quad F(T_{12}^*) = \frac{1}{1 + \exp\left(-\frac{\pi}{\sqrt{3}} T_{12}^*\right)}.$$

Формула критерію Буша-Вінда визначається за модулями $|S_{12}^*|$ і $|T_{12}^*|$

$$V_3 = -2\ln\left(2(1 - F(|S_{12}^*|))\right) - 2\ln\left(2(1 - F(|T_{12}^*|))\right).$$

Це випадкові величини і з ймовірністю $P = 0,95$ їх значення порівнюються з граничними значеннями $V_0(n)$, які представлені в таблиці 3.

Таблиця 3

n	5	10	15	20	30	40
$V_0(n)$	7,92	8,65	8,97	9,05	9,24	9,30

Перевірка однорідності псевдовипадкових вибірок критеріями Андерсона та Буша-Вінда. Для проведення дослідження скористуємося двома гіпотезами. Перша гіпотеза: нульова гіпотеза (H_0) – вибірки X_1 та X_2 походять з одного й того ж розподілу, тобто X_1 та X_2 описуються нормальним розподілом з параметрами математичного очікування -нуль ($M[X]=0$), дисперсією одиниця ($D[X]=1$). Друга – альтернативна гіпотеза (H_1) – вибірки X_1 та X_2 походять із різних розподілів ймовірності. У якості вибірки X_1 – будемо використовувати вибірку псевдовипадкових величин, що описуються нормальним розподілом ймовірності, а у якості вибірки X_2 будемо використовувати вибірки Тьюкі з статистичними параметрами ($M[X]=1$, $D[X]=1$). Забруднені вибірки Тьюкі – це тип вибірок, у яких деяка частина даних походить від іншого розподілу (або з іншими параметрами) й порушує початкові припущення щодо однорідності вибірки. Термін "забруднення" часто використовується в контексті статистичних досліджень для позначення ситуації, коли в дані вибірки додаються спостереження, що відрізняються за розподілом від основної частини вибірки. Це може бути результатом помилок вимірювань, непередбачених зовнішніх факторів або неконтрольованих варіацій у даних. Забруднені вибірки Тьюкі – це певний тип забруднення даних, яке впливає на застосування статистичних методів для перевірки гіпотез, особливо при перевірці однорідності вибірок. Забруднення можуть виникати на будь-якому етапі роботи з даними. Їх своєчасне виявлення та корекція критично важливі для отримання вірних результатів. Використання якісного експерименту і суворих методів аналізу допомагає мінімізувати ризики у будь-якій галузі знань (соціальних, біологічних, медичних, тощо) [1, 7]. Проведемо дослідження та визначимо однорідність за допомогою Критерію Андерсона та Буша-Вінда. За гіпотезами H_0 та (H_1) стверджується, що критерії Андерсона та Буша-Вінда не повинні реагувати на порівняння вибірок з різними видами забруднення (стрибки, викиди, помилки вимірювання, зміщення вибірки, пропущені дані, дублювання даних, шум у даних). Для перевірки гіпотез, першу вибірку будемо формувати з нормальним законом розподілу з статистичними параметрами: $M[X_1]=0$, $D[X_1]=1$, а другу будемо моделювати за вибіркою Тьюкі, з параметрами

$M[X_2]=0$, $D[X_2]=1$. Тобто, забруднення вибірки X_2 призводить до зміни її статистичних параметрів, оскільки вона більше не буде мати нормального розподілу. Проведемо дослідження псевдовипадкових величин на різних за розміром вибірках $n=5,10,15,20,25,30,40,50,100$, де n – кількість елементів у вибірці. На рис. 1 – рис.4. представлені гістограми двох вибірок: перша – з нормальним розподілом ймовірності та друга – з різним відсотком забруднення (вибірка Тьюкі), а також наведено ймовірність прийняття рішення нульової гіпотези.

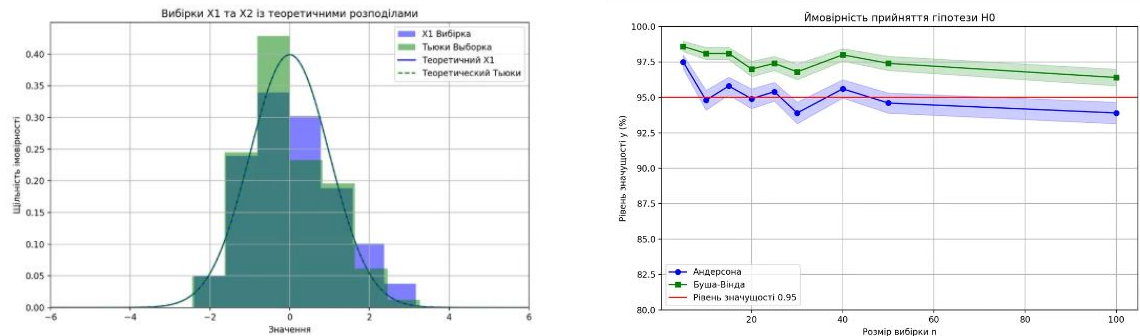


Рисунок 1 - Гістограма та ймовірність прийняття рішення щодо нульової гіпотези за відсутності внесення забруднення у вибірки (забруднення – 0%)

За результатом дослідження (рис.1) отримано, що ймовірність прийняття нульової гіпотези (при $P=0,95$) дуже висока і складає 0,94% за критерієм Андерсона, та за критерієм Буша Вінда – 97%. Це можна пояснити тим, що критерій Андерсона особливо чутливий. В даному випадку використовувалися псевдовипадкові вибірки з однаковими статистичними параметрами: $M[X]=0$, $D[X]=1$. На рис. 2 представлені гістограми двох вибірок: перша – з нормальним законом розподілу ймовірності та друга вибірка Тьюкі (на 5 % забруднена) зі статистичними параметрами $M[X_2]=1$ та $D[X_2]=1$.

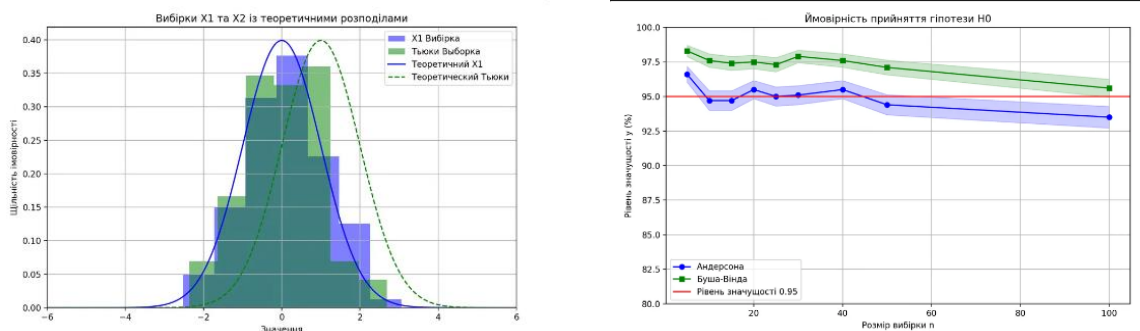


Рисунок 2 - Гістограма та ймовірність прийняття рішення щодо нульової гіпотези з внесенням забруднення у другу вибірку (забруднення – 5%)

На рис. 3 та рис. 4: представлено дослідження, з урахуванням того, що перша вибірка має нормальний закон розподілу, а друга – забруднена та складає 15% та 20% відповідно.

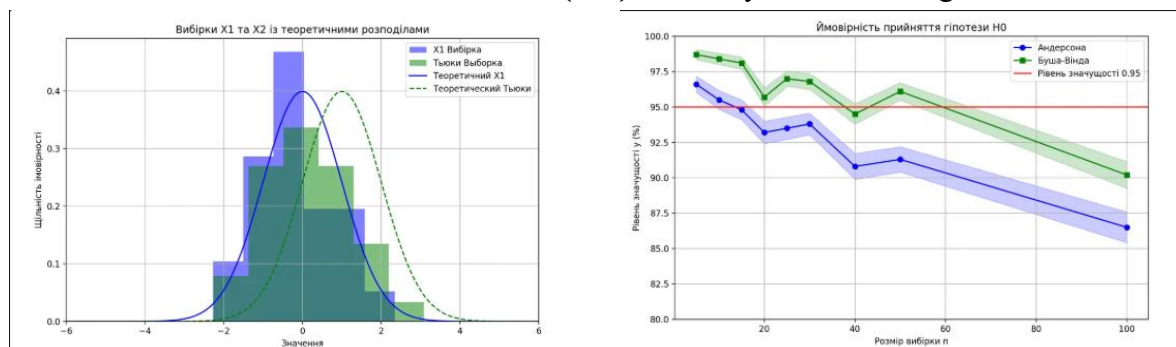


Рисунок 3 - Гістограма та ймовірність прийняття рішення щодо нульової гіпотези з внесенням забруднення у другу вибірку (забруднення – 15%)

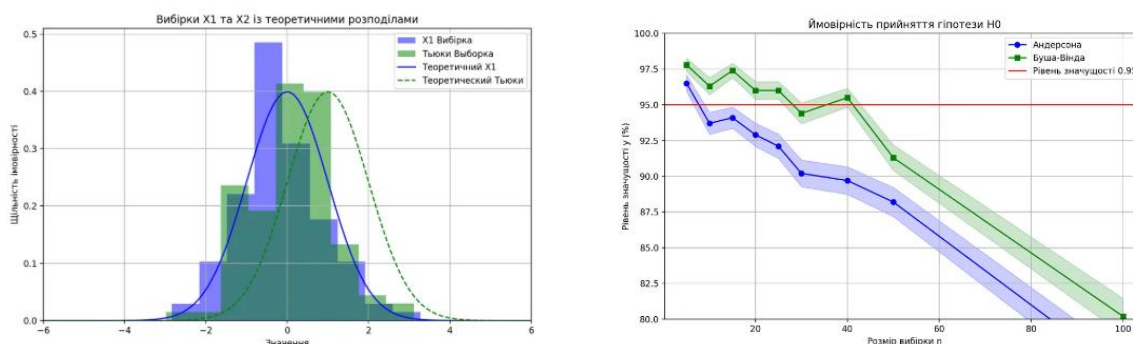


Рисунок 4 - Гістограма та ймовірність прийняття рішення щодо нульової гіпотези з внесенням забруднення у другу вибірку (забруднення – 20%)

Проведені дослідження показали, що кожен з критеріїв має свої переваги та обмеження залежно від специфіки вибірок та їх статистичних параметрів. Критерій Андерсона демонструє високу точність при роботі з великими вибірками з нормальним та рівномірним розподілом ймовірності. Критерій Буша-Вінда дозволяє проводити аналіз вибірки з різними характеристиками розподілу ймовірності, але за результатами дослідження визначено, що може бути прийняття помилкового рішення, щодо нульової гіпотези, навіть за умов дослідження коротких вибірок. В такому випадку рекомендується використовувати додаткові методи для оцінки однорідності.

Висновки. У статті проведено дослідження двох критеріїв: двох вибіркового критерію Андерсона та комбінованого критерію Буша-Вінда для визначення однорідності псевдовипадкових вибірок. Псевдовипадкові вибірки формувалися з урахуванням статистичних параметрів розподілу ймовірності (нормальний розподіл ймовірності та забруднені вибірки Тьюкі з різним ступенем забруднення). За результатами проведеного дослідження можна стверджувати, що чим більше вноситься відсоток забруднення (шум) у одну з вибірок, що порівнюються за критеріями Андерсона та Буша-Вінда, тим більше приймається рішення щодо альтернативної гіпотези – H_1 . Тобто критерії Андерсона та Буша-Вінда дійсно здатні виявити відмінність, яке викликане забрудненням і особливо це можемо бачити на вибірках $n \leq 40$. Для перевірки вибірок з невідомими статистичними параметрами найкраще підходить двох вибіркового критерій Андерсона.

Слід зазначити, що зі збільшенням відсотка забруднення ймовірність прийняття H_0 знижується для обох критеріїв, що свідчить про зростання ймовірності помилкового відхилення H_0 . Однак критерій Андерсона демонструє вищу стійкість до аномалій, особливо при малих вибірках ($n=20\div 40$), де розкид його результатів менший, порівняно з критерієм Буша-Вінда. Для великих вибірок ($n\geq 60$) вплив забруднення на обидва критерії зменшується, що узгоджується із законом великих чисел. Якщо вибірки у дослідженні $n\leq 100$, то критерій Андерсона є більш надійним для аналізу даних із невідомими статистичними параметрами. Для підвищення надійності результатів дослідження, в умовах, коли вибірки з невідомими статистичними параметрами – доцільно комбінувати статистичні критерії з методами очищення даних або використовувати робастні аналоги. 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. Kobzar, A. I. (2006). *Applied Mathematical Statistics: For Engineers and Researchers*. М.: Fizmatlit.– P. 816
3. GailF. FahoomeTwenty Nonparametric Statistics And Their Large Sample Approximations / *Journal of Modern Applied Statistical Methods*. No2, Wayne State University, 2002. p.248–268.
4. Hajek J. *Nonparametric Statistics*. Holden-Day, San Francisco, 1969. – P. 346.
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: ІМА, 2014. 556 с.
7. Рогач І.М., Керецман А.О., Сіткарь А.Д. Правильно вибраний метод статистичного аналізу – шлях до якісної інтерпретації даних медичних досліджень. Науковий вісник Ужгородського університету, серія «Медицина», випуск 2 (56), 2017 р. Режим доступу: <https://dspace.uzhnu.edu.ua/jspui/bitstream/lib/31008/1/25>.

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. Kobzar, A. I. (2006). *Applied Mathematical Statistics: For Engineers and Researchers*. М.: Fizmatlit.– R. 816

3. GailF. FahoomeTwenty Nonparametric Statistics And Their Large Sample Approximations / Journal of Modern Applied Statistical Methods. No2, Wayne State University, 2002. p.248–268.
4. Hajek J. Nonparametric Statistics. Holden-Day, San Francisco, 1969. – R. 346.
5. Tablytsi funktsii ta krytychnykh tochok rozpodiliv. Rozdily: Teoriia ymovirnostei. Matematychna statystyka. Matematychni metody v psykholohii. / uklad. M. M. Horoneskul. Kharkiv: UTsZU, 2009. 90s.
6. Turchyn V.M. Teoriia ymovirnostei i matematychna statystyka: navch. pos. Dnipropetrovsk: IMA, 2014. 556 s.
7. Rohach I.M., Keretsman A.O., Sitkar A.D. Pravylny vybranyi metod statystychnoho analizu – shliakh do yakisnoi interpretatsii danykh medychnykh doslidzhen. Naukovyi visnyk Uzhhorodskoho universytetu, seriia «Medytsyna», vypusk 2 (56), 2017 r. Rezhym dostupu: <https://dspace.uzhnu.edu.ua/jspui/bitstream/lib/31008/1/25>.

Received 14.05.2025.

Accepted 19.05.2025.

Testing the homogeneity of pseudorandom sample by the anderson and Bush-Wind criteria

The study of statistical homogeneity of samples is an important task in mathematical statistics, since it allows us to assess the reliability of experimental measurements and determine whether they belong to the same general population. This article considers two approaches to checking statistical homogeneity of samples: the two-sample Anderson criterion and the combined Bush-Wind criterion. Each of these criteria has its own characteristics, scope and effectiveness depending on the sample size and probability distribution being analyzed. The purpose of the study: to assess the effectiveness of two approaches to checking statistical homogeneity of samples: the two-sample Anderson criterion and the combined Bush-Wind criterion, as well as to determine the influence of sample size and statistical parameters of pseudo-random variables, and to identify the disadvantages and advantages of using these criteria. Research methods: the study considers the criteria for checking the homogeneity of pseudorandom variables: the two-sample Anderson criterion and the combined Bush-Wind criterion, which are used to determine the homogeneity of pseudorandom samples under conditions of unknown statistical parameters (probability distribution, mathematical expectation, variance. Scientific novelty of the study: it was determined that, if the samples under study are $n \leq 100$, then the Anderson criterion is more reliable for the analysis of pseudorandom variables with unknown statistical parameters, since it demonstrates higher resistance to anomalies, compared to the combined Bush-Wind criterion. It was determined that when studying small samples ($n=20 \div 40$) both criteria show a high probability of making the correct decision when analyzing pseudorandom variables for homogeneity with different statistical parameters. Practical novelty of the study: for the analysis of random variables with unknown statistical parameters for short samples, it is proposed to use the two-sample Anderson criterion and the combined criterion Bush-Wind, but to increase the reliability of the results in the study, in conditions of contamination, it is advisable to combine statistical criteria with data cleaning methods or use robust analogues. Areas of application: the study can be used in ar-

eas where it is necessary to assess the statistical homogeneity of samples, in experimental sciences, sociology, medicine and engineering, to assess the quality of measurements and identify possible deviations in the data.

Keywords: mathematical statistics, method, sample, probability distribution, Anderson criterion, Bush-Wind criterion.

Клименко Олексій Денисович - аспірант кафедри кібербезпеки і комп'ютерно-інтегрованих технологій, Дніпровський національний університет імені Олеся Гончара.

Klymenko Oleksii - postgraduate student of the Department of Cybersecurity and Computer-Integrated Technologies, Oles Honchar Dnipro National University.