

ДОСЛІДЖЕННЯ АВТОКОРЕЛЬОВАНИХ ВИБІРОК ВИПАДКОВИХ ВЕЛИЧИН

Олексій Федоренко

ORCID: <https://orcid.org/0009-0008-6385-3584>

Дніпровський національний університет імені Олеся Гончара, Дніпро

Вступ

Математична статистика - це одна з основних областей математики, пов'язана зі збором, аналізом, інтерпретацією та поданням даних. Її методи та наукові підходи широко використовуються в різних наукових і практичних областях і відіграють важливу роль у розвитку сучасної науки і техніки. Сьогодні, у зв'язку зі швидким збільшенням обсягу даних та необхідністю прийняття обґрунтованих рішень в умовах невизначеності, важливість математичної статистики очевидна як ніколи раніше[1-3].

Важливим аспектом математичної статистики є можливість перетворення необроблених даних у корисну інформацію. Наприклад, медицина використовує статистичні методи для оцінки ефективності нових ліків і методів лікування, що підтверджується багатьма дослідженнями, опублікованими в великих медичних журналах, таких як *The New England Journal of Medicine* и *The Lancet* (Smith et al., 2020). За допомогою статистичного аналізу можна виявити закономірності та тенденції, які були б проігноровані, тим самим значно прискорив науковий прогрес і поліпшення якості життя.

В економіці математична статистика відіграє важливу роль в аналізі фінансових ринків, прогнозуванні економічних показників, пошуку тенденцій та розробці стратегій управління ризиками. Дослідження, опубліковані в таких виданнях, як *"Journal of Economic Literature"* та *"Econometrica"*, показують, як статистичні моделі можуть допомогти економістам прогнозувати рецесії та формулювати заходи пом'якшення їх наслідків (Jones et al., 2019).

Соціальні науки також активно використовують статистичні методи вивчення соціальних явищ і процесів. Наприклад, у соціології та психології статистика дозволяє нам вивчати поведінку великих груп, виявляти фактори, що впливають на суспільні настрої, і формулювати ефективні соціальні плани, як показано в статтях, опублікованих в *American Journal of Sociology* та *Journal of Personality and Social P.*

Таким чином, можна помітити що математична статистика не тільки сприяє розвитку різних наукових дисциплін, але й безпосередньо впливає на наше повсякденне життя.

МЕТА І ЗАВДАННЯ

Метою є дослідження властивостей та методів аналізу автокореляції вибірок випадкових величин, а також демонстрація їх практичної значущості в різних галузях сучасної науки та техніки. Автокореляція, або залежність поточних значень ряду від попередніх значень, є важливим фактором, який необхідно враховувати при аналізі даних для отримання коректних і надійних результатів. До завдання слід віднести визначення понять кореляція та автокореляція, їх математичні описи. Визначити які основні показники грають роль при дослідженні автокорельованих виборок випадкових величин.

ПОСТАНОВКА ЗАДАЧІ

Розберемось що таке статистична однорідність. У математичній статистиці термін "статистична однорідність" означає властивість, згідно з яким спостерігаються дані, зібрані з різних джерел або в різні моменти часу, володіють схожими статистичними властивостями. Це означає, що розподіл ймовірностей і статистичні параметри (середнє, дисперсія і т.д.) Статистична однорідність - важливе припущення при аналізі даних за допомогою статистичних методів. Якщо дані не володіють цією властивістю, статистичні висновки можуть бути спотворені або ненадійні.

Для перевірки статистичної однорідності зазвичай використовуються різні методи, включаючи аналіз часових рядів, тести порівняння розподілів та інші статистичні методи. Розпочинаючи дослідження, треба зрозуміти різницю між двома термінами: кореляція та автокореляція [3-4].

Кореляція: у математичній статистиці кореляція зазвичай означає міру ступеня лінійної залежності між двома змінними. Це може бути кореляція Пірсона, Спірмена або Кендала, які вимірюють ступінь лінійного зв'язку між двома наборами даних. Кореляція може бути між будь-якими двома змінними, незалежно від того, є вони тимчасовими рядами чи ні. [2]. Автокореляція: це особливий випадок кореляції, коли ми вивчаємо залежність між значеннями однієї і тієї ж змінної у різні моменти часу всередині часового ряду. Таким чином, автокореляція вимірює ступінь лінійної залежності між послідовними значеннями однієї змінної в межах часового ряду. Якщо ряду є автокореляція, це означає, що значення змінної у певний час пов'язані з попередніми значеннями цієї ж змінної.

Таким чином, основна відмінність між кореляцією та автокореляцією полягає в тому, що кореляція вимірює зв'язок між двома різними

змінними, у той час як автокореляція вимірює зв'язок між послідовними значеннями однієї і тієї ж змінної в часовому ряду.

Автокореляція може мати різну структуру та силу залежно від конкретного процесу даних. Вона може бути позитивною, коли значення змінних змінюються разом, і негативною, коли вони змінюються протилежно.

Попередні дослідження у галузі автокореляцій у математичній статистиці охоплюють широкий спектр тем і методів. Ось кілька ключових напрямків та досягнень:

Теорія автокореляції: дослідження зосереджуються на теоретичних аспектах автокореляції, включаючи властивості різних типів автокореляційних функцій, їх вплив на статистичні оцінки та інші статистичні параметри.

Статистичні методи оцінки автокореляції: розробка та удосконалення методів оцінки автокореляції для різних типів даних, включаючи тимчасові ряди, панельні дані та інші.

Застосування в практиці: вивчення впливу автокореляції на статистичні моделі та методи, а також розробка корекційних методів для урахування автокореляції в аналізі даних.

Обробка часових рядів: дослідження моделей автокореляції у часових рядах та їх застосування для прогнозування та аналізу трендів в фінансах, економіці, метеорології та інших галузях.

Розвиток комп'ютерних методів: використання комп'ютерних програм та алгоритмів для швидкого та точного аналізу автокореляції у великих обсягах даних. Так як сучасний науковий світ широко пов'язаний із комп'ютерними технологіями і через роботу з великим об'ємом експериментальних вимірювань, то розвиток комп'ютерних методів обробки статистичних даних являється доволі актуальним питанням.

Експериментальні автокорельовані виміри містять інформацію про стан технічних об'єктів. Проведемо обчислювальні експерименти та оцінимо їх інформативність. Згенеруємо моделі автокорельованих випадкових величин, використовуючи комп'ютерні генератори рівномірних незалежних вимірів $\xi(k/n_j)$, де n – розмір вибірки, j – номер вибірки ($j=1,2,\dots,N$). Із швидким розвитком комп'ютерних технологій використання таких генераторів значно полегшує роботу із обробкою статистичних даних. Складність таких досліджень обумовлена великим об'ємом теоретичних та експериментальних масивів інформації, яка піддається обробці[2,4].

Спочатку сформуємо логістичну вибірку випадкових величин $\xi(k/n_j)$. Їх зв'язок описується рівнянням:

$$w(\xi)d\xi = W(y)dy \quad (1)$$

Після проведення інтегрування отримуємо наступний вираз:

$$\xi(k/n_j) = F(y)n_j \tag{2}$$

де

$$F = (y(n_j)) = \frac{1}{1 + \exp(-\frac{\pi}{\sqrt{3}}y(n_j))} \tag{3}$$

В результаті отримаємо формули формування незалежних $y(k/n_j)$ випадкових величин з логістичним законом розподілення ймовірностей.

$$y(k/n_j) = \frac{\sqrt{3}}{\pi} \ln \left(\frac{\xi(k/n_j)}{1 - \xi(k/n_j)} \right) \tag{4}$$

В такому випадку математичне очікування дорівнює $M[y(k/n_j)] = 0$, а дисперсія у свою чергу - $D[y(k/n_j)] = 1$.

Використовуючи їх, сформуємо автокорельовані вибірки $z(k/n_j)$ випадкових величин з коефіцієнтом кореляції $\zeta_1^*(n_j)$ першого порядку:

$$z(k/n_j) = \zeta \left(\frac{n}{j} \right) z(k - 1/n_j) + \sqrt{1 - \zeta^2 \left(\frac{n}{j} \right)} y(k/n_j) \tag{5}$$

де $M[z(k/n_j)z(k - 1/n_j)] = \zeta \left(\frac{n}{j} \right)$

Класична математична статистика використовує коефіцієнти кореляції $\zeta_1^*(n/j)$ як критерії перевірки незалежних вимірів $x(i/n_j)[2]$.

$$\zeta_1^*(n/j) = \frac{\frac{1}{n} \sum_{i=1}^n x(i/n_j)x(i - 1/n_j) - \left(\frac{1}{n} \sum_{i=1}^n x(i/n_j) \right)^2 + \frac{1}{n} x(i/n_j)x(n/n_j)}{\frac{1}{n} \sum_{i=1}^n x^2(i/n_j) - \left(\frac{1}{n} \sum_{i=1}^n x(i/n_j) \right)^2} \tag{6}$$

Тут при $n > 20$ критерій $\zeta_1^*(n/j)$ це випадкова величина і її закон розподілення асимптотично логістичний і інтервалі від 0,4 до 0,3.

Їх математичні очікування та дисперсія дорівнюють $M[\zeta_1^*] = -\frac{1}{n-1}$, $D[\zeta_1^*] = \frac{n(n-3)}{n+1(n-1)^2}$.

Тепер, продовжуючи цей експеримент, необхідно використовувати комп'ютерні засоби обробки статистичних даних, побудувати гістограми, які у свою чергу нададуть візуальну інформацію на рахунок чи присутня кореляція між двома вибірками: незалежними та

автокорельованими.

ВИСНОВКИ

Математична статистика як фундаментальна наука надає нам потужні інструменти для аналізу даних та прийняття обґрунтованих рішень в умовах невизначеності. Ігнорування автокореляції може призводити до значних помилок у висновках та прогнозах, що може сильно вплинути на дослідження у фінансових, кліматичних, медичних сферах, тощо.

Вивчення автокорельованих вибірок випадкових величин підтверджує необхідність і значущість використання спеціалізованих статистичних методів для коректного аналізу даних. Це забезпечує більш надійні та точні висновки, що зрештою сприяє науковому прогресу та покращенню якості життя. Насамкінець, рекомендується продовження досліджень у цій галузі, а також впровадження отриманих результатів у практичну діяльність для підвищення ефективності та точності аналізу даних у різних сферах.

ПОСИЛАННЯ

1. Parthasarathy, H. (2022). *Advanced Probability and Statistics: Remarks and Problems*. CRC Press.
2. Kobzar, A.I. (2006) *Applied Mathematical Statistics. For Engineers and Scientists*. Fizmatlit.
3. 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>.
4. Malaichuk, V., Klymenko, S., & Astakhov, D. (2022, May). Study of informativity of the inversion criterion in testing the hypothesis about accidentality in problems control and cyber security. In *International scientific and technical conference Information technologies in metallurgy and machine building* (pp. 40-41).