



<https://www.sworldjournal.com/index.php/swj/article/view/swj34-01-020>

DOI: 10.30888/2663-5712.2025-34-01-020

УДК 519.2+519.6

STATISTICAL ANALYSIS OF THE 2004 UKRAINIAN PRESIDENTIAL ELECTION FIRST-ROUND RESULTS FOR POLLING STATIONS IN LUHANSK REGION

СТАТИСТИЧНИЙ АНАЛІЗ ДАНИХ І ТУРУ ВИБОРІВ ПРЕЗИДЕНТА УКРАЇНИ-2004 НА ВИБОРЧИХ ДІЛЬНИЦЯХ ЛУГАНСЬКОЇ ОБЛАСТІ

Krykun H. Ivan / Крикун І.Г.

Ph.D. in Physics and Mathematics, Associate Professor / к.ф.-м.н., доцент

ORCID: 0000-0001-5468-512X

Institute of Applied Mathematics and Mechanics, NAS of Ukraine,

Sloviansk, Batiuka St., 19, 84116

& KROK University, Kyiv, Tabirna St., 30-32, 03113

Інститут прикладної математики і механіки НАН України,

Слов'янськ, Батюка, 19, 84116

& Університет КРОК, Київ, Табірна, 30-32, 03113

Анотація. В роботі розглядається підсумки I туру виборів Президента України 2004 року на виборчих дільницях Луганської області. Проведено статистичний аналіз оприлюднених результатів, використовуючи показники явки та голосів “проти всіх”. Для цього, згідно із раніше запропонованим автором методом, побудовано арктангенс-регресію цих даних та знайдено статистичні оцінки параметрів відповідного розподілу Коші. Згідно отриманих оцінок зроблено висновки щодо можливих маніпуляцій з виборчими підсумками.

Ключові слова: арктангенс-регресія, електоральні результати, розподіл Коші, статистична оцінка параметрів розподілу.

Вступ.

Електоральні результати демонструють прояв стохастичності в соціальній динаміці. Підсумковий розподіл голосів є функцією двох ключових змінних імовірнісного характеру: по-перше, флуктуацій електоральних преференцій і, по-друге, варіативності явки виборців. У ситуації, коли розрив у рівнях підтримки основних кандидатів є невеликим, саме показник електоральної активності (явка виборців) може набувати вирішального значення.

Однак, аналіз виборчих процесів у низці країн засвідчує, що електоральні маніпуляції можуть справляти суттєвий або навіть вирішальний вплив на підсумки голосування. Діапазон таких маніпуляцій охоплює спектр від відносно легітимних політичних технологій (як-от «чорний піар», цільова мобілізація електорату на підтримку певного кандидата, демотивація виборців інших



кандидатів або позиціонування «свого» кандидата як «меншого зла» для невизначеної частини виборців) до відвертих фальсифікацій (пряме вкидання бюлетенів в виборчі скриньки, пошкодження бюлетенів за кандидатів-суперників, підробка протоколів про підрахунок голосів).

Реалізацію легальних виборчих технологій (зокрема мотивація «своїх» або демотивація «чужих» виборців через соцмережі та контекстна політична реклама як інструмент такої мотивації) можна було спостерігати на недавніх виборчих процесах в таких країнах сталої демократії як Великобританія (так званий «Brexit», тобто референдум 2016 року за вихід із Євросоюзу; використані технології навіть описані у документальному фільмі «Brexit: The Uncivil War») та США (президентські вибори 2016, які закінчилися перемогою Д.Трампа). Також привертають увагу нещодавні висококонкурентні (і можна сказати навіть скандальні) вибори в Польщі, Молдові, Румунії, Туреччині, де є різноманітні свідчення про намагання міжнародних злочинців з Російської Федерації вплинути на результати виборів.

Прямі фальсифікації ж результатів виборів можна було спостерігати на деяких виборах в Україні (зокрема президентські вибори 2004 року, масові фальсифікації на яких на користь провладного кандидата В.Януковича не лише засвідчені висновками міжнародних незалежних спостерігачів [1], а і призвели до масових народних протестів, які отримали назву «Помаранчевої революції» та переголосуванню вирішального II туру) та фактично на всіх виборчих процесах в Російській Федерації принаймні протягом останніх 20-25 років.

Огляд публікацій.

Для виявлення можливих виборчих фальсифікацій дослідниками було запропоновано кілька підходів. Коротко охарактеризуємо їх.

Один з підходів полягає в тому, що дослідники вивчають частотний розподіл явки виборців або голосів на підтримку певного кандидата на дільничних виборчих комісіях (ДВК) одного або кількох виборчих округів. Якщо виборчі результати не зазнавали впливу, то цей розподіл очікується близьким до нормального розподілу. Відхилення ж згаданого частотного розподілу від



нормального розподілу – наприклад, помітні “важкі хвости” (тобто наявність значного числа ДВК із високою явкою або підтримкою певного кандидата) – є підставою для висновків про наявність маніпуляцій із виборчими даними. Також на подібних графіках явки або підтримки певного кандидата помітним є збільшення появи “круглих” чисел (кратних 5 або 10).

Другий підхід до статистичного аналізу виборчих результатів полягає в тому, що автори розглядають, як частка голосів за кожного кандидата від загальної кількості виборців корелює з явкою.

Третій основний підхід полягає в аналізі співвідношення голосів, поданих за кандидатів/партії, залежно від явки. Вперше він був використаний Дж. Б. Кіслінгом у 2004 році для аналізу виборів у Вірменії. Метод довів свою корисність у Росії для виявлення шахрайства на користь керівної партії.

Ще один підхід – це використання так званого “закону Бенфорда” (або “закону першої цифри”) для порівняння отриманої частоти появи у підсумковій виборчій документації першої (або другої) цифри із теоретичним розподілом.

Більш детальний огляд використовуваних статистичних підходів і прийомів виявлення виборчих маніпуляцій можна знайти в [2]. Опис деяких легальних технологій впливу на результат виборів наведено в [3]. В роботі [4] містяться результати статистичного аналізу виборчих результатів I туру виборів Президента України в 2004 році по усіх 35 округах Донецької та Луганської областей. В даній роботі проведено аналогічний [4], але більш детальний статистичний аналіз виборчих результатів по всіх 1762 дільничним виборчим комісіям Луганської області.

Використана методика.

Проте всі запропоновані статистичні підходи оцінки фальсифікацій на виборах наражаються на закономірну критику, адже, наприклад, відхилення емпіричного розподілу частот від нормального розподілу може бути викликане іншими причинами (це можуть бути регіональні відмінності в активності виборців чи в підтримці певних кандидатів).

Головна ідея даної роботи полягає у побудові, після певної обробки



статистичних даних, арктангенс-регресії отриманих даних та оцінці параметрів цієї арктангенс-регресії, згідно з результатами автора [5], [6].

Наведемо коротко результати з [5], [6].

Нагадаємо, що розподілом Коші з параметрами α та γ називають розподіл випадкової величини, що має щільність [7]

$$f(x) = \frac{1}{\pi} \frac{\gamma}{(x - \alpha)^2 + \gamma^2}, \quad \text{для } -\infty < x < +\infty$$

і, відповідно функцію розподілу

$$F(x) = \frac{1}{2} + \frac{1}{\pi} \operatorname{arctg} \frac{x - \alpha}{\gamma}$$

Параметри α та γ називають відповідно параметром розташування та масштабу.

Відомо [7], що розподіл Коші не має ні математичного сподівання, ні моментів вищих порядків. Розподіл Коші належить до розподілів з «важкими хвостами», для яких не виконується закон великих чисел. Через такі властивості розподілу Коші оцінити його параметри стандартними методами (метод моментів або метод максимальної вірогідності) не вдається.

Базуючись на тому, що функцією розподілу для розподілу Коші буде функція арктангенс, в [5, Пропозиція 2] було запропоновано простий підхід до оцінки невідомих параметрів розподілу Коші, заснований на використанні емпіричної арктангенс-регресії.

Добре відомий результат [7], що відношення двох стандартних нормальних випадкових величин з параметрами 0 та 1 є випадковою величиною, що має розподіл Коші з параметрами 0 та 1.

А оскільки багато виборчих показників (явка виборців, відсотки голосів за певних кандидатів, зіпсованих бюлетенів чи голосів “проти всіх”) є правдоподібним вважати такими, що мають випадкові відхилення, розподілені за нормальним законом розподілу. Таким чином, якщо ми знайдемо математичне сподівання та дисперсії цих величин, то зможемо центрувати та нормувати їх і побудувати відношення цих величин, що мало б мати розподіл Коші з параметрами 0 та 1.



Пропозиція 1. Після нормування двох (або більше) обраних виборчих показників складемо відношення двох з них. Далі проведемо статистичну обробку отриманих даних згідно [5, Пропозиція 3 та Лема 1] та отримаємо емпіричні оцінки параметрів розподілу. Потім порівняємо отримані оцінки з теоретичними параметрами розподілу Коші (тобто зі значеннями 0 та 1) і в разі помітних відмінностей спробувати оцінити ймовірність такого відхилення параметрів від теоретичних (за допомогою довірчих інтервалів [5, Лема 2]) та зробити статистичні висновки про правдоподібність отриманого результату.

Зауваження 1. Для кращого бачення картини, варто виділяти множину виборчих дільниць, щодо яких є свідчення про масові однотипні порушення (наприклад, вкидання бюлетенів за «свого» кандидата або псування бюлетенів за «чужого» кандидата) і аналізувати множину даних з цих дільниць.

Зауваження 2. Якщо є підстави вважати масовою фальсифікацію збільшення явки (через вкидання бюлетенів або багаторазове голосування), то варто обчислити і врахувати в подальшому аналізі підсумкові виборчі показники (% явки, голосів проти всіх та недійсних бюлетенів) без врахування даних з таких округів/виборчих дільниць.

Основні результати.

Історичний контекст виборів.

На виборах Президента України в 2004 році в Луганській області було створено 13 виборчих округів в складі 1462 дільничних виборчих комісії. Луганська область разом з Донецькою складають макрорегіон Донбас, що на тих виборах був базовим для провладного кандидата (В. Януковича), який, згідно з офіційними результатами [8] за підсумками I туру в цих двох областях набрав 39,5% від усіх голосів на свою підтримку, а саме 4,35 млн. із 11,0 млн. Зокрема, в Луганській області В. Янукович набрав 1,638 млн. голосів виборців.

На цих виборів були зафіксовані масові порушення на користь провладного кандидата В. Януковича [1]. Ці порушення фіксувалися і під час I туру, а особливого розмаху фальсифікації на виборах набули під час II туру цих виборів,



що призвело до масових народних протестів проти оприлюднених сфальсифікованих підсумків голосування у листопаді-грудні 2004 року (які отримали назву «Помаранчева Революція») та ухвалення Верховним Судом України судового рішення про переголосування вирішального II туру, що відбулося 26.12.2004 року. Особливо масштабними були фальсифікації саме в Донецькій та Луганській областях, про що свідчить і нижченаведена Таблиця 1.

Таблиця 1 – Статистичні підсумки голосування в Донецькій та Луганській областях

	I тур, кількість виборців в списку	Явка (голосів і % від списку I туру)		
		I тур	II тур	Переголосування II туру
Донецька область	3,685 млн.	2,878 млн. 78,1 %	3,712 млн. 100,7 %	3,144 млн. 85,3 %
Луганська область	1,946 млн.	1,472 млн. 75,6 %	1,754 млн. 90,1 %	1,638 млн. 84,2 %

Авторська розробка на підставі даних [8]

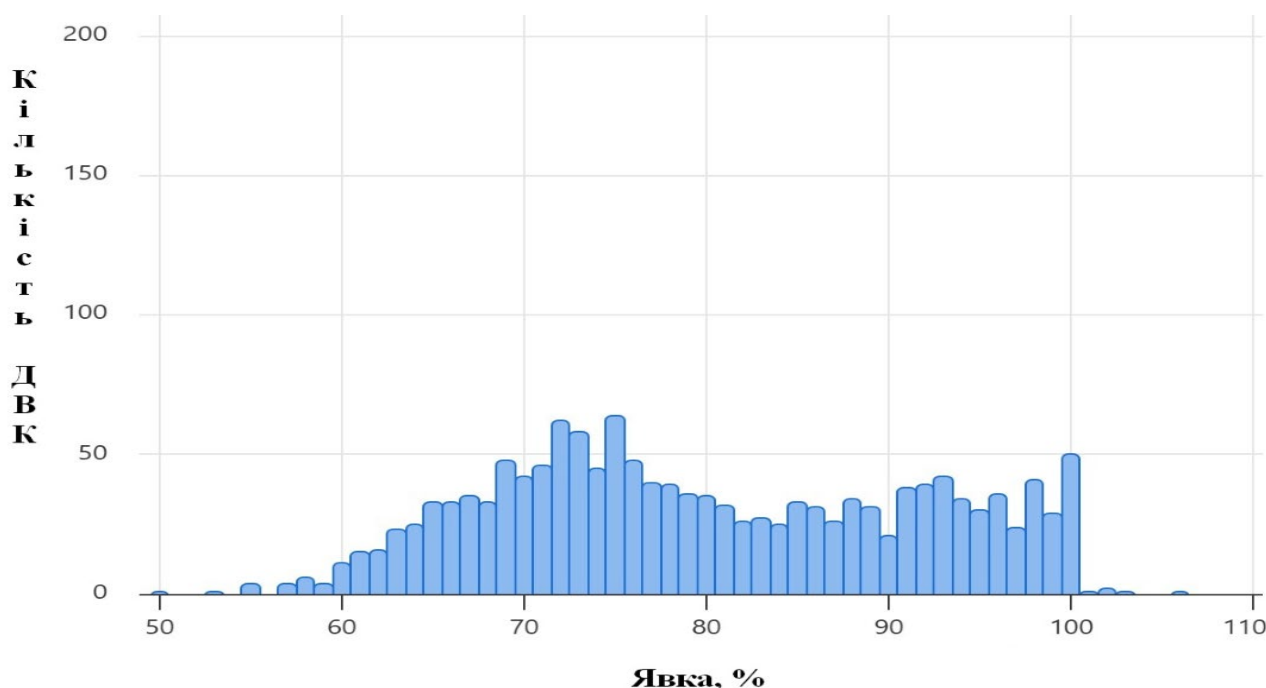


Рисунок 1 – Розподіл ДВК Луганської області за явкою в I турі



Джерело: авторська розробка з використанням інструменту [9]

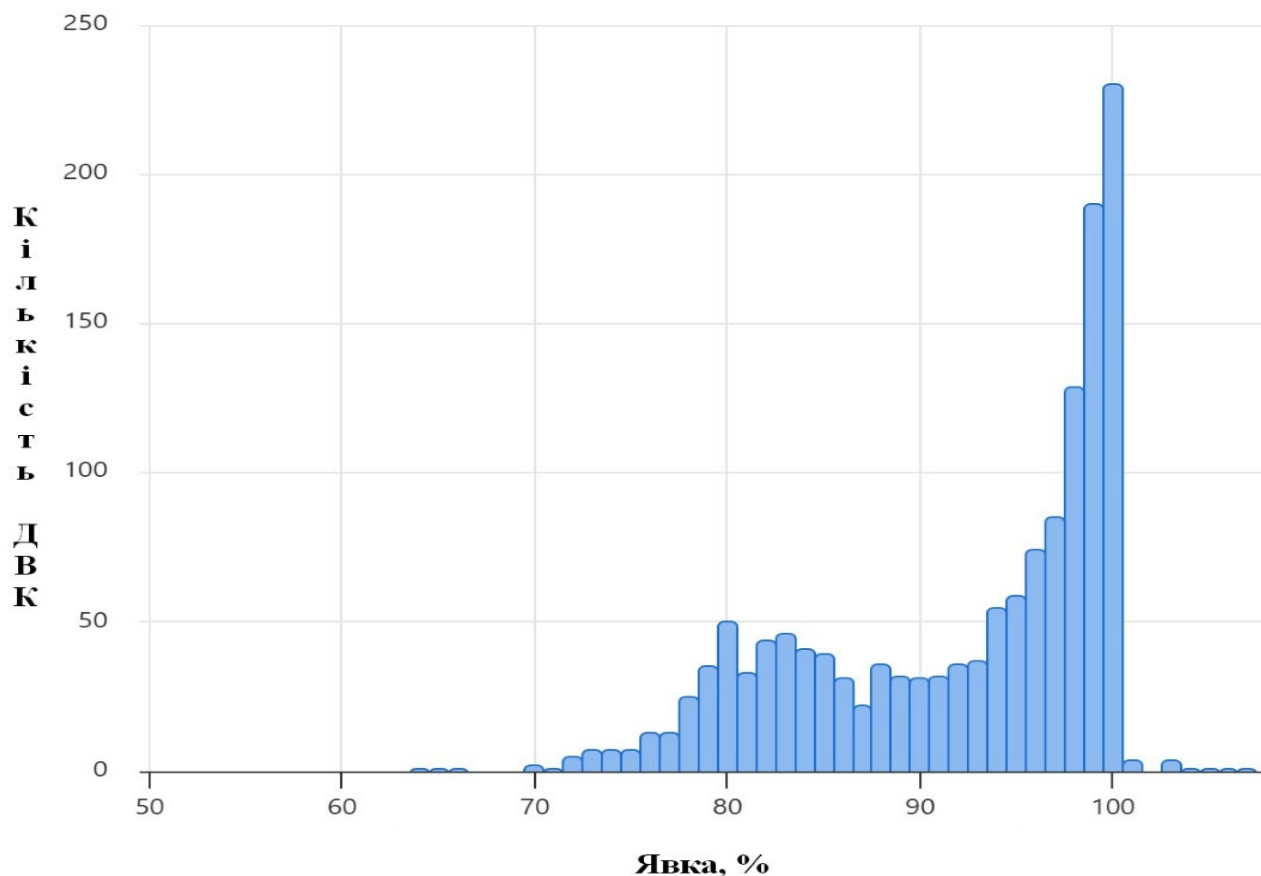


Рисунок 2 - Розподіл ДВК Луганської області за явкою в II турі

Джерело: авторська розробка з використанням інструменту [9]

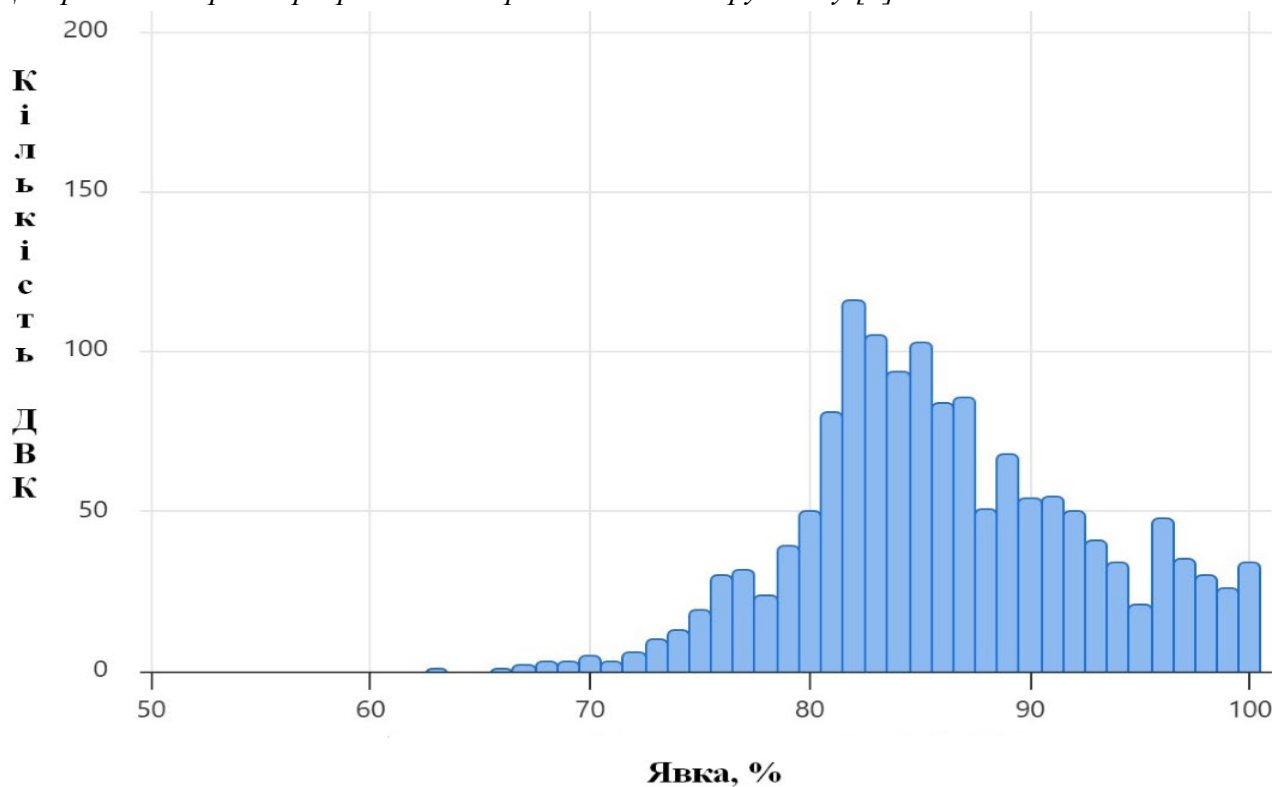


Рисунок 3 - Розподіл ДВК Луганської області за явкою в переголосуванні II туру



Джерело: авторська розробка з використанням інструменту [9]

Вже фантастична цифра явки під час проведення II туру виборів по всій Донецькій області у понад 100% (від кількості виборців у списку I туру) підтверджує твердження про масовість фальсифікацій.

Про наявність виборчих порушень свідчать також діаграми явки, що для дільничних виборчих комісій Донецької та Луганської областей наведено в [4], а окремо для ДВК Луганської області наведено нижче:

Оцінка параметрів розподілу Коші

Отже, по I туру голосування в Луганській області можемо припустити наявність фальсифікацій, пов'язаних із підвищенням явки виборців (через багаторазове голосування або пряме вкидання бюлетенів). Тому для аналізу оберемо 2 показника – відсоток явки виборців (який може бути завищеним) та відсоток голосів “проти всіх” (який може бути заниженим).

Далі, для уникнення впливу завищеної явки знайдемо вказані статистичні показники I туру цих виборів без врахування голосів з Донецької та Луганської областей. Маємо (згідно офіційних даних [8]):

Таблиця 2 – Статистичні підсумки голосування без врахування даних по Донецькій та Луганській областях

	Явка	“Проти всіх”
Середнє	74,502 %	2,027 %
Середньоквадратичне відхилення	6,810 %	1,363 %

Авторська розробка на підставі даних [8]

Далі, використовуючи дані з Таблиці 2, centruємо та нормалізуємо показники явки та голосів “проти всіх” з ДВК Луганської області та складемо їх відношення за формулою

$$x_i = \frac{\text{нормований та центрований \% явки на } i - \text{й ДВК}}{\text{нормований та центрований \% голосів проти всіх на } i - \text{й ДВК}}$$

Для отриманої вибірки x_i застосуємо процедуру [5, Пропозиція 2] та використаємо [5, формули (10), (11)]. Крім того, врахуємо [5, Пропозиція 3] та



відкинемо певну кількість найбільших та найменших значень з вибірки x_i .

Після вказаної обробки даних отримаємо таблицю оцінок параметрів:

Таблиця 3 – Статистичні оцінки параметрів за даними всіх 1762 ДВК Луганської області

Оцінка параметру	Відкинуто найбільших і найменших даних		
	по 1 %	по 10%	по 25%
α	-1,634623978	-1,50362973	-1,482665792
γ	1,773495061	2,06348137	2,302247941

Авторська розробка

Враховуючи висновки з [5, Пропозиція 3], згідно з якими нівелювання впливу “важких хвостів” розподілу Коші шляхом відкидання 10% або 25% найбільших і найменших вибірових даних дає найближчу до теоретичних значень статистичну оцінку параметрів розподілу Коші, можемо стверджувати, що оцінка параметру розташування α перебуває в межах $[-1,483; -1,503]$, а оцінка параметру масштабу γ перебуває в межах $[2,063; 2,302]$.

Правдоподібність виборчих результатів

Нагадаємо, що згідно [7] якщо вихідні дані про показники явки та голосування “проти всіх” справді мають нормальний розподіл, то теоретичні значення параметрів розподілу Коші дорівнюють: $\alpha=0$, $\gamma=1$.

Відповідно можемо констатувати, що емпіричні оцінки параметрів значно відрізняються від теоретичних, що свідчить про наявність помітного впливу на результати цих виборів.

Тому постає питання про правдоподібність отриманих виборчих даних, тобто наскільки ймовірною є поява подібних показників?

Для відповіді на це питання розглянемо вибірку із спостережень реалізації випадкової величини, що має розподіл Коші з невідомим параметром α і відомим параметром γ . Нами обчислено середнє вибірки $\bar{x} = -0,789$ і для простоти припустимо, що параметр γ є відомим. Тоді обчислимо теоретичну ймовірність того, що параметр α міститься в розширеному (порівняно з отриманими



оцінками) інтервалі, наприклад в інтервалі $[-1,55; -1,45]$.

Для цього використаємо таку властивість розподілу Коші [7]: якщо маємо вибірку з розподілу Коші з параметрами α і γ , то їх середнє арифметичне \bar{x} теж матиме розподіл Коші з тими ж параметрами.

Тоді отримуємо оцінку ймовірності (вважаючи параметр $\gamma=2$)

$$\begin{aligned} P\{\alpha \in [-1,55; -1,45]\} &= P\{-1,55 \leq \alpha \leq -1,45\} = \\ &= P\{-1,55 - \bar{x} \leq \alpha - \bar{x} \leq -1,45 - \bar{x}\} = \\ &= \frac{1}{\pi} \arctan \frac{-1,45 - \bar{x}}{\gamma} - \frac{1}{\pi} \arctan \frac{-1,55 - \bar{x}}{\gamma} = \\ &= \frac{1}{\pi} \arctan(-0,3305) - \frac{1}{\pi} \arctan(-0,3805) = 0,0141. \end{aligned}$$

Порівняємо з аналогічним результатом для $\gamma=2,3$:

$$P\{\alpha \in [-1,55; -1,45]\} = \frac{1}{\pi} \arctan(-0,2874) - \frac{1}{\pi} \arctan(-0,3309) = 0,0126.$$

І ймовірності для вузких інтервалів (для $\gamma=2$):

$$P\{\alpha \in [-1,52; -1,47]\} = 0,0071,$$

і для $\gamma=2,3$

$$P\{\alpha \in [-1,52; -1,47]\} = 0,0063.$$

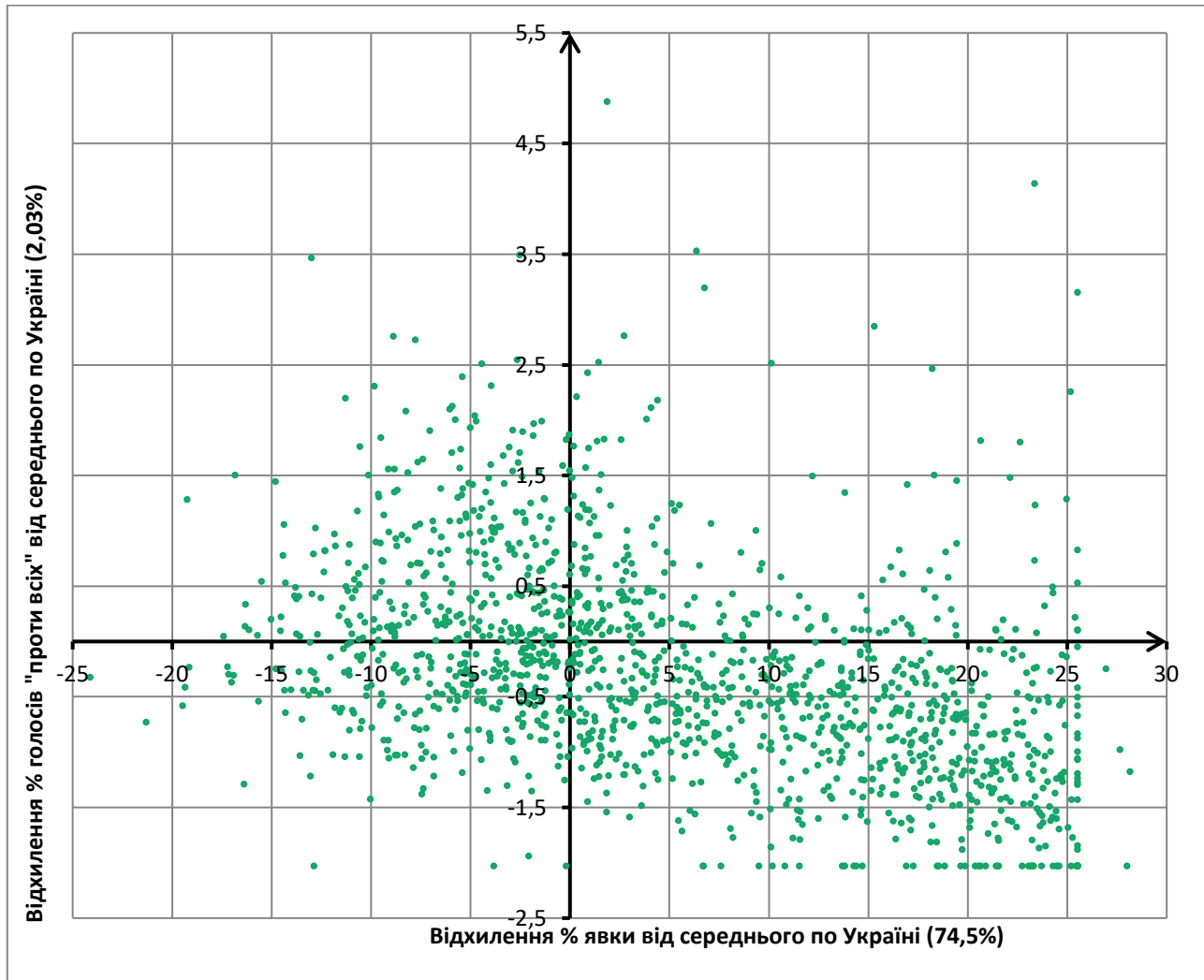
Отже, можемо стверджувати, що оприлюднені результати цих виборів є дуже мало ймовірними, що підтверджує зафіксовані спостерігачами масові порушення.

Інтерпретація отриманих результатів

Можемо бачити, що оцінки обох параметрів – і α , і γ – сильно відрізняються від теоретичних. Беручи до уваги властивості цих параметрів та інформацію з Рисунку 1, можемо припустити, що на цих виборах на значній частині ДВК були наявні грубі порушення (вкидання бюлетенів), і одночасно з тим, на значній частині ДВК подібних порушень не було – це призвело до того, що на графіку ДВК явка/“проти всіх” значна частина точок була зміщена в напрямку збільшення відсотку явки і одночасного зменшення відсотку “проти всіх” (зміщення точок вправо і вниз на Рисунку 4 нижче) – що і призвело до зміщення параметру α (параметр масштабу), збільшення параметру γ (параметр масштабу)



та виникнення двомодальності на Рисунку 1.



**Рисунок 4 – Відхилення показників явки та голосів “проти всіх” на ДВК
Луганської області**

Джерело: авторська розробка

Висновки.

Були розглянуті виборчі дані з 1762 ДВК, утворених в Луганській області під час I туру виборів Президента України 2004 року.

Було отримано статистичні висновки щодо наявності виборчих порушень на значній частині виборчих дільниць.

Фінансова підтримка.

Цю роботу частково підтримано грантом від Simons Foundation (SFI-PD-Ukraine-00017674, Krykun I. H.).



Вираз вдячності.

Автор висловлює глибоку вдячність славетним бійцям Сил Оборони України, що з 2014 року захищають автора та його рідних від кривавих російських терористів.

Також автор висловлює вдячність за допомогу зі статистичними даними проекту "Електоральна пам'ять" (ukr.vote) громадської організації "Український центр суспільних даних" (<https://socialdata.org.ua>) та його керівнику Сергію Васильченку, а також Роману Удоту, Співголови Правління Руху на захист прав виборців «Голос».

Література:

1. Заключний звіт Місії ОБСЄ/БДПІЛ зі спостереження за виборами. <https://www.osce.org/files/f/documents/8/2/14674.pdf>
2. Myagkov M. The Forensics of Election Fraud: Russia and Ukraine. New York: Cambridge University Press, 2009.
3. Krykun I. H., Pavlov M. S. Statistics of Electoral Systems and Methods of Election Manipulation // Journal of Social Mathematical & Human Engineering Sciences, 2023. – Vol.1, №1. – Pp. 11–21. DOI: 10.31586/jsmhes.2023.610
4. Krykun I. H. New Approach to Statistical Analysis of Election Results // International Journal of Mathematical, Engineering, Biological and Applied Computing, 2022. – Vol. 1 № 2. – Pp. 68–76. DOI: 10.31586/ijmebac.2022.466
5. Крикун І.Г. Арктангенс-регресія та оцінювання параметрів розподілу Коші // Український математичний вісник – 2020. – Т. 17, № 2.– С. 196-214.
6. Krykun I. H. The Arctangent Regression and the Estimation of Parameters of the Cauchy Distribution // Journal of Mathematical Sciences, 2020. – Vol. 249, Issue 5. – Pp. 739-753. DOI: 10.1007/s10958-020-04970-3
7. Krishnamoorthy K. Handbook of Statistical Distributions with Applications. New York: Chapman and Hall/CRC, 2006.
8. Центральна виборча комісія України, вибори Президента України 2004.



<https://cvk.gov.ua/pls/vp2004/wp0011>

9. Graphical data. <https://www.electoral.graphics/ru-ru/histogram-generator>

Abstract. *This study examines the results of the first round of the 2004 Presidential Elections of Ukraine at the polling stations in the Luhansk region. A statistical analysis of the published results was conducted using the indicators of voter turnout and the number of votes "against all" candidates. For this purpose, an arctangent regression of this data was constructed, according to the method previously proposed by the author, and statistical estimates of the parameters of the corresponding Cauchy distribution were found. Based on the estimates obtained, conclusions were drawn regarding possible manipulations with the election results.*

Key words: *arctangent regression, Cauchy distribution, electoral results, statistical estimation of distribution parameters.*

Статтю надіслано: 09.11.2025 року

© Крикун І.Г.