

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ  
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ  
імені ІГОРЯ СІКОРСЬКОГО»**

Приладобудівний факультет

Кафедра автоматизації експериментальних досліджень

«На правах рукопису»  
УДК 519.24

До захисту допущено:  
Завідувач кафедри  
\_\_\_\_\_ Володимир ЄРЕМЕНКО  
«\_\_» \_\_\_\_\_ 2020 р.

**Магістерська дисертація**

**на здобуття ступеня магістра**

**зі спеціальності: 152 «Метрологія та інформаційно-вимірювальна техніка»**

**на тему: «Дослідження ефективності робастних методів оцінювання  
середнього значення»**

Виконав :  
студент VI курсу, групи ПА-91мп  
Володарський Валентин Артемович

\_\_\_\_\_

Керівник:  
доцент, к.т.н.  
Добролюбова М.В.

\_\_\_\_\_

Консультант з “Розробки стартап-проектів”:  
доцент, д.е.н.  
Бояринова К.О.

\_\_\_\_\_

Рецензент:

\_\_\_\_\_

Засвідчую, що у цій магістерській  
дисертації немає запозичень з праць  
інших авторів без відповідних посилань.

Студент (-ка) \_\_\_\_\_

Київ – 2020 року

**Національний технічний університет України  
«Київський політехнічний інститут  
імені Ігоря Сікорського»**

Факультет – Приладобудівний

Кафедра інформаційно-вимірювальних технологій

Рівень вищої освіти – другий (магістерський) за освітньо-професійною (освітньо-науковою) програмою «Метрологія та вимірювальна техніка»

Спеціальність 152 «Метрологія та інформаційно-вимірювальна техніка»

ЗАТВЕРДЖУЮ  
Завідувач кафедри  
\_\_\_\_\_ Володимир ЄРЕМЕНКО

«\_\_\_» \_\_\_\_\_ 20\_\_ р.

**ЗАВДАННЯ**  
**на магістерську дисертацію студенту**  
**Володарський Валентин Артемович**  
\_\_\_\_\_  
(прізвище, ім'я, по батькові)

1. Тема дисертації Дослідження ефективності робастних методів оцінювання середнього значення \_\_\_\_\_

науковий керівник дисертації \_\_\_\_\_,  
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затверджені наказом по університету від «\_\_\_» \_\_\_\_\_ 20\_\_ р. № \_\_\_\_\_

2. Строк подання студентом дисертації 09.12.2020 р.

3. Об'єкт дослідження вплив випадкових величин на результати вимірювання

4. Предмет дослідження (вихідні дані – для магістерської дисертації за освітньо-професійною програмою) точність алгоритмів оцінювання середніх значень \_\_\_\_\_

5. Перелік завдань, які потрібно розробити створити вибірку з 9 елементів, мінімальне і максимальне значення якої є «викидами». Знайти середнє значення та СКВ при виключенні викидів за критерієм Граббса. Побудувати ітераційну процедуру визначення середнього значення та СКВ при абсолютному та відносному збільшенні максимального значення. Зробити висновки.

6. Перелік графічного (ілюстративного) матеріалу графічні зображення результатів дослідження \_\_\_\_\_

7. Орієнтовний перелік публікацій тези на всеукраїнську науково-практичну конференцію студентів, аспірантів та молодих вчених «ефективність інженерних рішень у приладобудуванні»

8. Консультанти розділів дисертації\*

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Стартап-проект	Бояринова К.О., к.е.н., доцент, викладач кафедри менеджменту		

9. Дата видачі завдання 15.09.2020 року.

Календарний план

№ з/п	Назва етапів виконання магістерської дисертації	Строк виконання етапів магістерської дисертації	Примітка
1	Огляд існуючих аналогічних рішень. Задача оцінювання середнього значення	03.09.2020 – 24.09.2020	
2	Робастні методи оцінювання середнього значення	25.09.2020 – 02.10.2020	
3	Розробка моделі для дослідження ефективності робастних методів оцінювання середнього значення	03.10.2020 – 23.10.2020	
4	Тестування розробленої моделі	24.10.2020 – 19.11.2020	
5	Оформлення стартап проєкту	20.11.2020 – 25.11.2020	
6	Висновки та оформлення роботи	26.11.2020 – 01.12.2020	

Студент

\_\_\_\_\_  
(підпис)

Валентин ВОЛОДАРСЬКИЙ

(ім'я, прізвище)

Науковий керівник дисертації

\_\_\_\_\_  
(підпис)

Марина ДОБРОЛЮБОВА

(ім'я, прізвище)

\* Консультантом не може бути зазначено наукового керівника

## РЕФЕРАТ

**Магістерська дисертація на тему:** «Дослідження ефективності робастних методів оцінювання середнього значення», 76 сторінок, 3 додатки, 12 літературних джерел.

**Об'єкт дослідження:** Способи та критерії виявлення аномальних результатів (промахів).

**Предмет дослідження:** Ефективність робастних методів та їх програмно-алгоритмічна організація, збіжність ітераційних алгоритмів

**Мета роботи:** Метою роботи є аналіз ефективності ітераційного робастного алгоритму визначення, на підставі вибіркового даних, параметрів закону розподілу, які відповідають правильності та відтворюваності методики випробування.

**Методи дослідження та апаратура:** Робота з інформаційними джерелами та літературою. Експериментальне дослідження з застосуванням комп'ютерного моделювання.

**Результати роботи та їхня новизна:** Створено алгоритм та його програмна реалізація, що дозволяє, при наявності у вибіркового даних викидів, статистично надійно оцінювати параметри розподілу. Показано, що робастні методи зберігають свою ефективність (збіжність алгоритму, число кроків ітерації) навіть у випадку, коли мають місце «масковані» викиди.

**Рекомендації щодо використання результатів роботи:** Розроблений алгоритм та його програмне забезпечення можуть бути використані у практиці опрацювання вибірок малих обсягів. Отримані результати будуть стійкими до викидів.

## ABSTRACT

**Master's thesis:** "Study of the effectiveness of robust methods of estimating the average value", 76 pages, 3 appendices, 12 references.

**The object of study:** Methods and criteria for detecting abnormal results (misses).

**Subject of research:** Efficiency of robust methods and their program-algorithmic organization, convergence of iterative algorithms

**Objective:** The aim of the work is to analyze the efficiency of the iterative robust algorithm for determining, on the basis of sample data, the parameters of the distribution law that correspond to the correctness and reproducibility of the test method.

**Methods and apparatus:** Work with information sources and literature. Experimental study using computer simulation.

**The results and their novelty:** An algorithm and its software implementation have been created, which allows, in the presence of sample emissions, to statistically reliably estimate the distribution parameters. It is shown that robust methods retain their efficiency (algorithm convergence, number of iteration steps) even in the case of "masked" emissions.

**Recommendations for the use of work:** The developed algorithm and its software can be used in the practice of processing small samples. The results obtained will be resistant to emissions.

## ЗМІСТ

ВСТУП	7
1. АНОРМАЛЬНІСТЬ РЕЗУЛЬТАТІВ ПРИ ЕКСПЕРИМЕНТАЛЬНИХ ДОСЛІДЖЕННЯ	10
1.1 Статистичні методи виявлення викидів	11
1.2 Критерій виявлення декількох результатів з надмірними похибками	13
2. ТОЧНІСТЬ ЛАБОРАТОРНИХ ВИПРОБУВАНЬ	17
2.1 Правильність та прецизійність лабораторних випробувань	19
3. РОБАСНІ МЕТОДИ ОЦІНЮВАННЯ	23
3.1 Властивості робастних методів	23
3.2 Розробка ітераційного алгоритму робастного оцінювання параметрів розподілу за вибірковими даними	29
3.2.1 Отримання первинних робастних оцінок	29
3.2.2 Перехід до робастної процедури	31
4. МОДЕЛЮЮЧИЙ ЕКСПЕРИМЕНТ	36
5. АНАЛІЗ ТА РОЗРОБКА СТАРТАПУ «AvCheck»	39
5.1 Опис ідеї проекту	39
5.2 Технологічний аудит проекту	42
5.3 Аналіз ринкових можливостей запуску стартап-проекту	43
5.4 Розроблення ринкової стратегії проекту	49
5.5 Розроблення маркетингової програми стартап-проекту	52
5.6 Висновки	55
ВИСНОВОК	57
СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ	58
ДОДАТОК А. ТАБЛИЦІ З РЕЗУЛЬТАТАМИ РОБОТИ РОБАСТНОГО МЕТОДУ З ЗМІНОЮ МЕДІАНИ	60
ДОДАТОК Б. ПРОГРАМНА РЕАЛІЗАЦІЯ РОБАСТНОГО АЛГОРИТМУ	63
ДОДАТОК В. СПИСОК ПУБЛІКАЦІЙ	66

## ВСТУП

Статистичне опрацювання та прийняття на їх підставі рішення базуються на припущенні про нормальність розподілу. Це в основному обумовлено тим, що є добре розроблена теорія статистичних висновків. Однак в ряді практичних завдань немає достатньо обсягу вихідних даних для побудови параметричних моделей. Ще в 60-і роки минулого століття видатні вчені в галузі математичної статистики на підставі результатів детальних досліджень встановили, що дані, які обробляються на підставі теорії про нормальність розподілу, як правило, в середньому близько 10 % містять грубі промахи (від 1 до 20 % ) як в явній, так і в прихованій формі.

Всі теоретичні передумови базуються на можливості проведення дослідів (здійснення спостережень) при одних і тих самих незмінних умовах. На практиці ж має місце просторово-часова мінливість умов проведення спостережень, в тому числі і мінливість самого об'єкту випробування. Базуючись на цих даних, видатний американський математик Тьюкі зробив висновок, що «нормальність – це міф і нормальності розподілу не було і ніколи не буде». Особливо дане твердження справедливо при малому обсязі вибірки. Причинами появи викидів в результатах дослідження можуть бути помилки, збої засобів вимірювальної техніки, що застосовуються при проведенні випробувань, недотримання правил проведення експерименту, помилки і промахи при оформленні результатів дослідження, зовнішні фактори та багато іншого.

З огляду на умовності закону розподілу, вид якого фактично є передбачуваною моделлю, якій повинні відповідати експериментальні дані, сама по собі реальна вибірка може мати деякі розбіжності з ідеалом (особливо при малих обсягах) – включати деякі значення, які підпорядковуються іншим законом, а не передбачуваному. Проте, параметричний підхід з припущенням, що закон розподілу відомий (повинен бути нормальним) настільки глибоко

увійшов в практику статистичної обробки даних, що недоцільно було від нього відмовлятися.

На практиці наявність в вибірках навіть невеликого числа спостережень, що різко виділяються, тобто є аномальними, здатна кардинально змінити результат статистичного дослідження, і значення отримані в кінцевому підсумку, стають недостовірними, а в деяких випадках взагалі перестануть нести будь-яку інформацію. Для того, щоб уникнути подібної ситуації, традиційно застосовують статистичні критерії, які дозволяють виділити, а потім і виключити аномальні значення, які називають «викидами» [1]. Даний підхід має право і є ефективним для вибірок великого обсягу. Однак для вибірок малого обсягу, що має місце при проведенні випробувань дорогих, унікальних об'єктів або при випробуваннях з руйнуванням, без відновлення, необхідно з певною обережністю користуватися даними способами. Це обумовлено тим, що при вибірках малого обсягу використовувані статистичні критерії втрачають чутливість до потенційних «викидів». Крім того, виключення будь-якого аномального результату з наявних експериментальних даних знижує статистичну надійність одержуваної оцінки. Наприклад, відношення вибіркового значення СКВ до його математичного сподівання при числі спостережень  $n = 4$  складає 42 %, а при  $n = 3$  воно стає рівним 52 %. Як можна зазначити, виключення одного результату з наявних даних при малих обсягах вибірки призводить до зменшення статистичної надійності приблизно на 10 %.

Ще складніша буде ситуація при проведенні випробувань, де основною умовою забезпечення точності отриманих результаті і довіри до них є відповідність параметрів «вибіркового» розподілу, генеральному розподілу [2].

Тому необхідно вводити у практику та розвивати нові статистичні методи, що дозволяють отримувати достовірні результати і при цьому не втрачати їх статистичну надійність.

Виходячи з наведеного вище можна прийти до висновку, що тема роботи є **актуальною**.



*Метою роботи* є дослідження ефективності роботи ітераційного робастного алгоритму, використовуючи вибіркові дані та параметри закону розподілу, які відповідають відтворюваності та правильності методики випробування.

*Об'єктом дослідження* даної роботи є способи та критерії виявлення аномальних результатів.

*Предмет дослідження* роботи – ефективність робастних методів та їх програмно-алгоритмічна організація, збіжність ітераційних алгоритмів.

Слід зазначити, що експериментальне дослідження в даній роботі проведене з використанням комп'ютерного моделювання.

**Результати роботи та їх новизна** полягають у створенні алгоритму, який дозволяє – при наявності у вибіркових даних викидів – статистично надійно оцінювати параметри розподілу, та його програмній реалізації. За допомогою розробленого алгоритму продемонстровано, що робастні методи зберігають свою ефективність (збіжність алгоритму, число кроків ітерації) навіть у випадку наявності «маскованих» викидів.

**Рекомендації щодо використання результатів роботи:** Розроблений алгоритм та його програмне забезпечення можуть бути використані при опрацюванні вибірок малих обсягів. Отримані результати будуть стійкими до викидів.

## **1 АНОМАЛЬНІСТЬ РЕЗУЛЬТАТІВ ПРИ ЕКСПЕРИМЕНТАЛЬНИХ ДОСЛІДЖЕННЯХ**

Головною та вирішальною ознакою методики виконання вимірювань (випробувань) є гарантія необхідної точності вимірювань. Раніше відповідно до вимог нормативних документів кожному результату дослідження приписувалася похибка, розрахована за результатами метрологічної атестації. На сучасному етапі, коли випробування можуть проводитися в різних лабораторіях і навіть в різних країнах для атестації методики випробувань при заданих умовах, приймаючи до уваги багатостадійність етапів проведення випробувань, цього недостатньо. Необхідно оцінити та нормувати показники точності методики, які враховували б можливі впливи умов проведення випробувань, характерні особливості при організації експерименту в лабораторії. Потрібні нові підходи до нормування показників на основі статистичних методів.

Кожному типу експерименту відповідає свій математичний апарат, відповідно до якого проводиться опрацювання даних експерименту і отримання значень показників якості, які характеризують методику випробування в цілому.

Для всіх типів експериментів має бути забезпечено всебічний та кваліфікований виклад методів, критеріїв та процедур первинної обробки даних, що є засобом отримання надійних оцінок статистичних показників точності.

Аналіз отриманих даних повинен включати:

- формальний аналіз вихідних і проміжних даних з метою знаходження викидів і квазівикидів;
- неформальне вивчення кожного викиду і квазівикиду із залученням додаткової інформації про можливі впливаючі фактори, які супроводжували відповідні дані;

– прийняття рішень про видалення викидів або їх збереження в масиві оброблюваних даних.

### 1.1 Статистичні методи виявлення викидів

Наявність в вибірках навіть невеликого числа спостережень, які різко виділяються, здатне кардинально змінити результат статистичного дослідження, і значення, отримані в кінцевому підсумку, стають недостовірними, а в деяких випадках взагалі перестануть нести будь-який здоровий глузд. Для того, щоб уникнути подібної ситуації, традиційно застосовують статистичні критерії, які дозволяють виділити, а потім і виключити аномальні дані, які називають викидами.

Якщо розподіл результатів вимірювання нормальний, то такою оцінкою є середнє арифметичне. Але середнє арифметичне дуже чутливе до відхилень від прийнятої моделі розподілу, що виявляються у вигляді промахів (результатів вимірювання, які мають надмірні похибки). Окремі промахи виявляють за допомогою відповідних критеріїв і вилучають з вибірки.

Правила оцінки аномальності результатів вимірювання класифікують за наявністю апіорної інформації щодо відомих генерального середнього і генерального середнього квадратичного відхилення (СКВ). Є три критерії оцінки аномальності [1]:

1. *Критерій оцінки аномальності результатів вимірювання, якщо генеральне середнє  $\mu$  і генеральне СКВ  $\sigma$  не відомо.* Для впорядкованої вибірки результатів вимірювань  $x_1 \leq x_2 \leq \dots \leq x_n$  обчислюють вибіркове середнє  $\bar{x}$  і вибіркове СКВ

$$\bar{x} = \sum_{i=1}^n x_i / n ; S = \sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 / (n - 1)} \quad (1)$$

Щоб оцінити належність  $x_n$  чи  $x_1$  до певної нормальної сукупності і прийняти рішення про вилучення чи залишення  $x_n(x_1)$  у складі вибірки, знаходять відношення, яке є статистикою (критерієм) Граббса [2]:

$$G_n = \frac{x_n - \bar{x}}{s} \quad (2)$$

для максимального значення впорядкованого ряду, чи

$$G_1 = \frac{\bar{x} - x_1}{s} \quad (3)$$

для мінімального значення.

Результат порівнюють з табличним значенням  $G_{кр}$  для певного об'єму вибірки  $n$  і прийнятої ймовірності  $\alpha$  помилкового прийняття рішення. Якщо розрахункове значення більше табличного, то підозрюваний в аномальності результат вимірювання вилучають, якщо інакше – залишають.

2. *Критерій оцінки аномальності результатів вимірювання, якщо відомо генеральне СКВ  $\sigma$  і не відомо генеральне середнє  $m$ .* Для впорядкованої вибірки результатів вимірювань  $x_1 \leq x_2 \leq \dots \leq x_n$  обчислюють  $\bar{x}$  і відношення

$$t_n = \frac{x_n - \bar{x}}{\sigma}; \quad t_1 = \frac{\bar{x} - x_1}{\sigma}, \quad (4)$$

яке порівнюють з табличним значенням при певному об'ємі вибірки і критичним значенням  $\alpha$  [1].

Якщо  $t_n \geq h$  ( $t_1 \geq h$ ), то результат вимірювання  $x_n$  ( $x_1$ ) вилучають, якщо інакше, то залишають.

3. *Критерій оцінки аномальності результатів вимірювання, якщо відомо генеральне середнє  $m$  і генеральне СКВ  $\sigma$ .* Для оцінки аномальності результатів вимірювання  $x_n$  чи  $x_1$ , якщо  $x_1 \leq x_2 \leq \dots \leq x_n$ , обчислюємо значення коефіцієнта.

$$V_n = \frac{x_n - m}{\sigma} \text{ чи } V_1 = \frac{m - x_1}{\sigma} \quad (5)$$

І порівнюємо його з табличним критичним, взятим з відповідної таблиці. Якщо  $t_n \geq h$  ( $t_1 \geq h$ ), то  $x_n$  ( $x_1$ ) вважають аномальним і вилучають з вибірки, якщо інакше, то залишають.

Подані вище критерії можна використовувати для цензурованої вибірки, тобто вибірки з вилученим результатом, що його підозрюють в аномальності.

Замість першого критерію аномальності використовують критерій:

$$U_n^{\cdot} = \frac{x_n - \bar{x}_{n-1}}{S_{n-1}}; \bar{x}_{n-1} = \frac{1}{n-1} \sum_{i=1}^{n-1} x_i^{\cdot}, \quad (6)$$

$$S_{n-1} = \sqrt{\frac{1}{n-2} \sum_{i=1}^{n-1} (x_i^{\cdot} - \bar{x}_{n-1})^2}, \quad (7)$$

де  $x_i^{\cdot}$  - результат цензурованої вибірки,  $\bar{x}_{n-1}$ ,  $S_{n-1}$  - середнє і вибіркове СКВ цензурованої вибірки.

Замість другого критерію аномальності використовують критерій

$$t_n^{\cdot} = \frac{x_n - \bar{x}_{n-1}}{\sigma} \quad (8)$$

У цьому разі значення  $t_n^{\cdot}$  порівнюють зі значенням  $h$  для цензурованої вибірки.

## 1.2 Критерій виявлення декількох результатів з надмірними похибками

Критерій виявлення аномальних результатів, який було запропоновано Граббсом і Смирновим, використовують для одного екстремального результату вимірювання. Якщо за вирішальним правилом підозрілий результат є «викидом», його виключають з вибірки і процедуру повторюють з результатами, що залишилися.

Процедуру цензурування закінчують, якщо за вирішальним правилом отримано висновок, що аномальних результатів немає. У такому разі виникає проблема, пов'язана з, так званим, «маскувальним ефектом» [3]. Підозрілі результати часто групуються близько один до одного створюючи групу у віддаленні від основної маси результатів. Виникає питання, що робити з критерієм Граббса-Смірнова, який нечутливий до таких результатів.

Ефективність цих процедур пов'язана з тим, що перехід від усієї маси результатів  $x_{(1)} \dots x_{(N)}$  (де  $x_{(i)}$  – член вибірки після ранжування) до  $x_{(1)} \dots x_{(N-1)}$  дуже змінює значення критерію, якщо є результат з надмірною похибкою.

Цього не виникає, якщо є угруповання аномальних результатів. Приклад з [3], для ілюстрації «маскувального ефекту» наведено нижче.

Перевіримо наявність аномального результату у вибірці: 10, 11, 10, 12, 11, 11, 11, 17. Підозрілий результат, відносно якого формулюють критерій - 17.

За результатами обчислень отримуємо

$$\bar{x} = 11,63; S = 1,22; G_n = \frac{17-11,63}{1,22} = 4,4. \quad (9)$$

Табличне значення для  $\alpha = 0,05$  становить 2,03. Значить результат 17 необхідно вилучити.

Якщо таких результатів два (вибірка 10, 11, 10, 12, 11, 11, 11, 16, 17), то за результатами обчислень отримуємо:

$$\bar{x} = 12,1; S = 2,6; G_n = \frac{17-12,1}{2,6} = 1,89. \quad (10)$$

Табличне значення для  $\alpha = 0,05$  вже буде 2,11. Таким чином, так званий «маскувальний ефект» ще одного аномального результату зумовлює висновок, що аномальних результатів немає.

Процедури виявлення  $k$  ( $k > 1$ ) екстремальних результатів запропонували американські статистики Г. Тітсен і Г. Мур [4]. Процедури розраховано на нормальний розподіл результатів вимірювань.

Вирішальне правило для виключення  $k$  найбільших результатів будується на значенні величини

$$L_k = \frac{\sum_{i=1}^{N-k} (x_{(i)} - \bar{x}_k)^2}{\sum_{i=1}^N (x_{(i)} - \bar{x})^2}, \quad (11)$$

де  $x_{(1)} \leq \dots \leq x_{(N)}$  – варіаційний ряд для вибірки  $x_1 \dots x_N$ .

$$\bar{x}_k = \frac{1}{(N-k)} \sum_{i=1}^{N-k} x_{(i)}, \quad (12)$$

де  $\bar{x}_k$  – середнє арифметичне  $N - k$  результатів, що залишилися після вилучення  $k$  найбільших результатів.

Такого самого типу величину (позначимо її  $\widetilde{L}$ ), що і (11), тільки із заміною чисельника на  $N - k$  найбільших значень, використовують у критерії виключення  $k$  найменших результатів:

$$\widetilde{L}_k = \frac{\sum_{i=k+1}^N (x_{(i)} - \underline{x}_k)^2}{\sum_{i=1}^N (x_{(i)} - \bar{x})^2}, \quad (13)$$

де  $\underline{x}_k$  – середнє арифметичне  $N - k$  вимірювань, що залишились після вилучення  $k$  найменших результатів:

$$\underline{x}_k = \frac{1}{(N-k)} \sum_{i=k+1}^N x_{(i)}. \quad (14)$$

Якщо підозрілими є як найбільші, так і найменші члени вибірки, то для їх виявлення Тітьєн і Мур пропонують таке правило. За вибіркою  $x_1, \dots, x_N$  обчислюють середнє арифметичне  $\bar{x}$ . Після цього обчислюють абсолютні відхилення  $r_1 = |x_1 - \bar{x}|, r_2 = |x_2 - \bar{x}|, \dots, r_N = |x_N - \bar{x}|$ . Проранжуємо  $r_i$  і введемо позначення:  $z_i$  – результат, що відповідає абсолютному відхиленню  $r_i$ , яке є  $i$ -м по порядку. Це означає, що  $z_1$  – це найближчий результат до  $\bar{x}$ , а  $z_N$  – найвіддаленіший результат від  $\bar{x}$ . Для перевірки гіпотези виключення  $k$  найбільших за модулем відхилень результатів використовують величину

$$E_k = \frac{\sum_{i=1}^{N-k} (z_{(i)} - \bar{z}_k)^2}{\sum_{i=1}^N (z_{(i)} - \bar{z})^2}, \quad (15)$$

де  $\bar{z}_k = \frac{1}{(N-k)} \sum_{i=1}^{N-k} z_{(i)}$  – середнє арифметичне з  $N - k$  вимірювань, що залишились після виключення  $k$  екстремальних результатів,  $\bar{z}$  – середнє арифметичне всієї вибірки.

Наведено приклад вибірки: 0,916; 0,944; 1,292; 1,428; 1,452; 1,542; 1,604; 1,632. Два найменші результати 0,916 і 0,944 підозрюють в аномальності. Якщо використати критерій Граббса-Смірнова, то за результатами обчислень отримуємо:  $\bar{x} = 1,349$ ;  $S = 0,29$ ;  $G_n = \frac{1,349-0,916}{0,29} = 1,49$ . Для  $\alpha = 0,05$  визначаємо граничне значення  $h = 2,03$  і робимо висновок, що результат не підлягає вилученню. Наявність другого екстремального значення створює «маскувальний ефект».

Якщо ж скористаємося критерієм Тітьєна – Мура, то отримаємо  $\widetilde{L}_k = 0,144$ . Критичне значення  $C_{0,05} = 0,146$ . Це означає, що обидва результати можуть бути вилученими з вибірки. Отже, критерій Тітьєна – Мура дозволяє запобігти «маскувальному ефектові».

З проведеного аналізу можна зробити висновок, що наведені методи дозволяють виключити вплив промахів на обчислене середнє значення, як

оцінки величини, що вимірюється. Але чим менше елементів у вибірці, тим менше буде статистична надійність обчисленого середнього.



## 2 ТОЧНІСТЬ ЛАБОРАТОРНИХ ВИПРОБУВАНЬ

Головною, вирішальною ознакою методики виконання випробувань є гарантованість необхідної точності вимірювань. Раніше відповідно до вимог нормативних документів кожному результату дослідження приписувалася похибка, розрахована при метрологічній атестації.

На сучасному етапі, коли випробування можуть проводитися в різних лабораторіях і навіть в різних країнах для атестації методики випробувань при заданих умовах, включаючи і багатостадійну підготовку проби, цього недостатньо. Необхідно оцінити, а потім і нормувати показники точності методики, які враховували б можливі впливи умов проведення випробувань, характерні особливості при організації експерименту в лабораторії [5].

На етапі проведення випробувань основна роль відводиться методикам їх виконання.

На розсіювання результатів спостережень може впливати безліч факторів. Серед них можна виділити п'ять основних [2]:

- оператор;
- використовуване обладнання;
- калібрування;
- навколишнє середовище;
- час, що пройшов між вимірюваннями.

Варто також розглядати і стандартний зразок, тому що його параметри є складовими умов випробувань і можуть вплинути на результат при оцінюванні можливої залежності характеристики точності від значення параметра (рівня випробувань).

Аналітичним шляхом таке завдання вирішити майже неможливо, так як відсутня математична модель, що враховує зв'язок характеристик об'єкта випробувань, умов і режимів роботи об'єкта випробувань з вихідною вимірюваною величиною. Єдиним шляхом, що дозволяє вирішити цю задачу, є проведення міжлабораторних спільних випробувань (МСВ), при яких

враховуються особливості організації проведення випробувань в лабораторіях, можливі поєднання впливових величин в межах заданих (передбачуваних) норм. Сучасні тенденції до встановлення та оцінювання характеристик точності результатів випробувань полягають у відході від ідеології призначення їх допустимих меж і затвердження розробленої методики на підставі вимірів організації розробника.

Міжлабораторний експеримент дозволяє моделювати реальну різноманітність умов та можливостей проведення випробувань даної величини, що має місце в конкретному регіоні. Як результат такого дослідження є об'єктивне визначення дисперсії відтворюваності. Надалі будь-які висновки щодо прийнятності показників правильності і прецизійності можуть проводитися лише при відомому значенні відтворюваності. Зіставлення повинні бути засновані не на виявленні різниці між двома дисперсіями або між двома середніми значеннями, а на застосуванні критеріїв значимості, що використовують поняття приналежності результатів випробувань до однієї або різних множин.

Оскільки випробування однотипної продукції, виконані за єдиною методикою, можуть здійснюватися декількома лабораторіями в різних умовах, що знаходяться в заданих межах, оснащеними прописаним обладнанням з різними (індивідуальними) технічними характеристиками, а дослідження проводяться за участю різних операторів і в різний час, то результати, природно, будуть різними. При цьому в кожній лабораторії повинна виконуватися умова повторюваності.

Тому міжлабораторний експеримент фактично є фізичною моделлю реалізації процедури випробувань відповідно до методики його проведення, створюваний на основі співдружності лабораторій, що мають близький професійний рівень і спеціалізуються в даному виді випробувань.

Точність у цьому випадку виступає як інтегральний якісний показник [2], що характеризується правильністю й прецизійністю, складові яких наведені на рисунку 1.

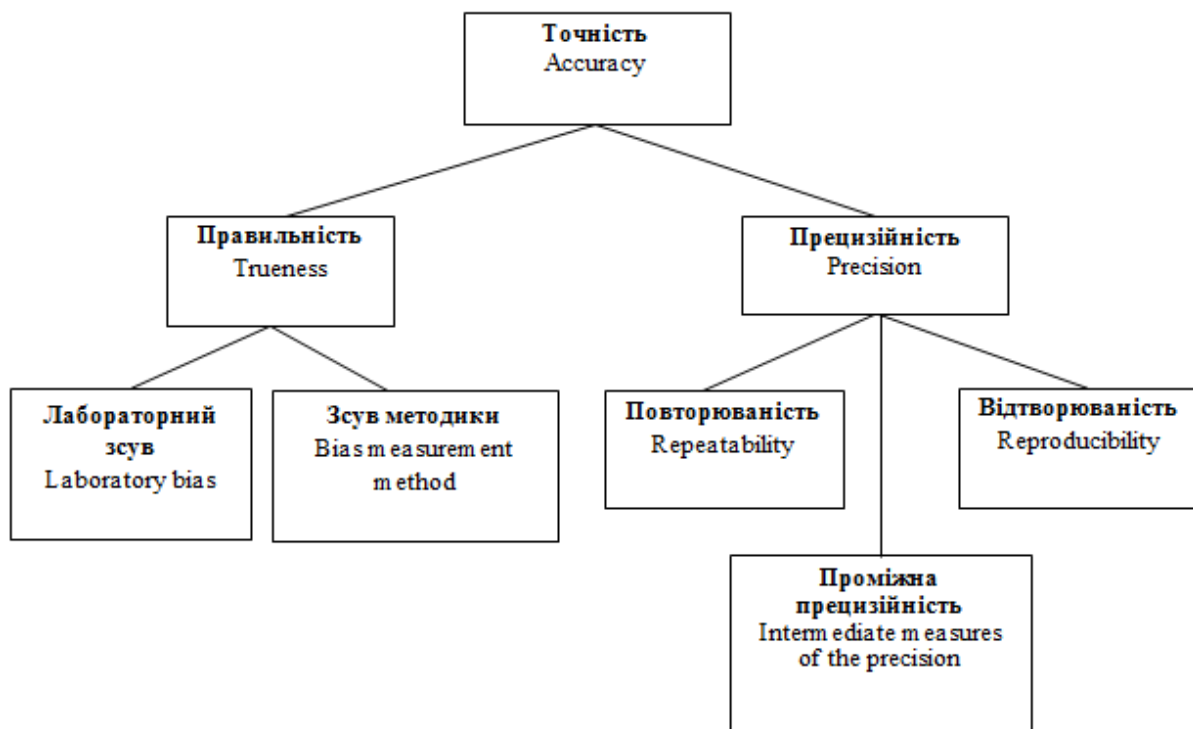


Рисунок 1 – Точність та її складові

В свою чергу правильність характеризується лабораторним зсувом та зсувом методики, а поняття прецизійності диференціюється для різних умов виконання експерименту й називається відповідно до них або повторюваністю, або проміжною прецизійністю, або відтворюваністю результату (методики).

## 2.1 Правильність та прецизійність лабораторних випробувань

**Правильність** результатів випробувань характеризує ступінь близькості середнього значення до істинного та обумовлена недосконалістю методики, недосконалістю формування або встановлення умов випробувань, конструктивними особливостями засобів виміральної техніки й випробувального устаткування тощо. Мірою правильності є систематична похибка, а оцінкою систематичної похибки є зсув (bias).

Складовими правильності є лабораторний зсув, що характеризує індивідуальні особливості організації процедури випробування в лабораторії, та зсув методики, що характеризує її правильність.

Оскільки випробування є складною експериментальною процедурою, що включає етап виділення інформативних параметрів або величин, що характеризують властивості об'єкта, то правильність методики буде залежати від якості реалізації цього етапу. Наприклад, при хімічному аналізі зсув результату може виникати, якщо застосовувана методика випробування не в змозі витягти весь компонент або якщо присутність одного компонента перешкоджає визначенню іншого, близького за властивостями компонента.

**Прецизійність.** Для одержання результату випробувань необхідно реалізувати процедуру вимірювання, яка супроводжується впливом випадкових величин, що призводить до розсіювання результатів спостережень навколо деякого середнього.

Характеристикою розсіювання результатів вимірювань або, іншими словами, їх стабільності, є прецизійність – друга, так звана експериментальна, складова точності.

Прецизійність характеризує близькість незалежних результатів випробувань, отриманих у заданих умовах, а показником прецизійності є стандартне (середньоквадратичне) відхилення результатів випробувань, виконаних у цих умовах. Причому, в якості результатів можуть розглядатися як результати одноразових спостережень, так і середні значення декількох спостережень. Для характеристики прецизійності часто використовують відносне стандартне відхилення, що є величиною безрозмірною, тому й більш наочною.

Прецизійність залежить тільки від випадкових факторів і не зв'язана ні із істинним значенням, ні із заданим. При цьому особлива увага приділяється умовам випробувань, які саме впливають на мінливість результатів.

**Повторюваність.** Як відомо, стійкість результатів вимірювань спостерігається в їхній повторюваності при незмінних умовах випробувань [6]. Тому для забезпечення умов повторюваності необхідно забезпечити сталість під час проведення експерименту всіх п'яти перерахованих вище факторів. Якщо дослідження проводяться в тій самій лабораторії на ідентичних

випробуваних об'єктах, тим самим оператором на тій самій установці за малий, у порівнянні зі змінами навколишнього середовища, проміжок часу, то розбіжності між спостереженнями будуть визначатися впливом лише випадкових величин. У більшості випадків ці розбіжності добре апроксимуються нормальним законом.

Показником повторюваності є стандартне відхилення  $\sigma_r$ , великої кількості результатів випробувань, отриманих в умовах повторюваності. Якщо випробування проводяться тільки однією лабораторією, то поняття повторюваність і прецизійність ідентичні.

**Відтворюваність.** Випробування того самого об'єкта, виконані за однією методикою, можуть здійснюватися декількома лабораторіями у різних, але таких, що перебувають у заданих межах, умовах, оснащеними устаткуванням з різними (індивідуальними) технічними характеристиками, а досліди проводяться за участю різних операторів та у різний час. Результати, отримані у такий спосіб (за умов відтворюваності), можуть відрізнятися. Цій властивості повторних випробувань відповідає особлива характеристика – відтворюваність результатів випробувань – прецизійність в умовах відтворюваності. При цьому в кожній лабораторії має виконуватися умова повторюваності.

Як показник відтворюваності використовується стандартне відхилення відтворюваності – стандартне відхилення результатів вимірювань, отриманих в умовах відтворюваності і при стабільних характеристиках однорідних об'єктів.

Будь-які висновки щодо прийнятності та відповідності при порівнянні не тільки лабораторій, а й методик за показниками правильності і прецизійності можуть здійснюватися лише при відомому значенні відтворюваності. Причому порівняння повинно базуватись на виявленні відмінностей між двома дисперсіями або між двома середніми значеннями, а не застосуванні критеріїв значимості, що використовують поняття приналежності результатів випробувань до однієї або різних множин.

При проведенні спільного міжлабораторного експерименту можлива відмінність в допустимих межах між організацією випробувального процесу в

різних лабораторіях. Це призводить до того, що будуть відрізнятися і лабораторні складові зсуву результату  $B$ . Для характеристики можливих змін у сукупності, тобто для нормування можливих відмінностей значень лабораторних зсувів вводиться міжлабораторна дисперсія

$$Var(B) = \sigma_L^2, \quad (1)$$

що обумовлено впливом можливих поєднань розбіжностей між операторами, обладнанням і організацією проведення випробувань в лабораторіях.

При чому, чим більше експериментальних даних використовується при визначенні міжлабораторної дисперсії, тим більшу статистичну надійність буде мати  $\sigma_L^2$ , яке у подальшому використовується як нормована величина.

Тому виключення лабораторних результатів, які на підставі критеріїв, що були розглянуті вище, не допустимо. Крім того, на практиці дуже часто мають місце розподіли з «важкими хвостами», коли ймовірність великих похибок перевищує допустиму для нормального розподілу. «Важкі хвости» бажано оцінювати такими методами, що дають достатню точність за наявності відхилень від прийнятої моделі. Такі методи, що називаються робастними, чи стійкими, розробляють у сучасній математичній статистиці і використовують на практиці.

### **3 РОБАСТНІ МЕТОДИ ОЦІНЮВАННЯ**

На практиці наявність в вибірках навіть невеликого числа спостережень, що різко виділяються серед інших даних, здатна кардинально змінити результат статистичного дослідження. При цьому, значення, отримані в кінцевому підсумку, стають недостовірними, а в деяких випадках взагалі перестануть нести будь-який здоровий глузд. Для того, щоб уникнути подібної ситуації, традиційно застосовують статистичні критерії, які дозволяють виділити, а потім і виключити аномальні дані, які називають «викидами» [7]. Даний підхід має право на існування і ефективний для вибірок великого обсягу.

При проведенні міжлабораторних випробувань при визначенні показників прецизійності (повторюваності та відтворюваності) методики випробувань, обсяг даних обмежений. Тому виключення викидів приведе до неправильного нормування повторюваності та відтвореності методики. А це може привести до недостовірних результатів випробувань або неправильних рішень при акредитації лабораторії.

#### **3.1 Властивості робастних методів**

При атестації характеристик точності методики випробувань, як вже відзначалося, проводять спільний експеримент, для участі в якому залучаються лабораторії, які мають відповідну професійну підготовку. При спільному експерименті існує припущення, що всі залучені лабораторії мають однакову повторюваність. На практиці ж часто виявляється, що деякі лабораторії мають гіршу повторюваність і цьому є об'єктивні причини. При визначенні за результатами вимірювання значення фізичної величини застосовують один з розглянутих вище статистичних методів виключення викидів. Але при нормуванні показників прецизійності методики випробувань або при застосуванні її при оцінюванні професійного рівня лабораторії це робити

неможливо. Покажемо це на чисельному прикладі, скориставшись даними, наведеними в [2].

Було проведено міжлабораторний експеримент, в якому брало участь 9 лабораторій. За результатами випробувань однорідних зразків отримані середні значення для кожної лабораторії:

24,140 20,155 19,500 20,300 20,705 17,570 20,100 20,940 21,185.

На підставі цих даних можна обчислити оцінки центру розподілу  $x_0 = 20,511$  та середнього квадратичного відхилення  $S_0 = 1,727$ .

Попередній аналіз представлених даних показує, що підкреслені елементи ряду (мінімальне та максимальне значення) різко відрізняються від сусідніх елементів. Перевірка крайніх елементів по критерію Граббса-Смірнова показує, що це є викиди. При класичному підході їх необхідно виключити, як аномальні.

Для семи елементів  $n = 7$  маємо  $x_7 = 20,41$  та  $S_7 = 0,501$ .

Як бачимо, середнє змінилося не суттєво, а дисперсія, яка визначає відтвореність методики «покращилась» більше ніж в 2 рази. Це могло б створити великі проблеми при акредитації лабораторій.

Таким чином, вибіркова дисперсія більше чутлива до викидів, ніж вибіркове оцінки середнього значення.

Робастні методи дозволяють наблизитись до реальних значень без втрати статистичної надійності, тобто на підставі усіх наявних даних.

Вся статистична обробка та прийняття на підставі отриманих значень рішення базуються на припущенні про нормальність розподілу. Однак в ряді практичних завдань немає достатньо обсягу вихідних даних для побудови параметричних моделей. Ще в 60-і роки минулого століття видатні вчені в галузі математичної статистики на підставі результатів детальних досліджень встановили, що дані, які обробляються на підставі теорії про нормальність розподілу, як правило, в середньому близько 10 % містять грубі промахи (від 1 до 20 %) як в явній, так і в прихованій формі [7].



Вихідною передумовою математичної статистики є той факт, що можливо багато разів проводити досліди (здійснення спостережень) при одних і тих самих незмінних умовах. На практиці ж має місце просторово-часова мінливість умов проведення спостережень, в тому числі і мінливість самого об'єкта випробування. Виходячи з цього американський математик Тьюки прийшов до висновку [8], що «нормальність – це міф і нормальності розподілу не було і ніколи не буде». Особливо дане твердження справедливо при малому обсязі вибірки. Причинами появи викидів в результатах дослідження можуть бути помилки, збої засобів вимірювальної техніки, що застосовуються при проведенні випробувань, недотримання правил проведення експерименту, помилки і промахи при оформленні результатів дослідження, зовнішні фактори і багато іншого.

Оскільки фактично висувається гіпотеза про вид закону розподілу, то вважається, що експериментальні дані мають відповідати цьому закону. В дійсності, сама по собі реальна вибірка може мати деякі розбіжності з ідеалом (особливо при малих обсягах), тобто містити деякі значення, які підпорядковуються іншому розподілу, а не передбачуваному. Проте, параметричний підхід з припущенням, що закон розподілу відомий (повинен бути нормальним) настільки глибоко увійшов в практику статистичної обробки даних, що недоцільно від нього відмовлятися.

Модель засмічення характеризується розтягнутими «хвостами» щільності ймовірності. У схемі з засміченням є середня стійка частина розподілу, яка обумовлена звичайними малими складовими похибки, і розтягнуті «хвости».

Деяка центральна частина розподілу експериментальних даних відповідає припущенням про розподіл генеральної сукупності. Для даних, які знаходяться в цій частині розподілу доцільно використовувати метод найменших квадратів (МНК). Модульний критерій, запропонований Лапласом, є більш стійким до викидів, ніж МНК, тобто дає найкращий результат при найбільш несприятливому розподілі.

Як витікає з рисунку 2,  $\rho(x) = x^2$  для значень  $|x| \leq c$ , та при великих абсолютних значеннях  $|x|$  функція втрат буде приймати значення  $\rho(x) < x^2$ .

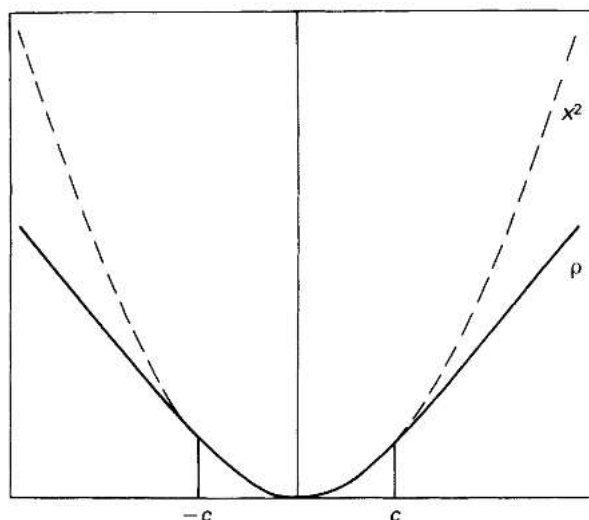


Рисунок 2 – Робастна функція втрат  $\rho(\varepsilon)$

Виходячи з вищенаведених міркувань, при побудові робастних методів робиться «симбіоз» – для деякої центральної групи використовується метод найменших квадратів, а починаючи з певної межі, для зменшення впливу викидів, але зі збереженням даних, застосовується модульний критерій.

Щоб зменшити чутливість до викидів в [9] запропоновано мінімізувати функціонал

$$\sum_{i=1}^n \rho(x_i - \mu), \quad (1)$$

де  $\rho(\varepsilon)$  – функція втрат, що дозволяє менш «строго» підходити до відбраковування викидів, які віддалені від центра.

Константа  $c$  регулює ступінь робастності і значення її залежить від ступеня «засмічення». Так, при «засміченні» 1 %,  $c = 2$ , а при «засміченні» 5 %,  $c = 1,4$ . Зазвичай вибирають значення  $c = 1,5$ .

Відповідно до обраного критерію необхідно провести модифікацію наявних даних, а саме:

$$x_i^* = \begin{cases} x_i & \text{при } |x_i - \hat{\mu}| < c\sigma; \\ \hat{\mu} - c\sigma & \text{при } x_i \leq \hat{\mu} - c\sigma; \\ \hat{\mu} + c\sigma & \text{при } x_i \geq \hat{\mu} + c\sigma, \end{cases} \quad (2)$$

де  $x_i^*$  – попередньо ранжовані  $x_i$  в порядку зростання.

Проведені дослідження [9] показали, що найкращими властивостями з точки зору стійкості до викидів має середина інтервалу, що знаходиться між вибірковими кuartилями. Тому в якості початкової оцінки центру розподілу, стійкої до викидів, береться вибірка медіана (рис. 3). У якості стійкого (стабільного) береться інтерквартильний інтервал (interquartile range) – різниця між значеннями третього  $p = 3/4$  і першого  $p = 1/4$  кuartилей.

Інтерквартильний інтервал є характеристикою розкиду розподілу.

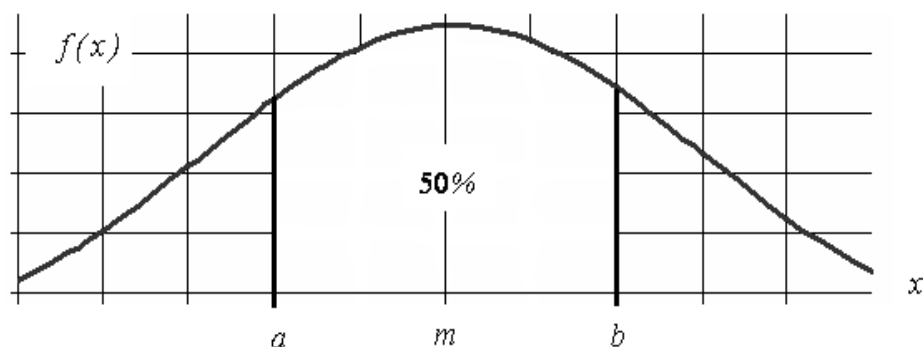


Рисунок 3 – Визначення інтерквартильного інтервалу,  
 $a$  і  $b$  –відповідають ординатам першого и третього кuartилей

У припущенні про можливий законі розподілу довжина інтервалу однозначно відповідає дисперсії цього розподілу.

Генеральне стандартне відхилення  $\sigma$ , яке зустрічається у формулі (2), зазвичай не відомо. Як початкова оцінка масштабу переходу від повного розподілу до усіченого вводиться медіана абсолютних відхилень MAD (median absolute deviation)

$$MAD_n = med\{|x_i - M_n|\} \quad (2)$$

де,  $M_n = med\{x_i\}$ ,  $x_i$  – елемент виборки, а індекс  $n$  відповідає числу елементів у виборці.

Для того, щоб встановити взаємозв'язок між параметрами «усіченого» і передбачуваного генерального розподілу сукупності, тобто виконати умову масштабування, необхідно здійснити перерахунок СКВ ( $\sigma$ ), скориставшись коригувальним коефіцієнтом, який визначають з вихідної щільності розподілу.

$$f(x) = \frac{d}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}} \quad (3)$$

При «неусіченому» розподілі  $d = 1$ , при інтерквартильному «усіченому» розподілі

$$d = \frac{0,5}{\Phi\left(\frac{b-m}{\sigma}\right) - \Phi\left(\frac{a-m}{\sigma}\right)} \quad (4)$$

де  $\Phi$  – інтегральна нормована функція нормального розподілу. Скориставшись таблицями для нормального розподілу [10], знаходять значення

$$d = \frac{1}{0,6745} = 1,4826 \approx 1,483 \quad (5)$$

Таким чином, первинна оцінка СКВ  $s^*$ , для вибірки обсягу  $n$ , знаходиться на підставі нормованого інтерквартильного розмаху і складає

$$s^* = 1,483 \cdot MAD_n \quad (6)$$

Отже, для реалізації критерію стійкості до викидів на підставі обраної функції втрат, необхідно визначити медіану вихідного масиву даних. Потім, скориставшись виразом (6), обчислити значення, яке разом з обраним  $c = 1,5$  дозволить встановити початкову (вихідну) границю переходу від методу найменших квадратів до методу найменшого модуля (МНМ), а саме

$$\varphi = cs^* \quad (7)$$

На початковому етапі в якості оцінки центру розподілу необхідно брати медіану вихідного ряду, яка більш стійка до викидів, ніж середнє значення. На наступних етапах ітераційної процедури використовується уточнене середнє. Обчислення проводяться по відношенню до «усіченого» розподілу, в якому при  $c = 1,5$  знаходиться 86,6 % повного розподілу. Тому, для визначення уточненого

СКВ для «неусіченого» нормального розподілу, проводиться масштабування шляхом введення коефіцієнта 1,134 [11].

### 3.2. Розробка ітераційного алгоритму робастного оцінювання параметрів розподілу за вибірковими даними

Для демонстрації реалізації робастного алгоритму, запропонованого Хьюбером [12], скористаємося наведеними у [2] даними, які були отримані при проведенні міжлабораторного спільного експерименту при залучені  $n=9$  лабораторій:

24,140 20,155 19,500 20,300 20,705 17,570 20,100 20,940 21,185

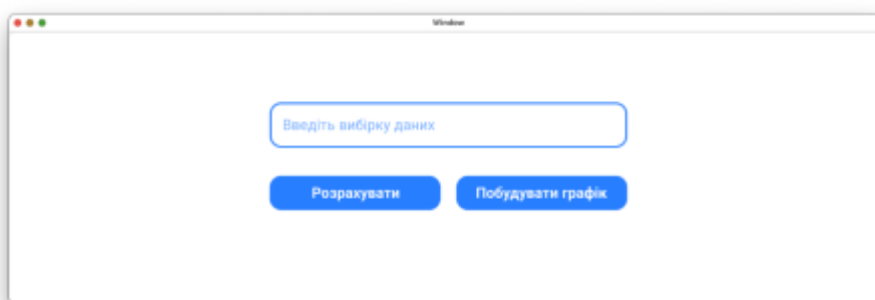


Рисунок 4 – початковий інтерфейс для роботи з алгоритмом

Як було сказано вище, підкреслені елементи (мінімальний та максимальний), виходячи зі статистичних критеріїв, наприклад Грабса-Смірнова, мають бути виключеними. Але, при використанні робастної процедури опрацювання даних сумісного експерименту, використовуємо усі наявні дані і будемо їх «вганяти» в модель, яка визначається інтерквартильним інтервалом. Центром розсіювання на первинному етапі, його оцінкою, вибирається медіана вхідних даних, які були у лабораторіях.

#### 3.2.1 Отримання первинних робастних оцінок

Проведемо ранжування даних (від меншого до більшого) і занесемо у стовпчик «0» ( $j = 0$ ) Таблиці 1. Тут і в подальшому будемо позначати результат, який після ранжування, у чисельному ряді займає  $i$  – порядкове місце, як  $x_i^*$ .

Обчислимо середнє значення

$$\bar{x}_0^* = \frac{1}{n} \sum_{i=1}^9 x_{i(0)} = 20.511 \quad (1)$$

та середнє квадратичне відхилення (СКВ) для цього стовпчика:

$$S_0^* = \sqrt{\frac{1}{n-1} \sum_{i=1}^9 (x_{i(0)}^* - \bar{x}_0^*)^2} = 1.727 \quad (2)$$

Занесемо обчислені початкові значення середнього та СКВ у стовпчик  $j=0$  Таблиці 1. Вони не будуть відповідати дійсності, оскільки було встановлено, що наявні дані включають викиди. Щоб знайти, на підставі наявних експериментальних даних параметри методики випробувань - зміщення В та можливе розсіювання лабораторних результатів, застосуємо робастну процедуру. Це дозволяє використовувати усі результати, отримані у лабораторіях і тим самим, виходячи з умов проведення спільного експерименту, забезпечити статистичну надійність.

Було показано, що стійкою до викидів оцінкою центру розподілу, є вибіркова медіана. При застосуванні робастного методу саме медіану використовують як початкове наближення до центру розподілу.

Аналіз ранжируваного чисельного ряду, представленого у стовпчику  $j=0$ , дав результат, що медіаною ряду є  $x_{5(0)}^* = 20,300$ , тобто

$$\text{med}(x_{i(0)}^*) = x_{5(0)}^* = 20,300 \quad (3)$$

$$j s^* = 1,483 \cdot MAD_9 \quad (4)$$

де коефіцієнт 1,483 враховує перехід від «усіченого» розподілу до повного.

Для знаходження  $MAD_9$  – медіани абсолютних відхилень  $\Delta_i$ , необхідно обчислити абсолютні відхилення елементів стовпчика  $j=0$  від їх медіани

$$\Delta_i = |x_{i(0)}^* - x_{5(0)}^*| \quad (5)$$

Отримаємо ряд абсолютних різниць:

$$2,73 \quad 0,80 \quad 0,20 \quad 0,145 \quad 0 \quad 0,405 \quad \underline{0,640} \quad 0,889 \quad 3,84.$$

Звідки витікає, що  $MAD_9 = 0,640$  (підкреслене значення, яке дозволяє визначити приблизне значення стійкої оцінки СКВ):

$$s^*_{(0+1)} = 1,483 \cdot 0,640 = 0,949 \quad (6)$$

Занесемо знайдене значення  $s_{(0+1)}^*$  в останній рядок стовпчика  $j=0$  – нове, вже стійке до викидів, значення СКВ, яке використовується при переході до уточнюючого наступного кроку робастної ітераційної процедури.

### 3.2.2 Перехід до робастної ітераційної процедури

Обираємо, як рекомендоване значення,  $c=1,5$  та при  $s_{(0+1)}^* = 0,949$ , визначимо за виразом (7) розділу 3.1 точку переходу від методу найменших квадратів МНК до методу найменших модулів МНМ:

$$\varphi_1 = 1,5 \cdot 0,949 = 1,424.$$

Потім знаходимо граничні значення переходу від МНК до МНМ:

$$\text{- нижню } x_0^* - \varphi_1 = 20,300 - 1,424 = 18,876; \quad (7)$$

$$\text{- верхню } x_0^* + \varphi_1 = 20,300 + 1,424 = 21,724. \quad (8)$$

Порівняємо значення  $x_{i(0)}^*$  ( $i = 1, 2, \dots, 9$ ) для стовпчика  $j = 0$  з цими граничними значеннями, представленими виразами (7) та (8).

Оскільки

$$x_{1(0)}^* = 17,570 < 18,876, \quad (9)$$

то замість  $x_{1(0)}^*$  у стовпчику  $j=1$  (перший крок ітерації) заноситься  $x_{1(1)}^* = 18,876$ , тобто приймається значення нижньої границі.

Оскільки у якості робастного центру розподілу береться медіана вихідних даних, серед яких  $x_{9(1)}^*$  приблизно дорівнює розрахованому верхньому граничному значенню, то у стовпчику  $j=1$  найбільшому елементу приписується значення верхньої границі, тобто  $x_{9(0)}^* = 21,185$ .

Всі інші дані зі стовпчика  $j=0$  переносяться у стовпчику  $j=1$  без зміни, оскільки для них виконується середнє співвідношення.

Переходимо до опрацювання даних, які наведені у стовпчику  $j=1$  – переходимо до першого кроку ітерації. Обчислюємо середнє значення для модифікованих даних:

$$\bar{x}_1^* = \frac{1}{9} \sum_{i=1}^9 x_{i(0)} = 20.387,$$

яке використовується для визначення СКВ даних у стовпчику  $j=1$ :

$$S_1^* = \sqrt{\frac{1}{9-1} \sum_{i=1}^9 (x_{i(1)}^* - \bar{x}_1^*)^2} = 0.869$$

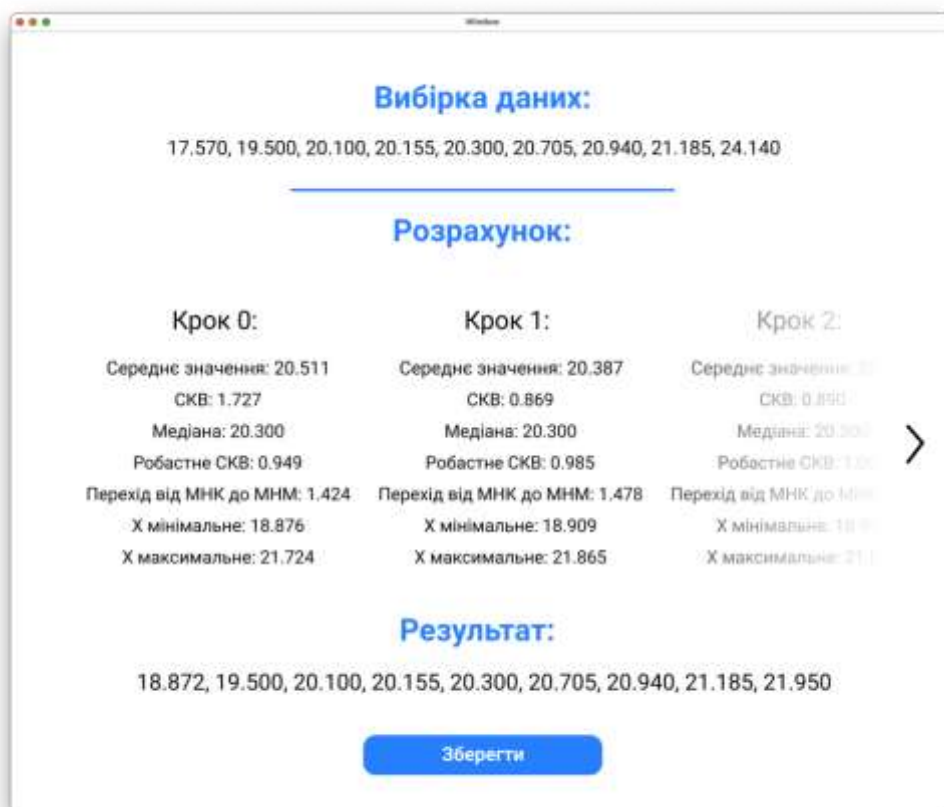


Рисунок 5 – відображення процесу роботи алгоритму та результату

Таблиця 1 – Покрокове визначення середнього значення та СКВ

Номер ітерації $j$	0	1	2	3	4
$I$	2	3	4	5	6
$\varphi_j$		1,424	1,478	1,514	1,539
$x_j^* - \varphi_j$ $x_j^* + \varphi_j$		18,876 21,724	18,909 21,865	18,893 21,921	18,872 21,950
$x_{1(j)}^*$	17,570	18,876	18,909	18,893	18,872
$x_{2(j)}^*$	19,500	19,500	19,500	19,500	19,500
$x_{3(j)}^*$	20,100	20,100	20,100	20,100	20,100



Продовження таблиці 1

1	2	3	4	5	6
$x_{4(j)}^*$	20,155	20,155	20,155	20,155	20,155
$x_{5(j)}^*$	20,300	20,300	20,300	20,300	20,300
$x_{6(j)}^*$	20,705	20,705	20,705	20,705	20,705
$x_{7(j)}^*$	20,940	20,940	20,940	20,940	20,940
$x_{8(j)}^*$	21,185	21,185	21,185	21,185	21,185
$x_{9(j)}^*$	24,140	21,724	21,865	21,921	21,950
Середнє $\bar{x}_j^*$	20,511	20,387	20,407	20,411	20,412
Стандартне відхилення $\bar{S}_j^*$	1,727	0,869	0,890	0,905	0,916
Нове $\bar{x}_{j+1}^*$	20,300	20,387	20,407	20,411	20,412
Нове $\bar{S}_{(j+1)}^*$	0,949	0,985	1,009	1,026	1,039

Обчислимо «нове» (робастне) значення СКВ, отримане на першому кроці ітерації:

$$s_{(1+1)}^* = 1,134 \cdot s_1^* = 1,134 \cdot 0,869 = 0,985, \quad (10)$$

тобто результат першого робастного наближення. Оскільки розходження між  $s_{(0+1)}^*$  та  $s_{(1+1)}^*$  достатньо велике (тобто «підгонка» розподілу ще не закінчилася), переходимо до другого кроку ітерації – стовпчик  $j=2$  (Табл. 1).

Спочатку обчислюємо значення точки, що «розділяє» області застосування МНК та МНМ:

$$\varphi_2 = 1,5 \cdot s_{(1+1)}^* = 1,5 \cdot 0,985 = 1,478,$$

а також відповідно нижню та верхню границі переходу від МНК до МНМ:

$$\bar{x}_{1-}^* - \varphi_2 = 20,387 - 1,478 = 18,909$$

$$\bar{x}_{1+}^* - \varphi_2 = 20,387 + 1,478 = 21,865$$

Вид розподілу, після першого кроку ітерації, було уточнено і це надало можливість розширити границі переходу від МНК до МНМ – «підгону» до нормального закону, якому відповідає інтерквартильний інтервал, визначений на підставі  $MAD_9$ . На підставі цього у стовпчику  $j=2$  маємо модифіковані

значення  $x_{1(2)}^* = 18,909$  та  $x_{9(2)}^* = 21,865$ . Інші елементи стовпчика не модифікуються.

Для даних, наведених у стовпчик  $j = 2$ , обчислимо середнє:

$$\bar{x}_2^* = \frac{1}{9} \sum_{i=1}^9 x_{i(2)} = 20.407,$$

яке використовується для визначення СКВ даних у стовпчику  $j=2$ :

$$S_2^* = \sqrt{\frac{1}{9-1} \sum_{i=1}^9 (x_{i(2)}^* - \bar{x}_2^*)^2} = 0.890.$$

Обчислимо «нове» (робастне) значення СКВ, отримане на другому кроці ітерації, яке характеризує «неусічений» розподіл:

$$S_{(2+1)}^* = 1,134 \cdot S_2^* = 1,134 \cdot 0,890 = 1,009,$$

тобто результат другого робастного наближення (другого кроку ітераційної процедури).

Оскільки розходження з робастною оцінкою, отриманою на другому кроці, ще залишається великим, переходимо до третього етапу, який починається з знаходження

$$\varphi_3 = 1,5 \cdot s_{(2+1)}^* = 1,5 \cdot 0,985 = 1,514,$$

а також відповідно нижню та верхню границі переходу від МНК до МНМ:

$$\bar{x}_2^* - \varphi_3 = 18,893 \text{ та } \bar{x}_2^* + \varphi_3 = 21,921$$



Рисунок 6 – Після збереження результату можливо побудувати графік або продовжити роботу

Ітераційна процедура закінчується, коли різниця між робастними середніми та СКВ на поточному і попередньому кроці не стане величиною другого порядку малості. Результати розрахунку наведені у відповідних стовпчиках Таблиці 1. Як показує аналіз даних, середнє значення досягає сталого стану після другого кроку, а СКВ – після третього.

#### 4 МОДЕЛЮЮЧИЙ ЕКСПЕРИМЕНТ

Проведений аналіз та результати, наведені у Таблиці 1, дозволяють зробити висновок, що абсолютне значення крайніх елементів вибірки практично не впливають на робастне оцінювання параметрів закону розподілу. Це пояснюється тим, що абсолютне медіанне відхилення визначається по відношенню до медіани вибірових даних і, при незмінному  $x_{7(0)}^* = 20,940$ , завжди буде  $MAD_9 = 0,640$ . Це пояснюється тим, що «сусідні» абсолютні значення різниць у впорядкованому чисельному ряді будуть  $\Delta_6 = 0,405$  та  $\Delta_8 = 0,889$  і при будь-якому значенні  $x_{6(0)}^*$  і виконанні нерівності  $0,405 \leq x_{7(0)}^* \leq 0,889$ ,  $x_{6(0)}^* = 20,705$  буде медіаною абсолютних медіанних відхилень.

Проаналізовано випадок, коли впорядкований числовий ряд включає два близько розташованих елементи, які є викидами. При цьому має місце, так званий, замаскований ефект викидів. Застосування критерія Тітена – Мура дозволяє запобігти «маскувальному ефектові» і виключити викиди. Але знайдені оцінки параметрів розподілу суттєво відрізняються від істинних значень.

Проаналізуємо, як в цьому випадку буде себе поводити ітераційний робастний алгоритм.

Для цього введемо у стовпчик  $j=0$  таблиці 2 модифіковане значення  $x_{2(0)}^* = 18,250$ .

Таблиця 2 – Експеримент з наявністю «маскувального ефекту»

Номер ітерації $j$	0	1	2	3	4
$\varphi_j$		1,424	1,868	2,132	2,306
$x_j^* - \varphi_j$ $x_j^* + \varphi_j$		18,380 21,116	18,105 22,369	17,928 22,540	17,812 22,656
$x_{1(j)}^*$	17,570	18,380	18,105	17,929	17,812
$x_{2(j)}^*$	18,250	18,250	18,250	18,250	18,250
$x_{3(j)}^*$	20,100	20,100	20,100	20,100	20,100
$x_{4(j)}^*$	20,155	20,155	20,155	20,155	20,155
$x_{5(j)}^*$	20,300	20,300	20,300	20,300	20,300
$x_{6(j)}^*$	20,705	20,705	20,705	20,705	20,705
$x_{7(j)}^*$	20,940	20,940	20,940	20,940	20,940
$x_{8(j)}^*$	21,185	21,185	21,185	21,185	21,185
$x_{9(j)}^*$	24,140	21,116	21,369	22,540	22,656
Середнє $\bar{x}_j^*$	20,372	20,248	20,237	20,234	20,234
Стандартне відхилення $\bar{S}_j^*$	1,863	1,098	1,253	1,355	1,424
Нове $\bar{x}_{j+1}^*$	20,300	20,248	20,237	20,234	20,234
Нове $\bar{S}_{((j+1))}^*$	0.949	1,245	1,421	1,537	1,615

Проаналізувавши Таблицю 2, можна зробити висновок, що ітераційний алгоритм «працює» і в даному випадку – він залишається стійким навіть при наявності замаскованих викидів. Але це робиться за рахунок збільшення числа кроків ітераційної процедури.

Проаналізуємо, як впливає модифікація значення медіани числового ряду на ефективність (число кроків збіжності) ітераційного алгоритму. Для цього

будемо здійснювати зміну значення медіани вихідного впорядкованого ряду від  $x_{4(0)}^* = 20,155$  до  $x_{6(0)}^* = 20,705$ . У наведених в додатку таблицях, розглядаються випадки, коли медіана приймає значення 20,250; 20,450; 20,600. За результатами експерименту побудована залежність кроків ітерації до асимптотичного стійкого стану і від абсолютного значення медіани (рис. 4).

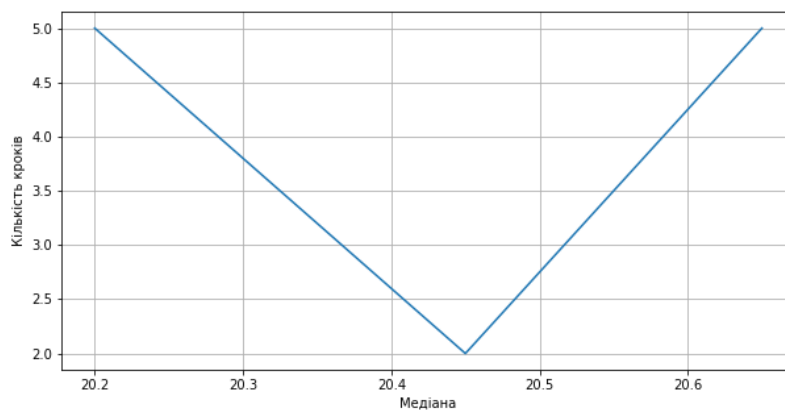


Рисунок 4 – Залежність збіжності ітераційного процесу від зміни значення медіани

Із залежності, наведеної на рисунку, можна прийти до заключення, що мінімальне число кроків ітерації буде для випадку, коли значення медіани відповідає половині суми попереднього та наступного елементів чисельного ряду.

## 5 АНАЛІЗ ТА РАЗРОБКА СТАРТАП ПРОЕКТУ «AvCheck»

### 5.1 Опис ідеї проекту

На сьогоднішній день у всьому світі розробляють дуже багато нових технологій, що спрощують нам життя або допомагають виконувати небезпечну роботу, досліджувати космос чи океани. Кожна з нових розробок проходить певну кількість тестів різної складності для отримання впевненого результату надійності роботи самої розробки та процес затвердження її для подальшого впровадження у використання людиною. Але не кожна розробка може бути протестована у достатній мірі, бо сам експеримент може бути досить дорогим або небезпечним для додаткових тестувань. Це накладає деякі обмеження щодо висновків, бо вихідна вибірка даних достатньо мала та може містити аномальні результати або так звані «викиди».

Для забезпечення виявлення викидів та зменшення їх впливу на середнє значення тестування, був розроблений алгоритм, що забезпечить досягнути досить точного результату вимірювання при недостатньо великій вибірці вихідних даних.

Ідея полягає у створенні онлайн ресурсу для розповсюдження ліцензійного доступу для компаній різного розміру та напрямку для отримання більш точного результату тестування їх продукту при досить обмежених можливостях додаткових тестів.

Спектр послуг:

- Обробка викидів у виборці результатів, навіть замаскованих.
- Встановлення більш точного загального результату вимірювань.
- Додаткова допомога у подальшому аналізі даних.

Таблиця 5.1 – Опис ідеї стартап проекту

Зміст ідеї	Напрямки застосування	Вигоди для користувача
Система обробки вибірки малих обсягів з наявністю різних аномальних результатів (викидів)	Обробка та демонстрація даних	Отримання точних значень без впливу викидів на результат середнього значення
		Допомога у подальших вимірюваннях та тестах продукту

Веб-система з можливістю обробки та візуалізації даних є достатньо технологічною розробкою, що повинна бути спроектована таким чином, щоб мати можливість стати бізнес проектом. Актуальна можливість в створенні рентабельної підприємницької діяльності завдяки доцільності та високій інноваційності системи. Тому пропонується стартап проект, що має можливість для розширення, який використовує найновітніші технології та має низку застосувань.

Саме тому одним з важливих розділів магістерської дисертації є розробка стартап проекту. Для цього розглядаються рішення для розробки перспективного проекту з ринковою актуальністю, перспективністю, фінансовим аналізом, планом для просування пропозиції інвесторам, маркетинговим плануванням, організованістю, аналізом ризиків і можливостей.

Для більш детального ознайомлення з тим, де використовується, які проблеми вирішує проект та який час необхідний для його здійснення, в таблиці 5.2. представлено інформаційну картку проекту.



Таблиця 5.2 – Інформаційна картка проекту

Назва проекту	AvCheck
Автор проекту	Володарський Валентин Артемович
Коротка анотація	Система статистичного оцінювання результатів тестування або дослідження дозволяє оцінювати середнє значення використовуючи робастні алгоритми, що досить стійкі до наявності аномальних результатів вимірювання. Проста у використанні та швидка при обробці даних.
Термін реалізації проекту	6 місяців
Необхідні ресурси	Фінансові: придбання веб-хостингу та наймання команд розробників. Інтелектуальні: патентування методу, промисловий дизайн, маркетингова стратегія, реклама.
Опис проблеми, яку вирішує	Вирішує гостру проблему в необхідності точності вимірювань при малому обсязі вхідної вибірки.

Наступним кроком є проведення аналізу техніко-економічних переваг ідеї порівняно з конкурентами:

- проводимо порівняльний аналіз показників для власної ідеї визначено показники, що мають а) гірші значення (W, слабкі); б) аналогічні (N, нейтральні) значення; в) кращі значення (S, сильні), наведені у таблиці 5.3.
- визначаємо коло конкурентів, що вже існують на ринку та проводимо збір інформації щодо показників для ідеї власного проекту та проектів конкурентів відповідно до визначеного вище переліку
- визначаємо перелік техніко-економічних властивостей ідеї

Таблиця 5.3 – Визначення сильних, слабких та нейтральних характеристик ідеї проекту

№ п/п	Техніко-економічні характеристики ідеї	(потенційні) товари/концепції конкурентів			W (слабка сторона)	N (нейтральна сторона)	S (сильна сторона)
		Мій проект	Trianz	Фріланс			
1.	Вартість розробки	16500	43000	21000	-	-	+
2.	Час розробки	6 місяців	12 місяців	6 місяців	-	+	
3.	Технології	Сучасні	Відносно сучасні	Відносно сучасні	-	-	+
4.	Відомість	Відсутня	Велика	-	+	-	-

Проаналізувавши таблицю 5.3, можна побачити, що наведених характеристик дві є сильною стороною проекту, одна нейтральна та одна слабка. Тому можна зробити висновок, що в даному проекті домінують сильні сторони, що показує позитивну можливість даного проекту стати в майбутньому успішним.

## 5.2 Технологічний аудит проекту

Таблиця 5.4 – Технологічна здійсненність проекту

№ п/п	Ідея проекту	Технології її реалізації	Наявність технологій	Доступність технологій
1.	Алгоритм для статистичної обробки вибірки з малим обсягом даних та можливої наявності викидів	Низькорівнева мова програмування C++, зручна для виконання сучасними технологіями автоматизації	+	+
		Високорівнева мова програмування Python для аналізу даних	+	+
Обрана технологія реалізації ідеї проекту: Друга технологія				

Після аналізу технологічної реалізації ідеї проекту, можна зробити висновок, що програмно реалізація проекту можлива, але необхідно визначитися з мовою програмування. Загалом досить вигідно використовувати мову програмування C++, як більш швидкодіючу, проте, після додаткового аналізу в мережі Інтернет, було виявлено, що при використанні певних бібліотек для високорівневої мови програмування Python, може значно оптимізувати та пришвидшити її роботу. На сьогоднішній день є досить багато готових рішень для більш простого написання алгоритму, без громіздких підпрограм, тому було вирішено для аналізу та відладки алгоритму, та подальшого впровадження було обрано високорівневу мову програмування Python.

### 5.3. Аналіз ринкових можливостей запуску стартап-проекту

У цьому розділі визначаються ринкові можливості, які можна використати при впровадженні проекту на ринку та можливі загрози, які можуть перешкодити успішній реалізації стартапу.

Для початку проведемо аналіз динаміки розвитку ринку, наявності попиту та обсягу.

Таблиця 5.5 – Попередня характеристика потенційного ринку стартапу

№ п/п	Показники стану ринку (найменування)	Характеристика
1	Кількість головних гравців, од	4
2	Загальний обсяг продаж, грн/ум. од	628540
3	Динаміка ринку (якісна оцінка)	Стабільно зростаюча
4	Наявність обмежень для входу (вказати характер обмежень)	Відсутня
5	Специфічні вимоги до стандартизації та сертифікації	Відповідають
6	Середня норма рентабельності в галузі (або по ринку), %	68.9%

Розглянувши таблицю стає зрозуміло, що є певна кількість конкурентів, але сама динаміка ринку є зростаючою. Певну загрозу становлять компанії гіганти та інноваційні компанії.

Наступним кроком буде визначення потенційних груп клієнтів, їх особливості, та створення приблизного переліку вимог до системи.

Таблиця 5.6 – Характеристика потенційних клієнтів стартап-проекту

№ п/п	Потреба, що формує ринок	Цільова аудиторія (цільові сегменти ринку)	Відмінності у поведінці різних потенційних цільових груп клієнтів	Вимоги споживачів до товару
1.	Базова потреба, яка задовольняється це необхідність в скороченні витрат додаткові випробування та отримання точного середнього значення.	Великі та малі виробничі підприємства. Науково-дослідні лабораторії. Університети.	Головною перевагою є те, що замовник буде отримувати додаткову допомогу у подальших тестуваннях.	Більш ефективний, а отже вигідний спосіб контролю за якістю продукції.

На даний момент на ринку існує необхідність забезпечення контролю якості більше ефективними способами для зменшення витрат виробництв, як великих так і малих.

Таблиця 5.7 – Фактори ризику впровадження стартап-проекту

№ п/п	Фактор	Зміст загрози	Можлива реакція компанії
1.	Фінансовий	Малий стартовий капітал	Пошук нових інвесторів
2.	Підприємство невелике і нове	Малий ринок послуг	Реклама, реклама у інтернеті
3.	Конкуренти	Можливість бути поглинутим «великим гравцями»	Новизна та ефективність продукції
4.	Не обізнаність на ринку	Невідомі підходи до клієнтів	Залучення більш досвідчених експертів

Проаналізувавши таблицю 5.7 можна зазначити, що найбільш впливовим фактором ризику впровадження стартап-проекту є те, що підприємство є новим

і маловідомим для клієнтів. Потрібно вдало та якнайшвидше реагувати на кожну в потенційних проблем. В якості протидії усім цим факторам ризику запропоновані контрдії, що можуть покращити та значно стабілізувати становище проекту в тяжких ситуаціях. В якості результату можна вважати проект надійним.

Але поряд із колом загроз існують і певні можливості (таблиця 5.8).

Таблиця 5.8 – Фактори можливостей

№ п/п	Фактор	Зміст можливості	Можлива реакція компанії
1.	Залучення нових кваліфікованих людей до роботи	Індивідуальних підхід до кожного клієнта	Покращення якості продукту
2.	Притік нових клієнтів	Збільшення кількості клієнтів	Подальше збільшення команди
3.	Використання нових технологій	Доступна, більш якісна продукція	Новизна

Розглянувши перелічені можливості в таблиці 5.8, та порівнявши їх з усіма факторами ризику з таблиці 5.7, можна зробити висновок, що стартап-проект перекидає потенційні ризики своїми можливостями.

Також розуміння своїх конкурентів є дуже важливим при появі на ринку, їх швидкості адаптації, перспектив продукту та можливостей. Необхідно враховувати їх потенційний вплив на загальний ринок. Тому розглянемо таблицю 5.9.

Таблиця 5.9 – Ступеневий аналіз конкуренції на ринку

Особливості конкурентного середовища	В чому проявляється дана характеристика	Вплив на діяльність підприємства (можливі дії компанії, щоб бути конкурентоспроможною)
Складність	Жорстка боротьба за кожного клієнта	Різний підхід до пропозицій
Глобальний ринок	Розгляд конкурентів можливий, як у межах країни, так і в рамках сусідніх держав	Подальше розширення взаємодії з клієнтами інших держав
Спеціалізованість	В великому досвіді роботи з клієнтами	Залучення більш досвідчених експертів за межами держави

Завдяки незначній конкуренції на ринку, ми маємо певну можливість експериментувати у впровадженні нашого продукту у випадку переваги над клієнтами завдяки якості та індивідуальному підході до кожного з наших клієнтів. В таблиці 5.9 приведені дії для поліпшення конкурентного становища стартап-проекту.

Після аналізу конкуренції проведемо більш детальний огляд умов конкуренції в галузі. Стартап-проекту необхідна розробка успішної стратегії, яка забезпечить переваги над конкурентами та допоможе закріпити позиції на ринку.

Таблиця 5.10 – Аналіз конкуренції в галузі за М. Портером

Складові аналізу	Прямі конкуренти в галузі	Потенційні конкуренти	Постачальники	Клієнти	Товари-замінники
	Trianz	ITransition	Trianz	Великі та середні підприємства	-
Висновки:	Потужний конкурент.	Конкурент в майбутньому	Використання своїх наробіток щодо обробки даних	Необхідно зайняти ринок малих і середніх підприємств. В подальшому захопити частину великих підприємств	-

Проаналізувавши таблицю 5.10 можна зробити висновок, що для того щоб успішно закріпитися на ринку потрібно постійно нарощувати клієнтів з малих та середніх підприємств, які б зверталися регулярно, розвивати маркетингову компанію та в подальшому намагатися перейняти клієнтів з великих підприємств, вдало використовувати сильні сторони проекту.

Грунтуючись на основі вищесказаного можна переходити до розгляду факторів конкурентоспроможності проекту. Необхідно пам'ятати, що підприємство пропонує кваліфіковану допомогу клієнту та високу якість. В наведеній нижче таблиці 5.11 продемонстровано фактори які забезпечать позитивні результати на конкурентному ринку.

Таблиця 5.11 – Обґрунтування факторів конкурентоспроможності

№ п/п	Фактор конкурентоспроможності	Обґрунтування (наведення чинників, що роблять фактор для порівняння конкурентних проектів значущим)
1.	Ціна товару	За рахунок простоти кінцевого продукту, ціна буде невелика.
2.	Якість товару	Якість товару на високому рівні забезпечується внутрішнім алгоритмом, що і є ключовим у цьому проекті.
3.	Час розробки	Клієнтів цікавить мінімальний час розробки
4.	Інформаційне забезпечення	Інформаційне забезпечення дозволить швидко повідомляти клієнтам про нові оновлення та іншу важливу інформацію.
5.	Політичний стан у країні	Впливає на можливість розширення підприємства в регіонах, на співпрацю з іншими галузями

Таблиця 5.12 – Порівняльний аналіз сильних та слабких сторін проекту

№ п/п	Фактор конкурентоспроможності	Бали 1-20	Рейтинг товарів-конкурентів у порівнянні з стартап-проектом						
			-3	-2	-1	0	+1	+2	+3
1	Ціна	19	+						
2	Час розробки	19		+					
3	Якість товару	20	+						
4	Інформаційне забезпечення	14				+			

Проаналізувавши таблиці 5.11 та 5.12 можна зробити висновки, що фактори конкурентоспроможності проекту мають позитивний внесок. Головною перевагою є ціна та якість продукту.

Останнім етапом ринкового аналізу можливостей впровадження проекту є складання SWOT – аналізу (матриці аналізу сильних (Strength) та слабких (Weak) сторін, можливостей (Opportunities) та загроз (Troubles), пов'язаних з його здійсненням).

Таблиця 5.13 – SWOT - аналіз стартап-проекту

Сильні сторони: 1. Ціна послуг та продукції. 2. Якість товару. 3. Час розробки 4. Індивідуальний підхід до кожного клієнта.	Слабкі сторони: 1. Підприємство нове, мале та невідоме. 2. Відсутність довіри до підприємства у існуючих клієнтів. 3. Відсутність достатньої кількості фінансування
Можливості: 1. Доступна та якісна продукція. 2. Індивідуальний підхід до кожного клієнта.	Загрози: 1. Можливість бути витісненим «великим гравцем». 2. Недостатньо коштів. 3. Малий ринок послуг.

На основі аналізу факторів загроз та факторів можливостей маркетингового середовища був складений перелік ринкових загроз та можливостей. Ринкові можливості та ринкові загрози є прогнозованими результатами впливу факторів.

Визначені альтернативи аналізуються з точки зору строків та ймовірності отримання ресурсів (табл. 5.14).

Таблиця 5.14 – Альтернативи ринкового впровадження стартап-проекту

№ п/п	Альтернатива (орієнтовний комплекс заходів) ринкової поведінки	Ймовірність отримання ресурсів	Строки реалізації
1	Стратегія нейтралізації ринкових загроз сильними сторонами стартап-проекту	Середня	2 місяці
2	Вкладення більшого відсотку прибутку в розробку нових алгоритмів та розширення ринку	Середня	1 рік
3	Стратегія компенсації слабких сторін стартап-проекту наявними ринковими можливостями	Середня	4 місяці

З наведених альтернатив в таблиці 5.14 обираємо стратегію компенсації ринкових загроз сильними сторонами стартап-проекту.



## 5.4 Розроблення ринкової стратегії проекту

Розроблення ринкової стратегії першим кроком передбачає визначення стратегії охоплення ринку: опис цільових груп потенційних споживачів.

Таблиця 5.15 – Вибір цільових груп потенційних споживачів

№ п/п	Опис профілю цільової групи потенційни х клієнтів	Готовність споживачів сприйняти продукт	Орієнтов ний попит в межах цільової групи (сегменту)	Інтенсивніс ть конкуренції в сегменті	Простота входу у сегмент
	Великі промислові виробництв а	Майже на кожному етапі виготовлення продукції є необхідність контролю вихідної продукції чи напівфабрикату	Попит надзвичайно великий	Конкуренці я не інтенсивна	Вхід до сегменту є досить усередненим, адже на даний момент на ринку немає достатньо великих гравців які представляют ь загрозу.
	Малі промислові виробництв а			Конкуренці я не інтенсивна	
Які цільові групи обрано: Під час аналізу потенційних груп споживачів було прийнято рішення що компанія буде працювати із малими промисловими виробництвами.					

Виходячи з проведеного аналізу потенційних груп споживачів було обрано в якості цільової групи малі промислові виробництва. Вони є більш вибагливими до якості контролю на попередніх етапах, на відміну від великих підприємств, адже кількість товару менша, а отже за якістю необхідно слідкувати краще.

Для роботи в обраному сегменті ринку необхідно сформулювати базову стратегію розвитку.

Таблиця 5.16 – Визначення базової стратегії розвитку

№ п/п	Обрана альтернатива розвитку проекту	Стратегія охоплення ринку	Ключові конкурентоспроможні позиції відповідно до обраної альтернативи	Базова стратегія розвитку*
1.	Стратегія спеціалізації. Концентрація на потребах одного цільового сегменту	Відповідати цільовим потребам ринку краще ніж конкуренти. Диференційований маркетинг	Однією з найважливіших позицій є виготовлення продукції за прийнятною ціною та якістю її на вищому рівні	Стратегія спеціалізації

Проаналізувавши таблицю 5.16, для даного стартап-проекту була обрана стратегія яка є найбільш вдалою, це стратегія базового розвитку – стратегія спеціалізації. Передбачає концентрацію на потребах одного цільового сегменту, без прагнення охопити увесь ринок. Мета тут полягає в задоволенні потреб вибраного цільового сегменту краще, ніж конкуренти.

Таблиця 5.17 – Визначення базової стратегії конкурентної поведінки

№ п/п	Чи є проект «першопрохідцем» на ринку?	Чи буде компанія шукати нових споживачів, або забирати існуючих у конкурентів?	Чи буде компанія копіювати основні характеристики товару конкурента, і які?	Стратегія конкурентної поведінки*
1.	Ні	Компанія буде охоплювати сегмент нових клієнтів, та надаватиме потреби існуючим за необхідності.	Ні	Стратегія наслідування лідеру

Проаналізувавши таблицю 5.17 було обрано стратегію наслідування лідеру тому що підприємство є невеликим та з обмеженою кількістю коштів і не буде здатна до розширення певний час. Стратегія наслідування лідеру найчастіше має місце у випадку олігополії, коли підприємство прагне уникнути боротьби, У подальшому стратегія компанії буде змінюватися.

На основі вимог споживачів з обраного сегменту до постачальника і продукту, а також в залежності від стратегії розвитку та стратегії конкурентної поведінки розробляємо стратегію позиціонування яка визначається у формування ринкової позиції, за яким споживачі мають ідентифікувати проект.

Таблиця 5.18 – Визначення стратегії позиціонування

№ п/п	Вимоги до товару цільової аудиторії	Базова стратегія розвитку	Ключові конкурентоспроможні позиції власного стартап-проекту	Вибір асоціацій, які мають сформувати комплексну позицію власного проекту (три ключових)
1.	Висока якість та точність товару	Стратегія спеціалізації	Вища якість ніж у конкурентів, за прийнятну ціну	Якість. Ціна. Індивідуальний підхід

Результатом даного підрозділу є система рішень щодо ринкової поведінки компанії, вона визначає в якому напрямі буде працювати компанія на ринку.

На основі вимог споживачів з обраних сегментів до постачальника та до продукту (табл. 5.6), а також в залежності від обраної базової стратегії розвитку (табл. 5.16) та стратегії конкурентної поведінки (табл. 5.17) розробляється стратегія позиціонування (табл. 5.18). що полягає у формуванні ринкової позиції (комплексу асоціацій), за яким споживачі мають ідентифікувати торгівельну марку/проект. Обраною стратегією позиціонування є:

- на основі специфічних відчутних характеристик.

### 5.5. Розроблення маркетингової програми стартап-проекту

Розробка маркетингової концепції товару, який отримає споживач є першим кроком маркетингової програми. У таблиці 5.19 підсумовуємо результати аналізу конкурентоспроможності товару.

Таблиця 5.19. Визначення ключових переваг концепції потенційного товару

№ п/п	Потреба	Вигода, яку пропонує товар	Ключові переваги перед конкурентами (існуючі або такі, що потрібно створити)
1.	Необхідність в ресурсі для обробки вибірки контролю виробництва	Ефективний алгоритм, який забезпечує точність середнього значення навіть з наявністю замаскованих викидів	Якісна продукція, яка працює ефективніше, за привабливою ціною

Висновком до таблиці 5.19 є те, що клієнти хочуть отримати доступний ресурс, що полегшить кількість затрат на тестування їх продукції та дасть точне середнє значення загальному дослідженню.

Таблиця 5.20 – Опис трьох рівнів моделі товару

Рівні товару	Сутність та складові		
I. Товар за задумом	Алгоритм оцінювання середнього значення		
II. Товар у реальному виконанні	Властивості/характеристики	М/Нм	Вр/Тх /Тл/Е/Ор
	1. Ціна	М	Е
	2. Якість	Нм	Тх
	3. Ефективність	Нм	Тх
	4. Простота	Нм	Тл
	5. Підтримка	Нм	Ор
	Якість: відповідає нормам ДСТУ		
	Пакування: продукція постачається за допомогою онлайн ресурсу		
	Марка: AvCheck		
	До продажу: ліцензійний ключ до ресурсу		
III. Товар із підкріпленням	Після продажу: Підтримка та модифікації продукту за заявою клієнта. Підтримка клієнта доки він користується продукцією. Підтримка у подальших тестуваннях та дослідженнях клієнта.		
	За рахунок чого потенційний товар буде захищено від копіювання: патент на метод		

Кінцевим продуктом має стати онлайн ресурс, на якому встановлено програмне забезпечення, розроблене підприємством у вигляді алгоритму, яке поставляється замовнику з ліцензійним ключем, випробним терміном, постійною підтримкою. Відповідає нормам ДСТУ.

Необхідно зазначити основні витрати проекту, станом на початок основної діяльності (в 0-й рік реалізації проекту) за таблицею 5.21, приведеною нижче.

Таблиця 5.21 – Загальні початкові витрати проекту

№ з/п	Стаття витрат	Обсяги витрат в 0-й рік, тис. грн.
1.	Проведення НДДКР	20
2.	Розробка проектних матеріалів і ТЕО	5
3.	Робоче проектування і прив'язка проекту	-
4.	Витрати на придбання обладнання та устаткування та пристроїв	8,8
5.	Витрати на придбання нематеріальних активів	10
6.	Одноразові виплати, зокрема гарантуючим і страховим організаціям	6,3
7.	Витрати на перед виробничі маркетингові дослідження і створення збутової мережі	16
8.	Витрати, пов'язані з діяльністю команди	184
9.		
	<b>Разом</b>	<b>245,1</b>

Варто зазначити та визначити обсяг поточних загальногосподарських витрат, необхідний для реалізації проекту.

Таблиця 5.22 – Планові загальногосподарські витрати

№ з/п	Стаття витрат	Витрати за період, тис. грн.		
		1-й рік	2-й рік	3-й рік
1.	Витрати на обладнання, устаткування та пристрої	25,8	-	-
2.	Витрати на придбання нематеріальних активів	14	-	-
3.	Витрати на персонал (на відрядження, соціальні заходи тощо)	184	188	195
4.	Витрати на збут	-	-	-
5.	Витрати на просування та рекламу	12	12	15
6.	Оплата юридичних послуг	5	-	-
	<b>Разом:</b>	<b>240,8</b>	<b>200</b>	<b>210</b>

Після визначення основних витрат, наступним кроком є визначення цінових меж, якими необхідно керуватися при встановленні ціни на потенційний товар, це передбачає деякий аналіз цін товарів конкурентів, та доходів споживачів продукту.

Таблиця 5.23 – Визначення меж встановлення ціни

№ п/п	Рівень цін на товари- замінники	Рівень цін на товари-аналоги	Рівень доходів цільової групи споживачів	Верхня та нижня межі встановлення ціни на товар/послугу
1.	-	5600-8200 грн	Не можливо визначити	4600-7300 грн/од.

Було визначено приблизні загальні витрати проекту у 0-й рік реалізації, а також загальногосподарські витрати на перший, другий, третій роки відповідно. Завдяки інформації з таблиць 5.21 та 5.22, було проведене порівняння цін з іншими схожими товарами на ринку, а також визначено верхню та нижню межу можливого встановлення ціни на товар. Ціновий діапазон не є широким. Рівень доходів цільової споживчої групи дізнатися неможливо.

Таблиця 5.24 – Формування системи збуту

№ п/п	Специфіка закупівельної поведінки цільових клієнтів	Функції збуту, які має виконувати постачальник товару	Глибина каналу збуту	Оптимальна система збуту
1.	Нова закупівля	Організація доступу до онлайн ресурсу	1	Через мережу Інтернет

Висновок: В якості каналу збуту, підприємство буде працювати як перший канал, а саме торгівля завдяки мережі Інтернет.

Таблиця 5.25 – Концепція маркетингових комунікацій

№ п/п	Специфіка поведінки цільових клієнтів	Канали комунікацій, якими користуються цільові клієнти	Ключові позиції, обрані для позиціонування	Завдання рекламного повідомлення	Концепція рекламного звернення
1	Необхідність у ресурсі для контролю якості виробництва.	Інтернет	Якість Ціна Індивідуальний підхід	Показати переваги та ефективність даного товару у порівнянні з іншими схожими	Інтернет-оголошення таргетовані на малі підприємства

Інтернет був обраний як основний шлях розповсюдження реклами відносно стартап-проекту, що вказує на ефективність товару у порівнянні з конкурентами.

## 5.6. Висновки

В даному розділі було розроблено стартап-проект з впровадження алгоритму для оцінювання середнього значення малих вибірок з наявністю викидів. Так як продаж за ліцензійним договором виключно програмного забезпечення є не достатньо прибутковим на сьогоднішній день, було прийнято рішення створити підприємство на базі ФОП, яке буде займатися виготовленням кінцевого продукту з програмним забезпеченням, а також постійною підтримкою клієнтів.

У даному розділі було проведено аналіз таких важливих складових стартап-проекту як ідея, проведено технологічний аудит проекту, визначення сильних та слабких характеристик. Виконано аналіз ринкових можливостей запуску стартап-проекту, де досліджено основні фактори можливостей та потенційних загроз, зроблений SWOT-аналіз, порівняні сильні та слабкі сторони проекту. Окрім того, розроблена ринкова стратегія проекту з вибором цільових груп потенційних споживачів, визначення базової та подальшої стратегій розвитку, стратегії позиціонування. Розроблена маркетингова

програма стартап-проекту, де описані три основні рівні моделі товару, сформована концепція маркетингових комунікацій та система збуту. Також була визначені загальні початкові витрати проекту та планові загальногосподарські витрати.



## ВИСНОВОК

Статистичне оцінювання вимірювань є досить важливою частиною кожного дослідження. В даній роботі був проведений аналіз ефективності ітераційного робастного алгоритму визначення, на підставі вибірових даних, параметрів закону розподілу, які відповідають правильності та відтворюваності методики випробування.

Робастні методи дозволяють оцінити статистично надійні параметри закону розподілу на підставі вибірових значень з так званим «викидом». Після проведення експерименту та аналізу отриманих даних, було виявлено, що число кроків збіжності ітераційного алгоритму залежить від абсолютного відхилення екстремальних значень числового ряду, тобто, чим більше це значення, тим більше кроків необхідно, щоб досягнути сталого стану. Також, на підставі проведеного експерименту можна відзначити те, що число кроків ітерації залежить від «розташування» медіани по відношенню до попереднього та наступного елементів чисельного ряду результатів вимірювання і буде мінімальним, коли значення медіани вибірки буде дорівнювати полусумі цих елементів.

## СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Ціделко В.Д., Яремчук Н.А. Невизначеність вимірювання: Монографія. – К.:ІВЦ «Політехніка», 2002. – 176 с.
2. Точність (правильність і прецизійність) методів та результатів вимірювання.: (ISO/IEC 5725:1994, IDT): ДСТУ ГОСТ ИСО 5725:2005. – [Чинний від 2006-07-01]. – К.: Держспоживстандарт України, 2006. –162 с. – (Національний стандарт України).
3. Смоляк С.А., Титаренко Б.П. Устойчивые методы оценивания. – М.:Статистика, 1980. – 181 с.
4. Tietjen G, Moore H. Some Grabb's type statistics for the detection of several outliers. – Technometrics, 1972, vol. 14.
5. Регламент Європейського Парламенту та Ради № 765/2008 від 9 липня 2008 р.
6. Метрологічне забезпечення вимірювань і контролю: навч. посіб./ [Володарський Є.Т. Кухарчук В.В., Поджаренко В.О., Сердюк Г.Б.]. – Вінниця: Велес, 2001. – 219 с.
7. Sarhan Ahmed E., Greenberg Bernard G., editors: Contributions to order statistics (Вклады в порядковые статистики).– John Wiley & Sons, 1962, pp. 482.
8. Tukey John W. Exploratory Data Analysis (Разведочный анализ данных). Addison-Wesley. 1978.
9. Analyst Robust statistics – How Not to Reject Outliers (Робастная статистика – Как не отвергать выбросы). December 1989, vol. 114
10. Большев Л.Н. Таблицы математической статистики / Л.Н. Большев, Н.В.Смирнов – М.: Наука, 1983. – 416 с.
11. Volodarsky E., Warsza Z.L., Examples of robust estimation with small number of measurements. Progress in Automation, Robotics and Measuring Techniques, Vol. 3 “Measuring Techniques and Systems”. (ISBN 978-3-319-15834-1), Vol. 352 of series: Advances in Intelligent Systems and Computing (ISSN 2194-5357) Springer 2015, 285–291

12. Хьюбер Дж. П. Робастность в статистике: Пер. с англ. – М: Мир, 1984. – 304 с.

## ДОДАТОК А. ТАБЛИЦІ З РЕЗУЛЬТАТАМИ РОБОТИ РОБАСТНОГО МЕТОДУ З ЗМІНОЮ МЕДІАНИ

Таблиця з медіаною, що приймає значення 20.250

Номер ітерації $j$	0	1	2	3	4
$\varphi_j$		1,534	1,560	1,574	1,584
$x_j^* - \varphi_j$ $x_j^* + \varphi_j$		18,716 21,784	18,811 21,931	18,823 21,971	18,819 21,987
$x_{1(j)}^*$	17,570	18,716	18,811	18,823	18,819
$x_{2(j)}^*$	19,500				
$x_{3(j)}^*$	20,100				
$x_{4(j)}^*$	20,155				
$x_{5(j)}^*$	20,250				
$x_{6(j)}^*$	20,705				
$x_{7(j)}^*$	20,940				
$x_{8(j)}^*$	21,185				
$x_{9(j)}^*$	24,140	21,784	21,931	21,971	21,987
Середнє $\bar{x}_j^*$	20,505	20,371	20,397	20,403	20,405
Стандартне відхилення $\bar{S}_j^*$	1,728	0,917	0,925	0,931	0,936
Нове $\bar{x}_{j+1}^*$	20,250	20,371	20,397	20,403	20,405
Нове $\bar{S}_{(j+1)}^*$	1,023	1,040	1,049	1,056	1,061

Таблиця з медіаною, що приймає значення 20.450

Номер ітерації $j$	0	1	2	3	4
$\varphi_j$		1,090	1,252	1,358	1,430
$x_j^* - \varphi_j$ $x_j^* + \varphi_j$		19,360 21,540	19,185 21,689	19,076 21,792	19,004 21,864
$x_{1(j)}^*$	17,570	19,360	19,185	19,076	19,004
$x_{2(j)}^*$	19,500				
$x_{3(j)}^*$	20,100				
$x_{4(j)}^*$	20,155				
$x_{5(j)}^*$	20,450				
$x_{6(j)}^*$	20,705				
$x_{7(j)}^*$	20,940				
$x_{8(j)}^*$	21,185				
$x_{9(j)}^*$	24,140	21,540	21,689	21,792	21,864
Середнє $\bar{x}_j^*$	20,527	20,437	20,434	20,434	20,434
Стандартне відхилення $\bar{S}_j^*$	1,725	0,736	0,798	0,840	0,869
Нове $\bar{x}_{j+1}^*$	20,450	20,437	20,434	20,434	20,434
Нове $\bar{S}_{(j+1)}^*$	0,727	0,835	0,905	0,953	0,985

Таблиця з медіаною, що приймає значення 20.600

Номер ітерації $j$	0	1	2	3	4
$\varphi_j$		1,113	1,273	1,374	1,443
$x_j^* - \varphi_j$ $x_j^* + \varphi_j$		19,487 21,713	19,214 21,760	19,088 21,836	19,014 21,900
$x_{1(j)}^*$	17,570	19,487	19,214	19,088	19,014
$x_{2(j)}^*$	19,500				
$x_{3(j)}^*$	20,100				
$x_{4(j)}^*$	20,155				
$x_{5(j)}^*$	20,600				
$x_{6(j)}^*$	20,705				
$x_{7(j)}^*$	20,940				
$x_{8(j)}^*$	21,185				
$x_{9(j)}^*$	24,140	21,713	21,760	21,836	21,900
Середнє $\bar{x}_j^*$	20,544	20,487	20,462	20,457	20,455
Стандартне відхилення $\bar{S}_j^*$	1,725	0,749	0,808	0,848	0,876
Нове $\bar{x}_{j+1}^*$	20,600	20,487	20,462	20,457	20,455
Нове $\bar{S}_{(j+1)}^*$	0,742	0,849	0,916	0,962	0,993

## ДОДАТОК Б. ПРОГРАМНА РЕАЛІЗАЦІЯ РОБАСТНОГО АЛГОРИТМУ

```

import numpy as np

c = 1.5
const = 1.134
iterations = 5
measurement_results = np.sort(np.array([24.140, 20.155, 18.250,
20.300, 20.705, 17.570, 20.100, 20.940, 21.185]))

def round(value):
    return np.around(value, decimals = 3)

def calculate_standart_deviation(measurements, average):
    values = []
    for measure in measurements:
        values.append(np.square((measure - average)))
    return round(np.sqrt(np.sum(values)/(measurements.size - 1)))

def get_average(values):
    return round(np.average(values))

def calculate_delta(measurements, median):
    values = []
    for measure in measurements:
        value = np.absolute(round(measure - median))
        values.append(value)
    return values

def get_first_robust_standart_deviation(measurements, median):
    delta = np.median(calculate_delta(measurements, median))
    return round(1.483 * delta)

def calculate_min_max(x_average, fi):
    x_min = round(x_average - fi)

```

```

x_max = round(x_average + fi)
return x_min, x_max

def check_min_max(measurements, x_min, x_max):
    measurements[0] = x_min
    measurements[-1] = x_max
    return measurements

def check_min_max_zero_step(measurements, x_min, x_max):
    if measurements[0] < x_min:
        measurements[0] = x_min
    if measurements[-1] > x_max:
        measurements[-1] = x_max
    return measurements

def get_median(values):
    return round(np.median(values))

def operate_with_values(values):
    measures = values
    for i in range(0, iterations):
        average = get_average(measures)
        measure_std = calculate_standart_deviation(measures,
average)
        robust_measure_std = round(1.134 * measure_std)
        fi = round(c * robust_measure_std)
        x_min, x_max = calculate_min_max(average, fi)
        measures = check_min_max(measures, x_min, x_max)

def robust_method(values):
    measure_avg = get_average(values)
    measure_median = get_median(values)
    measure_std = calculate_standart_deviation(
        values, measure_avg
    )
    first_robust_std = get_first_robust_standart_deviation(

```



```

        values, measure_median
    )
    first_fi = round(c * first_robust_std)
    x_min, x_max = calculate_min_max(
        measure_median, first_fi
    )
    inputted_measurements = check_min_max_zero_step(
        values, x_min, x_max
    )
    operate_with_values(inputted_measurements)

def default_robust_method_experiment():
    robust_method(measurement_results)

def robust_expereriment_with_different_medians():
    medians = np.arange(20.200, 20.700, 0.05)
    for median in medians:
        values = measurement_results
        values[4] = round(median)
        robust_method(values)

def main():
    default_robust_method_experiment()
    robust_expereriment_with_different_medians()

main()

```

## ДОДАТОК В. СПИСОК ПУБЛІКАЦІЙ

**УДК 519.24**

*В.А. Володарський, студент гр. ПА-91мн, А.А. Помилуйко, студент гр. ПА-91мн КІІ ім. Ігоря Сікорського*

### ДОСЛІДЖЕННЯ ЕФЕКТИВНОСТІ РОБАСТНИХ ПРОЦЕДУР

**Анотація.** В статті досліджується ефективність стійких до викидів ітераційних алгоритмів визначення за вибірковими даними параметрів генерального розподілу.

**Ключові слова:** викиди, відтворюваність, робастність, абсолютне медіанне відхилення, ітерація.

### ВСТУП

На сучасному етапі, коли випробування можуть проводитися в різних лабораторіях і навіть в різних країнах, для атестації методики випробувань при заданих умовах, проводять спільний міжлабораторний експеримент, за результатами якого нормуються показники точності – зміщення та відтвореність [1]. Наявність у вибірці навіть невеликого числа спостережень, які різко виділяються і називаються промахи, здатне кардинально змінити результат статистичного дослідження. Прوماхи виявляють за допомогою відповідних критеріїв та вилучаються з вибірки [2].

Оскільки випробування однотипної продукції, виконані за єдиною методикою, можуть здійснюватися декількома лабораторіями в різних умовах, результати, природно, будуть дещо відрізнятися. Врахування впливових величин аналітичним шляхом практично неможливо. Єдиним підходом для вирішення цієї задачі є проведення міжлабораторних спільних випробувань. Міжлабораторний експеримент фактично є фізичною моделлю реалізації методики з залученням лабораторій, що мають близький професійний рівень і спеціалізуються в даному виді випробувань. При спільному експерименті існує припущення, що всі залучені лабораторії мають однакову повторюваність [1]. Але, з об'єктивних причин, це не завжди виконується.

При такому підході, при нормуванні показників точності методики випробувань, застосування статистичних критеріїв виключення викидів, як це робиться у вимірюванні, не можливо, бо це приведе до невірного визначення показників.

Для виключення впливу викидів при оцінюванні показників точності методики випробувань застосовуються робастні методи [3], які надають можливість використовувати всі наявні експериментальні дані.

## ПОСТАВНОВКА ЗАДАЧІ

Метою роботи є аналіз ефективності ітераційного робастного алгоритму, на підставі вибірових даних, параметрів закону розподілу, які відповідають правильності та відтворюваності методики випробування.

## ОСНОВНА ЧАСТИНА

Вибіркові дані, у загальному випадку, можуть мати деяку розбіжність з передбачуваним розподілом (особливо при малих обсягах) – містити деякі значення, що підпорядковуються іншому розподілу.

Вихідним є те, що деяка центральна частина розподілу експериментальних даних відповідає розподілу генеральної сукупності [4]. Тому при опрацюванні даних, які знаходяться в цій частині розподілу, доцільно використовувати метод найменших квадратів (МНК). Але метод найменших модулів (МНМ) є більш стійким до викидів, ніж МНК, тобто дає найкращий результат при найбільш несприятливому розподілі.

Виходячи з цього, при створенні робастних методів робиться «симбіоз» – для деякої центральної частини розподілу використовується МНК, а для іншої частини для зменшення впливу викидів, але зі збереженням наявних даних, застосовується МНМ. Граничне значення границі переходу від МНК до МНМ відповідає  $\varphi = c\sigma$ . Константа  $c$  регулює ступінь робастності і її значення залежить від ступеня «засмічення» [3]. Зазвичай обирають значення  $c = 1,5$ .

Найбільш стійким до викидів є інтервал, що знаходиться між вибіровими кuartілями. У припущенні про можливий закон розподілу довжина інтервалу однозначно відповідає дисперсії цього розподілу. У якості початкової оцінки центру розподілу береться вибірова медіана. Як початкова оцінка при переході

від повного розподілу до усіченого, береться медіана абсолютних відхилень  $MAD$  (*median absolute deviation*).

$$MAD_n = med\{|x_i - M_n|\}. \quad (1)$$

де,  $M_n = med\{x_i\}$ ,  $x_i$  – елемент вибірки, а індекс  $n$  відповідає числу елементів.

Первинна оцінка СКВ, яка є стійкою до викидів для вибірки з  $n$  елементів, знаходиться на підставі нормального інтерквартильного розмаху і складає  $S_{(0+1)}^* = 1,483 \cdot MAD_n$ . Константа 1,483 використовується для нормування при переході від «інтерквартильного» сегменту до генерального розподілу. Значення  $S_{(0+1)}^*$  використовується при переході до робастної ітераційної процедури уточнення параметрів розподілу [4].

Перед початком цієї процедури визначається точка переходу від МНК до МНМ

$$\varphi_1 = 1,5 \cdot S_{(0+1)}^*. \quad (2)$$

та нижнього і верхнього граничних значень

$$x_{1\min}^* = x_0^* - \varphi_1; \quad x_{1\max}^* = x_0^* + \varphi_1. \quad (3)$$

Молодший елемент ранжируваного ряду вихідних даних порівнюють з  $x_{1\min}^*$ , а старший елемент порівнюють з  $x_{1\max}^*$ . Якщо молодший елемент менше нижнього граничного значення, то йому присвоюється значення  $x_{1\min}^*$ . У випадку, коли старший елемент буде більше верхнього граничного значення, то замість нього в ряд вводиться  $x_{1\max}^*$ . Всі ж інші елементи ряду залишають без зміни. Приходимо до модифікованого чисельного ряду, для якого обчислюється середнє  $\bar{x}_1^*$ , яке є уточненим значенням центру розподілу, та СКВ «усіченого» розподілу  $S_1^*$  (так зване «старе»), яке використовується для обчислення уточненого на першому кроці, робастного СКВ («нового»)  $S_{(1+1)}^* = 1,134 \cdot S_1^*$ . Константа 1,134 дозволяє перерахувати СКВ, обчислене для «усіченого» розподілу до генерального [5].

Другий крок ітераційної процедури починаємо з обчислення граничних значень  $x_{2\min}^*$  та  $x_{2\max}^*$ . Для цього застосовуються співвідношення (3), в яких використовується  $\bar{x}_1^*$ .

Ітераційна процедура продовжується, поки розходження між параметрами розподілу на поточному і попередньому кроці не стане менше заданого значення.

Для прикладу реалізації робастного ітераційного алгоритму скористаємося ранжированими даними, наведеними у [1].

Таблиця 1. Покрокове визначення середнього значення та СКВ

Номер ітерації $j$	0	1	2	3	4
$\varphi_j$		1,424	1,478	1,514	1,539
$x_j^* - \varphi_j$ $x_j^* + \varphi_j$		18,876 21,724	18,909 21,865	18,893 21,921	18,872 21,950
$x_{1(j)}^*$	17,570	18,876	18,909	18,893	18,872
$x_{2(j)}^*$	19,500	19,500	19,500	19,500	19,500
$x_{3(j)}^*$	20,100	20,100	20,100	20,100	20,100
$x_{4(j)}^*$	20,155	20,155	20,155	20,155	20,155
$x_{5(j)}^*$	20,300	20,300	20,300	20,300	20,300
$x_{6(j)}^*$	20,705	20,705	20,705	20,705	20,705
$x_{7(j)}^*$	20,940	20,940	20,940	20,940	20,940
$x_{8(j)}^*$	21,185	21,185	21,185	21,185	21,185
$x_{9(j)}^*$	24,140	21,724	21,865	21,921	21,950
Середнє $\bar{x}_j^*$	20,511	20,387	20,407	20,411	20,412
Стандарт не відхилення $\bar{s}_j^*$	1,727	0,869	0,890	0,905	0,916
Нове $\bar{x}_{j+1}^*$	20,300	20,387	20,407	20,411	20,412
Нове $\bar{s}_{(j+1)}^*$	0,949	0,985	1,009	1,026	1,039

Дані були отримані при проведенні міжлабораторного спільного експерименту при залученні  $n = 9$  лабораторій (підкресленням позначено найменший та найбільший лабораторні результати):

24,140 20,155 19,500 20,300 20,705 17,570 20,100 20,940 21,185

В таблиці 1 представлені дані для  $j$ -ого кроку ітерації.

Аналіз результатів, які були отримані при моделюванні, дозволяє зробити висновок, що значення граничних елементів вибірки практично не впливають на робастне оцінювання параметрів закону розподілу. Ефективність алгоритму пояснюється тим, що абсолютне медіанне відхилення визначається по відношенню до медіани  $x_{5(0)}^* = 20,300$  вибірки.

Також проаналізовано випадок, коли впорядкований числовий ряд включає два близько розташованих елементи, які є викидами. При цьому має місце, так званий, замаскований ефект викидів. Застосування критерія Тітьєна-Мура дозволяє запобігти «маскувальному ефектові» і виключити викиди. Але знайдені оцінки параметрів розподілу відрізняються від істинних значень.

Для оцінювання ефективності алгоритму при наявності викидів з «маскувальним ефектом» у стовпчику  $j = 0$  таблиці 1, проведена модифікація  $x_{2(0)}^* = 18,250$ . Проведений аналіз показав, що ітераційний алгоритм «працює» і, в даному випадку, він залишається стійким навіть при наявності замаскованих викидів.

## СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

- [1] Точність (правильність і прецизійність) методів та результатів вимірювання. Частина 2. Основний метод визначення повторюваності та відтворюваності стандартного методу вимірювань: (ISO/IEC 5725-2:1994, IDT) ДСТУ ИСО 5725-2:2005. – [Чинний від 2006.07.01]. – К.: Держспоживстандарт України, 2006. – 48 с. – (Національний стандарт України).
- [2] Ціленко, В.Д. Невизначеність вимірювання: монографія / В.Д. Ціленко, Н.А. Яремчук. – К.: ІВЦ «Політехніка», 2002. – 176 с.
- [3] Sarhan Ahmed E., Greenberg Bernard G.: Contributions to order statistics (Вклады в порядковые статистики). – John Wiley & Sons, 1962, pp. 482.
- [4] Хьюбер, П. Робастность в статистике: монографія / П. Хьюбер ; пер. з англ. И. Моховой, В. Хохлова – М. : Мир, 1984, – 304с.

- [5] Odporna ocean dokładności metod pomiarowych / E. Volodarsky, L. Kosheva, Z. Warsza // Pomjari, awtomatica, control. – 2012. – №4. – P. 396-401.

***Наук. керівник – к.т.н., доц. Добролюбова М.В.***

**УДК 681**

*Помилуйко А.А., студент гр. ПА-91мп, Володарський В.А., студент гр. ПА-91мп*

КПІ ім. Ігоря Сікорського

**МЕТОДИ ПОБУДОВИ ГРАДУЮВАЛЬНОЇ ХАРАКТЕРИСТИКИ**

**Анотація.** В статті представлені матеріали щодо переваги вибору методу найменших квадратів при побудові градуовальної характеристики.

**Ключові слова:** метод найменших квадратів, градуовальна характеристика, лінійна залежність.

**ВСТУП**

В аналітиці найчастіше зустрічаються випадки у потребі побудови лінійної залежності. [1] Для забезпечення правильного результату аналізу побудова градуовальної характеристики має важливе значення. Однак, всі результати вимірювань характеризуються деякою невизначенністю, а також дані, які були отримані для побудови градуовальної характеристики, теж являються важливими. Ці дані мають розкид відносно прямої, і так часто буває, що пряму проводять так, що точки були “розкидані” відносно прямої більш-менш рівномірно. Використання статичних методів дозволяє визначити найбільш правдиве положення прямої.

**ПОСТАНОВКА ЗАДАЧІ**

Мета статті – показати переваги вибору методу найменших квадратів при побудові градуовальної характеристики.

**ОСНОВНА ЧАСТИНА**

В якості основного принципу використовують метод найменших квадратів (МНК). Його суть полягає у тому, що набір даних, що являється експериментальним, найкращим чином описує саме та пряма, для котрої сума квадратів відхилень експериментальних значень від розрахованих є мінімальною.

Для побудови градуовальної характеристики потрібно вибрати вид залежності. Оскільки на практиці найчастіше зводять градування до знаходження лінійної залежності, то можна прийняти такий вид залежності за



основу. Таким чином, можна вважати, що величини  $X$  та  $Y$  – лінійно зв'язані. Така залежність може бути виражена так:

$$Y = a + bX \quad (1)$$

Саме рівняння (1) є лінійною регресією. Величина  $X$  – незалежною змінною, а  $Y$  – залежною змінною, величини  $a$  і  $b$  – вважаються коефіцієнтами рівняння регресії.

Але можуть траплятися випадки, коли величина  $X$  має похибку, і нею можна знехтувати. Також може бути і так, що така величина може прийматися до відома із малою похибкою. Тож можна сказати, що дані можуть бути описані таким чином:

$$Y_i = a + bX_i + \varepsilon_i \quad (2)$$

де  $a$  – це відрізок, що відсікається на осі ординат,  $b$  – це кутовий коефіцієнт рівняння регресії,  $\varepsilon_i$  - випадкова похибка.

Також слід брати до уваги таку проблему, яка може зустрічатися в кількісному аналізі, як грубі промахи. Цей фактор може впливати на результати вимірювань, тому необхідно враховувати при виборі методу побудови градуовальної характеристики. Та необхідно враховувати оптимальний спосіб оцінки параметрів  $a$  та  $b$ .

Метод найменших квадратів (МНК) є широко розповсюдженим методом при виборі побудови градуовальної характеристики. [2] Оцінки параметрів  $a$  і  $b$  в такому методі розраховують з умови мінімуму суми квадратів вертикальних відрізків, що являються відстаннями від експериментальних точок до шуканої прямої. Головними передумовами для МНК є дані, які підкоряються нормальному закону розподілу та є незалежними між собою. Можна привести відомі вирази для оцінок параметрів  $a$  і  $b$ :

$$a = \frac{(\sum Y_i - b \sum X_i)}{m} \quad (3)$$

$$b = \frac{m \sum X_i Y_i - \sum X_i \sum Y_i}{m \sum X_i^2 - (\sum X_i)^2} \quad (4)$$

Застосування даного методу є цілком доцільним при малих похибках визначеної величини  $X_i$  та нормальному розподілу величини  $Y$ .

Також як альтернативний варіант, можна запропонувати метод “найглибшої” регресії. Такий метод представляє собою пошук класу прямих, які при обертанні навколо покривають якнайбільш максимальну кількість експериментальних точок. Якщо така усова виконується, то потрібно зробити вибір середніх прямої із сукупної кількості цих прямих. Можна засвідчити, що основною перевагою даного методу є стійкість до грубих промахів. На жаль, появу грубих промахів при проведенні кількісного аналізу не можна ігнорувати. Однак, після відкидання таких промахів є можливість використання усіх переваг МНК.

Властивості оцінок МНК [3]:

- Незміщенність оцінок параметрів, тобто математичне сподівання оцінок рівняється істинній величині параметрів;
- Ефективність оцінок МНК, тобто мають мінімальні похибки;
- Оцінки параметрів розподілені за нормальним законом;

Перш за все, необхідно визначити можливу кількість грубих промахів, або таку кількість промахів, яка може з'явитися при виконання градуювання. Взагалі, загальне число вимірювань  $m$  для даного методу майже ніколи не перевищує 12 (часто проводять 4-5 градуювальних зразків). Тому можна стверджувати той факт, що ймовірність появи великої кількості промахів мала.

Дійсно, при налагодженій процедурі вимірювань ймовірність грубого промаху для одиничного вимірювання не перевищує 0,02 (це оцінка зверху; насправді ця величина зазвичай менше). При цьому ймовірність появи одного

грубого промаху під час градування в цілому менше 0,22, а двох - менше 0,05. Іншими словами, ймовірністю появи двох грубих промахів відразу в ході градування можна знехтувати.

Для виявлення грубих промахів при регресійному аналізі найчастіше використовуються методи, засновані на аналізі залишків [4]

$$e_i = Y_i - \hat{Y}_i \quad (5)$$

де  $\hat{Y}_i$  - точки на знайдений прямій, відповідні  $X_i$ . Коректно розраховувати «студентизовані» залишки

$$r_i = \frac{e_i}{[S_{Y/X}(1-h_i)^{1/2}]} \quad (6)$$

де  $s_{Y/X}$  - стандартне відхилення залишків  $e_i$  та

$$h_i = 1/m + \frac{(X_i - \bar{X})^2}{(m-1)s_X^2}, \quad s_X^2 = \sum (X_i - \bar{X})^2 / m-1 \quad (7)$$

До величин  $r_i$  можна застосувати всі методи виявлення грубих помилок, розроблені для нормально розподілених сукупностей. Іншим способом є використання відстані Кука  $d_i$ , статистики, яка показує ступінь зміни коефіцієнтів рівняння регресії при виключенні окремої точки

$$d_i = \frac{e_i^2 h_i}{s_{Y/X}^2 (1-h_i)^2} \quad (8)$$

Величини  $d_i$  завжди більше нуля і в нормі не перевищують. Якщо  $d_i$  приймає значення більше 1, то відповідну величину розглядають як грубий промах. Після відкидання грубого промаху (якщо він був достовірно виявлений) ніщо не заважає застосовувати МНК для знаходження оцінок  $a$  і  $b$ .

Слід пам'ятати, що метод найменших квадратів справджується тільки для лінійних рівнянь, або для таких рівнянь, що припускають можливість зведення до лінійних рівнянь.

## ВИСНОВКИ

Таким чином, для оцінки коефіцієнтів градувальної характеристики при кількісному аналізі буде доцільним використовувати метод найменших квадратів (МНК) після відкиду грубих промахів. Однак слід пам'ятати, що потрібно перевіряти виконання основних положень для кожної методики.

## СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

- [1] Т.І. Ахметова “Статистика в кількісному аналізі”, 2013, ст. 16-20
- [2] [http://www.geol.univ.kiev.ua/lib/zhukov\\_n\\_n/MC\\_5-6.pdf](http://www.geol.univ.kiev.ua/lib/zhukov_n_n/MC_5-6.pdf)
- [3] <http://chemistry.univer.kharkov.ua/files/lin-mnk.pdf>
- [4] <https://files.stroyinf.ru/Data2/1/4293768/4293768858.pdf>

*Наук. Керівник – д. т. н., проф. Володарський Є.Т.*



**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ**

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ  
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ  
імені ІГОРЯ СІКОРСЬКОГО»**

**ПРИЛАДОБУДІВНИЙ ФАКУЛЬТЕТ**

***XII Всеукраїнська науково-практична конференція студентів,  
аспірантів та молодих вчених***

# **"ПОГЛЯД У МАЙБУТНЄ ПРИЛАДОБУДУВАННЯ"**

15-16 травня 2019 р.  
м. Київ, Україна

**Збірник праць конференції**



КИЇВ 2019

*XII Всеукраїнська науково-практична конференція студентів, аспірантів та молодих вчених «ПОГЛЯД У МАЙБУТНЄ ПРИЛАДОБУДУВАННЯ», 15-16 травня 2019 року, КПІ ім. Ігоря Сікорського, м. Київ, Україна*

*Б.М. Шуба, студент гр. ВВ-81мд, д.т.н., проф. Єременко В.С.*  
**СИСТЕМА АНАЛІЗУ І СИНТЕЗУ СИГНАЛІВ ІЗ ЗАДАНИМИ**  
**ХАРАКТЕРИСТИКАМИ** ..... - 465 -

## **СЕКЦІЯ 10. АВТОМАТИЗАЦІЯ ЕКСПЕРИМЕНТАЛЬНИХ ДОСЛІДЖЕНЬ**

*В.А. Володарський, студент гр. ВМ-51-а, к.т.н., доц. Добролюбова М.В.*  
**СИСТЕМА УПРАВЛІННЯ СОНЯЧНОЮ БАТАРЕЄЮ НА БАЗІ**  
**ОДНОКООРДИНАТНОГО СОНЯЧНОГО ТРЕКЕРУ** ..... - 470 -

*Лівадіна А.Ю., студент гр. ВА-71-мн*  
**АВТОМАТИЗОВАНА СИСТЕМА ІДЕНТИФІКАЦІЇ ОСОБИСТОСТІ ЗА**  
**ЗОБРАЖЕННЯМ НА ОСНОВІ МЕТОДУ EIGENFACE** ..... - 474 -

*Д.М. Семенко, студент гр. ВМ-51-1, О.В. Стаценко, к.т.н., доц.*  
**АВТОМАТИЗАЦІЯ КЕРУВАННЯ ВЕНТИЛЯЦІЄЮ ОФІСНИХ ПРИМІЩЕНЬ** ... - 478 -

*О.В. Токаренко, студент гр. ВМ-51-1, к.т.н., доц. Ю.С. Шумков*  
**ПОХИБКА ДИСТАНЦІЙНОГО ВІМІРЮВАННЯ ІМІТАНСУ** ..... - 481 -

*В.О. Хамич, студент гр. ВМ-51-1*  
**ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ ДЛЯ GPS-МОНІТОРИНГУ** ..... - 485 -



УДК 62-67

*В.А. Володарський, студент гр. ВМ-51-а, к.т.н., доц. Добролюбова М.В.  
КПІ ім. Ігоря Сікорського*

### **СИСТЕМА УПРАВЛІННЯ СОНЯЧНОЮ БАТАРЕЄЮ НА БАЗІ ОДНОКООРДИНАТНОГО СОНЯЧНОГО ТРЕКЕРУ**

*Анотація.* Представлені матеріали, що описують склад та принципи функціонування розробленої системи управління сонячною батареєю на базі однокоординатного сонячного трекеру для підвищення ефективності видобутку енергії.

*Ключові слова:* відновлювана енергія, галіоенергетика, фотоелектричні модулі, система стеження за Сонцем.

#### **ВСТУП**

Станом на теперішній час в багатьох країнах світу спостерігається постійне зростання або відчутне коливання цін на нафтопродукти, газ та відповідну сировину. Це, безумовно, є підставою для розгляду способів отримання так званої відновлюваної (альтернативної) енергії. Відновлювана енергетика представляє собою комплекс найбільш перспективних технологічних рішень – вітро-сонячні системи, геліосистеми, теплові насоси та багато іншого – щодо генерації традиційної електричної або теплової енергії від енергій Сонця, вітру та природного стан ґрунту на поверхні Землі [1]. На даний час ці технології не достатньо поширені у порівнянні з традиційними способами, але становлять інтерес внаслідок абсолютної екологічності у виробництві без будь-яких шкідливих відходів та малого ризику заподіяння шкоди навколишньому середовищу, простоті в експлуатації, тривалого часу роботи систем, мінімального сервісного та технічного обслуговування, доступності і невичерпності ресурсів, економічної ефективності. Слід зазначити, що електрозабезпечення, гаряче водопостачання та опалення, які базуються на технологіях, пов'язаних з відновлюваною енергетикою, можливі без підключення до централізованої системи та мають достатньо високий рівень захисту від відключень електромереж та постійних перепадів напруги.

Наша країна безперечно має дуже великі перспективи щодо впровадження цього життєво необхідного та важливого ресурсу на споживчий ринок як для приватних осіб, так і для великих компаній, заводів і підприємств. За останні роки держава все більше сприяє розвитку даного перспективного напрямку, підтримуючи його дотаціями та необхідною законодавчою базою – успішно функціонує ринок мережних станцій, всім відомий «зелений тариф», проводиться безліч виставок та форумів, розробляються програми на рівні державних, обласних та районних бюджетів по впровадженню даних технологій на соціальних об'єктах [2].

Серед десяти найбільш багатообіцяючих джерел енергії майбутнього варто звернути увагу на сонячну енергію або галіоенергетику. Сонячні електростанції – одні з найпоширеніших на планеті, оскільки вони, виробляючи електричний струм, використовують невичерпне джерело енергії – сонячне світло. У процесі вироблення електрики, вони не завдають жодної шкоди довкіллю. Першим практичним застосуванням сонячної батареї було живлення орбітальних супутників та інших космічних апаратів, але на сьогоднішній день більшість



фотоелектричних модулів використовують на промислових підприємствах та для видобутку електроенергії в побуті, а при необхідності ще й тепла для обігріву житлових приміщень та подачі гарячої води [3].

Отже, метою статті є представлення підходів щодо розробки системи управління сонячною батареєю на базі однокоординатного сонячного трекеру для підвищення ефективності видобутку енергії.

## ОСНОВНІ РЕЗУЛЬТАТИ

Результати дослідження з оцінки використання сонячного випромінювання для задач енергетики свідчать про економічну неефективність та комерційну непривабливість фотоелектричних модулів (ФМ) в умовах сучасної цінової політики. Досягнути зниження вартості виробництва електроенергії можливо двома способами: підвищенням ефективності видобутку енергії або зниженням вартості ФМ.

Зниження вартості ФМ реалізується шляхом здешевлення матеріалів галіоприймача або шляхом здешевлення виробництва (автоматичне виробництво ФМ).

Підвищення ефективності видобутку енергії, в свою чергу, можливо через:

- використання двосторонніх фотоелектричних підсилювачів, що незначно підвищують ефективність при значних підвищеннях цін;
- додавання концентраторів, які збільшують фото-віддачу, але при цьому спостерігається нагрівання фотоелементів, що негативно позначається на ККД та тривалості використання;
- застосування багатопшарових фотоприймачів на гетеропереходах, але їх вартість вища ніж у звичайних, тому відношення ціна/ефективність фактично не змінюється;
- оснащення фотоелементів системою стеження за Сонцем, що дає найбільш прийнятне відношення ціна/ефективність.

Останній варіант на даному етапі розвитку сонячної енергетики найбільш актуальний. Кут падіння сонячних променів є головною величиною, що впливає на добуту потужність ФМ, – навіть при найбільш ефективній стаціонарній установці ФМ програш в видобутку потужності складає до 50 %, ніж при безперервному орієнтуванні на Сонце. Використовуючи системи стеження, можна змінювати кут нахилу ФМ протягом дня таким чином, щоб зберігати прямий кут падіння сонячних променів на його поверхню, що дозволяє збільшити кількість потужності, яка надходить, і, як наслідок, кількість потужності, що виробляється. Даний спосіб підходить для вже працюючих сонячних електростанцій – потрібно лише змінити опорну конструкцію.

До основного складу розроблюваної системи управління увійшли:

1. Кроковий двигун для обертання сонячної батареї 28BYJ-48, який має достатню точність для сонячного трекеру, меншу ціну, ніж сервопривід і високу надійність.

2. Мікросхема ULN2003, що йде у комплекті з двигуном та використовується як драйвер.



3. Мікроконтролер Arduino, зокрема плата Arduino Nano, яка застосовується для управління процесом обертання крокового двигуна, роботою датчиків і визначення точки максимальної потужності. Перевагами даного мікроконтролера є простота інтерфейсу, швидкість обробки даних, набір портів вводу/виводу тощо.

4. Блок живлення, що реалізується підключенням через USB до джерела напруги 220 В.

Структурна схема системи управління сонячною батареєю на базі однокординатного сонячного трекера представлена на рисунку 1.



Рисунок 1. Система управління сонячною батареєю на базі однокординатного сонячного трекера

Оскільки система повороту сонячної батареї базується на обертанні валу крокового двигуна, при її першому запуску необхідно встановлювати точку відліку. Тобто, необхідна система, яка допоможе відслідковувати стан валу. Для цієї цілі необхідно використати будь-який датчик, при спрацюванні якого обертання валу зупиниться в потрібному положенні, і це положення можна буде прийняти за нульове (точка відліку). Точка відліку встановлюється програмним шляхом, відлік кроків двигуна ведеться від даного положення. Для встановлення точки відліку в розроблюваній системі використовується геркон SF1 і розташований на конструкції тримача сонячної батареї магніт [4]. Вал, обертаючись проти годинникової стрілки, «підводить» магніт до геркона, в результаті чого останній замикається і подає керуючий сигнал на мікроконтролер Arduino з метою зупинки вала двигуна і фіксації нульового положення. Порт аналого-цифрового перетворювача мікроконтролера очікує замикання геркона SF1 і прийому сигналу. До тих пір, поки на даний порт не подасться ніяких сигналів, він буде видавати випадкові значення перетворення, в результаті власних шумів і наведень із зовнішнього середовища. Для того, щоб цього уникнути, порт A0 підключається на землю через резистор R5 в 10 кОм, який використовується для обмеження струму. Без використання резистора в результаті «спрацювання» геркона мікроконтролер замкнеться своїм внутрішнім джерелом живлення безпосередньо на землю і вийде з ладу як при

короткому замиканні. Після зупинки обертання валу, від прийнятої точки відліку відбувається наступний етап руху крокового мотору з метою пошуку максимуму вироблення. Керуючі сигнали з цифрових портів D4-D7 мікроконтролера надходять на порти IN1-IN4 драйвера крокового двигуна ULN2003. В свою чергу, отримуючи керуючі сигнали, які представляють собою логічні сигнали з високим або низьким значенням напруги, драйвер подає струм на обмотки статора крокового мотору 28BYJ-48, змушуючи обертатися ротор в потрібному напрямку на заданий кут з певною швидкістю. Потужність сонячної батареї складає 3 Вт, а напруга, що нею видається – до 15 В на холостому ході при яскравому сонячному світлі. Точка максимального видобутку енергії визначається по напрузі резистора R4, що зчитується АЦП з порту мікроконтролера. Залежно від опору навантаження значення, вимірювані АЦП, можуть перевищувати гранично допустимі значення. Також, оскільки АЦП мікроконтролера працює в діапазоні напруг від 0 до 4,1 В, використовується дільник напруги з резисторів R1-R4 номіналами по 12 кОм. Конденсатор C1, ємністю 1 мкФ, підключений паралельно до резистора R4, використовується як згладжувальний фільтр, щоб зменшити пульсацію напруги з сонячної батареї, яка має місце бути при використанні системи в приміщенні із штучним освітленням.

### **ВИСНОВОК**

Отже, завдяки розробленій структурній та розрахованій принциповій схемі, можливо виділити декілька переваг використання однокоординатного сонячного трекера, як способу генерації альтернативної енергії – мінімальний набір необхідних елементів та відносно мала вартість надають можливість розробити систему, що дозволить збільшити приріст видобувної енергії, як мінімум, на рівні 40 % у співвідношенні зі стаціонарними сонячними модулями. В перспективі планується розробка двокоординатного сонячного трекера на базі одноплатного комп'ютеру Raspberry Pi, що підвищить ефективність використання системи відслідковування сонця та дозволить виправити деякі недоліки однокоординатного трекера.

### **СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ**

- [1] Альтернативні джерела енергії / Вікіпедія. – Режим доступу: <https://uk.wikipedia.org/wiki/> – 15.04.2019 р.
- [2] Зелений тариф 2019 / Sunsay Energy. – Режим доступу: [https://sunsayenergy.com/green?gclid=EAlaIqobChMly4WX19Lk4QIVy5QYCh1f8Q8XEAAAYASAAEgJ1L\\_D\\_BwE](https://sunsayenergy.com/green?gclid=EAlaIqobChMly4WX19Lk4QIVy5QYCh1f8Q8XEAAAYASAAEgJ1L_D_BwE) – 15.04.2019 р.
- [3] Галіоенергетика: загальні відомості, основні поняття, визначення / Helpiks – Режим доступу: <https://helpiks.org/3-51264.html> – 16.04.2019 р.
- [4] Геркони в схемах на МК / Nauchebe – Режим доступу: <http://nauchebe.net/2011/03/gerkony-v-sxemax-na-mk/> – 10.03.2019 р.

*Наук. керівник – к.т.н., доц. Добрялюбова М.В.*