

Національна академія наук України
Інститут проблем безпеки атомних електростанцій

А. Д. Скорбун

**КОМП'ЮТЕРНА (ОБЧИСЛЮВАЛЬНА) СТАТИСТИКА
МЕТОДИ МОНТЕ-КАРЛО В ЗАДАЧАХ ОЦІНКИ
НЕВИЗНАЧЕНОСТІ РЕЗУЛЬТАТІВ ВИМІРЮВАНЬ**

Монографія

Чорнобиль 2024

УДК 519.245:539.16](02)

С49

Рецензенти:

д-р фіз.-мат. наук С. В. Гобелков,
канд. фіз.-мат. наук І. В. Жиганюк,
канд. техн. наук М. І. Панасюк

За редакцією акад. НАН України
А. В. Носовського

*Рекомендовано до друку вченою радою
Інституту проблем безпеки АЕС НАН України*

Скорбун А. Д.

С49 Комп'ютерна (обчислювальна) статистика. Методи Монте-Карло в задачах оцінки невизначеності результатів вимірювань: монографія / А. Д. Скорбун; НАН України, Ін-т проблем безпеки АЕС. – Чорнобиль (Київ. обл.): Ін-т проблем безпеки АЕС, 2024. – 172 с.

ISBN 978-617-14-0330-7

Представлено способи обробки даних методами комп'ютерної (computational) статистики. Радіаційні вимірювання, багаторічні моніторингові дослідження міграції радіонуклідів, оптимізація відбору проб і визначення в них вмісту радіонуклідів та компонентів хімічного складу, визначення кристалічних фаз речовин методами рентгенівської дифракції та інші проблеми досліджень у Чорнобильській зоні потребували використання статистичного аналізу. Показано, що методи комп'ютерної статистики успішно можна використовувати для вирішення наукових та практичних задач.

Для науковців та інженерно-технічних працівників, які працюють у галузі ядерної фізики, атомної енергії, радіоекології, радіаційного матеріалознавства та моніторингових досліджень, а також для студентів та аспірантів фізичних, хімічних і геологічних спеціальностей університетів.

УДК 519.245:539.16](02)

ISBN 978-617-14-0330-7

© А. Д. Скорбун, 2024

*Назарчику, не бійся сумніватися:
не такі вони вже непорушні,
ці підвалини будівлі фізики.*



Автор виконує гамма-спектрометричні
вимірювання під час будівництва в радіоактивно
забруднених ґрунтах котлованів для фундаментів
нового безпечного конфайнмента («Арки»)
Чорнобиль, 2010

Зміст

Перелік скорочень	6
Передмова	7
1. ПРО ДЕЯКІ ПРОБЛЕМИ СТАТИСТИЧНОГО АНАЛІЗУ ДАНИХ ...	11
1.1. Основні принципи і припущення	11
1.2. Про поняття невизначеності результатів вимірювань	14
1.3. Принципи комп'ютерної статистики.....	18
1.3.1. Метод перестановок (resampling, permutation test). Аналіз ступеня відмінності параметрів двох вибірок (їхніх середніх, дисперсій чи інших статистик).....	20
1.3.2. Аналіз кореляцій	26
1.3.3. Джекнайф (jackknife) метод	32
1.3.4. Оцінка ймовірності появи певного значення серед елементів вибірки	32
1.3.5. Метод бутстрапу (bootstrap).....	34
2. ПРИКЛАДИ ЗАСТОСУВАННЯ КОМП'ЮТЕРНОЇ СТАТИСТИКИ ДО АНАЛІЗУ РЕАЛЬНИХ ЗАДАЧ	44
2.1. Оцінка надійності статистичного аналізу шумів реакторного обладнання на основі теорії гіпервипадкових явищ.....	45
2.1.1. Вступ	45
2.1.2. Опис експериментальних даних.....	47
2.1.3. Аналіз даних на гіпервипадковість	48
2.1.4. Результати аналізу.....	51
2.1.5. Обговорення і висновки	52
2.2. Кореляційний метод виділення слабких сигналів з фону	55
2.2.1. Постановка розрахункової задачі	57
2.2.2. Кореляційний метод обробки гамма-спектрів	60
2.2.2.1. Тест перестановок	60
2.2.2.2. Перевірка отриманих результатів на реальних спектрах.....	67
2.2.2.3. Можливості методу по визначенню інтенсивності ліній.....	77
2.2.2.4. Висновки до пошуку гамма-ліній	78
2.2.3. Метод ідентифікації кристалічних фаз з даних рентгенівської дифракції від багатофазних зразків з низьким їхнім відносним вмістом	79
2.2.3.1. Кореляційний метод аналізу дифрактограм (пошуку фаз)	80
2.2.3.2. Демонстраційні розрахунки.....	85
2.2.3.3. Аналіз даних рентгенівської дифракції	85
2.2.3.4. Модельна дифрактограма.....	86

2.2.3.5. Приклад з даними рентгенівської дифракції.....	89
2.2.3.6. Новий метод ідентифікації фаз в РФА.....	90
2.2.4. Метод оцінки вмісту кристалічних фаз у багатофазних зразках за даними рентгенівської дифракції	91
2.2.4.1. Дані рентгенівської дифракції для аналізу.....	93
2.2.4.2. Кореляційний аналіз на основі методів Монте-Карло	93
2.2.4.3. Розрахунок.....	94
2.2.4.4. Параметр взаємної кореляції даних рентгенівської дифракції..	95
2.2.4.5. Алгоритм оцінки вмісту кристалічних фаз	97
2.2.4.6. Формули для розрахунку вмісту фаз у зразках	98
2.2.4.7. Тестування методу	100
2.2.4.8. Оцінка невизначеності методу	101
2.2.4.9. Висновки до методу оцінки вмісту кристалічних фаз у багатофазних зразках.....	102
2.2.5. Резюме до розділу про кореляційний аналіз	103
2.3. Залежність невизначеності результату від невизначеності вхідних параметрів на прикладі оцінки вмісту контейнера з РАВ.....	107
2.3.1. Методика розрахунків	108
2.3.2. Постановка задачі	109
2.3.3. Обговорення результатів	113
2.3.4. Резюме до розділу про контейнер	115
2.4. Аналіз впливу нерегулярності даних	117
2.5. Оцінка гауссовості даних за допомогою модифікованого методу нормованого розмаху (R/S analysis)	121
2.5.1. Критерій відхилення від нормальної послідовності	124
2.5.2. Застосування до експериментальних даних.....	126
2.6. Визначення похибки вимірювання при невідомій аналітичній залежності функції виходу від параметрів входу.....	129
2.6.1. Постановка задачі	129
2.6.2. Метод розв'язування і оцінка його точності	133
2.6.3. Оцінка необхідної кількості вимірювань.....	134
2.6.4. Висновки.....	134
2.7. Використання методу джекнайф (jackknife) для оцінки невизначеності середнього.....	136
Статистичне моделювання.....	138
2.8. Метод геохімічної статистики для пошуку кореляційних зв'язків між хімічним складом та концентраціями ^{90}Sr в підземних водах	141
Додаток А.....	151

ПЕРЕЛІК СКОРОЧЕНЬ

ДСТУ	– державний стандарт України
ЗВТ	– засіб вимірювальної техніки
ЛПВМ	– лавоподібні паливовмісні матеріали
МВА	– мінімальна активність, яку можна достовірно визначити
МДА	– мінімальна активність, яку можна детектувати
ОСГИ	– образцовый спектрометрический гамма-источник
ПЕД	– потужність еквівалентної дози
РАВ	– радіоактивні відходи
РФА	– рентгенівський фазовий аналіз
ТРВ	– тверді радіоактивні відходи
ЧАЕС	– Чорнобильська атомна електростанція
DL	– decision level – поріг виявлення (МДА)
DT	– decision threshold – поріг прийняття рішення (МВА)
ISO	– international standard organization
R/S аналіз	– rescaled range analysis – метод нормованого розмаху
u	– стандартна невизначеність
U	– розширена невизначеність
u^2 ($\equiv \sigma^2$)	– дисперсія

ПЕРЕДМОВА

Кілька причин спонукали автора написати цю книжку.

Перша причина – основана на проблемах, що виникли при розв'язанні численних задач, пов'язаних з дослідженням характеристик радіоактивного забруднення навколо ЧАЕС, у яких автор брав участь. Діяльність у зоні відчуження, особливо при роботах по перетворенню об'єкта "Укриття" в екологічно безпечну систему, супроводжувалася великою кількістю різноманітних радіаційних вимірювань. Усі пов'язані з цим розглянуті нижче питання – гамма-спектрометрія низькоактивних проб (на рівні МДА), оцінка статистично достатньої необхідної кількості проб для одержання надійного результату, питання оцінки невизначеності результатів вимірювань, аналіз характеристик нейтронних потоків від ЛПВМ зруйнованого 4-го енергоблоку ЧАЕС тощо – вимагали статистичної обробки, при виконанні якої виявилася непридатність стандартних методів математичної статистики для цих задач. У результаті було розроблено нові підходи або неочевидне застосування та модифікація відомих підходів до наших задач. Знайомство з цими методами, на думку автора, має бути корисним для практиків.

Друга причина – абсолютне незнайомство знайомих мені спеціалістів з методами так званої комп'ютерної (computational) статистики, яка була широко застосована для розв'язання згаданих вище задач. Існує дві категорії спеціалістів, яким важко сприйняти необхідність застосування нових методів статистики. Перша – це спеціалісти без відповідної математичної підготовки (наприклад, біологи, соціологи), які повністю покладаються на те, що написано в широко розповсюджених підручниках чи керівництвах. Про потрібні їм методи комп'ютерної статистики та необхідність їхнього застосування їм просто немає звідки навіть почути. У той час як саме їм, у зв'язку із специфікою їхніх задач (короткі вибірки з великим розкидом даних), комп'ютерна статистика буде дуже корисна. Друга – це підготовлені інженери, які займаються розробкою нової апаратури, та, відповідно, спеціалісти-метрологи. Вони у своїй роботі змушені користуватися лише нормативними документами, де весь статистичний аналіз побудовано на гауссовому розподілі, і відійти від цього вони просто не мають права. Навіть у нормативних документах можна знайти приписку, що, якщо розподіл даних відрізняється від гауссового, треба використовувати інші методи аналізу даних, проте не вказується, які. Отже, знайомство з новими підходами також має бути для них корисним.

Третя причина – статистичний аналіз довгих рядів вимірювань має свої особливості, які також не входять у стандартні прийоми статистичної обробки. У той же час довгі ряди вимірювань зараз є широко розповсюдженим типом даних і застосування описаних тут прийомів може допомогти витягнути з них додаткову корисну інформацію.

Можливо, останні дві причини пов'язані з недостатнім знайомством нашої наукової спільноти з англійською літературою. Тому що розвиток

нових статистичних методів обробки даних широко обговорюється в сучасній англомовній літературі. Правда, якщо бути об'єктивним, в основному обговорюється розвиток цих підходів. Використання методів комп'ютерної статистики, у всякому разі у статтях в області інтересів автора, також не зустрічалось.

Статистичний аналіз по своїй суті є спрямованим на допомогу досліднику у прийнятті рішення [1, с. 9; 2]. Тому ставлення до результатів статистичної обробки завжди буде критичним у випадках, коли за такими результатами робиться висновок про наявність чи відсутність шуканого ефекту, і аналіз невизначеностей у цій ситуації набуває особливого значення. Можна вказати близькі до нас випадки, коли оцінка невизначеності результату має не тільки формальне, але й принципове значення. Усі радіологічні вимірювання підпадають під закон України "Про метрологію та метрологічну діяльність" і повинні обов'язково супроводжуватись оцінкою їхньої невизначеності. Проте на практиці до цього часто ставляться формально в силу значних труднощів таких оцінок і неможливості контролю їхньої правильності. І хоча у значній кількості випадків для прийняття рішень досить мати якісні дані про рівні радіаційного забруднення, у ряді випадків одержання точних цифр є принциповим. Прикладом першого випадку може бути оцінка загальної кількості похованої у сховищах активності, яка, хоч і є обов'язковою, робиться дуже приблизно в силу об'єктивних труднощів такої оцінки (частину з яких описано в даній книзі), а результат мало впливає на прийняття управлінських рішень. Прикладом другого випадку може бути задача віднесення РАВ до певної категорії: якщо в результаті вимірювань виявляється, що РАВ знаходяться на межі між середньо- та високоактивними, рішення про віднесення їх до середньо- чи високоактивних (яке пов'язано з питанням про точність отриманих результатів) різко змінює організаційні і фінансові витрати на поводження з ними і тому вимагає надійно встановлених, метрологічно обґрунтованих меж результатів вимірювань. Аналогічно питання про присутність чи відсутність радіонукліда у пробі може бути принциповим для задач екології, де результати вимірювань часто знаходяться на межі чутливості методу. Про важливість вимірювань в області малих концентрацій свідчить те, що вони регламентуються спеціальним міжнародним стандартом ISO 11929. Тому завдання одержання обґрунтованих оцінок невизначеності результатів радіаційних вимірювань є актуальним і, додамо, у цілому не розв'язаним.

Оскільки книжка носить методичний характер, вона побудована наступним чином. У першій частині обговорюються основні завдання і підходи класичної математичної статистики, необхідні для розуміння проблем, що аналізуються, а також розглянуто основні прийоми комп'ютерної статистики. У другій частині розглянуто окремі нетривіальні задачі з різних прикладних галузей науки, де була використана комп'ютерна статистика (методи Монте-Карло), і які демонструють неможливість їхнього розв'язання іншими способами. У додатку А наводяться основні формальні поняття

математичної статистики, необхідні для розуміння проблем, що обговорюються в книзі.

Література наводиться лише для ілюстрації висловлених тверджень, вибрана на смак автора і не претендує на повноту.

Слід вказати також, що поява комп'ютерної статистики викликана потребами практики, які не могли бути задоволеними без використання комп'ютерів. Не викликає сумнівів продовження використання зараз і в осяжному майбутньому стандартної математичної статистики в таких областях, наприклад, як контроль якості продукції, де контролюють зміну середнього або дисперсії результатів вимірювань. Проте аналіз коротких вибірок з великим розкидом даних або пошук інформації, прихованої у шумових сигналах, уже потребують інших підходів статистичного аналізу.

Для підготовлених читачів хочеться зробити таке застереження. Методи математичної статистики виникли з потреб практики, а саме з необхідності мати спосіб якимось чином характеризувати дані, що мають вигляд випадкових та шумоподібних. У результаті і стандартна математична статистика, і розглянута у цій книжці комп'ютерна статистика намагаються «найкращим» способом зробити деякі усереднені оцінки таких даних. Цей підхід досить добре зарекомендував себе в багатьох практичних задачах, наприклад в аналізі розкиду однотипних вимірювань, на чому у свою чергу будуються методи контролю якості продукції.

Але є області досліджень, коли методи класичної математичної статистики (виведені на сьогоднішній день у ранг аксіом й уведені в нормативні документи) є абсолютно непридатними.

Так, наприклад, у теорії гіпервипадковості І. І. Горбаня [3] розвивається ідея, що при всіх вимірюваннях (особливо природних процесів) час від часу трапляються зовнішні впливи, які спотворюють класичну статистику, а отже порушується базове положення теорії ймовірностей і математичної статистики, що дані повинні бути незалежними й однаково розподіленими.

У теорії С. Ф. Тімашева [4] вважається, що будь-який природний процес проходить через періоди сплесків і спокійного розвитку (різної тривалості), що теж порушує базовий принцип незалежності, випадковості й однакового розподілу даних.

Нарешті, існують погляди (С. Е. Шноль [5]), що в природі немає випадкових явищ узагалі, що всьому є причина, до якої можна докопатися.

Якщо згадати відому приказку «за деревами не видно лісу», то методи статистики призначені характеризувати саме ліс, не розглядаючи окремі дерева. При цьому вони тільки нівелюють приховані в шумових сигналах закономірності.

Отже, використовуючи ті чи інші методи математичної статистики, слід пам'ятати, що вони мають свої області застосування, де їхнє надійне використання перевірено практикою (не плутати із ситуацією, коли результат вважається правильним, тому що було використано загальноприйнятий метод), і при аналізі даних бути свідомим щодо тих областей, до яких належать

як власне дані, так і запропоновані методи їхнього аналізу, де ці підходи не працюють.

Автор вдячний колегам за моральну підтримку, Н. Покотилу за допомогу в підготовці рисунків, редактору Л. М. Троян, а особливо Б. М. Сплошному за обговорення проблем оцінки невизначеностей з точки зору метрології та аналіз нормативних документів.

Чорнобиль – Київ – Буковець 2017 – 2021

1. ПРО ДЕЯКІ ПРОБЛЕМИ СТАТИСТИЧНОГО АНАЛІЗУ ДАНИХ

Комп'ютерна або обчислювальна статистика (computational statistics) – це не аналіз даних за допомогою комп'ютера-калькулятора. Це аналіз даних з використанням методів Монте-Карло (методу випадкових випробувань), який практично неможливо зробити без використання комп'ютера. Не дивлячись на те, що комп'ютерна статистика давно ввійшла в наукову практику, для публікації робіт, які розвивають ці методи, видаються спеціалізовані журнали (Computational Statistics and Data Analysis, Computational Statistics, Journal of Statistical Computation and Simulation та ін.), у широку практику, а тим більше в нормативні документи вона не ввійшла; і в керівництвах (підручниках), при викладанні питань використання комп'ютерних методів у статистиці розглядаються можливості пакетів програм з математичної статистики, тобто обробки даних на основі стандартних підходів теорії ймовірностей.

Ця ж книга присвячена розгляду принципово інших підходів до тих задач по аналізу даних, які не може розв'язати математична статистика, яку в подальшому інколи будемо називати класичною (математичною) статистикою, щоб підкреслити відмінність від комп'ютерної статистики.

1.1. Основні принципи і припущення

Усі вимірювання породжують ряди чисел – результатів вимірювання деякої фізичної величини. У багатьох випадках це будуть випадкові числа в тому сенсі, що результат одного вимірювання не залежить від попереднього. Так, інтуїтивно можна вважати, що випадковими будуть результати вимірювань розміру однієї і тієї ж деталі, результати вимірювань розмірів у партії однотипних деталей, зріст кількох людей, які пройшли повз спостерігача, число автомобілів, що проїхали біля світлофору за певний відрізок часу, і т. ін. Чи можна знайти щось спільне в таких числах і якимось чином їх характеризувати? Також інтуїтивно зрозуміло, що результати таких вимірювань в основному будуть коливатися в певних межах і лише зрідка будуть сильно відхилятися від цих меж: несподівано пройде карлик або баскетболіст. Так приходять до того, що групи об'єктів, тобто групи чисел як результатів вимірювань, можна характеризувати деяким середнім значенням, функцією розподілу числа людей по зросту (більш формально ці поняття введені в додатку А). Якщо розглядати результати вимірювання параметрів однієї і тієї ж деталі, то різниця між середнім значенням всієї серії вимірювань й окремим результатом називається похибкою окремого вимірювання. І тоді можна розвинути деякий математичний спосіб обробки таких даних (тобто наборів чисел), які дадуть відповідь на те, як часто зустрічаються великі відхилення від середнього, які значення зустрічаються найчастіше тощо. Тобто отримати деякі характеристики такого набору чисел. Для згаданих вище підрахунків кількості автомобілів також можна порохувати такі відхилення від

середнього, і хоча це вже буде не похибка результату, проте математично процедура аналізу не буде відрізнятися і можна також отримати відповіді на ті самі запитання: як часто зустрічаються великі відхилення від середнього, які значення зустрічаються найчастіше тощо.

Математичний апарат для аналізу таких випадкових чисел називається теорією ймовірностей, яка дає відповіді на поставлені вище запитання. Схема процедур теорії ймовірностей така: **якщо взяти набір чисел, які мають такі й такі математичні властивості**, то параметри таких наборів чисел (їхні функції розподілу, середні і т. ін.) можна розрахувати за такими й такими формулами. Прикладом може бути отримання закону Пуассона [6, § 4]: «Виведення закону Пуассона базується на наступних основних припущеннях:

1. Випадкова величина, яка нас цікавить, може приймати лише цілі додатні значення.

2. Якщо довжина інтервалу t прямує до нуля, то ймовірність p_1 також прямує до нуля як нескінченно мала першого порядку, а ймовірності p_2, p_3 і т. д. прямують до нуля як нескінченно малі більш високого порядку.

3. Події, які відносяться до інтервалів, які не перекриваються, статистично незалежні.

Якщо ці умови виконано, то відповідна величина розподілена за законом Пуассона. Порушення хоча б однієї з них веде до неможливості застосування закону Пуассона до даної задачі.»

На відміну від теорії ймовірностей, математична статистика має справу з **числами, які отримано в результаті вимірювань**. Математична статистика також ставить собі за мету знайти спосіб охарактеризувати ці експериментальні набори чисел. При цьому вона покладається на теорію ймовірностей. І тут ми натикаємося на перший підводний камінь: для того щоб використати результати теорії ймовірностей, частіше за все **припускають** з певних міркувань (чи аргументують з тим чи іншим ступенем надійності), що **числа, отримані в результаті вимірювань, мають саме такі й такі математичні властивості**. Тобто якщо в дійсності результати вимірювань відхиляються від таких властивостей або відповідність таким властивостям була визначена помилково, то й використання відповідних формул математичної статистики не буде мати теоретичного обґрунтування. (Інша справа, якою буде величина отриманої при цьому помилки, можливо, і не дуже велика.)

В основі теорії (і, особливо, практики) математичної статистики при аналізі результатів вимірювань лежить віра в те, що випадкові дані частіше за все розподілені за законом Гаусса (нормальний розподіл). У свою чергу ця віра покладається на так звану центральну граничну теорему теорії ймовірностей ([7, с. 180]; [8, с.187]; [9, с. 55]), за якою сумарний вплив багатьох чинників, незалежно від їхніх окремих функцій розподілу, дає в результаті гауссовий розподіл. У результаті, яку б книжку з математичної статистики не взяти, після уважного осмислення можна зробити висновок, що переважну більшість описаних у ній результатів отримано на основі використання гауссового

розподілу. Так, у чудовій книзі Закса [1] описано велику кількість задач по обробці даних, переважно для медиків і біологів, і практично всі способи обробки базуються на припущенні, що розподіл є гауссовим. Навіть непараметричні методи типу Колмогорова - Смирнова у кінцевому результаті використовують розподіл χ^2 , який, у свою чергу, є похідним від нормального розподілу. У [10] у змісті є прямо розділ «гауссовий і пов'язані з ними розподіли (t-розподіл, χ^2 , ...)». У [7] – книзі для астрономів з обробки результатів вимірювань – є лише коротеньке згадування про існування розподілів, відмінних від гауссового, але робота з такими даними глибоко не аналізується. У [8] також у подробицях розглядаються способи аналізу даних для технічних задач, проте теж лише коротко вказується як перевірена практикою аксіома, що результати вимірювань (наприклад, розмірів деталей) розподілені за гауссовим законом. Відповідно всі довідкові таблиці – це таблиці для гауссового розподілу. Більше того, мало хто знає, що формула для звичайного коефіцієнта кореляції (Пірсона) отримана у припущенні гауссового розподілу [7, с. 329], а для негауссових даних розроблено коефіцієнт кореляції Спірмена [1, с. 368]. І, нарешті, усі нормативні документи, як ДСТУ [11], так і міжнародне керівництво з розрахунку невизначеностей [12], також усі розрахунки дають для гауссового розподілу. Там можна лише зустріти приписку дрібним шрифтом: «...якщо розподіл не гауссовий, то треба використовувати інші методи». Про підходи до використання цих «інших» методів можна знайти в [13, с. 113].

Не можна сказати, що автор першим звернув увагу на таку ситуацію (що гауссівський підхід є універсальним у задачах аналізу даних). Ця тема також є предметом багатьох публікацій. Наведемо лише характерні назви двох з них [14, 15]: Орлов А. И. «Часто ли распределение результатов наблюдений является нормальным?»; Дубровский С. А. и др. «О ненормальности нормального закона распределения (вообще и в приложении к металлургии в частности)». У роботі [16, розд. 5.1] наведено результати критичного аналізу даних метрологічних вимірювань і зроблено висновок про те, що використання гауссового розподілу далеко не завжди є допустимим. У [17] аналізується ця ситуація і прямо говориться, що «... ймовірнісна теорія похибок дійсно задумувалася Гауссом і Лапласом як фізична теорія, але в якості фізичної випробування часом не витримала, однак не була і відкинута, оскільки багато разів доводила свою корисність.»

Проте ситуація з підручниками з математичної статистики не змінюється і згадувань про існування сумнівів чи хоча б прямих вказівок на те, що використана теорія базується на певних припущеннях, у них не зустріти. Дозволю собі навести такий приклад. Мій колега, еколог, звернувся до мене з питанням про аналіз його даних по вимірюванню просторового забруднення великих індустріальних територій. Характерною особливістю таких даних є великий розкид їхніх значень, але всі дані вважаються правильними. Це була велика робота, і обробку даних він віддав на кафедру математики. Ті виконавці йому сказали, що у його даних є багато «викидів» (outliers) і для подальшої

обробки дані треба почистити від цих викидів відповідно до відомої формули: відношення $[(\text{відхилення від середнього})/(\text{середньоквадратичне відхилення})]$ не повинно перевищувати певного критерію. Як читач уже здогадується, це критерій викидів по відношенню до гауссового розподілу. Після такої чистки від даних залишилася ледве половина, при цьому вони перестали бути репрезентативними, тому що фактично кожне вимірювання характеризувало деяку точку на карті.

Широку дискусію з цього питання було проведено журналом *Statistical Science* [18], у якій взяли участь провідні світові спеціалісти з математичної статистики. Обговорювалося поставлене Лео Брейманом запитання: чому аналізуються не реальні дані, а приписані їм функції розподілу і чому ситуація не змінюється, не дивлячись на те, що всі професіонали-статистики розуміють цю проблему?

Повний аналіз літератури з цього питання не є предметом даної роботи, тому на підтримку сказаного наведемо лише яскраву назву однієї із статей [19]: **"Signal processing: the future is non-Gaussian, non-linear and non-stationary"** ("Обробка сигналів: майбутнє є негауссовим, нелінійним і нестационарним").

Якщо дані не гауссові, то які вони? Радикальну відповідь пропонує, наприклад, теорія гіпервипадковості І. І. Горбаня [3], у якій припускається-стверджується, що будь-які результати вимірювань (особливо природних процесів) через вплив непередбачуваних явищ час від часу будуть змінювати свої характеристики, тобто реальні процеси є нестационарними і нестійкими у статистичному сенсі, що означає, що результати вимірювань залежать від часу (тривалості) вимірювань. При цьому гауссовість може виникнути лише як окремий випадок на короткому відрізку часу. Такий підхід є розширенням теорії ймовірностей, для нього розвинуто спеціальну математику. Але він підходить для обробки достатньо довгих регулярних рядів вимірювань. Проблема ж аналізу не гауссових коротких вибірок з великим розкидом даних залишається.

Таким чином постає питання, як подолати вказані труднощі аналізу даних? Перед тим, як перейти до розгляду відповідних підходів, розглянемо питання про невизначеність результатів вимірювань, тому що це і є власне проблемою, яка потребує вирішення.

1.2. Про поняття невизначеності результатів вимірювань

Математична статистика, як і будь-яка наука, вимагає для свого використання її поглибленого вивчення та досить високого рівня математичної підготовки, що в дійсності не є просто. Один з провідних спеціалістів з цього питання навіть опублікував статтю про керівництво для викладачів [20]. Але ми спробуємо спростувати цю сентенцію стосовно до її практичного використання. Будь-яка книга з математичної статистики починається якщо не з доведення, то хоча б перераховування теорем теорії ймовірностей, потім виведення формул для розрахунку статистичних параметрів, а потім деяких

прикладів їхнього використання. При цьому як основний результат розглядається розрахунок імовірності реалізації тієї чи іншої події (див., наприклад, [21]).

Проте якщо відкинути необхідність уміти довести ці теореми, відкинути необхідність уміти обґрунтувати всі формули, тобто просто повірити в їхню правильність і залишити лише бажання ними скористуватись (що й реалізується при використанні готових пакетів прикладних програм з математичної статистики), а після цього уважно подивитися на ті задачі, які розв'язує математична статистика (див., наприклад, [1]), то, як не дивно, виявиться, що всі вони зводяться до двох-трьох запитань: чи відрізняються між собою чимось дві вибірки, наприклад, чи відрізняються між собою їхні середні значення або дисперсії σ^2 , або інші параметри; чи існує кореляція між двома вибірками. Сюди не входить навіть питання, чому дорівнюють середнє і дисперсія, тому що ці дві величини розраховуються за стандартними формулами, які не залежать від особливостей статистичних характеристик вибірки (хоча самі ці поняття й породжені теорією ймовірностей), тобто розрахунок їх не потребує спеціальних знань статистичного аналізу.

У чому ж тоді питання? Це **другий прихований підводний камінь статистичного аналізу даних**. Стандартна невизначеність (похибка) результату вимірювань (тобто середньоквадратичне відхилення σ) розраховується однозначно. А от установлення меж достовірності результату – довірчого інтервалу – часто зводиться до правила трьох сігм (σ), що вже ґрунтується на певних припущеннях. Мається на увазі широко розповсюджена віра, що якщо результат знаходиться в межах трьох сігм, тобто відмінність його від деякого середнього значення не перевищує $\pm 3\sigma$, то він гарантовано надійний, бо відмінність не перевищує межі допустимих статистичних флуктуацій. Але це твердження обґрунтовано на припущенні гауссового розподілу результатів, для якого практично всі, а формально 99,73 % елементів вибірки лежать у діапазоні $\pm 3\sigma$ (див. додаток А та [8]). Але якщо дані не гауссові, тобто для інших функцій розподілу, використання критерію 3σ уже втрачає своє обґрунтування.

Таким чином проблеми статистичного аналізу виявилися зведеними до задачі коректного порівняння вибірок, що у свою чергу означає необхідність коректного розрахунку числової оцінки ступеня їхньої відмінності чи подібності, що тісно пов'язано з аналізом точності (чи невизначеності) результатів. Як окремий випадок, можна оцінювати відмінність експериментальних даних від теоретичної вибірки.

У принципі, весь цивілізований світ офіційно замість поняття «похибка результату вимірювань» використовує поняття «невизначеність результату вимірювань». Хоча при розрахунках похибки і невизначеності результат часто той самий, розуміння різниці між цими двома поняттями є надзвичайно важливим у філософському плані, як для власне аналізу результатів вимірювань, так і для обґрунтування необхідності використання методів комп'ютерної статистики, що буде розглянуто нижче. Тісно пов'язаними з цим

є також підходи інтервального аналізу. Тому розглянемо постановку задачі про оцінку невизначеності в подробицях.

У 1993 р. Міжнародна організація по стандартизації, Міжнародний комітет мір і ваг та інші прийняли документ під назвою «Guide for expression uncertainty in measurements» [12] («Керівництво з оцінки невизначеностей вимірювань», далі – «Керівництво з невизначеностей...»), у якому встановлюється поняття «невизначеності» результатів вимірювань замість «точності» або «похибки» вимірювань та підходи до оцінки цієї величини. У 2001 р. Міждержавною радою із стандартизації, метрології і сертифікації було прийнято міждержавний документ [22], учасником якого є і Україна, який є спрощеним роз'ясненням [12]. Проте цей документ ще й досі широко не використовується в Україні, скоріше за все тому, що діючі нормативні метрологічні документи орієнтують на оцінки похибки результатів вимірювань. Наприклад, у діючих документах [23, 24] в якості нормованих метрологічних характеристик зазначено похибки вимірювань. Натомість вимоги до компетентності вимірювальних лабораторій, що встановлюються відповідно до загальних вимог [25], установлюють, що атестовані лабораторії (будь-яка структура, що виконує вимірювання у сфері державного метрологічного нагляду, тобто і радіаційні вимірювання) повинні: а) мати і використовувати методики оцінювання невизначеності вимірювань; б) мати процедури контролю якості для здійснення поточного контролю (моніторингу) достовірності випробувань, що проводяться.

На сьогодні згідно із Законом України «Про метрологію і метрологічну діяльність» дозволяється використання оцінок похибки, проте для результатів вимірювань, що направляються закордонним організаціям, необхідно давати оцінку невизначеностей. Задля справедливості вкажемо, що, не дивлячись на існування [12], проблема переходу на використання невизначеностей у науковій літературі продовжує дискутуватися, питання оцінки невизначеностей широко обговорюються [2, 25 – 29], існує навіть неприйняття цієї ідеології.

Опис проблеми

У цілому мова йде про поступову заміну поняття «похибки результату вимірювань» поняттям «невизначеності» результату вимірювань.

Документи [12, п. 2.2.3; 21, п. 3.1] установлюють такі основні поняття для опису результатів вимірювань:

невизначеність результатів вимірювань: «параметр, пов'язаний з результатом вимірювань, який характеризує розкид (dispersion) значень, які обґрунтовано можуть бути приписані величині, яка вимірюється»;

стандартна невизначеність (u_c): «невизначеність результату вимірювань, виражена у вигляді середньоквадратичного відхилення»;

сумарна стандартна невизначеність (u): «стандартна невизначеність результату вимірювань, отриманого через значення інших величин, яка

дорівнює кореню квадратному суми членів, причому члени є дисперсіями або коваріаціями цих інших величин, зважених у відповідності до того, як результат вимірювань змінюється при зміні цих величин»;

розширена невизначеність (U): «величина, яка визначається як інтервал результатів вимірювань, який, як очікується, включає в себе значну частину розподілу значень, які обґрунтовано можуть бути приписані величині, яка вимірюється».

Необхідність переходу з оцінки похибок на оцінки невизначеностей викликана рядом причин. Є причини прикладного характеру, у першу чергу те, що коли в статті (документі) вказується точність (похибка) результатів, то частіше за все неможливо зрозуміти, як вона одержана, а при поглибленому аналізі виявляється, що зроблена оцінка точності враховує не всі можливі складові похибки результатів і т. ін. Як результат, часто оцінки похибок результатів вимірювань надаються таким чином, що їх важко в подальшому врахувати в нових вимірюваннях. Але є й причини фундаментального характеру. Перехід до аналізу невизначеностей базується на постулатах, що ні точне значення, ні точна похибка оцінки його величини не можуть бути визначені [12 GUM, розділ D5.2]: **«Невизначеність результатів таким чином є відображенням того факту, що для даної величини, яка вимірюється, і для даного результату вимірювання немає єдиного значення, а є нескінченне число значень, розсіяних навколо результату, який узгоджується з усіма спостереженнями і даними, а також із знанням фізичного світу, і який з певним ступенем впевненості може бути приписаний величині, яка вимірюється».** Тобто ми змушені розглядати і результат, і похибку його вимірювання як невизначену величину, для оцінки якої можна вказати лише границі інтервалу, у межах яких вона знаходиться із заданою ймовірністю (із ступенем впевненості). Але залишається невідомим, в якій точці цього інтервалу вона знаходиться. Такий підхід принципово відрізняється від прийнятого в теорії похибок, де вважається, що точне значення існує, хоча експериментально визначити його неможливо (див., наприклад, [7]). Ідея про існування істинного значення є стандартом і в математичній статистиці, де оцінюється відхилення середніх значень від невідомого, але (як вважається) існуючого точного значення [8, с. 86; 1, с. 71].

Отже, слід ще раз підкреслити, що проблема аналізу даних полягає не стільки у визначенні середнього значення результату вимірювань, скільки у визначенні меж, у яких воно може знаходитися. У дійсності природа появи цих величин різна. Середнє значення вимірюваної величини є результатом **моделі об'єкта** вимірювань, а невизначеність результату визначається вибраним **методом вимірювань**. Пояснимо це відомим анекдотичним прикладом. Якщо хтось вирішив виміряти відстань до Місяця за допомогою жердини, то в результаті своїх спроб він одержить нижню межу відстані до Місяця: вона більша, ніж довжина жердини плюс-мінус точність визначення довжини жердини, яку можна зробити досить високою. Якщо провести такі ж вимірювання на високій горі, то нижню межу відстані до Місяця можна

збільшити до кількох кілометрів. Точність визначення вимірюваної відстані залишиться такою ж, а шукане значення зміниться в тисячі разів. Отже, наша задача статистичного аналізу полягає в коректній оцінці меж, у яких може знаходитися результат, тобто розширеної невизначеності результату.

Керівництво з невизначеності [12] розглядає дві різні ситуації оцінки невизначеності результатів вимірювань. Перша – це коли існує достатня статистична інформація (багаторазові вимірювання) – так звана оцінка по методу А. Друга – це коли невизначеність результату встановлюється виходячи із загальних міркувань, наприклад, з відомостей про невизначеність подібних вимірювань, що проводилися в іншому місці (так звана оцінка по методу В). Можливо також об'єднання цих підходів.

Але звернемо увагу, що розрахунок невизначеності за [12] також неявно містить деякі характеристики гауссового розподілу, зокрема аналіз «викидів» та коефіцієнти охоплення. Отже, залишається актуальним запитання, чи не можна запропонувати інші способи оцінки невизначеності результату (тобто області значень U , яка описує результат), які, по-перше, були б вільні від використання *припущень* про вигляд (гауссовість) функції розподілу, і, по-друге, використовували для таких оцінок лише наявні експериментальні результати.

1.3. Принципи комп'ютерної статистики

У попередньому розділі було розглянуто різні аспекти статистичного аналізу даних, з яких випливає необхідність перегляду існуючих підходів до такого аналізу. Сформулюємо їх стисло як систему положень, які повинні враховуватися при будь-яких підходах до статистичної обробки даних.

1. Не існує такого ключового для математичної статистики поняття, як наближене середнім *точно* значення, а існують лише межі, в яких знаходиться результат вимірювання. (Дійсно, якщо подивитися неупередженим поглядом, то хто і коли бачив точне значення? На цих позиціях стоїть і інтервальний аналіз [30, 31].)

2. Аналіз даних зводиться до порівняння двох вибірок. Сюди для повноти картини ми віднесемо і порівняння з модельними даними для задач визначення деяких характеристик однієї вибірки, наприклад її функції розподілу (перевірка гіпотез). Навіть тоді, коли оцінюються параметри конкретної вибірки, це завжди робиться для (можливо, майбутнього) порівняння з іншими даними.

3. Коректний аналіз невизначеності результату є необхідним не лише в силу вимог нормативних документів. У рамках нових підходів без цього просто немає сенсу результат роботи по п. 1. Більше того, у рамках задачі порівняння даних неявно приховано, що їхні межі невизначеності перекриваються. (Без усякого статистичного аналізу результатів вимірювань ясно, що розміри автомобіля і мурашки достовірно різні.) Тому необхідно ці межі оцінювати надійно.

4. Існує проблема коректного статистичного аналізу даних, що не розподілені за законом Гаусса: як впливає з попереднього розгляду, у більшості випадків потрібні для цього формули чи таблиці відсутні.

Нижче буде розглянуто підхід (комп'ютерна статистика), який відповідає поставленим у пп. 1 – 4 вимогам. При цьому мається на увазі, що метод порівняння повинен бути вільним від припущень про функцію розподілу, а покладатися лише на таблиці експериментальних даних.

В основі такого підходу лежать методи Монте-Карло. Ми будемо вважати, що принципи комп'ютерної статистики, які буде викладено нижче, достатньо обґрунтовані, у тому числі на контрольних задачах показано, що її результати повністю збігаються [20, 32, 35, 36, 37] з результатами теорії ймовірностей і математичної статистики. Отже, результати, одержані шляхом моделювання методами Монте-Карло, тобто шляхом усереднення й порівняння результатів по багатьох змодельованих вибірках, на сьогоднішній день не тільки не вимагають підтвердження шляхом порівняння з формулами теорії ймовірностей, а, як стверджують деякі дослідники [32], мають значно більшу доказову силу. Хоча б тому, що результати, одержані в рамках теорії ймовірностей, так чи інакше базуються на висновках, отриманих при переході до нескінченності, на використанні теорем великих чисел тощо, що рідко зустрічається в реальній практиці, у всякому випадку в розглянутій у даній роботі області практичних задач. Так, у статистичній обробці даних польових радіаційних вимірювань існує проблема великого розкиду результатів, причиною якого є нерівномірність забруднення. Такі дані аж ніяк не мають підстав описуватися нормальним розподілом [33]. Більше того, у багатьох практичних випадках генеральна вибірка має у принципі обмежене число елементів. Такі проблеми не вдається розв'язати стандартними методами за допомогою пакетів прикладних програм типу «Статистика». Точніше, навряд чи можна обґрунтувати правомірність застосування цих програм до таких задач. Крім того, не слід забувати, що існують достатньо обґрунтовані підстави вважати, що реальні дані є гіпервипадковими [3], тобто з часом їхні характеристики змінюються. Це означає, що спроби отримати в експерименті нескінченну генеральну сукупність наштовхуються не тільки на необхідність нескінченно тривалих вимірювань, а й на теоретичну заборону отримати такі дані, оскільки протягом вимірювань їхні характеристики зміняться (дані перестають бути незалежними й однаково розподіленими, див. також додаток А).

Отже, розглянемо прийоми так званої комп'ютерної статистики для розв'язання наших задач. Якщо не зважати на різновиди і модифікації, то грубо методи комп'ютерної статистики можна звести до кількох прийомів. Перший – це так званий бутстрап (bootstrap). Це метод, який дозволяє «розмножувати» вибірки. Вважаючи наявну з малим числом елементів вибірку правильною, тобто такою, яка відображає властивості генеральної сукупності, методом бутстрапу можна згенерувати додатково потрібне число вибірок з такими ж статистичними властивостями, що дає можливість зробити для неї надійний

статистичний аналіз невизначеності результату оцінки характеристик вибірки. Другий – метод перестановок (permutation test). У цьому методі так чи інакше елементи вибірок переставляються місцями, що дає змогу виявити найбільш імовірні комбінації між ними і, відповідно, малоймовірні, тобто зробити висновок про ймовірність наявності у вибірках того чи іншого ефекту, наприклад кореляції. Третій метод базується на аналізі змінності параметрів вибірки, якщо з неї видалити декілька випадково вибраних елементів. У вітчизняній літературі він хоча і широко відомий [34], але відповідного підходящого терміну не має. В англомовній літературі він має назву jackknife method (у дивному перекладі метод «складного ножа»).

Слід ще раз підкреслити, що всі ці методи дають результати, які повністю збігаються з результатами теорії ймовірностей і математичної статистики у випадку контрольних задач, де використання формул математичної статистики є теоретично обґрунтованим, а також використовують, наприклад, такі її результати, як генерація комп'ютером випадкових модельних чисел.

1.3.1. Метод перестановок (resampling, permutation test). Аналіз ступеня відмінності параметрів двох вибірок (їхніх середніх, дисперсій чи інших статистик)

Виходячи з проведеного в попередньому розділі обговорення, ще раз підкреслимо, що ніякий поглиблений статистичний аналіз не веде до підвищення точності результату в тому сенсі, що ми будемо точніше знати середнє значення. Уся статистична обробка має за мету покращити чи зробити більш коректно, з більшою достовірністю, оцінку невизначеності результату, тобто, за вказаним вище принципом, точніше встановити межі, у яких обґрунтовано знаходиться результат.

Нагадаємо, що для аналізу ми маємо лише дві вибірки, тобто дві таблиці чисел (одна з них може бути модельною). І можемо для них порахувати середні значення і дисперсії (незалежно від того, які статистичні властивості мають такі набори чисел [12]). Більше ніякої інформації про ці дані (читай, гіпотез про функції розподілу і т. ін.) немає, а отже, немає ніяких формул для статистичних розрахунків. Ставиться задача зробити висновок про те, чи відрізняються в них середні значення і дисперсії (середньоквадратичні відхилення). Звичайно, можна розглядати й будь-які інші статистики (характеристики) цих вибірок.

На рис. 1.1 показано зліва дві модельні вихідні вибірки (№ 1 і № 2).

Рисунок демонструє, що будь-який обмін елементами між вибірками (перестановки їх місцями) веде до зміни величини оцінюваного параметра (середнього або дисперсії). Елементи вибірок вибрано так, щоб їхні середні значення були без сумніву різними. Величини елементів цих вибірок вважаються випадковими, просто для наочності подальшого аналізу їх було упорядковано по зростанню. Уявимо, що ці два набори чисел є моделлю

	№ 1	№ 2		№ 3	№ 4
	1	11		1	11
	2	12		12	2
	3	13		3	13
	4	14		4	14
	5	15		5	15
	6	16		6	16
	7	17		17	7
	8	18		8	18
	9	19		9	19
	10	20		10	20
Середнє					
=	5,5	15,5		7,5	13,5
σ					
=	3,03	3,03		4,74	5,60
Різниця					
середніх R_0 =	10,0		R_1 =	6,0	
Різниця					
для σ =	0,0			0,9	

Рис. 1.1. Залежність різниць середніх і середньоквадратичних відхилень для двох вибірок залежно від перестановки місцями їхніх елементів.

результатів вимірювань – випадковими вибірками, і при подальших вимірюваннях вони можуть змінюватися, а отже зміняться їхні середні значення.

Оскільки ставиться запитання, чи відрізняються між собою їхні середні значення, то шуканим параметром у нашому аналізі візьмемо різницю між середніми значеннями цих вибірок

З рис. 1.1 видно, що різниця між середніми для вихідних вибірок дорівнює 10. Тепер дещо модифікуємо наші початкові вибірки, помінявши місцями їхні елементи у другому і сьомому рядах. Тепер у нас утворилися дві вибірки (№ 3 і № 4), області значень яких перекриваються (рис. 1.2). Відповідно їхні характеристики теж змінилися.

На основі цього розгляду можна зробити такі висновки. Перший висновок: якщо в нас є дві вибірки із своїми середніми і дисперсіями, то для них можна порахувати різницю цих величин, а обмін елементами між цими вибірками веде до зміни величин цієї різниці. Другий висновок: якщо вибірки гарантовано різні за своїми параметрами (середніми і дисперсіями), то обмін елементами між ними веде *лише* до зменшення різниць між шуканими параметрами. Це елементарні, але важливі для подальшого аналізу результати.

Чи можна у цій ситуації говорити, що вибірки № 3 і № 4 є різними? Для відповіді на це запитання спочатку потрібно визначитися з критеріями, за якими оцінюються – чи різними є вибірки, чи однаковими.

Відповідь на це запитання також виявляється далеко не очевидною. Виявляється, що для цього немає строгих математичних критеріїв.

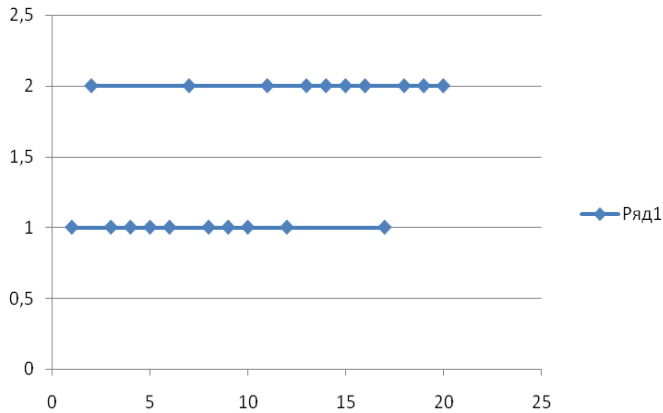


Рис. 1.2. Дві числові осі, на які нанесено діапазони значень вибірок № 3 і № 4.

По-перше, можна нагадати, що статистичний аналіз у дійсності не дає відповіді на такі запитання. Усе, що ми одержуємо в результаті статистичного аналізу, це підтвердження або не підтвердження із **заданим рівнем достовірності** тієї гіпотези, яка перевірялася, тобто в нашому випадку, чи відрізняються вибірки за тим чи іншим вибраним параметром: дисперсією, середнім чи будь-яким іншим моментом функції розподілу. Це ситуація, яку в [17] назвали позалогічним (тобто не основаним на деякому формальному критерії) рішенням: експериментатор завжди приймає рішення про достовірність результату, виходячи з деяких своїх внутрішніх уподобань (варто нагадати відоме серед експериментаторів правило, що експериментальна крива повинна бути одержана мінімум тричі, щоб результат можна було вважати скільки-небудь надійним).

При такому підході до аналізу того, чи відрізняються між собою вибірки (результати вимірювань), тобто коли усвідомлено, що строгих формальних критеріїв для цього не існує, можна поставити таке запитання. Оскільки результат (на основі аналізу випадкових вибірок) отримується лише статистично, то наскільки він є випадковим, оскільки ми маємо справу з випадковими числами? І як при збільшенні чи зменшенні числа вимірювань (числа елементів вибірок) цей результат, тобто висновок, зміниться? У методах комп'ютерної статистики відповідь на це запитання, тобто статистична основа для прийняття рішення (на основі моделювання методами Монте-Карло це буде продемонстровано нижче на конкретних прикладах) виглядає приблизно так: *у результаті виконаного статистичного аналізу показано, що конкретна реалізація результатів вимірювань (вибірка чисел) могла реалізуватися випадково, наприклад у 20 випадках із 100 (якщо повторити вимірювання сто разів). Більше ніякої інформації в дійсності статистичний аналіз не дає.* А далі однаково потрібно приймати «позалогічне» рішення про те, чи можна вважати проаналізовані вибірки однаковими, якщо ймовірність того, що вони однакові, дорівнює 80 % чи будь-якій іншій частці від 100. Проілюструємо суть питання і відповідь на нього прикладом: один мій знайомий мас-спектрометрист говорив, що він вважає, що елемент у пробі присутній, якщо

відповідна лінія з'являється приблизно у 5 спектрах з 15 повторних вимірювань. Відповідно хтось інший може мати інший критерій достовірності результату.

Алгоритм методу перестановок

Отже, розглянемо конкретний алгоритм пошуку відмінностей між двома вибірками так званим методом перестановок (permutation test). Задачами такого типу є вимірювання при різних зовнішніх умовах, порівняння характеристик контрольної групи пацієнтів і тих, які приймали ліки, порівняння режимів харчування тварин чи обробки посівів у сільському господарстві тощо.

Переформулюємо постановку задачі так: оскільки елементи вибірок № 1 і № 2 є випадковими числами (на рис. 1.1 вони упорядковані по зростанню лише для наочності), необхідно встановити, чи не могла б оцінена різниця R_0 між їхніми середніми значеннями утворитися випадково, і в інших вимірюваннях буде отримано інший, навіть протилежний по знаку результат. Ясно, що при цьому припускається, що результати цих інших вимірювань належать до тієї ж генеральної сукупності.

В основу методу перестановок покладено встановлений вище факт (див. рис. 1.1 і 1.2), що при перестановці місцями елементів вибірок завжди змінюється різниця між шуканими статистиками. На цих рисунках для прикладу спеціально було змодельовано такі вибірки, щоб усі зміни були наочними. Те, що елементи вибірок для наочності були впорядковані по зростанню, для даного розгляду не має значення.

Далі потрібно більш строго сформулювати гіпотезу, яка перевіряється (тобто чи відрізняються між собою вибірки), тобто **припустимо**, що вибірки № 1 і № 2 породжено одним і тим же процесом, тобто вони є двома незалежними вибірками з однієї і тієї ж генеральної сукупності і за статистичними властивостями вони є однаковими: при відборі з неї 10 елементів перший раз були випадково вибрані числа, що утворили вибірку № 1, а другий раз – вибірку № 2. І хоча вони виглядають різними, але, у принципі, процес відбору є випадковим і вибірки № 1 і № 2 є еквівалентними, тобто відібраними з однієї і тієї ж генеральної сукупності. Будемо перевіряти, як може змінюватися різниця між середніми значеннями вибірок у припущенні випадкового вибору їхніх елементів із генеральної сукупності. Отже, припускаємо, що деякі з елементів, що утворюють вибірку № 1, могли з таким же успіхом потрапити у вибірку № 2 і навпаки. Таку ситуацію ілюструють вибірки № 3 і № 4. Іншими словами, за допомогою методу перестановок перевіряємо так звану нульову гіпотезу, яка полягає у припущенні, що вибірки (їхні середні значення) є статистично однаковими.

Спочатку підрахуємо для кожної з вибірок (№ 1 і № 2) середні значення і різницю цих середніх значень R_0 . Ставиться задача оцінити, наскільки випадковою є величина R_0 , яка буде характеристикою відмінності вибірок.

У вибірках № 3 і № 4 місцями переставлено лише два елементи. Ясно, що вибір цих елементів було спеціально зроблено випадково і для наочності. Процедура ж у загальному виглядає таким чином. Об'єднаємо обидві вибірки в одну і з такої об'єднаної вибірки виберемо випадково по 10 елементів, утворивши таким чином варіант вибірок № 3 і № 4. Далі повторимо цю операцію по перестановці місцями (permutation) елементів вибірок № 1 і № 2 (тобто випадково вибираючи з об'єднаної вибірки дві підвибірки) багато разів, кожен раз підраховуючи різницю середніх R_i . Якщо вихідні вибірки № 1 і № 2 були еквівалентними, то величини R_i повинні приймати випадкові значення, як більші, так і менші від R_0 , яке мали вихідні вибірки № 1 і № 2. При достатньо великій кількості N таких операцій по перестановці вибірок місцями утвориться масив значень $R_i - R_0$, для якого можна побудувати гістограму і таким способом оцінити величину розкиду значень R_i . Результат виконання такої процедури по відношенню до вибірок № 1 і № 2 і для вибірок № 3 і № 4, коли число перестановок дорівнювало 100 000, показано на рис. 1.3, а, б.

У першому випадку, коли вибірки (№ 1 і № 2) не перекриваються, не було жодної комбінації, коли було б отримано $R_i > R_0$, тобто в усіх випадках $(R_i - R_0) < 0$. Для вибірок № 3 і № 4 вже отримано випадки, коли $R_i > R_0$ (див. рис. 1.3, б): на гістограмі з'явилися від'ємні значення. Отже, ми бачимо, що коли вибірки гарантовано різні (коли величини всіх елементів однієї вибірки менші від будь-якого елемента другої вибірки, то завжди $R_i < R_0$, якщо ж області значень вибірок перекриваються (див. рис. 1.2), то з'являється деяке число випадків n , коли $R_i > R_0$. «Позалогічно» [17] робиться висновок, тобто приймається, що відношення $P = n/N$ є мірою відмінностей між вибірками. В англомовній літературі ця величина називається Р-критерій (P-value).

Отже, ми отримали спосіб оцінки достовірності гіпотези про рівність чи нерівність середніх значень результатів двох вимірювань, виходячи лише з наявних результатів вимірювань, без будь-яких попередніх, явних чи неявних, припущень про їхні властивості.

Потрібно лише пам'ятати, що як і будь-який спосіб аналізу він має свої межі застосування. При великих відмінностях між середніми значеннями вибірок метод перестає працювати, тому що при якійсь величині різниці R_0 між середніми n стає рівним нулю і при подальшому збільшенні R_0 P перестає змінюватися, тобто при $n = P = 0$ вони перестають бути чутливими до подальших змін. Якщо звернутися до гістограм на рис. 1.3, а, б, то це означає, що метод перестає працювати, коли при збільшенні різниці між середніми гістограми перестають перекриватися. Що стосується малих відмінностей між вибірками, то там вступає в дію інший ефект. Можна перевірити, генеруючи на комп'ютері серії по дві однакові за параметрами випадкові вибірки, що для них R_i теж будуть статистично розподілені в деякому діапазоні (див. рис. 1.3, в, г). На рис. 1.3, в показано гістограму розкиду середнього для багаторазової генерації вибірок з тими самими середніми і дисперсіями, що й у вибірці № 2.

Ширина цього діапазону буде залежати від характеристик вибірок (функції розподілу, дисперсії і т. ін.) і тому повинна перевірятися для конкрет-

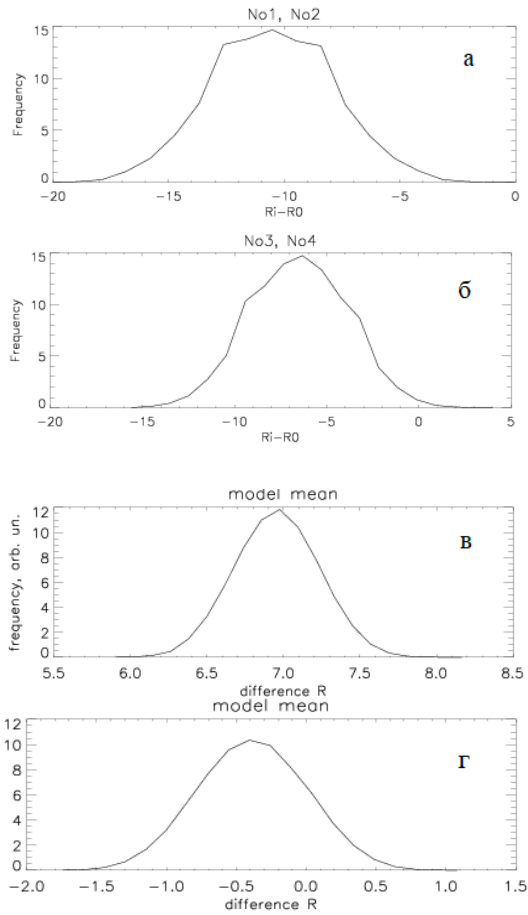


Рис. 1.3. Розкиди середніх значень та різниць між ними:
 а – розкид різниць середніх для вибірок № 1 і № 2; б – розкид різниць середніх для вибірок № 3 і № 4; в – гістограма розкиду значень середнього для багаторазової генерації рівномірно розподілених випадкових чисел з параметрами (середнє та стандартне відхилення) вибірки № 1 з рис. 1.1;
 г – гістограма розкиду значень величини R при порівнянні середнього двох випадкових вибірок з параметрами (середнє та стандартне відхилення) вибірки № 2.

них ситуацій. Якщо результат, отриманий для конкретної експериментальної вибірки, буде знаходитися в межах цього діапазону, вибірки повинні вважатися однаковими (говорячи строго, повинна бути прийнята нульова гіпотеза, що вибірки між собою не відрізняються за величиною середнього). На рис. 1.3, г показано результат порівняння різниць середніх двох послідовних випадкових вибірок з тими ж параметрами вибірки № 2. Ці два графіки демонструють діапазон розкиду відповідних параметрів, у межах якого зміни відмінностей мають вважатися недостовірними.

Алгоритм

процедури по встановленню відмінності між двома вибірками за допомогою методів Монте-Карло. Перевіряється, чи однаковими є середні значення вибірок.

Метод перестановок (permutation test)

1. Розраховуємо для вибірок, що підлягають аналізу, їхні середні значення і різницю між цими середніми значеннями R_0 .
2. Висуваємо нульову гіпотезу, що ці дві вибірки, які мають число елементів n_1 та n_2 відповідно, є *статистично однаковими*, тобто є вибірками з однієї і тієї ж генеральної сукупності.
3. Об'єднуємо ці дві вибірки в одну.
4. Переставляємо місцями елементи об'єднаної вибірки і випадковим чином розділяємо її на дві вибірки з числом елементів n_1 і n_2 .
5. Розраховуємо для цих нових вибірок різницю R_i між їхніми середніми.
6. Повторюємо процедуру отримання R_i N разів, отримуючи таким способом масив значень R_i розміром у N елементів.
7. Оцінюємо число випадків n , коли $R_i > R_0$.
8. Величина відношення $P = n/N$ є мірою відмінності між вибірками.
9. Результат такого *статистичного* аналізу формулюється так: щодо вибраної характеристики (середнього в даному випадку) ці вибірки могли утворитися випадково при повторних вимірюваннях у n випадках з N .

Примітки: 1) указану перевірку на відмінність між вибірками може бути виконано за будь-яким параметром (будь-яким моментом) вибірок. І висновок про відмінність може виявитися різним для різних параметрів;

2) прийняття на основі отриманої величини P рішення про те, чи правильною є нульова гіпотеза, тобто чи є різними середні цих двох вибірок, не входить до компетенції математичної статистики;

3) у загальному випадку повинно виконуватися $N > n_1, n_2$ і у випадку великих вибірок (наприклад, довгих рядів вимірювань) метод перестає працювати, оскільки в реальності вимагає занадто багато комп'ютерного часу.

1.3.2. Аналіз кореляцій

Принцип пошуку кореляцій між двома вибірками дуже подібний до описаного вище алгоритму пошуку відмінності між вибірками. Нехай у нас є дві, для наочності упорядковані по зростанню, вибірки № 1 та № 2 (рис. 1.4).

Перемножимо попарно перші елементи цих вибірок, другі і т. д. Результат дано у третій колонці рисунка: сума цих добутків дорівнює 935. Тепер розглянемо вибірки № 3 і № 4. Вибірка № 3 ідентична вибірці № 1, а вибірка № 4 – це вибірка № 2, у якій переставлено місцями елементи № 2 і № 6. Для неї сума попарних добутків менша і дорівнює 919. Легко перевірити на модельних вибірках в Excel, що максимально можлива сума попарних добутків отримується у випадку, коли вибірки повністю скорельовані, тобто їхні елементи упорядковані по зростанню. Тепер розглянемо вибірки № 3 і № 4.

Вони вже не є повністю скорельованими й існує можливість у процесі

№ 1	№ 2	Попарний добуток	№ 3	№ 4	Попарний добуток
1	11	11	1	11	11
2	12	24	2	16	32
3	13	39	3	13	39
4	14	56	4	14	56
5	15	75	5	15	75
6	16	96	6	12	72
7	17	119	7	17	119
8	18	144	8	18	144
9	19	171	9	19	171
10	20	200	10	20	200
Сума добутоків $S_0 =$		935	$S =$		919

Рис. 1.4. Дані, з яких видно, що будь-яка перестановка місцями елементів однієї вибірки веде до зміни величини суми попарних добутоків.

багаторазових перестановок отримати варіант, коли всі елементи вибірки № 4 будуть краще упорядкованими по зростанню, ніж вихідна, у тому числі в результаті перестановок отримати вибірки № 1 і № 2, для яких сума попарних добутоків S_i буде більшою, ніж початкове значення $S_0 = 935$. Число випадків n , коли S_i буде більшою від S_0 є числом випадків, коли $S_i > S_0$ могло утворитися випадково. Отже, для не повністю скорельованих вибірок відношення $P = n/N$ уже не буде апіорі дорівнювати нулю, а буде показувати ступінь відхилення від повної скорельованості.

Цей факт дає змогу побудувати алгоритм оцінки ступеня кореляції між вибірками, також оснований на методі перестановок і побудові гістограм.

Сформулюємо постановку задачі. Нехай вимірювалися дві різні характеристики (X та Y) деякого об'єкта при зміні деякого параметра, наприклад температури. У результаті утворилися дві вибірки (дві таблиці чисел x_i та y_i). При прямій лінійній функціональній залежності $Y = aX$ вибірки будуть повністю скорельованими і при використанні звичайного методу пошуку кореляцій шляхом побудови графіка в координатах (x_i, y_i) буде одержано пряму лінію, тангенс кута нахилу якої дорівнює a . Але якщо кожне з таких вимірювань має випадкову статистичну складову ε , то в результаті отримується $(x_i + \varepsilon_{xi})$ та $(y_i + \varepsilon_{yi})$. У цьому випадку точки будуть розкидані навколо апроксимуючої прямої, а величина відхилень від прямої буде мірою нескорельованості величин x_i та y_i . Проілюструємо це прикладом.

Згенеруємо на комп'ютері дві вибірки випадкових чисел (X та Y) з однаковими параметрами. Тобто між цими двома вибірками ніякої кореляції

не може бути за визначенням. Побудуємо тепер графік, по одній осі якого відкладаємо x_i , а по іншій – y_i . Результат показано на рис. 1.5, а.

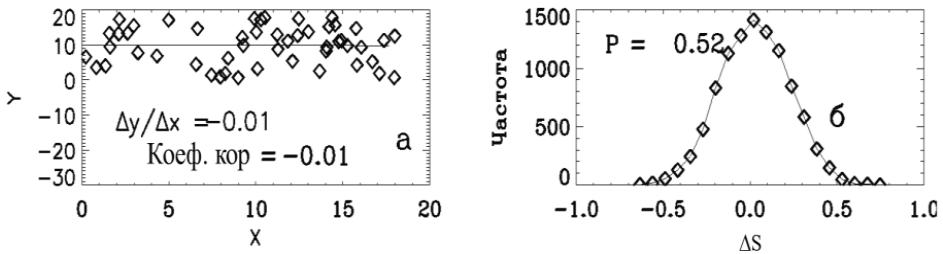


Рис. 1.5. Модельна вибірка для повністю випадкових даних.

Далі використаємо розглянутий вище підхід методу перестановок. Метод докладно описано у [35, 36, 37], і в нашому випадку полягає в наступному. Парно перемножимо елементи $x_i; y_i$ обох вибірок і підсумуємо результати множення $S_0 = \sum x_i y_i$. Суть методу базується на тому, що для впорядкованих за зростанням (повністю скорельованих) вибірок така сума буде найбільшою, а для антискорельованих (коли одна вибірка впорядкована по зростанню, а інша по зменшенню) – найменшою із можливих. Потім випадковим чином переставимо місцями елементи в одній з вибірок, перемножимо їх і знову підрахуємо суму результатів множення. Повторимо цю операцію N разів, кожного разу запам'ятовуючи результат S_k ($k = 1, 2, 3 \dots N$), і побудуємо гістограму розподілу отриманих сум.

Частка (так звана величина P) значень S_k , отриманих у результаті багаторазових перестановок, які будуть більше суми S_0 для початкової комбінації вибірок, буде показувати, наскільки випадковою є кореляція між двома вибірками, що аналізуються. Але для зручності краще будувати гістограму величин $\Delta S = (S_0 - S_k)$, Тоді точка S_0 виявляється зсунутою в нуль і випадковими будуть комбінації елементів, яким відповідає значення $\Delta S < 0$. (Можна використовувати також величину $\Delta S = (S_k - S_0)$; тоді мірою відхилення буде $\Delta S > 0$.) Цей результат показано на рис. 1.5, б. Як бачимо, імовірність отримати початкове значення S_0 , випадково дорівнює приблизно половині випадків (як при підкиданні монети). На графік винесено також результати оцінки скорельованості вибірок шляхом прямого розрахунку коефіцієнта кореляції, а також нахил проведеної через модельні точки прямої лінії. Усі три способи – розрахунок коефіцієнта кореляції Пірсона, нахил прямої та метод перестановок – дають той самий результат: проаналізовані вибірки повністю не скорельовані, що й було закладено в даний приклад.

У першому випадку вибірки X і Y являли собою незалежні набори випадкових чисел. Для таких вибірок коефіцієнт кореляції близький до нуля (див. рис. 1.5, а). Пряма, проведена через точки графіка, по одній осі якого відкладено значення X , а по іншій Y , іноді може показати слабкий нахил як у

бік кореляції, так і антикореляції. Ці відхилення від нуля визначають точність методу в області малих відмінностей між вибірками.

Виконаємо для цієї вибірки операцію підрахунку сум S_k для різних перестановок і від кожної такої суми S_k віднімемо початкове значення $S_k - S_0$, установлюючи таким чином точку відліку в нуль. На отриманій для таких випадкових вибірок гістограмі значень $\Delta S = S_k - S_0$ центр розподілу знаходиться поблизу нуля (див. рис. 1.5, б). Це свідчить про те, що принаймні в половині випадків початкова сума може реалізуватися при випадковій комбінації переставлених місцями елементів вибірок. Частка випадків P , коли випадкова сума виявилася більшою від початкової, винесена на графік. Цей графік показує, що при перестановках більші чи менші за S_0 суми реалізуються практично з рівною ймовірністю (як при підкиданні монети), тобто й кореляції між цими двома вибірками немає.

Розглянемо тепер ситуацію скорельованих вибірок. Створимо з випадкових чисел вибірку X , потім утворимо вибірку $Y = aX + \xi$, де ξ – випадкове число, що імітує розкид даних. Можна бачити, що хоча на графіку рис. 1.6, а точки мають деякий розкид, коефіцієнт кореляції дорівнює 0,91, і на рис. 1.6, б, на відміну від рис. 1.5, б, усі суми значно більші від нуля.

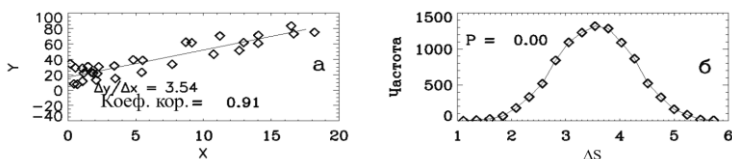


Рис. 1.6. Модельна вибірка з кореляцією даних.

Отже, для даного випадку при випадкових перестановках немає жодної комбінації, яка дала б початковий результат. Це є свідченням того, що вихідні вибірки досить сильно скорельовані.

Отриманий результат є очікуваним, тому що ми вже знаємо, що сума попарних перемножень максимальна для повністю скорельованих вибірок і перестановка елементів місцями тільки погіршує кореляцію, тобто сума стає тільки меншою. Графічно цей результат представлено на рис. 1.6, б, P -критерій $= 0$.

Тепер викладений метод можна проілюструвати на реальних експериментальних даних. Розглянемо метеорологічні дані вимірювань швидкості вітру (метеостанція м. Чорнобиль). Регламентом передбачається вимірювання середньої швидкості вітру за певний проміжок часу і паралельно вимірювання максимального значення швидкості вітру за цей же проміжок часу. Можна було б думати, що між ними існує пряма кореляція, тому що більша максимальна швидкість вітру веде до збільшення середньої швидкості, оскільки максимальне значення також враховується при підрахунку середнього. При стандартному підході кореляцію між цими двома величинами можна отримати з графіка, на якому по одній осі відкладається середня

швидкість вітру, а по іншій – максимальна. На рис. 1.7, а наведено графіки середньої і максимальної швидкостей за деякий відрізок часу.

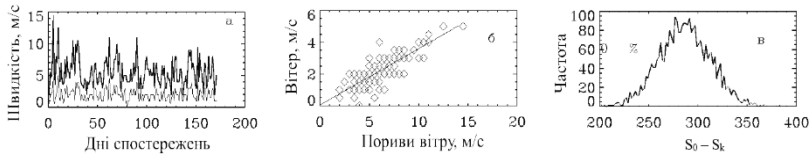


Рис. 1.7. Зіставлення середньої швидкості та максимальних поривів вітру: а – дані метеостанції; б – кореляційні залежності між ними; в – гістограма розподілу величини $S_0 - S_k$.

На рис. 1.7, б дано кореляційний графік взаємозалежності між цими величинами. З цього графіка слід зробити висновок, що хоча між цими величинами і є кореляція, їхній взаємозв'язок не такий уже й однозначний, як очікувалося (точки не лежать на одній прямій). І можна поставити питання про причини таких розбіжностей: чи вони викликані статистичною природою такого явища, як вітер, чи, може, існують інші фактори впливу. Застосуємо тепер до цих даних метод перестановок (див. рис. 1.7, в). Як ми знаємо з рис. 1.5 і 1.6, відсутність кореляції веде до появи від'ємних значень на гістограмах методу перестановок. Цього немає на рис. 1.7, в, отже мусимо зробити висновок, що всі середні швидкості й максимальні значення є скорельованими. (Причини ж існуючого розкиду навколо кореляційної прямої було проаналізовано в роботі [38] уже з використанням інших методів і показано, що відсутність прямого зв'язку між величинами середньої і максимальної швидкості вітру ймовірно викликано тим, що максимальні значення швидкості вітру зумовлено іншими процесами, ніж тими, що забезпечують середню швидкість вітру). Отже, як і для випадку порівняння середніх значень, можна сформулювати оснований на монте-карлівському тесті перестановок алгоритм оцінки скорельованості двох вибірок.

Резюме до використання методу перестановок

У розглянутих вище прикладах число елементів вибірки дорівнювало 10. Оскільки у процесі аналізу переставлялися місцями елементи ряду з 10 елементів, то число можливих (різних) комбінацій дорівнювало числу перестановок (permutations) за формулами комбінаторики $C_{10} = 3\,628\,800$. В описаній вище процедурі на рис. 1.3 число перестановок елементів місцями N було 100 000, що значно менше числа можливих комбінацій C_{10} . Утворення гістограми з максимумом означає, що деякі з комбінацій повторювалися кілька разів, що й дало можливість її побудувати. Отже, з практичної точки зору мінімальне число перестановок (число циклів у програмі) у даному методі повинно бути таким, щоб подальше його збільшення не призводило до зміни форми гістограми. Якщо число елементів вибірки досить велике (довгий часовий ряд вимірювань), то N може сягати десятків тисяч, якщо не мільйонів.

У результаті час роботи програми на побутових сучасних комп'ютерах у деяких задачах автора сягав близько 40 годин. Це свідчить про те, що для аналізу довгих рядів вимірювань треба шукати більш економні алгоритми, ніж описані вище, наприклад застосовувати описану процедуру обробки до частини ряду. Результати застосування описаного методу перестановок (permutation test) до задачі по визначенню ступеня скорельованості двох модельних вибірок повністю збігаються з класичними результатами теорії ймовірностей (див., наприклад, [39], де крім того показано, що він не чутливий до вигляду функції розподілу). Більше того, у ньому використовуються, наприклад, такі результати теорії ймовірностей, як генерація комп'ютером випадкових модельних чисел.

Алгоритм

процедури по встановленню ступеня скорельованості двох вибірок з однаковим числом елементів за допомогою методів Монте-Карло

Оцінка кореляцій за допомогою методу перестановок (permutation test)

1. Маємо дві вибірки, які підлягають аналізу, і позначимо їх як V_1 і V_2 .
2. Розраховуємо для вибірок V_1 і V_2 попарно добутки їхніх елементів і розраховуємо суму цих добутків S_0 .
3. Висуваємо нульову гіпотезу, що ці дві вибірки, які мають число елементів $n_1 = n_2$ відповідно, не є скорельованими, тобто зміни величин елементів однієї вибірки ніяк не пов'язані із змінами величин відповідних елементів іншої вибірки.
4. Переставляємо місцями елементи однієї з вибірок (наприклад, V_2) випадковим чином і позначимо цю вибірку V_{2i} .
5. Попарно перемножуємо елементи вибірок V_1 і V_{2i} і розраховуємо суму S_i для парних перемножень елементів.
6. Повторюємо процедуру отримання S_i N разів, отримуючи таким способом масив значень S_i розміром у N елементів.
7. Оцінюємо число випадків n , коли $S_i > S_0$.
8. Величина відношення $P = n/N$ є мірою нескорельованості двох вибірок. Відповідно величина $1-P$ є мірою скорельованості вибірок.
9. Результат такого статистичного аналізу формулюється так: стосовно ступеня скорельованості ці вибірки (V_1 і V_2) могли б утворитися випадково при повторних N вимірюваннях у n випадках з N .

Примітки: 1) прийняття на основі отриманої величини P висновку про те, чи правильною є нульова гіпотеза, тобто чи є скорельованими ці дві вибірки, не входить до компетенції математичної статистики. Рішення приймає експериментатор, якого задовольняє або не задовольняє величина ймовірності, з якою приймається або відкидається нульова гіпотеза;

2) рекомендована в літературі величина $N = 1000$ в дійсності придатна лише для коротких вибірок (десятки елементів). У загальному випадку повинно виконуватися $N \gg n_1, n_2$ і у випадку довгих рядів вимірювань метод перестановок працювати, оскільки в реальності вимагає занадто багато комп'ютерного часу. У всякому випадку автору доводилося використовувати і $N > 1000\ 000$.

1.3.3. Джекнайф (jackknife) метод

Цей метод є більш відомим, тому що він вимагає значно меншого числа розрахунків і використовувався ще в докомп'ютерні часи, хоча в україно-російськомовній літературі і не має власної назви [див., наприклад, 34]. Ідея методу дуже проста. Нехай маємо результати вимірювань, наприклад забруднення деякої території. Тобто по площі цієї території було взято деяку кількість проб N , у результаті чого утворилася вибірка чисел $X = (x_1, x_2, \dots, x_i, \dots, x_N)$. Потрібним результатом є середнє забруднення (або ж середній вміст корисних елементів у задачах геологічного пошуку) території, тобто середнє значення вибірки X . Для цієї вибірки X можна порахувати також середньоквадратичне відхилення (σ_0). Але, як уже говорилося, ця величина (σ_0) характеризує межі, у яких з певною ймовірністю буде знаходитися результат наступного вимірювання X_{N+1} . Нас же цікавить точність (невизначеність) оцінки середнього \bar{X} .

Для початку припустимо, що N достатньо велике, а число проб на практиці могло б бути і трохи більшим, і трохи меншим. Тоді алгоритм оцінки невизначеності результату (оцінки середнього в даному випадку) буде таким. Видалимо з повної вибірки X деяке x_i , де i – довільний номер елемента вибірки, і підрахуємо середнє для такої вкороченої вибірки. Оскільки вибірка є набором випадкових чисел, то отримане середнє значення, так само, як і середнє значення для повної вибірки \bar{X}_0 , буде випадковим числом. Таким способом, викидаючи багато разів довільно вибране значення x_i , можна отримати відповідне число значень \bar{X}_i . Середньоквадратичне відхилення σ_x для вибірки з елементами \bar{X}_i може бути мірою розкиду цих середніх значень, тобто мірою невизначеності *середнього* забруднення території.

У другій частині книги (розділ 2.7) пропонується розвиток цього методу, який дає змогу визначити мінімальну кількість проб, необхідну для оцінки середнього значення вибірки, при допустимій величині похибки (невизначеності) результату.

1.3.4. Оцінка ймовірності появи певного значення серед елементів вибірки

На відміну від попередніх задач, де ми функцією розподілу не цікавилися, тут розглядаються вибірки з відомою функцією розподілу. Це задача урни з відомим складом кульок, і оцінюється ймовірність витягнути ту чи іншу комбінацію кульок при випадковому відборі з урни певного числа елементів. Такі задачі рідко зустрічаються у практиці експериментальних вимірювань, проте вони є улюбленою задачею математичної статистики [21] і розв'язок їх методом Монте-Карло тут наводиться лише для повноти картини.

Отже, нехай ми маємо урну з білими і чорними кульками. Загальне число кульок N , серед них чорних кульок k . Одне з типових питань: який шанс (яка ймовірність) витягнути чорну кульку, якщо відібрати з урни випадковим

Алгоритм

процедури розрахунку невизначеності оцінки величини деякого параметра вибірки за допомогою методів Монте-Карло

Метод джекнайф (jackknifetest)

1. Маємо вибірку випадкових чисел (результатів вимірювань) $X = (x_1, x_2, \dots, x_i, \dots, x_N)$ розміром у N елементів, для якої оцінюються її параметри (статистики) θ (середнє, дисперсія, асиметрія тощо). Розраховуємо початкове значення потрібної статистики θ_0 . Ставиться задача визначити довірчий інтервал для θ_0 .

2. Видаляємо з вибірки довільно вибраний i -й елемент і для цієї вкороченої вибірки розраховуємо θ_i .

3. Повторюємо процедуру, видаляючи кожен раз інший елемент з вихідної вибірки. При великому числі повторів, що є необхідним для побудови гістограми, ширина якої буде мірою розкиду θ , номер елемента вибирається випадково і ці номери можуть повторюватися.

4. У результаті отримуємо уже вибірку можливих значень потрібної статистики θ_i .

5. Розкид цих значень $\Delta\theta$ і є тим інтервалом, у якому можуть знаходитися значення θ_i при наступних вимірюваннях. Для зручності будується гістограма значень θ_i .

6. Результат такого *статистичного* аналізу формулюється так: оскільки параметр θ_0 є випадковою величиною, при наступних (повторних) вимірюваннях можна очікувати, що він з вибраним ступенем довіри буде приймати вибране значення з діапазону $\Delta\theta$.

Примітки: а) зауважимо, що величина довірчого інтервалу $\Delta\theta$ немає нічого спільного з величиною середньоквадратичного відхилення σ_0 для вихідної вибірки X ; б) видаляти можна і не по одному елементу, а по два, три і т. д.

способом m кульок? У класичному варіанті методом Монте-Карло задача розв'язується таким способом.

Формуємо масив з N елементів, серед яких є задане число білих, чорних, чи інших кульок. Вважаємо, що кожна кулька має свій номер, а кожному номеру відповідає свій колір. Випадковим способом з існуючих N елементів вибираємо m елементів і фіксуємо, чи є серед них хоча б одна чорна кулька. При випадковому розташуванні кульок для цього можна використати такий алгоритм. Оскільки кульки розташовані випадково, то можна вибрати перші m номерів $(0, \dots, m)$. Далі «перемішаємо» вихідний масив, тобто переставимо місцями випадковим способом усі N елементів і виберемо перші m номерів. Повторимо цю операцію багато разів Q і зафіксуємо число n випадків, коли серед вибраних m елементів трапилася хоча б одна чорна кулька. Відношення $P = n/Q$ буде шуканою ймовірністю отримати потрібний результат: при однократному відборі m кульок оцінити шанс знайти серед них хоча б одну чорну.

Звернемо увагу на те, що: 1) це так звана задача з поверненням: вибрані кульки повертаються назад і наступний відбір m кульок знову проводиться з незмінної вихідної вибірки (з незмінними N і k). Однак при використанні методу Монте-Карло немає ніякої проблеми з розрахунками ймовірностей і для задачі без повернення. Просто кожний черговий цикл розрахунків треба виконувати з новими N і n ; 2) при «випадковому виборі» кульок вважалось, що кожна кулька має однакову ймовірність бути вибраною, тобто використовується поняття рівномірного розподілу з теорії ймовірностей. Хоча задачу легко ускладнити, збільшивши число кольорів чи задавши інші спеціальні умови задачі.

1.3.5. Метод бутстрапу (bootstrap)

Розглянутий у цьому розділі метод настільки незвичний, що при першому знайомстві викликають асоціації з бароном Мюнхаузеном, який сам себе за волосся витягнув з болота. Словник поруч з іншими дає такий переклад з англійської: *bootstrap* – *використовувати існуючий варіант системи або процесу для створення нового варіанта*.

Як і інші розглянуті у цій книжці прийоми комп'ютерної статистики, бутстрап призначено для скільки-небудь обґрунтованого встановлення меж надійності результату. Метод бутстрапу полягає в можливості розмножувати вибірку (у кінцевому результаті збільшувати число проаналізованих варіантів), виходячи лише з даних наявної вибірки. На перший погляд це видається неможливим і тому цей прийом асоціюється з дією барона Мюнхаузена.

Ідею підходу можна уявити з наступного. Очевидно, що статистичні властивості достатньо великої вибірки аналогічні властивостям генеральної сукупності, з якої вона була відібрана. Якщо відібрана вибірка достатньо велика, випадкова підвибірка з неї також буде мати властивості генеральної сукупності і, відповідно, властивості своєї попередниці, з якої вона була відібрана. На перший погляд очевидно, що з коротких вибірок уже не можна робити такий відбір. Проте показано, що навіть короткі вибірки несуть інформацію про генеральну сукупність або, що те саме, про функцію розподілу цієї генеральної сукупності. Трансформуючи наявну вибірку описаним нижче методом бутстрапу, можна створити нову вибірку, яка буде мати аналогічні статистичні характеристики. Можна згенерувати велику кількість таких вибірок, розраховувати для них потрібний параметр (середнє, дисперсію тощо) і таким способом оцінити його статистичний розкид. Цей розкид і буде розширеною невизначеністю цього параметра.

Зауважимо, що в описаному вище методі джекнайф робиться ідейно те саме: з наявної вибірки шляхом видалення з неї одного випадково вибраного елемента створюється серія підвибірок, що дає змогу оцінити статистичний розкид її параметрів.

Звичайно, з точки зору здорового глузду, застосування бутстрапу до коротких вибірок видається сумнівним. Математично строге обґрунтування методу статистичного бутстрапу було зроблено в роботі [40], а спрощене роз'яснення можливостей і обмежень можна знайти в доступній статті [20], де наведено також великий список літератури. Зауважимо також, що, як і інші методи комп'ютерної статистики, можливості бутстрапу легко перевіряються без усякого знання теорії. Якщо хтось освоїв ідею методу і створив за вказаним нижче алгоритмом програму, то правильність методу елементарно перевіряється на контрольних прикладах, наприклад на згенерованих комп'ютером вибірках. Більше того, така процедура контрольних перевірок рекомендується як для перевірки створених програм, так і для контролю за отриманими результатами аналізу експериментальних даних. (Нагадаємо, що до отриманих статистичних оцінок не слід ставитися формально, вони є лише підставою для прийняття рішення. А результат статистичного аналізу, відповідно до «Керівництва з невизначеностей...», повинен **"узгоджуватися з усіма спостереженнями і даними, а також із знанням фізичного світу, щоб його з певним ступенем впевненості можна було приписати величині, яка вимірюється."**)

Для демонстраційного аналізу вихідною є вибірка № 2 з рис. 1.1, яка складається з 10 елементів, представлених на рис. 1.8, що демонструє процес формування бутстрапних вибірок та змінність їхнього середнього значення. У першій колонці даються порядкові номери цих елементів. Ставиться таке запитання. Якщо вибірка № 2 утворилася випадково (повторимо, що вона упорядкована по зростанню лише для наочності і для подальшого аналізу порядок розташування її елементів не має значення), то чи не можна оцінити, виходячи лише з даних вибірки, тобто з наявної таблиці з 10 чисел у колонці 2, наскільки великим може бути розкид параметрів вибірки, наприклад її середнього, при гіпотетичних повторних вимірюваннях. Відповідь отримується шляхом застосування процедури бутстрапу. Спочатку генерується 10 випадкових чисел, від одиниці до десяти, по числу елементів вихідної вибірки. У процесі генерування такого ряду числа можуть і будуть повторюватися. Вони наведені в третій колонці рис. 1.8. Ці числа потім співвідносяться з номерами в колонці 1. Відповідні цим номерам числа з другої колонки утворюють потрібну бутстрапну вибірку зі своїм середнім (колонка 4). В останніх двох колонках рис. 1.8 наведено результат повторення процедури формування бутстрапної вибірки. Тепер слід повторити цю процедуру кілька тисяч разів (у [37] рекомендується 1000 повторів, у [20] – 10 000). У дійсності число повторів залежить від розміру вихідної вибірки і легко підбирається експериментально: із збільшенням числа повторів форма гістограми повинна перестати змінюватися. Далі будується гістограма утворених бутстрапних середніх значень. Ширина гістограми дає уявлення про невизначеність (область можливих значень) середнього цих вибірок.

Проілюструємо описаний алгоритм прикладом. Візьмемо для аналізу вибірки № 1, 2, 3, 4 з рис. 1.1. Сформулюємо статистичну задачу наступним

чином. Як уже говорилося, середні значення і стандартне відхилення (дисперсію) можна розраховувати незалежно від будь-якого статистичного аналізу вибірок, і це будуть характеристики даної вибірки. Але, оскільки ці вибірки – результат однократного вимірювання, виникає кілька запитань.

№ п/п	Вибірка № 2	Бутстрапні № п/п	Бутстрап вибірки № 2	Бутстрапні № п/п	Бутстрап вибірки № 2
1	11	1	11	2	12
2	12	1	11	3	13
3	13	1	11	3	13
4	14	4	14	4	14
5	15	4	14	5	15
6	16	5	15	9	19
7	17	7	16	9	19
8	18	8	18	10	20
9	19	9	19	10	20
10	20	10	20	10	20
Середнє	15,5		14,9		16,5

Рис. 1.8. Результат аналізу модельних вибірок із рис. 1.1 методом бутстрапу.

Наприклад, хоча очевидно, що вибірки № 1 і № 2 мають достатньо різні середні значення, поставимо питання: оскільки ці вибірки є (або ми розглядаємо їх як такі) випадковими вибірками з деякої генеральної сукупності (наприклад, нескінченного набору цілих чисел), то наскільки зміняться їхні параметри при наступному відборі і наскільки при цьому зміняться відмінності між їхніми середніми і т. ін. І відповідь на це запитання потрібно отримати з аналізу самих цих вибірок. Бутстрап дає можливість відповісти на такі запитання таким способом. Виходячи з того, що бутстрапні вибірки мають ті ж статистичні властивості, що і вихідна вибірка, то шукані, наприклад, середні значення також будуть належати до тієї ж генеральної сукупності. Відповідно розкид розрахованих методом бутстрапу параметрів буде характеризувати можливий розкид шуканого параметра цієї вихідної вибірки. Результати такого аналізу наведено на рис. 1.9, де показано гістограми розкиду можливих середніх значень при гіпотетичних повторних вимірюваннях.

Як і очікувалося (тому що вихідна ідея бутстрапу полягає в тому, що вихідна вибірка й отримані з неї бутстрапні є однаковими у статистичному сенсі), середні значення вибірок близькі до значень у максимумі гістограм. Гістограми дають уявлення про потенційно можливий розкид середнього. Перший висновок з цього аналізу – гістограми не перекриваються, отже різниця між середніми значеннями вибірок № 1 і № 2 не входить у межі

статистичного розкиду. Тобто їхні середні значення достовірно різні. Друге – гістограми вибірок № 3 і № 4 перекриваються лише на хвостах. Отже, тут уже існує невелика ймовірність (яка приблизно дорівнює відношенню площі самої

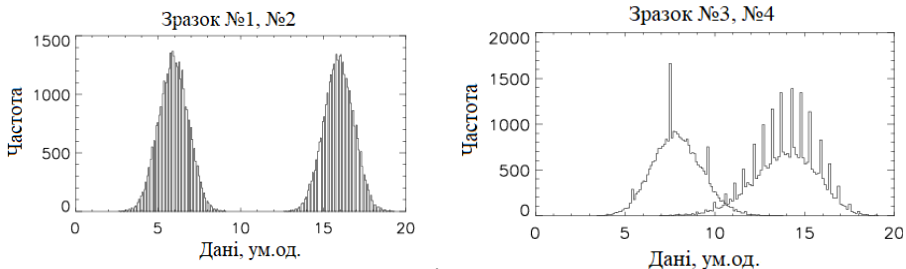


Рис. 1.9. Результат аналізу модельних вибірок з рис. 1.1 методом бутстрапу. Демонстрація змінності середнього значення.

гістограми) отримати однакові значення середніх значень вибірок № 3 і № 4 випадково. Однак слід звернути увагу на зміну форми бутстрапних гістограм. На них з'явилися помітні «викиди», що свідчить про сильну неоднорідність вибірки. Дійсно, нагадаємо, що вибірки № 3 і № 4 були утворені з гарантовано різних вибірок № 1 і № 2 шляхом обміну між ними кількома елементами, які явно виглядають як «викиди» по відношенню до вихідних вибірок № 1 і № 2. Ставлення до цього також залежить від експериментатора: коли при аналізі автором методом бутстрапу результатів одного біологічного експерименту з'явилися такі гістограми, було прийнято рішення про зміну методики виконання експерименту, оскільки такі викиди не мали фізичного обґрунтування.

Резюме для процедури бутстрапу

Ставиться задача по одній наявній вибірці зробити висновок про невизначеність оцінки її параметрів (статистик).

У черговий раз підкреслимо, що її стандартне відхилення на це запитання не відповідає. Воно дає лише межі, у яких очікується величина наступного (гіпотетичного одинадцятого) елемента такої вибірки при продовжуванні вимірювань.

Методом бутстрапу, на основі лише наявних елементів вибірки, яка аналізується, генерується з використанням прийомів Монте-Карло серія вибірок, статистичні властивості яких будуть аналогічними статистичним властивостям вихідної вибірки. Для кожної з цих бутстрапних вибірок розраховується потрібний параметр (яким може бути і стандартне відхилення). Розкид значень цього параметра і є відповіддю на поставлене запитання.

Зауважимо, що в описаних методах комп'ютерної статистики існує також задача ефективних алгоритмів маніпуляцій з вибірками, оскільки методи комп'ютерної статистики вимагають багато комп'ютерного часу [41].

Алгоритм

процедури розрахунку невизначеності оцінки величини деякого параметра вибірки за допомогою методів Монте-Карло

Метод бутстрап (bootstrap method)

1. Маємо вибірку випадкових чисел (результатів вимірювань) $X = (x_1, x_2, \dots, x_i, \dots, x_N)$ розміром у N елементів, для якої оцінюються її параметри (статистики) θ (середнє, дисперсія, асиметрія тощо). Розраховуємо початкове значення потрібної статистики θ_0 . Ставиться задача визначити довірчий інтервал для θ_0 .
2. Зафіксуємо порядок розміщення елементів у вибірці, перенумерувавши їх від 1 до N .
3. Генеруємо випадковий ряд цілих чисел, який має розмір у N елементів, а значення змінюються від 1 до N . Тобто відбираємо N елементів з гіпотетичної генеральної вибірки, яка має нескінченну кількість елементів, що приймають значення від 1 до N в однаковій кількості. При цьому деякі значення (з діапазону $1..N$) можуть зустрічатися кілька разів, а деякі не трапляються ні разу.
4. Формуємо нову (бутстрапну) вибірку, елементи якої дорівнюють елементам вихідної вибірки X відповідно до номерів ряду, згенерованого в попередньому пункті.
5. Для цієї бутстрапної вибірки розраховуємо θ .
6. Повторюємо процедуру формування бутстрапної вибірки по пп. 3, 4, 5. (У деяких керівництвах рекомендується число повторів до 10 000 разів). Для кожної такої бутстрапної вибірки розраховуємо θ_i .
7. У результаті отримуємо уже вибірку можливих значень потрібної статистики θ_i .
8. Розкид цих значень $\Delta\theta$ і є тим інтервалом, в якому можуть знаходитися значення θ при наступних вимірюваннях величини X .
9. Результат такого *статистичного* аналізу формулюється так: оскільки параметр θ_0 є випадковою величиною, при наступних (повторних) вимірюваннях можна очікувати, що він з вибраним ступенем довіри буде приймати значення з діапазону $\Delta\theta$.

Примітка. Зауважимо, що величина довірчого інтервалу $\Delta\theta$ немає нічого спільного з величиною середньоквадратичного відхилення σ_0 для вихідної вибірки X .

Післямова

Ця книжка була задумана як практичний посібник для бажаючих зробити деякі статистичні оцінки і звернути увагу цих бажаючих, що методи «класичної» статистики не завжди є тим, що насправді треба. І автор завжди дивувався, після заглиблення в проблему статистичних розрахунків, чому в монографіях чи підручниках так мало уваги приділяється питанням практичних розрахунків. Тепер же, після завершення цієї роботи, стало

зрозумілим, що в цілому задача аналізу складається з трьох тісно пов'язаних частин. Починаючи цікавитися статистичним аналізом, пересічний дослідник знаходить більш-менш підходящу книжку (наприклад, [1, 8, 10, 21]), де більш-менш описано практичні прийоми статистичних розрахунків, і починає їх робити – це і є перша частина задачі. Однак швидко виявляється, що для виконання справжньої задачі, тобто аналізу даних за допомогою статистичних методів, потрібно все-таки розуміти математичні основи цієї науки.

Але математичні основи теорії ймовірностей виходять далеко за межі знань і потреб початківця, тому автори відповідних керівництв просто змушені логікою викладання матеріалу присвячувати час складним питанням власне теорії ймовірностей. Більше того, через те, що комп'ютерна статистика є маловідомим методом статистичного аналізу, рецензенти статей у наукових журналах вимагають, щоб роз'яснення використаних прийомів було включено в статтю. Отже, та сама логіка викладання матеріалу привела до необхідності написання першої частини цієї книги.

Однак виявляється, що в реальній практиці є ще проблема зрозуміти і сформулювати, у чому власне полягає завдання статистичного аналізу конкретних даних. Тобто що завдання полягає не просто в тому, щоб порахувати середнє, дисперсію чи кореляцію. А в тому, щоб зрозуміти, що в дійсності завжди треба порівнювати зроблені оцінки з іншими: порівнювати дані за різні роки, порівнювати з теорією або навіть щоб хтось інший міг порівняти свої результати з опублікованими. Тобто виявляється, що всі статистичні задачі так чи інакше треба уміти переформулювати в задачу порівняння двох вибірок: або двох експериментальних серій вимірювань, або експериментальної і теоретичної вибірок. Наведені вище алгоритми якраз і маніпулюють з двома вибірками. Останнє неявно реалізується й тоді, коли експериментальні дані аналізуються за допомогою класичної статистики: при цьому йде порівняння з теоретичними припущеннями про характер даних (властивості чисел), що аналізуються. Тобто про закладене в методи розрахунків *припущення* про функцію розподілу.

Комп'ютерна (обчислювальна) статистика намагається ці проблеми обійти і пропонує алгоритми розрахунків, не основані на вказаних підходах чи теоремах теорії ймовірностей. Усі запропоновані й чітко прописані вище у відповідних таблицях-вставках алгоритми зводяться до прямих монте-карлівських розрахунків без необхідності знати для цього функцію розподілу. Але тут постає інша проблема: треба уміти зводити статистичну задачу до порівняння двох вибірок. Тобто не казати, що потрібен розрахунок середнього, дисперсії, кореляції чи інших характеристик набору даних. А формулювати задачу приблизно так: у мене є результати експерименту (набір чисел), з яких я хочу зробити такий чи такий висновок, тобто відповісти на такі чи такі фізичні-практичні запитання. Для цього мені необхідно порівняти ці таблиці чисел. А для порівняння треба порахувати і проаналізувати відмінності між цими двома наборами чисел. Як показує практика, автору ще ні разу не вдалося отримати завдання щось порахувати, висловлене в такому формулюванні.

Це реальна проблема, яка і є третьою частиною сформульованої вище триєдиної задачі: пояснити читачу, як зводити його задачі до проблеми порівняння двох вибірок. На жаль, уже після обговорення написаного виявилось, що в книжці немає простих прозорих прикладів постановки задач для комп'ютерної статистики. Наведені нижче у другій частині книги приклади використання комп'ютерної статистики не охоплюють багатьох типових задач і, скоріше, демонструють приклади її використання для крайніх випадків, коли, на перший погляд, там статистичної задачі немає зовсім. Найбільшого, чого вдалося досягти, на мій погляд, це усвідомлення читачами того, що існує метод, який а) по-перше, відкидає класичні устої типу центральних граничних теорем, що само по собі викликає запитання-сумніви; б) по-друге, дозволяє отримати прозорі результати там, де інші підходи просто не дають відповіді на такі запитання.

Для ілюстрації процесу зведення задачі до порівняння двох вибірок розглянемо наступний приклад (книжка все-таки розрахована на підготовленого читача). Радіоактивні відходи, включаючи відпрацьоване паливо або продукти аварії на ЧАЕС, зберігаються у відповідних контейнерах. Ці контейнери необхідно контролювати на наявність нейтронів вимушеного поділу, щоб запобігти виникненню самопідтримуваної ланцюгової реакції. Радіоактивні відходи містять багато ізотопів, розпад яких і дає основну масу імпульсів у системі контролю нейтронного випромінювання від контейнера. Проте при наявності в контейнері ізотопів, що діляться, час від часу будуть з'являтися також так звані нейтрони вимушеного поділу, коли в процесі поділу ядра одночасно вилітає кілька нейтронів. Зараз такі нейтрони фіксують за допомогою спеціальної техніки – «часових воріт», які фіксують лише нейтрони, що приходять лише в короткий проміжок часу, коли «часові ворота» відкриті. Ставиться запитання, чи не можна виявити такі нейтрони шляхом статистичного аналізу? Відповідь буде «ні», тому що «звичайні» детектори мають таке апаратне обмеження, як наявність мертвого часу, який є набагато більшим від коротких інтервалів часу між появами нейтронів вимушеного поділу і не сприймає короткі інтервали, чому й була створена техніка «часових воріт». Однак володіючи описаними методами комп'ютерної статистики, які позиціонуються як набагато чутливіші, знову поставимо запитання, чи не можна виявити в нейтронних потоках присутність нейтронів вимушеного поділу, використовуючи методи комп'ютерної статистики?

Ідея розрахунків базується на наступному. Існують нейтрони спонтанного поділу, які формують ряд вимірювань і який має свою функцію розподілу. Поява серед них нейтронів вимушеного поділу, які мають свою функцію розподілу, веде не лише до появи коротких інтервалів між імпульсами, а й до спотворення основної функції розподілу. Тому ставиться задача відрізнити функцію розподілу від нейтронів спонтанного поділу від функції розподілу для вимірювань, коли в системі йдуть обидва процеси. Функцію розподілу для нейтронів спонтанного поділу можна взяти або як теоретичну (оскільки фізика процесу вважається відомою), або визначивши її з вимірювань від

лабораторного джерела нейтронів, де процеси вимушеного поділу відсутні за визначенням. Таким чином задача зведена до пошуку відмінності між двома вибірками результатів вимірювань від двох джерел. У такій постановці вона дуже близька до задачі визначення мінімальної активності, яку ще можна помітити, і яка регулюється серією міжнародних стандартів ISO 11929 (див. розділ 2.2 у другій частині книги). Звичайно, метод має свою межу чутливості, яка в основному визначається співвідношенням між числом нейтронів спонтанного й вимушеного поділу: не можна зафіксувати один, два чи три нейтрони на фоні інших. Це статистична задача, отже на властивості вибірок накладаються певні умови.

Без наведення доказів скажемо, що таку задачу (знаходження відмінностей між вибірками для нейтронів лише спонтанного поділу і з додавкою нейтронів вимушеного поділу) було успішно розв'язано для деяких практичних ситуацій.

В інтернеті можна знайти готові пакети програм для обчислювальної статистики. Але користуватися ними читач, який не знає принципів методу і не вміє зводити свою задачу до порівняння двох вибірок, звичайно, не зможе. Можна сподіватися, що викладене у цій частині може бути вступом до сфери статистичного аналізу потоку нейтронів, які контролюються звичайними детекторами.

СПИСОК ЛІТЕРАТУРИ

1. *Закс Л.* Статистическое оценивание. – М.: Статистика, 1976. – 598 с.
2. *Левин С. Ф.* Руководство по выражению неопределённости измерения: проблемы, нереализованные возможности и ревизия. Ч. 2. Вероятностно-статистические проблемы // Измерительная техника. – 2018. – № 4. – С. 7 – 12.
3. *Горбань И. И.* Теория гиперслучайных явлений: физические и математические основы. – Киев: Наукова думка, 2011. – 318 с. – ISBN 978-966-00-1093-2
4. *Тимашев С. Ф.* Фликкер-шумовая спектроскопия. Информация в хаотических сигналах. – М.: Физматлит, 2007. – 248 с.
5. *Shnoll Simon E.* On the Cosmophysical Origin of Random Processes. Open Letter to the Scientific Community on the Basis of Experimental Results Obtained During 1954–2014 // Progress in Physics. – Vol. 10 (2014). Letters to Progress in Physics. Issue 4 (October).
6. *Гольданский В. И., Куценко А. В., Подгорецкий М. И.* Статистика отсчетов при регистрации ядерных частиц. – М.: Гос. изд-во физ.-мат. лит., 1959. – 411 с.
7. *ЩигOLEV Б. М.* Математическая обработка наблюдений. – М.: Наука, 1969. – 344 с.
8. *Смирнов Н. В., Дунин-Барковский И. В.* Курс теории вероятностей и математической статистики для технических приложений. – М.: Наука, 1969. – 363 с.
9. *Худсон Д.* Статистика для физиков. – М.: Мир, 1970. – 296 с.
10. *Поллард Дж.* Справочник по вычислительным методам статистики – М.: Финансы и статистика, 1982. – 344 с.

11. *ДСТУ ГОСТ 8.207:2008*. Государственная система обеспечения единства измерений. Прямые измерения с многократными наблюдениями. Методы обработки результатов наблюдений. Основные положения.
12. *Guide to expression of uncertainty in measurements / GUM:1993*, First edition, 1993 ISO. – Geneva, 1995.
13. *Шишкин И. Ф.* Метрология, стандартизация и управление качеством: Учебное пособие для вузов. – М.: Изд-во стандартов, 1990. – 342 с.
14. *Орлов А. И.* Часто ли распределение результатов наблюдений является нормальным? // Заводская лаборатория. – 1991. – Т. 55, № 7. – С. 64 – 66.
15. *Дубровский С. А., Толстова Т. А.* О ненормальности нормального закона распределения (вообще и в приложении к металлургии в частности) // Всерос. науч.-техн. конф. "Современная металлургия начала нового тысячелетия". 29 окт. – 2 нояб. 2001 г. – Липецк, ЛГТУ. Докл. № 12-2.
16. *Новицкий П. В., Зограф И. А.* Оценка погрешностей результатов измерений. – Ленинград: Энергоатомиздат, 1985. – 248 с.
17. *Алимов Ю. И., Кравцов Ю. А.* Является ли вероятность "нормальной" физической величиной? // УФН. – 1992. – №:162 (7). – С. 149–182.
18. *Breiman Leo.* Statistical Modeling: The Two Cultures // *Statistical Science*. – 2001. – Vol. 16, No. 3. – P. 199–231.
19. *McLaughlin S.* Signal processing: the future is non-Gaussian, non-linear and non-stationary // *Workshop on Modern Methods for Data Intensive Modelling of Nonlinear Dynamics. Systems*. 12-13th Sept. 2002, Hamilton Institute, NUI Maynooth.
20. *Hesterberg Tim.* What Teachers Should Know about the Bootstrap: Resampling in the Undergraduate Statistics Curriculum // arXiv:1411.5279v1 [stat. OT] 19 Nov 2014.
21. *Гмурман В. Е.* Руководство к решению задач по теории вероятностей и математической статистике. – М.: Высшая школа, 2000. – 400 с.
22. *Рекомендации по межгосударственной стандартизации РМГ 43 – 2001.* Применение «Руководства по выражению неопределенности измерений». Межгосударственный совет по стандартизации, метрологии и сертификации. – Минск: ИПК "Издательство стандартов", 2003.
23. *ГОСТ 8.009-84.* Нормируемые метрологические характеристики средств измерений.
24. *МИ 1317-86* Результаты и характеристики погрешности измерений. Формы представления. Способы использования при испытаниях образцов продукции и контроле их параметров.
25. *ДСТУ ISO/IEC 17025-2001.* Загальні вимоги до компетентності випробувальних та калібрувальних лабораторій.
26. *Kirkup L.* A guide to GUM // *Eur. J. Phys.* – 2002. – Vol. 23. – P. 483 – 487.
27. *Захаров І.* Взаємне перерахування похибок та невизначеностей вимірювань // Стандартизація, сертифікація, якість. – 2005. – № 5. – С. 49 – 56.
28. *Захаров І.* Оценивание неопределенности измерений при проведении калибровок // Метрологія та прилади. – 2007. – № 1. – С. 31 – 42.
29. *Чепела В., Никитин Д.* Методи расчета неопределенности измерений в метрологии // Метрологія та прилади. – 2008. – № 4. – С. 51 – 54.
30. *Шокин Ю. И.* Интервальный анализ. – Новосибирск: Наука, 1981. – 112 с.
31. *Ferson S., Ginzburg L., Krenovich V. et al.* Exact bounds of finite populations of interval data // *Reliable computing*. – 2005. – Vol. 11, No. 3. – P. 207 – 233.

32. *Tanizaki H.* Power comparison of nonparametric tests: Small sample properties from Monte Carlo experiments // *Journal of Applied Statistics.* – 1997. – Vol. 24, No. 5. – P. 603 – 632.
33. *Матерон Ж.* Основы прикладной геостатистики. – Москва: Мир, 1968 г., серия "Науки о земле", т. 12. – 408 с.
34. *Прерис А. М.* Общая методика подсчета запасов месторождений твердых полезных ископаемых. – К.: Головное изд-во издательского объединения "Вища школа", 1978. – 106 с.
35. *Hesterberg T., Moore D. S., Monaghan S. et al.* Bootstrap methods and permutation tests. – N.Y.: W. H. Freeman. – 2005. – P. 14.2 – 14.69.
36. *Moore D. S., McCabe G. P., Craig B. A.* Introduction to the Practice of Statistics (W. H. Freeman, 2014) [ISBN: 1464158932]
37. *Simon J. L.* Resampling: The new statistics. Duxbury Press, 1993.
38. *Скорбун А. Д., Огородников Б. И.* Анализ нестабильности ветра по результатам измерений на метеостанции Чернобыль в 2000 – 2010 гг. // *Метеорология и гидрология.* – 2014. – № 5. – С. 54 – 64.
39. *Sakaori F.* Permutation test for equality of correlation coefficients in two populations // *Communications in Statistics - Simulation and Computation.* – 2002. – Vol. 31. – Issue 4. – P. 641 – 651.
40. *Efron Bradley, Tibshirani Robert J.* An Introduction to the Bootstrap, 1993, Springer-Science+Business Media, B.V. – 452 p.
41. *Christina Chatzipantsiou, Marios Dimitriadis, Manos Papadakis and Michail Tsagris.* Extremely efficient permutation and bootstrap hypothesis test using R. 2018. [https://arxiv.org/ftp/arxiv/papers/1806/1806.10947.pdfs](https://arxiv.org/ftp/arxiv/papers/1806/1806.10947.pdf)

2. ПРИКЛАДИ ЗАСТОСУВАННЯ КОМП'ЮТЕРНОЇ СТАТИСТИКИ ДО АНАЛІЗУ РЕАЛЬНИХ ЗАДАЧ

У другій частині книги розглядаються приклади застосування методів Монте-Карло до аналізу невизначеності результатів. Розгляд ґрунтується на опублікованих роботах, метою яких були різні не пов'язані між собою дослідження, проте в кожній з них виникало питання про достовірність отриманих результатів (статистичного) аналізу даних. Для відповіді на такі запитання не досить оцінки дисперсії результатів, а часто таку оцінку й неможливо зробити. Часто самі дані, що аналізуються, вимагають не кількісної оцінки у вигляді відхилення від деякого середнього значення $Q \pm \Delta Q$, а якісної відповіді на запитання, чи відрізняються між собою дані, що аналізуються. Скоріше за все це тому, що самі дані є не стаціонарними вибірками, а змінними в часі рядами.

Хочеться звернути увагу читача на те, що в усіх наведених нижче прикладах ставиться не питання про оцінку похибки результату, а саме про оцінку невизначеності результату. Нагадаємо, що відповідно до аналізу "Керівництва з невизначеностей" (див. першу частину книги) *"Невизначеність результатів таким чином є відображенням того факту, що для даної величини, яка вимірюється, і для даного результату вимірювання немає єдиного значення, а є нескінченне число значень, розсіяних навколо результату, який узгоджується з усіма спостереженнями і даними, а також із знанням фізичного світу, і який з певним ступенем впевненості може бути приписаний величині, яка вимірюється."* Отже, у проаналізованих нижче випадках, у всякому разі на цих, перших, етапах досліджень, шукається якісна відповідь на запитання: у яких межах скільки-небудь надійно знаходиться результат. Оскільки в усіх випадках аналізується шумовий сигнал, в основі використаного підходу покладено аксіоматичне припущення, що можна визначити деякі характеристики дійсно випадкових шумів і знайти відхилення від таких в експериментальному сигналі. Практично в усіх випадках для аналізу невизначеності результатів використовується метод перестановок з комп'ютерної статистики (permutation test).

В аналізі шумів реакторного обладнання аналізується шумоподібний сигнал регулярних вимірювань температури охолоджувача першого контуру. Показано, що поняття середнього та середньоквадратичного відхилення не може бути характеристикою такого змінного в часі сигналу. Для його аналізу використано методи теорії гіпервипадковості.

Аналіз невизначеності оцінок активності в контейнері з РАВ базується на переборі всіх можливих варіантів розрахунків.

Аналіз розкиду результатів при аналізі систем з невідомою чи занадто складною функціональною залежністю результату від факторів зовнішнього впливу також базується на переборі всіх можливих варіантів розрахунків.

Для пошуку слабких сигналів з малим відношенням сигнал/шум (у гамма-спектрах чи дифрактограмах) використовується метод пошуку кореляцій з тестовим сигналом, також на основі комп'ютерної статистики.

Широко відомий на практиці метод оцінки розкиду даних шляхом видалення з вибірки кількох випадково вибраних елементів (в англійській літературі *jackknife*) розвинуто до оцінки необхідного числа вимірювань (елементів вибірки) для досягнення заданої невизначеності.

Перебирання варіантів видалення елементів теж виконується методом Монте-Карло.

2.1. Оцінка надійності статистичного аналізу шумів реакторного обладнання на основі теорії гіпервипадкових явищ

У даному розділі проаналізовано статистичні властивості довгих часових рядів даних, які отримуються у процесі безперервного контролю за станом обладнання атомної станції. Основний інтерес в даному випадку становить не власне використаний метод аналізу і не виявлені особливості роботи реактора як такі, а запропонований [1] спосіб оцінки надійності виконаного статистичного аналізу, оснований на використанні методів Монте-Карло.

2.1.1. Вступ

Одним із завдань безпечної експлуатації реакторної установки є надійний контроль за значеннями параметрів технологічного процесу. Допустимі чи контрольні рівні параметрів встановлюються проектом реакторної установки, технологічним регламентом безпечної експлуатації, інструкціями з експлуатації та ін. У процесі експлуатації фактичні показання засобів вимірювання контролюються операторами. На сьогоднішній день більшість параметрів контролюється технічними засобами автоматизованої системи внутрішньореакторного контролю. За їх допомогою значення контрольованих параметрів вимірюються з певною частотою, проходять обробку, відображаються на моніторах, використовуються для обчислень, зберігаються в архівах та ін. Це дає можливість їх розглядати як матеріал для більш глибокого аналізу з використанням математичної статистики і подальшого практичного застосування результатів для підвищення надійності контролю за станом технологічного процесу, станом обладнання і, як додаткову інформацію, для прийняття рішень. Усе це безпосередньо впливає на підвищення безпечної та надійної експлуатації ядерної установки.

Застосування методів статистики призводить до існування певних невизначеностей результатів аналізу, пов'язаних з тим, що статистичний аналіз завжди базується на припущеннях про функцію розподілу (зауважимо, що деякі методи статистичного аналізу настільки широко розповсюджені і стали стандартом, що про закладені в них припущення давно забули). Як буде показано в даній роботі, застосування сучасних статистичних підходів дає змогу підвищити надійність такого аналізу.

Застосування комп'ютерних методів контролю роботи обладнання дозволяє також уже практично розглядати й інший аспект задачі. Перевищення

встановлених контрольних рівнів є сигналом аварійної ситуації. Проте в багатьох випадках контрольні рівні значно перевищують нормальний робочий рівень. І контроль за зміною нормального рівня роботи (що можна робити, аналізуючи постійний потік даних від контрольної апаратури) дозволяє виявляти небезпечні тенденції у зміні ситуації "на далеких підступах" до контрольного аварійного рівня [2]. Наприклад, такі підходи розглядаються і вже практично використовуються в системах регулювання на основі методів нечіткої логіки [див., наприклад, 3], а не статистики.

Статистичне оцінювання даних, фактично, є результатом намагання спрогнозувати розвиток ситуації. Стандартна процедура виглядає так: на основі попередніх вимірювань чи спеціальних досліджень визначається значення контрольованого параметра, яке характеризує нормальний стан роботи обладнання. Далі визначаються межі, у яких зафіксовано флуктуації цього параметра (наприклад, середнє значення та його середньоквадратичне відхилення). На основі цих даних, на смак дослідника, вибирається математична модель об'єкта – його функція розподілу. А далі на основі аналізу функції розподілу розраховується ймовірність отримати в експерименті те чи інше значення параметра.

Велика подібність у багатьох випадках експериментальних кривих до гауссового розподілу (дзвоноподібні гістограми), математична розвинутість методу та традиція, уведена в нормативні документи, призвели до того, що в переважній більшості випадків (включаючи вимоги нормативних документів [див., наприклад, 4] у якості математичної моделі випадкових сигналів вибирається гауссовий (нормальний) розподіл. Однак існує багато як теоретичних [5, 6], так і експериментальних [7, 8] посилок для того, щоб вважати, що гауссовий розподіл у багатьох випадках є неадекватною моделлю результатів вимірювань. Усе це веде до появи паліативних підходів, коли, розуміючи неправильність гауссової моделі по суті, дослідники доводять, що сама по собі формула в межах прийнятної точності адекватно описує експеримент [9] або аналізуються цензуровані вибірки.

Альтернативою описаній ситуації є розвиток підходів по обробці експериментальних даних без опори на функції розподілу [10]. Більше того, розроблено теорію аналізу довгих рядів даних, основою якого є повна відмова від ідеї Гаусса, що наявність багатьох випадкових чинників, які впливають на вимірюваний сигнал, призводить до усереднення їхнього сумарного впливу. Це так звана теорія гіпервипадкових явищ [11], у якій висувається ідея (аксіома) про те, що, якщо в природі й існує рафінований випадковий процес, який описується конкретною функцією розподілу, то завжди при достатньо довгих вимірюваннях знайдуться непередбачувані зовнішні впливи, які при реальних вимірюваннях цю функцію розподілу спотворюють, а вибірка втрачає статистичну стійкість. До таких систем з очевидністю належать, наприклад, ядерний реактор, у якому на спонтанні процеси поділу накладаються ланцюжки вимушеного поділу, або така складна динамічна

система, як реакторна установка в цілому з її позитивними і негативними зворотними зв'язками.

У даній роботі довгі часові ряди вимірювань деяких параметрів обладнання атомної станції аналізуються саме з використанням методів (і, відповідно, формул) теорії гіпервипадкових явищ [11]. При цьому запропоновано деякі нові прийоми для оцінки надійності результатів.

2.1.2. Опис експериментальних даних

Для аналізу використовувалася 10-хвилинна вибірка даних вимірювання температур теплоносія в гарячій та холодній частинах першої циркуляційної петлі реакторної установки. Періодичність/інтервал вимірювання становив приблизно 0,5 с. Для обробки були використані дані таких точок вимірювання (параметри):

YA11T27 – температура теплоносія на виході з реактора першої петлі головного циркуляційного контуру (засіб вимірювання – термопара);

YA12T01 – температура теплоносія на вході в реактор першої петлі головного циркуляційного контуру (засіб вимірювання – термопара);

YA11T31 – температура теплоносія на виході з реактора першої петлі головного циркуляційного контуру (засіб вимірювання – термометр опору);

YA12T32 – температура теплоносія на вході в реактор першої петлі головного циркуляційного контуру (засіб вимірювання – термометр опору).

Уявлення про характер даних, що аналізуються, дає рис. 2.1.1, на якому представлено гістограми реальних даних і модельних.

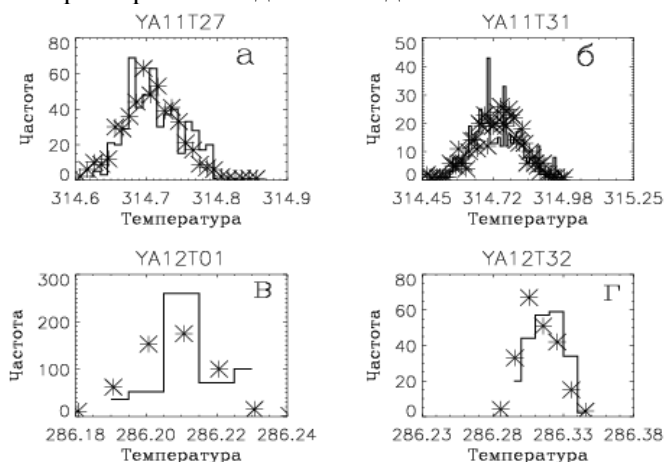


Рис. 2.1.1. Гістограми для експериментальних даних (суцільні лінії) і для змодельованих нормальних вибірок з тими ж параметрами відповідно (зірочки).

Модельні дані – це згенеровані комп'ютером нормально розподілені числа з такими ж, як у відповідних їм реальних даних, числом елементів вибірки, середнім і дисперсією. Можна бачити, по-перше, що для реальних експериментальних даних гістограми значно більш зубчаті, а, по-друге, для

нижніх двох вибірок експериментальні дані взагалі приймають всього по 5 - 6 значень, що навряд чи дозволяє моделювати їх неперервним розподілом.

2.1.3. Аналіз даних на гіпервипадковість

Методи теорії ймовірностей базуються на центральних граничних теоремах, які стверджують, що при збільшенні числа елементів вибірки дисперсія результату прямує до нуля. На відміну від цього теорія гіпервипадковості, ґрунтуючись на тому, що існує багато експериментальних прикладів того, що похибка при збільшенні числа вимірювань не перестає зменшуватись [11, 12], аналізує саме поведінку дисперсії середнього значення при збільшенні числа елементів вибірки. І якісною ознакою гіпервипадкових властивостей вибірки є зростання або значно уповільнене зменшення дисперсії середнього при збільшенні розмірів вибірки, що трактується як втрата вибіркою статистичної стійкості. А при втраті даними статистичної стійкості їхні статистичні оцінки стають неспроможними (залежними від числа елементів вибірки). Для аналізу спроможності даних, виходячи з ідей гіпервипадковості, у нашій роботі використовувалися такі формули [11]:

$X = x_1, x_2, \dots, x_n$ – часовий ряд результатів вимірювання – вихідні дані;
накопичуване середнє вихідного ряду даних

$$y_N = \frac{1}{n} \sum_{i=1}^n x_i (n = \overline{1, N}); \quad (2.1.1)$$

накопичувана дисперсія вихідного ряду даних

$$D_x = \frac{1}{n-1} \sum_{i=1}^n (x_i - y_N)^2; \quad (2.1.2)$$

середня дисперсія вихідного ряду даних

$$D_y = \frac{1}{N^2} \sum_{i=1}^n D_x; \quad (2.1.3)$$

незміщена дисперсія накопиченого середнього

$$\overline{D}_{Y_N} = \frac{1}{N-1} \sum_{n=1}^N (Y_n - \overline{m}_{Y_N})^2 \quad (2.1.4)$$

накопичене середнє флуктуації середнього

$$\overline{m}_{Y_N} = \frac{1}{N} \sum_{n=1}^N Y_n. \quad (2.1.5)$$

Зміна з часом величини \overline{D}_{Y_N} і є предметом подальшого аналізу.

У теорії вводяться також величини, які є кількісною мірою гіпервипадкових властивостей, а саме статистичної нестійкості ряду даних: коефіцієнт γ_N , який характеризує абсолютний рівень статистичної нестійкості

$$\gamma_N = \frac{M[\overline{D}_{Y_N}]}{ND_{Y_N}}, \quad (2.1.6)$$

де $M[*]$ – оператор математичного очікування;

коефіцієнт μ_N , який пов'язаний з параметром γ_N і характеризує співвідношення між складовими вибірки, які можна статистично прогнозувати (описуються певною функцією розподілу), і ефектами непередбачуваних впливів, які не можна статистично прогнозувати (не описуються певною функцією розподілу):

$$\mu_N = \sqrt{\frac{\gamma_N}{1+\gamma_N}}, \quad (2.1.7)$$

Щоб мати можливість порівнювати різні вибірки між собою, необхідно ввести одиницю статистичної нестійкості.

Для коефіцієнта γ_N в якості одиниці вимірювання статистичної нестійкості пропонується величина γ_{0N} , яка відповідає некорельованому ряду відліків з постійною дисперсією $D_{x_n} = D_x$ і нульовим математичним очікуванням при фіксованому значенні N . Коефіцієнт γ_{0N} розраховується аналітично як

$$\gamma_{0N} = \frac{N+1}{(N-1)N} C_N - \frac{2}{N-1}, \quad (2.1.8)$$

де $C_N = \sum_{n=1}^N \frac{1}{n}$.

Використовуючи одиницю вимірювань γ_{0N} , вводиться відносна величина

$$h_N = \frac{\gamma_N}{\gamma_{0N}}, \quad (2.1.9)$$

це коефіцієнт, який характеризує абсолютний рівень статистичної нестійкості в одиницях γ_{0N} .

Відносний рівень нестійкості μ_N в одиницях статистичної нестійкості можна охарактеризувати як

$$L_N = \frac{\gamma_N - \gamma_{0N}}{\gamma_N} = \frac{h_N - 1}{h_N}, \quad (2.1.10)$$

що являє собою відношення коефіцієнта статистичної нестійкості, розрахованого для непрогнозованої частини розглянутого процесу, до

коефіцієнта статистичної нестійкості всього процесу. Границі вимірювань параметра $h_N - [0, \infty)$, а параметра $L_N - (-\infty, 1]$.

Ці коефіцієнти є безрозмірними величинами.

Хоча описані вище коефіцієнти введені в теорії [11] досить математично строго, однак при погляді на графічні результати аналізу (див. нижче) виникає запитання, чи не можна було б отримати такий же результат для випадкового ряду? Для відповіді на це запитання запропоновано два якісних способи оцінювання відмінності аналізованого ряду від випадкового. У першому способі багато разів генерується комп'ютером ряд випадкових нормально розподілених чисел з тими ж значеннями середнього, дисперсії та кількості елементів, що й в експериментальному. Гауссовий ряд вибрано з тих же причин, що й завжди: див. рис. 2.1.1 видно, що дані мають дзвоноподібний розподіл і в найпростішому варіанті можуть моделюватися функціями, що мають подібний вигляд. Цей ряд пропускається через ті ж програми гіпервипадкової обробки. Таким чином ми можемо оцінити відмінність наших результатів від випадкового ряду з нормальним розподілом. Ця процедура повторюється багато разів, і після нанесення всіх результатів на один графік одержимо зону можливих варіантів для випадкових чисел.

Другий підхід базується на необхідності відмовитись при аналізі від використання даних, породжених невідповідними реальності модельними функціями розподілу, тобто аналізувати лише реальні дані. Для цього використовується така ідея. Якщо, відповідно до гіпотези гіпервипадкового устрою нашого світу [1], за яким реальні явища є сумою багатьох процесів, а, відповідно, результати реальних вимірювань є сумою даних з дуже різними статистичними характеристиками, що веде до появи відхилень від основоположного принципу незалежності між елементами вибірки, то можна думати, що отриманий експериментальний зовні випадковий ряд має деякі внутрішні закономірності. У цьому випадку перестановка місцями елементів вибірки випадковим способом повинна привести до зникнення цих закономірностей, що у свою чергу приведе до зміни величини введених вище коефіцієнтів гіпервипадковості. Обробка вибірки з переставленими місцями елементами використаним у роботі методом аналізу на гіпервипадковість, повторення цієї процедури багато разів і нанесення результатів на один графік дасть, як і в першому підході, зону величин коефіцієнтів гіпервипадковості, яка буде рівнем фону або шуму для даного методу. Зауважимо, що на відміну від першого підходу до визначення фонові чутливості методу на основі аналізу випадкових рядів, у другому підході в результаті багаторазових перестановок може реалізуватись ситуація створення ряду із ще більше упорядкованою структурою. Число таких упорядкованих реалізацій дає ймовірність того, що експериментальний ряд є випадковим.

2.1.4. Результати аналізу

Для ретельної демонстрації проведеного аналізу було використано набір даних вимірювання температури першої петлі головного циркуляційного контуру (YA11T31). Результати представлено на рис. 2.1.2 і 2.1.3.

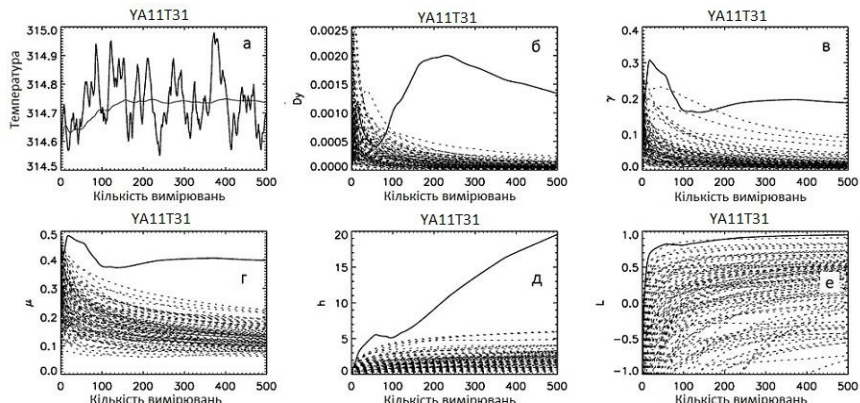


Рис. 2.1.2. Аналіз регулярних вимірювань температури теплоносія на виході з реактора першої головної циркуляційної петлі (YA11T31) на гіпервипадковість. Суцільні лінії – експериментальні дані, пунктир – результати комп'ютерного моделювання: *a* – часовий ряд вимірювань з показаним накопиченим середнім значенням; *б* – дисперсія накопиченого середнього; *в*, *г*, *д*, *е* – зміна у часі параметрів γ , μ , h та L відповідно (пояснення в тексті). Модельні дані – аналіз рядів псевдовипадкових чисел.

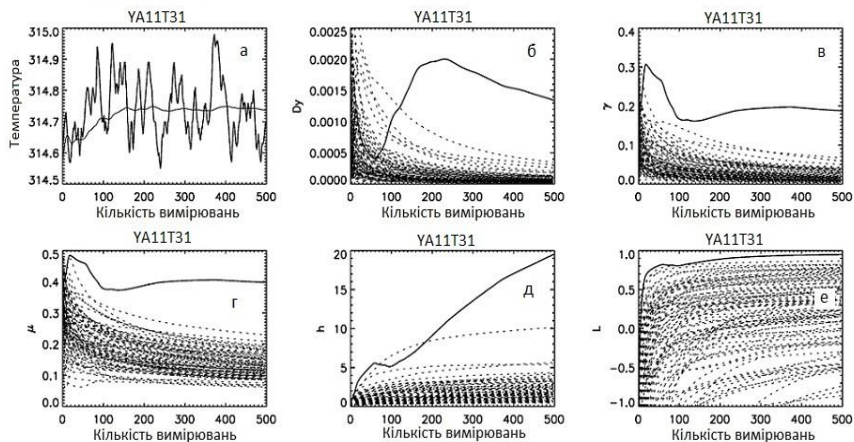


Рис. 2.1.3. Те саме, що й на рис. 2.1.2, але моделювання фонових значень (пунктир) виконано шляхом перемішування експериментальних даних (permutation test).

Суцільні лінії – експериментальні дані, пунктир – описане вище комп'ютерне моделювання. На рис. 2.1.2, а показано дані, що підлягають аналізу: послідовний ряд значень температури – параметра, який

досліджується. Середня лінія – накопичене середнє значення температури. Невеликі флуктуації температури зовні виглядають випадковими.

На рис. 2.1.2, б показано хід дисперсії накопиченого середнього значення температури.

Відповідно до гіпотези, що перевіряється, про гіпервипадковість нашої вибірки, дисперсія середнього не спадає до нуля, як це демонструють модельні дані. Крім того, починаючи з деякого розміру вибірки ($\sim 100 \div 150$ елементів) експериментальні значення лежать помітно вище модельних (на рисунки нанесено близько 100 реалізацій).

Ще більш помітно відмінність експериментальних даних від модельних випадкових проявляється для коефіцієнтів, які характеризують гіпервипадковість вибірки (див. рис. 2.1.2, в – е). Серед них найбільш яскраво демонструє гіпервипадковість вибірки коефіцієнт h_N : його хід із зростанням числа елементів вибірки якісно відрізняється (зростає) від ходу для випадкових чисел. Нагадаємо, що коефіцієнт h_N характеризує абсолютний рівень статистичної нестійкості в одиницях γ_{0N} . З рис. 2.1.2, д очевидно, що статистична стійкість даних з ростом вибірки зменшується.

На рис. 2.1.3 продемонстровано ті ж самі експериментальні дані, але в якості модельних вибірок використано саму експериментальну вибірку з переставленими місцями елементами.

Припускається, що якщо ряд даних, незалежно від їхньої функції розподілу, є випадковим, то перестановка місцями його елементів не змінить його статистичних властивостей. Хоча можна дискутувати з приводу недоліків і переваг кожного з підходів до вибору модельних даних, якісно відповідь на поставлене запитання в обох випадках залишається незмінною: поведінка експериментальних даних істотно відрізняється від поведінки модельних (тобто дійсно випадкових) даних.

2.1.5. Обговорення і висновки

Питання про невизначеність результатів аналізу в даній задачі взагалі звучить не як питання про величину невизначеності, а як питання про відмінність результатів від випадкових. Для відповіді на нього було запропоновано два підходи.

У першому пропонується порівнювати результати аналізу на гіпервипадковість експериментальних даних з точно таким аналізом деяких модельних даних. В якості модельних даних було взято згенеровані на комп'ютері нормально розподілені ряди даних з такою ж дисперсією і середнім, як в експериментальних. Підставою для використання для порівняння нормально розподілених даних є дзвоноподібний вигляд розподілу експериментальних даних.

У другому підході, щоб зовсім не звертатися до використання невідомих функцій розподілу, застосовано метод перестановок (permutation test). Тут використано ідею, що перестановка місцями елементів вибірки не змінить її

статистичні властивості, якщо дані випадкові. Якщо ж у них є внутрішня упорядкованість, вона повинна зникнути при перестановці елементів вибірки місцями (випадково, звичайно).

Існує ряд причин, основною з яких є скінченність числа елементів вибірки, з яких не досягаються асимптотичні значення параметрів модельних вибірок. Це означає, що результати для різних вибірок з однієї і тієї ж генеральної сукупності будуть дещо відрізнятися між собою. Щоб знайти межі таких відмінностей, запропоновано спосіб, оснований на методах Монте-Карло: багаторазова генерація випадкових вибірок з однаковими параметрами і порівняння їх з реальними експериментальними даними. Спосіб є наочним і дає недвозначну відповідь на запитання про можливу випадковість отриманих результатів.

Наведені на рис. 2.1.3 результати моделювання методом перестановок місцями елементів експериментальної вибірки свідчать, що одержати експериментальну криву випадковим шляхом не вдається: зона модельних кривих в усіх випадках лежить нижче експериментальної кривої. Це само по собі є свідченням того, що елементи в експериментальній вибірці розташовані не випадково і є незалежним (на додаток до аналізу на гіпервипадковість) свідченням про те, що вибірка має внутрішню структуру, яка веде до втрати статистичної стійкості.

СПИСОК ЛІТЕРАТУРИ

1. *Скорбун А. Д., Стадник С. М., Котеленець В. Г.* Статистичний аналіз шумів реакторного обладнання на основі теорії гіпервипадкових явищ // Проблеми безпеки атомних електростанцій Чорнобиля. – 2016. – Вип. 26. – С. 36 – 43.
2. *Турбаевский В. В.* Совершенствование системы контрольных уровней радиационных параметров на атомных электростанциях // Ядерная радиационная безопасность. – 2012. – № 1(53). – С. 25 – 29.
3. *Жиравок А. И.* Нечеткие множества и их использование для принятия решений // Соросовский образовательный журнал. – 2001. – Т. 7. – № 2. – С. 109 – 115.
4. *ДСТУ ГОСТ 8.207:2008.* Государственная система обеспечения единства измерений. Прямые измерения с многократными наблюдениями. Методы обработки результатов наблюдений. Основные положения.
5. *Тутубалин В. Н.* Вероятность, компьютеры и обработка результатов эксперимента // Успехи физических наук. – 1993. – Т. 163. – № 7. – С. 93 – 109.
6. *Breiman Leo.* Statistical modeling: the two cultures // Statistical science. – 2001. – Vol. 16. – No 3. – P. 199 – 231.
7. *Новицкий П. В., Зограф И. А.* Оценка погрешностей результатов измерений. – Ленинград: Энергоатомиздат. – 1985. – 248 с.
8. *Дубровский С. А., Толстова Т. А.* О ненормальности нормального закона распределения (вообще и в приложении к металлургии в частности) // Всероссийская научно-техническая конференция "Современная металлургия начала нового тысячелетия. 29 октября – 2 ноября 2001 г. Липецк. Доклад № 12-2.

9. *Котеленец В. Г.* Диагностирование элементов 1-го контура ядерной энергетической установки с реактором типа ВВЭР по теплогидравлическим параметрам: Автореф. дис. канд. т. н. / НПО ЦКТИ, Санкт – Петербург, 1992.
10. *Simon J. L.* Resampling: The new statistics. – Duxbury Press, 1993.
<http://www.resample.com/content/text/index.shtml>
11. *Горбань И. И.* Статистическая неустойчивость физических процессов // Радиозлектроника. – 2011. – Т. 54. – № 9. – С. 40 – 52.
12. *Эльясберг П. Е.* Измерительная информация: сколько ее нужно? Как ее обрабатывать? – М.: Наука. Гл. ред. физ.-мат. лит. – 1983. – 208 с.

2.2. Кореляційний метод виділення слабких сигналів з фону

Крім задач, у яких необхідно безпосередньо розраховувати невизначеність, методи комп'ютерної статистики можна використовувати для задач, у яких отримується якісний результат. Тобто невизначеність так чи інакше оцінюється, проте довести результат до формальної кількісної оцінки стандартної невизначеності – це окрема задача майбутнього.

У даному розділі розглядається застосування методу пошуку кореляцій до задачі пошуку прихованих у шумах ліній як гамма-спектрів, так і спектрів порошкової рентгенівської дифракції. Тут основний результат полягає не стільки в тому, щоб дати кількісну оцінку параметрам спектра, скільки в тому, щоб на фоні шумів знайти лінію взагалі. З точки зору обробки спектра це є еквівалентним задачі визначення мінімальної активності (мінімальний сигнал), присутність якої апаратура може виявити. У реальності це просто перший крок при аналізі спектрів (див. нижче).

Загальноприйнятий метод визначення параметрів лінії у гамма-спектрі, особливо при його автоматизованій обробці, який базується на аналізі похідних від функції форми спектра, не може виділяти лінії, інтенсивність яких знаходиться на рівні фону. Фактично на практиці лінію, яка ідентифікується програмним забезпеченням спектрометра, візуально видно. При зниженні інтенсивності лінії до рівня фону цей метод перестає працювати і лінія не детектується, можливо також хибне детектування відсутніх ліній.

Фактичне прирівнювання нулю концентрації радіонуклідів у зразках природного середовища, коли ясно, що вони в реальності там присутні, але в малих концентраціях, може серйозно спотворити результати при статистичній обробці даних. Див., наприклад, [1, 2], де розглядається питання врахування відсутніх даних при статистичній обробці всього набору даних.

Усе це свідчить про потребу мати метод обробки гамма-спектрів, який дозволяв би виконувати аналіз спектра в усьому діапазоні інтенсивностей ліній, до мінімальної активності, яку ще можна детектувати (МДА), включно.

Виявилось, що розроблений метод має загальне значення з точки зору пошуку сигналів заданої форми на фоні шумів і його можна ефективно використати також у рентгенівському фазовому аналізі [3].

З точки зору поставлених у даній книзі задач оцінки невизначеностей результатів вимірювань розглянута задача пошуку слабких ліній на фоні шумів виявилось можливим переформулювати в задачу пошуку відмінностей між двома вибірками (з наявністю лінії і без) та оцінки достовірності таких відмінностей у випадку, якщо вони будуть виявлені. У результаті показано, що застосування методів комп'ютерної статистики (permutation test) є принципово необхідним для використання методів кореляційного аналізу у даній задачі.

Для розуміння методів розв'язання поставленої задачі вимагаються вельми специфічні знання, тому для спрощення розгляду вкажемо, що задача розв'язувалася за такою схемою. 1. Спочатку показано, що задачу можна звести до статистичного порівняння двох вибірок. 2. Далі, оскільки потрібно знайти

сигнал, якого практично не видно у шумах, обґрунтовується можливість використання методів кореляційного аналізу, користуючись тим, що форма корисного сигналу відома. 3. Для оцінки достовірності, з якою робиться висновок про наявність шуканого корисного сигналу у шумах, використовується описаний у першій частині метод комп'ютерного пошуку кореляцій. 4. Показано, що використання методів комп'ютерної статистики дає можливість, на відміну від використання простої оцінки коефіцієнта кореляції Пірсона, оцінювати також відносну площу ліній (тобто інтенсивність корисного сигналу).

Виділення слабких гамма-спектрометричних сигналів з фону. Опис проблеми

У ряді випадків при дослідженнях радіоактивності навколишнього середовища доводиться мати справу із зразками дуже низької активності. При вимірюваннях методом гамма-спектрометрії із зменшенням активності рано чи пізно постає питання про наявність у спектрі лінії того чи іншого радіонукліда взагалі [4]. Такі випадки зустрічаються, наприклад, при спробі визначити глибину, на яку промігрували радіонукліди з поверхні ґрунту. Ясно, що при цьому виникає задача обґрунтовано заявити, що у взятій з певної глибини пробі вже немає шуканих радіонуклідів.

Принцип методу гамма-спектрометрії полягає в наступному. Детектор з під'єднаною до нього електронікою може реєструвати прихід гамма-кванта на детектор, генеруючи при цьому імпульс, амплітуда якого пропорційна енергії гамма-кванта. Факт появи такого імпульсу і його амплітуда запам'ятовуються. З технічних причин, в основному через обмежену роздільну здатність апаратури, весь діапазон чутливості спектрометра по відношенню до енергії гамма-квантів розділяється на деяке число діапазонів-каналів. Для сцинтиляційних детекторів це, як правило, 1024 канали. Отже, у відповідний канал потрапляють (запам'ятовуються) імпульси, амплітуда яких знаходиться в межах даного каналу (діапазону енергій). Таким чином ясно, що гамма-спектр це гістограма ряду чисел – амплітуд послідовно зареєстрованих імпульсів. А канал відповідає карману гістограми (див. додаток А). Якщо не звертати увагу на тонкі ефекти, то слід вказати, що хоча при розпаді радіонукліда випромінюється гамма-квант строго визначеної енергії система детектування може приписати йому дещо іншу енергію, наприклад через те, що до нього в момент реєстрації може додатися інший гамма-квант від ефектів розсіяння або додатися, або віднятися шумовий сигнал електроніки. У результаті від строго монохроматичного джерела гамма-квантів реєструються гамма-кванти в деякому діапазоні енергій. І замість того, щоб давати сигнал лише в одному каналі, від такого джерела реєструються сигнали у кількох сусідніх каналах, і в гамма-спектрі замість одного вертикального стовпчика гістограми реєструється так звана лінія дзвоноподібної форми, яка ще називається піком повного поглинання (приклад гамма-спектрів див. нижче). Ширина цієї лінії визначає роздільну здатність спектрометра, тобто його

здатність розділяти дві сусідні лінії з близькими енергіями. Вважається, що лінія гамма-спектра має гауссову форму. Отже, у загальному гамма-спектр від кількох радіонуклідів (з гамма-квантами різних енергій) – це багатомодальна гістограма з піками гауссової форми.

Крім ліній від радіонуклідів, у реальному спектрі існує також фон, тобто шумове заповнення каналів гамма-спектрометра. Не вдаючись у природу появи цього явища, укажемо лише, що розподіл числа фонових імпульсів по каналах вважається гауссовим.

Призначення гамма-спектрометра – визначити енергію гамма-квантів від кожного нукліда (за яку приймається енергія в максимумі лінії) та площу цієї лінії, яка пропорційна при інших фіксованих умовах інтенсивності випромінювання від цього джерела. Гамма-спектр можна описати аналітично як суму гауссових функцій плюс випадковий фон і деякий тренд. Програмне забезпечення сучасних гамма-спектрометрів діє за таким алгоритмом: знаходять максимуми цієї функції (нулі перших похідних), тобто положення ліній на шкалі енергій, потім визначають ширину цих ліній і таким чином їхню площу.

Проблеми виникають тоді, коли інтенсивності ліній наближаються до інтенсивності шумів, лінія втрачає гладку форму і визначити похідні стає проблематичним. У даному розділі, на відміну від стандартних аналітичних методів обробки спектра, запропоновано метод виділення слабких ліній на фоні шумів у гамма-спектрах шляхом використання статистичної обробки: оскільки, як було вказано, гамма-спектри – це гістограми для деякого (зазвичай невідомого) ряду чисел, то задачу можна звести до статистичного аналізу цих гістограм як спеціальних вибірок з фіксованим положенням їхніх елементів. Тоді можна з використанням описаних у першій частині прийомів комп'ютерної статистики порівнювати, наприклад, гістограми фону і випромінювання від джерела і знаходити відмінності, тобто появу чи відсутність ліній порівняно з фоном.

Розроблений метод [5] оснований на розрахунку кореляції між контрольним сигналом з відомими параметрами та сигналом, який спостерігається, з використанням методу перестановок.

2.2.1. Постановка розрахункової задачі

Зараз при стандартній обробці гамма-спектрів методом аналізу похідних визначається положення ліній та їхня площа [див., наприклад, 6]. Активність проби пропорційна площі лінії, тобто кількості імпульсів, зафіксованих у каналах, що відповідають даній лінії.

Схему задачі знаходження присутності сигналу у фоні наведено на рис. 2.2.1, де показано рівень фону p_0 і рівні сигналів над ним, які треба знайти. Величині МДА відповідає рівень p_1 , а сигналу, інтенсивність якого достатня, щоб надійно визначити його параметри (мінімальна активність, яку можна визначити – МВА, див. нижче), відповідає рівень p_2 .

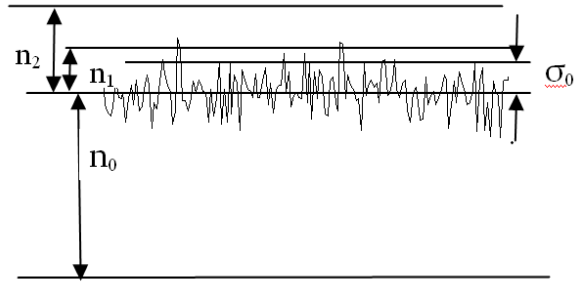


Рис. 2.2.1. Схема задачі розрахунку МДА, МВА.

У загальноприйнятому підході для оцінки МДА пропонується [7 – 10] статистично порівнювати кількість імпульсів, зафіксованих у каналах відповідно для фону і для корисного сигналу + фон. Тобто виконувати статистичний аналіз двох вибірок. Це зовсім інша задача, яка не має нічого спільного з методом похідних і практично використовується лише для якісних оцінок наявності сигналу. При цьому власне активність не визначається. Схема цього підходу показана на рис. 2.2.1. Середнє значення фону дорівнює n_0 . Треба знайти мінімальний рівень сигналу n_1 , який можна було б виділити над фоном. З рисунка ясно видно, що рівень фону n_0 не має ніякого значення, а має значення лише розкид його значень у різних каналах, тобто його дисперсія.

У такому підході задача була розв'язана [7] як стандартна перевірка статистичної гіпотези про знаходження із заданою ймовірністю відмінності параметрів двох вибірок (фон без сигналу і з сигналом) з нормальними розподілами – необхідно визначити кількість імпульсів сигналу над фоном n_0 при заданому рівні достовірності. Величини n_0 , n_1 , n_2 – це середні значення по багатьох серіях вимірювань. При цьому було одержано формули [11]:

$$n_1 = k_{1-\alpha} \sigma_0 \quad \text{та} \quad n_2 = n_1 + k_{\beta} \sigma \quad (2.2.1)$$

відповідно для границі детектування (МДА) та для границі розпізнавання (МВА), де n_1 , n_2 – число відліків у каналах піка повного поглинання, $k_{1-\alpha}$ - $(1-\alpha)$ -й (правосторонній), а k_{β} - β -й (лівосторонній) квантилі стандартного нормального розподілу, σ_0 – середньоквадратичне відхилення для фонового, σ – для повного сигналу. Цей результат означає, що потрібно багато разів провести вимірювання фону і знайти для нього середнє значення та σ_0 . Перевищення середнього значення фону на величину n_1 при вимірюваннях корисного сигналу є рівнем МДА. Аналогічно перевищення фонового сигналу на величину n_2 є рівнем МВА.

Якщо прийняти, що рівню МДА відповідає довірчий інтервал 0,05, а рівню МВА – 0,95, то тоді $k_{1-\alpha} = k_{\beta} = 1,645$ [12].

Для оцінки можливостей спектрометра по вимірюванню слабких сигналів (ISO 11929:2000) введено у практику як стандарт дві характеристики: а) поріг ухвалення рішення DT (decision threshold – аналог МДА) – найнижчий рівень детектування, коли наявність сигналу може бути зафіксована з певною

ймовірністю, скажімо, 5 % (якісне визначення); б) границя детектування DL (detection limit – аналог МВА) – верхній рівень детектування, коли сигнал гарантовано із заданою імовірністю (скажімо, 95 %) може й бути вимірним (кількісне визначення параметрів сигналу).

Проте до цього підходу існують і претензії (див., наприклад, [13]), продовжуються спроби його розвитку [14, 15]. Необхідно підкреслити, що величина n_2 відповідає інтенсивності сигналу, коли його вже (за самим визначенням цієї величини) надійно видно у спектрі і програмне забезпечення спектрометра може визначити його параметри. Що стосується n_1 , то для даного спектрометра вона фактично є абстрактною характеристикою (фону) спектрометра, а не реальним мінімальним рівнем детектування активності. Але головним недоліком підходу, на наш погляд, є те, що в ньому розрахунки МДА виконуються не для конкретних сигналів, а заздалегідь, виходячи з відомого фону установки, а формули записано, виходячи з певних припущень про його функцію розподілу.

Оскільки в нашій проблемі гамма-спектрометрії ми переходимо від обробки спектра до статистичного аналізу числа імпульсів у каналах спектрометра, то можливо розглянути й інші, ніж у роботі [7] статистичні підходи до вирішення цієї задачі. Тому було запропоновано принципово новий підхід, який полягає у використанні кореляційних методів виділення гамма-ліній із шумів у припущенні, що корисний сигнал+фон має додаткову статистичну інформацію порівняно з фоном. Метод, який пропонується для детектування слабких гамма-ліній у шумах базується на розрахунку кореляції між контрольним сигналом з відомими параметрами і спостережним сигналом, з використанням для статистичної обробки методу перестановок (permutation test). Як виявилось, запропонований метод дозволяє надійно детектувати лінію в ситуації, коли стандартним програмним забезпеченням вона ще не визначається. Тобто дозволяє підвищити надійність прийняття рішення про присутність чи відсутність сигналу.

Спосіб виділення корисного сигналу з використанням корелятора широко відомий у радіотехніці та оптиці [16, 17] і базується на пошуку кореляцій між експериментальним сигналом і тестовим сигналом, параметри якого відомі і близькі до очікуваного сигналу. Ідея використання кореляційних методів у гамма-спектрометрії також не нова, у всякому випадку автору відомо про пропозицію його використання, зроблену у звіті [18]. Але по суті кореляційний метод в гамма-спектрометрії сьогодні не використовується, точніше, використовується лише як допоміжний засіб.

Стан використання кореляційних методів на сьогодні ілюструється технічною документацією на гамма-спектрометричну систему контролю РАВ фірми Фаворит-Плюс [19]. Спочатку знаходять (кореляційні) сигнали в каналах, які відповідають шуканим лініям, і таким способом визначають положення і форму ліній. Але потім більш ретельно за рахунок збільшення часу вимірювань у цих каналах намагаються знайти лінії і звичайними методами визначити їхні площі. Зауважимо, що при цьому задача

повноцінного аналізу спектра кореляційним методом навіть не ставиться. Кореляційний метод розглядається лише як спосіб фільтрації шумів і видалення лінійного тренда. Тобто в кінцевому результаті кореляційний метод як такий не використовується. Причина цього ясна: кореляційний метод з використанням звичайного коефіцієнта кореляції не дає можливості оцінювати площі ліній у гамма-спектрі. Тому в реальній практиці гамма-спектрометрії для аналізу ліній використовується метод похідних. Однак розвиток методів комп'ютерної статистики дозволяє знову повернутись до розгляду можливостей використання кореляційних методів у гамма-спектрометрії для розв'язання спеціальної задачі: проведення вимірювань на рівні МДА.

2.2.2. Кореляційний метод обробки гамма-спектрів

Для автоматизованого методу пошуку сигналу в гамма-спектрі необхідно задати критерії пошуку. Загальними є наступні посилки: сигнал (пік, лінія) має гауссову форму, а фон – нормальний розподіл [6, 11, 19, 20].

У нашому підході застосовується розрахунок кореляції між експериментальним і тестовим сигналом і ступінь кореляції додатково аналізується статистичними методами. При розрахунках також виходили з того, що лінія випромінювання має гауссову форму, а ширина її відома і залежить від енергії. Що стосується фону, то в даній роботі для теоретичного аналізу були використані модельні псевдовипадкові нормально розподілені вибірки, хоча в подальшому при розрахунках та аналізі реальних спектрів припущення про форму розподілу фону не використовується.

Тобто в основі методу лежать такі посилки: а) гамма-лінія має гауссову форму; б) відповідно додаткові до фону імпульси корисного сигналу заповнюють канали вибраної лінії не рівномірно – їх більше в каналах навколо положення центроїда; в) розраховується кореляція між лінією в спектрі і тестовою лінією також гауссової форми; г) для аналізу використовується статистичний метод перестановок (*permutation test*, тест перестановок) [21, 22, 23], за допомогою якого розраховується ступінь кореляції між тестовою лінією і лінією, яка аналізується.

2.2.2.1. Тест перестановок

Дамо загальну схему методу (рис. 2.2.2). Спектр розглядаємо як вибірку з m елементів, де m – число каналів. У спектрі виділяється діапазон каналів (вікно) розміром n , ширина якого відповідає ширині лінії, центроїд якої знаходиться посередині вибраного діапазону n в k_0 -му каналі спектра. Кількість імпульсів у кожному каналі y_i , $i = 1, 2, \dots, n$, утворює вибірку для подальшого аналізу.

Далі утворюємо модельну вибірку z_i з n елементів-каналів, які заповнюються так, щоб утворити гауссову лінію – значення z_i максимальні в

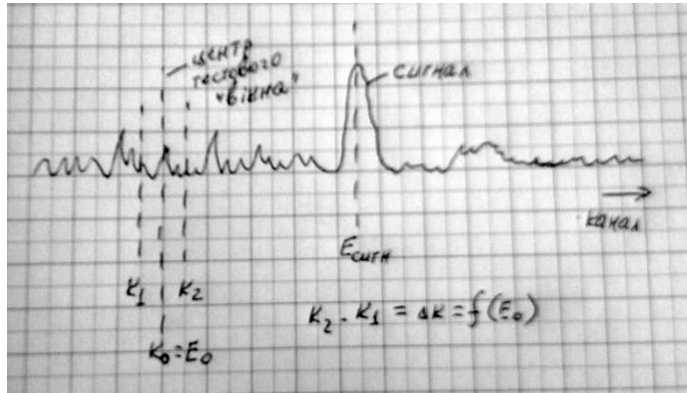


Рис. 2.2.2. Схема сканування вхідного сигналу тестовою лінією (див. також рис. 2.2.20).

середині вибірки і симетрично спадають до країв. Потім для цих двох вибірок розраховуємо кореляцію за допомогою тесту перестановок. Результат розрахунків – гістограма можливих значень коефіцієнта кореляції. За коефіцієнт кореляції c_k приймається значення в максимумі гістограми, яке присвоюється k_0 – му елементу нової вибірки розміром m . Далі вікно пересувається вправо на один канал, процес повторюється і розраховується елемент c_{k+1} . У результаті повного проходу по спектру утворюється нова вибірка з m елементів $c_k, k = 1, 2, \dots, m$, яку ми назвемо кореляційним спектром.

Суть такого кореляційного методу така. Нехай ми маємо послідовність з n каналів аналого-цифрового перетворювача $x_1, x_2, \dots, x_0, \dots, x_n$. Задамо в них спектральну лінію у вигляді числа відліків в i -му каналі

$$y_i = A_0 \exp(-(x_i - x_0)^2 / 2\sigma^2) + \varepsilon_i, \quad (2.2.2)$$

де σ^2 – дисперсія, а ε_i – випадкова компонента. Задамо тестову лінію також у вигляді гауссової функції

$$z_k = A_1 \exp(-(x_k - x_c)^2 / 2\sigma^2), \quad (2.2.3)$$

$k = 1, 2, \dots, m, c = m/2$ – середина вибірки, $m \sim \sigma \ll n$.

Виберемо тепер підряд m елементів з послідовності y_i : $Q = y_j, \dots, y_{j+m}$ і утворимо суму

$$S_{j0} = \sum_{k=j}^{j+m} y_k z_k. \quad (2.2.4)$$

Легко перевірити, що ця величина буде максимальна, коли $x_c = x_0$ (див. опис пошуку кореляцій методом комп'ютерної статистики у першій частині).

Коли величина ε стає порівняною з корисним сигналом, наявність сигналу вже не можна визначити візуально. У цих випадках постає питання про ступінь достовірності прийняття рішення про наявність сигналу. Її можна оцінити на основі застосування статистичної обробки. Якщо при фіксованих z_k випадковим чином переставити місцями елементи вибірки y_k , знову

підрахувати суму S_j і повторити цю процедуру багато разів, то відсоток P значень S_j , які будуть перевищувати S_{j0} , буде ймовірністю того, що початкова сума S_{j0} утворилася за рахунок випадкового співвідношення між y_k та z_k (так званий Р-критерій [22 – 24]). Для зручності можна змістити точку відліку в нуль, утворивши величину $\Delta S = S_{j0} - S_j$. Тоді від'ємні значення ΔS будуть свідчити про можливість виникнення кореляції при випадкових реалізаціях z_k . Типову гістограму розподілу значень ΔS для модельних вибірок, створених за формулами (2.2.2) і (2.2.3), показано на рис. 2.2.3 а.

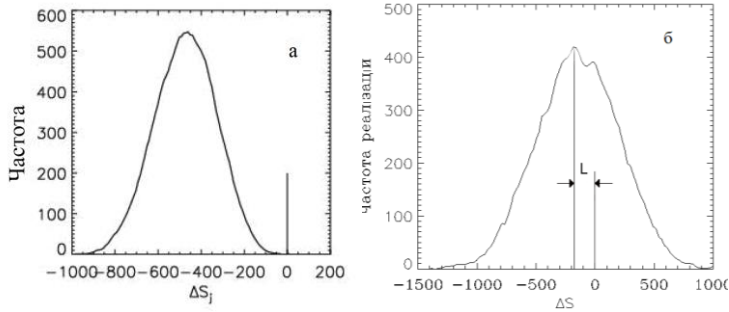


Рис. 2.2.3. Гістограма розподілу відхилень величини S_j від експериментальної величини S_{j0} при багаторазовому перемішуванні однієї з вибірок: а – при повній скорельованості вибірок, коли Р-критерій практично дорівнює нулю; б – при неповній кореляції, коли Р-критерій малий, але величина L відмінна від нуля.

Зазвичай за міру кореляції в тесті перестановок приймають Р-критерій. Але тест перестановок не годиться для оцінки ступеня кореляції сильно скорельованих вибірок [22], оскільки Р стає рівним нулю. Тому в якості параметра оцінки було введено й використано величину $L \equiv \Delta S$ (див. рис. 2.2.3, б) зміщення максимуму гістограми вліво чи вправо від величини $\Delta S = 0$. ΔS є прямою мірою кореляції (див. формулу (2.2.4)) і тому також може слугувати в якості характеристики збігу тестової лінії і експериментального сигналу: чим більше L , тим більшою є кореляція між двома вибірками.

Якщо просканувати спектр y_i тестовою функцією z_k , зміщуючись послідовно на один канал і виконуючи описаний вище тест перестановок, можна замість y_i утворити нові вибірки, ставлячи у відповідність кожному значенню x_i величину L_{ii} . У розглянутому випадку вибірка L буде мати особливість (максимум) у точці x_0 . Аналізуючи такі особливості, можна приймати з визначеною ймовірністю рішення про наявність чи відсутність корисного сигналу.

Утворимо тепер модельний спектр фону шляхом генерування ряду нормально розподілених псевдовипадкових чисел і приписування кожному каналу випадкового числа. Проскануємо його тестовою функцією як описано вище, утворюючи послідовність чисел L_i . Ширина піків у гамма-спектрах залежить від їхньої енергії. Тому при пошуку кореляції ширина тестової

функції повинна дорівнювати ширині піка в спектрі. На рис. 2.2.4 показано одержану експериментально залежність ширини піка від енергії для конкретного сцинтиляційного детектора.

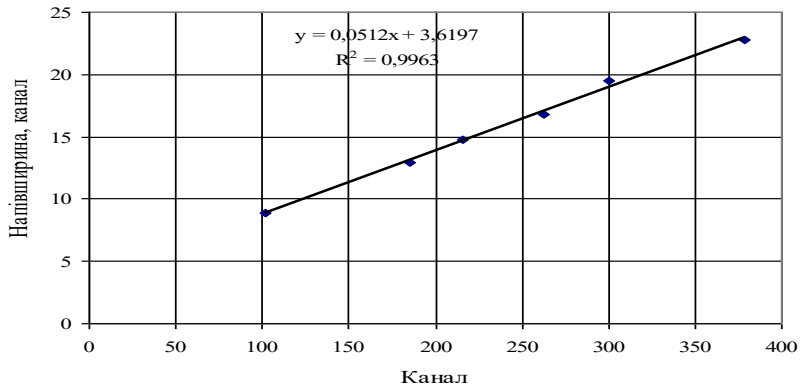


Рис. 2.2.4. Залежність ширини піка від його енергії для піків ^{152}Eu для використаного в роботі сцинтиляційного детектора.

Ця залежність була закладена в програму моделювання: при скануванні спектра тестовою функцією (4.3) її ширина змінювалась за вказаним на рис. 2.2.4 лінійним законом.

На рис. 2.2.5 показано результати розрахунків для модельного фонового сигналу, просканованого тестовою функцією, коли фон має як нормальний, так і рівномірний розподіл. Якщо обрізати від'ємні значення нижньої кривої, які свідчать про наявність антикореляцій і нас не цікавлять, а також згладити її, одержимо результат обробки сигналу, придатний для подальшого аналізу. Відзначимо дещо несподіваний результат: наявність у випадковій послідовності чисел досить помітних кореляцій з тестовою функцією. Це означає, що навіть у чисто фоновому сигналі можуть утворюватися комбінації, які при однократному вимірюванні спектра помилково можна прийняти за наявність сигналу. Діапазон зміни L (у даному випадку приблизно від -15 до +10) треба вважати фоном, перевищення якого при кореляційній обробці спектра буде свідчити про наявність сигналу.

На рис. 2.2.6 показано результат обробки модельного сигналу, коли до випадкового фону додавався корисний сигнал (з центром у 100-му каналі). Сумарний сигнал утворювався таким способом: а) генерується вибірка з певного числа нормально розподілених імпульсів із заданою дисперсією; б) будується її гістограма, тобто модельна корисна лінія; в) ця гістограма додається до фону в каналах заданої лінії. Тобто до імпульсів фону додається певне число імпульсів, але вже у вигляді гауссової лінії. Зауважимо, що ця лінія набуває правильної гауссової форми лише при досить великій кількості імпульсів.

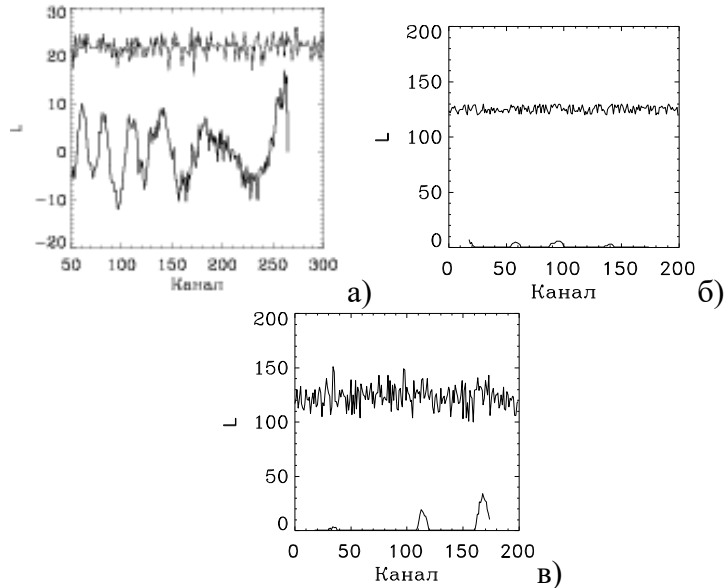


Рис. 2.2.5. Фонівий сигнал і результат його кореляційної обробки. а) фон з гауссовим розподілом (угорі) та його кореляційний сигнал L (внизу). б) фон з рівномірним розподілом; в) фон з гауссовим розподілом після обрізування від'ємних значень L і згладжування. Верхні лінії – сигнал, нижні – показник кореляції $L > 0$.

На рис. 2.2.6 показано результат обробки модельного сигналу, коли до випадкового фону додавався корисний сигнал (з центром у 100 – му каналі). Сумарний сигнал утворювався таким способом: а) генерується вибірка з певного числа нормально розподілених імпульсів із заданою дисперсією; б) будується її гістограма, тобто модельна корисна лінія; в) ця гістограма додається до фону в каналах заданої лінії. Тобто до імпульсів фону додається певне число імпульсів, але вже у вигляді гауссової лінії. Зауважимо, що ця лінія набуває правильної гауссової форми лише при досить великій кількості імпульсів.

На рис. 2.2.6 зліва направо і зверху вниз: фон як набір нормально розподілених випадкових чисел, корисний сигнал у вигляді гауссової лінії, загальний сигнал як сума імпульсів фону і корисного сигналу у відповідних каналах, тестова лінія як гауссова лінія з такою ж дисперсією, як і дисперсія корисного сигналу, результат тесту перестановок і кореляційний спектр як результат обробки кореляційної кривої (заміна від'ємних значень нульовими і згладжування).

У кореляційному спектрі спостерігається п'ять ліній: чотири паразитних і корисний сигнал у 100-му каналі. Рівень паразитних піків є рівнем фону в кореляційному спектрі. Підкреслимо, що такий паразитний пік може з'явитися

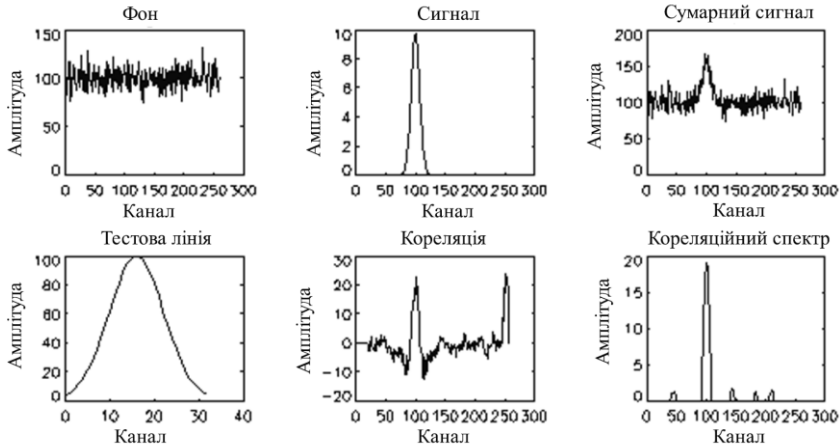


Рис. 2.2.6. Результати кореляційної обробки модельного сигналу з фоном.

в каналах корисного сигналу і додатися до його інтенсивності. Тобто за корисний сигнал у кореляційному спектрі потрібно приймати сигнал, інтенсивність якого перевищує рівень фонових піків.

Одним з підходів до пошуку корисного сигналу можна вважати сигналом те, що перевищує максимально можливий фоновий пік. Визначення цієї величини для нашої моделі було зроблено за допомогою такої процедури.

Генерувалася послідовність випадкових чисел довжиною 100 елементів, для якої виконувалась кореляційна обробка, як на рис. 2.2.6. З одержаної послідовності L_i вибиралося і запам'ятовувалося найбільше значення. Процедура повторювалася багато разів і таким чином утворювалася вибірка максимальних елементів. Далі при виборі рівня фону можна піти двома шляхами. Можна в якості рівня фону вибрати максимальний елемент. Тоді гарантовано все, що вище цього рівня, можна вважати сигналом. Проте такі великі значення фону зустрічаються нечасто, і хоча збільшення рівня фону підвищує надійність виявлення сигналу, проте означає втрату чутливості. Можна також установити рівень фону, виходячи із статистичних міркувань, тобто вибрати певний довірчий інтервал і допустити можливість хибного виявлення сигналу.

Можливо використати також звичайний коефіцієнт кореляції Пірсона q , який дається виразом [24]:

$$q = \frac{\frac{1}{n} \sum_{i=1}^n (z_i - \bar{z})(y_i - \bar{y})}{\sigma_y \sigma_z} \quad (2.2.5)$$

Цей коефіцієнт розраховується паралельно з коефіцієнтом L на кожній перестановці і за величину кореляції було прийнято значення q у максимумі гістограми. У такий спосіб одержується найбільш імовірне значення кореляцій трьома методами в одному й тому ж наборі даних показано на рис. 2.2.7.

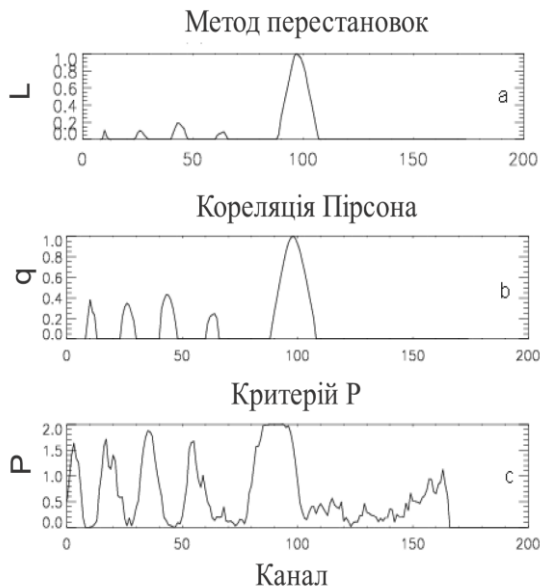


Рис. 2.2.7. Порівняння можливостей різних параметрів кореляції для однієї і тієї ж вибірки.

Зверху вниз: коефіцієнт кореляції, одержаний методом перестановок з критерієм кореляції L , одержаний за допомогою тесту перестановок коефіцієнт кореляції Пірсона q , коефіцієнт кореляції, одержаний методом перестановок з критерієм кореляції P . Із рис. 2.2.7 очевидно, що P -критерій абсолютно не підходить для задачі аналізу спектрів. Як звичайний коефіцієнт кореляції, так і метод на основі комп'ютерної статистики надійно детектують сигнал у 100-му каналі. Проте спостерігаються також паразитні піки, які детектуються навколо реального піка у 100-му каналі, вони є шумами кореляційного методу. Ці паразитні піки майже вдвічі менші у випадку використання тесту перестановок з величиною L в якості параметра кореляції. У свою чергу це означає, що метод перестановок з критерієм L має найвищу чутливість серед цих трьох методів.

Порівняння кореляційного і звичайного статистичного підходів

Для порівняння статистичних оцінок за формулою (2.2.1) з результатами кореляційної обробки було виконано такі розрахунки. Будемо генерувати фоновий сигнал, додавати до нього корисний сигнал контрольованої інтенсивності, виконувати кореляційну обробку і перевіряти, чи перевищує число імпульсів у каналах сигналу заданий фоновий рівень (2.2.1), тобто перевіряти, чи зафіксовано сигнал. Повторимо цю операцію багато разів і підрахуємо число випадків, коли було зафіксовано сигнал. Відношення цієї

величини до загальної кількості повторів дає величину ймовірності зафіксувати сигнал при даному відношенні сигнал/фон. У нашій моделі величини n_1 і n_2 трохи змінюються від серії до серії розрахунків, оскільки вони є середніми значеннями для даної серії, тому ймовірність детектування сама є ймовірнісною величиною.

На рис. 2.2.8 показано ймовірність детектування сигналу для різних співвідношень сигнал/фон, що в даному випадку для зручності порівняння з формулою (2.2.1) означає відношення: (число імпульсів корисного сигналу в повному сигналі)/ σ_0 .

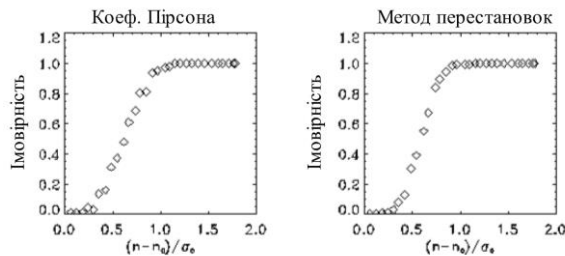


Рис. 2.2.8. Ймовірність виявлення сигналу залежно від співвідношення сигнал/фон. Зліва – коефіцієнт кореляції Пірсона. Справа – коефіцієнт кореляції з методу перестановок.

Для DT відповідно до формули (2.2.1) це число дорівнює 1,645, а для DL удвічі більше, оскільки число імпульсів чистого сигналу в кожному циклі розрахунків було постійним, а $\sigma_0 = \sigma$. З рис. 2.2.8 очевидно, що коли відношення сигнал/фон дорівнює 1,645, ймовірність задетектувати сигнал кореляційним методом значно перевищує рівень DT, рівний 0,05, і досягає 1 (100 %), і тим більше гарантовано досягає 100 % для рівня DL.

Для порівняння розрахунки виконано двома методами. На рис. 2.2.8 зліва показано результат розрахунку коефіцієнта кореляції Пірсона, а справа – коефіцієнт кореляції з тесту перестановок.

І хоча, як було показано вище (див. рис. 2.2.7), коефіцієнт кореляції Пірсона більш чутливий до паразитних кореляцій у фоновому сигналі, у масштабі даних рисунків це мало помітно. Недоліком тесту перестановок у свою чергу можна вважати те, що він вимагає багато комп'ютерного часу.

2.2.2.2. Перевірка отриманих результатів на реальних спектрах

Розроблений метод було перевірено на реальних спектрах. На рис. 2.2.9 – 2.2.11 представлено спектри: фоновий спектр спектрометра, спектр KCl (^{40}K) і джерела ОСГИ ^{152}Eu . У верхній частині рисунків представлено спектр, як його дає програмне забезпечення спектрометра. У тому числі можна бачити виділені, тобто ідентифіковані спектрометром лінії. У нижній частині наведено результат кореляційної обробки цього спектра. Ці рисунки демонст-

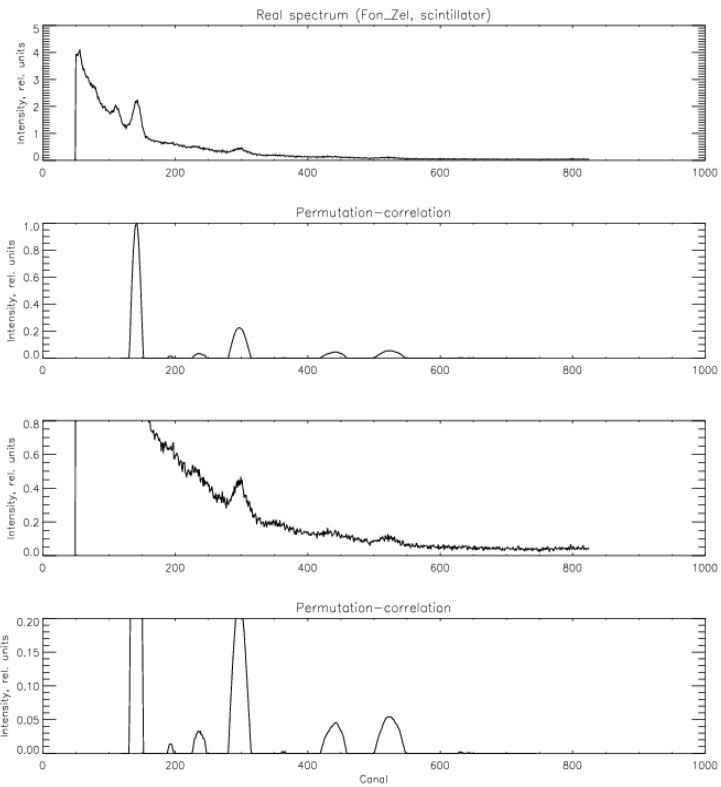
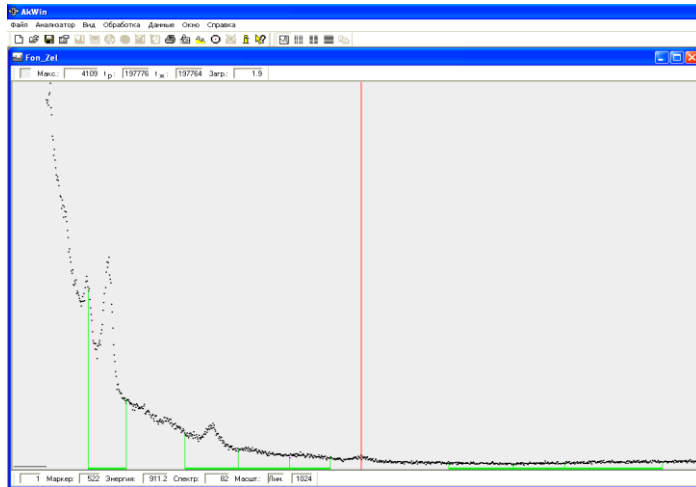
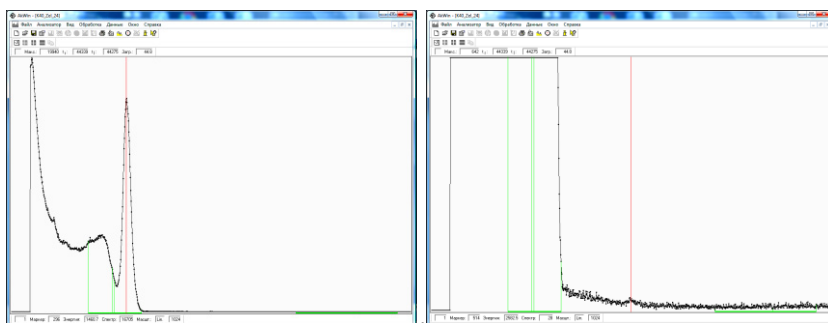
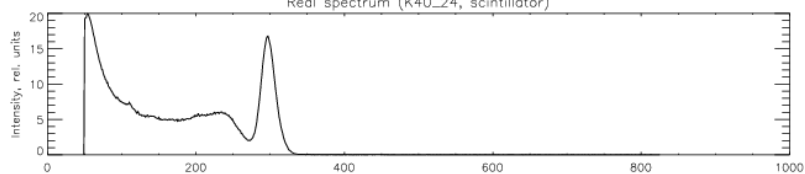


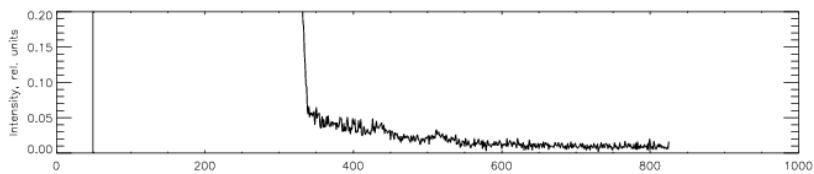
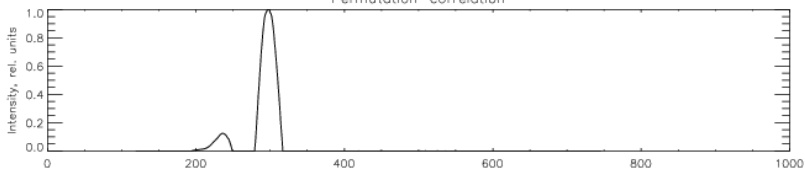
Рис. 2.2.9. Фоновый спектр використаного спектрометра.



Real spectrum (K40_24, scintillator)



Permutation-correlation



Permutation-correlation

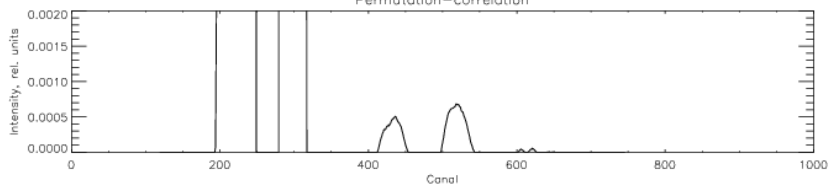


Рис. 2.2.10. Спектр ^{40}K .

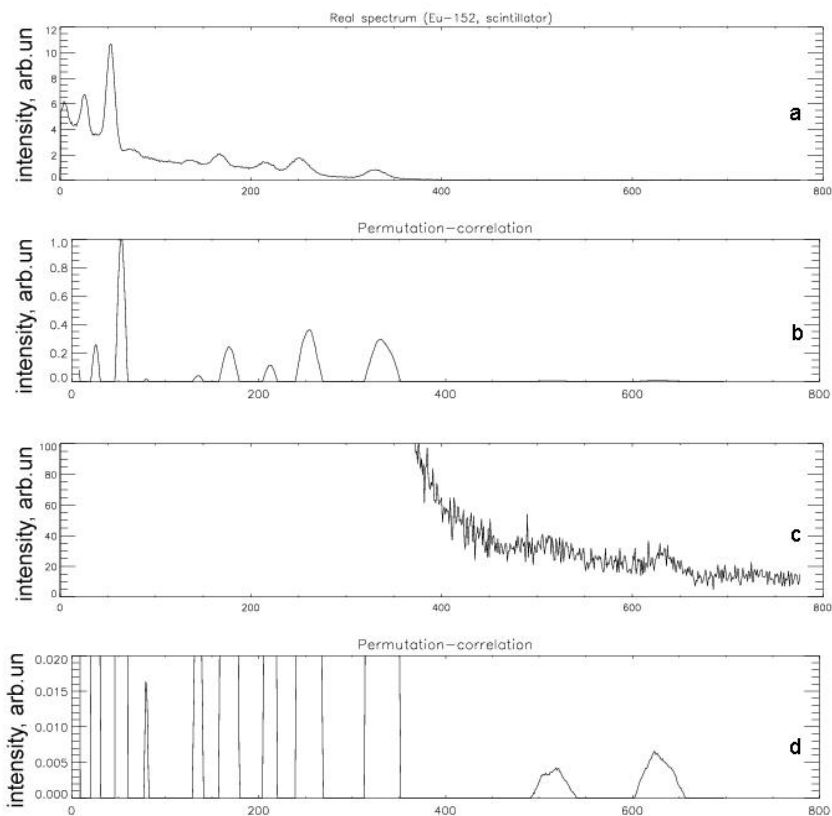


Рис. 2.2.11. Кореляційна обробка гамма-спектра від стандартного джерела ^{152}Eu : a – сцинтиляційний спектр ^{152}Eu ; b – його кореляційний спектр; c, d – ті самі спектри в більшому масштабі.

рують, що всі лінії досліджених радіонуклідів, які видно на реальних спектрах, також присутні на кореляційних спектрах. Рис. 2.2.9 – це фоновий спектр даного спектрометра, з яким будуть порівнюватись наступні спектри. Тривалість набору спектра 197 776 с. На цьому спектрі присутній пік ^{137}Cs у ~ 130 каналі, лінія ^{40}K у каналі 296. Крім того, у фоновому спектрі можна бачити слабкі лінії в 506 і 425 каналах. Ці лінії через їхню малу інтенсивність не ідентифікуються програмним забезпеченням спектрометра, але надійно виявляються на кореляційному спектрі. Ми ідентифікуємо лінію у 506 каналі як ^{208}Tl з енергією 2 615 кеВ.

Рис. 2.2.10 демонструє можливості розробленого методу на прикладі реального спектра джерела ^{40}K . На цьому рисунку і на наведених нижче показано загальний спектр у вигляді фото з екрана у двох масштабах. Один з них – загальний спектр, другий – той же самий спектр у більшому масштабі, щоб продемонструвати останні канали з фоном (слабкі фонові лінії і шум). На

цьому спектрі видно лінію ^{40}K у каналі 296, і фон із слабкими лініями у 506 і 425 каналах. Ці лінії, як і на рис. 2.2.9, не виділяються програмним забезпеченням спектрометра, але їх видно як лінії на кореляційному спектрі.

На рис. 2.2.11, а представлено спектр джерела ОСГИ ^{152}Eu , знятого на спектрометрі СЕГ – 001М виробництва фірми "АтомКомплексПрилад". Це джерело має багато добре видимих стандартних ліній і було використано, щоб продемонструвати, що кореляційний метод також детектує їх. Рис. 2.2.11, б – його кореляційний спектр. Видно, що в кореляційному спектрі всі піки надійно детектуються. Оскільки в даному масштабі рисунків слабку (високоенергетичну) частину спектра видно не дуже добре, на рис. 2.2.11, с, d показано ті самі спектри в більшому масштабі. "Підозрілі" області в районі 500 і 450 каналів стають ясно видимими на кореляційному спектрі.

Аналіз реальних спектрів і обговорення результатів

На рис. 2.2.12 - 2.2.16 показано гамма-спектри проби з природним вмістом радіонуклідів, яка була надана МАГАТЕ для контрольних перевірок.

Проблема визначення присутності чи відсутності сигналу, тобто проблема МДА, так чи інакше зводиться до одного й того ж питання: чи є відмінності між фоновим спектром і спектром проби, чи ні. Як добре відомо (див. наприклад [26]), відповідь на це запитання залежить у тому числі від часу набору спектра. Зауважимо також, що ISO 11929-4:2000 вимагає, щоб "добуток середньої частоти повторення імпульсів на тривалість вимірювання був достатнім, щоб дозволити апроксимацію реального часо-імпульсного розподілу нормальним стандартним розподілом, і що втрати за рахунок "мертвого" часу незначні".

Проведені нами дослідження цього питання якраз і демонструються на наведених нижче спектрах. Одержаний практичний висновок, оснований не на припущеннях про форми розподілу і т. ін., звучить так: час набору спектра повинен бути таким, щоб кількість імпульсів у каналах відображала статистичні властивості генеральної вибірки. Спектри було отримано при тривалості набору 200, 500, 1000, 1500 і 660 000 с. Останній спектр має такий же вигляд, як представлений на рис. 2.2.9 фоновий спектр, тобто його можна вважати базовим, з яким можливо порівнювати інші спектри. Питання, яке постає, – це яка мінімально можлива тривалість набору спектра дасть правильний результат. З рис. 2.2.12 – 2.2.16 випливає, що задовільний результат одержується вже при тривалості набору спектра 1500 с. У той же час програмне забезпечення спектрометра лінії з такою статистикою ідентифікувати ще не може. Порівнюючи спектри із зростаючою тривалістю набору можна бачити, що при поганій статистиці, коли деякі з каналів взагалі не заповнені, кореляційний метод демонструє хибне виявлення неіснуючих ліній. Ці лінії поступово зникають, коли статистика спектра покращується. З чого випливає вказаний вище зрозумілий критерій: число імпульсів у каналах повинно бути статистично достовірним. На практиці набір спектра можна

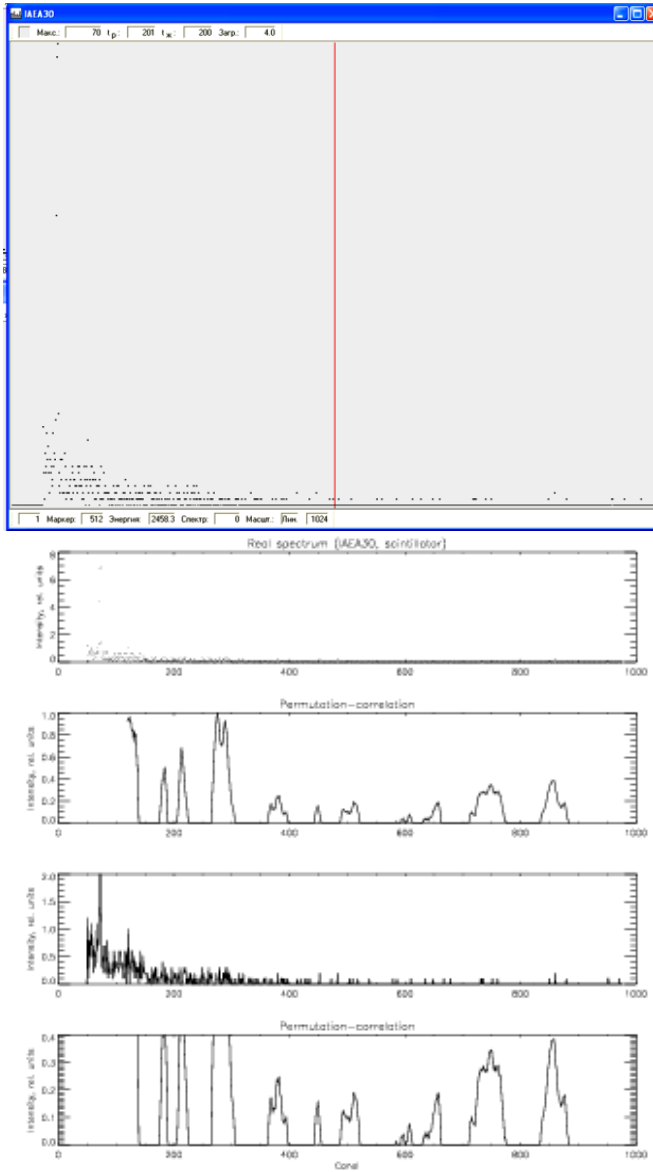


Рис. 2.2.12. Сцинтиляційний спектр зразка МАГАТЕ за час набору 200 с.

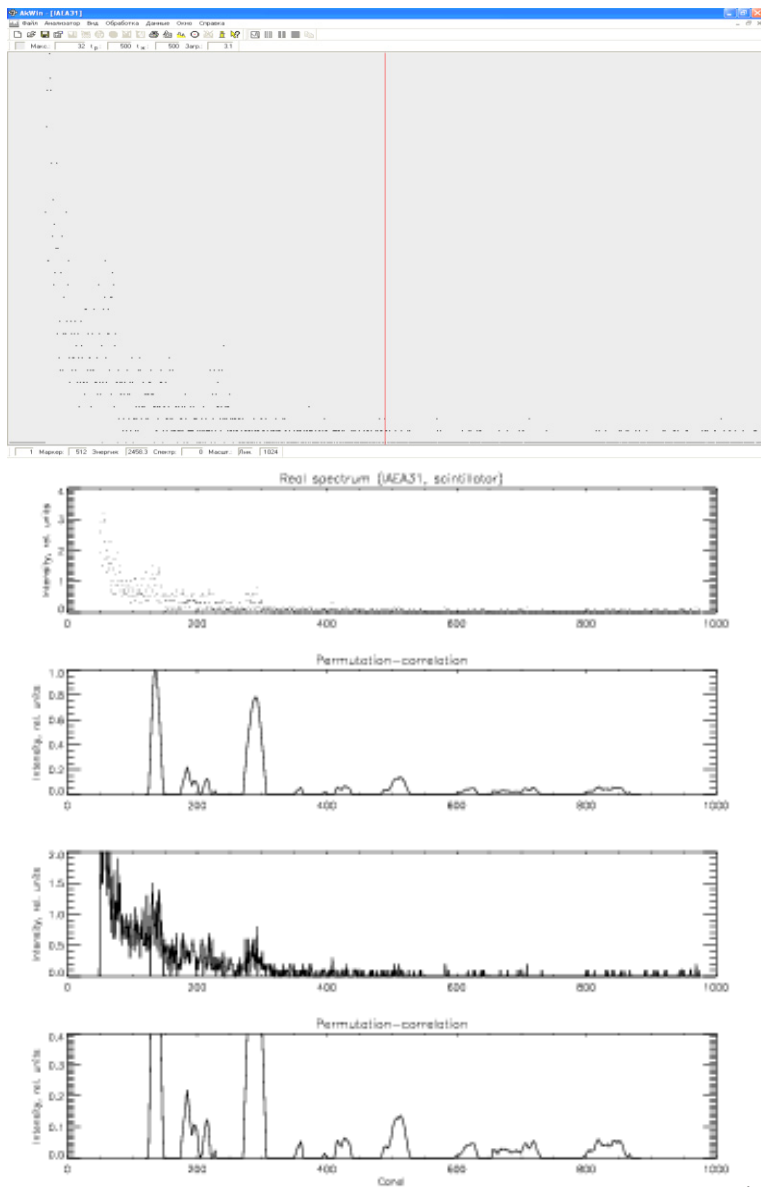


Рис. 2.2.13. Сцинтиляційний спектр зразка МАГАТЕ за час набору 500 с.

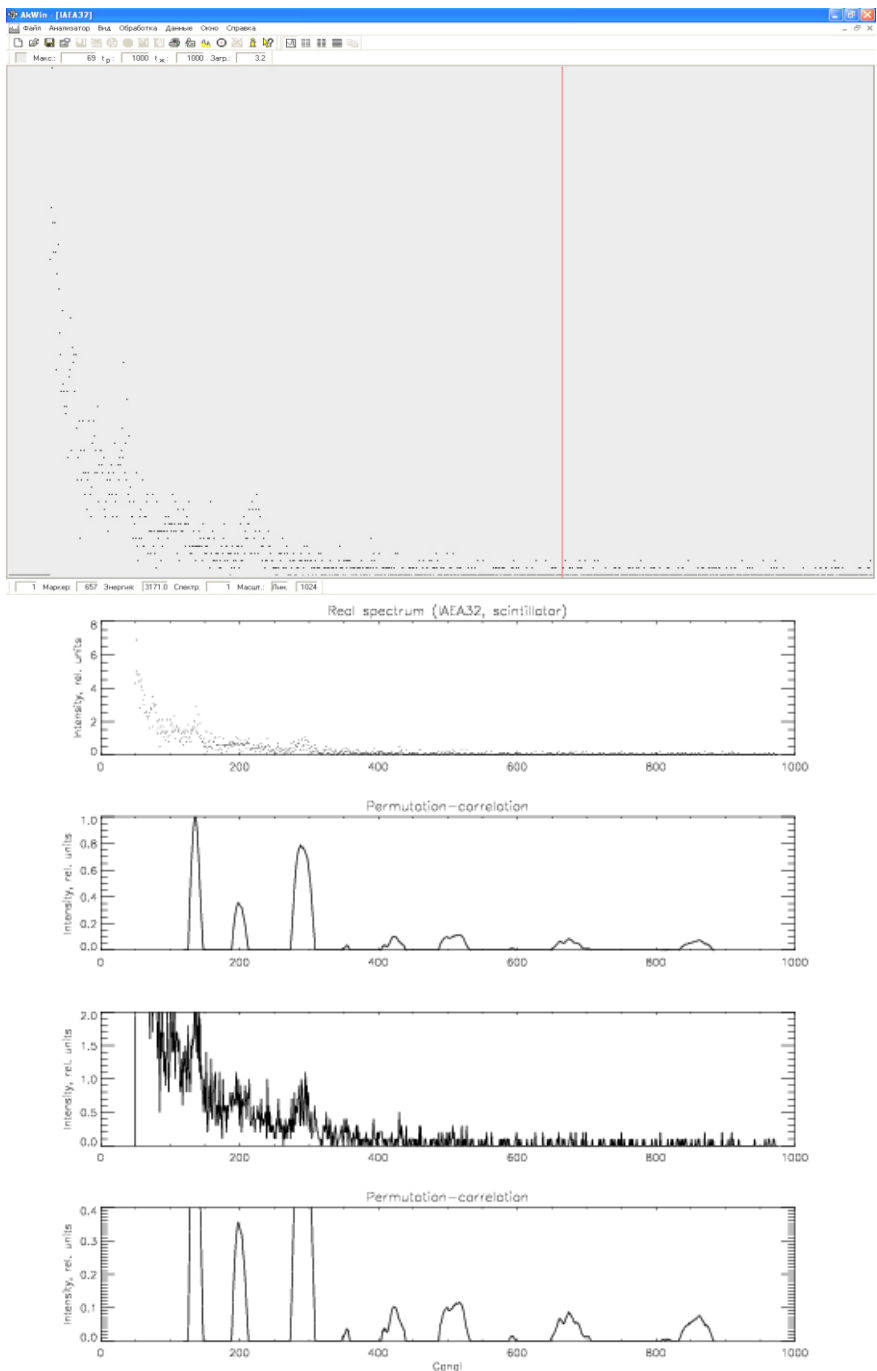


Рис. 2.2.14. Сцинтиляційний спектр зразка МАГАТЕ за час набору 1000 с.

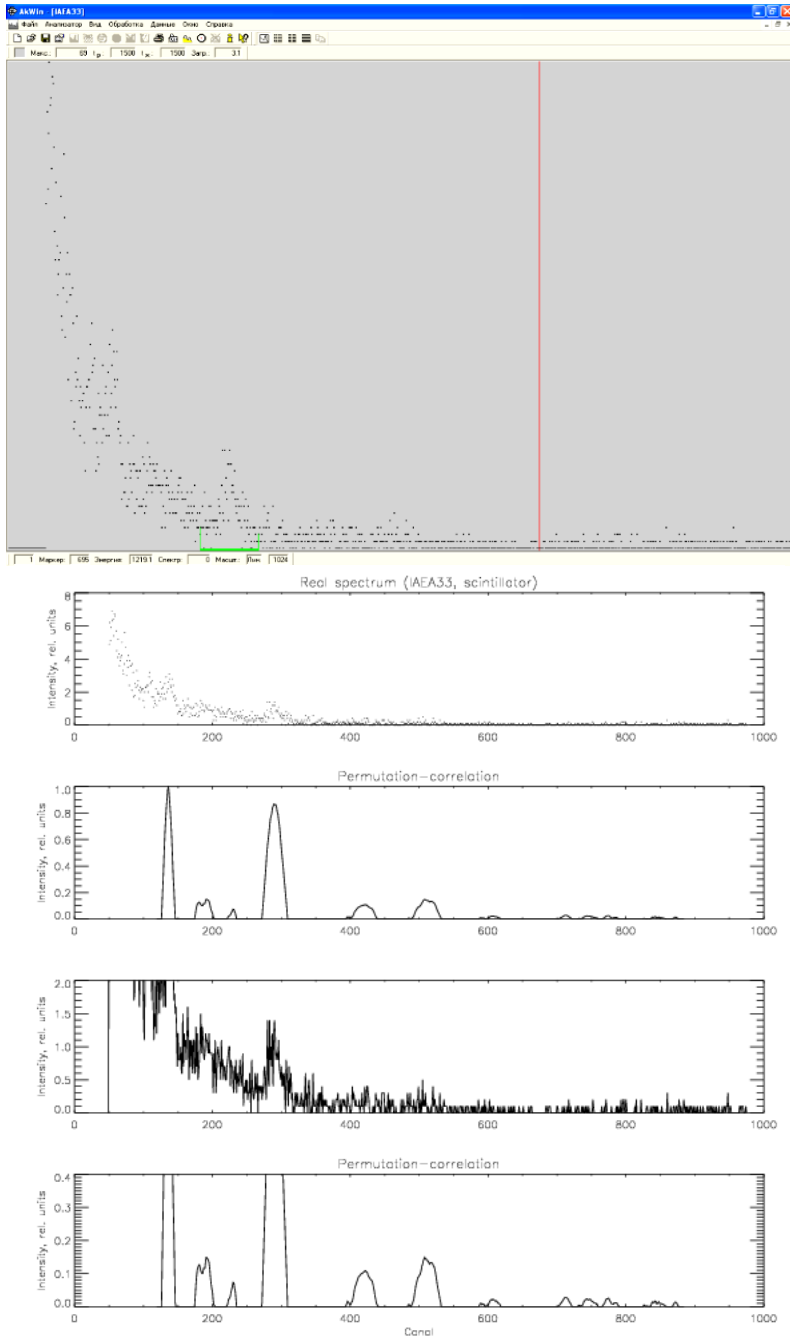


Рис. 2.2.15. Сцинтиляційний спектр зразка МАГАТЕ за час набору 1500 с.

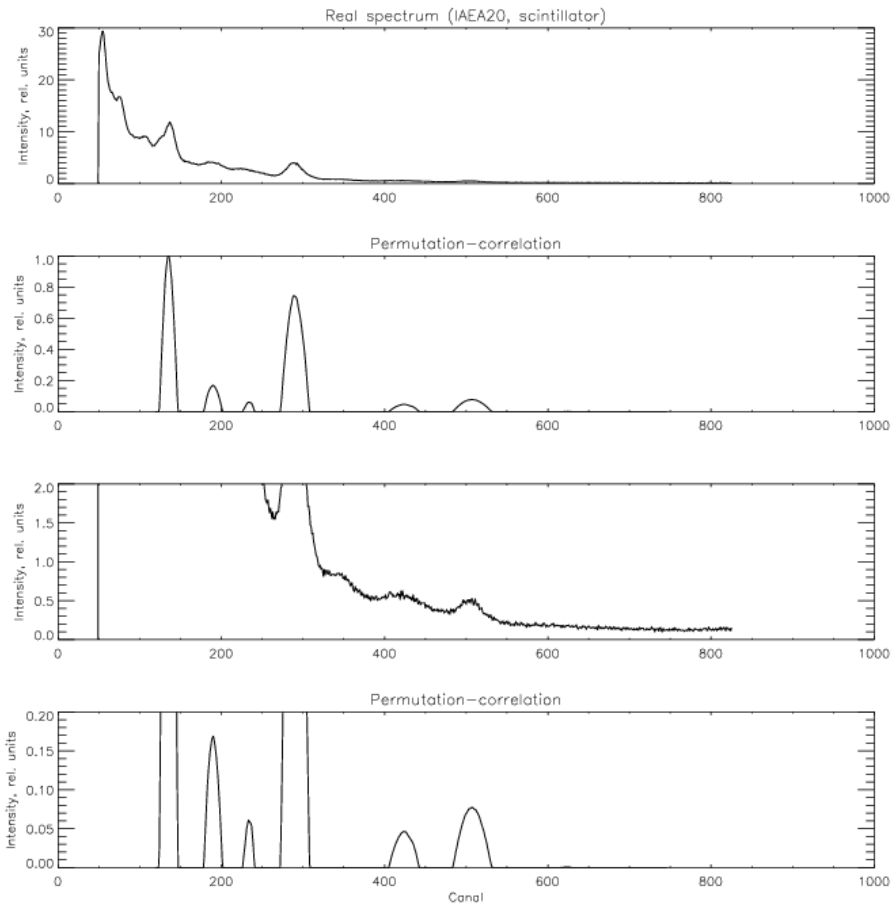
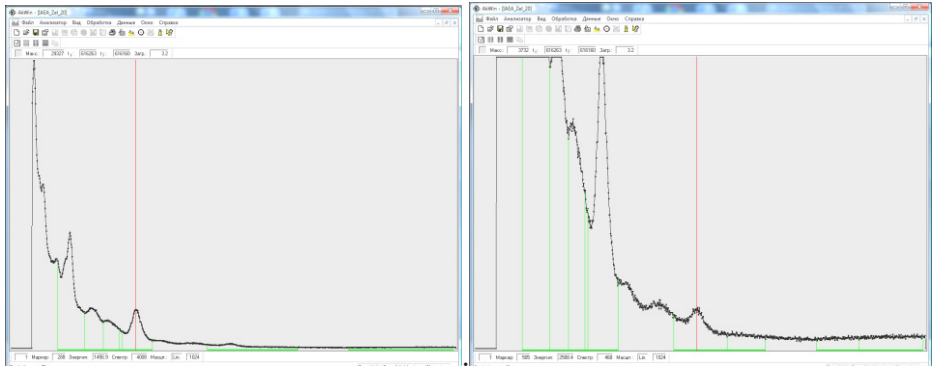


Рис. 2.2.16. Сцинтиляційний спектр зразка МАГАТЕ за час набору 660 000 с.

припиняти, коли кореляційний спектр перестане змінюватись, як, наприклад, на рис 2.2.15.

2.2.2.3. Можливості методу по визначенню інтенсивності ліній

Весь зроблений вище аналіз і результати на рис. 2.2.8 показують, що кореляційний метод з використанням методу перестановок виявився досить чутливим при пошуках відповіді на запитання, чи присутня лінія у даних каналах спектра. З рис. 2.2.16 видно, що в кореляційному спектрі наявні всі лінії вихідного (початкового) спектра, навіть ті слабкі лінії (~ 520-й, 620-й канали), які програмним забезпеченням спектрометра не ідентифікуються. Питання, яке зараз нас цікавить, – чи можна з кореляційного спектра робити кількісні висновки (оскільки в області слабких сигналів програмне забезпечення спектрометра лінії не ідентифікує і, відповідно, інформації про їхню площу не дає) про те, як співвідносяться площі відповідних ліній початкового і кореляційного спектрів, чи відображає кореляційний спектр лише ступінь кореляції між реальними і тестовою лініями, чи він відображає також їхні інтенсивності.

Обидва розглянуті методи підрахунку коефіцієнта кореляції дають інформацію про ступінь кореляції між двома вибірками. Проте між коефіцієнтом кореляції Пірсона і одержаним в результаті тесту перестановок є принципова різниця. Коефіцієнт кореляції Пірсона вводиться виразом (2.2.5) і величина ρ змінюється в інтервалі $(0 \dots 1)$, а кореляція, одержана в результаті тесту перестановок, пропорційна виразу (2.2.4), тому при зміні площі лінії, що є еквівалентним домноженню вибірки Y на деякий коефіцієнт, вираз (2.2.5) залишається незмінним, у той час як вираз (2.2.4) збільшується на цей коефіцієнт.

Перевірку того, як співвідносяться площі звичайного і кореляційного спектрів, було зроблено на основі експериментальних спектрів. Для такого аналізу було одержано серію спектрів ^{137}Cs (662 keV) на сцинтиляційному спектрометрі. Спектри було одержано в одному експерименті для різних тривалостей набору спектра. Кожний спектр було оброблено кореляційним методом.

На рис. 2.2.17 показано співвідношення між площами лінії ^{137}Cs , які було одержано програмним забезпеченням спектрометра із звичайного спектра, і площами ліній, підрахованими з його кореляційного спектра. Між площами існує пряма пропорційність. Це означає, що, одержавши калібрувальну залежність, як на рис. 2.2.17, можна визначати активність ^{137}Cs в пробі навіть тоді, коли програмним забезпеченням спектрометра лінія вже не детектується, проте її видно на кореляційному спектрі. У той же час кореляційний коефіцієнт Пірсона для всіх спектрів рис. 2.2.17 один і той же.

Зауважимо, що на даному етапі розглянутий метод обробки гамма-спектрів не замінює існуючі методи, основані на аналізі похідних, у першу

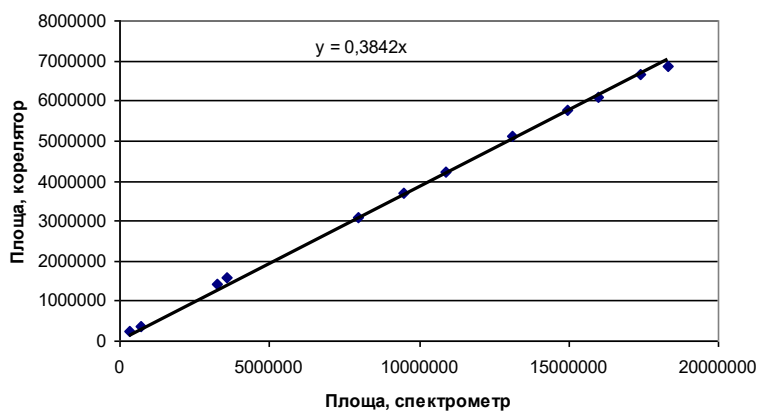


Рис. 2.2.17. Співвідношення між площами ліній реального і кореляційного спектрів.

чергу тому, що він є ефективним лише для окремо розташованих ліній. Дві близько розташовані лінії, які можна розглядати як одну уширену, ширина якої не відповідає очікуваній для даного каналу, кореляційний метод може взагалі не помітити. Це питання може бути предметом подальших досліджень, тим більше що рекомендації ISO [11] по статистичній обробці спектрів для пошуку ліній у шумах, також справедливі лише для окремо розташованих ліній.

2.2.2.4. Висновки до пошуку гамма-ліній

Запропоновано метод пошуку ліній у гамма-спектрах з поганою статистикою шляхом кореляційної обробки – розрахунку кореляції між тестовою лінією і відповідною ділянкою спектра. Показано, що застосування методів комп'ютерної статистики (перестановок методом Монте-Карло) дозволяє надійно визначати присутність корисного сигналу з інтенсивністю на рівні шуму. Метод вільний від будь-яких припущень про функції розподілу і дозволяє в єдиному циклі обробки визначати характеристики ліній до рівня порога детектування по ISO 11929-1:2000. Підвищення чутливості не тільки полегшує проблему аналізу низькоактивних проб, але й є перспективним для інших застосувань, наприклад скорочення часу вимірювань в умовах обмежень часу на такі вимірювання [27].

Стандартизований підхід [7, 11] передбачає можливість вимірювання фону при гарантованій відсутності сигналу – вимірювання при відсутності та наявності зразка в спектрометрі і т. ін. Але при вимірюванні, наприклад, зразків навколишнього середовища, при виконанні спектрометричного гамма-каротажу провести вимірювання фону окремо від сигналу неможливо. Навіть якщо детектор чистий і шумами електронної частини спектрометра можна

знехтувати, то фон забезпечується комптонівськими гамма-квантами від радіонуклідів з більш високою енергією випромінювання, ніж досліджувана ділянка спектра. Цей фон не може бути вимірний окремо від корисного сигналу і його можна ввести у формули для розрахунку DT і DL тільки як результат деяких оцінок. Тому вказані вище (4.1) оцінки DT можна розглядати лише як абстрактну характеристику чутливості (чистоти детектора) спектрометра як приладу, а не як реальну можливість зафіксувати присутність чи відсутність слабого сигналу у спектрі з наявними іншими лініями, що є принциповим при польових вимірюваннях. ISO 11929-1:2000 пропонує визначення фону з двох сторін навколо лінії, яка аналізується. У запропонованому в цьому розділі методі багаторазові вимірювання не обов'язкові і фон видаляється автоматично в процесі обробки спектра. Метод є кількісним і дозволяє проводити прямі вимірювання площі ліній аж до рівня DT без будь-яких припущень про статистичні властивості сигналів, крім гауссової форми ліній. Хоча, власне, і це припущення не є обов'язковим: хоча тестова лінія має гауссову форму, при обробці реальних спектрів ніякі припущення про їхні статистичні властивості не робляться. Важливо лише, щоб тестова лінія мала форму, близьку до форми експериментальної лінії. При відповідному калібруванні спектрометра розроблений кореляційний метод дозволяє визначати параметри ліній у ситуації, коли стандартні методи статистики результатів не дають.

Хоча весь аналіз зроблено на прикладі скінтіляційних спектрів, немає ніяких перешкод для їхнього застосування до напівпровідникових спектрів. Але для них залежність ширини лінії від енергії інша і її також треба визначати з експерименту.

Кореляційний метод добре працює для окремих ліній (дві близько розташовані лінії разом не мають гауссової форми і теоретичної ширини для даної енергії, що є необхідною умовою для застосування кореляційного методу). Але те ж саме можна сказати і про звичайний метод: ISO 11929-3(2000) передбачає можливість використання підходу (2.2.1) також лише для окремих ліній. З іншого боку, відсутність або мала інтенсивність у кореляційному спектрі лінії, яку ясно видно у звичайному спектрі, указує на уширення цієї лінії.

Одержаний результат (на сьогоднішній день якісний) показує, що час набору спектра також має обмеження знизу і повинен бути достатнім, щоб виконувались закладені у формули статистичні посилки, тобто всі канали повинні бути заповненими хоча б 10 імпульсами кожний.

2.2.3. Метод ідентифікації кристалічних фаз з даних рентгенівської дифракції від багатофазних зразків з низьким їхнім відносним вмістом

Практика якісного РФА свідчить, що застосування цього методу до аналізу багатофазних (більше 8 – 10 фаз) зразків, особливо з малим (менше 1 – 3 %) вмістом фаз, наштовхується на великі труднощі [28]. Ця проблема є

наслідком того, що на дифрактограмах таких матеріалів з високою щільністю розташовані десятки або сотні рефлексів, утворені невідомою кількістю кристалічних фаз з низьким вмістом. Проблема ускладнюється, якщо в досліджуваному зразку є фази (наприклад, алюмосилікати), що дають на дифрактограмі десятки рефлексів середньої та, переважно, низької інтенсивності. Візуальне порівняння експериментальних даних рентгенівської дифракції з даними таблиць баз даних (COD, ASTM та ін.) не дає можливості перебрати сотні можливих варіантів для ідентифікації цих фаз. Застосування програм типу [29] для ідентифікації фаз у таких специфічних зразках також виявляється не ефективним. Таким чином, дифрактограми таких багатофазних зразків містять складну комбінацію рефлексів від багатьох фаз, у яких, крім усього, деякі низькоінтенсивні рефлекси можуть бути відсутніми та можуть з'явитися додаткові рефлекси внаслідок суперпозиції деяких рефлексів від різних сполук.

У роботі [28] було показано, що методи комп'ютерної статистики дозволяють виділити слабкі за інтенсивністю лінії (на рівні шумів) у дифрактограмах матеріалів з низьким вмістом кристалічних фаз. Як результат, візуально дифрактограма різко ускладнюється, на ній проявляється велика кількість низькоінтенсивних ліній, що стає своєрідною перешкодою на етапі ідентифікації цих фаз через вищенаведені причини.

Авторами роботи [28] також запропоновано метод пошуку окремих ліній у “зашумлених” даних рентгенівської дифракції на основі статистичного аналізу. У цьому розділі розвинуто метод автоматизованого виявлення серед даних рентгенівської дифракції від досліджуваного матеріалу внесків окремих фаз із малим їхнім відносним вмістом у досліджуваному матеріалі.

2.2.3.1. Кореляційний метод аналізу дифрактограм (пошуку фаз)

У практиці РФА існують емпіричні критерії визначення кристалічних фаз із даних рентгенівської дифракції, які залишаються актуальними дотепер [30, 31] і які, фактично, можна використати як алгоритм для автоматизованого методу ідентифікації фаз.

1. Положення (кути 2θ) рефлексів (від 3 до 6 рефлексів) на дифрактограмі повинні збігатися з положеннями рефлексів передбачуваної кристалічної фази, узяті з відповідних баз даних, наявність якої у зразку ми намагаємось установити.

2. Ці 3–6 рефлексів мають бути найінтенсивнішими серед інших можливих рефлексів цієї ж фази в даних рентгенівської дифракції досліджуваного зразка.

3. Ці 3–6 рефлексів повинні бути розташовані (за їхньою відносною інтенсивністю) у тій же послідовності, в якій вони розташовані для тієї самої фази в базі даних рентгенівської дифракції.

Як бачимо, усі вищенаведені критерії засновані на візуальному порівнянні експериментальних дифрактограм з дифрактограмами на основі

даних рентгенівської дифракції з баз даних. Однак насправді цей підхід непридатний для матеріалів, у яких присутні більше 4–6 фаз i , відповідно, 20–90 рефлексів. Оскільки окремі рефлекси з різних фаз можуть збігатися, важко визначити, чи є конкретний рефлекс середньої інтенсивності одним із 6 рефлексів з фази з найбільшим змістом, чи це основний (тобто перший) рефлекс з фази з меншим змістом у досліджуваному зразку. Ситуація ускладнюється, якщо врахувати, що як положення рефлексів (кут 2θ), так і їхні інтенсивності мають експериментальну похибку.

Важко визначити фазу, коли через помилку експерименту інтенсивність рефлексу (яка є середнім у серії з 6 рефлексів однієї фази) дорівнює або перевищує інтенсивність першого рефлексу іншої фази. Відсутність окремих рефлексів малої інтенсивності на фоні шуму також можна розглядати як ситуацію, коли реальна дифракційна картина лише з певною ймовірністю має свою структуру рефлексів.

Відомі програми обробки даних РФА [29, 31] дають задовільні результати у випадках дослідження матеріалів з невеликою кількістю фаз і дають практично неприйнятні результати у випадках, коли інтенсивності рефлексів знаходяться на рівні “шумів”. Таким чином, стає зрозумілим, що для процесу ідентифікації фаз у матеріалах зі значною їхньою кількістю доцільно створювати та використовувати спеціалізовані математичні програмні продукти з алгоритмами, побудованими на методах обчислювальної статистики, що дозволить ідентифікувати фази у вищевказаних випадках.

На відміну від традиційного способу обробки даних рентгенівської дифракції, описаного вище [32], який по суті є детермінованим, у цьому розділі розглядається можливість використання ймовірнісних методів. Справді, при традиційному методі обробки даних рентгенівської дифракції також перевіряється виконання певних умов: наявність певної кількості рефлексів, співвідношення їхніх відносних інтенсивностей тощо.

Оскільки положення рефлексів (кути 2θ) відповідних кристалічних фаз відомі точно, особливо для тих, що наведені в базах даних, на перший погляд може здатися, що необхідно порівнювати інтенсивності рефлексів для заданих кутів. Однак насправді положення рефлексів у РФА визначаються з деякою похибкою. Значення відносних інтенсивностей рефлексів також мають свою невизначеність ($\pm 15 - 25\%$), яку важко врахувати традиційним методом ідентифікації. Тобто і тут результат отримано з певною достовірністю. І, отже, неявно передбачається, що дані певною мірою є випадковими.

Усе це дає підстави думати про необхідність, на відміну від сформульованих вище трьох критеріїв, застосування методів аналізу для ідентифікації фаз у даних рентгенівської дифракції, які б давали можливість оцінювати ймовірність правильності зробленої ідентифікації (у більш строгому статистичному формулюванні – оцінку ймовірності помилки першого роду – прийняття неправильної гіпотези [33] про ідентифікацію фази). Стандартні програмні продукти типу [29], що застосовуються до багатофазних матеріалів, дають великий за обсягом масив передбачуваних

фаз, що вимагає неприпустимих витрат часу на їхній аналіз. Тому логічно сформулювати проблему підвищення ефективності процесу ідентифікації за допомогою статистичних методів.

Для вирішення задачі пошуку раніше встановленої комбінації рефлексів серед багатьох інших рефлексів на експериментальній дифрактограмі пропонується шукати кореляцію між вибірками експериментальних даних (на основі яких побудована експериментальна дифрактограма) і даних рентгенівської дифракції, узятих із різних даних експериментів порошкової дифракції. Розрахунки базуються на використанні методів обчислювальної статистики для визначення кореляцій [5, 22, 28]. По-перше, давайте переформулюємо проблему в термінах статистики.

1. Дифрактограма в цілому або якась її частина (вибраний діапазон кутів) буде розглядатися як статистична вибірка. У цьому випадку кути беруть без пропусків регулярно з кроком вимірювання, скажімо, в $0,05^\circ$. Кожне значення кута відповідає певному значенню інтенсивності, тобто ці пари значень $\{2\theta_i, I_i\}$ утворюють регулярну вибірку пар значень.

2. З бази даних [34] взято таблицю кутів і відносних інтенсивностей, що відповідають передбачуваній фазі, наявність якої необхідно визначити на експериментальній дифрактограмі. У цих табличних даних є лише окремі кути та відповідні значення відносних інтенсивностей. Тому їх необхідно перетворити у форму, яку має експериментальна дифрактограма. Для цього створюється вибірка такого ж розміру, як і експериментальна вибірка в попередньому пункті, тобто з такими ж значеннями кутів, але з нульовою інтенсивністю. Далі нульові значення інтенсивностей для табличних (з бази даних [34]) кутів замінюються на відповідні табличні значення відносних інтенсивностей з тієї ж бази даних. Таким чином створюється масив інтенсивностей для табличних кутів з тим самим розміром, що й експериментальний.

3. Бази даних надають конкретне значення кута для відповідного рефлексу. Але це лише величина кута в максимумі рефлексу, яка, до того ж, може мати різні значення за даними різних авторів (як приклад див. аналіз ситуації в [28]). У практиці РФА прийнятною вважається різниця між експериментальними значеннями кутів і відповідними значеннями кутів з баз даних на величину $0,004 - 0,02^\circ$. У методі, описаному нижче, значення інтенсивностей експериментальних і табличних (з баз даних рентгенівської дифракції) зразків для відповідних кутів перемножуються. Якщо кути відрізняються на обраний крок $0,05^\circ$, то результат множення дорівнюватиме нулю, оскільки в дослідних чи табличних зразках вони матимуть нульові значення інтенсивностей.

4. Для врахування ширини рефлексів у вибірці, створеній з бази даних, необхідно одиничні табличні значення для певних кутів перетворити у рефлекс певної ширини. Окремого розгляду потребує питання про форму рефлексу, побудованого за даними експериментальної рентгенівської дифракції [35]. Однак для наших цілей створення методу пошуку серед рефлексів

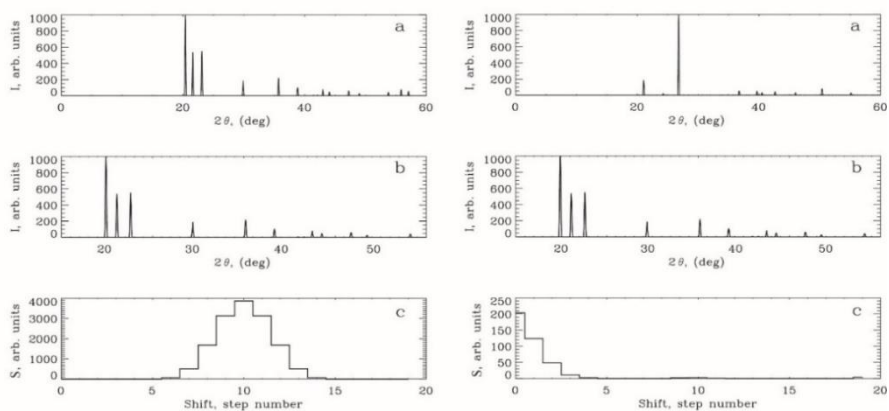
експериментальної дифрактограми, що відповідають шуканій кристалічній фазі, форма рефлексів не має принципового значення. Таким чином, один із найпростіших способів “розширити” табличний рефлекс, побудований на основі даних із баз даних рентгенівської дифракції, полягає в тому, щоб взяти значення інтенсивностей рефлексів з експериментальної вибірки даних.

5. Після створення таких вибірок з експериментальних даних і даних баз даних рентгенівської дифракції COD або International Center for Diffraction Data (ICDD) для потрібної фази необхідно розглянути завдання їхнього статистичного порівняння, а саме розрахувати ступінь кореляції між ними. Відразу зазначимо, що розрахунок звичайного коефіцієнта кореляції (Пірсона або інших) не дає відповіді на поставлені питання. Тому запропоновано метод розрахунку ступеня кореляції між створеними вибірками, який базується на методах обчислювальної статистики [22, 35]. Ці методи, у свою чергу, засновані на використанні методів Монте-Карло. Поглиблене обґрунтування необхідності застосування підходів обчислювальної статистики для нашої задачі аналізу закономірностей, алгоритмів і формул можна знайти в [5, 22, 28, 35]. Цей метод аналізу базується на кореляційних розрахунках. Отже, якщо взяти дві ідентичні вибірки дифракційних даних, ці дані будуть повністю корельовані. Особливістю пошуку кореляції в запропонованому методі є те, що дифракційні дані фази, наявність якої встановлено на складній експериментальній дифрактограмі, не є зразком, який необхідно сканувати по різних ділянках експериментальної дифрактограми – загальний діапазон кутів однаковий для обох моделей. Питання полягає в тому, наскільки скорельовані інтенсивності для даного набору кутів. На перший погляд здається, що після створення двох вибірок для порівняння, як описано вище, можна просто обчислити коефіцієнт кореляції між ними.

6. Однак прямий розрахунок ступеня кореляції між двома вибірками даних, відображених на двох дифрактограмах, виявився неефективним в тому сенсі, що отриманий коефіцієнт кореляції нема з чим порівнювати. Ми а ргіогі визнаємо, що з різних причин, як перерахованих вище, так і з простих причин, що в “зашумлених” дифракційних даних співвідношення інтенсивностей рефлексів спотворюються шумом, порівнянням за інтенсивністю з табличними даними для рефлексів. З цієї причини ми не можемо сподіватися на 100 % кореляцію. І немає можливості відкалібрувати результати обчислень. Тому була запропонована така методика. Експериментальний набір даних і набір даних з баз даних рентгенівської дифракції розглядаються як два зразки з однаковими кутами, але різними інтенсивностями для цих кутів. Зрушимо набір даних інтенсивностей дифракційної картини шуканої фази з бази даних на певну кількість позицій вліво по осі абсцис, на якій відкладено значення кутів (у наведених нижче прикладах 10 кроків), які були обрані, що відповідає зсуву $10 \times 0,05 = 0,5$ град) (рис. 2.2.18). Тепер, використовуючи згадані вище методи комп’ютерної статистики, ми обчислюємо кореляцію між експериментальним набором дифракційних даних і даними бази даних, що відповідають потрібній фазі, щоразу на один крок зсуваючись до «правильної»

відповідності кутів, загалом на 20 кроків. У випадку, коли ми розраховуємо кореляцію двох ідентичних наборів даних, ми отримуємо максимальну кореляцію на 10-му кроці, яка почне зменшуватися далі зі зростанням відмінностей між кутами. Зверніть увагу, що при деякому зсуві кутів дифракційних картин також можлива помітна кореляція, але ми знаємо (оскільки це включено в метод розрахунку), що справжній результат виходить лише на 10-му кроці.

Особливістю оцінки ступеня кореляції методами обчислювальної статистики в нашому випадку є те, що в даному випадку виходить не коефіцієнт кореляції в частинах одиниці, а ймовірність правильного прийняття нульової гіпотези про ідентичність набору даних рентгенівської дифракції з баз даних і набору експериментальних даних. При цьому, у силу специфіки методу, результат розрахунків виходить у вигляді деяких відносних одиниць. Це ще не дає можливості порівняти відносний вміст окремих кристалічних фаз між собою в зразку, але дає можливість достовірно встановити наявність даної фази в одному зразку за ступенем кореляції між наборами даних експериментальних і табличних моделей.



1. Пошук тридиміту на дифрактограмі цього ж тридиміту: а – модельна” дифрактограма тридиміту; б – дифрактограма тридиміту з бази даних; с – результуюча гістограма кореляції між ними.

2. Результат того ж пошуку тридиміту в α -кварці.

Рис. 2.2.18. Результати пошуку рефлексів з бази даних в модельних дифракційних даних тридиміту та α -кварцу.

На цей факт вказує відмінна від нуля кореляція між експериментальним набором даних і вибраним набором даних фази з бази даних. Якщо хтось хоче автоматизувати процес прийняття рішення про ідентифікацію фази, йому також доведеться встановити критерій для цього. На сьогоднішній день не розроблено неемпіричних методів для порівняння таких наборів даних. У

нашому випадку це означає, що експериментатор сам вирішує, чи надійно рефлекс на кроці 10 з'являється на гістограмі.

2.2.3.2. Демонстраційні розрахунки

Спочатку продемонструємо вищезазначене на прикладі розрахунку ступеня кореляції між вибірками з бази даних рентгенівської дифракції. Зверніть увагу, що значення інтенсивності для цих даних нормалізовано таким чином, що максимальна інтенсивність рефлексу встановлена на рівні 1000. Як приклади оксиди кремнію тридиміт і кристобаліт були обрані довільно. Для початку, щоб продемонструвати роботу алгоритму розпізнавання, порівняємо ту саму вибірку даних рентгенівської дифракції тридиміту з бази даних COD (див. рис. 2.2.18, 1, а). Ми розглянемо один масив даних, який відображається як дифрактограма (рисунок 2.2.18, 1, а), отриманий експериментально, тоді як він, показаний як дифрактограма на рис. 2.2.18, 1, b, буде “тестом”.

На рис. 2.2.18, 1, c показано результат аналізу, який побудований як результат обчислення ступеня кореляції між двома масивами даних рентгенівської дифракції за допомогою алгоритму “сканування”, як описано вище. Ступінь кореляції обчислюється у відносних одиницях, оскільки вона залежить від абсолютних значень інтенсивностей у вибірках, що порівнюються між собою згідно з [22]. А про сам факт кореляції свідчить той факт, що на кроці 10 є пік, висота якого перевищує всі інші й задана в цих відносних одиницях. Як і очікувалося, максимальна кореляція спостерігається, коли зміщення між кутами дорівнює нулю, тобто на 10-му кроці сканування.

На цьому ж рисунку показано гістограму ступеня збігу дифрактограми α -кварцу та дифрактограми кристобаліту. Ця ситуація, коли дифрактограми порівнюваних сполук відрізняються, повинна продемонструвати, що в такому очевидному випадку ступінь кореляції дорівнює нулю — немає піка на кроці 10. Підкреслимо, що неможливо порівняти висоти піків для різних фаз на цьому етапі розрахунків: як було сказано, вони залежать від значень інтенсивностей рефлексів у вихідних вибірках даних рентгенівської дифракції.

Тому на рис. 2.2.18, 1 чітко видно, що між “експериментальним” (а) і табличним (b) масивом даних (у даному випадку ідентичні), існує надійна кореляція, яка проявляється наявністю піка на гістограмі (c) на 10-му кроці. І явно видно, що немає кореляції між різними масивами даних рентгенівської дифракції, відображених на дифрактограмах (α -кварцу і тридиміту) на рис. 2.2.18, 2.

2.2.3.3. Аналіз даних рентгенівської дифракції

Тепер перевіримо запропонований спосіб на більш складному прикладі. Цей приклад також є типовим, тому що якщо ми тестуємо метод, ми повинні знати правильну відповідь перед початком роботи. Тому ми беремо табличні дифрактограми тридиміту, кристобаліту, діоксиду урану та дифрактограму α -

кварцу. Давайте об'єднаємо ці чотири дифрактограми (додаючи масиви даних) в одну і застосуємо наш метод, щоб знайти в ній ті чотири фази, з яких вона складається. Результат показаний на рис. 2.2.19, 2.2.20, 2.2.21, 2.2.22. Для всіх варіантів пошуку на кроці 10 сканування є пік, тобто знайдено всі фази, які були включені в загальну дифракційну картину.

2.2.3.4. Модельна дифрактограма

Оскільки представлені в цьому розділі результати є методичними за своєю суттю, то результати запропонованого методу необхідно показати на прикладах, де результати відомі заздалегідь. Але таких даних практично немає: у базах даних рентгенівської дифракції є дані тільки про окремі кристалічні фази, але не про багатофазні матеріали.

Виявлено, що знайти тестовий приклад складної експериментальної дифрактограми для демонстрації реальних можливостей запропонованого методу не так просто. Так, у ЛПВМ зруйнованого 4-го енергоблоку ЧАЕС кристалічні фази розподілені в неоднорідній і нерівноважній склофазі.

Унаслідок цього, рефлекси оксиду урану на дифрактограмах виявляються

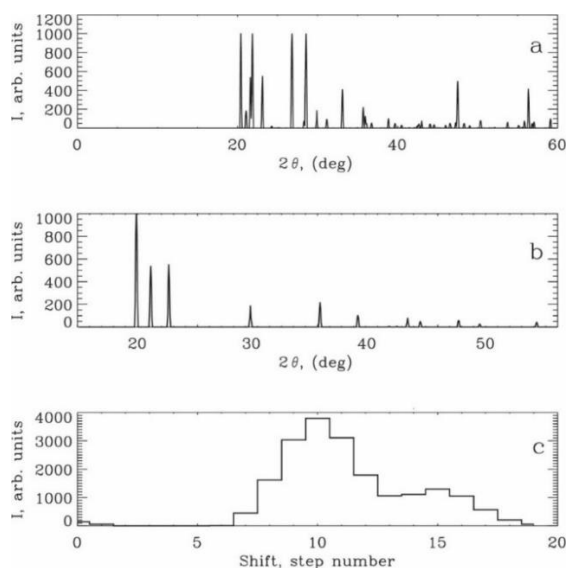


Рис. 2.2.19. Результати пошуку рефлексів фази тридиміту на “модельній” дифрактограмі: а – “модельна” дифрактограма; б – дифрактограма тридиміту, побудована на основі даних з бази даних рентгенівської дифракції; с – результуюча гістограма, на основі коефіцієнтів кореляції між цими двома масивами дифракційних даних.

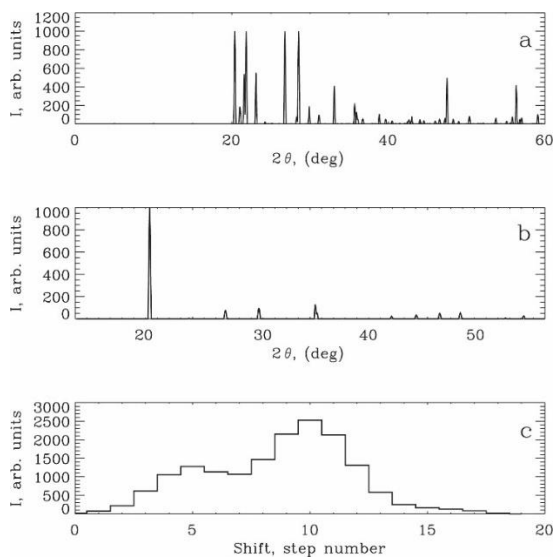


Рис. 2.2.20. Результати пошуку кристобаліту на "модельній" дифрактограмі: а – "модельна" дифрактограма; б – дифрактограма кристобаліту, побудована з використанням даних рентгенівської дифракції з бази даних COD; с – результуюча гістограма, на основі коефіцієнтів кореляції між цими двома масивами дифракційних даних.

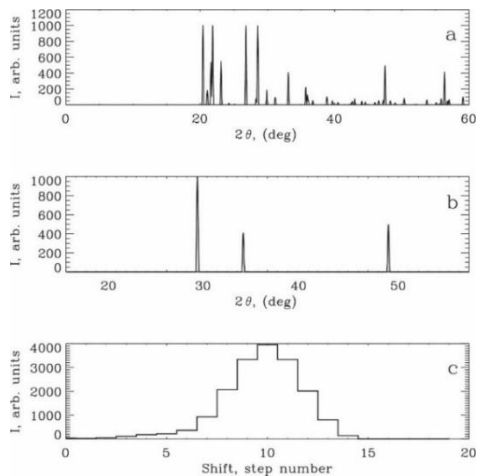


Рис. 2.2.21. Результати пошуку рефлексів фази діоксиду урану (з бази даних COD) на "модельній" дифрактограмі: а – "модельна" дифрактограма; б – дифрактограма діоксиду урану; с – результуюча гістограма на основі коефіцієнтів кореляції між цими двома масивами дифракційних даних.

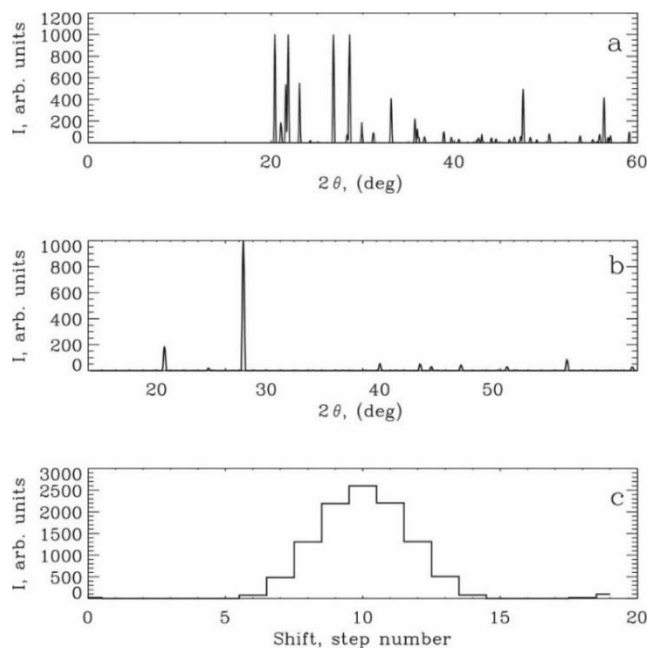


Рис. 2.2.22. Результати пошуку рефлексів α -кварцу (COD) на “модельній” дифрактограмі: a – “модельна” дифрактограма; b – дифрактограма α -кварцу; c – результуюча гістограма на основі коефіцієнтів кореляції між цими двома масивами дифракційних даних.

зміщеними по кутах від табличних значень внаслідок деформації ґратки оксиду урану через механічне стиснення включень оксиду урану склофазою [3].

Таким чином, дифрактограми таких матеріалів відрізняються від табличних даних і не можуть використовуватися як тести, оскільки самі вимагають нестандартної ідентифікації. Однак є бажання продемонструвати, як метод працює на реальних даних.

Таким чином, єдиний спосіб отримати складну дифрактограму для тестування - це змоделювати її шляхом підсумовування значень інтенсивностей рефлексів при відповідних значеннях кутів. Експериментальну дифрактограму коричневої кераміки ЛПВМ 4-го енергоблоку ЧАЕС прийнято за базову. На цій дифрактограмі спостерігаємо “шумову доріжку” з кількома рефlekсами високої інтенсивності, які виділяються на фоні шумів. Показано, що коричнева кераміка ЛПВМ містить щонайменше 8 кристалічних фаз [3].

Як додаток до цієї дифрактограми також було обрано експериментальну дифрактограму зі стандарту α -кварцу, з якої видалено найінтенсивніший рефлекс ($2\theta = 25^\circ$). У такій конструкції моделі інтенсивність максимальної лінії (друга лінія $2\theta = 21,05^\circ$ α -кварцу, див. рис. 2.2.23, b) була зроблена приблизно на 0,1 від максимальної лінії базової дифракційної картини, щоб перевірити чутливість нашого методу до слабких рефлексів.

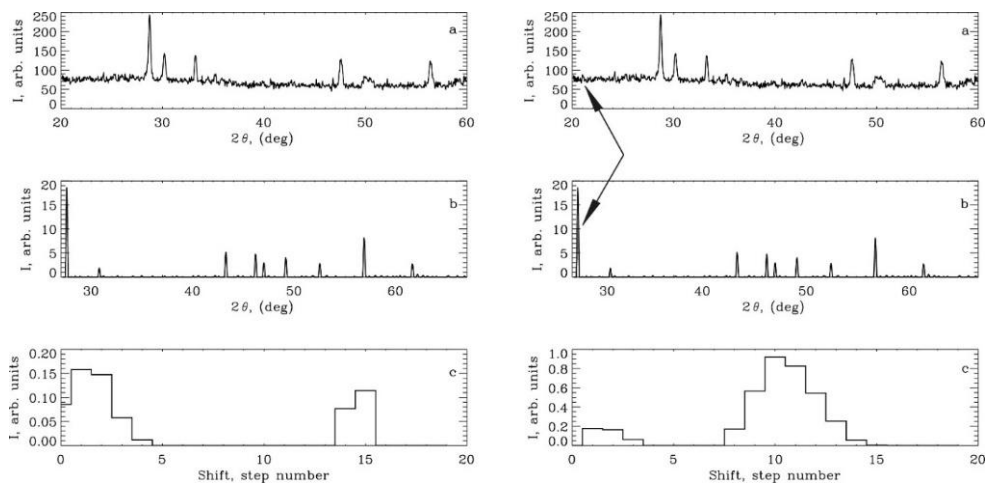


Рис. 2.2.23. Результати того самого пошуку α -кварцу в сумі α -кварцу та коричневої кераміки. Пошук α -кварцу на дифрактограмі коричневої кераміки: а – дифрактограма коричневої кераміки; б – дифрактограма α -кварцу з бази даних; с – результуюча гістограма кореляції між ними. Результат того ж пошуку α -кварцу в сумі α -кварцу та коричневої кераміки.

Потім ці дві дифрактограми об'єднали в одну (див. рис. 2.2.23, а). У такій комбінованій дифракційній картині лінію від $2\theta = 21,05^\circ$ майже не видно (див. стрілку), а слабші лінії α -кварцу не завжди видно над шумом, але ми пам'ятаємо, що вони присутні там.

Результат нашого пошуку фази α -кварцу в такій дифракційній картині показано на рис. 2.2.23, с. Гістограма з чітким піком на 10 кроці свідчить про те, що розвинений у цьому розділі метод обробки даних рентгенівської дифракції задовільно працює в такій складній ситуації.

2.2.3.5. Приклад з даними рентгенівської дифракції

Вище ефективність методу була продемонстрована на модельних даних, які можна розглядати як приклад відносно складних закономірностей. Однак набагато більший інтерес являє питання про можливості методу для аналізу даних зі значно більшим набором фаз, сумарний ефект рефлексів від яких уже схожий на шум.

Як такий експериментальний приклад також було взято дифрактограму так званих коричневих ЛПВМ 4-го енергоблоку ЧАЕС з рис. 2.2.23, але без додавання α -кварцу. Зовні це виглядає як шумова смуга з кількома рефlekсами, які можна віднести до оксиду урану. Однак добре відомо, що ці лави мають складний фазовий склад [3]. За різними даними [3, 36] встановлено, що ці лави повинні містити оксид цирконію, наявність якого ми спробуємо перевірити розробленим методом. Результат розрахунку ступеня

кореляції між дифрактограмами коричневої кераміки та кубічного ZrO_2 наведено на рис. 2.2.24. На цьому рисунку видно, що наявність кубічного ZrO_2

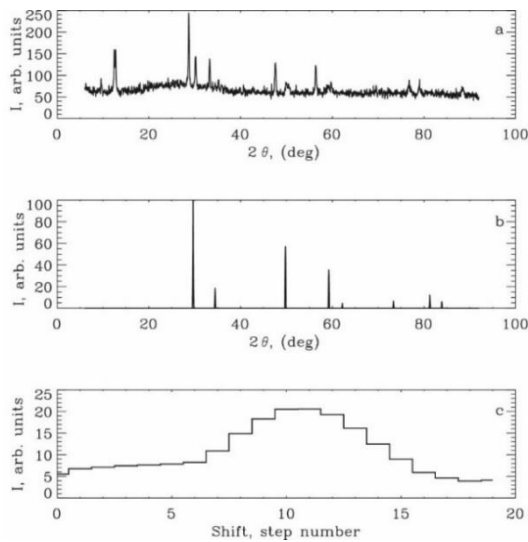


Рис. 2.2.24. Результати пошуку кубічного оксиду цирконію на дифрактограмі коричневої кераміки: а – дифрактограма коричневої кераміки; б – дифрактограма кубічного оксиду цирконію; с – результуюча гістограма на основі коефіцієнтів кореляції між цими двома масивами дифракційних даних.

достовірно встановлено за наявності піка на сходинці 10 пошуку кореляцій. У цьому випадку на тлі дифракційної картини візуально видно лише перший з усіх рефлексів ZrO_2 . Зверніть увагу, що значення кореляції в максимумі гістограми на рис. 2.2.24 дорівнює 20 відносних одиниць, що значно менше ніж відповідні значення кореляцій на рис. 2.2.19–2.2.21 для модельних дифрактограм. Це повністю відповідає очікуванням, оскільки очікувані рефлекси від ZrO_2 не видно на експериментальній дифракційній картині на тлі фонового шуму. Це означає, що його концентрація дуже мала.

2.2.3.6. Новий метод ідентифікації фаз в РФА

Розроблено метод ідентифікації фаз на дифрактограмах, який базується на використанні статистичного підходу. Метод використовує розрахунок кореляції між реальною та табличною дифрактограмами.

Запропонований у представленій роботі метод ідентифікації кристалічних фаз у багатофазних зразках дозволяє:

а) надійно ідентифікувати наявність певних кристалічних фаз у багатофазних зразках;

б) метод має підвищену чутливість і дає змогу визначати фази відносно

низького вмісту в зашумлених даних рентгенівської дифракції – надійно визначати фази, інтенсивність рефлексів яких сумірна з фоном;

в) на цьому етапі розроблений метод ідентифікації фаз у зашумлених структурах відповідає лише на питання про наявність цієї фази у складній структурі. Він дає однозначний, хоча й досі не кількісний критерій наявності тієї чи іншої фази: наявність піка на гістограмі для коефіцієнта кореляції, який знаходиться у правильному положенні (у нашому випадку на 10-му кроці зсуву кута). Водночас підкреслимо, що представлені результати беззаперечно свідчать про надійність методу там, де інші методи взагалі не працюють, а саме коли рефлекси практично непомітні на фоні шумів. Тобто, як кінцевий результат, використання запропонованого методу визначення кореляцій підвищує надійність ідентифікації фаз.

2.2.4. Метод оцінки вмісту кристалічних фаз у багатофазних зразках за даними рентгенівської дифракції

Процедуру кількісного аналізу даних рентгенівського фазового аналізу багатофазних зразків можна умовно розділити на дві частини. Спочатку по дифрактограмі ідентифікуються фази, а потім за співвідношеннями інтенсивностей рефлексів оцінюється вміст виявлених фаз. Загальноприйнята процедура аналізу стає складною, якщо не неможливою, у двох крайніх випадках: коли число виявлених рефлексів стає дуже великим (сотні рефлексів) та коли рефлекси важко виділити на фоні шумів, що веде до можливості втрати істинних чи помилкової ідентифікації неіснуючих рефлексів. Однак навіть у випадку двох фаз оцінка вмісту фаз на сьогоднішній день вимагає порівняння інтенсивності основного рефлексу експериментальної дифрактограми з еталонними калібрувальними кривими.

У попередньому розділі розглядаються проблеми аналізу даних рентгенівської дифракції від зразків багатофазних матеріалів при низькому вмісті фаз. Класичні методи РФА основані на ймовірнісному підході визначення тих чи інших параметрів дифрактограм, який у свою чергу базується на очевидній посиленні, що в реальності складні дифрактограми аналізуються з певним ступенем надійності, особливо у випадках зашумлених дифрактограм з низьким вмістом фаз, коли лінії накладаються одна на одну, ховаються в шумах тощо. Усе це дає підстави розглядати такі випадки як ситуацію, коли в реальній дифрактограмі відома з баз даних таблична дифрактограма деякої фази реалізується лише з певною ймовірністю.

У роботі [28] було розглянуто проблему ідентифікації ліній дифрактограми для випадку, коли інтенсивності цих ліній знаходяться на рівні шумів. Цю задачу було вирішено шляхом використання кореляційного аналізу: пошуку кореляції між вибраною ділянкою дифрактограми і тестовою лінією, параметри якої близькі до параметрів ліній експериментальної дифрактограми. Ключовим у цьому підході є розрахунок ступеня кореляції на основі методів обчислювальної статистики [5, 22]. Цей метод дозволив у кілька

разів підвищити можливість визначати наявність ліній для даного кута та, відповідно, підвищити надійність оцінки інтенсивності ліній, амплітуда яких знаходиться на рівні шумів.

У попередньому розділі було запропоновано ймовірнісний підхід до ідентифікації вже не окремих ліній, а окремих фаз, які у такій задачі розглядаються як набір значень інтенсивностей для заданих кутів. Цю задачу розв'язано шляхом пошуку кореляцій між двома вибірками, одна з яких є експериментальною дифрактограмою, а інша – отримана з бази даних дифрактограма деякої сполуки, присутність чи відсутність якої у даному зразку треба встановити. Ступінь кореляції також розраховується на основі методів обчислювальної статистики [22]. Розроблений у [28] кореляційний метод у випадку зашумлених дифрактограм виявився набагато ефективнішим, ніж звичайна процедура розшифровки дифрактограм, яка полягає в «ручному» порівнянні отриманої експериментальної дифрактограми з інформацією з баз даних. Отже, другу задачу, а саме задачу ідентифікації фаз, особливо фаз низького вмісту, на дифрактограмах з великим числом фаз також можна вважати розв'язаною.

Проте задача кількісного визначення вмісту фаз у вказаних вище умовах низького вмісту і зашумлених дифрактограм залишається проблематичною. Звичайна процедура аналізу для визначення вмісту фаз полягає в порівнянні інтенсивностей найбільш інтенсивних ліній різних фаз та використанні еталонів [38, 39]. При великій кількості малоінтенсивних ліній від фаз низького вмісту та відсутності еталонів ця процедура виявляється не ефективною через велику ймовірність неправильної ідентифікації слабких ліній.

У подальшому показано результати, отримані у роботах [5, 22, 28]. У цьому розділі метод ідентифікації фаз на основі автоматизованого кореляційного аналізу, включаючи програмне забезпечення для ідентифікації фаз, розширено на задачу кількісної оцінки вмісту окремих фаз.

Усі дифрактограми в базах даних нормуються так, щоб інтенсивність максимального рефлексу дорівнював тисячі. Тому отримані в експерименті інтенсивності рефлексів з наборів даних рентгенівської дифракції є відносними величинами. У результаті зміни експериментальних інтенсивностей рефлексів у дійсності з табличними не можна порівнювати і для кількісних оцінок використовуються калібрувальні криві, отримані з аналізу еталонних зразків.

Ідея запропонованого підходу базується на тому, що для кожної кристалічної фази, тобто для відповідного їй масиву дифракційних даних, можна ввести деякий характерний параметр, за яким інтенсивності рефлексів різних масивів дифракційних даних можна порівнювати між собою. В якості такого параметра було запропоновано використовувати суму інтенсивностей рефлексів нормованого масиву даних рентгенівської дифракції для данної кристалічної фази з бази даних. У поєднанні з тим, що значення інтенсивностей прив'язано до певних кутів, такий параметр є однозначною

характеристикою набору даних рентгенівської дифракції даної, наперед визначеної кристалічної фази.

2.2.4.1. Дані рентгенівської дифракції для аналізу

У даній роботі для розробки і демонстрації можливостей методу використано дані рентгенівської дифракції з баз даних. Початкові спроби авторів розробляти метод на основі власних експериментальних дифрактограм наштовхнулися на проблему. Для демонстрації можливостей методу його потрібно демонструвати на повністю і надійно ідентифікованих кристалічних фазах і відповідних масивах даних рентгенівської дифракції, щоб заздалегідь знати правильну відповідь, з якою буде порівнюватись результат розробленого методу. Такий результат можна отримати, наприклад використовуючи еталони. Але при створенні еталонів однаково з'являється проблема надійності створених еталонів. А в кінцевому результаті отримані експериментально дані рентгенівської дифракції від еталонів знову ж таки аналізуються шляхом порівняння з даними рентгенівської дифракції з баз даних. Тому в даній роботі дифракційні дані від багатозначних зразків для аналізу розробленим методом створювалися як проста комбінація даних рентгенівської дифракції від кількох кристалічних фаз, узятих з баз даних. А вже еталонні дані рентгенівської дифракції використано для демонстрації працездатності методу.

2.2.4.2. Кореляційний аналіз на основі методів Монте-Карло

Поставлену задачу визначення наявності та кількісного вмісту тих чи інших кристалічних фаз в методах рентгенівської порошкової дифракції можна звести до задачі порівняння двох вибірок: експериментальної дифрактограми і дифрактограми з бази даних для сполуки, присутність якої перевіряється. З точки зору статистики дифрактограма являє собою дві колонки цифр: колонка регулярних кутів і колонка відповідних їм інтенсивностей ліній. Якщо порівняння інтенсивностей ліній різних дифрактограм проводиться для одних і тих же кутів, колонки кутів можна далі не розглядати, і тоді задачу розшифровки дифрактограми можна переформулювати як задачу порівняння інтенсивностей для відповідних кутів, тобто як задачу порівняння парних вибірок чисел, які через невизначеності оцінок інтенсивностей рефлексів, через можливе накладання ліній і навіть можливу відсутність деяких ліній можна розглядати як задачу порівняння двох випадкових (у вказаному сенсі) парних вибірок – як задачу для математичної статистики.

Розрахунки оснований на використанні методів обчислювальної (англ. computational) статистики [22], які дозволяють з підвищеною достовірністю визначати наявність кореляції між вибірками. Цей метод оцінки кореляцій методами Монте-Карло було модифіковано для можливості розрахунку кореляцій між добре скорельованими вибірками [5], що, як виявилось, дало

змогу не просто розраховувати ступінь кореляції між двома вибірками, а й отримати принципово новий результат: результат розрахунку кореляцій виявився чутливим до абсолютних значень елементів вибірок, тобто до інтенсивностей ліній. У статті [28] метод було розширено на аналіз даних рентгенівської дифракції, для пошуку слабких за інтенсивністю рефлексів, а в попередньому розділі успішно застосовано для ідентифікації фаз у даних рентгенівської дифракції від багатофазних матеріалів на основі пошуку кореляцій між експериментальними і тестовими вибірками даних.

У методах обчислювальної статистики [21, 22], основаних на використанні методів Монте-Карло, кореляція між двома вибірками розраховується не як аналог коефіцієнта Пірсона у вигляді числа від нуля до одиниці, а як імовірність прийняти гіпотезу про наявність чи відсутність кореляції між ними на основі аналізу розкиду величини числового показника ступеня кореляції, отриманого методами обчислювальної статистики. При використанні модифікованого методу обчислювальної статистики [5, 28] показник кореляції розраховується у відносних одиницях, бо залежить від параметрів вибірок.

2.2.4.3. Розрахунок

Якщо (на початковому етапі аналізу) не звертати увагу на необхідність врахування різноманітних поправок, то на перший погляд може здатися, що відношення інтенсивностей максимальних рефлексів може слугувати за потрібний параметр для порівняння двох різних фаз. Далі в це відношення вносяться необхідні поправки, отримані, наприклад, з вимірювання еталонних сумішей. Однак якраз через велику кількість факторів, які можуть впливати на це відношення, існуюча практика визначення співвідношення фаз стикається з великими труднощами, а аналіз цих факторів практично є обов'язковою процедурою підготовки дифракційних даних (див., наприклад, метод Рітвелда [37]). На практиці такі роботи вимагають створення еталонів, що далеко не завжди є можливим. Окремо стоїть також питання про невизначеність результату такої оцінки.

Отже, в якості параметра, який характеризує фазу, потрібно використовувати більш однозначні характеристики даних рентгенівської дифракції. Наприклад, як у даній роботі, використати описану нижче кореляційну характеристику вибірки (дифракційні дані розглядаються як упорядкована по кутах вибірка чисел). Мається на увазі, що інформація (кути чи міжплощинні відстані та відповідні їм величини інтенсивностей) про всі можливі фази уже знаходиться в базах даних (кристалографічні задачі зараз не розглядаються). На основі підходів обчислювальної статистики можна ввести деяку кореляційну характеристику, яка однозначно характеризує дані рентгенівської дифракції.

2.2.4.4. Параметр взаємної кореляції даних рентгенівської дифракції

Розглянемо табличні, з нормованою на 1 000 інтенсивністю, дані рентгенівської дифракції. Уведемо позначення: число регулярних кутів, для яких отримано дані рентгенівської дифракції, тобто розмір вибірки позначимо через N . Числа (інтенсивності для кожного кута) в експериментальній вибірці позначимо як g_i , $i = 1, 2, \dots, N$, а значення інтенсивностей у тестових даних рентгенівської дифракції з бази даних відповідно через b_i . В обчислювальній статистиці [22] ступінь кореляції між двома вибірками є пропорційною величині $S = \sum_i^N g_i, b_i$ а отже залежить від значень елементів вибірок, для різних дифракційних даних буде різним і тому в подальшому дається у відносних одиницях. Ще раз підкреслимо, що нас цікавить не просто величина S , а пошук найбільш імовірного значення S при аналізі наших вибірок, з урахуванням можливих невизначеностей (похибок) значень як b_i , так і g_i .

Легко перевіряється (див. також [5, 28]), що величина S буде максимальною при порівнянні ідентичних, тобто повністю скорельованих вибірок. Ступінь самоскорельованості S_{max} для даної вибірки (фази) буде різним для різних фаз. І хоча в запропонованому підході результат оцінки ступеня кореляції при таких розрахунках отримується у відносних одиницях, проте для деякого конкретного алгоритму розрахунку результат буде однозначним і може порівнюватися для різних фаз. Цей момент проілюстровано на рис. 2.2.25 та 2.2.26 для взятих з бази даних Crystallography Open Database (COD) [34] дифрактограм нантокіту і еріохальциту. Ці дані рентгенівської дифракції взято з бази даних, де прийнято нормувати інтенсивність максимальної лінії на 1000.

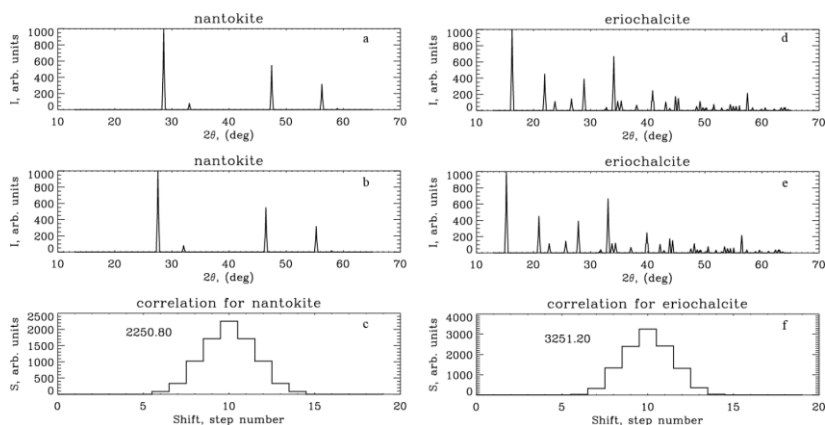


Рис. 2.2.25. Результат “сканування” дифрактограм нантокіту та еріохальциту такими ж дифрактограмами, інтенсивність яких нормовано на 1000. Зліва нантокіт, справа – еріохальцит. Результуючі гістограми ступеня збігу мають очікуваний пік на 10-му кроці. Значення S_{max} винесено на рисунок.

Звернемо увагу, що хоча в табличних даних інтенсивності основних ліній для різних сполук однакові ($= 1000$), величини S_{max} у них різні. Саме цей факт і є основою для подальших розрахунків. Будь-які зміни інтенсивностей ліній експериментальної дифрактограми деякої сполуки призведуть у результаті до зменшення величини S порівняно з S_{max} . Демонстрацію цього наведено на рис. 2.2.26, де показано результат розрахунку ступеня кореляції S між двома вибірками для випадку, коли обидві вибірки однакові за структурою (дифрактограма нантокіту з бази даних COD 96-901-3925 [34]), проте мають різні інтенсивності. На рис. 2.2.26, а наведено дифрактограму нантокіту, яку ми приймаємо за експериментальну. Інтенсивність ліній на ній вдесятеро менша, ніж інтенсивність ліній цієї ж дифрактограми, нормованої на 1 000, яку наведено на рис. 2.2.26, б і яку ми розглядаємо як тестову. Результат розрахунку кореляції показано на рис. 2.2.26, с. Максимальне значення S_{max} для тестового розрахунку для нантокіту, як вказано вище, дорівнює 223,9. У наведеному на рис. 2.2.26 випадку, при зменшенні інтенсивності всіх ліній в експериментальній дифрактограмі вдесятеро, значення S , тобто ступінь кореляції між цими дифрактограмами, також зменшилося вдесятеро і дорівнює 22,39. Цей результат, тобто отримане зменшення S порівняно з S_{max} , відкриває шлях до оцінки вмісту деякої кристалічної фази в багатофазних матеріалах.

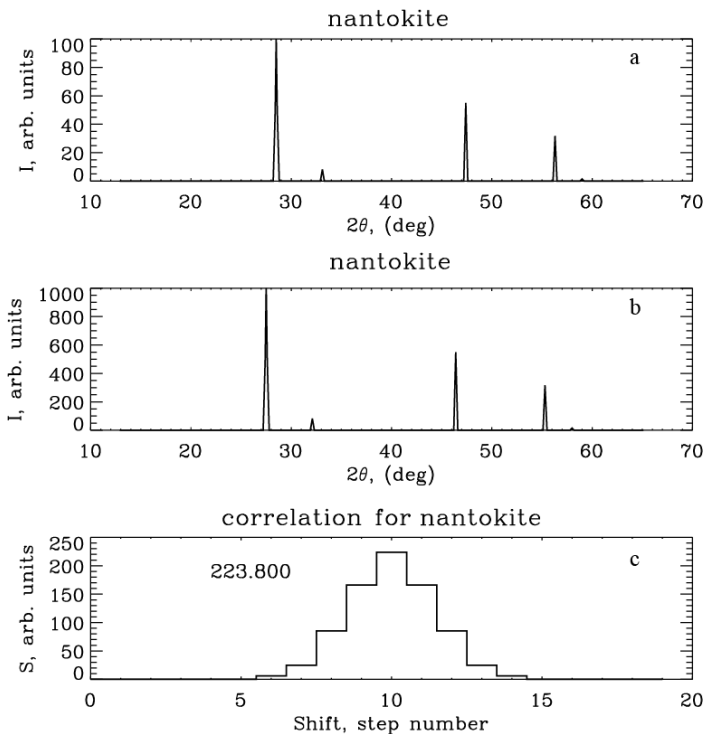


Рис. 2.2.26. Результат сканування модельної дифрактограми нантокіту, інтенсивність якої у 10 разів менша, еталонною дифрактограмою нантокіту. Відповідно зменшилася і величина S .

2.2.4.5. Алгоритм оцінки вмісту кристалічних фаз

Спочатку сформулюємо умови задачі і припущення, за яких проводяться розрахунки. Уявімо схему вимірювань (рис. 2.2.27), коли зразок (кювета з порошком) розділена на дві рівні частини.

Кожна з них є окремою фазою. Це наближення класичної ситуації аналізу двофазної системи з однаковими коефіцієнтами поглинання. У такій схемі експерименту немає взаємного впливу фаз. Відповідно загальна дифрактограма буде сумою дифрактограм від окремих фаз у рівному співвідношенні їхнього вмісту.

1. У роботі розглядається можливість порівнювати між собою лише відносний вміст ідентифікованих фаз. Тобто наявність сигналів, схованих у шумах, чи внесок аморфної фази не враховуються.

2. Якщо фази знаходяться в рівних співвідношеннях, то за всіх умов співвідношення інтенсивностей чи площ ліній повинно відображати це співвідношення.

3. Поправки типу самопоглинання, впливу текстури тощо на цьому етапі не розглядаються, оскільки вони беруться не з результатів конкретного вимірювання (конкретної дифрактограми), а з інших міркувань (з аналізу калібрувальних кривих, теоретичних оцінок тощо).

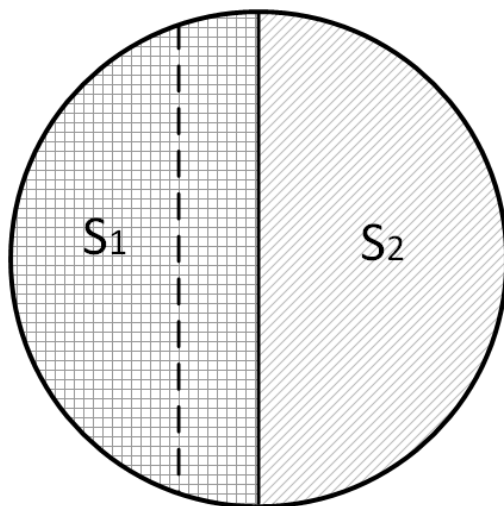


Рис. 2.2.27. Схема вимірювань модельного еталону.
Пунктир – поділ кювети у відношенні 1:2.

4. Сумарний відносний вміст всіх ідентифікованих фаз повинен дорівнювати одиниці.

5. Потрібно враховувати відношення сигнал/шум: у теорії сигнал може бути як завгодно малим, а в реальності існує межа – мінімально допустимий сигнал/шум повинен бути більше одиниці.

6. Вважається [38, 39], що інтенсивність рефлексів від даної фази в ідеальному варіанті є пропорційною вмісту даної фази. Якщо є дві фази з вмістом по 50 % (див. рис. 2.2.27), то очевидно, що співвідношення експериментальних інтенсивностей повинно відображати це 50 % співвідношення вмісту фаз. Перемішування зразка, якщо коефіцієнти поглинання є рівними, не змінює ситуацію.

7. У роботі пропонується для характеристики окремих дифрактограм використовувати таку величину, як сума інтенсивностей всіх ліній їхньої еталонної дифрактограми $S_{1,2}^{(0)} = \sum_i^n I_i$, де n – число ліній еталонної дифрактограми. Очевидно, що ця величина, як і інтенсивність основної лінії, є пропорційною вмісту фази. Якщо, наприклад, на рис. 2.2.27 перемістити лінію розділу так, щоб співвідношення площ стало не 1 : 1, а 1 : 2 (пунктир), це повинно однозначно відобразитись у співвідношенні величин S_{expr} для фаз розглянутого зразка: $S_{expr1}/S_{expr2} = \text{const}$ повинно змінитися на $S_{expr1}/S_{expr2} = \text{const}*(1:2)$.

8. У попередньому розділі для характеристики реальної дифрактограми була введена величина S_{max} , яка для даної сполуки (даної фази) є константою. Якщо розглянути дифрактограму, отриману в схемі, показаній на рис. 2.2.27, то при рівному співвідношенні фаз $S_1/S_2 = 1$, а $S_{max1}/S_{max2} = \text{const}$ і для розглянутих фаз нантокіту та еріокальциту $\text{const} = S_{max1}/S_{max2} = 2239/3229 \cong 0,69$.

9. Відповідно, якщо співвідношення вмісту фаз на рис. 2.2.27 замінити на показане пунктиром співвідношення 1:2, то сигнал від другої фази буде удвічі більшим ($S_1/S_2 = 1/2$), а $S_{expr1}/S_{expr2} = \text{const}*(1/2) \cong 0,69 * 0,5$.

2.2.4.6. Формули для розрахунку вмісту фаз у зразках

Вище (див. рис. 2.2.26) було продемонстровано, що запропонований метод дозволяє розрізнити еталонні вибірки для однакових фаз з різним вмістом. Тепер перевіримо можливості методу на модельних сумішах фаз. Візьмемо табличні дифрактограми нантокіту та еріокальциту, нормовані на 1000. Створимо модельну дифрактограму суміші цих фаз, об'єднавши їх в одну. Нагадаємо, що проблему ідентифікації окремих фаз у таких сумішах статистичними методами було розв'язано в [28]. Зараз же нас цікавить можливість кількісного визначення відносного вмісту в ній цих двох фаз, з яких вона складається, який ми збираємося оцінювати за величиною S .

Розглянемо модельну ситуацію суміші двох фаз у рівних концентраціях, коли інтенсивності найбільш інтенсивних ліній кожної фази нормовані на 1 000, як це показано на рис. 2.2.25. На верхніх графіках рис. 2.2.28 показана сумарна дифрактограма суміші цих двох фаз, у середині – таблична дифрактограма фази, яка розшукується в суміші (нантокіт чи еріокальцит), а внизу – результат пошуку відповідної фази.

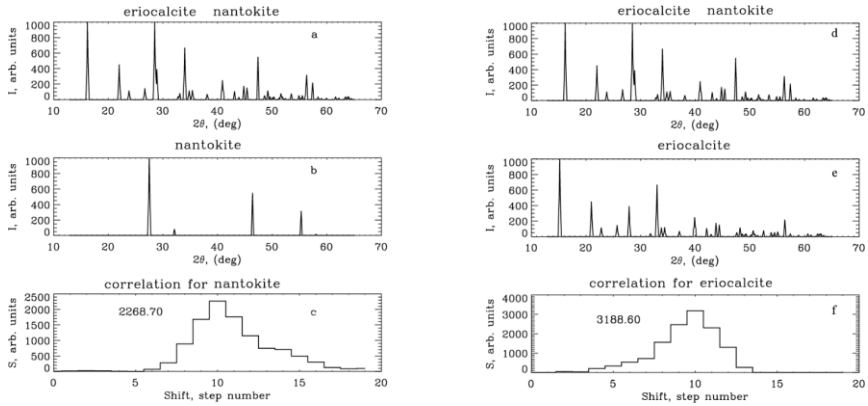


Рис. 2.2.28. Результат аналізу дифрактограми модельної суміші двох сполук (див. текст) у співвідношенні 50 – 50 %. Зліва – пошук фази нантокіту, справа – пошук еріохальциту. а, d – модельна (таблична) дифрактограма суміші; b, e – дифрактограма фази, яка розшукується; c, f – гістограма ступеня кореляції експериментальної фази з табличною.

Оскільки максимальні лінії обох фаз однакові і дорівнюють 1000, то величини S_{expr}/S_{max} для кожної фази будуть дорівнювати одиниці (див. рис. 2.2.25). У даному випадку сумарна величина відносних вмістів для двох ($k = 2$) фаз $S_{calc} = \sum_i^k \frac{S_{expr}^{(i)}}{S_{max}^{(i)}}$. Але в модель закладено, що ці дві фази знаходяться в рівних пропорціях, вміст кожної фази дорівнює 1/2. Тому правильне співвідношення для окремих фаз отримується як $(S_{expr}^{(i)}/S_{max}^{(i)})/S_{calc}$.

Отже, при довільних (і невідомих) співвідношеннях між фазами $\sum_i^k \frac{S_{expr}^{(i)}}{S_{max}^{(i)}} = S_{calc}$, а для кожної фази її внесок становить $(S_{expr}^{(i)}/S_{max}^{(i)})/S_{calc}$.

Приклад такого аналізу для модельних дифрактограм суміші двох фаз з однаковим вмістом (див. рис. 2.2.28) дано в табл. 2.2.1.

Таблиця 2.2.1. Результати пошуку співвідношення фаз у модельних зразках

Фаза	S_{max}	S_{expr}	S_{expr}/S_{max}	S_{calc}	$(S_{expr}^{(i)}/S_{max}^{(i)})/S_{calc}$
Нантокіт	2250,8	2268,7	1,008	1,989	0,5068 (50,68 %)
Еріохальцит	3251,2	3188,6	0,981	1,989	0,4932 (49,32 %)

З останньої колонки видно, що точність розрахунків є не гіршою 0,5 % (у модель було закладено співвідношення фаз 50 : 50 %).

Отже, величини S_{expr}/S_{max} дають відносне значення вмісту для двох вибраних фаз. Зразок може складатися з багатьох фаз, які можна

ідентифікувати, а також з неідентифікованих кристалічних та/або аморфних фаз.

Загальний вміст всіх кристалічних та аморфних фаз визначається за формулою

$$S = \sum_i^N \frac{S_{expr}^{(i)}}{S_{max}^{(i)}} + B \equiv 1,$$

де доданок B містить внески від неідентифікованих кристалічних фаз, внесок гало від аморфної фази тощо.

Однак проведений аналіз є коректним для будь-якої пари фаз і не залежить від вмісту інших. Це можна легко зрозуміти, уявивши, що зображення фаз на рис. 2.2.26 є лише частиною більшого зразка. Величина B оцінюється вже з інших міркувань, поза межами даного підходу. У рамках розробленого методу величина B не враховується, а оцінюється лише співвідношення вмісту ідентифікованих фаз.

Якщо число ідентифікованих фаз буде більше двох, процедура аналізу загальної дифрактограми не змінюється: за описаним алгоритмом розраховуються співвідношення вмістів будь-яких двох фаз.

2.2.4.7. Тестування методу

Як уже вказувалося, працездатність підходу до задачі ідентифікації фаз шляхом розрахунку кореляцій на основі методів Монте-Карло обґрунтовано і продемонстровано в [28]. Вище показано, що метод працює у випадку “ідеальних” табличних дифрактограм. Залишилося показати можливості цього методу для оцінки вмісту фаз у багатофазних матеріалах на реальних даних. Оскільки вважається, що це можна зробити, лише маючи калібровочну криву, отриману на основі еталонів, розроблений метод оцінки вмісту фаз треба перевіряти шляхом аналізу дифрактограм каліброваних сумішей. Для цього було використано еталонну суміш нантокіту і еріокальциту у пропорціях 70 % : 30 %. Графічні результати аналізу наведено на рис. 2.2.29.

Розрахунки для базових дифрактограм для оцінки S_{max} див. на рис. 2.2.25. Результати розрахунків зведено в табл. 2.2.2.

Таблиця. 2.2.2. Результати пошуку співвідношення фаз в експериментальних зразках

Фаза (70 : 30 %)	S_{max}	S_{expr}	S_{expr}/S_{max}	S_{calc}	$(S_{expr}^{(i)}/S_{max}^{(i)})/S_{calc}$
Нантокіт	2250,8	902,8	0,401	0,65	0,617 (61,7 %)
Еріокальцит	3251,2	821,8	0,253	0,65	0,389 (38,9 %)

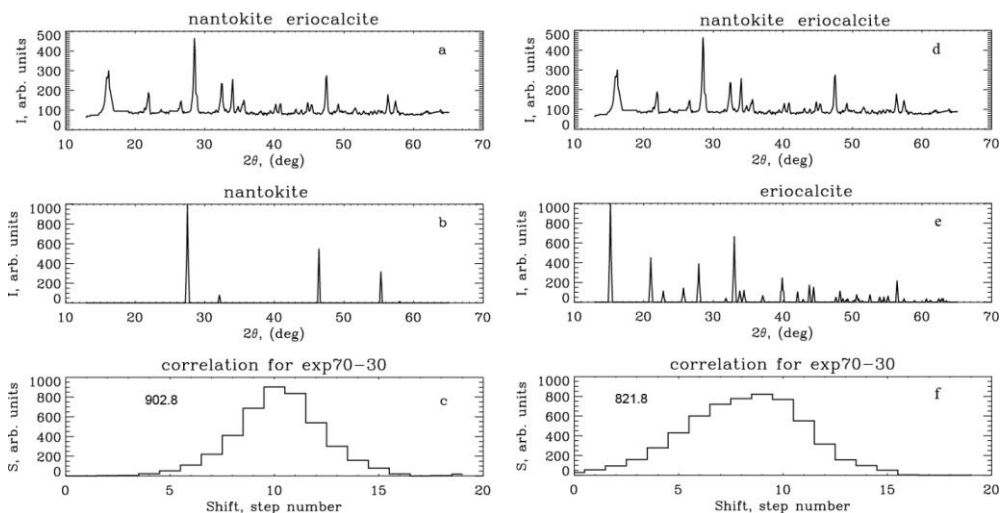


Рис. 2.2.29. Результат аналізу фазового вмісту дифрактограми суміші двох сполук (див. текст) у співвідношенні 70 – 30 %. Зліва – пошук фази нантокіту, справа – пошук еріокальциту. a, d - експериментальна дифрактограма суміші; b, e - дифрактограма фази, яка розшукується; c, f - гістограма ступеня кореляції даних рентгенівської дифракції від експериментальної фази з даними відповідної фази з бази даних COD.

Отриманий результат – вміст натокіту складає 60 %, а вміст еріокальциту 40 % (замість очікуваних 70 : 30 %). Розбіжність можна пояснити або неточністю приготування сумішей, або ж тим, що це є реальна чутливість методу при невідомих особливостях експерименту. Для остаточної відповіді потрібен аналіз метрологічних еталонів.

2.2.4.8. Оцінка невизначеності методу

Оцінимо невизначеність алгоритму розрахунків i , відповідно, власне методу. Необхідність такої оцінки випливає з того, що в реальності для конкретної дифрактограми можна лише, виходячи з попереднього досвіду, приписати точність, з якою визначається інтенсивність основної лінії. Що стосується слабких ліній, то для них на основі окремої дифрактограми скільки-небудь обґрунтовано оцінити невизначеність оцінки їхньої інтенсивності практично неможливо.

Невизначеність результату статистичних розрахунків видно з табл. 2.2.1, відмінність результату від закладеного в модель результату приблизно 0,5 %.

Приймемо, що невизначеність оцінки інтенсивності всіх рефлексів однакова і довільно вибрана рівною $\pm 5\%$. Тобто дифрактограма є вибіркою чисел, експериментальні величини яких знаходяться в інтервалі $\pm 5\%$ навколо деякого невідомого “точного” значення. У цій ситуації за міру невизначеності результату оцінки вмісту фази пропонується оцінити діапазон можливих значень S_{expr} , які змінюються при зміні інтенсивностей рефлексів. Це можна

зробити на основі використання методів Монте-Карло. А саме, до кожної лінії табличної дифрактограми випадковим чином додається значення $\pm\Delta$ з діапазону $\pm 5\%$. Для такої модифікованої вибірки розраховується S_{expr} . Ця процедура повторюється кілька тисяч разів. Далі можна побудувати гістограму розподілу отриманих значень S_{expr} і за правильний результат прийняти найбільш імовірне значення. Розкид навколо нього (ширина гістограми) дає невизначеність отриманої оцінки. Виявилось, що ширина її на половині висоти також дорівнює приблизно $\pm 5\%$, проте слід зауважити, що це вже максимальна величина, яка враховує змінність всіх рефлексів дифрактограми.

Невизначеність за рахунок таких ефектів, як наявність поглинання, текстури та інших експериментальних факторів, не залежить від описаного методу і може бути взята з літератури.

Застереження. Продемонстрована працездатність запропонованого методу аналізу дифрактограм на основі вищенаведеного алгоритму не повинна створювати ілюзію стовідсоткової достовірності отриманого результату. Розробка описаного методу кореляційного аналізу була стимульована проблемами ідентифікації та оцінки вмісту фаз у зразках з їхнім малим вмістом. У цьому випадку різко зростають відносні похибки визначення інтенсивностей ліній аж до можливої помилкової ідентифікації фаз. Відповідно зростає ймовірність помилкового визначення вмісту даної фази через те, що невизначеність оцінки $S_{експ}$ визначається невизначеністю оцінки інтенсивності відповідних ліній. Оскільки оцінки вмісту обов'язково повинні супроводжуватись оцінками їхніх невизначеностей, ми рекомендуємо, як обов'язковий елемент алгоритму по п. 1, до початку розрахунків перевіряти можливий внесок добре відомих практикам факторів впливу (накладання ліній, самопоглинання тощо) і при описі результатів оцінки вмісту окремих фаз обов'язково, як цього вимагає [12], вказувати прямим текстом, внесок яких факторів було перевірено.

Надзвичайно важливого значення у цьому методі має ефект спотворення інтенсивностей ліній за рахунок накладання ліній від різних фаз, внесок якого може набагато перевищувати прийняті 5% . Тому розрахунки рекомендується проводити шляхом поступових ітерацій. Спочатку визначати наявність тієї чи іншої фази; потім перевірити, що співвідношення інтенсивностей (особливо достатньо інтенсивних ліній) відповідає базовій дифрактограмі. При необхідності потрібно внести відповідні поправки (шляхом віднімання внесків ліній інших фаз), повторити розрахунки після внесення поправок і лише потім приймати рішення про правильність оцінки S_{expr} .

2.2.4.9. Висновки до методу оцінки вмісту кристалічних фаз у багатофазних зразках

1. Запропоновано метод оцінки відносного вмісту кристалічних фаз у даних рентгенівської дифракції від багатофазних матеріалів. Ідея методу

полягає в тому, що можна ввести параметр, який однозначно характеризує дифрактограму.

2. Задачу ідентифікації фаз та оцінки їхнього відносного вмісту можна звести до задачі ймовірнісного аналізу інтенсивностей рефлексів.

3. Метод базується на розрахунках ступеня кореляції між експериментальною дифрактограмою та дифрактограмою шуканої фази з баз даних. Уведено параметр, оснований на розрахунку парних добутків дифрактограм, за яким їх можна порівнювати. Принциповим моментом є те, що розрахунки виконуються з використанням комп'ютерної статистики (permutation test).

4. Розрахунки робляться на основі деяких припущень, наприклад про відсутність внеску ефекту самопоглинання, і тому результат розрахунків за запропонованим методом вимагає ретельного аналізу правильності закладених наближень. Похибка статистичних розрахунків є $\pm 0,5\%$. Але правильність оцінки відносного вмісту окремих фаз залежить від правильності оцінки інтенсивностей ліній дифрактограми, внесення необхідних поправок у конкретних випадках тощо. Крім того, це повинен бути багатостадійний процес, бо спочатку треба визначитися з гіпотетичним фазовим складом, створити відповідні базові дифрактограми тощо. І лише потім розрахувати відносний вміст фаз за рекомендаціями підрозділу 2.2.4.6 «Формули для розрахунку вмісту фаз у зразках».

5. Менше з тим показано, що запропонований метод, хоча б грубо в межах указаної невизначеності, дозволяє ідентифікувати фази і визначати їхній відносний вміст, відмовившись від використання еталонів (за умови правильності закладених наближень: рівності коефіцієнтів поглинання, відсутності текстури тощо).

6. Запропонований метод можна використовувати також як метрологічний прийом для перевірки якості еталонів.

2.2.5. Резюме до розділу про кореляційний аналіз

На основі методів Монте-Карло запропоновано процедуру по виявленню ліній у гамма- чи інших спектрах у ситуаціях, коли відношення сигнал/шум для таких ліній дуже низьке, порядку одиниці. Метод виявився набагато чутливішим порівняно із стандартними процедурами по пошуку ліній у спектрах. У результаті вдалося не тільки виявляти лінії, коли їхня інтенсивність настільки низька, що їх узагалі не видно на фоні шумів, а й з більшою точністю оцінювати їхню інтенсивність. Усе це дозволило робити кількісний аналіз уже на рівні, коли стандартні процедури ISO дозволяють робити лише якісні висновки. Тобто робити метрологічний аналіз уже на рівні МДА. У роботі [37] показано, що створений у рамках цього напрямку кореляційний метод аналізу даних рентгенівської дифракції дозволяє впевнено ідентифікувати відображення при рівні співвідношення сигнал/шум до 0,4. Водночас при рівнях цього співвідношення від 0,4 до 0,2 необхідний

додатковий аналіз відображень для встановлення достовірності отриманих даних про наявність відображення. У практиці РФА існують емпіричні критерії визначення кристалічних фаз із даних рентгенівської дифракції, які, фактично, можна використати як алгоритм для автоматизованого методу ідентифікації фаз.

1. Від 3 до 6 ліній на дифрактограмі повинні збігатись з лініями ймовірної кристалічної фази, взятої з відповідних баз даних, наявність якої у зразку ми намагаємось встановити.

2. Ці 3 – 6 ліній повинні бути найбільші за інтенсивністю серед інших можливих ліній цієї ж фази в даних для досліджуваного зразка.

3. Ці 3 – 6 ліній повинні бути розташовані за інтенсивністю в тій самій послідовності, в якій вони розташовані для цієї ж фази у відповідній базі даних.

Як бачимо, всі вони базуються на візуальному порівнянні експериментальних дифрактограм з інформацією, наведеною в базах даних. Проте в дійсності цей підхід дає збої при наявності більше 4 – 6 фаз, та відповідно 15 – 35 ліній. Важко розібратися - конкретна середня лінія, це одна з 6 ліній фази з найбільшим вмістом, чи це основна лінія іншої фази, з меншим вмістом у матеріалі. Це ускладнюється, коли положення ліній (2θ) має експериментальну похибку.

Ідентифікація фаз ускладнюється у випадках, коли накопичення похибок у визначенні інтенсивності рефлексів, призводить до того, що інтенсивності середніх за інтенсивностями рефлексів перевищує інтенсивність першого за інтенсивністю рефлексу. Це призводить до помилки в визначенні певної кристалічної фази. Зникнення окремих низькоінтенсивних рефлексів на фоні шумів також може призводити до зменшення ймовірності ідентифікації конкретної кристалічної фази.

СПИСОК ЛІТЕРАТУРИ

1. *Fievet B. , Vedova C. D.* Dealing with non-detect values in time-series measurements of radionuclide concentration in the marine environment // J. Environ. Radioact. – 2010. – Vol. 101. – Issue 1. – P. 1 – 7.

2. *Dennis Helseland, Lopaka (Rob) Lee.* Analysis of Environmental Data With Nondetects. Statistical Methods for Censored Environmental Data // Continuing Education Workshop at the Joint Statistical Meetings. American Statistical Association. Seattle, Washington, August 2006.

3. *Gabielkov S.V., Zhyganiuk I.V., Kudlai V.G., Nosovsky.V., Parkhomchuk P.E., Chikolovets S.O., Shcherbin V.M.* Phase composition of brown ceramics of lava-like fuel-containing materials of the object "Shelter" ChNPP // Nucl. Phys. At. Energy 20, 388 (2019).

4. *Агеев В. А., Вурічек С. Д., Лашко А. П. та ін.* Нижні порогові рівні детектування ^{241}Am напівпровідниковими гамма-спектрометрами в пробах, забруднених внаслідок аварії на ЧАЕС – НАН України – Чорнобиль, 1998. – 26 с. – (Препр./ НАН України. МНТЦ "Укриття", 98-2).

5. *Скорбун А. Д., Панасюк М. І.* Виділення слабких ліній у гамма-спектрах // Проблеми безпеки атомних електростанцій і Чорнобиля. – 2008. – Вип. 9. – С. 125 – 130.
6. *Спектрометрическая система Genie 2000, версия 2.1.* Руководство пользователя / Фирма Canberra Industries, 2002.
7. *Currie L. A.* Limits for qualitative detection and quantitative determination / Application to radiochemistry // Analytical Chemistry. – 1968. – Vol. 40. – P. 586 – 593.
8. *Каденко І. М., Плюйко В. А.* Основи методів оцінок статистичних даних і функцій їх розподілу. – К.: Київ. ун-т, 2003. – 145 с.
9. *Определение активности естественных радионуклидов в объектах окружающей среды.* (Методическое пособие). – К.: Научно-производственное предприятие "АКП", 1992. – 59 с.
10. *Исаев А. Г., Бабенко В. В., Казимиров А. С. и др.* Минимальная детектируемая активность. Основные понятия и определения // Проблеми безпеки атомних електростанцій і Чорнобиля. – 2010. – Вип. 13. – С. 103 – 110.
11. *ISO 11929-3:2000.* Determination of the detection limit and decision threshold for ionizing radiation measurements. Part 3: Fundamentals and application to counting measurements by high resolution gamma spectrometry, without the influence of sample treatment.
12. *Guide to expression of uncertainty in measurement / GUM:1993, First edition, 1993* ISO. – Geneva, 1995.
13. *Luca A.* Some considerations about MDA and ISO 11929-3(2000) / ICRM Gamma Spectrometry Working Group. Workshop. Paris, Laboratoire National Henri Becquerel, 23 – 24 February 2009. http://www.nucleide.org/ICRM_GSWG.htm/
14. *Luca A. and Morel J.* Influence of the background approximation methods on the analysis of γ -ray spectra // Applied Radiation and Isotopes, 2004. – Vol. 60. – Issues 2 – 4. – P. 233 – 237.
15. *Yoram NIR-EL.* Minimum detectable activity in gamma-ray spectrometry: statistical properties and limits of applicability // Health physics. – 2001. – Vol. 80. – No. 2. – P. S22 – S25.
16. *Рытов С. М., Кравцов Ю. А., Татарский В. И.* Введение в статистическую радиофизику. Часть 2. Случайные поля. – М.: Наука, 1978. – 464 с.
17. *Ланге Ф.* Корреляционная электроника. Основы и применение корреляционного анализа в современной технике связи, измерений и регулирования. – Л.: Судпромгиз, 1963. – 447 с.
18. *Разработка методики гамма-спектрометрического анализа технологических сред АЭС: (Отчет) / Обнинский филиал МИФИ.* – 1983. – 62 с.
19. *Руководство по эксплуатации радиационного монитора для контроля автотранспорта с РАО РМАТ-01-ФДП "Фаворит плюс".* – К., 2007.
20. *Соловьева С. Л.* Программный комплекс для моделирования гамма-спектрометрических экспериментов и методика обработки гамма-спектров / Автореф. дис. ... канд. физ.-мат. наук. – М., 2008.
21. *Simon J. L.* Resampling: The new statistics. – Duxbury Press, 1993. <http://www.resample.com/content/text/index.shtml>
22. *Hesterberg T., Moore D. S., Monaghan S. et al.* Bootstrap methods and permutation tests. – N.Y.: W. H. Freeman. – 2005. – P. 14.2 – 14.69.
23. *Davison A. C., Hinkley D. V.* Bootstrap methods and their application. – Cambridge: University Press, 1997. – 592 p.
24. *Худсон Д.* Статистика для физиков. – М.: Мир, 1970. – 296 с.

25. Tanizaki H. On Small Sample Properties of Permutation Tests: Independence Test between Two Samples // International Journal of Pure and Applied Mathematics. – 2004. – Vol. 13. – No. 2. – P. 235 – 243.
26. Lee M. H., Jung E. C., Song K. S. Determination of a Minimum Detectable Activity through a Measurement of Pu Isotopes in Environmental Samples // Bull. Korean Chem. Soc. – 2008. - Vol. 29, No. 6. – P. 1162 – 1166.
27. Jarman K. D. et al. Indirect estimation of radioactivity in containerized cargo //– Radiation Measurements, 2011. – Vol. 46, Issue 1. – P. 10-20.
28. Skorburn A.D., Gabelkov S.V., Zhyganiuk I.V., Kudlai V.G., Parkhomchuk P.E., Chikolovets S.A.. Method of X-Ray Diffraction Data Processing for Multiphase Materials with Low Phase Contents // Ukr. J. Phys. – 2019. – Vol.64.–P.862.
29. Brandenburg K., Putz H. Program for Phase Identification from Powder Diffraction "Match!", Crystal Impact, Bonn, Germany, 2021.
30. Dinnebier R.E., Billinge S.J.L. Overview and Principles of Powder Diffraction, in International Tables for Crystallography (2019), Vol. H, ch. 1.1, pp. 2–23.
31. Gilmore C.J., Kaduk J.A., Schenk H. Survey of computer programs for powder diffraction, in International Tables for Crystallography (2019), Vol. H, ch. 6.1, pp. 698–715.
32. Toby B.H. Rietveld refinement, in International Tables for Crystallography (2019), Vol. H, Ch. 4.7, pp. 465–472.
33. Pollard J.H. A Handbook of Numerical and Statistical Techniques (Cambridge University Press, 1977).
34. Cédric Duée et al. Impact of heterogeneities and surface roughness on pXRF, pIR, XRD and Raman analyses: Challenges for on-line, real-time combined mineralogical and chemical analyses on drill cores and implication for "high speed" Ni-laterite exploration. Journal of Geochemical Exploration V. 198, 1 (2019).
35. Rozhenko N.M., Grigor'ev O.N., and Kartuzov V.V. A Method of Analysis of the Shapes of X-Ray Diffraction Lines Not Requiring a Transition to the Space of an Object, Reports Natl. Acad. Sci. Ukr. 1, 47 (2019).
36. Gabelkov S.V., Zhyganiuk I.V., Dolin V.V., Skorburn A.D., Kudlai V.G., Parkhomchuk P.E., Slyvinsky V.M., Chikolovets S.O. Phase composition of lava-like fuel-containing materials of Unit 4 of the Chornobyl NPP. Black ceramics Journal of Nuclear Materials, 579, 154392 (2023). DOI: 10.1016/j.jnucmat.2023.154392
37. Toby B. H. Rietveld refinement. *International Tables for Crystallography* (2019). Vol. H, Chapter 4.7, pp. 465–472.
38. Gabelkov S. V., Zhyganiuk I. V., Skorburn A. D. et al. Possibilities of the X-ray diffraction data processing method for detecting reflections with intensity below the background noise component. *Powder Diffraction.*, 2024, Vol. 39, No. 3, September, pp. 132 - 143. 2024:1-12. doi:10.1017/S0885715624000241.
39. Mudryi S. I., Kulik Yu. O., Yakimovich A. S. X-ray structural analysis in material science. Textbook. Lviv Franko University, 2017, 226 p. (in Ukrainian).
40. Madsen I. C. and Scarlett N. V. Y. Powder Diffraction: Theory and Practice, ed. R. E. Dinnebier and S. J. L. Billinge, The Royal Society of Chemistry, 2008, ch. 11, pp. 298–331.

2.3. Залежність невизначеності результату від невизначеності вхідних параметрів на прикладі оцінки вмісту контейнера з РАВ

У даному прикладі розглянуто застосування методів Монте-Карло для деякої практичної задачі. Але зараз нас повинні цікавити не самі задачі, а пов'язане з їхнім розв'язуванням запитання: наскільки отримана невизначеність розв'язку могла бути результатом такого фактора, як неточність вихідних даних, тобто наскільки отриманий результат може бути випадковим. Прийом, використаний для отримання відповіді на поставлене запитання, якщо не зважати на тонкощі, у загальному вигляді так: давайте зробимо розрахунки для всіх можливих варіантів і порівняємо результати. Мається на увазі, що в даному випадку існують способи розрахунку потрібного результату (існує відповідна формула), або ж існує функціональна залежність результату від вхідних параметрів.

У випадку, коли формула функціональної залежності відома, існує і формула оцінки похибки результату, яка вимагає знання частинних похідних від цієї відомої функції. Однак при цьому залишається проблема оцінки невизначеності результату, пов'язана з неточним знанням величини вхідних параметрів. До такого випадку належить розглянута нижче задача оцінки точності визначення активності ТРВ гамма-спектрометрами з теоретичним розрахунком ефективності [1]. Тому в цьому розділі проаналізовано джерела виникнення похибок і запропоновано спосіб оцінки похибки, що виникає при неточному знанні характеристик об'єкта. Виконано розрахунки похибки для випадків, коли параметри задачі задано інтервалами чи у вигляді нечітких величин. Причому проблема визначення такої похибки не залежить від її абсолютної величини: навіть коли її можна зменшити шляхом звуження інтервалів невизначеності параметрів, але повинен існувати спосіб її оцінки.

Дистанційні методики визначення активності ТРВ базуються на однаковому принципі: встановлюється коефіцієнт K перерахунку ПЕД, виміряної від контейнера з ТРВ, в активність (тобто кількість радіонуклідів), яка міститься в цьому контейнері, а саме:

$$\text{Активність} = \text{ПЕД} \cdot K. \quad (2.3.1)$$

Тобто вимірюється активність лише гамма-випромінюючих нуклідів. Такий принцип діє бездоганно, коли ТРВ є однорідним матеріалом відомого складу з однорідним розподілом активності в його об'ємі й ці характеристики не змінюються від контейнера до контейнера. Однак у реальності контейнер може бути заповнений досить різними матеріалами – від будівельного сміття до радіоактивно забрудненого металу, що не дає можливості точно встановлювати параметри ТРВ. Це об'єктивні труднощі і вихід з них знаходять шляхом введення різних категорій речовин, які складають ТРВ, та різного вмісту в них радіонуклідів. Деякі діючі інструкції передбачають, що оператор повинен визначити склад ТРВ і залежно від цього вибрати дані для розрахунку

(пропонуються два варіанти речовинного складу - бетон, залізо). І немає ніяких роз'яснень щодо пов'язаних з таким моделюванням можливих похибок визначення коефіцієнтів K , які для кожної з указаних ситуацій різні й залежать від точності визначення густини, складу радіонуклідів тощо.

Такий метод визначення характеристик ТРВ є відомим [2], як відомі і його недоліки: при визначенні активності реальних ТРВ похибки залежать від точності виконання закладених у розрахунки припущень. Наприклад, у роботах [3, 4] можлива похибка подібних методів взагалі оцінюється в 100 %, оскільки на практиці в контейнер може бути завантажено що завгодно.

На сьогодні найбільш розвинутим підходом для цієї задачі є використання сучасних спектрометрів з теоретичним розрахунком їхньої ефективності. Наявність гамма-спектрометричних систем із можливістю теоретично визначати ефективність спектрометра для конкретного випадку вимірювань, а також можливість оцінки похибки, яку дає ця операція, кардинально змінили рівень достовірності при оцінці характеристик ТРВ. Основною вимогою при цьому є те, що ТРВ при вимірюванні активності повинні знаходитись у контейнерах, тобто мати визначені розміри.

У нашому випадку коефіцієнт K , який пов'язує показання приладу з активністю вимірюваного об'єкта, і є ефективністю спектрометра. Ця величина залежить від геометрії вимірювань: форми об'єкта, відстані до нього тощо. Наприклад, ясно, що виміряна для даного об'єкта (контейнера) величина ПЕД у формулі (2.3.1) буде різною при вимірюваннях на різній відстані від нього. Сучасні гамма-спектрометричні комплекси дають можливість теоретичного розрахунку ефективності для різних геометрій вимірювання.

Ще раз підкреслимо, що цей коефіцієнт (ефективність) залежить від геометрії вимірювань (поле від циліндричного та прямокутного контейнерів буде різним), конструкції детектора тощо, тому всі ці параметри повинні бути строго визначені для кожної задачі: не можна визначити активність, яка міститься в «чорному ящику», шляхом вимірювання створюваного ним гамма-поля. Проте це можна зробити при наявності певної інформації про його вміст.

2.3.1. Методика розрахунків

Розглянемо докладніше складові для розв'язання задачі.

Геометрія вимірювань. Загальноприйнятий на сьогоднішній день підхід ґрунтується на тому, що більшість реальних ситуацій можна звести до обмеженого числа геометрій (куля, циліндр, паралелепіпед тощо), для яких можна написати точні формули. Наприклад, збільшуючи відстань між вимірювальним приладом і контейнером, можна одержати геометрію точкового джерела. А контейнер, заповнений РАВ, легко зводиться до геометрії паралелепіпеда чи циліндра.

Схему розгляду задачі показано на рис. 2.3.1, де показано, що заповнення матеріалом у контейнері не описується простою формою. Із рисунка випливає,

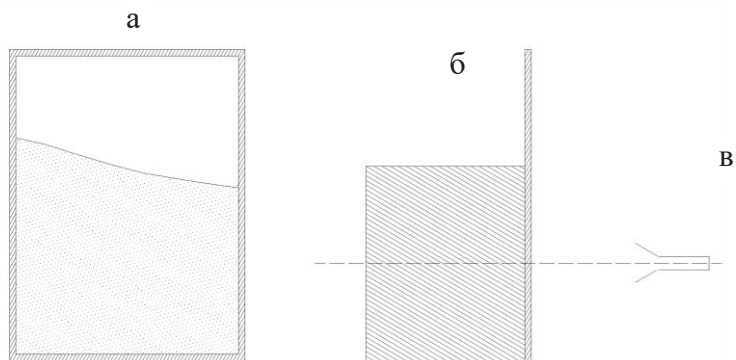


Рис. 2.3.1. Схема вимірювань для контейнера:
 а – частково заповнений контейнер;
 б – модель контейнера для розрахунку ефективності; в – детектор.

в тому числі, що помилка у визначенні форми заповнення контейнера буде зменшуватися із збільшенням відстані між контейнером і детектором.

Матеріал заповнення контейнера. Другим параметром, який не піддається точному визначенню, є коефіцієнт поглинання в матеріалі, яким заповнено контейнер. Підхід фірми MicroShield (зараз такий підхід є загальноприйнятим), наприклад, у цьому випадку полягає в тому, що потрібно визначити відносний внесок кожного хімічного елемента в загальний об'єм речовини в контейнері, а далі, відповідно до одержаного ефективного атомного номера, розрахувати коефіцієнт поглинання (при цьому питання про правильність визначення співвідношення кількостей, наприклад металу й неметалевих матеріалів, що містяться в контейнері, залишається відкритим). Таким чином, визначення перерахункового коефіцієнта (ефективності спектрометра) є можливим і для ситуації з дуже неоднорідними матеріалами. Тобто навіть для насипного будівельного сміття можна задавати перерахунковий коефіцієнт, оскільки такий матеріал можна описувати усередненими параметрами – густиною і речовинним складом. Але точність визначення активності при цьому залежить від точності визначення вмісту контейнера (точності задавання зазначених параметрів).

2.3.2. Постановка задачі

Отже, сучасний рівень гамма-спектрометричних вимірювань уже дозволяє робити обґрунтовані оцінки їхньої точності. Фактично, якщо для кожного контейнера є можливість визначити ефективність спектрометра, то постає стандартна задача оцінки точності цієї процедури як для лабораторного гамма-спектрометра. У термінах метрології задача метрологічної атестації в

такий спосіб зводиться до оцінки похибки даної методики визначення активності. Похибка буде залежати від ряду складових:

- гетерогенність матеріалу в контейнері;
- похибки вимірювань власне приладом (це показник, що зазначений у свідоцтві про метрологічну перевірку спектрометра, або похибка, яку розраховує програма);
- похибки розрахунків ефективності (повинна бути вказана у технічній документації на прилад або взята з літератури), що описує точність розрахунків методом Монте-Карло;
- похибки визначення кожного з параметрів, що вводяться в програму (формулу);
- алгоритму розрахунку активності.

У даному розділі розглядається похибка, яка визначається останніми двома пунктами.

Формула, що пов'язує активність джерела і створене ним поле, має такий вигляд:

$$M = CA\rho_1 \int_V \frac{\exp(-\mu_1 r_1 \rho_1 - \mu_2 r_2 \rho_2 - \mu_3 r_3 \rho_3) dV}{(r_1 + r_2 + r_3)^2}, \quad (2.3.2)$$

де A – активність; C – коефіцієнт пропорційності; M – вимірюваний параметр гамма-поля (ПЕД, площа піка тощо); μ_1, μ_2 і μ_3 – коефіцієнти поглинання для густини 1 г/см^3 у матеріалі РАВ, стінках контейнера (залежать від ефективного атомного номера, тобто від хімічного складу цих матеріалів) та в повітрі; r_1, r_2 і r_3 – шлях, який випромінювання проходить відповідно в матеріалі РАВ, стінці контейнера та в повітрі; ρ_1 – густина матеріалу ТРВ; ρ_2 – густина стінки контейнера; ρ_3 – густина повітря; V – об'єм джерела (контейнера). Відстань між детектором і контейнером d не входить у формулу, але використовується при розрахунку $r_1 + r_2 + r_3$.

Оскільки густина матеріалу в контейнері $\rho_1 = m/V$, де m – маса РАВ, а V – їхній об'єм, який у свою чергу визначається як функція розміру (для конкретності – для контейнера у формі паралелепіпеда з розмірами t і висотою заповнення h) $V = t \cdot t \cdot h$, то формула (1.5) набуває вигляду

$$M = C \frac{Am}{htt} \int_V \frac{\exp(-\mu_1 r_1 \frac{m}{htt} - \mu_2 r_2 \rho_2 - \mu_3 r_3 \rho_3) dV}{(r_1 + r_2 + r_3)^2}, \quad (2.3.3)$$

звідки видно, що результат нелінійно залежить від параметрів.

Як впливає з вищенаведеного, похибка результату вимірювання залежить від точності визначення параметрів, необхідних для розрахунку ефективності спектрометра. Тобто від густини матеріалу РАВ, ефективних атомних номерів матеріалу РАВ і матеріалу контейнера, відстані між

детектором і контейнером, розмірів контейнера, а також точності вимірювань спектра даним приладом. Сумарна похибка не визначається сумою похибок визначення вказаних величин. Оскільки при розрахунку активності вони функціонально зв'язані, то сумарна похибка визначається дещо іншим способом, як похибка для функції багатьох змінних (диференціальна формула оцінки похибки) [5]. Це досить складний спосіб, проте головна проблема в тому, що таким методом похибка визначається для конкретних значень параметрів. Тобто таким способом можна оцінити похибку при вибраних значеннях змінних, а потім повторювати цю процедуру в усьому діапазоні зміни їхніх значень. Оскільки практично це неможливо зробити при неперервних діапазонах зміни величин, а визначення похибки для крайніх значень діапазонів зміни (невизначеності) параметрів не означає визначення крайніх меж похибки, виникає необхідність застосування спеціальних методів.

Тут пропонується підхід, який полягає у прямих розрахунках активності на основі підходів Монте-Карло шляхом перебору різних комбінацій значень параметрів у заданих діапазонах їхньої зміни з наступним статистичним аналізом одержаних результатів. Ідея розрахунку така. Параметри моделі визначаються одні з більшою, інші з меншою точністю не тільки як результат вимірювань, але й як результат припущень про, наприклад, склад ТРВ. Тому для розрахунків необхідно вказати можливий діапазон змін кожної із величини, які входять у розрахункову формулу. Якщо неможливі прямі вимірювання, ці діапазони визначаються методом експертних оцінок [6], у нашому випадку виходячи з очікуваної геометрії вимірювань, літературних даних та практичного досвіду експертів. Для прикладу вкажемо, що в інструкції по експлуатації гамма-спектрометричної системи ISOCS фірми Canberra [7] розглянуто розрахунок активності для випадку, коли контейнер „заповнено на три чверті”. Це автоматично означає, що висота заповнення була експертом визначена „на око”, і найкраще, що можна про неї сказати, це що похибка її визначення не менша $1/8$ висоти контейнера.

У табл. 2.3.1 наведено середні значення параметрів та інтервали їхніх змін, використані в подальшому при модельному розрахунку. Ці величини були вибрані при аналізі практичної ситуації: частково заповнений радіоактивними відходами 200 – літровий контейнер зі стороною t .

Підкреслимо, що при цьому ми не задаємося, наприклад, конкретним складом ТРВ, а просто задаємо можливий діапазон зміни коефіцієнта поглинання для вибраних "модельних" матеріалів.

При виконанні вимірювань спектрометром з теоретичним розрахунком ефективності оператор уводить у комп'ютер указані в табл. 2.3.1 параметри. При цьому дані вводяться як точні, незалежно від точності їх визначення. У результаті комп'ютер спектрометра видає результат визначення активності. Це є експериментальний результат. Якщо ввести інші параметри із заданих інтервалів їхньої зміни, одержимо інший, але так само експериментальний результат.

Таким чином, якщо операцію розрахунків повторити багато разів, можна

Таблиця 2.3.1. Параметри розрахунків, визначені з використанням експертних оцінок

Параметр	Позначення	Середня величина	Діапазон варіювання параметра
Площа спектра	M	80 відн. од.	± 15 відн. од.
Відстань між контейнером і детектором	d	398 см	± 2 см
Маса ТРВ	m	190000 г	±100 г
Коефіцієнт поглинання ТРВ	μ_1	0,077319 см ⁻¹	± 0,000372 см ⁻¹
Коефіцієнт поглинання матеріалу контейнера (заліза)	μ_2	0,076110 см ⁻¹	± 0,000001 см ⁻¹
Розмір контейнера	t	40 см	± 0,1 см
Товщина стінки контейнера	r ₂	0,29 см	± 0,01 см
Висота заповнення контейнера	h	65 см	± 15 см

одержати уяву про похибку визначення активності, пов'язану з неточністю визначення параметрів. Припустивши, що запропонований у даній роботі метод розрахунку активності еквівалентний тому, який використовується в спектрометрі, можна зауважити, що одержані результати еквівалентні експериментальним.

У результаті задача виявляється зведеною до ситуації, коли параметри розрахунків не мають певних значень, а задаються інтервалами можливих значень. Існує кілька підходів до виконання розрахунків для таких задач. Інтервальна математика [8, 9] та теорія планування експериментів [10] дають можливість розрахувати крайні діапазони значень шуканої величини *A*, проте цього не досить, і застосування методу Монте-Карло дає більше інформації. Нарешті, оскільки ряд параметрів задається на основі експертних оцінок, їх можна розглядати як нечіткі дані і скористатися прийомами нечіткої логіки [11].

Було виконано два види розрахунків. У першому в заданих інтервалах (див. табл. 2.3.1) з рівною ймовірністю випадково вибиралися значення параметрів (інтервальні розрахунки). Далі для різних комбінацій параметрів з указаних діапазонів їхніх змін розраховувалася активність *A* для кожної комбінації. Для збільшення статистики та усунення можливих особливостей,

пов'язаних із генератором випадкових чисел, такий розрахунок повторювався ще кілька десятків разів (до $50 \times 30\,000$ випадкових комбінацій).

У другому випадку вважалося, що найбільш імовірним є значення параметра із середини діапазону, а крайні значення менш імовірні (нечіткі множини). Для простоти розрахунків функція приналежності була вибрана гауссовою [12] (рис. 2.3.2).

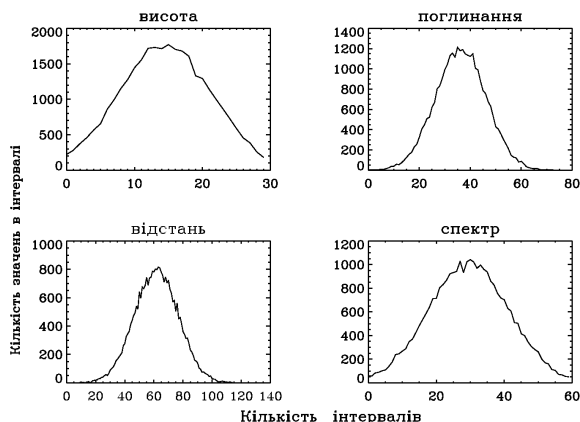


Рис. 2.3.2. Розподіл значень параметрів при гауссовій імовірності їх вибору для висоти заповнення контейнера, коефіцієнта поглинання, відстані між контейнером і детектором та площі спектра.

Тобто при розрахунках значення параметрів випадково вибиралися із заданого інтервалу не рівномірно, а більш часто вибиралися значення із середини діапазону.

Правильність формул і розрахунків перевірялись шляхом порівняння з аналогічними розрахунками програмою MicroShield фірми GroveEngineering при кількох конкретних значеннях параметрів. Результати збігаються з точністю в кілька відсотків. Статистична похибка розрахунків використаним методом Монте-Карло залежить від обсягу вибірки і при кількості точок більше 50 000 становить менше 1 %.

2.3.3. Обговорення результатів

На рис. 2.3.3 показано одержаний масив результатів розрахунку активності при вибраних інтервалах параметрів моделі (див. табл. 2.3.1) у вигляді гістограм для рівномірного вибору значень параметрів і для гауссового розподілу вибраних значень. Інтервали зміни активності в обох випадках практично збігаються, хоча криві розподілу значно відрізняються. Закладена в методику розрахунку ідея про статистичне перебирання параметрів у межах їхніх інтервалів невизначеності та вигляд гістограми на рис.1.3, а приводять до думки про можливість оцінки шуканої похибки методами математичної статистики (аналізом повторних вимірювань).

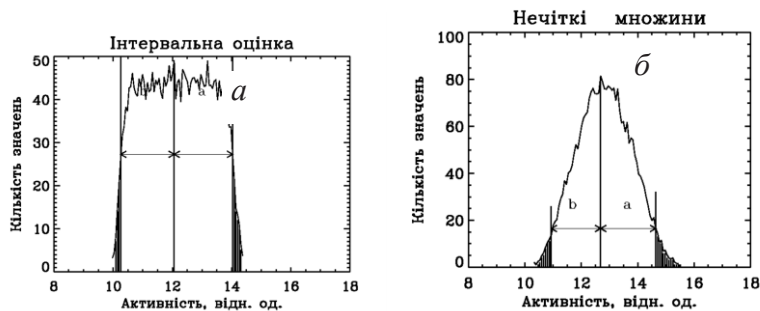


Рис. 2.3.3. Функції розподілу розрахованих значень активності для заданого гауссовою функцією вибору значень в інтервалі (а) та рівномірного вибору (б).

Зробимо зауваження, які покажуть, що формально (теоретично) цього не можна зробити. Перше: одержані значення активності не є випадковими величинами. Гістограми з максимумами одержуються тому, що деякі комбінації параметрів можуть давати близькі результати. Одержані результати розрахунків знаходяться в певному діапазоні і за нього вийти не можуть, яким би великим не був обсяг вибірки. Друге: обидва використані підходи вибору комбінацій параметрів – рівномірний чи з гауссовою функцією розподілу здаються однаково можливими (обґрунтованими). Вибір іншої функції розподілу (наприклад, трикутної, яка широко використовується в теорії нечітких даних) або зміна довільно вибраних параметрів використаних гауссових функцій призведе до зміни виду гістограм з рис. 2.3.3, тобто вид гістограм залежить від способу описування даних, хоча самі інтервали при цьому не змінюються. Третє: розподіли типу одержаного на рис. 2.3.3, б взагалі в нормативній документації по оцінці похибок не розглядаються. Четверте: крайні значення розрахованої активності зустрічаються менш часто, ніж інші. Проте апріорі невідомо, які комбінації параметрів дають такий результат і немає об'єктивного способу вирішити, чи така комбінація параметрів малоймовірна з точки зору їхнього вибору при проведенні експерименту.

Шлях до оцінки похибки в нестандартних ситуаціях дає „Керівництво з невизначеності...” (див. частину 1), в якому рекомендуються неформалізовані способи оцінки невизначеності у випадках, коли неможливе застосування статистичних методів (повторного вимірювання). У цьому випадку рекомендується представляти невизначеність (похибку) у вигляді меж відхилення значення вимірюваної величини від її оцінки. Якщо ж припустити, як це прийнято робити на практиці, що крайні значення малоймовірні, то можна їх відкинути і підрахувати похибку (табл. 2.3.2) як відношення половини інтервалу решти значень розрахованої активності до середнього, медіани або найбільш імовірного (у максимумі гістограми). Така оцінка показана на рис. 2.3.3. Задамо довірчий інтервал, наприклад 95 %. Проведемо на рисунку горизонтальну лінію, нижче якої будуть знаходитись вибрані

Таблиця 2.3.2. Розрахунок похибок для показаних на рис. 2.3.3 результатів

Метод	Число комбінацій	Найбільш імовірне значення активності відн. од.	Ліва похибка, %	Права похибка, %	Довірчий інтервал, %
Висота заповнення контейнера 65 ± 15 см					
Інтервали	10 × 3000	12,1	16,7	17,5	93,6
Нечіткі дані	10 × 3000	12,8	13,8	15,4	94,6

приблизно 5 % значень, які найбільше відхиляються від середнього зліва і справа (заштриховані області). Ці точки входять у межі 5 % малоімовірних значень і з подальшого розгляду викидаються. Відповідно вертикальними лініями, проведеними через точки перетину гістограми (яка тепер розглядається як функція щільності ймовірності) з горизонтальною лінією, відмічено крайні значення активності, у межах яких проводяться подальші розрахунки. Прийmemo за найбільш імовірне для цих точок максимум A_{\max} (можна було б вибрати й точку середньоарифметичного значення або медіану) – загальний інтервал невизначеності при цьому не міняється. Горизонтальними стрілками вказано максимальні відхилення від середнього вліво та вправо. Лівою похибкою (у відсотках) назвемо величину $(a/A_{\max}) \cdot 100$, а правою похибкою величину $(b/A_{\max}) \cdot 100$. У випадку несиметричних гістограм ліві і праві похибки можуть помітно відрізнятись.

Як видно, при такому підході величина похибки залежить від способу описування даних. Усе це разом означає, що при заданій довірчій імовірності використання середньоквадратичного відхилення, яке дає значно меншу оцінку похибки, є некоректним і найбільш правильним було б в якості похибки просто вказувати можливі межі оцінок активності.

2.3.4. Резюме до розділу про контейнер

Ні в нормативних документах, ні в спеціальній літературі не вказано способу визначення похибки результату для розглянутої в даній роботі ситуації. У той же час радіаційні вимірювання вимагається виконувати за методиками з обов'язковим визначенням похибки вимірювань. Методом Монте-Карло на основі перебору всіх можливих комбінацій параметрів визначено межі, в яких лежить результат розрахунків та проаналізовано можливість статистичного аналізу цих результатів. Показано, що для реальних геометрій (контейнерів з РАВ) пов'язана з цим похибка визначення активності, яку зазвичай не враховують, становить не менше 25 – 30 % залежно від прийнятих наближень.

Проведений аналіз показує, що пов'язана з неточним знанням параметрів розрахунку похибка не може бути визначена шляхом збільшення числа вимірювань та загальноприйнятими методами розрахунку похибок.

Запропонований спосіб оцінки похибки вимірювання активності при використанні гамма-спектрометричних систем із можливістю теоретичного розрахунку ефективності, яка виникає через неточне визначення параметрів ТРВ у контейнері, і проведений за допомогою цього способу аналіз показує, що ця похибка взагалі є інтервальною величиною. Указану в табл. 2.3.2. її величину порядку $\pm 20\%$ можна вважати консервативною оцінкою для конкретного об'єкта – стандартного контейнера.

При інших співвідношеннях між значеннями параметрів та інтервалами їхньої зміни вона може значно змінитися.

Зауважимо також, що це оцінка складових невизначеності лише для двох з перерахованих вище джерел невизначеності. Загальну невизначеність оцінки активності в контейнері потрібно зробити відповідно до процедури, описаної в «Керівництві з невизначеності...».

СПИСОК ЛІТЕРАТУРИ

1. *Панасюк М. І. Скорбун А. Д., Сплошной Б. М.* Про точність визначення активності твердих радіоактивних відходів гамма-спектрометрами з теоретичним розрахунком ефективності // Проблеми Чорнобиля. – 2008. – Вип. 1. – С. 39 – 42.
2. *Гусев Н. Г.* Гамма-метод и его применение на практике // Дозиметрические и радиометрические методики. – М.: Атомиздат, 1996. – С. 412 – 422.
3. *Гришмановский В. И., Комодов А. А., Лызлов А. Ф.* Оценка радиоактивности твердых отходов // Дозиметрический контроль при работе с радиоактивными веществами и источниками ионизирующих излучений. – М.: Атомиздат, 1980. – Т. 1. – С. 262 – 266.
4. *Исаков А. П. Лебедева Т. Г., Мартынюк Ю. Н. и др.* Проблемы паспортизации радиоактивных отходов на предприятиях ЯТЦ // АНРИ. – 2006. – № 2. – С. 47 – 51.
5. *ЩигOLEV Б. М.* Математическая обработка наблюдений. – М.: Наука, 1969. – С. 344.
6. *Орлов А. И.* Экспертные оценки // Заводская лаборатория. – 1996. – Т. 62, № 1. – С. 54 - 60.
7. *Программа калибровки ISOCS. Версия 4.0.* Руководство пользователя. – Canberra, 2003.
8. *Шокин Ю. И.* Интервальный анализ. – Новосибирск: Наука, 1981. – 112с.
9. *Ferson S., Ginzburg L., Krenovich V. et al.* Exact bounds of finite populations of interval data. Reliable computing. – 2005. - Vol. 11, No. 3. – P. 207 – 233.
10. *Асатурян В. И.* Теория планирования эксперимента. – М.: Радио и связь, 1983. – С. 249.
11. *Кофман А.* Введение в теорию нечетких множеств. – М.: Радио и связь, 1982. – 432 с.
12. *Kala Zdenek, Omissshore Abayomi.* Comparison of fuzzy set theory and stochastic method results in application to the analysis of the ultimate load-carrying capacity of steel members with imperfections // VIII Int. conf. on Computational Plasticity (COMPLAS VIII). E. Ofiate and D.R.J. Owen (Eds). CIMNE. – Barcelona, 2005.

2.4. Аналіз впливу нерегулярності даних

Як неодноразово згадувалося, класична математична статистика розроблена для даних (вибірок, рядів чисел), які є незалежними й однаково розподіленими. Однак існує багато вимірювань (і пов'язаних з ними фізичних задач), де результати цим вимогам не відповідають. До них належать, наприклад, метеорологічні вимірювання, у яких параметри (температура, тиск, швидкість вітру) є неперервними функціями. У всякому випадку, оскільки атмосфера описується як суцільне середовище, логічно припустити, що коли зробити, наприклад, щосекундні вимірювання температури, а потім узяти усереднені щохвилинні дані, то добова крива зміни температури буде плавною, а величина кожного наступного значення буде залежати від попереднього. А зміни між сусідніми значеннями будуть невеликими порівняно з добовими. У такому наближенні дані перестануть бути незалежними. Більше того, статистичною ознакою не випадковості таких даних є також існування закономірних змін типу добових, сезонних і т. ін. (бо ми можемо передбачити, що температура передбачувано зміниться через дванадцять годин). Отже, існують регулярні ряди даних, які зовні можуть виглядати як випадкові (шумоподібні), проте вони мають деякі внутрішні закономірності. Для них розраховувати параметри класичної математичної статистики звичайно можна, проте з очевидними застереженнями. Очевидно також, що для аналізу таких рядів даних потрібно розвивати методи аналізу, які виявляли б їхню не випадкову структуру.

Одним з таких методів є фрактальний аналіз, який розвинувся із спостережень Харста (Hurst) за рівнем води в річці Ніл [1]. Результатом фрактального аналізу в нашому випадку буде показник Харста H . Теоретично показано, що для випадкового ряду з нормально розподіленими числами $H = 0,5$, а при появі залежності між даними H стає більшим від $0,5$, для природних процесів $H \sim 0,7 - 0,8$ [1].

Явно чи неявно при проведенні аналізу методом нормованого розмаху Харста припускається, що дані не є випадковими, тобто вони породжені деяким процесом так, що між ними існує залежність. Відповідно постає питання, чи можна виявити цю залежність, якщо дані будуть неповні: недостатня частота опробувань або вимірювання проводилися нерегулярно тощо. У нашому випадку питання виникло з такої практичної задачі. Проф. Б. І. Огородніков багато років проводив вимірювання концентрації аерозолів усередині об'єкта «Укриття». Вимірювання були нерегулярні в тому сенсі, що проводились кілька разів на тиждень і навіть не кожний місяць. Виникло запитання, чи можна в таких, по-суті випадкових, вимірюваннях знайти якісь закономірності. Оскільки було показано, що концентрації аерозолів пов'язані зі швидкістю вітру в навколишньому середовищі [2], було вирішено розглянути статистичні властивості швидкості вітру, дані про яку існують на метеостанції «Чорнобиль», сподіваючись, що їх можна перенести на вимірювання аерозолів. У цих дослідженнях було показано, що швидкість

вітру має фрактальні властивості, а величина H залежить від частоти проведення вимірювань (3 год, доба тощо). У результаті виникло таке запитання: наскільки представницькою є величина H , отримана для наших рядів вимірювання швидкості вітру, наскільки вона залежить від регулярності вимірювань, тривалості вимірювань тощо.

З точки зору питань, що розглядаються в даній книзі, це означає: якою є невизначеність отриманого результату оцінки H , від чого вона залежить і наскільки загальним є отриманий результат. Отже, було зроблено спробу з'ясувати, наскільки спотворюються результати розрахунків H при нерегулярності вимірювань.

Зрозуміло, що якщо випадково видаляти певне число елементів ряду, то рано чи пізно ряд втратить свої властивості, якщо він не випадковий. Стосовно коефіцієнта Харста H це питання перевірялося (не дуже ретельно) у [3], де було показано для модельних даних на послідовності із заданим H , з якої було викинуто неперервний інтервал приблизно в чверть від загальної послідовності, що різні методи розрахунку H по-різному на це реагують, тобто в них різна чутливість до такого процесу. Подібну оцінку, але для ситуації, коли з реальних даних (багаторічні вимірювання індексу Північно-Атлантичного коливання (NAO)) було викинуто лише а) близько 5 % випадково вибраних значень і б) 5 % останніх значень було зроблено в [4], де одержано, що коефіцієнт Харста при цьому зберігає своє значення. Питання про значимість коефіцієнта Харста при нерегулярних вимірюваннях розглянуто також у [5, 6].

Нас цікавила така задача: яке відхилення від істинного H (тобто отриманого для повного ряду даних) можна очікувати, якщо вимірювання проводяться нерегулярно впродовж певного інтервалу часу. Для цього з повної серії вимірювань середньодобових швидкостей вітру випадковим чином видаляли деяке число елементів і для такої укороченої вибірки підраховували H . Можна уявити два способи розгляду таких вибірок з пропущеними даними. Перший – не звертати увагу на існування пропущених даних і розглядати наявні дані як неперервний ряд. При цьому для випадкового ряду не повинно бути ніяких змін. Другий – при розрахунках враховувати, у які проміжки часу дані було втрачено. Нижче описано цей другий підхід.

Опис методу Харста див. у розділі 2.5. Процедура розрахунків при наявності пропущених вимірювань була така. Оскільки є безперервний ряд послідовних щоденних вимірювань швидкості вітру, то кожному вимірюванню присвоювався послідовний номер. Спочатку розглядається безперервний ряд послідовних **номерів** вимірювань, розбитий на деяке число інтервалів рівної тривалості τ . Потім з цього повного ряду випадковим чином вибирається задане число елементів, яким присвоюється значення 0. Це означає, що дані про вимірювану величину (швидкості вітру) для цих днів невідомі й у наявному ряді вимірювань для цих номерів дані видалялися. У результаті утворюються інтервали, в яких число елементів менше, ніж τ . При великому числі видалених елементів можуть утворюватися інтервали, в яких

всі елементи рівні 0. Такі інтервали з підрахунків викидалися. Потім для кожного інтервалу визначалися R , S та середній нормований розмах для даного τ .

Щоб оцінити невизначеність у підрахунку коефіцієнта Харста, пов'язану з випадковістю вибору номерів пропущених елементів з повного ряду при заданій їхній кількості, процедура підрахунку для заданого числа пропущених елементів повторювалася кілька разів, при цьому елементи вибиралися щоразу випадковим чином. Результати таких розрахунків представлено на рис. 2.4.1, б, де показано зміну H залежно від P – частки елементів, що залишилися у вибірці після видалення з неї їхнього деякого числа, по відношенню до повного ряду. Цей графік демонструє вельми складну і цікаву залежність величини H від кількості пропущених даних. Навіть у випадку одного і того ж числа опущених елементів, за рахунок їхнього випадкового вибору H виявляється різним і точки на графіку утворюють деяку смугу, ширина якої залежить від P і є мірою невизначеності оцінки H . Можна зробити такі висновки: 1) із зменшенням числа обраних для аналізу елементів аж до випадковим чином взятих $P = 0,2$ частин із загального числа елементів ряду отриманий коефіцієнт Харста відхиляється від базового в межах $\pm 15\%$, хоча і зберігає високе значення – основну характеристику, яка є показником наявності довгострокової пам'яті. Отже, зберігається, як і для повного ряду, властивість $H > 0,5$; 2) із зростанням частки викинутих значень до приблизно 50% середня величина H зростає, а при подальшому збільшенні починає падати; 3) невизначеність в оцінці H за рахунок випадковості вибору пропущених елементів поступово зростає зі збільшенням їхньої частки і досягає приблизно $\pm 10\%$; 4) якщо взяти менше 0,2 елементів від їхнього загального числа, H різко падає до значень $\sim 0,5 \pm 0,07$, тобто інформація про довгострокові ефекти втрачається повністю.

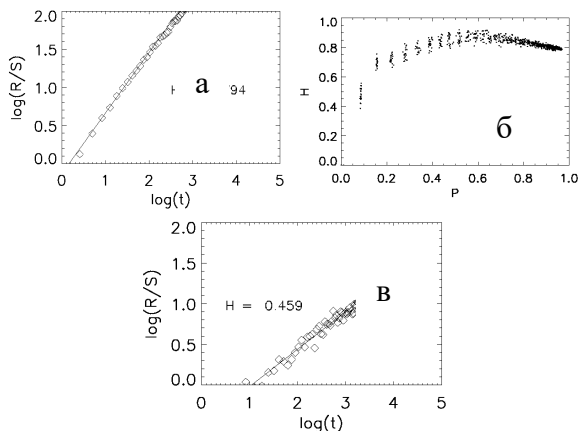


Рис. 2.4.1. Аналіз даних спостережень середньої швидкості вітру методом нормованного розмаху Харста: а – аналіз регулярного ряду спостережень; б – залежність H від частки збережених даних P ; в – аналіз для даних на дати вимірювань концентрації аерозолів.

Дати проведення вимірювань концентрації радіоактивних аерозолів в об'єкті «Укриття» можна вважати випадковою вибіркою з повного ряду дат вимірювань середньої швидкості вітру. Якщо за чотири роки (2003–2006) число можливих щоденних спостережень склало 1461, а число відібраних проб становить близько 200, то для цих вимірювань $P = 200/1461 \approx 0,14$. Розрахунок H для середньої швидкості вітру для дат пробовідбору, виконаний за описаним вище алгоритмом, зображено на рис. 2.4.1, в. У повній відповідності з наведеними на рис. 2.4.1, б результатами розрахунків, отриманий коефіцієнт Харста склав близько 0,5 (див. рис. 2.4.1, в), тобто інформація про довготривалі залежності зникла і ряд набув властивостей випадкового.

Висновки

Виявлене високе значення H для середньої швидкості вітру свідчить про сильну залежність поточного стану системи від попередніх значень (процес з довготривалою пам'яттю). Тобто ряд є далеким від випадкового.

Для вихідного ряду H визначається як нахил прямої на рис. 2.4.1, а і можна думати, що точність визначення нахилу досить хороша.

Проте проведений аналіз показав, що при штучно створених пропусках регулярності можуть виникати зміни в оцінці H , які знаходяться в межах приблизно $H = 0,7 - 0,9$, що відповідає відносній невизначеності $\pm 12\%$.

При пропусках, число яких сягає 80 %, ряд повністю втрачає свої початкові властивості й оцінка величини H втрачає сенс.

Зроблений аналіз змінності, тобто невизначеності отримуваних оцінок, можливо було отримати лише з використанням методів Монте-Карло, тобто випадкових випробувань.

СПИСОК ЛІТЕРАТУРИ

1. Федер Е. Фракталы. – М.: Мир – 1991 – 260 с.
2. Скорбун А. Д., Огородников Б. И. Анализ долговременных характеристик скорости ветра в районе Чернобыльской АЭС и возможность прогноза выброса аэрозолей из объекта “Укрытие” // Метеорология и гидрология. – 2010. – № 8. – С. 49 – 60.
3. Chamoli A., Bansal A. R., Dimri V. P. Wavelet and rescaled range approach for the Hurst coefficient for short and long time series // Computers and Geosciences. – 2007. – Vol. 33. – Issue 1. – P. 83 – 93.
4. Collette C., Auslous M. Scaling analysis and evolution equation of the North Atlantic oscillation Index fluctuations // arXiv:nlin.CD/0406068.
5. Sergei Katsev, Ivan L'Heureux. Are Hurst exponents estimated from short or irregular time series meaningful? // Computers&Geosciences. – 2003. – Vol. 29. – P. 1085 – 1089.
6. Kärner O. On nonstationarity and antipersistence in global temperature series // Journal of geophysical research. – 2002. – Vol. 107. – No. D20. – P. 4415 – 4425.

2.5. Оцінка гауссовості даних за допомогою модифікованого методу нормованого розмаху (R/S analysis)

У книзі багато разів згадувалася думка, що гауссовий розподіл далеко не завжди правильно апроксимує експериментальні дані. Але й перевірка гіпотези про гауссовість розподілу в дійсності робиться з певними наближеннями і при коротких вибірках відомі процедури, основані на критеріях порівняння із заданим розподілом, є не дуже точними. У цьому розділі описується прийом, який опирається не на формули, а на порівняння результатів роботи деякого алгоритма над згенерованою комп'ютером гауссовою вибіркою і над експериментальними даними.

Розглянемо набір результатів вимірювань, кількість яких дорівнює N . Якщо перенумерувати ці числа, тобто присвоїти їм порядковий номер, отримаємо випадкову вибірку із зафіксованим порядком розташування елементів. Таку вибірку можна розглядати як ряд послідовних регулярних вимірювань деякої (випадкової) величини в часі і проаналізувати можливість існування в такому ряді певних закономірностей. Метод для такого аналізу, який розглядається в даному розділі, називається фрактальним аналізом Харста (Hurst) [1]. Результатом такого аналізу є деяке число, яке називається коефіцієнтом Харста. Теоретично показано, що для нормально розподілених випадкових вибірок цей коефіцієнт дорівнює 0,5. Ми скористаємося цим результатом, щоб на основі модифікації фрактального аналізу Харста розробити спосіб оцінки відхилення деякого ряду даних від гауссового (нормально розподіленого), оцінюючи відхилення результату від величини 0,5.

Пропонується така схема розгляду. У припущенні, що комп'ютер генерує нормально розподілені випадкові послідовності з $\langle a \rangle = 0$, $\sigma = 1$, виконаємо потрібне нам статистичне моделювання на модельних вибірках, а потім порівняємо цей результат з результатом аналогічного аналізу нашої експериментальної вибірки.

Хоча R/S аналіз був розвинутий для аналізу часових рядів, ми застосуємо його до модельної вибірки (модельного ряду випадкових чисел), вважаючи, що кожне вимірювання зроблено через одиничний інтервал часу. R/S аналіз використаємо у вигляді, описаному в роботі [2]. Уся вибірка з n елементів розбивається на ряд з k інтервалів довжиною $\tau = n/k$. Для кожного із цих інтервалів розраховуються

$$X(n, \tau_k) = \sum_{i=1}^n (a_i - \langle a \rangle_{\tau_k}),$$
$$R(\tau_k) = \max X(n, \tau_k) - \min X(n, \tau_k),$$

$$S(\tau_k) = \frac{1}{\tau_k} \left[\sum_{i=1}^{\tau_k} (a_i - \langle a \rangle_{\tau_k})^2 \right]^{1/2},$$

$$\langle a \rangle_{\tau_k} = \frac{1}{\tau_k} \sum_{i=1}^{\tau_k} a_i. \quad (2.5.1)$$

Далі підраховуються середні значення цих величин по всіх інтервалах

$$R(\tau) = \frac{1}{k} \sum_k R(\tau_k) \quad S(\tau) = \frac{1}{k} \sum_k S(\tau_k). \quad (2.5.2)$$

Основою методу є одержане Харстом [цит. по 3] співвідношення

$$R(\tau)/S(\tau) \sim \tau^H \quad (2.5.3)$$

і було теоретично показано, що для гауссового процесу $H = 0,5$. На перший погляд, це можна було б використати для перевірки нормальності розподілу вибірки. У роботі [4] цей факт було використано для перевірки якості генераторів псевдовипадкових чисел і показано, що в згенерованих комп'ютерами послідовностях цим методом легко фіксуються кореляції, тобто на коротких послідовностях спостерігаються відхилення від $H = 0,5$. У вказаній роботі припускається, що цей ефект пов'язаний з особливостями генерації псевдовипадкових чисел комп'ютером, тому ми будемо вважати це експериментальним фактом і враховувати в подальшому при аналізі точності моделювання як існуючу статистичну властивість послідовності багатьох псевдовипадкових вибірок.

Модифікований R/S метод. Проблема полягає в тому, що нас цікавить аналіз коротких вибірок, застосування до яких R/S є проблематичним [5] (як завжди у статистиці теоретичний результат отримується при переході до нескінченності). Ми проаналізуємо можливість формальної побудови функції R/S за таким алгоритмом. Починаючи з перших елементів вибірки в якості першого інтервалу виберемо перші три елементи і далі створюємо інтервали τ послідовно збільшуючи їх на одиницю. Тобто

$\tau_1 = 1, 2, 3,$ де 1, 2, 3 – номери елементів вибірки

$\tau_2 = 1, 2, 3, 4$

$\tau_3 = 1, 2, 3, 4, 5$

....

$\tau_n = 1, 2, 3, \dots n$

Далі розраховуємо для кожного з них R/S (без усереднення по інтервалах).

При спробі застосувати цей розрахунок до короткої (кілька десятків елементів) вибірки виявилось, що для окремо взятої згенерованої нормальної послідовності закон Харста не виконується, проте виконується в середньому

для суми таких послідовностей. На рис. 2.5.1 зліва для прикладу показана серія таких R/S залежностей і середнє значення H .

Для різних вибірок підраховане таким способом значення H може змінюватись від 0,3 до 0,7, хоча середнє значення $H = 0,48$ близьке до теоретичного $H = 0,5$.

Це відображає малопоширене, але класичне положення теорії ймовірностей, що очікуване теоретичне значення параметра розподілу, наприклад середнього, одержується як середній результат по багатьох вибірках, а кожна окрема вибірка може досить значно відрізнятись від теоретичної.

Однак виявляється, що для згенерованої нормальної вибірки, упорядкованої по зростанню, тобто варіаційного ряду, ця залежність (пряма лінія) відтворюється досить точно з $H \sim 0,95$ для всіх вибірок (див. рис. 2.5.1,

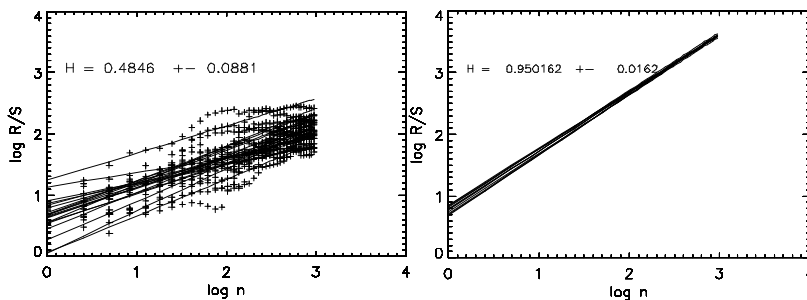


Рис. 2.5.1. Зліва – криві Харста для серії нормальних вибірок; справа – те саме для вибірок, упорядкованих по зростанню.

справа). Причому закономірність витримується як для великих вибірок (десятки тисяч елементів), так і для малих, число елементів яких становить 20 – 40. Безпосередня перевірка за допомогою критерію χ^2 показала, що із 10 000 вибірок 550 ± 20 не є нормальними, як при порівнянні з теоретичним розподілом ($\langle a \rangle = 0$, $\sigma = 1$), так і при визначенні $\langle a \rangle$ і σ за даними вибірки. Цей результат також не залежить від розміру вибірки. Тобто і для таких вибірок значення H знаходяться в межах, показаних на рис. 2.5.1, справа.

Виконаємо тепер аналогічний аналіз для вибірки, яка містить значення, що „випадають” із нормального розподілу. Додамо постійну величину b до кількох останніх (найбільших) елементів досліджуваного варіаційного ряду згенерованої на комп’ютері нормальної послідовності (або зробимо те саме для деякої частини випадковим чином відібраних елементів вибірки). Починаючи з деякого b у верхній частині залежності R/S чітко фіксується відхилення від прямої лінії (див. рис. 2.5.2). При цьому описана вище перевірка на нормальність за допомогою критерію χ^2 показує, що вибірка все ще залишається нормальною.

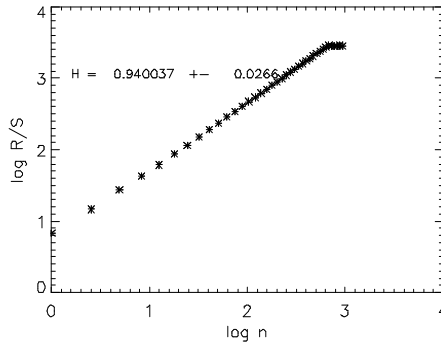


Рис. 2.5.2. Залежність Харста для упорядкованої нормальної вибірки, яка містить значення, що „випадають” із нормального розподілу.

2.5.1. Критерій відхилення від нормальної послідовності

Виявлене відхилення точок на графіку рис. 2.5.2 від стандартної прямої може бути використано для визначення наявності відхилення від нормальності.

Візьмемо серію модельних вибірок і проаналізуємо їх за допомогою модифікованого R/S аналізу. Тобто будемо послідовно генерувати нормально розподілені вибірки, підраховувати для них критерій χ^2 , та одержимо залежність R/S для упорядкованої вибірки так, як показано на рис. 2.5.2. Для більш наочної демонстрації наявності відхилення цієї залежності від прямої побудуємо для неї другу похідну. Типовий результат для 9 таких довільних нормальних послідовностей показано на рис. 2.5.3, де на кожному графіку наведено також результат перевірки по критерію χ^2 .

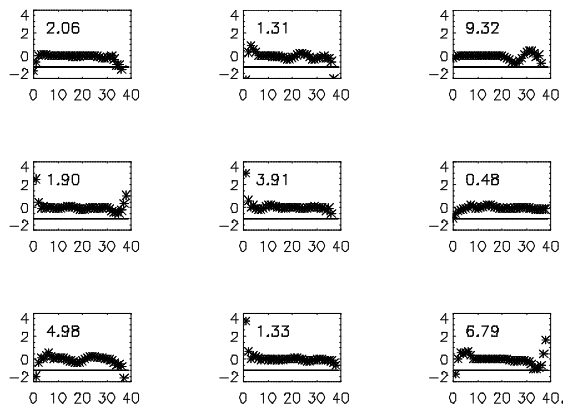


Рис. 2.5.3. Друга похідна від R/S залежності Харста для упорядкованих нормальних вибірок. Критерій нормальності $\chi^2 = 7,81$.

Цей рисунок демонструє межі відхилення по критерію χ^2 для різних вибірок, включаючи ті, для яких не виконується критерій нормальності. (Нагадаємо, що вище було показано, що серед згенерованих на комп'ютері нормальних вибірок близько 5 % не задовольняють критерію нормальності за χ^2) З рис. 2.5.3 можна зробити два висновки: 1) навіть для нормальних вибірок спостерігаються особливості на краях кривої; 2) розраховані значення χ^2 як для нормальних, так і для вибірок, які за цим критерієм не є нормальними, ніяк не корелюють із виглядом другої похідної від функції R/S. Тобто, аналогічно ситуації з критерієм χ^2 , для теоретично нормальних вибірок (псевдовипадкових послідовностей) у поведінці другої похідної також спостерігаються особливості, яких не повинно бути, якби упорядкована по зростанню вибірка була ідеальною прямою (див. рис. 2.5.1). У подальшому аналізі ці відхилення будуть розглядатися як шум, на фоні якого потрібно виділяти справжні ефекти відхилення від нормальності.

На рис. 2.5.4 показана серія таких кривих для упорядкованих модельних нормальних вибірок, згенерованих в інтервалі $[0...1]$, для кожної з яких послідовно величина трьох останніх (найбільших) значень збільшується на 0,3. Тобто зліва направо зверху вниз на верхньому лівому графіку до трьох останніх значень додано 0, а для нижнього правого – 2,7.

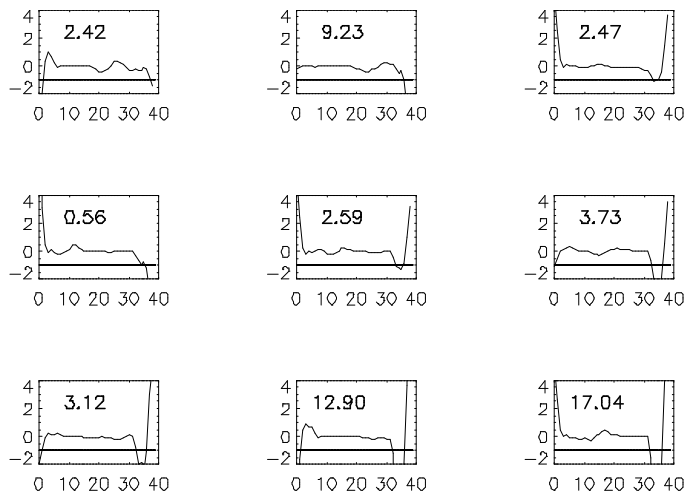


Рис. 2.5.4. Друга похідна від залежності Харста для упорядкованих нормальних вибірок, які містять значення, що „випадають” із нормального розподілу. Критерій нормальності $\chi^2 = 7,81$.

Особливістю останніх кривих (які спеціально показані суцільними лініями) є явне утворення мінімуму в правій частині кривих при зростанні останніх елементів вибірки (що й очікувалось із вигляду графіка на рис. 2.5.2). Цей факт може бути використаний для вибору емпіричного критерію наявності відхилення від нормальної кривої: кількості точок, які знаходяться нижче деякого критерію. У нашому випадку таким критерієм було довільно вибрано

значення другої похідної, рівне -1 . Як видно з рис. 2.5.4, навіть для теоретично нормальної кривої можлива наявність від'ємних значень другої похідної нижче вибраного рівня -1 . Тому були розраховані середні по 1000 вибірок кількості точок нижче лінії -1 для різних b . Паралельно підраховувалось число вибірок, які не відповідають рівню нормальності по критерію χ^2 .

Відповідні графіки наведено на рис. 2.5.5.

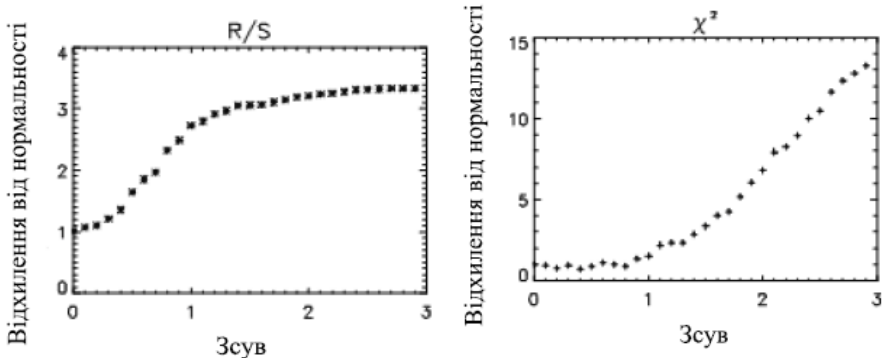


Рис. 2.5.5. Порівняння чутливості R/S та χ^2 аналізів до відхилення вибірки від нормальності.

Як видно, R/S аналіз починає відчувати відхилення останніх елементів від нормального закону починаючи приблизно з $b = 0,5$. При значеннях $b \geq \approx 1,5$ крива виходить на насичення. Але це не означає відсутності чутливості до відхилення. Просто всі статистично можливі точки опинилися нижче вибраного критерію і на графіку рис. 2.5.4 спостерігається стабільний мінімум.

На противагу цьому критерій χ^2 з $b \geq \approx 1,5$ тільки починає відчувати відхилення останніх елементів від нормальної послідовності.

Спотворення гауссової вибірки додаванням константи b зроблено для більшої наочності. Аналогічний результат одержано, коли розглядається бімодальна вибірка, утворена з двох зсунутих гауссових вибірок.

2.5.2. Застосування до експериментальних даних

Застосуємо тепер описаний метод до експериментальної вибірки 1 (з табл. 2.7.1, наведеної в розділі 2.7 джекнайф). Для того щоб було більш чітко видно, з яких точок починається відхилення від нормальності на кривій R/S, рисунок зроблено не в логарифмічному масштабі.

На рис. 2.5.6 ясно видно, що 10 крайніх точок відхиляються від нормальної залежності (від прямої лінії). Рисунок зроблено для всієї вибірки. Якщо просто відкинути одну крайню точку, одержимо результат для вибірки, яка по критерію χ^2 визначається як нормальна. У той же час по даним запропонованого способу аналізу ще 9 точок не відповідають нормальному розподілу, що підтверджується і виглядом гістограми, яка явно має правосторонню асиметрію.

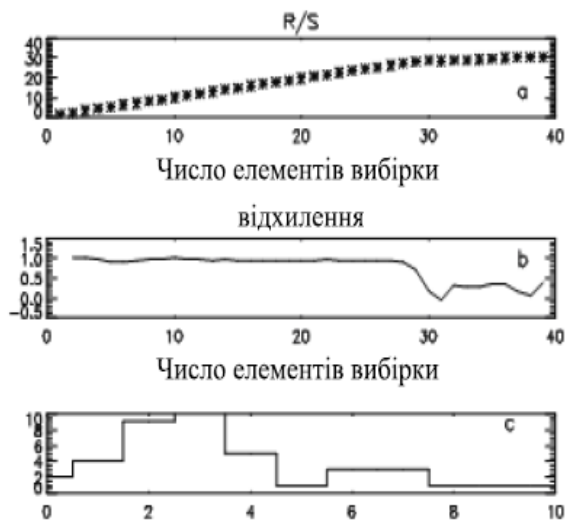


Рис. 2.5.6. Статистичний R/S аналіз експериментальної вибірки. Верхній рисунок – $R(\tau)$, середній – похідна від верхньої кривої, нижній – гістограма.

Висновки

Розроблено метод пошуку "викидів" з нормального розподілу, оснований на використанні модифікованого методу R/S аналізу, основою якого є упорядкування вибірки перед обробкою по зростанню (перетворення її у варіаційний ряд).

Показано, що модифікований R/S аналіз, коли він застосовується до нормальної вибірки з "викидами", є значно більш чутливим порівняно з методом χ^2 при оцінці відхилення вибірки від нормальності. Для вибірок, які мають явно бімодальний розподіл (складаються з двох гауссових кривих), запропонований метод дозволяє виявляти бімодальність при значно менших відстанях між центрами розподілів, ніж методом χ^2 .

Застосування розробленого методу до експериментальної вибірки результатів вимірювання активності неоднорідно забруднених ґрунтів, яка має досить великий розкид значень показало, що не дивлячись на те, що гістограма експериментальної вибірки має явну правосторонню асиметрію, по критерію χ^2 ці дані визначаються як нормальні чи логнормальні, хоча в дійсності такими не є. Тому звичайний статистичний аналіз до них треба застосовувати з обережністю. У той же час застосування R/S аналізу явно демонструє, що майже третина точок не відповідає нормальному розподілу і є всі підстави вважати її бімодальною. Хоча модифікований метод R/S аналізу було перевірено за допомогою моделювання з використанням комп'ютерних псевдовипадкових послідовностей, для яких є можливим відхилення від теоретично очікуваних залежностей, застосування цього методу до результатів

експериментальних вимірювань не залежить від будь-яких особливостей псевдовипадкових послідовностей.

Зауважимо, що запропоноване застосування R/S аналізу не дозволяє робити ні математично, ні фізично обґрунтовані висновки про функцію розподілу. Скоріше, він дозволяє формально описати вибірку як суму двох функцій, одна з яких є гауссовою.

Одержаний результат демонструє також, що підсвідоме бажання звести, наприклад, за допомогою відкидання „грубих похибок” або логарифмування вибірку до нормального розподілу, щоб одержати „обґрунтований” спосіб розрахунку середнього значення чи похибки його визначення, означає просто ігнорування реальної картини розподілу даних вимірювань. Ще раз підкреслимо, що розглянута ситуація є типовою для екологічних вимірювань, тобто апріорі припускати гауссовість розподілу небезпечно. Для прикладу вкажемо [6], де без будь-яких підстав припускається нормальність розподілу результатів радіаційних польових вимірювань і викидаються або ігноруються результати, що "випадають" із загального ряду по критерію нормальності розподілу.

СПИСОК ЛІТЕРАТУРИ

1. *Hurst H. E.* Long-Term Storage Capacity of Reservoirs // Transactions of the American Society of Civil Engineering. – 1951. – Vol. 116. – P. 770 – 799.
2. *Федер Е.* Фракталы. – М.: Мир, 1991. – 260 с.
3. *Кроновер Р. М.* Фракталы и хаос в динамических системах. – М.: Постмаркет, 2000. – С. 352.
4. *Gammel B. M.* Hurst's rescaled range statistical analysis for pseudorandom number generator, used in physical simulations // Phys. Rev. E. – 1988. – Vol. 58, No. 2. – P. 2586 – 2597.
5. *Sergei Katsev and Ivan L'Heureux.* Are Hurst exponents estimated from short or irregular time series meaningful? // Computers & Geosciences. – November 2003. – Vol. 29. – Issue 9. – P. 1085 – 1089.
6. *Ганжа Д. Д., Назаров О., Б., Сплошной Б. М.* Радиологічна зйомка територій у мірілі від 1 : 500 до 1 : 5000 // Проблеми безпеки атомних електростанцій і Чорнобиля. – 2008. – Вип. 9. – С. 144 – 151.

2.6. Визначення похибки вимірювання при невідомій аналітичній залежності функції виходу від параметрів входу

Схеми процедур для підрахунку невизначеності результатів вимірювань, відповідно до «Керівництва з невизначеностей...», починаються з опису функціональної залежності результату від факторів впливу. Це тому, що сумарна похибка результату залежить від впливу (похибки) кожного фактора. Але існують ситуації, коли функціональна (аналітична) залежність результату вимірювань від ряду параметрів невідома чи складно визначається, хоча і відомо, що в принципі вона існує. Це є характерним, наприклад, при аналізі роботи складних вимірювальних комплексів з багатьма факторами впливу. У результаті немає можливості скористатись для оцінки похибки наведеною нижче стандартною формулою, яка в теорії похибок впливає з виду функціональної залежності.

Було показано, що в таких випадках задачу можна звести до розв'язку системи лінійних рівнянь з не точно заданими параметрами. При цьому задача виявляється оберненою, некоректною, а розв'язок у принципі виявляється неточним, тобто таким, що належить до певної області значень.

У цьому розділі пропонується розв'язування такої задачі методом Монте-Карло. Показано, що запропонований метод дає змогу не тільки одержати розв'язок такої погано обумовленої задачі, але й одержати діапазон його допустимих значень. Результати розв'язку такої системи рівнянь даються у вигляді гістограм можливих значень цих невідомих. Виявилось, що максимум такої гістограми відповідає шуканому значенню невідомої величини, а ширина гістограми є характеристикою невизначеності результату.

2.6.1. Постановка задачі

Отже, сформулюємо задачу так. Є, наприклад, достатньо складний вимірювальний комплекс, для якого результат вимірювання залежить від багатьох чинників (факторів впливу), наприклад від режимів роботи окремих підсистем. Тобто відомо, які параметри впливають на результат, але як врахувати їхній вплив невідомо. Що відомо, так це експериментальний результат вимірювання при відомих значеннях параметрів впливу.

Для задачі, коли є формула, що пов'язує результат з факторами впливу, похибка розраховується через оцінку частинних похідних цієї функції багатьох аргументів (формула 2.6.2). Коли функціональна залежність невідома (тобто не можна прямо знайти ці частинні похідні), задачу оцінку похибки вимірювань можна звести до задачі знаходження частинних похідних на основі виразу для її диференціала [1]. Нижче буде показано, що вона відноситься до класу некоректних задач, для яких не існує ні точних розв'язків, ні методів оцінки точності результату. І при розв'язуванні її можливі грубі помилки. Запропоновано метод вирішення проблеми, оснований на

статистичній обробці результатів розрахунків, проведених з використанням методів Монте-Карло.

Використаємо таку схему розгляду проблеми. Виберемо відому задачу з відомою аналітичною функцією залежності результату від параметрів впливу і розрахуємо для неї шукану невизначеність результату за стандартними формулами. Потім уявимо, що ця функція невідома, а відомо лише про наявність таких параметрів впливу, застосуємо для розв'язання задачі методи Монте-Карло [2] і порівняємо результати.

Продовжуючи [1], розглянемо проблему оцінки похибки вимірювання при невідомій аналітичній залежності функції виходу від параметрів входу на прикладі добре відомої з підручників [3] задачі бомбометання, в якій функція виходу X відома, а саме:

$$X(h, C, v) = v \sqrt{\frac{2h}{g} (1 - 1,8 \cdot 10^{-5} Ch)} \quad , \quad (2.6.1)$$

де v – швидкість літака, h – його висота, C – поправочний коефіцієнт, g – прискорення вільного падіння. Ці величини відомі з певною точністю. Їхні середні значення і стандартні відхилення є відповідно: швидкість $v_0 = 152$ м/с, $\sigma_v = 1$ м/с; висота $h_0 = 3\,950$ м, $\sigma_h = 40$ м; коефіцієнт $C_0 = 1$, $\sigma_c = 0,1$. Ставиться задача, знаючи неточність заданих параметрів h , C , v , знайти похибку оцінки X – величини відносу бомби від точки її скидання.

Стандартна формула для оцінки похибки для функції багатьох змінних – диференціальна формула оцінки похибки [4] – у нашому випадку має вигляд

$$\varepsilon = \sqrt{\left(\frac{\partial X}{\partial h}\right)^2 \sigma_h^2 + \left(\frac{\partial X}{\partial v}\right)^2 \sigma_v^2 + \left(\frac{\partial X}{\partial C}\right)^2 \sigma_c^2} \quad . \quad (2.6.2)$$

Прямим диференціюванням формули (2.6.1) одержимо

$$\frac{\partial X}{\partial h} = 0,429; \quad \frac{\partial X}{\partial v} = 26,4; \quad \frac{\partial X}{\partial C} = -307. \quad (2.6.3)$$

Розрахована за формулою (2.6.2) з наведеними вище (2.6.3) параметрами похибка $\varepsilon = 31,47$.

Тобто для оцінки похибки необхідно знати частинні похідні функції X . При невідомому вигляді цієї функції їх пропонується [1] визначати з формул для її скінченних приростів [5, с. 390]:

$$\begin{aligned} \Delta X_1 &= \frac{\partial X}{\partial v} \Delta v_1 + \frac{\partial X}{\partial h} \Delta h_1 + \frac{\partial X}{\partial C} \Delta C_1, \\ \Delta X_2 &= \frac{\partial X}{\partial v} \Delta v_2 + \frac{\partial X}{\partial h} \Delta h_2 + \frac{\partial X}{\partial C} \Delta C_2, \\ \Delta X_3 &= \frac{\partial X}{\partial v} \Delta v_3 + \frac{\partial X}{\partial h} \Delta h_3 + \frac{\partial X}{\partial C} \Delta C_3. \end{aligned} \quad (2.6.4)$$

Задаючи прирости (параметри впливу) Δv , Δh , ΔC у формулі (2.6.4), які можуть приймати довільні значення за умови їхньої малості порівняно із самою змінною, відповідні їм величини ΔX_i знаходять з експерименту.

Можна задатися наведеними в перших трьох рядках табл. 2.6.1 значеннями Δh_i , Δv_i , ΔC_i і розрахувати для них за формулою (2.6.1) ΔX_i .

Система рівнянь (2.6.4) є системою трьох рівнянь з трьома невідомими,

Таблиця 2.6.1. Значення параметрів розрахунків

Номер експерименту	Значення факторів впливу		
	Δh_i	Δv_i	ΔC_i
1	40	3	0,15
2	40	2	0,10
3	40	1	0,20
4	-80,0	2,0	0,15
5	-40,0	1,0	0,10
6	-120,0	3,0	0,20

яка елементарно розв'язується. Одержаний результат $\frac{\partial X}{\partial h} = 0.4375$,

$\frac{\partial X}{\partial v} = 26,1$, $\frac{\partial X}{\partial C} = -314$ близький до наведених вище теоретичних значень частинних похідних (2.6.3), які одержуються прямим диференціюванням формули (2.6.1).

На перший погляд здається, що задачу справді розв'язано. Однак слід пам'ятати, що параметри системи (2.6.4) Δh_i , Δv_i , ΔC_i в межах своїх змін можуть приймати довільні значення, тобто не тільки ті, які використано в табл. 2.6.1. Хоча одержаний для вказаних приростів результат і відхиляється від "точного" на $\sim 2\%$, залишилось невідомим, яке може бути максимальне відхилення при інших їхніх комбінаціях. І, наприклад, якщо систему рівнянь (2.6.4) скласти з параметрів, які наведені в рядках 4 - 6 табл. 2.6.1, то така система не має розв'язку в звичайному розумінні взагалі – її матриця сингулярна. Не важко підібрати прирости з діапазону їхніх можливих значень, при яких результат не буде мати нічого спільного з результатами формули (2.6.3).

Таким чином, використання описаного вище методу для визначення частинних похідних може дати неконтрольований результат. Продемонструємо цю проблему прямими розрахунками. Будемо випадковим чином генерувати прирости Δv_i , Δh_i , ΔC_i , $i = 1, 2, 3$ з діапазонів значень їхніх $\pm \sigma_i$, розраховувати для них ΔX_i і розв'язувати систему (2.6.4) стандартним методом Крамера [6]. Одержані значення частинних похідних показано у

вигляді гістограм їхнього розподілу на рис. 2.6.1. Відповідно розраховані паралельно за формулою (2.6.2), в яку підставлялися ці змінні значення частинних похідних, похибки показано на рис. 2.6.2. При цих розрахунках значення приростів бралось з машинною точністю, що дозволило при автоматизованому методі розв'язування уникнути ситуацій, коли детермінант системи точно дорівнював нулю і машина зупиняється, і таким чином провести розрахунки і побудувати графіки. Якщо брати прирости з точністю вимірювань, тобто до двох знаків після коми, провести такий аналіз методом

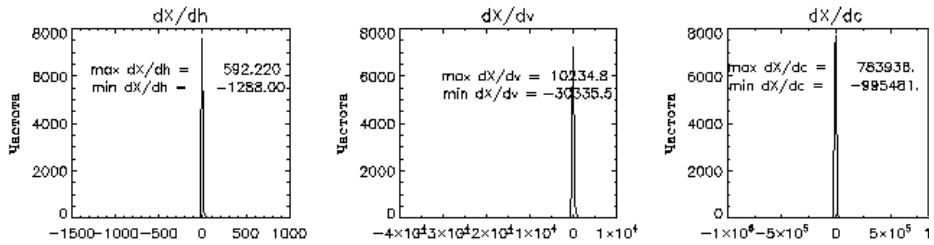


Рис. 2.6.1. Гістограми розподілу значень частинних похідних для випадкових значень приростів змінних Δv_i , Δh_i , ΔC_i при розв'язуванні методом Крамера. Указано також максимальні і мінімальні значення шуканих похідних, які були присутні у вибірці.

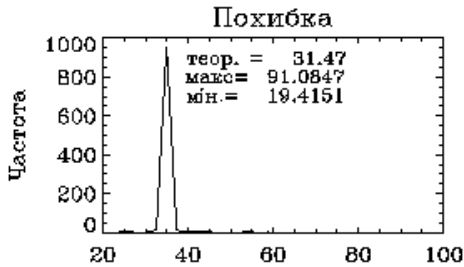


Рис. 2.6.2. Гістограми розподілу значень похибки для випадкових значень частинних похідних при розв'язуванні методом Крамера.

Крамера практично неможливо, оскільки комбінації параметрів, які дають нульовий детермінант, зустрічаються досить часто. На графіках по осі абсцис, наприклад для dX/dh , показано діапазон значень від -1500 до 1000. Це означає, що серед розрахованих значень зустрічаються і такі величини, хоча і рідко, тоді як максимуму

гістограми відповідає значення $\frac{\partial X}{\partial h} \cong 0,43$, яке фактично дорівнює

теоретичному з (2.6.3). Одержані значення частинних похідних мають досить великий розкид, що веде до великих розкидів розрахованої за формулою (2.6.2) похибки при поодиноких вимірюваннях (три вимірювання для розв'язання системи рівнянь (2.6.4)). Проте отримані гістограми мають максимум при значеннях, близьких до теоретичних значень формули (2.6.3), наведених вище. Це відкриває шлях до вирішення проблеми: потрібно з використанням методу Монте-Карло коректно розв'язати систему рівнянь (2.6.4) і з максимумів отриманих гістограм визначити похибки, які будуть аналогом значень аналітичних частинних похідних у формулі (2.6.3).

2.6.2. Метод розв'язування і оцінка його точності

Проблема полягає в тому, що система рівнянь (2.6.4) записана для так званої прямої задачі: розрахунків величини ΔX за відомими частинними похідними, а знаходження частинних похідних з системи рівнянь (2.6.4) є вже оберненою задачею. Оскільки величини приростів Δv_i , Δh_i , Δc_i можуть приймати довільні значення, система рівнянь (2.6.4) має нескінченне число рівнянь для трьох невідомих. Така задача відноситься до класу некоректних задач [7], у яких невеликим змінам параметрів можуть відповідати великі зміни результату, що ми і спостерігаємо. Близькі до нуля значення детермінанта в даній ситуації фактично означають його рівність нулю в межах точності розрахунків, а систему рівнянь (2.6.4) не можна розв'язувати методом Крамера. Тобто її немає сенсу розв'язувати й іншими методами, які дозволяють обходити сингулярності, тому що в дійсності отримані при цьому конкретні розв'язки можуть не мати фізичного значення (у некоректних задачах одержують деяку область можливих розв'язків і найбільш правильний серед них визначають з додаткових міркувань).

Але, як уже стало ясно, нам потрібно отримати не окремі розв'язки при конкретних значеннях параметрів, а набір розв'язків, які ми представимо у вигляді гістограми, і тому для розв'язування системи рівнянь (2.6.4) з використанням комп'ютера використаємо метод врахування сингулярних значень (singular value determination (SVD)) [8]. Оскільки, як відомо [7, 8], некоректні задачі розв'язуються наближено, ми ставимо за мету не тільки дати спосіб оцінки частинних похідних, а й оцінити межі похибки, обчисленої за формулою (2.6.2) за умови, що значення частинних похідних задаються в певних діапазонах. Тобто оцінити найбільш імовірне значення результату і межі його невизначеності.

Розрахунки, проведені за описаною вище схемою, але не методом Крамера, а методом SVD, дали результат, аналогічний показаному на рис. 2.6.2, проте у цьому методі розв'язуванню піддається система типу (2.6.4) при будь-яких комбінаціях параметрів. Підкреслимо ще раз, що наявність самого по собі розв'язку системи рівнянь (2.6.4) ще не означає розв'язання задачі, тому що, як було показано вище, через некоректність задачі результат може значно відхилитися від правильного розв'язку (див. рис. 2.6.1) і лише перебір великого числа параметрів і побудова гістограм дає можливість знайти правильний результат. Максимуми гістограм, як уже зазначалося, відповідають теоретичним значенням, тому найпростіший шлях вирішення задачі знаходження шуканих похідних – узяти їхні значення максимумах гістограм.

Результат побудови гістограм розв'язків, одержаних методом SVD, наведено на рис. 2.6.3.

Одержані в максимумах гістограм такі значення:

$$\frac{\partial X}{\partial h} = 0.43, \quad \frac{\partial X}{\partial v} = 26,42, \quad \frac{\partial X}{\partial c} = -309,14 \text{ у межах } 2\% \text{ збігаються}$$

з теоретичними результатами, що наведені у формулах (2.6.3). Підкреслимо при цьому, що оскільки ці величини визначаються з максимуму гістограми, можна вважати, що невизначеність цих величин визначається шириною гістограми.

2.6.3. Оцінка необхідної кількості вимірювань

Для побудови гістограм потрібен досить великий масив даних, тобто велике число вимірювань для складання системи рівнянь (2.6.4). Але на практиці реально виконується обмежене число вимірювань; тоді скільки їх треба, щоб отримати якісну гістограму?

За результатами вимірювань можна скласти систему типу (2.6.4), але з більшою кількістю рівнянь, яка стає системою багатьох рівнянь з трьома невідомими. Оскільки при розглянутих розрахунках похідних використовується по три рівняння, то для незалежних розрахунків можна використовувати різні комбінації по три рівняння з розширеної системи (2.6.4). Виявилось, що якісні гістограми з чітким максимумом утворюються, коли вже система (2.6.4) містить 13 – 15 рівнянь. Одержане при цьому відхилення від табличних даних формули (2.6.3) не перевищує 2 %. Зауважимо, що оскільки при цьому гістограма вимальовується вже достатньо чітко, подальше збільшення числа рівнянь (числа вимірювань) не веде до покращення результату. Але яка невизначеність цього результату? Межі невизначеності можна встановити, просто оцінивши (програмно, за даними розрахунків) ліві і праві значення гістограм на заданому рівні, скажімо, у межах 95 %. Тобто відкинувши ті максимальні й мінімальні значення елементів гістограми, що в сумі дають 5 %.

2.6.4. Висновки

Запропоновано метод оцінки невизначеності (похибки) вимірювань для ситуації, коли результат є функцією багатьох змінних, проте сам вигляд функціональної залежності невідомий. Показано, що в цій ситуації визначення частинних похідних, знання яких є необхідними для розрахунку похибки, веде до некоректної математичної задачі, розв'язок у принципі виявляється неточним, тобто таким, що належить до певної області значень.

Показано, що одержати коректний результат можливо шляхом застосування статистичної обробки результатів розрахунків, виконаних із використанням моделювання на основі методів Монте-Карло.

Показано, що запропонований метод дає змогу не тільки одержати розв'язок такої погано обумовленої задачі, але й одержати діапазон його допустимих значень. Результати розв'язку такої системи рівнянь даються у вигляді гістограм можливих значень цих невідомих. Виявилось, що максимум гістограми відповідає шуканому значенню невідомої величини, а ширина гістограми є характеристикою невизначеності результату.

Як результат, отримано не тільки найбільш імовірне значення та межі його невизначеності, а й оцінено мінімальне число вимірювань, необхідне для одержання надійних даних, що неможливо було отримати іншим способом.

Описаний підхід легко перенести на просту задачу аналізу радіаційних вимірювань. При аналізі невизначеності результату методики вимірювань деяким приладом у стандартному підході враховується п'ять факторів впливу: паспортна похибка приладу, нестабільність температури навколишнього середовища, нестабільність джерела живлення, нестабільність вологості, змінність геометрії вимірювань. Причому діапазон змін впливу цих факторів є відомим (занесені в паспорт). В описаному підході вважається, що таких даних немає. Тоді для оцінки невизначеності результатів вимірювань таким приладом можна провести 10 – 20 вимірювань, контрольовано в різних комбінаціях змінюючи величини цих факторів впливу. І для цих результатів побудувати і розв'язати систему рівнянь 2.6.4.

СПИСОК ЛІТЕРАТУРИ

1. *Ленюк Г. К.* Визначення похибки вимірювання при невідомій аналітичній залежності функції виходу від параметрів входу // Укр. метрол. журн. – 2008. – № 4. – С. 59 – 60.
2. *Скорбун А. Д., Сплошной Б. М.* Коментар до статті про визначення похибки вимірювання при невідомій аналітичній залежності функції виходу від параметрів входу // Укр. метрол. журн. – 2010. – № 2. – С. 57 – 59.
3. *Вентцель Е. С.* Теория вероятностей. – М.: Наука, 1969. – 576 с.
4. *Щиголов Б. М.* Математическая обработка наблюдений. – М.: Наука, 1969. – С. 344.
5. *Фихтенгольц Г. М.* Курс дифференциального и интегрального исчисления. Т. 1. – М.: Наука, 1966. – 607 с.
6. *Шилов Г. Е.* Введение в теорию линейных пространств. – М., Л.: Гос. изд-во техн.-теор. лит., 1952. – 384 с.
7. *Тихонов А. Н., Арсенин В. Я.* Методы решения некорректных задач. – М.: Наука, 1988. – 287 с.
8. *Голуб Дж., Лоун Ч. Ван.* Матричные вычисления. – М.: Мир, 1999. – 548 с.

2.7. Використання методу джекнайф (jackknife) для оцінки невизначеності середнього

Проілюструємо описаний у першій частині прийом джекнайф для визначення необхідного числа навісок на конкретному прикладі оцінки середньої забрудненості території, на якій було зроблено деяке число вимірювань або, що є еквівалентним, визначення невизначеності оцінки вмісту у пробі потрібної компоненти (загальної активності в даному випадку). Прийом було модифіковано і розширено, а саме показано, що для даної вибірки, у припущенні, що вона відображає загальну картину, можна побудувати зону, яку можна назвати коридором похибок і яка буде відображати розкид даних для даного числа проб (для даного числа елементів вибірки).

Наша задача – проілюструвати метод джекнайф на реальних даних. Було відібрано зразок керна із свердловини, яка проходить через шар радіоактивного ґрунту, для якого треба визначити загальну активність і межі невизначеності результату. Підкреслимо, що іншої проби (інший kern) взяти не можна й оцінку треба зробити по наявній пробі. Крім того, kern у цілому проаналізувати не можна, можна аналізувати лише невелику пробу (навіску) з нього. Опис задачі є таким.

1. Результати лабораторних аналізів проб досліджених техногенних ґрунтів локальної зони об'єкта "Укриття" часто мають широкий розкид значень. Математичне моделювання неоднорідно забруднених окремими "гарячими" частками ґрунтів показало, що функція розподілу для вибірок з таких проб залежить від моделі й тому немає підстав вважати, що вони повинні описуватися нормальним або логнормальним розподілами. Отже, до них не можна застосовувати звичайні прийоми оцінки параметрів розподілу, розроблені у припущенні, що розподіл є відомим.

2. Необхідно запропонувати метод визначення необхідної кількості проб для коректного визначення питомої активності зразка в цілому, а також оцінки похибки (невизначеності) результату. Метод не повинен залежит від виду функції розподілу щільності ймовірності.

Весь зразок керна було розділено на 32 навіски, які будемо вважати генеральною сукупністю. Отже, це задача, де у принципі не можна розвинути теорію, основу на переході до нескінченності, як це робиться в теорії ймовірностей. Значення питомої активності всієї проби (визначене як сумарна активність для всіх навісок A_0) відоме, проте нас цікавить, якою є величина невизначеності цього результату і чи не можна зменшити кількість навісок, тобто об'єм проби і число аналізів, але щоб при цьому не змінилася відома нам її питома активність. Критерієм допустимості такої операції при цьому може бути прийнятне відхилення питомої активності скороченої проби від істинного значення. Результати лабораторних аналізів наведено в табл. 2.7.1.

Питома активність кожної навіски відома і всі вони утворюють вибірку, для якої ми зафіксуємо порядок розташування навісок. Починаючи з першого

номера вибірки № 1 з табл. 2.7.1, будемо послідовно по одній складати результати вимірювань активності й маси для кожної навіски й визначати їхню сумарну питому активність. Ясно, що якщо об'єм зразка досить великий, а число проб, на які він розділений, також досить велике, то, починаючи з деякого числа проб, середня питома активність буде наближатися до визначеного вище її справжнього значення A_0 .

За результатами цих розрахунків побудуємо графік, на якому по горизонталі відкладено сумарну кількість навісок, а по вертикалі - їхню накопичену середню питому активність. Тобто побудуємо кумулятивну криву. У розглянутій (див. табл. 2.7.1) вибірці № 1 пробам присвоєно порядковий номер. У такому вигляді ці результати можна розглядати як ряд послідовних вимірювань. Оскільки навіски могли бути пронумеровані в іншому порядку, побудуємо такі ж графіки для декількох незалежних випадкових вибірок (тобто випадковим способом перенумеруємо навіски й знову будемо складати, починаючи з першого номера). На рис. 2.7.1, а зображено такі залежності для досить великої кількості таких перенумерованих вибірок.

Таблиця 2.7.1. Проаналізовані вибірки значень активності проб ґрунтів. Вибірка № 1 упорядкована по зростанню питомої активності й пронумерована

Вибірка № 1								Вибірка № 2
Номер навіски	Маса, г	Активність, Бк	Питома активність, Бк/г	Номер навіски	Маса, г	Активність, Бк	Питома активність, Бк/г	Завантаження електронметра, імпл/с
1	2,03	$1,08 \cdot 10^6$	$5,30 \cdot 10^5$	17	2,06	$1,69 \cdot 10^6$	$8,20 \cdot 10^5$	2681,7
2	2,11	$1,24 \cdot 10^6$	$5,90 \cdot 10^5$	18	2,19	$1,86 \cdot 10^6$	$8,5 \cdot 10^5$	2682,8
3	2,18	$1,31 \cdot 10^6$	$6,00 \cdot 10^5$	19	2,08	$1,79 \cdot 10^6$	$8,60 \cdot 10^5$	2686,6
4	2,05	$1,31 \cdot 10^6$	$6,40 \cdot 10^5$	20	2,14	$1,84 \cdot 10^6$	$8,60 \cdot 10^5$	2688,3
5	2,06	$1,36 \cdot 10^6$	$6,60 \cdot 10^5$	21	2,22	$1,91 \cdot 10^6$	$8,60 \cdot 10^5$	2697,9
6	2,09	$1,40 \cdot 10^6$	$6,70 \cdot 10^5$	22	2,13	$1,94 \cdot 10^6$	$9,10 \cdot 10^5$	2700,8
7	2,14	$1,46 \cdot 10^6$	$6,80 \cdot 10^5$	23	2,13	$1,98 \cdot 10^6$	$9,30 \cdot 10^5$	2704,2
8	2,13	$1,45 \cdot 10^6$	$6,80 \cdot 10^5$	24	2,09	$2,03 \cdot 10^6$	$9,70 \cdot 10^5$	2707,7
9	2,21	$1,50 \cdot 10^6$	$6,80 \cdot 10^5$	25	2,06	$2,06 \cdot 10^6$	$1,00 \cdot 10^6$	2711,6
10	2,12	$1,46 \cdot 10^6$	$6,90 \cdot 10^5$	26	2,15	$2,15 \cdot 10^6$	$1,00 \cdot 10^6$	2713,2
11	2,06	$1,46 \cdot 10^6$	$7,10 \cdot 10^5$	27	2,1	$2,31 \cdot 10^6$	$1,10 \cdot 10^6$	2716,7
12	2,12	$1,53 \cdot 10^6$	$7,20 \cdot 10^5$	28	2,02	$2,22 \cdot 10^6$	$1,10 \cdot 10^6$	2722,5
13	2,07	$1,53 \cdot 10^6$	$7,40 \cdot 10^5$	29	2,07	$2,69 \cdot 10^6$	$1,30 \cdot 10^6$	2723,3
14	2,14	$1,63 \cdot 10^6$	$7,60 \cdot 10^5$	30	2,19	$3,29 \cdot 10^6$	$1,50 \cdot 10^6$	2723,8
15	2,13	$1,64 \cdot 10^6$	$7,70 \cdot 10^5$	31	2,11	$3,38 \cdot 10^6$	$1,60 \cdot 10^6$	2725,8
16	2,22	$1,80 \cdot 10^6$	$8,10 \cdot 10^5$	32	2,09	$3,76 \cdot 10^6$	$1,80 \cdot 10^6$	2726,2

Як видно, криві мають тенденцію до усереднення результату з ростом числа проб, однак при малому числі навісок отримана середня питома активність може досить істотно відхилитися від справжнього значення.

Криві на рис. 2.7.1, а заповнюють у площині рисунка обмежений простір, який можна розглядати як область допустимих значень накопиченої питомої активності проби. Якщо для заповненої кривими області взяти крайні верхні й нижні значення для кожного значення кількості навісок то одержимо дві криві, простір між якими можна назвати областю можливих відхилень для заданого числа навісок, або областю похибок (див. рис. 2.7.1, б). Як видно з рисунка,

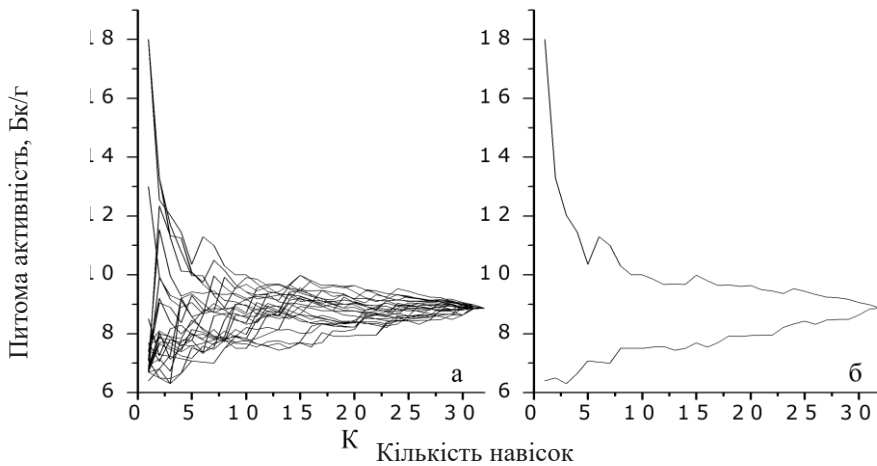


Рис. 2.7.1. Залежність питомої активності від кількості навісок та порядку їхнього підсумовування.

починаючи приблизно з 12 – 15 навісок, збільшення їхнього числа слабо впливає на кінцевий результат, не дивлячись на великий розкид результатів.

Статистичне моделювання

У даному дослідженні з самого початку припускалося, що розглянута вибірка повністю характеризує зразок і для неї відомо "справжнє" значення питомої активності, одержане як сумарна активність усіх 32 навісок, поділена на їхню загальну масу. Асимптотична збіжність кривих на рис. 2.7.1 ще не свідчить про те, що відібраний зразок керна дійсно є репрезентативним, тобто характеризує весь kern: якісно картина, зображена на рис 2.7.1, характерна для будь-яких вибірок. Надмірність вибірки характеризується тим, скільки членів випадкової послідовності можна відкинути без помітної зміни її статистичних властивостей. Тому проведемо наступний модельний експеримент із застосуванням методу джекнайф. З вибірки № 1, "перемішаної" так, як вказувалося вище, випадковим чином видалимо один з елементів. Для такої

скороченої вибірки визначимо значення питомої активності. Повторимо таку оцінку, послідовно перемішуючи (перенумеровуючи) вибірку № 1 і видаляючи випадковим чином вибрані один, два, три і т.д. елементи. (Підкреслимо, що в цьому експерименті, як і в зображеному на рис. 2.7.1, кожен раз беруться до уваги всі елементи вибірки № 1. Щоб підкреслити різницю між похибкою окремого вимірювання (середньоквадратичне відхилення) і похибкою оцінки середнього значення (тобто невизначеністю оцінки шуканого значення загальної активності проби) також застосуємо статистичне моделювання. На рис. 2.7.2, а показано результат розрахунку питомої активності для кожної вкороченої вибірки, отриманий шляхом багаторазової зміни послідовності її елементів.

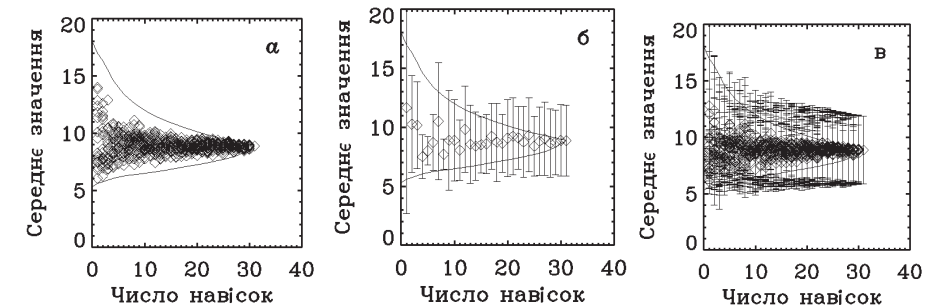


Рис. 2.7.2. Можливі реалізації середніх значень для вибірки № 1, яка випадковим чином скорочується послідовно на один елемент.

Область теоретично можливих значень на рис. 2.7.2, а показано суцільними лініями - гіпотетичними ситуаціями, коли реалізується вибірка в строго спадаючому або зростаючому порядку. Видно, що в реальності більш імовірним є групування результатів у центральній частині області можливих значень. Тобто межі можливого розкиду значень обмежено чорною областю.

Цей результат принципово відрізняється від звичайної статистичної оцінки похибки для вибірки № 1 як середньоквадратичного відхилення від середньоарифметичного значення. Побудуємо графік для однієї з реалізацій послідовності елементів вибірки й нанесемо "вуса" середньоквадратичної похибки (див. рис. 2.7.2, б). Середньоквадратичне відхилення значно перевищує встановлені допустимі межі похибки, що знаходяться між двома суцільними лініями (строго кажучи, ці межі є смугами з шириною, яка дорівнює похибці вимірювання кожного елемента, одержаній при аналізі навісок вибірки № 2). Особливо чітко це видно для всієї вибірки з 32 елементів. Для неї значення питомої активності (якщо знехтувати похибкою вимірювання, яка приблизно однакова для всіх елементів вибірки) було визначено точно і не залежить від розкиду значень елементів вибірки. Зобразимо те ж саме, що і на рис. 2.7.2, а, але із нанесеними для кожної точки "вусами" (рис. 2.7.2, б, в). Зона похибок, що окреслюється "вусами", значно перевищує зону змінності, установлену на рис. 2.7.2, а, і яка є більш

правильною оцінкою похибки результатів вимірювань питомої активності сумарної проби. Причина цього полягає в тому, що в нашій задачі аналізується невизначеність оцінки середнього.

Зауважимо, що графіки на рис. 2.7.2 одержано як результат маніпуляцій з єдиною експериментальною вибіркою. Його можна вважати графічним методом "джекнайфу" і бутстрапу, які, використовуючи статистичне моделювання, дозволяють зробити оцінку похибки визначення середніх значень параметрів вибірки.

Метод дає змогу визначити мінімальну кількість проб, необхідну для оцінки середньої питомої активності зразка ґрунту, на основі критерію допустимої величини похибки результату. Тобто, якщо на рис. 2.7.2, а провести вертикальну лінію для значення 10 навісок, переріз графіка буде визначати, в яких межах буде знаходитися результат. Якщо ж користуватися середньоквадратичним відхиленням як оцінкою розкиду, коридор похибок буде визначатися зоною «вусів» на рис. 2.7.2, в.

Метод може бути корисним, наприклад при плануванні повторних робіт по аналізу забрудненості території, на якій уже були зроблені такі вимірювання. З такого аналізу можна прийняти рішення про збільшення або зменшення числа необхідних вимірювань.

Що стосується аналізу уже проведених вимірювань, коли було отримано результати для N проб, то невизначеність результату буде визначатися шириною коридору похибок для $N-1$ проб на графіку типу 2.7.2, а.

2.8. Метод геохімічної статистики для пошуку кореляційних зв'язків між хімічним складом та концентраціями ^{90}Sr в підземних водах

За багаторічними результатами спостережень у рамках радіогідроecологічного моніторингу на проммайданчику ЧАЕС, об'ємні активності ^{90}Sr у пробах підземних вод із спостережних свердловин фіксуються в широких межах значень від 1-2 до 400 – 3800 Бк/л [1]. Надходження радіостронцію до водоносного горизонту відбувається за рахунок інфільтрації атмосферних опадів через радіоактивно забруднені ґрунти, втрат радіоактивно забрудненої води з приміщень зруйнованого 4-го енергоблоку ЧАЕС та з водонесучих комунікацій. Упродовж десятиліть, що пройшли після аварії, радіоекологічна обстановка у водоносному горизонті на проммайданчику ЧАЕС стабілізувалась у частині надходження радіонуклідів, розповсюдження їх у водоносному горизонті, балансу між сорбційними та десорбційними процесами. Тому випадки різкого зростання у 200 – 500 разів концентрацій радіостронцію у водоносному горизонті, найбільш вірогідно, пов'язані зі зміною геохімічних умов у підземних водах, у тому числі балансу сорбції/десорбції, а не з додатковим інтенсивним його надходженням з невідомих джерел.

Враховуючи те, що зв'язок між сорбційними/десорбційними властивостями ґрунтів і міграцією іонів є відомим явищем, ми висуваємо такі припущення:

підвищення концентрацій радіонуклідів у підземних водах пов'язані з формуванням умов, що забезпечують зниження сорбційної здатності ґрунтів, а не з появою додаткових джерел надходження радіонуклідів;

зниження сорбційної здатності ґрунтів до ^{90}Sr може відбуватись унаслідок впливу

окремих складових хімічного складу підземних вод та їхнього кислотно-лужного стану.

Інакше кажучи, повинен простежуватись тісний кореляційний зв'язок між ^{90}Sr та макрохімічним складом підземних вод. Але прямий розрахунок коефіцієнта кореляції Пірсона за допомогою програмного комплексу Excel дав результати в інтервалі від 0 до 0,5 у більшості проаналізованих випадків, що повинно свідчити про відсутність кореляції чи її низьку ймовірність, а відповідно, і відсутність взаємного впливу окремих елементів хімічного складу та концентрацій ^{90}Sr в підземних водах. Проте на деяких отриманих нами графіках (рис. 2.8.1) наявність кореляції видно і без статистичної обробки. Причини, з яких не підтверджується тісний зв'язок між концентраціями ^{90}Sr та деякими основними іонами і величиною рН при розрахунку коефіцієнта кореляції Пірсона, очевидно, полягають у тому, що результати спостережень не підпорядковуються нормальному розподілу (розподіл Гаусса). Передумовами використання коефіцієнта кореляції Пірсона є [2, 3]: а) усі спостереження взаємно незалежні; б) спостереження мають нормальний закон розподілу. Однак розподіл даних, що наведені на рис. 2.8.1,

аж ніяк не можна вважати незалежними й однаково розподіленими (тобто які мають однаковий закон розподілу). А отже до них не можна застосовувати підходи стандартної математичної статистики.

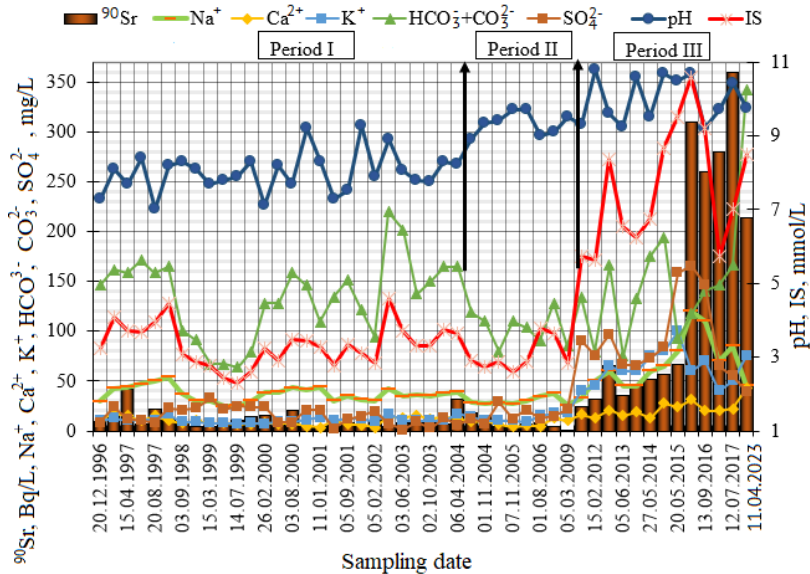


Рис. 2.8.1. Динаміка об'ємної активності ^{90}Sr , величини рН та компонентів хімічного складу підземних вод із свердловини 2-Г.

Таким чином, для кількісної оцінки наявності шуканих кореляційних зв'язків виникла необхідність звернутись до інших методів визначення тісноти кореляційних залежностей між концентраціями радіостронцію та основними іонами і значеннями рН.

Вплив окремих складових хімічного складу на формування високої міграційної здатності ^{90}Sr вивчається в ряді робіт [4, 5, 6, 7, 8]. Але по ряду факторів літературні дані суперечливі і по окремих положеннях не підтверджуються моніторинговими дослідженнями, що виконуються на проммайданчику ЧАЕС [9, 10, 11].

У роботі [4] показано, що за рахунок сорбційних процесів швидкість міграції ^{90}Sr у водоносному горизонті еолового піску уповільнюється до $\approx 9\%$ швидкості потоку ґрунтових вод ($K_d \approx 2$ мл/г). Але сорбційні властивості ґрунтів, що складають водоносний горизонт можуть значно знижуватись. Так, у роботі [5] стверджується, що присутність у водних розчинах іонів Na^+ і Ca^{2+} істотно знижує коефіцієнт розподілу (K_d) і адсорбцію радіонуклідів ^{137}Cs і ^{90}Sr .

Для монтморилоніту та каоліну сорбція радіонуклідів ^{137}Cs та ^{90}Sr є максимальною при рН водного середовища більше 4 [6]. Підвищення концентрацій катіонів Na^+ і Ca^{2+} , які характерні для природних водних середовищ, призводить до зниження сорбції ^{137}Cs та ^{90}Sr , що зумовлено конкурентним впливом цих катіонів. Також підтверджено, що рухливість ^{90}Sr зростає при підвищенні вмісту Ca^{2+} у воді. Для мінералів, що модифіковані

гуміновими кислотами (ГК) та $\text{Fe}(\text{OH})^{2+}$ максимальне значення K_d спостерігається в області рН 4 – 8, що свідчить про участь ГК і гідроксиду заліза (III) у зв'язуванні радіонуклідів.

Вивчення залежності ступеня поглинання ^{90}Sr різними ґрунтами і гірськими породами від рН в роботі [7] показало, що в значних кількостях стронцій починає сорбуватися ґрунтами при рН > 7. Максимальне поглинання для більшої частини вивчених ґрунтів і гірських порід було отримано при рН = 10 – 11. При менших значеннях рН ступінь поглинання різко знижувалася до мінімальної. Адсорбція відбувається переважно за механізмом іонного обміну. Також у роботі [8] вказано, що «прогресуюче зниження в останні роки концентрацій основних катіонів (зокрема Са) і збільшення рН у підземних водах приводить до суттєвого зростання сорбційних коефіцієнтів розподілу стронцію-90 в водоносному горизонті». Однак слід зазначити, що на протигагу цьому в роботі [9] стверджується, що концентрація урану у пробах ґрунтових вод підвищується у 2 – 12 разів в умовах сильнолужного середовища при значенні рН вище 9 – 10. Також, у роботі [10] концентрації ^{90}Sr при рН більше 9,5 – 10 збільшуються в 200 – 500 разів. У роботі [11] зроблено висновок, що за даними вивчення розподілу і форм знаходження радіонуклідів у підземних водах одним із чинників формування високих концентрацій ^{90}Sr в сильнолужному середовищі є десорбція ^{90}Sr із грубодисперсних частинок ґрунтів та з колоїдів. Вірогідно, основним чинником процесу десорбції, а значить і запобігання сорбції радіостронцію ґрунтами, є висока концентрація в розчині гідроксиду OH^- . За даними термодинамічного моделювання концентрація стронцію у вигляді сполуки SrOH^+ при «переході» із слаболужного до сильнолужного середовища збільшується в 57 разів. Таким чином, можливо підвищена водна міграція радіостронцію з сильнолужними підземними водами відбувається у вигляді комплексної сполуки SrOH^+ , але ці результати протирічать багатьом літературним даним.

Таким чином, за літературними даними, результати спостережень впливу хімічного складу підземних вод на міграційну здатність радіонуклідів досить суперечливі. Застосування розрахунків коефіцієнта кореляції Пірсона для вирішення цих задач теж не дає однозначного результату. Тому для кількісної оцінки наявності кореляційних зв'язків виникла необхідність звернутись до методів комп'ютерної статистики.

Отже, метою представленого розділу є демонстрація можливості застосування методів комп'ютерної статистики (метод геохімічної статистики) та математичне визначення впливу концентрацій окремих основних іонів хімічного складу підземних вод і їхнього кислотно-лужного стану на підвищення об'ємних активностей ^{90}Sr .

На рис. 2.8.2 також візуально простежується зв'язок між концентраціями Ca^{2+} та ^{90}Sr .

Задача полягає в тому, щоб кількісно оцінити взаємний зв'язок концентрацій ^{90}Sr , величини рН та деяких основних іонів, що впливають на міграційну здатність радіостронцію з підземними водами.

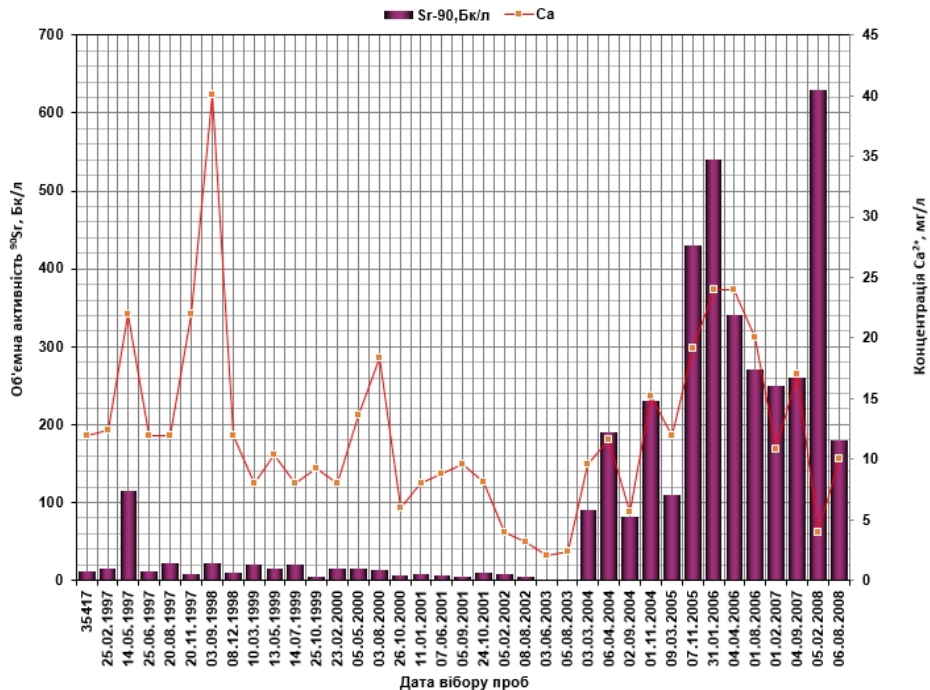


Рис. 2.8.2. Графіки динаміки об'ємної активності ⁹⁰Sr та концентрацій Ca²⁺ у пробах підземних вод із свердловини 4-Г.

На сьогоднішній день для аналізу таких даних розроблено методи комп'ютерної статистики [12, 13], в основі яких лежить підхід, що працювати треба з експериментальними даними, а не з формулами теорій, які їх апроксимують [14]. Методи комп'ютерної статистики базуються на використанні методів Монте-Карло. Великою перевагою методу є те, що він дозволяє врахувати в моделі елемент випадковості без аналізу функції розподілу. Крім того, метод є робастним по відношенню до зміни різних параметрів, таких як розподіл випадкової величини, наявності «викидів» тощо.

Дані, які аналізуються, являють собою ряди багаторічних моніторингових вимірювань концентрацій K⁺, Na⁺, Ca²⁺, величини рН та об'ємних активностей ⁹⁰Sr у пробах підземних вод із свердловин 1-1А, 1-2А, 2-Г, 4-Г, 21-1А.

В якості прикладу (табл. 2.8.1) наводиться фрагмент фактичних даних, які отримані по свердловині 4-Г в процесі моніторингових досліджень. Ставиться задача перевірити наявність кореляції між об'ємною активністю ⁹⁰Sr, концентраціями іонів та величиною рН за допомогою методу комп'ютерної статистики. Такий підхід отримав назву – метод геохімічної статистики [1].

У комп'ютерній статистиці приймається, що ступінь кореляції між двома

Таблиця 2.8.1. Фрагмент результатів спостережень по свердловині 4-Г

Na, мг/л	K ⁺ , мг/л	Ca ²⁺ , мг/л	pH	⁹⁰ Sr, Бк/л
27,60	15,50	12,00	6,8	12
40,00	43,50	12,40	7,28	16
55,00	51,00	22,00	6,6	115
77,50	39,00	12,00	6,8	12
36,40	35,50	12,00	6,2	22
44,60	40,00	22,00	6,75	8
57,60	33,30	40,10	7	22

вибірками пропорційна сумі попарних добуток значень, з яких складаються ці вибірки. Алгоритм розрахунку розглянемо на прикладі табл. 2.8.1: будемо шукати кореляцію для фактичних даних моніторингових досліджень, наприклад між колонками для ⁹⁰Sr та Ca²⁺.

1. Розраховуються попарні добутки значень для ⁹⁰Sr та Ca²⁺ в кожному рядку в табл. 2.8.1.

2. Сума таких добутоків по всіх рядках позначається S₀ (табл. 2.8.2).

Таблиця 2.8.2. Фрагмент визначення суми попарних добутоків фактичних моніторингових даних S₀

Ca, мг/л	⁹⁰ Sr, Бк/л	Попарні добутки Ca ²⁺ та ⁹⁰ Sr
12,00	12	144
12,40	16	198,4
22,00	115	2530
12,00	12	144
12,00	22	264
22,00	8	176
40,10	22	882,2
Сума (S ₀)		4338,6

3. Припустимо, що наші дані повністю нескорельовані, тобто випадкові, і не пов'язані між собою. Тоді очевидно, що для будь-яких перестановок місцями значень концентрацій ⁹⁰Sr та Ca²⁺ між рядками повинні отримуватись рівноцінні випадкові результати значення суми попарних їхніх добутоків.

4. Якщо повторити цю процедуру розрахунку суми попарних добутоків, які позначимо S_k, переставляючи числа в колонці місцями багато тисяч разів, отримаємо велику вибірку значень S_k (табл. 2.8.3).

5. Далі утворимо масив різниць між цими сумами (S₀ - S_k).

6. Можна побудувати гістограму значень цих різниць (S₀ - S_k). Якщо числа в розглянутих колонках були випадкові і не пов'язані між собою, гістограма буде мати дзвоноподібну форму з максимумом в районі нуля.

Таблиця 2.8.3. Приклад визначення суми попарних добутків даних, що моделюються S_k

Ca, мг/л	^{90}Sr , Бк/л	Попарні добутки Ca^{2+} та ^{90}Sr
12	16	12
12,4	115	12,4
22	12	22
12	22	12
12	8	12
22	22	22
40,1	22	40,1
Сумма (S_{k1})		3608

7. Якщо величини в рядках попарно скорельовані, максимум гістограми буде зсунуто відносно нуля.

8. Якщо збільшення концентрацій Ca^{2+} або величини рН буде супроводжуватися зростанням об'ємних активностей ^{90}Sr , то максимум гістограми буде зсуватися в бік додатних значень ($S_0 - S_k$). При повній скорельованості фактичних значень концентрацій Ca^{2+} та ^{90}Sr у двох колонках, завжди ($S_0 > S_k$) і від'ємні значення ($S_0 - S_k$) не зможуть утворитися ні при яких перестановках чисел у колонці місцями. Таким чином, за допомогою комп'ютерного моделювання в цьому випадку відображається процес зменшення сорбції ^{90}Sr ґрунтами під впливом, наприклад, Ca^{2+} або рН.

9. Відповідно в ситуації повністю антискорельованих вибірок, коли зростанню значень в одній колонці відповідає зменшення значень у другій колонці, максимум гістограми буде зміщений у бік від'ємних значень, наближуючись до значення -1.

10. Якщо вибірки не повністю скорельовані, максимум гістограм, як уже вказувалося, буде зсунутий відносно нуля. Якщо максимум гістограми буде знаходитися в області ($S_0 - S_k$) > 0 (див. рис. 2.8.4, нижче), дані вважаються скорельованими. Ступінь скорельованості визначається співвідношенням числа від'ємних значень ($S_0 - S_k$) до їхнього загального числа, тобто до повного числа перестановок. У літературі ця величина має назву Р-критерій. При зростанні скорельованості вибірок число від'ємних значень ($S_0 - S_k$) буде зменшуватися і Р буде прямувати до нуля. Тобто чим ближче до нуля значення Р при аналізі даних за допомогою даного методу, тим більша скорельованість спостерігається між вибірками.

Повторимо, що у випадку нескорельованих вибірок значення ($S_0 - S_k$) будуть розташовані навколо нуля, тобто $P = 0,5$ (50 % значень від'ємних ($S_0 - S_k$) на 50 % значень позитивних).

Якщо максимум гістограми буде знаходитися в області ($S_0 - S_k$) < 0 (див. рис. 2.8.6, нижче), дані вважаються антискорельованими. Тобто відображається процес підвищення сорбції ^{90}Sr ґрунтами під впливом параметра, кореляція ^{90}Sr з яким аналізується. У цьому випадку ступінь антискорельованості (оберненої кореляції) визначається співвідношенням

числа додатних значень ($S_0 - S_k$) до їхнього загального числа, тобто до повного числа перестановок. Якщо співвідношення P буде значно менше, ніж 0,5, то в такому випадку можна говорити про обернену кореляцію між парами, що аналізуються. Для того щоб відрізнити ситуацію, в якій спостерігається пряма кореляція, від ситуації з оберненою кореляцією, у подальшому тексті P -критерій та величину ступеня кореляції $K = 1 - P$ для оберненої кореляції будемо давати зі знаком мінус.

Розглянемо вищенаведене на прикладі свердловини 1-2А (рис. 2.8.3, 2.8.4, 2.8.5).

На гістограмах по вертикалі – кількість повторювання однакових значень ($S_0 - S_k$), по горизонталі – значення ($S_0 - S_k$), які були отримані в процесі перестановок.

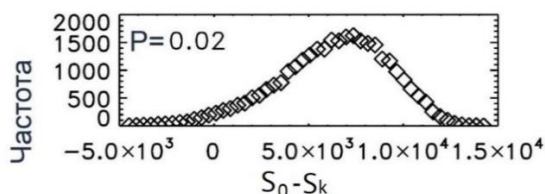


Рис. 2.8.3. Розрахунок кореляції між вибірками даних Ca^{2+} та ^{90}Sr .

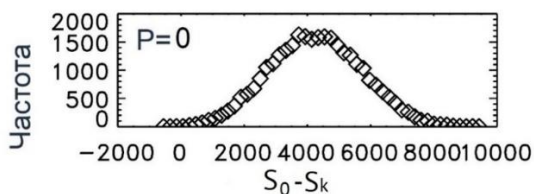


Рис. 2.8.4. Розрахунок кореляції між вибірками даних ^{90}Sr та рН.

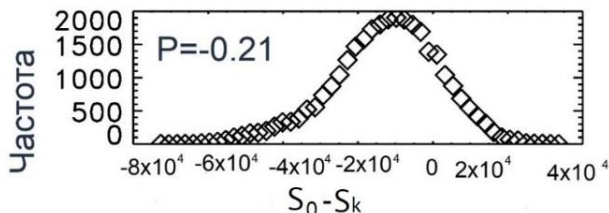


Рис. 2.8.5. Розрахунок кореляції між вибірками даних K^+ та ^{90}Sr .

Гістограма (див. рис. 2.8.3 і 2.8.4) даних вибірок дзвоноподібна, схожа на гауссову, однак не є такою тому, що часто має великий «хвіст» ліворуч від нуля.

Гістограма на рис. 2.8.5 є прикладом оберненої кореляції, коли зростання одного параметра (K^+) супроводжується зменшенням іншого параметра (^{90}Sr).

Як зазначалось вище, якщо дані не скорельовані, максимум такої гістограми для величини ($S_0 - S_k$) буде близький до нуля. Гістограма визначення кореляції між об'ємною активністю ^{90}Sr та концентрацією Ca^{2+} (див. рис. 2.8.3) має максимум у позитивній області і відносно «далеко» від нуля. Значення P

дорівнює – 0,02. Тобто ми спостерігаємо тісну кореляційну залежність об'ємних активностей ^{90}Sr від концентрацій Ca^{2+} зі ступенем кореляції $K=1-P=0,98$. У даному аналізі нуль – це точка відліку, яку ми беремо для оцінки величини зміщення максимально скорельованих пар даних відносно випадкових.

Аналогічний результат отримано при аналізі впливу рН на величину концентрації ^{90}Sr (див. рис. 2.8.4). Тільки величина P в цьому випадку дорівнює 0. А відповідно ступінь кореляції цих пар даних дорівнює 1.

На рис. 2.8.5 максимум гистограми знаходиться ліворуч від 0. А наведена на рисунку 2.8.5 величина $P = -0,21$ свідчить, що комбінації пар даних у цих вибірках мають обернену кореляцію. Ступінь антикореляції дорівнює значенню $-0,79$.

Також був проведений кореляційний аналіз зв'язку між концентраціями ^{90}Sr та величиною окислюваності (вмістом органічної речовини) (рис. 2.8.6) і, як результат, ми маємо зворотну кореляцію. Це підтверджується величиною $P = -0,08$ (див. рис. 2.8.6). Тобто ступінь кореляції K дорівнює $-0,92$, що свідчить про зворотний вплив розчинної органічної речовини на формування високих концентрацій ^{90}Sr в пробах підземних вод у даному випадку.

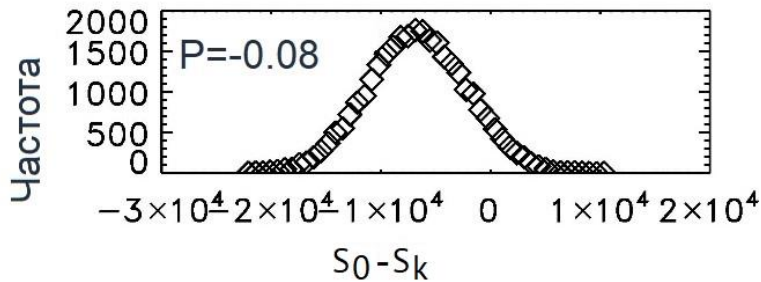


Рис. 2.8.6. Розрахунок кореляції між вибірками даних ^{90}Sr та величиною окислюваності по свердловині 4-Г.

Кореляційний зв'язок між іонною силою та об'ємними активностями ^{90}Sr представлено на рис. 2.8.7. Гістограма має два максимуми, що вказує на різну сукупність даних.

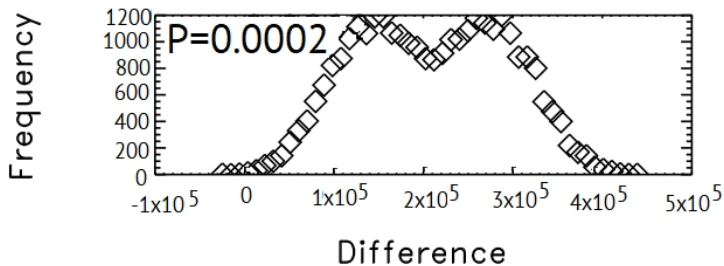


Рис. 2.8.7. Розрахунок кореляції між вибірками даних ^{90}Sr та іонною силою підземних вод по свердловині 21-1А.

Різна сукупність даних створилася через зміну напрямку руху підземних вод [15] у 2014 – 2015 рр. під впливом виведення водоймища-охолоджувача з експлуатації. Зміна напрямку руху підземних вод привела до того, що по свердловині з 2015 р. спостерігається вплив інших джерел забруднення, що і підтверджується формою кореляційної гістограми.

Зауважимо також ще раз, що подібний аналіз на основі методів комп'ютерної статистики можна застосовувати до будь-яких даних, тобто негауссових, гауссових, з великим розкидом, з викидами і малим числом елементів у вибірках та ін.

Результати розрахунків, представлені в таблицях, свідчать, що існує зв'язок між рН та концентрацією ^{90}Sr у наведених свердловинах. Подібний же зв'язок спостерігається між концентраціями іонів кальцію та радіостронцію, що видно з результатів статистичних аналізів (табл. 2.8.4.).

Таблиця 2.8.4. Результати розрахунку ступеня кореляції (К) між результатами моніторингових досліджень

Свердловина	pH	Na	Ca	K
1-2A	1	-0,99	0,98	-0,79
2-G	1	1	1	1
4-G	1	1	0,93	1
1-1A	1	0	1	0,51
21-1A	1	0,99	1	0,97

СПИСОК ЛІТЕРАТУРИ

1. *Kovalenko, I. O., Panasiuk, M. I., Skorburn et al.* (2021). Correlation between chemical composition and ^{90}Sr concentrations in groundwater of the Chernobyl NPP industrial site. *Journal of Environmental Radioactivity*, 240, 106756. <https://doi.org/10.1016/j.jenvrad.2021.106756>
2. *Ліхоузова Т. А.* Теорія імовірностей та математична статистика. Курс лекцій – Київ: Київ. політехн. ін-т ім. Ігоря Сікорського, 2018. – 300 с. – Режим доступу: https://ela.kpi.ua/bitstream/123456789/22404/1/TIMS_kurs_lek.pdf
3. *Щиголов Б. М.* Математическая обработка. – Москва: Гос. изд-во физ. мат. лит., 1962. – 344 с.
4. *Dewiere L.* ^{90}Sr migration to the geosphere from a waste burial in the Chernobyl exclusion zone / L. Dewiere, D. Bugai, C. Grenier, V. Kashparov, N. Ahamdach // *Journal of Environmental Radioactivity*. – 2004. – Vol. 74. – P. 139–150.
5. *Бахвалов А. В.* Биогеохимическое поведение ^{90}Sr в наземных и водных экосистемах [Електронний ресурс] / А. В. Бахвалов, Г. В. Лаврентьева, Б. И. Сынзыныс // *БИОСФЕРА*. – 2012. – Режим доступу до ресурсу: <https://www.elibrary.ru/item.asp?id=17852063>.

6. *Пшинко Г. Н.* Сорбция-десорбция ^{137}Cs и ^{90}Sr в процессах миграции и дезактивации / Г. Н. Пшинко, С. А. Кобец, В. М. Федорова // *Ядерная энергетика та довілля.* – 2014. – № 2 (4). – С. 32–39.
7. *Путилина В. С., Галицкая И. В., Юганова Т. И.* Сорбционные процессы при загрязнении подземных вод тяжелыми металлами и радиоактивными элементами. Стронций: Аналит. обзор. – Новосибирск: ГПНТБ СО РАН, 2013. – 95 с. – Библиогр.: С. 5–6, 87–92 (105 назв.). – (Сер. Экология; Вып. 101).
8. *Бугай Д. О.* Гідрохімічні реабілітаційні процеси у ґрунтових водах ПТЛРВ Рудий ліс / Д. О. Бугай, Є. Ю. Ткаченко, Н. Ван Меер, С. М. Курило, Ю. І. Кубко // *Чорнобильський науковий вісник.* – 2010. – № 2 (36). – С. 3–13.
9. *Панасюк М. І.* Закономірності розподілу урану в підземних водах проммайданчика ЧАЕС / М. І. Панасюк, І. А. Литвин // *Ядерна фізика та енергетика* – 2017. Т. 18, № 1. – С. 56 – 62.
10. *Литвин І. А.* Забруднення ^{90}Sr підземних вод території об'єкта «Укриття» Чорнобильської атомної електростанції / Литвин І. А., Панасюк М. І., Левін Г. В., Онищенко І. П. // *Проблеми безпеки атомних електростанцій і Чорнобиля, 2016* – Вип. 26. – С. 85 – 89.
11. *Панасюк М. І.* Можливість використання термодинамічного моделювання формування високих концентрацій ^{90}Sr в сильнолужних підземних водах / Панасюк М. І., Матросов Д. Т., Стоянов О. І., Левін Г. В. // *Проблеми безпеки атомних електростанцій і Чорнобиля.* – 2018. – Вип. 31. – С. 120 – 126.
12. *Hesterberg T.* Bootstrap methods and permutation tests / T. Hesterberg, D. S. Moore, S. Monaghan et al. – N.Y.: W. H. Ereeman. – 2005. – P. 14.2 – 14.69.
13. *D.S. Moore, G.P. McCabe, B.A. Craig.* Introduction to the Practice of Statistics (W.H. Freeman, N.Y., 2014) [ISBN: 1464158932].
14. *Breiman. L.* Statistical modeling: the two cultures / L. Breiman // *Statistical Science, 2001, Vol. 16, No. 3, 199–231.*
15. *Kovalenko, I. O., Panasiuk, M. I., Sosonna, N. V., Khan, V. E.-I., Buzynnyi, M. G., Koliabina, I. L., & Onyshchenko, I. P.* (2024). Factors influencing the increased ^{90}Sr radioisotope migration in highly alkaline groundwater at Chornobyl NPP site. *Journal of Environmental Radioactivity, 275, 107431.* <https://doi.org/10.1016/j.jenvrad.2024.107431>

Додаток А

У цьому розділі для довідок наведено основні терміни і поняття математичної статистики, знання яких необхідно для розуміння проблем, які обговорюються в основному тексті. А також додано у стислому вигляді роздуми про те, якою статистикою і чому треба користуватись.

Основні поняття і терміни математичної статистики

Математична статистика, яка намагається оцінити властивості наборів одержаних експериментально випадкових чисел, базується на теорії ймовірностей [1 - 5], яка має справу з ідеальними випадковими числами, які, у свою чергу, мають задані властивості-характеристики. Перехід від математичної теорії ймовірностей до математичної статистики настає тоді, коли набору експериментальних (як вважається, випадкових) чисел приписуються ті чи інші властивості, які вже описано в теорії ймовірностей. А далі використовуються формули теорії ймовірностей, розроблені для цього випадку (див., наприклад, Вентцель [3], § 11.4, 11.6).

Поняття випадкової величини вводиться на рівні інтуїції. Випадкова величина – це числова характеристика випадкової події, тобто події, яка в заданих умовах може трапитися, а може й не трапитися. Наприклад, результат вимірювання може прийняти деяке конкретне значення, а може виявитися іншим. Завжди існує область визначення випадкової величини: розмір вимірюваної деталі не може бути меншим і, відповідно, більшим деякого конкретного значення.

Однак випадковість появи деякої величини серед ряду інших величин ще не означає, що для набору таких (випадкових) величин не можна знайти характеристики, за якими їх можна відрізнити від інших випадкових величин. У більшості випадків виявляється, що для досить великих наборів випадкових чисел розраховані для них, наприклад, середнє значення чи відхилення окремих значень від нього є достатньо стабільними і, отже, можуть бути характеристиками таких наборів чисел.

Коли соціологи роблять свої прогнози, вони виходять з результатів вибіркового опитувань. При цьому вважається, що коли число опитаних є «достатньо великим», то можна очікувати, що й інше, паралельне опитування (при умові, звичайно, збереження умов опитування) дасть аналогічні результати, а отже можна очікувати, що ці результати можна розширити і на всю генеральну вибірку (на повну групу людей). Те саме стосується і, наприклад, контрольних вимірювань якості продукції: при виконанні тих самих умов відбору зразків правильно організоване вибіркоче вимірювання дає змогу оцінювати якість всієї виробленої продукції.

Так приходять до поняття генеральної сукупності об'єктів аналізу (випадкових величин) і ймовірнісного способу оцінки характеристик вибірок з неї.

Випадкова величина. Отже, величину вважаємо випадковою, якщо вона

приймає значення з деякої визначеної області значень і немає способу точно спрогнозувати її наступне значення. Проте і тут існують певні закономірності в тому, яке значення прийме ця випадкова величина при наступному вимірюванні, тобто можна з експерименту визначити, як часто випадкова величина приймає те чи інше значення із своєї області визначення. Наприклад, якщо взяти випадково велику групу людей, то виявиться, що значна їхня частина буде мати приблизно однаковий зріст, а число людей, зріст яких буде помітно більший чи менший, ніж в основній групі, буде зменшуватися зі збільшенням відмінності від зросту основної групи. Отже, якщо з такої групи об'єктів (людей у даному випадку) випадково вибирати кілька разів по одній особі, то ясно, що особи з основної частини будуть вибиратися частіше. Так ми приходимо до висновку, що можна знайти спосіб якимось чином характеризувати випадкову величину чи їхню сукупність.

Генеральна сукупність. Нехай ми маємо випадкову величину. Наприклад, нехай ми маємо величезний ящик (урну), заповнену однаковими пластмасовими кульками. Кульки перенумеровані від одиниці до, наприклад, 90. Число кульок N достатньо (поняття «достатньо» буде прояснено нижче, тому що ця величина залежить від конкретної ситуації) велике. У теорії ймовірностей формули часто отримуються при спрямуванні N до нескінченності. Але номери на кульках продовжують залишатися в діапазоні від 1 до 90 (області визначеності випадкової величини). Тоді в цьому випадку для задач математичної статистики генеральною сукупністю буде загальна сукупність перенумерованих кульок, що знаходяться в урні, або ж відповідна їм загальна сукупність чисел від 1 до 90, де ці числа (від 1 до 90) повторюються багато разів.

Вибірка з генеральної сукупності. Кульки в урні ретельно перемішані і є спосіб випадково (тобто номер на вибраній у черговий раз кульці не залежить від того, який номер був на попередній вибраній кульці) діставати їх з ящика і записувати номери. У дійсності на практиці стоїть задача визначити, чим можна характеризувати генеральну сукупність, наприклад, яке в ній співвідношення числа кульок з тим чи іншим номером і т. ін. (про поняття характеристик генеральної сукупності див. вище), на основі обмеженого числа вимірювань. Прямий шлях – вибрати всі кульки і порахувати, скільки з них мають той чи інший номер. Але при нескінченно великій кількості кульок цього зробити не можна у принципі – виконання такої процедури буде вимагати нескінченного часу. Генеральна сукупність може бути обмеженою, тобто мати скінченне число елементів. Наприклад, скінченною є кількість деталей, що виробляються на токарному станку деякого підприємства. У цьому випадку можна було б, у принципі, контролювати всі деталі, але організувати тотальний контроль кожної деталі може бути економічно не вигідним і бажано визначати якість всієї партії деталей на основі поодиноких вимірювань. В обох випадках мова йде про те, щоб, провівши деяку скінченну кількість вимірювань, зробити висновок про те, що являє собою генеральна сукупність. Практика показує, що такі висновки дійсно

можна робити, проаналізувавши деяку частину генеральної сукупності, яка називається вибіркою (з генеральної сукупності).

Ймовірність – це числова оцінка того, як часто з групи об’єктів при випадковому відборі зустрінеться об’єкт з тією чи іншою характеристикою.

Можливість аналізу випадкових величин базується на тому експериментальному факті, що при достатньо великому числі елементів вибірки її статистичні властивості наближаються до статистичних властивостей генеральної сукупності. Це можна зрозуміти, якщо згадати математичний аналіз, у якому приймається, що половина від нескінченної сукупності має також нескінченне число елементів. Тому обидві половини (тобто генеральну сукупність розділено на дві вибірки) повинні будуть мати однакові статистичні властивості. Нехай ми маємо описану вище генеральну сукупність (кульки в урни) з достатньо великим числом елементів N . Відберемо з неї (випадковим способом) m елементів (кульок) і порахуємо, скільки серед них буде елементів з деяким номером, наприклад 34. Нехай число таких елементів буде n . Утворимо величину $P = n/m$. Якщо число відібраних з генеральної сукупності елементів m буде наближатися до N , то число елементів n з номером $k = 34$ буде наближатися до свого істинного значення $n_{k\max}$, тобто до його конкретної кількості в генеральній сукупності з N елементів, і при $m = N$ буде точно дорівнювати цьому істинному значенню. Величина $P = n_k/m$ називається ймовірністю знайти елемент з номером k (витягнути кульку з номером k) серед інших елементів генеральної сукупності. Ця величина із збільшенням m прямує до свого точного значення $n_{k\max}/N$.

Можна дати цьому поняттю більш строге математичне визначення. Однак це вже область теорії ймовірностей і бажаючі можуть познайомитися з нею, наприклад, по [1 – 4].

Функція розподілу. Функція розподілу відповідає на запитання: як часто або з якою ймовірністю при витягуванні кульок з урни будуть зустрічатися кульки з номерами елементів, меншими деякої величини. Якщо взяти вибірку з m елементів, перенумерувати її елементи і позначити їх як $X = (x_1, x_2, x_3, \dots, x_i, \dots, x_m)$ (де великою літерою X позначається вибірка в цілому), то можна ввести функцію $F_X(x) = P\{X < x_i\}$, де $P\{X < x_i\}$ означає ймовірність.

Функція розподілу – математичне поняття і її можна ввести і для нескінченної вибірки, тобто побудувати відповідну теоретичну залежність. Приклади функції розподілу для рядів, розподілених за різними законами чисел, наведено на рис. А1.

Теорія ймовірностей має справу тільки з такими наборами випадкових величин, які можна охарактеризувати конкретною функцією розподілу [4, 128]. І саме це дозволяє характеризувати числовими характеристиками ряди (вибірки) випадкових чисел.

Функція щільності ймовірності. З функцією розподілу тісно пов’язана інша, у прямому значенні похідна від неї функція – функція щільності ймовірності, або на спрощеному жаргоні – щільність розподілу. Вона відповідає на запитання, якими є співвідношення між кількістю різних

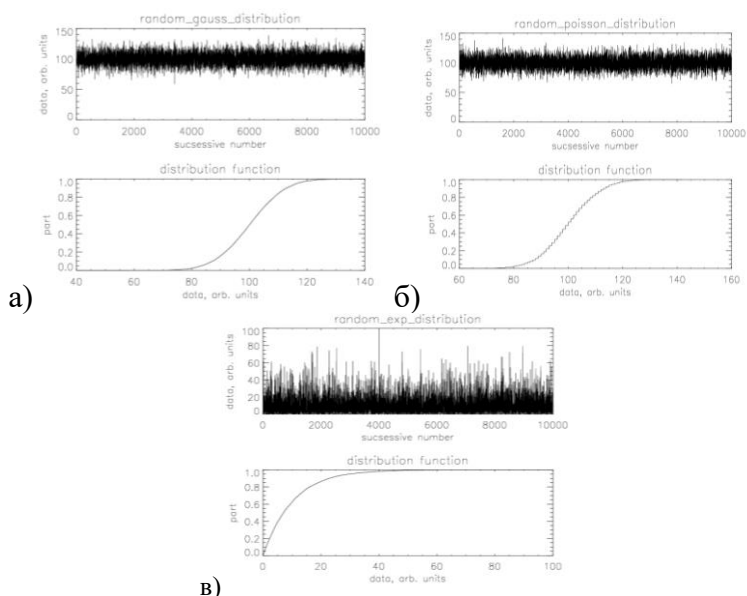


Рис. А1. Комп'ютерні модельні ряди чисел, розподілених за нормальним (гауссовим) законом (а), за законом Пуассона (б), за експоненціальним законом (в), та їхні функції розподілу.

елементів у генеральній вибірці, або, якщо взяти описану вище генеральну вибірку з кульками в урні, яким є співвідношення між величинами n_k . Для цього область визначення розбивається на деяке число піддіапазонів або карманів, або інтервалів (bin), або класів: у літературі зустрічаються всі ці терміни. У нашому випадку область визначення випадкової величини, тобто діапазон її можливих значень, є від 1 до 90, і його можна розбити на 90 піддіапазонів або на 10, або на будь-яке інше число. Вибір величини піддіапазону – окреме питання, на яке немає однозначної відповіді. У літературі зустрічаються певні рекомендації, але вони не мають теоретичного обґрунтування [6]. На практиці вибір ширини кармана визначається умовами задачі, особливостями набору даних і т. ін. Для початківців скажемо, що для початку можна вибрати число карманів порядку 10 та 20, порівняти графіки і зробити висновок про оптимальний розмір кармана для даної задачі.

Оскільки у розглянутому випадку величини, що аналізуються, є цілими числами, то величину піддіапазонів бажано вибрати так, щоб вони містили лише цілі числа. Отже, для нашого випадку цілих чисел виберемо величину піддіапазону рівною 10, тобто розіб'ємо весь діапазон значень номерів кульок на 9 частин. Якщо тепер зробити ящик з 9 секціями, підрахувати, скільки в генеральній сукупності (або вибірці з неї розміром у m елементів) елементів-кульок з відповідними номерами знаходиться в діапазоні 1 – 10, 11 – 20, ... 81 – 90 і покласти їх у відповідні секції-кармани, то одержимо щось подібне до показаного на рис. А2.

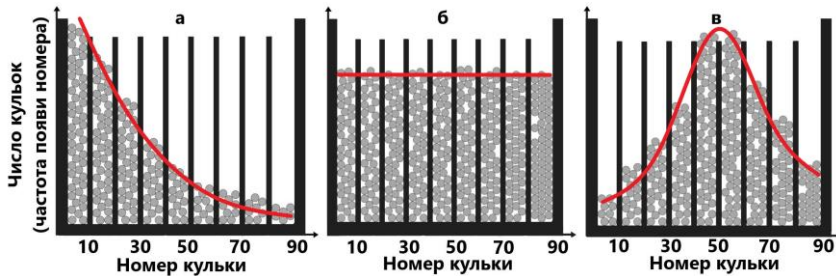


Рис. А2. Приклади розподілу номерів кульок по карманах гістограми залежно від їхніх співвідношень у генеральній сукупності: а – число кульок із збільшенням номера – зменшується за експоненціальним законом (експоненціальний розподіл); б – число кульок однакове для всіх номерів (рівномірний розподіл); в – число кульок з великими або малими номерами зменшується в міру віддалення від деякого центру (аналог гауссового та інших розподілів дзвоноподібної форми).

Показані там пунктиром апроксимуючі криві називаються гістограмою, яка і є характеристикою щільності розподілу. Гістограми, тобто функції щільності розподілу, є першими похідними від функції розподілу. У принципі, можна придумати будь-яку функцію розподілу, похідна від якої дасть відповідну гістограму [7] і навпаки. На рис. А2 наведено, як приклад, широко відомі функції щільності розподілу, що зустрічаються в даній книжці (їхні функції розподілу наведено на рис. А1). У довіднику [8] наведено властивості 24 розподілів. Для повноти картини вкажемо, що всі необхідні таблиці для гауссового (нормального) розподілу можна знайти в роботі [1].

Побудова гістограми щільності розподілу є найпростішим способом зробити первинну оцінку даних. Тому на рис. А3 і А4 для демонстрації показано залежність функції щільності розподілу від величини кармана для двох зовні дуже подібних розподілів: Пуассона і Гаусса. З цих рисунків видно, що при неправильному виборі розміру кармана можна отримати хибне уявлення про дані.

Числові характеристики вибірок. Якщо для даної вибірки функція розподілу $F_X(x)$ є відомою, то на її основі можна розрахувати деякі числові характеристики вибірки, за якими, власне, і відрізняють одну вибірку від іншої. В англійській літературі вони називаються статистиками вибірки, а у вітчизняній – характеристиками вибірки, параметрами вибірки, а також статистиками:

r-й момент

$$\mu'_r = \int_{-\infty}^{+\infty} x^r f(x) dx;$$

r-й центральний момент $\mu_r = \int_{-\infty}^{+\infty} x^r f(x - x_0) dx;$

математичне очікування (перший момент)

$$\mu_1 = \int_{-\infty}^{+\infty} x f(x) dx ;$$

дисперсія (другий центральний момент)

$$\mu_2 = \int_{-\infty}^{+\infty} x^2 f(x - x_0)^2 dx ;$$

стандартне відхилення $\sigma = \sqrt{\mu_2}$: додатний квадратний корінь із дисперсії;

коефіцієнт асиметрії $\eta_3 = \mu_3/\sigma^3$;

коефіцієнт ексцесу $\eta_4 = \mu_4/\sigma^4$;

коефіцієнт варіації σ/μ_1 .

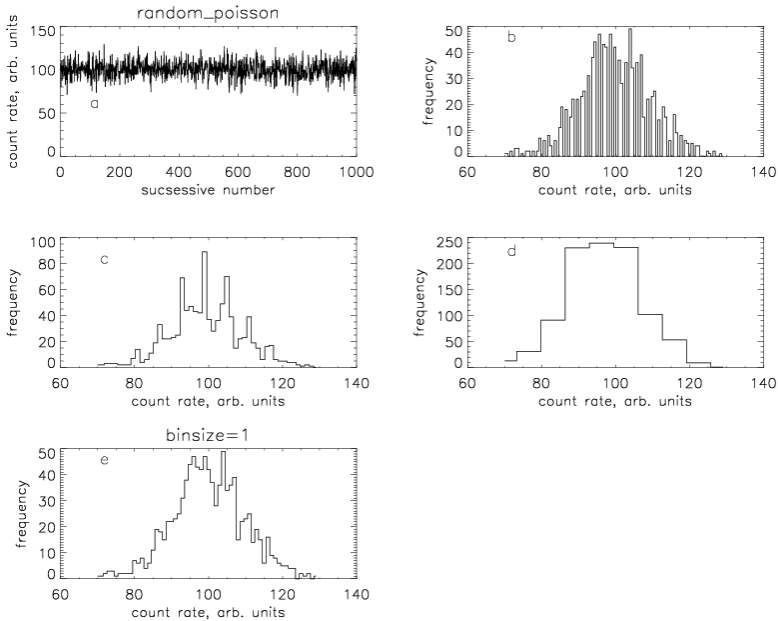


Рис. А3. Гістограми модельного ряду випадкових чисел, розподілених за законом Пуассона: а – послідовний ряд пуассонівських чисел. Число елементів ряду – максимальне значення на горизонтальній осі; б – гістограма, коли весь ряд розбито на 100 карманів; с – гістограма, коли весь ряд розбито на 50 карманів; d – гістограма, коли весь ряд розбито на 10 карманів; e - гістограма, коли розмір кармана дорівнює одиниці.

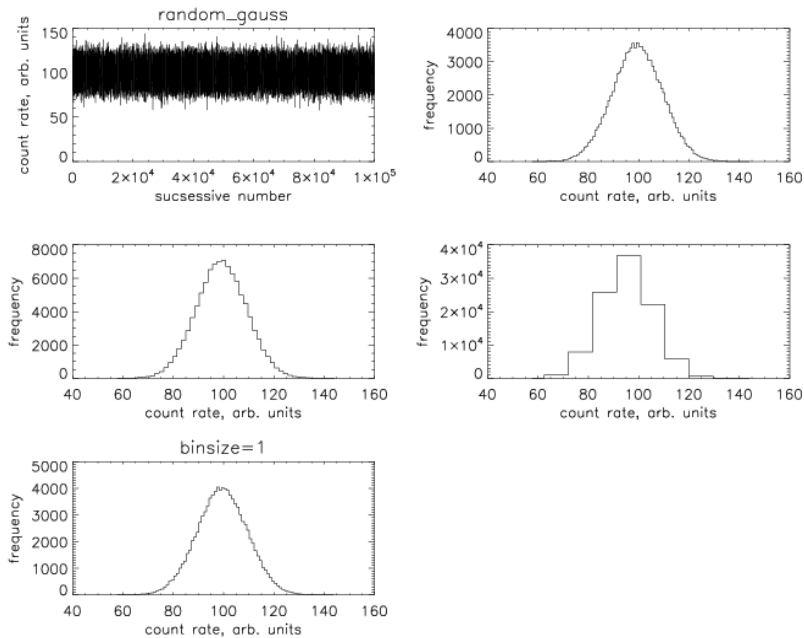


Рис. А4. Гістограми модельного ряду випадкових чисел, розподілених за законом Гауса: а – послідовний ряд гауссівських чисел. Число елементів ряду – максимальне значення на горизонтальній осі; б – гістограма, коли весь ряд розбито на 100 карманів; с – гістограма, коли весь ряд розбито на 50 карманів; d – гістограма, коли весь ряд розбито на 10 карманів; е – гістограма, коли розмір кармана дорівнює одиниці.

Основні стандартизовані визначення (терміни) з математичної статистики в англійському і українському варіантах з ISO 3534-1:1993 та ISO 3534-1:2006

Визначення (Definitions)

С.2.1

Ймовірність (probability)

Дійсне число від 0 до 1, прив'язане до деякої події, яке показує відносну частоту появи цієї події в довгому ряду вимірювань, або ступінь нашої віри в те, як часто буде траплятися така подія.

Ймовірність події А позначається $\text{Pr}(A)$ або $P(A)$.

С.2.2

Випадкова змінна (random variable)

Змінна величина, яка може приймати будь-яке значення із заданого ряду величин і з якою асоціюється функція розподілу.

Випадкова змінна, яка може приймати лише ізольовані значення, називається «дискретною».

Змінна, яка може приймати будь-яке значення із скінченного або нескінченного інтервалу, називається «неперервною».

C.2.3

Розподіл ймовірностей (випадкової змінної)

(Probability distribution (of a random variable))

Функція, яка дає ймовірність того, що випадкова змінна приймає будь-яке задане значення або належить до заданого ряду значень.

Примітка: ймовірність повного ряду значень випадкової змінної дорівнює 1.

C.2.4

Функція розподілу (distribution function)

Функція, яка дає, для деякого значення x , ймовірність того, що випадкова змінна X буде меншою або рівною x .

$$F(x) = \Pr(X \leq x)$$

C.2.5

Функція щільності ймовірності (для неперервної випадкової змінної)

Probability density function (for a continuous random variable)

Похідна (якщо вона існує) від функції розподілу

$$f(x) = dF(x)/dx$$

Примітка: $f(x)dx$ є «елемент ймовірності».

$$f(x)dx = \Pr(x < X < x + dx)$$

C.2.6

Масова функція ймовірності (probability mass function)

Функція, яка дає, для кожного значення дискретної випадкової змінної X , ймовірність p_i того, що випадкова змінна дорівнює x_i :

$$p_i = \Pr(X = x_i)$$

Центрована випадкова змінна (Centered random variable)

Випадкова змінна, (математичне) очікування якої дорівнює нулю.

Примітка: якщо випадкова змінна X має очікування, рівне x_0 , то відповідна центрована змінна є $(X - x_0)$

C.2.11

Дисперсія (випадкової змінної, або розподілу ймовірності)

Variance (of a random variable or of a probability distribution)

Математичне очікування квадрата відхилення центрованої випадкової змінної від її математичного очікування:

$$V(X) \equiv \sigma^2 = E \{[X - E(X)]^2\}$$

C.2.12

Стандартне відхилення (випадкової змінної, або функції розподілу)

Standard deviation (of a random variable or of a probability distribution):

Додатний квадратний корінь з дисперсії:

$$\sigma = \sqrt{V(X)}$$

C.2.13

Центральний момент порядку r

Central moment of order r

(див. вище)

C.2.14

Нормальний (гауссовий, Лапласа - Гаусса) розподіл

(Normal distribution, Laplace - Gauss distribution)

Розподіл ймовірності неперервної випадкової змінної X , функція щільності ймовірності якої є

$$f(x) = \frac{1}{\sigma(2\pi)^{1/2}} \exp\left[\frac{-(x - x_0)^2}{2\sigma^2}\right]$$

Для $-\infty < x < +\infty$.

Примітка: x_0 є очікування, а σ є стандартне відхилення нормального розподілу.

C.2.15

Характеристика (characteristic)

Властивість, яка допомагає ідентифікувати або розрізнати окремі елементи даної популяції.

Примітка: характеристика може бути або кількісною (для змінних), або якісною (для ознак).

C.2.16

Популяція (population)

Повний набір елементів, які аналізуються.

Примітка: у випадку випадкової змінної вважається, що визначення функції розподілу описує саме цю змінну.

C.2.17

Частота (frequency)

Число появ подій або величин спостережень даного типу, які потрапляють у даний клас (карман гістограми).

Примітка: звернемо увагу, що термін «частота» є нормативним у статистиці, поряд з «частотою» як характеристикою коливних процесів (*авт.*).

С.2.18

Розподіл частоти (frequency distribution)

Емпіричне співвідношення між величинами та їхніми частотами або їхніми відносними частотами.

Примітка: цей розподіл може бути представленим графічно у вигляді гістограми.

С.2.19

Арифметичне середнє (arithmetic mean) average

Сума значень, поділена на число цих значень.

Деякі нестандартизовані поняття

Квантиль – одна з числових характеристик функції розподілу випадкових величин. Квантилем x_p для даної упорядкованої по зростанню випадкової вибірки $X = (x_1, x_2, x_3, \dots, x_p, \dots, x_m)$ називається таке число x_p , серед елементів вибірки, для якого відношення числа елементів $x_i < x_p$ становить заданий наперед процент від загального числа елементів вибірки. Наприклад, медіана – це квантиль, який відповідає ймовірності 0,5, тобто половина значень вибірки буде меншою від медіани.

Поняття квантиля виникає, наприклад, при оцінці довірчих інтервалів. Якщо ми хочемо, щоб отриманий результат x_p було отримано з достовірністю $\alpha\%$, серед отриманих результатів вимірювань величини вибірки $X = (x_1, x_2, x_3, \dots, x_p, \dots, x_m)$ число елементів $x_i > x_p$ повинно становити α_m .

Квантилі нормального розподілу

- 0,25-квантиль називається **першим (або нижнім) квантилем**;
- 0,5-квантиль називається **медіаною** або **другим квантилем**;
- 0,75-квантиль називається **третім (або верхнім) квантилем**.

Статистична стійкість

За визначенням І. І. Горбаня [9] статистично нестійкими є ряди, у яких дисперсія не зменшується з ростом числа елементів вибірки (з часом). У цьому випадку їхні статистичні оцінки стають **неспроможними**, тобто залежними від часу

Стаціонарність. Стаціонарними у **вужькому** значенні називаються ряди, у яких функція розподілу не змінюється з часом. Стаціонарними у **широкому** значенні називаються ряди, у яких з часом не змінюються середнє і дисперсія.

Оцінка невизначеності і довірчий інтервал. Статистичні оцінки за самим своїм визначенням є приблизними оцінками реальної ситуації. Відповідно виникає питання про надійність таких оцінок, яке можна

переформулювати в питання про величину довірчого інтервалу: інтервалу, в якому із заданою ймовірністю знаходиться оцінюваний параметр вибірки. Прикладом параметра, який використовується для оцінки довірчого інтервалу є середньоквадратичне відхилення σ – число, яке характеризує величину розкиду результатів і дає можливість оцінити, яка частка результатів знаходиться у межах $\pm \sigma$. Зауважимо, що існує типова помилка, яка криється в тому, що отримане для вибірки X середньоквадратичне відхилення σ_x , тобто розкид значень елементів вибірки, приписується середньому значенню, у той час як добре відомо [10, § 72], що це σ_x характеризує похибку **окремого вимірювання**, тобто межі, в яких буде знаходитися результат наступного вимірювання. Похибка середнього визначається іншою формулою і буде значно меншою, тому що вже при числі вимірювань порядку 20 додавання ще одного числа (тобто збільшення числа елементів вибірки до 21) мало змінює її середнє. У припущенні нормального розподілу похибка середнього буде приблизно σ_x/\sqrt{m} , де m – число елементів вибірки. Спосіб оцінки невизначеності величини середнього значення для вибірки з довільним (невідомим) розподілом на основі методів комп'ютерної статистики розглянуто в розділах «бутстрап» та «джекнайф».

Часові ряди випадкових чисел. Крім просто ряду великого числа вимірювань (багаторазові повтори вимірювання розміру деталі, наприклад), можуть бути також послідовні вимірювання в часі. Наприклад, типовим є вимірювання числа розпадів радіоактивного елемента за певний проміжок часу або числа автомобілів на дорозі. Такі ряди є функцією часу і їхні характеристики можуть змінюватися з часом (вночі інтенсивність руху стає меншою). Для таких рядів усереднені по вибірці статистичні характеристики вже можуть бути не представницькими, а інтерес будуть являти саме зміни цих характеристик з часом.

Поняття про набори випадкових чисел (вибірки), які не мають функції розподілу. Набори «випадкових» чисел, які не мають функції розподілу, тобто побудований за вказаними вище правилами графік функції їхнього розподілу не описується певною аналітичною функцією, випадковими не вважаються [9]. Але такі дані існують і їх потрібно аналізувати. Мається на увазі, що згадані вище способи оцінок, оснований на гауссовому розподілі, у цій ситуації застосовувати не можна, і відповідних потрібних таблиць немає. А що робити? (Див. першу частину книги про методи комп'ютерної статистики).

Деякі зауваження про використання гауссового розподілу

Використання нормального розподілу є настільки широко розповсюдженим, що необхідно зробити з цього приводу кілька застережень. Не дивлячись на те, що в літературі розглядається велика (сотні) кількість розподілів, у реальній практиці використовується досить обмежене їхнє

число. І серед них нормальний розподіл займає пануючі позиції. Цьому є певні підстави.

Так, практично аксіоматичним є твердження, що похибки результатів вимірювань, тобто відхилення окремого вимірювання від їхнього середнього значення, розподілені за нормальним законом. Воно базується на тому, що багатьма дослідженнями показано, що гістограми похибок вимірювань дійсно мають дзвоноподібний вигляд. При будь-якому аналізі результатів на практиці завжди робиться припущення (висувається гіпотеза), що результати (у нашому випадку гістограму (див. рис. А2)) можна апроксимувати тією чи іншою функцією. Ясно, що для кульок на рис. А2, в спроба апроксимації функцією Гауса є цілком допустимою. А от слово «цілком» у даній ситуації є вже «позалогічним» (див. [11]) твердженням і фактично означає, що експериментатор з тих чи інших міркувань це твердження приймає. Тобто прийняття рішення про те, що результати експерименту розподілені за нормальним законом ґрунтується не на тому, що це надійний математичний результат, а на деяких інших міркуваннях, а математична статистика лише підтримує це рішення.

Результати вимірювань повинні обов'язково супроводжуватися вказуванням похибки (невизначеності), з якою вони отримані. Оскільки існує велика експериментальна практика, що реальні дані в області аналізу похибок дійсно більш-менш задовольняють гіпотезі про їхній нормальний розподіл, ця ідея була внесена в нормативні документи, тобто набула статусу обов'язкової для виконання (ДСТУ, Керівництво з невизначеностей, Закон про метрологію). А вже невизначеність результатів розраховується на основі гіпотези про нормальність розподілу.

Другим наріжним каменем гауссового підходу до аналізу даних є твердження центральних граничних теорем про те, що при виконанні деяких умов про властивості цих чисел (наприклад, у них повинен існувати третій момент), для суми багатьох незалежних випадкових величин сумарний їхній розподіл буде гауссовим незалежно від того, який розподіл має кожна з цих величин. На основі цього робиться висновок, що оскільки на практиці існує багато випадкових чинників, що впливають на результат вимірювання, то ці результати будуть мати нормальний розподіл. Осць як це пропонують у роботі [12, с. 160]: «Приклад 5.2. В цеху виготовляють кульки для підшипників. Кожна кулька повинна мати діаметр D , однак у силу ряду причин, неминучих в умовах масового виробництва, фактичний діаметр кульок відрізняється від D Оскільки відхилення діаметра кульки від номіналу складається із суми відхилень з різних причин у відповідності з центральною граничною теоремою, будемо вважати відхилення діаметра кульки випадковою величиною з середнім 0 Знайдемо ймовірність того, що i -та кулька виявиться бракованою.» Цей приклад яскраво демонструє ті проблеми використання класичної математичної статистики, позбутися яких намагається комп'ютерна статистика. Для того щоб розробити методику виконання вимірювань при контролі діаметра кульок, робиться ряд припущень: що результати вимірювань

є випадковими величинами, які мають розподіл; що існує багато зовнішніх факторів впливу, які впливають на діаметр кульки, що результати вимірювань є числами з властивостями, які задовольняють вимогам центральної граничної теореми; що ідея про усереднений вплив багатьох факторів дійсно реалізується на практиці. У результаті робиться висновок, а фактично припущення, що діаметри кульок розподілені за нормальним законом. І все це робиться для того, щоб мати можливість застосовувати до цих даних методи статистичного аналізу, які основані на нормальному розподілі і введені в нормативні документи.

Автор не чув про існування експериментальних досліджень, де б такі фактори впливу були перелічені, де було б досліджено, що вони саме так, випадково, впливають і дають у результаті нормальний розподіл відхилень.

Отже, шукаємо не там, де загубили, а там, де є освітлення: аналізуємо не самі експериментальні дані, а функцію розподілу, яка їх апроксимує, притому з невідомою точністю.

Основні теореми теорії ймовірностей і проблеми їхнього використання

У теорії ймовірностей існує кілька теорем, які можна вважати ключовими для математичної статистики [1 – 4].

1. Теореми великих чисел.

Це ряд теорем, де доводиться, що за виконання певних умов із збільшенням числа вимірювань (тобто із збільшенням числа елементів вибірки) середнє значення буде прямувати і в результаті як завгодно близько наближатися до «істинного» значення середнього генеральної вибірки.

2. Центральна гранична теорема.

Центральна гранична теорема стверджує, що сума деякого числа незалежних випадкових чисел, незалежно від їхнього розподілу, буде мати гауссовий розподіл. Ця теорема моделює широко розповсюджене твердження, що при виконанні вимірювань результат вимірювань залежить від великого числа малих випадкових впливів, тому результати будуть розподілені за законом Гаусса. На цьому будується теорія похибок. На її основі побудовано алгоритм генерування випадкових нормально розподілених чисел: якщо взяти суму деякого числа (досить уже 12 доданків) випадкових рівномірно розподілених чисел, то сумарні числа будуть розподілені за законом Гаусса.

Однак звернемо увагу, що ця теорема є правильною *при виконанні певних вимог* до складових. Наприклад, вони повинні бути незалежними тощо. А її застосування в математичній статистиці також базується на *припущенні*, що в реальності при виконанні вимірювань існує багато незалежних *малих* впливів, сумарний результат дії яких веде до того, що результат має гауссовий розподіл. Читач може спробувати сам відповісти на прохання хоча б перелічити ці гіпотетичні впливи, не кажучи вже про кількісну оцінку кожного з них.

В основному тексті згадувалося, що існує авторитетна критика такого широкого застосування описаного підходу [Зограф, Орлов, Тутубалін, Дубровський]. Ми ж наведемо тут кілька цитат інших авторів на цю тему.

Застереження до закону великих чисел

- “...не можна вимагати точності від нормального наближення при оцінці ймовірностей поодиноких подій ...” (Боровков [13], с. 171).

- “Застосування розглянутих теорем закону великих чисел на практиці обмежено умовою незалежності випробувань чи випадкових величин.” (Щиголев [10], с. 170).

- “Питання про можливість застосування закону великих чисел до явищ природи є достатньо складним. Твердження, подібні зробленим нами у деяких прикладах, повинні даватися з належними застереженнями і перевірятися систематичними спостереженнями.” (Щиголев [10], с. 170).

- “...дослідження лінійної кореляції є допустимим, якщо величини, які розглядаються, мають нормальний розподіл.” (Щиголев [10], с. 328).

Отже, використання цих теорем у реальній практиці може бути обмеженим через невідповідність реальних даних вимогам, закладеним у ці теореми. Як результат, нагадаємо про існування книги Ельясберга [14] «Измерительная информация: сколько ее нужно? Как ее обрабатывать?». У цій книзі розглядається проблема аналізу результатів вимірювання характеристик природних процесів. Це змінні величини, але вони не є випадковим «шумом», мають різкі «викиди» тощо. Тобто не підпадають під дію перелічених теорем.

Висновок: рафіновані вимірювання, для яких виконуються умови дії вказаних теорем, якщо і є, то зустрічаються дуже рідко, виконуються лише в ретельно організованих вимірюваннях з контролем зовнішніх умов тощо. Результати лише *апроксимуються* гауссовим розподілом. Якщо дослідника задовольняє ступінь апроксимації, він може звертатись до теорії ймовірностей, як до теоретичної бази свого аналізу. Якщо ж ні, то існують альтернативні підходи. Один з них – описана тут комп’ютерна статистика, яка обходиться без згаданих припущень. У книзі С. Ф. Тимашева [15] також відкидається ідея про те, що реальні дані відповідають вимогам теорем теорії ймовірностей.

Ідеї ще одного, не менш радикального підходу, коротко описано нижче.

Теорія гіпервипадкових явищ

Класична теорія ймовірностей розглядає набори незалежних однаково розподілених чисел (iid – independent identically distributed), тобто чисел, які належать до однієї і тієї ж генеральної сукупності. Якщо у такому наборі зустрічаються числа, які явно випадають із загального ряду (наприклад, величина яких у десятки разів більша від середнього), вони розглядаються як «викиди» (outliers) і викидаються з розгляду. Існують критерії виявлення таких значень у вибірці (які, звичайно, базуються на властивостях гауссового розподілу) і які, строго кажучи, виходять з того, що такі значення належать до іншої генеральної сукупності. Однак можливі ситуації, коли саме такі

«викиди» являють інтерес, наприклад при пошуку поодиноких явищ на фоні загальної картини.

Теорія гіпервипадкових явищ розглядає ситуацію, коли дані являють собою суміш кількох (багатьох) генеральних сукупностей, виходячи з того (таке твердження можна знайти практично у всіх монографіях з теорії ймовірностей), що частіше за все немає обґрунтованих аргументів приписувати даним певну функцію розподілу, і така процедура завжди є наближенням, яке більш-менш підтверджується на практиці [16 – 18].

Методи теорії ймовірностей базуються на центральних граничних теоремах, які стверджують, що при збільшенні числа елементів вибірки дисперсія результату прямує до нуля. На відміну від цього теорія гіпервипадкових явищ, опираючись на тому, що існує багато експериментальних прикладів, коли дисперсія при збільшенні числа вимірювань не перестає зменшуватись [14], аналізує саме поведінку дисперсії середнього значення при збільшенні числа елементів вибірки. І якісною ознакою гіпервипадкових властивостей вибірки є зростання або значно уповільнене зменшення дисперсії середнього при збільшенні розмірів вибірки, що трактується як втрата вибіркою статистичної стійкості (тобто в даних з'являється залежність від часу). А при втраті даними статистичної стійкості уже їхні статистичні оцінки стають неспроможними, тобто залежними від часу. Для аналізу спроможності даних, виходячи з ідей гіпервипадковості, запропоновано такі формули [19]:

$X = x_1, x_2, \dots, x_n$ – часовий ряд результатів вимірювання – вихідні дані;
накопичуване середнє вихідного ряду даних

$$Y_N = \frac{1}{n} \sum_{i=1}^n x_i \quad (n=1, N); \quad (1)$$

накопичуване середнє середнього вихідного ряду даних

$$\bar{m}_{Y_N} = \frac{1}{N} \sum_{n=1}^N Y_n; \quad (2)$$

накопичувана дисперсія вихідного ряду даних

$$D_x = \frac{1}{n-1} \sum_{i=1}^n (x_i - y_N)^2; \quad (3)$$

накопичувана середня дисперсія вихідного ряду даних

$$D_y = \frac{1}{N^2} \sum_{i=1}^N D_x; \quad (4)$$

незміщена дисперсія флуктуації накопичуваного середнього

$$\bar{D}_{Y_N} = \frac{1}{N-1} \sum_{n=1}^N (Y_n - \bar{m}_{Y_N})^2. \quad (5)$$

У теорії вводяться також величини, які є кількісною мірою

гіпервипадкових властивостей, а саме статистичної нестійкості ряду даних:

коефіцієнт γ_N , який характеризує абсолютний рівень статистичної нестійкості

$$\gamma_N = \frac{M[\overline{D_{Y_N}}]}{ND_{Y_N}}, \quad (6)$$

де $M[*]$ – оператор математичного очікування;

коефіцієнт μ_N , який пов'язаний з параметром γ_N і характеризує співвідношення між складовими вибірки, які можна статистично прогнозувати (які описуються певною функцією розподілу), і ефектами непередбачуваних впливів, які не можна статистично прогнозувати (які не описуються певною функцією розподілу):

$$\mu_N = \sqrt{\frac{\gamma_N}{1+\gamma_N}}. \quad (7)$$

Щоб мати можливість порівнювати різні вибірки між собою, необхідно ввести одиницю статистичної нестійкості.

Для коефіцієнта γ_N в якості одиниці вимірювання статистичної нестійкості пропонується величина γ_{0N} , яка відповідає нескорельованому ряду відліків із постійною дисперсією $D_{x_n} = D_x$ і нульовим математичним очікуванням при фіксованому значенні N . Коефіцієнт γ_{0N} розраховується аналітично:

$$\gamma_{0N} = \frac{N+1}{(N-1)N} C_N - \frac{2}{N-1}, \quad (8)$$

де $C_N = \sum_{n=1}^N \frac{1}{n}$.

Використовуючи одиницю вимірювань γ_{0N} , вводиться відносна величина

$$h_N = \frac{\gamma_N}{\gamma_{0N}}, \quad (9)$$

це коефіцієнт, який характеризує абсолютний рівень статистичної нестійкості в одиницях γ_{0N} .

Відносний рівень нестійкості μ_N в одиницях статистичної нестійкості можна охарактеризувати величиною

$$L_N = \frac{\gamma_N - \gamma_{0N}}{\gamma_N} = \frac{h_N - 1}{h_N}, \quad (10)$$

він являє собою відношення коефіцієнта статистичної нестійкості, розрахованого для непрогнозованої частини розглянутого процесу, до коефіцієнта статистичної нестійкості всього процесу. Границі вимірювань параметра $h_N - [0, \infty)$, а параметра $L_N - (-\infty, 1]$.

Ці коефіцієнти є безрозмірними величинами.

Зауважимо від себе, що параметр h_N видається найбільш чутливим для пошуку ознак гіпервипадковості в рядах даних. Приклад використання теорії гіпервипадковості для аналізу даних див. в основному тексті.

Філософські, тобто нестрогі застереження-зауваження

Автор припускає, що читач, який дійшов до цього абзацу, саме тому і читає цю книжку, що відчуває у своїй практичній діяльності, що стандартні методи математичної статистики його не задовольняють. І причина цього зовсім не в недоліках теорії: теорія ймовірностей і основана на ній математична статистика є замкнутою і математично строгою будовою. У відповідних монографіях завжди присутні доведення відповідних теорем, і це виховує довіру до результатів математичної статистики, які сприймаються як факт, який є самодостатнім. При цьому далеко не в усіх книжках можна знайти роз'яснення, що в дійсності математична статистика використовується дослідником лише для того, щоб перевірити деяку гіпотезу. Наприклад, широко розповсюдженою задачею є встановлення взаємозв'язку між двома явищами (шляхом пошуку кореляцій між відповідними двома рядами вимірювань). Припустимо навіть, що цей дослідник чітко розуміє, що наявність кореляції ще не означає наявність взаємозв'язку між явищами. Проте типовою (з досвіду автора) є така картина. Розрахувавши кореляцію, дослідник шукає у відповідних книжках відповідь інших, шанованих авторів, на запитання про те, яка величина коефіцієнта кореляції свідчить, що факт кореляції є більш-менш достовірним і знаходить у них, скажімо, типове твердження, що якщо коефіцієнт кореляції є більшим, ніж 0,6-0,7, то факт кореляції між двома явищами надійно встановлено. Якщо отриманий при аналізі даних коефіцієнт кореляції виявився у вказаних межах 0,6-0,7, то далі (**увага!**) про величину отриманого коефіцієнта кореляції вже ніхто не згадує, факт існування кореляції=взаємозв'язку вважається встановленим. У той час як правильне твердження (і ця думка постійно проводиться в усіх прийомах розглянутої в книжці комп'ютерної статистики) виглядає так: перевірялася гіпотеза, що дві величини є пов'язаними між собою, а результати виконаного статистичного аналізу з ймовірністю 0,6-0,7, цій гіпотезі **не протирічать**. Тобто математична статистика не доводить шуканий факт взаємозв'язку, а всього лише показує основу на статистичних розрахунках ступінь нашої довіри до цього висновку. Або інакше: у результаті статистичного аналізу дослідник отримує всього лише деяку підказку для прийняття рішення про те, як ставитися до отриманих результатів.

Крім того, слід було б також поставити питання про те, чому не було досягнуто успіху у решті 30 – 40 % випадків. Так, знаменитий дослід Флейшмана і Понса, у якому було виявлено явище холодного синтезу в електрохімічних процесах, погано (тобто далеко не у 100 % повторних дослідів у різних лабораторіях) відтворювався, що дало підстави визнати його помилковим, а самих дослідників брехунами. І лише коли хтось провів систематичну перевірку, то виявилось, що спеціальні електроди, які використовувалися в цих експериментах, є далеко не однаковими за своїми (невідомо якими) параметрами, і, умовно кажучи, із 100 дослідів із 100 різними електродами успіх було отримано лише приблизно у 30 випадках.

Висновок: рафіновані вимірювання, для яких виконуються умови можливості застосування вказаних теорем теорії ймовірностей, зустрічаються дуже рідко, виконуються лише в ретельно організованих вимірюваннях з контролем зовнішніх умов тощо. Результати лише апроксимуються гауссовим розподілом. Якщо дослідника задовольняє ступінь апроксимації, він може звертатись до теорії ймовірностей, як до теоретичної бази свого аналізу. Якщо ж ні, то існують альтернативні підходи. Один з них – комп'ютерна (обчислювальна) статистика, яка обходиться без згаданих припущень. У книзі С. Ф. Тимашева [15] також відкидається ідея про те, що реальні дані відповідають вимогам теорем теорії ймовірностей. Теорія гіпервипадкових явищ І. І. Горбаня [9], основується на тому, що існує багато експериментальних прикладів того, що дисперсія при збільшенні числа вимірювань не перестає зменшуватись, аналізує саме поведінку дисперсії середнього значення при збільшенні числа елементів вибірки і стверджує, що наш світ є не гауссовим, а гіпервипадковим.

СПИСОК ЛІТЕРАТУРИ ДО ДОДАТКА А

1. Смирнов Н. В., Дунин-Барковский И. В. Курс теории вероятностей и математической статистики для технических приложений. – М.: Наука, 1969. – 363 с.
2. Закс Л. Статистическое оценивание. – М.: Статистика, 1976. – 598 с.
3. Вентцель Е. С., Овчаров Л. А. Теория вероятностей и ее инженерные приложения. – М.: Наука. Гл. ред. физ.-мат. лит. – 1988. – 480 с.
4. Гнеденко Б. В. Курс теории вероятностей. – М.: Наука, 1969. – 400 с.
5. Вознесенский В. Л. Первичная обработка экспериментальных данных. – Изд-во «Наука», Ленинград. отд-ние. – Ленинград, 1969 – 83 с.
6. Новицкий П. В., Зограф И. А. Оценка погрешностей результатов измерений. – Ленинград: Энергоатомиздат, 1985. – 248 с.
7. Власов А. А. Статистические функции распределения. – М.: Наука, глав. ред. физ.-мат. лит., 1966. – 357 с.
8. Гастингс Н., Пикок Дж. Справочник по статистическим распределениям. – М.: Статистика, 1980. – 94 с.

9. *Горбань И. И.* Теория гиперслучайных явлений [Электронный ресурс]. – К.: ИПМ-МС НАН Украины, 2007. – 184 с. – Режим доступа: <http://ifsc.ualr.edu/jdberleant/intprob/>.
10. *ЩигOLEв Б. М.* Математическая обработка наблюдений. – М.: Наука, 1969. – 344 с.
11. *Алимов Ю. И., Кравцов Ю. А.* Является ли вероятность “нормальной” физической величиной? УФН 162 (7) 149–182 (1992).
12. *Колемаев В. А., Староверов О. В., Турундаевский В. Б.* Теория вероятностей и математическая статистика. – М.: Высшая школа, 1991. – 400 с.
13. *Боровков А. А.* Теория вероятностей. – М.: Наука, 1976. – 352 с.
14. *Эльясберг П. Е.* Измерительная информация: сколько ее нужно? Как ее обрабатывать? – М.: Наука. Гл. ред. физ.-мат. лит., 1983. – 208 с.
15. *Тимашев С. Ф.* Фликкер-шумовая спектроскопия. Информация в хаотических сигналах. – М.: Физматлит, 2007. – 248 с.
16. *Орлов А. И.* Часто ли распределение результатов наблюдений является нормальным? // Заводская лаборатория. – 1991. – Т. 55, № 7. – С. 64 – 66.
17. *Дубровский С. А., Толстова Т. А.* О ненормальности нормального закона распределения (вообще и в приложении к металлургии в частности) // Всерос. науч.-техн. конф. “Современная металлургия начала нового тысячелетия”. 29 окт. – 2 нояб. 2001 г. Липецк, ЛГТУ. Докл. № 12-2.
18. *Breiman Leo.* Statistical Modeling: The Two Cultures // *Statistical Science*. – 2001. – Vol. 16, No. 3. – С. 199–231.
19. *Горбань И. И.* Статистическая неустойчивость физических процессов // *Радиоэлектроника*. – 2011. – Т. 54. – № 9. – С. 40 - 52.

Наукове видання

Скорбун Анатолій Дмитрович

**КОМП'ЮТЕРНА (ОБЧИСЛЮВАЛЬНА) СТАТИСТИКА
МЕТОДИ МОНТЕ-КАРЛО В ЗАДАЧАХ ОЦІНКИ НЕВИЗНАЧЕНОСТІ
РЕЗУЛЬТАТІВ ВИМІРЮВАНЬ**

Монографія

Формат 70 × 100/16. Ум. Друк. арк. 13,9.
Наклад 50 пр. Зам. №19-3313.

Інститут проблем безпеки АЕС НАН України
Київська обл., 07270, м. Чорнобиль, вул. Кірова, 36а
Свідоцтво суб'єкта видавничої справи ДК №2114 від 25.02.2005

Віддруковано в друкарні ПП «Видавництво «Фенікс»
Свідоцтво суб'єкта видавничої справи ДК №271 від 07.12.2009 р.
03067, м. Київ, вул. Шутова, 136

Для нотаток

Для нотаток
