

Порівняльний аналіз теоретичних методів статистичної обробки експериментальних даних

ДОЛІШНЯ Н.Б., ПІНДУС Н.М.

Івано-Франківський університет нафти і газу

Здійснено порівняльний аналіз теоретичних методів статистичної обробки експериментальних даних, запропоновано скоригувати алгоритм обробки результатів опосередкованого вимірювання витрати, враховуючи перевірку нормальності, виключення промахів та перевірку відтворюваності дослідів.

Осуществлен сравнительный анализ теоретических методов статистической обработки экспериментальных данных, предложено скорректировать алгоритм обработки результатов косвенного измерения расхода, учитывая проверку нормальности, исключения промахов и проверку воспроизводимости опытов.

The comparative analysis of theoretical methods of statistical processing of experimental data has been completed. It is proposed to adjust the processing algorithm of indirect measurement floemetering results including verification of normality, excluding failures, and checking the reproducibility of experiments.

Постановка. Відомо, що на шляху проектування і створення нової зразкової техніки або ж вдосконалення уже існуючих на сьогоднішній день засобів виміральної техніки наукові дослідження є чи не найпершим кроком, в процесі якого дослідник має змогу відкрити нові закони та закономірності, аналізуючи роботу того чи іншого приладу, знайти наукове обґрунтування випадковим процесам, що супроводжують роботу приладу. Разом з тим на основі теоретичних методів дослідження підібрати раціональний та результативний експериментальний метод, зваживши фізичні процеси, які лежать в основі математичної моделі роботи приладу, та вплив зовнішніх факторів. Такий класичний підхід очевидний для більшості випадків, але він може бути не завжди ефективним у випадку, якщо кількість результатів вимірювання є обмеженою або відтворити вимірально експеримент важко чи зовсім неможливо. Це, власне, і є характерним при застосуванні методів та засобів вимірювання витрати та кількості газу.

Аналіз останніх досліджень. Аналізуючи опубліковані результати досліджень, присячених методам обробки експериментальних даних можна стверджувати, що більшість результатів вимірювання представляють собою вибірку з деякої генеральної сукупності, яка, як правило, містить аномальні спостереження. Аномальні спостереження у вибірці появляються за рахунок грубих похибок при реєстрації вимірювань, випадкових імпульсних завад, збоїв устаткування, вимірювання у неправильних одиницях. Якщо дані різко виділяються на фоні ряду спостережень, то вони можуть бути виключені з вибірки на попередньому етапі аналізу результатів проведеного вимірювання, з врахуванням фізичної суті вимірюваної величини. Результати, які знаходяться у так званій сумнівній зоні, розпізнаються складніше і вимагають застосування спеціальних статистичних процедур для виявлення аномальних спостережень. Проте навіть після виявлення таких аномальних результатів не можна вважати

аналіз точності виконання результатів вимірювання завершеним і правильним без ґрунтовних пояснень отриманих результатів.

Автоматичне видалення аномальних спостережень без встановлення причин їх виникнення виправдане і можливе лише тоді, коли створена модель спостереження добре вивчена і випробувана, тобто заздалегідь відома її поведінка. Проте загалом такий метод використовується рідко. В основному використовують статистичні критерії, за допомогою яких встановлюється закон розподілу спостережень і за відповідним алгоритмом обробки результатів експерименту, проводять видалення аномальних результатів, оскільки дуже часто не має достатньої інформації про якість вимірювання або ж ця інформація не є надійною. Підсумовуючи вищесказане, робимо висновок, що у випадку статистичної обробки набору даних отриманих в ході вимірального експерименту обов'язковим є виконання трьох послідовних етапів:

1. *Перевірка нормальності закону розподілу* є обов'язковою оскільки в основі процедур математичної статистики лежить нормальний закон розподілу;
2. *Перевірка на наявність промахів* є необхідною оскільки результати спостереження, які різко відрізняються від інших спотворюють кінцевий результат вимірювання;
3. *Перевірка відтворюваності дослідів* здійснюється для того, щоб підтвердити відсутність часової залежності та відносну стабільність зовнішніх умов при проведенні вимірального експерименту.

Формулювання мети статті. Метою статті є аналіз існуючих методів обробки експериментальних даних з метою оцінки можливості їх застосування при проведенні експериментальних досліджень в галузі витративимірювання.

Виклад основного матеріалу. Сучасна теоретична метрологія дає пояснення основних властивостей критеріїв, розподілів статистик, потужності критеріїв, їх переваги та недоліки. Принагідно варто зауважити, що не існує чітких відповідей на питання, який з існуючих статистичних критеріїв є найбільш коректним та потужним для вибірок

різного об'єму. Тому актуальною є задача розробки стійких методів обробки експериментальних спостережень.

Перевірка нормальності закону розподілу. Більшість існуючих критеріїв видалення аномальних

результатів спостереження спирається на їхню приналежність до нормального закону. Це говорить про те, що перш ніж приступати до перевірки результатів на аномальність, необхідно підтвердити нормальність закону



Рис. 1. Статистичні критерії математичної статистики, що використовуються для опрацювання результатів експериментальних досліджень

розподілу вибірки. Яким чином перевіряється нормальність? Перевірка статистичної гіпотези про припустимий закон невідомого розподілу [1] проводиться в такий же спосіб, як і перевірка гіпотез про параметри розподілу. Це відбувається за допомогою спеціально підібраної величини – критерію згоди (критерій перевірки гіпотези про допустимий закон розподілу) [2]. В числі найбільш застосовуваних в метрологічній практиці критеріїв згоди варто відзначити критерій Пірсона, критерій Колмогорова, складений критерій, критерій Мизеса-Смирнова ω^2 та ін. У практиці перевірки гіпотез про відповідність закону розподілу рекомендують застосовувати хоча б два методи перевірки для підвищення вірогідності прийняття рішення [3]. Особливо це важливо, коли характерні риси закону розподілу перебувають на його “хвостах”, що не завжди можна виявити при малому обсязі випробувань. На рис. 1 наводяться деякі статистичні критерії та область їх застосування, котрі приймалися до уваги авторами при аналізі існуючих на сьогоднішній день статистичних підходів обробки експериментальних даних. Звісно, кількість таких критеріїв, що подається в [4-7] є значно більшою, аніж це показано на рис. 1. Проаналізувавши ряд наукових праць в області математичної статистики, коротко зупинимося на деяких характерних особливостях для

кожного з критеріїв перевірки нормальності закону розподілу поданих на рис. 1:

- критерій Персона – доцільність та коректність його застосування доведена при великому об’ємі дослідних вибірок;
- складений критерій включає два незалежні критерії і використовується при перевірці гіпотези про нормальність розподілу вибірки результатів спостережень при умові $10 \leq n \leq 50$, де n – кількість елементів вибірки;
- критерій Колмогорова оснований на порівнянні емпіричної функції розподілу і теоретичної функції розподілу вибірки (може застосовуватися як критерій однорідності не лише як критерій згоди);
- критерій перевірки на симетричність призначений для перевірки гіпотез про симетричність спостережуваного закону розподілу при об’ємах вибірки $8 \leq n \leq 5000$;
- критерій перевірки на ексцес передбачений для використання при об’ємах вибірок $8 \leq n \leq 5000$;
- критерій Шапіро-Уїлка базується на аналізі лінійної комбінації відмінностей порядкових статистик, передбачено застосування критерію при об’ємі вибірки $8 \leq n \leq 50$;
- критерій Еппса-Паллі базується на порівнянні емпіричної і теоретичної характеристичних функцій. У стандарті передбачена доцільність застосування у випадку, коли $8 \leq n \leq 200$.

Перевірка на наявність промахів. У вітчизняній літературі можна знайти багато прикладів перевірки аномальності

льності результатів, проте дослідженню критеріїв, які призначені саме для таких цілей, не надається належної уваги [8]. Аналізуючи праці вітчизняних та іноземних науковців [3, 12, 13, 16], можна зробити висновок, що актуальним є дослідження названих критеріїв з точки нюансів коректності і доцільності застосування [9]. Для прикладу, науковці сусідніх держав говорять про застосування критерію Шовене для перевірки аномальності результатів спостереження для малих вибірок [10]. При цьому попередньо висувається нульова гіпотеза про приналежність результату, що підлягає перевірці на аномальність, вибірці (тобто результат не є аномальним), в той час, як альтернативна гіпотеза говорить про аномальність результату спостереження і не приналежність його вибірці, що підлягає статистичній обробці. Якщо розрахункове значення статистики вибраного критерію є більшим за встановлене граничне, що береться зі статистичних таблиць, то нульова гіпотеза не підтверджується, тобто цей результат – аномальний, і повинен бути виключений при подальшій обробці результатів спостережень. Критичні значення критерію Ірвіна та Романовського розраховані для достатньо великих вибірок (до 1000 та до ∞ для першого та другого відповідно); критерій варіаційного розмаху згадується в [11] як один із найпростіших методів перевірки аномальності результату; застосування критерію Діксона можливе лише для вибірок, об'єм яких знаходиться в межах $5 \leq n \leq 150$, в той час як табличні значення критерію Райта розраховано для кількості результатів спостережень більше 20 ... 50 та доцільність його застосування обґрунтована тоді, коли відома величина генерального середньоквадратичного відхилення; Критерій Шовене [10-12] застосовується для законів, що не суперечать нормальному, і ґрунтується на визначенні числа очікуваних результатів спостережень, які мають настільки ж великі похибки, як і підозрілий; як відомо з наукових праць та літературних джерел [4,6,13] критерій Смірнова використовується для перевірки однорідності вибірок, проте у [11] пропонується для вибірок обсягом $n \geq 25$ або при відомих значеннях генеральних середнього та середньоквадратичного відхилення застосовувати даний критерій для перевірки на відсутність аномальних результатів. Особливої уваги вартий критерій Граббса, котрий набув широкого застосування завдяки тому, що дає змогу перевірити одночасно кілька значень у вибірці на аномальність. На сьогоднішній день відомі наступні модифікації критерію Граббса:

- для перевірки на один промах (максимального або мінімального значення у вибірці);
- для перевірки перевірки на два промахи (одночасно двох найбільших або найменших значень та найбільшого і найменшого одночасно);
- для перевірки на три промахи (одночасно трьох мінімальних і трьох максимальних вибірових значень).

Відомо також, що отримано критичні точки критерію Граббса для вибірок з експоненціальним законом розподілу, а також виведена модифікація

критерію Граббса для перевірки на аномальність одночасно п'яти значень [14].

Оскільки критеріальні вимоги, що визначають межу, за якою знаходяться "грубі" результати спостережень у різних авторів різні, то перевірку слід виконувати відразу за кількома критеріями (рекомендується використовувати не менше трьох, із згаданих вище). Остаточний висновок про приналежність "підозрілих" результатів розглянутої сукупності спостережень слід робити за результатами, отриманими за більшістю з критеріїв.

Перевірка відтворюваності дослідів. Обробку результатів активного експерименту зазвичай проводять на основі регресійного аналізу. Проте для використання регресійного аналізу необхідне виконання наступних попередніх умов:

- похибки експерименту повинні бути розподілені за нормальним законом (вище уже наводилися критерії за допомогою яких забезпечується перевірка нормальності);
- досліді повинні бути відтворюваними.

Відповідно, перед тим як переходити до активного експерименту необхідно провести серію дослідів на об'єкті, який підлягає ретельному вивченню, для того, щоб підтвердити або ж заперечити зазначені вище попередні умови. Перша умова перевіряється за допомогою критеріїв згоди зазвичай, а перевірку другої умови забезпечують критерій Кохрена, Бартлетта та Фішера. Зупинемося на деяких особливостях згаданих критеріїв. Для перевірки гіпотези про приналежність двох вибірових дисперсій до однієї генеральної сукупності (їх однорідності), а відповідно і рівноточності серій вимірювання (показник відтворюваності дослідів у експерименті) використовується критерій Кохрена [3, 15]. Для перевірки гіпотези за допомогою критерію Кохрена необхідні результати декількох серій паралельних дослідів. У кожній із серій число результатів вимірювання повинна бути однаковою. Зазвичай, кількість серій не є більшою аніж 2-3, а кількість дослідів в серії теж може бути невеликою (достатньо навіть двох). Критерій Кохрена пропонується стандартом ISO 5725 для кількісної оцінки міжлабораторних спостережень. До переваг варто віднести простоту в обчисленні, та поряд має місце обмеження на число степенів вільності дисперсій, критерій виявляє відхилення лише в більшу сторону. Коли мова йде про застосування критерію Бартлетта, то характеристикою відтворюваності результатів серій (груп) вимірювань як і в попередньому випадку є вибірові дисперсії [4]. Варто зауважити, що відтворюваність і збіжність є показниками прецизійності вимірювання і кількісно виражаються через міжгрупову і внутрішньогрупову дисперсію. Критерій Бартлетта використовується для перевірки гіпотези про відтворюваність дослідів в тих випадках, коли маємо результати декількох серій паралельних дослідів, однак число дослідів в цих серіях різне [4,11,16]. До переваг даного критерію можна віднести той факт, що при його застосуванні не вимагається рівне число ступеней вільності дисперсій (тобто вибірки можуть бути різного об'єму), критерій ідентифікує двохстороннє відхилення як в найменшу так і в найбільше сторону. Разом з тим до ряду недоліків слід віднести:

- складність розрахунків;

- число степенів вільності будь-якої дисперсії повинне бути більше трьох;
- критерій дуже чутливий до відхилень від нормального розподілу даних.

Порівнюючи переваги та недоліки критеріїв Бартлетта і Кохрена, можна побачити, що вони є взаємодоповнюючими і повинні використовуватися спільно[4]. Критерій Фішера заснований на додаткових припущеннях про незалежність і нормальності вибірок даних. Здебільшого критерій Фішера стосується однофакторного дисперсійного аналізу і перед дослідником стоїть вимога присутності не менше як трьох рівнів факторів та не менше двох спостережень на кожному рівні і кількість цих спостережень на кожному рівні однакова. Як і в попередніх критеріях, до вибірок, що підлягають аналізу, висувається вимога необхідності нормального закону розподілу величин. Для забезпечення правильності обробки результатів спостережень, перевірки відхилень розподілів ймовірностей від нормального, коректного визначення основних методів розрахунку, оцінки точності методів і результатів вимірювання, розроблено ряд нормативних документів, ГОСТів та рекомендацій щодо їх застосування. Для прикладу стандарт [17] описує принципи планування експериментів, а також основний алгоритм кількісної оцінки прецизійності методів вимірювання на практиці; разом з тим цей стандарт міжнародного зразка є основою для прийняття рішень при проведенні міжлабораторних експериментів.

Висновки

На основі розглянутих критеріїв можемо зробити висновок, що використання кожного із них повинне бути попередньо обгрунтоване, необхідно вивчити поведінку вибраного критерію в залежності від умов його застосування, а також в залежності від поставленої перед дослідником задачі підібрати критерій згоди, однорідності та розсіювання, зваживши особливості застосування. Зокрема, при вимірюванні витрати газу, зважаючи на кореляційні зв'язки між параметрами процесу вимірювання, впливовими факторами та результатом, необхідним є вивчення їх взаємоповедінки та дослідження закону розподілу кожного зокрема і результативної ознаки. Для цього доцільним є застосування підходів статистичної ідентифікації, а саме параметричний та не параметричний метод встановлення виду закону розподілу. В основі кожного з таких підходів стоїть формування репрезентативної вибірки та побудова статистичного розподілу.

ЛІТЕРАТУРА

1. Эмпирическое исследование мощности критерия Саркади и его модификация / Золотухина Л.В.,

- Винник Е.В // Заводская лаборатория. 1985. - Т. 51. - №1. – С. 51-55.
2. Гмурман В.Е. Теория вероятностей и математическая статистика: Учебное пособие для вузов. – Изд. 7-е, стер. – М.: Высш. шк., 2001. – 479 с.
3. Володарський Є.Т., Кошева Л.О. статистична обробка даних: Навч. посібник. – К.: НАУ, 2008. – 308с.
4. Джонсон Н., Лион Ф.. Статистика и планирование эксперимента в технике и науке, М.:Мир, 1980. – 610с.
5. Курс лекцій з дисципліни «Статистика». Частина 1. Теорія статистики: В.П. Сторожук, О.В. Кустовська, Є.І. Ткач, І.М. Шост та ін.; За ред. Є.І. Ткача –Тернопіль: Економічна думка, 2006. – 224 с
6. М.Дж.Кендалл, А.Стьюарт. Статистические выводы и связи (пер.англ.). – М.:Наука, 1976.
7. Барковський В.В, Барковська Н.В., Лопатін О.К. Теорія ймовірностей та математична статистика: Навч.посібник. – К.: Центр навчальної літератури, 2006. – 424с.
8. Логістичний аналіз та моделювання розподілу вагонопотоків вантажного митного комплексу / А. М. Пасічник, Д. М. Козаченко, С. С. Кравчук // Вісник АМСУ. – 2009. – № 2. – С. 25–33
9. Исследование характеристик рядов ливневого стока малых рек бассейна верхнего Днестра [Электронный ресурс] / Б.В. Киндюк // Культура народов Причерноморья. — 2002. — N39. — С. 9-13
10. Тейлор Дж. Введение в теорию ошибок. М.: Мир, 1985.
11. Третьяк Л.Н. Обработка прямых измерений с многократными наблюдениями: Учебное пособие – Оренбург: ИПК ОГУ, 2002. – 60 с.
12. Викулин И.М. Оценка пригодности результатов измерений и исключение аномальных значений / Викулин И.М., Горбачев В.Э., Коробицын Б.В., Криськив С.К.// Наукові праці ОНАЗ ім. О.С. Попова, 2007, №2. – С.106-111
13. Лемешко Б.Ю. О сходимости распределений статистик и мощности критериев однородности Смирнова и Лемана-Розенблатта / Лемешко Б.Ю., Лемешко С.Б // Измерительная техника. 2005. № 12. – С.9-14
14. Сафарян Г.Г. Підвищення достовірності вимірювань показників якості електричної енергії з урахуванням ефекту кореляції даних: дис. ... кандидата техн. наук: 05.01.02 / Сафарян Григорій Гагікович. – Харків, 2008. – 223с.
15. Дерффель К. Статистика в аналитической химии. М. :Мир. 1994. 268 с
16. Специальные критерии согласия для малой выборки при оценке надежности нефтепромышленного оборудования. Ишемгузин И.Е., Атнагулов А.Р., Зотов А.Н. Нефтегазовое дело, 2008
17. ИСО 5725:1994 «Точность (правильность и прецизионность) методов и результатов измерений. Часть 1-6. Использование значений точности на практике»

