

ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

УДК 616. 07:519.248(045)

Е.Т. ВОЛОДАРСЬКИЙ¹, Л.О КОШЕВА²

1. Національний технічний університет України «Київський політехнічний інститут», Київ

2. Національний авіаційний університет, Київ

ІТЕРАЦІЙНА ПРОЦЕДУРА ОЦІНЮВАННЯ ЗАЛЕЖНОСТІ ПРЕЦИЗІЙНОСТІ РЕЗУЛЬТАТУ ВИПРОБУВАНЬ ВІД РІВНЯ ДОСЛІДЖУВАНОЇ ВЕЛИЧИНИ

Анотація. Розглянуто підхід до оцінювання залежності прецизійності результату від рівня досліджуваної величини, оснований на ітераційній процедурі, що дозволяє отримати адекватну модель регресії з гетероскедастичністю впливів.

Ключові слова: випробування, ітераційна процедура, прецизійність, зважений метод найменших квадратів.

Аннотация. Рассмотрен подход оценивания зависимости прецизионности результата от уровня исследуемой величины, основанный на итерационной процедуре, которая позволяет получить адекватную модель регрессии с гетероскедастичностью возмущений.

Ключевые слова: испытания, итерационная процедура, прецизионность, взвешенный метод наименьших квадратов

Abstract. The approach of estimation precision of the result from the level of the investigated quantity, which is based on the iterative procedure, which allows to get an adequate regression model with heteroscedasticity disturbances.

Keywords: the tests, iterative procedure, pretsyziynist weighted least squares method.

Вступ

Точність результатів випробувань у перше чергу залежить від прецизійності поточних результатів, яка характеризується дисперсією повторюваності (коли спостерігається тільки вплив випадкових величин) та дисперсією відтворюваності, що залежить від можливої розбіжності між випробувальним обладнанням, умовами проведення експерименту та кваліфікацією операторів у лабораторіях. Як впливає з міжнародного стандарту [1] невизначеність результатів вимірювання при випробуваннях визначається дисперсією відтворюваності стандартизованої методики. Тому для забезпечення порівнянності результатів, що отримують за даною методикою у будь-якій лабораторії необхідно визначити дисперсію відтворюваності, яка включає в себе міжлабораторну дисперсію та дисперсію повторюваності. Для встановлення цих характеристик у методиках проводиться спільний лабораторний експеримент, за результатами якого нормується можливе розсіювання результатів та їх зміщення, достовірність оцінювання якого зв'язано з дисперсією. Як показала практика, в багатьох випадках прецизійність результату випробувань не є постійною величиною – її значення залежить від рівня (значення) досліджуваної величини у зразках випробувань, що необхідно враховувати. Саме від способу врахування впливових величин залежить достовірність результатів.

Постановка проблеми

Коли вимірювана величина змінюється у відносно вузькому діапазоні, то припускається, що сумарний вплив випадкових величин носить адитивний характер, а зміна оцінки дисперсії при досліді визначається обмеженим обсягом вибірових даних. При зміні вимірюваної величини в широкому динамічному діапазоні, як це зазвичай відбувається при випробуваннях, таке припущення не завжди виконується. Тому необхідно за результатами експерименту визначити аналітичну залежність, яка враховує зміну дисперсії від рівня досліджуваної величини, і в подальшому використовувати її для оцінювання точності поточних результатів.

У випадку, коли прецизійність випробувань залежить від рівня досліджуваної величини, тобто коли отримані результати не є рівноточними, безпосереднє застосування методу найменших квадратів для отримання адекватної залежності для дисперсії дає не ефективні оцінки. У такому випадку доцільно застосовувати зважений метод найменших квадратів. Саме цей метод часто застосовується в ситуаціях, коли дисперсія результатів вимірювань не постійна на області значень незалежної змінної. Тоді можна приписати наявним результатам ваги, рівні зворотним величинам відповідної дисперсії, і отримати оцінки зважених найменших квадратів. Умова, за якою дисперсія впливів однакова для всіх спостережень результату називається умовою гомоскедастичності. Порушення умови гомоскедастичності впливів означає, що дисперсія залежить від значень факторів. Такі регресійні моделі називаються моделями з гетероскедастичністю впливів.

При аналізі експериментальних даних необхідно, по-перше, вирішити питання, існує чи ні регулярна залежність оцінок прецизійності від рівня досліджуваної величини і, по-друге, якщо така залежність існує, знайти відповідний аналітичний вираз, який дозволяє при проведенні дослідження обчислювати поточні показники прецизійності.

Мета

Показати підхід, який дозволить встановити адекватну залежність дисперсії результату вимірювання від значень оцінок вхідної величини, які визначені за експериментальними даними.

Основна частина

Сучасні тенденції до встановлення та оцінювання характеристик точності результатів випробувань полягають у відході від ідеології призначення їх припустимих меж та від ідеології приймання розробленої методики на підставі вимірювань організації-розроблювача, а вимагають нових підходів до їх нормування на основі статистичних методів. Раніше для оцінки метрологічних характеристик вимірювань при випробуванні об'єкта досить було провести внутрішньолабораторний експеримент. Але при цьому можна було говорити про точність результатів, отриманих однією лабораторією. Для зіставлення результатів, забезпечення довіри до них, цього не достатньо. Сучасний регламент атестації методик вимагає переходити до інших форм оцінювання характеристик методики, зокрема, до спільного міжлабораторного експерименту. Правильно організовані спільні експерименти з оцінювання характеристик точності (прецизійності та правильності) дозволяють визначити вірогідні характеристики повторюваності, проміжної прецизійності, відтворюваності й правильності методик, що стандартизуються або атестуються. Розглянутий підхід однаковий як для аналізу стандартного відхилення повторюваності, так і відтворюваності.

При проведенні спільного експерименту діапазон вхідних величин розбивають на q рівнів, для кожного j -го рівня необхідно мати сукупність однорідних зразків для того, щоб охопити діапазон можливих значень при випробуванні за методикою, що атестується [2]. Координатор, який проводить спільний експеримент з оцінювання залежності прецизійності від рівня досліджуваної величини відбирає p лабораторій відповідної кваліфікації, у які розсилаються по q зразків для кожного j -го рівня. Кожна з $i = \overline{1, p}$ лабораторій в умовах повторюваності проводить по n спостережень, на підставі яких для кожного j -го рівня визначається оцінка лабораторної дисперсії $S_{W_{ij}}^2$.

Для кожного j -го рівня після статистичної перевірки однорідності лабораторних результатів та вилучення викидів визначаються оцінки загального середнього

$$\hat{m}_j = \left(\sum_{i=1}^{p_j} \hat{m}_{ij} \right) / p$$

та оцінки дисперсії повторюваності

$$s_{rj}^2 = \left(\sum_{i=1}^{p_j} s_{W_{ij}}^2 \right) / p,$$

де \hat{m}_{ij} – середнє значення на j -му рівні у i -й лабораторії;

p_j – кількість лабораторій, результати яких не містять при проведенні експерименту на j -му рівні викидів та квазівикидів.

Наявні експериментальні дані, отримані при спільному експерименті, дозволяють за допомогою методу найменших квадратів [3] встановити приблизний аналітичний вираз для залежності дисперсії повторюваності від рівня досліджуваної величини. Найчастіше припускають, що дана залежність має вигляд $(a + b_m)$ з можливим значенням $a = 0$.

Отримавши приблизний вид аналітичної залежності, необхідно провести дослідження, залежить чи ні прецизійність від значення досліджуваної величини. Для цього загальний діапазон вимірювань необхідно розбити на декілька піддіапазонів, виходячи з умови однорідності дисперсії відтворюваності S_R^2 в кожному окремо взятому піддіапазоні, а потім розрахувати відношення максимальної дисперсії до

мінімальної $\frac{S_{RВ}^2}{S_{RH}^2}$, де індекси «в» та «н» відносяться відповідно до верхньої та нижньої меж

піддіапазонів. Якщо $\frac{S_{RВ}^2}{S_{RH}^2} \geq 2$ [4], то можна вважати, що прецизійність залежить від значення

досліджуваної величини, і можна переходити до встановлення точного виду аналітичної залежності, застосовуючи методи зваженого регресійного аналізу, а саме ітераційний метод.

Оскільки на відміну від класичного регресійного аналізу застосовуються не значення незалежної змінної m_j , а її оцінки \hat{m}_j , які є вибірковими значеннями випадкової величини, то необхідно вводити вагові коефіцієнти $w_{0j} = 1/s_{rj}^2$, які в подальшому уточнюються на кожному k -му кроці ітерації, ($k = 1, 2, \dots$).

Встановлення залежності розглянемо на прикладі дисперсії повторюваності.

На підставі умови

$$Q = \sum_{j=1}^q w_{0j} [s_{rj}^2 - (\hat{a}_0 + \hat{b}_0 \hat{m}_j)]^2 \rightarrow \min$$

одержимо оцінки коефіцієнтів:

$$\hat{a}_0 = \frac{\left(\sum_{j=1}^q w_{0j} \hat{m}_j^2 \right) \left(\sum_{j=1}^q w_{0j} s_{rj} \right) - \left(\sum_{j=1}^q w_{0j} \hat{m}_j \right) \left(\sum_{j=1}^q w_{0j} \hat{m}_j s_{rj} \right)}{\left(\sum_{j=1}^q w_{0j} \right) \left(\sum_{j=1}^q w_{0j} \hat{m}_j^2 \right) - \left(\sum_{j=1}^q w_{0j} \hat{m}_j \right)^2}, \quad (1)$$

$$\hat{b}_0 = \frac{\left(\sum_{j=1}^q w_{0j} \right) \left(\sum_{j=1}^q w_{0j} \hat{m}_j s_{rj} \right) - \left(\sum_{j=1}^q w_{0j} \hat{m}_j \right) \left(\sum_{j=1}^q w_{0j} s_{rj} \right)}{\left(\sum_{j=1}^q w_{0j} \right) \left(\sum_{j=1}^q w_{0j} \hat{m}_j^2 \right) - \left(\sum_{j=1}^q w_{0j} \hat{m}_j \right)^2}. \quad (2)$$

Маючи оцінки коефіцієнтів передбачуваної залежності, на наступному ($k + 1$) кроці обчислюють значення СКВ повторюваності у «точках» цієї залежності, які відповідають наявним оцінкам незалежної змінної \hat{m}_j , тобто знаходять уточнену оцінку (перше наближення) СКВ повторюваності $s_{rj_0} = \hat{a}_0 + \hat{b}_0 \hat{m}_j$ у j -х точках.

Виходячи із значення s_{rj_0} перераховують вагові коефіцієнти $w_{1j} = 1/s_{rj_0}^2$, які підставляють у вирази (1) і (2) замість w_{0j} , і знаходять уточнені в першому наближенні значення \hat{a}_1 та \hat{b}_1 .

У підсумку приходять до уточненого рівняння регресії, яке характеризує залежність дисперсії повторюваності від рівня досліджуваної величини $s_{rj_1} = \hat{a}_1 + \hat{b}_1 m$. Підставивши в цей вираз наявні значення \hat{m}_j , визначають уточнені значення оцінки повторюваності для j -го рівня, які використовуються для обчислення нових вагових коефіцієнтів $w_{2j} = 1/s_{rj_1}^2$, визначають \hat{a}_2 та \hat{b}_2 , і приходять до уточненого рівняння $s_{rj_2} = \hat{a}_2 + \hat{b}_2 m$.

Починаючи із другого кроку ітерації ($k = 2$) необхідно порівнювати оцінки СКВ у j -х точках для поточного k -го циклу й попереднього ($k - 1$). Для цього обчислюється різниця:

$$\left| s_{rjk} - s_{rj(k-1)} \right|,$$

за якою приймається рішення про припинення процедури уточнення коефіцієнтів регресійної моделі. При ухваленні рішення про достатню точність апроксимації одержуємо остаточний вираз:

$$s_r = \hat{a} + \hat{b} \hat{m},$$

за яким потім при проведенні випробувань визначають S_r для будь-якого отриманого поточного результату \hat{m} та його прецизійність. Як показує досвід, ітераційна процедура в багатьох випадках припиняється при $k = 3$.

Аналогічно наведеному оцінюється й характеристика відтворюваності.

Даний підхід може бути використаний при оцінюванні залежності повторюваності від рівня досліджуваної величини не тільки під час проведення спільних експериментів точності, але й і в одній окремо взятій лабораторії, що нерідко має місце у лабораторній практиці, особливо під час проведення валідації методики.

Для лабораторії надійність отриманих результатів буде залежати від кількості повторних вимірювань в лабораторіях n на кожному j -у рівні та стабільності характеристик стандартного зразка, застосовуваного при їх проведенні.

Розглянемо реалізацію ітераційної процедури на прикладі. Припустимо, що діапазон варіювання знаходиться в межах від 4,0 до 20,0 одиниць [1]. Для визначення аналітичної залежності було обрано п'ять ($q = 5$) рівнів проведення експерименту, які перекривають весь робочий діапазон, а саме ($j = \overline{1,5}$) 4,0; 8,0; 14,0; 16,0; 20,0. При цьому не потрібно задавати значення з високою точністю, тому що в подальшому для побудови регресійної залежності вихідними є оцінки цих значень, обчислені за результатами вимірювань. В табл.1 наведено обчислені для кожного рівня значення середніх \hat{m}_j та оцінок СКВ

S_{rj} .

Таблиця 1

j	1	2	3	4	5
\hat{m}_j	3,94	8,28	14,18	15,59	20,41
S_{rj}	0,092	0,1709	0,127	0,337	0,393
w_{0j}	118	31	62	8,8	6,5

Розраховують вагові коефіцієнти за формулою:

$$w_{0j} = \frac{1}{S_{rj}^2}$$

Так, для першого ($j=1$) рівня ваговий коефіцієнт становить $w_{01} = \frac{1}{(0,092)^2} = 118$. Отримані результати зводять до останнього рядка табл.1.

Далі за виразами (1) та (2) знаходять початкові оцінки коефіцієнтів залежності $\hat{a}_0 = 0,005$ та $\hat{b}_0 = 0,0090$, які використовуються для первинного уточнення оцінки СКВ. Так для третьої експериментальної «точки» ($j = 3$) отримаємо значення :

$$s_{r30} = 0,058 + 0,009 \times 14,18 = 0,185. \quad (3)$$

Аналогічно обчислюють уточнені значення СКВ для всіх експериментальних «точок», які використовуються для визначення перерахованих вагових коефіцієнтів w_{1j} , наведених у табл.2.

Таблиця 2

j	1	2	3	4	5
S_{rj}	0,093	0,132	0,185	0,197	0,240
w_{1j}	116	57	29	26	17

Значення w_{1j} підставляють у вирази (1) та (2) замість w_{0j} та обчислюють уточнені коефіцієнти залежності

$$s_{r1} = 0,030 + 0,0156 \hat{m} \quad (4)$$

Наступним кроком є обчислення за виразом (4) значень s_{rj1} для всіх точок залежності, в яких проводилися досліди, та визначення нових значень вагових коефіцієнтів $w_{2j} = \frac{1}{s_{r1j}^2}$, підстановка яких у вирази (1) та (2) дозволяє знайти значення $\hat{a}_2 = 0,032$ та $\hat{b}_2 = 0,0154$.

Обчислені за рівнянням

$$s_{r2} = 0,032 + 0,0154 \hat{m} \quad (5)$$

значення s_{rj2}^* наведені у нижній строчці табл. 3. (Символ * означає, що даний результат приймається як кінцевий)

Таблиця 3

j	1	2	3	4	5
s_{rj1}	0,092	0,159	0,251	0,273	0,348
w_{2j}	118	40	16	13	8
s_{rj2}^*	0,093	0,160	0,251	0,273	0,348

Порівняння значень s_{rj2}^* з відповідними значеннями s_{rj1} показує, що їх найбільше відносне відхилення для всіх рівнів не перевищує 0,1%.

Таким чином, залежність (5) може бути використана для оцінювання дисперсії повторюваності при проведенні вимірювань величини, значення якої змінюються в певному діапазоні вимірювань.

Висновки

При встановленні аналітичної залежності дисперсії результату вимірювання від значення (рівня) досліджуваної величини використовується зважений метод найменших квадратів, оскільки на відміну від класичного підходу, незалежні змінні задані з похибкою, бо визначені за дослідними даними як середнє значення для кожного рівня.

Застосування ітераційної процедури визначення оцінок коефіцієнтів регресії, коли за результатами попереднього обчислення оцінки дисперсії коригуються значення вагового коефіцієнту на наступному кроці, дозволяє отримати адекватну залежність дисперсії вимірювання від значення вхідної величини.

Список використаної літератури

1. Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation: ISO/TS 21748:2004. – (Міжнародний стандарт).
2. Точність (правильність і прецизійність) методів та результатів вимірювання. Частина 4. Основні положення та визначення: (ISO/IEC 5725-1:1994, IDT): ДСТУ ГОСТ ИСО 5725-1:2005. – [Чинний від 2006-07-01]. – К.: Держспоживстандарт України, 2006. – 31 с. – (Національний стандарт України).
3. Володарський Є.Т. Статистична обробка даних:[навч. посіб.] /Є.Т. Володарський, Л.О. Кошева. – К.: Книжкове видавництво НАУ, 2008. – 308 с.
4. Показники точності, правильності, прецизійності методик кількісного хімічного аналізу. Методи оцінювання: ДСТУ-Н РМГ 61:2006.– [Чинний від 2007-07-01]. – К.: Держспоживстандарт України, 2007. – 38 с.– (Національний стандарт України).

Стаття надійшла до редакції: 12.09.11.

Відомості про авторів

Володарський Євген Тимофійович, д.т.н., професор кафедри автоматизації експериментальних досліджень, Національний технічний університет України «Київський політехнічний інститут», просп. Перемоги, 37, Київ 03056, Україна, тел / факс (044) 4549800, e-mail: vet_1@voliacable.com.

Кошева Лариса Олександрівна, д.т.н., професор кафедри біокібернетики та аерокосмічної медицини, Національний авіаційний університет, просп. Космонавта Комарова, 1, Київ, 03058, Україна, тел / факс (044) 4067442, e-mail: arnis@ukrpost.net.