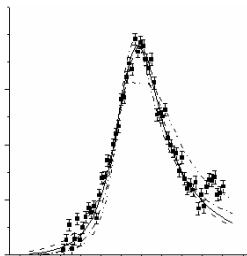
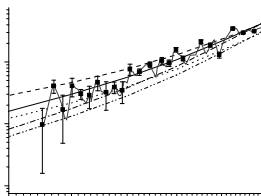
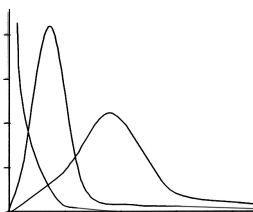
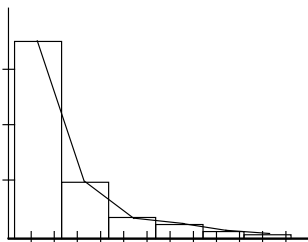


І.М.Каденко, В.А.Плюйко

ОСНОВИ МЕТОДІВ ОЦІНОК СТАТИСТИЧНИХ ДАНИХ ТА ФУНКЦІЙ ЇХ РОЗПОДІЛУ



КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ ТАРАСА ШЕВЧЕНКА

Каденко І.М., Плюйко В.А.

ОСНОВИ МЕТОДІВ ОЦІНОК СТАТИСТИЧНИХ ДАНИХ ТА ФУНКЦІЙ ЇХ РОЗПОДІЛУ

Навчальний посібник



УДК 591.21(075.8)
ББК 22.171я73
К13

Рецензент
д-р фіз.-мат. наук, проф. А. Т. Рудчик

*Рекомендовано до друку вченою радою фізичного факультету
(протокол № 7 від 21 січня 2001 року)*

Каденко І.М., Плюйко В.А.
К13 Основи методів оцінок статистичних даних та функцій їх розподілу :
Навчальний посібник. — К.: Видавничо-поліграфічний центр "Київський
університет", 2003. — 145 с.
ISBN 966-594-480-0

Викладено основи методів статистичної обробки результатів вимірів та способів обчислення надійності одержаних оцінок.

Для студентів, аспірантів та викладачів фізичних і фізико-технічних спеціальностей вищих навчальних закладів.

УДК 591.21(075.8)
ББК 22.171я73

© Каденко І.М., Плюйко В.А.
© Київський національний університет імені Тараса Шевченка,
ВПЦ "Київський університет", 2003

ЗМІСТ

Розділ 1

ОСНОВНІ ПОЛОЖЕННЯ ТЕОРІЇ ЙМОВІРНОСТЕЙ	7
1.1. Випадкові величини та їх функції розподілу	7
1.2. Числові характеристики випадкових величин.....	17
1.3. Наближене обчислення математичного сподівання та дисперсії для функції випадкових величин	25
1.4. Деякі приклади розподілів імовірностей.....	31

Розділ 2

ОСНОВИ МЕТОДУ ВИБІРКИ	47
2.1. Математична статистика та її завдання. Функція розподілу вибірки.....	47
2.2. Загальні вимоги до оцінок параметрів.....	49
2.3. Емпіричні функції розподілу та характеристики вибірок	53
2.4. Довірчі інтервали оцінок	63

Розділ 3

ОЦІНКИ ПАРАМЕТРІВ ФУНКЦІЇ РОЗПОДІЛУ	69
3.1. Класифікація оцінок за методами їх знаходження.....	69
3.2. Метод моментів.....	69
3.3. Метод максимальної правдоподібності.....	72
3.4. Об'єднання оцінок.....	78
3.5. Критерії оцінки граничних значень елементів вибірки. Вилучення аномальних похибок	85

Розділ 4

МЕТОД НАЙМЕНШИХ КВАДРАТІВ	91
4.1. Принцип найменших квадратів	91
4.2. Наближення функціями, лінійними за параметрами.....	97
4.3. Метод лінеаризації в МНК	103

Розділ 5

ОЦІНКИ РІВНІВ ДЕТЕКТУВАННЯ В ЯДЕРНО-ФІЗИЧНИХ ЕКСПЕРИМЕНТАХ МЕТОДОМ ПЕРЕВІРКИ СТАТИСТИЧНИХ ГІПОТЕЗ 109

- 5.1. Статистична гіпотеза, схема її перевірки
та загальне визначення рівнів детектування 109
- 5.2. Оцінка рівнів детектування при визначенні
кількості відліку подій 115
- 5.3. Асимптотичні оцінки граничних рівнів швидкості
відліку сигналів 118

Додаток

ПРОГРАМА FUMIL для ЗНАХОДЖЕННЯ ПАРАМЕТРІВ МЕТОДОМ НАЙМЕНШИХ КВАДРАТІВ 129

Рекомендована література 145

ПЕРЕДМОВА

Серед розділів математики, які широко використовуються в сучасній фізиці, важливу роль відіграють теорія ймовірностей і математична статистика. Ці галузі математики стали невід'ємною частиною сучасної теоретичної фізики при розгляді властивостей систем з мікрочастинок та процесів, які в них проходять. Наприклад, мікрооб'єкти, які досліджуються в нерелятивістській фізиці, описуються рівняннями квантової механіки відносно хвильової функції, квадрат якої інтерпретується як ймовірність знаходження системи в заданій точці простору.

Одночасно теорія ймовірностей і математична статистика стали важливим інструментом експериментальних досліджень, оскільки у фізиці насамперед мають справу з результатами вимірів, які є випадковими величинами. Появу випадкових результатів вимірів обумовлено або випадковою природою самого явища, що досліджується, або різними випадковими впливами, що неконтрольовано вносяться у процес вимірів.

Для випадковостей також є характерними і певні закономірності. Математичний апарат для вивчення таких закономірностей базується на теорії ймовірностей, яка є галуззю науки, що займається дослідженням закономірностей, притаманних випадковим величинам.

Якщо теорія ймовірностей вивчає закономірності випадкових явищ, базуючись на абстрактному описі дійсності, то математична статистика безпосередньо оперує результатами спостережень випадкового явища. Користуючись результатами теорії ймовірностей, математична статистика дозволяє не тільки оцінити значення характеристик, що досліджуються, але й виявити ступінь достовірності висновків, які отримані при обробці даних. Тобто математична статистика є розділом математики, що вивчає методи збирання, систематизації та оброблення результатів спостережень з метою виявлення статистичних закономірностей.

За останні роки відбулися суттєві зміни в методах оброблення результатів експериментальних досліджень. Математична статистика виявилася потужним

засобом отримання інформації з експериментальних даних. У сучасних дослідженнях майже не вдається безпосередньо вимірювати фізичні величини, що викликають інтерес. Експерименти, як правило, непрямі. У них спостерігаються величини, що складним чином залежать від характеристик, які досліджуються. Лише статистичний аналіз дозволяє робити надійні висновки відносно їх властивостей.

Широкому впровадженню математико-статистичних методів досліджень також сприяла поява електронних обчислювальних машин, зокрема персональних комп'ютерів. Створення зручних і потужних пакетів обчислювальних програм для обробки інформації за допомогою персональних комп'ютерів та їх використання, зокрема, у вбудованих у вимірювальну апаратуру системах повної автоматизованої обробки експериментальних даних, зробили ці методи більш доступними і наочними. Однак обробка результатів експериментів у фізиці потребує ясного уявлення про процес виміру, а також знань основних положень теорії ймовірностей і математичної статистики, вмілого використання методів обчислювальної математики і програмування.

У даному навчальному посібнику стисло викладено базові положення теорії ймовірностей та методи математичної статистики, що застосовуються при обробці експериментальних даних. Зокрема, розглянуто точкові та інтервальні методи оцінок невідомих параметрів, визначення довірчих інтервалів оцінок, критерії оцінки граничних значень елементів вибірки та вилучення аномальних похибок. Описано методи максимальної правдоподібності та найменших квадратів. Розглянуто алгоритм лінеаризації рівнянь методу найменших квадратів та його реалізація в комп'ютерній програмі *FUMILI*. Наведено приклади використання ймовірносних і математико-статистичних методів у задачах обробки експериментальних даних.

Розділ 1

ОСНОВНІ ПОЛОЖЕННЯ ТЕОРІЇ ЙМОВІРНОСТЕЙ

1.1. Випадкові величини та їх функції розподілу

Випадковою називається величина, яка може приймати будь-яке значення із множини всіх її можливих значень. Наприклад, випадковими величинами є, як правило, експериментальні дані, тому що похибка результатів кожного вимірювання може залежати від впливу багатьох випадкових факторів і не може бути записана у вигляді детермінованої функції від деяких відомих величин.

Усю сукупність значень, які може мати випадкова величина, називають генеральною сукупністю. Усяка обмежена частина її елементів, що розглядається окремо, називається вибіркою. Експериментальні дані, як правило, є вибірками значень випадкової величини. Конкретне значення x випадкової величини ξ називають реалізацією випадкової величини. Якщо випадкова величина набуває всіх значень із деякого числового інтервалу, то вона називається неперервною. Випадкові величини, які приймають тільки певні ізольовані одне від одного числові значення, називаються дискретними.

Ймовірність реалізації випадкової величини є кількісна міра, що показує, якою великою є можливість появи даної реалізації.

Існують три визначення поняття ймовірності: класичне, статистичне й аксіоматичне.

Математично строгим є аксіоматичне визначення ймовірності, яке було запропоновано 1933 р. Колмогоровим (1903–1987). Воно базується на введенні ймовірності як числової функції, що визначена у просторі подій і задовольняє деяким властивостям, які задаються системою аксіом.

Класичне і статистичне визначення ймовірності є більш фізичними і не такими абстрактними, як аксіоматичне. Класичне визначення ймовірності появи події будується як відношення числа результативних випробувань, сприятливих до події, до числа всіх випробувань (Паскаль (1623–1682), Ферма (1601–1665)). Останні вважаються однаково важливими і несумісними між собою.

Ширшим порівняно із класичним визначенням ймовірності й також наочним є статистичне визначення, запропоноване Мізесом (1883–1953). У цьому випадку припускається можливість багаторазового знаходження реалізацій випадкової величини шляхом їх вимірювання у стаціонарних умовах. На-

приклад, розглянемо визначення ймовірності дискретної випадкової величини. Уведемо вибірку відносну частоту появи її v -ї реалізації

$$P_v = m_v / M, \quad v = 1, \dots, N, \quad (1.1.1)$$

де M – повна кількість вимірювань (випробувань) m_v – кількість вимірювань (випробувань), у яких з'явилась v -та реалізація x_v випадкової величини ξ ; N – загальна кількість реалізацій. При скінченних M вибіркова частота P_v є випадковою, але при $M \rightarrow \infty$ вона прямує до стійкої і не випадкової границі

$$P_v = \lim_{M \rightarrow \infty} m_v / M. \quad (1.1.2)$$

Статистична частота P_v приймається як міра об'єктивної можливості появи v -ї реалізації випадкової величини й ототожнюється з імовірністю її появи. Існування границі (1.1.2) впливає із закону великих чисел, згідно з яким середні результати великого числа випадкових явищ асимптотично (при $M \rightarrow \infty$) стійкі, не залежать від окремих (індивідуальних) випадкових явищ та прямують до деякого скінченного значення (теорема Бернуллі).

Сукупність значень імовірностей $P_v \equiv P(\xi = x_v)$ для всіх можливих реалізацій $v = 1, \dots, N$ називається законом розподілу випадкової дискретної величини. Він є єдиною і повною характеристикою випадкової величини. З визначення (1.1.2) випливає, що ймовірність появи реалізації є меншою одиниці, а повна ймовірність появи випадкової величини є нормованою на одиницю

$$0 \leq P_v \equiv P(\xi = x_v) \leq 1, \quad \sum_{v=1}^N P_v = 1. \quad (1.1.3)$$

Кажуть, що задано закон розподілу випадкової дискретної величини, коли вказано перелік її можливих значень та відповідні значення ймовірностей. Такий закон, наприклад, може задаватися у вигляді так званої таблиці розподілу, перший рядок якої містить можливі значення (x_v) величини ξ , а другий – відповідно ймовірності P_v появи цих значень.

Таблиця 1.1

Таблиця розподілу випадкової величини ξ

ξ	x_1	x_2	...	x_M
P	P_1	P_2	...	P_M

Закон розподілу дискретної величини можна зобразити графічно, для цього на площині у прямокутній системі координат будують точки

$M_1(x_1, P_1), M_2(x_2, P_2), \dots, M_n(x_n, P_n)$, де x_i – можливі значення ξ , а P_i – відповідні ймовірності, і з'єднують їх відрізками прямих. Отриману фігуру називають многокутником розподілу.

Ймовірність $P(x_1 \leq \xi \leq x_2)$ знаходження реалізації неперервної випадкової величини в інтервалі $x_1 \leq \xi \leq x_2$ визначається за формулою

$$P(x_1 \leq \xi \leq x_2) = \lim_{M \rightarrow \infty} \frac{m_{12}}{M}, \quad (1.1.4)$$

де M – повна кількість вимірювань, а m_{12} – кількість знаходжень реалізацій ξ в даному інтервалі. У випадку, коли $x_1 \rightarrow -\infty$, а $x_2 = x$, ймовірність $P(-\infty \leq \xi \leq x)$, що розглядається як функція змінної x , називається функцією розподілу випадкової величини ($F_\xi(x)$), або інтегральним законом розподілу, тобто

$$F_\xi(x) \equiv P(-\infty \leq \xi \leq x) = \lim_{M \rightarrow \infty} \frac{m_x}{M_1}, \quad (1.1.5)$$

де m_x – кількість знаходжень реалізацій величини ξ в інтервалі $(-\infty, x)$. З визначення (1.1.5) випливає, що функція розподілу $F(x) \equiv F_\xi(x)$ є додатною і має такі властивості:

$$\begin{aligned} F(x_2) &\geq F(x_1), \text{ при } x_2 \geq x_1, \\ F(-\infty) &= 0, \quad F(+\infty) = 1; \\ P(x_1 \leq \xi \leq x_2) &= F(x_2) - F(x_1). \end{aligned} \quad (1.1.6)$$

Неперервну випадкову величину можна характеризувати ймовірністю її появи, віднесену до одиничного інтервалу реалізації. Таку величину називають щільністю $p(x)$ розподілу випадкової величини, а саме щільність $p(x)$ є ймовірністю появи реалізацій величини ξ в інтервалі $x \leq \xi \leq x + \Delta x$, віднесену до довжини Δx цього інтервалу при $\Delta x \rightarrow 0$:

$$p(x) = \lim_{\Delta x \rightarrow 0} \frac{1}{\Delta x} P(x \leq \xi \leq x + \Delta x) = \lim_{\substack{M \rightarrow \infty \\ \Delta x \rightarrow 0}} \frac{\Delta m_x}{M}, \quad (1.1.7)$$

де Δm_x – кількість реалізацій випадкової величини в даному інтервалі, а Δx – його довжина. Таким чином, щільність $p(x)$ визначає ймовірність

$p(x)dx$ того, що випадкова величина при вимірюванні (випробуванні) в однакових умовах набуде значень з інтервалу $(x, x + dx)$. З визначення (1.1.7) виходить, що вона є невід'ємною ($p(x) \geq 0$) і пов'язана з ймовірністю та функцією розподілу таким чином:

$$p(x) = \lim_{\Delta x \rightarrow 0} \frac{1}{\Delta x} P(x \leq \xi \leq x + \Delta x) = \lim_{\Delta x \rightarrow 0} \frac{1}{\Delta x} [F(x + \Delta x) - F(x)] = \frac{dF(x)}{dx},$$

$$P(x \leq \xi \leq x + \Delta x) = F(x + \Delta x) - F(x) = dF(x) = p(x)dx,$$

$$F(x) = P(-\infty \leq \xi \leq x) = \int_{-\infty}^x p(t)dt,$$

$$P(x_1 \leq \xi \leq x_2) = F(x_2) - F(x_1) = \int_{x_1}^{x_2} p(t)dt, \quad (1.1.8)$$

$$\int_{-\infty}^{+\infty} p(x)dx = P(-\infty \leq \xi \leq +\infty) = 1.$$

Зауважимо, що у фізиці функції, які визначаються формулами вигляду $dQ(x)/dx$ зазвичай називають густиною величини $Q(x)$, тобто щільність розподілу $p(x) \equiv dF(x)/dx$ випадкової величини є густиною її функції розподілу $F(x)$. Оскільки $F(x) = P(-\infty \leq \xi \leq x)$, то щільність $p(x)$ можна також інтерпретувати як густину ймовірності (густину розподілу ймовірності) випадкової величини. Єдина відмінність полягає в тому, що щільність розподілу завжди нормована на одиницю.

Очевидно, що за таблицею розподілу дискретної величини (табл. 1.1) можна також побудувати і функцію розподілу випадкової дискретної величини

$$F(x) = \sum_{x_v < x} P_v. \quad (1.1.9)$$

Цей запис означає, що підсумовування поширюється на ті індекси v , для яких $x_v < x$. Функція розподілу дискретної випадкової величини має розриви в кожній із точок x_v , причому згідно з визначенням (1.1.6) величина стрибка в точці x_v дорівнює ймовірності появи реалізації із значенням $x = x_v$:

$$F(x_v + 0) - F(x_v) = P(x = x_v). \quad (1.1.10)$$

Для скінченної множини можливих значень величини ξ графік $F(x)$ являє собою лінію у вигляді сходинок з інтервалами постійних значень між сусідніми значеннями x .

Користуючись узагальненими функціями, можна ввести і щільність розподілу випадкової дискретної величини згідно з формулою

$$p(x) = \sum_{v=1}^N P_v \delta(x - x_v), \quad (1.1.11)$$

де $\delta(x - x_v)$ – узагальнена δ -функція Дірака, така що

$$\delta(x - x_v) = \begin{cases} 0, & x \neq x_0, \\ \infty, & x = x_0, \end{cases} \quad \int_{-\infty}^{\infty} dx \delta(x - x_v) = 1. \quad (1.1.12)$$

Для узагальненої δ -функції є справедливим співвідношення

$$\int_{-\infty}^x dt \delta(t - x_v) = \theta(x - x_v) = \begin{cases} 1, & x \geq x_v, \\ 0, & x < x_v. \end{cases} \quad (1.1.13)$$

Урахувавши (1.1.13) та користуючись визначеннями (1.1.11), (1.1.8) щільності й функції розподілу, легко перевірити виконання умови (1.1.9):

$$F(x) \equiv \int_{-\infty}^x p(t) dt = \int_{-\infty}^x dt \sum_{v=1}^N P_v \delta(x - x_v) = \sum_{x_v < x} P_v. \quad (1.1.14)$$

Лінію, що зображує щільність розподілу, також називають кривою розподілу.

Дискретними і неперервними випадковими величинами не вичерпуються всі можливі типи випадкових величин. Так, випадкова величина може поводити себе в одних інтервалах як неперервна, а в інших – як дискретна.

Крім одномірних випадкових величин, існують багатовимірні випадкові величини або випадкові вектори.

Імовірність $P(\xi_1, \dots, \xi_n \in D)$ знаходження реалізації $x_1, \dots, x_n$ системи з n випадкових величин $\xi_1, \dots, \xi_n$ у n -вимірній області D може бути визначена статистичним чином, за аналогією з одновимірним випадком (1.1.2), у такий спосіб:

$$P(\xi_1, \dots, \xi_n \in D) = \lim_{M \rightarrow \infty} \frac{m(D)}{M}, \quad (1.1.15)$$

де $m(D)$ – кількість знаходження реалізацій $x_1, \dots, x_n$ системи в області D , а M – повне число випробувань.

Під функцією розподілу, або інтегральним законом розподілу, (F) системи випадкових величин розуміють імовірність знаходження реалізацій в області, що визначається нерівностями

$$-\infty \leq \xi \leq x_1, \dots, -\infty \leq \xi_n \leq x_n, \quad (1.1.16)$$

і функція розподілу розглядається як функція змінних $x_1, \dots, x_n$, що відповідають верхнім границям нерівностей (1.1.16), тобто

$$F(x_1, \dots, x_n) \equiv P(-\infty \leq \xi_1 \leq x_1, \dots, -\infty \leq \xi_n \leq x_n) = \lim_{M \rightarrow \infty} \frac{m(x_1, \dots, x_n)}{M} \quad (1.1.17)$$

з $m(x_1, \dots, x_n)$ як кількість знаходжень реалізації в даній області. Із визначення (1.1.17) можна довести, наприклад, такі властивості функції розподілу багатовимірної випадкової величини:

1) Функція розподілу є неспадною функцією за кожним з аргументів, тобто $F(x_1, \dots, x_n)$ монотонно зростає зі зростанням будь-якого зі своїх аргументів.

2) Функція розподілу прямує до одиниці, коли всі її аргументи прямують до нескінченності

$$F(+\infty, +\infty, \dots, +\infty) = 1. \quad (1.1.18)$$

3) Функція F прямує до нуля, якщо хоча б один із її аргументів прямує до $-\infty$

$$\lim_{x_k \rightarrow -\infty} F(x_1, x_2, \dots, x_n) = 0 \quad (1 \leq k \leq n) \quad (1.1.19)$$

за будь-яких значень x_i ($i \neq k$).

4) Функція

$$F(x_1, \dots, x_{k-1}, \infty, x_{k+1}, \dots, x_n) \equiv \lim_{x_k \rightarrow +\infty} F(x_1, \dots, x_{k-1}, x_k, x_{k+1}, \dots, x_n) \quad (1.1.20)$$

є функцією розподілу $(n-1)$ вимірної випадкової величини

$$(\xi_1, \xi_{k-1}, \xi_{k+1}, \dots, \xi_n).$$

Щільністю розподілу випадкового багатовимірного вектора $\vec{\xi} = (\xi_1, \dots, \xi_n)$, за аналогією з одновимірним випадком, називають функцію $p(x_1, \dots, x_n)$, що визначає відношення ймовірності появи реалізації вектора $\vec{\xi}$ в багатовимірній області значень $x_i \leq \xi_i \leq x_i + \Delta x_i$, ($i = 1, \dots, n$) до нескінченно малого значення об'єму даної області $V = \Delta x_1, \dots, \Delta x_n$ з $\Delta x_i \rightarrow 0$, тобто

$$\begin{aligned} p(x_1, \dots, x_n) &= \lim_{\substack{\Delta x_1 \rightarrow 0 \\ \Delta x_n \rightarrow 0}} \frac{1}{\Delta x_1 \Delta x_2 \dots \Delta x_n} P\left(\begin{matrix} x_1 \leq \xi_1 \leq x_1 + \Delta x_1 \\ \dots \\ x_n \leq \xi_n \leq x_n + \Delta x_n \end{matrix}\right) = \\ &= \lim_{\substack{M \rightarrow 0 \\ \Delta x_1, \dots, \Delta x_n \rightarrow 0}} \frac{\Delta m(x_1, \dots, x_n)}{M \Delta x_1 \dots \Delta x_n}, \end{aligned} \quad (1.1.21)$$

де $\Delta m(x_1, \dots, x_n)$ – кількість знаходжень реалізації в даній області.

З визначення (1.1.21) випливають співвідношення:

$$\begin{aligned}
 p(x_1, \dots, x_n) &\geq 0; \\
 P\left(x_1 \leq \xi_1 \leq x_1 + dx_1 \atop x_n \leq \xi_n \leq x_n + dx_n\right) &= p(x_1, x_2, \dots, x_n) dx_1 \dots dx_n, \\
 P(\bar{\xi} \in D) &= \int_D p(x_1, \dots, x_n) dx_1 \dots dx_n, \\
 F(x_1, \dots, x_n) &= \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} p(x'_1, \dots, x'_n) dx'_1, \dots, dx'_n, \\
 \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} p(x_1, \dots, x_n) dx_1, \dots, dx_n &= 1.
 \end{aligned} \tag{1.1.22}$$

Фактично, визначення (1.1.21) означає, що щільність розподілу $p(x_1, \dots, x_n)$ є частинною похідною від функції розподілу

$$p(x_1, \dots, x_n) = \frac{\partial^n F(x'_1, \dots, x'_n)}{\partial x'_1, \dots, x'_n} \bigg|_{\{x'_j\}=\{x_j\}} \tag{1.1.23}$$

Для дискретних випадкових величин вираз (1.1.23) треба розуміти в узагальненому сенсі аналогічно виразам (1.1.11)–(1.1.13).

У багатовимірному випадку, крім сумісної щільності $p(x_1, \dots, x_n)$, можна визначити також умовну щільність

$$p(x_k / x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n),$$

як щільність розподілу k -ї випадкової величини в багатовимірному просторі за умови, що всі інші величини мають фіксовані задані значення. Ця величина є невід'ємною і нормованою на одиницю

$$\begin{aligned}
 p(x_k / x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n) &\geq 0, \\
 \int_{-\infty}^{\infty} p(x_k / x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n) dx_k &= 1.
 \end{aligned} \tag{1.1.24}$$

За її допомогою можна знайти відповідні умовну ймовірність та умовну функцію розподілу

$$dP(x_k / x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n) = p(x_k / x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n) dx_k;$$

$$P(x_k' \leq \xi_k \leq x_k'' / x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n) = \int_{x_k'}^{x_k''} p(\bar{x}_k / x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n) d\bar{x}_k; \quad (1.1.25)$$

$$F(x_k / x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n) = \int_{-\infty}^{x_k} p\left(\frac{x_k'}{x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n}\right) dx_k'.$$

У загальному випадку умовна щільність розподілу k -ї випадкової величини залежить від значень усіх інших випадкових величин, якщо існують зв'язки між нею та іншими величинами. Якщо ж зміна значень $x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n$ не впливає на умовну ймовірність розподілу, то випадкова величина ξ_k називається незалежною.

Аналогічним чином можна визначити умовні щільності розподілів та умовні ймовірності при довільних часткових сукупностях випадкових величин, які формують дану систему випадкових величин.

Щільність розподілу якоїсь окремої випадкової величини, яку знайдено за довільних можливих значень інших випадкових величин системи, називається безумовною щільністю $p(x_k)$ розподілу даної випадкової величини. Вона пов'язана із щільністю розподілу таким інтегральним співвідношенням:

$$p(x_k) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} p(x_1, \dots, x_n) dx_1 \dots dx_{k-1} dx_{k+1} \dots dx_n. \quad (1.1.26)$$

Функція $p(x_k)$ задовольняє всім співвідношенням (1.1.8), що виконуються для щільності розподілу одновимірної випадкової величини.

Безумовну щільність розподілу можна також визначити і для будь-якої часткової сукупності випадкових величин системи згідно зі співвідношенням

$$p(x_1, \dots, x_k) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} p(x_1, \dots, x_n) dx_{k+1} \dots dx_n. \quad (1.1.27)$$

Відповідно до теореми множення ймовірностей сумісну щільність системи розподілу випадкових величин, пронумерованих у довільному порядку, можна подати у вигляді добутку умовних щільностей розподілу:

$$p(x_1, \dots, x_n) = p(x_1) p(x_2 / x_1) p(x_3 / x_1, x_2) \dots p(x_n / x_1, \dots, x_{n-1}). \quad (1.1.28)$$

Деякі висновки із цієї теореми:

1. Для незалежних випадкових величин умовні й безумовні щільності розподілів збігаються:

$$p(x_k / x_1, x_{k-1}, x_{k+1}) = p(x_k), \quad (1.1.29)$$

і тому сумісна щільність розподілу є добутком щільностей розподілів окремих випадкових величин

$$p(x_1, \dots, x_n) = \prod_{k=1}^n p(x_k). \quad (1.1.30)$$

Згідно з формулами (1.1.22), (1.1.30) маємо таке співвідношення між багатовимірною функцією розподілу та одновимірними функціями розподілів у випадку незалежних випадкових величин:

$$F(x_1, \dots, x_n) = \prod_{k=1}^N F(x_k). \quad (1.1.31)$$

2. Умовну та сумісну (спільну, об'єднану) щільності розподілів пов'язано співвідношенням

$$p(x_k / x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n) = \frac{p(x_1, \dots, x_n)}{\int_{-\infty}^{\infty} dx_k p(x_1, \dots, x_n)}. \quad (1.1.32)$$

3. Виконується теорема Байєса, яка зв'язує умовні й часткові сумісні щільності

$$\begin{aligned} P\left(\frac{x_1, \dots, x_e}{x_{e+1}, \dots, x_n}\right) &= \frac{P(x_1, \dots, x_e)}{P(x_{e+1}, \dots, x_n)} P\left(\frac{x_{e+1}, \dots, x_n}{x_1, \dots, x_e}\right) = \\ &= \frac{P\left(\frac{x_{e+1}, \dots, x_n}{x_1, \dots, x_e}\right) P(x_1, \dots, x_e)}{\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} P\left(\frac{x_{e+1}, \dots, x_n}{x_1, \dots, x_e}\right) P(x_1, \dots, x_e) dx_1, \dots, dx_e}. \end{aligned} \quad (1.1.33)$$

Нумерація випадкових величин у виразах (1.1.31)–(1.1.33) є довільною.

Розглянемо деякі властивості багатовимірних функцій розподілу і щільностей на прикладі двовимірної випадкової величини (ξ, η) з функцією розподілу $F(x, y)$. Властивість (1.1.20) дає змогу знайти функції розподілу окремих компонент (ξ, η) за допомогою формул

$$F_{\xi}(x) = \lim_{y \rightarrow \infty} F(x, y), \quad F_{\eta}(y) = \lim_{x \rightarrow \infty} F(x, y). \quad (1.1.34)$$

Двовимірною функцією розподілу пов'язана зі щільністю розподілу співвідношенням

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y p(x', y') dx' dy'. \quad (1.1.35)$$

Щільності розподілу $p_{\xi}(x)$, $p_{\eta}(y)$ компонентів двовимірної випадкової величини (ξ, η) легко знайти з виразів для щільності її двовимірного розподілу. Дійсно, згідно з формулою (1.1.34) функція розподілу випадкової величини ξ дорівнює

$$F_{\xi}(x) = \lim_{y \rightarrow +\infty} F(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} p(x', y') dx' dy'.$$

Звідки маємо

$$p_{\xi}(x) = \frac{dF_{\xi}(x)}{dx} = \int_{-\infty}^{+\infty} p(x, y') dy'. \quad (1.1.36)$$

Аналогічно

$$p_{\eta}(y) = \int_{-\infty}^{+\infty} p(x', y) dx'. \quad (1.1.37)$$

Часто постає питання про знаходження функції розподілу $F(z)$ і щільності розподілу $p(z)$ випадкової величини $\delta = \xi + \eta$, яка є сумою величин ξ , η із сумісними функцією розподілу $F(x, y)$ і щільністю $p(x, y)$. Знайдемо ці величини. Згідно з означенням функції розподілу маємо:

$$F(z) \equiv F_{\delta}(z) = P\{\delta < z\} = P(\xi + \eta < z). \quad (1.1.38)$$

Умова $\xi + \eta < z$ еквівалентна умові знаходження точки (ξ, η) у півплощині G площини xOy , де G визначається нерівністю $x + y < z$ (рис. 1.1).

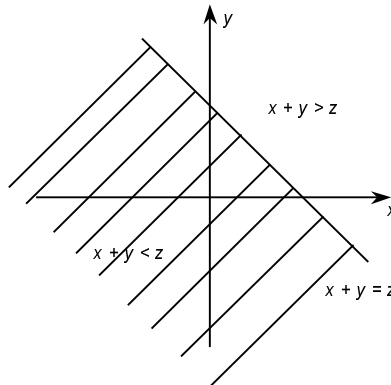


Рис. 1.1. Область визначення випадкової величини $\delta = \xi + \eta$

Тому для функції розподілу $F_{\delta}(z)$ маємо

$$F_{\delta}(z) = P\{(\xi, \eta) \in G\} = \int_G \int p(x, y) dx dy = \int_{-\infty}^{+\infty} dx \int_{-\infty}^{z-x} p(x, y) dy. \quad (1.1.39)$$

Зробимо у внутрішньому інтегралі заміну змінних $y = u - x$, після чого змінімо порядок інтегрування:

$$F_{\delta}(z) = \int_{-\infty}^{+\infty} dx \int_{-\infty}^z p(x, u-x) du = \int_{-\infty}^z du \int_{-\infty}^{\infty} p(x, u-x) dx. \quad (1.1.40)$$

Звідси отримуємо, що щільність розподілу $p(z)$ одновимірної випадкової величини $\delta = \xi + \eta$ визначається виразом

$$p(z) = \int_{-\infty}^{\infty} p(x, z-x) dx. \quad (1.1.41)$$

Нехай, зокрема, випадкові величини ξ і η незалежні. Тоді за формулою (1.1.30) $p(x, y) = p_{\xi}(x) p_{\eta}(y)$ і співвідношення (1.1.41) набуває вигляду

$$p(z) = \int_{-\infty}^{\infty} p_{\xi}(x) p_{\eta}(z-x) dx. \quad (1.1.42)$$

Цей вираз можна також записати як

$$p(z) = \int_{-\infty}^{\infty} p_{\xi}(z-y) p_{\eta}(y) dy. \quad (1.1.43)$$

Маємо відзначити, що функція $p(z)$, утворена з функцій $p_{\xi}(x)$ і $p_{\eta}(y)$ за формулами типу (1.1.42), (1.1.43), називається згорткою, або композицією цих функцій.

Вирази, подібні до (1.1.42) і (1.1.43), також легко вивести і для закону розподілу суми незалежних дискретних величин.

1.2. Числові характеристики випадкових величин

Найповнішу інформацію про характер і поведінку випадкової величини дає її функція розподілу або закон розподілу, записаний у довільній формі, наприклад, у формі щільності розподілу випадкової величини. Але при розв'язуванні багатьох задач теорії ймовірностей достатньо знати лише декілька чисел, які певною мірою характеризують випадкову величину (хоч і не так повно, як функція розподілу). Такі числа називаються кількісними характеристиками випадкової величини.

Найважливішими з кількісних характеристик є моменти розподілу випадкових величин. Вони є не випадковими величинами і визначаються інтегралами вигляду

$$\bar{\Phi} = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \Phi(x_1, \dots, x_n) p(x_1, \dots, x_n) dx_1, \dots, dx_n, \quad (1.2.1)$$

де $\Phi(x_1, \dots, x_n)$ – деякі функції, що залежать від типу моменту. Іноді такі величини також називають теоретичними моментами. Для дискретних випадкових величин інтеграл (1.2.1) розглядається в узагальненому значенні зі щільністю розподілу у вигляді суми δ -функцій вигляду (1.1.11).

Найважливішими є моменти першого та другого порядків. Початковим моментом першого порядку випадкової величини ξ_k називають таку величину:

$$\bar{x}_k \equiv M[x_k] = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} x_k p(x_1, \dots, x_n) dx_1 \dots dx_n = \int_{-\infty}^{\infty} x_k p(x_k) dx_k. \quad (1.2.2)$$

Початковий момент також називається математичним сподіванням, або теоретичним середнім значенням випадкової величини. Він характеризує положення центра розкиду (розсіювання) реалізацій у просторі $x_1, \dots, x_n$ сукупності випадкових величин.

Для моментів двох незалежних випадкових величин x_1, x_2 і невідповідних величин a_1, a_2 мають місце співвідношення

$$\begin{aligned} M[a_i] &= a_i; \quad M[a_i x_j] = a_i M[x_j]; \quad i, j = 1, 2; \\ M[a_1 x_1 + a_2 x_2] &= a_1 M[x_1] + a_2 M[x_2]; \\ M[x_1 x_2] &= M[x_1] M[x_2], \end{aligned} \quad (1.2.3)$$

що впливають із законів додавання та множення щільностей розподілу ймовірностей незалежних випадкових величин і означення математичного сподівання. Аналогічні співвідношення є справедливими також і для n незалежних величин. Зокрема, для n незалежних величин з однаковими середніми значеннями маємо

$$M\left[\sum_{i=1}^n x_i\right] = nM[x_i]. \quad (1.2.4)$$

Якщо при будь-яких значеннях $x_1, \dots, x_n$ щільність розподілу задовольняє умові

$$p(x_1^0 + x_1, \dots, x_n^0 + x_n) = p(x_1^0 - x_1, \dots, x_n^0 - x_n), \quad (1.2.5)$$

де $\{x_j^0\}$ – координати максимуму функції $p(\{x_j\})$, то її називають с и м е т р и ч н о ю . У випадку симетричних щільностей виконується співвідношення

$$M[x_j] = \bar{x}_j = x_j^{(0)}, j = 1, \dots, n. \quad (1.2.6)$$

Якщо всі середні значення $\bar{x}_j$ дорівнюють нулю ($\bar{x}_j = 0$), то система відповідних випадкових величин $\xi_1, \dots, \xi_n$ називається центрованою.

Центральні незмішані моменти другого порядку, або дисперсії випадкових величин, вводяться для характеристики відхилень випадкових величин від їх середніх значень. Вони визначаються як математичне сподівання квадрата відхилення цієї величини від її математичного сподівання

$$D[x_k] \equiv D_k \equiv R_{kk} = M\left[(x_k - \bar{x}_k)^2\right] \equiv \int_{-\infty}^{+\infty} (x_k - \bar{x}_k)^2 p(x_1 \dots x_n) dx_1 \dots dx_n. \quad (1.2.7)$$

Замість дисперсії часто розглядають квадратний корінь:

$$\sigma_k \equiv \sqrt{D_k}. \quad (1.2.8)$$

Цю величину називають середньоквадратичним відхиленням (стандартним відхиленням, або стандартом). Як дисперсія, так і середньоквадратичне відхилення, є мірою розкиду значень випадкової величини навколо її математичного сподівання. Очевидно, що ці величини завжди невід'ємні.

З визначення математичного сподівання випливає, що дисперсія має такі властивості:

1) При обчисленні дисперсії добутку не випадкової величини a на випадкову x не випадковий множник виноситься у квадраті за знак дисперсії:

$$D[ax] = M\left[(ax - M[ax])^2\right] = \overline{(ax - \overline{ax})^2} = a^2 \overline{(x - \bar{x})^2} = a^2 D[x]. \quad (1.2.9)$$

2) Дисперсія не випадкової величини дорівнює нулю:

$$D[a] = 0, \quad (1.2.10)$$

оскільки $\bar{a} = a$.

3) Дисперсія випадкової величини дорівнює різниці між математичним сподіванням квадрата цієї величини і квадратом її математичного сподівання:

$$D[x] = M[x^2] - (M[x])^2 \equiv \overline{x^2} - (\bar{x})^2. \quad (1.2.11)$$

Дійсно

$$D[x] = \overline{(x - \bar{x})^2} = M(x^2 - 2x\bar{x} + \bar{x}^2) = \overline{x^2} - 2(\bar{x})^2 + \bar{x}^2 \equiv \overline{x^2} - (\bar{x})^2.$$

4) Дисперсія суми двох незалежних випадкових величин дорівнює сумі їх дисперсій (закон додавання дисперсій):

$$D[x + y] = D[x] + D[y]. \quad (1.2.12)$$

Дійсно, оскільки для незалежних величин $M[xy] = M[x] \cdot M[y]$ (див.(1.2.3)), то

$$\begin{aligned} D[x + y] &= \overline{(x + y)^2} - (\bar{x} + \bar{y})^2 = M[x^2 + 2xy + y^2] - (\bar{x}^2 + 2\bar{x}\bar{y} + \bar{y}^2) = \\ &= M[x^2] + 2\bar{x}\bar{y} + M[y^2] - \bar{x}^2 - 2\bar{x}\bar{y} - \bar{y}^2 = \overline{x^2} - \bar{x}^2 + \overline{y^2} - \bar{y}^2 \equiv D[x] + D[y]. \end{aligned}$$

5) Якщо випадкові величини $\{x_j\}, j = 1, \dots, n$ попарно незалежні, то

$$D\left[\sum_{j=1}^n x_j\right] = \sum_{j=1}^n D[x_j]. \quad (1.2.13)$$

Доведення цього твердження є прямим узагальненням доведення з п. 4).
Оскільки

$$M[x_i x_j] = M[x_i] M[x_j],$$

то

$$\begin{aligned} D\left[\sum_{j=1}^n x_j\right] &= \overline{\left(\sum_j x_j\right)^2} - \left(\sum_j \bar{x}_j\right)^2 = M\left(\sum_j x_j^2 + 2\sum_{j < k} x_j x_k\right) - \left(\sum_j \bar{x}_j\right)^2 = \\ &= \sum_j \overline{x_j^2} + 2\sum_{j < k} \overline{x_j x_k} - \sum_j (\bar{x})^2 - 2\sum_{j < k} \bar{x}_j \bar{x}_k = \sum_j \overline{x_j^2} - \sum_j (\bar{x}_j)^2 = \\ &= \sum_j \left(\overline{x_j^2} - (\bar{x}_j)^2\right) \equiv \sum_{j=1}^n D[x_j]. \end{aligned}$$

Для дослідження залежності між двома випадковими величинами використовують центральні змішані моменти другого порядку, які визначаються співвідношеннями

$$\begin{aligned} R_{x_k x_{k'}} &\equiv R_{kk'} = \text{cov}(x_k, x_{k'}) = M[(x - \bar{x}_k)(x - \bar{x}_{k'})] = \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x_k - \bar{x}_k)(x_{k'} - \bar{x}_{k'}) p(x_1, \dots, x_n) dx_1, \dots, dx_n, \end{aligned} \quad (1.2.14)$$

$$k, k' = 1, \dots, n, k \neq k'.$$

Ці величини звичайно називають коефіцієнтами коваріації, а нормовані коефіцієнти коваріацій називають коефіцієнтами кореляцій:

$$\rho_{kk'} \equiv \rho_{x_k x_{k'}} = \frac{R_{kk'}}{\sqrt{R_{kk} R_{k'k'}}} \equiv \frac{R_{x_k x_{k'}}}{\sqrt{D_{x_k} D_{x_{k'}}}}. \quad (1.2.15)$$

Для коефіцієнта кореляції будь-яких випадкових величин ξ , η виконується нерівність

$$-1 \leq \rho_{\xi, \eta} \leq 1. \quad (1.2.16)$$

Коефіцієнт кореляції між випадковими величинами ξ і η дорівнює +1 або -1 тоді й тільки тоді, коли ξ і η зв'язані лінійною залежністю

$$\eta = a\xi + c. \quad (1.2.17)$$

При цьому знак коефіцієнта кореляції визначається знаком множника a

$$\rho_{\xi, \eta} = \text{sign}(a) = \begin{cases} +1, & a > 0 \\ -1, & a < 0 \end{cases}. \quad (1.2.18)$$

Якщо випадкові величини ξ і η є незалежними, то їх коефіцієнт кореляції дорівнює нулю:

$$\rho_{\xi, \eta} = 0, \quad (1.2.19)$$

оскільки для незалежних ξ та η незалежними є і величини $(\xi - \bar{\xi})$ і $(\eta - \bar{\eta})$,

$$\rho_{\xi, \eta} \sim (\xi - \bar{\xi})(\eta - \bar{\eta}) = M[\xi - \bar{\xi}] M[\eta - \bar{\eta}] = 0.$$

Випадкові величини ξ і η , між якими коефіцієнт кореляції дорівнює нулю, $\rho_{\xi, \eta} = 0$, називають некорельованими. Таким чином, незалежні величини завжди некорельовані, але з некорельованості двох випадкових величин, взагалі кажучи, не випливає їх незалежність. Згідно з (1.2.17) та (1.2.18) між некорельованими величинами гарантована лише відсутність лінійних зв'язків. Але нелінійні зв'язки можуть існувати. Наведемо приклад. Нехай ξ є рівномірно розподілена випадкова величина на відрізку $[-1, 1]$, а η – випадкова величина вигляду $\eta = \xi^2$. Знайдемо коефіцієнти кореляції між ними. Відповідно до визначень (1.1.8), (1.2.2) маємо для щільності розподілу (з умови нормування) та середнього значення величини ξ

$$p(x) = \frac{1}{2}, \quad x \in [-1, 1]; \quad \bar{\xi} = \frac{1}{2} \int_{-1}^1 x dx = 0.$$

Звідки знаходимо

$$R_{\xi, \eta} = M[(\xi - \bar{\xi})(\eta - \bar{\eta})] = M[\xi(\eta - \bar{\eta})] = M[\xi\eta] = \frac{1}{2} \int_{-1}^1 x x^2 dx = 0. \quad (1.2.20)$$

Отже $\rho(\xi, \eta) = 0$, хоча випадкові величини ξ і η зв'язані функціональною залежністю.

Але є також ряд випадків, коли з некорельованості величин випливає і їх незалежність, тобто буде виконуватись співвідношення $\rho(\xi, \eta) = \rho(\xi)\rho(\eta)$ для щільностей ρ . Так буде, наприклад, у випадку нормально розподілених величин, для яких їх некорельованість еквівалентна незалежності.

Матриця других центральних моментів розподілу, яка складена з дисперсій (головна діагональ) та коефіцієнтів коваріації, називається матрицею коваріацій:

$$\hat{D} = \|\hat{D}_{ij}\| = \begin{bmatrix} R_{11}, & \dots, & R_{1n} \\ \cdot & \dots & \cdot \\ \cdot & \dots & \cdot \\ \cdot & \dots & \cdot \\ R_{n1}, & \dots, & R_{nn} \end{bmatrix}. \quad (1.2.21)$$

Часто користуються кореляційною матрицею, що складається з коефіцієнтів кореляції (1.2.15), і тому має вигляд матриці з одиничними діагональними елементами

$$\hat{\rho} \equiv \|\rho_{ij}\| = \begin{bmatrix} 1 & , & \rho_{12}, & \dots, & \rho_{1n} \\ \rho_{21}, & 1 & , & \dots, & \rho_{2n} \\ \cdot & \cdot & & \dots, & \cdot \\ \cdot & \cdot & & \dots, & \cdot \\ \rho_{n1}, & \rho_{n2}, & \dots, & & 1 \end{bmatrix}. \quad (1.2.22)$$

Якщо всі випадкові величини $\{\xi_j\}$, $j = 1, \dots, n$ є незалежними, то матриця коваріацій – діагональна, а кореляційна матриця – одинична. Як і у випадку двох величин, обернене твердження гарантує лише лінійну незалежність випадкових величин.

Легко показати, що дисперсія випадкової суми $\sum a_j x_j$, де a_j – детерміновані величини, визначається елементами коваріаційної матриці:

$$D\left[\sum_{j=1}^n a_j x_j\right] = \sum_{i=1}^n \sum_{j=1}^n \hat{D}_{ij} a_i a_j. \quad (1.2.23)$$

Додавання невинескових доданків до випадкових величин x_j не змінює матрицю кореляцій цих величин.

Крім початкових моментів першого порядку (математичне сподівання) і другого порядку (дисперсія, коефіцієнти коваріації), у теорії ймовірностей та

її застосуваннях використовують також інші кількісні характеристики. Розглянемо деякі з них на прикладі одновимірних випадкових величин. Моментом k -го порядку випадкової величини ξ відносно точки a називається число (якщо воно існує), яке визначається виразом

$$v_k(a) = M[(\xi - a)^k], \quad (1.2.24)$$

де $k \geq 0$ – ціле, а a – дійсне число.

Якщо $a = 0$, то, як уже відзначалося раніш (див. (1.2.2)), моменти називаються початковими. Звичайно початковий момент k -го порядку позначають як v_k . Очевидно, що

$$v_1 \equiv M[\xi] \equiv \bar{x}, \quad v_k = \overline{x^k}. \quad (1.2.25)$$

Центральними моментами називаються моменти (1.2.24), що обчислюються відносно середнього значення, тобто – точки $a = M[\xi]$. Якщо позначити центральний момент k -го порядку як μ_k , то

$$\mu_k = M[(\xi - v_1)^k]. \quad (1.2.26)$$

Легко побачити, що $\mu_1 = 0$; $\mu_2 = D[\xi]$.

Між центральними і початковими моментами існує простий зв'язок. Дійсно, розкривши вираз у правій частині формули (1.2.25), дістаємо, наприклад, для перших чотирьох моментів:

$$\begin{aligned} \mu_1 &= 0, \quad \mu_2 = v_2 - v_1^2, \\ \mu_3 &= v_3 - 3v_2v_1 + 2v_1^3, \\ \mu_4 &= v_4 - 4v_3v_1 + 6v_2v_1^2 - 2v_1^4. \end{aligned} \quad (1.2.27)$$

Ці перші моменти дуже широко використовуються у статистиці.

У випадку симетричного розподілу кожний центральний момент непарного порядку (у випадку його існування) дорівнює нулю:

$$\mu_{2k+1} = M[(\xi - M[\xi])^{2k+1}] = 0, \quad k = 0, 1, 2 \dots \quad (1.2.28)$$

Тому кожний відмінний від нуля центральний момент непарного порядку характеризує ступінь асиметрії розподілу. Найпростішим із таких моментів є μ_3 , але для зручності замість нього беруть безрозмірну величину

$$\gamma = \frac{\mu_3}{(D\xi)^{\frac{3}{2}}} \equiv \frac{\mu_3}{\sigma_\xi^3}. \quad (1.2.29)$$

Цю величину називають коефіцієнтом асиметрії.

Для вивчення властивостей випадкових величин А.М.Ляпунов у 1900 р. ввів характеристичні функції. Наприклад, характеристичною функцією одновимірної випадкової величини x називають функцію

$$C_x(t) \equiv M[e^{itx}] = \int_{-\infty}^{\infty} e^{itx} p(x) dx, \quad (1.2.30)$$

яка є перетворенням Фур'є для щільності розподілу. Функція також є математичним сподіванням функції $\exp(itx)$, де t – дійсне число. З властивостей інтегралу Фур'є (1.2.30) випливає формула для щільності розподілу

$$p(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} C_x(t) e^{-itx} dt, \quad (1.2.31)$$

яка має вигляд оберненого перетворення Фермі від характеристичної функції.

Значення функції $C_x(t)$ дорівнює одиниці в нулі ($t = 0$), а її похідні $C_x^{(r)}(t=0)$ за x з точністю до коефіцієнта (i^r) збігаються з початковими моментами

$$\begin{aligned} C_x(0) &= \int_{-\infty}^{\infty} p(x) dx = 1, \quad C_x^{(1)}(0) = i \int x p(x) dx = i\bar{x} = i v_1, \\ C_x^{(r)}(0) &= i^r \int x^r p(x) dx = i^r \bar{x}^r \equiv i^r v_r. \end{aligned} \quad (1.2.32)$$

Якщо розкласти функцію $C_x(t)$ у ряд Маклорена, то маємо

$$C_x(t) = C_x(0) + t C_x^{(1)}(0) + \dots + \frac{t^r}{r!} C_x^{(r)}(0) = 1 + i t \bar{x} + \dots + \frac{(it)^r}{r!} \bar{x}^r + \dots \quad (1.2.33)$$

Тому, коли цей ряд збігається, то за значеннями початкових моментів можна знайти характеристичну функцію, і, отже, згідно з (1.2.31), щільність розподілу $p(x)$.

Характеристична функція широко застосовується в задачах із сумами незалежних випадкових величин. У цьому випадку $C_x(t)$ для суми незалежних величин $\xi = \sum_j \xi_j$ дорівнює добутку відповідних характеристичних функцій:

$$C_\xi(t) \equiv \int e^{it \sum_j x_j} p(\{x_j\}) dx_1 \dots dx_n = \prod_{j=1}^n C_{\xi_j}(t). \quad (1.2.34)$$

Зазначимо, що в застосуваннях теорії ймовірностей часто використовується корінь x_p рівняння

$$F(x_p) = p, \quad F(x_p) \equiv P\{X < x_p\}, \quad (1.2.35)$$

тобто таке число, що ймовірність появи значень випадкової величини X менших за x_p дорівнює p . Розв'язок x_p рівняння (1.2.35) називається квантилем рівня (порядку) p . Квантилі також називаються за значенням ймовірності p , коли остання записана у відсотках, так $x_p \in p/100$ -м квантилем наприклад, $x_{0,4} - 40$ -й квантиль. Якщо відомо два квантиля x_p і x_q , то ймовірність знаходження випадкової величини в інтервалі $x_q \leq X \leq x_p$ визначається за формулою

$$P\{x_q \leq X \leq x_p\} \equiv F(x_p) - F(x_q) = p - q, \quad (1.2.36)$$

на якій і базується використання квантилей.

Ймовірність q появи значень випадкової величини X більших за деяке значення x_0 визначається квантилем x_{1-q} і $x_0 \equiv x_{1-q}$,

$$P\{X \geq x_{1-q}\} = q. \quad (1.2.37)$$

Квантиль порядку $1/2$ називається медіаною розподілу. Випадкова величина з однаковою ймовірністю може приймати значення як більші, так і менші за медіану. Квантилі x_p і x_{1-p} називаються симетричними.

1.3. Наближене обчислення математичного сподівання та дисперсії для функції випадкових величин

У багатьох випадках буває необхідним знайти математичне сподівання $M[U]$ і дисперсію $D[U]$ величини U , що є функцією $U = y(\xi_1, \dots, \xi_n)$ випадкових величин $\{\xi_i\}$, кожна з яких має відоме математичне сподівання μ_i і дисперсію σ_i^2 , величини ξ_i, ξ_j також можуть бути корельованими з коефіцієнтами коваріації R_{ij} :

$$\mu_i = M[\xi_i]; \sigma_i^2 = D[\xi_i] = M[(\xi_i - \mu_i)^2], \quad R_{ij} \equiv \text{cov}(\xi_i, \xi_j) = M[(\xi_i - \mu_i)(\xi_j - \mu_j)]. \quad (1.3.1)$$

Така ситуація, наприклад, виникає за так званих непрямих вимірів величин. Точне обчислення середнього значення $M[U]$ і дисперсії $D[U]$ неможливе, оскільки закон розподілу функції U може бути дуже складним, тому зазвичай застосовують наближені методи їх обчислення. Для визначення $M[U]$ і $D[U]$ у випадку малих очікуваних середніх відхилень значень випадкових величин від їх математичних сподівань функцію $U = y(\{\xi_j\})$ можна лінеаризувати та обмежитися лише лінійними за відхиленням $(\xi_i - \mu_i)$ членами її розкладу в ряд Тейлора:

$$U = y(\{\xi_j\}) \cong U_0 \equiv y(\{\mu_i\}) + \sum_{i=1}^n \frac{\partial y}{\partial \mu_i} (\xi_i - \mu_i). \quad (1.3.2)$$

Після цього значення $M[U]$ і $D[U]$ можна ототожнити з відповідними величинами для лінійної функції U_0 :

$$\begin{aligned} \mu_u &\equiv M[U] = M[U_0] = \mu_{u_0}, \\ \sigma_u^2 &\equiv D[U] \equiv D[U_0] \equiv M[(U - \mu_u)^2]. \end{aligned} \quad (1.3.3)$$

Іноді для такої процедури вживається термін перенесення похибок.

У загальному випадку функції $U = y(\{\xi_i\})$ багатьох змінних, використовуючи визначення (1.3.1), розклад (1.3.2) та загальні теореми (див. 1.2) для обчислень математичного сподівання і дисперсії, маємо такі вирази як їх оцінки:

$$\begin{aligned} \mu_u &\cong \mu_{u_0} = y(\{\mu_i\}), \\ \sigma_u^2 &\cong \sigma_{u_0}^2 = \left(\frac{\partial y}{\partial \mu_1} \right)^2 \sigma_1^2 + \left(\frac{\partial y}{\partial \mu_2} \right)^2 \sigma_2^2 + \dots + \left(\frac{\partial y}{\partial \mu_n} \right)^2 \sigma_n^2 + \\ &+ 2 \left(\frac{\partial y}{\partial \mu_1} \right) \left(\frac{\partial y}{\partial \mu_2} \right) R_{1,2} + \dots + 2 \left(\frac{\partial y}{\partial \mu_{n-1}} \right) \left(\frac{\partial y}{\partial \mu_n} \right) R_{n-1,n}. \end{aligned} \quad (1.3.4)$$

Часто співвідношення для дисперсії записують у матричному вигляді за допомогою матриці коваріацій (1.2.21)

$$\hat{D} \equiv \left\| \hat{D}(\{\xi_j\}) \right\| = \left\| \begin{array}{cccc} \sigma_{\xi_1}^2, & R_{\xi_1, \xi_2}, & R_{\xi_1, \xi_3}, & \dots, R_{\xi_1, \xi_n} \\ R_{\xi_2, \xi_1}, & \sigma_{\xi_2}^2, & R_{\xi_2, \xi_3}, & \dots, R_{\xi_2, \xi_n} \\ \cdot & \cdot & \cdot & \cdot \\ R_{\xi_n, \xi_1}, & \dots & \dots & \sigma_{\xi_n}^2 \end{array} \right\|, \quad (1.3.5)$$

1.3. НАБЛИЖЕНЕ ОБЧИСЛЕННЯ МАТЕМАТИЧНОГО СПОДІВАННЯ ТА ДИСПЕРСІЇ ДЛЯ ФУНКЦІЇ ВИПАДКОВИХ ВЕЛИЧИН

що описує властивості других центральних моментів величин $\{\xi_j\}$. Неважко бачити, що вираз для дисперсії має вигляд

$$D_u \equiv \sigma^2 \cong D_{u_0} \equiv \sigma_{u_0}^2 \equiv \hat{F} \hat{D} \hat{F}^T, \quad (1.3.6)$$

де $\hat{F}$ – матриця-рядок перших похідних з елементами $\frac{\partial y}{\partial \mu_i}$:

$$\hat{F} = \left(\frac{\partial y}{\partial \mu_1}, \dots, \frac{\partial y}{\partial \mu_n} \right), \quad (1.3.7)$$

а F^T – відповідна їй транспонована матриця (матриця-стовпчик):

$$\hat{F}^T = \begin{pmatrix} \frac{\partial y}{\partial \mu_1} \\ \dots \\ \frac{\partial y}{\partial \mu_n} \end{pmatrix}. \quad (1.3.8)$$

Якщо випадкові величини незалежні або некорельовані, то $R_{\xi_i, \xi_j} = 0$, і маємо таке співвідношення для середньоквадратичного відхилення незалежних величин:

$$\sigma_u \cong \sigma_{u_0} = \left[\sum_{j=1}^n \left(\frac{\partial y}{\partial \mu_j} \sigma_j \right)^2 \right]^{1/2}. \quad (1.3.9)$$

Зазначимо, що значення середньоквадратичних відхилень ніколи не перевищує суми абсолютних відхилень:

$$\sigma_{u_0} \leq \sum_{j=1}^n \left| \frac{\partial y}{\partial \mu_j} \right| \sigma_{\xi_j}.$$

Дослідимо, у якому випадку можна використовувати метод перенесення похибок. Спочатку розглянемо функцію однієї змінної: $U = y(\xi)$. Припустимо, що останню можна розкласти в ряд Тейлора значення величини ξ , що відповідає її математичному сподіванню μ , тобто біля точки $\xi = \mu$,

$$U = y(\xi) = y(\mu) + \frac{dy}{d\mu}(\xi - \mu) + \frac{1}{2} \frac{d^2 y}{d\mu^2}(\xi - \mu)^2 + \frac{1}{6} \frac{d^3 y}{d\mu^3}(\xi - \mu)^3 + \dots, \quad (1.3.10)$$

де $\frac{d^j y}{d\mu^j}$ – похідна j -го порядку від функції $y(\xi)$ у точці $\xi = \mu$.

Якби різниця $\Delta\xi = (\xi - \mu)$ була заданим числом, то питання про збіжність ряду (1.3.10) могло бути вирішеним звичайним для математичного аналізу чином, а саме, треба було б дослідити, скільки членів ряду потрібно урахувати для забезпечення заданої точності при даному значенні $\Delta\xi$. Але у даному випадку такий підхід неможливий, оскільки як різниця $\xi - \mu$ є випадковою величиною, яка, наприклад, для нормального розподілу може приймати будь-яке значення. Тому можливість збереження в ряду (1.3.10) тільки лінійних членів при обчисленні математичних сподівань може бути доведена виключно методами теорії ймовірностей. Підхід для визначення можливості використання лінійного наближення для U , який не потребує знання закону розподілу, базується на оцінці похибок обчислення перших моментів, що отримані при використанні лінійного наближення.

Збережемо в ряду Тейлора (1.3.10) тільки два перших члена ("лінеаризуємо функцію"), поклавши

$$U = y(\xi) = y(\mu) + \frac{dy}{d\mu}(\xi - \mu). \quad (1.3.11)$$

У цьому наближенні математичне сподівання і дисперсія легко обчислюються, і згідно з (1.2.3), (1.2.9) маємо

$$\begin{aligned} \mu_u &\equiv M[U] \cong y(\mu), \\ \sigma_u^2 &= M[(\xi - \mu)^2] \cong \left(\frac{dy}{d\mu}\right)^2 \sigma_\xi^2. \end{aligned} \quad (1.3.12)$$

Для визначення можливості використання оцінок (1.3.12) обчислимо моменти функції, зберігши в розкладі (1.3.10) також і квадратичний член:

$$U = y(\xi) \cong y(\mu) + \frac{dy}{d\mu}(\xi - \mu) + \frac{1}{2} \frac{d^2 y}{d\mu^2} (\xi - \mu)^2. \quad (1.3.13)$$

У цьому випадку обчислення приводять до таких значень моментів

$$\begin{aligned} \mu_u &\cong y(\mu) + \frac{1}{2} \frac{d^2 y}{d\mu^2} \sigma_\xi^2, \\ \sigma_u^2 &\cong \left(\frac{dy}{d\mu}\right)^2 \sigma_\xi^2 + \frac{dy}{d\mu} \frac{d^2 y}{d\mu^2} \mu^3 + \frac{1}{4} \left(\frac{d^2 y}{d\mu^2}\right)^2 (\mu^4 - \sigma_\xi^4). \end{aligned} \quad (1.3.14)$$

Порівнюючи формули (1.3.12), (1.3.14), бачимо, що похибки $\delta\mu$ і δD обчислення математичного сподівання і дисперсії при використанні лінеаризованої функції U наближено визначаються за формулами

$$\delta\mu \cong \frac{1}{2} \frac{d^2 y}{d\mu^2} \sigma_\xi^2, \quad \delta D \cong \left(\frac{dy}{d\mu} \right) \left(\frac{d^2 y}{d\mu^2} \right) \mu^3 + \frac{1}{2} \left(\frac{d^2 y}{d\mu^2} \right) (\mu^4 - \sigma_\xi^4). \quad (1.3.15)$$

У теорії непрямих вимірювань необхідність використання формул (1.3.15) звичайно не виникає, оскільки можливість використання лінеаризованих функцій при аналізі похибок не викликає сумнівів із фізичних міркувань. Слід тільки мати на увазі, що лінеаризація неможлива у випадку, коли похідна $dy/d\mu$ дорівнює нулю або прямує до нескінченності. У першому випадку розклад (1.3.10) не має лінійного члену і першим наближенням буде заміна функції $y(\xi)$ не лінійним, а квадратичним виразом. В іншому випадку в точці $\xi = \mu$ функцію не можна розкласти в ряд Тейлора. Наприклад, для функцій $y(\xi) = \cos \xi$ і $y(\xi) = \operatorname{ctg} \xi$ при $\mu = 0$ лінеаризація неприпустима: для першої функції – тому що $y'(\mu) = -\sin \mu = 0$, а для другої – тому що $y'(\mu) = -1/\sin^2 \mu = -\infty$.

Як і в одномірному випадку, можливість лінеаризації випадкової функції багатьох змінних для знаходження моментів може бути перевірена після їх обчислення з урахуванням членів розкладу функції в ряд Тейлора. Як і раніше, лінеаризація є неприпустимою, коли хоча б одна з похідних дорівнює нескінченності при $\xi_i = \mu_i$, $i = 1, \dots, n$. Однак рівність нулю деяких із перших похідних ще не вказує на незастосовність методу лінеаризації в багатомірному випадку, а лише потребує детальнішого розгляду оцінки похибки, яка виникає.

Розглянемо використання методу перенесення похибок на прикладі обчислення перших моментів функцій двох випадкових змінних.

$$U_1 = y_1(\xi_1, \xi_2) = \frac{\xi_1}{\xi_2}, \quad (1.3.16)$$

$$U_2 = y_2(\xi_2, \xi_2) = \xi_1 \xi_2. \quad (1.3.17)$$

Вважаємо, що матриця кореляцій відома і має вигляд

$$\hat{D}(\xi_1, \xi_2) = \begin{vmatrix} \sigma_1^2 & \sigma_1 \sigma_2 \rho_{12} \\ \sigma_1 \sigma_2 \rho_{12} & \sigma_2^2 \end{vmatrix}, \quad (1.3.18)$$

а середні значення $\mu_1 = M[\xi_1]$ і $\mu_2 = M[\xi_2]$ відмінні від нуля.

Тоді для математичних сподівань функцій U_1, U_2 отримуємо вирази

$$\mu_{u_1} = \frac{\mu_1}{\mu_2}; \quad \mu_{u_2} = \mu_1 \mu_2. \quad (1.3.19)$$

Матриці перших похідних $\hat{F}_1, \hat{F}_2$ відповідно для функцій U_1 і U_2 мають вигляд

$$\hat{F}_1 = \begin{pmatrix} 1 \\ \mu_2 \end{pmatrix}, \quad \hat{F}_2 = (\mu_2, \mu_1), \quad (1.3.20)$$

тому згідно з (1.3.12) і (1.3.13) отримуємо для дисперсій

$$\begin{aligned} \sigma_{u_1}^2 &= \frac{\mu_1^2}{\mu_2^2} \left[\frac{\sigma_1^2}{\mu_1^2} + \frac{\sigma_2^2}{\mu_2^2} - 2\rho_{12} \frac{\sigma_1}{\mu_1} \frac{\sigma_2}{\mu_2} \right], \\ \sigma_{u_2}^2 &= \mu_1^2 \mu_2^2 \left[\frac{\sigma_1^2}{\mu_1^2} + \frac{\sigma_2^2}{\mu_2^2} - 2\rho_{12} \frac{\sigma_1}{\mu_1} \frac{\sigma_2}{\mu_2} \right]. \end{aligned} \quad (1.3.21)$$

У багатьох випадках непрямих експериментів, як і у попередніх прикладах, функції $y \equiv U$ мають вигляд добутків деяких простих функцій від ξ_j . У цьому разі простіше обчислити не похідні $\partial U / \partial \mu$, а спершу – похідні від логарифма $\ln U$ і далі перейти до похідних $\partial U / \partial \mu$. Дійсно, позначимо

$$\psi = \ln U \equiv \ln y(\{\xi_j\}). \quad (1.3.22)$$

Диференціюючи (1.3.22) за ξ_i , знаходимо

$$\left. \frac{\partial y}{\partial \xi_i} \right|_{\xi_i = \mu_i} \equiv \frac{\partial y}{\partial \mu_i} = y(\mu) \frac{\partial \psi}{\partial \mu_i}. \quad (1.3.23)$$

Наприклад, у випадку незалежних величин, підставивши це співвідношення у формулу (1.3.9), маємо:

$$\sigma_u^2 \cong \mu_u^2 \sum_{j=1}^n \left(\frac{\partial \psi}{\partial \mu_j} \sigma_{\xi_j} \right)^2. \quad (1.3.24)$$

Цей вираз часто є зручнішим за (1.3.9), оскільки похідні $\partial \psi / \partial \mu_j$ обчислюються простіше.

Розглянемо приклад, який пояснює застосування формули (1.3.24), а саме, обчислимо середньоквадратичне відхилення моменту інерції F деякої кульки, що обертається відносно деякої осі, якщо її власним моментом інерції відносно центра можна знехтувати. Вважаємо, що відстань R від центра кульки до осі обертання задається із середньоквадратичною похибкою σ_R , а маса кульки визначається за формулою

$$m = \frac{\pi}{6} \rho d^3,$$

де густина речовини кульки ρ задається із середньоквадратичною похибкою σ_ρ , діаметр (d) кульки – з похибкою σ_d . Математичні сподівання R , ρ та d дорівнюють відповідно μ_R , μ_ρ , μ_d .

Як відомо з класичної механіки, момент інерції визначається за формулою

$$F = mR^2 = \frac{\pi}{6} R^2 \rho d^3.$$

Для обчислення дисперсії вводимо функцію

$$\psi = \ln F = 2 \ln R + \ln \rho + 3 \ln d + \ln \left(\frac{\pi}{6} \right).$$

Після її диференціювання отримуємо

$$\frac{\partial \psi}{\partial \mu_R} = \frac{2}{\mu_R}, \quad \frac{\partial \psi}{\partial \mu_\rho} = \frac{1}{\mu_\rho}, \quad \frac{\partial \psi}{\partial \mu_d} = \frac{3}{\mu_d}.$$

Підставляючи ці величини в (1.3.24), знаходимо

$$\sigma_F = \mu_F \sqrt{4 \left(\frac{\sigma_R}{\mu_R} \right)^2 + \left(\frac{\sigma_\rho}{\mu_\rho} \right)^2 + 9 \left(\frac{\sigma_d}{\mu_d} \right)^2}, \quad \mu_F = \frac{\pi}{6} \mu_R^2 \mu_\rho \mu_d^3.$$

1.4. Деякі приклади розподілів імовірностей

Як уже зазначалося раніше, властивості випадкових величин повністю визначаються їх функціями (щільностями) розподілу. Наведемо декілька прикладів щільностей розподілів випадкових величин, які часто вживаються при статистичній обробці експериментальних даних.

Рівномірний (прямокутний) розподіл в інтервалі a, b . Щільність розподілу є сталою, тому згідно з умовою нормування має вигляд

$$p(x) = \begin{cases} \frac{1}{b-a}, & x \in [a, b], \\ 0, & x \notin [a, b]. \end{cases} \quad (1.4.1)$$

Для функції розподілу отримуємо

$$F(x) = \int_{-\infty}^x p(t) dt = \begin{cases} 0, & x < a, \\ \frac{x-a}{b-a}, & a < x < b, \\ 1, & x > b. \end{cases} \quad (1.4.2)$$

Графіки цих функцій наведено на рис. 1.2.

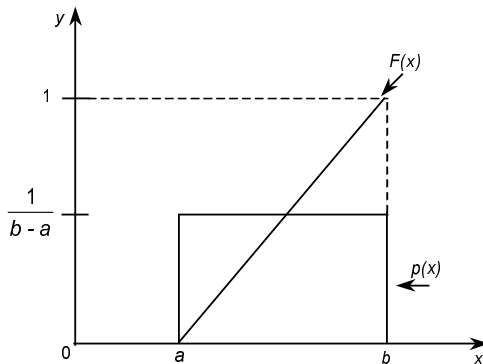


Рис. 1.2. Функція розподілу та щільність для рівномірно розподілених випадкових величин

Для прямокутного розподілу математичне сподівання та дисперсія дорівнюють відповідно:

$$\bar{x} = \mu = \int_a^b \frac{x}{b-a} dx = \frac{a+b}{2} \quad \text{середина відрізка,}$$

$$D(x) = \sigma^2 = \frac{1}{b-a} \int_a^b \left(x - \frac{a+b}{2} \right)^2 dx = \frac{(b-a)^2}{12}, \quad \sigma = \frac{b-a}{2\sqrt{3}}. \quad (1.4.3)$$

Коефіцієнт асиметрії дорівнює нулю:

$$\gamma \equiv \frac{\mu_3}{\sigma^3} = 0. \quad (1.4.4)$$

Імовірність знаходження випадкової рівномірно розподіленої величини ξ в інтервалі $[\alpha, \beta]$, де $\alpha, \beta \in [a, b]$ лінійно залежить від довжини інтервалу. Дійсно,

$$P(\alpha < \xi < \beta) \equiv F(\beta) - F(\alpha) = \frac{\beta - \alpha}{b - a}. \quad (1.4.5)$$

За допомогою (1.4.5) обчислимо ймовірність знаходження значень випадкової величини ξ з рівномірним розподілом на відрізку $\mu - \sigma, \mu + \sigma$. Згідно з (1.4.3) цей відрізок міститься в інтервалі $[a, b]$ і тому, використовуючи (1.4.3) і (1.4.5), маємо

$$P(\xi \in [\mu - \sigma, \mu + \sigma]) = \frac{2\sigma}{b-a} = \frac{1}{\sqrt{3}} \cong 0,5773, \quad (1.4.6)$$

тобто рівномірно розподілена випадкова величина знаходиться в інтервалі $[\mu - \sigma, \mu + \sigma]$ з ймовірністю 57,73.

Згідно із загальною термінологією, якщо значення величини ξ з ймовірністю P належить проміжку (x_1, x_2) , то останній називається $100 * P$ - м довірчим інтервалом, а x_1 і x_2 - його верхньою і нижньою границями відповідно. Імовірність P при цьому називають $100 * P\%$ - м рівнем надійності (значущості) параметра x . Таким чином, для випадкової величини з рівномірним розподілом відрізок $[\mu - \sigma, \mu + \sigma]$ є 57,73 - м довірчим інтервалом.

Нормальний розподіл, або розподіл за законом Гаусса. Він грає фундаментальну роль у теорії похибок. Параметрами розподілу є математичне сподівання μ та середньоквадратичне відхилення σ , а його щільність має вигляд

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad \bar{x} \equiv \mu, \quad \overline{(x-\bar{x})^2} \equiv \sigma^2. \quad (1.4.7)$$

На рис. 1.3 показано нормальні розподіли випадкової величини при $\mu = 0$ та $\sigma = 1$ і 2 .

Похибки, що обумовлені багатьма випадковими факторами з малими внесками, підпорядковуються нормальному розподілу. Такі характеристики, як середньоарифметичне значення вибірки та середньоквадратичне відхилення цього значення також асимптотично підпорядковуються нормальному розподілу, навіть якщо самі випадкові величини і не розподілені за нормальним законом.

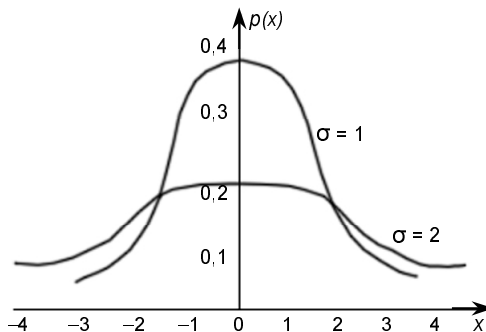


Рис. 1.3. Щільність нормального розподілу з $\mu = 0$, $\sigma = 1$ і 2

Нормальний розподіл зі щільністю (1.4.7) часто називають розподілом типу $N(\mu, \sigma^2)$. Якщо випадкова величина має розподіл $N(\mu, \sigma^2)$, то, очевидно, величина

$$u = \frac{x - \mu}{\sigma} \quad (1.4.8)$$

матиме нормальний розподіл типу $N(0,1)$ з нульовим середнім значенням та одиничним середньоквадратичним відхиленням, тобто її щільність дається виразом

$$p(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}. \quad (1.4.9)$$

Величину u (1.4.8) називають стандартною нормально розподіленою випадковою величиною.

Центральні моменти нормального розподілу задовольняють рекурентному співвідношенню

$$\mu_r = (r-1)\sigma^2\mu_{r-2}. \quad (1.4.10)$$

Тобто парні моменти залежать лише від середньоквадратичного відхилення σ :

$$\begin{aligned} \mu_2 &= \sigma^2, \quad \mu_4 = 3\sigma^4, \quad \mu_6 = 15\sigma^6, \\ \mu_r &= (r-1)!!\sigma^r, \end{aligned} \quad (1.4.11)$$

а всі непарні моменти дорівнюють нулю, оскільки розподіл $N(\mu, \sigma^2)$ є симетричним.

Імовірність знайти значення випадкової величини розподілу $N(\mu, \sigma^2)$ в інтервалі (x_1, x_2) дорівнює

$$P(x_1 < \xi < x_2) \equiv \int_{-\infty}^{x_2} dx p(x) - \int_{-\infty}^{x_1} dx p(x) = \frac{1}{2} \left[\Phi\left(\frac{x_2 - \mu}{\sigma}\right) - \Phi\left(\frac{x_1 - \mu}{\sigma}\right) \right], \quad (1.4.12)$$

де функція

$$\Phi(z) = \frac{2}{\sqrt{2\pi}} \int_0^z e^{-\frac{t^2}{2}} dt = -\Phi(z) \equiv \operatorname{erf}\left(\frac{z}{\sqrt{2}}\right) \quad (1.4.13)$$

є так званим інтегралом імовірності або функцією Лапласа, а $\operatorname{erf}(x)$ – функція помилок.

Оскільки $\Phi(-\infty) = -1$, то з (1.4.12) випливає такий вираз для функції розподілу нормальної величини:

$$F(x) = P(-\infty < \xi < x) = \frac{1}{2} \left[\Phi\left(\frac{x - \mu}{\sigma}\right) + 1 \right]. \quad (1.4.14)$$

Зазначимо, що в багатьох довідниках замість таблиці значень інтегралу ймовірності $\Phi(x)$ подається таблиця значень функцій розподілу $F(x)$ випадкових величин, що підпорядковуються стандартному нормальному закону $N(0,1)$:

$$F(x) = \frac{1}{2} [\Phi(x) + 1]. \quad (1.4.15)$$

Згідно з (1.4.12) і (1.4.13) імовірність того, що ξ набуває значень з відрізка $[\mu - t\sigma, \mu + t\sigma]$, тобто, що виконується нерівність $|\xi - \mu| \leq t\sigma$ ($t > 0$), дорівнює

$$P(\mu - t\sigma \leq \xi \leq \mu + t\sigma) \equiv P(|\xi - \mu| \leq t\sigma) = \Phi(t). \quad (1.4.16)$$

У випадку $t = 1, 2, 3$ за допомогою таблиць функцій Лапласа дістаємо:

$$\begin{aligned} P(|\xi - \mu| \leq \sigma) &= \Phi(1) \approx 0,68, \\ P(|\xi - \mu| \leq 2\sigma) &= \Phi(2) \approx 0,95, \\ P(|\xi - \mu| \leq 3\sigma) &= \Phi(3) \approx 0,997. \end{aligned} \quad (1.4.17)$$

Таким чином, відрізок $[\mu - 3\sigma, \mu + 3\sigma]$ є 99,7 -м довірчим інтервалом випадкової величини ξ .

Число 0,997 мало відрізняється від одиниці, тому подію з імовірністю 0,997 можна вважати практично достовірною, й останню рівність у (1.4.17) формують як так зване правило трьох сигм: практично достовірно, що нормально розподілена випадкова величина може відхилитися від свого математичного сподівання не більше ніж на потроєне середньоквадратичне відхилення. Цим правилом широко користуються на практиці.

Експериментатори іноді подають дані у вигляді $\mu \pm \Delta$, де величина Δ визначена як півширина щільності нормального розподілу і є відхиленням випадкової величини від її середнього значення, що приводить до зменшення у два рази значення її щільності розподілу. Тобто для визначення Δ маємо співвідношення

$$\frac{p(\mu)}{2} = p(\mu \pm \Delta). \quad (1.4.18)$$

З урахуванням (1.4.7) отримуємо $-\Delta^2 / 2\sigma^2 = -\ln 2$ і

$$\Delta = \sigma\sqrt{\ln 4} \approx 1,39\sigma. \quad (1.4.19)$$

Оскільки імовірність знаходження значень величини ξ в інтервалі $[\mu - 1,39\sigma, \mu + 1,39\sigma]$ дорівнює

$$P[|\xi - \mu| \leq 1,39\sigma] = \Phi(1,39) \approx 0,836, \quad (1.4.20)$$

то інтервал $[\mu - \Delta, \mu + \Delta] \equiv [\mu - 1,39\sigma, \mu + 1,39\sigma]$ вміщує 83,6 значень випадкової величини.

χ^2 -розподіл. Щільність цього розподілу залежить від одного параметра n і має вигляд

$$p(X) = \frac{X^{(n-2)/2}}{2^{n/2} \Gamma\left(\frac{n}{2}\right)} e^{-X/2}, \quad 0 \leq X < \infty, \quad (1.4.21)$$

де $\Gamma(z)$ – гама-функція:

$$\Gamma(n+1) = n!, \quad \Gamma\left(n + \frac{1}{2}\right) = \frac{(2n-1)!!}{2^n} \Gamma\left(\frac{1}{2}\right), \quad \Gamma\left(\frac{1}{2}\right) = \pi^{1/2} \approx 1,77245.$$

Параметр n приймає цілі додатні значення і називається числом ступенів свободи.

Функція (1.4. 21) є щільністю розподілу випадкової величини

$$X \equiv \chi_n^2 = \sum_{i=1}^n x_i^2, \quad (1.3.22)$$

що є сумою квадратів n незалежних випадкових величин x_i , які мають нормальний розподіл $N(0,1)$ із нульовим середнім значенням та одиничним середньоквадратичним відхиленням. На рис. 1.4 показано χ^2 -розподіл при $n = 1, 5$ і 10 .

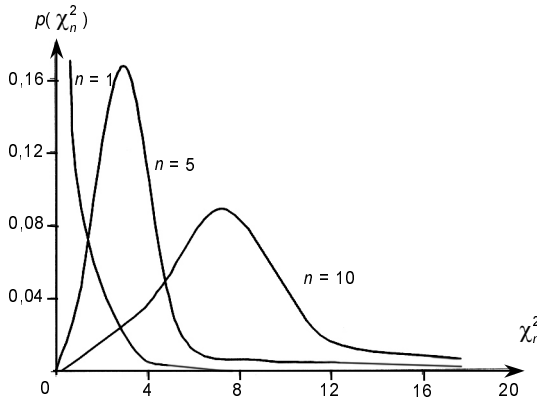


Рис. 1.4. Графік χ_n^2 -розподілу

Середнє значення величини $X \equiv \chi_n^2$ дорівнює n , а дисперсія – $2n$:

$$M[\chi_n^2] = \overline{\chi_n^2} = n, \quad D[\chi_n^2] = 2n. \quad (1.4.23)$$

Щільність $P(\chi_n^2)$ є асиметричною функцією з коефіцієнтом асиметрії:

$$\gamma_3 = 2\sqrt{2/n}. \quad (1.4.24)$$

Зі збільшенням n (практично при $n \geq 30$) χ_n^2 -розподіл (1.4.21) прямує до нормального з параметрами $N(n, 2n)$. Очевидно, що виконується співвідношення

$$\frac{\chi_n^2}{n} \xrightarrow{n \rightarrow \infty} 1, \quad (1.4.25)$$

і дисперсія χ_n^2/n прямує до нуля.

Для χ^2 -розподілу виконується теорема додавання, згідно з якою сума X , що складається з k випадкових незалежних величин $X_{n_i} \equiv \chi_{n_i}^2$, $i = 1 \div k$, кожна з яких має $\chi_{n_i}^2$ -розподіл з n_i ступенями свободи, підпорядковується χ_n^2 -розподілу з кількістю ступенів свободи n , що дорівнює сумі кількості ступенів свободи кожного із доданків:

$$X_n = \sum_{i=1}^k X_{n_i}, \quad n = \sum_{i=1}^k n_i. \quad (1.4.26)$$

Існує також і обернена теорема – теорема розкладу χ_n^2 -розподілу на суми квадратів величин, лінійно зв'язаних із деякими випадковими величинами u_i . Ця теорема формулюється таким чином.

Нехай існують суми квадратів

$$X = \sum_{j=1}^k l_j^2, \quad (1.4.27)$$

де k випадкових величин l_i є лінійними функціями незалежних змінних u_i , кожна з яких нормально розподілена з параметрами $N(0,1)$. Якщо між вели-

чинами u_i існують m лінійних співвідношень (зв'язків), то кількість ступенів свободи суми X буде дорівнювати

$$n = k - m . \quad (1.4.28)$$

Таким чином, якщо внаслідок лінійних перетворень сума квадратів n випадкових величин $u_1, \dots, u_n$ може бути розбита на k сум квадратів $X_{n_1}, \dots, X_{n_k}$ з $n_1, \dots, n_k$ ступенями свободи, тобто

$$X \equiv \chi_n^2 = \sum_{i=1}^n u_i^2 = \sum_{j=1}^k X_{n_j}, \quad (1.4.29)$$

то необхідною і достатньою умовою того, що величини $X_{n_1}, \dots, X_{n_k}$ будуть незалежними й описуватимуться χ^2 розподілами з $n_1, \dots, n_k$ ступенями свободи, є виконання рівності

$$\sum_{j=1}^k n_j = n . \quad (1.4.30)$$

Ця теорема виявляється дуже корисною для дослідження властивостей різних випадкових величин.

Розподіл Стьюдента (t -розподіл). Якщо дійсна випадкова величина $- \infty \leq t \leq \infty$ має щільність розподілу вигляду

$$p(t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}} \equiv S_n(t), \quad (1.4.31)$$

то кажуть, що вона розподілена за законом Стьюдента або задовольняє t -розподілу з n ступенями свободи. Середнє значення величини t дорівнює

нулю, а його дисперсія $-\frac{n}{n-2}$:

$$M[t] = 0; \quad D[t] = \frac{n}{n-2}. \quad (1.4.32)$$

Приклад t -розподілу з $n = 5$ подано на рис. 1.5. Розподіл Стюдента є симетричним відносно $t = 0$. При $n = 1$ він переходить до розподілу Коші (1.4.35), а при $n \rightarrow \infty$ прямує до нормального розподілу $N(0; 1)$.

Уперше цей розподіл запропонував англійський математик В.Госсет (псевдонім Стюдент (S), 1908). Наприклад, щільність розподілу (1.4.31) має випадкова величина

$$t = (\bar{x} - \mu) \frac{\tilde{s}}{\sqrt{n-1}}, \quad (1.4.33)$$

де $\bar{x} = \sum_{i=1}^n x_i / n$ – вибіркове середнє значення вибірки обсягу n випадкової величини X , що розподілена за нормальним законом із математичним сподіванням μ , а $\tilde{s}$ – незсунене середньоквадратичне відхилення вибіркового середнього (див. 2.3):

$$\tilde{s}^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2. \quad (1.4.34)$$

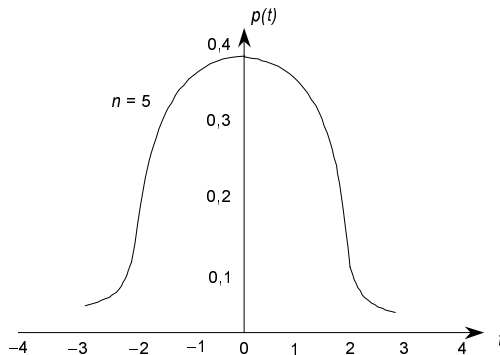


Рис. 1.5 Розподіл Стюдента з $n = 5$

Розподіл Коші, або розподіл Брейта – Вігнера. Його визначено за всіх дійсних значень $-\infty \leq x \leq \infty$ та задано симетричною щільністю розподілу вигляду

$$p(x) = \frac{1}{\pi} \frac{1}{1+x^2} \equiv p_k(x). \quad (1.4.35)$$

Розподіл Коші відноситься до так званих паталогічних, оскільки математичне сподівання для випадкових величин з цього розподілу, строго кажучи, не визначено, а всі інші моменти нескінченні. На рис. 1.6 порівнюються розподіл Коші та нормальний розподіл $N(0; 1)$ при $x \geq 0$.

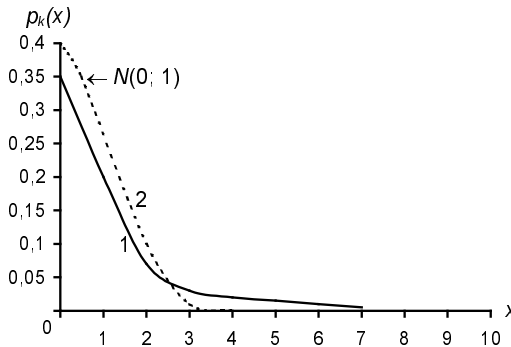


Рис. 1.6. Порівняння крив розподілу Коші (крива 1) та стандартного нормального розподілу (2) типу $N(0; 1)$

У фізиці розподіл Коші вигляду

$$p(x) \equiv \frac{1}{\pi\Gamma} \left[\frac{1}{1 + \left(\frac{x - x_0}{\Gamma} \right)^2} \right] \equiv \frac{1}{2\pi} \frac{\Gamma}{\Gamma^2 + (x - x_0)^2}, \quad (1.4.36)$$

де x_0 і Γ – параметри розташування та масштабу, називають розподілом Брейта – Вігнера. Величина x_0 визначає максимум або моду розподілу, а Γ є напівшириною на половині висоти. Дійсно, згідно з визначенням напівширини Δ на половині висоти маємо

$$\frac{1}{2} p_{\max} = p(x_0 + \Delta); \quad p_{\max} = p(x_0). \quad (1.4.37)$$

Звідки з урахуванням (1.4.35) отримуємо

$$\frac{1}{2} \frac{1}{\pi \Gamma} = \frac{1}{\pi \Gamma} \left(\frac{1}{1 + \left(\frac{\Delta}{\Gamma} \right)^2} \right)$$

і знаходимо, що $\Gamma \equiv \Delta$.

Розподіл Коші приймає властивості звичайного розподілу після його модифікації за допомогою процедури обмеження інтервалу визначення. Така найпростіша модифікація ідеальних розподілів обумовлена тим фактом, що область можливих значень вимірювань ніколи не буває нескінченною. Звичайно значення деякої фізичної величини X лежать у середині інтервалу з кінцевими межами A , B і тому, з погляду фізичних уявлень, щільність ідеального розподілу можна модифікувати таким чином:

1. Вважати, що вона відмінна від нуля лише на деякому дуже великому, але скінченному інтервалі $[A, B]$

2. Перенормувати щільність розподілу згідно з умовою нормування в інтервалі $[A, B]$

$$p(x) \rightarrow \frac{p(x)}{\int_A^B p(x) dx} = \frac{p(x)}{F(B) - F(A)}. \quad (1.4.38)$$

Хоча така процедура й ускладнює вирази, але іноді вона дуже корисна, як, наприклад, у випадку розподілу Коші $p_k(x)$. Дійсно, моменти цього розподілу нескінченні, але якщо інтервал визначення щільності розподілу Коші обмежити фіксованими значеннями на інтервалі від $-A$ до A ,

$$p(x) = \frac{p_k(x)}{\int_{-A}^A \frac{1}{\pi} \frac{1}{1+x^2} dx} = \frac{1}{2 \arctg A} \frac{1}{1+x^2}, \quad (1.4.39)$$

то математичне сподівання та дисперсія будуть скінченними і дорівнюватимуть

$$M[x] = \frac{1}{2 \operatorname{arctg} A} \int_{-A}^A \frac{x}{1+x^2} dx = 0;$$

$$D[x] = \frac{1}{2 \operatorname{arctg} A} \int_{-A}^A \frac{x^2}{1+x^2} dx = \frac{A}{\operatorname{arctg} A} - 1.$$

При $A \rightarrow \infty$ маємо $D[x] \rightarrow \infty$, тобто розподіл Коші спадає настільки повільно, що дисперсія може бути довільно великою при достатньо великому значенні границі інтервалу A .

Розподіл Пуассона. На відміну від попередніх виразів для щільностей неперервних випадкових величин розподіл Пуассона є розподілом значень дискретної додатної випадкової змінної n . Згідно з визначенням дискретну випадкову величину $\xi = n$ розподілено за законом Пуассона, якщо ймовірність її появи визначається функцією

$$P(\xi = n) \equiv P(n, \mu) = \frac{\mu^n \exp(-\mu)}{n!}. \quad (1.4.40)$$

У цьому випадку математичне сподівання і дисперсія збігаються та дорівнюють параметру μ

$$M[n] = D[n] = \mu. \quad (1.4.41)$$

Дійсно,

$$M[n] \equiv M[\xi = n] = \sum_{n=0}^{\infty} n P_n = e^{-\mu} \sum_{n=1}^{\infty} \frac{n \mu^n}{n!} = e^{-\mu} \mu \sum_{n=0}^{\infty} \frac{\mu^n}{n!} = \mu$$

$$D[n] = M[n^2] - (M[n])^2 = \mu^2 + \mu - \mu^2 = \mu.$$

Розподіл Пуассона – асиметричний із коефіцієнтом асиметрії

$$\gamma_3 = \frac{1}{\sqrt{\mu}}. \quad (1.4.42)$$

Функція розподілу для випадкової величини, розподіленої за законом Пуассона, очевидно, має вигляд

$$F(x) \equiv P(n < x) = \sum_{n=0}^x \frac{\mu^n \exp(-\mu)}{n!}, \quad F(+\infty) = 1. \quad (1.4.43)$$

На рис. 1.7 показано функції ймовірності розподілу Пуассона для деяких значень μ . За $\mu \rightarrow \infty$ розподіл Пуассона прямує до нормального розподілу.

Розподіл Пуассона визначає ймовірність появи n подій у заданому інтервалі часу t за умови, що події є незалежними і виникають зі сталою середньою швидкістю $\lambda = \mu/t$:

$$P_n(t) = P(n, \mu = \lambda t) \equiv \frac{(\lambda t)^n e^{-\lambda t}}{n!}. \quad (1.4.44)$$

Дійсно, обчислимо ймовірність появи таких подій, що виконуються за таких умов:

1. Ймовірність появи події в момент часу t не залежить від попереднього часу і кількості подій, що відбувалися раніш.
2. Ймовірність появи окремої події протягом малого інтервалу часу $\delta t \ll 1$ зростає зі сталою швидкістю (ймовірністю в одиницю часу) λ , тобто поява окремої події в інтервалі часу $(t, t + \delta t)$ дорівнює $\lambda \delta t$.
3. Ймовірність появи двох або більшої кількості подій за той же проміжок часу $(t, t + \delta t)$ дорівнює нулю.

Тепер обчислимо за таких припущень ймовірність того, що в інтервалі $(0, t)$ відбудеться n подій. Для цього розглянемо зміну ймовірності $P_r(t)$ появи r подій протягом малого часу $\delta t \ll 1$, а потім складемо диференціальне рівняння для даної ймовірності. Очевидно, за проміжок часу $(0, t + \delta t)$ не відбувається жодної події, якщо їх не було в інтервалах $(0, t)$ і $(t, t + \delta t)$, тобто, ураховуючи незалежність подій, маємо

$$P_0(t + \delta t) = P_0(t)(1 - \lambda \delta t). \quad (1.4.45)$$

Звідки,

$$\frac{P_0(t + \delta t) - P_0(t)}{\delta t} = -\lambda P_0(t). \quad (1.4.46)$$

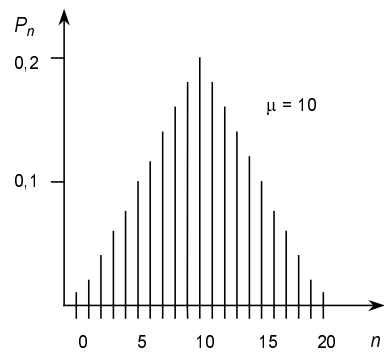
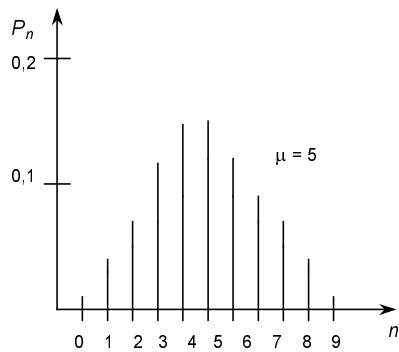
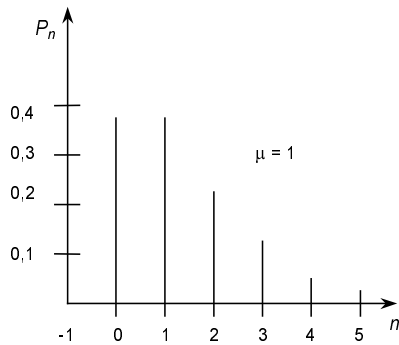


Рис. 1.7. Функції ймовірностей для розподілів Пуассона за $\mu = 1, 5$ та 10

Спрямувавши δt до нуля, отримуємо диференційне рівняння для ймовірності відсутності події

$$\frac{dP_0(t)}{dt} = -\lambda P_0(t). \quad (1.4.47)$$

Розв'язок цього рівняння має вигляд $P_0(t) = Ae^{-\lambda t}$. При $t = 0$ маємо $P_0(0) = 1$. Тобто $A = 1$ і

$$P_0(t) = e^{-\lambda t}. \quad (1.4.48)$$

При $r = 1$ маємо дві можливості – подія відбувається або в інтервалі $(0, t)$, або в інтервалі $(t, t + \delta t)$. Звідки при $\delta t \ll 1$ знаходимо

$$\begin{aligned} P_1(t + \delta t) &= P_1(t)(1 - \lambda \delta t) + P_0(t)\lambda \delta t, \\ \frac{P_1(t + \delta t) - P_1(t)}{\delta t} &= -\lambda P_1(t) + \lambda P_0(t). \end{aligned} \quad (1.4.49)$$

При $\delta t \rightarrow 0$ отримуємо диференційне рівняння

$$\frac{dP_1(t)}{dt} = -\lambda P_1(t) + \lambda P_0(t). \quad (1.4.50)$$

З урахуванням (1.4.47) розв'язок цього рівняння має вигляд

$$P_1(t) = \lambda t e^{-\lambda t}. \quad (1.4.51)$$

У загальному випадку диференційне рівняння набуває вигляду

$$\frac{dP_n(t)}{dt} = -\lambda P_n(t) + \lambda P_{n-1}(t), \quad (1.4.52)$$

розв'язком якого є функція розподілу Пуассона (1.4.44).

Умови 1–3, що визначають розподіл Пуассона з великою точністю, виконуються в багатьох процесах і явищах. Наприклад, у процесах радіоактивного розпаду ядер функція $P_n(t)$ визначає ймовірність появи n частинок при їх вильоті протягом фіксованого часу t (зі сталою розпаду λ).

Розділ 2

ОСНОВИ МЕТОДУ ВИБІРКИ

2.1. Математична статистика та її завдання.

Функція розподілу вибірки

Аналізом масових явищ і процесів за даними, що знаходяться зі спостережень над ними, займається математична статистика. Головна мета математичної статистики полягає в отриманні статистичних висновків, які відносяться не до окремих випробувань, а являють собою твердження про загальні характеристики явищ (закони розподілів, імовірності їх появи, математичні сподівання тощо), як правило, у припущенні незмінності умов, що породжують досліджуване явище.

Математична статистика базується на теорії ймовірності й вивчає методи, які дають змогу за результатами випробувань робити певні ймовірнісні висновки про властивості системи. З погляду теорії ймовірності головним завданням математичної статистики є оцінка характеристик генеральної сукупності випадкових величин за їх вибірковими значеннями.

Згідно з визначенням випадковою вибіркою, або просто вибіркою обсягу n , називається послідовність $x_1, x_2, \dots, x_n$ із n незалежних випадкових величин з одним і тим самим розподілом, який збігається з розподілом випадкової величини X , що досліджується. Інакше кажучи, випадкова вибірка обсягу n – це результат n послідовних і незалежних спостережень випадкової величини X , усі можливі значення якої формують генеральну сукупність.

Конкретний набір вибіркових значень $x_1, x_2, \dots, x_n$ можна інтерпретувати як реалізацію (одну з багатьох) багатовимірної випадкової величини $\vec{X}$ з незалежними компонентами $X_1, X_2, \dots, X_n$, які мають однакову функцію розподілу $F_0(z_i)$, що відповідає генеральній сукупності. Точну функцію розподілу, її параметри та числові характеристики зазвичай називають теоретичними на відміну від вибіркових (емпіричних) значень, які знаходяться за вибірковими даними. У випадку так званої вибірки з поверненням, коли склад генеральної сукупності не змінюється, багатовимірною випадковою величиною

$\bar{X}$ має сумісну теоретичну функцію розподілу у вигляді добутку функцій розподілу $F_0(z_i)$:

$$F(z_1, z_2, \dots, z_n) = P(x_1 < z_1, x_2 < z_2, \dots, x_n < z_n) = \prod_{i=1}^n P(x_i < z_i) = \prod_{i=1}^n F_0(z_i) \quad (2.1.1)$$

На практиці вибірка з генеральної сукупності часто здійснюється так, що після кожного експерименту досліджений елемент не повертається до генеральної сукупності перед наступним вибором (наприклад, так здійснюється контроль якості й така сама ситуація має місце в ядерних реакціях на тій самій мішені). При цьому, очевидно, не можна вважати, що всі експерименти проводяться в однакових умовах і що результати незалежні один від одного, оскільки після кожного експерименту склад генеральної сукупності змінюється. Така вибірка називається вибіркою без повернення, на відміну від простої випадкової вибірки, або вибірки з поверненням. Але якщо генеральна сукупність достатньо велика, а вибірка за своїм обсягом становить лише незначну її частку, то різниця між цими двома типами вибірок незначна і нею можна нехтувати та вважати, що сумісна функція розподілу вибірки задається (2.1.1).

Усі висновки про генеральну сукупність проводяться тільки з аналізу властивостей вибірок. Точність та надійність (достовірність) статистичних висновків визначаються не для даної конкретної вибірки $x_1, \dots, x_n$, тобто конкретної реалізації багатовимірної випадкової величини $X_1, X_2, \dots, X_n$, а для всього можливого набору вибірок. Тому головні задачі математичної статистики полягають у розробці та використанні методів оцінки загальних характеристик сукупності випадкових величин, перевірці статистичних гіпотез, виявленню взаємних зв'язків між характеристиками об'єктів тощо.

Усі задачі можна розбити на дві великі групи: **параметричні й непараметричні**. Наприклад, у попередньому розділі було розглянуто величини теорії ймовірності, що залежали від параметрів, а саме щільності та функції розподілу ймовірностей. Такі величини мали вигляд деяких визначених функцій

$\Phi(\theta_1, \theta_2, \dots, \theta_m)$ з деякими параметрами $\{\theta_j\}$. У цьому випадку для визна-

чення розподілів треба знайти оцінки для цих параметрів. Так, наприклад, у багатьох випадках на основі центральної граничної теореми можна зробити висновок, що випадкову величину розподілено нормально. Оскільки функція розподілу в цьому випадку залежить від двох параметрів, а саме, математичного сподівання μ та дисперсії σ^2 , то для визначення функції розподілу досить оцінити за результатами випробувань лише ці два параметри. Такі задачі, коли є відомими вигляд функції розподілу та область змін їх параметрів, називаються **параметричними**. Коли ж немає ніякої уяви навіть про вигляд щільності розподілу, не кажучи вже про можливі значення параметрів, ідеться про **непараметричні задачі**.

2.2. Загальні вимоги до оцінок параметрів

Оцінкою або статистикою деякого параметра θ теоретичного розподілу (наприклад, математичного сподівання, дисперсії тощо) називають таку функцію $\theta^* \equiv h$ від значень випадкових величин, яка в певному статистичному розумінні мало відрізняється від θ :

$$\theta^* = h(x_1, \dots, x_n). \quad (2.2.1)$$

Оцінки, що визначаються одним числом, називаються **точковими**. Оцінка вигляду (2.2.1) є точковою. При вибірці малого обсягу випадкових величин $\{x_j\}$ точкова оцінка може значно відрізнятись від оцінюваного параметра, тобто приводити до грубих помилок.

При невеликому обсязі вибірки слід користуватися так званими **інтервальними оцінками**. Інтервальною називають оцінку, що характеризується двома числами, які визначають кінцеві значення інтервалу, де одночасно можуть знаходитися величини θ і θ^* . Інтервальні оцінки дозволяють встановити точність і надійність (достовірність) оцінок у деякому статистичному розумінні. Слід зазначити, що абсолютно точно визначити параметр θ в результаті випробувань в принципі не можливо вже хоча б тому, що θ є константою, а оцінка $\theta^* = h(\{x_j\})$ як функція вибірки, тобто як функція випадкової величини, є випадковою величиною.

Кількісно міру близькості θ^* до справжнього значення θ , тобто похибки при заміні θ на θ^* або міру розсіювання θ^* відносно θ , цілком природно характеризувати величиною середньоквадратичної похибки

$$\sigma^2 \equiv M[(\theta^* - \theta)^2] \equiv \overline{(\theta^* - \theta)^2}. \quad (2.2.2)$$

Серед усіх оцінок з однаковою дисперсією $D[\theta^*]$ мінімальну міру розсіювання відносно $\theta^* = h(x_1, \dots, x_n)$ мають оцінки, для яких $M[\theta^*] = \theta$. Дійсно, оскільки θ є детермінованою величиною, то маємо

$$\begin{aligned} \sigma^2 &= M[(\theta^* - \theta)^2] = M[(\theta^* - \bar{\theta}^*) - (\bar{\theta}^* - \theta)]^2 = M[(\theta^* - \bar{\theta}^*)^2] - \\ &- 2M[(\theta^* - \bar{\theta}^*)(\bar{\theta}^* - \theta)] + M[(\bar{\theta}^* - \theta)^2] = M[(\theta^* - \bar{\theta}^*)^2] + (\bar{\theta}^* - \theta)^2 \quad (2.2.3) \\ &\equiv D[\theta^*] + (M[\theta^*] - \theta)^2 = D[\theta^*]. \end{aligned}$$

Оцінку θ^* називають незсуненою оцінкою параметра θ , якщо виконується рівність

$$M[\theta^*] = \theta, \quad (2.2.4)$$

або $M[\theta^* - \theta] = 0$. Згідно з (2.2.3) середній квадрат відхилення значення незсуненої оцінки від точного значення величини дорівнює дисперсії значення оцінки:

$$\sigma_n^2 \equiv M[(\theta^* - \theta)^2] = D[\theta^*] \equiv M[(\theta^* - M[\theta^*])^2]. \quad (2.2.5)$$

Незсуненість оцінки можна трактувати так: при багаторазовому використанні незсуненої оцінки $\theta^* \equiv h\{x_j\}$ як значення θ (тобто при багаторазовій заміні θ на величину θ^* , що обчислена з використанням різних вибірок $\{x_j\}$) середнє значення похибки $\theta^* - \theta$ буде дорівнювати нулю.

Тобто використання незсунених оцінок дає змогу уникнути систематичних похибок одного знаку. Часто перевірити незсуненість оцінки можна навіть тоді, коли є невідомим розподіл імовірності. Коли математичне сподівання оцінки не збігається зі значенням параметра, що оцінюється, то оцінка називається зсуненою.

Незсунені оцінки мають мінімальну міру розсіювання $\sigma_n^2 \equiv \frac{2}{n}$ відносно θ порівняно з оцінками, для яких $M[\theta^*] \neq \theta$. Разом із тим, можуть існувати різні незсунені оцінки одного і того ж самого параметра. Із сукупності таких оцінок мінімальне відхилення σ_n^2 (2.2.5) будуть мати ті, які мають мінімальну дисперсію. Незсунену оцінку θ^* параметра θ називають його ефективною оцінкою (найкращою незсуненою оцінкою), якщо θ^* має мінімально можливу дисперсію

$$D[\theta^*] = \min_{\theta^* \in M[\theta^*] = \theta} D[\theta^*] \equiv D_{\inf}[\theta^*], \quad (2.2.6)$$

де символ "min" означає операцію взяття точної нижньої границі.

Такі оцінки ще називають оцінками мінімальної дисперсії, оскільки вони мають мінімально можливе розсіювання σ_n^2 відносно оцінюваного параметра.

Природно виникає питання: наскільки малою може бути дисперсія $D_{\text{inf}}[\theta^*]$ ефективної оцінки. Виявляється, коли сукупність розподілів вибірки $\{x_i\}$ досить регулярно залежить від оцінювання параметра θ , тоді можна вказати нижню межу дисперсій усіх незсунених оцінок параметра. Знайти таку межу дозволяє нерівність Крамера – Рао. Порівнюючи дисперсію даної оцінки з нижньою межею дисперсій незсунених оцінок, можна з'ясувати, наскільки оцінка близька до найкращої можливої. Теорема Крамера – Рао, або нерівність Крамера – Рао, визначає не тільки нижню межу дисперсій, але й дає загальне параметричне співвідношення для щільностей розподілу, для яких ця межа може реалізовуватися. У випадку неперервних розподілів зі щільністю $p(\vec{x}; \theta) \equiv p(x_1, x_2, \dots, x_n; \theta)$, що залежить від параметра θ , ця теорема формулюється таким чином.

Теорема (нерівність) Крамера – Рао Нехай щільність розподілу p є двічі диференційованою функцією за θ і $\theta^* = h(\vec{x})$ – незсунена і гладка в деякому розумінні* оцінка параметра θ . Тоді дисперсія θ^* задовольняє нерівності

$$D[\theta^*] \geq \frac{1}{I(\theta^*)}, \quad (2.2.7)$$

де $I(\theta^*)$ – кількість інформації за Фішером,

$$I(\theta^*) \equiv M \left[\left(\frac{\partial}{\partial \theta^*} p(\vec{x}; \theta^*) \right)^2 \right]. \quad (2.2.8)$$

Причому знак рівності у формулі (2.2.7) досягається тоді й тільки тоді, коли похідну від p можна подати у вигляді

$$\frac{\partial}{\partial \theta} p(\vec{X}; \theta) = c(\theta) (h(\vec{X}) - \theta), \quad (2.2.9)$$

де $c(\theta)$ – деяка функція від θ .

Ця теорема має такі два наслідки:

* При формулюванні цієї теореми відзначено лише загальні риси поведінки функцій p та h . Точні умови на ці функції можна знайти у стандартних підручниках з математичної статистики.

1. Якщо для деякої оцінки $\tilde{\theta} = \theta^*$ нерівність Крамера – Рао перетворюється на рівність

$$D[\tilde{\theta}] = \frac{1}{I(\tilde{\theta})}, \quad (2.2.10)$$

то $\tilde{\theta}$ є ефективною оцінкою параметра θ .

2. Якщо деяка функція $h(\vec{x})$ і щільність $p(\vec{x}; \theta)$ зв'язані умовою (2.2.9), то $\theta^* = h(\vec{x})$ є незсуненою та ефективною оцінкою параметра θ .

Можна розглядати не одну оцінку $\theta^* = h(\vec{x}) \equiv h(x_1, \dots, x_m)$, побудовану за деякою вибіркою $\vec{x} = (x_1, \dots, x_m)$, а послідовність оцінок

$$\theta_n^* = h(x_1, \dots, x_n), \quad n \geq m, \quad (2.2.11)$$

з різною кількістю n елементів у вибірці. Оцінку θ_n^* (фактично послідовність оцінок) називають спроможною (слухною, консистентною) оцінкою параметра θ , якщо при $n \rightarrow \infty$ з ймовірністю близької до одиниці абсолютне значення різниці між θ_n^* та θ виявляється меншим за будь-яке достатньо мале число ε , тобто

$$P\left(|\theta_n^* - \theta| < \varepsilon\right) \rightarrow 1, \quad \varepsilon \rightarrow +0. \quad (2.2.12)$$

Співвідношення (2.2.12) є визначенням збіжності за ймовірністю величин θ_n^* і θ і зазвичай записується у вигляді

$$\theta_n^* \xrightarrow{P} \theta. \quad (2.2.13)$$

Якщо існує деяка неперервна функція $(\{\theta_i\})$ параметрів $\{\theta_i\}$, і значення $\{\theta_i^*\}$ є спроможними оцінками цих параметрів, то $(\{\theta_i^*\})$ також є спроможною оцінкою функції $(\{\theta_i\})$, тобто

$$(\{\theta_i^*\}) \xrightarrow{P} (\{\theta_i\}) \quad \text{при} \quad \theta_i^* \xrightarrow{P} \theta_i. \quad (2.2.14)$$

Зазначимо, що саме при використанні спроможних оцінок є виправданим збільшення числа елементів вибірки, оскільки саме в цьому випадку при збільшенні обсягу зменшується ймовірність можливої похибки оцінки невідомого параметра.

2. . Емпіричні функції розподілу та характеристики вибірок

Вигляд невідомої теоретичної функції розподілу випадкових величин і її характеристики можна знайти за допомогою емпіричної функції розподілу вибірки, яку будують у припущенні однакової ймовірності P_i появи кожного з елементів вибірки $x_1, \dots, x_n$:

$$P_i \equiv P(X = x_i) = P_i^* \equiv \frac{1}{n} \quad (2.3.1)$$

як частоту появи елементів вибірки

$$F_n^*(z) \equiv P(-\infty < x_i \leq z) = \sum_{x_i \leq z} P(X = x_i) = \frac{1}{n} \sum_{k=1}^n \theta(z - x_i) \equiv \frac{v}{n}, \quad (2.3.2)$$

де v – кількість елементів зі значеннями $x_i \leq z$, $\theta(z - x_i)$ – функція-сходінка.

Часто випробування x_i розміщують за їх зростанням, тобто спершу знаходять $x_1^* = \{x_i\}$, а далі наступні за величиною значення. Послідовність неспадних величин $x_1^*, x_2^*, \dots, x_n^*$, отримана після упорядкування вибірки, називається варіаційним рядом послідовності $x_1, x_2, \dots, x_n$, а величина x_i^* – варіантом або елементом варіаційного ряду. Очевидно, що за допомогою варіаційного ряду емпіричну функцію розподілу можна подати у вигляді

$$\begin{aligned} F_n^*(z) &= \frac{\text{кількість індексів } k \text{ для яких } x_k^* \leq z}{n} = \\ &= \begin{cases} 0, & \text{при } z < x_1^*, \\ \frac{(k-1)}{n}, & \text{при } x_{k-1}^* \leq z < x_k^*, \\ 1, & \text{при } z \geq x_n^*. \end{cases} \end{aligned} \quad (2.3.3)$$

Тобто $F_n^*(z)$ є функцією у вигляді сходінки зі стрибками, які дорівнюють $1/n$ у точках варіаційного ряду. Якщо ж l точок варіаційного ряду збігаються і дорівнюють x_i^* , то стрибок у точці x_i^* дорівнює l/n .

З визначень (2.3.2), (2.3.3) функції $F_n^*(z)$ випливає, що вона дійсно має всі властивості функції розподілу, а саме:

1. Її значення лежать в інтервалі $[0, 1]$

2. $F_n^*(z)$ – неспадна функція

3. Якщо x_1^* – найменший варіант, то $F_n^*(z) = 0$ при $z < x_1^*$, і якщо x_k^* – найбільший варіант, то $F_n^*(z) = 1$ при $z > x_k^*$.

Різниця між теоретичною функцією розподілу $F(z)$ та емпіричною функцією розподілу $F_n^*(z)$ генеральної сукупності полягає в тому, що теоретична функція визначає ймовірність події, а емпірична функція визначає частоту появи цієї ж події. З теореми Бернуллі виходить, що при $n \rightarrow \infty$ вибіркова функція розподілу $F_n^*(z)$ прямує за ймовірністю до теоретичної

$$F_n^*(z) \xrightarrow{P} F(z). \quad (2.3.4)$$

Звідси випливає доцільність використання емпіричної функції розподілу для наближеного подання теоретичної (інтегральної) функції розподілу генеральної сукупності. Якщо обсяг вибірки достатньо великий, то значення емпіричної функції розподілу в кожній точці x є наближеним значенням (тобто оцінкою) теоретичної функції розподілу в цій точці. Згідно з (2.3.4) $F_n^*(z)$ є спроможною оцінкою $F(z)$. Функцію $F_n^*(z)$ часто називають статистичним аналогом для $F(z)$. Існують різні критерії визначення близькості функцій $F_n^*(z)$ і $F(z)$ та міри їх відхилення при скінченному обсязі вибірки.

Зокрема, для скінченної вибірки обсягу n можна показати, що з надійністю 95% має місце таке співвідношення (як наслідок теореми Колмогорова)

$$F(z) - \frac{1}{\sqrt{n}} \leq F_n^*(z) \leq F(z) + \frac{1}{\sqrt{n}}, \quad (2.3.5)$$

і тому, наприклад, при $n = 100$ емпірична функція розподілу відрізняється в будь-якій точці z від теоретичної не більш ніж на 0,13. На рис. 2.1 зображено емпіричний розподіл вибірки, що має такі значення варіантів $x_i^* : 2, 5, 10$ і кількості їх появ $v_i = 12, 1, 0$ відповідно. Звідки маємо $n = 12 + 1 + 0 = 13$

$$F_n^*(z) = \begin{cases} 0, & z < 2, \\ 0,2, & 2 \leq z < 5, \\ 0,5, & 5 \leq z < 10, \\ 1, & z \geq 10. \end{cases}$$

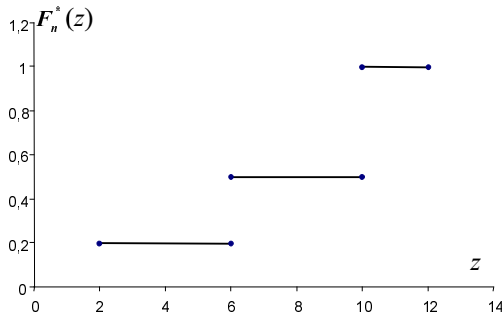


Рис. 2.1. Приклад емпіричної функції розподілу вибірки

У ситуації, коли випадкова величина неперервна або дискретна з дуже малим інтервалом між своїми значеннями, то оцінюють не саму функцію розподілу, а її щільність. Для наочної оцінки щільності часто використовують два способи побудови: 1) гістограм 2) полігона частот.

Для оцінки щільності розподілу елементи вибірки поділяють на групи, утворюючи так звану групувану вибірку. Для цього вибирають найменший $x_{\min}$ та найбільший $x_{\max}$ елементи вибірки і знаходять "розкид" вибірки $\tilde{\Delta}$ як різницю $x_{\min}$ і $x_{\max}$:

$$\tilde{\Delta} = x_{\max} - x_{\min} . \quad (2.3.6)$$

Далі вибирають кількість груп (розрядів) m , на які хочуть розбити вибірку, та визначають ширину інтервалу групи h за формулою

$$h = \frac{\tilde{\Delta}}{m} . \quad (2.3.7)$$

У кожному з отриманих інтервалів з даною шириною h підраховують число елементів n_i ($i = 1, 2, \dots, m$) початкової вибірки – так звану чисельність розряду (кількість елементів групи). Причому ті елементи, які випадково попадають на границю між двома сусідніми інтервалами, ураховуються і в лівому і в правому розрядах, припускаючи, що в кожний із них потрапляє однакова кількість елементів. Потім обчислюють частоту $\tilde{P}_i$ появи випадкової величини у групі:

$$\tilde{P}_i = \frac{n_i}{n}, \quad n = \sum_{i=1}^m n_i . \quad (2.3.8)$$

Кінцевий результат подібного розбиття на групи для зручності подають у вигляді таблиці (див. табл. 2.1). В останньому рядку таблиці наводяться ймовірності появи величини в кожній групі, які обчислюються за формулами

$$P_i \equiv P_T(x_i) = \int_{x_{\min} + (i-1)h}^{x_{\min} + ih} p_T(x) dx, \quad (2.3.9)$$

для деякого теоретичного закону щільності розподілу $p_T(x)$. Величина $\tilde{p}_i \equiv \tilde{P}_i / h$ є вибірковою щільністю і при $n \rightarrow \infty$, тобто при $h \ll 1$, збігається за ймовірністю до $p_T(x)$.

Таблиця 2.1

Характеристики згрупованої вибірки

Характеристики групи	Номер групи 1	2	...	...	M
Границі групи	$x_{\min} \div x_{\min} + h$	$x_{\min} + h \div x_{\min} + 2h$	...	...	$x_{\min} + (m-1)h \div x_{\max}$
Середина i -го інтервалу x_i	$\tilde{x}_1$	$\tilde{x}_2$	...	...	$\tilde{x}_m$
Чисельність групи n_i	n_1	n_2	...	...	n_m
Частота групи $\tilde{P}_i = \frac{\tilde{n}_i}{n}$	$\tilde{P}_1 = \frac{n_1}{n}$	$\tilde{P}_2 = \frac{n_2}{n}$	...	...	$\tilde{P}_m = \frac{n_m}{n}$
Теоретична ймовірність $P_T(x_i)$ появи елементів групи	P_1	P_2	...	...	P

Результати розбиття вибірки на групи подають графічно у вигляді гістограми частот, якою називають фігуру у вигляді сходинок, що складається з прямокутників, основою яких є інтервали довжини h , а висотою – відношення $\tilde{p}_i \equiv \tilde{P}_i / h$. Прямокутники будуються на кожному i -му інтервалі.

Згідно з формулами (2.3.8) та (2.3.9) при достатньо малому h гістограма зображує статистичний аналог щільності розподілу випадкової величини. У методі гістограм невідома щільність розподілу апроксимується кусково-сталю функцією. Якщо щільність $p_T(x)$ – достатньо гладка функція, то вона точніше апроксимується кусково-лінійною функцією. Графічне зображення такого наближення називають полігоном частот. Полігон частот є ломаною лінією, яка будується подібно гістограмі, але замість того, щоб проводити верхні грані прямими лініями, з'єднуються середини всіх верхніх сто-

рін. Побудований таким чином кусково-лінійний графік також є статистичним аналогом теоретичної щільності розподілу випадкових величин.

На рис. 2.2 подано гістограму та полігон частот щільності вибірки обсягу $n = 100$ з $x_{\min} = 0,2$, $x_{\max} = 12,2$ та $m = 6$ (тобто $h = 2$) згруповану згідно з даними табл. 2.2.

Таблиця 2.2
Характеристики гістограм та полігону частот для рис. 2.2.

N	1	2	3	4	5	6
Границі	0,2÷2,2	2,2÷4,2	4,2÷6,2	6,2÷8,2	8,2÷10,2	10,2÷12,2
Середина x_i	1,2	3,2	5,2	7,2	9,2	11,2
Чисельність n_i	70	20	4	3	2	1
Частота $\hat{P}_i \equiv n_i / 100$	0,7	0,2	0,04	0,03	0,02	0,01
Вибіркова щільність $\hat{p}_i \equiv \hat{P}_i / 2$	0,35	0,1	0,02	0,015	0,01	0,005

Для випадкових величин із гладкою щільністю розподілу полігон частот $\tilde{p}(\tilde{x}_i)$ дозволяє оцінювати теоретичну щільність p_T краще ніж гістограма. Зокрема, згідно з теоремою Смірнова визначена з полігону частот щільність $\tilde{p}(\tilde{x}_i)$ (для вибірки обсягу n з груповою довжиною h) знаходиться в такому інтервалі

$$p_T(x) - t_\alpha \sqrt{\frac{p_T(x)}{nh}} \leq \tilde{p}(x) \leq p_T(x) + t_\alpha \sqrt{\frac{p_T(x)}{nh}}, \quad (2.3.10)$$

де однозначно визначене число t_α залежить від обсягу вибірки n , кількості груп m , на яку розбито елементи, та ймовірності α , за якою розглядається близькість $\tilde{p}(x)$. Наприклад, для вибірок з $100 \leq n \leq 1000$, $5 \leq m \leq 12$ і близькості $\tilde{p}(x)$ до $p_T(x)$ з ймовірністю $90 \leq \alpha \leq 99$ значення t_α знаходяться в інтервалі $2, 0 \leq t_\alpha < 2$, зокрема, для $n = 100$, $m = 5$ і $\alpha = 5$ величина $t_\alpha = 2$.

Користуючись теоремою Смірнова (2.3.10), можна легко перевірити узгодження полігону частот для вибірки з табл. 2.2 (рис. 2.2) зі щільністю розподілу вигляду

$$p_T(x) = \begin{cases} 0,5 e^{-\frac{(0,2-x)^2}{2}} & ; 0,2 \leq x \leq 12,2; \\ 0 & ; x > 12,2. \end{cases}$$

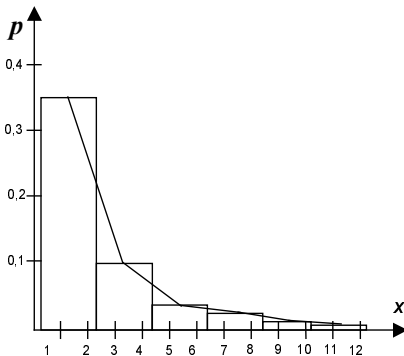


Рис. 2.2. Гістограма та полігон частот вибірки з табл. 2.2

Вибираючи кількість розрядів m при групуванні вибірки, треба користуватися такими міркуваннями. По-перше, кількість груп має бути достатньо великою для того, щоб можна було мати уяву про характер поведінки кривої щільності розподілу. Звичайно вважають, що повинно бути не менше ніж вісім груп. По-друге, слід прагнути, щоб у кожний розряд потрапляла достатньо велика кількість елементів, оскільки при маленьких значеннях n_i випадкові відхилення частот $\tilde{P}$ від імовірностей P_T будуть значні й спотворять результати аналізу, що базується на методах гістограм та полігону частот.

Зазначимо, що багато із сучасних ефективних методів оцінювання щільності невідомої функції розподілу базується на використанні так званих "ядерних оцінок". У цьому випадку як статистичний аналог теоретичної щільності розподілу використовують випадкову функцію вигляду

$$p_n(x) = \frac{1}{na_n} \sum_{i=1}^n k\left(\frac{x-x_i}{a_n}\right) \quad (2.3.11)$$

з деяким вибраним ядром $k(x)$ та набором чисел $a_n > 0$ і x_i . Отримано багато теоретичних результатів, що дозволяють при великому значенні n наближено оцінювати ймовірність відхилень оцінок вигляду (2.3.11) від теоретичної щільності.

Раніше ми бачили (див. 1.4), що функції розподілу та щільності, як правило, залежать від параметрів θ які, у свою чергу, є деякими функціоналами від цих функцій наприклад, у випадку одного параметра така залежність має вигляд

$$\theta \equiv (F(z, \theta)) \equiv \tilde{\sim}(p(x, \theta)). \quad (2.3.12)$$

Різні характеристики систем, такі як математичне сподівання і моменти, також визначаються функціоналами типу (2.3.12). Існує багато різних оцінок параметрів. Вибірковим, або емпіричним значенням параметра θ називають таку оцінку θ_n^* , яка будується заміною у визначенні параметра θ теоретичної функції розподілу F на вибіркову $F_n^*(z)$ (див. (2.3.1), (2.3.2)):

$$\theta_n^* = (F_n^*(z)) \equiv \tilde{(p^*)}, \quad p^*(x) = \sum_{i=1}^n \frac{1}{n} \delta(x - x_i). \quad (2.3.13)$$

Тут p^* – вибіркова щільність розподілу випадкової величини.

Тому згідно з (2.3.1) і (2.3.2), наприклад, вибіркове значення математичного сподівання $M[x] \equiv \mu_1$ (вибіркове середнє значення ($\bar{x}$)) дорівнює

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (2.3.14)$$

і збігається із середньоарифметичним значенням елементів вибірки. Аналогічно, вибіркове значення дисперсії, тобто вибіркове значення другого початкового моменту $M[(x_i - M[x])^2]$, (вибіркова дисперсія (s^2)) має вигляд

$$s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \langle x \rangle)^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - (\langle x \rangle)^2. \quad (2.3.15)$$

Оскільки як вибіркова функція, так і вибіркові характеристики є функціями випадкових величин, то вони і самі є випадковими величинами. Із закону великих чисел випливає, що оцінки (2.3.14) і (2.3.15) є спроможними оцінками для математичного сподівання ($\bar{x}$) і теоретичної дисперсії (σ^2), відповідно,

$$\begin{aligned} \langle x \rangle &\xrightarrow{P} \bar{x} \equiv M[x], \\ s^2 &\xrightarrow{P} \sigma^2 = D[x] \equiv M[(x_i - M[x])^2]. \end{aligned} \quad (2.3.16)$$

Вибіркове середнє значення $\langle x \rangle$ є також незсуненою оцінкою математичного сподівання $\bar{x}$. Дійсно, оскільки кожна з випадкових величин x_i визначається однаковою функцією розподілу $F(x)$, то $M[x_i] \equiv M[x] \equiv \bar{x}$ і

$$M[\langle x \rangle] = M\left[\frac{1}{n} \sum_{i=1}^n x_i\right] = \frac{1}{n} \sum_{i=1}^n M[x_i] = M[x] \equiv \bar{x}. \quad (2.3.17)$$

Однак виявляється, що вибіркова дисперсія s^2 є тільки зсуненою оцінкою для теоретичної дисперсії σ^2 . Перевіримо це твердження. Введемо позначення $y_i = x_i - \bar{x}$, тоді

$$M[y_i] = M[x_i] - \bar{x} = 0; \quad M[y_i^2] = M[(x_i - \bar{x})^2] = \sigma^2. \quad (2.3.18)$$

Користуючись цими виразами та враховуючи (2.3.15), маємо

$$\begin{aligned} M[s^2] &= \frac{1}{n} M \left[\sum_{i=1}^n (x_i - \langle x \rangle)^2 \right] = \frac{1}{n} M \left[\sum_{i=1}^n (y_i - \langle y \rangle)^2 \right] = \\ &= \frac{1}{n} M \left[\sum_{i=1}^n y_i^2 \right] - \frac{2}{n} M \left[\sum_{i=1}^n y_i \langle y \rangle \right] + \frac{1}{n} M [n \langle y \rangle^2] = \\ &= \frac{1}{n} \sum_{i=1}^n M[y_i^2] - 2M[\langle y \rangle^2] + M[\langle y \rangle^2] = \sigma^2 - M[\langle y \rangle^2]. \end{aligned} \quad (2.3.19)$$

Оскільки x_i та x_j незалежні при $i \neq j$, то y_i і y_j не корелюють $M[y_i y_j] = 0$ і

$$\begin{aligned} M[\langle y \rangle^2] &= M \left[\left(\frac{y_1 + y_2 + \dots + y_n}{n} \right)^2 \right] = \frac{1}{n^2} M \left[\sum_{i=1}^n y_i^2 + 2 \sum_{i < j} y_i y_j \right] = \\ &= \frac{1}{n^2} \sum_{i=1}^n M[y_i^2] = \frac{1}{n^2} n \sigma^2 = \frac{\sigma^2}{n}. \end{aligned} \quad (2.3.20)$$

Підставивши цей вираз у (2.3.19), остаточно дістанемо

$$M[s^2] = \sigma^2 - \frac{\sigma^2}{n} = \frac{n-1}{n} \sigma^2. \quad (2.3.21)$$

Отже, $M[s^2] \neq \sigma^2$, тобто s^2 є зсуненою оцінкою для σ^2 . Разом із тим, користуючись формулою (2.3.21), можна дістати незсунену оцінку. Такою оцінкою дисперсії є величина

$$\tilde{s}^2 = \frac{n-1}{n} s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \langle x \rangle)^2. \quad (2.3.22)$$

Справді, за допомогою формули (2.3.21) отримуємо

$$M[\tilde{s}^2] = \frac{n}{n-1} M[s^2] = \sigma^2. \quad (2.3.23)$$

При $n \rightarrow \infty$ обидві величини s^2 і $\tilde{s}^2$ прямують до однакової границі й тому оцінка $\tilde{s}^2$ є також спроможною. Коефіцієнт $\frac{n}{n-1}$ називають поправкою Бесселя.

Якщо для одержання незсуненої оцінки будь-якого параметра обсяг вибірки має бути нескінченним ($n \rightarrow \infty$), то кажуть, що така оцінка є асимптотично незсуненою. Згідно з (2.3.21) вибіркова дисперсія s^2 є асимптотично незсуненою оцінкою теоретичної дисперсії.

Відповідно до визначення (2.3.13) та формул (1.2.2), (1.2.24) вибіркові початковий (v_r^*) та центральний (μ_r^*) моменти будь-якого порядку задаються виразами

$$v_r^* = \frac{1}{n} \sum_{i=1}^n x_i^r, \quad \mu_r^* = \frac{1}{n} \sum_{i=1}^n (x_i - \langle x \rangle)^r. \quad (2.3.24)$$

Усі вибіркові моменти є спроможними оцінками відповідних теоретичних моментів. Вибіркові початкові й центральні моменти мають такі ж самі властивості, як і теоретичні моменти. Так, зокрема, вибірковий перший центральний момент завжди дорівнює нулю

$$\mu_1^* = \frac{1}{n} \sum_{i=1}^n (x_i - \langle x \rangle) = \frac{1}{n} \sum_{i=1}^n x_i - \langle x \rangle = 0. \quad (2.3.25)$$

Також зберігаються всі співвідношення між центральними та початковими моментами, наприклад,

$$\mu_2^* = \frac{1}{n} \sum_{i=1}^n (x_i - \langle x \rangle)^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - 2\langle x \rangle \frac{1}{n} \sum_{i=1}^n x_i + \langle \bar{x} \rangle^2 = v_2^* - \langle v_1^* \rangle^2. \quad (2.3.26)$$

Усі вибіркові центральні моменти μ_r^* (2.3.24) є тільки зсуненими оцінками теоретичних центральних моментів μ_r . Незсунені оцінки $\tilde{\mu}_r^*$ для μ_r мають такий вигляд:

$$\tilde{\mu}_r^* = \frac{n^{r-1}}{(n-1)(n-2)\dots(n-r+1)} \mu_r^*. \quad (2.3.27)$$

Виявляється, що для вибірок обсягу $n \leq r$ отримати незсунену оцінку для $(r+1)$ -го моменту взагалі не можливо.

Відзначимо, що середньовибіркове значення $\langle y \rangle$ випадкових величин y_i з формул (2.3.18) є відхиленням середньовибіркового значення $\langle x \rangle$ від його математичного сподівання $M[x] \equiv \bar{x}$:

$$\langle y \rangle \equiv \frac{1}{n} \sum_{i=1}^n y_i = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}) = \langle x \rangle - \bar{x}. \quad (2.3.28)$$

Тому співвідношення (2.3.20) визначає квадрат стандартного відхилення (дисперсію) $\sigma_{\langle x \rangle}^2 \equiv D[\langle x \rangle] = M[(\langle x \rangle - \bar{x})^2]$ середньовибіркового (середньоарифметичного) значення для вибірки з n незалежних випадкових величин

$$\sigma_{\langle x \rangle}^2 = M[(\langle x \rangle - \bar{x})^2] = M[\langle y \rangle^2] = \frac{\sigma^2}{n}, \quad (2.3.29)$$

або

$$\sigma_{\langle x \rangle} \equiv \sqrt{D[\langle x \rangle]} = \frac{\sigma}{\sqrt{n}}, \quad \sigma \equiv \sqrt{D[x_i]}. \quad (2.3.30)$$

Величину $\sigma_{\langle x \rangle}$ у теорії похибок зазвичай називають середньоквадратичною похибкою середньоарифметичного значення. Формула (2.3.30) означає, що середньоарифметична похибка середньоарифметичного значення суми незалежних величин, що вимірюються, дорівнює середньоквадратичній похибці окремого виміру поділеного на квадратний корінь з кількості вимірів.

Зазначимо, що обчислення середньоарифметичного значення за формулою (2.3.14) займає мало часу, і тому є ефективним, лише за так званої - обробці результатів вимірювань. У цьому випадку спочатку повністю виконується запланований експеримент, отримуються всі експериментальні дані й обробляються вже після експерименту. При - обробці даних мож-

на успішно використовувати вираз $\langle x \rangle = \sum_{i=1}^n x_i / n$, для обчислення якого

необхідно буде виконати $n+1$ арифметичну операцію. Однак дуже часто необхідно обробляти результати вимірювань паралельно з їх отриманням (- обробка), зокрема, якщо результати обробки використовуються для безперервного керування експериментом з метою його корекції. При - обробці дані надходять безперервно для загальної кількості спостережень n

для обчислення середнього значення $\langle x \rangle^{(n)} = \sum_{i=1}^n x_i / n$ за n кроків необ-

хідно виконати $(n+3)n/2$ арифметичні операції. Не важко перевірити, що проста видозміна (2.3.14), а саме використання еквівалентного виразу

$$\langle x \rangle^{(n)} = \langle x \rangle^{(n-1)} + \frac{1}{n} (x_n - \langle x \rangle^{(n-1)}), \quad (2.3.31)$$

де $\langle x \rangle^{(n)} = \sum_{i=1}^n x_i / n$ – середньовибіркове значення за n кроків, зменшує

(для $n > 3$) кількість операцій для обчислення $\langle x \rangle^{(n)}$ до $3n$, що є суттєвим при обробці великих обсягів інформації.

4. Довірчі інтервали оцінок

Вибіркові оцінки θ^* характеристик розподілу деякої величини θ є випадковими і можуть приймати різні значення. Тому ніколи не можна категорично стверджувати, що значення θ^* може задовольняти нерівність $|\theta - \theta^*| < \varepsilon$, де

$\varepsilon > 0$ деяке додатне число. Можна лише казати про моменти відхилення θ^* від θ та про ймовірність β , з якою ця нерівність виконується. З погляду теорії ймовірностей достовірність отриманих оцінок параметра θ визначається ймовірністю $\beta \in [0; 1]$,

$$\beta = P(|\theta - \theta^*| \leq \varepsilon) = 1 - \alpha, \quad (2.4.1)$$

з якою оцінка може знаходитися в інтервалі

$$\theta - \varepsilon \leq \theta^* \leq \theta + \varepsilon, \quad \varepsilon > 0. \quad (2.4.2)$$

Проміжок $(\theta^* - \varepsilon, \theta^* + \varepsilon)$ називають 100β $100 (1 - \alpha)$ -м довірчим інтервалом (інтервалом надійності). Величина $\varepsilon \equiv \varepsilon(\beta)$ є півшириною довірчого інтервалу, а $\theta - \varepsilon, \theta + \varepsilon$ – його нижня та верхня межі. Ймовірність β називається довірчою ймовірністю, що означає, з якою ймовірністю інтервал $\theta^* - \varepsilon, \theta^* + \varepsilon$ може містити (кажуть накрити) значення параметра θ . Значення $\beta = 1 - \alpha$ також називають рівнем надійності, або достовірності. Часто замість запису у вигляді довірчого інтервалу (2.4.2) використовується запис

$$\theta^* \pm \varepsilon,$$

де $\varepsilon \equiv \varepsilon(\beta)$ називається похибкою визначення параметра θ з рівнем надійності (з імовірністю) β .

Оскільки величина θ^* – випадкова, то довірчий інтервал також є випадковою величиною, і його значення залежить від закону розподілу θ^* . Побудова довірчих інтервалів для параметрів розподілу називається інтервальними оцінками цих параметрів.

Довірчу ймовірність (2.4.1) можна зв'язати з функцією розподілу $F_{\theta^*}(\xi)$ та щільністю розподілу p_{θ^*} оцінки θ^* , що є випадковою величиною. Згідно із загальними положеннями теорії ймовірностей (див. розділ 1) маємо

$$\begin{aligned} P(\theta - \varepsilon \leq \theta^* \leq \theta + \varepsilon) &= F_{\theta^*}(\theta + \varepsilon) - F_{\theta^*}(\theta - \varepsilon) = \\ &= \int_{-\infty}^{\theta + \varepsilon} p_{\theta^*}(x) dx - \int_{-\infty}^{\theta - \varepsilon} p_{\theta^*}(x) dx = \int_{\theta - \varepsilon}^{\theta + \varepsilon} p_{\theta^*}(x) dx. \end{aligned} \quad (2.4.3)$$

Тому довірна ймовірність (2.4.1), довірчий інтервал (2.4.2) і щільність розподілу оцінки зв'язані співвідношенням

$$\int_{\theta - \varepsilon}^{\theta + \varepsilon} p_{\theta^*}(x) dx = \beta = 1 - \alpha, \quad (2.4.4)$$

яке і визначає півширину ε інтервалу надійності при заданій довірчій ймовірності β .

Таким чином, для побудови довірчих інтервалів спершу треба знайти закон розподілу параметра θ^* і далі визначити півширину інтервалу як розв'язок нелінійного рівняння (2.4.4).

Необхідно зазначити, що ймовірність $\beta = P(|\theta^* - \theta| \leq \varepsilon)$ іноді невірно трактують (і тому записують) як ймовірність $P(\theta^* - \varepsilon \leq \theta \leq \theta^* + \varepsilon)$ того, що значення θ знаходиться в інтервалі $\theta^* - \varepsilon \leq \theta \leq \theta^* + \varepsilon$. У дійсності ж величина θ є деяким детермінованим і фіксованим числом, і тому вона або лежить у зазначеному інтервалі, або не лежить, тобто ймовірність $P(\theta^* - \varepsilon \leq \theta \leq \theta^* + \varepsilon)$ дорівнює або нулю або одиниці. Довірчий інтервал $|\theta^* - \theta| \leq \varepsilon$ має такий зміст: якщо здійснювати багато серій випробувань, то приблизно у $100\beta\%$ випадків одержані інтервали будуть містити θ і лише у $100(1 - \beta)\% \equiv 100\alpha\%$ випадків значення θ і знаходиться за межами довірчого інтервалу.

Знайдемо довірчий інтервал $\mu^* - \varepsilon, \mu^* + \varepsilon$ за оцінки математичного сподівання $\mu = M[x]$ вибіркоvim середнім $\mu^* = \sum_{i=1}^n x_i / n$ з елементами вибірки

x_i що розподілені за нормальним законом із відомою дисперсією σ^2 . Згідно із загальними правилами побудови законів розподілів сума нормально розподілених величин

$$\mu^* = \frac{1}{n} \sum_{i=1}^n x_i \quad (2.4.5)$$

розподілена за нормальним законом зі щільністю

$$p_{\mu^*}(x) = \frac{1}{\sqrt{2\pi D[\mu^*]}} e^{-\frac{(x - M[\mu^*])^2}{2D[\mu^*]}}, \quad (2.4.6)$$

що визначається двома параметрами – математичним сподіванням $M[\mu^*]$

та її дисперсією $D[\mu^*] \equiv M[(\mu^* - M[\mu^*])^2]$. Обчислимо ці параметри:

$$\begin{aligned} M[\mu^*] &= \frac{1}{n} \sum_{i=1}^n M[x_i] = \mu, \\ D[\mu^*] &= \frac{1}{n^2} \sum_{i=1}^n D[x_i] = \frac{\sigma^2}{n} \equiv \sigma_{\mu^*}^2. \end{aligned} \quad (2.4.7)$$

Звідки знаходимо, що щільність розподілу (2.4.7) випадкової величини μ^* має вигляд

$$p_{\mu^*}(x) = \frac{\sqrt{n}}{\sqrt{2\pi}\sigma} e^{-\frac{(x - \mu)^2 n}{2\sigma^2}}. \quad (2.4.8)$$

У результаті співвідношення (2.4.4) для визначення значення половини довірчого інтервалу ε приймає вигляд

$$\begin{aligned} \beta &= P(\mu - \varepsilon \leq \mu^* \leq \mu + \varepsilon) = \frac{1}{\sqrt{2\pi}\sigma / \sqrt{n}} \int_{\mu - \varepsilon}^{\mu + \varepsilon} e^{-\frac{(x - \mu)^2}{2\sigma^2/n}} dx = \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\varepsilon\sqrt{n}/\sigma}^{\varepsilon\sqrt{n}/\sigma} e^{-\frac{u^2}{2}} du = 1 - 2F\left(-\frac{\varepsilon}{\sigma}\sqrt{n}\right) = \Phi\left(\frac{\varepsilon}{\sigma}\sqrt{n}\right), \end{aligned} \quad (2.4.9)$$

де була використана заміна змінних $U = (x - \mu) \sqrt{n} / \sigma$. Тут $F(z) \equiv [\Phi(z) + 1] / 2$ – функція розподілу стандартної нормальної величини, а $\Phi(z)$ – функція Лапласа (1.3.13):

$$\Phi(z) = \frac{2}{\sqrt{2\pi}} \int_0^z e^{-\frac{t^2}{2}} dt.$$

Таким чином, остаточно знаходимо співвідношення між довірчим інтервалом ε і довірчою ймовірністю β

$$\Phi\left(\frac{\varepsilon}{\sigma/\sqrt{n}} \equiv \frac{\varepsilon}{\sigma_{\mu}^*} = t\beta\right) = \beta. \quad (2.4.10)$$

Розв'язки цього рівняння вже розглядалися раніше у 1.3. Зокрема, значенням інтервалів

$$\varepsilon = \frac{\sigma}{\sqrt{n}}, \frac{2\sigma}{\sqrt{n}}, \frac{3\sigma}{\sqrt{n}}, \frac{4\sigma}{\sqrt{n}} \Rightarrow t_{\beta} = 1, 2, 3, 4 \quad (2.4.11)$$

відповідають довірчі ймовірності

$$\beta = 0,683; 0,954; 0,9973; 0,9999373. \quad (2.4.12)$$

Як вже відзначалося в 1.3, зазвичай вважають, що ймовірність 0,9973 практично дорівнює одиниці, тому похибка визначення μ не може вийти за інтервал $3\sigma_{\mu}^*$, який відповідає цій ймовірності. Тобто, як правило, результати вимірювань подають у вигляді

$$\mu = \mu^* \pm 3\sigma_{\mu}^* \equiv \mu^* \pm 3 \frac{\sigma}{\sqrt{n}} \quad (2.4.13)$$

і кажуть, що похибка "не може вийти за значення трьох сигм". Незважаючи на те, що ймовірність виконання нерівності $|\mu - \mu^*| \leq 3\sigma_{\mu}^*$ велика, не слід забувати, що, по-перше, ця ймовірність отримана для нормального закону розподілу і, по-друге, що співвідношення (2.4.13) отримано достатньо штучно. Якщо взяти $\beta = 0,999937$ (що відповідає відміні від одиниці довірчій ймовірності на величину 0,000063 замість 0,0027 у випадку $\beta = 0,9973$), то отримаємо вже не "правило 3σ " а "правило 4σ " і ширина довірчого інтервалу зросте приблизно на 30%. Крім того, довірчий інтервал (2.4.13) було отримано у припущенні, що дисперсія σ^2 розподілу є відомою, а це, як правило, не так.

Інтервал надійності оцінки математичного сподівання, як середньоарифметичного значення, у випадку будь-якого розподілу при великих обсягах (n) вибірок легко знайти за допомогою центральної граничної теореми. Згідно з цією теоремою розподіл середньовибіркових значень $\mu^* = \sum_i^n x_i / n$ незалежних випадкових величин при $n \gg 1$ є нормальним незалежно від конкретного виду розподілу величин x_i . Тому якщо дисперсія σ^2 відома, то згідно з (2.4.10) довірчий інтервал обчислення математичного сподівання як середнього значення за вибіркою дорівнює

$$[\mu^* - \frac{\sigma}{\sqrt{n}} t_\beta, \mu^* + \frac{\sigma}{\sqrt{n}} t_\beta], \quad (2.4.14)$$

де константа t_β визначається довірчою ймовірністю β і є розв'язком нелінійного рівняння

$$\Phi(t_\beta) = \beta. \quad (2.4.15)$$

Як правило, значення дисперсії σ^2 невідоме, тому його часто замінюють незсунутою оцінкою (2.3.22): $\tilde{s}^2 = \sum_{i=1}^n \frac{(x_i - \mu^*)^2}{n-1}$. При цьому для рівня надійності $\beta = \Phi(t_\beta)$ використовують інтервал

$$\left[\mu^* - \frac{\tilde{s}}{\sqrt{n}} t_\beta, \mu^* + \frac{\tilde{s}}{\sqrt{n}} t_\beta \right]. \quad (2.4.16)$$

Але якщо обсяг вибірки недостатньо великий ($n < 30$), то такий інтервал надійності не заслуговує значної довіри тому, по-перше, що розподіл μ^* є лише асимптотично нормальним, а, по-друге, що було зроблено заміну невідомої дисперсії на її оцінку $\tilde{s}^2$.

Детальні обчислення показують, що у випадку, коли величини $\{x_i\}$ є вибіркою величини, яка розподілена за нормальним законом із невідомою дисперсією, то довірчий інтервал для математичного сподівання μ при довірчій ймовірності β задається формулою

$$\left[\mu^* - \frac{\tilde{s}}{\sqrt{n}} t_\beta(n), \mu^* + \frac{\tilde{s}}{\sqrt{n}} t_\beta(n) \right]. \quad (2.4.17)$$

Тут значення $t_\beta(n)$ знаходиться з рівняння

$$\beta = \int_{-t_{\beta}(n)}^{t_{\beta}(n)} S_{n-1}(x) dx, \quad (2.4.18)$$

де функція

$$S_{n-1}(x) = \frac{\Gamma(\frac{n}{2})}{\sqrt{\pi(n-1)}\Gamma(\frac{n-1}{2})} \left(1 - \frac{x^2}{n-1}\right)^{-\frac{n}{2}} \quad (2.4.19)$$

є розподілом Стюдента (1.4.31) з $(n-1)$ -м ступенем свободи. Наприклад, при $\beta = 0,99$ і $n = 10 \div 30$ значення $t_{\beta}(n)$ знаходяться в інтервалі $3,25 \div 2,75$ відповідно (у цьому випадку для нормального розподілу $t_{\beta} \approx 2,75$). При $n \geq 30$ розподіл Стюдента (2.4.19) практично не відрізняється від нормального з параметрами $\mu = 0$ і $\sigma = 1$, тому в цьому випадку можна використовувати довірчий інтервал (2.4.16) з t_{β} згідно з (2.4.15).

Як приклад використання співвідношення (2.4.17) визначимо довірчий інтервал довжини відрізка, який відповідає довірчій ймовірності $\beta = 0,99$, якщо внаслідок шести вимірів середньоарифметичне значення вимірів є $\mu^* = 131,75$ мм, а оцінка дисперсії похибок – $\hat{s}^2 = 0,625$ мм².

З формули (2.4.17) маємо для довірчого інтервалу

$$\varepsilon = \frac{\tilde{s}}{\sqrt{n}} t_{\beta}(n), \quad (2.4.20)$$

де $t_{\beta}(n)$ визначається за таблицями функції розподілу Стюдента. Для числа ступенів свободи $k = n - 1 = 5$ і довірчої ймовірності $\beta = 0,99$ з таблиць маємо

$$t_{\beta}(n) = 4,03, \text{ звідки отримуємо } \varepsilon = \frac{0,625 \cdot 4,03}{\sqrt{6}} \approx 1,03 \text{ мм. Якщо в цьому ви-}$$

падку ми використовували для оцінки ε нормальний розподіл, тобто використали (2.4.14), (2.4.15) з $t_{\beta}(n) = t_{\beta} \approx 2,8$ у (2.4.20), то знайдене значення було б приблизно в 1,5 рази меншим. Це явище збільшення ширини інтервалу надійності для малих обсягів вибірок завжди має місце при достатньо великих довірчих ймовірностях і пов'язано з тим, що щільність розподілу Стюдента із зростанням аргументу значно повільніше наближується до нуля ніж у випадку нормального розподілу. Тобто "правило 3σ " не виконується для вибірок малого обсягу з $n < 30$ і невідомою дисперсією.

Розділ

ОЦІНКИ ПАРАМЕТРІВ ФУНКЦІЇ РОЗПОДІЛУ

.1. Класифікація оцінок за методами їх знаходження

Методи знаходження оцінок характеристик досліджуваних величин можна поділити на дві великі групи. У першій групі методів оперують із функціями розподілу величин, що спостерігаються, причому загальний вигляд функції розподілу вважається відомим, а невідомими є тільки деякі числові параметри, від яких вона залежить. Серед цих методів найвідомішими є такі методи: максимальної правдоподібності; моментів; байєсовський метод знаходження оцінок; умовних математичних сподівань. Як часткові випадки цих методів (за певних припущень відносно функції розподілу спостережуваних величин) виступають такі методи: найменших квадратів (хоча він має і самостійне формулювання); найменших додатно-визначених форм, що узагальнює метод найменших квадратів; найменших модулів тощо. Виявляється, що існують оцінки, які мало змінюються за зміни виду функції розподілу. Такі оцінки величин називаються стійкими, або робастними.

При використанні другої групи методів знаходження оцінок інформація про функції розподілу не потрібна. Тут в основі кожного з методів лежить формальна обчислювальна схема, яка використовується при отриманні оцінок. Таким чином можна створити значну кількість методів. Наприклад, з такого погляду можна розглядати методи найменших квадратів, найменших модулів, мінімакса та ін. Якщо зробити деякі припущення про характер функцій розподілу, то майже кожен метод другої групи можна розглядати як наслідок або окремий випадок методів першої групи, і тоді ці методи отримують теоретичне (з погляду теорії ймовірностей) обґрунтування. Якщо ж не робити ніяких припущень, то ці методи залишаються досить формальними, і завжди може виникнути питання про точність отриманих з їх допомогою результатів, їхню поведінку в умовах, коли похибки вимірів підпорядковуються тому чи іншому закону розподілу ймовірностей.

.2. Метод моментів

У 2.3 зазначалося, що зі збільшенням числа спостережень емпіричні (вибіркові) функції розподілу як заведено мало відрізняються від теоретичних функцій розподілу, і ця різниця прямує до нуля з ймовірністю як заведено близькою до одиниці. Тому за значної кількості спостережень вибіркові й теоретичні моменти

будуть близькими один до одного. Зокрема, як ми бачили у 2.3, початкові й центральні вибіркові моменти є спроможними оцінками відповідних початкових і центральних теоретичних моментів. На цьому і базується метод моментів, що був запропонований Пірсоном для визначення параметрів функції розподілу. Цей метод оцінки невідомих параметрів заданого розподілу полягає в заміні теоретичних моментів розподілу, що досліджується, на відповідні емпіричні моменти.

А саме, у методі моментів Пірсона вважається відомим загальний вигляд функції розподілу випадкових величин, і для даного закону розподілу спершу обчислюють загальні співвідношення для початкових теоретичних моментів

$v_r = \overline{x^r}$, тобто знаходять їх вирази через невідомі параметри $\{\theta_j\}$ розподілу у вигляді співвідношень

$$v_r = (\{\theta_j\}, v_k), \quad (3.2.1)$$

а далі замінюють моменти v_r на їх вибіркові оцінки $v_r^* = \sum_n x^r / n$ і шукають розв'язки $\{\theta_j^*\}$ отриманих рівнянь

$$v_r^* = (\{\theta_j^*\}, v_k^*) \quad (3.2.2)$$

і тим самим знаходять оцінки параметрів $\{\theta_j\}$. Деякі з теоретичних моментів можуть не залежати від невідомих параметрів. У такому випадку до отриманих рівнянь додають співвідношення, що знаходять із визначення початкових моментів більш високого порядку, так щоб кількість рівнянь дорівнювала кількості невідомих параметрів.

Розглянемо приклади знаходження оцінок методом моментів. Спершу обчислимо оцінки v_1^* , s^2 параметрів та σ^2 нормального розподілу $N(\mu, \sigma^2)$ зі щільністю (1.4.7):

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

Для цього розподілу початкові моменти $v_1 = \bar{x}$ та $v_2 = \overline{x^2}$ дорівнюють (див. (1.2.27), (1.4.10))

$$v_1 = \mu, \quad v_2 = \mu^2 + \sigma^2. \quad (3.2.3)$$

Тому, згідно з методом моментів оцінки v_1^* , s^2 параметрів μ , σ є розв'язками системи рівнянь

$$v_1^* = \mu, \quad (v_2^*)^2 + s^2 = v_2^*, \quad (3.2.4)$$

де

$$v_1^* = \frac{1}{n} \sum_i x_i, \quad v_2^* = \frac{1}{n} \sum_{i=1}^n x_i^2.$$

З рівнянь (3.2.4) маємо

$$\begin{aligned} v_1^* &= v_1^* \equiv \frac{1}{n} \sum_i x_i, \\ s^2 &= v_2^* - (v_1^*)^2 = \frac{1}{n} \sum_i x_i^2 - (v_1^*)^2 = \\ &= \frac{1}{n} \sum_i \left[(x_i - v_1^*) + v_1^* \right]^2 - (v_1^*)^2 = \frac{1}{n} \sum_i (x_i - v_1^*)^2. \end{aligned} \quad (3.2.5)$$

Таким чином, оцінки параметрів v_1^* і s^2 збігаються з вибірковими середнім значенням та середньоквадратичним відхиленням s^2 . Остання величина є зсуненою оцінкою теоретичного значення середньоквадратичного відхилення (див. 2.3).

Тепер знайдемо оцінку θ^* параметра θ розподілу Релея, який за визначенням має щільність

$$p(x) = \frac{x}{\theta} e^{-\frac{x^2}{2\theta}}, \quad x \geq 0. \quad (3.2.6)$$

З такою щільністю розподілена, наприклад, випадкова величина, що є відношенням випадкових значень відстані $\sqrt{x^2 + y^2}$ точки у двовимірному просторі x, y з однаково нормально розподіленими значеннями координат x, y до математичного сподівання значення цієї відстані. Легко побачити, що перший та другий початкові моменти для цього розподілу зв'язані з параметром θ співвідношеннями

$$s_1 = \sqrt{\theta \frac{\pi}{2}}, \quad s_2 = 2\theta. \quad (3.2.7)$$

Користуючись методом моментів, маємо з першого рівняння формули (3.2.7) таку оцінку:

$$\theta^* = \frac{2}{\pi} (s_1^*)^2 = \frac{2}{\pi} \left(\frac{1}{n} \sum_i x_i \right)^2, \quad (3.2.8)$$

а з другого рівняння (3.2.7) отримуємо

$$\theta^* = \frac{s_2^*}{2} = \frac{1}{2n} \sum_i x_i^2. \quad (3.2.9)$$

Тобто маємо дві різні оцінки для одного і того ж самого параметра.

Цей приклад демонструє невизначеність методу моментів, оскільки в загальному випадку його використання може привести до різних результатів.

1.1. Метод максимальної правдоподібності

Загальним і важливим як у теоретичному відношенні, так і в застосуваннях методом побудови точкових оцінок є метод максимальної правдоподібності, розроблений Фішером. Ідея метода така. Розглянемо вибірку $\vec{x} = (x_1, x_2, \dots, x_n)$, яка отримана з генеральної сукупності, що розподілена за законом $p(\vec{x}, \vec{\theta})$ з деяким набором параметрів $\vec{\theta} = \{\theta_1, \dots, \theta_m\}$. Якщо випадкова величина є неперервною, то $p(\vec{x}, \vec{\theta})$ є щільністю її розподілу, а якщо випадкова величина є дискретною, то під $p(\vec{x}, \vec{\theta})$ необхідно розуміти ймовірність. За різних значень параметрів $\vec{\theta}$ маємо різні закони $p(\vec{x}, \vec{\theta})$ розподілу випадкової величини. Згідно з методом максимальної правдоподібності вважається, що дана вибірка $\vec{x}$ була отримана в силу того, що вона найбільш імовірна, і тому задача зводиться до побудови розподілу $p(\vec{x}, \vec{\theta})$ з такими значеннями параметрів $\vec{\theta}$, для яких ймовірність появи даної вибірки $\vec{x} = (x_1, x_2, x_3, \dots, x_n)$ буде максимально можливою. Таким чином, виникає задача: при заданій вибірці $\vec{x} = (x_1, x_2, x_3, \dots, x_n)$ визначити такі параметри $\vec{\theta}$, щоб сумісна щільність розподілу $p(\vec{x}, \vec{\theta})$ (у неперервному випадку) або сумісна ймовірність $P(\vec{x}, \vec{\theta})$ (у дискретному випадку) були максимальними.

Сумісний закон розподілу $p(\vec{x}, \vec{\theta})$, коли він розглядається лише як функція параметрів $\vec{\theta}$ при фіксованих значеннях вибірки $\vec{x} = (x_1, x_2, x_3, \dots, x_n)$, було названо Фішером функцією правдоподібності:

$$L(\vec{\theta}) = p(\vec{x}, \vec{\theta}). \quad (3.3.1)$$

Метод максимальної правдоподібності полягає в тому, що за оцінку параметрів $\vec{\theta}$ вибирається точка $\vec{\theta}^* \equiv \{\theta_j^*\}$, у якій функція правдоподібності досягає найбільшого значення. Інакше кажучи, оцінкою параметра θ вважається розв'язок рівняння

$$L(\vec{\theta}^*) = \max_{\vec{\theta}' \in \Omega} L(\vec{\theta}'), \quad (3.3.2)$$

якщо такий розв'язок існує, де Ω – замкнута множина всіх можливих значень $\vec{\theta}$.

Функції $L(\vec{\theta})$ та $l(\vec{\theta})$ досягають найбільшого значення в одній і тій самій точці, а обчислювати точку, у якій функція $l(\vec{\theta})$ досягає найбільшого значення, часто зручніше, саме тому цю функцію і використовують. Логарифм від функції правдоподібності $l(\vec{\theta})$ також називається логарифмічною функцією правдоподібності. Якщо функція $L(\vec{\theta})$ диференційована за $\{\theta_j\}$, то для того, щоб розв'язати рівняння (3.3.2), потрібно знайти стаціонарні точки $l(\vec{\theta})$ як розв'язки рівнянь, що відповідають необхідним умовам екстремуму,

$$\frac{\partial}{\partial \theta_i} l(\{\theta_k\}) \Big|_{\vec{\theta}=\vec{\theta}^*} = 0, \quad i = 1, \dots, m. \quad (3.3.3)$$

Система нелінійних рівнянь (3.3.3) називається рівняннями правдоподібності. Корні цієї системи $\vec{\theta}^*$ можуть відповідати максимуму, мінімуму функції $l(\vec{\theta})$ або бути точками перегину. Тому також необхідно перевірити, чи відповідають отримані розв'язки максимуму функції правдоподібності. Якщо система (3.3.3) не має розв'язків в області Ω , то це може означати, що розв'язки $\vec{\theta}^*$ знаходяться на межі області припустимих значень.

Оскільки вектор $\vec{x}$ є вибіркою і компоненти x_i є незалежними однаково розподіленими випадковими величинами, то $L(\vec{\theta})$ є сумісним законом розподілу вигляду (2.1.1), а саме:

$$L(\vec{\theta}) = p(\vec{x}, \vec{\theta}) = p(x_1, \vec{\theta}) p(x_2, \vec{\theta}) \dots p(x_n, \vec{\theta}), \quad (3.3.4)$$

$$l(\vec{\theta}) = \sum_{i=1}^n l(x_i, \vec{\theta}).$$

Метод максимальної правдоподібності, як правило, потребує значно складніших обчислень порівняно з методом моментів і в деяких випадках, наприклад для нормального розподілу $N(\mu, \sigma^2)$ та розподілу Пуассона, приводить до таких

самих результатів, що і метод моментів, але у загальному випадку оцінки, отримані методом максимальної правдоподібності, мають ряд, принаймні теоретичних, переваг порівняно з оцінками за іншими методами. Якщо закони розподілу випадкової величини мають деякі регулярні властивості, наприклад такі, як у випадку використання теореми Крамера – Рао (див. 2.2), то оцінки $\vec{\theta}^*$ метода максимальної правдоподібності мають такі властивості:

1. Розв'язок $\vec{\theta}^*$ рівнянь правдоподібності (3.3.3) з $1/L$ за (3.3.4) є однозначним.

2. Отримана оцінка $\vec{\theta}^* \equiv \vec{\theta}_n^*$ з використанням вибірки обсягу n є спроможною ($\vec{\theta}_n^* \xrightarrow{P} \vec{\theta}$), хоч і може бути зсуненою.

3. Оцінка максимальної правдоподібності є асимптотично ефективною.

4. Величина $\vec{\theta}^* \equiv \vec{\theta}_n^*$ асимптотично нормально розподілена, тобто при великих обсягах вибірок її функція розподілу близька до нормальної і має найменшу дисперсію ($\sigma_{\vec{\theta}^*}^2$) порівняно з іншими асимптотичними оцінками.

Згідно з теоремою Крамера – Рао дисперсія обернено пропорційна кількості інформації за Фішером, і відповідно до (2.2.8) у випадку одного параметра його дисперсія ($\sigma_{\vec{\theta}^*}^2$) дорівнює

$$\sigma_{\vec{\theta}^*}^2 = \frac{1}{M \left[\left(\frac{d1}{d\theta} \right)^2 \right] \Big|_{\vec{\theta}=\vec{\theta}^*}} = - \frac{1}{M \left[\frac{d^2 1}{d\theta^2} \right] \Big|_{\vec{\theta}=\vec{\theta}^*}}. \quad (3.3.5)$$

Якщо в логарифмічній функції правдоподібності можна відокремити (факторизувати) залежність від θ , то отримаємо

$$\sigma_{\vec{\theta}^*}^2 = - \frac{1}{\frac{d^2 1}{d\theta^2} \Big|_{\vec{\theta}=\vec{\theta}^*}}, \quad (3.3.6)$$

і у виразі (3.3.5) зникає необхідність явного обчислення математичного сподівання. Коли формулою (3.3.6) користуватися не можна, то дисперсія оцінки обчислюється згідно з (3.3.5) числовим способом. Зауважимо, що співвідношення (3.3.6) виконується для нормального, біноміального розподілів та розподілу Пуассона.

У багатовимірному випадку методом максимальної правдоподібності можна обчислити як дисперсії оцінок, так і кореляційні моменти параметрів: $D_{ij} \equiv M[(\theta_i^* - \theta_i)(\theta_j^* - \theta_j)]$. Для спрощення запису відповідних виразів використовують матрицю коваріацій або похибок $D(\vec{\theta})$ (1.2.21) з елементами

$$\|D(\vec{\theta})\|_{ij} = M[(\theta_i^* - \theta_i)(\theta_j^* - \theta_j)]. \quad (3.3.7)$$

Ця матриця є симетричною, а її діагональні елементи дорівнюють дисперсіям σ_i^2 відповідних випадкових величин θ_i :

$$\|D(\vec{\theta})\|_{ii} = \sigma_i^2. \quad (3.3.8)$$

Недіагональні елементи матриці (3.3.7) можна записати через середньоквадратичні відхилення σ_i та коефіцієнти кореляції ρ_{ij} у вигляді

$$\|D\|_{ij} = \sigma_i \sigma_j \rho_{ij}. \quad (3.3.9)$$

Якщо випадкові величини некорельовані, то матриця похибок є діагональною.

Матриця похибок для оцінки за методом максимальної правдоподібності є оберненою до матриці з елементами:

$$\|D^{-1}\|_{ij} = -M \left[\frac{\partial^2 \ln L}{\partial \theta_i \partial \theta_j} \right]_{\vec{\theta} = \vec{\theta}^*}. \quad (3.3.10)$$

Як і у випадку одномірних оцінок (3.3.6), якщо залежність від параметрів факторизується, то елементи оберненої матриці D^{-1} мають вигляд

$$\|D^{-1}\|_{ij} = - \left(\frac{\partial^2 \ln L}{\partial \theta_i \partial \theta_j} \right)_{\vec{\theta} = \vec{\theta}^*} = \left(\frac{\partial \ln L}{\partial \theta_i} \frac{\partial \ln L}{\partial \theta_j} \right)_{\vec{\theta} = \vec{\theta}^*}. \quad (3.3.11)$$

Остання рівність у (3.3.11) виконується за умови незалежності θ_i і θ_j , що часто має місце.

Розглянемо деякі приклади. Застосуємо метод максимальної правдоподібності до оцінки параметра μ розподілу Пуассона. Як ми бачили в 1.3 цей параметр також визначає і середнє значення цього розподілу. Імовірність розподілу Пуассона дається виразом (1.4.40):

$$P(x_i) = \frac{\mu^{x_i}}{x_i!} e^{-\mu},$$

тому його логарифмічна функція розподілу згідно з (3.3.4) має вигляд

$$l L(x, \mu) = \sum_{i=1}^n l P(x_i, \mu) = \sum_{i=1}^n [x_i \ln \mu - \mu - l(x_i)]. \quad (3.3.12)$$

Оцінки μ^* параметра μ шукаємо з рівняння правдоподібності (3.3.3)

$$\left. \frac{\partial l L}{\partial \mu} \right|_{\mu=\mu^*} = \sum_{i=1}^n \left(\frac{x_i}{\mu^*} - 1 \right) = \sum_{i=1}^n \frac{x_i - \mu^*}{\mu^*} = 0. \quad (3.3.13)$$

Звідки отримуємо $\sum_{i=1}^n (x_i - \mu^*) = 0$ і

$$\mu^* = \frac{1}{n} \sum_{i=1}^n x_i = \langle x \rangle. \quad (3.3.14)$$

Оскільки друга похідна

$$\left. \frac{\partial^2 l L}{\partial \mu^2} \right|_{\mu=\mu^*} = -\frac{1}{\mu^{*2}} \sum_{i=1}^n x_i = -\frac{n}{\mu^*} \quad (3.3.15)$$

є від'ємною, то з μ^* згідно з (3.3.14) функція розподілу має максимальне значення. Тобто оцінкою правдоподібності для математичного сподівання випадкових величин, що розподілені за законом Пуассона, є середньоарифметичне значення μ^* елементів вибірки. Для дисперсії μ^* маємо

$$\sigma_{\mu^*}^2 = \frac{\mu^*}{n}. \quad (3.3.16)$$

Тепер знайдемо оцінки за методом правдоподібності для параметрів μ і σ^2 нормального закону розподілу, тобто математичне сподівання μ і дисперсію σ^2 випадкової величини, що розподілена за нормальним законом зі щільністю

$$p(x_i, \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i - \mu)^2}{2\sigma^2}}.$$

Згідно з (3.3.4) логарифмічна функція розподілу в даному випадку має вигляд

$$\begin{aligned} l L(\mu, \sigma^2) &= \sum_{i=1}^n l P(x_i, \mu, \sigma^2) = \\ &= -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln \sigma^2 - \sum_{i=1}^n \frac{(x_i - \mu)^2}{2\sigma^2}. \end{aligned} \quad (3.3.17)$$

Оцінки параметрів μ і $\theta \equiv \sigma^2$, а саме μ^* , $\theta^* \equiv s^2$, знаходимо із системи рівнянь правдоподібності (3.3.3):

$$\left. \frac{\partial L}{\partial \mu} \right|_{\substack{\mu=\mu^* \\ \sigma^2=\theta^*}} = \sum_{i=1}^n \frac{x_i - \mu^*}{\theta^*} = 0; \quad (3.3.18)$$

$$\left. \frac{\partial L}{\partial \theta} \right|_{\substack{\mu=\mu^* \\ \sigma^2=\theta^*}} = -\frac{n}{2\theta^*} + \sum_{i=1}^n \frac{(x_i - \mu^*)^2}{2\theta^*} = 0. \quad (3.3.19)$$

Звідки

$$\begin{aligned} \sum_{i=1}^n x_i - n\mu^* &= 0; \\ -n + \frac{1}{\theta^*} \sum_{i=1}^n (x_i - \mu^*)^2 &= 0. \end{aligned} \quad (3.3.20)$$

Отже, оцінками правдоподібності математичного сподівання і дисперсії випадкових величин, що розподілені за нормальним законом, є відповідно середньоарифметичне значення та середня вибіркова дисперсія:

$$\mu^* = \frac{1}{n} \sum_i x_i = \langle x \rangle, \quad s^2 \equiv \theta^* = \frac{1}{n} \sum_{i=1}^n (x_i - \mu^*)^2. \quad (3.3.21)$$

Побудуємо тепер матрицю похибок для знайдених оцінок. Оскільки елементи оберненої матриці до матриці похибок мають вигляд

$$\begin{aligned} \|D^{-1}\|_{11} &= \left(-\frac{\partial^2 L}{\partial \mu^2} \right) \bigg|_{\substack{\mu=\mu^* \\ \theta=\theta^*}} = \frac{n}{\theta^*}, \\ \|D^{-1}\|_{12} &= \left\| D^{-1} \right\|_{21} = -\frac{\partial^2 L}{\partial \mu \partial \theta} \bigg|_{\substack{\mu=\mu^* \\ \theta=\theta^*}} = \\ &= \frac{1}{\theta^{*2}} \sum_{i=1}^n (x_i - \mu^*) = \frac{n\langle x \rangle}{\theta^{*2}} - \frac{n\langle x \rangle}{\theta^{*2}} = 0, \\ \|D^{-1}\|_{22} &= -\frac{\partial^2 L}{\partial \theta^2} \bigg|_{\substack{\mu=\mu^* \\ \theta=\theta^*}} = \end{aligned}$$

$$= -\frac{n}{2\theta^{*2}} + \frac{1}{\theta^{*2}} \sum_i (x_i - \mu^*)^2 = \frac{n}{2\theta^{*2}}, \quad (3.3.22)$$

то матриця похибок є

$$D = \begin{pmatrix} \frac{\theta^*}{n} & 0 \\ 0 & \frac{2\theta^{*2}}{n} \end{pmatrix}, \quad (3.3.23)$$

і тому дисперсії оцінок дорівнюють:

$$D[\mu^*] = \frac{\theta^*}{n} = \frac{s^2}{n}, \quad D[\theta^* \equiv \sigma_*^2] = \frac{2\theta^{*2}}{n} = \frac{2s^4}{n}. \quad (3.3.24)$$

Необхідно відзначити, що існує також ще одна особливість оцінок за методом максимальної правдоподібності, а саме, якщо деяка величина θ^* є оцінкою параметра θ , то оцінка (θ^*) функції $\ln L(\theta)$ за методом правдоподібності є (θ^*) , тобто $(\theta^*)' = (\theta^*)$.

3.3. Об'єднання оцінок

До цього часу розглядалися задачі, у яких вибірка характеризувалася однаковими значеннями математичного сподівання і дисперсії для всіх її елементів. Практичний інтерес має випадок, коли елементи вибірки характеризуються однаковим математичним сподіванням, але різними дисперсіями. Така ситуація називається нерівноточними вимірами. Вона, наприклад, виникає: 1) при об'єднанні серій вимірювань однієї і тієї ж величини приладами з різною точністю (або різними авторами, в різних місцях) 2) коли точність вимірювального приладу залежить від значень вимірювальної величини. Задача полягає в побудові за існуючими оцінками найбільш точної. Якщо ігнорувати нерівноточність ряду вимірів і обробляти його як рівноточний ряд, то інформація про вимірювальну величину буде використана неправильно. Якщо ж намагатися виправити положення, відкидаючи результати тих вимірів, ступінь довіри до яких є меншою ніж до інших, то частина інформації, що вміщується в деяких вимірах, буде загублена. У правильному підході повинні бути використані всі виміри, але з урахуванням їхньої точності.

Математично ця задача формулюється таким чином. Нехай дана вибірка нерівноточних вимірів $\{x_k\}$ ($k = 1, \dots, N$) обсягу N , що складається з n

підгруп вибірок $\{x_j\} (j = 1, \dots, n_i)$ обсягу n_i $1: \sum_{i=1}^n n_i = N$, і кожна з вибірок із n_i елементів є реалізацією випадкової величини ξ_i з набору випадкових величин $\{\xi_i\} (i = 1, \dots, n)$, які мають однакові математичні сподівання μ та різні дисперсії σ_i^2 :

$$M[\xi_i] = \mu, \quad D[\xi_i] = \sigma_{\xi_i}^2, \quad i = 1, \dots, n, \quad (3.4.1)$$

де серед середньоквадратичних відхилень σ_i^2 можуть бути й однакові.

Необхідно знайти найбільш точні оцінки (μ^* , $D^* = \sigma_{\mu^*}^2$) для математичного сподівання μ і дисперсії $D = \sigma_{\mu}^2$ випадкової величини, що є сумою n випадкових величин ξ_i і реалізацією якої є об'єднана вибірка $x_1, \dots, x_N$ обсягу N .

Розглянемо випадок незалежних вимірів. Для кожної незалежної випадкової величини ξ_i обчислюємо незсунену оцінку (x_i^*) її математичного сподівання $\mu_i^{(0)} \equiv M[\xi_i]$ у вигляді середнього вибіркового значення з елементів її реалізації $\{x_j\} (j = 1, \dots, n_i)$:

$$x_i^* = \frac{1}{n_i} \sum_{j \in n_i} x_j. \quad (3.4.2)$$

Оцінку μ^* шукаємо у вигляді лінійної комбінації незсунених оцінок x_i^* математичних сподівань випадкових функцій ξ_i

$$\mu^* = \sum_{i=1}^n w_i x_i^*, \quad (3.4.3)$$

де величини w_i є "вагами вимірів". Вони мають бути вибрані з урахуванням нерівноточності окремих вимірів так, щоб перевагу мали результати тих вимірів, які були виконані з найбільшою точністю. Оцінки з такими властивостями повинні бути незсуненими та ефективними (див. 2.2). З вимоги незсуненості величини μ^* маємо

$$M[\mu^*] = M\left[\sum_i w_i x_i^*\right] = \sum_i w_i \mu = \mu. \quad (3.4.4)$$

Тому ваги w_i повинні задовольняти умову нормування

$$\sum_{i=1}^n w_i = 1. \quad (3.4.5)$$

Умова ефективності означає, що дисперсія μ^* повинна мати мінімальне значення

$$D[\mu^*] = \min_{\mu} D[\mu]. \quad (3.4.6)$$

Оскільки випадкові величини ξ_i незалежні, то дисперсія суми (3.4.2) дорівнює сумі дисперсій доданків, і співвідношення (3.4.6) зводиться до умови:

$$D[\mu^*] \equiv \sum_{i=1}^n w_i^2 \sigma_i^2 = \min_{\mu} D[\mu]. \quad (3.4.7)$$

де σ_i^2 – дисперсія середньовибіркового значення $x_i^* : \sigma_i^2 = \sigma_{\xi_i}^2 / n_i$.

Тобто оцінки нерівноточних вимірів будуть незуненими й ефективними, якщо використовуються ваги, для яких квадратична форма $\sum w_i^2 \sigma_i^2$ мінімальна і виконується співвідношення (3.4.5). Одночасно ці умови можна виконати, попередньо переписавши за допомогою (3.4.5) один із коефіцієнтів w_i у (3.4.7) через інші та прирівнявши частинні похідні за w_i до нуля або використавши метод невизначених множників Лагранжа. Останній полягає в тому, що замість безпосереднього знаходження мінімуму виразу $\sum w_i^2 \sigma_i^2$ розглядається сума цього виразу та виразу $\sum w_i - 1$, який повинен дорівнювати нулю внаслідок накладених зв'язків (3.4.5), помноженого на невизначений множник λ . Далі знаходиться безумовний екстремум отриманої суми і підбирається невизначений множник таким чином, щоб виконувалося умова нормування (3.4.5) для знайдених величин. Таким чином, при застосуванні цього методу задача зводиться до визначення безумовного мінімуму виразу

$$\sum_{i=1}^n w_i^2 \sigma_i^2 - 2\sigma^2 \left(\sum_{i=1}^n w_i - 1 \right) \equiv T(\{w_i\}) = \min_{w_i} T(\{w_i\}), \quad (3.4.8)$$

де невизначений множник Лагранжа $\lambda = -2\sigma$.

Прирівнюючи нулю частинні похідні цього виразу за w_i , знаходимо значення коефіцієнтів w_i :

$$w_i = \frac{\sigma^2}{\sigma_i^2}, \quad (3.4.9)$$

при яких вираз (3.4.8) має екстремум. Оскільки вираз (3.4.8) є багаточленом другого ступеня відносно w_i з додатними значеннями коефіцієнтів при старших членах, то згідно із загальними теоремами математичного аналізу співвідношення (3.4.9) забезпечують його мінімум.

Таким чином одержуємо таке правило: для того, щоб оцінка математичного сподівання нерівноточних вимірів була ефективною, необхідно, щоб ваги вимірів були обернено пропорційними квадратам середньоквадратичних відхилень (дисперсій) окремих спостережень

$$w_1 : w_2 : \dots : w_n = \frac{1}{\sigma_1^2} : \frac{1}{\sigma_2^2} : \dots : \frac{1}{\sigma_n^2}. \quad (3.4.10)$$

Підставивши w_i з (3.4.9) в умову нормування (3.4.5), одержимо

$$\sigma^2 = \left(\sum_{i=1}^n \frac{1}{\sigma_i^2} \right)^{-1} \quad (3.4.11)$$

і знаходимо остаточний вираз для вагових коефіцієнтів w_i :

$$w_i = \frac{\sigma^2}{\sigma_i^2} = \frac{1}{\sigma_i^2} \left(\sum_{j=1}^n \frac{1}{\sigma_j^2} \right)^{-1}. \quad (3.4.12)$$

Формулу для дисперсії оцінки μ^* знаходимо після підстановки співвідношення (3.4.12) у ліву частину виразу (3.4.7):

$$\begin{aligned} \sigma_{\mu^*}^2 &= D[\mu^*] = \sum_i w_i^2 \sigma_i^2 = \sum_i w_i (w_i \sigma_i^2) = \\ &= \sigma^2 \sum_i w_i = \sigma^2 = \frac{1}{\sum_{i=1}^n \frac{1}{\sigma_i^2}}. \end{aligned} \quad (3.4.13)$$

Тобто величина $\sigma_{\mu^*}^2$ є стандартним відхиленням оцінки μ^* математичного сподівання μ за відомих дисперсій. Оскільки фактор $1/\sigma_i^2$, що визначає внесок даного виміру, є квадратом відповідної похибки σ_i цього виміру, то будь-який

вимір, який є суттєво менш точним ніж інші, дає значно менший внесок у середнє значення μ^* і дисперсію $\sigma_{\mu^*}^2$. Наприклад, якщо який-небудь вимір в чотири рази є менш точним ніж інші, то його вага у 16 разів менша від значень інших вагових коефіцієнтів і в багатьох випадках цим виміром можна знехтувати.

Усі ці формули при однакових значеннях дисперсій $\sigma_i = \sigma_x$ переходять у відповідні формули (див. 2.3) для рівноточних вимірів. Дійсно, якщо покласти у (3.4.12) $\sigma_i = \sigma_x$, то всі ваги будуть дорівнювати $1/n$, тобто оцінкою μ^* (3.4.3) стає середньоарифметичне значення елементів вибірки. Формула (3.4.13) для квадрата середньоквадратичного відхилення у цьому випадку приймає вигляд

$$\sigma_{\mu^*}^2 = D[\mu^*] = \frac{1}{n} \sigma_x^2, \quad (3.4.14)$$

що знов-таки збігається з формулою (2.3.29) для дисперсії середньоарифметичного значення незалежних величин.

Якщо похибка знаходження середнього вибіркового значення (тобто випадкова величина $\mu - \mu^*$) розподілена за нормальним законом із відомими дисперсіями, то оцінка μ^* також буде нормально розподіленою величиною з дисперсією, що задається (3.4.13). У цьому випадку півширина ε довірчого інтервалу, де оцінка μ^* з імовірністю β не відрізняється від свого точного значення, визначається за формулою

$$\Phi\left(\frac{\varepsilon}{\sigma_{\mu^*}}\right) = \beta, \quad (3.4.15)$$

яка збігається з (2.4.10). У (3.4.15) $\Phi(z) \equiv 2F(z) - 1$ – функція Лапласа (1.4.13), а $F(z)$ – функція розподілу стандартної нормальної величини (1.4.14).

У випадках, коли точні значення дисперсій невідомі, а відомі тільки їхні оцінки, то ваги w_i в (3.4.12) обчислюють, використовуючи емпіричні оцінки $\sigma_{x_i}^2$

(див. (2.3.22), (2.3.29)) дисперсій σ_i^2 :

$$\sigma_i^2 \Rightarrow \sigma_{x_i}^2 = \frac{\tilde{s}_i^2}{n_i}, \quad \tilde{s}_i^2 = \frac{1}{n_i - 1} \sum_{j \in n_i} (x_j - x_i^*)^2, \quad (3.4.16)$$

$$w_i = w_i^* \equiv \frac{g_i^*}{g_i}, \quad g^* = \sum_{i=1}^n g_i^*, \quad g_i^* = \frac{1}{\sigma_{x_i}^2}.$$

Відзначимо, що ваги w_i^* є випадковими величинами, тому обчислення оцінки σ_{μ}^{2*} дисперсії за виразом (3.4.13), який був отриманий у припущенні сталих значень ваг, з підстановкою оцінок $\sigma_{x_i}^{2*}$, замість дисперсій математичних сподівань σ_i^2 , є тільки деяким наближенням. Для більш точного обчислення σ_{μ}^{2*} використовують наступний рецепт (див., наприклад, книгу Т.А.Агеяна, 45). Обчислюють дві наближені оцінки σ_g^2, σ_w^2 для дисперсії σ_{μ}^{2*} середньовибіркового значення μ^* . Величина σ_g^2 обчислюється згідно з (3.4.13) із заміною дисперсій математичних сподівань σ_i^2 на $\sigma_{x_i}^{2*}$:

$$\sigma_g^2 = 1 / \sum_{i=1}^n \frac{1}{\sigma_{x_i}^{2*}} \equiv 1 / g^*, \quad (3.4.17)$$

а σ_w^2 є середньозваженою вибірковою оцінкою вигляду

$$\sigma_w^2 = \frac{1}{n-1} \sum_{i=1}^n w_i^* (x_i^* - \mu^*)^2. \quad (3.4.18)$$

Вираз (3.4.18) для оцінки дисперсії можна отримати з методу максимальної правдоподібності для випадкових величин розподілених за нормальним законом при сталих значеннях w_i^* . Ця оцінка є незсуненою й іноді називається формулою Бесселя для дисперсії нерівноточного ряду вимірювань.

Після обчислення σ_g^2, σ_w^2 їхні значення порівнюють між собою. Якщо виявляється, що $\sigma_g > \sigma_w$, то за стандартне відхилення середньозваженого значення μ^* приймається величина

$$\sigma_{\mu}^* = \frac{1}{2}(\sigma_g + \sigma_w). \quad (3.4.19)$$

Коли ж виконується умова $\sigma_w > \sigma_g$, то стандартне відхилення ототожнюють з σ_w :

$$\sigma_{\mu}^* = \sigma_w. \quad (3.4.20)$$

У конкретних розрахунках значень оцінок математичного сподівання (μ^*) та дисперсії (σ_{μ}^2) вирази (3.4.3) і (3.4.18) зазвичай обчислюють, користуючись деяким умовним початком відліку c . Для цього їх переписують у вигляді

$$\mu^* = \frac{1}{g^*} \sum_{i=1}^n g_i^* (x_i^* - c) + c,$$

$$\sigma_w^2 = \frac{\bar{\sigma}_w^2}{g^*}, \quad \bar{\sigma}_w^2 = \frac{1}{n-1} \left[\sum_{i=1}^n g_i^* (x_i^* - c)^2 - g^* (\mu^* - c)^2 \right], \quad (3.4.21)$$

де g_i^* – ненормована вага (3.4.16) середнього значення i -ї групи вимірів, а c – значення, що визначає умовний початок відліку, вибирається близьким до значення μ^* , або на око, або як середнє значення $\sum_i x_i^* / n$. Еквівалентність виразів (3.4.21) та (3.4.3), (3.4.18) перевіряється дуже легко, але співвідношення (3.4.21) більш зручні в розрахунках.

Як приклад використання формул (3.4.16)–(3.4.21) обчислимо оцінки μ^* і σ_{μ}^2 для об'єднаної серії із трьох вимірів маси металевої кульки. Остання вимірювалася у грамах вагами різної точності (табл. 3.1.). Значення x_i^* , $\sigma_{x_i}^2$ і g_i^* в табл. 3.1 було обчислено за формулами (3.4.2), (3.4.16) з $n_i = 2$.

Таблиця 3.1

Результати серій вимірювань маси кульки

Вимір	Серія		
	1	2	3
1	20,1	21,2	22
2	21,1	22,2	24
x_i^*	20,6	21,7	23
$\sigma_{x_i}^2$	0,25	0,25	1
g_i^*	4	4	1

Таблиця 3.2
*Допоміжні величини, необхідні для обчислення маси кульки
та дисперсії її вимірювань*

i	x_i^*	g_i^*	$x_i^* - c$	$g_i^* (x_i^* - c)$	$g_i^* (x_i^* - c)^2$
1	20,6	4	-1,4	-5,6	7,84
2	21,7	4	-0,3	-1,2	0,36
3	23	1	1	+1	1
		5		-5,8	9,2

Оцінки μ^* і σ_{μ}^2 просто обчислити, якщо скласти табл. 3.2, де вибрано $c = 22$ як значення, що близьке до середнього значення $\sum x_i^* / 3$ і тому до величини μ^* .

Підставляючи значення з останнього рядка табл. 3.2 у співвідношення (3.4.18), маємо ($n = 3$)

$$\mu^* = 22 + \frac{1}{g^*} \sum_{i=1}^3 g_i^* (x_i^* - 22) = 20,84$$

$$\sigma_g^2 = \frac{1}{g^*} = 0,2 ; \quad \sigma_w^2 = \frac{1}{5 \cdot 2} \left[\sum_{i=1}^3 g_i^* (x_i^* - 22)^2 - 5(20,84 - 22)^2 \right] = 0,25 .$$

Оскільки $\sigma_w > \sigma_g$, то $\sigma_{\mu}^* = \sigma_w = 0,5$.

3. . Критерії оцінки граничних значень елементів вибірки. Вилучення аномальних похибок

При аналізі результатів вимірів деякої величини, що подані у вигляді варіаційного ряду, тобто впорядкованої вибірки її елементів зі значеннями $x_{i+1} \geq x_i$,

$$x_1, x_2, \dots, x_{\max},$$

іноді виявляється, що граничні елементи цього ряду порівняно далеко стоять від сусідніх елементів. У цьому випадку виникає підозра, що граничні елементи вибірки вміщують грубі похибки і з'являється бажання відкинути ці сумнівні елементи. Однак, якщо в ході експерименту не було жодної причини для сумніву у правильності вимірів, наприклад збою в роботі обладнання, то не слід априорі вважати граничні елементи помилковими лише тому, що вони лежать далеко від інших значень елементів ряду. Великі відхилення можуть виникати за рахунок

природного розсіяння результатів вимірів і в цьому випадку їх вилучення приведе не до поліпшення, а до погіршення кінцевого результату. Такий результат, наприклад, мав місце в Англії у III ст. за обробки даних геодезичних зйомок. Точні значення довготи вдалося знайти лише після урахування результатів вимірів, які спочатку було відкинуто як сумнівні.

Складність аналізу грубих промахів полягає в тому здавалося б парадоксальному висновку, що чим більшим є обсяг вибірки, тим з більшою ймовірністю необхідно очікувати великих відхилень від середніх значень ("викидів"). Дійсно, як було показано у 1.3, 2.4 для нормально розподілених величин, ймовірність α появи окремого виміру зі значенням x_i , що відрізняється від свого середнього значення $\bar{x}$ більш ніж на 2σ і 3σ , є малою при великих обсягах вибірки, і величина $\alpha = 1 - \beta$ наближено дорівнює 0,05 та відповідно 0,003. Але ймовірність P появи такого значення зростає зі збільшенням числа вимірів. Так, наприклад, при числі незалежних вимірів k ймовірність того, що всі виміри значно не відхиляються від свого середнього значення є рівною $\beta^k \approx 1 - k\alpha$, і ймовірність P відхилення від середнього значення після k вимірів буде дорівнювати $P = 1 - \beta^k \approx k\alpha$. Наприклад, якщо виконано 100 вимірів, то ймовірність того, що один із вимірів буде мати значення поза інтервалом $|x_i - \bar{x}| \geq 3\sigma$ дорівнюватиме $P = 100 \cdot 0,003 = 0,3$, тобто $\approx 30\%$, і такі значення вже не можна відкидати. Таким чином, поява великих відхилень – це не аномалія, а статистична закономірність. Тому статистичний критерій відкидання грубих помилок (різких викидів) будується з урахуванням особливості всієї сукупності спостережень. Побудова такого критерію відноситься до задач перевірки статистичних гіпотез.

Оскільки результати вимірювань є випадковими, то за основу критерію відкидання грубих помилок має бути прийнята ймовірність α появи спостережуваного відхилення і те, що це відхилення виникає випадково серед результатів вимірів, які не містять викидів. Якщо ймовірністю α не можна знехтувати порівняно з деяким критичним значенням α_0 , то відхилення, яке викликає сумнів, не можна вважати аномальним і його треба враховувати при обробці результатів. Якщо ж ймовірність α настільки мала, що, скажімо, $\alpha \leq \alpha_0$, де $\alpha_0 \ll 1$, то можливість природного виникнення такого відхилення можна не враховувати, тобто його можна вважати аномальним і відкинути. Як завжди в задачах даного типу вибір граничної ймовірності α_0 , а саме такої, що ймовірності зі значенням $\alpha \leq \alpha_0$ можна вважа-

ти малими, має декілька довільний характер, однак при $\alpha_0 \ll 1$ це, як правило, відносно мало позначається на кінцевому результаті.

Розглянемо обчислення ймовірності α появи відхилення заданого значення для вибірки з елементів генеральної сукупності, розподіленою за нормальним законом, у двох випадках, коли дисперсія σ_x^2 генеральної сукупності є відомою і невідомою. Вважаємо, що в обох випадках математичне сподівання невідоме. У першій ситуації як міру відхилення найбільшого елемента вибірки від оцінки математичного сподівання можна взяти відносне відхилення Y максимального елемента вибірки:

$$Y = \frac{x_{\max} - \mu^*}{\sigma_x}, \quad \mu^* = \frac{1}{n} \sum_{j=1}^n x_j, \quad (3.5.1)$$

де $x_{\max}$ – найбільший елемент вибірки, а μ^* – оцінка математичного сподівання, обчислена за всіма елементами вибірки.

Міру аномальності значення максимального члена $x_{\max}$ варіаційного ряду $\{x_i\}$ визначає величина ймовірності α того, що максимальне відносне відхилення Y перевищує деяке значення λ :

$$\alpha = P(Y \geq \lambda) \equiv P(x_{\max} \geq \mu^* + \lambda \sigma_x). \quad (3.5.2)$$

А саме, використовується такий критерій: якщо величина $\alpha = P(Y \geq \lambda)$ не перевищує деяке значення α_0 , то елементи вибірки з відносною довжиною більшою за λ можна вважати елементами генеральної сукупності з довірчою ймовірністю $\beta_0 = 1 - \alpha_0$, тобто

$$P\left(\frac{x_{\max} - \mu^*}{\sigma_x} \geq \lambda\right) \equiv \alpha \leq \alpha_0 \equiv 1 - \beta_0. \quad (3.5.3)$$

Ймовірність α_0 задається з деяких додаткових міркувань і називається рівнем значущості критерію відбору граничних значень (3.5.3).

Розглянемо схему обчислення ймовірності (α) того, що максимальний елемент вибірки перевищує її середнє значення на $\lambda \sigma_x$. Згідно з нормальним законом розподілу елементів вибірки маємо для ймовірності (3.5.3) варіаційного ряду $\{x_i\}$

$$\alpha = k \int_{-\infty}^{x_2} dx_1 \int_{-\infty}^{x_3} dx_2 \dots \int_{-\infty}^{x_{\max}} dx_{\max-1} \int_{-\infty}^{\mu^* + \lambda \sigma_x} dx_{\max} \frac{1}{(2\pi)^{\frac{n}{2}} \sigma_x^n} e^{-\frac{1}{2\sigma_x^2} \sum_{j=1}^n (x_j - \mu^*)^2}, \quad (3.5.4)$$

де область інтегрування впорядкована у відповідності до співвідношень між членами варіаційного ряду $x_{\max} \geq x_{\max-1} \geq \dots \geq x_2 \geq x_1$, $x_{\max} - \mu^* \geq \lambda \sigma_x$ множник k перед інтегралом ураховує випадковий порядок появи елементів упорядкованої вибірки. За допомогою заміни змінних $\xi_j = x_j / \sigma_x$ можна вилучити середньоквадратичне відхилення σ_x як з підінтегральної функції, так і з границь інтегрування, тому значення інтегралу (3.5.4) буде залежати лише від числа λ та обсягу вибірки n і визначатися виразом вигляду

$$\alpha = \varphi(\lambda, n), \quad (3.5.5)$$

де φ – відома функція.

Ці співвідношення зв'язує величини α та λ , і тому за його допомогою можна знайти залежність дозволеного відносного відхилення $\lambda_{\max}$ найбільшого елемента вибірки від оцінки математичного сподівання в одиницях середньоквадратичного відхилення та обсягу вибірки:

$$\lambda_{\max} = \psi(\alpha = \alpha_0, n). \quad (3.5.6)$$

Функція φ протабульована і деякі значення величин α , λ , n , що задовольняють (3.5.4), (3.5.5), наведено у табл. 3.3. Звідки видно, що відносне відхилення зі зростанням обсягу вибірки збільшується.

Таблиця 3.3

Відносне значення $\lambda = (x_{\max} - \mu^*) / \sigma_x$ найбільшого елемента вибірки,

що відповідає ймовірності $\alpha = P(x_{\max} - \mu^* \geq \lambda \sigma_x)$

α	N		
	5	10	20
0,10	1,84	2,20	2,50
0,05	2,08	2,44	2,73
0,01	2,57	2,93	3,21

У випадку, коли дисперсія σ_x^2 генеральної сукупності невідома, як міру відхилення елемента $x_{\max}$ впорядкованої вибірки від оцінки μ^* математичного сподівання природно вибрати відхилення Z відносно незсуненої оцінки дисперсії $\sigma_{\mu^*}^2$:

$$Z = \frac{x_{\max} - \mu^*}{\sigma_{\mu^*}}, \quad \sigma_{\mu^*}^2 = \frac{1}{n-1} \sum_{j=1}^n (x_j - \mu^*)^2. \quad (3.5.7)$$

Для нормальної генеральної сукупності й у цьому випадку вдається отримати формулу для визначення ймовірності α відносного відхилення елементів вибірки на деяку довжину λ :

$$\alpha = P(Z \geq \lambda) \equiv P(x_{\max} - \mu^* \geq \lambda \sigma_{\mu^*}), \quad (3.5.8)$$

а критерій відбору максимального елемента має вигляд

$$P\left(\frac{x_{\max} - \mu^*}{\sigma_{\mu^*}} \geq \lambda\right) \equiv \alpha \leq \alpha_0 \equiv 1 - \beta_0. \quad (3.5.9)$$

Імовірність α у формулах (3.5.8), (3.5.9) є функцією лише λ та n і можна знайти залежність λ від α та n у вигляді деякої (однозначно визначеної) функції

$$\lambda = \tilde{\varphi}(\alpha, n). \quad (3.5.10)$$

Значення λ для деяких значень α та n наведено в табл. 3.4. Унаслідок симетрії нормального закону розподілу ймовірності значних відхилень у менший бік дорівнюють імовірностям відповідних відхилень у більший бік:

$$P(x_{\max} \geq \mu^* + \lambda \sigma_{\mu^*}) = P(x_1 \equiv x_{\min} \leq \mu^* - \lambda \sigma_{\mu^*}). \quad (3.5.11)$$

Тому табл. 3.3 і 3.4 можна використовувати для оцінки аномальних відхилень вибірки як у більший, так і в менший бік.

Для ілюстрації критерію відкидання аномальних похибок розглянемо такий приклад. Пятиразові виміри довжини хвилі спектральної лінії дали такі результати (в одиницях Ангстрем, Å):

$$x_1 = 5340, x_2 = 5370, x_3 = 5390, x_4 = 5410, x_5 = 5500.$$

Таблиця 3.4

Відносне значення $\lambda = (x_{\max} - \mu^*) / \sigma_{\mu^*}$ **найбільшого елемента вибірки,**
що відповідає ймовірності $\alpha = P(x_{\max} - \mu^* \geq \lambda \sigma_{\mu^*})$

α	N		
	5	10	20
0,1	1,6	2,03	2,38
0,05	1,67	2,10	2,56
0,025	1,72	2,29	2,71

Оцінити, чи можна вважати значення 5500 Å аномальним, якщо: 1) вважати, що середньоквадратичне відхилення вимірів є відомим і дорівнює $\sigma_x = 50 \text{ Å}$ 2) середньоквадратичне значення вимірів є невідомим. Як ймовірність α критерію відбракування великих відхилень взяти значення $\alpha_0 = 0,1$.

Спершу знаходимо середнє вибіркове значення: $\mu^* = \sum_{i=1}^5 x_i / 5 = 5402$, далі – абсолютне значення відхилення, що передбачається як аномальне: $x_5 - \mu^* = 98$. Використаємо умови (3.5.3) та (3.5.9) для відбракування великих відхилень.

У першому випадку маємо таке значення для відносного відхилення максимального елемента: $\lambda = (x_5 - \mu^*) / \sigma_x = 1,96$. Для цього значення λ при $n = 5$ знаходимо ймовірність α появи такого відхилення (див. табл. 3.3):

$$\alpha = P(x_5 - \mu^* \geq \lambda \sigma_x) = 0,075.$$

Це значення менше α_0 , тому вимір x_5 слід вважати нормальним і залишити.

У другому випадку оцінку σ_{μ^*} обчислюємо за формулою

$$\sigma_{\mu^*} = \sqrt{\frac{\sum_{i=1}^5 (x_i - \mu^*)^2}{4}} = 60,5.$$

Тому маємо, $\lambda = (x_5 - \mu^*) / \sigma_{\mu^*} = 1,62$. За допомогою табл. 3.4 отримуємо $\alpha = P(x_5 - \mu^* \geq \lambda \sigma_{\mu^*}) = 0,086$. Тобто і в цьому випадку x_5 треба залишити.

Розділ

МЕТОД НАЙМЕНШИХ КВАДРАТІВ

1. Принцип найменших квадратів

Головне завдання вимірів у фізиці полягає в тому, щоб із різних результатів вимірів, що подаються у вигляді чисел, отримати деякі закономірності або взаємозв'язки. Якщо між величинами, що вимірюються, існує залежність у вигляді деякої відомої функції параметрів, то для знаходження оцінок цих параметрів часто використовують метод найменших квадратів (МНК). Він, зокрема, дає змогу знаходити параметри і у випадку, коли їх кількість не збігається з кількістю вимірів. У цілому опис експериментальних даних деякими кривими належить до так званого регресивного аналізу, а апроксимуюча функція називається кривою регресії.

Розглянемо МНК на прикладі двох величин y та x , які зв'язані функціональним співвідношенням

$$y = f(x; \{\theta_k\}), \quad k = 1 \div m, \quad (4.1.1)$$

де вигляд функції відомий, а параметри θ_k ($k = 1, 2, \dots, m$) необхідно визначити з експерименту. Припустимо, що для знаходження $\{\theta_k\}$ виконано n вимірів (y_i) значень функцій при n значеннях (x_i) аргументу x . Якби виміри не супроводжувалися похибками, то кожна пара чисел y_i, x_i задовольняла б співвідношення (4.1.1), тобто виконувалася б рівність

$$y_i = f(x_i, \{\theta_k\}) \equiv f_i(\{\theta_k\}), \quad i = 1 \div n, \quad k = 1 \div m. \quad (4.1.2)$$

У дійсності ж через похибки вимірів ці співвідношення не задовольняються, оскільки вони є рівняннями, що виконуються у відсутності похибок, то їх називають умовними рівняннями. Різницю лівої і правої частин виразу (4.1.2) у присутності похибок вимірів називають нев'язкою, або лишком i -го виміру

$$\xi_i = y_i - f(x_i, \{\theta_k\}) \equiv y_i - f_i(\{\theta_k\}). \quad (4.1.3)$$

Сукупність усіх рівнянь типу (4.1.3), які зазвичай записують у вигляді

$$y_i = f_i(\{\theta_k\}) + \xi_i, \quad i = 1 \div n, \quad (4.1.4)$$

називають фундаментальною системою рівнянь МНК. Оскільки для кожного виміру точність і відображення реальних зв'язків (за допомогою співвідношення $f_i(\{\theta_k\})$) різні, тому кожному рівнянню (4.1.4) ставлять у відповідність деяке число w_i , яке визначає цінність даного виміру і яке називається його "вагою".

Зазначимо, що іноді у фундаментальну систему рівнянь також включають співвідношення вигляду

$$\varphi_\ell(\{\theta_k\}) = 0, \ell = 1 \div L, \quad (4.1.5)$$

які ставлять жорсткі обмеження на параметри і які не базуються на результатах вимірів. Наприклад, якщо визначаються кути θ_1 , θ_2 та θ_3 трикутника на площині, то сума вимірюваних значень кутів повинна дорівнювати 180° і тому має бути справедливим співвідношення

$$\varphi(\theta_1, \theta_2, \theta_3) = \theta_1 + \theta_2 + \theta_3 - 180^\circ = 0.$$

Якщо співвідношення (4.1.5) мають наближений характер і виконуються лише з точністю до похибок моделі, то точнішим записом реально існуючих зв'язків буде вираз

$$0 = \varphi_\ell(\{\theta_k\}) + \xi_\ell, \ell = 1 \div L. \quad (4.1.6)$$

У цьому випадку можна не розрізняти два типи співвідношень (4.1.4) і (4.1.6) та вважати рівняння (4.1.6) додатковими рівняннями нової фундаментальної системи рівнянь вигляду (4.1.4) з $y_\ell = 0$ і поставити йому у відповідність достатньо великий ваговий коефіцієнт.

У методі найменших квадратів оцінки параметрів (МНК-оцінки) шукають з умови мінімальності суми S квадратів нев'язок, взятих із відповідними ваговими коефіцієнтами w_i :

$$МНК \Rightarrow \min_{\{\theta_k\}} S(\{\theta_k\}) \equiv \min_{\{\theta_k\}} \left[\sum_{i=1}^n w_i (y_i - f_i(\theta_1, \theta_2, \dots, \theta_m))^2 \right]. \quad (4.1.7)$$

Сума квадратів S є функцією невідомих параметрів θ_k , і тому, якщо всі функції $f_i(\{\theta_k\})$ є диференційованими, необхідною умовою мінімуму S є рівність нулю всіх її перших частинних похідних за параметрами, тобто МНК-оцінки $\{\theta_k^*\}$ параметрів $\{\theta_k\}$ є розв'язками системи рівнянь

$$\frac{\partial S}{\partial \theta_j} = -2 \sum_{i=1}^n w_i \left[y_i - f_i(\{\theta_k\}) \right] \frac{\partial f_i}{\partial \theta_j} \bigg|_{\{\theta_k\} = \{\theta_k^*\}} = 0, j = 1 \div m. \quad (4.1.8)$$

Система (4.1.8) називається системою нормальних рівнянь. Вона складається з m рівнянь, тобто з кількості рівнянь, що збігається з кількістю невідомих, хоча кількість вимірів n може і не збігатися з m . Можливість отримання нормальної системи рівнянь обумовлена тільки диференційованістю функцій f_i і ніяк не пов'язана з лінійністю цих функцій. При нелінійних зв'язках розв'язок системи (4.1.8) може бути неоднозначним та пов'язаним зі значними обчислювальними труднощами. Часто при отриманні МНК-оцінок використовуються загальні методи мінімізації функцій, такі як метод перебору значень аргументу та градієнтні методи. У загальному випадку метод найменших квадратів може не забезпечувати отримання найкращих з погляду теорії ймовірностей оцінок, але його застосування виправдовується як простотою обчислювальної схеми, так і тим, що в багатьох випадках він дійсно дає найкращі оцінки.

Розглянемо один із таких випадків. По-перше, припустимо, що похибками вимірів значень аргументів x_i можна знехтувати порівняно з похибками y_i (тобто розглядаємо величини x_i як детерміновані) і, по-друге, будимо вважати, що при вимірах значень y_i систематичні похибки відсутні, а випадкові похибки є нормально розподіленими величинами з нульовим середнім значенням і дисперсією σ_y^2 . У цьому випадку для кожної точки x_i функція $f(x_i) = f_i$ у формулі (4.1.1) є математичним сподіванням випадкової величини $y(x_i) = y_i$:

$$M[y(x)] \equiv \mu = f(x; \{\theta_k\}), \quad x = x_i, \quad (4.1.9)$$

що розподілена за нормальним законом із функцією розподілу

$$p(y) \equiv p(y; M[y(x)], \sigma_y^2) = p(y; \mu \equiv f(x, \{\theta_k\}), \sigma_y^2) \equiv \frac{e^{-\frac{(y-\mu)^2}{2\sigma_y^2}}}{\sigma_y \sqrt{2\pi}}. \quad (4.1.10)$$

Задача побудови рівнянь для обчислення параметрів $\{\theta_k\}$ буде збігатися із задачею оцінки параметрів математичного сподівання з вибірки y_i випадкової величини з відомим законом розподілу (4.1.10). З погляду теорії ймовірності найточніші оцінки для математичного сподівання можна отримати за допомогою методу максимальної правдоподібності з умови максимуму логарифмічної функції правдоподібності

$$\ln L = \prod_{i=1}^n \ln p(y_i; \mu = f(x_i, \{\theta_k\}); \sigma_y^2), \quad (4.1.11)$$

яка, згідно з (4.1.10), має вигляд

$$\ln L = -\frac{1}{2} \sum_{i=1}^n \frac{(y_i - f_i(\{\theta_k\}))^2}{\sigma_y^2} - n \ln(2\pi\sigma_y). \quad (4.1.12)$$

Умова максимуму функцій правдоподібності, як видно з (4.1.12), буде еквівалентною умові мінімуму квадратичної форми

$$S(\{\theta_k\}) \equiv \sum_{i=1}^n w_i (y_i - f_i(\{\theta_k\}))^2 \quad (4.1.13)$$

з ваговими коефіцієнтами

$$w_i = \frac{1}{2\sigma_y^2}. \quad (4.1.14)$$

Вираз (4.1.13) збігається із квадратичною формою S методу найменших квадратів (4.1.7) з вагами згідно з (4.1.14). Статистичні вагові коефіцієнти вигляду $w_i \propto 1/\sigma_y^2$ відображають той очевидний факт, що найвагомішим повинен бути внесок тих емпіричних точок, що мають найменшу дисперсію. Таким чином, у розглянутому випадку МНК повністю еквівалентний методу максимальної правдоподібності, тому МНК-оцінки параметрів також будуть одними з найточніших з погляду теорії ймовірності (див. 3.3).

У загальному випадку підхід за МНК є самостійним методом оцінки параметрів, який не потребує інформації про вигляд функції розподілу ймовірності. Можна показати, що МНК-оцінки величин θ_k у випадку лінійної залежності функцій f від параметрів мають мінімальні дисперсії.

Якщо експериментальні дані попарно корельовані між собою, то МНК-оцінки параметрів апроксимуючої функції $f(x, \{\theta_k\})$ шукають з умови мінімуму квадратичної форми вигляду

$$\Phi(\{\theta_k\}) = \sum_{i=1}^n \sum_{j=1}^n [y_i - f_i(\{\theta_k\})] w_{ij} [y_j - f_j(\{\theta_k\})], \quad (4.1.15)$$

де вагові множники є матричними елементами деякої симетричної матриці W . Система нормальних рівнянь для оцінки параметрів тепер має вигляд

$$\frac{\partial \Phi}{\partial \theta_l} = -2 \sum_{i=1}^n \sum_{j=1}^n \frac{\partial f_i}{\partial \theta_l} w_{ij} [y_j - f_j] \Big|_{\{\theta_k\} = \{\theta_k^*\}} = 0, \quad l = 1 \div m. \quad (4.1.16)$$

У випадку, коли матриця вагових множників $\hat{W}$ діагональна ($w_{ij} = w_i \delta_{ij}$), квадратична форма (4.1.15) приймає вигляд виразу (4.1.7), (4.1.13), а система нормальних рівнянь (4.1.16) переходить у систему (4.1.8).

Розглянемо вибір матриці вагових множників.

У загальному випадку, як і для некорельованих вимірів, природно вважати, що члени квадратичної форми (4.1.15), які містять значення y_k з великими дисперсіями, повинні давати менший внесок у суму, ніж доданки з малим значенням дисперсії. Це пов'язано з тим, що чим точніше визначена величина, тим менше значення її дисперсії. У випадку, коли похибками аргументів можна знехтувати, таку поведінку коефіцієнтів w_{ij} можна забезпечити, виравши матрицю вагових множників $\hat{W}$ обернено пропорційною матриці коваріацій експериментальних даних $\hat{D}(y)$ (1.2.2), (1.3.11):

$$\hat{W}(y) = \sigma^2 \hat{D}^{-1}(y), \quad (4.1.17)$$

де

$$\hat{D} \equiv |\hat{D}| = \begin{vmatrix} \sigma_1^2 & \sigma_1 \sigma_2 \rho_{12} & \dots & \sigma_1 \sigma_n \rho_{1n} \\ \sigma_1 \sigma_2 \rho_{12} & \sigma_2^2 & \dots & \sigma_2 \sigma_n \rho_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ \sigma_1 \sigma_n \rho_{1n} & \cdot & \cdot & \sigma_n^2 \end{vmatrix}. \quad (4.1.18)$$

Тут σ_i^2 – дисперсія значень y_i , ρ_{ij} – коефіцієнт кореляції між значеннями y_i , y_j , а σ^2 – деякий сталий коефіцієнт, який вибирають з умов нормування вагових коефіцієнтів w_{ij} . Очевидно, що використання у виразах для квадратичних форм вагових коефіцієнтів, які відрізняються тільки загальним множником, ніяк не змінює вигляд нормальних рівнянь, а тому й оцінки параметрів за МНК. Наприклад, у відсутність кореляцій ($\rho_{ij} = 0$, $w_{ij} = w_i \delta_{ij}$) використовують як ненормовані вагові коефіцієнти, подібні до (4.1.14), а саме:

$$w_i = \frac{1}{2\sigma_i^2}, \quad (4.1.19)$$

так і нормовані вагові множники вигляду

$$w_i = \frac{\sigma^2}{\sigma_i^2}, \quad \sum_{i=1}^n w_i = 1, \quad (4.1.20)$$

де

$$\sigma^2 = \left(\sum_{j=1}^n \frac{1}{\sigma_j^2} \right)^{-1}. \quad (4.1.21)$$

Множники w_i , що задаються (4.1.20) і (4.1.21), збігаються з ваговими коефіцієнтами (3.4.12), які формують середнє значення випадкової величини при нерівноточних вимірах (див. 3.4).

Квадратична форма S ((4.1.7), (4.1.13)) з вагами вигляду (4.1.19) називається χ^2 -квадратом () та позначається як

$$\chi^2 \equiv \sum_{i=1}^n \left(\frac{y_i - f(x_i; \{\theta_k\})}{\sigma_i} \right)^2. \quad (4.1.22)$$

Вираз (4.1.22) позначається символом χ^2 тому, що при нормальному розподілі значень y_i величина (4.1.22) розділена за законом χ^2 . На практиці використовують емпіричні оцінки для середньоквадратичних відхилень σ_i і при необхідності їх уточнюють у подальших наближеннях.

Перша друкована робота, у якій використовувався МНК, належить Лежандру, яку він опублікував 1805 р. і присвятив визначенню орбіт комет. Однак концепцію МНК-оцінки вперше було використано вісімнадцятирічним Гауссом (1795) для аналізу даних топографічних зйомок, а пізніше він регулярно використовував його також в астрономічних обчисленнях, але матеріали не опубліковував. Учений вважав, що ідея, яка лежить в основі методу, використовувалась за багато років до нього. Коли Гаусс 1809 р. опублікував опис своїх астрономічних обчислень, то Лежандр звинуватив його у крадіжці ідеї підходу за МНК. Лежандр не знав, що Гаусс використовував МНК вже протягом десяти років, а Гаусс не став це йому доводити.

Складання умовних рівнянь та розв'язок нормальних рівнянь може бути достатньо трудомісткою справою. Зафіксовано своєрідні рекорди. Наприклад, виконуючи топографічну зйомку міста Ганновера, Гаусс склав і розв'язав систему із 300-ми умовними рівняннями із 55-ма невідомими. Це зайняло в нього приблизно три місяці безперервної роботи. При проведенні топографічної зйомки Індії (до другої світової війни) було складено близько 2500 умовних рівнянь із 88-ми невідомими. Обчислювальні роботи, які вручну виконувались співробітниками спеціального обчислювального бюро, продовжувалися приблизно два роки. Поява комп'ютерів дозволила набагато перевершити ці рекорди і у багато разів скоротити час обчислення. Наприклад, уже в 70-х рр. минулого сторіччя можна було знайти розв'язок системи лінійних алгебраїчних рівнянь з 10^5 невідомими.

2. Наближення функціями лінійними за параметрами

Спочатку розглянемо випадок наближення даних $y(x_i) \equiv y_i$ ($i = 1, \dots, n$) функцією $f(x; \theta) \equiv \theta g(x)$, яка лінійно залежить тільки від одного параметра θ . З урахуванням похибок вимірювань експериментальні дані задовольняють співвідношення вигляду (4.1.4), а саме:

$$y_i = \theta g_i + \xi_i, \quad g_i = g(x_i), \quad (4.2.1)$$

де ξ_i – випадкова похибка i -го виміру, $g(x)$ – деяка детермінована функція величини x , похибками якої можна знехтувати. Також вважаємо, що при вимірах значень y_i систематичні похибки відсутні, і тому математичне сподівання похибок дорівнює нулю: $M[\xi_i] = 0$. Дисперсію i -ї похибки позначаємо σ_i^2 : $D[\xi_i] \equiv M[(\xi_i - M[\xi_i])^2] = M[\xi_i^2] = \sigma_i^2$. Для простоти припустимо, що експериментальні дані є некорельованими.

Система нормальних рівнянь (4.1.8) зводиться до рівняння

$$\sum_{i=1}^n w_i [y_i - \theta^* g_i] g_i = 0, \quad (4.2.2)$$

де θ^* – оцінка параметра θ . Звідки маємо таке співвідношення для обчислення оцінки:

$$\theta^* = \frac{\sum_{i=1}^n w_i g_i y_i}{\sum_{i=1}^n w_i g_i^2}. \quad (4.2.3)$$

Оскільки в рівнянні (4.2.1) функція g , параметр θ і величини x_i є не випадковими, то виконується співвідношення

$$M[y_i] = \theta g_i, \quad D[y_i] \equiv M[(y_i - M[y_i])^2] = M[\xi_i^2] = \sigma_i^2. \quad (4.2.4)$$

З виразу для математичного сподівання виміру маємо, що оцінка (4.2.3) є незсуненою:

$$M[\theta^*] = \frac{\sum_{i=1}^n w_i g_i M[y_i]}{\sum_{i=1}^n w_i g_i^2} = \theta. \quad (4.2.5)$$

Із загальних правил обчислення дисперсій знаходимо, що дисперсія оцінки (4.2.3) зв'язана з дисперсією величин y_i співвідношенням

$$\sigma_{\theta^*}^2 \equiv D[\theta^*] \equiv M\left[\left(\theta^* - M[\theta^*]\right)^2\right] = \frac{\sum_{i=1}^n [w_i g_i]^2 \sigma_i^2}{\left[\sum_{j=1}^n w_j g_j^2\right]}, \quad (4.2.6)$$

де σ_i^2 – дисперсії окремих значень експериментальних даних. Якщо використовувати нормовані вагові множники вигляду $w_i = \frac{\sigma^2}{\sigma_i^2}$, $\sum_{i=1}^n w_i = 1$,

$\sigma^2 = \left(\sum_{j=1}^n \frac{1}{\sigma_j^2}\right)^{-1}$, див. (4.1.20), (4.1.21), то вираз (4.2.6) приймає вигляд

$$\sigma_{\theta^*}^2 = \left[\sum_{j=1}^n \left(\frac{g_j}{\sigma_j}\right)^2\right]^{-1}. \quad (4.2.7)$$

Тепер розглянемо наближення експериментальних даних функцією $f(x, \{\theta_k\})$, що лінійно залежить від m параметрів θ_l :

$$f(x, \{\theta_k\}) = \sum_{l=1}^m \theta_l \psi_l(x), \quad (4.2.8)$$

де $\psi_l(x)$ – задані функції.

Припустимо, що, по-перше, при вимірах експериментальних значень функції f відсутні систематичні похибки, а випадкові похибки (невязки) ξ_j розподілені за нормальним законом з однаковою дисперсією σ_y^2 ($D[\xi_i] = M[\xi_i^2] = \sigma_y^2$) і, по-друге, що виміри значень аргументів x_j виконуються без похибок. З урахуванням похибок вимірювань експериментальні дані задовольняють співвідношення вигляду (4.1.4):

$$y_j = \sum_{l=1}^m \theta_l \psi_l(x_j) + \xi_j, \quad j = 1, \dots, n. \quad (4.2.9)$$

Запишемо систему нормальних рівнянь. Для цього диференціюємо суму квадратів невязок ξ_j^2 :

за θ_r , маємо

де параметри θ_i замінено їх оцінками θ_i^* , оскільки система нормальних рівнянь визначає саме оцінки. Змінивши порядок підсумовування у другому компоненті (4.2.11), перепишемо систему рівнянь (4.2.11) у вигляді

Ця система рівнянь є лінійною системою відносно невідомих θ_l^* . Кількість рівнянь системи (4.2.12) збігається з кількістю параметрів і функцій $\psi_l(x)$. Для спрощення запису введемо позначення

де d_{lr} – невідповідні й симетричні відносно своїх індексів числа, а V_r – нормально розподілені випадкові величини, оскільки вони є лінійними комбінаціями випадкових величин y_i .

З урахуванням позначень (4.2.13) система нормальних рівнянь переписується у такому розгорнутому вигляді:

Розв'язок цієї системи рівнянь і визначає оцінки параметрів θ_l^* . Згідно з правилом Крамера маємо

99

Тут Δ – детермінант системи (4.2.14), а A_{lr} – детермінант алгебраїчного доповнення, а саме A_{lr} є детермінантом матриці, що знаходиться з матриці d_{lr} після відкидання l -го рядка та r -го стовпчика, і який береться зі знаком $(-1)^{l+r}$

$$\Delta = \begin{vmatrix} d_{11} & d_{12} & \dots & d_{1m} \\ d_{21} & d_{22} & \dots & d_{2m} \\ \dots & \dots & \dots & \dots \\ d_{m1} & d_{m2} & \dots & d_{mm} \end{vmatrix}; A_{lr} = (-1)^{l+r} \begin{vmatrix} d_{11} & \dots & d_{1r-1} & d_{1r+1} & \dots & d_{1m} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ d_{l-11} & \dots & d_{l-1r-1} & d_{l-1r+1} & \dots & d_{l-1m} \\ d_{l+11} & \dots & d_{l+1r-1} & d_{l+1r+1} & \dots & d_{l+1m} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ d_{m1} & \dots & d_{mr-1} & d_{mr+1} & \dots & d_{mm} \end{vmatrix}. \quad (4.2.16)$$

З формули (4.2.15) випливає, що величини θ_l^* розподілено за нормальним законом, оскільки вони є лінійними виразами нормально розподілених випадкових величин V_e . Можна також показати, що оцінки θ_l^* є незсуненими. Ураховуючи формули (4.2.13), (4.2.15), $D[y_i] \equiv M[(y_i - M[y_i])^2] = M[\xi_i^2] = \sigma_y^2$ і загальні правила обчислення дисперсій суми величин (див. 1.2), знаходимо такий вираз для дисперсії оцінок:

$$\sigma_{\theta_l^*}^2 \equiv D[\theta_l^*] = \frac{A_{ll}}{\Delta} \sigma_y^2, \quad l = 1, \dots, m. \quad (4.2.17)$$

Оцінки коефіцієнтів θ_l^* з різними індексами не є незалежними, оскільки для них існує відмінний від нуля коефіцієнт коваріації

$$R_{lr} \equiv \text{cov}(\theta_l^*, \theta_r^*) = M[(\theta_l^* - M[\theta_l^*])(\theta_r^* - M[\theta_r^*])]. \quad (4.2.18)$$

Використовуючи вирази (4.2.13), (4.2.15), $D[y_i] = \sigma_y^2$, після ряду алгебраїчних перетворень отримуємо співвідношення

$$R_{lr} = \frac{A_{lr}}{\Delta} \sigma_y^2, \quad (4.2.19)$$

які переходять у (4.2.17) при $l = r$.

Якщо дисперсія σ_y^2 похибки вимірів функції $y(\{\theta_k\})$ є відомою, то формула (4.2.17) дозволяє обчислити довірчі інтервали для оцінок θ_l^* , оскільки

їх математичні сподівання дорівнюють θ_l , а дисперсія визначається (4.2.17).

Ураховуючи нормальний закон розподілу для θ_l^* та згідно з розглядом у 2.4, маємо співвідношення, що пов'язує довірчу ймовірність β та півширину ε довірчого інтервалу:

$$\beta = P\left\{|\theta_l^* - \theta| \leq \varepsilon\right\} = \Phi\left(\frac{\varepsilon}{\sigma_{\theta_l^*}}\right), \quad (4.2.20)$$

де Φ – функція Лапласа (1.4.13).

Якщо дисперсія окремих вимірів σ_y^2 є невідомою, то її незсунена оцінка дисперсії визначається за формулою

$$\tilde{\sigma}_y^2 = \frac{1}{n-m} \sum_{j=1}^n \left[y_j - \sum_{l=1}^m \theta_l^* \psi_l(x_j) \right]^2, \quad (4.2.21)$$

де m – кількість невідомих параметрів.

Можна показати, що випадкова величина

$$T_l = (\theta_l^* - \theta_l) / \sqrt{\tilde{D}}, \quad (4.2.22)$$

де $\tilde{D} = \tilde{\sigma}_y^2 A_{ll} / \Delta$ визначається формулою (4.2.15), у якій замість σ_y^2 стоїть її оцінка $\tilde{\sigma}_y^2$, підпорядковується закону розподілу Стюдента з $(n-m)$ ступенями свободи. Тому довірчий інтервал для оцінки θ_l^* можна записати у вигляді (див. 2.4)

$$\left(\theta_l^* - t_\beta(n-m)\sqrt{\tilde{D}}, \theta_l^* + t_\beta(n-m)\sqrt{\tilde{D}} \right), \quad (4.2.23)$$

де числа $t_\beta(n-m)$ знаходяться відповідно з довірчою ймовірністю β і числом $(n-m)$ ступенів свободи за таблицями закону розподілу Стюдента.

У загальному випадку функції, якими апроксимують експериментальні дані, є нелінійними відносно параметрів θ_l що приводить до складних обчислень, але в деяких випадках вдається перейти до лінійних виразів, що значно спрощує розрахунки. Перехід до лінійних виразів можна виконати одним із двох способів. Наприклад, часто функцію $f(x, \theta_1, \dots, \theta_m)$ вдається лінеаризувати за деяких значень $\{\theta_k^{(0)}\}$ і далі за допомогою МНК знайти оцінку

для різниці $(\theta_l^{(0)} - \theta_l)$. Іноді функцію f вдається звести до лінійної за допомогою деяких перетворень наприклад, коли f має вигляд експоненти, то МНК використовують для логарифма від f .

Зазвичай, коли МНК використовують для оцінки параметрів, то у вагових коефіцієнтах ураховують лише невизначеності, пов'язані з похибками результатів вимірів, а не тих величин, що використовуються як аргументи. Разом із тим, в експериментах значення незалежної змінної x також є відомим не абсолютно точно, а з деякою невизначеністю. Простіший метод урахування похибок аргументів базується на перенормуванні експериментальних даних за допомогою апроксимації кривої, що досліджується, поліномом другого ступеня за середнього значення аргументу. Так, наприклад, в одновимірному випадку маємо

$$f(x; \bar{\theta}) = f(\bar{x}_i; \bar{\theta}) + (x - \bar{x}_i) f'_x(\bar{x}_i) + \frac{1}{2} (x - \bar{x}_i)^2 f''_x(\bar{x}_i), \quad (4.2.24)$$

де $\bar{x}_i$ – середнє значення аргументу в точці i , а $f'_x(\bar{x}_i)$, $f''_x(\bar{x}_i)$ – похідні першого та другого порядків за аргументом. Аргументи x_i є випадковими величинами, що розподілені за деяким законом, тому обчислимо середнє значення і дисперсію функції (4.2.30) за розподілом x_i . Після усереднення (4.2.24) за x маємо

$$\bar{f}(x; \bar{\theta}) = f(\bar{x}_i; \bar{\theta}) + \frac{1}{2} \sigma_{x_i}^2 f''_x(\bar{x}_i; \bar{\theta}) \equiv \bar{y}_i, \quad (4.2.25)$$

де $\sigma_{x_i}^2$ – дисперсія значень аргументу. Величини $\bar{y}_i$ враховують похибки аргументу і мають зміст нових експериментальних значень функції, що використовуються в МНК:

$$\bar{y}_i = y_i + \frac{1}{2} \sigma_{x_i}^2 f''_x(\bar{x}_i; \bar{\theta}). \quad (4.2.26)$$

Тобто величина $\bar{y}_i$ є перенормованим експериментальним значенням функції в точці $\bar{x}_i$. Дисперсію $\sigma^2(\bar{f}) \equiv \sigma^2(\bar{y}_i)$ цих значень можна знайти з формули (4.2.24), і вона дорівнює

$$\sigma^2(\bar{y}) = \sigma_y^2 + \sigma_{x_i}^2 \left[f''_x(\bar{x}_i; \bar{\theta}) \right]^2, \quad (4.2.27)$$

де σ_y^2 – дисперсія значень функції.

Оцінки параметрів МНК за формулами (4.2.26), (4.2.27), що наближено враховують похибки аргументу, можна знайти, мінімізуючи квадратичну форму

вигляду (4.1.13), (4.1.14) із заміною значень y_i , σ_y на величини $\bar{y}_i$ і $\sigma^2(\bar{y})$ відповідно. Труднощі цього методу полягають у залежності функції $\bar{y}_i$ (4.2.26) і дисперсії (4.2.27) від параметрів $\bar{\theta}$, які треба визначити. Тому для його реалізації необхідно користуватись методом послідовних наближень наприклад, використовуючи як $\bar{\theta}$ у формулах (4.2.26) та (4.2.27) спершу нульове наближення з подальшим поступовим його уточненням.

4.3. Метод лінеаризації в МНК

Метод лінеаризації для обчислення оцінок параметрів за МНК із наближеним урахуванням других похідних від функції χ^2 було розроблено в Дубні С.Н.Соколовим та І.М.Сіліним і реалізовано останнім у вигляді програми *FUMILI*, що написана алгоритмічною мовою програмування *Фортран*. Даний метод широко використовується при обробці експериментальних даних.

Розглянемо загальний алгоритм обчислень. Нехай ми маємо n вимірів експериментальних значень $y_1 \pm \sigma_1, \dots, y_n \pm \sigma_n$, які хочемо підігнати за МНК гладкими теоретичними функціями $f_1(\bar{x}_1; \bar{\theta}), \dots, f_n(\bar{x}_n; \bar{\theta})$, де $\bar{\theta} \equiv \{\theta_1, \dots, \theta_m\}$ – параметри, що оцінюються, а $x_j = \{x_{j,1}, \dots, x_{j,k}\}$ – багатовимірні координати експериментальних точок. Аналітичні вирази функцій f_j можуть бути як однаковими, так і різними для всіх значень аргументів. Згідно з МНК необхідно знайти мінімум суми квадратів нев'язок. Функція, що мінімізується в реалізації МНК, яка використовується у програмі *FUMILI*, має такий вигляд:

$$S(\bar{\theta}) = \frac{1}{2} \sum_{j=1}^n \left(\frac{f_j(\bar{x}_j; \bar{\theta}) - y_j}{\sigma_j} \right)^2. \quad (4.3.1)$$

При розробці алгоритму автори ставили за мету прискорити обробку експериментальних даних для складних функцій $f_j(\bar{x}; \bar{\theta})$, які потребують багато часу для обчислень. Від інших алгоритмів подібного роду алгоритм *FUMILI* відрізняється, перш за все, способом обмеження кроку у процесі виконання ітерацій.

Функція $S(\theta)$ може мати багато локальних мінімумів. Лише в рідкісних випадках, наприклад при лінійних відносно параметрів функціях $f_j(\theta)$, може бути доведено існування одного мінімуму. У загальному випадку за до-

помогою даного методу можна знайти будь-який локальний мінімум. Метод лінеаризації при пошуку мінімуму (θ^*) базується на тому спостереженні, що в його околі ($\theta - \theta^*$) будь-яку функцію можна наблизити квадратичною функцією і $S(\theta)$ з (4.3.1) можна записати у вигляді

$$S(\theta) \cong S(\theta^*) + \sum_{i,r=1}^n \frac{\partial^2 S}{\partial \theta_i \partial \theta_r} (\theta_i - \theta_i^*) (\theta_r - \theta_r^*), \quad (4.3.2)$$

оскільки в мінімумі $\partial S / \partial \theta_i = 0$. Далі припускають, що функція $S(\theta)$ достатньо гладка, тому її можна наблизити квадратичною функцією і в деякому околі мінімуму, а саме, для деяких значень $\bar{\theta}^{(0)}$ біля θ^* використовують розклад:

$$S(\theta) \cong S(\theta^{(0)}) + \sum_{i=1}^n \frac{\partial S}{\partial \theta_i} (\theta_i - \theta_i^{(0)}) + \sum_{i,r=1}^n \frac{\partial^2 S}{\partial \theta_i \partial \theta_r} (\theta_i - \theta_i^{(0)}) (\theta_r - \theta_r^{(0)}), \quad (4.3.3)$$

де похідні обчислюються при $\bar{\theta} = \bar{\theta}^{(0)}$. Крім того, близько мінімуму також нехтують нелінійністю функцій f_j і при обчисленні других похідних функції S залишають лише перші похідні $\frac{\partial f_j}{\partial \theta_r}$:

$$Z_{ir} \equiv \frac{\partial^2 S}{\partial \theta_i \partial \theta_r} = \sum_{j=1}^n \frac{1}{\sigma_j^2} \frac{\partial f_j}{\partial \theta_i} \frac{\partial f_j}{\partial \theta_r} + \sum_{j=1}^n \frac{f_j - y_j}{\sigma_j^2} \frac{\partial^2 f_j}{\partial \theta_i \partial \theta_r} \cong \sum_{j=1}^n \frac{1}{\sigma_j^2} \frac{\partial f_j}{\partial \theta_i} \frac{\partial f_j}{\partial \theta_r}. \quad (4.3.4)$$

Використання цього наближення дуже спрощує як аналітичне, так і числове диференціювання, а в деяких випадках при числових знаходженнях похідних зменшує похибки обчислень.

Умова мінімуму виразу (4.3.3) біля точки $\bar{\theta}^{(0)}$ приводить до такої системи нормальних рівнянь:

$$\frac{\partial S}{\partial \theta_i} = \frac{\partial S(\bar{\theta}^{(0)})}{\partial \theta_i} + \sum_r Z_{ir} \cdot (\theta_r - \theta_r^{(0)}) = 0, \quad (4.3.5)$$

де згідно з (4.3.1) похідні $\frac{\partial S(\bar{\theta}^{(0)})}{\partial \theta_i}$ також є функціями перших похідних від f_j :

$$\frac{\partial S(\theta^{(0)})}{\partial \theta_i} = \sum_{j=1}^n \frac{1}{\sigma_j^2} \frac{\partial f_j}{\partial \theta_i} (f_j(x_j; \theta^{(0)}) - y_j). \quad (4.3.6)$$

Позначивши величину $\theta_r - \theta_r^{(0)}$ через $\Delta\theta_r$, а $\frac{\partial S}{\partial \theta_i}$ через g_i , розв'язок системи рівнянь (4.3.5) має вигляд

$$\Delta\theta_r = -\sum_{i=1}^m g_i Z_{ir}^{-1}, \quad g_i = \frac{\partial S}{\partial \theta_i}, \quad (4.3.7)$$

де Z_{ir}^{-1} – елементи матриці, що обернена до Z .

Згідно з цими рівняннями, вибравши деяке початкове значення $\theta^{(0)}$, маємо оцінку для значення координати точного мінімуму

$$\theta^* = \theta^{(0)} + \Delta\theta, \quad \Delta\theta = \left\{ \Delta\theta_1 = \theta_1^* - \theta_1^{(0)}, \dots, \Delta\theta_m = \theta_m^* - \theta_m^{(0)} \right\}. \quad (4.3.8)$$

Зі співвідношень (4.3.3), (4.3.4) видно, що такий метод є абсолютно точним лише для функцій, що лінійно залежать від параметрів. У нелінійному випадку дуже важливо правильно вибрати значення нульового наближення $\theta^{(0)}$ таким, щоб воно лежало близько мінімуму функції S . Якщо функції f_j є нелінійними, то величина з правого боку виразу (4.3.8) буде лише першим наближенням θ^{**} до точного значення оцінки:

$$\theta^{(1)} = \theta^{(0)} + \Delta\theta. \quad (4.3.9)$$

Далі необхідно обчислити значення $S_1 \equiv S(\bar{\theta}^{(1)})$. Якщо виявиться, що дана величина більше $S_0 \equiv S(\theta^{(0)})$, то це може означати (при правильному виборі початкового наближення), що нелінійні ефекти достатньо великі, а область квазілінійності $f_j(\theta)$ погано оцінена, тому треба якимось чином зменшити крок $\Delta\theta$. Для цього в *FUMILI* використовується специфічний апарат обмеження кроку і вводиться масив обмежувачів ($PL(i)$) початкових кроків. Вони вибираються з тих міркувань, щоб при зміні θ_i в деяких фіксованих межах $\pm b_i$ функції $f_j(\theta)$ були в середньому лінійними (точніше нелінійна частина їх приросту була б за модулем менша ніж лінійна). Якщо крок $\Delta\theta$ виходить за межі паралелепіпеда W_0 з центром θ^{**} та з довжинами боків $2b_i$, то знаходиться точка перетинання вектором $\Delta\theta$ поверхні W_0 , яка і використовується як перше наближення θ^{**} . Якщо і в цьому ви-

падку виявляється, що в дуже близькому околі близько мінімуму $S_1 > S_0$, то треба зменшити значення b_i і тим самим пересунути $\theta^{(n)}$ ближче до $\theta^{(n-1)}$. У дійсності ж треба також зменшити крок і в тому випадку, коли S_1 мало зменшується порівняно з S_0 . Це дає змогу подавити погано загасаючі коливання біля мінімуму у випадку сильної нелінійності. Якщо є необхідність, то виконується повторне зменшення кроку, але яке не повинно виконуватися надмірно багато разів. Зокрема, при деяких невдалих значеннях $\theta^{(n)}$ можуть виникнути великі втрати точності обчислень S_0 і $\Delta\theta$, через які значення S_1 ніколи не буде менше S_0 . Тому, як правило, обмежуються максимум двома дробленнями кроку і переходять до наступної ітерації. У *FUMILI* є параметр $N1$ який задає максимальне число дроблень. Дроблення виконується при значенні $N1 > 1$, якщо виконується рівність

$$S_0 - S_1 < \frac{1}{2}(S_0 - \tilde{S}_{\min}), \quad (4.3.10)$$

де $\tilde{S}_{\min}$ – значення функції, що квадратично інтерполює функції $S(\theta)$ при заданих S_0 і S_1 у точці $\tilde{\theta} = (\tilde{\theta}^{(0)} + \tilde{\theta}^{(1)})/4$. Умову (4.3.10) знайдено за допомогою обчислювальних експериментів, і вона фактично визначає сильну нелінійність функції f_j на інтервалі від $\tilde{\theta}^{(0)}$ до $\tilde{\theta}^{(1)}$.

Якщо зменшення кроку виявилось непотрібним (умова (4.3.10) не виконується), а мінімуми функції S не знайдено, то необхідно перейти до другої ітерації. Вона знаходиться з першої заміною у відповідних виразах індексів 0 та 1 на відповідно 1 і 2, тобто загальна ітераційна схема має вигляд

$$\begin{aligned} \theta^{(n)} &= \theta^{(n-1)} + \Delta\theta^{(n)}; \Delta\theta^{(n)} \equiv \{\Delta\theta_1^{(n)}, \dots, \Delta\theta_m^{(n)}\}, \\ \Delta\theta_r^{(n)} &= -\sum_{i=1}^m g_i \left(\theta^{(n-1)} \right) \cdot \left[Z_{ir} \left(\theta^{(n-1)} \right) \right]^{-1}, \end{aligned} \quad (4.3.11)$$

де

$$\begin{aligned} g_i(\vec{R}) &\equiv \frac{\partial S(\vec{R})}{\partial \theta_i} = \sum_{j=1}^n \frac{1}{\sigma_j^2} \frac{\partial f_j(\vec{R})}{\partial \theta_i} \left[f_j(\vec{x}_j, \vec{R}) - y_j \right], \\ Z_{ir}(\vec{R}) &\equiv \frac{\partial^2 S(\vec{R})}{\partial \theta_i \partial \theta_r} = \sum_{j=1}^n \frac{1}{\sigma_j^2} \frac{\partial f_j(\vec{R})}{\partial \theta_i} \frac{\partial f_j(\vec{R})}{\partial \theta_r}, \quad \vec{R} \equiv \tilde{\theta}^{(n-1)}. \end{aligned} \quad (4.3.12)$$

Зазначимо, що в загальному випадку схема (4.3.11), (4.3.12) обчислення мінімуму деякої функції S називається методом градієнтного спуску.

У програмі *FUMILI* максимальне число ітерацій задається значенням параметра $N3$. Для прискорення збіжності ітерацій іноді слід збільшувати розміри b_i області квазілінійності. Тому, якщо у попередній ітерації (взагалі кажучи, у декількох попередніх ітераціях) дроблення кроку не було, то, як правило, b_i збільшуються в чотири рази. Обчислені у *FUMILI* величини b_i є мірою нелінійності $f_j(\theta)$ і тому вміщують інформацію, яка необхідна для числового диференціювання цих функцій. На практиці гарні результати дає вибір значення $0,01b_i$, як кроку числового диференціювання,

$$\frac{\partial f_j(\theta)}{\partial \theta_i} = \frac{f_j(\theta + h) - f_j(\theta)}{h_i}, \quad h_i = 0,01 b_i, \quad (4.3.13)$$

за параметром θ_i . Саме такий вираз використовується у підпрограмі

I M програми *FUMILI* при числовому знаходженні похідних в середині інтервалу зміни параметра θ_i .

Критерій припинення ітераційного процесу (4.3.11)–(4.3.12) базується на тому факті, що обчислювати параметри з точністю, яка набагато перевищує статистичну точність, немає сенсу. А саме, оскільки оцінкою коваріаційної матриці похибок є матриця Z^{-1} , то як критерій припинення ітерацій використовується умова

$$\left| \Delta \theta_i^{(n)} \right| \leq \varepsilon \sqrt{Z_{ii}^{-1}}, \quad i = 1 \div m, \quad (4.3.14)$$

де Z^{-1} значення оберненої матриці до Z у точці $\theta^{(n-1)}$. Обчислення припиняються, якщо виконується ця нерівність, або коли число ітерацій досягло свого максимального значення $N3$. Точність, з якою знаходиться отримане значення $S_{\min} = S(\theta^{(n)})$, оцінюється як $g\Delta\theta^{(n)}/2$. Дану оцінку було отримано за лінійної екстраполяції $f_j(\theta)$ і подальшого розрахунку $S_{\min}$ як функції лінійно-екстрапольованих значень $f_j(\theta)$. На практиці оцінка виявляється дуже близькою до точної.

У *FUMILI* також можна використовувати загальні обмежувачі області зміни параметрів. Вони мають вигляд $\theta_{i_{\min}} \leq \theta_i \leq \theta_{i_{\max}}$, де $\theta_{i_{\min}}, \theta_{i_{\max}}$ – конста-

нти, що знаходяться в масивах $MI()$ $M()$ відповідно. Зокрема, такі обмежувачі дають змогу виключити області виродження, що відповідають нульовим значенням визначника матриці Z . У випадку, коли обмежувачі введені, то при побудові паралелепіпеда W , який визначає зміну кроку ітерації, ураховують не тільки b_i , але і границі $\theta_{i_{\min}}$, $\theta_{i_{\max}}$, і якщо $\Delta\theta^{(0)} + b_i > \theta_{i_{\max}}$, то відповідною границею W стає $\theta_{i_{\max}}$; аналогічно і для $\theta_{i_{\min}}$. Обчислюються тільки такі компоненти приросту $\Delta\theta_j^{(n)}$, що лежать у заданому інтервалі, а ті параметри, для яких $\Delta\theta_i^{(n)}$ перевищує W , фіксуються значеннями на границі $\theta_{i_{\min}}$ або $\theta_{i_{\max}}$.

Приклад застосування програми *FUMILI* разом із коротким описом величин, що в ній використовуються, наведено у додатку.

Розділ

ОЦІНКИ РІВНІВ ДЕТЕКТУВАННЯ В ЯДЕРНО ФІЗИЧНИХ ЕКСПЕРИМЕНТАХ МЕТОДОМ ПЕРЕВІРКИ СТАТИСТИЧНИХ ГІПОТЕЗ

1. Статистична гіпотеза схема її перевірки та загальне визначення рівнів детектування

У спектрометричних експериментах ядерної фізики зі спостереження деякої кількості подій (далі ядерно-фізичні експерименти) при прийнятті рішення щодо наявності деякої властивості (присутності корисного ефекту) зазвичай користуються методами перевірки статистичних гіпотез.

Статистична гіпотеза є деяким твердженням відносно значення одного чи декількох параметрів розподілу, або відносно самого вигляду розподілу. Гіпотеза, що розглядається як більш важливе твердження, коли воно відхиляється, називається нульовою гіпотезою і позначається як H_0 . Альтернативною гіпотезою H_1 є гіпотеза, що з нею конкурує, тобто базується на протилежному, суперечному твердженні. Нульовою гіпотезою, порівняно з альтернативною, має бути та гіпотеза, яку помилково відхилити більш небезпечно.

Критерій перевірки статистичної гіпотези є процедура розробки рішення про її прийняття або відхилення. Перевірка статистичної гіпотези, наприклад H_0 , відносно деяких властивостей елементів вибірки $x_1, \dots, x_N$ деякої випадкової величини X , полягає в тому, що використовується спеціально складена вибіркова характеристика (статистика критерію) $\xi(x_1, \dots, x_N)$, яка будується з елементів вибірки з відомим точним або наближеним розподілом $p(X)$. Далі згідно цього розподілу визначається критичне значення ξ^* , таке що при слушності гіпотези H_0 ймовірність

$$P(\xi \geq \xi^*) = \quad (5.1.1)$$

є малою такою, що подію $\xi \leq \xi^*$ можна вважати практично неможливою. Тому, якщо в деякому конкретному випадку виявляється відхилення $\xi \leq \xi^*$, то гіпотеза H_0 відхиляється, в той час як поява значень $\xi \leq \xi^*$ вважається сумісною з гіпотезою H_0 , яка в цьому випадку приймається. Величина ξ також називається критичною статистикою.

Таким чином, множина всіх можливих значень критичної статистики ξ розбивається на дві підмножини, що не перетинаються, а саме – критичну область (область відхилення гіпотези) $S \equiv \xi \leq \xi^*$ і область допустимих значень (область прийняття гіпотези) $\bar{S} \equiv \xi > \xi^*$. Якщо значення статистики критерію ξ фактично спостерігається, потрапляє у критичну область S , то гіпотезу H_0 відхиляють. Тобто критичною областю (або областю не прийняття гіпотези) критерію є та частина (вибіркового) простору значень критичної статистики, що приводить до відхилення гіпотези. Рівнем значущості критерію є ймовірність того, що цей критерій приведе до відхилення нульової гіпотези у випадку її істинності, тобто він є ймовірністю того, що при виконанні нульової гіпотези результати перевірки будуть знаходитися у критичній області.

При перевірці статистичних гіпотез, загалом, можливі два типи помилок, що зв'язані з прийняттям рішення. Відхилення нульової гіпотези H_0 у випадку її істинності називається помилкою першого роду. Прийняття нульової гіпотези, коли вона є хибною, називається помилкою другого роду. У цьому випадку в дійсності слушною є гіпотеза H_1 , але прийнято рішення про її відхилення, тобто про прийняття гіпотези H . Ймовірність зробити помилку першого роду позначають як α і, очевидно, вона визначається співвідношенням (5.1.1). Згідно визначення, що було наведено вище, ця ймовірність відкинути гіпотезу H_0 , коли вона є слушною, є рівнем значущості критерію. Ймовірність $(1 - \alpha)$ не зробити помилки першого роду, тобто прийняти гіпотезу H_0 , коли вона є вірною, іноді називають оперативною характеристикою критерію.

Ймовірність зробити помилку другого роду, тобто прийняти гіпотезу H_0 , коли вона є помилковою зазвичай позначають β . Ймовірність $(1 - \beta)$ не зробити помилку другого роду, тобто відхилити гіпотезу H_0 , коли вона є хибною, називається потужністю (або функцією потужності) критерію.

Якщо результати перевірки нульової гіпотези H_0 знаходяться у критичній області, то ця гіпотеза відхиляється і приймається альтернативна гіпотеза H_1 . Тому критична область має знаходитися там, де вона відповідає альтернативній гіпотезі. У загальному випадку, залежно від вигляду конкуруючої гіпотези H_1 , критичну область вибирають не тільки як правосторонню, тобто як $S \equiv \xi \geq \xi^*$ у формулі (5.1.1), але як лівосторонню і двосторонню. А саме, границі критичних областей ξ^* за заданого рівня значущості α визначаються з таких співвідношень: 1) для правосторонньої критичної області – за формулою(5.1.1), тобто

$$P(\xi \geq \xi^*) = \alpha,$$

для лівосторонньої критичної області

$$P(\xi < \xi^*) = \alpha, \quad (5.1.2)$$

для двосторонньої критичної області

$$P(\xi < \xi^*) = P(\xi \geq \xi^*) = \alpha/2. \quad (5.1.3)$$

Імовірності α і похибок першого і другого роду (типу) однозначно визначаються вибором критичної області. Критична область має бути такою, що для заданого рівня значущості α потужність критерію $(1 - \beta)$ була б максимальною. Задача побудови такої критичної області (або, як кажуть, побудови найпотужнішого критерію) для простих гіпотез розв'язується за допомогою теореми Неймана – Пірсона. Безперечно є бажаним зробити значення α і β як можна меншими. Однак виявляється, що за фіксованого обсягу вибірки це зробити неможливо.

Принцип перевірки статистичної гіпотези не дає логічного доведення її слушності або хибності. Прийняття гіпотези H_0 порівняно з альтернативною H_1 не означає, що ми впевнені в абсолютній правильності нульової гіпотези або що її твердження є найкращим та єдино можливим. Воно лише означає, що гіпотеза H_0 не має суперечностей із вибірковими даними. Таку ж властивість поряд із H_0 можуть мати й інші гіпотези більш того, можливо, що при збільшенні обсягу вибірки або при порівнянні гіпотези H_0 з деякою альтернативною H_2 гіпотеза H_0 буде відхилена. Тобто прийняття гіпотези H_0 необхідно розцінювати не так, що її твердження є раз і назавжди встановленим істинним фактом, а лише як досить правдоподібне твердження, що не

суперечить спостереженням. Зазвичай вважається, що можна знехтувати подіями, ймовірність появи яких менша за $\alpha = 0,0$ хоча іноді треба враховувати події, що можуть з'явитися з імовірністю $\alpha = 0,001$ і менше.

Схема перевірки гіпотез, що описана вище, базується на припущенні про закон розподілу генеральної сукупності, з якого визначається розподіл критичної статистики. Критерії перевірки таких гіпотез називаються *параметричними*. Якщо закон розподілу генеральної сукупності є невідомим, то відповідні критерії отримали назву *непараметричних*. Природно, що непараметричні критерії мають значно меншу потужність ніж параметричні. Це означає, що для збереження однакової потужності у випадку непараметричного критерію необхідно мати значно більший обсяг спостережень.

За своїм змістом статистичні гіпотези підрозділяють на декілька основних типів, а саме про: 1) рівність числових характеристик генеральних сукупностей 2) числові значення параметрів 3) закон розподілу 4) однорідність вибірок (тобто належність їх одній і тій же генеральній сукупності).

У спектрометричних експериментах зі спостереження якоїсь кількості подій вважається, що події є реалізацією деякої випадкової величини X , яка розподілена за законом Пуассона (1.4.44) тобто вважається, що події є незалежними і виникають із постійною швидкістю. Зазвичай визначають або кількість подій, або швидкість реєстрації кількості подій. Випадкову величину, генеральною сукупністю якої є кількість подій, позначаємо як N , а елементи її вибірки (тобто події, що спостерігаються) – як n_i ($X \equiv N, x_i \equiv n_i$). Випадкову величину, реалізацією якої є швидкість лічби (відліку) появи подій, позначаємо як R , а елементи її вибірки (швидкість спостереження подій) – як r_i

($X \equiv R, x_i \equiv r_i$). Крім потоку основних частинок, що дають змогу дослідити властивості зразку, на експериментальні прилади потрапляють сторонні частинки, що утворюють фон. Останній, наприклад, може з'являтися через сторонні домішки у зразку, до якого також можна віднести космічне випромінювання. В ядерно-фізичних експериментах зазвичай виконують дві серії вимірів: фонові виміри без зразка ($X \equiv X_0, x_i \equiv x_i^{(0)}$) і виміри зі зразком, властивості якого досліджуються ($X \equiv X_b, x_i \equiv x_i^{(b)}$).

Корисний ефект ($X \equiv X_n, x_i \equiv x_i^{(n)}$), що вивчається, знаходять (розраховують) як різницю повного і фонових сигналів (кількості подій): $X \equiv X_n = X_b - X_0$, $x_i \equiv x_i^{(n)} = x_i^{(b)} - x_i^{(0)}$.

Зазвичай у ядерно-фізичних експериментах приймають і досліджують дві статистичні гіпотези:

- 1) нульова гіпотеза H_0 – корисний ефект відсутній
- 2) альтернативна гіпотеза H_1 – корисний ефект присутній.

При перевірці гіпотез як критерію статистики (критичної статистики) ξ використовують (вибіркові чи теоретичні) середні значення ($\xi \equiv M[X], < x = \sum_{i=1}^N x_i / N$) або для кількості подій $v = M[N], < n$, або для швидкості $\rho = M[R], < r$ відліку (лічби) подій. Тоді прийняття нульової гіпотези означає перевірку твердження, що середнє значення корисного ефекту ($\xi_n \equiv M[X_n], < x^{(n)}$) дорівнює нулю, а альтернативної гіпотези – $\xi_n = a$, де $a \neq 0$ – деяка стала тобто

$$\begin{aligned} H_0 \quad \xi_n^* = 0 = \xi_0^*, \\ H_1 \quad \xi_n^* = a \neq 0, \end{aligned} \quad (5.1.4)$$

де ξ_n^* – критична амплітуда середнього значення корисного сигналу (критична границя між областями прийняття гіпотез H_0 і H_1).

У цьому випадку ймовірність α зробити помилку першого роду (α -помилку) є ймовірністю того, що експериментатор хибно зареєструє (зареєстрував) відсутній ефект (тобто прийме фон за корисний сигнал), а ймовірність зробити похибку другого роду (β -помилку) є ймовірністю того, що він не зареєструє (не зареєстрував) присутній корисний сигнал.

Як зазначалося вище, при статистичній перевірці статистичних гіпотез визначають критичну границю між областями прийняття гіпотез H_0 і H_1 . Критична границя і значення a , яке відповідає їй та заданій ймовірності помилки другого роду β , отримали у випадку ядерного лічильного експерименту назву рівнів (границь): детектування та розпізнавання, а саме, використовують такі визначення:

- 1) Границя розпізнавання $x_n^* \equiv \xi_n^*(H = H_0)$ корисного ефекту з наперед заданою ймовірністю α помилки першого роду (є критичною границею прийняття гіпотези H_1) – це таке (середнє) значення сигналу від зразка, при

якому слухна гіпотеза про те, що зареєстровано лише фоновий ефект, а отже корисний ефект відсутній, приймається з імовірністю $1 - \alpha$.

2) Границя детектування $\xi_n^* \equiv \xi_n^*(H = H_1)$ корисного ефекту з наперед заданою ймовірністю помилки другого роду β (є критичною границею прийняття гіпотези H_1) – таке (середнє) значення корисного ефекту, при якому слухна гіпотеза про те, що корисний ефект присутній, буде прийнята з імовірністю $1 - \alpha$.

Ситуація, коли нульова гіпотеза буде помилково прийнята, є можливою, якщо, унаслідок недостатнього обсягу вибірки (наприклад, малого часу вимірювань), недоліків методики експерименту та інших причин, корисний сигнал є меншим або рівним границі розпізнавання.

Введення та використання поняття границі розпізнавання в експерименті, пов'язано з існуванням фонового ефекту, а отже характеризує середовище, у якому існує корисний ефект. Необхідність введення границі детектування пов'язано з тим, що вимірюється повний ефект, а потім вже на основі отриманих фонового і повного сигналів оцінюється корисний сигнал (який, у свою чергу, є лише експериментальною оцінкою корисного ефекту), а отже границя детектування характеризує здатність при вимірюванні зареєструвати присутність корисного ефекту. Тому границя детектування, зокрема, використовується для встановлення вимог до чутливості вимірювання в галузі радіаційної безпеки, а саме, вимірювання має бути проведено таким чином, щоб границя детектування в межах цього процесу була нижчою за деякий регулюючий рівень, встановлений законодавчими документами або іншими вимогами (наприклад, допустимими рівнями забруднення).

За інших незмінних умов вимога зменшення β призводить до збільшення границі розпізнавання, а якщо при цьому вимагати, щоб границя детектування залишалась незмінною, – до збільшення α . Якщо ж при зменшенні β вимагати незмінності або зменшення α , то збільшиться і границя детектування, а отже даний вимірювальний процес може стати непридатним для вимірювання слабких ефектів або ефектів, для яких встановлені регулюючі рівні. Єдиним способом зменшити одночасно ймовірності α і β , не зменшуючи чутливості даного вимірювального процесу, є збільшення вибірки, яке досягається за рахунок збільшення часу виміру, тобто за рахунок збільшення витрат на вимірювання.

Якщо істинне значення корисного ефекту буде меншим за границю детектування, але більшим за границю розпізнавання, то зберігається ймовірність його реєстрації, але при цьому збільшується ймовірність зробити помилку другого роду.

.2. Оцінка рівнів детектування при визначенні кількості відліку подій

Визначення рівнів детектування можна розглядати як завдання з перевірки гіпотез про математичне сподівання (див., наприклад, розв'язок такої задачі у книзі В.А.Фігуріна і В.В. Оболонкіна, 11.3). Уперше вирази для рівнів детектування у спектрометричних експериментах у такому підході було проаналізовано Каррі (, 1968). Дотримуючись роботи Каррі, позначимо границі розпізнавання і детектування n_n^* і v_n^* відповідно. У припущенні нормального розподілу (середнього значення) спостережного сигналу за відомої дисперсії були отримані наступні вирази для границі розпізнавання (середнього значення) сигналу та границі детектування

$$n_n^* = k_{1-\alpha} \cdot \sigma_0, \quad (5.2.1)$$

$$v_n^* = n_n^* + k_\beta \sigma, \quad (5.2.2)$$

де $k_{1-\alpha}$ – $(1-\alpha)$ -й (правосторонній), а k_β – -й (лівосторонній) квантили стандартного нормального розподілу $N(0,1)$ σ_0 і σ – середньоквадратичні відхилення фоновому і повного сигналу.

Вираз (5.2.1) можна отримати з умови перевірки нульової гіпотези про відсутність корисного сигналу (тобто рівність його математичного сподівання нулю $v = M[N] = 0$) із заданим рівнем значущості α у припущенні нормального розподілу середнього значення $x \equiv < n$ спостережного сигналу з відомою дисперсією σ_0^2 . Дійсно, нехай випадкові значення $x \equiv < n$ розподілені за нормальним законом $N(0, \sigma_0^2)$, тобто розподілені нормально з дисперсією σ_0^2 і нульовим математичним сподіванням $M(< n) = 0$. Тоді згідно з визначенням границі розпізнавання маємо

$$1 - \beta = P(< n \leq n_n^*) = \int_{-\infty}^{n_n^*} \frac{1}{\sqrt{2} \sigma_0} \cdot e^{-\frac{x^2}{2\sigma_0^2}} dx =$$

$$= \frac{1}{\sqrt{2}} \int_{-\infty}^{\frac{n_n^*}{\sigma_0}} e^{-z^2/2} dz \equiv \frac{1}{\sqrt{2}} \int_{-\infty}^{k_{1-\alpha}} e^{-z^2/2} dz. \quad (5.2.3)$$

Звідки знаходимо (5.2.1). Значення квантилі $k_{1-\alpha}$ стандартного нормального розподілу $N(0,1)$ для деяких значень ймовірності наведено у табл. 5.1.

Таблиця 5.1

Значення квантилі $k_{1-\alpha}$ стандартного нормального розподілу $N(0,1)$ для заданої ймовірності α

α	$= 1 - \alpha$	$k_{1-\alpha}$
0,1586	0,8414	1,000
0,1000	0,9000	1,282
0,0500	0,9500	1,645
0,0250	0,9750	1,960
0,0228	0,9772	2,000
0,0100	0,9900	2,326
0,0050	0,9950	2,576
0,0014	0,9986	3,000
0,0010	0,9999	3,090

Вираз (5.2.2) для границі детектування можна отримати з умови перевірки альтернативної гіпотези з урахуванням того, що в цьому випадку математичне сподівання середнього значення корисного сигналу дорівнює v_n^* у припущенні розподілу середнього значення $x \equiv < n$ корисного спостережного сигналу з відомою дисперсією σ^2 за нормальним законом $N(v_n^*, \sigma^2)$. Згідно визначення границі детектування знаходимо співвідношення

$$\begin{aligned}
 &= P(< n \leq n_n^*) = \int_{-\infty}^{\frac{n_n^*}{\sigma}} \frac{1}{\sqrt{2}} \cdot e^{-\frac{(x-v_n^*)^2}{2\sigma^2}} dx = \\
 &= \frac{1}{\sqrt{2}} \int_{-\infty}^{\frac{n_n^* - v_n^*}{\sigma}} e^{-z^2/2} dz \equiv \frac{1}{\sqrt{2}} \int_{-\infty}^{k_{\beta}} e^{-z^2/2} dz. \quad (5.2.4)
 \end{aligned}$$

Ця формула означає, що величина $(n_n^* - v_n^*) / \sigma$ є квантилем k_β .

Відзначимо, що на практиці припущення про нормальність розподілу середньовибіркових значень $\bar{X} = \sum_{i=1}^N n_i / N$ досить добре виконується, оскільки відомо, що незалежно від вигляду розподілу для випадкової величини N , що відповідає виборці $n_1, \dots, n_N$, при $N \geq 0$ розподіл для $\bar{X}$ практично збігається з нормальним.

Формулу (5.2.1) можна переписати у більш детальному вигляді. Ураховуючи те, що кількість відліків розподілено за законом Пуассона, маємо для дисперсії (середнього значення) повного сигналу

$$\sigma^2 = M(\bar{X}) = v_b = v_n + v_0, \quad (5.2.5)$$

де v_b, v_n, v_0 – середні значення відповідно повного, корисного та фонового ефектів. Якщо фоновий сигнал визначався з m експериментів, то його дисперсія має вигляд

$$\sigma_0^2 = M(\bar{X}) = M(\bar{X}) / m \equiv v_0 = v_0 / m. \quad (5.2.6)$$

У цьому випадку дисперсія випадкової величини повного корисного ефекту дорівнює

$$\begin{aligned} \sigma^2 &= \sigma^2(\bar{X}) \equiv \sigma^2(\bar{X}) = \\ &= \sigma^2(\bar{X}) + \sigma^2(\bar{X}) = (v_n + v_0) + v_0 / m. \end{aligned} \quad (5.2.7)$$

Тоді з рівнянь (5.2.1), (5.2.2) і (5.2.7) знаходимо кінцевий вираз для границі розпізнавання:

$$n_n^* = k_{1-\alpha} (v_0 + v_0 / m)^{1/2}. \quad (5.2.8)$$

Після підстановки (5.2.7) з $v_n = v_n^*$ у формулу (5.2.3) отримуємо такий вираз для границі детектування

$$v_n^* = n_n^* + k_\beta \sigma = n_n^* + k_\beta \left(v_n^* + v_0 + \frac{v_0}{m} \right)^{1/2} = n_n^* + k_\beta \left(v_n^* + \sigma_0^2 \right)^{1/2}. \quad (5.2.9)$$

З урахуванням (5.2.8) знаходимо таке співвідношення між границями розпізнавання та детектування

$$v_n^* = n_n^* + \frac{k_\beta^2}{2} \left(1 + \left(1 + \frac{n_n^*}{k_\beta^2} + \frac{n_n^{*2}}{k_{1-\alpha}^2 k_\beta^2} \right)^{1/2} \right). \quad (5.2.10)$$

5.3. Асимптотичні оцінки граничних рівнів швидкості відліку сигналів

Проаналізуємо можливість статистичних оцінок рівнів детектування і можливість використання властивостей "великих вибірок" у випадку ядерно-фізичних експериментів з достатньо великою тривалістю вимірів. Нехай проводиться один за одним i експериментів ($i \in [1, N]$) з однаковим часом виміру t , і в кожному з них маємо n_i відліків. Тоді для середньоарифметичних значень кількості ($\langle n \rangle$) і швидкості ($\langle r \rangle$) відліків маємо

$$\langle n \rangle = \frac{1}{N} \sum_{i=1}^N n_i = \frac{n_{\Sigma}}{N}, \quad \langle r \rangle = \frac{1}{N} \sum_{i=1}^N \frac{n_i}{t} = \frac{n_{\Sigma}/N}{t},$$

$$\langle r \rangle \xrightarrow{N \rightarrow \infty} M\left(\frac{1}{N} \sum_i \frac{n_i}{t}\right) = M(R) = \rho, \quad (5.3.1)$$

де $R = N/t$ – випадкова величина, що визначає швидкість лічби, а ρ – її математичне сподівання. Очевидно, що при достатньо великих N , в одиночному експерименті із часом лічби $T = tN$ середня швидкість є такою ж, як і у формулі (5.3.1)

$$r = \frac{n_{\Sigma}}{T} = \frac{n_{\Sigma}/N}{t}, \quad r \xrightarrow{T \rightarrow \infty} M(R_T) = \rho, \quad (5.3.2)$$

де R_T визначає випадкову швидкість лічби в одиночному експерименті як і кількість відліків їхня швидкість також буде розподілена за законом Пуассона. Отже, для ядерного лічильного експерименту з достатньо великим часом виміру ми можемо використовувати результати математичної статистики з розгляду асимптотичної поведінки емпіричних розподілів випадкових величин. Звідки згідно з центральною граничною теоремою та її наслідками можна вважати, що відхилення середньої швидкості лічби корисного сигналу ($r_b - r_0$) від її теоретичного математичного сподівання ($\rho_b - \rho_0$) асимптотично розподілено за нормальним законом $N(0,1)$:

$$\frac{(r_b - r_0) - (\rho_b - \rho_0)}{\sigma} \sim N(0,1), \quad (5.3.3)$$

де $\sigma \equiv \sigma(r_b - r_0)$ – асимптотична дисперсія випадкової величини корисного ефекту.

Розглянемо визначення рівнів детектування при дослідженні слушності нульової гіпотези H_0 , тобто ситуації коли оцінений з експериментальних даних корисний сигнал є меншим за граничний рівень і середнє значення корисного ефекту v_n приймається рівним нулю. Проаналізуємо два можливих випадки виконання нульової гіпотези (І.М.Каденко, А.Ю.Мелешин, 2001). У першій ситуації вважаємо, що нульова гіпотеза H_0 є вірною. Тоді маємо $\rho_n = \rho_b - \rho_0 = 0$, і тому в цьому випадку за законом $N(0, 1)$ розподілена різниця швидкостей $r_n^* = r_b - r_0$:

$$\frac{(r_b - r_0)}{\sigma_0} \sim N(0, 1), \quad (5.3.4)$$

де $\sigma_0 \equiv \sigma(r_b - r_0)|_{\rho_n=0}$ – середньоквадратичне відхилення швидкості відліку за нульового значення її математичного сподівання. Найбільше значення величини $r_n^* = r_b - r_0$, при якому ще приймається правильне рішення про справедливість гіпотези H_0 , відповідає ймовірності зробити помилку першого роду з ймовірністю α . Отже, вона повинна задовольняти співвідношення

$$1 - \alpha = P\left(\frac{(r_b - r_0)}{\sigma_0} \leq \frac{r_n^*}{\sigma_0} = k_{1-\alpha}\right), \quad (5.3.5)$$

і тому є границею розпізнавання. Згідно з цим виразом маємо для границі розпізнавання швидкості лічби:

$$r_n^* = k_{1-\alpha} \sigma_0. \quad (5.3.6)$$

Асимптотичну дисперсію корисного сигналу у правій частині (5.3.6) знаходимо, прирівнюючи її відповідній теоретичній дисперсії за нульової величини математичного сподівання корисного ефекту:

$$\sigma_0^2 \equiv \sigma^2(\rho_b - \rho_0)|_{\rho_n=0} = \sigma^2 \left(\frac{N_b}{t_b} - \frac{N_0}{t_0} \right) \Big|_{\rho_n=0} = \frac{v_b}{t_b^2} + \frac{v_0}{t_0^2}, \quad (5.3.7)$$

де t_b , t_0 – тривалості експерименту і виміру фону відповідно v_0 – кількість відліку за час виміру фону, а v_b – повна середня кількість відліку за повний час тривалості експерименту:

$$v_b = v_0 t_b / t_0 . \quad (5.3.8)$$

Використаємо середнє значення n_b повного сигналу (за нульового значення його математичного сподівання) для оцінки v_0 . Вважаємо, що статистичну оцінку для v_0 можна подати у вигляді

$$v_0 = aN_b + bN_0 , \quad (5.3.9)$$

де коефіцієнти задовольняють умові її незсуненості та ефективності (2.2.6), тобто знаходяться з таких співвідношень:

$$M(v_0) = av_0 \frac{t_b}{t_0} + bv_0 = v_0 , \quad (5.3.10)$$

$$\sigma^2(v_0) = a^2 v_0 \frac{t_b}{t_0} + b^2 v_0 \rightarrow \min . \quad (5.3.11)$$

Звідки отримуємо вираз для коефіцієнтів

$$a = b = \frac{t_0}{t_b + t_0} , \quad v_0 = \frac{(N_b + N_0)t_0}{t_b + t_0} \quad (5.3.12)$$

і співвідношення

$$v_0 = \frac{(n_b + n_0)t_0}{t_b + t_0} \quad v_0 . \quad (5.3.13)$$

для статистичної оцінки v_0 як найкращої оцінки v_0 .

З урахуванням формул (5.3.7), (5.3.8), (5.3.13) для дисперсії σ_0^2 маємо

$$\sigma_0^2 = \frac{v_0}{t_0} \left(\frac{1}{t_b} + \frac{1}{t_0} \right) = \frac{(n_b + n_0)}{t_b t_0} = \frac{r_b}{t_0} + \frac{r_0}{t_b}, \quad (5.3.14)$$

де $r_b = n_b / t_b$ і $r_0 = n_0 / t_0$ – середні швидкості лічби повного і фонових сигналів за нульового значення математичного сподівання швидкості корисного сигналу.

Найбільше значення r_b , що задовольняє гіпотезі H_0 , задовольняє рівність $r_n^* = r_b - r_0$ (див. (5.3.5)), тобто $r_b = r_n^* + r_0$. Тоді

$$\sigma_0^2 \equiv \sigma^2(r_b - r_0) \Big|_{\rho n=0} = \frac{r_n^* + r_0}{t_0} + \frac{r_0}{t_b} = \frac{r_n^*}{t_0} + r_0 \left(\frac{1}{t_b} + \frac{1}{t_0} \right). \quad (5.3.15)$$

Підставляючи (5.3.15) в (5.3.6), заходимо:

$$r_n^* = k_{1-\alpha} \sqrt{\frac{r_n^*}{t_0} + r_0 \left(\frac{1}{t_b} + \frac{1}{t_0} \right)}. \quad (5.3.16)$$

Звідки отримуємо остаточний вираз для границі розпізнавання r_n^* :

$$\begin{aligned} r_n^* &= k_{1-\alpha} \left(\frac{k_{1-\alpha}}{2t_0} + \sqrt{\left(\frac{k_{1-\alpha}}{2t_0} \right)^2 + r_0 \left(\frac{1}{t_b} + \frac{1}{t_0} \right)} \right) = \\ &= \frac{k_{1-\alpha}^2}{2t_0} \left(1 + \sqrt{1 + r_0 \left(\frac{2t_0}{k_{1-\alpha}} \right)^2 \left(\frac{1}{t_b} + \frac{1}{t_0} \right)} \right). \end{aligned} \quad (5.3.17)$$

Оскільки $r_0 (2t_0 / k_{1-\alpha})^2 / t_0 = n_0 / (k_{1-\alpha} / 2)^2$, то у випадку великої кількості відліків фону ($n_0 \gg (k_{1-\alpha} / 2)^2$) значення для границі розпізнавання можна наближено обчислити за формулою

$$r_n^* \simeq k_{1-\alpha} \sqrt{r_0 \left(\frac{1}{t_b} + \frac{1}{t_0} \right)}. \quad (5.3.18)$$

Відзначимо, що співвідношення (5.3.18) рекомендується німецьким стандартом IN 25482-1:1989 для розрахунку границі розпізнавання при проведенні лічильних вимірів без урахування впливу приготування проб.

Тепер розглянемо випадок, коли гіпотеза H_0 є хибною, але приймається. У такій ситуації при обробці експериментальних результатів припускаються помилки другого роду і помилкову нульову гіпотезу не відкидають навіть за наявності відмінного від нуля середнього значення корисного ефекту. Нехай математичне сподівання корисного ефекту є $\rho_n = \rho_b - \rho_0 = \rho_n^*$, тоді згідно з (5.3.3) відносне відхилення $((r_b - r_0) - \rho) / \sigma$ розподілено за нормальним законом $N(0,1)$:

$$\frac{(r_b - r_0) - \rho_n^*}{\sigma} \sim N(0,1). \quad (5.3.19)$$

де $\sigma \equiv \sigma(r_b - r_0) \Big|_{\rho_n = \rho_n^*}$ – середньоквадратичне відхилення швидкості відліків

з математичним сподіванням $\rho_n = \rho_n^*$.

Найбільше значення різниці $r_b - r_0 = r_n^* \Big|_{\rho_n = \rho_n^*}$, для якого ще буде прийняте помилкове рішення про справедливість гіпотези H_0 , відповідає ймовірності зробити помилку другого роду:

$$= P \left(\frac{(r_b - r_0) - \rho_n^*}{\sigma} \leq \frac{r_n^* \Big|_{\rho_n = \rho_n^*} - \rho_n^*}{\sigma} = k_\beta \right). \quad (5.3.20)$$

Звідки знаходимо такий вираз для границі детектування ρ_n^* :

$$\rho_n^* = r_n^* \Big|_{\rho_n = \rho_n^*} - k_\beta \sigma = r_n^* \Big|_{\rho_n = \rho_n^*} + k_{1-\beta} \sigma \quad (5.3.21)$$

Отже, як границя розпізнавання (5.3.6), так і границя детектування (5.3.21) визначаються середньоквадратичним відхиленням асимптотичного

нормального розподілу (асимптотичним стандартним відхиленням випадкової величини корисного ефекту).

Для обчислення границі детектування за формулою (5.3.21) необхідно оцінити границю розпізнавання $r_n^* \Big|_{\rho_n = \rho_n^*}$ з відмінним від нуля математичним сподіванням корисного сигналу. Це можна зробити за допомогою виразу, аналогічного (5.3.6), з тією різницею, що за оцінки асимптотичної дисперсії у виразі вигляду (5.3.14) замість r_b використаємо його оцінку $r_b^* = \rho_n^* + r_0$, тобто маємо

$$r_n^* \Big|_{\rho_n = \rho_n^*} = k_{1-\alpha} \sqrt{\frac{\rho_n^*}{t_0} + r_0 \left(\frac{1}{t_b} + \frac{1}{t_0} \right)}. \quad (5.3.22)$$

Далі, для обчислення ρ_n^* згідно з (5.3.21), знаходимо асимптотичну оцінку σ для теоретичної дисперсії з точного виразу для теоретичної дисперсії середнього значення корисного сигналу

$$\begin{aligned} \sigma^2 \left(\frac{N_b}{t_b} - \frac{N_0}{t_0} \right) \Big|_{\rho_n = \rho_n^*} &= \left(\frac{v_b}{t_b^2} + \frac{v_0}{t_0^2} \right) \Big|_{\rho_n = \rho_n^*} = (v_b = v_n^* + v_0 \frac{t_b}{t_0}) = \\ &= \frac{v_n^*}{t_b^2} + \frac{v_0}{t_0} \left(\frac{1}{t_b} + \frac{1}{t_0} \right) \end{aligned} \quad (5.3.23)$$

із заміною точної середньої швидкості v_0/t_0 лічби фонового ефекту на його оцінку r_0 , що є середньою швидкістю лічби фонового сигналу, яка отримана у фоновому експерименті. У результаті маємо таку асимптотичну оцінку для теоретичної дисперсії

$$\sigma^2 \equiv \sigma^2 (r_b - r_0) \Big|_{\rho_n = \rho_n^*} = \frac{v_n^*}{t_b^2} + r_0 \left(\frac{1}{t_b} + \frac{1}{t_0} \right) = \frac{\rho_n^*}{t_b} + r_0 \left(\frac{1}{t_b} + \frac{1}{t_0} \right). \quad (5.3.24)$$

Після підстановки формул (5.3.22) і (5.3.24) у (5.3.21) знаходимо нелінійне рівняння визначення границі детектування ρ_n^* :

$$\rho_n^* = k_{1-\alpha} \sqrt{\frac{\rho_n^*}{t_0} + r_0 \left(\frac{1}{t_b} + \frac{1}{t_0} \right)} + k_{1-\beta} \sqrt{\frac{\rho_n^*}{t_b} + r_0 \left(\frac{1}{t_b} + \frac{1}{t_0} \right)}. \quad (5.3.25)$$

Розв'язок цього рівняння можна знайти методом ітерацій, наприклад, починаючи з $\rho_n^* = 0$.

У випадку однакової тривалості експерименту і виміру фону $t_b = t_0 = t$, рівняння (5.3.25) значно спрощується

$$\rho_n^* = (k_{1-\alpha} + k_{1-\beta}) \sqrt{\frac{\rho_n^*}{t} + r_0 \frac{2}{t}}. \quad (5.3.26)$$

Звідки знаходимо

$$r_n^* = \frac{(k_{1-\alpha} + k_{1-\beta})^2}{2t} + (k_{1-\alpha} + k_{1-\beta}) \sqrt{\left(\frac{k_{1-\alpha} + k_{1-\beta}}{2t} \right)^2 + r_0 \frac{2}{t}}. \quad (5.3.27)$$

За великої кількості відліків фону ($n_0 \rightarrow (k_{1-\alpha} + k_{1-\beta})^2 / \epsilon$) оцінка (5.3.25) для границі детектування приймає вигляд

$$r_n^* = (k_{1-\alpha} + k_{1-\beta}) \sqrt{r_0 \frac{2}{t}}. \quad (5.3.28)$$

Відзначимо, що формули (5.3.6), (5.3.21) для границь розпізнавання і детектування були отримані з використанням властивості (5.3.3) асимптотичної нормальності емпіричних розподілів випадкових величин, що належать великій вибірці. У результаті значення границь розпізнавання та детектування визначалися асимптотичними значеннями дисперсій відповідних випадкових величин, і тому вони мають значення асимптотичних оцінок відповідних граничних рівнів сигналів для вибірок великого обсягу.

Необхідно також зазначити, що, користуючись різними виразами теорії великих вибірок, можна отримати різні вирази, які будуть збігатися для нескінченно великих обсягів вибірок. Це пов'язано з різною точністю виконання різних асимптотичних виразів у випадку їх використання для вибірок обмеженого обсягу.

Знайдемо, наприклад, вираз для граничних рівнів сигналів за допомогою теореми про перетворення статистик, згідно з якою, якщо випадкова величи-

на асимптотично розподілена за нормальним законом, то і довільна неособлива функція від неї буде також розподілена за нормальним законом з перенормованими середнім значенням і дисперсією (див., наприклад, С.Р.Рао, 1968). Нехай $f(r_b - r_0)$ – деяка диференційована і відмінна від нуля (у відсутність середнього значення швидкості лічби корисного сигналу) функція, тоді відповідно до властивості (5.3.3) асимптотичної нормальності таких швидкостей і вищезгаданої теореми маємо, що випадкові величини $f(r_b - r_0)$ асимптотично розподілені за нормальним законом типу

$$f(r_b - r_0) \sim N \left(f(\rho_b - \rho_0), \left[\frac{df(\rho_b - \rho_0)}{d(\rho_b - \rho_0)} \right]^2 \sigma^2(R_b - R_0) \right), \quad (5.3.29)$$

де R_b, R_0 – випадкові величини, що визначають швидкості відліків повного і фоновому сигналів $R_b \equiv N_b / t, R_0 \equiv N_0 / t$.

Оскільки $\rho_b - \rho_0 = \rho_n$, то (5.3.29) можна переписати у вигляді

$$f(r_b - r_0) - f(\rho_n) \sim N \left(0, \left[\frac{df(\rho_n)}{d\rho_n} \right]^2 \sigma^2(R_b - R_0) \right). \quad (5.3.30)$$

Звідки маємо, що за умови

$$\frac{df(\rho_n)}{d\rho_n} \sigma(R_b - R_0) = 1 \quad (5.3.31)$$

різниця випадкових величин $f(r_b - r_0) - f(\rho_n)$ асимптотично розподіляється за стандартним нормальним законом $N(0, 1)$:

$$f(r_b - r_0) - f(\rho_n) \sim N(0, 1). \quad (5.3.32)$$

У такому випадку досить просто проаналізувати слухність виконання нульової гіпотези. Зокрема, з властивості (5.3.32) випливає, що коли $\rho_n = \rho_b - \rho_0 = 0$, то гіпотеза H_0 має бути відкинutoю, якщо виконується умова

$$f(r_b - r_0) - f(0) \leq k_{1-\alpha}. \quad (5.3.33)$$

Рівність у цьому співвідношенні має місце для $r_b - r_0 = r_n^*$:

$$f(r_n^*) - f(0) = k_{1-\alpha}. \quad (5.3.34)$$

Звідки отримуємо для границі розпізнавання

$$r_n^* = f^{-1}(k_{1-\alpha} + f(0)). \quad (5.3.35)$$

З властивості (5.3.32) також випливає, що з $\rho_n = \rho_b - \rho_0 = \rho_n^*$ гіпотеза H_0 повинна бути прийнятою, якщо виконується співвідношення:

$$f(r_b - r_0) - f(\rho_n^*) \leq k_\beta. \quad (5.3.36)$$

Рівність у цьому виразі досягається за $r_b - r_0 = r_n^*$:

$$f(r_n^*) - f(\rho_n^*) = k_\beta. \quad (5.3.37)$$

Звідки знаходимо загальний вираз для границі детектування:

$$\rho_n^* = f^{-1}(k_{1-\beta} + f(r_n^*)). \quad (5.3.38)$$

З урахуванням (5.3.34), отримуємо:

$$\rho_n^* = f^{-1}\left(k_{1-\beta} + f\left(r_n^* \middle|_{\rho_n = \rho_n^*}\right)\right) = f^{-1}(k_{1-\alpha} + k_{1-\beta} + f(0)). \quad (5.3.39)$$

Тепер знайдемо загальний вираз для функції $f(\rho_n)$ зі співвідношення (5.3.31): $df(\rho_n)/d\rho_n = 1/\sigma(R_b - R_0)$. Спочатку записуємо співвідношення, що визначає теоретичну дисперсію середнього значення корисного сигналу. У результаті маємо

$$\begin{aligned}\sigma^2(R_b - R_0) &= \sigma^2\left(\frac{N_b}{t_b} - \frac{N_0}{t_0}\right) = \frac{v_b}{t_b^2} + \frac{v_0}{t_0^2} = \\ &= \frac{v_n}{t_b^2} + \frac{v_0}{t_0} \left(\frac{1}{t_b} + \frac{1}{t_0}\right) = \frac{\rho_n}{t_b} + \rho_0 \left(\frac{1}{t_b} + \frac{1}{t_0}\right),\end{aligned}\quad (5.3.40)$$

де враховано рівність $v_b = v_n + v_0 t_b / t_0$. Після підстановки (5.3.40) у співвідношення (5.3.31) і подальшого інтегрування, знаходимо:

$$f(\rho_n) = \int \frac{dx}{\sqrt{\frac{x}{t_b} + \rho_0 \left(\frac{1}{t_b} + \frac{1}{t_0}\right)}} = 2t_b \sqrt{\frac{\rho_n}{t_b} + \rho_0 \left(\frac{1}{t_b} + \frac{1}{t_0}\right)}.\quad (5.3.41)$$

Використовуючи формулу (5.3.39), отримуємо рівняння для границі детектування ρ_n^* :

$$2t_b \sqrt{\frac{\rho_n^*}{t_b} + \rho_0 \left(\frac{1}{t_b} + \frac{1}{t_0}\right)} = k_{1-\alpha} + k_{1-\beta} + 2t_b \sqrt{\rho_0 \left(\frac{1}{t_b} + \frac{1}{t_0}\right)}.\quad (5.3.42)$$

Звідки знаходимо кінцевий вираз для границі детектування:

$$\rho_n^* = \rho_n^{(0)} + (k_{1-\alpha} + k_{1-\beta}) \sqrt{r_0 \left(\frac{1}{t_b} + \frac{1}{t_0}\right)},\quad (5.3.43)$$

де

$$\rho_n^{(0)} = \frac{1}{t_b} (k_{1-\alpha} + k_{1-\beta})^2,\quad (5.3.44)$$

і як теоретичне середнє значення фонового ефекту ρ_0 використано його оцінку r_0 , що відповідає спостереженому середньому значенню швидкості лічби фонового сигналу.

Зауважимо, що за великої кількості відліків фону ($n_0 \gg (k_{1-\alpha} + k_{1-\beta})^2$) внесок першого доданку у виразі (5.3.41) є малим і ним можна знехтувати.

Відзначимо, що німецьким стандартом IN 25482-1:1989 для розрахунку границі детектування також рекомендовано вираз вигляду (5.3.43), але зі значенням $\rho_n^{(0)}$, що дорівнює

$$\rho_n^{(0)} = -\frac{1}{k_{1-\alpha} + k_{1-\beta}} \left(\frac{1}{t_b} + \frac{1}{t_0} \right). \quad (5.3.45)$$

Цю відмінність можна вважати не суттєвою, ураховуючи асимптотичний характер оцінки (5.3.43). У випадку однакової тривалості експерименту і виміру фону формули для рівнів детектування (5.3.27) та (5.3.43) з (5.3.45) збігаються.

Таким чином, за великої кількості відліків фону рівень детектування можна обчислювати за формулою

$$\rho_n^* = (k_{1-\alpha} + k_{1-\beta}) \sqrt{r_0 \left(\frac{1}{t_b} + \frac{1}{t_0} \right)}. \quad (5.3.46)$$

Цей вираз збігається з нульовою ітерацією нелінійного рівняння (5.3.25).

Як висновок відзначимо, що оцінки рівнів розпізнавання та детектування середніх значень сигналів та їхніх швидкостей лічби дають змогу дослідити й оптимізувати експериментальні методики та обладнання для надійного визначення результатів експерименту.

ДОДАТОК

ПІДПРОГРАМА ДЛЯ ЗНАХОДЖЕННЯ ПАРАМЕТРІВ МЕТОДОМ НАЙМЕНШИХ КВАДРАТІВ

Нижче наведено текст набору комплексу програм оцінки параметрів за МНК з використанням підпрограми *FUMILI*. Розглянуто приклад знаходження параметрів $\beta = (\beta_1, \beta_2, \beta_3)$ функції вигляду

$$f(x, y) = \beta_1 + \beta_2 x + \beta_3 x^2. \quad (A1)$$

Експериментальні дані було обчислено за формулою

$$y_j = 1 + 5 x_j, \quad j = 1 \div 20, \quad x = 1 \div 20, \Delta x = 1, \quad (A2)$$

а середньоквадратичні відхилення вважалися рівними $\sigma_j = 0,1 y_j$. Початкові значення параметрів $\beta_1, \beta_2, \beta_3$ бралися рівними нулю.

У результаті обчислень отримано такі оцінки:

$$\beta_1 = 1,00; \quad \beta_2 = 5,00; \quad \beta_3 = 0,122 \cdot 10^{-6},$$

що повністю узгоджуються зі значеннями ($\beta_1 = 1; \quad \beta_2 = 5; \quad \beta_3 = 0$), що були використані в (A2) як експериментальні дані.

Текст програм

С ГОЛОВНА ПРОГРАМА ЗНАХОДЖЕННЯ ПАРАМЕТРІВ ЗА МНК

С

С ВЕЛИЧИНИ, ЩО ЗАДАЮТЬСЯ ЗА ДОПОМОГОЮ COMMON БЛОКІВ:

С COMMON /A/A(20)/PL/PL(20)/EXDA/EXDA(600)/NED/N,NS

С COMMON /AU/AMAX(20)/AL/AMIN(20)

С N - кількість даних (значень функції $y_j, j = 1 \div$), що
С наближуються;

С NS = K + 2, K - розмірність багатовимірного простору векторного
С аргументу

С $x_j = \{x_{j,1}, \dots, x_{j,K}\}$ - функції, що досліджується;

С в одновимірному випадку NS = 1 + 2 = 3;

С A(20) - масив параметрів $\beta = \{\beta_1, \dots, \beta_M\}$, що оцінюються МНК:

ДОДАТОК

С спочатку в масив потрібно занести значення нульового наближення,
С а після розрахунків у масиві знаходяться результати обчислень;
С $PL(20) (\equiv b_i)$ - абсолютні значення початкових обмежувачів кроків;
С параметр $A(i)$ обчислюється за умови $PL(i) > 0$;
С при $PL(i) < 0$ або $= 0$ параметр $A(i)$ = початковому значенню = CONST;
С $EXDA(600)$ - масив значень функції y_j , її похибок σ_j та
С значень компонент векторного аргументу $\vec{x}_j$;
С $EXDA(1)$ - експериментальне значення функції y_1 у першій
С K -вимірній точці $\vec{x}_1 = \{x_{1,1}, \dots, x_{1,K}\}$;
С $EXDA(2)$ - похибка (середньоквадратичне відхилення σ_1) значення y_1 ,
С обов'язково $\sigma_1 > 0$ і $\neq 0$;
С $EXDA(2+1=3)$ - значення першого компоненту $x_{1,1}$ аргументу в першій
С точці;
С $EXDA(2 + 2 = 4), \dots, EXDA(2 + K = NS)$ - значення компонент $x_{1,2}, \dots, x_{1,K}$;
С $EXDA(NS + 1)$ - експериментальне значення функції y_2 у другій
С K -вимірній точці $\vec{x}_2 = \{x_{2,1}, \dots, x_{2,K}\}$ і т. д. до N -ї точки;
С $AMAX(20)$ - максимально можливе значення параметрів $A(i) \equiv \theta_i$;
С $AMIN(20)$ - мінімально можливе значення параметрів $A(i) \equiv \theta_i$;
С
С АРГУМЕНТИ ПІДПРОГРАМИ FUMILI
С M - кількість параметрів θ_i , що знаходяться;
С $N2$ - кількість успішних ітерацій, після яких крок
С зміни параметрів вибирається автоматично;
С $N1$ - число, що вказує, у скільки разів можна змінити
С значення PL за одну ітерацію: рекомендується $N2 = 1$; $N1 = 2$;
С $N3$ - максимальне число ітерацій;
С EPS - початкова точність наближення експериментальних даних;
С $AKAPPA$ - точність наближення, що отримана в кінці обчислень;
С якщо $AKAPPA < EPS$ або $AKAPPA = 0$ обчислення закінчуються;
С $ALAMBD$ - очікувана зміна χ^2 після останньої ітерації;
С MC -параметр, що вказує причину закінчення роботи
С (внутрішня діагностика FUMILI);
С IT -параметр, що керує друком інформації про хід обчислень
С FUMILI (інформація записується у файл FUMERR.DAT):
С $IT = 1$ - запис результатів після кожної ітерації (значення A та похибки);
С $IT = i$ - запис результатів після i -ї ітерації;
С $IT < 0$ або $IT = 0$ - запис проміжних результатів є відсутнім;
С S - подвоєне значення суми χ^2 для всіх точок;

**ПІДПРОГРАМА FUMILI ДЛЯ ЗНАХОДЖЕННЯ ПАРАМЕТРІВ
МЕТОДОМ НАЙМЕНШИХ КВАДРАТІВ**

```
C
C РЕЗУЛЬТАТИ ОБЧИСЛЕНЬ
C результати обчислень параметрів знаходяться у масиві A(20)
C у COMMON /A/A(20) ( $\theta_i = A(i)$ );
C величини, що визначають точності обчислень є аргументами
C підпрограми FUMILI (див. вище);
C обчислені та експериментальні значення функцій,
C похибки та значення  $\chi^2$  наводяться у файлі FUMALL.DAT,
C що формується в підпрограмі ERRORF(M)
C
C SUBROUTINE FUMILI MAIN
      COMMON /A/A(20)/PL/PL(20)/EXDA/EXDA(600)/NED/N,NS
      COMMON /AU/AMAX(20)/AL/AMIN(20)
      DIMENSION x0(200,1), YE(200), DYE(200)
      REAL A0(3)/3*0./
      OPEN (10,file='FUM.DAT', status='new')
      N = 20      !кількість точок
      NZ=50      !максимальне число ітерацій
      M=3        !кількість параметрів, які оцінюються МНК
      K=1        !розмірність вектора-рядка аргументу функції
      NS=K + 2
      N1=2
      N2=1
      IT=0       !керує видачею результатів
      EPS=0.001
      DO 11 i=1,M
        AMIN(I)= -0.1e+21
        AMAX(I)=0.1e+21
      1 A(I)=A0(i)
C формування значень експериментальних даних та їх похибок:
      DO 11 i=1,20
        x0(i,1)=i
        YE(i)=(1+5*x0(i,1))
        DYE(i)=YE(i)*0.1
11 CONTINUE
C формування масиву EXDA значень функції, її похибки та аргументу
      DO 21 j=1,N
        k1=NS*(j-1)+1
        k11=k1+1
        EXDA(k1)=YE(j)
        EXDA(k11)=DYE(j)
        NX=NS-2
        DO 22 k0=1,NX
          k2=k11+k0
          EXDA(k2)=x0(j,k0)
22 CONTINUE
21 CONTINUE
      CONP1=0.1
      DO 2 J=1,M
```

ДОДАТОК

```

      PL(J)=CONP1
2  CONTINUE
      CALL FUMILI(S,M,N1,N2,N3,EPS,AKAPPA,ALAMBD,IT,MC)
      CALL ERRORF(M)
      WRITE(10,100) M,EPS,AKAPPA, ALAMBD,S,K,N3,N,MC
      WRITE(10,101) CONP1
      WRITE(10,102)
      WRITE(10,103) (A0(ie),A(ie),ie=1,M)
100  FORMAT(10x,'FUMILI ESTIMATIONS',
*/8x,'M=',i3,2x,'EPS=',e10.3,2x,'AKAPPA=',e10.3,2x,'ALAMBD=',f7.3,
*/5x,'S=2*chi**2=',e10.3,2x,'K=',i3,2x,'N3=',i4,
*/8x,'N=',i4,2x,'MC=',i4)
101  FORMAT(5x,'CONP1=',f7.3,/)
102  FORMAT(4x,'-- A0(i) -- A(i) --')
103  FORMAT(2(1x,e10.3))
      CLOSE(10)
      STOP
      END

C
      REAL FUNCTION FUNCT(X)
C*****
C обчислення функції, що апроксимує експериментальні дані
C у K-вимірній точці X
C*****
      COMMON /A/A(20)
      DIMENSION X(1)
      FUNCT = A(1)+A(2)*X(1)+A(3)*X(1)**2
      RETURN
      END

C
      SUBROUTINE ARITHM(Y)
C*****
C обчислення значень (Y) апроксимуючої функції та її похідних (DF)
C у K-вимірній точці X(1),...,X(K);
C DF(J) - частинна похідна Y за параметром A(J);
C у випадках, коли є можливість, похідні краще спочатку знаходити
C аналітично, а потім обчислювати за даною програмою;
C A(20) - масив значень параметрів
C*****
      COMMON /DF/DF(20)/A/A(20)/X/X(10)
      COMMON /ENDFLG/ENDFLG,NA,INDFLG(5)
      COMMON /PL/PL(20)/AU/AMAX(20)/AL/AMIN(20)
      DATA RP/1.E-07/
C 10*MAXIMAL RELATIVE PRECISION ON THE COMPUTER
      F=FUNCT(X)
      DO 6 I=1,NA
      DF(I)=0.
      IF (PL(I)) 6,6,1
1  AI=A(I)

```

**ПІДПРОГРАМА FUMILI ДЛЯ ЗНАХОДЖЕННЯ ПАРАМЕТРІВ
МЕТОДОМ НАЙМЕНШИХ КВАДРАТІВ**

```

HI=0.01*PL(I)
PI=RP*ABS(AI)
IF(HI.LE.PI)HI=PI
A(I)=AI+HI
IF(A(I)-AMAX(I))5,5,2
2 A(I)=AI-HI
HI=-HI
IF(A(I)-AMIN(I))3,5,5
3 A(I)=AMAX(I)
HI=AMAX(I)-AI
IF(AMIN(I)-AI+HI)4,5,5
4 A(I)=AMIN(I)
HI=AMIN(I)-AI
5 DF(I)=(FUNCT(X)-F)/HI
A(I)=AI
6 CONTINUE
Y=F
RETURN
END

```

C

```

SUBROUTINE FUMILI(S,M,N1,N2,N3,EPS,AKAPPA,ALAMBD,IT,MC)
C*****
C Підпрограма оцінки параметрів за МНК
C M - кількість параметрів  $\theta_i$ , що знаходяться;
C N2 - кількість успішних ітерацій після яких крок
C зміни параметрів вибирається автоматично;
C N1 - число, що вказує у скільки разів можна змінити
C значення PL за одну ітерацію: рекомендується N2 = 1; N1 = 2;
C N3 - максимальне число ітерацій;
C EPS - початкова точність наближення експериментальних даних;
C AKAPPA - точність наближення, що отримана в кінці обчислень;
C якщо AKAPPA < EPS або AKAPPA = 0, то обчислення закінчуються;
C ALAMBD - очікувана зміна  $\chi^2$  після останньої ітерації;
C MC-параметр, що вказує причину закінчення роботи
C (внутрішня діагностика FUMILI);
C IT-параметр, що керує друком інформації про хід обчислень
C FUMILI (інформація записується у файл FUMERR.DAT):
C IT = 1 - запис результатів після кожної ітерації (значення A та
похибки);
C IT = i - запис результатів після i-i ітерації;
C IT < 0 або IT = 0 запис проміжних результатів відсутній;
C S - подвоєне значення суми  $\chi^2$  для всіх точок;
C*****
COMMON/ENDFLG/ENDFLG,NA,INDFLG(5)
COMMON/A/A(20)/G/G(20)/Z/Z(600)
*/PL/PL(20)/SIGMA/SIGMA(20)/PLU/PL(20)/R/R(20)/DA/DA(20)
*/AU/AMAX(20)/AL/AMIN(20)/Z0/Z0(600)
DIMENSION B(20)
EQUIVALENCE(PL,B)

```

ДОДАТОК

```
DATA RP/1.E-07/
C 10*MAXIMAL RELATIVE PRECISION NUMBERS ON THE COMPUTER
  INDFLG(3)=0
37  NN2=0
    N=M
    FIXFLG=0.
    ENDFLG=0.
    INDFLG(2)=0
    IFIX1=0
    FI=0.
    NN3=0
    DO 38 I=1,N
      R(I)=0.
      IF (EPS.GT.0.) SIGMA(I)=0.
38  PL(I)=PL0(I)
C NEW ITERATION
  1  NN1=1
    T1=1.
  2  S=0.
    NO=0
    DO 101 I=1,N
      G(I)=0.
      IF (PL0(I)) 101,101,100
100  NO=NO+1
      IF (PL(I)) 101,101,300
300  PL0(I)=PL(I)
101  CONTINUE
      NN0=NO*(NO+1)/2
      IF (NN0.LT.1) GO TO 1000
      DO 102 I=1,NN0
102  Z(I)=0.
1000 CONTINUE
      NA=M
      INDFLG(1)=0
      CALL SGZ(M,S)
      SP=RP*ABS(S)
      IF (NN0.LT.1) GO TO 1001
      DO 6 I=1,NN0
  6   Z0(I)=Z(I)
1001 CONTINUE
      IF (NN3) 4,4,39
39  CONTINUE
      IF (NN1-NN1) 30,30,4
30  T=2.*(S-OLDS-GT)
      IF (INDFLG(1)) 32,29,32
29  CONTINUE
      IF (ABS(S-OLDS).LE.SP.AND.-GT.LE.SP) GO TO 4
      IF (0.59*T+GT) 4,31,31
31  T=-GT/T
      IF (T-0.25) 32,33,33
```

**ПІДПРОГРАМА FUMIL1 ДЛЯ ЗНАХОДЖЕННЯ ПАРАМЕТРІВ
МЕТОДОМ НАЙМЕНШИХ КВАДРАТІВ**

```

32   T=0.25
33   GT=GT*T
      T1=T1*T
      NN2=0
      DO 3 I=1,N
      IF (PL(I).LE.0.) GO TO 3
      A(I)=A(I)-DA(I)
      PL(I)=PL(I)*T
      DA(I)=DA(I)*T
      A(I)=A(I)+DA(I)
3    CONTINUE
      NN1=NN1+1
      GO TO 2
4    IF (INDFLG(1).EQ.0) GO TO 28
      ENDFLG=-4.
      GO TO 402
28   K1=1
      K2=1
      I1=1
      DO 208 I=1,N
      IF (PL0(I)) 208,208,200
200  IF (PL(I).EQ.0.) PL(I)=PL0(I)
      IF (PL(I)) 201,201,202
201  K1=K1+I1
      GO TO 207
202  IF (.NOT. (A(I).GE.AMAX(I).AND.-G(I).GT.0..OR.A(I).LE.AMIN(I).
+AND.-G(I).LT.0.)) GO TO 1202
      PL(I)=0.
      GO TO 201
1202 DO 206 J=1,I
      IF (PL0(J)) 206,206,203
203  IF (PL(J)) 205,205,204
204  Z(K2)=Z0(K1)
      K2=K2+1
205  K1=K1+1
206  CONTINUE
207  I1=I1+1
208  CONTINUE
C PUT Z(I,I)
      I1=1
      L=I1
      DO 5 I=1,N
      IF (PL(I)) 5,5,103
103  R(I)=Z(L)
      I1=I1+1
      L=L+I1
5    CONTINUE
      N0=I1-1
      CALL MCONV(N0)
      IF (INDFLG(1)) 1005,1104,1005

```

ДОДАТОК

```
1005  INDFLG(1)=0
      INDFLG(2)=1
      GO TO 1118
1104  CONTINUE
      I1=1
      DO 110 I=1,N
      DA(I)=0.
      IF (PL(I)) 110,110,104
104   L1=1
      DO 109 L=1,N
      IF (PL(L)) 109,109,105
105   IF (I1-L1) 106,106,107
106   K=L1*(L1-1)/2+I1
      GO TO 108
107   K=I1*(I1-1)/2+L1
108   DA(I)=DA(I)-G(L)*Z(K)
      L1=L1+1
109   CONTINUE
      I1=I1+1
110   CONTINUE
      AFIX=0.
      IFIX=0
      I1=1
      L=I1
      DO 117 I=1,N
      IF (PL(I)) 117,117,111
111   SIGI=SQRT(ABS(Z(L)))
      R(I)=R(I)*Z(L)
      IF (EPS) 113,113,112
112   SIGMA(I)=SIGI
113   IF (.NOT.(A(I).GE.AMAX(I).AND.DA(I).GT.0..OR.A(I).LE.AMIN(I).
+AND.DA(I).LT.0.)) GO TO 116
114   AKAP=ABS(DA(I)/SIGI)
      IF (AKAP-AFIX) 116,116,115
115   AFIX=AKAP
      IFIX=I
      IFIX1=I
116   I1=I1+1
      L=L+I1
117   CONTINUE
      IF (IFIX) 118,119,118
118   PL(IFIX)=-1.
1118  CONTINUE
      FIXFLG=FIXFLG+1.
      FI=0.
      GO TO 4
119   CONTINUE
      ALAMBD=1.
      AKAPPA=0.
      IMAX=0
```

**ПІДПРОГРАМА FUMIL1 ДЛЯ ЗНАХОДЖЕННЯ ПАРАМЕТРІВ
МЕТОДОМ НАЙМЕНШИХ КВАДРАТІВ**

```

      DO 129 I=1,N
      IF (PL(I)) 129,129,120
120  BM=AMAX(I)-A(I)
      ABI=A(I)+B(I)
      ABM=AMAX(I)
      IF (DA(I)) 121,121,122
121  BM=A(I)-AMIN(I)
      ABI=A(I)-B(I)
      ABM=AMIN(I)
122  BI=B(I)
      IF (BI-BM) 124,124,123
123  BI=BM
      ABI=ABM
124  IF (ABS(DA(I))-BI) 127,127,125
125  AL=ABS(BI/DA(I))
      IF (ALAMBD-AL) 127,127,126
126  IMAX=I
      AIMAX=ABI
      ALAMBD=AL
127  AKAP=ABS(DA(I)/SIGMA(I))
      IF (AKAP-AKAPPA) 129,129,128
128  AKAPPA=AKAP
129  CONTINUE
      GT=0.
      AMB=1.E18
      IF (ALAMBD) 131,131,130
130  AMB=0.25/ALAMBD
131  CONTINUE
      DO 14 I=1,N
      IF (PL(I)) 14,14,15
15  IF (NN2-N2) 18,18,16
16  IF (ABS(DA(I)/PL(I))-AMB) 18,17,17
17  PL(I)=4.*PL(I)
      T1=4.
18  DA(I)=DA(I)*ALAMBD
      GT=GT+DA(I)*G(I)
14  CONTINUE
501  IF (-GT.GT.SP.OR.T1.GE.1..OR.ALAMBD.GE.1.) GO TO 503
502  ENDFLG=-1.
503  CONTINUE
      IF (ENDFLG) 402,500,500
500  IF (AKAPPA-ABS(EPS)) 135,400,400
135  IF (FIXFLG) 404,403,404
403  ENDFLG=1.
      GO TO 402
404  IF (ENDFLG) 402,136,1402
1402  IF (IFIX1) 402,402,401
405  IF (FI-FIXFLG) 401,401,136
400  IF (FIXFLG) 405,401,405
401  FI=FI+1.

```

ДОДАТОК

```

      ENDFLG=0.
402  CONTINUE
      IF (ENDFLG.EQ.0..AND.NN3.GE.N3) ENDFLG= -3.
      IF (ENDFLG.GT.0..AND.INDFLG(2).GT.0) ENDFLG= -2
      CALL MONITO (S,M,NN3,IT,GT,AKAPPA,ALAMBD,T1)
      IF (ENDFLG) 20,21,20
136  ENDFLG=1.
      FIXFLG=0.
      IFIX1=0
      DO 137 I=1,M
137  PL(I)=PL0(I)
      INDFLG(2)=0
      GO TO 4
      21  ENDFLG=0.
      23  DO 24 I=1,N
      A(I)=A(I)+DA(I)
      24  CONTINUE
      IF (IMAX) 133,133,132
132  A(IMAX)=AIMAX
133  CONTINUE
      OLDS=S
      NN2=NN2+1
      NN3=NN3+1
      GO TO 1
      20  MC=ENDFLG
      RETURN
      END

C
      SUBROUTINE MONITO(S,M,NN3,IT,GT,AKAPPA,ALAMBD,T1)
      COMMON /ENDFLG/ENDFLG,NA,INDFLG(5)/A/A(20)/R/R(20)
      COMMON /SIGMA/SIGMA(20)/PL/PL(20)/PLU/PL0(20)
      OPEN (12,file='FUMERR.DAT',status='replace')
      2  IF (IT) 30,5,3
      3  IF (NN3) 6,6,4
      4  IF (M) 5,6,6
      5  IF (ENDFLG) 6,13,6
      6  S1=2.*S
      GT1=2.*GT
      I1=6
      IF (M.GT.6) I1=5
      IF (M.GT.12) I1=4
      IF (M.GT.24) I1=1
      WRITE (12,7) I1,NN3,S1,GT1,AKAPPA,ALAMBD,T1
      PRINT 7,I1,NN3,S1,GT1,AKAPPA,ALAMBD,T1
      7  FORMAT(2x,I1,15X,I3,5H ITER,2x,3H2S=,E11.4,2x,3HGT=,E10.2//
+15x,4HKAP=,E8.2,2x,5HLAMB=,E8.2,2x,3HT1=,E8.2///
+17x,10HParameters,5x,6HErrors,10x,7HFactors/)
      DO 12 I=1,M
      IF (PL(I)) 10,10,8
      8  IF (PL0(I)) 18,16,15

```

**ПІДПРОГРАМА FUMILI ДЛЯ ЗНАХОДЖЕННЯ ПАРАМЕТРІВ
МЕТОДОМ НАЙМЕНШИХ КВАДРАТІВ**

```

15  CONTINUE
    WRITE (12,9) I,A(I),SIGMA(I),R(I)
C   PRINT 9,I,A(I),SIGMA(I),R(I)
9   FORMAT(10x,I3,E13.5,3H+/-,E10.4,6x,E10.3)
    GO TO 12
16  CONTINUE
    WRITE (12,17) I,A(I),SIGMA(I),R(I)
C   PRINT 17,I,A(I),SIGMA(I),R(I)
17  FORMAT(10x,I3,E13.5,3H+/-,E10.4,6x,E10.3,10H ON BOUND)
    GO TO 12
18  IF (PL0(I).GE.-1.) GO TO 16
    WRITE (12,19) I,A(I)
C   PRINT 19,I,A(I)
19  FORMAT(10x,I3,E13.5,12H+/- INFINITE)
    GO TO 12
10  CONTINUE
    WRITE (12,11) I,A(I)
C   PRINT 11,I,A(I)
11  FORMAT(10x,I3,E13.5,6x,17HFIXED PARAMETER)
12  CONTINUE
30  NM=-IT          !DEBUG! was: NM=-IT
13  NM=M+1          !DEBUG! was: NM=N+1
    IF (ENDFLG) 20,21,21
20  I=-ENDFLG
    GO TO (31,32,33,34),I
21  RETURN
31  CONTINUE
    WRITE (12,41)
C   PRINT 41
41  FORMAT(49H0S WANTS NOT DECREASE. MAY BE DERIVATIVES ARE
BAD//)
    GO TO 21
32  CONTINUE
    WRITE (12,42)
C   PRINT 42
42  FORMAT(45H0INFINIT ERROR ESTIMATION CAN NOT BE SKIPPED//)
    GO TO 21
33  CONTINUE
    WRITE (12,43)
C   PRINT 43
43  FORMAT('0ITERATION LIMIT IS EXEDED'//)
    GO TO 21
34  CONTINUE
    WRITE (12,44)
C   PRINT 44
44  FORMAT(48H0NEGATIVE OR ZERO PROBABILITY CAN NOT BE
SKIPPED//)
    GO TO 21
    CLOSE (12)
    END

```

```

SUBROUTINE MCONV(N)
COMMON /Z/Z(600)/PLU/PL0(20)/R/R(20)
COMMON /ENDFLG/ENDFLG,NA,INDFLG(5)
DATA RP/1.E-07/
C 10.*MAXIMAL RELATIVE PRECISION OF NUMBER ON THE COMPUTER
DATA AM/0.1E 38/
C MAXIMAL NUMBER ON THE COMPUTER
N1=N
IF (N1.LT.1) GO TO 1001
APS=SQRT(AM/FLOAT(N1))
AP=1./APS**2
IR=0
DO 1 I=1,N1
28 IR=IR+1
IF (PL0(IR)) 28,28,29
29 CONTINUE
NI=I*(I-1)/2
II=I*(I+1)/2
K=N+1
IF (Z(II).LE.RP*ABS(R(IR)).OR.Z(II).LE.AP) GO TO 1002
1003 CONTINUE
Z(II)=1./SQRT(Z(II))
NL=II-1
31 CONTINUE
IF (NL-NI) 100,100,30
30 Z(NL)=Z(NL)*Z(II)
IF (ABS(Z(NL)).GE.APS) GO TO 1004
NL=NL-1
GO TO 31
100 IF (I-N1) 2,1001,1001
2 K=K-1
NK=K*(K-1)/2
NL=NK
KK=NK+I
D=Z(KK)*Z(II)
C=D*Z(II)
L=K
3 LL=NK+L
LI=NL+I
Z(LL)=Z(LL)-Z(LI)*C
L=L-1
NL=NL-L
IF (L-I) 20,20,3
20 L=L-1
IF (L) 4,4,21
21 LL=NK+L
LI=NI+L
Z(LL)=Z(LL)-Z(LI)*D
L=L-1

```

**ПІДПРОГРАМА FUMIL1 ДЛЯ ЗНАХОДЖЕННЯ ПАРАМЕТРІВ
МЕТОДОМ НАЙМЕНШИХ КВАДРАТІВ**

```

      IF (L) 4,4,21
4      Z(KK)=-C
      IF (K-I-1) 1,1,2
1      CONTINUE
1001    CONTINUE
      IF (N.LT.1) GO TO 1000
      DO 5 I=1,N
      DO 5 K=I,N
      LL=K*(K+1)/2
      NL=K*(K-1)/2
      N2=N
      D=0.
      K1=K
      DO 6 L=K1,N2
      LI=NL+I
      LK=NL+K
      D=D+Z(LI)*Z(LK)
      NL=NL+L
6      CONTINUE
10      CONTINUE
      KI=K*(K-1)/2+I
      Z(KI)=D
5      CONTINUE
1000    CONTINUE
      RETURN
1002    PL0(IR)=-2
      R(IR)=0.
      INDFLG(1)=IR
      GO TO 1000
1004    K=I+NL-II
      IR=0
      DO 1006 I=1,K
1005    IR=IR+1
      IF (PL0(IR)) 1005,1005,1006
1006    CONTINUE
      GO TO 1002
      END

C      SUBROUTINE ERRORF(M)
      COMMON /A/A(20)/DF/DF(20)/Z/Z(600)/EXDA/EXDA(600)
      COMMON /NED/NED(2)/X/X(10)
      OPEN (9,file='FUMALL.DAT',status='replace')
      WRITE (9,4)
C      PRINT 4
4      FORMAT(3x,2HNN,3x,8HFunction,1x,3H+/-,3x,6HErrors,5x,
+4HHI-2,4x,' Experimental errors ',1x,'Arguments'/)
      NS=NED(1)
      DO 1 J=1,NS
      K1=NED(2)*(J-1)+1
      FUN=EXDA(K1)

```

```

        SIG=EXDA(K1+1)
        NX=NED(2)-2
        DO 2 K=1,NX
            K2=K1+1+K
            X(K)=EXDA(K2)
2       CONTINUE
            CALL ARITHM(Y)
            XI=((FUN-Y)/SIG)**2
            CALL SCAL(M,ER)
            ER=SQRT(ER)
            WRITE(9,5) J,Y,ER,XI,FUN,SIG,X(1)
C       PRINT 5,J,Y,ER,XI,FUN,SIG,X(1)
5       FORMAT(1x,I4,1x,E10.4,1x,3H+/-,E10.4,1x,E10.4,1x,3E10.4)
            IF(NX.eq.1) GO TO 1
            WRITE(9,55) (i,x(i),i=2,NX)
C       PRINT 55,(i,x(i),i=2,NX)
55      FORMAT(57x,2Hx(i1,2h)=,e10.4)
1       CONTINUE
            CLOSE(9)
            RETURN
        END
C
        SUBROUTINE SGZ(M,S)
        COMMON /G/G(20)/Z/Z(600)/A/A(20)/DF/DF(20)/NED/NED(2)
+ /EXDA/EXDA(600)/X/X(10)/PL/PL(20)
        COMMON /ENDFLG/ENDFLG,NA,INDFLG(5)
        K=NED(1)
        K2=1
        DO 2 L1=1,K
            K1=K2
            NX=NED(2)-2
            IF(INDFLG(3)) 7,8,7
7         NX=NED(2)
            K1=K1-2
8         CONTINUE
            DO 6 I=1,NX
                KI=K1+1+I
                X(I)=EXDA(KI)
6         CONTINUE
            CALL ARITHM(Y)
10      SIG=EXDA(K2+1)
            Y=Y-EXDA(K1)
            S=S+((Y/SIG)**2)/2.
11     CONTINUE
            N=0
            DO 5 J=1,M
                IF(PL(J)) 4,4,3
3         N=N+1
            DF(N)=DF(J)/SIG
            G(J)=G(J)+DF(N)*(Y/SIG)

```

**ПІДПРОГРАМА FUMIL1 ДЛЯ ЗНАХОДЖЕННЯ ПАРАМЕТРІВ
МЕТОДОМ НАЙМЕНШИХ КВАДРАТІВ**

```

4  CONTINUE
5  CONTINUE
    L=1
    IF (N.LT.1) GO TO 1000
    DO 1 I=1,N
    DO 1 J=1,I
    Z(L)=Z(L)+DF(I)*DF(J)
1   L=L+1
1000 CONTINUE
    K2=K2+NED(2)
2   CONTINUE
    RETURN
    END

C
    SUBROUTINE SCAL(M,ER)
    COMMON /Z/Z(600)/DF/DF(20)/PLU/PL(20)
    ER=0.
    N=0
    DO 5 J=1,M
    IF (PL(J)) 8,8,7
7   N=N+1
    DF(N)=DF(J)
8   CONTINUE
5   CONTINUE
    IF (N.LT.1) GO TO 1000
    DO 1 I=1,N
    DO 1 L=1,N
    IF (I-L) 2,2,3
2   K=L*(L-1)/2+I
    GO TO 4
3   K=I*(I-1)/2+L
4   ER=ER+Z(K)*DF(I)*DF(L)
1   CONTINUE
1000 CONTINUE
    RETURN
    END

```

РЕЗУЛЬТАТИ ОБЧИСЛЕНЬ

Результати обчислень що роздруковуються операторами з головної програми

```
FUMILI ESTIMATIONS
  M= 3 EPS= 0.100E-02 AKAPPA= 0.161E-04 ALAMBD= 1.000
  S=2*chi**2= 0.204E-08 K= 1 N3= 50
  N= 20 MC= 1
  CONPI= 0.100

-- A0(i) -- A(i) --
0.000E+00 0.100E+01
0.000E+00 0.500E+01
0.000E+00 0.122E-06
```

Величини у файлі даних FUM LL

NN Function		+/- Errors	HI-2	Experimental	errors	Arguments
1	0.6000E+01	+/-0.5481E+00	0.1011E-10	0.6000E+01	0.6000E+00	0.1000E+01
2	0.1100E+02	+/-0.4734E+00	0.7516E-10	0.1100E+02	0.1100E+01	0.2000E+01
3	0.1600E+02	+/-0.5927E+00	0.1151E-09	0.1600E+02	0.1600E+01	0.3000E+01
4	0.2100E+02	+/-0.7693E+00	0.1394E-09	0.2100E+02	0.2100E+01	0.4000E+01
5	0.2600E+02	+/-0.9355E+00	0.1555E-09	0.2600E+02	0.2600E+01	0.5000E+01
6	0.3100E+02	+/-0.1073E+01	0.1669E-09	0.3100E+02	0.3100E+01	0.6000E+01
7	0.3600E+02	+/-0.1178E+01	0.1898E-09	0.3600E+02	0.3600E+01	0.7000E+01
8	0.4100E+02	+/-0.1254E+01	0.1948E-09	0.4100E+02	0.4100E+01	0.8000E+01
9	0.4600E+02	+/-0.1311E+01	0.1987E-09	0.4600E+02	0.4600E+01	0.9000E+01
10	0.5100E+02	+/-0.1360E+01	0.2020E-09	0.5100E+02	0.5100E+01	0.1000E+02
11	0.5600E+02	+/-0.1419E+01	0.2246E-09	0.5600E+02	0.5600E+01	0.1100E+02
12	0.6100E+02	+/-0.1509E+01	0.2253E-09	0.6100E+02	0.6100E+01	0.1200E+02
13	0.6600E+02	+/-0.1653E+01	0.2258E-09	0.6600E+02	0.6600E+01	0.1300E+02
14	0.7100E+02	+/-0.1866E+01	0.2598E-09	0.7100E+02	0.7100E+01	0.1400E+02
15	0.7600E+02	+/-0.2157E+01	0.2580E-09	0.7600E+02	0.7600E+01	0.1500E+02
16	0.8100E+02	+/-0.2528E+01	0.2564E-09	0.8100E+02	0.8100E+01	0.1600E+02
17	0.8600E+02	+/-0.2977E+01	0.2841E-09	0.8600E+02	0.8600E+01	0.1700E+02
18	0.9100E+02	+/-0.3500E+01	0.2812E-09	0.9100E+02	0.9100E+01	0.1800E+02
19	0.9600E+02	+/-0.4095E+01	0.3057E-09	0.9600E+02	0.9600E+01	0.1900E+02
20	0.1010E+03	+/-0.4757E+01	0.3019E-09	0.1010E+03	0.1010E+02	0.2000E+02

Величини у файлі FUMERR.DAT

```
6      6 ITER 2S= 0.4070E-08 GT= -0.81E-08
      KAP=0.16E-04 LAMB=0.10E+01 T1=0.10E+01

Parameters      Errors      Factors
1  0.10000E+01  +/-0.8307E+00  0.336E+01
2  0.50000E+01  +/-0.4326E+00  0.140E+02
3  0.12231E-06  +/-0.2864E-01  0.915E+01
```

РЕКОМЕНДОВАНА ЛІТЕРАТУРА

1. Азекян Т.А. Основы теории ошибок для астрономов и физиков. – М.: Наука, 1968. – 148 с.
2. Ватутин В.В., Телевинова Т.М., Чистяков В.П. Вероятностные методы в физических исследованиях. – М.: Наука, 1985. – 208 с.
3. Гайдышев И. Анализ и обработка данных: специальный справочник. – СПб.: Питер, 2001. – 752 с.
4. Гнеденко Б.В. Курс теории вероятностей. – М.: Наука, 1987. – 400 с.
5. Гришин В.К., Живописцев Ф.А., Иванов В.А. Математическая обработка и интерпретация физического эксперимента. – М.: Изд-во МГУ, 1988. – 320 с.
6. Єжов С.М. Теорія ймовірностей, математична статистика і випадкові процеси. – К.: ВПЦ "Київський університет", 2001. – 166 с.
7. Колмаев В.А., Староверов О.В., Турундаевский В.Б. Теория вероятностей и математическая статистика. – М.: Высш. шк., 1991. – 400 с.
8. Мудров В.В., Кушко В.Л. Методы обработки измерений. – М.: Сов. радио, 1976. – 192 с.
9. y F y N
77:
-
, 1997 :
10. Пустыльник Е.И. Статистические методы анализа и обработки наблюдений. – М.: Наука, 1968. – 288 с.
11. Рудчик А.Т. Обробка експериментальної інформації на ЕЦОМ. – К.: Наук. думка, 1976. – 324 с.
12. Свешников А.А. Основы теории ошибок. – Л., Изд-во ЛГУ, 1972. – 124 с.
13. Скороход А.В. Елементи теорії ймовірностей та випадкових процесів. – К., Вища школа, 1975. – 296 с.
14. Статистические методы в экспериментальной физике : Пер. с англ. Под ред. А.А.Тяпкина. – М.: Атомиздат, 1976. – 335 с.
15. Турчин В.М. Математична статистика. – К.: Академія, 1999. – 240 с.
16. Худсон Д. Статистика для физиков. – М.: Мир, 1970. – 296 с.
17. Шефтель З.Г. Теорія ймовірностей. – К.: Вища школа, 1977. – 157 с.
18. Фигурин В.А., Оболонкин В.В. Теория вероятностей и математическая статистика. – Минск: Новое знание, 2000. – 206 с.
19. L. – 1968. – 40. – . 586.
20. Рао С.Р. Линейные статистические методы и их применения. – М.: Наука, 1968. – 206 с.
21. Каденко І.М., Мелешин А.Ю. УФЖ – 2001. – 45. – . 813.

Навчальне видання

Каденко Ігор Миколаєвич
Плюйко Володимир Андрійович

ОСНОВИ МЕТОДІВ ОЦІНОК СТАТИСТИЧНИХ ДАНИХ ТА ФУНКЦІЙ ЇХ РОЗПОДІЛУ

Навчальний посібник

Редактор *О.В.Литвинова*
Технічний редактор *А.В.Правдивець*

Оригінал-макет виготовлено Видавничо-поліграфічним центром "Київський університет"



Підписано до друку 26.09.03. Формат 60х84^{1/16}. Вид. № 492. Гарнітура Arial. Папір офсетний.
Друк офсетний. Наклад 100. Ум. друк. арк. 9,1. Обл.-вид. арк. 11,4. Зам. № 23-1452.

Видавничо-поліграфічний центр "Київський університет"
01033, Київ, б-р Т. Шевченка, 14, кімн. 43,
☎ (38044) 239 3222; (38044) 239 3172; (38044) 234 0105; факс (38044) 234 22 90.

Свідоцтво внесено до державного реєстру ДК № 1103 від 31.10.02.