ЛЕКЦІЯ 3
СТАТИСТИЧНИЙ АНАЛІЗ ВИПАДКОВИХ ПОМИЛОК

Як вже було сказано, основними методами експериментальних досліджень прийнято вважати ***спостереження*** та ***експеримент*** (див. лекцію 1). Метою ж експерименту є кількісне та якісне вивчення певних властивостей досліджуваного об'єкту (явища), виявлення взаємозв'язків між ними. Ці дослідження виконуються на основі ***вимірювань*** відповідних фізичних величин. І найважливішою характеристикою якості вимірювань є їх ***достовірність***.

Достовірність вимірювань – це характеристика, яка визначає міру довіри до отриманих результатів вимірювань. За даною характеристикою вимірювання поділяються на достовірні й недостовірні. *Достовірність вимірювань залежить від того, чи відома ймовірність відхилення результатів вимірювання від істинного значення вимірюваної величини.* Якщо ж імовірність вимірювань не визначена, то їх результати, як правило, не використовуються. Достовірність вимірювань обмежена ***похибкою вимірювань***.

Як відомо, ***один з найкращих способів оцінити достовірність вимірювань*** *полягає в тому, щоб повторити їх кілька разів, а потім порівняти між собою різні отримані значення.*

3.1 Прямі та непрямі вимірювання

Якщо величина вимірюється безпосередньо приладом, таке вимірювання називають ***прямим***. У сучасній науці багато понять введені в результаті розвитку теорії, і їх безпосередньо виміряти не можна. Вони обчислюються за теоретичними формулами, до яких входять величини, виміряні безпосередньо. Такі вимірювання називаються ***непрямими вимірюваннями***, а їх помилки визначаються помилками, що входять до формул визначення цих величин.

Наприклад, для визначення площі прямокутника, вимірюють його довжину l та висоту h і потім розраховують його площу S за формулою $S = lh$.

Аналогічно, найочевидніший спосіб визначення швидкості v тіла полягає в тому, щоб виміряти шлях d, пройдений тілом, та витрачений на це час t, а потім обчислити v за формулою $v = d/t$.

Таким чином, *майже всі важливі вимірювання включають два етапи, що складаються з простих вимірювань та подальших розрахунків*.

Якщо вимірювання включає ці два етапи, то й оцінювання похибок також включає їх. Спочатку потрібно оцінити похибки у величинах, які вимірюються безпосередньо, а потім визначити, як ці похибки «розповсюджуються» в розрахунках і призводять до виникнення похибки в остаточному результаті.

Як відомо, якщо вимірювання $x,\ldots, z$ проведені з відповідними похибками $\Delta_x, \ldots, \Delta_z$, і ці похибки є незалежними та випадковими, то загальна формула визначення похибки Δ_q функції кількох змінних $q(x,\ldots,z)$ має вигляд

$$\Delta_q = \sqrt{\left(\frac{\partial q}{\partial x}\Delta_x\right)^2 + \ldots + \left(\frac{\partial q}{\partial z}\Delta_z\right)^2}\ .$$

3.2 Випадкові та систематичні помилки

Експериментальні похибки, які можна виявити за допомогою багаторазових вимірювань (*і обробити статистичними методами*), називаються ***випадковими помилками (похибками)***, а ті, що не можна виявити таким способом (*і до яких статистичні методи **не можна** застосувати*), – ***систематичними помилками (похибками).***

Випадкові помилки виникають через безліч одночасно діючих джерел завад. *Вони виявляються лише при багаторазових вимірюваннях.* Це помилки, які можна піддати обробці за допомогою методів теорії ймовірностей та математичної статистики. Їх непередбачуваність, таким чином, зводиться до мінімуму.

Систематичні помилки повторюються з досліду в дослід і мають однакове значення. З них виділяють: *поправки* (уточнюють теорію, постійні впливи тощо), *помилки невідомого походження* (виникають через недостатньо розроблену теорію, складний експеримент) і *клас точності приладів*. *Найчастіше клас точності приладів вважається основним джерелом систематичних помилок.*

Важливим типом помилок є ***промахи***, тобто грубі помилки, що виникли в ході експерименту. Їх треба вміти відокремити від нормальних вимірювань, і основним способом їх уникнення є увага і ретельність.

Приклад 3.1. Припустимо спочатку, що ми вимірюємо час одного оберту диска, що рівномірно обертається. Одним з джерел помилок буде час нашої власної реакції при запуску та зупинці секундоміра. Якби цей час реакції завжди був однаковим, то два запізнювання, обумовлені реакцією, компенсували б один одного. Але фактично час нашої реакції змінюється. Ми можемо більше затриматися при запуску і таким чином недооцінити час оберту або ж більше затриматися при зупинці секундоміра і в цьому випадку переоцінити час. Оскільки обидві можливості є різноймовірними, то ефект є випадковим. При багаторазовому повторенні вимірювання ми іноді переоцінюватимемо час, а іноді – недооцінюватимемо. Таким чином, *змінний час нашої реакції проявиться у відмінності отриманих результатів*. Аналізуючи розкид в результатах методами статистики, ми можемо отримати дуже достовірну оцінку помилки цього типу (*випадкової помилки*).

З іншого боку, якщо наш секундомір постійно відстає, то все виміряні значення часу будуть недооцінені, і жодна кількість повторень (з тим же секундоміром) не виявить цього джерела помилок. Помилка такого типу є ***систематичною***, оскільки *вона завжди зсуває наш результат в однин бік*. (Якщо секундомір відстає, ми завжди недооцінюємо час, якщо поспішає – завжди переоцінюємо.) *Систематичні помилки не можна виявити статистичними методами*.

Майже всім вимірюванням притаманні як випадкові, так і систематичні похибки.

Типові джерела випадкових похибок – це невеликі помилки спостерігача, завади, які впливають на апаратуру (подібні до механічних вібрацій), проблеми визначення тощо. Можливо, найбільш очевидною причиною систематичної помилки є розкалібрування приладів (подібно до секундоміру, що відстає, в якого стрілка до початку вимірювань не була встановлена на нуль).

Різницю між випадковими і систематичними помилками не завжди можна ясно визначити. Наприклад, при зміні положення нашої голови відносно типового стрілочного приладу (наприклад, звичайного годинника) результати зчитування будуть змінюватися. Цей ефект називається *паралаксом*, і він призводить до того, що правильне зчитування зі шкали можливе тільки у випадку, коли наша голова розташована точно перед стрілкою. Навіть якщо експериментатор є дуже акуратним, він не зможете розташувати свої очі завжди точно перед стрілкою; отже, його вимірювання міститимуть малі похибки, пов'язані з паралаксом, і ця похибка буде, ймовірно, випадковою.

З іншого боку, необережний експериментатор, який поставить стрілочний прилад збоку від себе і забуде про вплив паралаксу, внесе систематичну помилку в усі свої вимірювання.

Таким чином, один і той же ефект (наприклад, паралакс) може призвести до випадкових похибок в одному випадку і до систематичних – в іншому.

Урахування випадкових помилок відрізняється від урахування систематичних помилок. *Статистичні методи дають достовірну оцінку випадкових похибок і вказують на певний спосіб їх зменшення.*

Систематичні ж похибки важко оцінити і навіть виявити. Людина, що проводить експеримент, повинна вміти передбачати можливі джерела систематичних помилок і забезпечити, щоб всі систематичні помилки були значно меншими необхідної точності. Для цього необхідна, наприклад, повірка стрілочних приладів за прийнятими стандартами, їх виправлення, або, якщо необхідно, їх заміна на досконаліші.

3.3 Середнє значення та стандартне відхилення

Нехай нам треба виміряти деяку величину x і ми виявили всі джерела систематичної помилки та зменшили їх вплив до нехтовно малого рівня. Оскільки всі джерела помилок, що залишилися, є випадковими, ми зможемо виявити їх, багато разів повторюючи вимірювання.

Можна було б, наприклад, виконати вимірювання 5 разів і отримати результати (опустимо одиниці вимірювання): 71, 72, 72, 73, 71.

Перше питання, яке можна задати: якщо є 5 виміряних значень, то що ми повинні прийняти за найкращу оцінку $x_{найкр}$ нашої величини x? Представляється розумним, щоб нашою найкращою оцінкою було ***середнє значення*** – середнє $\bar{x}$ п'яти знайдених значень: $x_{найкр} = \bar{x} = \dfrac{71+72+72+73+71}{5} = 71.8$.

Тут дріб – просте визначення середнього x для даних чисел.

У загальному випадку припустимо, що ми проводимо N вимірювань величини x (використовуючи одну й ту ж апаратуру і метод вимірювання) і отримуємо N значень: $x_1, x_2, \ldots, x_N$. Тоді найкращою оцінкою величини x буде середнє значення від $x_1, \ldots, x_N$, тобто

$$x_{найкр} = \bar{x} = \frac{x_1 + x_2 + \ldots + x_N}{N} = \frac{\sum x_i}{N} \tag{3.1}$$

Розглянемо поняття стандартного відхилення. ***Стандартне відхилення*** результатів вимірювань $x_1, x_2, \ldots, x_N$ – це оцінка середньої похибки результатів вимірювань $x_1, x_2, \ldots, x_N$, яке визначається наступним чином.

Якщо прийняти, що середнє $\bar{x}$ – це найкраща оцінка величини x, то логічно розглянути різницю $x_i - \bar{x} = d_i$. Ця різниця, часто звана ***відхиленням (або залишком)*** x_i від $\bar{x}$, показує, *наскільки результат i-го вимірювання* x_i *відрізняється від середнього значення* $\bar{x}$. Якщо відхилення d_i є дуже малими, то результати наших вимірювань близькі один до одного і, ймовірно, дуже точні. Якщо ж деякі з відхилень великі, то наші вимірювання є, очевидно, не дуже точними.

Приклад 3.2. Обчислимо відхилення для вибірки з п'яти результатів вимірювань 71, 72, 72, 73, 71. Результати розрахунку наведені в табл.3.1.

Таблиця 3.1 – Обчислення відхилень

Номер вимірювання i	Виміряне значення x_i	Відхилення $d_i = x_i - \bar{x}$	d_i^2
1.	71	–0.8	0.64
2.	72	0.2	0.04
3.	72	0.2	0.04
4.	73	1.2	1.44
5.	71	–0.8	0.64
	$\bar{x}$ =71.8	$\bar{d}$ = 0.0	$\sum d_i^2$ =2.8

З таблиці видно, що відхилення не всі однакові; d_i є малим, якщо в i-му результаті вимірювання x_i виявляється близьким до $\bar{x}$, але d_i є великим, якщо значення x_i є далеким від $\bar{x}$. Зазначимо також, що деякі з d_i є додатними, а деякі – від'ємними, оскільки деякі x_i повинні бути більшими за середнього значення $\bar{x}$, а деякі – меншими.

Щоб оцінити достовірність результатів вимірювань x_1, … x_5 в середньому, можна було б спробувати усереднити відхилення d_i. Як показує табл.3.1, середнє значення відхилень дорівнює нулю. Насправді так буде у випадку будь-якого набору результатів вимірювань x_1, x_2, …, x_N, оскільки вже саме визначення середнього значення $\bar{x}$ призводить до того, що $d_i = x_i - \bar{x}$ є іноді позитивними, а іноді негативними, щоб $\bar{d}$ дорівнювало нулю. Тому, ***середнє відхилень – це не найкраща характеристика достовірності результатів вимірювань x_1, x_2, …, x_N.***

Кращий спосіб обійти цю неприємність – *піднести до квадрату всі відхилення, які в цьому випадку утворюватимуть набір додатних чисел, а потім усереднити ці числа.* Якщо ми тепер дістанемо квадратний корінь з отриманого результату, то отримаємо величину, яка вимірюється в тих же одиницях, що і сама величина x. Це число, як відомо, називається ***стандартним відхиленням x_1, x_2, …, x_N*** і позначається σ_x:

$$\sigma_x = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(d_i)^2} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2}\ . \tag{3.2}$$

Виходячи з цього позначення, ***стандартне відхилення*** можна описати як середньоквадратичне відхилень результатів вимірювань $x_1, x_2, \ldots, x_N$. Тому його ще називають ***середньоквадратичним відхиленням***.

Розрахуємо σ_x для п'яти вимірювань з табл.3.1, результати також занесемо до цією ж таблиці.

Склавши числа d_i^2, з четвертого стовбця табл.3.1 та розділивши суму на n=5, ми отримаємо величину σ_x^2, яка називається ***дисперсією вимірювань***:

$$D = \sigma_x^2 = \frac{1}{N}\sum_{i=1}^{N} d_i^2 = \frac{2.8}{5} = 0.56\ . \tag{3.3}$$

Дістаючи квадратний корінь, визначаємо стандартне відхилення: $\sigma_x \approx 0.7$.

Таким чином, середня похибка результатів п'яти вимірювань 71, 72, 72, 73, 71 наближено дорівнює 0.7.

Має місце також *альтернативне визначення стандартного відхилення:*

$$\sigma_x = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(d_i)^2} = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2}\ . \tag{3.4}$$

Це визначення призводить до трохи більшого значення, ніж у (3.2), і це дещо компенсує недооцінювання похибки в результатах вимірювань $x_1, x_2, \ldots, x_N$, особливо у випадку, коли кількість вимірювань N є малою.

Хоча різниця між двома виразами (3.2) і (3.4) чисельно майже завжди є незначною, *краще використовувати більш консервативне (тобто те, що призводить до більшого значення) визначення (3.4).*

Таким чином, *стандартне відхилення σ_x характеризує середню похибку результатів вимірювань $x_1, x_2, \ldots, x_N$, за якими воно було обчислене*.

3.4 Стандартне відхилення як похибка одиничного вимірювання

Якщо результати наших вимірювань розподілені нормально, і якщо ми повторюватимемо вимірювання x дуже багато разів (завжди з тією ж апаратурою), то приблизно 70% (*точніше 68.28%*) результатів наших вимірювань перебуватимуть в межах σ_x від $\bar{x}$ (тобто 70% результатів наших вимірювань належатимуть інтервалу $\bar{x} \pm \sigma_x$).

Приклад 3.3. Нехай, у нас є коробка з однаковими пружинами і потрібно виміряти їх коефіцієнти пружності k.

Оскільки пружини однакові, то є достатні підстави очікувати однакової похибки у будь-якому вимірюванні (для будь-якої пружини). Тоді для кожної наступної пружини нам потрібно провести лише одне вимірювання.

Нехай проведено 10 вимірювань для першої пружини – і отримані наступні значення k (в Н/м): 86, 85, 84, 89, 86, 88, 88, 85, 83, 85.

За цими даними ми обчислюємо $\bar{k}$=85.9 Н/м і за визначенням (3.4) отримуємо σ_k=*1.9 Н/м ≈ 2 Н/м.*

Таким чином, похибка будь-якого окремого вимірювання k складає приблизно 2 Н/м. Якщо ми тепер зробимо одне вимірювання для другої пружини і отримаємо результат k=71 Н/м, то можна прийняти, що σ_k=2 Н/м, і стверджувати з 70%-вою імовірністю, що $k_{другої\ пружини}$=71±2 Н/м.

3.5 Стандартне відхилення середнього

Отже, якщо $x_1, x_2, \ldots, x_N$ – результати N вимірювань однієї величини x, то наша найкраща оцінка величини x є їх середнім $\bar{x}$: $x_{найкр} = \bar{x}$.

Стандартне відхилення σ_x характеризує середню похибку окремих вимірювань $x_1, x_2, \ldots, x_N$. Однак наш результат $x_{найкр}=\bar{x}$ є розумною комбінацією всіх N вимірювань, і тому він буде надійнішим, ніж будь-яке з окремих вимірювань. Тобто похибка в нашому результаті $x_{найкр}=\bar{x}$ дорівнює стандартному відхиленню σ_x, поділеному на $\sqrt{N}$. Ця величина називається ***стандартним відхиленням середнього*** і позначається $\sigma_{\bar{x}}$:

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{N}}. \tag{3.5}$$

Таким чином, ґрунтуючись на N виміряних значеннях $x_1, x_2, \ldots, x_N$, ми можемо сформулювати наш кінцевий результат для значення величини x як: *(значення x)*$=x_{найкр}\pm\sigma_{\bar{x}}$, де $x_{найкр}=\bar{x}$ – середнє від $x_1, x_2, \ldots, x_N$, а $\sigma_{\bar{x}}$ – стандартне відхилення середнього (3.5). Тоді *(значення x)*$=\bar{x}\pm\frac{\sigma_x}{\sqrt{N}}$.

Повернімося до десяти результатів вимірювань коефіцієнту пружності k однієї пружини (приклад 3.3). Як ми вже знаємо, середнє цих значень є $\bar{k}$ =85.9 Н/м, стандартне відхилення дорівнює $\sigma_k=1.9$ *Н/м*. Отже, стандартне відхилення середнього дорівнює $\sigma_{\bar{k}}=\frac{\sigma_k}{\sqrt{10}}=$ 0.6 Н/м, і наш остаточний результат, ґрунтований на цих десяти вимірюваннях, для коефіцієнту пружності пружини дорівнюватиме $k=85.9\pm0.6$ *Н/м.*

Розглянемо приклади застосування стандартного відхилення середнього.

Приклад 3.4. Нехай потрібно дуже точно виміряти площу A прямокутної пластинки розміром 2.5 см×5 см. Зроблено по десять вимірювань довжини l та ширини b пластинки. Отримані результати наведені в табл.3.2.

Таблиця 3.2 – Довжина та ширина (мм)

	Виміряні значення					Середнє значення $\bar{x}$	Стандартне відхилення σ_x	Стандартне відхилення середнього $\sigma_{\bar{x}}$
l	24.25	24.26	24.22	24.28	24.24	$\bar{l}=24.245$	$\sigma_l=0.019$	$\sigma_{\bar{l}}=0.006$
	24.25	24.22	24.26	24.23	24.24			
b	50.36	50.35	50.41	50.37	50.36	$\bar{b}=50.368$	$\sigma_b=0.024$	$\sigma_{\bar{b}}=0.008$
	50.32	50.39	50.38	50.36	50.38			

За десятьма отриманим значенням l легко обчислити середнє значення $\bar{l}$, стандартне відхилення σ_l та стандартне відхилення середнього $\sigma_{\bar{l}}$, що наведено в табл.3.2. Аналогічно можна обчислити $\bar{b}$, σ_b та $\sigma_{\bar{b}}$.

Нашою найкращою оцінкою для довжини є середнє значення $\bar{l}$, а її похибка – стандартне відхилення середнього $\sigma_{\bar{l}}$, таким чином, остаточний результат для *l* буде таким: *l*=24.245±0.006 мм (або 0.025 %).

Аналогічно результат для *b* має вигляд: *b*=50.368±0.008 мм (або 0.016%).

Далі обчислюємо середні значення $\bar{l}$ та $\bar{b}$ з їх похибками, які дорівнюють стандартним відхиленням відповідних середніх значень.

Тоді наша найкраща оцінка для площі дорівнює: *A*=*lb*, а її відносна похибка визначається квадратичною сумою відносних похибок в *l* та *b* (припускаючи, що помилки незалежні), як у непрямому вимірюванні:

$$A=(24.245 \text{ мм} \pm 0.025\%)\times(50.368 \text{ мм} \pm 0.016\%) = 1221.17 \text{ мм}^2 \pm 0.03\% =$$

$$= 1221.2 \pm 0.4 \text{ мм}^2.$$

Приклад 3.5. *Розглянемо випадок, коли не можна застосовувати звичайні статистичні методи до результатів прямих вимірювань, але можна – до остаточних результатів*. Припустимо, що ми хочемо виміряти коефіцієнт пружності *k* пружини, визначаючи період коливань для деякої маси *m*, прикріпленій до її кінця. З елементарної механіки відомо, що період таких коливань дорівнює $T = 2\pi\sqrt{\frac{m}{k}}$. Таким чином, вимірюючи *T* та *m*, ми можемо знайти *k* за формулою $k = \frac{4\pi^2 m}{T^2}$.

Найпростіший спосіб визначити *k* – взяти одну масу *m*, відому із заданою точністю, і провести кілька акуратних вимірювань *T*. Однак, з різних причин, кращим було б вимірювання *T* для різних мас *m*. (Наприклад, в цьому випадку разом з вимірюванням *k* ми могли б перевірити, що $T \sim \sqrt{m}$). Тоді можна було б отримати ряд значень, подібних до наведених у двох перших рядках табл.3.3.

Таблиця 3.3 – Вимірювання коефіцієнту пружності пружини *k*

Маса *m*, кг	0.513	0.581	0.634	0.691	0.752	0.834	0.901	0.950
Період *T*, с	1.24	1.33	1.36	1.44	1.50	1.59	1.65	1.69
$k = \frac{4\pi^2 m}{T^2}$	13.17	12.97	тощо	…	…	…	…	…

Вочевидь, немає сенсу усереднювати значення різних мас, наведених у першому рядку (а також і значення періодів у другому рядку), оскільки вони не є різними результатами вимірювань однієї величини.

Ми також нічого не зможемо сказати про похибки наших вимірювань, якщо порівнюватимемо ці різні значення *m*.

С іншого боку, кожне значення *m* можна об'єднати з відповідним значенням періоду *T* і обчислити *k*, як показано в останньому рядку табл.3.3. *Наші дані для k в останньому рядку є результатами вимірювань одної величини і, таким чином, до них застосовуються статистичні методи*. Зокрема, найкращою оцінкою *k* буде середнє значення $\bar{k}$ =13.16 Н/м, а похибкою – стандартне відхилення середнього $\sigma_{\bar{k}}$ = 0.06 Н/м (вам потрібно буде заповнити табл.3.3 до кінця, розрахувавши всі *k*). Таким чином, остаточний результат, ґрунтований на даних табл.3.3, матиме вигляд: *коефіцієнт пружності пружини k=13.16 ± 0.06 Н/м.*

Контрольні питання

1. Прямі та непрямі вимірювання.
2. Випадкові та систематичні помилки.
3. Середнє значення та стандартне відхилення.
4. Стандартне відхилення як похибка одиничного вимірювання.
5. Стандартне відхилення середнього.