

«Наука починається тоді, коли починають вимірювати.

У природі міра і вага – головні засоби пізнання»

Дмитро Менделєєв

ЗАГАЛЬНИЙ ВСТУП

1. **Знання** про природу хемічних та фізико-хемічних явищ, процесів і закономірностей набувають шляхом **обсервації**¹, а також вивчення, порівняння досвіду, експерименту тощо.

2. У **широкому** смислі під **обсервацією** розуміють процес планомірного накопичення і нагромадження фактів, результатів при спостереженні або експерименті, або шляхом опису та вивчення реального хемічного об'єкту.

3. **Як висловився Жюль Анрі Пуанкаре** (Poincare) (1854 – 1912 рр.): «Наука складається з фактів, як будинок з цеглин, однак просте нагромадження фактів схоже на науку не більше, ніж купа цеглин на будинок».

Власне накопичення, нагромадження фактів – наукового «будівельного» матеріалу (а не створення самої наукової «будівлі») – ще не становить зміст науки. Тільки зведені в певну систему факти, дані, результати опису, обстежень, експерименту набувають визначеного смислу і можуть бути приведені до форми, з якої можна витягнути сховану у фактах інформацію.

4. На шляху наукового пізнання є два підходи:

- емпіричний²;
- теоретичний.

На емпіричному рівні – це є опис спостережень, результатів експерименту, їх узагальнення, встановлення певних числових співвідношень між окремими характеристиками явища.

На теоретичному рівні на підставі таких узагальнень формують певну **гіпотезу** – припущення для пояснення цього явища. Якщо вдалося гіпотезу перевірити і за допомогою досліду встановити її правдивість, з високим рівнем ймовірності, то ця гіпотеза стає законом. Закон як достовірне знання отримує форму певного математичного співвідношення між параметрами, якими це явище описане.

5. Вимоги до дослідника. При планомірному накопиченні наукових даних дослідника, експериментатор повинен:

- вивчити апріорну³ інформацію;
- спланувати весь експеримент та його етапи;
- розробити логічну процедуру і стратегію дослідження;
- створити фізичні та математичні моделі об'єкта вивчення – об'єкта дослідження;
- вибрати методи аналізу і обробки експериментальних даних;
- зробити науково-обґрунтовані висновки та розробити рекомендації тощо.

6. Зміст науки хемометрії. Система таких накопичених і підданих аналізу знань складає зміст науки хемометрії – обробки і моделювання експерименту в хемії і хемотехнології.

7. Науковий фундамент хемометрії. З формальної точки зору «хемометрія» уявляє собою сукупність математичних методів:

- теорія ймовірностей;
- математична логіка;
- математична статистика;
- теорії математичного планування експерименту;
- математичні методи пошуку оптимальних рішень тощо.

Ці методи застосовують, прикладають, описують на формалізованій математичній мові хемічний об'єкт досліджень.

8. Курс лекцій складається з таких частин:

- основних означень;
- законів розподілу випадкових величин і оцінок їх параметрів;
- основ теорії помилок;
- дисперсійної аналізу;
- кореляційної аналізу;
- регресійної аналізу;
- математичних методів планування експериментів;
- математичних методів пошуку оптимальних рішень;
- основ промислового експерименту.

9. **Предметом вивчення хемометрії** з кількісної та якісної сторін є хемічний, фізико-хемічний або хеміко-технологічний об'єкт вивчення. При цьому зазвичай маються на увазі не одиничні (принаймні розгляд і вивчення одиничних явищ теж можливо), а групові, масові явища процеси та об'єкти вивчення, до яких можна застосовувати теорію ймовірностей, статичні закони та методи.

1. Основні означення

1.1. Величина

Основне поняття, з яким ми зустрічаємось на кожному кроці у будь-якій природонауковій або технологічних галузях людської діяльності, - це поняття «величини».

Під **величиною** розуміють все те, що може бути виміряно і виражено числом або числами. У конкретних питаннях природничих наук і технологій зустрічаються з величинами різної природи (концентрація реагента, температура, швидкість хемічної реакції тощо).

Математичні положення і закони формують, абстрагуючись від конкретної природи величин, беручи до уваги лише їх числові значення. У відповідності до цього у математиці розглядають величину взагалі, відволікаючись від її фізичного, хемічного або біологічного смислу, тобто ведуть мову про абстрактність методів. Тому виникає проблема переведення мови експериментатора на математичну (абстрактну) мову об'єкта вивчення та об'єкта дослідження, а після отримання математичної моделі – навпаки.

Для вивчення біологічних явищ необхідно ввести поняття змінної величини.

Змінною називають величину, яка набуває різних чисельних значень.

Сталою називають таку величину, яка зберігає однакове чисельне значення.

Будь-який процес з кількісного боку характеризується взаємозміною кількох змінних величин. Тому зв'язок між змінними величинами може бути детермінованим (функціональним), а при дії додатково випадкової величини, і стохастичним.

1.2. Ознаки. Класифікація ознак

1. Розрізняють: **одиначні спостереження** і **одиначні результати експерименту** та **систематичну сукупність** – **множини** відносно однорідних, але

індивідуально різних одиниць, об'єднаних для сумісного, групового вивчення. Поняття статистичної сукупності базується на принципі якісної або кількісної однорідності її складу.

2. Розрізняють означення **статистичного комплексу**, що базується на принципі **сукупності різнорідних груп**, об'єднаних для сумісного комплексного вивчення. При цьому кожна група, яка входить до комплексу, повинна складатися із однорідних елементів.

3. Питання про форму об'єднання даних експериментатор повинен вирішувати самостійно у залежності від об'єкту вивчення та мети дослідження .

4. Результати спостережень, що об'єднали у статистичну сукупність або статистичний комплекс, уявляють собою певну **систему**, яка не зводиться до простої, механічної суми складових її одиниць або груп. У **статистичних сукупностях або комплексах існує певний внутрішній зв'язок** між частинами і цілим, одиночним і загальним, які виказують себе в статистичних закономірностях масових явищ.

5. Властивість, якою один предмет дослідження відрізняється від іншого, називається **ознакою**.

6. Характерною властивістю хемічних ознак є **варіювання величини** ознак у певних межах при переході від однієї одиниці спостережень до іншої. Ці **коливання величини однієї ознаки**, які спостерігаються в масі однорідних членів статистичної сукупності, носять назву **варіацій**⁴, а окремі числові значення варійованої ознаки називають **варіантами**⁵ (в англійській літературі числові значення ознак називають **датами**).

7. **Класифікація ознак**. Всі хемічні ознаки варіюються, але не всі піддаються прямому (безпосередньому) вимірюванню (опосередковані ознаки).

Ознаки розділяють на:

- кількісні;
- якісні (атрибутивні).

Якісні ознаки не піддаються безпосередньому вимірюванню або розрахункам, а визначаються за наявністю їх властивостей окремих членів групи, що вивчаються, наприклад, «+», «-», (+1, -1): блакитний – синій, голубий – сірий і т.ін.

Кількісні ознаки піддаються прямому вимірюванню або розрахунку. Їх розділяють на:

- метричні (мірні);
- меристичні (лічильні).

Поділ ознак на якісні і кількісні до певної міри є умовним.

8. Якщо результати спостережень об'єднують у групи, що протиставляються одна одній, їх **варіювання** (на візнаку від звичайної зміни) називають **альтернативними ознаками**, а спостереження – альтернативними.

1.3 Подія як результат досліду

1. Будь-якому закономірному процесові супутником є **випадкові відхилення**.

Точно передбачити кінцевий результат окремого процесу або досліду практично неможливо, так як випадкові явища⁶ кожен раз протікають по-різному. Але при багатократному повторенні досліду виявляються певні закономірності, які дозволяють прогнозувати кінцевий результат більшості дослідів.

2. **Результат** будь-якого досліду, який здійснюється при певних умовах, є **подія**.

3. Подія може мати якісну або кількісну характеристику.

4. **Подію**, яка внаслідок певного досліду, повинна неодмінно відбутися називають **вірогідною**. Прикладом вірогідної події є витягування чорної кульки із комірки, у якій знаходяться лише чорні кульки.

5. Вірогідній події протилежна **неможлива подія**, тобто така подія, яка у даних умовах досліду не може відбутися, наприклад, витягування білої кульки із комірки із лише чорними кульками. Китайський вираз філософа: «Неможливо знайти чорну кішку в темній кімнаті, коли там її немає» розкриває суть неможливої події.

6. Подія, яка внаслідок даного досліду може відбутися, а може ні, називається **випадковою**. Прикладом випадкової події є витягування білої кульки із комірки, в якій знаходяться білі і чорні кульки. Тут результат випробувань точно передбачити неможливо.

7. Випадкові події називаються **несумісними**, якщо поява однієї з них викликає появу іншої.

Наприклад, полімерний зразок при термомеханічних випробуваннях або зруйнується або ні. Одночасна поява цих двох подій неможлива.

8. У тих випадках, коли поява однієї події не виключає можливості появи іншої, випадкові події називають **сумісними**.

9. Якщо в результаті даного досліду обов'язково відбудеться одна із кількох подій, то такі випадкові події називаються **єдино можливими**.

Так, дві події – руйнування або неруйнування зразка при термомеханічних випробуваннях – є **єдино можливими випадковими подіями**.

10. Результатом низки дослідів можуть бути такі події, кожна з яких не є об'єктивно більш можливою за іншу. Такі події називаються **рівноможливими випадковими подіями**, наприклад, випадання герба або решки при киданні монети.

11. Випадкові події можуть бути залежними чи незалежними.

Дві події називають **незалежними**, якщо ймовірність однієї події не залежить від настання або ні іншої події.

Дві події називають **залежними**, якщо ймовірність появи однієї з них залежить від настання або ні іншої.

12. Події A_1, A_2, \dots, A_n уявляють собою **повну систему випадкових подій**, якщо результатом досліду може з'явитися одна і тільки одна з цих подій. Кожна подія, яка може відбутися у випробуванні, називається **елементарним випадком**. Всі ті випадки, за яких настає подія, яку ми очікуємо внаслідок досліду, називають **сприятливою подією**.

1.4. Об'єкт вивчення та об'єкт дослідження

На стан **об'єкта вивчення** (рис. 1.1) впливають: вхідна змінна X , яка піддається контролю та керуванню. Вона містить значну кількість (k) вхідних параметрів $x_1, x_2, \dots, x_i, \dots, x_k$ (факторів), кожен з яких, у свою чергу, можна варіювати на m_k рівнях; W – змінна, контрольована, але некерована, яка теж містить W_L змінних, діючих систематично чи (та) випадково; Z – змінна,

неконтрольована та некерована, яка теж містить Z_R змінних, діючих випадково на об'єкт вивчення.

Стан об'єкта вивчення характеризується вихідним параметром (функцією відгуку, функцією цілі, параметром оптимізації) Y , яка має складові $y_1, y_2, \dots, y_j, \dots, y_m$, які всебічно описують об'єкт вивчення. Вхідних та вихідних параметрів, а також рівнів їх варіювання занадто багато, щоби провести процес дослідження об'єкта вивчення.

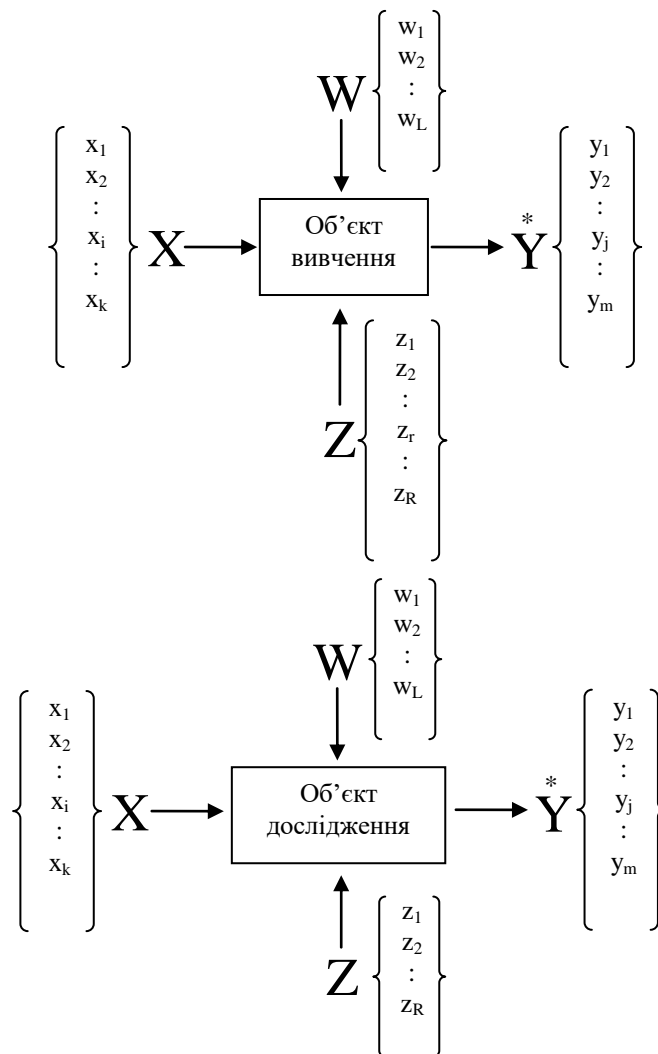


Рис 1.1. Об'єкт вивчення та об'єкт дослідження: X – вхідна змінна (контрольована, керована); Y – вихідна змінна; W – контрольована, але не керована змінна; Z – неконтрольована і некерована змінна.

Тому, за певними правилами та процедурами їх зменшують (відсіють) кількість вхідних параметрів, їх рівні варіювання, вихідні параметри та можливо, і

Z_r , які систематично впливають на об'єкт вивчення) і, таким чином, переходять до моделі об'єкта вивчення – об'єкта дослідження (рис 1.1.), для якої $k \ll K$, $m \ll M$, $l \ll R$.

1.5. Ймовірність

1. Ступінь можливості (число) реалізації випадкової події називають **ймовірністю**.

2. Якщо дослід зводиться до певної системи випадків, то під ймовірністю події A розуміють відношення кількості випадків, що сприяють даній події, до загального числа випадків всіх несумісних, єдиноможливих і рівноможливих випадків:

$$p(A) = \frac{m}{N},$$

де $p(A)$ – ймовірність події A ;

m – кількість випадків, що сприяє події A ;

N – загальне число випадків.

Наприклад, в комірці знаходиться 4 кульки білого кольору і 6 кульок чорного. Ймовірність того, що з $N=10$ кульок (можливих випадків) може бути витягнута біла кулька $m=4$, дорівнює:

$$p(A) = \frac{4}{10} = 0,4 \text{ (40\%)}. \quad (1.1)$$

3. Із визначення ймовірності витікають наступні властивості:

ймовірність **достовірної** події $p(A)=1$ або (100%), бо кожний можливий випадок є сприятливим $m=N$;

ймовірність **неможливої** події $p(A)=0$ (0%), тобто немає жодної події, що сприяє події A ($m=0$);

ймовірність випадкової події є позитивне число:

$$0 < p(A) < 1 \quad (1.2)$$

$$0\% < p(A) < 100\% \quad (1.3)$$

4. У більшості випадків ймовірність події A не можна знайти аналітичним шляхом. І тому ймовірності дають оцінку на основі результатів дослідів за допомогою **накопиченої частоти випадкової події**.

Хай реалізують n дослідів, в кожному з яких може з'явитися або ні подія A . У результаті цих дослідів подія A з'явилася i разів (число спостережень), тоді **накопичена частота події A , є статистичним аналогом ймовірності:**

$$W(A) = \frac{i}{n} \quad (1.4)$$

$$W(A) \xrightarrow{\text{імовірність}} p(A) \quad (1.5)$$

Наприклад, з 50 відібраних для аналізу фіксаналів технічний контроль виявив 5 з відхиленням від ТУ, тоді накопичена частота бракованих фіксаналів становить:

$$W_1(A) = \frac{i_1}{n_1} = \frac{5}{50} = \frac{1}{10} = 0,1 \text{ (10\%)} \quad (1.6)$$

Повторюючи серію з n дослідів багаторазово, будемо отримувати для накопиченої частоти події A (відхилення від ТУ) різні значення, але близькі для числа, що є ймовірністю події A :

$$W_2(A) = \frac{i_2}{n_2} = 0,12; \quad W_3(A) = \frac{i_3}{n_3} = 0,09; \dots \quad (1.7)$$

При зростанні кількості дослідів відхилення $W(A)$ від $p(A)$ буде зустрічатися рідко і при $N \rightarrow \infty$ $W(A) \approx p(A)$.

¹**обсервація** [лат. observation – спостереження, обстеження; система заходів за спостереженнями, обстеженнями].

²**емпіричний** [від гр. emperiа – дослід] Це в теорії пізнання є напрямок, який визначає дослід джерелом знань. При цьому зміст знання може бути представлено як опис цього дослідження або зведено до нього.

³**апріорний, апіорі** [від лат. a priori, букв. такого, що йде попереду, який передуює, виступає, відбувається попереду знання], який передуює дослідженню, експерименту і є незалежним від нього.

⁴**варіація** [від лат. variatio – зміна, коливання];

⁵**варіант** [від лат. varians, variantis – змінний, мінливий].

⁶**Закономірності** у випадкових явищах розглянуті в теорії ймовірностей, а **методи аналізу** результатів масових дослідів – у математичній статистиці.