Основи наукових досліджень в теплоенергетиці та гідроенергетиці

Лекція 8
Статистичні похибки та їх усунення

План лекції

1. Основні поняття вимірювання
2. Похибки вимірювань
3. Інтервальна оцінка за допомогою довірчої вірогідності
4. Встановлення мінімальної кількості вимірювань
5. Усунення грубих похибок

Рекомендована додаткова література

1. Чейлитко, А. О. Формування теплофізичних властивостей елементів конструкцій теплового захисту шляхом створення прогнозованих пористих структур : монографія / А. О. Чейлитко. — Запоріжжя : ЗДІА, 2017. — 318 с.
2. Ільїн С.В. Енергоаудит. Навчально-методичний посібник по курсу: «Енергоаудит» для слухачів курсів підвищення кваліфікації центру безперервної освіти. /Укл.: Ільїн С.В., Чейлитко А.О., Мних І.М. – Запоріжжя, 2018. – 130 с.
3. Семененко М.П. Методы обработки и анализа измерений в научных исследованиях. - Киев; Донецк: Вища школа, 1983. - 240 с.
4. Опря А. Т. Статистика. Навч. посіб. – К.: Центр учбової літератури, 2012. – 448 с.
5. Беляев Ю.К Статистические методы обработки результатов испытаний на надежность. - М.: Знание, 1982

Основні поняття

Розрізняють три класи вимірів:

* *Особливо точні* - еталонні виміри з максимально можливою точністю.
* *Високоточні* - виміри, погрішність яких не повинна перевищувати заданих значень.
* *Технічні* - виміри, в яких погрішність визначається особливостями засобів виміру.

Визначення

Точність виміру - це міра наближення виміру до дійсного значення величини.

Відхилення показання приладу від дійсного значення вимірюваної фізичної величини є ***похибкою приладу***.

Достовірність виміру показує міру довіри до результатів виміру, тобто вірогідність відхилень виміру від дійсних значень.

Методи вимірювань — фізичні явища, на яких засноване вимірювання.

Засіб вимірювання — технічний пристрій, який використовується для вимірювань і має нормовані метрологічні характеристики.

Похибки результатів вимірювань

- *абсолютна* $b_{\text{а}} = \Delta x = \pm\,(x_в - x_д)$

де $x_в$ - свідчення приладу (номінальне значення вимірюваної величини); $x_д$ - дійсне значення виміряної величини точнішім методом.

- відносна , %

$$b_{\text{від}} = \pm \frac{x_в - x_д}{x_д} \cdot 100\%$$

- приведена ,

$$b_{\text{пр}} = \pm \frac{x_в - x_д}{x_н} \cdot 100\%$$

де $x_н$ - нормоване значення: верхня межа вимірювання приладу, номінальне значення фізичної величини, що вимірюється, діапазон вимірювання, довжина шкали приладу.

Послідовні операції при вимірюванні

1. вибір методу вимірювання;
2. вибір, перевірка і установка приладів;
3. спостереження свідчень приладів і відлік;
4. обчислення шуканої величини з результатів вимірювань, аналіз та обчислення похибок.

Клас точності приладу

Точність приладу характеризується сумарною погрішністю. Засоби виміру діляться на класи точності залежно від похибок, що припускаються.

Клас точності приладу є узагальненою характеристикою точності приладу, яка визначає межі допустимих основної та додаткової похибки.

Умовні позначення класу точності:

1,5 – клас точності приладу в діапазоні вимірювання;

- клас точності приладу від одного показання;

- клас точності приладу від інтервалу вимірювань;

клас точності приладу від довжини шкали.

Похибки вимірів

Систематична

- складова похибки виміру, що залишається постійною або що закономірно змінюється при повторних вимірах однієї і тієї ж величини.
- Причиною появи такою похибки можуть бути несправність вимірювальної апаратури, невдосконаленість метода вимірювання, неправильна установка вимірювальних приладів

Випадкова

- це складова похибки вимірювання, яка змінюється випадковим чином, хаотично, нерегулярно при повторних вимірюваннях однієї і тієї ж величини в однакових умовах. Наявність випадкових похибок виявляється при проведенні черги вимірювань цієї величини, коли можна признати, що результати вимірювань не співпадають друг з другом.

Груба (надмірна, промах)

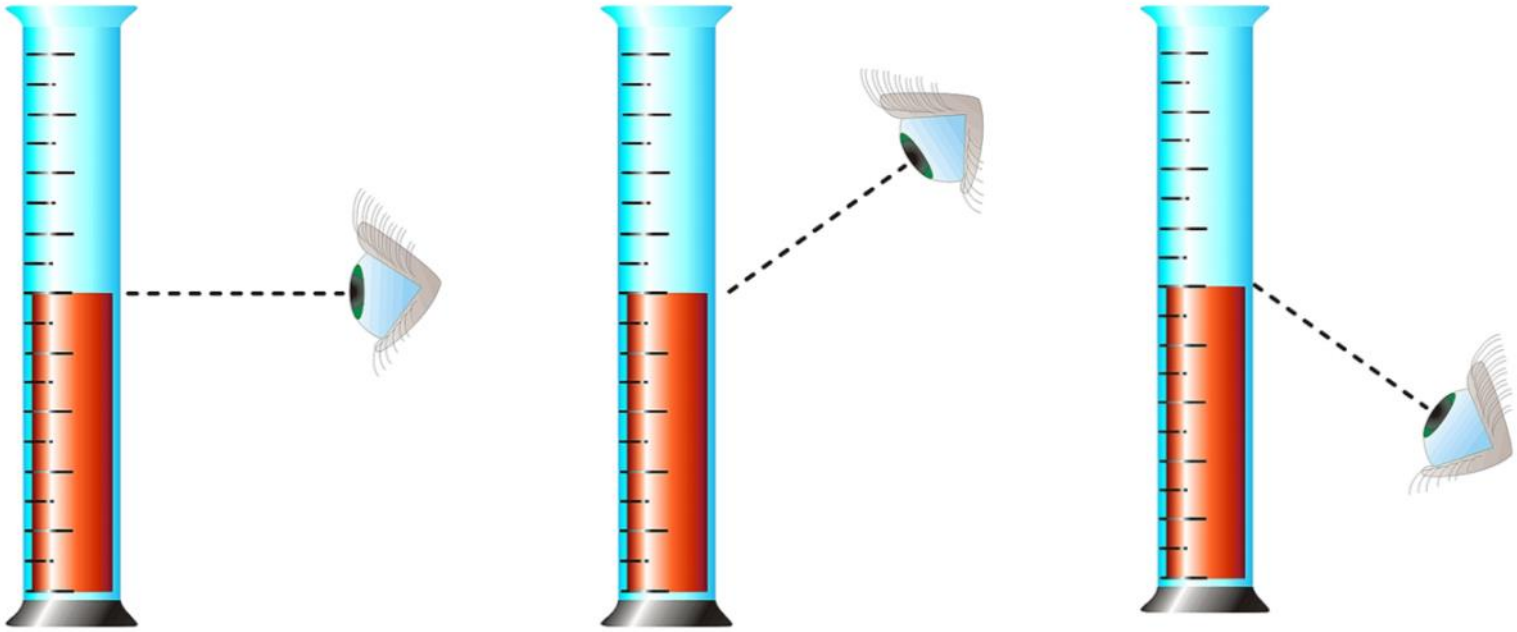
- це випадкова похибка, що суттєво перевищує очікувану за даних умов. Причиною грубої погрішності можуть бути короткочасна зміна живлючої напруги, неправильний відлік показань приладу, похибки експериментатора тощо.

Систематичні похибки

Випадкові похибки

Випадкову складову похибки виключити або зменшити неможливо, але можна врахувати її вплив, зробивши статистичну обробку ряду вимірювань. У цьому випадку доводиться виконувати багаторазові вимірювання однієї і тієї ж величини.

Випадкова помилка вимірювання залежить від експериментатора!!!



Теорія випадкових помилок

Теорія випадкових помилок дозволяє розв’язати дві основні задачі: оцінити точність і надійність вимірювання при даній кількості вимірів; визначити мінімальну кількість вимірів, що гарантує необхідну (задану) точність і надійність вимірювання. Разом з цим виникає необхідність виключити грубі помилки ряду, визначити достовірність одержаних даних та ін.

Інтервальна оцінка за допомогою довірчої вірогідності

Для великої вибірки і нормального закону розподілу загальною оцінною характеристикою виміру є дисперсія Д і коефіцієнт варіації

$$Д = \frac{1}{n-1} \cdot \sum_{1}^{n} \left(x_i - \bar{x}\right)^2 \qquad K_s = \frac{\sigma}{\bar{x}}$$

Дисперсія характеризує однорідність виміру. Чим вище Д, тим більше розкид вимірів. Коефіцієнт варіації характеризує мінливість. Чим він вищий, тим більше мінливість вимірів відносно середніх значень.

Дисперсія

Дисперсія вибірки обсягом n визначається як:

$$s_x^2 = \frac{(x_1 - \overline{X})^2 + (x_2 - \overline{X})^2 + \ldots + (x_n - \overline{X})^2}{n-1}, \qquad (2.3)$$

або $s_x^2 = \frac{\sum (x_i - \overline{X})^2}{n-1}$, де $\overline{X}$ – середнє арифметичне вибірки.

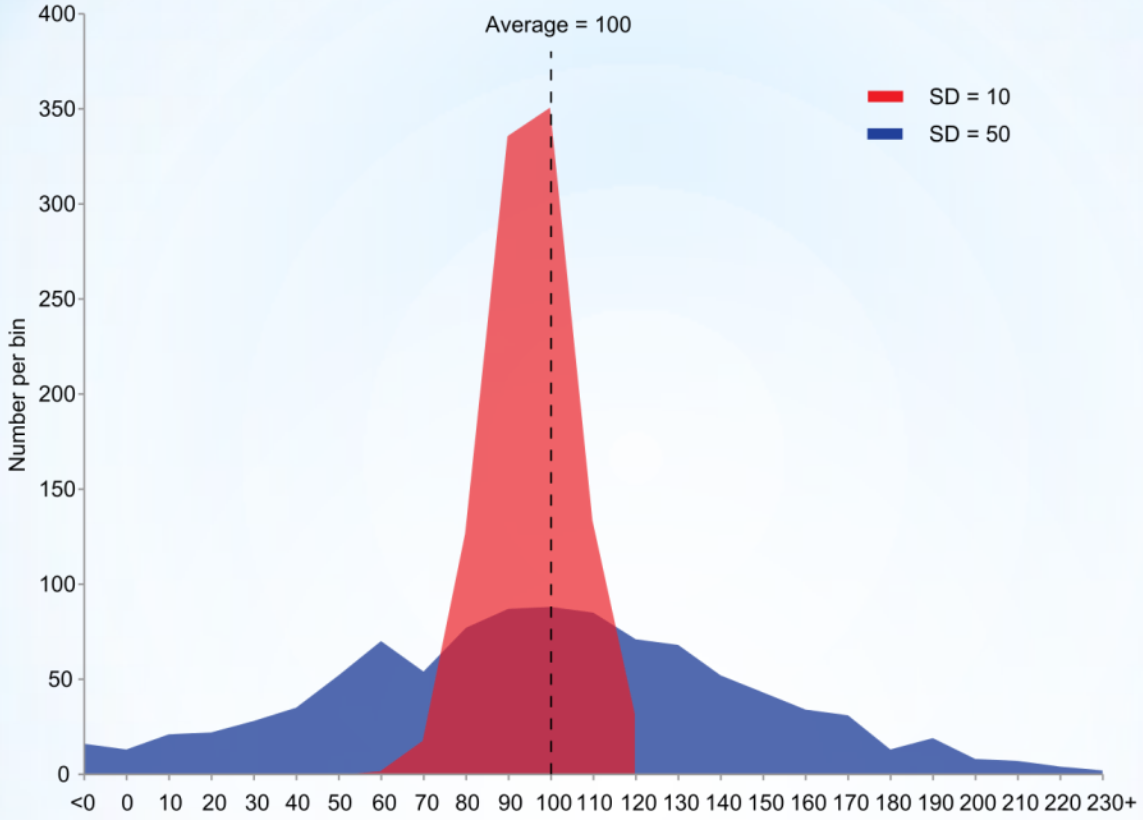
Дисперсія вибірки s^2_x, що розрахована за цією формулою, є незміщеною оцінкою свого генерального параметра σ^2_x завдяки внесенню поправки Бесселя $n/(n–1)$, тобто:

$$s_x^2 = \frac{\sum (x_i - \overline{X})^2}{n} \cdot \frac{n}{n-1} = \frac{\sum (x_i - \overline{X})^2}{n-1}. \qquad (2.4)$$

Різницю $n–1$ називають *числом степенів вільності* k – кількість об'єктів або значень у складі обмеженої статистичної сукупності, які можуть вільно варіювати. Якщо обмежень вільності варіації існує декілька (v), то число степенів вільності дорівнюватиме $k= n–v$ (де v – грецька літера «ню»).

Чисельник формули дисперсії можна перетворити у такий спосіб:

Дисперсія та математичне сподівання



*Однакове середнє, різні дисперсії

Стандартне відхилення

Стандартне відхилення вибірки визначається як $s_x = \sqrt{s_x^2}$. (2.8)

Стандартне відхилення генеральної сукупності $\sigma_x = \sqrt{\sigma_x^2}$. (2.9)

Коефіцієнт варіації V_x використовується у разі порівняльної оцінки різноякісних середніх величин і визначається (у тому числі у %) як відношення стандартного відхилення до середнього арифметичного:

$$V_x = s_x / \overline{X} \cdot 100\% \, . \qquad (2.10)$$

Найбільш імовірне значення вимірюваної величини ($x_{вимір}$) дорівнює середньому арифметичному значень, одержаних у результаті вимірювань:

$$x_{вимір} = x_{сер} = \frac{x_1 + x_2 + \ldots + x_N}{N}.$$

Де:

x_1 — результат першого вимірювання,

x_2 — другого вимірювання,

x_N — N-го вимірювання

Випадкова абсолютна похибка ($X_{вип}$) — *середня помилка, одержана в результаті всіх вимірювань*,— обчислюється за формулою:

$$\Delta x_{вип} = \sqrt{\frac{\left(x_1 - x_{сер}\right)^2 + \left(x_2 - x_{сер}\right)^2 + \ldots + \left(x_N - x_{сер}\right)^2}{N}}.$$

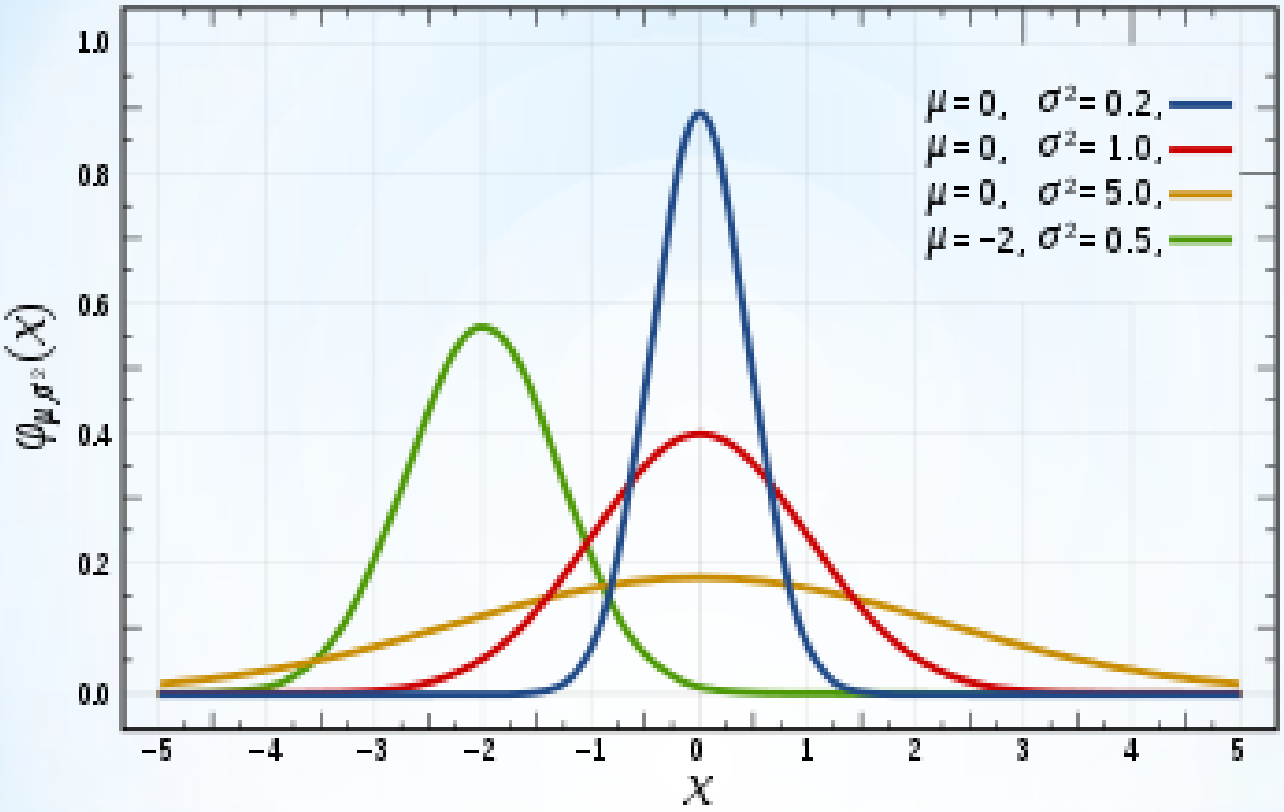
* *нормальний закон розподілу Гауса* - **щільність розподілу вірогідності, при якому вірогідність визначається рівнянням**

$$P(x) = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot \exp\left[-\frac{(x - X)^2}{2 \cdot \sigma^2} \right]$$

де X - дійсне значення вимірюваної величини.

Закон Гауса

- для великої кількості вимірювань різні за модулем похибки зустрічаються однаково часто;
- малі за модулем похибки зустрічаються частіше, ніж великі, тобто імовірність появи похибки тим менша, чим більше її абсолютне значення;
- похибки вимірювань становлять неперервний ряд значень.



Нормальне розподілення Гауса

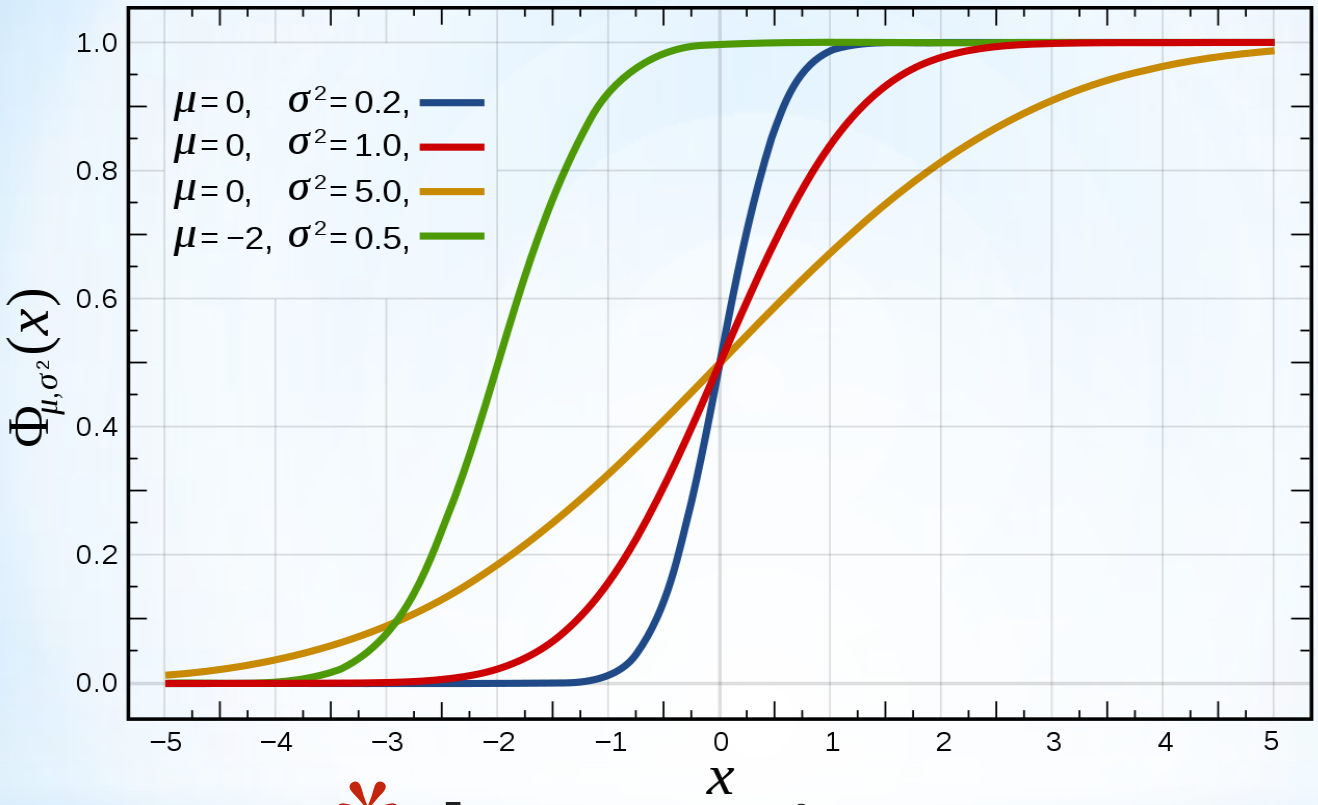
Інтервальна оцінка за допомогою довірчої вірогідності

Довірчим називається інтервал значень x_i, в який потрапляє дійсне значення $x_д$ вимірюваної величини із заданою вірогідністю.

Довірча вірогідність описується наступним вираженням

$$P_д = P(a < m(x) < b) = \frac{1}{2} \cdot \left[\Phi\left(\frac{b - \bar{x}}{\sigma} \right) - \Phi\left(\frac{a - \bar{x}}{\sigma} \right) \right]$$

де $\Phi(t)$ - функція Лапласа, аргументом якої є відношення відхилення μ до середньоквадратичному σ, тобто $t = \mu / \sigma$, де $\mu = b - \bar{x} = -(a - \bar{x})$, t - гарантійний коефіцієнт. Чисельне значення $\Phi(t)$ наводяться в таблицях.



*Функція розподілу ймовірностей

t	Φ(t)	t	Φ(t)	t	Φ(t)
0,00	0,0000	0,75	0,5467	1,50	0,8664
0,05	0,0399	0,80	0,5763	1,55	0,8789
0,10	0,0797	0,85	0,6047	1,60	0,8904
0,15	0,1192	0,90	0,6319	1,65	0,901 1
0,20	0,1585	0,95	0,6579	1,70	0,9109
0,25	0,1974	1,00	0,6827	1,75	0,9199
0,30	0,2357	1,05	0,7063	1,80	0,9281
0,30	0,2357	1,05	0,7063	1,80	0,9281
0,35	0,2737	1,10	0,7287	1,85	0,9357
0,40	0,3108	1,15	0,7419	1,90	0,9426
0,45	0,3473	1,20	0,7699	1,95	0,9488
0,50	0,3829	1,25	0,7887	2,00	0,9545
0,55	0,4177	1,30	0,8064	2,25	0,9756
0,60	0,4515	1,35	0,8230	2,50	0,9876
0,65	0,4843	1,40	0,8385	3,00	0,9973
0,70	0,5161	1,45	0,8529	4,00	0,9999

* Чисельні значення інтегральної функції Лапласа

Виконано 30 вимірювань міцності покриття труб.

При цьому середній модуль пружності покриття E_c =170 МПа. Значення середньоквадратичного відхилення σ =3,1 МПа. Визначити точність і достовірність експерименту.

* Необхідну точність вимірювань визначимо для різних рівнів довірчої вірогідності, взявши відповідно значення arg Φ(t) за таблицею для $P_д$ = 0,9011; 0,9545; 0,9973;

μ = ±3,1 • 1,65 = ± 5,1; ±3,1 • 2,0 = ± 6,2; ±3,1 • 3 = ± 9,3МПа.

* Необхідно визначити достовірність вимірювань для встановленого довірчого інтервалу μ =± 7 МПа.

* Для t = μ / σ =7/3,1=2,26 визначаємо $P_д$ = 0, 9761. Це означає, що в заданий довірчий інтервал із 100 вимірювань не потрапляють тільки два-три.

Рівень значущості

Значення $1-\Phi(t)$ називають рівнем значущості. З нього виходить, що при нормальному законі розподілу погрішність, що перевищує довірчий інтервал, зустрічатиметься один раз з $n_и$ вимірів

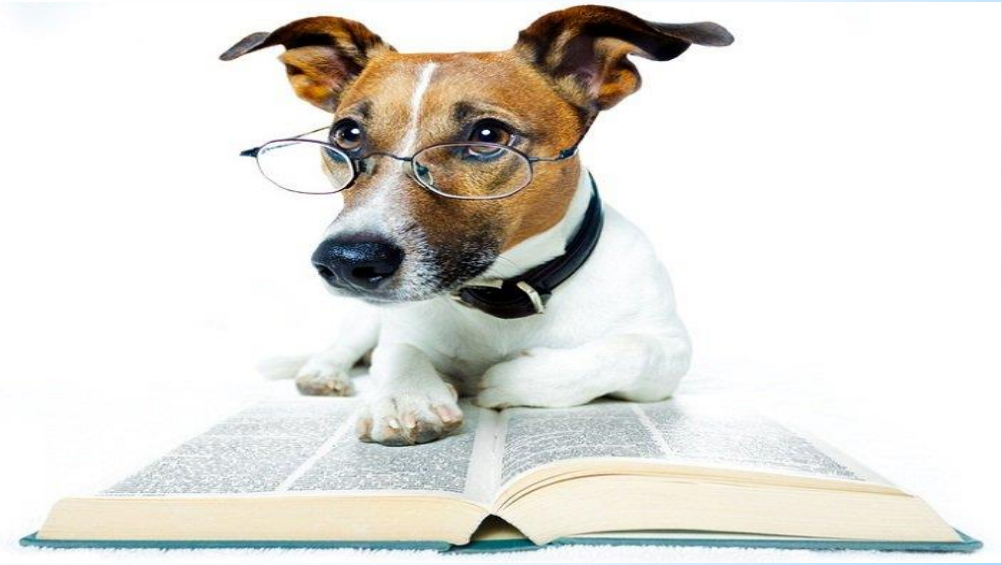
$$n_и = \frac{P_д}{1-P_д}$$

або інакше доводиться бракувати один з $n_и$ вимірів.

для $P_д$ = 0,9

n = 0,9/(1 - 0,9)= 9 вимір.

Для $P_д$ 0,95 і 0,9973, відповідно 19 і 367 вимірювань.



Встановлення мінімальної кількості вимірів

Задача зводиться до встановлення мінімального об'єму вибірки (числа вимірів) N_{min} при заданих значеннях довірчого інтервалу 2μ і довірчої вірогідності. При виконанні вимірів необхідно знати їх точність Δ, яку зазвичай характеризують σ_0 - середньоарифметичне значення середньоквадратичного відхилення

$$\sigma_0 = \sigma / \sqrt{n}\,;\ \Delta = \sigma_0 \cdot \sqrt{x}$$

Значення σ_0 часто називають середньою помилкою. Довірчий інтервал помилки виміру Δ визначається аналогічно, як і для вимірів $\mu = t \cdot \sigma_0$.

Мінімальну кількість вимірів можна визначити по наступній формулі

$$N_{\min} = \frac{\sigma^2 \cdot t^2}{\sigma_0^2} = \frac{K_в^2 \cdot t^2}{\Delta^2}\ ,$$

тут $K_в$ - коефіцієнт варіації, %;

Δ - точність вимірів, %.

Послідовність обчислення мінімальної кількості вимірів

1. Проводять попередній експеримент з кількістю вимірів *n*, яке складає залежно від трудомісткості досліду від 20 до 50.

2. Обчислюють середньоквадратичне відхилення σ.

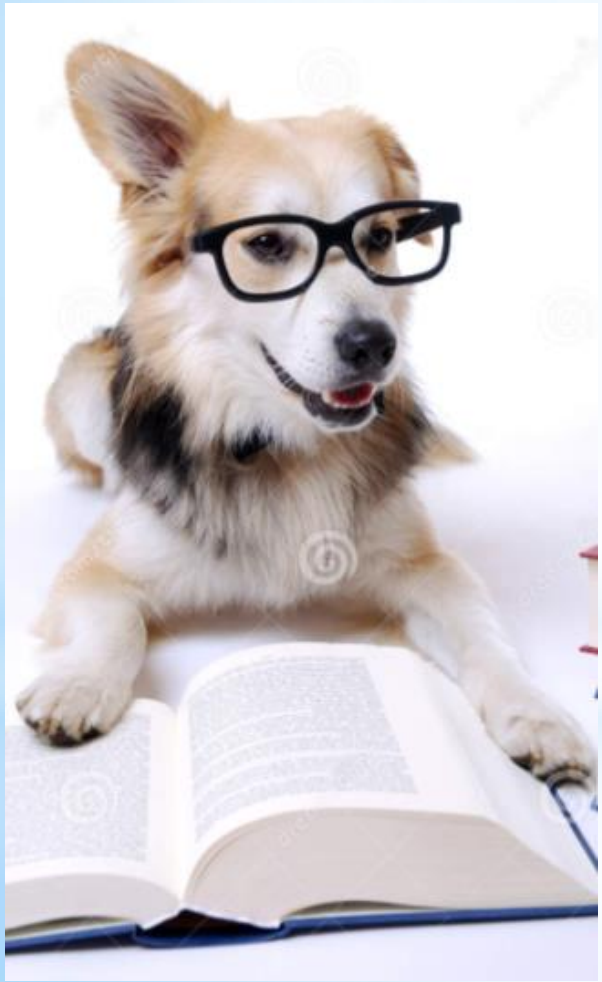
3. Відповідно до поставлених задач експерименту встановлюють необхідну точність вимірів , μ, яка має бути не менше точності приладу.

4. Встановлюють нормоване відхилення *t*, значення якого зазвичай задають; воно залежить також від точності методу. Наприклад, при великій точності вимірів можна прийняти *t=3,0*, при малій - *t=1,0*.

5. Визначають N_{min}. Надалі число вимірів не має бути менше розрахованого значення.

Згідно з тимчасовою інструкцією вимагається виконувати 25 вимірювань; Обчислене значення σ = 0,4 м.
Відхилення параметра, що допускається ± 0,1 м.
Необхідно визначити, з якою достовірністю комісія оцінює даний параметр.

$$t = \sqrt{n} \cdot \frac{\Delta}{\sigma};$$



t = 1,25 Рд = 0,79.
Похибка, що перевищує довірчий інтервал 0,2 м, траплятиметься один раз з 0,79/(1 - 0,79)= 3,76, тобто з 4 вимірювань.
Обчислимо мінімальну кількість вимірювань з довірчою вірогідністю Рд, що дорівнює 0,9 і 0,95.
Nmin = 0,42·1,652/0,12 = 43 вимірювання при Рд= 0,90 і
64 вимірювання при Рд = 0,95.

t	Φ(t)	t	Φ(t)	t	Φ(t)
0,00	0,0000	0,75	0,5467	1,50	0,8664
0,05	0,0399	0,80	0,5763	1,55	0,8789
0,10	0,0797	0,85	0,6047	1,60	0,8904
0,15	0,1192	0,90	0,6319	1,65	0,901 1
0,20	0,1585	0,95	0,6579	1,70	0,9109
0,25	0,1974	1,00	0,6827	1,75	0,9199
0,30	0,2357	1,05	0,7063	1,80	0,9281
0,30	0,2357	1,05	0,7063	1,80	0,9281
0,35	0,2737	1,10	0,7287	1,85	0,9357
0,40	0,3108	1,15	0,7419	1,90	0,9426
0,45	0,3473	1,20	0,7699	1,95	0,9488
0,50	0,3829	1,25	0,7887	2,00	0,9545
0,55	0,4177	1,30	0,8064	2,25	0,9756
0,60	0,4515	1,35	0,8230	2,50	0,9876
0,65	0,4843	1,40	0,8385	3,00	0,9973
0,70	0,5161	1,45	0,8529	4,00	0,9999

* Чисельні значення інтегральної функції Лапласа

Згідно з тимчасовою інструкцією вимагається виконувати 2 вимірювань;
Обчислене значення σ = 0,1 м.
Відхилення параметра, що допускається ± 0,1 м.
Необхідно визначити, з якою достовірністю комісія оцінює даний параметр.

$$t = \sqrt{n} \cdot \frac{\Delta}{\sigma};$$

t = 1,4 Рд = 0,8385

σ = 0,1 м. Відхилення ± 0,1 м.
n=4

t = 2 Рд = 0,9545

Виключення грубих помилок ряду

1. Правило трьох сигм. Розкид випадкових величин від середнього значення не перевищує $x_{max,\ min} = \bar{x} \pm 3 \cdot \sigma$

2. Метод, що базуються на використанні довірчого інтервалу

За наявності грубих помилок критерії їх появи

$$\beta_1 = \frac{x_{max} - \bar{x}}{\sigma \cdot \sqrt{\frac{n-1}{n}}};\ \beta_2 = \frac{\bar{x} - x_{min}}{\sigma \cdot \sqrt{\frac{n-1}{n}}},$$

де x_{max}, x_{min} - відповідно найбільше і найменше значення з n вимірів.

3. на використанні критерію Романовського

Задаються довірчою вірогідністю $P_д$, і по таблиці залежно від n знаходять коефіцієнт q. Обчислюють гранично допустиму абсолютну помилку окремого виміру $\varepsilon_{пр} = \sigma \cdot q$.

Якщо $\bar{x} - x_{max} > \varepsilon_{пр}$, то вимір x_{max} виключають з ряду спостережень.

Методика глибокого аналізу виключення помилок ряду

1. Після здобуття експериментальних даних у виді статистичного ряду заздалегідь виключають систематичні помилки.

2. Аналізують ряд в цілях виявлення грубих помилок і промахів:

- встановлюють підозрілі значення x_{max} і x_{min};
- визначають середньоквадратичне відхилення σ;
- обчислюють критерії β_1 і β_2 і зіставляють з β_{max} і β_{min} або $\varepsilon_{пр}$;
- виключають при необхідності із статистичного ряду x_{max} або x_{min} і отримують новий ряд з нових членів;

3. Обчислюють середньоарифметичне $\bar{x}$, погрішності окремих вимірів $\bar{x} - x_i$ і середньоквадратичне очищеного ряду σ.

4. Знаходять середньоквадратичне σ_0 серії вимірів, коефіцієнт варіації $K_в$.

5. При великій вибірці задаються довірчою вірогідністю $P_д = \Phi(t)$ або рівнянням значущості *1*- і по таблиці визначають *t*. При малій вибірці ($n \le 20$) залежно від прийнятої довірчої вірогідності $P_д$ і числа членів ряду *n* приймають коефіцієнт Стьюдента $\alpha_{ст}$; для великої вибірки по формулі $Д = \frac{1}{n-1} \cdot \sum_1^n \left(x_i - \bar{x}\right)^2$; $K_в = \frac{\sigma}{\bar{x}}$, а для малої вибірки по формулі $\mu_{ст} = \sigma_0 \cdot \alpha_{ст}$ визначають довірчий інтервал.

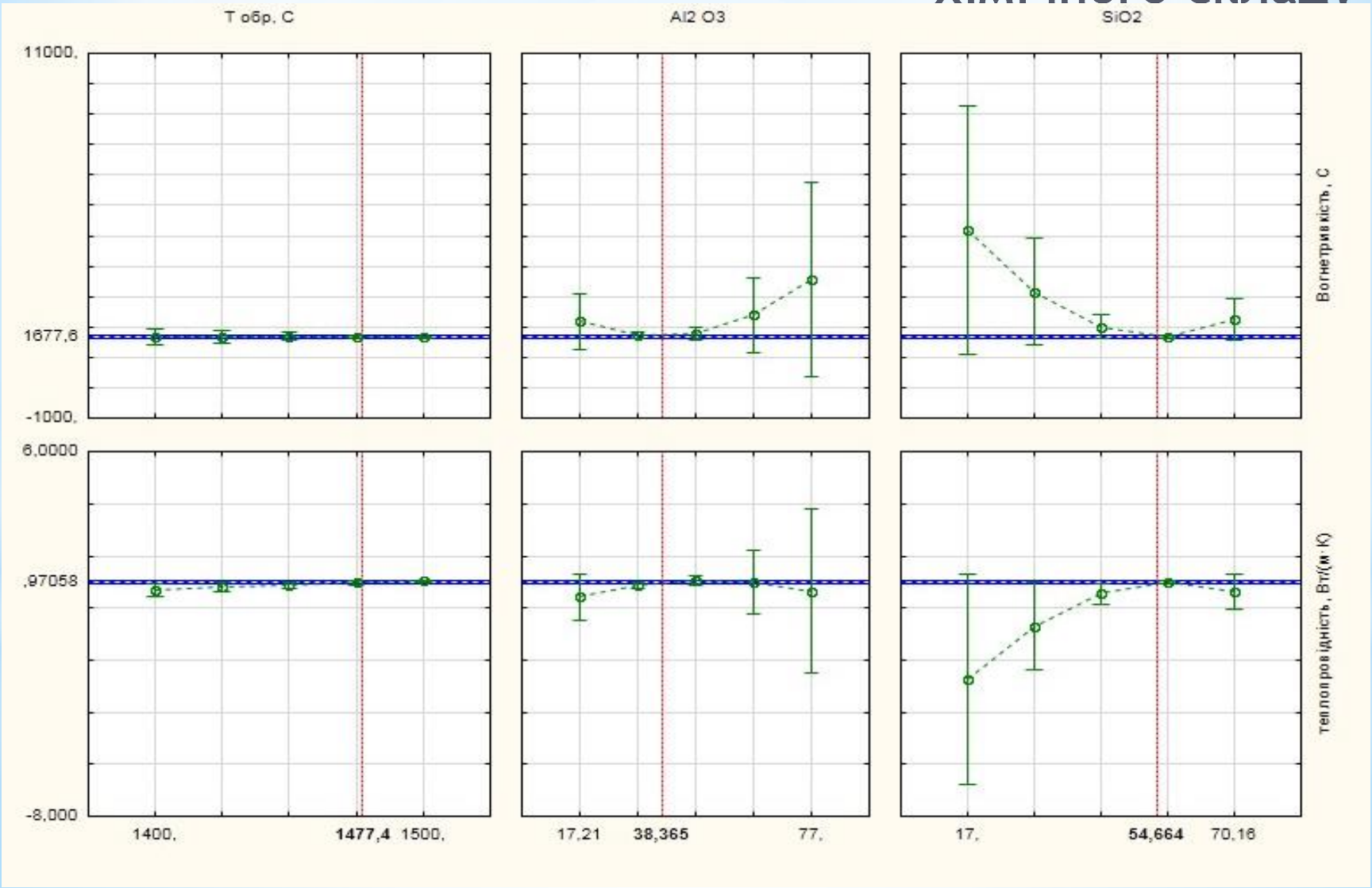
Методика глибокого аналізу виключення помилок ряду

6. Встановлюють дійсне значення досліджуваної величини.

7. Оцінюють відносну погрішність результатів серії вимірів при заданій довірчій вірогідності $P_д$

$$\delta = \frac{\sigma_0 \cdot \alpha_{cm}}{\bar{x}} \cdot 100\%$$

* Графіки залежності вогнетривкості та теплопровідності від температури обробки та хімічного складу



Дякую за увагу!

Лекцію підготував

доктор технічних наук, доцент,академік Європейської науково-освітньої академії,

Чейлитко Андрій Олександрович

cheilytko@i.ua