

Міністерство освіти і науки, молоді та спорту України
Криворізький державний педагогічний університет
Фізико-математичний факультет
Кафедра фізики

СТАТИСТИЧНА ОБРОБКА РЕЗУЛЬТАТІВ
ЕКСПЕРИМЕНТАЛЬНИХ ВИМІРЮВАНЬ

Кривий Ріг – 2011

УДК 53.088

ББК 22.3я7

Статистична обробка результатів експериментальних вимірювань : методичні рекомендації / укл. В. П. Ржепецький. – Кривий Ріг : Криворізький державний педагогічний університет, 2011. – 48 с.

Укладач: Ржепецький Василь Петрович, кандидат фізико-математичних наук, доцент кафедри фізики КДПУ.

Рецензент: Мулявко Валерій Іванович, доктор технічних наук, професор кафедри фізики КТУ.

Затверджено на засіданні кафедри фізики

Протокол № 6 від 16.12.2010 р.

Затверджено радою фізико-математичного факультету
Криворізького державного педагогічного університету

Протокол № 4 від 30.12.2010 р.

ББК 22.3я7

© КДПУ, 2011

ЗМІСТ

Вступ.....	4
Розділ 1. Похибки вимірювань.	
1.1. Основні поняття процесу вимірювання.....	5
1.2. Похибки вимірювань.....	8
Розділ 2. Обчислення похибок.	
2.1. Обчислення похибок при прямих вимірюваннях.....	10
2.2. Обчислення похибок при непрямих вимірюваннях...	18
Розділ 3. Знаходження кривих, що зображують зв'язок між досліджуваними величинами.	
3.1. Графічне представлення даних експерименту.....	23
3.2. Апроксимація методом найменших квадратів.....	25
Розділ 4. Деякі застосування статистичних методів до педагогічних досліджень.	
4.1. Наочне представлення даних.....	29
4.2. Встановлення кореляції між результатами сумісних вимірювань.....	32
4.3. Перевірка статистичних гіпотез.....	37
Література.....	42
Додатки.....	43

ВСТУП

Результати будь-яких вимірювань, як би ретельно і на якому б науковому рівні вони не проводились, завжди містять деякі похибки. Теорія похибок – це наука, що вивчає і оцінює похибки вимірювань. Вона дає можливість досліднику визначити значення похибок в його вимірюваннях і, якщо це необхідно, зменшити їх. Аналіз похибок є важливою частиною будь-якого наукового експерименту, тому теорія похибок повинна бути невід’ємною частиною всіх курсів з експериментальних наук. В даній роботі описані основні ідеї та поняття математичної обробки результатів вимірювань та наведені алгоритми розрахунків при прямих та непрямих вимірюваннях. Студенти, які оволоділи навичками статистичної обробки результатів вимірювань, зможуть в майбутньому вірно організувати цю роботу в школі. В роботі також наведені деякі застосування статистичних методів обробки результатів до педагогічних досліджень.

РОЗДІЛ 1

ПОХИБКИ ВИМІРЮВАНЬ

1.1. Основні поняття процесу вимірювання

Вимірюванням називають знаходження значень фізичної величини дослідним шляхом з допомогою спеціальних технічних засобів [5]. В найпростішому випадку **виміряти** яку-небудь фізичну величину означає **порівняти її з іншою однорідною фізичною величиною, взятою за одиницю вимірювання.**

Результат вимірювання фізичної величини A подають у вигляді:

$$A = \{A\}[A], \quad (1.1)$$

де $\{A\}$ - числове значення величини A , $[A]$ - одиниця її вимірювання.

Істинне значення фізичної величини $A_{іст}$, яке ідеально характеризувало б властивості об'єкту, як правило, невідоме. Експериментальним шляхом можна знайти **дійсне** значення фізичної величини A_d , яке настільки наближається до істинного, що може бути використане замість нього. Ступінь наближення дійсних значень до істинних залежить від досконалості технічних **засобів вимірювань**, які при цьому використовуються.

До **засобів вимірювання** відносяться **вимірювальні прилади, міри і вимірювальні системи та установки**, що з них складаються.

Вимірювальними приладами називаються засоби вимірювання, з допомогою яких можна безпосередньо вирахувати значення величин, що вимірюються (динамометр, амперметр, манометр та ін.).

Мірами називають засоби вимірювання, що відтворюють фізичні величини заданих розмірів (гирі, лінійки, нормальні елементи, зразкові опори, котушки індуктивності і т.д.)).

За способом знаходження числового значення фізичної величини вимірювання ділять на **прямі, непрямі, сукупні та спільні** [3].

Під час **прямих** вимірювань шукане значення величини знаходять безпосередньо з досліду – прямим порівнянням вимірюваної величини з **мірами** (наприклад, вимірювання довжини лінійкою, вимірювання маси на рівноплечих терезах), або з допомогою вимірювального приладу, проградуйованого в одиницях вимірювання (наприклад, вимірювання температури термометром, вимірювання сили струму амперметром).

При *непрямих* вимірюваннях шукане значення величини A знаходять з допомогою обчислень, використовуючи результати прямих вимірювань інших фізичних величин, з якими величина A зв'язана відомою функціональною залежністю (формулою):

$$A = f(x_1, x_2, \dots, x_n) \quad (1.2)$$

Наприклад, густину тіла можна обчислити, використавши результати прямих вимірювань його маси і об'єму $\left(\rho = \frac{m}{V}\right)$, опір провідника можна знайти з закону Ома, якщо відомі результати прямих вимірювань сили струму в провіднику і напруги на його кінцях $\left(R = \frac{U}{I}\right)$.

В залежності від вибору методу вимірювань значення деяких фізичних величин можна визначити шляхом як прямих, так і непрямих вимірювань. Наприклад, густину рідини можна виміряти ареометром, а опір провідника визначити омметром.

Сукупні та *спільні* вимірювання дають можливість знайти значення величин x_1, x_2, \dots, x_n , які не можна спостерігати безпосередньо, за результатами вимірювань інших величин y_1, y_2, \dots, y_m , які є їх функціями:

$$y_j = \varphi_j(x_1, x_2, \dots, x_n), j = 1, 2, \dots, m \quad (1.3)$$

Після прямих вимірювань значень величин y_j результати цих вимірювань підставляють в систему (1.3), розв'язок якої дозволить знайти значення величин x_1, x_2, \dots, x_n . Очевидно, що для розв'язання системи (1.3) необхідно виконання умови $m \geq n$.

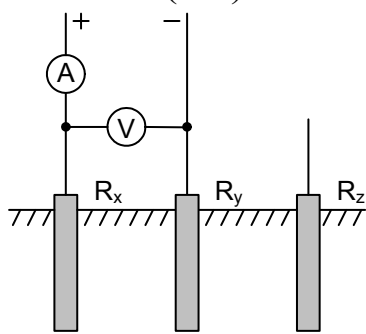


Рис. 1.1

При *сукупних* вимірюваннях безпосередньо вимірюють значення різних комбінацій кількох однойменних величин, кожен з яких зокрема виміряти неможливо. Наприклад, при вимірюванні опору заземлення (рис. 1.1) попарно вимірюють опори трьох заземлень – одного основного (робочого) R_x і двох допоміжних R_y і R_z :

$$R_1 = R_x + R_y; \quad R_2 = R_x + R_z; \quad R_3 = R_y + R_z.$$

Розв'язавши цю систему рівнянь, знаходять опори R_x , R_y і R_z .

При **спільних** вимірюваннях знаходять залежність між декількома неоднорідними величинами. Наприклад, треба знайти залежність опору провідника від температури, яка наближено описується формулою:

$$R = R_0(1 + \alpha t^0), \quad (1.4)$$

де R_0 - опір при температурі 0^0C , α - температурний коефіцієнт опору. Щоб знайти R_0 і α , слід знайти два значення опору R_1 і R_2 при двох значеннях температури t_1^0 і t_2^0 , а потім розв'язати систему рівнянь:

$$\begin{cases} R = R_0(1 + \alpha t_1^0) \\ R = R_0(1 + \alpha t_2^0) \end{cases} \quad (1.5)$$

За результатами спільних вимірювань в багатьох випадках будують графіки залежностей між величинами.

1.2. Похибки вимірювань

Недосконалість вимірювальних приладів і наших органів відчуттів, а також дія багатьох факторів, що спотворюють вимірювання, приводять до того, що результат кожного вимірювання фізичної величини A_0 , не співпадає з її істинним значенням $A_{іcm}$. Різницю між результатом вимірювання та істинним значенням шуканої величини називають **абсолютною похибкою вимірювання**:

$$\Delta A = |A_{іcm} - A_0| \quad (1.6)$$

Оскільки $A_{іcm}$ в більшості випадків ми не знаємо, то точно визначити похибку вимірювань не можна. Можна лише наближено вказати інтервал можливих значень шуканої величини, всередині якого розташоване її істинне значення:

$$A_0 - \Delta A < A_{іcm} < A_0 + \Delta A \quad (1.7)$$

або:

$$A_{іcm} = A_0 \pm \Delta A \quad (1.8)$$

Практично **абсолютною похибкою вимірювань** називають саме величину ΔA в рівняннях (1.7) чи (1.8).

Якість результатів вимірювання зручніше характеризувати не абсолютною похибкою ΔA , а її відношенням до дійсного значення шуканої величини A_0 ; це відношення називають **відносною похибкою** і звичайно виражають в процентах:

$$E = \frac{\Delta A}{A_0} \cdot 100\% \quad (1.9)$$

Відносну похибку також називають **точністю**.

Похибки вимірювань ділять на три типи: **систематичні**, **випадкові** та **промахи**.

Систематичними похибками вимірювань називають похибки, які при багатократному вимірюванні однієї і тієї ж величини залишаються сталими, або змінюються за певним законом. Систематичні похибки вимірювань включають в себе **методичні** та **інструментальні похибки**.

Методичні похибки викликані недоліками методу вимірювань, недосконалістю теорії фізичного явища чи неточністю розрахункової

формули. Для зменшення цих похибок слід вдосконалювати методи вимірювань.

Інструментальні похибки, або **похибки приладу** викликані недосконалістю конструкції та неточністю виготовлення вимірювальних приладів (наприклад, невелика різниця в довжинах плеч важільних терезів, невідповідність центра шкали стрілочного приладу вісі обертання стрілки, та ін.). Зменшують інструментальні похибки, використовуючи більш досконалі і точні прилади. Однак повністю усунути інструментальну похибку неможливо.

Випадковими похибками називають похибки, абсолютне значення і знак яких змінюються при багатократних вимірюваннях однієї і тієї ж величини. Випадкові похибки викликаються багатьма факторами, що не піддаються обліку. Повністю позбавитися від цих похибок теж неможливо, але їх можна зменшити шляхом багатократного повторення вимірювань. При цьому має місце часткова компенсація випадкових відхилень результатів вимірювань в сторону збільшення і в сторону зменшення. Розрахунок випадкових похибок здійснюють методами теорії ймовірностей і математичної статистики.

Промах – це похибка, що суттєво перевищує ту, яку ми чекаємо при даних умовах вимірювання. Промах може бути зроблений внаслідок невірного запису показів приладу, помилки експериментатора при використанні приладу, несправності апаратури для вимірювання. Промахи виявляють, повторюючи вимірювання в інших умовах, або аналізуючи результати (якщо тільки одне вимірювання значно відрізняється від інших, то, цілком можливо, це вимірювання є промахом). Якщо промах виявлено, то його слід виключити з результатів, і, в разі необхідності, повторити вимірювання.

РОЗДІЛ 2

ОБЧИСЛЕННЯ ПОХИБОК

2.1. Обчислення похибок при прямих вимірюваннях.

Припустимо спочатку, що систематичні похибки відсутні і всі похибки можна вважати випадковими. В цьому випадку під час повторних вимірювань однієї і тієї ж величини ми одержимо результати, що трохи відрізняються один від одного:

$$A_1, A_2, A_3, \dots, A_i, \dots, A_n.$$

Якщо під час повторних вимірювань будуть одержані зовсім однакові числові значення, то це вкаже на недостатню чутливість вимірювального приладу (тобто значну похибку приладу). Результати, що значно відрізняються один від одного, також свідчать про неточність вимірювань і необхідність вдосконалення методики вимірювань [3].

Під час обробки одержаних результатів виникають два питання:

1) **Як одержати з цих n значень найкращу оцінку істинного значення шуканої величини?**

2) **Чому дорівнює похибка вимірювань?**

Відповіді на ці питання дає теорія імовірності (див., наприклад, [1,2]). Наведемо їх без доведення, обмежившись лише якісним поясненням.

Найкращою оцінкою значення величини, що вимірюється, є **середнє арифметичне** результатів вимірювань:

$$A = \bar{A} = \frac{\sum_{i=1}^n A_i}{n} \quad (2.1)$$

Для оцінки похибки вимірювань знайдемо відхилення кожного з вимірювань від середнього значення: $\Delta A_i = A_i - \bar{A}$. Серед ΔA_i будуть зустрічатись як додатні, так і від'ємні числа.

Стандартною або **середньою квадратичною похибкою окремого вимірювання** називають величину:

$$S = \sqrt{\frac{\sum_{i=1}^n \Delta A_i^2}{n-1}} \quad (2.2)$$

Знайдене середнє значення \bar{A} є кращою оцінкою A , ніж будь-яке з A_i . **Середня квадратична похибка середнього**

$$S_{\bar{A}} = \frac{S}{\sqrt{n}} = \sqrt{\frac{\sum_{i=1}^n \Delta A_i^2}{n(n-1)}} \quad (2.3)$$

Саме цю похибку ми й будемо обчислювати під час обробки результатів вимірювань.

Вирази (2.2) і (2.3) справедливі при $n > 30$. При малому числі дослідів n для одержання тієї ж довірчої імовірності (див. далі) слід збільшити довірчий інтервал, помноживши $S_{\bar{A}}$ на деякий коефіцієнт t_{α} , який називають **коефіцієнтом Стюдента** [3]. Значення цих коефіцієнтів наведено в додатку (табл. 3). Результат вимірювання записують у виді

$$A = \bar{A} \pm \Delta A, \quad (2.4)$$

де;

$$\Delta A = \Delta A_{\text{вип}} = t_{\alpha} S_{\bar{A}}.$$

Абсолютна похибка ΔA округляється до двох значущих цифр, якщо перша з них одиниця чи двійка, і до однієї значущої цифри у всіх інших випадках. В кінцевому результаті вимірювання \bar{A} останньою значущою цифрою повинна бути цифра того ж розряду, що і в абсолютній похибці. В проміжних результатах обчислень пишуть ще одну лишню цифру, що дає можливість точніше округлити кінцевий результат.

При великій кількості вимірювань постає питання, як записати результати вимірювання. Для упорядкування результатів їх можна розташувати в порядку збільшення числового значення; якщо якийсь із результатів зустрічається не один раз, то зручно вказати число появ цього результату.

Наприклад, в результаті вимірювань одержано наступні значення: 23, 24, 26, 28, 23, 24, 25, 24, 26, 25. Їх можна упорядкувати так: 23, 24, 24, 24, 25, 25, 26, 26, 26, 28.

Значення A_i	23	24	25	26	28
Число реалізацій n_i	1	3	2	3	1

Розрахунок середнього тепер стає зручнішим:

$$\bar{A} = \frac{\sum_i A_i n_i}{N} \quad (2.5)$$

Коефіцієнти n_i в (2.5) називають **ваговими коефіцієнтами** (множниками), а сума в (2.5) називається **зваженою сумою**.

Є ще один спосіб, який часто буває ще більш зручним. Замість того, щоб говорити, що результат $x = 24$ було одержано тричі, можна сказати, що результат $x = 24$ було одержано в $\frac{3}{10}$ випадків від повного числа вимірювань. Замість n_i вводиться відношення

$$F_i = \frac{n_i}{N}, \quad (2.6)$$

яке дає долю повного числа вимірювань N , з якою одержується результат x_i ; це відношення називають **частотою**. Частоти F_i характеризують **розподіл** результатів. Формула для середнього з використанням частоти стає ще простішою:

$$\bar{A} = \sum_i A_i F_i \quad (2.7)$$

В (2.7) припускається, що:

$$\sum_i F_i = 1 \quad (2.8)$$

Умова (2.8) називається **умовою нормування**.

Графічно розподіл результатів вимірювань можна представити на **гістограмі**. По вертикалі відкладають частоту F_i , а по горизонталі – значення x_i (рис. 2.1). Величина, зображена на гістограмі рис. 2.1, називається **дискретною**; результати вимірювань такої величини представлені цілими числами (наприклад оцінки студента на

екзаменах, число імпульсів за хвилину при вимірюванні радіаційного фону і т.д.).

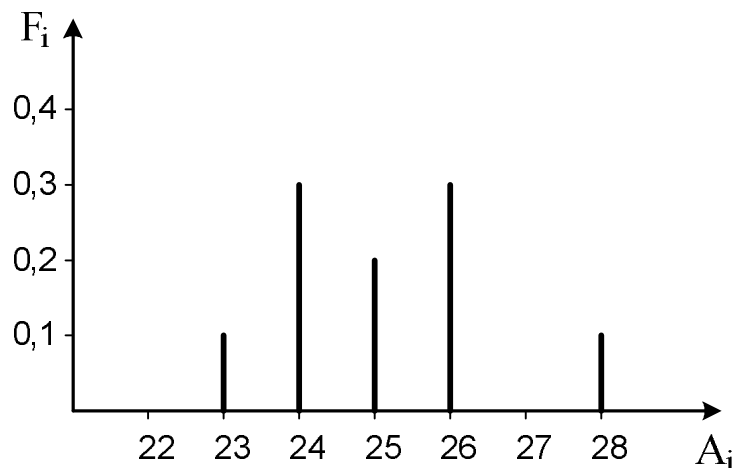


Рис. 2.1

Для побудови гістограми неперервної величини слід розбити діапазон можливих значень цієї величини на певне число інтервалів (або бінів) і підрахувати, скільки значень попадає в кожен бін. Вибір ширини інтервалу визначається числом вимірювань: чим більше число вимірювань, тим більш вузькими повинні бути біни. Зі збільшенням числа вимірювання гістограма наближається до деякого **граничного розподілу**, який зображається неперервною кривою. Ця крива визначає функцію, яку називають **функцією розподілу** $f(x)$ (рис. 2.2). Доля вимірювань, що попадає в інтервал між двома будь-якими значеннями a і b , дорівнює площі під кривою між $x = a$ і $x = b$. Дана площа – це визначений інтеграл $\int_a^b f(x)dx$. Отже, при дуже великій кількості вимірювань вказаний інтеграл визначає імовірність α того, що результат будь-якого одиничного вимірювання попаде в інтервал від $x = a$ до $x = b$:

$$\alpha = \int_a^b f(x)dx$$

Зрозуміло, що $\int_{-\infty}^{\infty} f(x)dx = 1$, тобто функція $f(x)$ **нормована**.

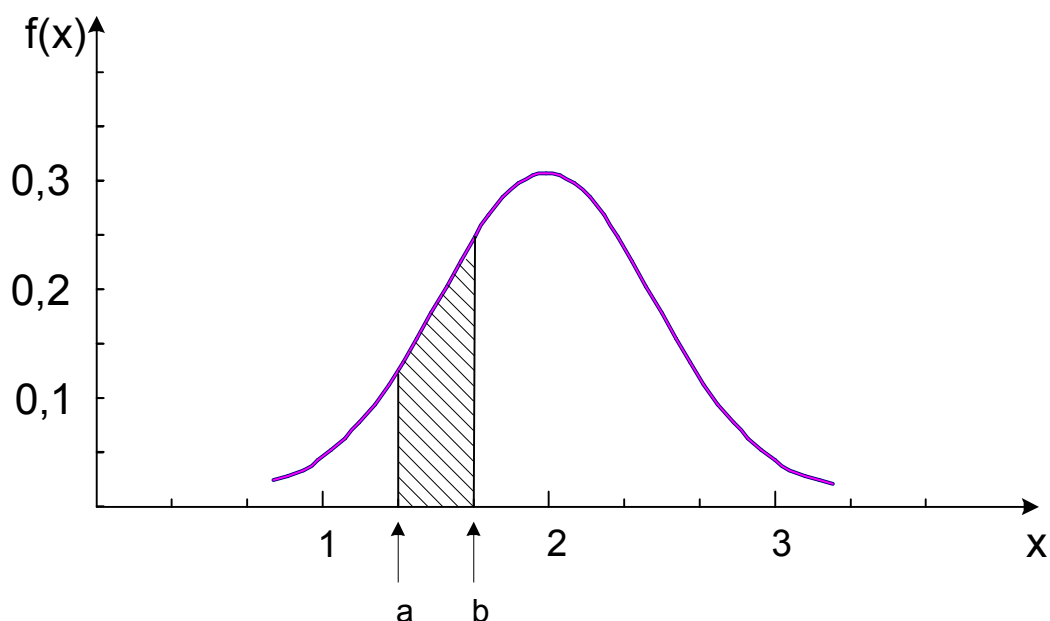


Рис.2.2

Різні види вимірювань мають різні граничні розподіли і не всі граничні розподіли мають вид симетричної кривої, зображеної на рис. 2.2. Проте дуже велика кількість вимірювань мають граничний розподіл у виді кривої рис. 2.2. Зокрема, це матиме місце при наявності в процесі вимірювання лише невеликих випадкових похибок. В математиці функція, графік якої має форму симетричної кривої, зображеної на рис. 2.2, називається **функцією нормального розподілу** або **функцією Гауса**. Нормована функція Гауса має вид:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-A)^2}{2\sigma^2}} \quad (2.9)$$

Параметр σ функції Гауса – це стандартна або середньоквадратична похибка окремого вимірювання (2.2), яку ми одержали б у випадку великої кількості вимірювань. Параметр σ називають також **стандартним відхиленням**. Малі значення σ приводить до розподілу типу гострого піка, що відповідає точним вимірюванням, а великі значення σ дають широкий розподіл, що відповідає вимірюванням з малою точністю (рис. 2.3).

Квадрат стандартного відхилення σ^2 називають **дисперсією**.

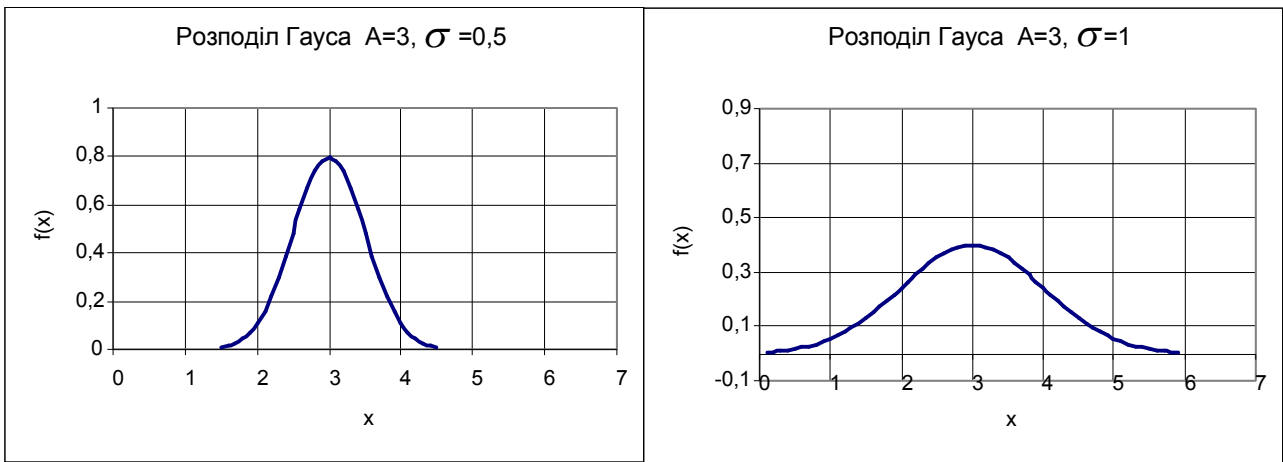


Рис. 2.3

Інтервал значень величини A , рівний $(\bar{A} - \Delta A, \bar{A} + \Delta A)$, називають **довірчим інтервалом**, а імовірність того, що істинне значення величини, яка вимірюється, попаде в довірчий інтервал, називають **довірчою імовірністю α** (або **надійністю результату**). Для довірчого інтервалу $(\bar{A} - \sigma, \bar{A} + \sigma)$ надійність результату $\alpha = 0,68$:

$$\alpha(\text{в межах } \pm \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \int_{A-\sigma}^{A+\sigma} e^{-\frac{(x-A)^2}{2\sigma^2}} dx = 0,68;$$

$$\text{для } (\bar{A} - 2\sigma, \bar{A} + 2\sigma) - \alpha = 0,95;$$

$$\text{для } (\bar{A} - 3\sigma, \bar{A} + 3\sigma) - \alpha = 0,997.$$

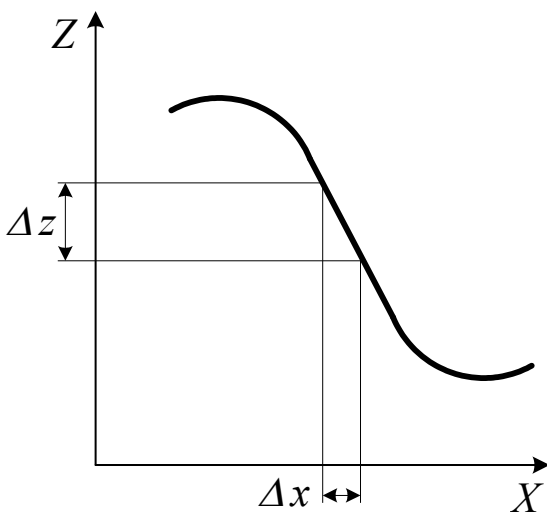


Рис. 2.4

Часто доводиться визначати яку-небудь величину z , графічно зв'язану з вимірюваною величиною x (рис. 2.4). Якщо величина x вимірюється з абсолютною похибкою Δx , то тим самим вноситься похибка Δz у значення величини z . Метод визначення похибки Δz зрозумілий з рисунка.

Як видно з виразу (2.3), збільшення числа вимірювань приводить до зменшення випадкової похибки вимірювань. Проте похибка вимірювання залежить ще й від систематичних похибок, якими ми до цього часу нехтували. Скільки вимірювань ми не виконали б невірною зробленою лінійкою, точний результат ми не одержимо. Але ж будь-яка лінійка виготовлена не цілком точно!

Яку ж кількість вимірювань слід виконати під час проведення експерименту? **Ніколи не слід обмежуватись однократним вимірюванням.** Завжди необхідно зробити повторне контрольне вимірювання. Якщо результати вимірювань співпали, то на цьому, як правило, можна зупинитись. Якщо ж між результатами є відмінності, то вимірювання слід виконати ще 2...3 рази, щоб зрозуміти, в чому причина: в тому, що одне з вимірювань було виконано невірне, чи в тому, що результати вимірювань відрізняються через випадкові похибки. В першому випадку слід просто відкинути невірне вимірювання, а в другому спробувати розібратись в причині відмінності результатів. Якщо ця причина може бути ліквідована шляхом регулювання приладу (змащення частин, в яких виникає тертя, усунення люфту і т. ін.), то це обов'язково слід зробити. Якщо ж усунути причину розходження результатів не вдається, то слід здійснити цілу серію повторних вимірювань, щоб зробити випадкову похибку достатньо малою (меншою систематичної чи меншою, ніж допустима похибка при необхідній в даній роботі точності вимірювань).

Прилади найчастіше нормуються по приведеній похибці, вираженій в процентах від верхньої межі вимірювань. Ця похибка називається **класом точності** і має значення: 0,1; 0,2; 0,5; 1,0; 1,5; 2,5; 4,0. Похибка приладу:

$$\Delta A_{\text{прил}} = \frac{k}{100} A_{\text{max}}, \quad (2.10)$$

де k – клас точності приладу, A_{max} – верхня межа його вимірювання. Оскільки похибка приладу визначає максимально можливе значення похибки, а похибки вимірювань характеризують середньоквадратичною похибкою, то, щоб одержати середньоквадратичну похибку приладу слід $\Delta A_{\text{прил}}$ поділити на 2:

$$\Delta A_{\text{сист}} = \frac{1}{2} \Delta A_{\text{прил}} \quad (2.11)$$

Якщо клас точності приладу не вказано, то вважають, що **систематична похибка дорівнює половині ціни найменшої поділки шкали приладу.**

В загальному випадку при наявності як випадкових, так і систематичних похибок, похибка вимірювання величини A

$$\Delta A = \sqrt{(\Delta A_{\text{вип}})^2 + (\Delta A_{\text{сист}})^2} \quad (2.12)$$

і визначається з точністю біля 20 %. Тому, якщо випадкова і систематична похибки відрізняються більш, ніж у два рази, то можна вважати, що ΔA дорівнює більшій з них. Нехай, наприклад, $\Delta A_{\text{вип}} = 0,5 \Delta A_{\text{сист}}$; тоді:

$$\Delta A = \sqrt{(0,5 \Delta A_{\text{сист}})^2 + (\Delta A_{\text{сист}})^2} = \sqrt{1,25 (\Delta A_{\text{сист}})^2} = 1,12 \Delta A_{\text{сист}} \approx \Delta A_{\text{сист}}$$

Іноді результат одного вимірювання (чи однієї серії вимірювань) значно відрізняється від інших. В цьому випадку експериментатор повинен вирішити, чи є цей аномальний результат вимірювання наслідком грубої помилки (тобто промахом) і з цієї причини його слід відкинути, чи цей результат повинен враховуватись разом з іншими. Відповідь на це питання дати не так просто; проблеми відкидання даних викладені, наприклад, в [1,2], ми ж обмежимося наступною рекомендацією. Обчисліть стандартне відхилення $S_{\bar{A}}$ (2.3) з врахуванням всіх значень величини, що вимірюється, потім сумнівний результат A_k порівняйте зі значенням $3S_{\bar{A}}$; якщо $A_k > \bar{A} + 3S_{\bar{A}}$, чи $A_k < \bar{A} - 3S_{\bar{A}}$, то цей результат слід відкинути. Після цього необхідно виконати перерахунок похибки $3S_{\bar{A}}$. Відкидання даних виконується тільки один раз.

2.2. Обчислення похибок при непрямих вимірюваннях.

Нехай для визначення величини A використовується відома функціональна залежність її від інших величин x_1, x_2, \dots, x_n :

$$A = f(x_1, x_2, \dots, x_n).$$

Величини x_1, x_2, \dots, x_n можуть бути трьох типів:

1). Величини, що визначаються шляхом прямих вимірювань; їх значення подаються в стандартній формі (2.4):

$$x_1 = \bar{x}_1 \pm \Delta x_1, \quad \bar{x}_2 = \bar{x}_2 \pm \Delta x_2, \quad \text{і т. д.}$$

2). Дані установки, відомі з попередніх (тарувальних) вимірювань; ці величини також повинні бути задані в формі (2.4), в іншому випадку вважають, що задане без вказівки похибки значення виміряне з точністю до половини одиниці останнього десяткового розряду.

3). Табличні величини і константи – їх бажано брати з таблиць з такою кількістю значущих цифр, щоб похибка округлення була на порядок менша похибки вимірювання величин першого і другого типу; в цьому випадку похибкою табличних величин і констант можна знехтувати.

Найкращим значенням величини A при непрямих вимірюваннях буде

$$\bar{A} = f(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n), \quad (2.13)$$

де \bar{x}_i – середні значення величин, одержані з прямих вимірювань. Межу довірчого інтервалу похибки непрямого вимірювання знаходять за формулою:

$$\Delta A = \sqrt{\sum_{i=1}^n \left(\frac{\partial f}{\partial x_i} \Delta x_i \right)^2}, \quad (2.14)$$

де Δx_i відповідають одному й тому ж значенню довірчої імовірності α . Довірча імовірність результату буде не менше α .

Середнє арифметичне значення величини \bar{A} і похибку результату вимірювання можна визначити двома способами. Для простоти міркувань припустимо, що A – функція однієї змінної x :

$$A = f(x) \quad (2.15)$$

В результаті вимірювань одержано ряд значень x_i . Знайдемо $\bar{x} = \frac{\sum x_i}{n}$, підставимо це значення в (2.15) і одержимо $\bar{A} = f(\bar{x})$. Але можна для кожного зі значень x_i обчислити $A_i = f(x_i)$, а потім визначити \bar{A} з співвідношення $\bar{A} = \frac{\sum A_i}{n}$. Відповідно двома способами можна знайти і похибку величини A : або визначити похибку величини x і скористатись співвідношенням $\Delta A = \frac{df}{dx} \Delta x$, або

використати ряд значень A_i : $\Delta A = \sqrt{\frac{\sum (\bar{A} - A_i)^2}{n(n-1)}}$. Якщо похибки

вимірювань малі порівняно з величиною, що вимірюється, то обидва способи дають досить близькі результати, тому неважливо, яким з них скористатися. Однак перший спосіб (тобто формули (2.13) і (2.14)) з точки зору розрахунків менш трудомісткий, тому йому й надамо перевагу.

Відносну похибку величини A обчислимо, скориставшись (2.14):

$$(\Delta A)^2 = \sum_{i=1}^n \left(\frac{\partial f}{\partial x_i} \Delta x_i \right)^2; \quad \left(\frac{\Delta A}{A} \right)^2 = \sum_{i=1}^n \left(\frac{1}{f} \frac{\partial f}{\partial x_i} \Delta x_i \right)^2;$$

оскільки: $\frac{1}{f} \frac{\partial f}{\partial x_i} = \frac{\partial \ln f}{\partial x_i}$, то :

$$E = \frac{\Delta A}{A} = \sqrt{\sum_{i=1}^n \left(\frac{\partial \ln f}{\partial x_i} \Delta x_i \right)^2} \quad (2.16)$$

Оскільки в більшості випадків (2.16) значно простіше, ніж (2.14), то звичайно спочатку за формулою (2.16) обчислюють відносну похибку E , а потім – межу довірчого інтервалу ΔA :

$$\Delta A = E \cdot \bar{A} \quad (2.17)$$

Для деяких найпростіших і в той же час найбільш поширених випадків формули розрахунку похибок наведено в додатку (таблиці 1 і 2).

Щоб швидко оцінити похибку методу (максимальну похибку), можна скористатись виразом:

$$E = \left| \partial \ln f(x_1, x_2, \dots, x_n) \right|. \quad (2.18)$$

Після приведення подібних всі доданки в (2.18) додають по модулю (найбільш несприятливий випадок). В якості ∂x_i беруть похибки приладів $\Delta A_{\text{прил}}$.

На закінчення перелічимо основні кроки, які слід виконувати при обробці результатів вимірювань.

Для прямих вимірювань.

1. Результати кожного вимірювання записати в таблицю.
2. Обчислити середнє арифметичне з n вимірювань:

$$\bar{A} = \frac{1}{n} \sum_{i=1}^n A_i.$$

3. Обчислити середньоквадратичну похибку:

$$S_A = \sqrt{\frac{1}{n(n-1)} \sum_{i=1}^n (A_i - \bar{A})^2}.$$

4. Якщо серед результатів вимірювань є один (або два), що значно відрізняється від інших, то слід перевірити, чи не є це вимірювання промахом.

5. Задати значення довірчої імовірності α і за табл. 3 визначити коефіцієнт Стюдента.

6. Обчислити випадкову похибку вимірювань $\Delta A_{\text{вип}} = t_\alpha \cdot S_A$ і порівняти її з систематичною $\Delta A_{\text{сист}} = \frac{1}{2} \Delta A_{\text{прил}}$.

Якщо ці похибки сумірні за значенням, то похибка вимірювання величини A дорівнює:

$$\Delta A = \sqrt{t_\alpha^2 S_A^2 + \left(\frac{t_\infty}{2}\right)^2 \Delta A_{\text{сист}}^2},$$

де t_∞ – коефіцієнт Стюдента для $n = \infty$

7. Записати результат вимірювань у вигляді:

$$A = \bar{A} \pm \Delta A \quad (\alpha = 0,95)$$

8. Обчислити відносну похибку вимірювань:

$$E = \frac{\Delta A}{A} 100\%$$

Для непрямих вимірювань.

1. Для кожної серії вимірювань величин, що визначають шукану величину, провести обробку, як описано вище. Довірча імовірність α задається однаковою для всіх величин.
2. Обчислити середнє значення шуканої величини (див.(2.13)):

$$\bar{A} = f(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n).$$

3. Користуючись формулами (2.14) і (2.16) (чи таблицями 1, 2 в додатку), знайти вирази для абсолютної та відносної похибки вимірювань:

$$\Delta A = \sqrt{\sum_{i=1}^n \left(\frac{\partial f}{\partial x_i} \Delta x_i \right)^2}, \quad E = \frac{\Delta A}{A} = \sqrt{\sum_{i=1}^n \left(\frac{\partial \ln f}{\partial x_i} \Delta x_i \right)^2}$$

Значення похідних $\frac{\partial f}{\partial x_i}$ обчислюються при $x_i = \bar{x}_i$.

4. Записати кінцевий результат у виді:

$$A = \bar{A} \pm \Delta A (\alpha = 0,95).$$

В шкільній лабораторії при виконанні лабораторних робіт для спрощення обробки результатів слід рекомендувати спочатку обчислювати відносну похибку по спрощеній формулі (2.18)

$$E = |\partial \ln f(x_1, x_2, \dots, x_n)|. \quad (2.18)$$

а потім обчислювати абсолютну похибку $\Delta A = E \cdot \bar{A}$.

Наведемо як приклад порядок розрахунку питомого опору провідника:

$$\rho = \frac{R \cdot S}{l}.$$

R вимірюють з допомогою амперметра і вольтметра: $R = \frac{U}{I}$; площа

поперечного перерізу $S = \frac{\pi d^2}{4}$, де d вимірюють мікрометром; l вимірюють лінійкою.

Розрахункова формула: $\rho = \frac{\pi d^2 U}{4l}$. Прологарифмуємо її та знайдемо диференціал: $\ln \rho = \ln \pi + 2 \ln d + \ln U - \ln 4 - \ln I - \ln l$;

$$\frac{d\rho}{\rho} = \frac{2dd}{d} + \frac{dU}{U} - \frac{dI}{I} - \frac{dl}{l}$$

Змінюємо знаки диференціалу d на знаки похибки Δ і додаємо по модулю:

$$E = \frac{\Delta\rho}{\rho} = \frac{2\Delta d}{d} + \frac{\Delta U}{U} + \frac{\Delta I}{I} + \frac{\Delta l}{l} \quad (2.19)$$

Можна показати [2], що відносна похибка, одержана за формулою (2.19), буде не меншою, ніж одержана при квадратичному додаванні (формула (2.16)). Користуючись наведеним прикладом, для кожної конкретної лабораторної роботи можна одержати спрощені формули для розрахунку похибки вимірювань.

РОЗДІЛ 3

ЗНАХОДЖЕННЯ КРИВИХ, ЩО ЗОБРАЖУЮТЬ ЗВ'ЯЗОК МІЖ ДОСЛІДЖУВАНИМИ ВЕЛИЧИНАМИ

3.1. Графічне представлення даних експерименту

В багатьох випадках досліджувана величина змінюється при зміні умов досліду і задача вимірювань полягає в знаходженні функціональної залежності, що найкращим чином описує закон зміни досліджуваної величини. Прикладами можуть бути: залежність часу падіння тіла від висоти (чи висоти від часу; теоретична залежність – $h = \frac{g}{2}t^2$); залежність видовження дроту від прикладеного навантаження ($F = k\Delta l$); опору металу від температури ($R = R_0(1 + \alpha t^o)$) і т. д.

Нехай величина y залежить тільки від величини x . В результаті вимірювань для кожного значення x_i одержані відповідні значення \bar{y}_i , для яких визначені довірчі інтервали Δy_i . Якщо значення x_i не можуть бути задані точно, а також вимірюються з деякою похибкою, то і для них відомі середні значення \bar{x}_i і відповідні довірчі інтервали Δx_i . Всі значення \bar{y}_i і \bar{x}_i можна розглядати, як координати точок на площині, що дає можливість представити результати графічно (рис. 3.1).

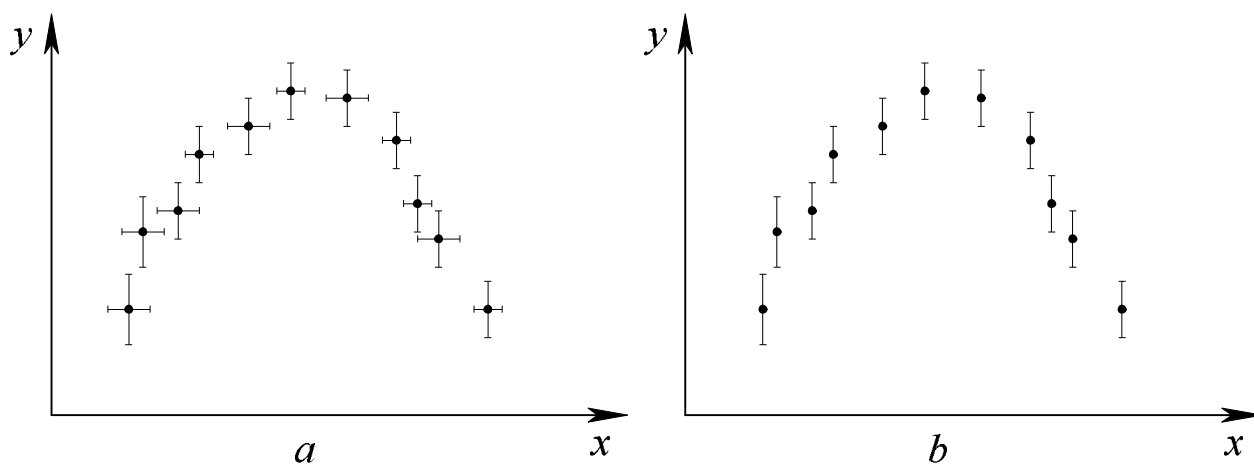


Рис. 3.1

Відрізки, зображені біля кожної точки, означають значення довірчого інтервалу (тобто, похибки вимірювання відповідної величини). Рисунок 3.1а відповідає випадку, коли обидві змінні величини мають похибки; рис. 3.1b – випадку, коли похибками Δx_i можна нехтувати ($\Delta x_i \approx 0$). Звичайно відкладають значення довірчого інтервалу для довірчої імовірності 0,68 або 0,95, яка повинна бути однаковою для всіх точок графіку і яку треба вказати. Значення довірчого інтервалу в різних точках можуть бути різними. Діаграми рис. 3.1 дають можливість встановити з певною імовірністю функціональну залежність, якою зв'язані величини x і y . Як правило, через експериментальні точки, ґрунтуючись на певних фізичних законах, проводять плавну криву, яка розумно описує одержані результати (наприклад так, як показано на рис. 3.2).

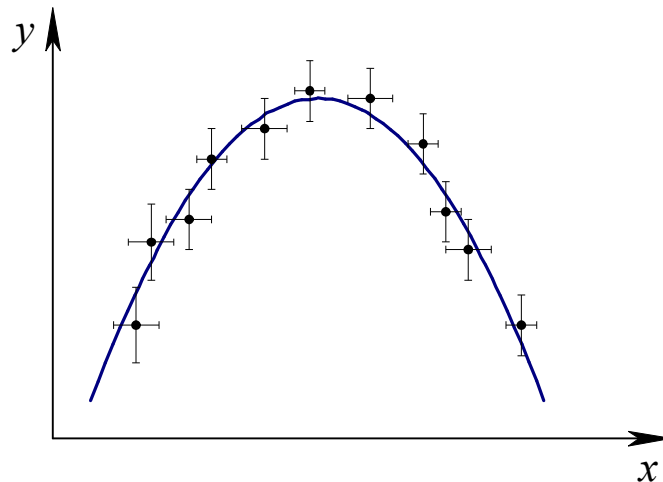


Рис. 3.2

Велика кількість фізичних величин зв'язані **лінійною залежністю** ($F = k\Delta l$, $R = R_0(1 + \alpha t^o)$, $v = v_0 + gt$ і т. д.), або можуть бути приведені до неї ($h = \frac{g}{2}t^2$, h лінійно залежить від t^2 ; $N = N_0e^{-\alpha t}$, якщо прологарифмувати цей вираз, то залежність лінеаризується: $\ln N = \ln N_0 - \alpha t$, $\ln N$ лінійно залежить від часу t).

Задача знаходження прямої, що найліпшим чином описує дану лінійну залежність може бути розв'язана з допомогою **лінійної регресії** або **апроксимації методом найменших квадратів**.

3.2. Апроксимація методом найменших квадратів

Для спрощення будемо вважати, що похибками у визначенні величини x можна знехтувати, а похибки у визначенні величини y однакові за числовими значеннями. Спочатку розглянемо залежність виду $y = kx$. Пряма в цьому випадку проходить через початок координат.

На рис. 3.3 показані експериментальні точки і пряма, яка, як ми вважаємо, найкраще описує цю залежність. Оскільки існують похибки вимірювань, то між експериментальними значеннями функції y_i та її теоретичними значеннями kx_i існують різниці δ_i .

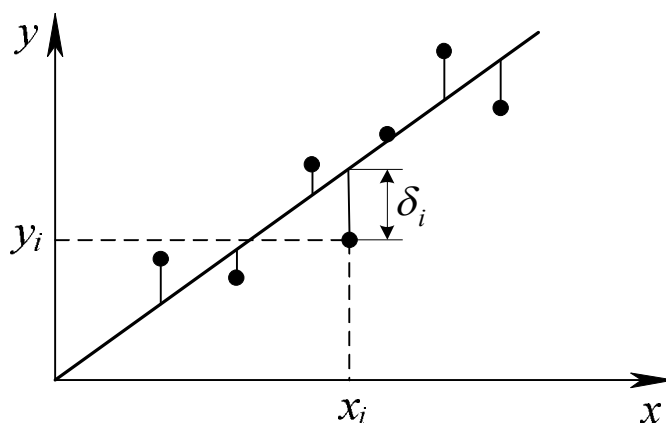


Рис. 3.3

В методі найменших квадратів припускається, що пряма буде найкращою, якщо сума квадратів різниць δ_i буде мати мінімальне значення. Мінімум цієї суми знайдемо, якщо прирівняємо похідну від неї до нуля:

$$\sum_{i=1}^n \delta_i^2 = \sum_{i=1}^n (kx_i - y_i)^2; \quad \frac{d}{dk} = 2 \sum (kx_i^2 - x_i y_i); \quad (kx_i^2 - x_i y_i) = 0;$$
$$k = \frac{\sum x_i y_i}{\sum x_i^2}. \quad (3.1)$$

Середньоквадратична похибка у визначенні k дорівнюватиме:

$$S_k = \sqrt{\frac{1}{n(n-2)} \left(\frac{\sum y_i^2}{\sum x_i^2} - k^2 \right)}. \quad (3.2)$$

В загальному випадку лінійна залежність описується рівнянням $y = ax + b$. Застосування методу найменших квадратів [2] дає наступні вирази для оцінки сталих a і b :

$$a = \frac{n(\sum x_i y_i) - (\sum x_i)(\sum y_i)}{\Delta}; \quad b = \frac{(\sum x_i^2)(\sum y_i) - (\sum x_i)(\sum x_i y_i)}{\Delta} \quad (3.3)$$

де $\Delta = n(\sum x_i^2) - (\sum x_i)^2$.

Середньоквадратична похибка у визначенні y_i дорівнює:

$$S_y = \sqrt{\frac{1}{n-2} \sum (y_i - ax_i - b)^2}, \quad (3.4)$$

де a і b даються виразом (3.3). Похибки сталих a і b дорівнюють:

$$S_a = S_y \sqrt{\frac{n}{\Delta}}; \quad S_b = S_y \sqrt{\frac{\sum x_i^2}{\Delta}} \quad (3.5)$$

Обчислення виразів (3.1)–(3.5) значно спрощується при використанні інженерних мікрокалькуляторів чи електронних таблиць Microsoft Office Excel.

Наведемо приклад застосування електронних таблиць для розв'язування задач на знаходження лінії регресії.

Нехай ми вивчаємо зв'язок між тиском газу і його температурою при сталому об'ємі. Для ідеального газу рівняння стану має вид:

$$pV = \nu RT \quad (3.6)$$

Будемо вважати, що похибка вимірювання температури значно більша похибки вимірювання тиску. Приведемо рівняння до виду $y = ax + b$:

$$T = \frac{V}{\nu R} p.$$

Дослідним шляхом вимірюється температура за шкалою Цельсія: $T = t^\circ + 273,15^\circ$. Після підстановки в попередній вираз:

$$t = \frac{V}{\nu R} p - 273,15 \quad (3.7)$$

Позначимо $a = \frac{V}{\nu R}$, $b = -273,15$. Стала b - це значення

температури, при якій тиск ідеального газу дорівнюватиме нулю; її називають *абсолютним нулем температури*.

При вимірюваннях температури і тиску одержали п'ять пар значень:

Табл. 3.1

№ n/n	$t^{\circ}C$	$p, \text{мм рт ст}$
1	-20	65
2	17	75
3	42	85
4	94	95
5	127	105

Визначимо за цими даними коефіцієнти a і b та їх похибки. Введемо в комірку A1 символ $t = y$, а в комірку B1 – $p = x$. В комірки A2 – A6 вводимо п'ять значень температури, а в комірки B2 – B6 п'ять значень тиску. В комірку A7 записуємо знак $=$. Наводимо курсор на знак f_x („Вставка функции”) і натискаємо ліву кнопку миші. З'являється діалогове вікно „Мастер функций”. Наводимо курсор на кнопку у вікні „Категория”, натискаємо ліву кнопку миші: випадає список функцій, в якому вибираємо (наводимо курсор) „Статистические” і клацаємо лівою кнопкою. У вікні „Выберите функцию” знаходимо „Линейн”, клацаємо лівою кнопкою, потім – „ОК”. На екрані з'явиться вікно „Аргументы функции”. У верхню строчку записуємо номери комірок, в яких містяться значення y : A2:A6; в наступну – x : B2:B6. Строчку „Конст.” можна пропустити, в строчку „Статистика” записуємо (російською): ИСТИНА. (Детальну інформацію по функції можна одержати, натиснувши Справка по этой функции). Натискаємо ОК. В комірці A7 з'являється число 3,71. Наводимо курсор \dagger на комірку A7, натискаємо на ліву кнопку і виділяємо чотири комірки: A7, A8, B7, B8. Натискаємо на клавішу F2 (в комірці A7 з'явиться набрана формула), потім, утримуючи натиснутими клавіші Ctrl+Shift, натискаємо Enter. У виділених комірках з'являються: A7 – значення сталої a ; A8 – значення

стандартного відхилення σ_a ; В7 – значення сталої b ; В8 – значення стандартного відхилення σ_b (див. табл. 3.2).

Таким чином, для абсолютного нуля температури одержане значення: $b = T_0 = -(263 \pm 18)^{\circ}C$, що задовільно узгоджується з відомим значенням $-273^{\circ}C$.

Табл. 3.2

	A	B
1	t=y	p=x
2	-20	65
3	17	75
4	42	85
5	94	95
6	127	105
7	3,71	-263,35
8	0,211266	18,20446

РОЗДІЛ 4

ДЕЯКІ ЗАСТОСУВАННЯ СТАТИСТИЧНИХ МЕТОДІВ ДО ПЕДАГОГІЧНИХ ДОСЛІДЖЕНЬ

4.1. Наочне представлення даних

Спостереження можуть охоплювати всіх членів сукупності, що вивчається, або обмежуватись дослідженням лише частини даної сукупності. Статистична сукупність, що досліджується і з якої роблять вибір, називається **генеральною сукупністю**. Відібрана з генеральної сукупності деяка частина одиниць, яку будуть безпосередньо досліджувати, називається **вибірковою сукупністю** або **вибіркою**.

В математичній статистиці замість слова "дані" вживають термін "**варіанти**", а числову характеристику варіанти називають "**ознакою**". Число об'єктів сукупності називають **об'ємом** сукупності.

Щоб вибірка найповніше відображувала структуру генеральної сукупності, потрібно проводити процес відбору випадковим чином, щоб забезпечити всім варіантам рівну можливість потрапити до складу вибірки. При дотриманні всіх вимог можна всі висновки, зроблені при дослідженні вибірки, перенести на генеральну сукупність. При вибірковому методі дослідження проводяться на невеликій частині сукупності – від 5 % до 15 %.

Покажемо на прикладах, як організувати дані вибіркової сукупності. Нехай в результаті проведеної контрольної роботи учні одержали такі оцінки:

Табл. 4.1

№ в списку	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Оцінка	6	4	4	10	3	7	7	7	5	2	8	8	7	8	4
№ в списку	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Оцінка	8	8	9	4	6	6	5	5	9	3	6	6	8	7	6

Таблицю 4.1 називають **рядом варіант** або **простим статистичним рядом**.

Першим кроком обробки даних є впорядкування варіант: запис даних у вигляді таблиці в зростаючому порядку. Впорядкування дає можливість швидко одержати додаткову інформацію про дані:

максимальне та мінімальне значення варіант, межі їхньої зміни. Таку таблицю називають *варіаційним рядом* (табл. 4.2):

Табл. 4.2

№ п/п	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Оцінка	2	3	3	4	4	4	4	5	5	5	6	6	6	6	6
№ п/п	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Оцінка	6	7	7	7	7	7	8	8	8	8	8	8	9	9	10

Об'єм таблиці значно зменшується, якщо ввести *частоту* варіанти n_i . Частота – це число, яке показує, скільки раз та чи інша варіанта зустрічається в таблиці даних. Ряд n_1, n_2, \dots, n_m називається *рядом частот*. Сума всіх частот повинна дорівнювати об'єму

вибірки n : $\sum_{i=1}^m n_i = n$. В таблиці 4.2 оцінка 2 зустрічається 1 раз, оцінка

3 – 2 рази, оцінка 4 – 4 рази і т.д. Зв'язок між рядом варіант і відповідними частотами називають *статистичним розподілом вибірки* або *статистичним рядом* (табл. 4.3):

Табл. 4.3

Оцінка	2	3	4	5	6	7	8	9	10
Частота	1	2	4	3	6	5	6	2	1

Статистичний ряд графічно подається *полігоном розподілу*. Щоб побудувати полігон розподілу, на вісі абсцис відкладають значення варіант, а на вісі ординат – відповідні їм частоти (рис. 4.1).

Нерідко замість частоти використовують *відносні частоти*: $W = \frac{n_i}{n}$.

Сума всіх відносних частот дорівнює одиниці: $\sum_{i=1}^m W_i = 1$.

Крім полігону розподілу, статистичний ряд можна подати у вигляді *гістограми* (рис. 4.2).

У випадку, коли досліджувана величина – неперервна і обсяг вибірки великий, результати вибірки подають *інтервальним рядом*. Для цього область реалізацій величини розбивають на k інтервалів і для кожного інтервалу визначають частоти. Кількість інтервалів $k \leq 5 \lg n$, а їхню довжину Δx_i найчастіше беруть однаковою. Здобутий ряд геометрично подається гістограмою. Для побудови її на вісі абсцис відкладають інтервали, а на них як на основах будують прямокутники, висота яких пропорційна частоті (відносній частоті) інтервалу. Гістограма дає певне уявлення про графік густини розподілу.



Рис. 4.1



Рис. 4.2

4.2. Встановлення кореляції між результатами сумісних вимірювань

Припустимо, що для знаходження функції $q(x, y)$ ми вимірюємо дві величини x та y декілька раз і одержуємо n пар даних $(x_1, y_1), \dots, (x_n, y_n)$. За результатами вимірювань обчислюємо \bar{x} , \bar{y} та дисперсії σ_x^2 і σ_y^2 :

$$\sigma_x^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}; \quad \sigma_y^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n-1} \quad (4.1)$$

Для n пар значень x та y можна знайти n значень величини q :

$$q_i = q(x_i, y_i), \text{ де } i = 1, 2, \dots, n.$$

Будемо вважати, що всі похибки малі, тобто всі числа x_1, \dots, x_n близькі до \bar{x} , а всі числа y_1, \dots, y_n – до \bar{y} . В цьому випадку справедливе наближення:

$$q_i = q(x_i, y_i) \approx q(\bar{x}, \bar{y}) + \frac{\partial q}{\partial x}(x_i - \bar{x}) + \frac{\partial q}{\partial y}(y_i - \bar{y}) \quad (4.2)$$

В рамках цього наближення середнє набуває виду:

$$\bar{q} = \frac{1}{n} \sum_{i=1}^n q_i = \frac{1}{n} \sum_{i=1}^n \left[q(\bar{x}, \bar{y}) + \frac{\partial q}{\partial x}(x_i - \bar{x}) + \frac{\partial q}{\partial y}(y_i - \bar{y}) \right],$$

з якого випливає, що середнє значення функції q дорівнює [2, с. 182]:

$$\bar{q} = q(\bar{x}, \bar{y}) \quad (4.3)$$

Дисперсія для n значень q_1, \dots, q_n визначається виразом:

$$\sigma_q^2 = \frac{\sum (q_i - \bar{q})^2}{n-1}$$

Підставивши в цей вираз (4.2) і (4.3), одержимо:

$$\sigma_q^2 = \frac{1}{n-1} \sum \left[\frac{\partial q}{\partial x}(x_i - \bar{x}) + \frac{\partial q}{\partial y}(y_i - \bar{y}) \right]^2 =$$

$$= \left(\frac{\partial q}{\partial x} \right)^2 \frac{\sum (x_i - \bar{x})^2}{n-1} + \left(\frac{\partial q}{\partial y} \right)^2 \frac{\sum (y_i - \bar{y})^2}{n-1} + 2 \frac{\partial q}{\partial x} \frac{\partial q}{\partial y} \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n-1}.$$

Суми в перших двох доданках – це дисперсії σ_x^2 і σ_y^2 (див. (4.1)); сума в третьому доданку – це так званий **змішаний другий момент** змінних x та y або **коваріація**:

$$\sigma_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n-1} \quad (4.4)$$

З врахуванням цих позначень вираз для дисперсії σ_q^2 матиме вид:

$$\sigma_q^2 = \left(\frac{\partial q}{\partial x} \right)^2 \sigma_x^2 + \left(\frac{\partial q}{\partial y} \right)^2 \sigma_y^2 + 2 \frac{\partial q}{\partial x} \frac{\partial q}{\partial y} \sigma_{xy} \quad (4.5)$$

Якщо значення x та y **незалежні**, то після багатьох вимірювань σ_{xy} буде наближатись до нуля і ми приходимо до відомого виразу для незалежних і випадкових похибок (див. (2.14)):

$$\sigma_q^2 = \left(\frac{\partial q}{\partial x} \right)^2 \sigma_x^2 + \left(\frac{\partial q}{\partial y} \right)^2 \sigma_y^2 \quad (4.6)$$

Якщо вимірювання x та y **не є незалежними**, то змішаний другий момент σ_{xy} не повинен дорівнювати нулю. Якщо σ_{xy} не дорівнює нулю, то говорять, що похибки в x та y **мають кореляцію**, і похибка σ_q повинна обчислюватись за виразом (4.5), а не (4.6). Проте можна показати, що і при наявності кореляції в похибках

$$\sigma_q \leq \left| \frac{\partial q}{\partial x} \right| \cdot \sigma_x + \left| \frac{\partial q}{\partial y} \right| \cdot \sigma_y, \quad (4.7)$$

тобто (4.7) є верхньою межею похибки в будь-якому випадку.

В попередньому розділі розглядалась лінійна регресія або апроксимація даних методом найменших квадратів Використання змішаного другого моменту дозволяє відповісти на питання, наскільки добре набір результатів вимірювань підтверджує гіпотезу про лінійну залежність між x та y .

Нехай ми одержали n пар значень $(x_1, y_1), \dots, (x_n, y_n)$ двох змінних, які, як ми вважаємо, повинні бути зв'язані лінійною залежністю виду $y = ax + b$. Підкреслимо, що x_i (як і y_i) в даному випадку не результати вимірювання однієї й тієї ж величини; це результати вимірювань n різних значень однієї змінної (наприклад, n різних висот, з яких кидали кульку, оцінки студентів за контрольну роботу і т. ін.). Ступінь підтвердження лінійної залежності між величинами x та y визначається **коефіцієнтом лінійної кореляції** (або просто **коефіцієнтом кореляції**):

$$r_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}, \quad (4.8)$$

де змішаний другий момент σ_{xy} і стандартні відхилення σ_x і σ_y визначаються формулами (4.1) і (4.4). Підставимо ці вирази в (4.8):

$$r_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} \quad (4.9)$$

Коефіцієнт кореляції може мати значення між $+1$ і -1 . Якщо r_{xy} близький до ± 1 , то точки лежать близько до деякої прямої; якщо r_{xy} близький до 0 , то точки не корельовані і не групуються поблизу прямої.

Кількісним критерієм значущості r_{xy} є імовірність того, що n вимірювань двох змінних x та y , кореляція між якими відсутня, дадуть значення r , не менше, ніж будь-яке задане значення r_0 : $P_n(|r| \geq |r_0|)$. Значення цих імовірностей наведено в табл. 4 **Значущою** вважається кореляція r_0 , якщо імовірність одержати коефіцієнт $|r| \geq |r_0|$ для некорельованих змінних менша 5% . (При $P_n < 1\%$ кореляція **високо значуща**).

Так, якщо для $n=10$ одержано $r_0 = 0,8$, то $P_n(|r| \geq 0,8) = 0,5\%$, тобто імовірність того, що між значеннями існує лінійна залежність, висока.

Наведемо приклад знаходження кореляції між оцінками за контрольну роботу з механіки і пропусками занять. Розв'яжемо цю задачу на прикладі потоку ФІ-06 (2007 р.) Використаємо електронні таблиці Microsoft Office Excel. (Застосування цих таблиць до розв'язування різних статистичних задач див., наприклад, в [8]).

Всього студентів 38. В комірки А1 – А38 заносимо число пропусків занять кожним зі студентів, в комірки В1 – В38 – їх оцінки за контрольну роботу з механіки (максимальна оцінка – 8 балів). В комірку С1 записуємо знак =. Наводимо курсор на знак f_x („Вставка функції“) і натискаємо ліву кнопку миші. З'являється діалогове вікно „Мастер функций“. Наводимо курсор на кнопку у вікні „Категория“, натискаємо ліву кнопку миші: випадає список функцій, в якому вибираємо (наводимо курсор) „Статистические“ і клацаємо лівою кнопкою. У вікні „Выберите функцию“ знаходимо „КОРРЕЛ“, клацаємо лівою кнопкою, потім – „ОК“. На екрані з'явиться вікно „Аргументы функции“. У масив 1 заносимо А1:А38, у масив 2 – В1:В38, потім натискаємо ОК. В комірці С1 з'являється цифра "- 0,46413" (див. табл. 4.1). Це – значення коефіцієнта кореляції r_{xy} . Знак "-" вказує на обернену залежність (більше пропусків – нижче оцінка). За таблицю 4 знаходимо імовірність того, що 38 вимірювань двох некорельованих змінних дадуть коефіцієнт кореляції $r \geq 0,46$. Для 40 вимірювань імовірність менша від 1,1 %, тобто кореляція високо значуща. Немає сумніву в тому, що погані оцінки на контрольній роботі пов'язані з пропусками занять.

Табл. 4.1

	А	В	С
1	4	0	-0,46413
2	1	1	
3	4	0	
4	1	8	
5	2	0,5	
6	1	4	
7	2	2,5	
8	2	1	
9	0	3	
10	1	3	
11	0	7	
12	0	4	
13	0	2,5	
14	2	0	
15	0	1	
16	3	1	
17	0	2	

18	0	7	
19	0	4	
20	0	7	
21	0	1	
22	1	0	
23	0	6	
24	2	4,5	
25	0	4	
26	1	2	
27	0	7	
28	1	1	
29	5	0	
30	0	1,5	
31	2	2	
32	0	3	
33	0	0,5	
34	0	0	
35	0	1	
36	0	3	
37	0	6	
38	0	6	

4.3. Перевірка статистичних гіпотез

Статистичними гіпотезами називають гіпотези про вид невідомого розподілу чи про невідомі параметри відомого розподілу [10]. Наприклад, статистичними будуть гіпотези:

- 1) Генеральна сукупність розподілена за нормальним законом;
- 2) Дві вибірки належать до однієї генеральної сукупності;
- 3) Дві вибірки належать до різних генеральних сукупностей;
- 4) Між двома вибірками існує кореляційна залежність і т. д.

Поряд з висунутою гіпотезою розглядають і гіпотезу, яка її заперечує. Висунуту гіпотезу називають **основною** або **нульовою**; її позначають H_0 . Гіпотезу, що заперечує основну, називають **конкуруючою** (**альтернативною**); її позначають H_1 . Висунута гіпотеза може бути правильною або неправильною, тому виникає необхідність її перевірки. Перевірка здійснюється статистичними методами, тому її теж називають **статистичною**. Статистична перевірка гіпотези може привести до помилок двох видів.

Помилка першого роду полягає в тому, що буде відкинута правильна гіпотеза.

Помилка другого роду полягає в тому, що буде обрана неправильна гіпотеза.

Імовірність здійснити помилку першого роду називають **рівнем значущості** і позначають α . Найчастіше рівень значущості обирають рівним 0,05 або 0,01. Якщо, наприклад, обрано рівень значущості рівний 0,05, то це означає, що в п'яти випадках зі ста ми ризикуємо допустити помилку першого роду (відкинути правильну гіпотезу). Інколи обирають третій рівень, рівний 0,001. Останній використовується рідко, оскільки він передбачає дуже високі вимоги до обґрунтування результатів експерименту.

Педагогічний експеримент полягає найчастіше в порівнянні результатів, наприклад, навчання експериментальної та контрольної груп учнів. В більшості випадків педагогічний експеримент є квазіекспериментом, оскільки на результати його впливає велика кількість неврахованих факторів. Крім того, майже завжди проводиться вибіркоче дослідження, коли вивчається лише 10-15 % членів сукупності. При формуванні вибірки практично неможливо дотриматись випадковості відбору, тому до початку експерименту слід виконати перевірку рівня підготовки та інших ознак у відібраних груп учнів і в якості контрольної вибрати групу з кращими

результатами перевірки.

В деяких випадках певну інформацію про досягнення учнів можна одержати при порівнянні результатів зрізів знань в одному й тому ж класі через певний проміжок часу.

В будь-якому разі з точки зору математичної обробки результатів ми повинні дати відповідь на питання: чи суттєво відрізняються між собою одержані два масиви чисел? Чи вправі ми зробити висновок, що результати експериментальної групи суттєво відрізняються від результатів контрольної?

Розглянемо порядок операцій, які виконуються в ході експерименту. Нульова гіпотеза – відмінностей між експериментальною та контрольною групами немає. Альтернативна – учні експериментальної групи показують більш високий рівень знань.

Підраховують середні арифметичні оцінок $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$.

Знаходять суму квадратів відхилень $SS_x = \sum (x_i - \bar{x})^2$ і обчислюють дисперсію σ_x^2 та стандартне відхилення σ_x :

$$\sigma_x^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}; \quad \sigma_x = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}.$$

Стандартне відхилення використовують для підрахунку **статистичної похибки середнього арифметичного**, яка виступає **мірою точності**, з якою вибірковий показник репрезентує генеральний параметр:

$$\sigma_{\bar{x}} = \sqrt{\frac{\sigma_x^2}{n}} = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n(n-1)}}.$$

Точність визначення вибіркової середньої визначається відношенням $\sigma_{\bar{x}}$ до \bar{x} . Це відношення позначають C_S :

$$C_S = \frac{\sigma_{\bar{x}}}{\bar{x}} \cdot 100\%$$

Точність визначення середніх показників вважається достатньою, якщо C_S не перевищує 15 %.

Для перевірки обраної гіпотези використовують **критерій достовірності**. Оскільки найчастіше в дослідженнях одержуємо сукупність даних, які можна описати нормальним законом розподілу, то скористаємося критерієм, що використовує значення середнього арифметичного і ширини довірчого інтервалу в залежності від обраного рівня значущості. Цей критерій називають **t-критерієм Стьюдента**:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{S_{(\bar{x}_1 - \bar{x}_2)}}, \quad (4.10)$$

де \bar{x}_1, \bar{x}_2 – середні арифметичні в експериментальній та контрольній групах, а $S_{(\bar{x}_1 - \bar{x}_2)}$ – стандартна похибка різниці середніх арифметичних, яка для $n_1 \neq n_2$ визначається за формулою:

$$S_{(\bar{x}_1 - \bar{x}_2)} = \sqrt{\sigma_{x_1}^2 + \sigma_{x_2}^2} = \sqrt{\frac{\sum (x_i - \bar{x}_1)^2 + \sum (x_i - \bar{x}_2)^2}{n_1 + n_2 - 2} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)} \quad (4.11)$$

Знайшовши фактичне значення t-критерія Стьюдента t_ϕ , порівнюємо його з теоретичним значенням t_{st} для обраного рівня значущості α і числа ступенів свободи $f = n_1 + n_2 - 2$ (див. табл. 5 в додатку). Якщо $t_\phi \geq t_{st}$, то відмовляються від нульової гіпотези і обирають альтернативну.

Наведемо приклад застосування t-критерія Стьюдента. В комірки A1-A11 заносимо результати тестування експериментальної групи, в B1-B9 – контрольної (див. табл. 4). В комірках A12 і B12 обчислюємо середні арифметичні СРЗНАЧ(A1:A11) та СРЗНАЧ(B1:B9), в комірках A13 і B13 – стандартне відхилення СТАНДОТКЛОН(A1:A11) та СТАНДОТКЛОН(B1:B9), в комірках A14 і B14 суми квадратів відхилень КВАДРОТКЛ(A1:A11) та КВАДРОТКЛ(B1:B9). В комірці C1 набираємо формулу =(A12-B12)/КОРЕНЬ(0,01122*(A14+B14)). Число 0,01122 – це

результат обчислення $\left(\frac{1}{n_1} + \frac{1}{n_2} \right) \cdot \frac{1}{n_1 + n_2 - 2}$ (в прикладі $n_1 = 11, n_2 = 9$).

В комірці C1 з'являється число 3,982072 (див табл. 4.2). Табличне значення критерія Стьюдента для $f = 18$ дорівнює 2,878 (табл. 5 в додатку) при рівні значущості $\alpha = 0,01$ (можливість зробити помилку в одному з ста випадків). Оскільки $t_\phi > t_{st}$, то ми відхиляємо нульову

гіпотезу і обираємо альтернативну, що учні експериментальної групи показують більш високий рівень знань.

Табл. 4.2

	A	B	C
1	12	13	3,982072
2	14	9	
3	13	11	
4	16	10	
5	11	7	
6	9	6	
7	13	8	
8	15	10	
9	15	11	
10	18		
11	14		
12	13,63636	9,444444	
13	2,460599	2,185813	
14	60,54545	38,22222	

Смисл t -критерія Стьюдента стає особливо зрозумілим, якщо переписати вираз (4.10) для випадку, коли $n_1=n_2=n$. Підставимо в (4.10) вираз (4.11), використавши цю умову.

$$t = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{\sum (x_i - \bar{x}_1)^2 + \sum (x_i - \bar{x}_2)^2}{2 \cdot (n-1)} \cdot \frac{2}{n}}} = \frac{|\bar{x}_1 - \bar{x}_2| \cdot \sqrt{n}}{\sqrt{\frac{\sum (x_i - \bar{x}_1)^2 + \sum (x_i - \bar{x}_2)^2}{(n-1)}}}$$

В знаменнику під коренем стоїть сума дисперсій:

$$t = \frac{|\bar{x}_1 - \bar{x}_2| \cdot \sqrt{n}}{\sqrt{\sigma_{x_1}^2 + \sigma_{x_2}^2}}$$

Отже, чим більша вибірка, чим більша різниця середніх арифметичних і чим менші дисперсії, тим надійніший висновок про значущу відмінність між результатами експериментальної і контрольної груп.

Слід зауважити, що, відкидаючи правильну нульову гіпотезу, ми з імовірністю α ризикуємо здійснити помилку першого роду. З іншого боку, припустимо, що значення критерію відповідає умові прийняття

нульової гіпотези. Помилково було б думати, що тим самим нульова гіпотеза доведена. Правильно говорити, що дані спостережень узгоджуються з нульовою гіпотезою і не дають підстави її відкинути.

На практиці для більшої впевненості в обраній гіпотезі її перевіряють іншими способами або збільшують об'єм вибірки.

ЛІТЕРАТУРА

- 1) Зайдель А. Н. Погрешности измерений физических величин / А. Н. Зайдель. – Л. : Наука, 1985. – 112 с.
- 2) Тейлор Дж. Введение в теорию ошибок : пер. с англ. / Дж. Тейлор – М. : Мир, 1985. – 272 с.
- 3) Рего К. Г. Метрологическая обработка результатов технических измерений : справ. пособие / К. Г. Рего. – Киев : Техника, 1987. – 128 с.
- 4) Лабораторный практикум по общей физике / под ред. Е. М. Гершензона и Н. Н. Малова. – М. : Просвещение, 1985. – 353с.
- 5) Чертов А. Г. Единицы физических величин / А. Г. Чертов. – М. : Высш. шк., 1977. – 287 с.
- 6) Нижник В. Г. Вимірювання фізичних величин та обчислення помибок / В. Г. Нижник. – Київ : Рад. шк., 1979. – 104 с.
- 7) Скалкова Я. Методология и методы педагогического исследования : пер. с чешск. / Я. Скалкова [и др.] – М. : Педагогика, 1989. – 224 с.
- 8) Барковський В. В. Теорія ймовірностей та математична статистика / В. В. Барковський, Н. В. Барковська, О. К. Лопатін. – Київ : ЦУЛ, 2002. – 448 с.
- 9) Волощенко А. Б. Теорія ймовірностей та математична статистика : навч.-метод. посібник для самост. вивч. дисц. / А. Б. Волощенко, І. А. Джалладова. – К. : КНЕУ, 2003. – 256 с.
- 10) Гмурман В. Е. Теория вероятностей и математическая статистика/ В. Е. Гмурман. – 4-е изд., доп. – М. : Высшая школа, 1972. – 368 с.

Додатки

Таблиця 1.

Визначення похибки функції однієї змінної

№ п/п	Вид функції $A = f(x)$	Абсолютна похибка ΔA	Відносна похибка $E = \frac{\Delta A}{A}$
1	$C \cdot x, C = const$	$C \cdot \Delta x$	$\frac{\Delta x}{x}$
2	x^n	$n \cdot x^{-(n-1)} \Delta x$	$n \cdot \frac{\Delta x}{x}$
3	$\sqrt[n]{x}$	$\frac{1}{n} \cdot \frac{\sqrt[n]{x}}{x} \cdot \Delta x$	$\frac{1}{n} \cdot \frac{\Delta x}{x}$
4	$e^{\frac{x}{C}}, C = const$	$e^{\frac{x}{C}} \cdot \frac{\Delta x}{C} *$	$\frac{\Delta x}{C}$
5	$B^{\frac{x}{C}}, C = const, B = const$	$B^{\frac{x}{C}} \cdot \ln B \cdot \frac{\Delta x}{C}$	$\ln B \cdot \frac{\Delta x}{C}$
6	$\ln x$	$\frac{\Delta x}{x}$	$\frac{\Delta x}{\ln x \cdot x}$
7	$\sin \alpha$	$\cos \alpha \cdot \Delta \alpha **$	$\operatorname{ctg} \alpha \cdot \Delta \alpha **$
8	$\cos \alpha$	$\sin \alpha \cdot \Delta \alpha **$	$\operatorname{tg} \alpha \cdot \Delta \alpha **$
9	$\operatorname{tg} \alpha$	$\frac{\Delta \alpha}{\cos^2 \alpha} **$	$\frac{2 \Delta \alpha}{\sin 2 \alpha} **$
10	$\operatorname{ctg} \alpha$	$\frac{\Delta \alpha}{\sin^2 \alpha} **$	$\frac{2 \Delta \alpha}{\sin 2 \alpha} **$

* Для досить великих значень C ($\bar{x} \ll |C|$).

** $\Delta \alpha$ вимірюється в *рад*.

Таблиця 2.

Визначення похибки функції декількох змінних

№ п/п	Вид функції $A = f(a, b, c)$	Абсолютна похибка ΔA	Відносна похибка $E = \frac{\Delta A}{A}$
1	$Pa \pm Qb$ ($P, Q - const$)	$\sqrt{P^2 \Delta a^2 + Q^2 \Delta b^2}$	$\frac{\sqrt{P^2 \Delta a^2 + Q^2 \Delta b^2}}{Pa \pm Qb}$
2	$a \cdot b$	$\sqrt{\bar{b}^2 \Delta a^2 + \bar{a}^2 \Delta b^2}$	$\sqrt{\frac{\Delta a^2}{\bar{a}^2} + \frac{\Delta b^2}{\bar{b}^2}}$
3	$\frac{a}{b}$	$\frac{\sqrt{\bar{b}^2 \Delta a^2 + \bar{a}^2 \Delta b^2}}{\bar{b}^2}$	$\sqrt{\frac{\Delta a^2}{\bar{a}^2} + \frac{\Delta b^2}{\bar{b}^2}}$
4	$a^\alpha \cdot b^\beta \cdot c^\gamma$ ($\alpha, \beta, \gamma - const$)	—	$\sqrt{\alpha^2 \frac{\Delta a^2}{\bar{a}^2} + \beta^2 \frac{\Delta b^2}{\bar{b}^2} + \gamma^2 \frac{\Delta c^2}{\bar{c}^2}}$
5	$\frac{a}{a \pm b}$	$\frac{\sqrt{\bar{b}^2 \Delta a^2 + \bar{a}^2 \Delta b^2}}{(\bar{a} \pm \bar{b})^2}$	$\frac{\sqrt{\bar{b}^2 \Delta a^2 + \bar{a}^2 \Delta b^2}}{\bar{a} \cdot (\bar{a} \pm \bar{b})}$
6	$a \cdot b + c$	$\sqrt{\bar{b}^2 \Delta a^2 + \bar{a}^2 \Delta b^2 + \Delta c^2}$	$\frac{\sqrt{\bar{b}^2 \Delta a^2 + \bar{a}^2 \Delta b^2 + \Delta c^2}}{\bar{a} \cdot \bar{b} + \bar{c}}$
7	$a \cdot (b + c)$	$\sqrt{(\bar{b} + \bar{c})^2 \Delta a^2 + \bar{a}^2 (\Delta b^2 + \Delta c^2)}$	$\frac{\sqrt{(\bar{b} + \bar{c})^2 \Delta a^2 + \bar{a}^2 (\Delta b^2 + \Delta c^2)}}{\bar{a} \cdot (\bar{b} + \bar{c})}$

Таблиця 3. Коефіцієнти Стьюдента

n	Довірча імовірність α			
	0,5	0,7	0,95	0,999
2	1,00	2,0	12,7	636,6
3	0,82	1,3	4,3	31,6
4	0,77	1,3	3,2	12,9
5	0,74	1,2	2,8	8,6
6	0,73	1,2	2,6	6,9
7	0,72	1,1	2,4	6,0
8	0,71	1,1	2,4	5,4
9	0,71	1,1	2,3	5,0
10	0,70	1,1	2,3	4,8
15	0,69	1,1	2,1	4,1
20	0,69	1,1	2,1	3,9
40	0,68	1,1	2,0	3,6
60	0,68	1,0	2,0	3,5
120	0,68	1,0	2,0	3,4
∞	0,67	1,0	2,0	3,3

Таблиця 4. Подана у процентах імовірність $P_n(|r| \geq |r_0|)$ того, що результати n вимірювань двох некорельованих змінних дадуть коефіцієнт кореляції $|r| \geq |r_0|$, як функція n і r_0 .
(Прочерк вказує на імовірність, яка менша, ніж 0,05 %)

n	r_0										
	0	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1
3	100	94	87	81	74	67	59	51	41	29	0
4	100	90	80	70	60	50	40	30	20	10	0
5	100	87	75	62	50	39	28	19	10	3,7	0
6	100	85	70	56	43	31	21	12	5,6	1,4	0
7	100	83	67	51	37	25	15	8,0	3,1	0,6	0
8	100	81	63	47	33	21	12	5,3	1,7	0,2	0
9	100	80	61	43	29	17	8,8	3,6	1,0	0,1	0
10	100	78	58	40	25	14	6,7	2,4	0,5	-	0
11	100	77	56	37	22	12	5,1	1,6	0,3	-	0
12	100	76	53	34	20	9,8	3,9	1,1	0,2	-	0
13	100	75	51	32	18	8,2	3,0	0,8	0,1	-	0
14	100	73	49	30	16	6,9	2,3	0,5	0,1	-	0
15	100	72	47	28	14	5,8	1,8	0,4	-	-	0
16	100	71	46	26	12	4,9	1,4	0,3	-	-	0
17	100	70	44	24	11	4,1	1,1	0,2	-	-	0
18	100	69	43	23	10	3,5	0,8	0,1	-	-	0
19	100	68	41	21	9,0	2,9	0,7	0,1	-	-	0
20	100	67	40	20	8,1	2,5	0,5	0,1	-	-	0
25	100	63	34	15	4,8	1,1	0,2	-	-	-	0
30	100	60	29	11	2,9	0,5	-	-	-	-	0
35	100	57	25	8,0	1,7	0,2	-	-	-	-	0
40	100	54	22	6,0	1,1	0,1	-	-	-	-	0
45	100	51	19	4,5	0,6	-	-	-	-	-	0
	0	0,05	0,1	0,15	0,2	0,25	0,3	0,35	0,4	0,45	
50	100	73	49	30	16	8,0	3,4	1,3	0,4	0,1	
60	100	70	45	25	13	5,4	2,0	0,6	0,2	-	
70	100	68	41	22	9,7	3,7	1,2	0,3	0,1	-	
80	100	66	38	18	7,5	2,5	0,7	0,1	-	-	
90	100	64	35	16	5,9	1,7	0,4	0,1	-	-	
100	100	62	32	14	4,6	1,2	0,2	-	-	-	

Таблиця 5. Критичні точки t – критерія Стьюдента.
Довірчі границі для t з f степенями свободи

f	Двосторонні границі			
	5 %	2 %	1 %	0,1 %
1	12,710	31,820	63,660	636.600
2	4,303	6,965	9,925	31,600
3	3,182	4,541	5,841	12,920
4	2,776	3,747	4,604	8,610
5	2,571	3,365	4,032	6,869
6	2,447	3,143	3,707	5,959
7	2,365	2,998	3,499	5,408
8	2,306	2,896	3,355	5,041
9	2,262	2,821	3,250	4,781
10	2,228	2,704	3,169	4,587
11	2,201	2,718	3,106	4,437
12	2,179	2,681	3,055	4,318
13	2,160	2,650	3,012	4,221
14	2,145	2,624	2,977	4,140
15	2,131	2,602	2,947	4,073
16	2,120	2,583	2,921	4,015
17	2,110	2,567	2,898	3,965
18	2,101	2,552	2,878	3,922
19	2,093	2,539	2,861	3,883
20	2,086	2,528	2,845	3,850
21	2,080	2,518	2,831	3,819
22	2,074	2,508	2,819	3,792
23	2,069	2,500	2,807	3,767
24	2,064	2,492	2,797	3,745
25	2,060	2,485	2,787	3,725
26	2,056	2,479	2,779	3,707
27	2,052	2,473	2,771	3,690
28	2,048	2,467	2,763	3,674
29	2,045	2,462	2,756	3,659
30	2,042	2,457	2,750	3,646
40	2,021	2,423	2,704	3,551
50	2,009	2,403	2,678	3,495
60	2,000	2,390	2,660	3,460
80	1,990	2,374	2,639	3,415
100	1,984	2,365	2,626	3,389
200	1,972	2,345	2,601	3,339
500	1,965	2,334	2,586	3,310
∞	1,960	2,326	2,576	3,291
f	2,5 %	1 %	0,5 %	0,05 %
	Односторонні границі			

Таблиця 6. Деякі статистичні функції електронних таблиць Excel

ДИСП	$\sigma_x^2 (S_x^2) = \frac{\sum (x_i - \bar{x})^2}{(n-1)} \quad (\text{вибірка})$
ДИСПР	$\sigma_x^2 = \frac{\sum (x_i - \bar{x})^2}{n} \quad (\text{генер. сукупність})$
ДОВЕРИТ	=ДОВЕРИТ(альфа; станд_откл; размер); напр. (0,05; 2.5; 50). Для $\alpha=0,05$: $\bar{x} \pm 1,96 \left(\frac{\sigma}{\sqrt{n}} \right)$.
КВАДРОТКЛ	$\sum (x_i - \bar{x})^2$
КОВАР	$\text{Cov}(X, Y) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$
КОРРЕЛ	$r_{xy} = \frac{\text{Cov}(X, Y)}{\sigma_x \cdot \sigma_y}; \quad r_{xy} = \frac{\sum (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$
ЛИНЕЙН	$y = mx + b$
НАКЛОН	$b = \frac{\sum (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$
СРЗНАЧ	$\frac{\sum x_i}{n}$
СРОТКЛ	$\frac{\sum x_i - \bar{x} }{n}$
СТАНДОТКЛ	$\sigma_x (S_x) = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}} \quad (\text{вибірка})$
СТАНДОТКЛП	$\sigma_x = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n}} \quad (\text{генер. сукупність})$