

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
Київський національний університет будівництва і архітектури

**МОДЕЛЮВАННЯ ТА ПРОГНОЗУВАННЯ  
СТАНУ ДОВКІЛЛЯ**

Методичні вказівки  
до виконання індивідуальної роботи  
для здобувачів першого (бакалаврського)  
рівня вищої освіти за спеціальностями  
101 «Екологія», 183 «Технології захисту  
навколишнього середовища»

Київ 2024

УДК 504.064.2.001.18:519.711

М74

Укладачі: О. А. Котовенко, канд. техн. наук, доцент  
О. Ю. Мірошніченко, старш. викладач

Рецензент Л. О. Василенко, канд. техн. наук, доцент

Відповідальний за випуск Т. М. Ткаченко, д-р техн. наук, професор

*Затверджено на засіданні кафедри технологій захисту  
навколишнього середовища та охорони праці, протокол № 6 від  
16 січня 2024 року*

В авторській редакції

**Моделювання** та прогнозування стану довкілля : методичні вказівки  
М74 до виконання індивідуальних робіт /уклад.: О.А. Котовенко,  
О.Ю. Мірошніченко. – Київ : КНУБА, 2024. – 44 с.

Містять теоретичні відомості, які необхідні для виконання  
індивідуальних робіт, приклади за всіма методами обробки даних, а  
також рекомендації щодо їх виконання

Призначено для здобувачів першого (бакалаврського) рівня  
вищої освіти за спеціальностями 101 «Екологія» та 183 «Технології  
захисту навколишнього середовища».

## ЗМІСТ

Загальні положення.....	4
ЧАСТИНА 1. Методи первинної обробки детермінованих даних спостережень.....	5
1.1. Апроксимація функцій.....	5
Постановка задачі.....	5
Вибір вигляду емпіричної формули.....	5
Визначення параметрів емпіричної формули.....	12
1.2. Інтерполювання функцій.....	21
Постановка задачі.....	21
Інтерполяційний багаточлен Лагранжа.....	21
Інтерполяційні методи Ньютона.....	22
Оцінки погрешностей інтерполяційних формул.....	26
ЧАСТИНА 2. Первинна статистична обробка даних спостережень.....	29
2.1. Статистичний розподіл.....	30
2.2. Полігони та гістограми частот.....	32
2.3. Емпірична функція розподілу випадкової величини.....	35
2.4. Визначення чисельних характеристик статистичного розподілу....	37
СПИСОК ЛІТЕРАТУРИ.....	42

## Загальні положення

У наш час методи математичного моделювання застосовуються в різних галузях науки, зокрема в екології і в технологіях захисту навколишнього середовища. Сьогодні неможливо вирішувати складні проблеми, що виникають в процесі дослідження стану біосфери та її компонентів, а також процесів, що проходять в цій складній системі, без сучасних методів інформатики, системного аналізу, методів фізичного і математичного моделювання і на їхній основі прогнозування станів цих складних систем.

Вивчення дисципліни «Моделювання і прогнозування стану довкілля» надає змогу засвоїти наукове обґрунтування суми знань про методи математичної постановки та методи вирішення задач природокористування, прогнозування стану навколишнього середовища, еволюційного розвитку техногенезу та їм подібних, що виникають і будуть виникати у екології в зв'язку з антропогенним навантаженням на навколишнє середовище та дають змогу підійти до екологічних проблем з погляду системного аналізу.

Метою індивідуальних робіт є набуття студентами практичних навичок із застосування методів первинної математичної обробки даних, які одержані внаслідок спостережень за станом навколишнього середовища або лабораторних експериментів та формування базової інформації для подальшого моделювання та прогнозування.

Методичні вказівки складаються з двох частин.

Частина 1 включає в себе методи для опрацювання даних спостережень, які є детермінованими, і складається з методів побудови апроксимаційних та інтерполювальних функцій за табличною залежністю.

Частина 2 включає в себе методи первинної статистичної обробки даних спостережень за станом навколишнього середовища або даних, що одержані у результаті лабораторних експериментів.

## ЧАСТИНА 1

### Методи первинної обробки детермінованих даних спостережень

#### 1.1. Апроксимація функцій

##### Постановка задачі

Нехай у результаті експерименту отримані дані, що виражають залежність деякої величини  $y$  від незалежної змінної  $x$ . Представимо їх у вигляді табл. 1.1.

Таблиця 1.1

$x$	$x_1$	$x_2$	...	$x_n$
$y$	$y_1$	$y_2$	...	$y_n$

На основі цих даних, зазвичай, можемо вирішувати такі задачі:

1. За відомою формулою, що виражає функціональну залежність  $y$  від  $x$ , потрібно визначити чисельне значення параметрів, що входять у формулу. Наприклад, у випадку лінійної залежності  $y$  від  $x$ , формула має вигляд:

$$y = ax + b,$$

де  $a$  і  $b$  – параметри, що підлягають визначенню.

2. Якщо невідомий характер функціональної залежності між величинами  $y$  і  $x$ , потрібно знайти її аналітичне вираження у вигляді залежності  $y = f(x)$ , яку прийнято називати *емпіричною формулою*.

Побудова емпіричної формули поділяється на два основних етапи:

- 1) вибір загального виду емпіричної формули;
- 2) визначення найкращих (у визначеному смислі) її параметрів (частіше така задача вирішується за допомогою методу найменших квадратів).

##### Вибір виду емпіричної формули

**Графічний спосіб.** Для визначення виду емпіричної формули необхідно за даними табл. 1.1 побудувати точковий графік (більш точно на міліметровому папері) таким чином: на координатній площині будуюмо точки, абсциси яких – значення аргументу  $x$ , ординати – значення  $y$ . Потім, дивлячись на точковий графік, проводимо (на око) плавну лінію так, щоби точки  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$  графіка були близькі до неї і розташовувалися з обох боків від неї (деякі точки можуть виявитися на лінії). Це і буде графік функції, що приблизно виражає залежність між величинами  $x$  і  $y$ , а рівняння побудованої в такий спосіб наближеної лінії, буде шуканою емпіричною формулою. Якщо наближена лінія, – пряма (рис. 1.1), то

природно вважати, що залежність між  $x$  і  $y$  – *лінійна*, й емпірична формула має вигляд (1.1). У випадку нелінійної залежності (наближена лінія – крива, рис. 1.2) вигляд формули можна визначити шляхом порівняння побудованої кривої з відомими кривими (окремі неправильності при цьому ігноруються).

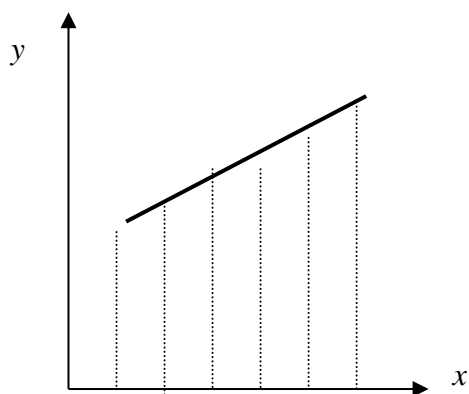


Рис. 1.1 Лінійна залежність

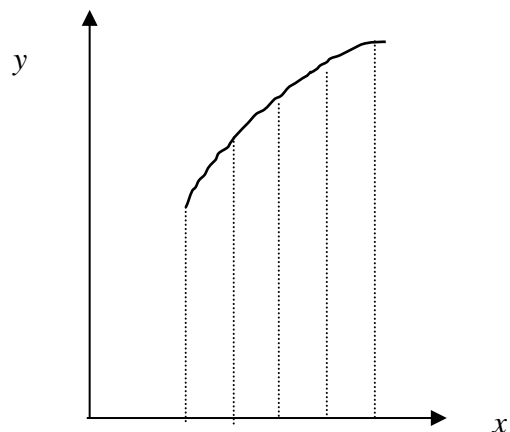


Рис. 1.2. Нелінійна залежність

**Приклад 1.1.** За даними табл. 1.2, користуючись графічним способом, знайти загальний вигляд емпіричної формули, що виражає залежність питомого об'єму  $V$  (м<sup>3</sup>/кг) насиченої водяної пари, від тиску  $p$  (атм).

Таблиця 1.2

$p$	3,304	1,457	0,903	0,382	0,245	0,167	0,119
$V$	0,5	1,2	2,0	5,0	8,0	12,0	17,0

**Рішення.** Позначимо  $x=p$ ,  $y=V$ . Нанесемо на координатну площину точки  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$  (рис. 1.3).

Ми бачимо, що точки розташовуються приблизно на кривій, що схожа на область гіперболи, і асимптотично наближується до осі  $0x$ . Тому емпіричну формулу запишемо у вигляді:

$$y = a + \frac{b}{x}.$$

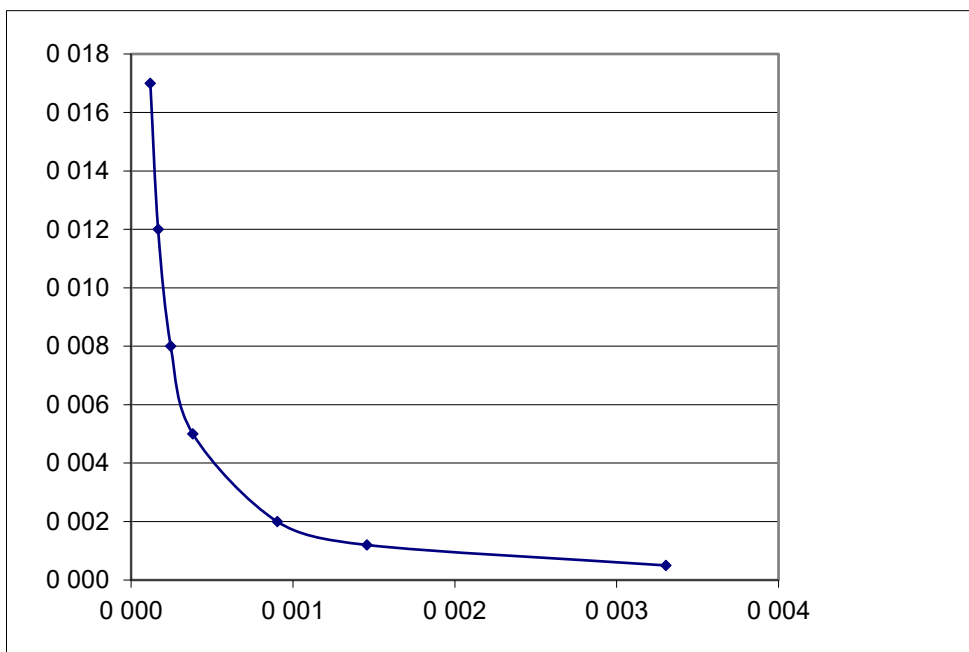


рис. 1.3. Графік функціональної залежності  $V$  від  $T$

**Аналітичний спосіб.** Цей спосіб ґрунтується на використанні вибору загального вигляду емпіричної формули за деякими аналітичними критеріями.

При цьому будемо вважати, що вихідні дані  $x_i, y_i$  ( $i=1,2, \dots, n; n > 3$ ) позитивні і  $x_1 < x_2 < \dots < x_n$ . Якщо ж усі  $x_i < 0$  (чи усі  $y_i < 0$ ), то досить розглянути таблицю значень  $(-x_i, y_i)$  (чи відповідно  $(x_i, -y_i)$ ). Аналогічно якщо  $x_i < 0$  і  $y_i < 0$ , досить побудувати емпіричну формулу для таблиці  $(-x_i, -y_i)$ .

У випадку, коли деякі зі значень  $x_i, y_i$  позитивні, а деякі – негативні, завжди можна виконати заміну змінних, підібравши такі числа  $M > 0$  і  $N > 0$ , при яких  $\xi_i = M + x_i > 0$ ,  $\eta_i = N + y_i > 0$ . Тоді задача зводиться до перебудування емпіричної формули для позитивних значень  $\xi_i, \eta_i$  ( $i = 1,2,3, \dots, n$ ).

#### *Емпіричні формули з двома параметрами*

Аналізуючи й описуючи закономірності хімічних і фізико-хімічних процесів і явищ емпіричну формулу найбільш часто приходиться вибирати серед функцій:

I)  $y = ax + b$  – лінійна;

II)  $y = ab^x$  – показова;

III)  $y = \frac{1}{ax + b}$  – дрібно-раціональна;

IV)  $y = a \ln x + b$  – логарифмічна;

V)  $y = ax^b$  – степенева, причому, якщо  $b > 0$  – залежність *параболічна*,  
якщо  $b < 0$  – залежність *гіперболічна*,

VI)  $y = a + \frac{b}{x}$  – *гіперболічна*;

VII)  $y = \frac{x}{ax + b}$  – дрібно-раціональна

Щоби визначити, яка з цих функцій найкраще описує експериментальні дані, необхідно виконати такі обчислення.

1. За даними табл. 1.1. знайти:

$$x_{\text{ар}} = \frac{x_1 + x_n}{2} \quad (\text{середнє арифметичне значень } x_1, x_n);$$

$$x_{\text{геом}} = \sqrt{x_1 \cdot x_n} \quad (\text{середнє геометричне});$$

$$x_{\text{гарм}} = \frac{2 \cdot x_1 \cdot x_n}{x_1 + x_n} \quad (\text{середнє гармонійне});$$

$$y_{\text{ар}} = \frac{y_1 + y_n}{2}, \quad y_{\text{геом}} = \sqrt{y_1 \cdot y_n}, \quad y_{\text{гарм}} = \frac{2y_1 \cdot y_n}{y_1 + y_n}.$$

2. За даними табл. 1.1, користуючись, наприклад, інтерполяційними формулами (або наближено побудованим графіком шуканої функції, навколо якого групуються експериментальні точки  $(x_i, y_i)$ ), знайти значення  $y_{\text{ар}}^*$ ,  $y_{\text{геом}}^*$ ,  $y_{\text{гарм}}^*$ , що відповідають знайденим у попередньому пункті значенням  $x_{\text{ар}}$ ,  $x_{\text{геом}}$ ,  $x_{\text{гарм}}$ . Причому, якщо, наприклад,  $x_{\text{ар}}$  (чи  $x_{\text{геом}}$ ,  $x_{\text{гарм}}$ ), збігається з табличним значенням  $x_i$  ( $i = 1, 2, 3, \dots, n$ ), то відповідне значення  $y_{\text{ар}}^*$  або  $y_{\text{геом}}^*$ ,  $y_{\text{гарм}}^*$  дорівнюватиме табличному значенню  $y_i$ , в іншому випадку  $y_{\text{ар}}^*$  можна визначити користуючись формулою лінійної інтерполяції :

$$y_{\text{ар}}^* = y_i + \frac{y_{i+1} - y_i}{x_{i+1} - x_i} \cdot (x_{\text{ар}} - x_i), \quad (1.2)$$

де  $x_{i+1}$ ,  $x_i$  – значення табл. 1.1, між якими знаходиться  $x_{\text{ар}}$

$$(x_i < x_{\text{ар}} < x_{i+1}), \quad y_i < y_{\text{ар}}^* < y_{i+1} \quad (i = 1, 2, 3, \dots, n-1).$$

### 3. Знайти величини:

$$\varepsilon_1 = |y_{ар}^* - y_{ар}|$$

$$\varepsilon_2 = |y_{ар}^* - y_{геом}|$$

$$\varepsilon_3 = |y_{ар}^* - y_{гарм}|$$

$$\varepsilon_4 = |y_{геом}^* - y_{ар}|$$

$$\varepsilon_5 = |y_{геом}^* - y_{геом}|$$

$$\varepsilon_6 = |y_{гарм}^* - y_{ар}|$$

$$\varepsilon_7 = |y_{гарм}^* - y_{гарм}|,$$

і серед них визначити мінімальне значення  $\varepsilon = \min\{\varepsilon_1, \varepsilon_2, \varepsilon_3, \dots, \varepsilon_7\}$ .

4. Зробити вибір емпіричної формули серед функцій I-VII. Шукана емпірична функція матиме вигляд:

I)  $y = ax + b$  – якщо  $\varepsilon = \varepsilon_1$ ;

II)  $y = ab^x$  – якщо  $\varepsilon = \varepsilon_2$ ;

III)  $y = \frac{1}{ax + b}$  – якщо  $\varepsilon = \varepsilon_3$ ;

IV)  $y = a \ln x + b$  – якщо  $\varepsilon = \varepsilon_4$ ;

V)  $y = ax^b$  – якщо  $\varepsilon = \varepsilon_5$ ;

VI)  $y = a + \frac{b}{x}$  – якщо  $\varepsilon = \varepsilon_6$ ;

VII)  $y = \frac{x}{ax + b}$  – якщо  $\varepsilon = \varepsilon_7$

Треба враховувати, що функції I-VII – монотонні, і, отже, експериментальні дані, що відповідають їм  $(x_i, y_i)$ , якщо  $\Delta x_i = x_{i+1} - x_i$  ( $i = 1, 2, 3, \dots, n-1$ ), повинні мати постійний знак  $\Delta y_i = y_{i+1} - y_i$  ( $i = 1, 2, 3, \dots, n-1$ ). У протилежному випадку залежності I-VII протипоказані.

Такий спосіб вибору емпіричної формули є досить грубим, тому що він не враховує поведінку всіх проміжних даних  $(x_i, y_i)$ . Крім того, може статися, що змінні  $x$  та  $y$  підкоряються іншій закономірності, і тоді вигляд емпіричної формули відрізнятиметься від функцій I-VII.

За даними табл. 1.1 легко переконатися в тому, що залежність між  $x$  та  $y$  лінійна, користуючись таким аналітичним критерієм:

якщо  $K_i = \Delta y_i / \Delta x_i = \text{const}$  ( $i = 1, 2, 3, \dots, n-1$ ;  $\Delta y_i = y_{i+1} - y_i$ ;  $\Delta x_i = x_{i+1} - x_i$ ), то точки  $(x_i, y_i)$ , мабуть, лежать точно на одній прямій лінії. Таким чином, якщо  $K_1 \approx K_2 \approx K_3 \approx \dots \approx K_{n-1}$ , то можна вважати, що формула  $y = ax + b$  приблизно представляє експериментальні дані.

**Приклад 1.2** Перевірити, використовуючи аналітичний критерій, чи є залежність між величинами  $x$  та  $y$ , значення яких представлені в табл. 1.3, лінійною.

Таблиця 1.3

$I$	1	2	3	4	5	6	7
$x_i$	2,5	4,5	6,5	8,5	10,5	12,5	14,5
$y_i$	5,6	7,8	9,9	12,2	14,3	16,5	18,7

**Рішення.** Неважко побачити, що  $\Delta x_i = x_{i+1} - x_i = 2$  ( $i = 1, 2, 3, \dots, 6$ ), тобто значення  $x_i$  рівновіддалені. Складемо таблицю різниць  $\Delta y_i$  (табл. 1.4).

Таблиця 1.4

$i$	1	2	3	4	5	6
$\Delta y_i$	2,2	2,1	2,3	2,1	2,2	2,2

Ми бачимо, що значення  $\Delta y_i$  практично постійні. Тому можна вважати, що залежність між  $x$  та  $y$  – лінійна, і відповідна емпірична формула має загальний вигляд (I).

#### Емпіричні формули з трьома параметрами

Найважливішими емпіричними формулами з трьома параметрами є:

$$y = ax^2 + bx + c \text{ – квадратична залежність} \quad (1.3)$$

$$y = ax^b + c \text{ – степенева залежність} \quad (1.4)$$

$$y = a \cdot e^x + c \text{ – показова залежність} \quad (1.5)$$

Розглянемо аналітичний критерій для квадратичної залежності. Будемо вважати, що  $\Delta x_i = x_{i+1} - x_i > 0$  ( $i = 1, 2, 3, \dots, n-1$ ), а послідовність  $y_1, y_2, \dots, y_n$  – монотонна, тобто  $\Delta y_i = y_{i+1} - y_i$  ( $i = 1, 2, 3, \dots, n-1$ ) зберігає постійний знак, або ж ця послідовність має єдиний екстремум, тобто різниця  $\Delta y_i$  лише один раз змінює знак.

Введемо розділені різниці першого і другого порядків:

$$[x_i, x_{i+1}] = \frac{\Delta y_i}{\Delta x_i}; \quad (1.6)$$

$$[x_i, x_{i+1}, x_{i+2}] = \frac{[x_{i+1}, x_{i+2}] - [x_i, x_{i+2}]}{x_{i+2} - x_i},$$

або 
$$[x_i, x_{i+1}, x_{i+2}] = \frac{\Delta \left( \frac{\Delta y_i}{\Delta x_i} \right)}{\Delta_i x_i}, \quad (1.7)$$

де 
$$\Delta_i x_i = x_{i+2} - x_i = \Delta x_i + \Delta x_{i+1}. \quad (1.8)$$

Доведено, що точки  $(x_i, y_i)$  ( $i = 1, 2, 3, \dots, n$ ), розташовані на параболі (див. рис. 1.3) тоді і тільки тоді, коли зберігають постійне значення всі розділені різниці другого порядку.

Зокрема, якщо значення  $x_1, x_2, \dots, x_n$  рівновіддалені, тобто  $\Delta x_i = x_{i+1} - x_i = \text{const}$  ( $i = 1, 2, 3, \dots, n - 1$ ), то для існування емпіричної квадратичної залежності (1.3) необхідно і досить, щоб була постійною друга різниця

$$\Delta^2 y_i = y_{i+2} - 2 \cdot y_{i+1} + y_i = \Delta y_{i+1} - \Delta y_i, \quad (i=1, 2, 3, \dots, n-2).$$

**Приклад 1.3.** Користуючись аналітичним критерієм, перевірити, чи існує квадратична залежність між величинами  $x$  та  $y$ , значення яких представлені в табл. 1.5.

Таблиця 1.5

$X$	6	11	17	23	28	32	37
$Y$	85,47	74,95	63,23	53,20	46,13	41,33	36,38

**Рішення.** Користуючись формулами (1.6) – (1.8), складаємо таблицю кінцевих різниць (табл. 1.6).

Тому що розділені різниці другого порядку практично постійні, можна вважати, що між величинами  $x$  та  $y$  має місце квадратична залежність.

Про існування залежностей (1.4) і (1.5) між величинами  $x$  і  $y$  можна судити, використовуючи або графічний спосіб, описаний вище, або метод вирівнювання.

Таблиця 1.6

$x$	$\Delta x$	$y$	$\Delta y$	$\Delta y/\Delta x$	$\Delta(\Delta y/\Delta x)$	$\Delta_1 x$	$\frac{\Delta(\Delta y/\Delta x)}{\Delta_1 x}$
6		85,47					
	5		-10,52	-2,10			
11		74,95			0,15	11	0,01
	6		-11,72	-1,95			
17		63,23			0,28	12	0,02
	6		-10,03	-1,67			
23		53,20			0,26	11	0,02
	5		-7,07	-1,41			
28		46,13			0,21	9	0,02
	4		-4,80	-1,20			
32		41,33			0,21	9	0,02
	5		-4,95	-0,99			
37		36,38					

### Визначення параметрів емпіричної формули

Якщо вид емпіричної формули обраний, то виникає задача визначення коефіцієнтів (параметрів), які входять у цю формулу.

У загальному вигляді ця задача ставиться в такий спосіб: нехай експериментальні дані (див. табл. 1.1) приблизно описуються формулою:

$$y = f(x, a_1, a_2, \dots, a_m), \quad (1.9)$$

де  $f$  – відома функція;  $a_1, a_2, \dots, a_m$  – невідомі постійні, число яких  $m$  звичайно менше числа точок  $(x_i, y_i)$ , тобто  $m < n$ . Потрібно визначити ці постійні.

Одним з методів рішення такої задачі є *метод найменших квадратів*.

Якщо в емпіричну формулу (1.9) підставити вихідні дані  $x_i, y_i$  ( $i=1,2,3,\dots, n$ ), то ліва частина формули не буде дорівнювати правій. Різниці (нев'язання):

$$\varepsilon_i = f(x, a_1, a_2, \dots, a_m) - y_i \quad (1.10)$$

називаються *відхиленнями* і являють собою відстані по вертикалі точок  $(x_i, y_i)$  від графіка емпіричної функції (1.9), узяті зі знаком «плюс» (+) чи зі знаком «мінус» (-).

Відповідно до методу найменших квадратів найкращими коефіцієнтами  $a_1, a_2, \dots, a_m$  вважаються ті, для яких сума квадратів відхилень

$$S(a_1, a_2, \dots, a_m) = \sum_{i=1}^n [f(x_i, a_1, a_2, \dots, a_m) - y_i]^2 \quad (1.11)$$

буде мінімальною.

Використовуючи необхідні умови екстремуму функції декількох змінних, одержуємо так звану нормальну систему для визначення коефіцієнтів  $a_i$  ( $i = 1, 2, 3, \dots, m$ ):

$$\frac{\partial S}{\partial a_1} = 0, \quad \frac{\partial S}{\partial a_2} = 0, \dots, \quad \frac{\partial S}{\partial a_m} = 0. \quad (1.12)$$

*Лінійна апроксимація*

Розглянемо застосування методу найменших квадратів для визначення параметрів найпростішого випадку. Нехай емпірична формула має вигляд (1.1). Використовуючи формулу (1.11) для даних табл. 1.1, запишемо

$$S(a, b) = \sum_{i=1}^n [ax_i + b - y_i]^2.$$

Нормальна система для визначення  $a, b$  буде мати вигляд:

$$\left. \begin{aligned} \frac{\partial S}{\partial a} &= 2 \sum_{i=1}^n [ax_i + b - y_i] \cdot x_i = 0 \\ \frac{\partial S}{\partial b} &= 2 \sum_{i=1}^n [ax_i + b - y_i] \cdot 1 = 0 \end{aligned} \right\}$$

Або після найпростіших перетворень одержимо:

$$\left. \begin{aligned} a \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i &= \sum_{i=1}^n x_i y_i \\ a \sum_{i=1}^n x_i + nb &= \sum_{i=1}^n y_i \end{aligned} \right\}. \quad (1.13)$$

Вирішивши систему (1.13) відомими методами (формули Крамера, метод Гауса та ін.), знайдемо коефіцієнти  $a$  і  $b$  шуканої емпіричної залежності.

Середньоквадратичну погрішність апроксимації визначаємо за формулою:

$$\delta = \sqrt{\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n-1}} \quad (1.14)$$

*Квадратична апроксимація*

Для емпіричної формули (1.3) на основі (1.11) запишемо

$$S(a, b, c) = \sum_{i=1}^n [ax_i + b - y_i]^2.$$

Одержимо тепер нормальну систему для визначення коефіцієнтів  $a, b, c$

$$\left. \begin{aligned} \frac{\partial S}{\partial a} &= \sum_{i=1}^n [ax_i^2 + bx_i + c - y_i] \cdot x_i^2 = 0 \\ \frac{\partial S}{\partial b} &= \sum_{i=1}^n [ax_i^2 + bx_i + c - y_i] \cdot x_i = 0 \\ \frac{\partial S}{\partial c} &= \sum_{i=1}^n [ax_i^2 + bx_i + c - y_i] \cdot 1 = 0 \end{aligned} \right\} \quad (1.15)$$

Після найпростіших перетворень система (1.15) буде:

$$\left. \begin{aligned} a \sum_{i=1}^n x_i^4 + b \sum_{i=1}^n x_i^3 + c \sum_{i=1}^n x_i^2 &= \sum_{i=1}^n x_i^2 y_i \\ a \sum_{i=1}^n x_i^3 + b \sum_{i=1}^n x_i^2 + c \sum_{i=1}^n x_i &= \sum_{i=1}^n x_i y_i \\ a \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i + nc &= \sum_{i=1}^n y_i \end{aligned} \right\} \quad (1.16)$$

Одержали систему трьох лінійних алгебраїчних рівнянь із трьома невідомими  $a, b, c$ . Її зручно вирішувати з використанням однієї з обчислювальних схем Гауса – схеми єдиного розподілу (це не виключає можливості застосування інших відомих методів).

Якщо емпірична формула має вигляд (II) – (VII) або (1.4) або (1.5), то перетворимо її за допомогою методу вирівнювання до лінійного виду (I), а потім застосуємо метод найменших квадратів.

Середньоквадратичну погрішність апроксимації визначаємо за формулою (1.14).

#### *Метод вирівнювання*

Нехай розглянуті способи дали змогу встановити, що між величинами існує нелінійна залежність, і знайдено приблизно загальний вигляд відповідної емпіричної формули. Перш ніж переходити до визначення чисельних значень коефіцієнтів в обраній формулі, необхідно, застосовуючи метод вирівнювання, переконатися в правильності зробленого припущення щодо її загального виду (іноді виявляється, що емпірична крива схожа на декілька кривих, рівняння яких різні).

Метод вирівнювання полягає в перетворенні апроксимуючої (наближуючої) функції таким чином, щоби перетворити її в лінійну функцію.

Розглянемо, наприклад, у системі координат  $xOy$  нелінійну залежність виду:

$$\psi(y) = a \cdot \varphi(x) + b, \quad (1.17)$$

де  $a, b$  – постійні,  $\psi(y), \varphi(x)$  – строго монотонні функції. Перейдемо до нової змінної  $X = \varphi(x)$ ,  $Y = \psi(y)$  так, щоб у новій системі координат  $XOY$  емпірична залежність (1.17) стала лінійною.

$$Y = aX + b. \quad (1.18)$$

Обчисливши значення  $X_i = \varphi(x_i)$ ,  $Y_i = \psi(y_i)$ , нанесемо точки  $(X_i, Y_i)$  на координатну площину  $XOY$ . Якщо ці точки розташовуються поблизу прямої лінії, то обрана емпірична формула (1.17) підходить для опису експериментальних даних.

Покажемо, як розглянуті нелінійні залежності (II) – (VII), перетворенням координат можна звести до лінійної.

Для показникової залежності  $y = ab^x$ , логарифмуючи, одержимо:

$$\lg y = x \lg b + \lg a.$$

Поклавши  $\lg y = Y$ ,  $x = X$ ,  $\lg b = B$ , матимемо рівняння прямої

$$Y = BX + A.$$

*Дрібно-раціональну функцію*

$$y = \frac{1}{ax + b}$$

перетворимо у лінійну в такий спосіб: Запишемо зворотну залежність

$$\frac{1}{y} = ax + b,$$

після чого введемо нові змінні  $X=x$ ,  $Y=1/y$ . Тоді одержимо залежність

$$Y = aX + b.$$

Для логарифмічної залежності  $y = a \lg x + b$  введемо нові змінні  $Y = y$ ,  $X = \lg x$ . Тоді одержимо лінійну залежність

$$Y = aX + B.$$

Маючи степеневу залежність  $y = b \cdot x^a$  ( $a > 0$ ,  $b > 0$ ,  $x > 0$ ,  $y > 0$ ), логарифмуємо її,  $\lg y = a \lg x + \lg b$ , та поклавши  $Y = \lg y$ ,  $X = \lg x$ ,  $B = \lg b$ , одержимо

$$Y = aX + b.$$

Гіперболічну залежність  $y = a + \frac{b}{x}$  перетворимо у лінійну, ввівши змінні  $Y=y$ ,  $X=1/x$ . Отримаємо

$$Y = a + bX.$$

Для дрібно-раціональної функції  $y = \frac{x}{ax + b}$  запишемо спочатку

зворотну залежність  $\frac{1}{y} = a + \frac{b}{x}$ , а потім введемо нові змінні  $Y=1/y$ ,  $X=1/x$ ,

після чого одержимо лінійну залежність

$$Y = a + bX.$$

Розглянемо тепер степеневу залежність  $y = ax^b + c$ .

Записавши її у вигляді  $y - c = ax^b$  і прологарифмувавши, одержимо  $\lg(y - c) = \lg a + b \lg x$ .

Звідси, беручи  $\lg(y - c) = Y$ ,  $\lg x = X$ , одержимо лінійну залежність

$$Y = bX + \lg a.$$

Якщо точки  $(X_i, Y_i)$ , де  $Y_i = \lg(y_i - c)$ ,  $X_i = \lg x_i$  нанести на координатну площину, і вони будуть розташовуватися на прямій (чи майже на прямій), то залежність (1.4) правдива. Для визначення значення  $Y_i$  необхідно знати постійну  $c$ . Одержимо формулу, що дає змогу знайти  $c$ . Запишемо середнє геометричне  $x_G = \sqrt{x_1 x_n}$ , де  $x_1, x_n$  крайні значення змінної  $x$ . Користуючись кресленням або

методом лінійної інтерполяції, знайдемо для  $x_s$  відповідне значення  $y_s$ . Припускаючи, що точки  $M_1(x_1y_1)$ ,  $M_s(x_sy_s)$ ,  $M_n(x_ny_n)$  розташовані на кривій (1.4), одержимо три рівності:

$$y_1 = ax_1^b + c, \dots, y_s = ax_s^b + c, \dots, y_n = ax_n^b + c.$$

Зводячи  $x_s = \sqrt{x_1x_n}$  у степінь  $b$  і помноживши на  $a$ , одержимо

$$ax_s^b = \sqrt{ax_1^b \cdot ax_n^b} \quad \text{або} \quad y_s - c = \sqrt{(y_1 - c)(y_n - c)}.$$

Вирішивши останню рівність відносно  $c$ , знайдемо:

$$c = \frac{y_1y_n - y_s^2}{y_1 + y_n - 2y_s}. \quad (1.19)$$

Подібним чином можемо перетворити показову залежність (1.5) у лінійну. Переносячи доданок  $c$  у ліву частину рівності (1.5) і логарифмуючи, одержимо:

$$\lg(y - c) = \lg a + (b \cdot M)x,$$

де  $M = 0,434229$ .

Таким чином,

$$Y = \lg a + (b \cdot M)x,$$

де  $Y = \lg(y - c)$ .

Для обчислення  $c$  використовуємо формулу (1.19), але  $y_s$  у цьому випадку відповідає середньому арифметичному  $x_s$

$$x_s = \frac{x_1 + x_n}{2}$$

і знаходиться за кресленням або лінійною інтерполяцією.

**Приклад 1.4.** Користуючись аналітичним способом, знайти загальний вид емпіричної формули, що відповідає табл. (1.2); визначити параметри методом найменших квадратів; оцінити точність отриманої емпіричної формули.

**Рішення.** З огляду на рис. 1.3, виберемо емпіричну формулу серед залежностей II- VII.

За даними табл.і 1.2 знаходимо:

$$x_{\text{ар}} = \frac{x_1 + x_n}{2} = \frac{3,304 + 0,119}{2} = 1,712;$$

$$x_{\text{геом}} = \sqrt{x_1 x_n} = \sqrt{3,304 \cdot 0,119} = 0,627;$$

$$x_{\text{гарм}} = \frac{2x_1 x_n}{x_1 + x_n} = \frac{2 \cdot 3,304 \cdot 0,119}{3,304 + 0,119} = 0,230;$$

$$y_{\text{ар}} = \frac{y_1 + y_n}{2} = \frac{0,5 + 17}{2} = 8,75;$$

$$y_{\text{геом}} = \sqrt{y_1 y_n} = \sqrt{0,5 \cdot 17} = 2,92;$$

$$y_{\text{гарм}} = \frac{2y_1 y_n}{y_1 + y_n} = \frac{2 \cdot 0,5 \cdot 17}{0,5 + 17} = 0,97.$$

Користуючись інтерполяційною формулою (1.2), знайдемо:

$$y^*_{\text{ар}} = 0,5 + \frac{1,2 - 0,5}{1,457 - 3,304} (1,712 - 3,304) = 1,103;$$

$$y^*_{\text{геом}} = 2 + \frac{5 - 2}{0,382 - 0,903} (0,627 - 0,903) = 3,589;$$

$$y^*_{\text{гарм}} = 8 + \frac{12 - 8}{0,167 - 0,245} (0,23 - 0,245) = 8,769.$$

Знайдемо величини:

$$\varepsilon_2 = |y^*_{\text{ар}} - y_{\text{геом}}| = |1,103 - 2,92| = 1,817$$

$$\varepsilon_3 = |y^*_{\text{ар}} - y_{\text{гарм}}| = |1,103 - 0,97| = 0,133$$

$$\varepsilon_4 = |y^*_{\text{геом}} - y_{\text{ар}}| = |3,589 - 8,75| = 5,161$$

$$\varepsilon_5 = |y^*_{\text{геом}} - y_{\text{геом}}| = |3,589 - 2,92| = 0,669$$

$$\varepsilon_6 = |y^*_{\text{гарм}} - y_{\text{ар}}| = |8,769 - 8,75| = 0,019$$

$$\varepsilon_7 = |y^*_{\text{гарм}} - y_{\text{гарм}}| = |8,769 - 0,97| = 7,799$$

$\varepsilon = \min \{ \varepsilon_2, \varepsilon_3, \dots, \varepsilon_7 \} = \varepsilon_6 = 0,019$ , то шукана емпірична формула має вигляд

$$y = a + \frac{b}{x}.$$

Ввівши нові змінні  $x = \frac{1}{x}$ ,  $Y = y$ , одержимо лінійну залежність

$$Y = bX + a.$$

Відповідні дані наведені в табл. 1.7

Таблиця 1.7

X	0,303	0,686	1,107	2,618	4,082	5,988	8,403
Y	0,5	1,2	2,0	5,0	8,0	12,0	17,0

Побудувавши в площині  $XOY$  точки  $(X_i, Y_i)$ , переконуємося, що вони лежать приблизно на одній прямій (рис. 1.4). Отже, вибір емпіричної формули зроблений правильно.

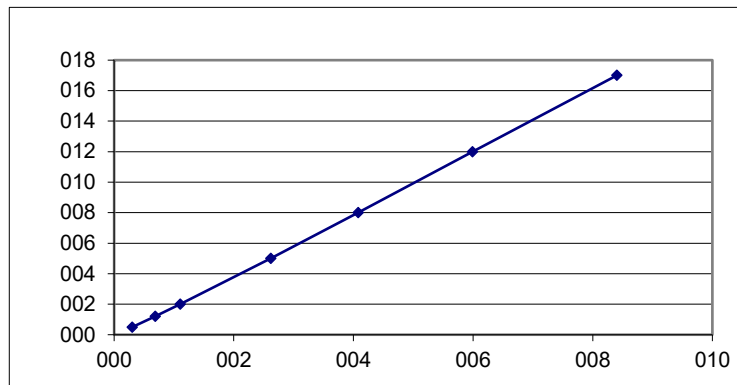


Рис. 1.4. Лінеаризована функція

Параметри  $a$  і  $b$  визначимо методом найменших квадратів. Для обчислення коефіцієнтів нормальної системи (1.13) складаємо табл. 1.8.

Таблиця 1.8

$i$	$X_i$	$Y_i$	$X_i^2$	$X_i Y_i$
1	0,303	0,500	0,092	0,152
2	0,686	1,200	0,471	0,823
3	1,107	2,000	1,225	2,214
4	2,618	5,000	6,854	13,090
5	4,082	8,000	16,663	32,656
6	5,988	12,000	35,856	71,856
7	8,403	17,000	70,610	142,851
<b><math>\Sigma</math></b>	<b>23,187</b>	<b>45,700</b>	<b>131,771</b>	<b>263,642</b>

Тепер запишемо нормальну систему рівнянь:

$$\left. \begin{aligned} 131,771b + 23,187a &= 263,642 \\ 23,187b + 7a &= 45,7 \end{aligned} \right\}$$

Знайдемо  $a$  і  $b$  за формулами Крамера:  $a = 0,001$ ,  $b = 1,97$

Отже, шукана емпірична формула матиме вигляд :

$$y = 0,001 + \frac{1,97}{x}. \quad (1.20)$$

Порівняння значень  $\bar{y}$ , отриманих за формулою з вихідними даними  $x$ , показано в табл. 1.9.

Таблиця 1.9

$x$	$y$	$\bar{y}$	$\varepsilon_2 = y - \bar{y}$	$\varepsilon^2$
0,304	0,5	0,6	-0,1	0,01
1,457	1,2	1,3	-0,1	0,01
0,903	2,0	2,2	-0,2	0,04
0,382	5,0	5,2	-0,2	0,04
0,2450	8,0	8,0	0,0	0,00
0,1670	12,0	11,8	0,2	0,04
0,1190	17,0	16,6	0,4	0,16
			$\Sigma$	0,3000

Середньоквадратичну погрішність апроксимації визначаємо за формулою (1.14), для розглянутого приклада одержимо:  $\delta \approx 0.02$ .

Така погрішність говорить про те, що знайдена емпірична формула (1.20) досить добре апроксимує представлені в таблиці 1.2 експериментальні дані.

## 1.2. Інтерполювання функцій

### Постановка задачі

Нехай функція  $y = f(x)$  на деякому відрізку  $[a, b]$  задана своїми значеннями  $y_0, y_1, \dots, y_n$ , у точках  $x_0, x_1, \dots, x_n$ , цього відрізка. Потрібно побудувати функцію  $F(x)$ , що належить відомому класу і набуває в точках  $x_0, x_1, \dots, x_n$ , таких же значень, що і  $f(x)$ , тобто таку, що

$$F(x_0) = y_0, F(x_1) = y_1, \dots, F(x_n) = y_n. \quad (1.21)$$

Таку функцію  $F(x)$  називають *інтерполюючою* функцією, а точки  $x_0, x_1, \dots, x_n$ , – *вузлами інтерполяції*.

Отриману інтерполяційну формулу  $y = F(x)$  зазвичай використовують для наближеного обчислення значень даної функції  $f(x)$  для значень аргументу  $x$ , відмінних від вузлів інтерполяції. Така операція називається *інтерполяцією функції  $f(x)$* . При цьому розрізняють *інтерполяцію у вузькому смислі*, якщо

аргумент  $x$  належить заданому відрізку  $[x_0, x_n]$  і *екстраполювання*, коли аргумент  $x$  перебуває за межами відрізка інтерполяції  $[x_0, x_n]$ .

Залежно від виду функції  $F(x)$  можуть бути різного типу інтерполяції: параболічна, тригонометрична і т. д. Найбільше поширення одержала *параболічна інтерполяція*, коли  $F(x)$  є багаточленом. Будемо позначати інтерполяційний багаточлен  $n$ -ої степені  $P_n(x)$ .

Можливі різні форми запису інтерполяційного багаточлена  $P_n(x)$ , що мають ті чи інші обчислювальні особливості. Розглянемо інтерполяційні багаточлени Лагранжа і Ньютона.

### ***Інтерполяційний многочлен Лагранжа***

Інтерполяційний багаточлен Лагранжа, що приймає у вузлах інтерполяції  $x_0, x_1, \dots, x_n$  відповідно значення  $y_0, y_1, \dots, y_n$  може бути записаний у такий спосіб:

$$\begin{aligned}
 P_n(x) = & y_0 \frac{(x-x_1)(x-x_2)(x-x_3)\dots(x-x_n)}{(x_0-x_1)(x_0-x_2)(x_0-x_3)\dots(x_0-x_n)} + \\
 & + y_1 \frac{(x-x_0)(x-x_2)(x-x_3)\dots(x-x_n)}{(x_1-x_0)(x_1-x_2)(x_1-x_3)\dots(x_1-x_n)} + \dots + \\
 & + y_i \frac{(x-x_0)(x-x_1)\dots(x-x_{i-1})(x-x_{i+1})\dots(x-x_n)}{(x_i-x_0)(x_i-x_1)\dots(x_i-x_{i-1})(x_i-x_{i+1})\dots(x_i-x_n)} + \dots + \\
 & + y_n \frac{(x-x_0)(x-x_1)(x-x_2)\dots(x-x_{n-1})}{(x_n-x_0)(x_n-x_1)(x_n-x_2)\dots(x_n-x_{n-1})}.
 \end{aligned} \tag{1.22}$$

Інтерполяційний багаточлен Лагранжа можна побудувати для будь-якого розташування вузлів інтерполяції.

**Приклад 1.5.** Об'єм 1 кг метану, якщо  $t=0^\circ\text{C}$  змінюється залежно від тиску таким чином:

$P$	673	473,1	361,6
$V$	60	65	70

Визначити об'єм 1 кг метану, якщо  $P=500$ .

**Рішення.** Оскільки число вузлів інтерполяції дорівнює трьом ( $x_0=p_0=673$ ;  $x_1=p_1=473,1$ ;  $x_2=p_2=361,6$ ), ступінь багаточлена Лагранжа буде дорівнює двом. Формула (1.22) набуде вигляду:

$$P_2(x) = y_0 \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)} + y_1 \frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)} + y_2 \frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)}$$

З огляду на те, що в нашому випадку  $y_0 = v_0 = 60$ ;  $y_1 = v_1 = 65$ ;  $y_2 = v_2 = 70$ ;  $x = p = 500$ , запишемо :

$$P_2(x) = 60 \frac{(500-473,1)(500-361,6)}{(673-473,1)(673-361,6)} + 65 \frac{(500-673)(500-361,6)}{(473,1-673)(473,1-361,6)} + 70 \frac{(500-673)(500-473,1)}{(361,6-673)(361,6-473,1)} = 64,03$$

Таким чином, об'єм 1 кг метану, якщо  $P = 500$  дорівнює 64,03.

### ***Інтерполяційні методи Ньютона***

Нехай функція  $y = f(x)$  задана своїми значеннями  $y_0 = f(x_0)$ ,  $y_1 = f(x_1)$ , ...,  $y_n = f(x_n)$  у рівновіддалених вузлах інтерполяції, тобто  $x_k - x_{k-1} = h$  — крок інтерполяції.

У цьому випадку доцільно користуватися інтерполяційними формулами Ньютона. Перш ніж їх розглядати, познайомимося з дуже важливим поняттям *кінцевих різниць*.

**Кінцеві різниці.** Назвемо *кінцевою різницею першого порядку* різницю між значеннями функції в сусідніх вузлах інтерполяції:

$$y_1 - y_0 = \Delta y_0, \quad y_2 - y_1 = \Delta y_1, \quad y_n - y_{n-1} = \Delta y_{n-1}.$$

З кінцевих різниць першого порядку можна одержати кінцеві різниці другого порядку :

$$\Delta^2 y_0 = \Delta y_1 - \Delta y_0, \quad \Delta^2 y_1 = \Delta y_2 - \Delta y_1, \quad \Delta^2 y_{n-2} = \Delta y_{n-1} - \Delta y_{n-2}.$$

*Кінцеві різниці  $n$ -го порядку* визначаються формулами

$$\Delta^n y_0 = \Delta^{n-1} y_1 - \Delta^{n-1} y_0, \quad \Delta^n y_1 = \Delta^{n-1} y_2 - \Delta^{n-1} y_1, \dots, \Delta^n y_{k-1} = \Delta^{n-1} y_k - \Delta^{n-1} y_{k-1}.$$

Кінцеві різниці різних порядків зручно розташовувати у формі таблиць двох видів: *горизонтальної* (табл. 1.10) або *діагональної* (табл. 1.11).

Таблиця 1.10

$x$	$y$	$\Delta y$	$\Delta^2 y$	$\Delta^3 y$
$x_0$	$y_0$	$\Delta y_0$	$\Delta^2 y_0$	$\Delta^3 y_0$
$x_1$	$y_1$	$\Delta y_1$	$\Delta^2 y_1$	
$x_2$	$y_2$	$\Delta y_2$		
$x_3$	$y_3$			

Таблиця 1.11

$x$	$y$	$\Delta y$	$\Delta^2 y$	$\Delta^3 y$
$x_0$	$y_0$			
		$\Delta y_0$		
$x_1$	$y_1$		$\Delta^2 y_0$	
		$\Delta y_1$		$\Delta^3 y_0$
$x_2$	$y_2$		$\Delta^2 y_1$	
		$\Delta y_2$		
$x_3$	$y_3$			

### Перша інтерполяційна формула Ньютона

Обчислення значень функції  $y = f(x)$  для значень аргументу, що лежать на початку таблиці, зручно проводити, користуючись *першою інтерполяційною формулою Ньютона*:

$$P_n(x) = y_0 + \frac{\Delta y_0}{h}(x - x_0) + \frac{\Delta^2 y_0}{2!h^2}(x - x_0)(x - x_1) + \frac{\Delta^3 y_0}{3!h^3}(x - x_0)(x - x_1)(x - x_2) + \dots + \frac{\Delta^n y_0}{n!h^n}(x - x_0)\dots(x - x_{n-1}). \quad (1.23)$$

На практиці часто використовують формулу (1.23) в іншому вигляді

$$P_n(x) = y_0 + q\Delta y_0 + \frac{q(q-1)}{2!}\Delta^2 y_0 + \frac{q(q-1)(q-2)}{3!}\Delta^3 y_0 + \dots + \frac{q(q-1)(q-2)\dots(q-n+1)}{n!}\Delta^n y_0, \quad (1.24)$$

де  $q = \frac{(x - x_0)}{h}$  являє собою *число кроків*, необхідних для досягнення точки  $x$ , виходячи з точки  $x_0$ .

Якщо у формулі (1.24) покласти  $n=1$ , то одержимо формулу лінійної інтерполяції:

$$P_1(x) = y_0 + q\Delta y_0. \quad (1.25)$$

Якщо  $n = 2$ , маємо формулу параболічного або квадратичного інтерполювання:

$$P_2(x) = y_0 + q\Delta y_0 + \frac{q(q-1)}{2!} \Delta^2 y_0. \quad (1.26)$$

Застосовуючи першу інтерполяційну формулу Ньютона, зручно користуватися горизонтальною таблицею кінцевих різниць (табл. 1.10), тому що в цьому випадку потрібні значення різниць функції знаходяться у відповідному горизонтальному рядку таблиці. Ступінь  $n$  полінома  $P_n(x)$  звичайно вибирають такою, щоб кінцева різниця  $\Delta^n y_i$  була постійна з заданим ступенем точності. За початкове значення  $x_0$  беруть найближче до заданого  $x$  табличне значення аргументу.

#### *Друга інтерполяційна формула Ньютона*

Для інтерполяції наприкінці таблиці звичайно застосовують другу інтерполяційну формулу Ньютона:

$$P_n(x) = y_n + \frac{\Delta y_{n-1}}{h}(x - x_n) + \frac{\Delta^2 y_{n-2}}{2!h^2}(x - x_n)(x - x_{n-1}) + \frac{\Delta^3 y_{n-3}}{3!h^3}(x - x_n)(x - x_{n-1})(x - x_{n-2}) + \dots + \frac{\Delta^n y_0}{n!h^n}(x - x_n)\dots(x - x_1). \quad (1.27)$$

Більш зручною для практичного використання є формула :

$$P_n(x) = y_n + q\Delta y_{n-1} + \frac{q(q+1)}{2!} \Delta^2 y_{n-2} + \frac{q(q+1)(q+2)}{3!} \Delta^3 y_{n-3} + \dots + \frac{q(q+1)(q+2)\dots(q+n-1)}{n!} \Delta^n y_0, \quad (1.28)$$

де  $q = \frac{(x - x_n)}{h}$ . При цьому як значення  $x_n$  беруть найближче до заданого  $x$  табличне значення аргумента.

### Екстраполювання функцій

Обидві інтерполяційні формули Ньютона можуть бути використані для *екстраполювання* функції, тобто для знаходження значень функції  $y$  для значень аргументів  $x$ , що лежать за границями таблиці (звичайно в межах кроку). Якщо  $x < x_0$  і  $x$  близько до  $x_0$ , то вигідно застосовувати першу інтерполяційну формулу Ньютона, причому тоді

$$q = \frac{(x - x_0)}{h} < 0.$$

Якщо ж  $x > x_n$ , і  $x$  близько до  $x_n$ , то зручніше користуватися другою інтерполяційною формулою Ньютона, причому

$$q = \frac{(x - x_n)}{h} > 0.$$

Таким чином, перша інтерполяційна формула Ньютона зазвичай використовується для *інтерполяції вперед і екстраполювання назад*, а друга інтерполяційна формула Ньютона, навпаки, – для *інтерполяції назад і екстраполювання вперед*.

**Приклад 1.6.** В'язкість води залежить від температури таким чином:

$t^{\circ}\text{C}$	10	13	16	19	22	
$\eta$ (H <sub>2</sub> O)	1.308	1.203	1.111	1.030	0.958	

Визначити яка в'язкість води, якщо:

а)  $t = 17$ ;

б)  $t = 12$

в)  $t = 9$ .

**Рішення.** Побудуємо горизонтальну таблицю кінцевих різниць (табл. 1.13), взявши позначення  $\eta = y$ ,  $t = x$ .

Таблиця 1.13

$x$	$y$	$\Delta y$	$\Delta^2 y$	$\Delta^3 y$
<u>10</u>	<u>1,308</u>	<u>-0,105</u>	<u>0,013</u>	<u>-0,002</u>
13	1,203	-0,092	0,011	-0,002
16	1,111	-0,081	0,009	
19	1,030	-0,072		
22	0,958			

Ми бачимо, що кінцеві різниці третього порядку постійні, тому ними обмежимося.

а) Оскільки значення  $x = t = 17$  розташоване ближче до кінця таблиці, скористаємося другою інтерполяційною формулою Ньютона (1.23). У якості  $x_n$  вибираємо вузол найближчий до розглянутої точки:  $x_n = 19$ . Тоді:

$$q = \frac{x - x_n}{h} = \frac{17 - 19}{3} = \frac{2}{3} \quad (x=t=17; \quad x_n=19; \quad h=3).$$

Підставивши з таблиці 1.13 у формулу (1.28) марковані курсивом різниці, одержимо

$$P_3(20) = 1.083.$$

б) Скористаємося першою інтерполяційною формулою Ньютона (1.24). Як  $x_0$  вибираємо вузол найближчий до точки  $x=12$ . Тоді:

$$q = \frac{(x - x_0)}{h} = \frac{12 - 10}{3} = \frac{2}{3}.$$

Підставивши у формулу (1.24) підкреслені різниці з таблиці 1.13, одержимо

$$P_3(12) = 1,237.$$

в) Скористаємося формулою (1.28) для екстраполювання назад. Візьмемо  $x_0 = 10$ ;  $x = 9$ ;  $h = 3$ . Тоді

$$q = \frac{x - x_0}{h} = \frac{9 - 10}{3} = \frac{1}{3}.$$

Підставивши у формулу (1.24) з табл. 1.13 підкреслені різниці, одержимо

$$P_3(9) = 1,34.$$

### **Оцінки погрешностей інтерполяційних формул**

Інтерполяційний багаточлен Лагранжа  $P_n(x)$  збігається з функцією  $y = f(x)$  у вузлах інтерполяції  $x_0, x_1, \dots, x_n$ . Про ступінь наближення  $P_n(x)$  до функції  $f(x)$  в інших точках можна судити по величині залишкового члена

$$R_n(x) = f(x) - P_n(x).$$

Якщо функція  $y = f(x)$  у розглянутій області  $a \leq x \leq b$  зміни  $x$ , що містить вузли інтерполяції  $x_0, x_1, \dots, x_n$ , має всі похідні  $f'(x), f''(x), \dots, f^{(n+1)}(x)$  до  $(n+1)$ -го порядку включно, то  $R_n(x)$  можна визначати за формулою:

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} (x-x_0)(x-x_1)\dots(x-x_n), \quad (1.29)$$

де  $\xi$  залежить від  $x$  і лежить усередині відрізка  $[a, b]$ .

Приймаючи, що  $M_{n+1} = \max |f^{(n+1)}(x)|$ ,  $a \leq x \leq b$ , одержуємо наступну оцінку погрішності інтерполяційного багаточлена Лагранжа:

$$|R_n(x)| \leq \frac{M_{n+1}}{(n+1)!} |(x-x_0)(x-x_1)\dots(x-x_n)|, \quad (1.30)$$

**Приклад 1.7.** З якою точністю можна обчислити  $\sqrt{117}$  за допомогою інтерполяційної формули Лагранжа для функції  $y = \sqrt{x}$ , вибравши вузли інтерполяції  $x_0=100$ ,  $x_1=121$ ,  $x_2=144$ ?

**Рішення.** Знайдемо:

$$y' = \frac{1}{2}x^{-\frac{1}{2}}, y'' = -\frac{1}{4}x^{-\frac{3}{2}}, y''' = \frac{3}{8}x^{-\frac{5}{2}}.$$

Звідси

$$M_3 = \max |y'''| = \frac{3}{8} \cdot \frac{1}{100^{\frac{5}{2}}} = \frac{3}{8} \cdot 10^{-5}, \quad \text{якщо } 100 \leq x \leq 144.$$

На підставі формули (1.30) одержимо :

$$|R_2| \leq \frac{3}{8} \cdot 10^{-5} \frac{1}{3!} |(117-100)(117-121)(117-144)| \approx 0.1 \cdot 10^{-2}.$$

Якщо вузли інтерполяції  $x_0, x_1, \dots, x_n$  – рівновіддалені, причому  $x_{i+1} - x_i = h$ , то, вважаючи що  $q = (x-x_0)/h$ , на підставі формули (1.29) одержимо залишковий член першої інтерполяційної формули Ньютона

$$R_n(x) = h^{(n+1)} \frac{q(q-1)(q-2)\dots(q-n)}{(n+1)!} f^{(n+1)}(\xi), \quad (1.31)$$

де  $\xi$  – деяке проміжне значення між вузлами інтерполяції  $x_0, x_1, \dots, x_n$  і розглянутою точкою  $x$ .

Аналогічно, поклавши у формулі (1.29)

$$q = \frac{(x - x_n)}{h},$$

одержимо залишковий член другої інтерполяційної формули Ньютона

$$R_n(x) = h^{(n+1)} \frac{q(q+1)(q+2)\dots(q+n)}{(n+1)!} f^{(n+1)}(\xi). \quad (1.32)$$

У практичних розрахунках аналітичний вид функції не завжди відомий. Тоді припускаючи, що в таблиці кінцевих різниць для функції  $y = f(x)$  різниці  $(n+1)$ -го порядку  $\Delta^{(n+1)}y$  майже постійні і  $h$  досить маленьке, а також з огляду на те, що

$$f^{(n+1)}(x) = \lim_{h \rightarrow 0} \frac{\Delta^{n+1}y}{h^{n+1}},$$

приблизно можна припустити

$$f^{(n+1)}(\xi) \approx \frac{\Delta^{n+1}y}{h^{n+1}}.$$

У цьому випадку залишковий член першої інтерполяційної формули Ньютона дорівнює:

$$R_n(x) \approx \frac{q(q-1)(q-2)\dots(q-n)}{(n+1)!} \Delta^{(n+1)}y_0,$$

а для другої інтерполяційної формули Ньютона залишковий член дорівнює:

$$R_n(x) \approx \frac{q(q+1)(q+2)\dots(q+n)}{(n+1)!} \Delta^{(n+1)}y_0.$$

## ЧАСТИНА 2

### Первинна статистична обробка даних спостережень

Випадковою називається величина, яка в результаті спостереження чи досліду набуває визначеного значення, невідомого яке саме.

Нехай для дослідження деякого явища, що характеризується випадковою величиною  $X$ , проводиться  $n$  спостережень, у кожному з яких  $X$  набуває визначеного значення.

Сукупність значень  $x_1, \dots, x_n$ , що набуває випадкова величина  $X$  в  $n$  спостереженнях, називається **простою статистичною сукупністю** чи простим статистичним рядом.

Звичайно проста статистична сукупність подається у вигляді таблиці. У першому ряду цієї таблиці вказують номер спостереження (досліду), у другому значення випадкової величини, що спостерігається.

Простий статистичний ряд являє собою первинну форму запису статистичного матеріалу і може бути оброблений різними способами.

Якщо необхідно дослідити достатньо велику сукупність однорідних об'єктів  $X_i$  відносно деякої кількісної або якісної ознаки, випадково вибирають з неї обмежену кількість об'єктів та піддають їх вивченню. Така сукупність випадково відібраних об'єктів називається **вибірковою сукупністю**, чи просто **вибіркою**. Та велика сукупність об'єктів, з якої проводиться вибір, називається **генеральною сукупністю**.

**Об'ємом сукупності** (вибіркової чи генеральної) називається кількість об'єктів цієї сукупності. Наприклад, об'єм вибірки, що подана у табл. 2.1 дорівнює 20.

В подальшому будемо розглядати проведені  $n$  дослідів (спостережень), як «вибірку з деякої чисто умовної генеральної сукупності», яка складається з нескінченної кількості можливих дослідів (спостережень), що можна було б провести в умовах.

Якщо одержана вибірка об'єктів  $n$ , у якій значення  $x_1$  випадкової величини  $X$  спостерігалось  $n_1$  раз;  $x_2$  -  $n_2$  раз, ...,  $x_k$  -  $n_k$  разів, при цьому

$$n_1 + n_2 + \dots + n_k = n.$$

Розташування значень  $x_1, x_2, \dots, x_n$  в порядку зростання називається ранжуванням статистичних даних. Значення  $x_1 \dots x_n$  називається **варіантами**.

Послідовність варіант записаних у порядку зростання називається **варіаційним (ранжованим) рядом**;  $n_1, \dots, n_k$  називаються **частотами варіант**  $x_1, \dots, x_n$ , відношення  $n_i/n = W_i$  ( $i = \overline{1, k}$ ) називається **відносною частотою** варіанти  $x_i$ .

## 2.1. Статистичний розподіл

**Статистичним розподілом вибірки** називається перелік варіант  $x_i$  та відповідних їм частот  $n_i$  або відносних частот  $W_i$  (його подають звичайно у вигляді таблиці).

Статистичний розподіл частот.

$x_i$	$x_1$	$x_2$	...	$x_k$
$n_i$	$n_1$	$n_2$	...	$n_k$

Статистичний розподіл відносних частот.

$x_i$	$x_1$	$x_2$	...	$x_k$
$W_i$	$W_1$	$W_2$	...	$W_k$

**Приклад 2.1.** В деякому джерелі, що спостерігається, у визначені відрізки часу відбиралися партії проб води. Шляхом відбору середньої проби з різних партій проб і її наступного аналізу були знайдені значення  $x_i$  випадкової величини  $X$  – вмісту деякого забруднюючого компонента в  $i$ -й пробі.

Завдання:

1. Записати простий статистичний ряд дослідження у вигляді таблиці.
2. Знайти статистичний розподіл частот і відносних частот за даними одержаної таблиці .

**Рішення**

1. Побудована таблиця має вигляд (табл. 2.1):

Таблиця 2.1

Простий статистичний ряд

$i$	$x_i$	$i$	$x_i$	$i$	$x_i$
1	2	8	2	15	9
2	5	9	5	16	6
3	6	10	5	17	5
4	7	11	2	18	9
5	7	12	6	19	9
6	9	13	2	20	9
7	5	14	5		

2. Випадкова величина  $X$  в 20-ти спостереженнях набула п'ять різних значень: 2, 5, 6, 7, 9. Значення спостерігались 2-4 рази; 5-6 разів; 6-3 рази; 7-2 рази; 9-5 разів. Запишемо статистичний розподіл частот у вигляді табл. 2.2.

Сума частот  $n_i$  дорівнює  $n = 20$  (об'єм вибірки). Визначивши відносні частоти  $W_i = n_i/n$  ( $i = 1,5$ ), запишемо їх у таблицю статистичного розподілу відносних частот.

Таблиця 2.2

**Статистичний розподіл частот.**

Варіанти $x_i$	2	5	6	7	9
Частоти $n_i$	4	6	3	2	5

Таблиця 2.3

**Статистичний розподіл відносних частот**

Варіанти $x_i$	2	5	6	7	9
Відносні частоти $W_i$	0,2	0,3	0,15	0,1	0,25

З табл. 2.3 бачимо, що

$$\sum_{i=1}^5 W_i = 1.$$

Статистичний розподіл можна також задати у вигляді послідовності інтервалів і відповідних їм частот (у якості частоти, відповідної інтервалу, беруть суму частот варіант, що потрапили у цей інтервал). Такий розподіл часто називають *інтервалом статистичного розподілу*.

Для одержання інтервального статистичного розподілу, необхідно весь діапазон значень  $x_i$  випадкової величини  $X$  розділити на інтервали і підрахувати кількість  $m_i$  значень  $x_i$ , що приходяться на  $i$ -й інтервал, а також знайти відносну частоту  $W_i = m_i/n$  ( $i = \overline{1,k}$ ), що відповідає  $i$ -му інтервалу. Кількість інтервалів не має бути надто великою чи надто малою. Практично групування в інтервали рекомендується, якщо  $n > 30$ ; якщо  $n > 100$  рекомендована кількість інтервалів буде 10–20, якщо  $n < 100$  буде 6–8. Довжини інтервалів можуть бути як однаковими, так і різними.

Для зручності обчислень віддають перевагу однаковій довжини інтервалів. Але, якщо дослідні дані розподілені нерівномірно, то у області найбільшої щільності розподілу можна обирати інтервали більш вузькі, ніж в області малої густоти. Інтервальні статистичні розподіли подають у вигляді табл. 2.4.

Таблиця 2.4

$l_i$	$x_1, x_2$	$x_2, x_3$	...	$x_i, x_{i+1}$	...	$x_k, x_{k+1}$
$m$	$m_1$	$m_2$	...	$m_i$	...	$M_k$

де  $l_i$  – позначення  $i$ -го інтервалу;  $x_i, x_{i+1}$  – границі  $i$ -го інтервалу;  $m_i$  – відповідна частота;  $k$  – кількість інтервалів.

Таку таблицю називають ще статистичним рядом.

## 2.2. Полігони та гістограми частот

**Приклад 2.2.** Проведено аналіз 50 проб води на вміст забруднюючої речовини. Результати аналізу подані у табл.2.5.

Таблиця 2.5

Інтервальний статистичний розподіл

$I_i$	3-8	8-13	13-18	18-23	23-28
$m_i$	5	10	25	6	4
$W_i=m_i/n$	0,1	0,2	0,5	0,12	0,08

де  $m_i$  – кількість проб, у яких вміст продукту відповідає  $i$ -му інтервалу;  $W_i=m_i/n$  ( $n=50$ ) – відносна частота, що відповідає  $i$ -му інтервалу.

Таблиця 2.6

Варіанти $x_i$	$x_1$	$x_2$	...	$x_k$
Частоти $n_i$	$n_1$	$n_2$	...	$n_k$

Нанесемо точки  $(x_1, n_1) \dots (x_k, n_k)$  на координатну площину та поєднаємо їх відрізками прямих; одержимо ламану, яка називається **полігоном частот**.

**Приклад 2.3.** Побудувати полігон частот за даними, наведеними у табл. 2.2.

**Рішення.** Відкладаємо на осі абсцис варіанти  $x_i$  а на осі ординат – відповідні їм частоти  $n_i$  і, з'єднавши точки  $(x_i, n_i)$  відрізками прямих, одержуємо полігон частот (рис. 2.1а)

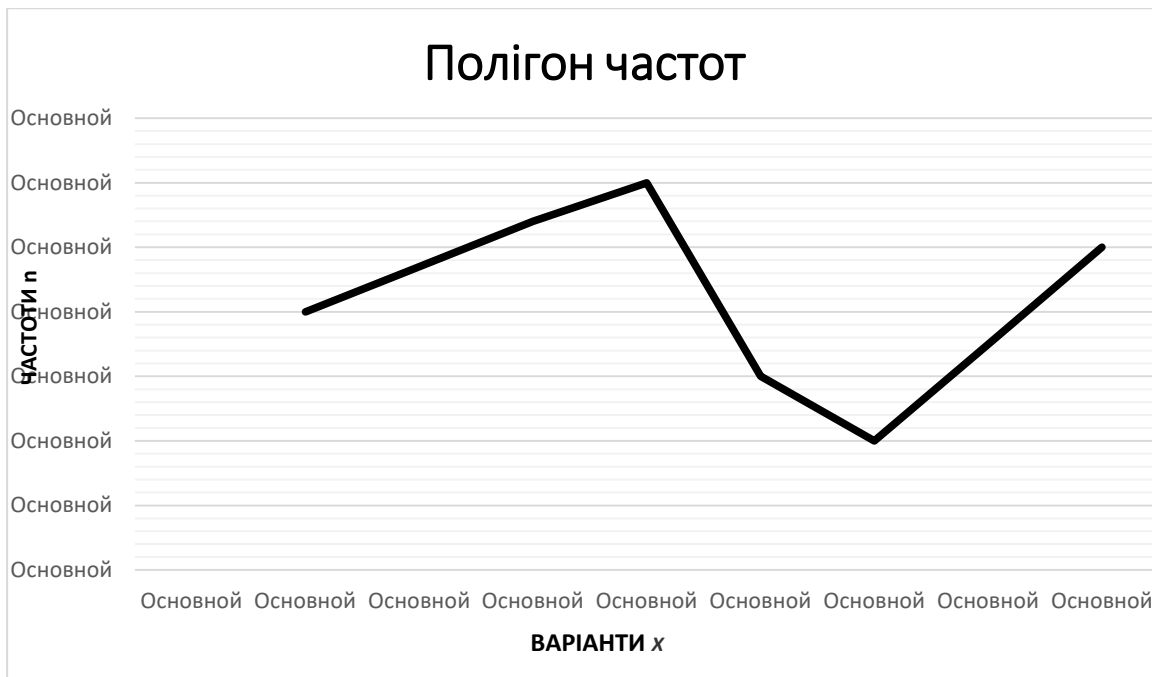


Рис. 2.1.а. Полігон частот за даними табл. 2.2

**Полігоном відносних частот** називають ламану, відрізки якої поєднують точки  $(x_1, W_1), \dots, (x_k, W_k)$ ; де  $W_i = n_i/n$ ,  $(i = \overline{1, k})$ .

**Приклад 2.4.** Побудувати полігон відносних частот за даними, наведеними у табл. 2.3.

**Рішення.** Відкладаємо на осі абсцис варіанти  $x_i$ , а на осі ординат – відповідні їх відносні частоти  $W_i$ . З'єднаємо точки  $(x_i, W_i)$  відрізками прямих і одержуємо шуканий полігон відносних частот. (рис. 2.1 б.)

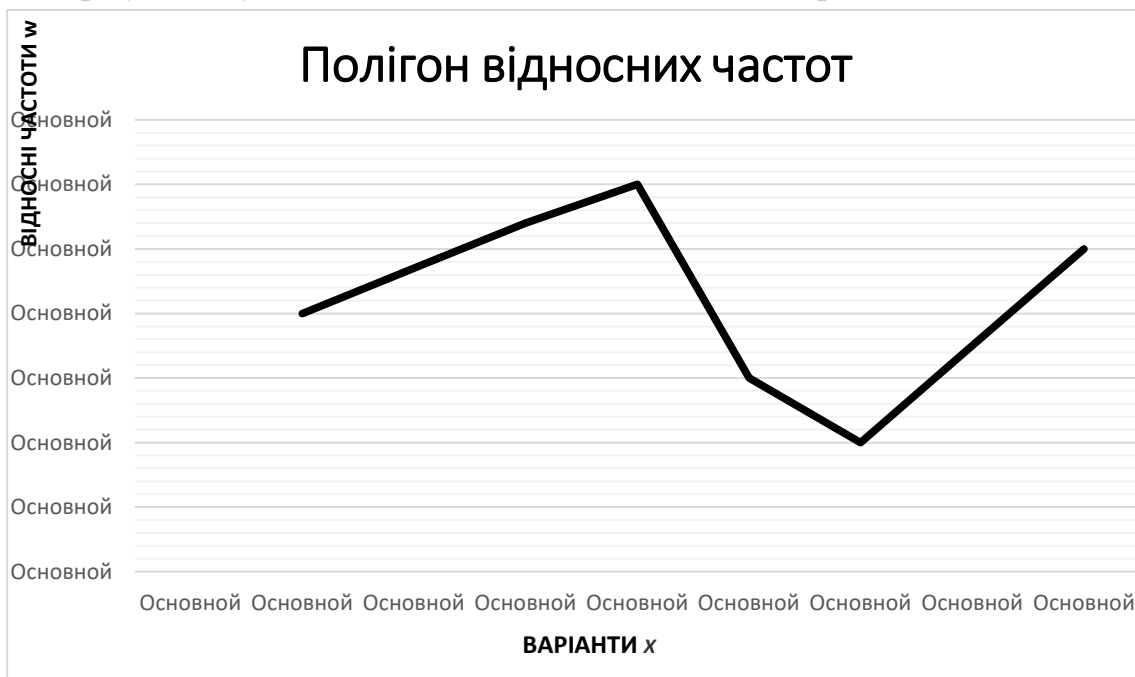


Рис. 2.1.б. Полігон відносних частот за даними табл. 2.3

Інтервальний статистичний розподіл графічно зображується у вигляді так званих гістограм.

**Гістограмою частот** називають ступеневу фігуру, яка складається з прямокутників, основою яких є інтервали довжини  $h$ , а висоти дорівнюють відношенню  $m_i/h$  (щільність частоти, де  $m_i$  – сума частот варіант, що потрапили до  $i$ -го інтервалу). Площа гістограми частот дорівнює об'єму вибірки  $n$ .

**Гістограмою відносних частот** називають ступеневу фігуру, що складається з прямокутників, основами яких є інтервали довжини  $h$ , а висоти рівні відношенню  $W_i/h$  (щільності відносної частоти). Площа гістограми відносних частот дорівнює сумі усіх відносних частот, тобто одиниці.

**Приклад 2.5.** Побудувати гістограми частот та відносних частот за даними табл. 2.5.

**Рішення.** На осі абсцис відкладаємо  $h = 5$ . Над цими інтервалами проводимо відрізки, які паралельні осі абсцис на відстані  $m_i/h$  для чого знайдемо:

$$m_1/h = 1; m_2/h = 2; m_3/h = 5; m_4/h = 1,2; m_5/h = 0,8.$$

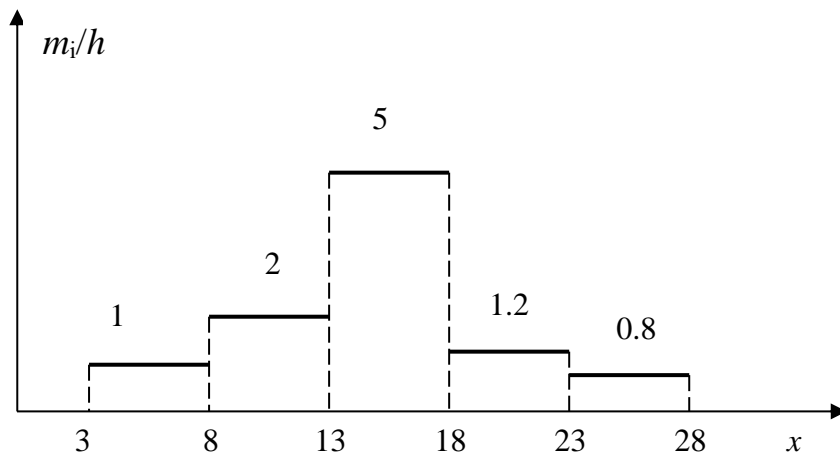


Рис. 2.2. Гістограма частот

Для побудови гістограми відносних частот знайдемо щільності:

$$W_1/h=0,02; W_2/h=0,04; W_3/h=0,1; W_4/h=0,024; W_5/h=0,0016.$$

Відкладаємо на осі абсцис задані інтервали  $h$ . Проведемо над цими інтервалами відрізками, які паралельні осі абсцис, на відстані  $W_i/h$ .

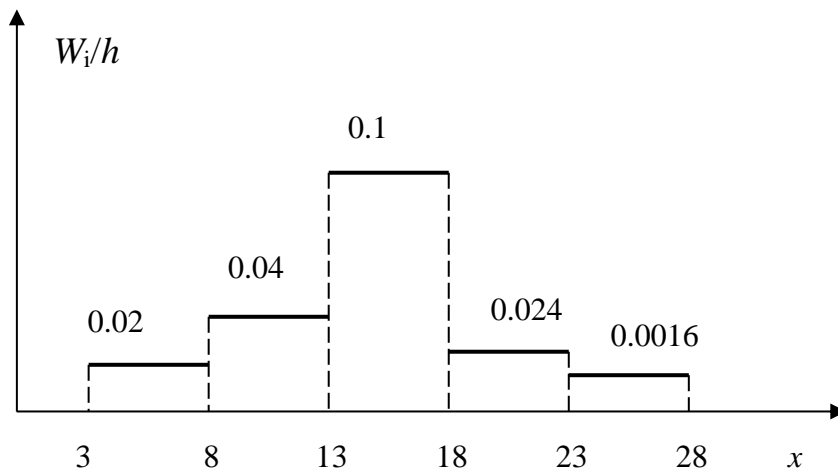


Рис. 2.3. Гістограма відносних частот

### 2.3 Емпірична функція розподілу випадкової величини

Одним із способів обробки статистичних даних є побудова емпіричної функції розподілу випадкової величини. Нехай відомо статистичний розподіл частот (табл.2.7). Позначимо  $n_k$  – кількість спостережень у яких  $X$  прийняла значення, менше ніж  $x$ ;  $n$  – загальна кількість досліджень (спостережень) – об’єм вибірки.

Таблиця 2.7

$x_i$	$x_1$	$x_2$	...	$x_k$
$n_i$	$n_1$	$n_2$	...	$n_k$

**Емпіричною функцією розподілу** (функцією розподілу вибірки) називають функцію  $F^*(x)$ , яка визначає для кожного значення  $x$  відносну частоту події  $X < x$ :  $F^*(x) = n_k/n$ .

**Властивості** емпіричної функції такі:

- 1) значення її належить відрізку  $[0,1]$ ;
- 2)  $F^*(x)$  – функція, що не зменшується;
- 3) якщо  $x_1$  – найменша варіанта, а  $x_k$  – найбільша, то  $F^*(x) = 0$ , якщо  $x < x_1$  та  $F^*(x) = 1$ , якщо  $x > x_k$ .

Графік емпіричної функції розподілу має ступеневий вигляд. Величини стрибків (скачків) у точках  $x_1$  і  $x_k$  дорівнюють частотам  $n_1, \dots, n_k$ .

Доведено, що у разі збільшення кількості дослідів  $n$  для будь-якого  $x$  відносна частота події  $X < x$  наближується до ймовірності цієї події. Відповідно, емпірична функція розподілу  $F^*(x)$  наближується до інтегральної функції  $F^*(x) = P(X < x)$ , яку часто називають **теоретичною функцією розподілу**. Графік  $F^*(x)$  у цьому випадку буде нескінченно наближуватися до плавної кривої  $F(x)$ . Таким чином емпірична функція розподілу вибірки служить для оцінки теоретичної функції розподілу генеральної сукупності.

**Приклад 2.6.** Побудувати емпіричну функцію по статистичному розподілу частот, поданому у табл. 2.2.

Таблиця 2.2

Статистичний розподіл частот

Варіанти $x_i$	2	5	6	7	9
Частоти $n_i$	4	6	3	2	5

**Рішення.** Найменша варіанта дорівнює 2, відповідно  $F^*(x) = 0$ , якщо  $x \leq 2$ . Значення  $X < 5$ , відповідно  $x_1 = 2$  спостерігалось 4 рази, тобто  $F^*(x) = 4/20 = 0,2$ , якщо  $2 < x \leq 5$ . Значення  $X < 6$ , відповідно  $x_1 = 2$  та  $x_2 = 5$  спостерігались  $4+6 = 10$  разів, тобто  $F^*(x) = 10/20 = 0,5$ , якщо  $5 < x \leq 6$ . Значення  $X < 7$ , відповідно  $x_1 = 2$ ,  $x_2 = 5$  та  $x_3 = 6$  спостерігались  $4 + 6 + 3 = 13$  разів, тобто  $F^*(x) = 13/20 = 0,65$ , якщо  $6 < x \leq 7$ . Значення  $X < 9$ , відповідно  $x_1 = 2$ ,  $x_2 = 5$ ,  $x_3 = 6$  та  $x_4 = 7$  спостерігались  $4+6+3+2=15$  разів, тобто  $F^*(x) = 15/20 = 0,75$ , при  $7 < x \leq 9$ . Оскільки  $x_5 = 9$  – найбільша варіанта, то  $F^*(x) = 1$ , якщо  $x > 9$ .

Запишемо емпіричну функцію і побудуємо її графік (рис.2.4):

$$F^*(x) = \begin{cases} 0, & \text{при } x \leq 2; \\ 0,2, & \text{при } 2 < x \leq 5; \\ 0,5, & \text{при } 5 < x \leq 6; \\ 0,65, & \text{при } 6 < x \leq 7; \\ 0,75, & \text{при } 7 < x \leq 9; \\ 1, & \text{при } x > 9. \end{cases}$$

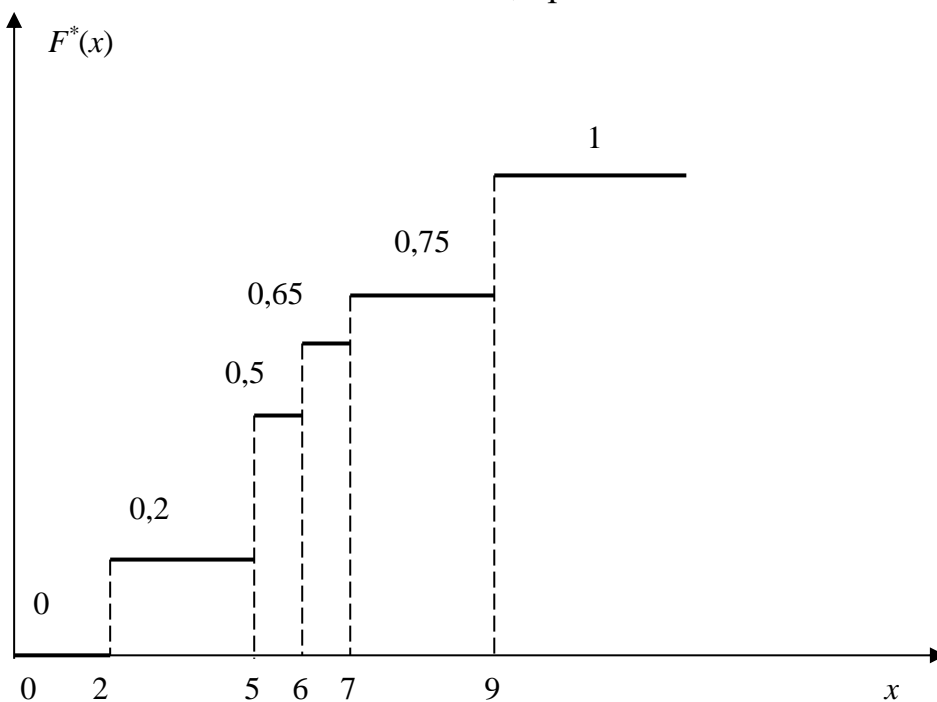


Рис.2.4. Графік емпіричної функції

## 2.4. Визначення чисельних характеристик статистичного розподілу

Нехай по вибірці об'єму  $n$  побудовано статистичний розподіл частот.

$x_i$	$x_1$	$x_2$	...	$x_k$
$n_i$	$n_1$	$n_2$	...	$n_k$

Знайдемо вибіркоче середнє:

$$\bar{x}_B = \frac{n_1 x_1 + \dots + n_k x_k}{n}, \quad (2.1)$$

де  $n = n_1 + n_2 + \dots + n_k$ .

**Вибіркове середнє** – це середнє зважене значення випадкової величини  $X$  з вагами, рівними відповідним частотам.

Характеристикою розсіяння даних спостереження  $x_i$  навкруги свого середнього значення  $\bar{x}_B$  є вибіркова дисперсія.

**Вибіркова дисперсія** – це середня зважена квадратів відхилень значень  $x_i$  від їх середнього значення  $\bar{x}_B$  з вагами, рівними відповідним частотам

$$D_B = \frac{\sum_{i=1}^k n_i (x_i - \bar{x}_B)^2}{n}. \quad (2.2)$$

**Вибірковим середнім квадратичним відхиленням (стандартом)** називають квадратний корінь з вибіркової дисперсії

$$\sigma_n = \sqrt{D_B}. \quad (2.3)$$

На практиці для обчислення дисперсії використовують таку формулу:

$$D_B = \bar{x}^2 - (\bar{x}_B)^2 \quad (2.4)$$

де  $\bar{x}^2 = \frac{n_1 x_1^2 + \dots + n_k x_k^2}{n}$  – середнє квадратів значень  $x_i$ , а  $(\bar{x}_B)^2$  квадрат середнього значення  $x_B$ , обчисленого за формулою (2.1).

### Приклад 2.7

За даними прикладу 2.1 (таблиця 2.2) знайти:

- середній вміст визначуваного компонента у ґрунті;
- дисперсію вмісту компонента у ґрунті.

Таблиця 2.2

Статистичний розподіл частот

Варіанти $x_i$	2	5	6	7	9
Частоти $n_i$	4	6	3	2	5

### Рішення

а) по статистичному розподілу частот (див. табл. 2.2), використовуючи формулу (2.1), знайдемо:

$$\bar{x}_B = \frac{2 \cdot 4 + 5 \cdot 6 + 6 \cdot 3 + 7 \cdot 2 + 9 \cdot 5}{20} = 5,75.$$

б) за формулою (2.2) або (2.4) знаходимо:

$$D_B = \frac{56,25 + 3,38 + 0,19 + 3,13 + 52,81}{20} = 5,79.$$

Таким чином, середній вміст компоненту в ґрунті дорівнює 5,75%; дисперсія вмісту компонента в ґрунті дорівнює  $D_B=5,79$ .

Нехай усі значення випадкової величини  $X$  розбиті на кілька груп. Розглянувши кожну групу як самостійну сукупність, можна знайти її середнє арифметичне.

**Груповою середньою** називається середнє арифметичне значень величини  $X$ , що належать цій групі.

Знаючи групові середні та об'єм групи, можна знайти загальну середню: **загальне середнє** ( $\bar{x}$ ) дорівнює середньому арифметичному групових середніх, зважених по об'єму груп.

### Приклад 2.8

Знайти групові та загальне середні вибіркової сукупності значень  $x_i$  з таблиці спостережень 2.2, що сформовані (розбиті) у дві групи і подані у вигляді табл. 2.8.

Таблиця 2.8

#### Дані спостережень

Групи	$j=1$		$j=2$		
	Значення	2	5	6	7
Частоти	4	6	3	2	5
Об'єм групи	$N_1=4+6=10$		$N_2=3+2+5=10$		

## Рішення

$$\bar{x}_1 = \frac{4 \cdot 2 + 6 \cdot 5}{10} = 3,8;$$

$$\bar{x}_2 = \frac{3 \cdot 6 + 2 \cdot 7 + 9 \cdot 5}{10} = 7,7;$$

знайдемо загальне середнє:

$$\bar{x} = \frac{N_1 \bar{x}_1 + N_2 \bar{x}_2}{N_1 + N_2} = 5,75.$$

**Груповою дисперсією** називають дисперсію значень випадкової величини  $X$ , що належить групі, відносно групової середньої

$$D_{i_{\text{гр}}} = \frac{\sum_{i=1}^k n_i (x_i - \bar{x}_j)^2}{N_j}, \quad (2.5)$$

де  $n_i$  – частота значень  $x_i$ ,  $j$  – номер групи,  $\bar{x}_j$  – групове середнє  $j$ -ї групи,

$N_j = \sum_{i=1}^k n_i$  – об'єм групи  $j$ .

## Приклад 2.9

Знайти групові дисперсії сукупності яка складається з двох груп за даними, які подані у табл. 2.8.

**Рішення.** Групові дисперсії будемо визначати за формулою (2.4), використовуючи середні  $\bar{x}_1$  та  $\bar{x}_2$ , що були знайдені у прикладі 2.8,

$$D_{1_{\text{гр}}} = \frac{4(2 - 3,8)^2 + 6(5 - 3,8)^2}{10} = 2,16;$$

$$D_{2_{\text{гр}}} = \frac{3(6 - 7,7)^2 + 2(7 - 7,7)^2 + 5(9 - 7,7)^2}{10} = 1,81.$$

Знаючи дисперсію кожної групи, можна знайти їх середнє арифметичне.

**Внутрішньо-груповою дисперсією** називають середнє арифметичне групових дисперсій, зважене за об'ємами груп:

$$D_{\text{вн гр}} = \frac{\sum N_j D_{j_{\text{гр}}}}{n}, \quad (2.6)$$

де  $N_j$  – об'єм  $j$  – ї групи;  $D_{j \text{ гр}}$  – групова дисперсія  $j$  – ї групи;  
 $n = \sum_{j=1}^k N_j$  – об'єм усієї сукупності;  $k$  – число груп.

### Приклад 2.10

Знайти внутрішньогрупову дисперсію за даними табл. 2.8.

**Рішення.** Внутрішньогрупову дисперсію знаходимо за формулою (2.6). Вона дорівнюватиме:

$$D_{\text{вн гр}} = \frac{10 \cdot 2,16 + 10 \cdot 1,81}{20} = 1,99.$$

Якщо відомі групові та загальна середні, то можна знайти дисперсію групових середніх відносно загальної.

**Міжгруповою дисперсією** називають дисперсію групових середніх відносно середньої.

$$D_{\text{між гр}} = \frac{\sum_{j=1}^k N_j (\bar{x}_j - \bar{x})^2}{n}, \quad (2.7)$$

де  $n = \sum_{j=1}^k N_j$  – об'єм усієї сукупності,  $\bar{x}_j$  – групова середня  $j$ -ї групи,  $N_j$  – об'єм групи  $j$ ,  $\bar{x}$  – загальна середня сукупності.

### Приклад 2.11

Знайти міжгрупову дисперсію за даними приклада 2.8 (див. табл. 2.8).

**Рішення.** Оскільки  $\bar{x}_1 = 3,8$ ;  $\bar{x}_2 = 7,7$ ;  $\bar{x} = 5,75$  (значення, які розраховані у попередніх прикладах) та використавши формулу (2.6), обчислюємо міжгрупову дисперсію.

$$D_{\text{між гр}} = \frac{10(3,8 - 5,75)^2 + 10(7,7 - 5,75)^2}{20} = 3,8.$$

Якщо сукупність складається з кількох груп, то загальна дисперсія дорівнює сумі внутрішньогрупової та міжгрупової дисперсій.

$$D_{\text{заг}} = D_{\text{між гр}} + D_{\text{вн гр}}. \quad (2.8)$$

$$D_{\text{заг}} = 3,8 + 1,99 = 5,79$$

Результати, одержані під час вирішення прикладів 2.10 та 2.11, показують, що ( $D_{\text{заг}}=5,79$ ;  $D_{\text{між гр}}=3,8$ ;  $D_{\text{вн гр}}=1,99$ ).

Крім розглянутих вже характеристик (вибіркове середнє та дисперсія) часто застосовуються інші характеристики.

**Медіаною**  $Me$  називають варіанту, яка поділяє варіаційний ряд на дві частини, які рівні щодо кількості варіант.

Якщо кількість варіант парна ( $n = 2k$ ), то  $Me = (x_k + x_{k+1})/2$ . Наприклад для ряду 4, 6, 8, 10 медіана дорівнює  $(6+8)/2 = 7$ . Якщо вона непарна ( $n = 2k + 1$ ), то  $Me = x_{k+1}$ . Наприклад для ряду 4, 6, 8, 10, 12, 14, 16 медіана дорівнює 10.

Найпростішою характеристикою розсіяння варіаційного ряду є **розмах варіювання**  $R$ , це різниця між найбільшою та найменшою варіантами. Для ряду з варіант 4, 6, 8, 10  $R=10-4=6$ .

## СПИСОК ЛІТЕРАТУРИ

1. *Заграй Я.М.*, Моделювання і прогнозування стану довкілля. навч. посіб./ Я.М. Заграй, О.А. Котовенко. – Київ : КНУБА, 2008. – 97 с.
2. *Заграй Я.М.*, Статистичний аналіз в екології : навч. посібник /Я.М. Заграй, О.А. Котовенко, В.О. Карасьова. – Київ : КНУБА, 2001. – 132 с.
3. *Брановицька С.В.* Обчислювальна математика та програмування. Обчислювальна математика в хімії і хімічній технології: підручник / С. В. Брановицька, Р. Б. Медведєв, Ю. Я. Фіалков. – Київ : ІВЦ Вид-во «Політехніка», 2004. – 220 с.

Для нотаток

Навчально-методичне видання

## **МОДЕЛЮВАННЯ ТА ПРОГНОЗУВАННЯ СТАНУ ДОВКІЛЛЯ**

Методичні вказівки  
до виконання індивідуальної роботи  
для здобувачів першого (бакалаврського)  
рівня вищої освіти за спеціальностями  
101 «Екологія», 183 «Технології захисту  
навколишнього середовища»

Укладачі КОТОВЕНКО Олена Андріївна,  
МІРОШНИЧЕНКО Олена Юріївна

Комп'ютерне верстання *А. П. Селівестрової*

Підписано до друку 10.01. 2025. Формат 60 × 84 <sup>1</sup>/<sub>16</sub>.

Ум. друк. арк. 2,66. Обл.-вид. арк. 2,75.

Електронний документ. Вид. № 196/III-24

Видавець і виготовлювач:

Київський національний університет будівництва і архітектури  
проспект Повітряних Сил, 31, Київ, Україна, 03037

Свідectво про внесення до Державного реєстру суб'єктів  
видавничої справи ДК № 808 від 13.02.2002