НАВЧАЛЬНО-МЕТОДИЧНА ЛІТЕРАТУРА

Міністерство освіти і науки України
Тернопільський національний технічний
університет імені Івана Пулюя

Кафедра інжинірингу
машинобудівних технологій

НАУКОВІ ДОСЛІДЖЕННЯ
І ТЕОРІЯ ЕКСПЕРИМЕНТУ

методичні вказівки
до виконання практичних робіт
для студенів факультету інженерії машин, споруд та
технологій
освітнього рівня «магістр»
зі спеціальності 131 «Прикладна механіка»

Тернопіль
2022

УДК 631.3.001

Наведено методичні вказівки до виконання практичних робіт, у кожній з яких викладено мету роботи, теоретичні відомості та передумови, порядок виконання роботи та аналізу отриманих результатів, а також контрольні запитання.

Затверджено методичною комісією факультету інженерії машин, споруд та технологій, протокол № 6 від 24.02.2022 р.

Укладачі: д.т.н., професор В.М. Барановський, д.т.н., професор Ч.В Пулька, к.т.н.. доц. Окіпний І.Б., к.т.н., доц. В.С. Сенчишин, к.т.н., доц. В.Р. Паньків

Рецензенти: д.т.н., проф. М.І. Пилипець, д.т.н., проф. О.Л. Ляшук

Навчальне видання
Наукові дослідження і теорія експерименту

Методичні вказівки
до виконання практичних робіт
для студенів факультету інженерії машин, споруд та технологій
освітнього рівня «магістр»
зі спеціальності 131 «Прикладна механіка»

Укладачі: Барановський Віктор Миколайович, Пулька Чеслав Вікторович, Окіпний Ігор Богданович, Сенчишин Віктор Степанович, Паньків Віталій Романович

ВСТУП

Вивчення технологій і робочих процесів технології машинобудування, розробка нових машин, технічних ліній і складних виробничих комплексів і практичне використання методів раціональної експлуатації технічних засобів вимагають від інженерно-наукових кадрів глибокого засвоєння наукових методів ідентифікації сучасних технологічних і виробничих процесів, а також сучасних виробничих об’єктів.

У даний час, коли розвиток автоматизації та механізації промислового виробництва вимагає сучасних, принципово нових підходів і технічних рішень, молоді спеціалісти повинні опанувати прогресивні методики та принципи методологій побудови, проведення та обробки і аналізу результатів експериментальних досліджень.

Принциповою особливістю системного підходу до проблем виробничого комплексу є прагнення врахувати весь спектр наявних факторів і чинників, які визначають подальший напрямок розвиток кожної галузі промисловості. Тому комплекс тісно пов’язаних між собою організаційних, економічних, технологічних та інших підходів, їх глибоке вивчення в поєднанні з моделюванням технологічних процесів і інженерних методів машинобудування є собою основою системного підходу до вирішення найбільш широких науково-технічних задач.

Програма навчальної дисципліни “Наукові дослідження і теорія експерименту” передбачає виконання практичних робіт з метою більш глибокого засвоєння методології та наукових методів і принципів проведення наукових досліджень, методики планування та обробки результатів експериментальних досліджень, що сприяє розвитку творчого мислення студентів та прийняття ними відповідних прогресивних рішень.

Задачі курсу – дати знання з основ організації наукових досліджень, поняття про експериментальні методи дослідження, методи побудови математичних моделей, планування і проведення багатофакторних експериментів та обробки і аналізу отриманих результатів.

Практичні роботи виконуються студентами самостійно на практичних заняттях під керівництвом викладача. У кожній роботі визначена її мета, викладені теоретичні передумови, порядок виконання і оформлення результатів та аналізу виконаних робіт.

При виконанні кожної практичної роботи студент спочатку вивчає теоретичні передумови, виконує необхідні математичні розрахунки за приведеними формулами, заповнює таблиці, будує відповідні графічні залежності і відповідним чином оформлює звіт, який потім пред’являється викладачу для перевірки та захисту. За кожну виконану роботу студент отримує відповідну оцінку.

Викладений матеріал може бути використаний як навчальний посібник для студентів вищих навчальних закладів

ПРАКТИЧНА РОБОТА № 1

МЕТОДИКА ПЛАНУВАННЯ ТА ПРОВЕДЕННЯ БАГАТОФАКТОРНОГО ЕКСПЕРИМЕНТУ

Мета роботи: вивчити методику планування, проведення та обробки результатів багатофакторного експерименту з використанням статистичних методів регресійного аналізу.

Теоретичні відомості

Експеримент є одним з основних методів проведення наукових досліджень. У більшості випадків експерименти є багатофакторними, тобто на об’єкт дослідження діє два і більше змінних вхідних факторів, або незалежних змінних величин, які впливають на вихідну величину (параметр оптимізації об’єкту дослідження).

Наприклад, дослідження залежності зміни продуктивності шнекового конвеєра, який є параметром оптимізації від таких факторів, як діаметра шнека, кроку спіральних витків, кута нахилу гвинтової лінії спіральних витків.

При цьому дослідження об’єкта (шнекового конвеєра) може здійснюватись класичним методом (послідовністю проведення однофакторних експериментів), або методом наукового планування експерименту.

Класичний метод, або класичний однофакторний експеримент – це послідовність проведених однофакторних експериментів, за яких всі незалежні змінні величини (фактори), крім однієї, приймаються постійними. Наприклад, при дослідженні впливу діаметра шнека, кроку спіральних витків, кута нахилу гвинтової лінії спіральних витків на зміну продуктивності шнекового конвеєра, в першому експерименті приймають діаметр шнека змінним в певних межах, а два останніх фактора постійними.

В наступному експерименті приймають змінним крок спіральних витків, а діаметр шнека та кут нахилу гвинтової лінії спіральних витків постійними і т.д.

Такий шлях дослідження зміни продуктивності шнекового конвеєра залежно від прийнятих факторів призводить до реалізації значної кількості експериментів. Крім того, при цьому не завжди коректним є припущення про можливість постійної стабілізації (підтримання на одному незмінному рівні) всіх змінних факторів за послідовного виділення одного з них, тобто неможливо визначити характер взаємодії факторів між собою і їх сумісний вплив на об’єкт дослідження.

Ефект взаємодії означає, що вплив одного фактора на параметр оптимізації залежить від того, яке значення приймають інші діючі фактори.

Вираз ***«планування експерименту»*** не означає організацію проведення експериментальних досліджень у загальноприйнятому розумінні, а передбачає виконання певного об’єму робіт за певними періодами часу.

Планування експерименту передбачає зміну всіх досліджуваних факторів за певним планом з урахуванням їх взаємодії.

Планування експерименту значно підвищує точність і зменшує об’єм експериментальних досліджень.

У загальному прикінцевою метою реалізації планованих факторних експериментів є розроблення математичних емпіричних моделей, які функціонально описують поведінку об’єкта дослідження залежно від зміни вхідних факторів.

В експериментальних дослідженнях об’єкт дослідження схематично можна уявити у вигляді так званого ***«чорного ящика»***, який може мати n $(X_1, X_2, ..., X_n)$ і k $(Z_1, Z_2, ..., Z_k)$ входів і m виходів $Y_1, Y_2, ..., Y_m$ (рис. 1).

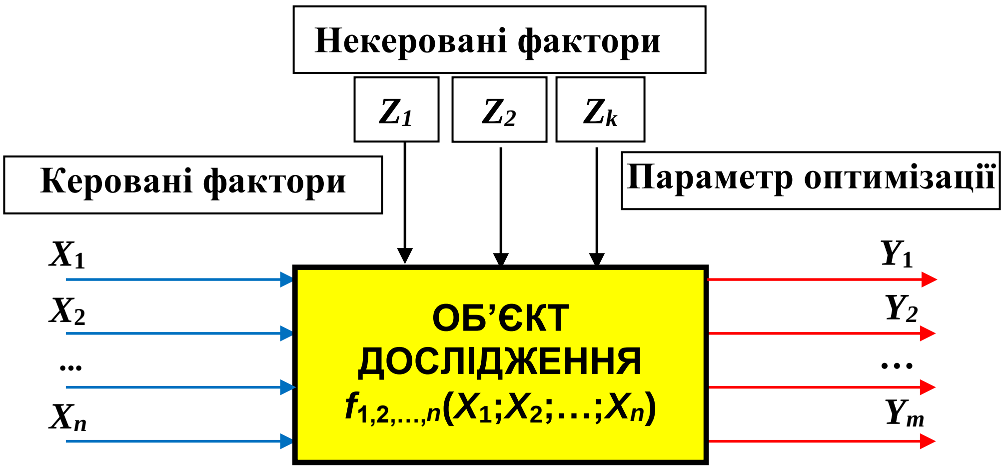


Рисунок 1 – Схема об’єкта дослідження («чорного ящика»)

Вхідними величинами можуть бути зовнішні, відносно об’єкта дослідження, впливи (некеровані фактори, наприклад, частота перемінного струму в мережі тощо, якими експериментатор не може керувати) та параметри самого об’єкта, або керовані фактори, наприклад, напруга живлення, величина струму тощо.

Вихідними величинами, або параметрами оптимізації можуть бути стани чи параметри – кількісні або якісні характеристики об’єкта. Наприклад, при дослідженні процесу роботи свердлильного верстата вхідними факторами можна приймати швидкість подачі різання та частоту обертання свердла тощо, а вихідними величинами, наприклад, сила свердління тощо.

Залежно від умов постановки задачі вихідна величина називається ***відгуком, функцією відгуку, параметром оптимізації***.

Як правило, аналітичний зв'язок між входом і виходом невідомий, а відомі тільки вхідні змінні незалежні фактори та визначені, або виміряні після проведення експерименту вихідні залежні величини.

Аналітичний зв'язок між входом і виходом, який описує взаємодію параметрів об'єкта дослідження, називають ***математичною моделлю*** об'єкта дослідження.

Моделювання – це процес представлення об'єкта дослідження (технічної системи) адекватною (подібною) йому технічною моделлю та подальшого проведення експериментів з цією моделлю для отримання інформації про об'єкт дослідження. При моделюванні математична модель є, як засіб, і як об'єкт дослідження, який знаходиться у відношеннях подібності до реального існуючого технічного об'єкта. Іншими словами, математична модель – це фізична або абстрактна система, яка адекватно уявляє собою технічний об'єкт дослідження.

Математична модель є абстрактною моделлю, яка описується мовою математичних співвідношень, законів, постулатів тощо, які відомі в галузі математики, фізики, хімії, природознавства тощо. Вона має форму функціональних залежностей між вхідними параметрами $\boldsymbol{f_{1,2,\dots,n}(X_1;X_2;\dots;X_n)}$, які задані експериментатором і вихідними функціями $\boldsymbol{Y}_1$, $Y_2,\dots,Y_n$, або відповідними концептуальними моделями. Якщо параметр оптимізація (вихідна функція) один, тоді функціональна залежність між вхідними параметрами $\boldsymbol{f_1(X_1;X_2;\dots;X_n)}$, які задані експериментатором і вихідною функцією $\boldsymbol{Y}_1$ відповідає концептуальній моделі, яку записано у загальному вигляді $\boldsymbol{Y}_1 = \boldsymbol{f_1(X_1;X_2;\dots;X_n)}$**.** Ці залежності конкретизують причинно-наслідкові зв'язки, які встановлено в концептуальній моделі, і характеризують їх кількісно.

При розроблені та дослідженні складних технічних систем застосовують методи аналітичного, чисельного, імітаційного, натурного й напівнатурного моделювання.

Аналітичні методи полягають у перетворенні символьної інформації, яка записана мовою математичного аналізу. При використанні аналітичних методів будується математична модель системи, що описує її фізичні властивості за допомогою математичних співвідношень, наприклад, у вигляді диференціальних або інтегральних рівнянь. Моделі такого типу називають аналітичними. Аналітична модель будується на основі понять, символіки і методів деякої теорії (наприклад, молекулярно-кінетичної теорії газів), яка визначає клас математичних аналогій, тобто основоположних математичних моделей. При аналітичному підході необхідні залежності виводяться з математичної моделі послідовним застосуванням математичних правил.

Натурним моделюванням називають проведення досліджень на реальному об'єкті (системі) з наступною обробкою результатів експерименту на основі теорії подібності. При функціонуванні технічної системи відповідно до поставленої мети вдається виявити закономірності проходження реального процесу. Необхідно відзначити, що такі різновидності натурного експерименту, як виробничий експеримент і комплексні випробування, мають високий ступінь достовірності. Методи натурного моделювання базуються на вимірюванні характеристик процесів, що проходять в реальних системах, й обробці результатів вимірювання з метою виявлення тих чи інших закономірностей, які цікавлять дослідника. Експериментальні дослідження дають найбільш достовірну інформацію, але вони носять частковий характер. Натурне моделювання може проводитися також на фізичних моделях, які моделюють реальні процеси.

Вихідну величину, що є невідомою функцією від зміни вхідних факторів, можна визначити як

$$Y = \varphi(x_1, x_2, ..., x_n). \quad (1)$$

У цьому випадку задачею планування експерименту може бути побудова математичної моделі об’єкта у вигляді аналітичного виразу

$$Y = f(X_1, X_2, ..., X_n), \quad (2)$$

який в області можливих значень факторів $X_1, X_2, ..., X_n$ досить точно збігається з невідомою залежністю (1).

У багатофакторному експерименті на об’єкт дослідження діє два і більше вхідних факторів, які в ході експерименту варіюють, тобто приймають різні значення своїх величин.

Значення величин, які фактори приймають при варіюванні в експерименті, називаються ***рівнями факторів***. Наприклад, в експерименті швидкість подачі різання знаходилась в межах V = 0,16 – 0,3 м/хв. У цьому випадку говорять, що фактор варіює на двох рівнях – менше значення ($V_н$ = 0,16 м/хв) називається нижнім рівнем варіювання фактора, а більше значення ($V_в$ = 0,3 м/хв) – верхнім рівнем варіювання фактора.

Фактори можуть варіювати як на двох, так і на трьох рівнях, при цьому третій рівень називають ***основним, або нульовим рівнем*** V_o. Нульовий рівень завжди приймає середнє значення суми верхнього та нижнього рівнів, тобто V_o = ($V_н$ + $V_в$) / 2 = (0,16 + 0,3)/2 = 0,46 / 2 = 0,23.

Експеримент, в якому реалізуються можливі комбінації рівнів факторів, називається ***планованим факторним експериментом*** і позначають як ПФЕ P^K, де: P – кількість рівнів варіювання фактором; k – кількість діючих факторів в експерименті, наприклад, ПФЕ 2^2 – двофакторний планований експеримент на двох рівнях варіювання факторами.

Побудова математичної моделі об'єкта за результатами реалізації планованого факторного експерименту здійснюється наступним чином.

Кількість дослідів N при плануванні ПФЕ у загальному випадку визначають за формулою

$$N = P^k . \tag{3}$$

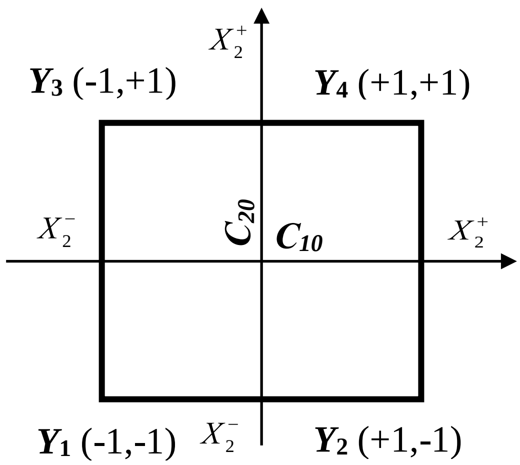


Рисунок 2 – Результати двофакторного експерименту для кодованих факторів

Взявши, наприклад, варіант, коли на об'єкт дослідження діють два фактори, які варіюють на двох рівнях, тобто двофакторний експеримент на двох рівнях варіювання факторами, нижньому рівні, який позначається знаком (– 1) та верхньому, який позначається знаком (+ 1), можемо записати: C_1^-, C_2^- – нижній рівень факторів, C_1^+, C_2^+ – верхній рівень факторів, де C_1, C_2 – діючі в експерименті змінні фактори в натуральних (фізичних) величинах.

При відтворені експерименту з комбінацією двох факторів на двох рівнях їх варіювання кількість дослідів згідно з формулою (2) буде дорівнювати $N = 2^2 = 4$.

При кожному можливому варіанті набору змінних факторів (у кожному окремому досліді) вимірюється значення функції відгуку **Y**, у результаті чого отримуємо чотири точки значень відгуку, які позначено на рис. 2: **Y_1**(-1, -1); **Y_2**(+1, -1); **Y_3**(+1, +1); **Y_4**(-1,+1).

Для кожного експерименту необхідно вибрати ***інтервал варіювання***, яким називається половина різниці між більшим (верхнім) і меншим (нижнім) значенням фактора. Значення кожного фактора C_1 і C_2 в центрі області експерименту називається його ***основним рівнем*** і помічається індексом (0), тобто **C_{10}** та **C_{20}** (рис. 2).

Для зручності запису плану експерименту і обробки експериментальних даних користуються кодованими значеннями факторів, які є безрозмірними величинами.

Кодовані і фізичні (натуральні) змінні зв'язані між собою співвідношенням

$$X_i = (C_i - C_{io}) / \lambda_i , \tag{4}$$

де C_i – значення фактора на одному з рівнів;

C_{io} – значення фактора на основному рівні;

λ_i – інтервал варіювання фактора.

Кодування факторів рівнозначно переносу початку координат у точку основного рівня факторів і зміні масштабу. Верхній рівень фактора позначають (+1), а нижній – (-1).

Тоді план експерименту можемо представити у вигляді плану-матриці, яку подано в табл. 1.

План експерименту для кодованих факторів задається графами 3 і 4. Перший рядок відповідає точці 1 (рис. 2) з координатами (-1, -1). У ньому записані значення координат $\boldsymbol{X_1}$ і $\boldsymbol{X_2}$ та результати вимірювання відгуку $\boldsymbol{Y}$ в точці $\boldsymbol{Y_1}$.

Для проведення повного факторного експерименту значення факторів розділяються на рівні, які задаються відповідними рядками плану. Після цього вимірюються значення відгуку.

Для підвищення точності серії вимірів можуть повторюватись кілька разів (паралельні досліди), результати яких наведені в табл. 2.

Таблиця 1 – План – матриця двофакторного експерименту типу ПФЕ2^2

Номер досліду	Рівні факторів				Відгук
U*	C_1	***(X_1)***	C_2	***(X_2)***	$\boldsymbol{Y}$
1	C_1^-	(-1)	C_2^-	(-1)	$\boldsymbol{Y_1}$
2	C_1^+	(+1)	C_2^-	(-1)	$\boldsymbol{Y_2}$
3	C_1^-	(-1)	C_2^+	(+1)	$\boldsymbol{Y_3}$
4	C_1^+	(+1)	C_2^+	(+1)	$\boldsymbol{Y_4}$

*U = 1, 2, 3,…, N.

Вибір математичної моделі об’єкта – неформалізований етап, який часто базується на інтуїтивних міркуваннях з урахуванням попередніх досліджень експериментатора.

Числові значення коефіцієнтів рівняння регресії вибраного емпіричного виразу визначають за результатами обробки отриманих експериментальних даних.

Математична модель об’єкта дослідження, тобто рівняння регресії, визначається змінними X_i та постійними параметрами (коефіцієнтами) β_i і в загальному випадку має вигляд

$$y = Y_U(x_1, x_2, ..., x_k, \beta_1, \beta_2, ..., \beta_r). \tag{5}$$

Таблиця 2 – Результати паралельних дослідів

Номер досліду	Рівні факторів		Відгук		
U	X_1	X_2	Y_{u1}	Y_{u2} …	Y_{un}
1	-1	-1	Y_{11}	Y_{12}	Y_{1n}
2	+1	-1	Y_{21}	Y_{22}	Y_{2n}
3	-1	+1	Y_{31}	Y_{32}	Y_{3n}
4	+1	+1	Y_{41}	Y_{42}	Y_{4n}

Модель відносно змінних X_i і параметрів β_i може мати лінійну залежність

$$y = \beta_o + \beta_1 x_1 + ... + \beta_k x_k \,, \tag{6}$$

а бути також нелінійною, тобто

$$y = \beta_o + \beta_1 x_1 + ... + \beta_k x_k + \beta_{12} x_1 x_2 + ... + \beta_{k+1} x_{k+1} x_k \,. \tag{7}$$

При визначенні коефіцієнтів моделі може вводитись фіктивна змінна x_o = +1 та так званий середній (нульовий) рівень варіювання фактора. У цьому випадку говорять, що фактори змінюються на трьох рівнях варіювання.

З урахуванням цього план-матриця, наприклад ПФЕ 3^2 на трьох рівнях варіювання факторів, тобто на основному (нульовому), верхньому та нижньому, буде матиме вигляд (табл. 3).

Таблиця 3 – План-матриця двофакторного експерименту на трьох рівнях варіювання факторів

Номер досліду	Фіктивна змінна	Рівні факторів		Відгук		
U	X_0	X_1	X_2	Y_{U1}	Y_{U2}	Y_{UN}
1	+1	-1	-1	Y_{11}	Y_{12}	Y_{1N}
2	+1	+1	-1	Y_{21}	Y_{22}	Y_{2N}
3	+1	0	-1	Y_{31}	Y_{32}	Y_{3N}
4	+1	-1	+1	Y_{41}	Y_{42}	Y_{4N}
5	+1	+1	+1	Y_{51}	Y_{52}	Y_{5N}
6	+1	0	+1	Y_{61}	Y_{62}	Y_{6N}
7	+1	-1	0	Y_{71}	Y_{72}	Y_{7N}
8	+1	+1	0	Y_{81}	Y_{82}	Y_{8N}
9	+1	0	0	Y_{91}	Y_{92}	Y_{9N}

Тоді математичну модель, наприклад виразу рівняння (6), можна записати у вигляді

$$Y = \beta_o x_o + \beta_1 x_1 + ... + \beta_k x_k. \tag{8}$$

Оскільки величини Y_U випадкові, то одержують неточне рівняння моделі, а його оцінку

$$Y = b_o x_o + b_1 x_1 + ... + b_k x_k. \tag{9}$$

У випадку ПФЕ 2^2, враховуючи взаємодію факторів (нелінійність типу добутку факторів), можна записати

$$Y = b_o x_o + b_1 x_1 + b_2 x_2 + b_{12} x_1 x_2, \tag{10}$$

де $b_1 x_1$, $b_2 x_2$ – члени, які характеризують вплив окремих факторів; $b_{12} x_1 x_2$ – вплив взаємодії факторів.

Ефект взаємодії означає, що вплив одного фактора, наприклад x_1, на вихідну величину (параметр оптимізації) Y залежить від того, яке значення приймають інші фактори.

За аналогією з формулою (10) математична модель об’єкта у випадку експерименту типу ПФЕ2^3 матиме вигляд

$$Y = b_o x_o + b_1 x_1 + b_1 x_2 + b_3 x_3 + b_{12} x_1 x_2 + b_{13} x_1 x_3 + b_{23} x_2 x_3. \tag{11}$$

Для ПФЕ 2^2, виходячи з моделі (10) та матриці планування з таблиці 3, можемо одержати:

$$Y_1 = b_o(+1) + b_1(-1) + b_2(-1) + b_{12}(+1);$$
$$Y_2 = b_o(+1) + b_1(+1) + b_2(-1) + b_{12}(-1);$$
$$Y_3 = b_o(+1) + b_1(-1) + b_2(+1) + b_{12}(-1);\ Y_4 = b_o(+1) + b_1(+1) + b_2(+1) + b_{12}(+1), \text{ або}$$
$$Y_1 = b_o - b_1 - b_2 + b_{12};$$
$$Y_2 = b_o + b_1 - b_2 - b_{12};$$
$$Y_3 = b_o - b_1 + b_2 - b_{12};$$
$$Y_4 = b_o + b_1 + b_2 + b_{12};$$
$$b_o = (Y_1 + Y_2 + Y_3 + Y_4)/4;$$
$$b_1 = (-Y_1 + Y_2 - Y_3 + Y_4)/4;$$
$$b_2 = (-Y_1 - Y_2 + Y_3 + Y_4)/4;$$
$$b_{12} = (Y_1 - Y_2 - Y_3 + Y_4)/4.$$

У загальному вигляді можна записати:

$$b_i = \frac{\sum_{u=1}^{N} x_{iu}\bar{y}_u}{\sum_{u=1}^{N} x_{iu}^2} = \frac{\sum_{u=1}^{N} x_{iu}\bar{y}_u}{N}; \quad (12)$$

$$b_{ij} = \frac{\sum_{u=1}^{N} x_{iu}x_{ju}\bar{y}_u}{N}, \quad (13)$$

де x_{iu} – значення кодованої змінної у відповідному стовпці плану експерименту; $\bar{y}_u$ – середній результат u-го досліду; u – порядковий номер досліду; i – номер фактора; j – номер фактора, відмінного від i-го.

Для експерименту типу ПФЕ 2^3 при вибраній математичній моделі згідно з залежністю (11) невідомі коефіцієнти рівняння регресії b_i, b_{ij}, і т.д. визначаються аналогічно формулам (12) та (13), а коефіцієнти b_{ijk} – за формулою

$$b_{ijk} = \frac{\sum_{u=1}^{N} x_{iu}x_{ju}x_{ku}\bar{y}_u}{N}. \quad (14)$$

Визначивши всі коефіцієнти регресії b_i і підставивши їх у відповідне рівняння, одержимо математичну модель досліджуваного об’єкта.

Після одержання математичної моделі об’єкта необхідно визначити статистичну значимість коефіцієнтів рівняння регресії (тобто можливість прийняття значення деякого коефіцієнта b_i рівним нулю), а також перевірити адекватність (відповідність) моделі реальному об’єкту.

Статистична значимість коефіцієнтів рівняння регресії b_i визначається у такій послідовності:

- дисперсія похибок дослідів у рядках плану ПФЕ

$$S_u^2 = \frac{1}{n-1}\sum_{j=1}^{n}(y_{ui} - \bar{y}_u)^2, \quad (15)$$

де n – кількість паралельних дослідів (повторності одного експерименту); j =1, 2,..., n.

- дисперсія відтворення досліду

$$S_y^2 = \frac{1}{N}\sum_{u=1}^{N} S_u^2, \quad (16)$$

- похибка відтворення

$$S_y = \sqrt{S_y^3}, \quad (17)$$

- умова значимості коефіцієнтів рівняння регресії

$$b_{i(jk)} > \frac{t_T S_y}{\sqrt{Nn}}, \tag{18}$$

де t_T – табличне значення коефіцієнта Ст'юдента, яке вибирається з таблиці залежно від ступеня відповідності f та рівня значимості α.

Ступень відповідності визначається

$$f = (n-g)N, \tag{19}$$

де g – число коефіцієнтів в рівнянні регресії.

Якщо умова значимості (18) не виконується, то такий коефіцієнт b_i рівняння регресії можна прийняти рівним нулю, а відповідний член x_i рівняння регресії виключити.

Перевірка адекватності, або рівня точності опису отриманою математичною моделлю реального процесу, або встановленої точності відхилення отриманих даних параметра оптимізації за рівнянням регресії від виміряного експериментального значення, здійснюється таким чином:

- визначається дисперсія адекватності

$$S_{ag}^2 = \frac{n}{N-g'} \sum_{u=1}^{N} (\bar{y}_u - \tilde{y}_u)^2, \tag{20}$$

де $N-g'$ – число степенів вільності дисперсії адекватності; g' – число значимих коефіцієнтів в рівнянні регресії; $\bar{y}_u$ – середнє значення відгуку в u-му досліді; $\tilde{y}_u$ – значення відгуку в u-й точці плану, обчислене за рівнянням регресії;

- визначається розрахунковий критерій відповідності Фішера F_p

$$F_p = \frac{S_{ag}^2}{S_y^2}, \tag{20}$$

де S_y^2 – дисперсія відтворення досліду (16);

- визначається табличний критерій Фішера F_T за заданим рівнем значимості α і двома степенями відповідності $f_{ag} = N - g$ та $f_y = N(n-1)$.

Умова адекватності

$$F_p < F_T. \tag{21}$$

<u>Порядок виконання роботи.</u>

<u>Приклад</u>. Провести планування трифакторного експерименту типу ПФЕ 2^3 та статистичну обробку результатів експериментів з визначення продуктивності шнекового конвеєра Q залежно від зміни трьох вхідних факторів, тобто розробити математичну емпіричну модель процесу на основі регресійного аналізу, при цьому визначити вид емпіричної моделі та її коефіцієнти регресії.

Діючі вхідні фактори експерименту:

- крок гвинтової лінії T;
- кутова швидкість шнека ω;
- діаметр шнека D.

Відгук (параметр оптимізації), тобто зміну продуктивності шнекового конвеєра залежно від вхідних факторів (кроку гвинтової лінії T; кутової швидкості шнека ω; діаметра шнека D) оцінюється секундною подачею, кг/с.

Характеристика вхідних змінних факторів та числове значення їх рівнів варіювання (нижнього та верхнього) наведено в табл. 4.

Таблиця 4 – Характеристика факторів та значення рівнів варіювання

№ фактора	Найменування	Розмірність	Рівнів факторів
1	Крок гвинтової лінії T, см	$\boldsymbol{x_1}$	2,0(-1) - 10,0(+1)
2	Кутова швидкість шнека ω, рад/с	$\boldsymbol{x_2}$	1,0(-1) - 6,0(+1)
3	діаметр шнека D, м	$\boldsymbol{x_3}$	0,3(-1) - 0,8(+1)

1. Рівняння математичної моделі об’єкта дослідження приймаємо у вигляді

$$Q = b_o x_o + b_1 x_1 + b_1 x_2 + b_3 x_3 + b_{12} x_1 x_2 + b_{13} x_1 x_3 + b_{23} x_2 x_3 . \qquad (22)$$

2. Об’єкт досліджується за допомогою трифакторного експерименту на двох рівнях варіювання факторами. Результати проведених досліджень наведені в табл. 5.

У цьому випадку, згідно з формулою (3), кількість дослідів $\boldsymbol{N} = 2^3 = 8$.

Таблиця 5. Результати багатофакторного експерименту

Номер досліду	Рівні факторів				Відгук, $\boldsymbol{Q}$			Середнє значення
					Повторюваність досліду, $\boldsymbol{n}$			
$\boldsymbol{u}$	$\boldsymbol{x_0}$	$\boldsymbol{x_1}$	$\boldsymbol{x_2}$	$\boldsymbol{x_3}$	**1**	**2**	**3**	$\boldsymbol{Q_{сер}}$

1	+	-	-	-	7.3	8.2	7.9	7.8
2	+	+	-	-	8.5	8.2	8.3	8.3
3	+	-	+	-	8.6	8.5	8.9	8.7
4	+	+	+	-	11.2	10.9	11.8	11.3
5	+	-	-	+	9.8	8.3	9.44	9.2
6	+	+	-	+	17.7	22.3	19.2	19.7
7	+	-	+	+	20.2	23.8	22.2	22.1
8	+	+	+	+	37.3	39.9	39.2	38.8

Кожен порядковий номер експерименту (досліду) повторюється 3 рази. Тому для кожної точки факторного простору одержано 3 значення відгуку *Q*.

3. За формулами (12), (13) і (14) обчислюємо коефіцієнти рівняння регресії:

$$b_0 = \frac{7.8+8.3+8.7+11.3+9.2+19.7+22.1+38.8}{8} = \frac{125.9}{8} = 15.74;$$

$$b_1 = \frac{-7.8+8.3-8.7+11.3-9.2+19.7-22.1+38.8}{8} = \frac{30.3}{8} = 3.79;$$

$$b_2 = \frac{-7.8-8.3+8.7+11.3-9.2-19.7+22.1+38.8}{8} = \frac{35.9}{8} = 4.49;$$

$$b_3 = \frac{-7.8-8.3-8.7-11.3+9.2+19.7+22.1+38.8}{8} = \frac{53.7}{8} = 6.71;$$

$$b_{12} = \frac{7.8-8.3-8.7+11.3+9.2-19.7-22.1+38.8}{8} = \frac{8.3}{8} = 1.04;$$

$$b_{13} = \frac{7.8-8.3+8.7-11.3-9.2+19.7-22.1+38.8}{8} = \frac{24.1}{8} = 3.01;$$

$$b_{23} = \frac{7.8+8.3-8.7-11.3-9.2-19.7+22.1+38.8}{8} = \frac{28.1}{8} = 3.51.$$

Таким чином, рівняння регресії, у кодованих величинах буде мати вигляд

$$Q = 15{,}74 + 3{,}79x_1 + 4{,}49x_2 + 6{,}71x_3 + 1{,}04x_1x_2 + 3{,}01x_1x_3 + 3{,}51x_2x_3. \qquad (23)$$

4. Визначається статистична значимість коефіцієнтів рівняння регресії:

- дисперсія похибок дослідів у рядках плану ПФЕ згідно з формулою (15), де: n = 3 – кількість паралельних дослідів:

$$S_1^2 = \frac{1}{3-1}\left[(7.3-7.8)^2 + (8.2-7.8)^2 + (7.9-7.8)^2\right] = \frac{1}{2}(0.25+0.16+0.01) = 0.21;$$

$$S_2^2 = 0.025;\ S_3^2 = 0.045;\ S_4^2 = 0.21;\ S_5^2 = 0.875;$$

$$S_6^2 = 5.505;\ S_7^2 = 3.255;\ S_8^2 = 1.855;$$

- дисперсія відтворення досліду згідно з формулою (16)

$$S_y^2 = 1/8(0{,}21+0{,}025+0{,}045+0{,}21+0{,}875+5{,}505+3{,}255+1{,}855) = \frac{1}{8}\cdot 11{,}98 = 1{,}497;$$

- похибка відтворення за формулою (17)

$$S_y = \sqrt{S_y^3} = \sqrt{1{,}497} = 1{,}223;$$

- умова значимості коефіцієнтів рівняння регресії перевіряється за умовою відповідності (18) де критерій відповідності t_T визначається із таблиці залежно від (19), де f = (3-1)* 8 = 16 та рівня значимості α = 0,05.

При цьому отримали табличне значення рівня t_T = 2,12.

Таким чином $\dfrac{t_T S_y}{\sqrt{Nn}} = \dfrac{2{,}12\cdot 1{,}223}{\sqrt{8\cdot 3}} = 0{,}53$.

Тоді: b_o = 15,74 > 0.53 – значимий; b_1 = 3,79 > 0.53 – значимий;

b_2 = 4,49 > 0.53 – значимий; b_3 = 6,71 > 0.53 – значимий;

b_{12} = 1,04 > 0.53 – значимий; b_{13} = 3,01 > 0.53 – значимий;

b_{23} = 3,51 > 0.53 – значимий.

5. Остаточно рівняння регресії прийнятої математичної моделі об'єкта дослідження матиме вигляд

$$Y = 15{,}74 + 3{,}79x_1 + 4{,}49x_2 + 6{,}71x_3 + 1{,}04x_1x_2 + 3{,}01x_1x_3 + 3{,}51x_2x_3.$$

6. Перевірка адекватності моделі:

- визначається дисперсія адекватності за формулою (20) де $\tilde{y}_1$ – обчислюється шляхом підстановки в остаточне рівняння регресії верхнього та нижнього значень рівнів вхідних змінних факторів (± 1) з плану-матриці ПФЕ2^3

$$\tilde{y}_1 = 15{,}74 - 3{,}79 - 4{,}49 - 6{,}71 + 1{,}04 + 3{,}01 + 3{,}51 = 8{,}31;$$

$$\tilde{y}_2 = 7{,}79;\ \tilde{y}_3 = 8.19;\ \tilde{y}_4 = 11.83;\ \tilde{y}_5 = 8.69;\ \tilde{y}_6 = 20.21;\ \tilde{y}_7 = 22.61;\ \tilde{y}_8 = 38.29;$$

$$S_{ag}^2 = \frac{3}{8-7}\begin{bmatrix}(7.8-8.31)^2 + (8.3-7.79)^2 + (8.7-8.19)^2 + (11.3-11.83)^2 + \\ +(9.2-8.69)^2 + (19.7-20.21)^2 + (22.2-22.61)^2 + (38.8-38.29)^2\end{bmatrix} =$$

$$= \frac{3}{8-7}\cdot 2.1 = 3\cdot 2.1 = 6.3;$$

- визначається розрахунковий критерій Фішера F_p за формулою (20)

$$F_p = \frac{S_{ag}^2}{S_y^2} = \frac{6.3}{1.497} = 4{,}21;$$

- визначається наближене значення критерію Фішера зі статистичної таблиці за умови $f_{ag} = N - g$ = 8 –7 = 1, $f_y = (n-1)N$ = 16 та рівні значущості α = 0,05. При цьому F_T = 4,5.

Умова адекватності F_p < F_T. Маємо F_p = 4.21 < F_T = 4,5, тобто вибрана математична модель адекватно (рівнозначно) описує поведінку досліджуваного об’єкта.

7. Перехід від кодованих (***x*₁, *x*₂, *x*₃**) величин вхідних факторів до їх натуральних значень (*T*; ω; *D*), або $x_1 \to T$, $x_2 \to \omega$, $x_3 \to D$ проводять за формулами:

$$x_i = \frac{X_i - X_{i0}}{\Delta X_i}, \tag{24}$$

де X_{i0} – натуральне значення *i*-го фактору на нульовому рівні;

ΔX_i – інтервал варіювання *i*-го фактору;

- нульовий рівень, або середнє значення кожного вхідного фактора, визначають за формулою

$$X_0 = \frac{X_{max} + X_{min}}{2}, \tag{25}$$

де X_0, X_{max}, X_{min} – відповідно, числове значення нульового, верхнього та нижнього рівня вхідного фактора;

- визначають інтервали варіювання факторами

$$\Delta X_i = \frac{X_{max} - X_{min}}{2}; \tag{26}$$

- у рівняння регресії (23) підставляють значення $\boldsymbol{x_i}$, які визначено згідно з рівнянням (24) – (26)

$$\begin{gathered} Q = 15{,}74 + 3{,}79\left(\frac{T-T_0}{\Delta T}\right) + 4{,}49\left(\frac{\omega-\omega_0}{\Delta\omega}\right) + 6{,}71\left(\frac{D-D_0}{\Delta D}\right) + \\ +1{,}04\left(\frac{T-T_0}{\Delta T}\right)\left(\frac{\omega-\omega_0}{\Delta\omega}\right) + 3{,}01\left(\frac{T-T_0}{\Delta T}\right)\left(\frac{D-D_0}{\Delta D}\right) + 3{,}51\left(\frac{\omega-\omega_0}{\Delta\omega}\right)\left(\frac{D-D_0}{\Delta D}\right) \end{gathered}; \tag{27}$$

- проводять перетворення та спрощення рівняння (27) та в кінцевому випадку отримують рівняння регресії, або емпіричну модель у натуральних величинах факторів.

8. Вихідні дані для проведення статистичних розрахунків взяти згідно з варіантом табл. 6.

9. Побудова 3D графічного відтворення одержаних результатів експериментальних досліджень у вигляді поверхні відгуку та її двомірного

перерізу з використанням прикладної програми для ПК Statistica:

- складається план-матриця планованого двофакторного експерименту типу ПФЕ 2^3 згідно з табл. 5.

- у графи план-матриці заносяться вихідні дані, які наведено згідно з варіантом завдання в табл. 6;

- будується поверхня відгуку та її двомірний переріз з застосуванням пакету прикладної програми для ПК.

- приклад побудованої поверхні відгуку та двомірного перерізу поверхні відгуку наведено на рис. 4.

Таблиця 5 – Вихідні дані план-матриці експерименту типу ПФЕ 2^3

	Var1	**Var2**	**Var3**	**Var4**
1	2	1	0,3	7,8
2	10	1	0,3	8,3
3	2	6	0,3	8,7
4	10	6	0,3	11,3
5	2	1	8	9,2
6	10	1	8	19,7
7	2	6	8	22,1
8	10	6	8	38,8

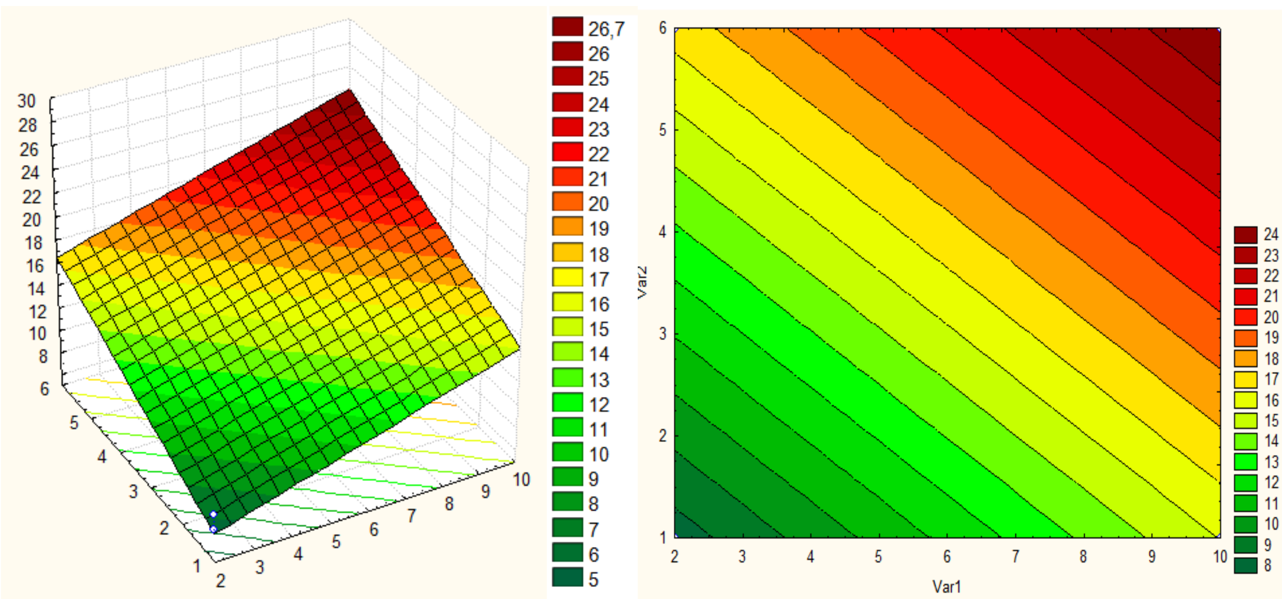


Рисунок 4 – Поверхня відгуку та двомірний переріз поверхні відгуку зміни продуктивності шнекового конвеєра як функція $Q = f(T_1;\omega)$

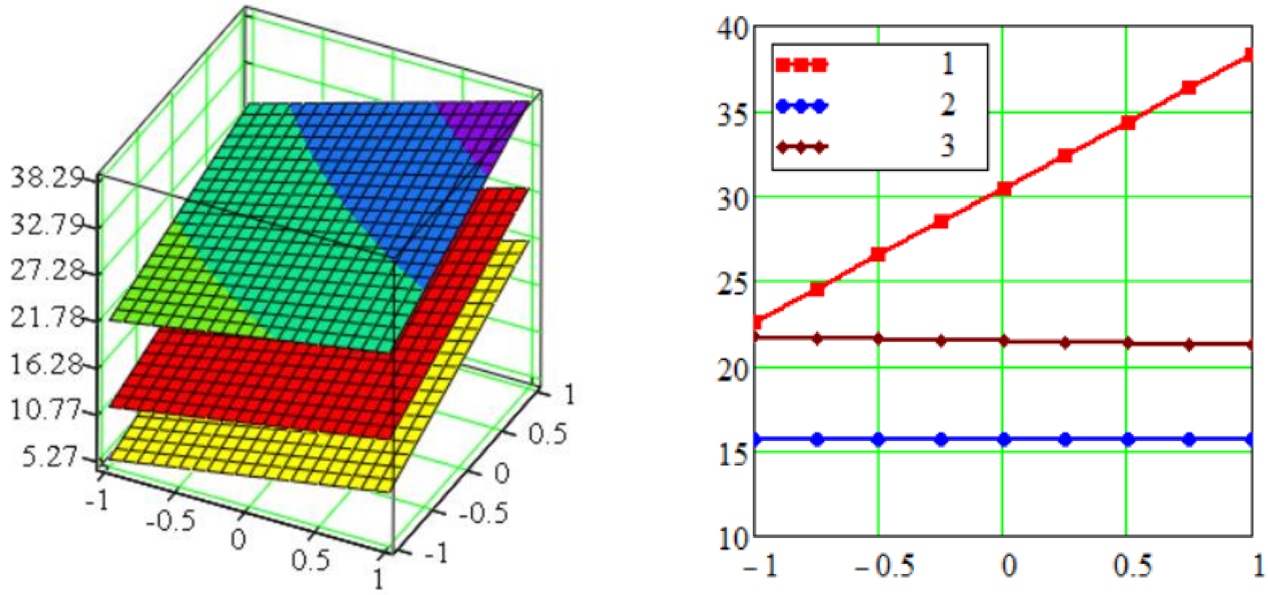
10. Побудова одержаних результатів експериментальних досліджень за

рівнянням регресії

$$Y = 15{,}74 + 3{,}79x_1 + 4{,}49x_2 + 6{,}71x_3 + 1{,}04x_1x_2 + 3{,}01x_1x_3 + 3{,}51x_2x_3$$

у вигляді 3D і 2D графічних залежностей з використанням прикладної програми для ПК MathCad.

Приклад побудованого графічного відтворення за рівнянням регресії $Y = 15{,}74 + 3{,}79x_1 + 4{,}49x_2 + 6{,}71x_3 + 1{,}04x_1x_2 + 3{,}01x_1x_3 + 3{,}51x_2x_3$ у вигляді 3D і 2D графічних залежностей з використанням прикладної програми для ПК MathCad



наведено на рис. 5.

Рисунок 5 – 3D і 2D графічні залежності

Контрольні запитання

1. Дати визначення поняттю “планування експерименту”.
2. Що таке математична модель об’єкта дослідження?
3. Поняття діючих факторів і рівнів їх варіювання.
4. Кодоване позначення факторів і натуральне значення факторів.
5. Що таке план-матриця експерименту?
6. Як визначається значимість коефіцієнтів регресії?
7. Поняття адекватності математичної моделі.
8. Критерії Ст’юдента і Фішера та їх визначення.
9. Дати визначення поняттю “емпірична модель”.
10. За якою формулою визначають загальну кількість експериментів?

Таблиця 6 – Варіанти вихідних даних для проведення розрахунків

В. 1			**В. 2**			**В. 3**		
1	2	3	1	2	3	1	2	3
7,2	8,1	7,1	9,1	7,9	7,9	17,8	16,9	17,1
8,1	8,2	7,9	12,6	15,8	14,3	13,1	13,1	14,2
8,3	8,5	8,4	8,4	8,9	9,1	8,7	9,1	9,1
10,9	9,9	12,1	12,5	9,8	11,7	12,5	11,2	12,4
8,9	8,3	9,1	9,8	9,2	9,8	9,1	9,2	9,0
16,9	22,3	19,1	20,2	20,1	20,4	17,3	18,1	17,6
20,3	23,7	22,1	17,6	18,1	18,3	17,2	17,9	17,9
35,3	39,9	37,4	23,2	24,7	23,9	15,4	15,8	15,5
В. 4			**В. 5**			**В. 6**		
1	2	3	1	2	3	1	2	3
6,9	8,2	7,3	13,1	14,2	13,9	6,3	7,2	6,9
8,5	8,2	7,9	10,7	11,8	15,4	7,5	7,3	8,1
8,7	8,1	8,3	9,3	9,4	9,5	7,6	7,5	7,9
12,2	14,1	13,9	11,4	13,2	12,8	10,2	9,9	10,8
5,5	5,9	6,1	9,1	10,3	9,9	9,1	8,3	9,3
10,2	11,4	13,7	17,1	16,4	15,8	15,4	16,2	15,9
18,5	19,2	18,9	20,1	22,4	22,1	19,8	19,1	20,3
18,9	19,3	19,1	32,3	31,9	33,1	25,7	29,4	24,9
В. 7			**В. 8**			**В. 9**		
1	2	3	1	2	3	1	2	3
15,2	14,8	13,9	10,1	9,9	9,3	12,9	12,1	12,6
7,6	7,4	7,9	9,2	9,0	9,0	15,8	15,4	16,1
9,3	8,5	8,7	11,7	11,5	11,4	10,3	10,2	9,9
10,1	11,3	11,5	9,8	9,3	10,1	11,4	11,5	11,3
7,1	7,2	7,4	13,4	13,9	14,0	9,7	8,7	9,1
19,2	19,3	19,0	18,5	19,1	18,9	13,4	13,5	13,4
17,4	18,5	17,9	26,1	25,9	24,9	17,8	19,1	18,5
21,2	23,7	24,1	31,3	32,1	31,2	27,4	26,9	26,9
В. 10			**В. 11**			**В. 12**		
1	2	3	1	2	3	1	2	3
7,2	8,7	7,9	9,4	9,2	9,4	13,7	12,1	12,1
9,1	8,3	8,5	9,1	8,9	8,7	8,4	9,5	9,5
7,6	8,3	7,9	5,1	5,4	5,5	7,6	7,8	7,6
11,4	9,9	10,8	9,9	9,9	10,3	8,1	8,5	8,5
9,7	8,2	9,3	9,8	9,7	9,5	10,2	9,9	11,7
15,4	13,7	13,9	11,9	12,5	11,8	15,5	16,1	15,9
20,1	23,7	22,4	13,3	14,2	14,1	17,2	17,2	16,9
В. 13			В. 14			В. 15		
1	1	2	1	2	1	2	2	3

8,4	7,9	8,0	7,9	8,0	7,9	8,0	12,1	12,1
9,1	9,1	7,4	9,1	7,4	9,1	7,4	9,5	9,5
8,6	8,8	8,1	8,8	8,1	8,8	8,1	7,8	7,6
10,1	10,4	11,7	10,4	11,7	10,4	11,7	8,5	8,5
8,7	9,9	9,8	9,9	9,8	9,9	9,8	9,9	11,7
13,3	18,1	20,1	18,1	20,1	18,1	20,1	16,1	15,9
18,1	19,7	22,5	19,7	22,5	19,7	22,5	17,2	16,9

ПРАКТИЧНА РОБОТА № 2

ВСТАНОВЛЕННЯ ВИГЛЯДУ ФУНКЦІОНАЛЬНОЇ ЕМПІРИЧНОЇ ЗАЛЕЖНОСТІ МІЖ ДВОМА ПЕРЕМІННИМИ ВЕЛИЧИНАМИ

Мета роботи: вивчити порядок визначення та підбору вигляду емпіричної залежності між двома змінними величинами отриманого експериментального масиву даних проведеного класичного експерименту.

Теоретичні відомості

При обробці результатів однофакторного (класичного або пасивного) експерименту, тобто отриманого масиву експериментальних даних, іноді на першому етапі необхідно визначити (встановити) та підібрати ***вигляд емпіричної залежності*** $Y_{\text{T}} = f(x)$ між двома перемінними експериментальними величинами (незалежною змінною ***X*** – вхідним фактором і залежною змінною ***Y*** – параметром оптимізації).

На другому етапі необхідно провести вирівнювання побудованої експериментальної (емпіричної) кривої за найбільш близькою до неї теоретичною кривою.

Розглянемо випадок, коли експеримент проводиться з метою встановлення вигляду функції щільності вірогідності.

Обробка результатів експерименту проводиться в такій послідовності:

- за дослідними даними будується емпірична крива;

- визначаються параметри емпіричного розподілення;

- висувається одна або декілька гіпотез про функцію щільності досліджувальної випадкової величини, виходячи із зовнішнього вигляду експериментальної (емпіричної) кривої, із значень її параметрів і технологічних факторів, які впливають на її вигляд;

- емпірична крива вирівнюється за одною або послідовно за декількома прийнятими теоретичними кривими;

- проводиться порівняння за одним із критеріїв узгодження емпіричної і теоретичної (вирівняної емпіричної) кривих;

- визначається і вибирається аналітична алгебраїчна функція, за якою будемо мати найкраще узгодження.

Відомо, що число теоретичних кривих щільності дуже велике. Тому при виборі підходящої функції необхідно в результат первинного аналізу звести до мінімуму число можливих теоретичних кривих.

Цій меті можуть служити знання умов або ознак, при яких необхідно чекати появи найбільш поширених в техніці функцій.

Вигляд залежності між двома перемінними величинами отриманого експериментального масиву даних, тобто незалежною вхідної змінною X – вхідним фактором і залежною змінною Y, яку вимірюють під час проведення експерименту, або параметром оптимізації може бути ***функціональною*** або ***стохастичною***.

Функціонально залежними є такі величини, у яких кожному значенню однієї величини відповідає цілком певне (як правило всього одне) значення другої величини, наприклад залежність зміни тиску від температури, величини зносу лемеша від часу або структури ґрунту і т.п.

Стохастично залежними називаються такі величини, у яких різним значенням однієї величини відповідають різні закони розподілення другої. Окремим випадком стохастичної залежності є корелятивна залежність, яка виникає в тому випадку, коли кожному значенню однієї величини відповідають різні середні значення другої величини, наприклад залежність зміни продуктивності шнекового конвеєра від зміни частоти обертання шнека, або діаметра, або кроку шнека тощо.

Для встановлення вигляду функціональної залежності класичний експеримент проводять таким чином, що для кожного значення одного вхідного фактора $\boldsymbol{X}$ визначається значення параметра оптимізації $\boldsymbol{Y}$, а результати заносять в табл. 1.

Таблиця 1 – Результати проведеного однофакторного експерименту

Кількість проведених дослідів, N	1	2	3	4	...	N
Вхідний фактор $\boldsymbol{X_i}$	$\boldsymbol{X_1}$	$\boldsymbol{X_2}$	$\boldsymbol{X_3}$	$\boldsymbol{X_4}$	**…**	$\boldsymbol{X_N}$
Параметр оптимізації $\boldsymbol{Y_i}$	$\boldsymbol{Y_1}$	$\boldsymbol{Y_2}$	$\boldsymbol{Y_3}$	$\boldsymbol{Y_4}$	**…**	$\boldsymbol{Y_N}$

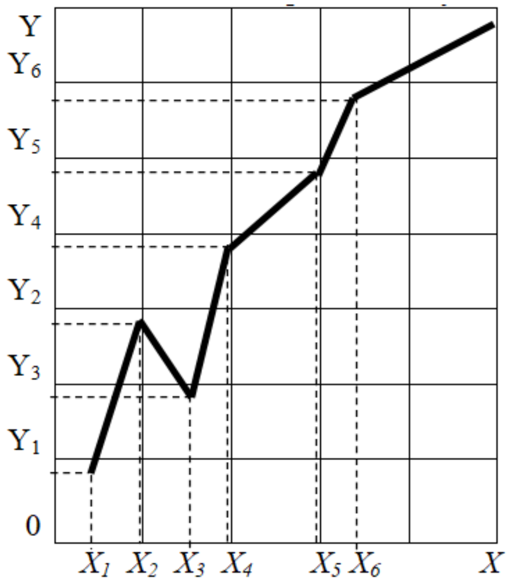


Рисунок 1 – Експериментальна ломана залежність

За цими даними будують графік експериментальної залежності між величинами ***X*** і ***Y*** у вигляді ломаної прямої лінії, приклад побудови якого наведено на рис. 1.

Отриману ломану лінію необхідно вирівняти за найбільш близькою до неї теоретичною кривою. Подібний алгебраїчний вираз, тобто математична залежність, називається ***емпіричною залежністю*** або ***формулою***, а сам процес підбору теоретичної залежності і вирівнювання побудованої експериментальної ломаної лінії за підібраною теоретичною (емпіричною) залежністю називається ***апроксимацією***, а сам графік функції – апроксимованим.

Нехай потрібно знайти функцію $Y_T = f(x)$, значення якої при $X = X_1, X_2, \ldots, X_N$ яка б як можна менше відрізнялися від експериментальних (емпіричних) значень $Y_1, Y_2, \ldots, Y_N$.

В основу рішення покладено принцип Лежандра, за яким сума квадратів відхилень емпіричних значень Y_T, які виміряні під час проведення експериментів повинна бути якнайменшою від значень Y_i, які визначені за встановленим алгебраїчним виразом, або функцією.

Тому що більшість функцій може бути представлена в вигляді багаточлена n-ої степені, то при вирівнюванні доцільно представляти залежність між перемінними величинами в вигляді параболи n-ої степені

$$y = a_o + a_1 x + a_2 x^2 + \ldots + a_N x^N,$$

де $a_o, a_1, a_2, \ldots, a_N$ – невідомі параметри, або невідомі коефіцієнти.

Для їх знаходження використовують інтерполяційну формулу Чебішева, яка має вигляд

$$y = k_0 q_0(x) + k_1 q_1(x) + k_2 q_2(x) + \ldots + k_\lambda q_\lambda(x). \qquad (1)$$

У цій формулі величина $\lambda \le N - 1$ характеризує порядок параболи, N – кількість значень незалежної перемінної ***X***, або кількість проведених дослідів однофакторного експерименту (членів ряду Чебішева), а аргументом є величина $X = \bar{u}$.

Середнє значення $\bar{u}$ визначають за формулою

$$\bar{u} = \frac{\sum u_i}{N} \qquad (2)$$

де $\sum u_i$ – сума всіх експериментальних значень X_i, тобто

$$\sum u_i = X_1 + X_2 + X_3 + ... + X_N. \qquad (3)$$

Послідовність обчислення і способи визначення невідомих коефіцієнтів k і q, які входять в інтерполяційну формулу (1), розглянемо на прикладі складеної табл. 2.

Вихідними даними для проведення розрахунків є колонки 1-3 табл. 2, де в колонці 1 наведено кількість проведених дослідів N в експерименті, в колонці 2 – значення параметра оптимізації $\boldsymbol{y}$ (функції), в колонці 3 – значення аргументу u.

Чисельне значення X_i (колонка 5) визначають за формулою

$$x_i = u_i - \overline{u}. \qquad (4)$$

Таблиця 2 – Вихідні дані, послідовність обчислення і способи визначення невідомих коефіцієнтів

N	y	u	Y_i^2	X_i	$y_i X_i$	X_i^2	$X_i^2 y_i$	X_i^3	X_i^4	$X_i^3 y_i$	X_i^5	X_i^6	${}_2f(u)$	${}_3f(u)$
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	2	1	4	-6	-12	36	72	-216	1296	-482	-7776	46656	2,9	2,1
2	10	3	100	-4	-40	16	160	-64	256	-640	-1024	4096	9,6	13,1
3	17	4	289	-3	-51	9	153	-27	81	-459	-243	729	16,3	20,6
4	33	6	1089	-1	-33	1	33	-1	1	-33	-1	1	35,6	37,7
5	51	7	2601	0	0	0	0	0	0	0	0	0	48,5	48,4
6	66	8	4356	+1	+66	1	66	1	1	66	1	1	63,5	61,2
7	96	10	9216	+3	+288	9	864	27	81	2592	243	729	99,8	95,3
8	120	11	14400	+4	+480	16	1920	64	256	7680	1024	4096	121,2	117,6
9	172	13	29584	+6	+1032	36	6192	216	1296	37152	7776	46656	170,3	175,3
Сума	567	63	61639		1730	124	9460	0	3268	-1564 +47490 45926	0	102964		

<u>Порядок виконання роботи</u>

1. Визначення параболи нульової степені

1.1. Підрахувати $\sum y_i =$ 567 і занести її в кінці колонки 2 табл. 2.

1.2. Визначити величину $\boldsymbol{k_0}$ за формулою

$$k_0 = \frac{\sum y_i}{N} = 567 / 9 = 63.$$

Для рівняння параболи нульового порядку $\boldsymbol{q_0 (x) = x^0 =}$ 1.

1.3. Визначити рівняння параболи нульового порядку і записати його, при цьому

$$_0\, f(x) = k_0 q_0(x) = 63.$$

Знаходимо величину

$$\sum{}_0 = \sum y_i^2 - \frac{(\sum y_i)^2}{N} = 61639 - 567^2 / 9 = = 25918.$$

1.4. Визначити основну похибку. Для цього знаходимо y_i^2 і заповнюємо колонку 4.

Вираховуємо, що

$$\sum \mathrm{y_i^2} = 61639.$$

Основну похибку визначаємо за формулою

$$\sigma_0 = \sqrt{\frac{\sum_0}{N-1}} = \sqrt{\frac{25918}{9-1}} = 56{,}9.$$

2. Визначення параболи першого порядку

2.1. Вирахувати за формулою (3) $\sum u_i =$ 63.

2.2. Визначити за формулою (2)

$$\bar{u} = \frac{\sum y_i}{N} = 63 / 9 = 7.$$

2.3. Заповнити колонку 5, при цьому вирахувати значення $\boldsymbol{X_i}$ згідно приведеної формули (4).

2.4. Заповнити колонку 6, при цьому вирахувати значення величини $\boldsymbol{X_i^2}$ та знайти значення $\sum X_i^2 = 124$.

2.5. Визначити рівняння параболи першої степені, при цьому обчислити k_1 за формулою

$$k_1 = \frac{\sum y_i x_i}{\sum x_i^2} = 1730 / 124 = 13{,}59.$$

Для параболи першого порядку величина $q_1(x) = x$.

Тому

$$k_1 q_1(x) = 13{,}59\,x.$$

2.6. Скласти (просумувати) рівняння параболи нульового порядку і першої степені, тобто

$$k_0 q_0(x) + k_1 q_1(x) = 63 + 13{,}95\,x.$$

2.7. Записати шуканий вираз в такій формі

$${}_1 f(x) = 63 + 13{,}95\,x.$$

2.8. Вирахувати основну похибку за формулою

$$\sigma_1 = \sqrt{\frac{\sum_1}{N-2}},$$

де $\sum_1 = \sum_0 - k_1^2 \sum x_i^2 = 25918 - 13{,}95^2 * 124 = 1784{,}5$.

Тоді $\sigma_1 = \sqrt{\frac{1784{,}5}{9-2}} = 15{,}9$.

Тому що σ_0 значно перевищує σ_1, то необхідно продовжити інтерполярування.

3. Визначення параболи другого порядку

1.3.1. Вирахувати добуток величин $x_i^2 y_i$ і заповнити значення колонки 8, при цьому знайти $\sum x_i^2 y_i = 9460$.

3.2. Вирахувати x_i^3 і заповнити колонку 9, при цьому $\sum x_i^3 = 0$.

3.3. Вирахувати x_i^4 і заповнити колонку 10, при цьому сума x_i^4 буде рівною: $\sum x_i^4 = 3268$.

3.4. Обчислити допоміжні величини $\boldsymbol{A}_2$, $\boldsymbol{b}_2$, $\boldsymbol{C}_2$, які необхідні для подальших розрахунків невідомих величин k_2 і $q_2(x)$ за наступними формулами:

$$A_2 = \frac{\sum x_i^2}{N} = 124/9; \;\; b_2 = \frac{\sum x_i^3}{\sum x_i^2} = \frac{0}{124} = 0\,;$$

$$C_2 = \sum x_i^4 - b_2 \sum x_i^3 - A_2 \sum x_i^2 = 3268 - 0 * 0 - 124/9 * (124) = 1559{,}56.$$

3.5. Вирахувати величину $k_2 q_2(x)$, при цьому визначити значення

$$k_2 = \frac{\sum x_i^2 y - k_0 \sum x_i^2 - k_1 \sum x_i^3}{C_2} = (9460 - 63*124 - 13{,}95*0)/1559{,}56 = 1{,}06;$$

$$q_2(x) = (x^2 - b_2 x - A_2) = (x^2 - 0 * x - \frac{124}{9}) = x^2 - 13{,}78\,.$$

Таким чином

$$k_2 q_2(x) = 1{,}06(x^2 - 13{,}78) = 1{,}06x^2 - 14{,}6.$$

3.6. Скласти (просумувати) $k_0 q_0(x) + k_1 q_1(x) + k_2 q_2(x)$, при цьому рівняння параболи другого порядку буде мати вигляд

$$_2 f(x) = 63 + 13{,}59x + 1{,}06x^2 - 14{,}6 = 1{,}06x^2 + 13{,}95x + 48{,}4. \qquad (5)$$

3.7. Обчислити основну похибку за формулою

$$\sigma_2 = \sqrt{\frac{\sum_2}{N-3}},$$

де $\sum_2 = \sum_1 - k_2^2 C_2 =$ 1784,5 – 1,06^2 * 1559,56 = 32,12.

Тоді

$$\sigma_2 = \sqrt{\frac{32{,}12}{9-3}} = 2{,}3\,.$$

Якщо отримане значення $\sigma_2 = 2{,}3$ рахувати достатньо малим, тоді на даному етапі проведення апроксимації можна обмежитися обчисленням параболи другого порядку і провести апроксимацію експериментальних даних за одержаною функцією (5).

3.8. Для проведення апроксимації переходимо від аргументу $x = u - \bar{u} = u - 7$ до аргументу u, підставивши в рівняння параболи 2-го порядку (5) значення $x = u - 7$.

Тоді в кінцевому отримаємо

$$_2f(u) = 1{,}06(u-7)^2 + 13{,}95(u-7) + 48{,}4\,,$$

або після перетворення

$$_2f(u) = 1{,}06u^2 + 0{,}89u + 2{,}74. \qquad (6)$$

4. Визначення параболи третього порядку

4.1. Обчислити добуток $x_i^3 y_i$ і заповнити колонку 11, при цьому знайти

$$\sum x_i^3 y = 45926.$$

4.2. Обчислити величину x_i^5 і заповнити колонку 12, при цьому знайти

$$\sum x_i^5 = 0.$$

4.3. Обчислити величину x_i^6 і заповнити колонку 13, при цьому знайти

$$\sum x_i^6 = 102964.$$

4.4. Обчислити допоміжні величини ***C*з, *D*з, *A*з, *b*з, *E*з**, які необхідні для подальших розрахунків невідомих величин k_3 і $q_3(x)$ за наступними формулами:

$$C_3 = \sum x_i^5 - b_2 \sum x_i^4 - A_2 \sum x_i^3 = 0 - 0 * 3268 - 124/9 * 0 = 0;$$

$$D_3 = \sum x_i^6 - b_2 \sum x_i^5 - A_2 \sum x_i^4 =$$
$$= 102964 - 0 * 0 - 124/9 * 3268 = 45025{,}778;$$

$$A_3 = \frac{C_2}{\sum x_i^2} = 1559{,}56/124 = 12{,}577;$$

$$b_3 = \frac{C_3}{C_2} - \frac{\sum x_i^3}{\sum x_i^2} = 0/1559{,}56 - 0/124 = 0;$$

$$E_3 = D_3 - b_3 C_3 - A_3 \sum x_i^4 = 45025{,}78 - 0 * 0 - 12{,}58 * 3268 = 3914{,}34.$$

4.5. Вирахувати величину $k_3 q_3(x)$, при цьому визначити

$$k_3 = \frac{\sum x_i^3 y_i - k_0 \sum x_i^3 - k_1 \sum x_i^4 - k_2 C_3}{E_3} =$$
$$= (45926 - 63 * 0 - 13{,}95 * 3268 - 1{,}06 * 0) / 3914{,}34 = 0{,}0862;$$

$$q_3(x) = (x - b_3)q_2(x) - A_3x = (x-0)(x^2 - 124/9) - 12{,}58x = x^3 - 26{,}36x.$$

Таким чином

$$k_3q_3(x) = 0{,}0862(x^3 - 26{,}36x) = 0{,}0862x^3 - 2{,}267x.$$

4.6. Скласти (просумувати) $k_0q_0(x) + k_1q_1(x) + k_2q_2(x) + k_3q_3(x)$, при цьому рівняння параболи третього порядку буде мати вигляд

$${}_3f(x) = 63 + 13{,}95x + 1{,}06x^2 - 14{,}6 + 0{,}0862x^3 - 2{,}267x,$$

або після спрощення

$${}_3f(x) = 0{,}0862x^3 + 1{,}06x^2 + 11{,}68x + 48{,}4. \qquad (7)$$

4.7. Обчислити основну похибку за формулою

$$\sigma_3 = \sqrt{\frac{\sum_3}{N-4}},$$

де $\sum_3 = \sum_2 - k_3^2E_3 =$ 32,12 – 0,0862^2 * 3914,34 = 3,15.

Тоді

$$\sigma_3 = \sqrt{\frac{3{,}15}{9-4}} = 0{,}79.$$

Тому що $\sigma_3 < \sigma_2$, то отже, вирівнювання за параболою 3-го порядку дає декілька краще приближення.

Величина σ_3 дуже мало відрізняється за своїм чисельним значенням від величини σ_2 і тому подальше збільшення порядку параболи недоцільно.

Тоді на даному етапі можна обмежитися обчисленням параболи третього порядку і провести апроксимацію експериментальних даних за одержаною функцією (7). Необхідно також відзначити, що в практичних випадках параболи вище 3-го порядку зустрічаються дуже не часто і дають практично несуттєве зменшення основної помилки.

4.8. Для проведення апроксимації переходимо від аргументу $x = u - \bar{u} = u - 7$, для цього виразимо аргумент x функції ${}_3f(x)$ через аргумент u, підставивши в рівняння параболи 3-го порядку (7) замість x як і раніш значення $x = u - 7$.

Тоді у кінцевому отримаємо

$${}_3f(u) = 0{,}0862(u-7)^3 + 1{,}06(u-7)^2 + 11{,}68(u-7) + 48{,}4,$$

або після перетворення

$$_3f(u) = 0{,}0862u^3 - 0{,}75u^2 + 9{,}51u - 10{,}99. \qquad (8)$$

Побудувати апроксимовані залежності:

- першого порядку $y1(x) = {}_1f(x) = 63 + 13{,}95x$;
- другого порядку $y2(x) = {}_2f(u) = 1{,}06u^2 + 0{,}89u + 2{,}74$;
- третього порядку $y3(x) = {}_3f(u) = 0{,}0862u^3 - 0{,}75u^2 + 9{,}51u - 10{,}99$.

Побудовані залежності наведено на рис. 2.

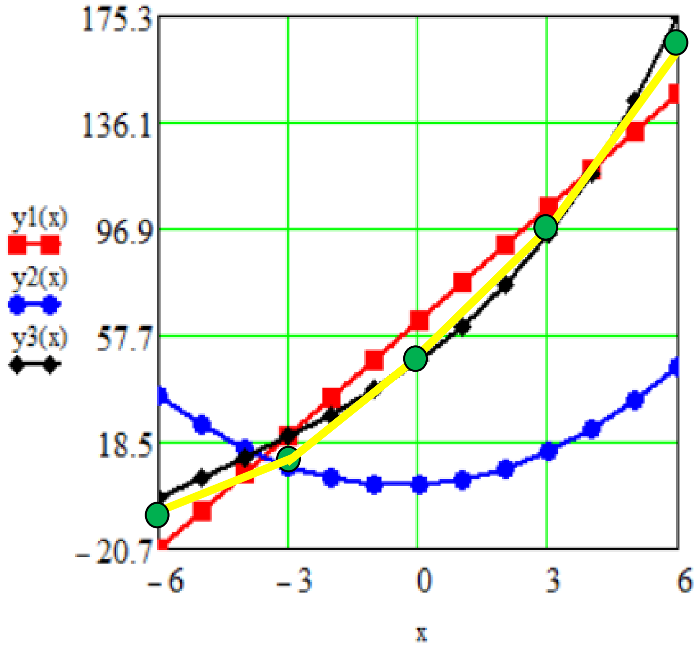


Рисунок 2 – Графічна інтерпретація процесу апроксимації ломаної лінії, яку побудовано за точками експериментальних даних

Порядок оформлення звіту

2.1. Записати вихідні дані для проведення розрахунків згідно табл. 1. Відповідний варіант завдання взяти згідно табл. 3.

2.2. Провести і записати математичні розрахунки згідно п. 1 “Порядок виконання роботи” та заповнити табл. 2.

2.3. Обчислити вирівняні значення параметра *y*, які вирахували за рівняннями отриманих парабол 2-го (6) та 3-го (8) порядку і заповнити колонки 14 та 15 табл. 2.

2.4. Побудувати графік експериментальної залежності в вигляді ломаної лінії згідно рис. 1.

2.5. На тих же координатних осях побудувати графіки отриманих парабол 1-го, 2-го та 3-го порядку.

Побудову графіків виконати в різних кольорах.

2.6. Зробити висновок за результатами проведеної апроксимації експериментальних даних.

Контрольні запитання

1. Які величини називаються функціонально залежними?
2. Дати визначення поняттю «апроксимація».
3. Яка мета апроксимації експериментального масиву даних?
4. Навести приклади двох функціонально перемінних величин.
5. Дати визначення принципу Лежандра, його призначення.
6. Поняття інтерполяції.
7. Загальна методика проведення інтерполяції експериментальних даних.
8. Як отримують експериментальний масив даних?

Таблиця 3 – Вихідні дані для проведення розрахунків

Вар.	**B1**		**B2**		**B3**		**B4**		**B5**		**B6**	
N	**y**	**u**	**y**	**u**	**y**	**u**	**y**	**u**	**y**	**u**	**y**	**u**
1	23	5	8	3	2	1	49	7	79	9	26	4
2	50	7	23	5	7	3	83	9	125	11	40	6
3	65	8	63	8	35	6	97	10	17	12	48	7
4	127	11	85	9	48	7	169	13	143	15	96	10
5	154	13	112	11	80	9	224	15	276	17	147	12
6	195	14	143	12	91	10	245	16	317	18	164	13
7	231	15	160	13	128	11	282	17	354	19	191	14
8	279	17	227	15	171	13	371	19	443	21	257	16
9	312	18	250	16	198	14	384	20	494	22	281	17
10	397	20	315	18	253	16	499	22	579	24	369	19
11	495	22	393	20	311	18	569	24	679	26	444	21
Вар.	**B7**		**B8**		**B9**		**B10**		**B11**		**B12**	
1	8	3	5	2	2	1	51	7	81	9	15	4
2	62	8	28	5	7	3	83	9	125	11	40	6
3	86	9	63	8	39	6	101	10	137	12	45	7
4	138	12	85	9	46	7	173	13	230	15	102	10
5	195	14	122	11	80	9	230	15	282	17	140	12
6	226	15	143	12	91	10	245	16	317	18	164	13
7	255	16	163	13	118	11	282	17	364	19	191	14
8	317	18	227	15	165	13	364	19	447	21	268	16
9	368	19	258	16	188	14	394	20	489	22	281	17
10	446	21	317	18	253	16	499	22	570	24	365	19
11	526	23	393	20	314	18	567	24	681	26	450	21
Вар.	**B13**		**B14**		**B15**		**B16**		**B17**		**B18**	
1	5	2	10	3	2	1	13	4	61	8	36	6
2	11	3	19	4	3	2	27	5	115	10	61	8
3	15	4	23	5	11	3	77	6	117	11	84	9
4	37	6	45	7	23	5	69	8	183	14	146	12
5	64	8	82	9	43	7	104	10	263	16	183	14
6	85	9	103	10	62	8	115	11	277	17	224	15
7	112	11	150	12	98	10	159	13	314	18	251	16
8	149	12	167	13	115	11	227	15	403	20	318	18

9	182	14	228	15	174	13	254	16	454	21	371	19
10	237	15	255	16	193	14	289	17	530	23	446	21
11	290	17	313	18	251	16	367	19	617	25	534	23

ПРАКТИЧНА РОБОТА № 3

ВИЗНАЧЕННЯ КОЕФІЦІЄНТА КОРЕЛЯЦІЇ ДВОХ СТОХАСТИЧНО ЗАЛЕЖНИХ ВИПАДКОВИХ ВЕЛИЧИН

Мета роботи: вивчити порядок обчислення коефіцієнта кореляції між випадковими величинами за вибірками невеликого та великого об’єму.

Теоретичні відомості

Випадковою величиною називають величину, яка в результаті проведеного експерименту (досліду) може приймати різні значення. Наприклад, нехай задано верхню та нижню граничну межу відхилення розміру виточеної деталі від заданого нульового рівня, яка знаходиться в межах ± 30 мкм. Кількість вимірів розміру деталей, які знаходяться вище і нижче граничних меж ± 30 мкм, є випадковим величинами.

Також випадковими величинами є кількість вимірів розміру деталей, який розташований в граничних межах, тобто в межах “-“ 30 мкм – 0 – “+” 30 мкм, при цьому випадкові величини можуть приймати будь-які значення в цих заданих граничних межах.

Випадкові величини, як правило позначають великими буквами, наприклад, $\boldsymbol{X}$. Значення випадкової величини, які вона приймає в результаті досліду, позначають малими буквами $\boldsymbol{x_1, x_2, \dots, x_n}$.

При великій кількості проведених дослідів кожне із можливих значень випадкової величини (розташування головок коренеплодів відносно поверхні ґрунту) $\boldsymbol{x_1, x_2, \dots, x_n}$ може зустрітися $\boldsymbol{m_1, m_2, \dots, m_n}$ раз. Ці числа називають частотами розподілу.

Якщо було проведено N дослідів, тобто $\sum_{i=1}^{n} m_i = N$, то відношення $\frac{m_i}{N}$ називають ***частістю*** або ***відносною частотою*** розподілу.

Сукупність, яка містить всі деталі, в яких вимірювали їх розмір називається ***генеральною сукупністю***, а відібрані із генеральної сукупності якась кількість N деталей утворюють вибірку об’єму N.

Випадкові величини можуть бути дискретними або безперервними. ***Дискретними*** випадковими величинами називають такі, які можуть приймати лише визначенні значення, наприклад 10, 15, 20 і т.д.

Безперервними випадковими величинами називають такі, які в деякому інтервалі (граничних межах) можуть приймати будь-які значення. Кількість вимірів деталей, розмір яких більше і менше граничних меж ± 30 мкм в різних вибірках із генеральної сукупності є дискретна випадкова величина, а чисельне значення відхилення розміру деталі від заданого – безперервна випадкова величина.

Дискретна випадкова величина задана, якщо є вірогідність кожного її значення (табл. 1).

Таблиця 1 – Дискретна випадкова величина

X	$\boldsymbol{x_1}$	$\boldsymbol{X_2}$	$\boldsymbol{x_3}$	**...**	$\boldsymbol{x_n}$
$\boldsymbol{P(X = x_i)}$	$\boldsymbol{P(x_1)}$	$\boldsymbol{P(x_2)}$	$\boldsymbol{P(x_3)}$	**…**	$\boldsymbol{P(x_n)}$

Будь-яка випадкова величина характеризується щільністю і інтегральною функцією розподілення випадкових величин.

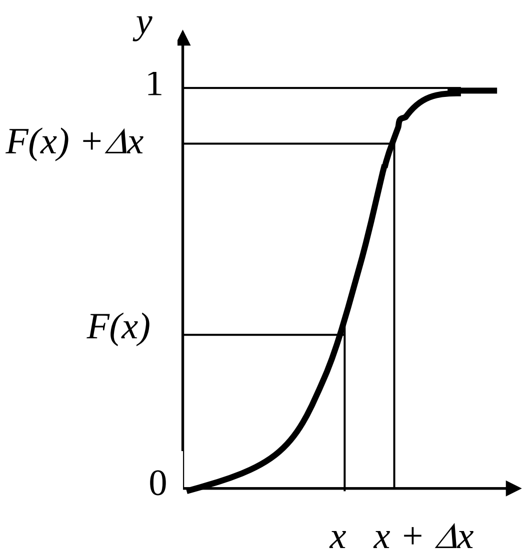


Рисунок 1 – Інтегральна функція розподілення

Якщо ***X*** – випадкова величина, а ***x*** – деяке її значення, то вірогідність того, що $\boldsymbol{X < x}$ дорівнює ***F(x) = P(X < x)***, де ***F(x)*** – деяка функція, яка називається ***інтегральною функцією розподілення*** (рис. 1).

На рис. 1 ***F(x)*** – ордината кривої в деякій точці ***x***.

При будь-якому ***x*** завжди $\mathbf{0} \le \boldsymbol{F(x)} \le 1$.

Щільність вірогідності ***φ(x)*** є границя відношення вірогідності того, що випадкова величина ***X*** прийме значення, яке знаходиться між ***x*** і ***x + Δx***, до величини інтервалу ***Δx*** при $\boldsymbol{\Delta x \to 0}$, тобто

$$\varphi(x) = \lim_{\Delta x \to 0} \frac{P(x \le X \le x + \Delta x)}{\Delta x}.$$

Функцію $\varphi(x)$ називають також ***диференціальним законом розподілення*** випадкових величин, при цьому $\varphi(x)$ і ***F(x)*** зв'язані відношенням

$$F(x) = P(X \prec x) = P(-\infty \prec X \prec x) = \int_{-\infty}^{x} \varphi(x) dx.$$

Будемо вважати, що випадкова величина задана теоретичним законом розподілення, якщо задані її інтегральний закон або її щільність вірогідності.

Випадкова величина задана емпіричним законом розподілення, якщо для кожного значення випадкової величини відома частота повторюваності або відносна частота, яка отримана із N дослідів (табл. 2).

Таблиця 2 – Емпіричний закон розподілення випадкової величини

Значення X	x_1	x_2	x_3	...	x_n
Частоти	m_1	m_2	m_3	...	m_n
Відносна частота	m_1/N	m_2/N	m_3/N	...	m_n/N

При $N \to \infty \dfrac{m_i}{N} \to P(x_i)$.

Будь-яке теоретичне розподілення характеризується величиною своїх основних параметрів: ***математичним сподіванням*** MX (центром групування) і ***дисперсією*** DX (величиною розсіювання).

Для дискретної випадкової величини (див. табл. 1) значення MX буде

$$MX = \sum_{i=1}^{n} x_i P(x_i); \quad DX = \sum_{i=1}^{n} P(x_i)(x_i - MX)^2.$$

Для безперервної випадкової величини, яка задана своєю щільністю вірогідності $\varphi(x)$, математичне сподівання і дисперсія дорівнюють

$$MX = \int_a^b x\varphi(x)dx; \quad DX = \int_a bx^2\varphi(x)dx - (MX)^2, \tag{1}$$

або

$$MX = \int_{-\infty}^{\infty} x\varphi(x)dx; \quad DX = \int_{-\infty}^{\infty} x^2\varphi(x)dx - (MX)^2. \tag{2}$$

Формули (1) застосовуються в тих випадках, коли випадкова величина приймає значення від $\boldsymbol{a}$ до $\boldsymbol{b}$ (в заданих граничних межах), а формули (2) – коли $\boldsymbol{x}$ змінюється від $-\infty$ до ∞.

Вираз $\sqrt{DX} = \sigma$ називається ***середнім квадратичним відхиленням*** або стандартом теоретичного випадкового розподілення.

Емпіричне розподілення характеризується середнім значенням $\bar{x}$

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i m_i}{\sum_{i=1}^{n} m_i} = \frac{\sum_{i=1}^{n} x_i m_i}{N}, \text{ або } \bar{x} = \frac{\sum_{i=1}^{N} x_i}{N}.$$

Середнє значення $\bar{x}$ характеризує центр групування значень випадкової величини. При достатньо великому N ($N \to \infty$) вибіркове значення $\bar{x}$ прямує за величиною до математичного сподівання, тобто $\bar{x} \approx MX$.

Величина розсіювання вибіркових значень навколо їх середнього значення характеризується ***емпіричною дисперсією*** S^2, яка дорівнює

$$S^2 = \frac{1}{N-1}\sum_{i-1}^{n} m_i (x_i - \bar{x})^2. \qquad (3)$$

Для $N \geq 25$ замість формули (3) для обчислення дисперсії використовують формулу

$$S^2 = \frac{1}{N}\sum_{i-1}^{a} m_i (x_i - \bar{x})^2 = a_2 - \bar{x}^2,$$

де $a_2 = \frac{\sum_{i-1}^{n} x_i^2 m_i}{N}$.

Вираз $\sqrt{S^2} = S$ називається ***середнім квадратичним відхиленням*** емпіричного випадкового розподілення.

Нехай маємо дві випадкові величини ***X*** і ***Y*** з заданими математичними сподіваннями MX і MY та середніми квадратичними відхиленнями σ_x і σ_y їх теоретичного розподілу.

Величина

$$\mu_{xy} = M\left[(X - MX)(Y - MY)\right] = M(XY) - MX \cdot MY = \mathrm{cov}(XY) \qquad (4)$$

має назву коваріації.

Пронормуємо випадкові величини ***X*** і ***Y***, тобто перейдемо до нових випадкових величин ***X′*** і ***Y′***, математичні сподівання яких дорівнюють нулю, а дисперсія – 1. Тоді

$$X' = \frac{X - MX}{\sigma_x}; \qquad Y' = \frac{Y - MY}{\sigma_y},$$

а ***cov (X′Y′)*** нормованих випадкових величин ***X′*** і ***Y′*** називається ***коефіцієнтом кореляції***, тобто

$$\rho_{xy} = \mathrm{cov}(X'Y') = M\left[\left(\frac{X - MX}{\sigma_x}\right)\left(\frac{Y - MY}{\sigma_y}\right) \right] = \frac{M(XY) - MX \cdot MY}{\sigma_x \sigma_y}, \qquad (5)$$

або

$$\rho_{xy} \frac{\mathrm{cov}(XY)}{\sigma_x \sigma_y} = \frac{\mu_{xy}}{\sigma_x \sigma_y}.$$

Коефіцієнт кореляції вказує на щільність зв'язку між двома випадковими величинами і вимірюється від –1 до +1. При прямій лінійній залежності, тобто коли з збільшенням значень $\boldsymbol{x_i}$ відповідно зростають значення $\boldsymbol{y_i}$ $\rho_{xy} = +1$, а при зворотній лінійній залежності, тобто коли з збільшенням $\boldsymbol{x_i}$ значення $\boldsymbol{y_i}$ зменшуються $\rho_{xy} = -1$. Якщо x і y незалежні величини, тоді $\rho_{xy} = 0$.

При $\rho_{xy} \neq 1$ кожному значенню $\boldsymbol{x_i}$ відповідає декілька значень $\boldsymbol{y_i}$.

Умовним середнім $\overline{X}_y$ називається середнє значення із величини $\boldsymbol{x_i}$ при даному значенні $\boldsymbol{y_i}$.

Умовним середнім $\overline{Y}_x$ називається середнє значення із величини y_i при даному значенні $\boldsymbol{x_i}$. Дві лінії, які з'єднують всі значення $\overline{X}_y$ і $\overline{Y}_x$, називаються ***лініями регресії***.

Коефіцієнт кореляції $\rho_{xy} > 0$ коли з збільшенням значень $\boldsymbol{x_i}$ значення умовних середніх $\overline{Y}_x$ збільшуються.

Коефіцієнт кореляції $\rho_{xy} < 0$ коли з збільшенням значень $\boldsymbol{x_i}$ значення умовних середніх $\overline{Y}_x$ зменшуються.

Якщо лініями регресії є прямі лінії, тоді кореляція називається ***прямолінійною***. В даній роботі обмежимося розглядом тільки прямолінійної кореляції.

Теоретичне обчислення коефіцієнта кореляції за формулою (5) в більшості випадків викликає багато утруднень. Тому для його визначення користуються результатами експериментальних даних.

Порядок виконання роботи

1. Визначення коефіцієнта кореляції за вибіркою невеликого об'єму

Послідовність обчислення розглянемо на прикладі складеної табл. 3.

Таблиця 3 – Таблиця для проведення розрахунків

N	$\boldsymbol{x_i}$	$\boldsymbol{y_i}$	$\boldsymbol{x_i y_i}$	$\boldsymbol{x_i^2}$	$\boldsymbol{y_i^2}$
1	0,90	-0,30	-0,2700	0,8100	0,0900
2	1,22	0,10	0,1220	1,4889	0,0100
3	1,32	0,70	0,9240	1,7424	0,4900
4	0,77	-028	-0,2156	0,5929	0,0784
5	1,30	-0,25	-0,3250	1,6900	0,0625
6	1,20	0,02	0,0240	1,4400	0,0004

7	1,32	0,37	0,4884	1,7424	0,1369
8	0,95	-0,70	-0,6650	0,9025	0,4900
9	1,45	0,55	0,7975	2,1025	0,3025
10	1,30	0,35	0,4550	1,6900	0,1225
11	1,20	0,32	0,3840	1,4400	0,1024
Сума	12,93	0,88	1,7193	15,6411	1,8856

Нехай необхідно виявити наявність кореляційного зв'язку між розмірами моделей і відливок до них. Шляхом вимірювання визначені відхилення від номіналів висот моделей для 11 зразків (x_i) і відливок (y_i).

Вихідними даними для проведення розрахунків є колонки 1-3 табл. 3, де в колонці 1 наведено кількість проведених дослідів N в експерименті, в колонці 2 – відхилення висот моделей $\boldsymbol{x_i}$, в колонці 3 – відхилення висот відливок $\boldsymbol{y_i}$.

1.1. Обчислити добуток $x_i y_i$ і заповнити колонку 4, при цьому знаходимо $\sum x_i y_i =$ 1,7193.

1.2. Обчислити x_i^2 і y_i^2 та заповнити відповідні колонки 5 і 6, при цьому $\sum x_i^2 = 15{,}6411$; $\sum y_i^2 = 1{,}8856$.

1.3. Визначити емпіричний коефіцієнт кореляції для вибірки невеликого об'єму за формулою:

$$r_{xy} = \frac{N\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{N\sum x_i^2 - (\sum x_i)^2} \cdot \sqrt{N\sum y_i^2 - (\sum y_i)^2}}. \tag{6}$$

1.4. Підставити знайдені суми в формулу (6), при цьому

$$r_{xy} = \frac{11 \cdot 1{,}7193 - 12{,}93 \cdot 0{,}88}{\sqrt{11 \cdot 15{,}6411 - (12{,}93)^2} \cdot \sqrt{11 \cdot 1{,}8856 - (0{,}88)^2}} = 0{,}76.$$

2. Визначення коефіцієнта кореляції за вибіркою великого об'єму (об'єм вибірки $N > 50$)

Вибіркою великого об'єму будемо вважати вибірку, в які декілька значень перемінних зустрічаються по 2 і більше разів.

Послідовність обчислення розглянемо на прикладі складеної табл. 4.

Нехай необхідно визначити коефіцієнт кореляції між випадковими величинами розмірів двох деталей, які обробляються одночасно на одному станку. Після обробки кожної пари деталей проводять вимірювання, результати якого заносять в протокол (див. табл. 4).

Вихідними даними для проведення розрахунків є кількість проведених дослідів (вимірювань) N та результати вимірювань розмірів кожної пари деталей ***X*** і ***Y***.

Таблиця 4 – Результати вимірювання двох деталей

N	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**…**	**100**
Деталь *X*	**21,867**	**21,845**	**21,871**	**21,878**	**21,847**	**21,867**	**21,867**	**...**	**21,867**
Деталь *Y*	**21,852**	**21,843**	**21,864**	**21,871**	**21,838**	**21,852**	**21,853**	**...**	**21,854**

2.1. Знайти в кожній строчці мінімальне і максимальне значення розмірів деталей ***X*** і ***Y***, тобто в нашому випадку 21,845; 21,878 та 21,838; 21,871.

2.2. Якщо різниця між цими значеннями відносно велика, то всі значення доцільно розбити на групи.

У нашому прикладі

21,878 – 21,845 = 0,033 і 21,871 – 21,838 = 0,033.

2.3. Об’єднати всі значення розмірів в групи з шириною інтервалу ***h*** = 0,002. Ширину інтервалу ***h*** вибирати в залежності від значення різниці між максимальним і мінімальним значеннями розмірів деталей.

2.4. Побудувати кореляційну таблицю (див. табл. 5), в якій привести визначені інтервали, середини інтервалів x_i і y_i та значення нових випадкових величин x_i' і y_i', які визначаються за формулами:

$$y_i' = \frac{y_i - y_0}{h_1}; \quad x_i' = \frac{x_i - x_0}{h_2}. \tag{7}$$

де h_1, h_2 – відповідно, величина інтервалів випадкових значень величин ***X*** і ***Y***;

x_0, y_0 – відповідно, середнє значення середин інтервалів випадкових значень величин ***X*** і ***Y***.

2.4.1. За x_0 і y_0 прийняти середнє значення середини інтервалів, тобто

$$x_0 = 0,5\,(X_{мін} + X_{макс}) = 0,5\,(21,845 + 21,878) = 21,8615;$$

$$y_0 = 0,5\,(Y_{мін} + Y_{макс}) = 0,5\,(21,838 + 21,871) = 21,8545.$$

У нашому випадку $h_1 = h_2 = 0,002$.

Перехід до нових випадкових величин доцільно робити в тих випадках, коли середини інтервалів x_i і y_i мають двозначні, або більші за двозначні значення.

2.4.2. Для заповнення кореляційної таблиці використати протокол вимірів деталей (табл. 4). Взяти перший результат виміру першої пари деталей ***X*** і ***Y*** (див. колонку 1 табл. 4), тобто 21,867 ... 21,852 і знайти в табл. 5 по горизонталі інтервал, який містить число 21,867, а по вертикалі – 21,852.

На перехресті цих координат поставити точку (в табл. 5 ця точка обведена кільцем).

Потім взяти другий результат вимірів другої пари деталей ***X*** і ***Y*** (див. колонку 2 табл. 4), тобто 21,845 ... 21,843 і знайти інтервал, який містить ці значення в табл. 5, і на перехресті координат поставити точку (в табл. 5 ця точка також обведена кільцем).

2.4.3. Виконати ті самі дії із всіма останніми парами вимірів деталей (вибірка *N*) і занести відповідні точки в відповідні прямокутники (перехрестя відповідних координат) кореляційної таблиці.

У результаті заповнення кореляційної таблиці (див. табл. 5) отримуємо частоти повторюваності n_{xy} всіх різних пар значень x_i і y_i.

2.5. Визначити емпіричний коефіцієнт кореляції за формулою

$$r_{xy} = \frac{N\sum_{x'}\sum_{y'} n_{x'y'} x'_i y'_i - (\sum_{x'} n_{x'} x')(\sum n_{y'} y'_i)}{\sqrt{\left[N\sum_{x'} n_{x'}(x'_i)^2 - (\sum n_{x'} x'_i)^2\right]\left[N\sum_{y'} n(y'_i)^2 - (\sum n_{y'} y'_i)^2\right]}}, \tag{8}$$

де N – число дослідів, або вибірка із генеральної сукупності;

$n_{x'y'}$ – частота повторюваності спільного наставання подій ***X*** і ***Y***.

2.5.1. Послідовність обчислення коефіцієнта кореляції r_{xy} наведена в строках 1 – 5 і колонках 1 – 3 табл. 5.

Значення $n_{x'}$ і $n_{y'}$ обчислюються як суми частот за всіма колонками і строками кореляційної таблиці (див. табл. 5).

Для цього необхідно:

а) обчислити $\sum_{x'} n_{x'} = \sum_{y'} n_{y'} = N = 100$, як суму значень 1-ої строчки $(A^* = \sum_{x'} n_{x'})$ і 1-ої колонки $(M^* = \sum_{y'} n_{y'})$.

Рівність $\sum_{x'} n_{x'}$ і $\sum_{y'} n_{y'}$ править контролем правильності обчислення $n_{x'}$ і $n_{y'}$;

<u>**Примітка до табл. 5:**</u>

$$A^* = \sum_{x'} n_{x'}; \quad B^* = \sum_{x'} n_{x'} x'; \quad C^* = \sum n_{x'} (x')^2;$$

$$D^* = \sum_{y'} \sum_{x'} n_{x'y'} y'; \quad E^* = \sum_{y'} \sum n_{x'y'} y' x'; \quad K^* = \sum_{y'} n_{y'} (y')^2;$$

$$L^* = \sum_{y'} n_{y'} y'; \qquad M^* = \sum_{y'} n_{y'}.$$

ж) визначити добуток значень $n_{x'y'}$ на відповідне значення y_i', сумуємо ці добутки і заповнюємо строчку 4.

Наприклад:

1∗ (-7) = -7;   2 ∗ (-8) = - 16;   1∗ (-6) = -6;   1∗ (-4) = -4.

1 (-5) + 3 (-4) + 1 (-3) + 2 (-1) + 1∗ 2 = -5 –12 – 3 – 2 + 2 = - 20 і т.д.;

Таблиця 5 – Кореляційна таблиця

Таблиця 5

Кореляційна таблиця

																				1	2	3
		x_i'	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8			
		Середина інтервалів x_i	21,8335	8405	8425	8445	8465	8485	8505	8525	8545	8565	8585	8605	8625	8645	8665	8685	8705			
y_i'	Середина інтервалів y_i	x / y	21,838-21,839	840 841	842 843	844 845	846 847	848 849	850 851	852 853	854 855	856 857	858 859	860 861	862 863	864 865	866 867	868 869	870 871	$n_{y'}$	$n_{y'}\,y'$	$n_{y'}\,(y')^2$
-8	21,8455	21,845-21,846			2															2	-16	128
-7	8475	847 848	1																	1	-7	49
-6	8495	849 850				1														1	-6	36
-5	8515	851 852						1												1	-5	25
-4	8535	853 854					1	3		1										5	-20	80
-3	8555	855 856						1		3	1									5	-15	45
-2	8575	857 858							1	2	1									4	-8	16
-1	8595	859 860						2		3	1	2								8	-8	16
0	8615	861 862							1			2	2	1						6	0	0
1	8635	863 864										1		2						3	3	3
2	8655	865 866						1		3	2		4	3	4	2				20	40	80
3	8575	867 868								2	5	1	2	4	6	2				22	66	198
4	8695	869 870									1		1	1	4	5				13	52	208
5	8715	871 872												1	3	3		1		8	40	200
6	8735	873 874																				
7	8755	875 876																				
8	8775	877 878																	1	1	8	64

1	n_{xy}	1		2	1	1	8	2	14	11	7	9	13	17	12		1	1	A*	100
2	$n_{x'}\,x'$	-8		-12	-5	-4	-24	-4	-14	0	7	18	39	68	60		7	8	B*	136
3	$n_{x'}\,(x')^2$	64		12	25	16	72	8	14	0	7	36	117	272	300		49	64	C*	1116
4	$\Sigma n_{x'y'}\,y'$	-7		-16	-6	-4	-20	-2	-8	17	6	18	31	57	45		5	8	D*	124
5	$\Sigma n_{x'y'}\,y'\,x'$	56		96	30	16	60	4	8	0	6	36	93	228	225		35	64	E*	957

M*	L*	K*
100	124	1140

з) просумувати всі значення цієї строчки, визначити

$$D^* = \sum_{y'} \sum_{x'} n_{x'y'} y'_i = 124.$$

Контролем правильності попередніх обчислень править рівність сум значень 4-ої строчки і 2-ої колонки, тобто $\sum_{y'} \sum_{x'} n_{x'y'} y'_i = \sum_{y'} n_{y'} y'$, при цьому

124 = 124;

е) значення 4-ої строчки помножити на x'_i і заповнити строчку 5.

Сума значень цієї строчки дорівнює $\sum_{y'} \sum_{x'} n_{x'y'} x'_i y'_i = 957$.

2.6. Обчислені значення сум підставити в формулу (8) і визначити емпіричне значення коефіцієнта кореляції, тобто

$$r_{xy} = \frac{100 \cdot 957 - 136 \cdot 124}{\sqrt{(100 \cdot 1116 - 136^2) \cdot (100 \cdot 1140 - 124^2)}} = \frac{78836}{95824} = +0{,}82. \cdot$$

2.7. Провести оцінку відмінності отриманого значення коефіцієнта кореляції r_{xy} від 0.

Для вирішення цієї задачі можна скористатися способом Фішера, згідно якого табличне значення функції ***Ф(t)*** розраховується за формулою $\Phi(t)_т = \frac{1}{\sqrt{2\pi}} \int_0^t e^{-\frac{t^2}{2}} dt \cdot$

Випадкова величина Z вибірки *N* визначається за формулою

$$Z = 0{,}5\{lq(l + r_{xy}) - lq(l - r_{xy})\} = 0{,}5lq \frac{l + r_{xy}}{l - r_{xy}}. \qquad (9)$$

При цьому вона із генеральної сукупності підпорядковується нормальному закону розподілення випадкових величин із середнім квадратичним відхиленням, яке визначається як

$$\sigma_z = \frac{1}{\sqrt{N - 3}}, \qquad (10)$$

де *N* – кількість проведених дослідів, або вибірка із генеральної сукупності.

Значення ***Z*** для різних значень коефіцієнта кореляції r_{xy} наводяться в спеціальних розроблених таблицях – в нашому випадку така таблиця додається викладачем (табл. А). Для цього необхідно:

а) знайти за табл. А випадкову величину ***Z*** – в нашому випадку при r_{xy} = 0,82; Z = 1,1568;

б) визначити середнє квадратичне відхилення σ_z за формулою 10), при цьому $\sigma_Z = 1/\sqrt{N-3} = 1/\sqrt{100-3} = 1/\sqrt{97} = 0{,}1015$;

в) визначити розрахункове значення коефіцієнта t_p за формулою $t_p = \frac{Z}{\sigma_z}$, при цьому $t_p = \frac{1{,}1568}{0{,}1015} = 11{,}4$;

г) за знайденим значенням t_p згідно таблиці Б, (яка додається викладачем) знайти ***Ф(t)ₘ***, тобто при t_p= 11,4 табличне значення функції $\boldsymbol{\Phi(t)_m}$ = 0,5.

Вірогідність того, що відхилення $P_{r_{xy}=0}$ визначеного коефіцієнта кореляції r_{xy} = 0,82 від 0 випадково дорівнює $P_{r_{xy}=0} = 0{,}5 - \Phi(t)$, тому $P_{r_{xy}=0} = 0$.

За рівень значимості $P_{r_{xy}=0}$, як правило приймають рівень значимості 0,05 або 0,01.

Якщо $P_{r_{xy}=0}$> 0,05; 0,01, то значення коефіцієнта кореляції r_{xy} можна вважати отриманим випадково, а досліджувальні випадкові величини некорелятивними.

2.8. Знайти інтервал, в якому з заданою надійністю знаходиться теоретичне значення коефіцієнта кореляції ρ_{xy} випадкових величин.

Тому що коефіцієнт кореляції r_{xy} є випадковою величиною, то іноді необхідно за емпіричним значенням r_{xy} оцінити теоретичне значення коефіцієнта кореляції ρ_{xy}, тобто знайти такий інтервал, в якому з заданою надійністю знаходиться значення ρ_{xy}.

Для цього необхідно:

а) задати надійність функції Ф(t), наприклад, прийняти надійність Ф(t) рівною 0,95 , тобто в цьому випадку величина функції Ф(t) буде рівною $\frac{1}{\sqrt{2\pi}}\int_0^t e^{-\frac{t^2}{2}}dt$ = 0,475;

б) знайти табличне значення коефіцієнта $t_т$ за табл. Б – при значенні $\Phi(t)_т$ = 0,475; $t_т$ = 1,96;

в) визначити довірчий інтервал випадкової величини $Z_{ген}$ генеральної сукупності випадкової величини для випадкової величини Z вибірки N із генеральної сукупності, який відповідає теоретичному значенню коефіцієнта кореляції ρ_{xy} за тотожністю:

$$Z - 1{,}96\sigma_z \le Z_{ген} \le Z + 1{,}96\sigma_Z,$$

при цьому

$$1{,}1568 - 1{,}96 * 0{,}1015 \le Z_{ген} \le 1{,}1568 + 1{,}96 * 0{,}1015, \text{ або}$$

$$1{,}0579 \le Z_{ген} \le 1{,}3557;$$

г) знайти табличне значення коефіцієнта кореляції r_{xy} за табл. А – при значенні $Z_1 = 1{,}0579$; $r_{xy_1} = 0{,}79$ і для $Z_2 = 1{,}3557$; $r_{xy_2} = 0{,}89$.

Таким чином, $0{,}79 \le \rho_{xy} \le 0{,}89$, тобто теоретичне значення коефіцієнта кореляції з вірогідністю 0,95 знаходиться в цьому інтервалі.

<u>Порядок оформлення звіту</u>

2.1. Записати вихідні дані для проведення розрахунків завдання за п.1.1 згідно табл. 3.

Відповідний варіант завдання взяти згідно таблиці 6.

2.2. Провести і записати математичні розрахунки згідно п. 1 "Порядок виконання роботи" і підпунктів 1.1.1 – 1.1.4.

Заповнити відповідні строки і колонки табл. 3.

2.3. Обчислити емпіричний коефіцієнт кореляції для вибірки невеликого об'єму за формулою (6).

2.4. Записати вихідні дані для проведення розрахунків завдання за п. 1.2 згідно табл. 4.

Відповідний варіант завдання взяти згідно таблиці 7.

2.5. Провести і записати математичні розрахунки згідно п. 1 "Порядок виконання роботи" і підпунктів 1.2.1 – 1.2.8.

Заповнити кореляційну таблицю (див. табл. 5).

2.6. Обчислити емпіричний коефіцієнт кореляції для вибірки великого об'єму за формулою (8).

2.7. Написати висновки за результатами виконаної роботи.

Контрольні запитання

1. Які величини називають випадковими? Навести приклад.
2. Яка різниця між генеральною сукупністю і вибіркою об'єму випадкових величин?
3. Поняття дискретних і безперервних випадкових величин. Навести приклади.
4. Що таке інтегральна функція розподілення?
5. Емпіричний закон розподілення випадкової величини.
6. Основні параметри теоретичного розподілення.

7. Коефіцієнт кореляції, його суть та призначення.

Таблиця 6 – Варіанти вихідних даних для проведення розрахунків

Вар.	**В1**		**В2**		**В3**		**В4**		**В5**		**В6**	
N	x_i	y_i	x_i	y_i	x_i	y_i	x_i	y_i	x_i	y_i	x_i	y_i
1	0,91	-0,31	0,92	-0,32	0,93	-0,33	0,94	-0,34	0,95	-0,35	0,96	-0,36
2	1,23	0,11	1,24	0,12	1,25	0,13	1,26	0,14	1,27	0,15	1,28	0,16
3	1,33	0,71	1,34	0,72	1,35	0,73	1,36	0,74	1,37	0,75	1,38	0,76
4	0,78	-0,29	0,79	-0,30	0,80	-0,31	0,81	-0,32	0,82	-0,33	0,83	-0,34
5	1,31	-0,26	1,32	-0,27	1,33	-0,28	1,34	-0,29	1,34	-0,30	1,35	-0,31
6	1,21	0,03	1,22	0,04	1,23	0,05	1,24	0,06	1,25	0,07	1,26	0,08
7	1,33	0,38	1,34	0,38	1,35	0,39	1,36	0,40	1,37	0,41	1,38	0,41
8	0,96	-0,71	0,97	-0,72	0,98	-0,73	0,99	-0,74	1,00	-0,75	1,01	-0,76
9	1,46	0,56	1,47	0,57	1,48	0,58	1,49	0,59	1,50	0,60	1,51	0,61
10	1,31	0,36	1,32	0,37	1,33	0,38	1,34	0,39	1,35	0,40	1,36	0,41
11	1,21	0,33	1,22	0,34	1,23	0,35	1,24	0,36	1,25	0,37	1,26	0,38
Вар.	**В7**		**В8**		**В9**		**В10**		**В11**		**В12**	
1	0,89	-0,29	0,88	-0,28	0,87	-0,27	0,86	-0,26	0,85	-0,25	0,84	-0,24
2	1,21	0,09	1,20	0,08	1,19	0,07	1,18	0,06	1,17	0,05	1,16	0,04
3	1,31	0,69	1,30	0,68	1,29	0,67	1,28	0,66	1,27	0,65	1,26	0,64
4	0,76	-0,27	0,75	-0,26	0,74	-0,25	0,73	-0,24	0,72	-0,23	0,71	-0,22
5	1,29	-0,24	1,28	-0,23	1,27	-0,22	1,26	-0,21	1,25	-0,20	1,24	-0,19
6	1,19	0,01	1,18	0	1,17	-0,01	1,16	-0,02	1,15	-0,03	1,14	-0,04
7	1,31	0,36	1,30	0,35	1,29	0,34	1,28	0,33	1,27	0,32	1,26	0,31
8	0,94	-0,69	0,93	-0,68	0,92	-0,67	0,91	-0,66	0,90	-0,65	0,89	-0,64
9	1,44	0,54	1,43	0,53	1,42	0,52	1,41	0,51	1,40	0,50	1,39	0,49
10	1,29	0,34	1,28	0,33	1,27	0,32	1,26	0,31	1,25	0,30	1,24	0,29
11	1,19	0,31	1,18	0,30	1,17	0,29	1,16	0,28	1,15	0,27	1,14	0,26
Вар.	**В13**		**В14**		**В15**		**В16**		**В17**		**В18**	
1	0,97	-0,37	0,98	-0,38	0,99	-0,39	1,00	-0,40	1,01	-0,41	1,02	-0,42
2	1,29	0,17	1,30	0,18	1,31	0,19	1,32	0,20	1,33	0,21	1,34	0,22
3	1,39	0,77	1,40	0,78	1,41	0,79	1,42	0,80	1,43	0,81	1,44	0,82
4	0,84	-0,35	0,85	-0,36	0,86	-0,37	0,87	-0,38	0,88	-0,39	0,89	-0,40
5	1,37	-0,32	1,38	-0,33	1,39	-0,34	1,40	-0,35	1,41	-0,36	1,42	-0,37
6	1,27	0,09	1,29	0,10	1,30	0,11	1,31	0,12	1,32	0,13	1,33	0,14
7	1,39	0,44	1,40	0,45	1,41	0,46	1,42	0,47	1,43	0,48	1,44	0,49
8	1,02	-0,77	1,03	-0,78	1,04	-0,79	1,05	-0,80	1,06	-0,81	1,07	-0,82
9	1,52	0,62	1,53	0,63	1,54	0,64	1,55	0,65	1,56	0,66	1,57	0,67
10	1,37	0,42	1,38	0,43	1,39	0,44	1,40	0,45	1,41	0,46	1,42	0,47
11	1,27	0,39	1,28	0,40	1,29	0,41	1,30	0,42	1,31	0,43	1,32	0,44

Таблиця 7 – Вихідні дані для визначення варіанту завдання**

N п	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Деталь X	15,656	15,657	15,645	15,678	15,659	15,673	15,645	15,648	15,672	15,669
Деталь Y	15,645	15,646	15,659	15,670	15,671	15,669	15,667	15,663	15,646	15,647
N п	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
Деталь X	15,645	15,646	15,647	15,648	15,649	15,650	15,651	15,652	15,653	15,654
Деталь Y	15,658	15,674	15,670	15,671	15,673	15,674	15,673	15,672	15,671	15,670
N п	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**
Деталь X	15,670	15,669	15,668	15,667	15,666	15,665	15,664	15,663	15,662	15,661
Деталь Y	15,654	15,655	15,656	15,657	15,658	15,659	15,660	1,661,	15,662	15,663
N п	**31**	**32**	**33**	**34**	**35**	**36**	**37**	**38**	**39**	**40**
Деталь X	15,661	15,662	15,663	15,664	15,665	15,667	15,668	15,669	15,670	15,671
Деталь Y	15,633	15,634	15,644	15,654	15,664	15,674	15,635	15,645	15,655	15,673
N п	**41**	**42**	**43**	**44**	**45**	**46**	**47**	**48**	**49**	**50**
Деталь X	15,672	15,662	15,652	15,673	15,663	15,652	15,643	15,645	15,555	15,655
Деталь Y	15,634	15,644	15,654	15,674	15,635	15,645	15,655	15,673	15,633	15,643
N п	**51**	**52**	**53**	**54**	**55**	**56**	**57**	**58**	**59**	**60**
Деталь X	15,678	15,668	15,664	15,658	15,648	15,646	15,656	15,662	15,675	15,645
Деталь Y	15,672	15,662	15,652	15,642	15,634	15,671	15,661	15,651	15,646	15,659
N п	**61**	**62**	**63**	**64**	**65**	**66**	**67**	**68**	**69**	**70**
Деталь X	15,673	15,663	15,653	15,644	15,677	15,667	15,657	15,647	15,670	15,660
Деталь Y	15,661	15,663	15,665	15,667	15,671	15,673	15,653	15,653	15,643	15,633
N п	**71**	**72**	**73**	**74**	**75**	**76**	**77**	**78**	**79**	**80**
Деталь X	15,645	15,674	15,654	15,664	15,646	15,656	15,666	15,676	15,677	15,646
Деталь Y	15,653	15,663	15,673	15,643	15,658	15,648	15,658	15,678	15,637	15,647
N п	81	82	83	84	85	86	87	88	89	90
Деталь X	15,649	15,659	15,678	15,668	15,669	15,648	15,659	15,658	15,677	15,676
Деталь Y	15,647	15,657	15,667	15,673	15,673	15,674	15,639	15,649	15,659	15,669
N п	91	92	93	94	95	96	97	98	99	100
Деталь X	15,676	15,666	15,656	15,646	15,641	15,651	15,661	15,671	15,648	15,664
Деталь Y	15,637	15,647	15,657	15,667	15,670	15,673	15,669	15,654	15,648	15,633

** – числові значення ***X***, ***Y*** пар 1-100 кожного варіанту встановлює викладач

ПРАКТИЧНА РОБОТА № 4

ВИРІВНЮВАННЯ ЕМПІРИЧНОЇ КРИВОЇ ЗАЛЕЖНИХ ПЕРЕМІННИХ ВЕЛИЧИН

Мета роботи: вивчити порядок проведення вирівнювання (апроксимації) кривих ломаних експериментальних залежностей за показовою, степеневою і логарифмічною функціями.

Теоретичні відомості

Після проведення експерименту дослідник отримує ряд статистичних величин (масив експериментальних даних), коли кожному значенню вхідного фактора $x_1;\ x_2;\ \dots\ ;\ x_N$ відповідає певне значення вихідної досліджуваної величини (функції) або параметру оптимізації $y_1;\ y_2;\ \dots\ ;\ y_N$ (дивитись. таблицю 1).

Таблиця 1 – Результати проведеного однофакторного експерименту

Кількість проведених дослідів N	1	2	3	4	...	N
Вхідний фактор X_i	X_1	X_2	X_3	X_4	…	X_N
Параметр оптимізації Y_i	Y_1	Y_2	Y_3	Y_4	…	Y_N

На підставі аналізу отриманого масиву експериментальних даних можливо підібрати таку математичну залежність $y = f(x)$, яка б описувала поведінку досліджуваних величин факторів та функції, тобто поведінку зміни досліджуваної функції від зміни вхідного фактора.

Подібний алгебраїчний вираз, тобто математична залежність, називається емпіричною залежністю або емпіричною формулою, а сам процес підбору - апроксимацією.

Емпіричні формули дають можливість оперативно визначати величину функції, тобто параметра оптимізації залежно від діючого фактора в широкому інтервалі його зміни. Бажано щоб емпіричні формули були простими, але якомога точніше описували процес, що досліджується.

Тому для підбору емпіричних формул використовують здебільшого елементарні функції, а для визначення коефіцієнтів та сталих цих функцій застосовують найбільш поширений спосіб найменших квадратів.

Процес підбору емпіричних формул складається з двох етапів:

І етап. Дані експерименту наносять на Декартову систему координат, з’єднують отримані точки ламаною лінією і підбирають орієнтовно вигляд математичної залежності (моделі, формули) об’єкта дослідження.

ІІ етап. Обчислюють невідомі коефіцієнти та сталі, що входять до обраної математичної формули.

Розглянемо методику визначення невідомих коефіцієнтів основних елементарних функцій.

Графіки основних елементарних алгебраїчних функцій наведені на рис. 1 та рис. 2.

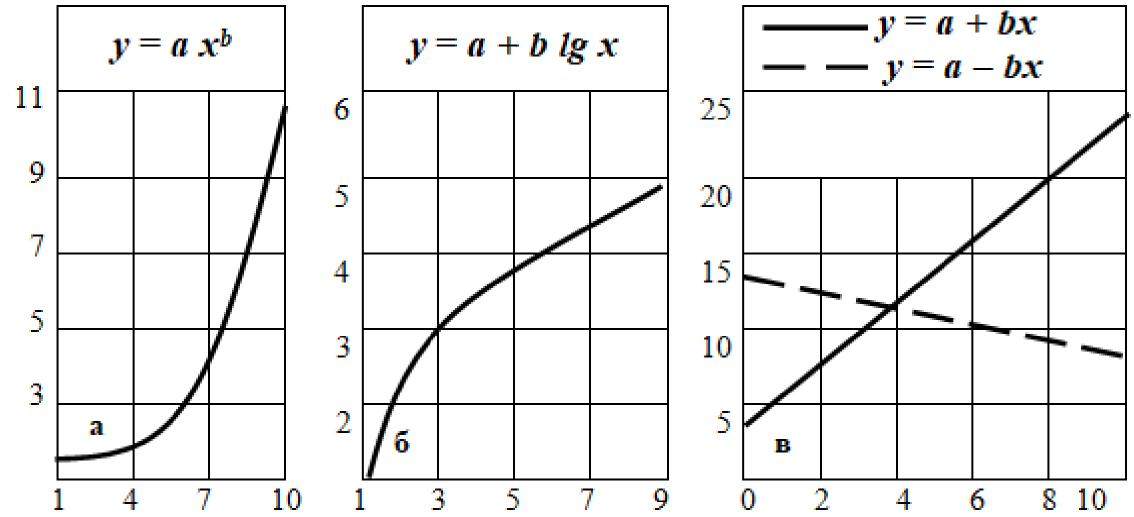


Рисунок 1 – Графічні залежності алгебраїчних функцій:
а – степеневої; б – логарифмічної; в – прямо пропорційної

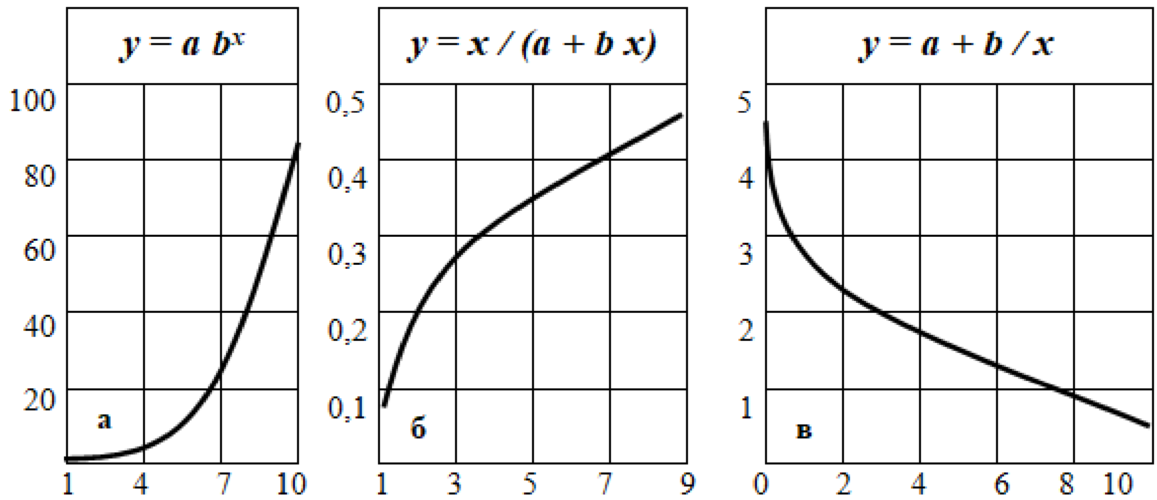


Рисунок 2 – Графічні залежності алгебраїчних функцій:
а – показової; б, в – зворотно пропорційної

В ряді випадків, для проведення апроксимації експериментального масиву даних, виникає необхідність підбору загальновідомих показової, степеневої і логарифмічної теоретичних функцій та визначення доцільності їх застосування за критерієм визначення основної похибки наближеного вирівнювання.

1). *Вирівнювання за показовою функцією*

Рівняння показової функції має вигляд:

$$y = ab^x, \text{ або } lg\, y = lg\, a + x\, lg\, b,$$

де a, b – невідомі постійні коефіцієнти.

Невідомі коефіцієнти a та b, що входять до емпіричної формули, визначаються за формулами:

$$lg\, a = \frac{\sum x_i^2 \sum lg\, y_i - \sum x_i \sum (x_i\, lg\, y_i)}{N \sum x_i^2 - (\sum x_i)^2}; \quad (1)$$

$$lg\, b = \frac{N \sum x_i\, lg\, y_i - \sum x_i \sum lg\, y_i}{N \sum x_i^2 - (\sum x_i)^2}. \quad (2)$$

2). *Вирівнювання за степеневою функцією*

Рівняння степеневої функції має вигляд

$$y = ax^b, \text{ або } lg\, y = lg\, a + b\, lg\, x,$$

де a, b – невідомі постійні коефіцієнти.

Невідомі коефіцієнти a та b, що входять до емпіричної формули, визначаються за виразами:

$$lg\, a = \frac{\sum lg^2 x_i \sum lg\, y_i - \sum lg\, x_i \sum lg\, x_i\, lg\, y_i}{N \sum lg^2 x_i - (\sum lg\, x_i)^2}; \quad (3)$$

$$b = \frac{N \sum lg\, x_i\, lg\, y_i - \sum lg\, x_i \sum lg\, y_i}{N \sum lg^2 x_i - (\sum lg\, x_i)^2}. \quad (4)$$

3). *Вирівнювання за логарифмічною функцією*

Рівняння логарифмічної функції має вигляд

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

де a, b – невідомі постійні коефіцієнти.

Невідомі коефіцієнти a та b, що входять до емпіричної формули визначаються за виразами:

$$a = \frac{\sum y_i \sum lg^2 x_i - \sum y_i \, lg\, x_i \sum lg\, x_i)}{N \sum lg^2 x_i \; - (\sum lg\, x_i)^2}; \tag{5}$$

$$b = \frac{N \sum y_i \, lg\, x_i - \sum y_i \sum lg\, x_i}{N \sum lg^2 x_i \; - (\sum lg\, x_i)^2}. \tag{6}$$

Коефіцієнти та сталі, які входять до останніх емпіричних формул визначаються із системи нормальних рівнянь, отриманих способом найменших квадратів і наведених в табл. 2.

Таблиця 2 – Вигляд нормальних рівнянь

Вихідне рівняння	Система нормальних рівнянь
$\boldsymbol{y = a + bx}$	$\boldsymbol{b\, \Sigma\, x_i^2 + a\, \Sigma\, x_i - \Sigma\, (x_i\, y_i) = 0}$ $\boldsymbol{b\, \Sigma\, x_i + a\, N - \Sigma\, y_i = 0}$
$\boldsymbol{y = a + b / x}$	$\boldsymbol{b\, \Sigma(1 / x_i^2) + a\, \Sigma\, (1 / x_i) - \Sigma\, (y_i / x_i) = 0}$ $\boldsymbol{b\, \Sigma\, (1/x) + aN - \Sigma y = 0}$
$\boldsymbol{y = x / (a + bx)}$	$\boldsymbol{a\, \Sigma(1 / x_i^2) + b\, \Sigma(1 / x_i) - \Sigma\, (1 / x_i\, y_i) = 0}$ $\boldsymbol{a\, \Sigma(1 / x_i) + b\, N - \Sigma\, (1 / y_i) = 0}$

Знак $\boldsymbol{\Sigma}$**,** наведений у формулах таблиці для визначення невідомих коефіцієнтів $\boldsymbol{a}$ та $\boldsymbol{b}$, означає суму експериментальних значень (вихідних даних для проведення розрахунків) від $\boldsymbol{i} = 1$ до $\boldsymbol{N}$, де $\boldsymbol{i}$ – порядковий номер досліду в експерименті; $\boldsymbol{N}$ – кількість проведених повторності (дослідів) в експерименті.

Розв'язати систему нормальних рівнянь відносно невідомих коефіцієнтів $\boldsymbol{a}$ та $\boldsymbol{b}$ можна виконати буд-яким загальноприйнятим способом.

При виконанні першого етапу може виникнути такий випадок , що для апроксимації експериментальних даних за своєю графічною побудовою (виглядом) орієнтовно підходять дві чи три функції.

Тому для оцінки ступеня точності вирівнювання (апроксимації), побудованої графічної залежності за результатами отриманих експериментальних даних, тією чи іншою вибраною апроксимуючою функцією, необхідно визначити величину так званої "нев'язки"

$$\boldsymbol{\Theta = \Sigma\varepsilon^2},$$

тобто сумарне середньоквадратичне відхилення теоретичних значень, обчислених за визначеною емпіричною формулою від експериментальних значень проведеного досліду для кожної з вибраних для проведення апроксимації формул.

Величина “нев’язки” визначається за формулою

$$\Theta = \sum_{i=1}^{N} \varepsilon^2 = \sum_{i=1}^{N} (y_{i_e} - y_{i_m})^2 , \quad (7)$$

де y_{ie} – експериментальні значення i-го досліду;

y_{im} – теоретичне значення i-го досліду, визначеного за одержаною емпіричною формулою.

Для подальшого використання залишають те емпіричне рівняння, для якого величина нев'язки” $\Theta = \Sigma\varepsilon^2$ найменша за числовим значенням, де знак Σ означає суму квадратів ε від i = 1 до N, де $\varepsilon = y_{i\,e} - y_{im}$.

Приклад. Підібрати емпіричну формулу для такого масиву експериментальних даних, де кількість повторності в експерименті дорівнює 8.

x	1	2	3	4	5	6	7	8
y	10	16	20	25	3О	35	50	80

Наносимо точки експериментальних даних на координатні осі і будуємо ламану лінію, тобто графік експериментальної залежності (рис 3).

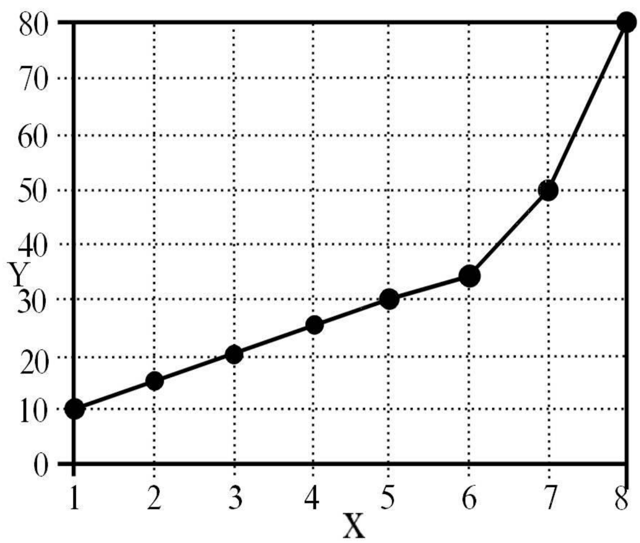


Рисунок 3 – Експериментальна залежність у вигляді ломаної прямої лінії

Орієнтовно, для описання поведінки побудованих експериментальних точок (експериментального масиву даних) обираємо дві функції: лінійну виду $y_1 = a + bx$ та другого порядку виду $y_2 = a + bx + cx^2$, тобто відповідно пряму лінію та параболу.

Після визначення методом найменших квадратів невідомих коефіцієнтів

a, ***b*** та ***c***, що входять до цих формул, записуємо одержані емпіричні формули:

$$\boldsymbol{y_1 = -5{,}1 + 8{,}5\,x};$$

$$\boldsymbol{y_2 = 17{,}6 - 5{,}1\,x + 1{,}5\,x^2}.$$

Перевіримо, яка з одержаних емпіричних формул $\mathbf{y_1 = -5{,}1 + 8{,}5\,x}$ та $\boldsymbol{y_2 = 17{,}6 - 5{,}1\,x + 1{,}5\,x^2}$, краще апроксимує (описує, вирівнює) побудовану експериментальну залежність, наведену на рис. 3.

Для цього визначимо показники, які наведені в табл. 3 і занесемо до неї результати цих обчислень.

Таблиця 3 – Результати апроксимації експериментальних даних.

N	Експериментальні дані X та $У$ і розрахунок вихідних даних				Апроксимація					
					$y_1 = -5{,}1 + 8{,}5\,x$			$y_2 = 17{,}6 - 5{,}1\,x + 1{,}5\,x^2$		
	x_i	y_i	$x_i\,y_i$	x_i^2	y_{im}	ε	ε^2	y_{im}	ε	ε^2
1	**1**	**10**	**10**	**1**	**3,4**	**6,6**	**43,6**	**14**	**4**	**16**
2	**2**	**15**	**30**	**4**	**11,9**	**3,1**	**9,6**	**13,4**	**1,6**	**2,56**
3	**3**	**20**	**60**	**9**	**19,5**	**0,4**	**0,16**	**15,8**	**4,2**	**17,6**
4	**4**	**25**	**100**	**16**	**23,4**	**3,9**	**15,2**	**21,2**	**3,8**	**14,4**
5	**5**	**30**	**150**	**25**	**37,1**	**7,1**	**50,4**	**29,6**	**0,4**	**0,16**
6	**6**	**35**	**210**	**36**	**45,5**	**10,5**	**110,2**	**41**	**6**	**36**
7	**7**	**50**	**350**	**49**	**54,4**	**4,4**	**19,4**	**55,4**	**5,4**	**29,1**
8	**8**	**80**	**640**	**64**	**62,9**	**17,1**	**282,4**	**72,8**	**7,2**	**51,8**
Σ	**36**	**265**	**1550**	**204**	**259,1**	**53,1**	**540,9**	**259,2**	**32,6**	**167,6**

Як бачимо, величина “нев’язки” $\boldsymbol{\Theta = \Sigma\varepsilon^2}$ для параболи становить 167,6, а для прямої – 540,9, тобто апроксимація параболою більш точно описує, або вирівнює експериментальну криву.

На рис. 4 зображено графік експериментальної залежності у вигляді ламаної лінії, точки якої побудовані за значеннями експериментального масиву даних та вибраної емпіричної залежності $\boldsymbol{y_2 = 17{,}6 - 5{,}1\,x + 1{,}5\,x^2}$, яка апроксимує експериментальну.

Рисунок 4 – Результати апроксимації експериментальних даних

Послідовність проведення математичних розрахунків і критерію доцільності проведення апроксимації, тією чи іншою функцією, розглянемо на наступних прикладах.

Порядок виконання роботи

1. Проведення обчислень за показовою функцією

Послідовність виконання розрахунків розглянемо на прикладі табл. 4.

Таблиця 4 – Вихідні дані і послідовність обчислень

N	x_i	y_i	lgy_i	$x_i lgy_i$	x_i^2	$lgy_i^{\prime}$	$y^{\prime}$	$(y_i - y_i^{\prime})^2$
1	2	3	4	5	6	7	8	9
1	1	5,59	0,7474	0,7474	1	0,7111	5,14	0,2025
2	4	6,27	0,7973	3,1892	16	0,8075	6,42	0,0225
3	5	6,30	0,7993	3,9965	25	0,8397	6,91	0,0372
4	7	8,25	0,9165	6,4155	49	0,9040	8,02	0,0529
5	9	8,88	0,9484	8,5356	81	0,9683	9,30	0,1764
6	12	11,21	1,0496	12,5952	144	1,0648	11,61	0,1600
7	13	13,62	1,1341	14,7433	169	1,0970	12,50	1,2544
Сума	51	60,12	6,3926	50,2227	485			2,2408

1.1. Занести в колонку 2 незалежні перемінні x_i, в колонку 3 – залежні y_i.

1.2. Обчислити і занести отримані дані в колонки 4 – 6.

1.3. Просумувати значення всіх колонок, заповнити строчку “сума” і отримати: $\sum x_i =$ 51; $\sum y_i =$ 60,12; $\sum lg\, y_i =$ 6,3926; $\sum x_i\, lg\, y_i =$ 50,2227; $\sum x_i^2 =$ 485.

1.4. Знайдені суми підставити в рівняння (1) і (2), тобто:

$$lg\, a = \frac{485 \cdot 6{,}3926 - 51 \cdot 50{,}2227}{7 \cdot 485 - 51^2} = 0{,}6789\,;$$

$$lg\, b = \frac{7 \cdot 50{,}2227 - 51 \cdot 6{,}3926}{7 \cdot 485 - 51^2} = 0{,}03216\,.$$

1.5. Знайти коефіцієнти ***a*** і ***b***, тобто ***a*** = 4,77; ***b*** = 1,08. Тоді записати вихідне рівняння отриманої емпіричної моделі, яке буде мати такий вигляд

$$y = 4{,}77 * 1{,}08^x.$$

1.6. Обчислити значення y_i^* вирівняної експериментальної кривої. Для цього в вираз $lg\, y_i^* = lg\, a + x_i\, lg\, b =$ 0,6789 + x_i * 0,03216 підставити значення $\boldsymbol{x_i}$ із колонки 2 і заповнити колонку 7.

1.7. Знайти значення y_i^* і заповнити колонку 8.

1.8. Знайти величину основної похибки. Для цього обчислити значення $(y_i - y_i^*)^2$, заповнити колонку 9 і знайти

$$\sum (y_i - y_i^*)^2 = 2{,}2408.$$

Основну похибку визначити за формулою

$$\sigma_n = \sqrt{\frac{\sum (y_i - y_i^*)^2}{N-1}}, \qquad (7)$$

де $\boldsymbol{N}$ – кількість проведених дослідів в експерименті.

Тоді

$$\sigma_n = \sqrt{2{,}2408 / (7-1)} = 0{,}61.$$

Цю величину (похибку) можна вважати малою і признавати вирівнювання задовільним, коли σ_n менше 0,1 $\bar{y}$.

2. Проведення обчислень за степеневою функцією

Послідовність виконання розрахунків розглянемо на прикладі табл. 5.

Таблиця 5 – Вихідні дані і послідовність обчислень

N	x_i	y_i	lgx_i	lgy_i	lg^2x_i	$lgx_i\, lgy_i$	lgy_i^*	y_i^*	$(y_i - y_i^*)^2$
1	2	3	4	5	6	7	8	9	10
1	1	5,59	0	0,7474	0	0	0,6774	4,76	0,6889
2	4	6,27	0,6021	0,7973	0,3625	0,4805	0,8651	7,33	1,1236
3	5	6,30	0,6990	0,7993	0,4886	0,5587	0,8953	7,86	2,4336
4	7	8,25	0,8451	0,9165	0,7142	0,7745	0,9409	8,73	0,2304
5	9	8,88	0,9542	0,9484	0,9105	0,9050	0,9749	9,44	0,3136
6	12	11,21	1,0792	1,0496	1,1861	1,1327	1,0139	10,35	0,7396
7	13	13,62	1,1139	1,1341	1,2408	1,2633	1,0247	10,68	8,7616
Сума	51	60,12	5,2935	6,3926	4,9027	5,1147			14,2913

2.1. Занести в колонку 2 незалежні перемінні $\boldsymbol{x_i}$, в колонку 3 – залежні $\boldsymbol{y_i}$.

2.2. Обчислити і занести отримані дані в колонки 4 – 7.

2.3. Просумувати значення всіх колонок, заповнити строчку “сума” і отримати:

$$\sum x_i = 51;$$

$$\sum y_i = 60{,}12;$$

$$\sum lg\, y_i = 6{,}3926;$$

$$\sum lg\, x_i = 5{,}2935;$$

$$\sum lg^2\, x_i = 4{,}9027;$$

$$\sum lgx_i\, lg\, y_i = 5{,}1147.$$

2.4. Знайдені суми підставити в рівняння (3) і (4), тобто визначити ***lga*** і коефіцієнт ***b***:

$$lg\, a = \frac{6{,}3926 \cdot 4{,}9027 - 5{,}2935 \cdot 5{,}1147}{7 \cdot 4{,}9027 - 5{,}2935^2} = 0{,}6774\,;$$

$$b = \frac{7 \cdot 5{,}1147 - 5{,}2935 \cdot 6{,}3926}{7 \cdot 4{,}9027 - 5{,}2935^{\,2}} = 0{,}3118\,.$$

2.5. Знайти коефіцієнти ***a***, тобто ***a*** = 0,1692.

Тоді записати вихідне рівняння отриманої емпіричної моделі, яке буде мати такий вигляд

$$y = 0{,}1692 * x^{0{,}3118}.$$

2.6. Обчислити значення y_i^* вирівняної експериментальної кривої.

Для цього в вираз $lg\, y_i^* = lg\, a + b\, lg\, x_i$ підставити значення ***lga*** і ***b***, тобто отримаємо

$$lg\, y_i^* = 0{,}6774 + 0{,}3118 * \, lg\, x_i\,. \tag{8}$$

Підставити рівняння (8) значення ***lgxi*** із колонки 4 і заповнити колонку 8.

2.7. Знайти вирівняні значення y_i^* і заповнити колонку 9.

2.8. Знайти величину основної похибки за формулою (7).

Для цього обчислити значення $(\, y_i - y_i^*\,)^2$, заповнити колонку 10 і знайти $\sum (\, y_i - y_i^*\,)^2 = 14{,}2913$.

Тоді $\sigma_c = \sqrt{14{,}2913 / (7 - 1)} = 1{,}54$.

Співставляючи значення σ_c з значенням основної похибки при вирівнюванні за показовою функцією σ_n, робимо висновок, що там помилка значно менша, тобто показова функція дає краще (більш точніше) приближення, ніж степенева.

3. Проведення обчислень за логарифмічною функцією

Послідовність виконання розрахунків розглянемо на прикладі табл. 6.

Таблиця 6 – Вихідні дані і послідовність обчислень

N	x_i	y_i	lgx_i	$y_i\, lgx_i$	lg^2x_i	$y_i^{/}$	$(y_i - y_i^{/})^2$
1	2	3	4	5	6	7	8
1	1	5,59	0	0	0	4,04	2,4025
2	4	6,27	0,6021	3,7752	0,3625	7,66	1,9321
3	5	6,30	0,6990	4,4037	0,4886	8,28	3,9204
4	7	8,25	0,8451	6,9721	0,7142	9,13	0,7744
5	9	8,88	0,9542	8,4573	0,9105	9,78	0,8100
6	12	11,21	1,0792	12,0978	1,1861	10,54	0,4489
7	13	13,62	1,1139	15,1713	1,2408	10,75	8,2369
Сума	51	60,12	5,2935	50,8774	4,9027		18,5252

3.1. Занести в колонку 2 незалежні перемінні x_i, в колонку 3 – залежні y_i.

3.2. Обчислити і занести отримані дані в колонки 4 – 6.

3.3. Просумувати значення всіх колонок, заповнити строчку "сума" і отримати: $\sum x_i = 51$; $\sum y_i = 60{,}12$; $y_i \sum lg\, x_i = 50{,}8774$; $\sum lg^2\, x_i = 4{,}9027$.

3.4. Знайдені суми підставити в рівняння (5) і (6), тобто визначити коефіцієнти ***a*** і ***b***:

$$a = \frac{60{,}12 \cdot 4{,}9027 - 50{,}8774 \cdot 5{,}2935}{7 \cdot 4{,}9027 - 5{,}2935^2} = 4{,}038;$$

$$b = \frac{7 \cdot 50{,}8774 - 60{,}12 \cdot 5{,}2935}{7 \cdot 4{,}9027 - 5{,}2935^2} = 6{,}017.$$

3.5. Знайти коефіцієнти ***a*** і ***b***, тобто ***a*** = 4,038; ***b*** = 6,017.

Тоді записати вихідне рівняння отриманої емпіричної формули (логарифмічної кривої), яке буде мати такий вигляд:

$$y_i^* = 4{,}04 + 6{,}017 * lg\, x_i.$$

3.6. Обчислити значення y_i^* вирівняної експериментальної кривої. Для цього в вираз y_i^* = 4,04 + 6,017 lgx_i підставити значення lgx_i із колонки 4 і заповнити колонку 7.

3.7. Знайти величину основної похибки за формулою (7).

Для цього обчислити значення $(y_i - y_i^*)^2$, заповнити колонку 8 і знайти

$$\sum (y_i - y_i^*)^2 = 18{,}5252.$$

Тоді $\sigma_л = \sqrt{18{,}5252/(7-1)} = 1{,}76.$

Співставляючи значення $\sigma_л$ з значенням основної похибки при вирівнювання за показовою $\sigma_п$ і степеневою σ_c функціями, робимо висновок, що в попередніх випадках помилка значно менша, тобто вирівнювання за логарифмічною функцією дає найгірші результати.

Порядок оформлення звіту

1. Записати вихідні дані для проведення розрахунків (експериментальний масив даних згідно з табл. 5).

2. Нанести експериментальні точки на координатну площину ХОУ. З'єднати їх відрізками прямих ліній (побудувати експериментальну ламану лінію).

3. Виходячи з вигляду експериментальної ламаної лінії, підібрати орієнтовно дві аналітичні залежності (функції), які приблизно описують побудований графік експериментальної залежності.

4. Для кожної з вибраних аналітичних залежностей визначити невідомі коефіцієнти в такій послідовності.

4.1. Визначити та записати систему нормальних рівнянь з табл. 2.

4.2. Розв'язати вибрану систему нормальних рівнянь відносно невідомих коефіцієнтів.

4.3. Підставити значення коефіцієнтів у відповідні формули.

Записати емпіричні формули в остаточному вигляді.

5. Встановити, яка з отриманих емпіричних формул краще апроксимує експериментальні дані. Для цього скласти таблицю за формою, подібною до форми табл. 3.

6. Шляхом порівняння "нев'язки" $\boldsymbol{\Theta = \Sigma\varepsilon^2}$ вибрати необхідну емпіричну залежність.

7. Побудувати за обраною емпіричною формулою теоретичну залежність $\boldsymbol{y = f(x)}$ на тих же координатних осях що і в п.2 (лінію виконати в іншому стилі чи кольорі, ніж експериментальну ламану).

Контрольні запитання

1. Дати визначення поняттю "масив експериментальних даних".

2. Апроксимація, її суть і призначення.
3. Для чого призначена система нормальних рівнянь?
4. Як визначається величина “нев’язки”?

Таблиця 5 – Вихідні дані для проведення розрахунків

Варіант 1

x	1	2	3	4	5	6	7	8	9	10
y	1.5	10	20	110	100	280	300	580	630	900

Варіант 2

x	1	2	3	4	5	6	7	8	9	10
y	5.5	4.03	2.66	2.75	2.11	2.36	1.77	2.16	1.67	1.9

Варіант 3

x	1	2	3	4	5	6	7	8	9	10
y	0.64	1.13	1.16	1.49	1.3	1.64	1.44	1.73	1.5	1.64

Варіант 4

x	1	2	3	4	5	6	7	8	9	10
y	5	10	9	15.6	15	20	25.3	24	31.4	32

Варіант 5

x	1	2	3	4	5	6	7	8	9	10
y	4.2	4.33	5.2	4.94	5.72	5.4	5.83	5.67	6.11	6.09

Варіант 6

x	1	2	3	4,	5	6	7	8	9	10
y	0.201	0.134	0.124	0.078	0.091	0.058	0.068	0.045	0.056	0.041

Варіант 7

x	1	2	3	4	5	6	7	8	9	10
y	1.5	8	12	20.8	21.5	45.6	31.4	67.1	53	92

Варіант 8

x	1	2	3	4	5	6	7	8	9	10
y	8	5.26	3.54	3.8	29	3.4	2.52	3.11	2.36	2.85

Варіант 9

x	1	2	3	4	5	6	7	8	9	10
y	0.34	0.58	0.63	0.84	0.76	0.96	0.87	1.01	0.93	1.06

Варіант 10

x	1	2	3	4	5	6	7	8	9	10
y	10.9	10	20.9	24	25.4	36.6	36.2	45.1	45.5	56

Варіант 11

x	1	2	3	4	5	6	7	8	9	10
y	0.89	1.68	2.57	2.4	3.4	2.8	3.36	4	3.6	4.29

Варіант 12

x	1	2	3	4	5	6	7	8	9	10

у	0.2	0.12	0.11	0.078	0.043	0.058	0.033	0.045	0,026	0.041

ПРАКТИЧНА РОБОТА № 5

ПАРАБОЛІЧНЕ ВИРІВНЮВАННЯ ЕМПІРИЧНОГО РОЗПОДІЛЕННЯ ЗА МЕТОДОМ ЧЕБІШЕВА

Мета роботи: вивчити порядок вирівнювання статистичних (емпіричних) кривих за способом найменших квадратів.

Теоретичні відомості

В деяких випадках для вирівнювання отриманого емпіричного розподілення, тобто побудованої за результатами проведеного експерименту емпіричної ломаної лінії (графічної залежності параметра оптимізації від вхідного фактора), доцільно не підбирати різні основні алгебраїчні функціональні залежності $y = f\ (x)$, а вирівнювати його за параболою тієї, чи іншої степені.

На практиці застосовують декілька способів параболічного вирівнювання емпіричних кривих. Найбільш поширеним і досконалим є спосіб Чебішева. За цим способом можна поступово підвищувати порядок (степінь) параболи, використовуючи в повному обсязі дані попередніх обчислень.

Нехай маємо емпіричну криву, яку побудовано за результатами проведеного експерименту і яку необхідно вирівняти за параболою тієї чи іншої степені.

Для цього необхідно знайти функцію $\mathrm{Y_T} = f(x)$, значення якої при $x = x_1, x_2, \ldots , x_N$ як можна менше відрізнялися від експериментальних (емпіричних) значень $Y_1, Y_2, \ldots , Y_N$.

Тому що більшість функцій може бути представлена в вигляді багаточлена *n*-ої степені, то при вирівнюванні доцільно представляти залежність між перемінними величинами в вигляді параболи *n*-ої степені $\mathrm{y} = \mathrm{a}_0 + a_1 x + a_2 x^2 + \ldots + a_N x^N$, де a_0, a_1, a_2, ..., a_N – невідомі параметри, або невідомі коефіцієнти.

Розглянемо випадок, коли значення незалежної змінної є рівновіддаленими, тобто $x_2 - x_1 = x_3 - x_2 = x_4 - x_3 = \ldots = x_N - x_{N-1}$.

Обчислення параболи необхідного порядку проводять в поступовому обчисленні ряду Чебішева за формулою

$$f(x)=\frac{\sum m_i}{N}+\frac{\sum m_i\phi_1(x_i)}{\sum \phi_1^2(x_i)}\phi_1(x)+\frac{\sum m_i\phi_2(x_i)}{\sum \phi_2^2(x_i)}\phi_2(x)+...+\frac{\sum m_i\phi_\lambda(x_i)}{\sum \phi_\lambda^2(x_i)}\phi_\lambda(x),$$

або

$$f(x)=k_0+k_1\phi_1(x)+k_2\phi_2(x)+...+k_\lambda\phi_\lambda(x), \qquad (1)$$

де k_0 – вільний член; $k_1, k_2, ..., k_\lambda$ – відповідно, невідомі коефіцієнти.

В цій формулі величина $\lambda \leq N-1$ характеризує порядок параболи, $\boldsymbol{N}$ – число інтервалів (кількість значень незалежної перемінної $\boldsymbol{x}$, або кількість проведених дослідів в експерименті, або членів ряду Чебішева).

Значення $\varphi_\lambda(x_i)$ і $\sum\varphi_\lambda^2(x_i)$ для того чи іншого значення $\boldsymbol{N}$ обчислені попередньо і наведені в таблицях коефіцієнтів поліномів Чебішева, які приведені в табл. В (табл. В надає викладач).

Ці таблиці дають можливість знаходити параболу до 5-го порядку.

Для зручності обчислень в табл. В дані цілі числа $\phi_1, \phi_2, \phi_3, ..., \phi_5$.

В назвах (заголовках) табл. Г (табл. Г надає викладач) вказані коефіцієнти C_1, C_2, C_3, ... , C_5 дозволяють виконувати перехід від дробових чисел до цілих.

Тому необхідно для отримання дійсних чисел $\dfrac{\sum C_\lambda\varphi_\lambda y_i}{\sum(C_\lambda\varphi_\lambda)^2}$, знайдене із таблиць помножити на $C_\lambda\phi_\lambda(x)$, тобто помножити на добуток коефіцієнта C_λ, який вказаний в назві λ-го стовпця таблиць (табл. Г) даного значення N на вираз $\phi_\lambda(x)$, який дорівнює:

$$\phi_1(x)=(x-\frac{N+1}{2});$$

$$\varphi_2(x)=\left[(x-\frac{N+1}{2})^2-\frac{N^2-1}{12}\right];$$

$$\varphi_3(x)=\left[(x-\frac{N+1}{2})^3-\frac{3N^2-7}{20}(x-\frac{N+1}{2})\right];$$

$$\phi_4(x)=\left[(x-\frac{N+1}{2})^4-\frac{1}{14}(3N^2-13)(x-\frac{N+1}{2})^2+\frac{3(N^2-1)(N^2-9)}{560}\right];$$

$$\varphi_5(x)=(x-\frac{N+1}{2})^5-\frac{5}{18}(N^2-7)(x-\frac{N+1}{2})^3+$$
$$+\frac{1}{1008}(15N^4-230N^2+407)(x-\frac{N+1}{2}).$$

Похибки наближення параболи обчислюють за формулою

$$\sigma_\lambda = \sqrt{\frac{\sum_\lambda}{N-(\lambda+1)}}, \text{ де } \sum{}_\lambda = \sum{}_{\lambda-1} - \frac{(\sum y_i C_\lambda \phi_\lambda)^2}{\sum (C_\lambda \phi_\lambda)^2} = \sum{}_{\lambda-1} - k_\lambda^2 \sum (C_\lambda \phi_\lambda)^2,$$

де k_λ – коефіцієнт при x^λ даної параболи $\sum{}_0 = \sum y_i^2 - \frac{(\sum y_i)^2}{N}$.

Послідовність проведення обчислення розглянемо на прикладі складеної табл. 1.

Вихідними даними для проведення розрахунків є колонки 1-2 табл. 1 в яких наведено числові значення, що отримані в результаті проведення експерименту.

Таблиця 1 – Вихідні дані і порядок проведення обчислень

3	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
m_i^2	$C_1\phi_1$	$m_iC_1\phi_1$	$C_2\phi_2$	$m_iC_2\phi_2$	$C_3\phi_3$	$m_iC_3\phi_3$	$C_4\phi_4$	$m_iC_4\phi_4$	$3m_i$
9	-13	-39	+13	+39	-143	-429	+143	+429	-
64	-11	-88	+7	+56	-11	-88	-77	-616	9,8
121	-9	-99	+2	+22	+66	+726	-132	-1452	20,0
400	-7	-140	-2	-40	+98	+1960	-92	-1840	23,2
729	-5	-135	-5	-135	+95	+2565	-13	-351	26,8
1296	-3	-108	-7	-252	+67	+2412	+63	+2268	26,3
841	-1	-29	-8	-232	+24	+696	+108	+3132	25,0
324	+1	+18	-8	-144	-24	-432	+108	+1944	22,2
289	+3	+51	-7	-119	-67	-1139	+63	+1071	18,5
289	+5	+85	-5	-85	-95	-1615	-13	-221	14,2
64	+7	+56	-2	-16	-98	-784	-92	-736	9,6
16	+9	+36	+2	+8	-66	-264	-132	-528	5,2
1	+11	+11	+7	+7	+11	+11	-77	-77	1,3
1	+13	+13	+13	+13	+143	+143	+143	+143	
4444		-368		-878		+3762		+4166	

1	2
x_i	m_i
1	3
2	8
3	11
4	20
5	27
6	36
7	29
8	18
9	17
10	17
11	8
12	4
13	1
14	1
Σ	200

Порядок виконання роботи

1. Визначення параболи нульового порядку

1.1. Визначити $\sum m_i$ і заповнити колонку 2 табл. 1, при цьому

$$\sum_{i=1}^{N} m_i = m_1 + m_2 + ... + m_N = 3 + 8 + ... + 1 = 200.$$

1.2. Обчислити середнє значення $\overline{m}$ за формулою

$$\overline{m} = \frac{\sum_{i-1}^{N} m_i}{N}, \text{ при цьому } \overline{m} = 200 / 14 = 14{,}29.$$

1.3. Скласти рівняння нульового порядку, тобто $_0 f(x) = 14{,}29$ (індекс перед $f(x)$ тут і в подальшому вказує на порядок кривої, тобто параболи).

1.4. Обчислити суму квадратів $(\sum_{i=1}^{N} m_i)^2$ і заповнити колонку 3 табл. 1, при цьому $(\sum_{i=1}^{N} m_i)^2$ = 4444.

1.5. Знайти суму квадратів різниць між експериментальними значеннями і визначеним рівнянням нульового порядку за формулою

$$\sum{}_0 = \sum_{i=1}^{N} m_i^2 - \frac{(\sum_{i=1}^{N} m_i)^2}{N}, \text{ при цьому } \sum{}_0 = 4444 - 200^2 / 14 = 1586{,}86.$$

1.6. Обчислити основну похибку за формулою

$$\sigma_0 = \sqrt{\frac{\sum{}_0}{N-1}} = \sqrt{\frac{1586{,}86}{14-1}} = 11{,}05.$$

2. Визначення параболи 1-го порядку

2.1. За додатком К знайти $C_1\phi_1$ для кількості дослідів N = 14 і заповнити колонку 4 табл. 1.

2.2. Обчислити добутки m_i на $C_1\phi_1$ і заповнити колонку 5 табл. 1.

2.3. Обчислити суму $\sum_{i=1}^{N} m_i C_1\phi_1$ = –368 і поділити її на суму квадратів $\sum (C_1\phi_1)^2$, яка взяти внизу першої колонки $2\phi_1$ додатку К для N = 14, тобто 910.

При цьому $\frac{\sum_{i=1}^{N} m_i C_1\phi_1}{\sum (C_1\phi_1)^2}$ = –368 / 910 = –0,4044.

2.4. Помножити, отриманий в п. 1.2.3 частковий вираз на $C_1(x - \frac{N+1}{2})$, при цьому коефіцієнт C_1 визначити за таблицею додатку К із якої видно, що в першій колонці $2\phi_1$ коефіцієнт C_1 при ϕ_1 дорівнює 2, тобто C_1 = 2 для N = 14.

$$\frac{\sum_{i=1}^{N} m_i C_1\phi_1}{\sum (C_1\phi_1)^2} C_1(x - \frac{N+1}{2}) = -0{,}4044 \cdot 2(x - \frac{14+1}{2}) = 6{,}066 - 0{,}8088\,x.$$

2.5. Знайдений результат прибавити до правої частини рівняння параболи нульового порядку ${}_0 f(x) = 14{,}29$ (див. п.1.1.3), тобто скласти (додати) між собою значення вільних членів

14,29 + 6,066 = 20,356.

2.6. Записати рівняння параболи першого порядку, тобто

$${}_1 f(x) = 20{,}356 - 0{,}8088\,x. \qquad (2)$$

2.7. Обчислити основну похибку за формулою $\sigma_1 = \sqrt{\frac{\sum_1}{N-2}}$, де $\sum_1 = \sum_0 - k_1^2 \sum (C_1\phi_1)^2$, а визначений коефіцієнт k_1 згідно виразів (1) і (2) буде дорівнювати k_1 = 0,8088.

Тоді: Σ_1 = 1586,86 – 0,8088^2 * 910 = 991,57; $\sigma_1 = \sqrt{\frac{991{,}57}{14-2}} = 9{,}1.$

3. Визначення параболи 2-го порядку

3.1. За додатком К знайти $C_2\phi_2$ для кількості дослідів N = 14 і заповнити колонку 6 табл. 1.

3.2. Обчислити добутки m_i на $C_2\varphi_2$ і заповнити колонку 7 табл. 1.

3.3. Обчислити суму $\sum_{i=1}^{N} m_i C_2 \phi_2$ = 878 і поділити її на суму квадратів $\sum (C_2 \phi_2)^2$, яка взяти внизу другої колонки $\frac{1}{2}\phi_2$ додатку К для N = 14, тобто 728, при цьому $\dfrac{\sum_{i=1}^{N} m_i C_2 \phi_2}{\sum (C_2 \phi_2)^2}$ = –878 / 728 = –1,206.

3.4. Помножити, отриманий в п. 1.3.3 частковий вираз на $C_2[\,(x-\frac{N+1}{2})^2-\frac{N^2-1}{12}\,]$, при цьому коефіцієнт C_2 визначити за таблицею додатку К із якої видно, що в другій колонці $\frac{1}{2}\phi_2$ коефіцієнт C_2 при ϕ_2 дорівнює 1/2, тобто C_2 = 0,5 для N = 14.

$$\frac{\sum_{i=1}^{N} m_i C_2 \phi_2}{\sum (C_2 \phi_2)^2} C_2[\ (x-\frac{N+1}{2})^2-\frac{N^2-2}{12}\ ]=$$

$$=-1{,}206\cdot 0{,}5[\ (x-\frac{14+1}{2})^2-\frac{14^2-1}{12}\ ]=-24{,}12+9{,}045x-0{,}603x^2.$$

3.5. Знайдений результат прибавити до правої частини рівняння (2) параболи першого порядку $_1 f(x)=20{,}356-0{,}8088x$ (див. п.1.2.6), тобто скласти (додати) між собою значення вільних членів 20,356 + (–24,12) = –3,764 і значення коефіцієнтів при x першої степені (–0,809X) + 9,045x = 8,236x .

3.6. Записати рівняння параболи другого порядку, тобто

$$_2 f(x)=-3{,}764+8{,}236x-0{,}603x^2. \qquad (3)$$

3.7. Обчислити основну похибку за формулою $\sigma_2=\sqrt{\frac{\sum_2}{N-3}}$, де $\sum_2=\sum_1-k_2^2\sum (C_2\phi_2)^2$, а визначений коефіцієнт k_2 згідно виразів (1) і (3) буде дорівнювати k_2 = 0,603.

Тоді

$$\Sigma_2 = 991{,}57 - 0{,}603^2 * 728 = 726{,}86;$$

$$\sigma_2=\sqrt{\frac{726{,}86}{14-3}}=8{,}1.$$

Похибка σ зменшилась, продовжуємо обчислення.

4. Визначення параболи 3-го порядку

4.1. За додатком К знайти $C_3\phi_3$ для кількості дослідів N = 14 і заповнити колонку 8 табл. 1.

4.2. Обчислити добутки m_i на $C_3\phi_3$ і заповнити колонку 9 табл. 1.

4.3. Обчислити суму $\sum_{i=1}^{N} m_i C_3\phi_3$ = +3762 і поділити її на суму квадратів $\sum (C_3\phi_3)^2$, яка взяти внизу третьої колонки $\frac{5}{3}\phi_3$ додатку К для N = 14, тобто 97240, при цьому $\frac{\sum_{i=1}^{N} m_i C_3\phi_3}{\sum (C_3\phi_3)^2} =$ +3762 / 97240 = = 0,0387.

4.4. Помножити, отриманий в п. 1.4.3 частковий вираз на $C_3[\ (x-\frac{N+1}{2})^3 - \frac{3N^2-7}{20}(x-\frac{N+1}{2})\ ]$.

При цьому коефіцієнт C_3 визначити за таблицею додатку К із якої видно, що в третій колонці $\frac{5}{3}\phi_3$ коефіцієнт C_3 при ϕ_3 дорівнює 5/3, тобто C_3 = 5/3 для N = 14.

$$\frac{\sum_{i=1}^{N} m_i C_3\phi_3}{\sum (C_3\phi_3)^2} C_3[\ (x-\frac{N+1}{2})^3 - \frac{3N^2-7}{20}(x-\frac{N+1}{2})\ ] =$$

$$= 0{,}0387 \cdot \frac{5}{3}[\ (x-\frac{14+1}{2})^3 - \frac{3\cdot 14^2-7}{20}(x-\frac{14+1}{2}\ ] =$$

$$= -13{,}158 + 9{,}011x - 1{,}451x^2 + 0{,}064x^3.$$

4.5. Знайдений результат прибавити до правої частини рівняння (3) параболи другого порядку $_2 f(x) = -3{,}764 + 8{,}236x - 0{,}603x^2$ (див. п. 1.3.6), тобто скласти (додати) між собою значення вільних членів (–3,764) + (–13,158) = –16,922, значення коефіцієнтів при x першої степені 8,236X + 9,011x = 17,247x і значення коефіцієнтів при x^2 другої степені

(–0,603x^2) + (–1,*451**x^2) = = –2,*054*x^2.

4.6. Записати рівняння параболи третього порядку, тобто

$$ {}_3 f(x) = -16{,}922 + 17{,}247x - 2{,}054x^2 + 0{,}0645x^3. \tag{4}$$

4.7. Обчислити основну похибку за формулою $\sigma_3 = \sqrt{\dfrac{\sum_3}{N-4}}$, де $\sum_3 = \sum_2 - k_3^2 \sum (C_3\phi_3)^2$, а визначений коефіцієнт k_3 згідно виразів (1) і (4) буде дорівнювати k_3 = 0,0645.

Тоді:

$$\Sigma_3 = 726{,}86 - 0{,}0645^2 * 97240 = 404{,}52;$$

$$\sigma_3 = \sqrt{\frac{404{,}52}{14-4}} = 6{,}4.$$

Похибка σ зменшилась, продовжуємо далі обчислення.

5. Визначення параболи 4-го порядку

5.1. За додатком К знайти $C_4\phi_4$ для кількості дослідів N = 14 і заповнити колонку 10 табл. 1.

5.2. Обчислити добутки m_i на $C_4\phi_4$ і заповнити колонку 11 табл. 1.

5.3. Обчислити суму $\sum_{i=1}^{N} m_i C_4\phi_4 = +4166$ і поділити її на суму квадратів $\sum (C_4\phi_4)^2$, яка взяти внизу четвертої колонки $\frac{7}{12}\phi_4$ додатку К для N = 14, тобто 136136.

При цьому

$$\frac{\sum_{i=1}^{N} m_i C_4\phi_4}{\sum (C_4\phi_4)^2} = +4166 / 136136 = 0{,}0306.$$

5.4. Помножити, отриманий в п. 1.5.3 частковий вираз на $C_4[\ (x - \frac{N+1}{2})^4 - \frac{1}{14}(3N^2 - 13)(x - \frac{N+1}{2})^2 + \frac{3(N^2-1)(N^2-9)}{560}\ ]$, при цьому коефіцієнт C_4 визначити за таблицею додатку К із якої видно, що в четвертій колонці $\frac{7}{12}\phi_4$ коефіцієнт C_4 при ϕ_4 дорівнює 7/12, тобто

$$C_4 = 7/12 \text{ для } N = 14.$$

$$\frac{\sum_{i=1}^{N} m_i C_4 \phi_4}{\sum (C_4\phi_4)^2} C_4 \Big[\Big(x - \frac{N+1}{2}\Big)^4 - \frac{1}{14}(3N^2 - 13)\Big(x - \frac{N+1}{2}\Big)^2 +$$

$$+ \frac{3(N^2-1)(N^2-9)}{560} \Big] =$$

$$= 0{,}0306 \cdot \frac{7}{12} \Big[\Big(x - \frac{14+1}{2}\Big)^4 - \frac{1}{14}(3 \cdot 14^2 - 13)\Big(x - \frac{14+1}{2}\Big)^2 +$$

$$+ \frac{3(14^2-1)(14^2-9)}{560} \Big] =$$

$$= -18{,}73 - 19{,}12x + 5{,}29x^2 - 0{,}5355x^3 + 0{,}01785x^4.$$

5.5. Знайдений результат прибавити до правої частини одержаного раніше рівняння (4) параболи третього порядку (див. п. 1.3.6) $_3 f(x) = -16{,}922 + 17{,}247x - 2{,}054x^2 + 0{,}0645x^3$, тобто скласти (додати) між собою значення:

- вільних членів (–16,922) + …+ 18,73 = 1,81;
- значення коефіцієнтів при x першої степені

$$17{,}25X + (-19{,}12X) = (-1{,}87X);$$

- значення коефіцієнтів при x^2 другої степені

$$(-2{,}054X^2) + 5{,}29X^2) = 3{,}24X^2;$$

- значення коефіцієнтів при x^3 третьої степені

$$0{,}0645X^3 + 0{,}5355X^3 = 0{,}471X^3.$$

5.6. Записати рівняння параболи четвертого порядку, тобто

$$_4 f(x) = 1{,}81 - 1{,}87x + 3{,}24x^2 - 0{,}471x^3 + 0{,}01785x^4. \qquad (5)$$

5.7. Обчислити основну похибку за формулою $\sigma_4 = \sqrt{\frac{\sum_4}{N-5}}$, де $\sum_4 = \sum_3 - k_4^2 \sum (C_4\phi_4)^2$, а визначений коефіцієнт k_4 згідно виразів (1) і (5) буде дорівнювати $k_4 = 0{,}01785$.

Тоді

$$\Sigma_4 = 404{,}52 - 0{,}01785^2 * 136136 = 361{,}4;$$

$$\sigma_4 = \sqrt{\frac{361{,}4}{14-5}} = 6{,}33.$$

Тому що σ_4 незначно відрізняється від σ_3, то обчислювати параболу 5-го порядку немає необхідності.

5.8. Обчислити значення ${}_3m_1$ і ${}_4m_i$ відповідно за рівняннями параболи 3-го (4) і 4-го (5) порядків шляхом підстановки в знайденні рівняння значень X = 1, 2, 3, 4, 5, ... , 14 та заповнити колонки 12 і 13 табл. 1.

Наприклад, рівняння параболи 4-го порядку має вигляд

$$ {}_4f(x) = 1{,}81 - 1{,}87x + 3{,}24x^2 - 0{,}471x^3 + 0{,}01785x^4. $$

Тоді значення ${}_4m_i$ будуть дорівнювати:

$${}_4m_i = 1{,}81 – 1{,}87 * 1 + 3{,}24 * 1^2 – 0{,}471 * 1^3 + 0{,}01785 * 1^4 = 2{,}7;$$

$${}_4m_2 = 1{,}81 – 1{,}87 * 2 + 3{,}24 * 2^2 – 0{,}471 * 2^3 + 0{,}01785 * 2^4 = 8{,}4;$$

$${}_4m_3 = 1{,}81 – 1{,}87 * 3 + 3{,}24 * 3^2 – 0{,}471 * 3^3 + 0{,}01785 * 3^4 = 14{,}1 \text{ і т. д.}$$

<u>Порядок оформлення звіту</u>

1. Записати вихідні дані для проведення розрахунків згідно табл. 1. Відповідний варіант завдання взяти згідно завдання, яке видає викладач.
2. Провести і записати математичні розрахунки згідно п. 1 “Порядок виконання роботи” та заповнити табл. 1.
3. Побудувати графік експериментальної залежності в вигляді ломаної лінії.
4. На тих же координатних осях побудувати графіки отриманих парабол 1-го, 2-го, 3-го та 4-го порядку. Побудову графіків виконати в різних кольорах.
5. Зробити висновок за результатами проведеної апроксимації експериментальних даних.

Контрольні запитання

1. Способи параболічного вирівнювання експериментального масиву даних (емпіричних кривих).
2. Записати у загальному вигляді функціональні залежності парабол першого, другого, третього та четвертого порядку.
3. Записати у загальному вигляді формулу, яка характеризує ряд Чебишева.
4. Що таке основна похибка, формула за якою її визначають.

ПРАКТИЧНА РОБОТА № 6

СТАТИСТИЧНА ОБРОБКА ОСЦИЛОГРАМ

Мета роботи: навчитись методам і порядку обробки осцилограм, отриманих при експериментальних дослідженнях випадкового процесу.

Теоретичні відомості

Більшість технологічних машин і агрегатів працюють в умовах, коли навантаження на їх робочі органи мають випадковий (непостійний і нерівномірний) характер. Наприклад, дія нерівномірності швидкості та глибини різання при точінні деталі, частоти обертання різця тощо.

Зміна дії випадкового процесу в часі чи просторі називається ***реалізацією***, а запис (фіксація результатів зміни) реалізації на реєструючому пристрої (моніторі, екрані і т.д.) – ***осцилограмою*** запису протікання процесу.

Якщо записувати один і той же процес декілька разів, то будемо мати ***ансамбль реалізацій*** випадкового процесу.

На рис. 1 подано ансамбль з трьох реалізацій випадкового процесу зміни нерівномірності швидкості точіння z_i за декілька проходів I_i різця по деталі, або об'єкта дослідження.

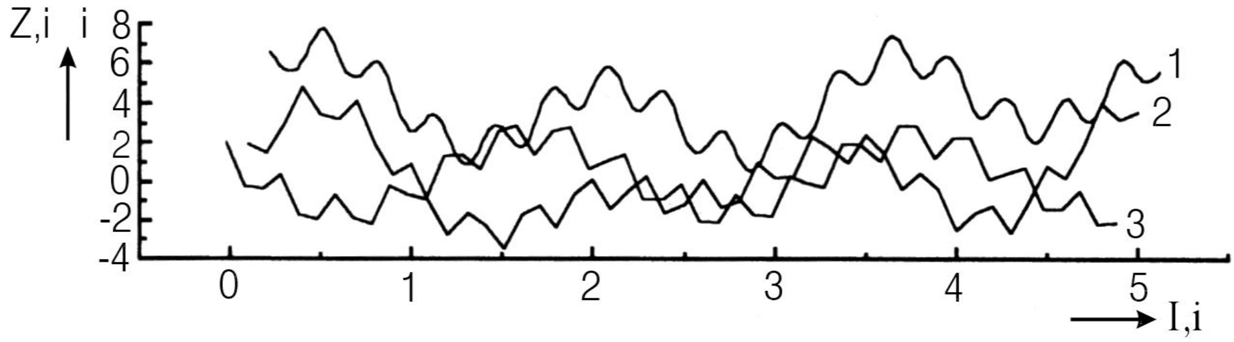


Рисунок 1 – Ансамбль реалізацій випадкового процесу точіння деталі

На перший погляд здається, що кожна з наведених ліній докорінно відрізняється від інших. Але водночас ми усвідомлюємо, що кожна з цих ліній характеризує нерівність швидкості різання під час точіння однієї і тієї ж деталі.

Тому за умов достатньої довжини реалізації кожна з них несе спільну загальну інформацію (реалізації мають спільний внутрішній стан) про випадковий процес.

Загальноприйнятими характеристиками випадкового процесу виступають:

- математичне сподівання m;

- середнє квадратичне відхилення σ;

- коефіцієнт варіації V.

Для обчислення цих величин треба мати конкретні значення ординат випадкової величини, тобто експериментальний масив даних. Тому необхідно провести ***дискретизацію*** осцилограми безперервного випадкового процесу. Суть методики проведення дискретизації пояснюється на рис. 2.

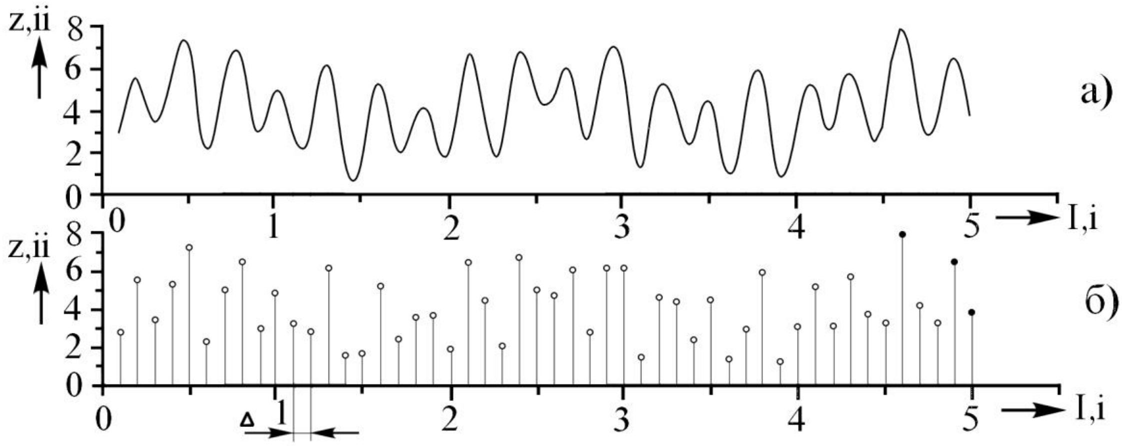


Рисунок 2 – Дискретизація реалізації випадкового процесу:

а – безперервний тип; б – дискретний тип

Будь-який випадковий процес характеризується випадковими величинами, які завжди присутні в ході проведення експерименту.

Випадковою величиною називають величину, яка в результаті досліду може приймати різні величини. Випадкові величини, як правило, позначають великими літерами, наприклад X, а значення випадкової величини, яке вона приймає в результаті досліду, – малими літерами $x_1, x_2, \ldots, x_n$.

При великій кількості досліджень кожне із можливих значень випадкової величини $x_1, x_2, \ldots, x_n$ може зустрітися $b_1, b_2, \ldots, b_n$ разів.

Ці числа називають ***частотами***.

Якщо всього було проведено N вимірів, тобто $\sum_{i=1}^{n} b_i = N$, то відношення b_i / N називають ***відносною частотою***.

Сукупність, яка містить всі досліджувані величини, називається ***генеральною сукупністю***.

Вибрані із генеральної сукупності N вимірів утворюють ***вибірку об'єму N***.

Дискретними випадковими величинами називають такі величини, які можуть приймати лише визначені фіксовані значення, наприклад, 0,1; 0,2 і т.д.

Безперервними випадковими величинами називають такі величини, які в деякому інтервалі можуть приймати будь-які значення.

Наприклад, кількість бракованих деталей у різних вибірках із генеральної сукупності є дискретною випадковою величиною, а розмір цих деталей – безперервною випадковою величиною.

Всяку безперервну випадкову величину можна задати у вигляді дискретної, якщо всі можливі її значення розбити на ***інтервали*** і задати ***вірогідність появи цих інтервалів*** (визначення ординат випадкового процесу).

Кількість відрізків $\boldsymbol{\Delta}$ вздовж осі абсцис (тобто величина $\boldsymbol{\Delta}$) визначається, виходячи з максимальної частоти перебігу випадкового процесу (при виконанні роботи користуватися рекомендаціями розділу "Порядок виконання роботи", пункт 2).

Після проведення дискретизації і вимірювання ординат випадкового процесу отримаємо ***масив експериментальних даних***. Якщо ці дані розподілити на ***класи***, то можна побудувати графічне зображення експериментальних розподілів, для чого будуються ***гістограми* і *полігони розподілу частот*** з'явлення тієї чи іншої величини випадкового процесу.

Для графічного відображення експериментального масиву даних будуються ***гістограми*** і ***полігони розподілу***.

Для випадкових величин дискретного типу застосовуються, як правило, полігони розподілу, а для випадкових величин безперервного типу – ***гістограми***.

Розподіл випадкового процесу на класи можна виконати за правилом Штюргеса.

Кількість класів випадкового процесу визначають за формулою:

$$\boldsymbol{k = 1 + 3{,}32\, lg n,} \qquad (1)$$

де $\boldsymbol{n}$ – кількість вибраних ординат в експериментальному масиві даних випадкового процесу.

Результат обчислень за формулою (1) округлюють до цілого числа.

Інтервал кожного класу визначають за наступною формулою

$$\delta = \frac{Z_{max} - Z_{min}}{k}, \qquad (2)$$

де $\boldsymbol{Z_{max}}$ та $\boldsymbol{Z_{min}}$ – відповідно максимальне та мінімальне чисельне значення виміряної ординати одного класу з експериментального масиву даних.

Полігони розподілів і гістограми будують за частотами розподілу експериментальних даних випадкового процесу.

Методика підрахування ***частот*** пояснюється табл. 1.

Таблиця 1 – Розбивка масиву експериментальних даних на класи та методика обчислення частот розподілу

Клас i*	Межі інтервалів класів	Середини інтервалів, z_i	Частота		
			абсолютна B_i	відносна b_i	накопичена $b_i^н$
1	від z_{min} до $z_{min}+\delta_{викл}$**				
2	від $z_{min}+\delta$ до $z_{min}+2\delta_{викл}$				
.....					
k	від $z_{min}+(k-1)\delta$ до $z_{max\ вкл}$***				

Примітка: * – i = 1, 2, ... k; ***викл* – виключно; ****вкл* – включно.

Середини інтервалів кожного визначеного класу обчислюють за формулою

$$\bar{Z}_i = 0{,}5\left(Z_{H_i} + Z_{B_i}\right), \tag{3}$$

де Z_{Hi} та Z_{Bi} – відповідно нижня та верхня межа i-го класу.

Абсолютна частота B_i – це кількість ординат (вимірів) з масиву даних, що потрапили до меж i-го класу.

Відносна частота b_i визначається за формулою

$$b_i = \frac{B_i}{n}. \tag{4}$$

Накопичена частота i-го класу $b_i^н$ складається з суми відносних частот i-го та всіх попередніх класів.

<u>Приклад</u> розбивки масиву експериментальних даних на класи та обчислення частот.

Таблиця 2 – Масив експериментальних вихідних даних

183	170	176	178	176	180	176	185	184	174	168	174	189
172	175	167	179	176	169	178	169	171	170	177	176	179
174	176	188	178	172	176	167	166	180	183	176	182	178
172	185	183	175	174	180	165	169	171	178	169	170	179
171	178	173	177									

Для побудови гістограми по осі абсцис відкладають у вибраному масштабі інтервали класів ***"від"*** і ***"до"***, а по осі ординат пропорційно частотам висоти прямокутників.

Гістограма зображує диференційний закон розподілення випадкової величини.

Для побудови полігону розподілу по осі абсцис відкладають значення випадкової величини, а по осі ординат – значення величини, пропорційної накопиченій частоті. Сума ординат дорівнює одиниці.

Всі данні для побудови беруть із табл. 3.

Таблиця 3. Таблиця прийнятих класів та знайдених за формулами частот

№ класу, *i*	Межі інтервалів класів	Середини інтервалів, Z_i	Частота		
			абсолютна B_i	відносна b_i	накопичена $b_i^н$
1	Від 165 до 169 викл.	167	5	0,089	0,089
2	Від 169 до 173 викл.	171	13	0,232	0,321
3	Від 173 до 177 викл.	175	15	0,268	0,589
4	Від 177 до 181 викл.	179	14	0,250	0,839
5	Від 181 до 185 викл.	183	5	0,089	0,928
6	Від 185 до 189 вкл.	187	4	0,072	1,00

Для побудови гістограми по осі абсцис відкладають у вибраному масштабі інтервали класів “від” і “до” , а по осі ординат пропорційно частотам висоти прямокутників.

Гістограма зображує диференційний закон розподілу випадкової величини.

Графічне зображення побудованих частот наведено на рис. 3.

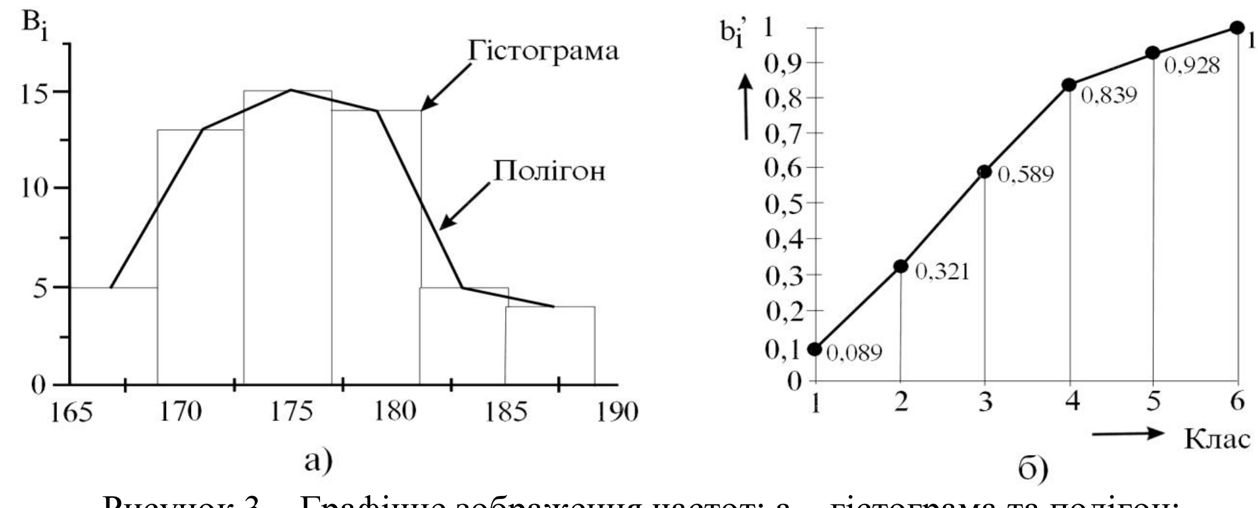


Рисунок 3 – Графічне зображення частот: а – гістограма та полігон; б – кумулятивна лінія

<u>Порядок виконання роботи</u>

1. На окремому аркуші робочого зошита (бажано на міліметровому папері) намалювати заданий викладачем варіант осцилограми.

2. Кількість відрізків Δ вздовж горизонтальної осі взяти в межах 40…50 штук.

3. Заміряти і записати величину кожної з ординат, тобто утворити масив даних.

4. Визначити кількість класів за формулою (1).

5. Визначити інтервал класів за формулою (2).

6. Провести розбивку отриманого експериментального масиву даних на класи та обчислити частоти і записати всі отримані результати згідно з наведеною табл. 1.

7. Побудувати гістограму, полігон та кумулятивну лінію.

8. Визначити математичне сподівання випадкового процесу за формулою

$$m = \frac{\sum\limits_{i=1}^{k} B_i \overline{Z}_i}{n}.$$

9. Вивчити середнє квадратичне відхилення випадкового процесу за формулою

$$\sigma = \sqrt{\frac{\sum\limits_{i=1}^{k} B_i (\overline{Z}_i - m)^2}{n}}.$$

10. Визначити коефіцієнт варіації (у відсотках) за формулою

$$V = \frac{\sigma}{m} 100\% .$$

11. Знайти середню помилку досліду згідно з наведеним виразом

$$\pm\, p = \frac{\sigma}{\sqrt{n}}.$$

Контрольні запитання

1. Пояснити суть термінів "реалізація", "ансамбль реалізацій", "осцилограма".

2. Як виконати дискретизацію випадкового процесу?

3. Пояснити порядок розподілу масиву даних на класи.

4. Як визначити абсолютну, відносну та накопичену частоти?

5. Як побудувати гістограму та кумулятивну лінію. Які висновки можна зробити про процес залежно від вигляду цих графіків?

6. Що таке полігон розподілу випадкової величини?

ПРАКТИЧНА РОБОТА № 7

ПЕРЕВІРКА ГІПОТЕЗИ ПРО НОРМАЛЬНИЙ РОЗПОДІЛ ЕКСПЕРИМЕНТАЛЬНИХ ДАНИХ

Мета роботи: закріпити знання з методів перевірки гіпотези відповідності емпіричного розподілу випадкової величини теоретичному та навчитись визначати адекватність емпіричного розподілу результатів експериментальних досліджень теоретичному.

Теоретичні відомості

Випадковою величиною називається величина, яка в результаті досліду може приймати будь-яке різне значення. Нехай випадкова величина задана теоретичним законом розподілу, при цьому для кожного значення випадкової величини відома частота $\boldsymbol{m}$ повторюваності або накопичена частота, яка отримана із $\boldsymbol{N}$ дослідів. При $\boldsymbol{N} \to \infty$ $\frac{m_i}{N} \to \boldsymbol{F(x_i)}$.

Основними параметрами теоретичних розподілів є математичне сподівання MX (центр групування) і дисперсія DX (величина розсіювання).

Для дискретної випадкової величини, коли випадкові величини знаходяться у визначення в визначеному інтервалі, у загальному випадку:

$$\boldsymbol{MX} = \sum_{i=1}^{n} x_i F(x_i), \text{ а } \boldsymbol{DX} = \sum_{i=1}^{n} F(x_i)(x_i - MX)^2,$$

при цьому $\sqrt{DX} = \boldsymbol{\sigma}$, що називається середнім квадратичним відхиленням теоретичного розподілу. Середнє значення теоретичного розподілу випадкової величини $\bar{x}$ характеризує центр групування значень випадкової величини.

У ряді задач виникає потреба в оцінці випадкового розходження між теоретичними і експериментальними параметрами розподілів, тобто між $\boldsymbol{MX}$ і $\bar{x}$, $\boldsymbol{\sigma}$ і S – експериментальним середнім квадратичним відхиленням та іншими критеріями оцінок.

Така необхідність виникає, наприклад, в наступних випадках:

1) до досліду теоретично розраховані величини $\boldsymbol{MX}$ і $\boldsymbol{\sigma}$. За результатами вибірки необхідно перевірити правильність теоретичних підрахунків;

2) виходячи із ряду технологічних та інших факторів, робиться припущення про чисельні значення параметрів розподілу.

Обґрунтування цього припущення піддається перевірці за допомогою експериментальних даних.

3) якщо вже відомі основні параметри і характеристики розподілу, то тоді робляться вибірки через певні проміжки часу. Вирахувані потім експериментальні характеристики зрівнюються з теоретичними.

Після того, як експериментальна крива вирівняна за теоретичною, необхідно знайти вірогідність того, що досліджувана експериментальна крива відповідає вибраному теоретичному закону.

Перевірка гіпотези про нормальний розподіл вибіркової сукупності випадкової величини необхідна для підтвердження або відхилення нульової гіпотези про розподіл випадкової величини та належність вибіркової сукупності до генеральної. В термін "гіпотеза" в даному випадку вкладається припущення, яке належить до об'єктивних властивостей явища, що вивчається, та підлягає перевірці.

Для нормального закону розподілу випадкової величини найбільш повно розроблені статистичні методи. Тому при відповідності емпіричного розподілу нормальному закону можна максимально використати всі статистичні методи обробки результатів експерименту. При відхиленні гіпотези про нормальний розподіл випадкової величини необхідно визначити закон, до якого ближче знаходиться розподіл результатів експерименту, та використати статистичні методи, розроблені для іншого закону розподілу.

Перевірка відповідності розподілу випадкової величини нормальному закону розподілу може бути проведена за кількома критеріями відповідності: критерієм узгодження Пірсона $\boldsymbol{\chi^2}$, критерієм Колмогорова $\boldsymbol{K(\lambda)}$, Романовського, ***САВ*** і т.д.

У загальному випадку щільність імовірності розподілу випадкової величини для нормального закону має вигляд

$$f(x)=F'(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \qquad (1)$$

де $\boldsymbol{F'(x)}$ – похідна функції розподілу випадкової величини;

$\boldsymbol{\sigma}$ – середнє квадратичне відхилення теоретичного розподілу;

$\boldsymbol{\mu = MX}$ – математичне сподівання теоретичного розподілу.

Перевірка відповідності експериментального розподілу теоретичному за критерієм Пірсона $\boldsymbol{\chi^2}$ полягає у визначенні величини $\boldsymbol{\chi^2}$ та співставленні її з критичним значенням $\boldsymbol{\chi^2_{кр} = \chi^2_к}$, яке знаходять з таблиць для відповідного значення імовірності $\boldsymbol{\alpha}$ (імовірність практично неможливих випадків) та кількості ступенів відносності u. Критерій Пірсона $\boldsymbol{\chi^2}$ визначається за формуло

$$\chi^2=\sum_{j=1}^{k}\frac{(\gamma_j-\gamma_{Tj})^2}{\gamma_{Tj}}, \qquad (2)$$

де γ_j – емпіричні абсолютні частоти потрапляння випадкової величини в певний інтервал;

k – кількість класів розрядних частот;

γ_{Tj} – теоретична частота.

Гіпотеза про розподіл приймається у випадку, коли $\chi^2_{розр} < \chi^2_{табл} = \chi^2_k$.

Критерій Пірсона χ^2 необхідно застосовувати в тих випадках, коли теоретичні значення параметрів функцій розподілу невідомі.

Він забезпечує мінімальну похибку в прийнятті неправильної гіпотези порівняно з іншими критеріями.

Критерій відповідності Колмогорова $K(\lambda)$ базується на знаходженні максимального абсолютного значення різниці між теоретичною $F_m(x)$ та емпіричною $F_e(x)$ функціями розподілу $D_m = max|F_m(x) - F(x)|$ і застосовується, якщо відомі теоретичні значення параметрів.

За величиною D_m знаходять емпіричне значення критерію Колмогорова $K(\lambda)$ за формулою

$$K(\lambda) = D_m\sqrt{n}\,. \tag{3}$$

Порівнявши експериментальні значення функції частот $F_m(x)$ зі значеннями теоретичної безперервної функції розподілу $F(x)$ за якою проводять вирівнювання, знаходять найбільше абсолютне значення різниці між ними $D_m^{(0)}$, при цьому складають вираз $\lambda_0 = D_m^{(0)}\sqrt{n}$.

Для цього з λ_0 знаходять значення виразу $1 - K(\lambda_0)$ за таблицями.

Якщо величина $1 - K(\lambda_0)$, знайдена за таблицями, менше для даного λ_0 при прийнятому рівні значимості, то нульова гіпотеза приймається.

Критерії Пірсона χ^2 та Колмогорова $K(\lambda)$ застосовуються для визначення відповідності розподілу випадкової величини для сукупності, яка має велику кількість значень об’єму вибірок (n > 50).

Для невеликих вибірок (n < 12) можна перевірити нормальність розподілу за величиною середнього абсолютного відхилення (***САВ***)

$$CAB = \frac{\sum|x_i - \bar{x}|}{n}, \tag{4}$$

де x_i – значення випадкової величини з ряду x_1; x_2; ... ; x_n;

$\bar{x}$ – середнє значення випадкової величини;

n – кількість значень випадкової величини.

Для вибірки, яка має приблизно нормальний розподіл повинен бути справедливим вираз

$$\left|\frac{CAB}{\overline{S}} - 0,7979\right| \langle \frac{0.4}{\sqrt{n}},$$

де $\overline{S}$ – експериментальне середньоквадратичне відхилення випадкової величини.

Порядок виконання роботи

1. Обчислити кількість класів $\boldsymbol{k}$ вибіркової сукупності випадкової величини за формулою

$$k \approx 1 + 3{,}32\ lgn.$$

Отримане число закруглити до цілого в більшу сторону.

2. Обчислити ширину кожного класу $\boldsymbol{d}$ для визначеної кількості класів вибіркової сукупності за формулою

$$d = \frac{x_{max} - x_{min}}{k} = \frac{R}{k}, \quad (5)$$

де $\boldsymbol{R}$ – статистичний розмах випадкової величини;

$\boldsymbol{x_{max}}$, $\boldsymbol{x_{min}}$ - відповідно максимальне та мінімальне значення випадкової величини.

3. Розбити вибіркову сукупність випадкової величини на класи, включаючи до них значення величин таким чином:

- до першого класу від $\boldsymbol{x_{min}}$ до ($\boldsymbol{x_{min}} + \boldsymbol{d}$);
- до другого класу від ($\boldsymbol{x_{min}} + \boldsymbol{d}$) до ($\boldsymbol{x_{min}} + \boldsymbol{2d}$);
- до третього класу від ($\boldsymbol{x_{min}} + \boldsymbol{2d}$) до ($\mathbf{x_{min}} + \boldsymbol{3d}$) і так далі.

Розбивку на класи слід виконати так, щоб одне і теж значення не було записано двічі.

4. Визначити середні значення для кожного $\boldsymbol{j}$-го класу за формулою

$$\overline{x}_j = \frac{x_{j\,max} - x_{j\,min}}{2}, \quad (6)$$

де $\boldsymbol{x_{jmax}}$ – максимальне значення випадкової величини, яка потрапила в певний клас;

$\boldsymbol{x_{jmin}}$ – мінімальне значення випадкової величини, яка потрапила в певний клас

5. За одержаними даними побудувати гістограму розподілу випадкової величини, відкладаючи по осі ординат $\boldsymbol{n_j}$ (кількість значень які потрапили в певний інтервал) для кожного $\boldsymbol{j}$-го класу з середнім значенням $\overline{x}_j$, а по осі абсцис середні значення величини для кожного $\boldsymbol{j}$-го класу та ширину цього

класу.

6. Перевірити, чи відповідає розподіл вибіркової сукупності випадкової величини нормальному закону за ***САВ***, враховуючи що:

- середнє значення вибіркової сукупності випадкової величини визначаються за формулою

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, i = 1,2,...,n; \tag{7}$$

- середньоквадратичне відхилення визначається за формулою

$$\overline{S} = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n-1}}. \tag{8}$$

Звіт

Звіт повинен вміщувати:

1. Перевірку нормальності розподілу випадкової величини за ***САВ***.
2. Побудовану гістограму частостей розподілу експериментальних даних для визначеного числа класів ***k***.
3. Розрахунки провести для ряду експериментальних даних згідно з варіантом, табл. 1), за критерієм ***САВ***.

Контрольні запитання

1. Що таке випадкова величина?
2. Назвіть основні параметри теоретичних розподілів.
3. Як визначається математичне сподівання та дисперсія?
4. Дати визначення критеріїв відповідності.
5. Для яких вибірок застосовується критерій САВ?

Таблиця 1 – Вихідні дані для проведення розрахунків

№	В.1	В.2	В.3	В.4	В.5	В.6	В.7	В.8	В.9	В.10	В.10	В.12
1	5.62	6.66	8.59	6.05	7.52	3.91	6.02	11.55	10.88	2.06	091	0.68
2	3.70	6.21	7.77	7.36	7.84	0.07	7.50	12.46	9.48	-0.21	5.46	1.02
3	4.51	4.57	8.57	6.77	7.26	0.23	10.57	12.27	8.84	-0.90	5.21	0.69
4	3.05	5.63	7.68	7.99	6.81	0.59	6.11	10.44	11.18	0.51	6.86	0.31
5	3.30	3,45	7.92	10.00	6.35	1.25	6.27	11.82	10.26	0.32	4.59	1.41
6	4.62	5.28	7.25	5.59	5.19	0.30	7.05	10.81	9.46	-0.83	2.48	0.94
7	6.04	3.88	8.57	6.17	6.12	0.19	7.06	9.897	9.28	0.32	3.56	0.99
8	3.82	6.48	6.66	9.60	8.17	0.47	8.25	10.59	9.21	-0.40	6.13	0.16
9	3.04	3.55	7.62	9.36	5.57	0.66	7.15	8.919	10.19	-3.69	3.41	0.81
10	6.361	5.22	7.98	7.99	5.46	0.36	7.58	8.667	10.24	1.49	7.21	0.74
11	3.03	7.47	6.45	6.08	7.16	0.08	6.48	9.098	10.32	-1.46	5.54	0.58
12	6.20	2.36	9.48	8.23	6.51	1.00	5.88	10.33	11.70	-0.61	4.66	0.83
13	4.38	3.33	7.86	6.43	5.35	0.12	6.92	12.22	9.49	0.27	1.83	0.55
14	2.06	7.37	7.39	7.39	5.26	0.61	7.72	9.158	9.57	0.26	4.25	4.60
15	5.09	6.21	9.05	6.22	5.13	1.27	7.07	9.505	9.76	-2.79	8.80	0.42
16	3.15	3.43	6.50	6.16	7.63	0.86	7.92	11.25	10.40	0.06	3.22	0.94
17	3.86	7.68	9.19	8.07	6.97	0.85	7.30	10.47	10.08	0.82	9.58	0.92
18	3.03	4.66	8.57	6.44	9.60	0.12	7.17	11.83	10.68	-0.59	5.56	1.06
19	4.72	3.66	8.97	7.73	5,05	1.20	9.34	10.18	10.18	0.90	0.92	0.35
20	3.82	4.45	7.29	7.50	6.54	2.71	6.68	9.48	9.484	-0.14	3.07	0.79
21	4.32	3.54	8.56	6.03	8.36	0.90	553	8.61	8.618	0.07	5.90	1.27
22	4.30	6.44	8.87	7.13	6.01	1.13	6.97	9.77	9.77	-2.81	3.14	0.75
23	4.04	5.79	7.92	4.98	6.70	0.22	5.89	10.04	10.04	-006	0.79	1.07
24	4.71	2.05	8.93	6.35	7.55	3.49						
25	5.07	4.35	7.19	6.59	6.70	0.91						

Список використаної літератури

1. Тушева, В. В. Основи наукових досліджень [Текст] : навч. посіб. для використання в навч. закл. / В. В. Тушева. – Харків : Федорко, 2014. – 407 с
2. Сукач, М. К. Основи наукових досліджень [Текст] : навч. посіб. Для студентів ВНЗ / М. К. Сукач. – Київ : Леся, 2014. – 211 с.
3. Основи наукової діяльності: методологія, організація, оформлення результатів [Текст] : навч. посіб. / В. М. Головій, Є. Ю. Кузькін, Л. В. Піддубна. – К.: «Хай-Тек Пресс», 2010. – 344 с
4. Основи наукових досліджень у прикладних задачах [Текст] : навч. посіб. для студентів ВНЗ / Л. О. Кривопляс-Володіна [та ін.] ; Нац. ун-т харч. технологій. – К. : Сталь, 2016. – 272 с.
5. Чмиленко, Ф.О. Посібник до вивчення дисципліни «Методологія та організація наукових досліджень» [Текст] / Ф.О. Чмиленко, Л.П. Жук. – Д.: РВВ ДНУ, 2014. – 48 с.
6. Шейко, В. М. Організація та методика науково-дослідницької діяльності [Текст]: підручник / В.М. Шейко, Н.М. Кушнаренко. − 5-те вид., стер. − К.: Знання, 2006. − 307 с.
7. Цехмістрова, Г.С. Основи наукових досліджень [Текст]: навч. посіб. / Г.С. Цехмістрова. − К.: Вид. дім «Слово», 2003. − 240 с.
8. Крушельницька, О. В. Методологія та організація наукових досліджень [Текст]: навч. посіб. / О.В. Крушельницька. − К.: Кондор, 2006. − 206 с.
9. Власов, К. П. Методы исследований и организация экспериментов [Текст] / К.П. Власов, А.А. Киселева. − Х.: Гуманит. центр, 2002. − 256 с.

ЗМІСТ