МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ МІСЬКОГО ГОСПОДАРСТВА імені О. М. БЕКЕТОВА

МЕТОДИЧНІ РЕКОМЕНДАЦІЇ

до проведення практичних занять, самостійного вивчення та виконання контрольної роботи з навчальної дисципліни

«ОСНОВИ НАУКОВИХ ДОСЛІДЖЕНЬ»

(для здобувачів першого (бакалаврського) рівня вищої освіти всіх форм навчання зі спеціальності 194 – Гідротехнічне будівництво, водна інженерія та водні технології)

Харків

ХНУМГ ім. О. М. Бекетова

2021

Методичні рекомендації до проведення практичних занять, самостійного вивчення та виконання контрольної роботи з навчальної дисципліни «Основи наукових досліджень» (для здобувачів першого (бакалаврського) рівня вищої освіти всіх форм навчання зі спеціальності 194 – Гідротехнічне будівництво, водна інженерія та водні технології) / Харків. нац. ун-т міськ. госп-ва ім. О. М. Бекетова ; уклад. : І. М. Чуб, В.О. Ткачов − Харків : ХНУМГ ім. О. М. Бекетова, 2021. – 56 с.

Укладачі: канд. техн. наук, доц. І. М. Чуб,
канд. техн. наук, доц. В. О. Ткачов

Рецензент

Г. І. Благодарна, кандидат технічних наук, доцент Харківського національного університету міського господарства імені О. М. Бекетова

Рекомендовано кафедрою водопостачання, водовідведення та очищення вод, протокол від № 1 від 2.09.2021.

ЗМІСТ

ВСТУП

Наукові дослідження є формою існування і розвитку науки. Процес наукового пізнання відрізняється особливою систематичністю і послідовністю. Науковий пошук завжди має організований і цілеспрямований характер специфічного дослідження.

Сьогодні до сфери науково-дослідної діяльності залучені сотні тисяч людей в усьому світі, результати їх досліджень стають безпосередньою продуктивною силою, в значною мірою визначають напрями та тенденції розвитку сучасного суспільства. Утворюються нові форми організації науки, формуються великі дослідницькі колективи, наука перетворилася на величезний, складний соціальний організм. У цьому зв’язку оволодіння знаннями основ наукового дослідження є обов’язковим для усіх фахівців.

Навчальна дисципліна «Основи наукових досліджень» необхідна для підготовки фахівців, які володітимуть навичками науково-дослідної роботи, знатимуть методологію та методику, планування та організацію наукових досліджень у галузі водної інженерії та водних технологій. Недостатнє знання ними сучасних методів математичного опрацювання та аналізу результатів експерименту викликає звичайно серйозні утруднення і призводить до застосування спрощених і недостатньо обґрунтованих прийомів. Це відноситься до питань добору емпіричних формул і оцінки їхніх параметрів, оцінки істинних значень величин, що вимірюються, і точності вимірів, дослідження кореляційних залежностей.

Рівень науково-дослідної діяльності, розвиток навичок самостійного творчого мислення є важливим фактором, який визначає інтелектуальний науковий потенціал і висоту духовного зростання країни, компетентність її кадрів, забезпечує можливість та потреби для постійного самостійного оновлення своїх знань і швидку адаптацію надалі до мінливих умов діяльності та розвитку нового в науці. Виконуючи практичні роботи, студенти закріплюють навички теоретичного та практичного застосування основних методів опрацювання й аналізу результатів експерименту до різноманітних питань водопідготовки.

ЗМІСТОВИЙ МОДУЛЬ 1

ОСНОВНІ ЕЛЕМЕНТИ НАУКОВОГО ПІЗНАННЯ. ТЕОРЕТИЧНІ ДОСЛІДЖЕННЯ

Тема 1 Методологічні основи наукових досліджень

Методологія – це вчення про систему методів наукового пізнання та перетворення реальної дійсності. В буквальному розумінні методологія - це вчення про метод. *Головною метою методології* є вивчення тих засобів, методів та прийомів наукового дослідження, за допомогою яких суб'єкт наукового пізнання одержує нові знання про реальну дійсність. *Предмет її вивчення* - це поняття і методи науки, їх сфера застосування.

Методологія науки може бути *загальною або конкретно-науковою*.

Загальна методологія науки досліджує закони розвитку наукового пізнання в цілому. Водночас методологія ґрунтується на законах окремих наук, особливостях пізнання конкретних процесів і проявляється у здійсненні теоретичних узагальнень, принципів методів дослідження окремих наук. Тому вона виступає і як конкретно-наукова.

Розвиток методології науки пов'язаний з розвитком методів наукового пізнання дійсності.

Метод (від грец. methodos - спосіб пізнання) – це спосіб, шлях пізнання та практичного перетворення реальної дійсності, система прийомів та принципів, що регулюють практичну та пізнавальну діяльність людей.

Таким чином, щодо наукового дослідження метод визначається як сукупність визначених правил, прийомів, способів і норм пізнання певного суб'єкта чи явища.

Типологія методів наукового дослідження

В сучасному наукознавстві успішно працює багаторівнева методологічна класифікація методів наукового пізнання, згідно з якою за ступенем спільності та сферою дії методи наукового пізнання поділяються на *загальні філософські, загальнонаукові, окремо наукові, дисциплінарні та міждисциплінарні методи дослідження*.

Загальні методи – це система принципів, прийомів, що мають загальний, універсальний характер, є абстрактними, суворо не регламентовані, не піддаються формалізації та математизації і не замінюють спеціальних методів (методів окремих наук).

Методи окремих наук – це сукупність способів та принципів пізнання, прийомів і процедур дослідження, що застосовуються в тій чи іншій науці.

Загальнонаукові методи дослідження можна класифікувати залежно від рівнів пізнання – *емпіричного або теоретичного*, на яких вони (методи) застосовуються.

На емпіричному рівні переважає живе споглядання (чуттєве пізнання), раціональний момент тут наявний, але має підпорядковане значення. Тому досліджуваний об'єкт відображається переважно з боку зовнішніх зв'язків та проявів, що доступні живому спогляданню. Збирання фактів, їх первинний опис, узагальнення, систематизація – характерні ознаки емпіричного пізнання. До основних методів, які використовуються на емпіричному рівні дослідження, можуть бути віднесені: ***спостереження, порівняння, вимірювання, експеримент, абстрагування, аналіз і синтез.***

Теоретичний рівень дослідження пов'язаний з більш глибоким аналізом фактів, з проникненням у сутність досліджуваних явищ, з пізнанням та формулюванням законів, тобто з поясненням реальної дійсності. До основних методів, які використовуються на теоретичному рівні дослідження, можуть бути віднесені: ***індукція і дедукція, ідеалізація, формалізація та інші***.

Спостереження – це цілеспрямоване, систематичне, планомірне, активне вивчення предметів та явищ реальної дійсності, що знаходяться в природному стані або в умовах наукового експерименту.

Під спостереженням також розуміють апробацію, обґрунтування висунутих гіпотез або проміжних результатів дослідження. Вчений використовує спостереження з метою збору наукових фактів для винайдення способу розв'язання проблеми (висування та доведення гіпотези).

Наукові факти – відбиті свідомістю факти дійсності, причому перевірені, осмислені та зафіксовані мовою науки у вигляді емпіричних суджень.

Порівняння – один із найбільш поширених методів пізнання, який дозволяє встановити подібність та розбіжність предметів та явищ. Недарма говорять, що «все пізнається в порівнянні». У результаті порівняння виявляється те загальне, що притаманне ряду об'єктів. Різновидом порівняння є аналогія.

Аналогія – метод наукового дослідження, завдяки якому досягається пізнання одних предметів і явищ на основі їх подібності з іншими.

Одним із різновидів методу аналогій є метод моделювання.

Моделювання – метод наукового пізнання, що ґрунтується на заміні предмета або явища, що досліджуються, на їх аналог - модель, що містить істотні риси оригіналу.

Вимірювання – це метод дослідження, за допомогою якого визначається кількістне значення деякої величини з використанням одиниці вимірювання об'єкта.

Експеримент – метод емпіричного дослідження, що базується на активному та цілеспрямованому втручанні суб'єкта у процес наукового пізнання явищ та предметів реальної дійсності шляхом створення контрольованих та керованих умов, що дозволяють виділяти визначені якості, зв'язки в об'єкті, що досліджується, та багатократно їх відтворювати.

Абстрагування – метод, який дає змогу переходити від конкретних питань до загальних понять і законів розвитку.

Зміст цього методу полягає в суттєвому відволіканні від несуттєвих властивостей, зв'язків, відносин, предметів та в одночасному виділенні, фіксуванні певних сторін цих предметів, які цікавлять дослідника.

Конкретизація – метод дослідження предметів у всій їх різноманітності, у якісній багатогранності реального існування на відміну від абстрактного вивчення предметів.

Метод сходження від абстрактного до конкретного є загальною формою руху наукового пізнання - це відображення дійсності в мислені. Згідно з цим методом процес пізнання ніби розпадається на два відносно самостійні етапи: *перший етап* - від чуттєво-конкретного до його абстрактних визначень; *другий етап* - сходження від абстрактних визначень об’єкта до конкретного у пізнанні.

Аналіз – метод дослідження, що полягає в уявному або практичному розчленуванні цілого на складові частини, кожна з яких аналізується окремо у межах єдиного цілого.

Синтез – метод вивчення об’єкта у його цілісності, у єдиному взаємному зв’язку його частин. У процесі наукових досліджень синтез пов’язаний з аналізом, оскільки дає змогу поєднати частини предмета (об’єкта чи явища), розчленованого в процесі аналізу, встановити їх зв’язок і пізнати предмет (об’єкт чи явище) як єдине ціле.

Сучасна наука є двигуном науково-технічного прогресу. Впровадження науки у виробництво виражається у підвищенні продуктивності праці, створенні нових машин й матеріалів, покращенні експлуатаційних показників, надійності й довговічності продукції, зниженні її собівартості (рис. 1.1).

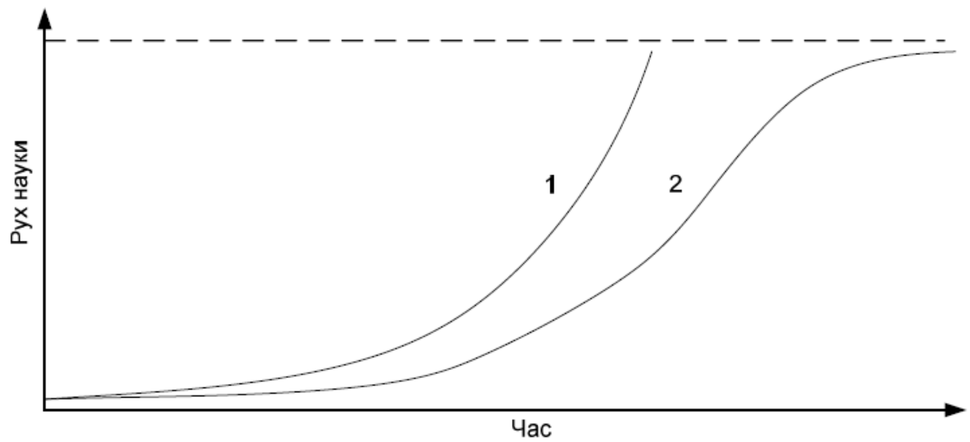


Рисунок 1.1 – Закономірності розвитку результатів наукових досліджень у часі:
1 – експонента; 2 – ймовірна крива

Кожна наука має певну сукупність методів проведення досліджень при вивченні власного предмета, яку можна класифікувати на такі групи:

– ***методи накопичування фактів,*** що мають відношення до об’єкта дослідження (спостереження, реєстрація, вимірювання);

– ***методи описування фактів*** або властивостей ідеалізованого об’єкта дослідження та факторів, що відбивають ці властивості, а також явищ (процесів), що досліджуються, розвиток яких визначається цими факторами;

– ***методи аналізу фактів, властивостей, факторів і явищ*** за різними показниками і критеріями (оцінка, зіставлення, порівняння, класифікація, впровадження, систематизація);

– ***методи обґрунтування*** наукових висновків, серед яких мають бути такі методи: побудови (синтезу), доведення, оцінки достовірності;

– ***методи вибору і обґрунтування*** наукових рекомендацій, у т.ч. методи побудови (синтезу), оцінки й оптимізації;

– ***методи інтерпретації та експериментальної перевірки*** висновків і рекомендацій;

– ***методи техніко-економічної оцінки*** рекомендацій.

У процесі розв’язання наукової проблеми вчений, як правило, самостійно шукає методи та способи її розв’язання. Всі прийняті методичні Розв’язання необхідно фіксувати у формі методик, які періодично переглядаються.

Методика дослідження – сукупність методів і прийомів правильного і цілеспрямованого вивчення явищ. При визначенні методики необхідно використовувати не тільки особистий досвід, але й досвід інших дослідників.

Обрану методику потрібно удосконалювати на основі критичного аналізу попередніх робіт і результатів їх впровадження в практику. Оскільки метод не являє собою щось незалежне від задач, об’єкта і умов дослідження, методи диференціюють та індивідуалізують.

Поняття наукової проблеми

Наукова проблема – це питання, що потребує наукового розв’язання; завдання для пошуку невідомого; сукупність нових діалектично складних теоретичних або практичних питань, які суперечать існуючим знанням або прикладним методикам у конкретній науці і потребують розв’язання за допомогою наукових досліджень.

Проблема в науці – це суперечлива ситуація, яка вимагає свого розв’язання. Така ситуація найчастіше виникає в результаті відкриття нових фактів, які явно не вкладаються в межі колишніх теоретичних уявлень, тобто коли жодна з теорій не може пояснити щойно виявлені факти.

Розв’язання проблеми не міститься в існуючому знанні та не може бути отримане шляхом перетворення наявної наукової інформації.

Правильна постановка та чітке формулювання проблеми не менш важливе, ніж її розв’язання. Вибір проблеми значною мірою визначає як стратегію дослідження взагалі, так і напрям наукового пошуку зокрема. По суті, мова йде про вміння відокремити головне від другорядного, про виявлення того, що поки не відоме науці з предмета дослідження, про усвідомлення того, що ми чогось не знаємо.

Джерелами наукових проблем є як практика, так і потреби власне науки (необхідність удосконалення методів наукового дослідження, уточнення категорійно-понятійного апарату тощо).

Контрольні питання

1. Яка мета наукових досліджень?
2. Які дослідження належать до фундаментальних?
3. Які дослідження належать до прикладних?
4. Які наукові праці належать до дослідно-конструкторських?
5. Які типи завдань можна розв`язувати в результаті виконання прикладних науково-дослідних робіт?
6. Назвіть і охарактеризуйте критерії економічної ефективності науково-дослідних тем.
7. Назвіть етапи виконання прикладної науково-дослідної роботи.
8. Назвіть етапи виконання дослідно-конструкторської розробки.
9. Охарактеризуйте інформатику як науку.
10. Перелічіть напрями розвитку інформаційних наук

Тема 2 Загальні відомості про помилки вимірювань

Похибка – кількісна характеристика неоднозначності результату вимірювання. Її оцінюють виходячи зі всієї інформації, накопиченої при підготовці і виконанні вимірювань. Цю інформацію обробляють для сумісного визначення остаточного результату вимірювання і його похибки. Остаточний результат не можна розцінювати як «дійсне значення» вимірюваної фізичної величини, оскільки в цьому немає сенсу через наявність похибки.

Похибка може бути виражена в одиницях вимірюваної величини x, - у такому разі вона позначається Δx і носить назву **абсолютної похибки.** Проте абсолютна похибка не відображає якості вимірювань: наприклад, абсолютна похибка 1 мм при вимірюванні розмірів приміщення свідчить про високу якість вимірювання, та ж похибка абсолютно неприйнятна при вимірюванні діаметру тонкого дроту.

Критерієм якості вимірювання є відношення абсолютної похибки до остаточного результату вимірювання

$$\delta x = \frac{\Delta x}{x}$$

Це відношення безрозмірне. Величину δx називають **відносною погрішністю** і використовують як в абсолютному, так і в відсотковому вираженні. Високій точності вимірювання відповідає мале значення відносної похибки.

Основні типи похибок:

1. Хиби або грубі похибки – виникають унаслідок несправності вимірювальних приладів або помилок в експерименті, зроблених через неувагу.

2. Приладна похибка – систематична похибка, присутня в результатах вимірювань, виконаних за допомогою будь-якого вимірювального приладу. Приладна похибка, зазвичай, невідома і не може бути врахована. Її можна оцінити тільки шляхом порівняння показів приладу з показами іншого, точнішого. Іноді результати спеціально проведеного порівняння наводять в паспорті приладу, проте здебільшого вказують максимально можливу похибку для приладів цього типу.

3. Модельна похибка. Будь-яке експериментальне дослідження, пов’язане з вимірюваннями, закладена модель. Модель містить фізичний опис досліджуваного об’єкта або процесу, який дозволяє скласти його математичний опис, а саме, набір функціональних співвідношень, що включають фізичні величини. Неправільно побудована модель, у якій не відобразилися важливі процеси або фактори, що впливають на результат вимірювань, також приводить до невідповідностей. Як наслідок, вимірювані в експерименті величини, що обчислюються за отриманими з моделі робочими формулами, містять похибки, які носять назву модельних погрішностей. До розряду модельних може бути віднесена похибка зважування на важельних вагах. Згідно закону Архімеда вага тіла і гирь зменшується через дію виштовхуючої сили повітря. Нагадаємо, що вага 1 м^3 повітря дорівнює приблизно 10 Н. Для того, щоб правильно знайти масу зважуваного тіла, знову ж таки, потрібно ввести поправки на втрату ваги гирями і самим тілом.

4. Випадкові похибки – при повторних вимірюваннях похибки цього типа показують свою випадкову природу. Виникають вони внаслідок безлічі причин, спільна дія яких на кожне окреме вимірювання неможливо врахувати або наперед встановити. Такими причинами можуть виявитися, наприклад, незначні коливання температури різних деталей і вузлів установки, скачки напруги, вібрації, турбулентні рухи повітря, тертя в механізмах, помилки прочитування показань приладів тощо. Єдино можливий спосіб об'єктивного обліку випадкових погрішностей полягає у визначенні їх статистичних закономірностей, що виявляються в результатах багатократних вимірювань. Розраховані статистичні оцінки вносять в остаточний результат вимірювання.

Однією з грубих помилок, які допускають студенти, є знаходження похибки вимірювання, як

$$\Delta x = x_{екс} - x_m ,$$

де $x_{екс}$ – набутого в процесі експерименту середнього значення величини

x_m – значення, узяте з довідника або розраховане виходячи з теоретичних уявлень. Метою експерименту є саме перевірка існуючих теорій і уточнення табличних значень.

З іншого боку, при виконанні учбових лабораторних робіт корисно порівняти отримані результати з довідковими табличними величинами і, у разі значної їх розбіжності, проаналізувати, які експериментальні фактори і модельні похибки могли привести до цього.

Операції з наближеними числами. Помилки вимірювання і міри точності

Майже всі вимірювання і математичні операції дають наближені значення шуканих величин. Складові наближеного числа можуть бути *вірними, сумнівними* і *невірними.*

Постулати:

1. Якщо похибка числа не вказана, то його абсолютна похибка дорівнює половині одиниці розряду останньої цифри.

2. Розряд старшої цифри похибки показує розряд сумнівної цифри в числі.

3. Як значущі цифри, можуть бути тільки правильні і сумнівні цифри.

4. Якщо похибка числа не вказана, всі цифри значущі.

5. Під значущими цифрами числа розуміють послідовність цифр без урахування місця коми, а для чисел менше одиниці – без урахування нуля перед комою і подальших нулів (табл. 1.1).

Таблиця 1.1 – Значущі цифри

Результат	Кількість значущих цифр	Граничні значення помилки	Δx	Максимальна помилка $(\Delta x/xi)\cdot 100$ %
3 3,5 3,55	1 2 3	2,5 – 3,5 3,45 – 3,55 3,545 – 3,555	0,5 0,05 0,005	16,7 1,4 0,14

Враховуючи, що інженерні розрахунки припускаються помилки 2–5 %, то не доцільно видавати результати з більш ніж двома значущими цифрами.

1 1,5 1,55	… … …	… … …	0,5 0,05 0,005	50% 3,3% 0,3%
0,1 0,15 0,155	… … …	… … …	0,05 0,005 0,0005	50% 3,3% 0,3%

Приклад 1.1. Кількість 216,8 + 2,3 варто записувати, як 216 + 2,3, а цифру 8 відкинути як неправильну, бо вже цифра 6 є сумнівною.

Округлення варто проводити до найближчого *парного* числа, причому при округленні відкидають всі цифри, що стоять праворуч від розряду, до якого проводиться округлення. Останню цифру, що залишилася, збільшують на одиницю, якщо найближча відкидана цифра дорівнює і більше 5, або не змінюють, якщо вона менше 5. Якщо ж відкидають тільки одну цифру 5 (або після неї йдуть нулі), *то останню* цифру, що *залишається,* збільшують на одиницю, якщо вона непарна, і залишають без зміни, якщо вона парна.

Приклад 1.2. Дано два значення: 50 ± 3 і 30 ± 2. Визначити величину відносної помилки складання, різниці, множення і ділення.

1. *Складання.* Дійсне значення лежить між 47 + 28 = 75 і 53 + 32 = 85. Відносна помилка суми дорівнює (85 - 75) / (85 + 75) = 10 / 160 = 0,0625 (6,25%).

2. *Віднімання.* Дійсне значення лежить між 47 - 32 = 15 і 53 - 28 = 25 («перехресне» віднімання, тобто максимальне значення одного числа віднімається з мінімального значення іншого і мінімальне значення одного числа – з максимального значення іншого).

Відносна помилка різниці:

(25 - 15) / (25 + 15) = 10 / 40 = 0,25 (±25 %).

3. Множення. Дійсне значення лежить в межах від 47·28 = 1316 до 53·32 = 1696. Відносна помилка добутку:

$$\frac{1316-50\cdot30}{50\cdot30}=\frac{1316-1500}{1500}=\frac{-184}{1500}=-0{,}123\ (-12{,}3\%);$$

$$\frac{1696-50\cdot30}{50\cdot30}=\frac{196}{1500}=0{,}131\ (13{,}1\%).$$

4. Ділення. Дійсне значення лежить між 47 / 32 = 1,469 і 53 / 28 = 1,893 («перехресне» ділення). Відносна помилка частного

$$\frac{1{,}469-50/30}{50/30}=\frac{1{,}469-1{,}667}{1{,}667}=-0{,}119\ (-11{,}9\%);$$

$$\frac{1{,}893-50/30}{50/30}=0{,}136\ (13{,}6\%).$$

Таким чином, зі всіх арифметичних операцій найбільшу помилку дає віднімання, а мінімальну – складання.

Приклад 1.3. Знайти наближену відносну помилку добутку: (40±10%)·(30±5%).

Розв'язання. Згідно з формулою (1) за [3, табл. 5.2] отримуємо: (1200 ± 10% ± 5%), або шукана величина буде дорівнювати 1200 ± 15 %.

Приклад 1.4. Знайти приблизну відносну помилку частки:

[(500 + 20)·(200 +15)] / (400 – 20).

Розв'язання. Цьому виразу відповідає частка (500 + 4%)·(200 + 7,5%) / (400 - 5%).

Згідно формул (1) і (6) з [3, табл. 5.2] приблизна відносна помилка дорівнює 4 + 7,5 + 5 = 16,5%.

Таким чином, остаточний результат:

(500 · 200/400) ± 16,5% = 250 ± 16,5% або 250 ± 41, тобто в межах 209 і 291.

Приклад 1.5. Довжина, ширина і висота цеглини рівні x_1, x_2, x_3 см з відносними помилками 0,01. Знайти максимальну абсолютну помилку об'єму.

Розв'язання. Згідно з формулою (2) за [3, табл. 5.2] максимальна помилка в об'ємі дорівнює 0,01 + 0,01 + 0,01 = 0,03 (3 %). Тоді максимальна абсолютна помилка дорівнює $0{,}03 \cdot x_1 \cdot x_2 \cdot x_3$ см^3.

Приклад 1.6. Фізико-хімічний аналіз складу, наприклад води, супроводжується систематичними і випадковими помилками. Таріровкою встановлено, що систематична помилка дорівнює 5%. Випадкові помилки *підкоряються нормальному закону з середнім квадратичним відхиленням σ = 15 %.*

Знайти:

а) імовірність визначення складу води з помилкою, що не перевищує по абсолютній величині 30 %;

б) імовірність того, що визначуваний фізико-хімічний склад води не перевершить істинного.

Розв'язання. Щільність імовірності випадкової помилки виглядає так:

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(\bar{x} - x_i)}{2 \cdot \sigma^2}\right] = \frac{1}{15\sqrt{2\pi}} \exp\left[-\frac{(\bar{x} - 2)}{450}\right].$$

Згідно з формулою (5.3.33) з [3] маємо

$$P(|x_i| < 30) = P(-30 < x_i < 30) = \left[\Phi\left(\frac{30+5}{15}\right) - \Phi\left(\frac{-30+5}{15}\right)\right] = [\Phi(2{,}33) - \Phi(-1{,}67)].$$

Оскільки $\Phi(-1) = -\Phi(1)$, то $P(|x_i| < 30) = [\Phi(2{,}33) + \Phi(1{,}67)]$. З [3] знаходимо Φ (2,33) = 0,490 1; Φ (1,67) = 0,452 5. Тоді P ($|x_i| <$ 30) = 0,4901 + 0,452 5 = 0,9426.

Імовірність того, що визначуваний фізико-хімічний склад води не перевищує істинного:

$$P(-\infty < x_i < 0) = [\Phi(0{,}5) + \Phi(\infty)],$$

Оскільки, з [3] Φ (0,5) = 0,191 5, звідки $P(-\infty < x_i < 0)$ = 0,5 + 0,191 5 = 0,691 5.

Висновок: Імовірність виконання умови (а) цього завдання приблизно дорівнює 94,3%, а умови *(б)*: 69,2%.

Методи виключення грубих помилок

Груба помилка може привести до появи як значення, що різко виділяється, так і значення, візуально не відмітного від основної маси спостережень. Ці помилки звичайно особливо добре помітні при розташуванні результатів розрахунків або експериментів у порядку убування значень або при побудові графіків («відскок крапок»).

Статистичні методи виявлення грубих помилок варто застосовувати лише тоді, коли додаткова інформація про якість вимірювань *або неповна, або ненадійна.*

У будь-якому випадку до виключення "підозрілих" значень з вибірки потрібно підходити з особливою обережністю.

Метод виключення при відомій σ. Нехай ми маємо вибірку $x_1, x_2, \ldots, x_n$, в якій є підозріле значення x^*. Алгоритм перевірки може бути наступним:

1. Підрахуємо середнє арифметичне значення $\bar{x}_0$ за формулою

$$\bar{x}_0 = \sum_{i=1}^{n} x_i / n \tag{1.1}$$

і середню квадратичну похибку σ з виразу

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(\bar{x} - x_i)^2, \sigma = \sqrt{\sigma^2} \quad (1.2)$$

виключивши з цих розрахунків значення x^*.

2. Знайдемо відношення

$$t = \left|x^* - \bar{x}\right| / \left(\sigma\sqrt{(n+1)/n}\right). \quad (1.3)$$

3. Визначимо вірогідність 1 - 2$\Phi(t)$ (додатки [3])

4. Якщо виконується умова

$$1 - 2\Phi(t) < q, \quad (1.4)$$

то вважається, що значення x^* підлягає вибраковуванню. Тут q – прийнятий в розрахунках рівень значущості, а n – об'єм вибірки. У інженерних розрахунках зазвичай приймається рівень значущості 0,1; 1,0; 5,0 %.

Вважають, що значення x^* містить грубу помилку з надійністю виведення $P = 1 - q$. Значення $t = t(P)$, для якого $1 - 2\Phi(t) = q$ і, значить, $2\Phi(t) = P$, називається *критичним значенням* відношення (1.3) при надійності P.

Приклад 1.7. У вибірці з 1 000 результатів незалежних вимірювань з середньою квадратичною помилкою σ = 14,2 виявлене одне значення x*, що "вискакує" = 120,2. Середнє з інших 1 000 – 1 = 999 результатів $\bar{x}$ = 84,5. Чи можна вважати, що значення, що «вискакує», містить грубу помилку і його потрібно виключити з подальших розглядів?

Розв'язання. З відношення (1.3) маємо:

$$\mathrm{t} = |84{,}5 - 120{,}2| / (14{,}2\sqrt{1001/1000}) \approx 2{,}50.$$

За [3] для t = 2,5 оцінюємо вірогідність 1 - 2$\Phi(t)$ = 0,01242 < 0,013. Отже, з надійністю виведення P > 0,987 (98,7%) можна рахувати, що значення x^* = 120,2 варто виключити. Якщо ж задовольняє тільки велика надійність (наприклад, P = 99 %), то виключати x^* з подальшого розгляду не можна.

Виключення помилок при невідомій σ. Якщо величина σ не відома наперед, то вона спочатку оцінюється приблизно за наслідками вимірювань [3, п. 5.4.10]. У подібному випадку замість σ використовують *емпіричний стандарт*

$$S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}. \quad (1.5)$$

Перевірку починають з визначення відношення

$$t(x^*) = \left|x^* - \bar{x}\right| / S. \quad (1.6)$$

Далі можливі варіанти:

1. Якщо

$$t(x^*) < 2/\left(3\sqrt{q}\right) = K, \quad (1.7)$$

де q – заданий рівень значущості, то сумнівне значення x^* залишають у вибірці.

2. Відношення (1.6) порівнюють з критичним значенням *tn(P)*. Якщо при заданому числі n прийнятних значень відношення (1.6) опиниться між двома

критичними значеннями при надійності P_1 і P_2 ($P_2 > P_1$), то з надійністю висновку більше за P_1 можна вважати, що значення, що «вискакує», містить грубу помилку.

3. Згідно методиці "вибраковування" по критерію Смирнова – Граббса, якщо невідомі a та σ, то можна скористатися такою величиною:

$$\overline{T}^* = \left| x^* - \overline{x} \right| / S^*, \tag{1.8}$$

де

$$S^* = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(x_i - \overline{x} \right)^2} \,. \tag{1.9}$$

Потім за наперед заданим рівнем значущості q знаходимо з [3] критичне значення C_q, що відповідає числу спостережень n. Якщо $T^* > C_q$, то сумнівне значення x^* значущо відхиляється від середнього x і може бути виключено як помилкове.

4. Д. Химмельблау рекомендує спостереження x^* відкидати, якщо виконується нерівність

$$\left| x^* - \overline{x} \right| / S^* > C, \tag{1.10}$$

де C – константа, яка знаходиться через t-критерій Ст'юдента (Додаток Е) по виразу:

$$\left[\frac{n \cdot C^2 (k + k_0 - 1)}{k \left[k + k_0 - (n \cdot C^2)/k \right]} \right]^{1/2} \approx t_{q=0,05}^{k+k_0-1}, \tag{1.11}$$

де $к = n - 1$ – кількість ступеня свободи оцінки дисперсії S^2; k_0 – будь-яке кількість додаткових ступенів свободи (зазвичай $k_0 = 0$).

Оцінка може проводитися багато разів. Середнє квадратичне відхилення S розраховується кожного разу за вибіркою, що залишилась.

Приклад 1.8. Є дані вибіркових спостережень за тривалістю повного осадження забруднень в судинах Лисенко. Тривалість окремих дослідів склала, хвил.:

10, 11, 11, 12, 12, 13, 13, 13, 14, 20.

Сумнівним здається останній результат $x_{10} = 20$. Чи є підстави його виключити? Перевіримо.

Варіант 1.

$$\overline{x} = \frac{10 + 11 \cdot 2 + 12 \cdot 2 + 13 \cdot 3 + 14}{9} = 12{,}1;$$

$$S = \sqrt{\frac{1}{9-1} \sum_{i=1}^{n} \left(x_i - 12{,}1 \right)^2} \approx 1{,}27;$$

$$t = (20 - 12{,}1)^2 / 1{,}27 = 6{,}22.$$

Приймаємо рівень значущості $q = 0{,}05$. Тоді за формулою (1.7) маємо

$$K = 2 / \left(3\sqrt{q} \right) = 2 / \left(3\sqrt{0{,}05} \right) = 2{,}98.$$

Умова (1.7) не виконується. Отже, за цим способом значення $x^* = 20$ не залишається у вибірці.

Варіант 2. За [3] критичне значення $t_n(P)$ для n = 10 навіть при P = 0,999 дорівнює 5,01, тобто t = 6,22 > t_n(0,999)= 5,01, та ми маємо право виключити значення x^* = 20 з виборки. З [3] випливає, що це значення можна було б не виключати тільки при n = 7 та P = 0,999, де t_7(0,999) = 6,37.

Варіант 3. За тих самих умов вибракування за критерієм Смірнова –Граббса показує:

– середня величина, що розрахована за формулою

$$\bar{x} = \frac{1}{10}\sum_{1}^{10} x_i = 12{,}9;$$

– оцінка середньоквадратичного відхилення S^* за формулою (1.9) дорівнює 2,62;

– за формулою (1.8) T^* = (20,0 - 12,9)/2,62 = 2,71;

– за [3] навіть при рівні значущості q = 0,025 та n = 10 $C_{0,025}$ = 2,414 < T^* = 2,71.

Таким чином, гіпотеза про однорідність вишукуваного ряду відхиляється та останнє значення x^* = 20 можна вважати нетиповими.

Для остаточної обробки цього ряду розрахуємо нові значення середньої величини

$$\bar{x} = \frac{1}{9}\sum_{1}^{9} x_i = 12{,}1$$

та оцінюємо середньоквадратичне відхилення S, але вже за формулою:

$$S = \sqrt{\frac{1}{9-1}\sum_{1}^{9}\left(x_i - \bar{x}\right)^2} \approx 1{,}27.$$

Варіант 4. З розрахунку за варіантом 1 маємо, що $\bar{x}$ = 12,1, а оцінка дисперсії за дев’ятьма значеннями, що залишилися, буде дорівнювати S = 1,27.

Для рівня значущості q = 0,05, n = 9, k = 9 – 1 = 8 за [3] маємо t = 2,31. Тоді за формулою (1.11) отримаємо:

$$\left[\frac{9\cdot C^2(8+0-1)}{8\left[8+0-(9\cdot C^2)/8\right]}\right]^{1/2} \approx 2{,}31^{(8+0-1)}.$$

Перевіривши, отримаємо: $C \approx 2{,}849$.

Розв’язання рівняння (1.11) методом підстановки складне. Відповідні алгебраїчні перетворення приводять (1.11) до такого вигляду

$$C = \left\{\frac{(k^2 + k\cdot k_0)\cdot t_q^{2(k+k_0-1)}}{n\left[k + k_0 - 1 + t_q^{2(k+k_0-1)}\right]}\right\}^{1/2}$$

або (при k_0 = 0)

$$C = \left\{\frac{k^2\cdot t_q^{2(k-1)}}{n\left[k - 1 + t_q^{2(k-1)}\right]}\right\}^{1/2}. \qquad (1.12)$$

Для наших умов за формулою (1.12) $C \approx 2{,}846$. Тоді, згідно з оцінкою (1.10), $\frac{20{,}0-12{,}1}{1{,}27} = 4{,}89 > C = 2{,}846$, тому спостереження $x_{10} = 20$ можна виключити з подальших розрахунків.

Контрольні питання

1. Що таке *похибка вимірювання*?
2. Чим абсолютна похибка відрізняється від відносної?
3. Що таке приладова (систематична) похибка?
4. Що таке модельна похибка?
5. Що таке випадкова похибка і які причини приводять до її появи?
6. Операції з наближеними числами.
7. Помилки вимірювання і міри точності.
8. Методи виключення грубих помилок.

Вихідні дані та завдання для індивідуального розв'язання видаються викладачем.

Тема 3 Системний підхід і системний аналіз

Система – це сукупність елементів, які певним чином взаємопов'язані, утворюють цілісність, а також взаємодіють між собою для виконання заданих цільових функцій. Система утворює особливу єдність з середовищем та є елементом «надсистеми». Зі свого боку й елементи системи можна розглядати як системи, якщо визначити інший критерій декомпозиції.

Елемент – це межа членування систем із погляду конкретного аспекту розгляду системи, розв'язання конкретного завдання, досягнення поставленої мети. Систему можна розчленовувати на елементи різними способами залежно від формулювання завдання, мети і її уточнення в процесі системного аналізу. Сукупність n ізольованих елементів ще не є системою. Для їх вивчення можна провести не більше n дослідних процедур. Водночас: для Дослідження системи із n взаємозв'язаних елементів необхідно вивчити n(n-1) зв'язків.

Способи опису систем Різноманітні описи системи відображають певні групи їх властивостей і дозволяють виявити впорядкованість, структурованість і функціональну організованість системи. Будь-яка система або об'єкт передусім цікаві своїм призначенням, місцем, яке вони займають серед інших систем і об'єктів в навколишньому світі, своєю функцією. Тому для характеристики системи передусім повинен бути одержаний функціональний опис (далі – ФО), який дозволяє оцінити призначення системи, її відношення до інших систем, її контакти з навколишнім світом, напрями можливих функціональних змін. Функціональний опис пов'язує зовнішні впливи на систему з її реакцією, відповіддю, поведінкою, дією на елементи системи. ФО може задаватися деяким оператором в алгебраїчній, логічній, диференціальній, інтегрально-диференціальній формі, який входить в скалярне, векторне або матричне рівняння. Оператор складається на основі вимірювання зовнішніх характеристик (принцип чорного ящика: вивчення зв'язку – вплив – реакція) або на основі знань про будову системи. ФО виходить із цільових функцій системи. Послідовність дій при виконанні системою деякої функції відображає зміст закону поведінки, яка залежить від процесів, які протікають від процесів, які протікають в середині системи (закони внутрішньої поведінки), і від процесів, в які залучена вся система в рамках метасистеми (закони зовнішньої поведінки). Кожна підсистема сама містить набір елементів, що виконують свої частки функції, тому закони внутрішнього функціонування системи одночасно є законами зовнішнього функціонування для будь-якої підсистеми цієї системи – закони зовнішнього функціонування першого нижнього рівня. Уявлення про будову системи дає її морфологічний опис (МО), що дозволяє виділити основні елементи, зв'язки, визначити тип структури. Можливими є три варіанти представлення морфологічного опису: з позицій теоретико-множинного підходу, у вигляді структури (найпоширеніші типи структур наведено на рис. 1.1) та у вигляді відповідної матриці суміжностей.

Структура системи може характеризуватися типами зв’язків, які в ній переважають (рис. 1.2).

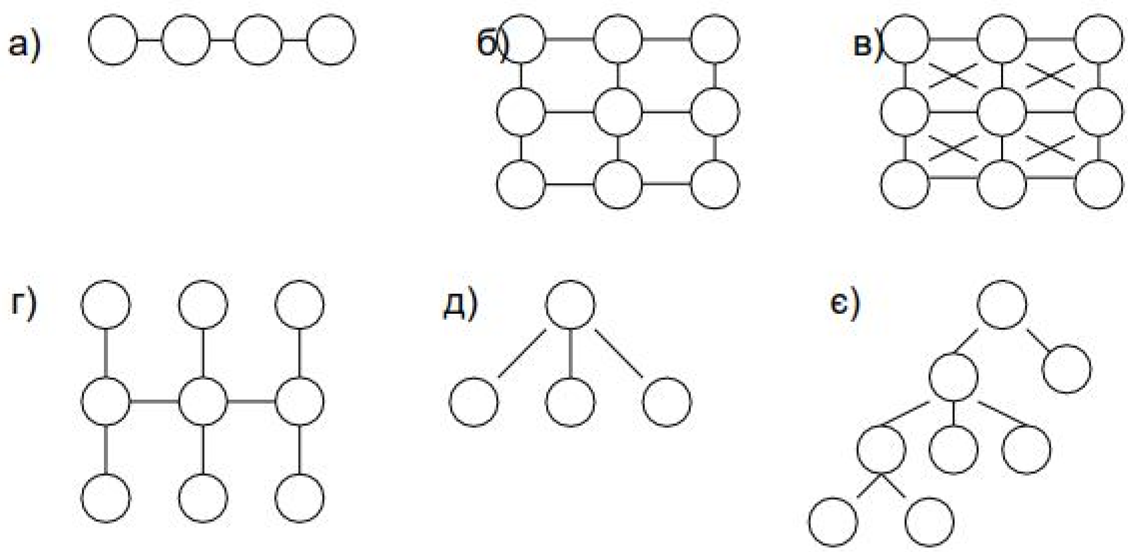


Рисунок 1.2 – Приклади основних типів структур системи а) – лінійна; б) – матрична; в) – мережева; г) – кістякова; д) – ієрархічна; є) – деревовидна

Найпростішими зв’язками є паралельні та послідовні. Залежно від характеру внутрішньої організації системи та зв’язків між елементами виокремлюються основні типи структур, які можна зобразити графічно, у вигляді опису (вербально), матриць або іншими способами (рис. 1.2).

Ці два види опису системи доповнюються третім – інформаційним описом (далі – ІО), що дозволяє судити про рівень її організації (дезорганізації), передбачити в імовірнісному розумінні реакції системи на той чи інший вплив. Сюди входить також характеристика інформаційних потоків, циркулюючих в системі і дані про алгоритми взаємодії елементів. Четвертий вид опису системи пов'язаний із характеристикою процесів зародження системи і еволюцією її розвитку в історичному плані – це генетико-прогностичний опис (далі – ГПО).

Контрольні питання

1. Що таке система?
2. Що таке елемент системи?
3. Що таке декомпозиція?
4. Охарактеризуйте функціональний опис системи.
5. Охарактеризуйте морфологічний опис системи.
6. Охарактеризуйте інформаційний опис системи.
7. Що таке структура? Які види ієрархічних структур вам відомі?
8. Як представити структуру у вигляді графа?

Вихідні дані та завдання для індивідуального розв'язання видаються викладачем.

ЗМІСТОВИЙ МОДУЛЬ 2

МЕТОДИ ЕКСПЕРИМЕНТАЛЬНИХ ДОСЛІДЖЕНЬ

Теми 4,5 Обробка результатів наукових досліджень за допомогою методів кореляційного та регресійного аналізів. Методи графічного зображення результатів експериментів

Теорія статистичного оцінювання розглядає два основні види оцінок: точкові та інтервальні.

Точковою оцінкою називають деяку функцію результатів спостереження $S_n^*(x_1, x_2, \ldots, x_n)$, значення якої за цих умов береться, як найбільше наближення до значення параметра S генеральної сукупності.

Проте при вибірці невеликого обсягу точкова оцінка S_n^* може істотно відрізнятися від дійсного значення параметра, тобто спричиняти грубі помилки, тому у разі малої вибірки зазвичай використовують інтервальні оцінки.

Інтервальною оцінкою називають кількісний інтервал (S_1^*, S_2^*), визначуваний за результатами вибірки, відносно якого можна стверджувати з визначеною, близькою до одиниці, ймовірністю, що він містить значення оцінюваного параметра генеральної сукупності.

Середні значення та їхні оцінки

Випадкова величина x повністю задається **щільністю ймовірності** $\rho(x)$ (інші назви – розподіл імовірності, розподіл величини *x).*

Середнє значення $\langle x \rangle$ вимірюваної величини x вказує центр розподілу, біля якого групуються результати окремих вимірювань:

$$\langle x \rangle = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Методи обчислення середніх

Приклад 2.1. Змішуються дві рідини з об'ємом 2 і 5 літрів з концентраціями якоїсь речовини 500 і 200 мг/л відповідно. Визначити концентрацію речовини в суміші.

Розв'язання. За формулою (2.1) маємо:

$$\bar{x} = \sum p_i x_i / \sum p_i = (2 \cdot 500 + 5 \cdot 200)/(2+5) = 285{,}7 \text{ мг/л}, \tag{2.1}$$

$$S_* = \sqrt{\sum p_i (x_i - \bar{x})^2 / \sum p_i}\ , \tag{2.2}$$

$$S_* = \sqrt{\frac{2(500-285{,}7)^2 + 5(200-285{,}7)^2}{2+5}} = 135{,}5 \text{ мг/л.}$$

Зважене стандартне відхилення визначається за формулою (2.3)

$$S = \sqrt{\frac{S_1^2(n_1-1) + S_2^2(n_2-1) + ... + S_k^2(n_k-1)}{\sum n_k - k}}. \tag{2.3}$$

Приклад 2.2. Нехай серія з *N* дослідів дала такі значення:

$n_1 = 10,\quad \bar{x}_1 = 10,\quad (S_1 = 2),\quad S_1^2 = 4,$

$n_2 = 12,\quad \bar{x}_2 = 9,\quad (S_2 = 3),\quad S_2^2 = 9,$

$n_3 = 6,\quad \bar{x}_3 = 7,\quad (S_3 = 1),\quad S_3^2 = 1.$

Потрібно визначити вибіркове зважене середнє $\bar{x}$, зважене стандартне відхилення і відношення стандартного відхилення до середнього значення (коефіцієнт варіації V).

Розв'язання.

$$\bar{x} = \frac{(10 \cdot 10 + 12 \cdot 9 + 6 \cdot 7)}{28} = 8{,}93;$$

$$S = \sqrt{\frac{4 \cdot (10-1) + 9 \cdot (12-1) + 1 \cdot (6-1)}{28-3}} = 2{,}37;$$

$$S^2 = 5{,}6.$$

Коефіцієнт варіації, що визначається за формулою (5.4.9) [3]:

$$V = S / \bar{x}, \quad \bar{x} \neq 0, \quad V = 2{,}37 / 8{,}93 = 0{,}265.$$

Приклад 2.3. Вибірка, приведена в таблиці 2.1. Визначити значення медіани Me і моди Mo.

Розв'язання. Оскільки медіана розміщується між двадцятим і двадцять першим членами ряду, а 2+3+5+6 = 16 і 2+3+5+6+10 = 26, то *Me* належить до п'ятого класу (інтервалу). Тоді при $x_{Me} = 9$, $b = 2$, $n = 40$, $\sum m_{Me-1} = 16$, $m_{Me} = 10$ за формулою маємо:

$$Me = \tilde{x} = x_{Me} + b \cdot \left(\frac{0{,}5 \cdot n - \sum m_{Me-1}}{m_{Me}} \right), \tag{2.4}$$

де x_{Me} – нижня межа класу (інтервалу), що містить медіану; b – ширина медіанного інтервалу; $\sum m_{Me-1}$ – сума накопичених частот, передуючих медіанному інтервалу; m_{Me} - частота медіанного інтервалу.

$$Me = \tilde{x} = 9 + 2 \cdot \left(\frac{0{,}5 \cdot 40 - 16}{10} \right) = 9{,}8.$$

Таблиця 2.1 – Розрахункова вибірка

Інтервал	Середина інтервалу	Частота m_i	Сума накопичених частот
1 <x <3	2	2	2
3 <x <5	4	3	5(2 + 3)
5 <x <7	6	5	10(5 + 5)
7 <x <9	8	6	16(10 + 6)
9 < x <11	10	10	26(16 + 10)
11 <x <13	12	9	35(26 + 9)
13 <x <15	14	5	40(35 + 5)
		$n = 40$	

Для нашого випадку середнє значення вибірки за формулою (2.1)

$$\bar{x} = \frac{(2\cdot 2 + 4\cdot 3 + 6\cdot 5 + 8\cdot 6 + 10\cdot 10 + 12\cdot 9 + 14\cdot 5)}{40} = 9{,}3\,.$$

Величину *Mo* підрахуємо за [3]:

$$Mo = x_{Mo} + b\cdot\frac{m_{Mo} - m_{Mo-1}}{2\cdot m_{Mo} - m_{Mo-1} - m_{Mo+1}}, \qquad (2.5)$$

де x_{Mo} – нижня межа інтервалу, що містить моду; b – величина модального інтервалу; m_{Mo} – частота модального інтервалу; m_{Mo-1} – частота інтервалу, передуючого модальному; m_{Mo+1} – частота інтервалу, наступного за модальним.

Теоретичні середні (моменти розподілу)

Третій центральний момент слугує для характеристики *асиметрії* (або *скошеності)* розподілу (рис. 2.2) і визначається за формулою

$$\mu_3 = \sum_1^n (x_i - \bar{x})^3/(n-1)\,. \qquad (2.6)$$

Четвертий центральний момент слугує для характеристики «крутизни» розподілу.

$$\mu_4 = \sum_1^n (x_i - \bar{x})^4/(n-1)\,. \qquad (2.7)$$

На практиці зручніше користуватися безрозмірними величинами, наприклад, у цьому випадку – *коефіцієнтом асиметрії* A_x

$$A_x = \mu_3 / S_x^3 \qquad (2.8)$$

та *ексцесом*

$$C_x = (\mu_4 / S_x^4) - 3\,. \qquad (2.9)$$

Приклад 2.4. За умовами прикладу 2.3 в нашому випадку маємо:

$$\sum_1^n (x_i - \bar{x})^3 = 2\cdot(2-9{,}3)^3 + 3\cdot(4-9{,}3)^3 + 5\cdot(6-9{,}3)^3 + 6\cdot(8-9{,}3)^3 +$$

$$+10\cdot(10-9{,}3)^3 + 10\cdot(10-9{,}3)^3 + 9\cdot(12-9{,}3)^3 + 5\cdot(14-9{,}3)^3 = -717{,}8.$$

Тоді μ_3 = -717,8/(40 - 1) = -18,4,

Звідки за формулою (2.8) при S = 3,34 A_x = -18,4/3,34^3 = -0,494.

Тоді μ_4 = 11577,4 / 39 = 296,8, звідки за формулою (2.9) маємо

$$C_x = (296,8/3,34^4)-3 = -0,615.$$

Помилка в оцінці показника асиметрії:

$$S_A = \sqrt{\frac{6\cdot(n-1)}{(n+1)\cdot(n+3)}}. \qquad S_A = \sqrt{\frac{6\cdot(40-1)}{(40+1)\cdot(40+3)}} = 0{,}364, \tag{2.10}$$

а розсіяння A_x і S_A визначаються дисперсіями:

$$S_{Ax} = \sqrt{6/(n+3)}, \tag{2.11}$$

$$S_{SA} = 2\cdot\sqrt{6/(n+5)}, \tag{2.12}$$

$$S_{Ax} = \sqrt{6/(40+3)} = 0{,}374, \qquad S_{SA} = 2\cdot\sqrt{6/(40+5)} = 5{,}480.$$

Якщо у формулах (2.8) та (2.9) замість 5 обрати S_* = 3,30, то A_{x*} = -0,512, а C_{x*} = -0,497.

Приклад 2.5. Нехай для сорока вимірювань, результати яких наведені в прикладі 2.3, відомо, що S = 3,30, $\bar{x}$ =9,3. Потрібно оцінити дійсне значення вимірюваної величини X з надійністю P = 0,95.

Розв'язання. За [3] для *P = 2Φ(t)* = 0,95 знаходимо t = 1,960. Отже, з надійністю 0,95 за залежністю (2.13) можна вважати, що довірча оцінка виглядає так:

$$|X-\bar{x}| < t(P)\cdot\sigma/\sqrt{n}, \tag{2.13}$$

$$|X-\bar{x}| = |X-9{,}3| < 1{,}960\cdot(3{,}3/\sqrt{40}) = 1{,}02,$$

тобто значення X лежить в інтервалі (8,28; 10,32) або $X = \bar{x} \pm 1{,}02$.

<u>*Приклад 2.6.*</u> За умовами попереднього прикладу за [3] маємо: t(0,95; 40 - 1) = 2,023. Тоді за формулою

$$|X-\bar{x}| < t(P;k)\cdot S/\sqrt{n-1}, \tag{2.14}$$

$$|X-\bar{x}| < 2{,}023\cdot(3{,}3/\sqrt{40-1}) = 1{,}07,$$

тобто. значення X лежить в інтервалі X = 9,3 ± 1,07.

За правилом трьох сигм [3] при невідомій σ

$$|X-\bar{x}| < 3S/\sqrt{n}, \tag{2.15}$$

а для нашого випадку

$$X = 9{,}3 \pm 3\cdot 3{,}3/\sqrt{40} \approx 9{,}3 \pm 1{,}6.$$

Приклад 2.7. У циліндрі з висотою стовпа води 0,7-1,0 м визначався час осадження окремих частинок піску. Розрахункові значення швидкості осадження (v) наведені в таблиці 2.2. Потрібно знайти розподіл статистичної ймовірності швидкості осадження (гідравлічної крупності) частинок піску фільтруючого матеріалу. Провести оцінку математичного очікування і средньоквадратичного відхилення. Оцінити найбільш вірогідну швидкість осадження частинок піску.

Розв'язання.

1. Діапазон зміни швидкостей осадження від $V_{\max}$= 36,1 до $V_{\min}$ =15,1 ділимо на n інтервалів. Кількість інтервалів можна вибирати за напівемпіричною формулою

$$n = 1 + 3{,}32 \cdot \lg N = 1 + 3{,}32 \cdot \lg 100 \approx 7{,}64\,, \qquad (2.16)$$

де N – кількість частинок, для яких проводилося визначення швидкостей осадження (у нашому випадку N = 100).

Кількість інтервалів n, одержане за формулою, округляється до найближчого цілого (приймаємо n = 7).

2. Визначається довжина інтервалу варіаційного ряду:

$$\Delta = (x_{\min} - x_{\max}) / n = (36{,}1 - 15{,}1) / 7 = 3{,}0\,.$$

Таблиця 2.2 – Результати визначення швидкостей осадження частинок піску

№ п.п.	V, см/с	№ п.п.	V, см/с	№ п.п.	V, см/с	№ п.п.	V, см/с	№ п.п.	V, см/с
1	2	3	4	5	6	7	8	9	10
1	22,4	21	22,4	41	23,2	61	25,0	81	24,1
2	21,0	22	21,0	42	27,1	62	24,1	82	21,7
3	25,0	23	26,0	43	30,9	63	31,0	83	21,7
4	22,4	24	21,0	44	24,1	64	25,0	84	28,3
5	15,1	25	34,2	45	20,3	65	23,2	85	26,0
6	23,2	26	32,5	46	20,1	66	24,1	86	22,4
7	16,7	27	31,0	47	19,7	67	24,1	87	22,4
8	19,1	28	31,0	48	21,7	68	29,6	88	26,0
9	26,0	29	26,0	49	25,0	69	28,3	89	21,7
10	24,1	30	28,3	50	24,1	70	27,1	90	21,7
11	23,2	31	28,3	51	26,0	71	27,1	91	21,7
12	27,1	32	21,0	52	26,0	72	28,3	92	26,0
13	19,7	33	29,5	53	24,1	73	23,2	93	27,1
14	19,7	34	28,3	54	21,7	74	26,0	94	21,7
15	21,0	35	26,0	55	25,0	75	26,0	95	22,4
16	26,0	36	25,0	56	24,1	76	25,0	96	22,4
17	32,5	37	23,2	57	29,5	77	36,1	97	30,9
18	28,3	38	34,2	58	25,0	78	30,9	98	28,3
19	24,1	39	23,2	59	31,0	79	28,3	99	27,1
20	26,0	40	20,3	60	27,1	80	27,1	100	26,0

3. Визначається кількість m числа частинок піску, швидкості яких потрапили у відповідні інтервали (x_{m-1}, x_m).

4. Визначається статистична імовірність (частість) попадання швидкості осадження у відповідний інтервал.

$$P_i = m_i / N\,. \qquad (2.17)$$

Результати розрахунків представлені в таблиці 2.3.

Таблиця 2.3 – Визначення частості попадання

Номер інтер-валу	Інтервали варіаційних рядів	Кількість попадань в інтервал m_i	Частотність, $P_i = m_i /N$	Середина Інтервалу, $\bar{x}_i$	Сумарна частотність
1	15,1 – 18,1	2	0,02	16,6	0,02
2	18,1 – 21,1	12	0,12	19,6	0,14
3	21,1 – 24,1	23	0,23	22,6	0,37
4	24,1 – 27,1	31	0,31	25,6	0,68
5	27,1 – 30,1	20	0,20	28,6	0,88
6	30,1 – 33,1	9	0,09	31,6	0,97
7	33,1 – 36,1	3	0,03	34,6	1,00

За результатами розрахунків побудуємо гістограму розподілу швидкості осадження частинок піску (рис. 2.4). З рисунка видно, що найбільш вірогідне значення швидкості осадження – це 25,0-26,0 см/с (25,6 см/с). Статистична імовірність появи даної випадкової величини складає 31,0 %.

Визначимо параметри статистичного розподілу:

– математичне очікування $\bar{x}$ і стандартне відхилення S по формулах (1.1) та (1.5)

$$S = \sqrt{\begin{gathered}0{,}02 \cdot (16{,}6 - 25{,}4)^2 + 0{,}12 \cdot (19{,}6 - 25{,}4)^2 + 0{,}23 \cdot (22{,}6 - 25{,}4)^2 + 0{,}31 \cdot (25{,}6 - 25{,}4)^2 + \\ + 0{,}20 \cdot (28{,}6 - 25{,}4)^2 + 0{,}09 \cdot (31{,}6 - 25{,}4)^2 + 0{,}03 \cdot (34{,}6 - 25{,}4)^2\end{gathered}} =$$

$$= \sqrt{15{,}715} = 3{,}96$$

$$S = 3{,}96 \text{ см/с} \approx 4{,}0 \text{ см/с}.$$

Згідно з правилом трьох сигм (3σ), математичне очікування швидкості осідання лежить в межах

$\bar{x} - \dfrac{3S}{\sqrt{n}} \le \bar{X} \le \bar{x} + \dfrac{3S}{\sqrt{n}}$, тобто в нашому випадку $24{,}2 \le \bar{X} \le 26{,}6$.

Величину математичного очікування можна оцінити і за формулою (2.15).

Порівняння дисперсій і середніх значень

Порівняння дисперсій

Порівняння двох дисперсій нормальних генеральних сукупностей

Приклад 2.8. За двома незалежними вибірками $n_1 = 10$ і $n_2 = 15$, витягнутих із нормальних сукупностей X_1 и X_2, знайдені виправлені вибіркові дисперсії S^2_{x1} = 0,86 и S^2_{x2} = 0,43. При q = 0,05 перевірити H_o: $D(X_1) = D(X_2)$ при конкуруючій гіпотезі H_1: $D(X_1) > D(X_2)$.

Розв'язання. $F_{спосте}$ = 0,86 / 0,43 = 2,0, оскільки H_1: $D(X_1) > D(X_2)$ - критична область правостороння. За [3] $F_{кр.}$ = 2,65 при q = 0,05; k_1 = 10-1 = 9; k_2 = 15-1 = 14.

Оскільки $F_{спостер.} < F_{кр.}$, то вибіркові генеральні дисперсії розрізняються незначно.

Приклад 2.9. Щільність зерен в шматку шихти (г/м[3]), узятої з двох незалежних вибірок X_1 і X_2, приведена нижче:

x_1	2,41	2,38	2,43	2,50	2,44	2,41	2,52
x_2	2,65	2,70	2,73	2,66	2,63	2,71	2,73

При q = 0,1 перевірити H_o: $D(X_1) = D(X_2)$ *при конкуруючій гіпотезі* H_1: $D(X_1) \neq D(X_2)$.

Розв'язання. За формулою (1.1) визначаємо $\bar{x}_i$, а за формулою (1.5) – S_i. Тоді $\bar{x}_1$ = 2,44; $\bar{x}_2$ = 2,69; S_1^2= 0,05; S_2^2 = 0,04. При цьому $F_{набл.}$ = 0,05 / 0,04 = 1,25. Критична область двостороння, оскільки конкуруюча гіпотеза має вигляд $D(X_1) - D(X_2)$. За цих умов за правилом № 2 [3, с. 104] при k_1 = 7-1 = 6 і k_2 = 7-1 = 6 і q_1= q / 2 = 0,1 / 2 = 0,05

Маємо: $F_{кр}$ = 4,28 > $F_{спостер.}$, тобто гіпотезу H_1 відкидаємо, оскільки обидва вимірювання були виконані з однаковою точністю. Можна переходити до порівняння середніх (див. приклади 2.12 и 2.14).

Порівняння виправленої вибіркової дисперсії з гіпотетичною генеральною дисперсією нормальної сукупності

Приклад 2.10. За трьома незалежними вибірками однакового обсягу п = 11 з нормальних генеральних сукупностей визначені вибіркові дисперсії: S_1^2= 0,25; S_2^2= 0,36; S_3^2= 0,49. Потрібно: а) при рівні значущості q = 0,05 перевірити нульову гіпотезу про однорідність дисперсій; б) оцінити генеральну дисперсію; в) розрахувати інтервальні оцінки дисперсій.

Розв'язання: А. Розрахункове значення критерію Кохрена:

$$G_{набл.} = S_{\max}^2 / \sum_{i=1}^{l} S_i^2 = \frac{0,49}{(0,25 + 0,36 + 0,49)} = 0,4454. \qquad (2.18)$$

За [3] при рівні значущості q = 0,05, кількості ступенів свободи k = 11 - 1 = 10 і кількості вибірок l = 3 критична точка $G_{кр.}$(0,05; 10; 3) = 0,6025. Оскільки $G_{набл.}$ < $G_{кр}$, вважаємо, що виправлені вибіркові дисперсії розрізняються незначуще (тобто дисперсії однорідні).

Б. Враховуючи однорідність дисперсій S_1^2 і S_2^2, як оцінку генеральної дисперсії приймаємо середню арифметичну виправлену дисперсію:

$$D_r = (0,25 + 0,36 + 0,49)/3 \approx 0,37. \qquad (2.19)$$

В. За [3] $\chi^2_{q/2}$ = 20,5, а $\chi^2_{1-q/2}$ = $\chi^2_{0,975}$ = 3,25. Тоді інтервальну оцінку дисперсії можна визначити з такої умови:

$$\frac{S_x^2 \cdot k}{\chi^2_{q/2}} \leq \sigma_x^2 \leq \frac{S_x^2 \cdot k}{\chi^2_{1-q/2}}, \qquad (2.20)$$

де σ_x^2 – дисперсія змінної *X;* $\chi^2_{q/2}$ – значення χ^2–розподілу для рівня значущості *q/*2; $\chi^2_{1-q/2}$ – значення χ^2– розподілу при рівні значущості (*1-q/*2); *k* = *n* - 1 – кількість ступенів свободи; S_x^2 – виправлена вибіркова дисперсія.

$$\frac{0{,}25\cdot 10}{20{,}5} \le \sigma_{x1}^2 \le \frac{0{,}25\cdot 10}{3{,}25};\quad 0{,}12 \le \sigma_{x1}^2 \le 0{,}77;\ 0{,}18 \le \sigma_{x2}^2 \le 1{,}11;\ 0{,}24 \le \sigma_{x3}^2 \le 1{,}51.$$

Порівняння двох середніх генеральних сукупностей, дисперсії яких відомі (великі незалежні вибірки пі > 30)

Приклад 2.11. Є дві незалежні вибірки п₁ = 50 і п₂ = 60, обрали на підставі нормальних генеральних сукупностей концентрації зважених речовин у річці до і після населеного пункту. Знайдені вибіркові середні: $\bar{x}_1$ =10 мг/л і $\bar{x}_2$ = 13 мг/л. Генеральні дисперсії відомі: $\sigma^2_{1(x_1)}$ = 8, $\sigma^2_{1(x_2)}$ = 10. Потрібно при рівні значущості q = 0,01 перевірити нульову гіпотезу Hₒ: $M(x_1) = M(x_2)$ проти гіпотези:

а) H₁: $M(x_1) \ne M(x_2)$; *б) H₁:* $M(x_1) < M(x_2)$.

Розв'язання: Знайдемо спостережуване значення критерію:

$$z_{набл.} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma^2_{(x_1)}}{n_1} + \frac{\sigma^2_{(x_2)}}{n_2}}} = \frac{10-13}{\sqrt{\frac{8}{50} + \frac{10}{60}}} = -5{,}25. \qquad (2.21)$$

За умовою (*а*) конкуруюча гіпотеза виглядає так: *H₁:* $M(x_1) \ne M(x_2)$, тому критична область двостороння. Знайдемо праву критичну точку з рівності

$$\Phi(z_{кр}) = (1-q)/2 = (1-0{,}01)/2 = 0{,}495. \qquad (2.22)$$

За таблицею функції Лапласа (див [3]) знаходимо $z_{кр}$ = 2,58. Оскільки $|z_{спостер.}| > z_{кр}$, то, відповідно до правила 1, нульову гіпотезу відкидаємо, тобто вибіркові середні розрізняються значущо.

За умовою (*б*) конкуруюча гіпотеза виглядає так: *H₁:* $M(x_1) < M(x_2)$. Знайдемо допоміжну точку з рівності

$$\Phi(z_{кр}) = (1-2q)/2 = (1-2\cdot 0{,}01)/2 = 0{,}490. \qquad (2.23)$$

За таблицею Лапласа (Додаток А) знаходимо $z_{кр}$ = 2,33. Оскільки $z_{спостер.}$= –5.25 < –2,33, то нульову гіпотезу [3] відкидаємо і вважаємо, що $M(x_1) < M(x_2)$.

Після знаходження $z_{спостер.}$ за формулою (2.21) задаємо бажану імовірність виведення *P* і за нею знаходимо в [3] відповідне значення *t(P)* (наприклад, при *P* = 0,99 знаходимо *t* = 2,576, а при *P* = 0,95 визначаємо *t* = 1,960). Тоді, якщо $|z_{спостер.}|$ > *t(P),* то розбіжність середніх значень можна вважати значущою з надійністю виведення *P.*

Приклад 2.12. Дано дві незалежні малі вибірки $n_1 = 16$ і $n_2 = 20$ з нормальної генеральної сукупності X_1 і X_2 з вибірковими середніми $\bar{x}_1 = 42$ і $\bar{x}_2 = 40$ та виправленими дисперсіями $S_{x1}^2 = 0{,}70$ і $S_{x2}^2 = 0{,}37$. Потрібно при рівні значущості q = 0,05 перевірити гіпотезу H_o: $M(x_1) = M(x_2)$ на противагу альтернативній гіпотезі:

$$а)\ H_1:\ M(x_1) \neq M(x_2); \quad б)\ H_1:\ M(x_1) > M(x_2).$$

Розв'язання. Виправлені дисперсії різні, тому спочатку перевіримо гіпотезу про рівність генеральних дисперсій, використовуючи критерій Фішера (див. приклади 2.8, 2.9).

Знайдемо відношення більшої дисперсії до меншої:

$$F_{розр.} = 0{,}70 / 0{,}37 = 1{,}89.$$

Беручи до уваги, що дисперсія S_{x1}^2 значно більша S_{x2}^2, приймаємо як конкуруючу гіпотезу H_1: $D(x_1) > D(x_2)$. Тут критична область – правостороння. За [3], при рівні значущості $q = 0{,}05$ і ступенях свободи $k_1 = n_1 - 1 = 16 - 1 = 15$ і $k_2 = n_2 - 1 = 20 - 1 = 19$ знаходимо критичну точку $F_{кр}$ (0,05; 15; 19) = 2,23.

Оскільки $F_{розр.} < F_{кр}$, то приймається гіпотеза H_0 про рівність генеральних дисперсій. Якщо гіпотеза про рівність дисперсій підтверджується, то можна порівняти середні.

Визначимо розрахункове значення критерію Ст'юдента за формулою

$$T = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{(n_1 - 1)\cdot S_{x1}^2 + (n_2 - 1)\cdot S_{x2}^2}} \cdot \sqrt{\frac{n_1 \cdot n_2 \cdot (n_1 + n_2 - 2)}{n_1 + n_2}}. \qquad (2.24)$$

Якщо

$$\left|T_{спост.}\right| < t_{двост.кр.}(q; k = n_1 + n_2 - 2), \qquad (2.25)$$

то гіпотеза H_0 приймається, якщо не виконується, то приймається конкуруюча гіпотеза H_1.

$$T_{\hat{o}\hat{i}\c{c}\eth.} = \frac{42 - 40}{\sqrt{(16-1)\cdot 0{,}70 + (20-1)\cdot 0{,}37}} \cdot \sqrt{\frac{16 \cdot 20 \cdot (16 + 20 - 2)}{16 + 20}} = 8{,}30.$$

За умовою (*а*) конкуруюча гіпотеза має вигляд H_1: $M(x_1) \neq M(x_2)$, тому при рівні значущості 0,05 і кількості ступенів свободи $k = n_1 + n_2 - 2)$ = 16 + 20 - 2 = 34 знаходимо критичну точку $t_{двост.кр.}$ (0,05; 34) = 2,032. Оскільки $T_{розр.} = 8{,}3 > t_{двост.кр.}$. = 2,032, тобто умова (2.25) не виконується, то вважаємо, що значення $\bar{x}_1$ і $\bar{x}_2$ відрізняються значно.

За умовою (*б*) конкуруюча гіпотеза виглядає, як H_1: $M(x_1) > M(x_2)$, тому при q =- 0,05 і k = 34 знаходимо критичну точку $t_{правост.кр.}$= 1,69. Оскільки умова (2.25) не виконується, вважаємо, що $\bar{x}_1$ дійсно більше $\bar{x}_2$.

Перевірка гіпотези про рівність середніх

Методи порівняння середніх значень, зазначені в прикладах 2.11 і 2.12, можна застосувати і для розв'язання задачі перевірки гіпотези про рівність середніх. Порівняння вибіркової середньої з гіпотетичної генеральної середньої нормальної сукупності дисперсія генеральної сукупності невідома.

Приклад 2.13. За вибіркою з обсягом п = 6, обраної за нормальною генеральною сукупністю, знайдене вибіркове середнє межі міцності при стискуванні зразків-балочок $\bar{x}$ = 431,1 кгс/см^2 і виправлене середнє квадратичне відхилення S = 4,71. Потрібно при рівні значущості q = 0,05 перевірити гіпотезу H_0: a = a_0 = 435 при конкуруючій гіпотезі H_1: $a \neq a_0$= 435.

Розв'язання: Знаходимо критерій перевірки *T* за формулою

$$T = (\bar{x} - a_0)\sqrt{n} / S, \tag{2.26}$$

де S - виправлене среднє квадратичне відхилення, що визначається за формулою (1.5), має розподілення Ст'юдента з $k = n - 1$ ступенями свободи.

$$T = \frac{(431{,}1 - 435{,}0)\sqrt{6}}{4{,}71} = -2{,}03.$$

При $k = n - 1$= 6 - 1 = 5 находимо критичну точку $t_{двост.кр.}$(0,05; 5) = 2,57. Оскільки виконуується умова

$$T_{спост.} < t_{двост.кр.}(q;\, k = n - 1), \tag{2.27}$$

то гіпотеза H_0: $a = a_0$, приймається, тобто приходимо до висновку, що вибіркове среднє $\bar{x}$ = 431,1 частвідрізняється від гіпотетичної генеральної середньої a_0 = 435.

Порівняння двох середніх нормальних генеральних сукупностей з невідомими дисперсіями (залежні вибірки)

Приклад 2.14. Знайдено такі межі міцності при розколюванні двох складів шихти, МПа:

x_1	20,1	19,9	20,3	18,3	21,0	20,7	20,3
x_2	22,7	23,0	22,9	22,9	21,3	22,4	21,0

При рівні значущості q = 0,05 встановити, значущо або незначущо розрізнюванні середні значення межі міцності. Передбачається, що розподіл нормально.

Розв'язання. Знайдемо різницю, віднімаючи з чисел першого рядка числа другої. Одержимо:

d_i: - 2,6; - 3,1; - 2,6; - 4,5; - 0,3; - 1,7; - 0,7.

Знайдемо вибіркову середню, враховуючи, що $\sum d_i$ = -15,5; $\bar{d}$ = -15,5/7 = -2,2. Тоді виправлене середнє квадратичне відхилення

$$S_d = \sqrt{\frac{\sum d_i^2 - \left[\sum d_i\right]^2 / n}{n - 1}} = 1{,}44, \tag{2.28}$$

а спостережуване значення критерію $T_{спост}$

$$T_{спост} = \bar{d}\sqrt{n} / S_d = -2{,}2\sqrt{7} / 1{,}44 = -4{,}03. \tag{2.29}$$

За [3] знаходимо критичну крапку

$$t_{двост.кр.}(0{,}05;\, 6) = 2{,}45 < |T_{спост.}| = -4{,}03. \tag{2.30}$$

Оскільки умова $|T_{спост.}| < t_{двост.кр.}$ не виконується, вважаємо, що середні значення вимірювань межі міцності відрізняються значно.

Перевірка гіпотези нормальності закону розподілу випадкової величини

Приклад 2.15. З урахуванням вихідних даних, наведених в прикладі 2.3, і подальших розрахунках в прикладі 2.4, маємо: n = 40; l = 7; A_x *= -0,494;* C_x *= -0,615;* A_x^* *= -0,512;* C_x^* *= -0,497; x = 9,3;* S^* *= 3,30; S = 3,25. Потрібно перевірити гіпотезу нормальності розподілу випадкової величини наближеним методом.*

Розв'язання. Помилку оцінки показника асиметрії A_x визначимо за формулою

$$S_{Ax} = \sqrt{\frac{6 \cdot (n-1)}{(n+1) \cdot (n+3)}} = \sqrt{\frac{6 \cdot (40-1)}{(40+1) \cdot (40+3)}} = \sqrt{\frac{6{,}39}{41{,}43}} = 0{,}133\,, \qquad (2.31)$$

а помилку оцінювання C_x - за формулою

$$S_{Cx} = \sqrt{\frac{24 \cdot (n-2) \cdot (n-3)}{(n-1)^2 \cdot (n+3) \cdot (n+5)}} = \sqrt{\frac{24 \cdot 38 \cdot 37}{39^2 \cdot 43 \cdot 45}} = 0{,}107\,. \qquad (2.32)$$

Оскільки $$\frac{|A_x = -0{,}494|}{(2...3)} \approx (0{,}165...0{,}267) > S_{Ax} = 0{,}133\,,$$

а $$\frac{|C_x = -0{,}615|}{(2...3)} \approx (0{,}205...0{,}308) > S_{Cx} = 0{,}107\,,$$

то можна поставити під сумнів нормальність закону розподілу досліджуваної величини.

Застосуємо критерій χ^2, для чого вихідні дані і частину розрахунків подамо в табличній формі.

Для застосування критерію χ^2 об'єднаємо крайні інтервали, щоб кількість даних в кожному інтервалі було не менше п'яти. Одержані дані представлені в перших двох стовпцях таблиці 2.4. Крайні інтервали обрано як нескінченні. У третьому стовпці підраховано відношення

$$t_i = \frac{x_i - x}{S} = \frac{x_i - 9{,}3}{3{,}25}\,. \qquad (2.33)$$

Таблиця 2.4 – Перевірка нормальності закону розподілення

Інтервали	m_i	t_i	$\Phi(t_i)$	p_i	m_i - np_i	$(m_i - np_i)2 / np_i$
-∞ ... 5	5	-1,323	-0,4071	0,0929	1,284	0,444
5 ... 7	5	-0,708	-0,2605	0,1466	-0,864	0,127
7 ... 9	6	-0,092	-0,0367	0,2238	-2,952	0,973
9 ... 11	10	0,523	0,1991	0,2358	0,568	0,034
11 ... 13	9	1,138	0,3725	0,1734	2,064	0,614
13 ... +∞	5	∞	0,5000	0,1275	-0,100	0,002
	Σ 40			1,0000		2,194 = χ_p^2

Отже, для правого кінця, наприклад першого інтервалу, це буде t_1 = (5 - 9,3)/3,25 = - 1,323. У четвертому стовпці наведено відповідні значення інтеграла імовірності $\Phi(t_i)$ з [3]. За значенням $\Phi(t_i)$ з п'ятому стовпці обчислена імовірність p_i, як різниця відповідних значень, наприклад $\Phi(t): p_i = \Phi(t_i) - \Phi(t_{i-1})$, p_2 = - 0,2605 - (-0,4071) = 0,1466. При обчисленні ймовірності p_1 враховано, що $\Phi(-\infty)$ = - 0,5. Порівняння χ^2_p= 2,194 с $\chi^2_{кр.}$, одержаних при кількості ступенів свободи k = 6 - 3 = 3 і рівні значущості q = 0,05, доводить, що немає підстави сумніватися в нормальності розподілу, оскільки $\chi^2_p = 2{,}194 < \chi^2_{кр.} = 7{,}8$.

Приклад 2.16. Визначення теоретичного закону розподілу. За умовами прикладу 2.6 визначити теоретичний закон розподілу швидкості осідання зерен піску фільтруючого матеріалу.

Попередній аналіз одержаної гістограми показує, що теоретичним законом розподілу може бути нормальний закон (закон Гауса).

Для перевірки гіпотези про те, що швидкість осідання підпорядкована нормальному закону розподілу, використовуємо один із найбільш поширених критеріїв – критерій Пірсона (χ^2), який визначається за формулою:

$$\chi^2 = \sum_{n=1}^{k} \frac{(m_i - nP_i)^2}{nP_i}, \tag{2.34}$$

де m_i – кількість попадань у відповідний інтервал; n – об'єм вибірки (у нашому випадку n = 100); P_i – імовірність потрапляння в *і-й* інтервал; k – кількість інтервалів.

Критерій χ^2 передбачає порівняння кількості елементів, що потрапили в *і-й* інтервал з вибірки об'ємом n з кількістю елементів, які в середньому потрапляли у цей інтервал з вибірки того ж самого об'єму, у разі правильності висунутої нами гіпотези про закон розподілу. Інакше кажучи, порівнюється відносна частота потрапляння в інтервал з гіпотетичною імовірністю потрапляння в нього.

Критерій Пірсону має кількість ступенів свободи, що дорівнює $n = k - l - 1$, де l - кількість оцінюваних параметрів в законі розподілу.

При нормальному законі розподілу оцінюється дисперсія і математичне очікування, тобто l = 2.

При використанні критерію χ^2 необхідно враховувати, що інтервали з кількістю елементів менше 10 необхідно об'єднувати з сусідніми. При цьому кількість ступенів свободи, природно, зменшується.

Раніше одержаний варіаційний ряд (табл. 2.3) перетворюється на новий (табл. 2.5).

Таблиця 2.5 – Перетворений варіаційний ряд

Номер інтервалу	Інтервали варіаційних рядів	Кількість попадань в інтервал, m	Частість, $P_i = m_i / n_i$
1	15,1 - 21,1	14	0,14
2	21,1 - 24,1	23	0,23
3	24,1- 27,1	31	0,31
4	27,1 - 30,1	20	0,20
5	30,1 - 36,1	12	0,12

Розрахунки, пов'язані з перевіркою гіпотези про нормальний розподіл, проведемо в таблиці 2.6.

Таблиця 2.6 – Перевірка гіпотези нормальності розподілу

m_i	t_{i-1}	t_i	$\Phi(t_{i-1})$	$\Phi(t_i)$	$P_i = \Phi(t_i) - \Phi(t_{i-1})$	$n \cdot \overline{P}$	$m_i - n\overline{P}$	$\left(m_i - n\overline{P}\right)^2$	$\left(m_i - n\overline{P}\right)^2 / nP$
14	-2,60	-1,09	-0,495	-0,362	0,133	13,3	0,7	0,49	0,037
23	-1,09	-0,33	-0,362	-0,133	0,229	22,9	0,1	0,01	0,00
31	-0,33	0,43	-0,133	0,166	0,299	29,9	1,1	1,21	0,04
20	0,43	1,19	0,166	0,383	0,217	21,7	-1,7	2,89	0,133
12	1,19	2,70	0,383	0,496	0,113	11,3	0,7	0,49	0,043
100						Σ=99,			Σ = 0,253

Тут t_i і t_{i-1} – нормовані випадкові величини, визначувані за формулою:

$$t = (x_i - \overline{x}) / \sigma ,$$

Наприклад, (15,1 - 25,4) / 3,96 = -10,3 / 3,96 = -2,6;

(21,1 - 25,4) / 3,96 = -4,3 / 3,96 = -1,09.

x_i; x_{i-1} – нормовані межі інтервалів варіаційних рядів; $\Phi(t_i)$; $\Phi(t_{i-1})$ – значення нормованих функцій Лапласа від t_i і t_{i-1}, визначаються по таблицях ([3] у додатках); $P_i = \Phi(t_i) - \Phi(t_{i-1})$ – гіпотетична імовірність потрапляння випадкової величини в інтервал (t_{i-1} - t_i); nP – гіпотетична частота попадання випадкової величини у відповідний інтервал.

Порівняння розрахункової величини критерію χ^2 = 0,253 з критичним його значенням $\chi^2_{кр.}$ = 5,99, одержаним для заданого рівня значущості q = 5 % і кількості ступенів свободи n = 5 - 2 - 1 =2 свідчить про те, що $\chi^2 < \chi^2_{кр.}$ і що швидкість осадження зерен фільтруючого матеріалу має нормальний розподіл ймовірності генеральної сукупності, який можна виразити такою формулою:

$$y = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x_i - \overline{x})^2}{2\sigma^2} \right]. \qquad (2.35)$$

Після підстановки $\overline{x}$ = 25,4 і σ = 3,96 отримаємо

$$y = \frac{1}{3{,}96\sqrt{2 \cdot 3{,}14}} \exp\left[-\frac{(x_i - 25{,}4)^2}{2 \cdot 3{,}96^2} \right] = 0{,}1 \cdot \exp\left[-0{,}032 \cdot (x - 25{,}4)^2 \right],$$

де y - щільність ймовірності розподілу випадкової величини; x - випадкова величина (швидкість осадження).

Контрольні питання

1. Що таке середні значення та їхні оцінки?
2. Що таке перетворений варіаційний ряд?
3. Що таке нормальний закон розподілу (закон Гауса)?
4. Як відбувається перевірка гіпотези про рівність середніх?

Тема 6 Вибір емпіричної залежності МНК

Обробка результатів експерименту. Експеримент проводиться відповідно до вибраної матриці планування, або плану. План є множиною з N точок *K- мірного* факторного простору.

На цьому просторі діє декілька функцій, що описують поведінку відгуків плану. Значення відгуків за допомогою кодування за шкалою бажаності зводяться до одного значення узагальненого відгуку.

Таким чином, внаслідок проведення експерименту N точкам факторного простору, співставляється N експериментальних точок узагальненої функції відгуку.

Сама функція невідома і представлена внаслідок проведення одного кроку планування своїми N значеннями. З цих N значень вибирається найбільш оптимальне. Вибір відбувається відповідно до перевірки гіпотез про відмінність двох середніх і однорідності дисперсій.

Отже, вибрана одна точка факторного простору, що доставляє локальний оптимум на безлічі N точок факторного простору. Об'єднавши гладкість і безперервність функції відгуку, ми можемо підібрати таку межу цієї оптимальної точки, що функція узагальненого відгуку наближатиметься гіперплощиною $\Psi(x_1, x_2, ..., x_K)$, заданою на факторному просторі. Різниці

$$\xi_i = y_i - \Psi_i, \quad i = 1,2,...,N$$

називаються незв'язаністю.

Величина незв'язаності визначається, як помилка експерименту, і як придатність лінійної моделі, пов'язаної величиною області експерименту. І обидві ці причини змішані.

Ступінь ефективності наближення функції узагальненого відгуку гіперплощиною визначається за допомогою оцінної функції, заданої на безлічі незв'язаності.

Як оцінні функції U, використовують суму квадратів або суму модулів незв'язаності, або мінімаксний критерій:

$$U = \sum_{i=1}^{N} \xi_i^2, \quad U = \sum_{i=1}^{N} |\xi_i|, \quad U = \min \max |\xi_i|$$

Загалом можуть бути використані й інші оцінки, наприклад, сума модулів кубів незв'язаності. Оцінки порівнюють за величиною похибки коефіцієнтів регресії і простотою обчислень. Щодо цього відношенні метод найменших квадратів найбільш оптимальний.

Метод найменших квадратів

Метод найменших квадратів приводить до мінімально можливої залишкової суми квадратів відгуку, або незв'язаності, і в цьому сенсі є оптимальною оцінкою коефіцієнтів моделі.

Функцію відгуку у сфері експерименту ми представляємо у вигляді лінійного рівняння за факторами:

$$y = b_0 + \sum_{j=1}^{M} b_j x_j$$

Таке рівняння називається рівнянням регресії. На кожному кроці планування експерименту обчислюються коефіцієнти b_j. Для кожного досліду рівняння регресії виконується приблизно, з точністю до незв'язаності:

$$y_i - b_0 - \sum_{k=1}^{M} b_k x_{ki} = \xi_i$$

Метод найменших квадратів полягає в тому, що коефіцієнти моделі b_j обирають з умови

$$U = \sum_{i=1}^{N} \xi_i^2 = \sum_{i=1}^{N} (y_i - b_0 - \sum_{k=1}^{M} b_k x_{ki})^2 = \min .$$

Для досягнення мінімуму функції $U(b_k)$ необхідно прирівняти всі приватні похідні до нуля: $\partial U / \partial b_k = 0$. Після простих перетворень одержуємо:

$$Nb_0 + \sum_{i=1}^{N} \sum_{k=1}^{M} b_k x_{ki} = \sum_{i=1}^{N} y_i \quad b_0 \sum_{i=1}^{N} x_{ki} + \sum_{i=1}^{N} \sum_{k=1}^{M} b_k x_{ki}^2 = \sum_{i=1}^{N} y_i x_{ki} .$$

Далі слід пригадати властивості матриці планування. Згідно з властивістю симетричності матриці планування сума, алгебри елементів будь-якого, вектор стовпця (будь-якого фактору) дорівнюють нулю:

$$\sum_{i=1}^{N} x_{ki} = 0$$

А за умовою нормування

$$\sum_{i=1}^{N} x^2{}_{ki} = N$$

Крім того, ми скористаємося ортогональністю матриці планування, а саме

$$\sum_{i=1}^{N} x_{ki} x_{mi} = 0, \quad k \neq m, \quad k, m = 1,2,..,M .$$

З урахуванням цих властивостей одержуємо відомі вирази для коефіцієнтів моделі плану:

$$b_0 = \frac{1}{N} \sum_{i=1}^{N} y_i , \quad b_k = \frac{1}{N} \sum_{i=1}^{N} y_i x_{ki} , \quad k = 1,2,..,M .$$

Формули для обчислення коефіцієнтів взаємодії виглядають аналогічно:

$$b_{km} = \frac{1}{N} \sum_{i=1}^{N} y_i x_{ki} x_{mi} , \quad k \neq m, \quad k, m = 1,2,..,M .$$

Таким чином, прості формули обчислення коефіцієнтів регресії – є слідство застосування методу МНК до ортогональної, симетричної і нормованої матриці планування

Кореляційний аналіз

Під кореляцією розуміється будь–який зв'язок між двома або декількома досліджуваними явищами. Кореляція може бути детерміністичною або випадковою (імовірнісною). Перший тип зв'язку визначається чіткими закономірностями, які описуються фізико-хімічними формулами.

При кореляційному аналізі перевіряється лише самий факт зв'язку, тобто статистична гіпотеза про відсутність або наявність зв'язку. Сама природа величин, між якими такий випадковий зв'язок передбачається, дозволяє судити про нього як про імовірнісний. Результат кореляційного аналізу також має статистичне спрямування, тому що висновок про наявність або відсутність зв'язку приймається за деякою наперед заданою довірчою імовірністю.

Прикладом задачі кореляційного аналізу може служити дослідження впливу температурного режиму на вихід якого-небудь хімічного продукту в складному технологічному процесі. При цьому зі збільшенням температури можливо не тільки підвищення швидкості досліджуваної реакції, але і протікання побічних реакцій, а також і зворотна реакція розкладання продукту. Тому зв'язок між температурою і виходом можна охарактеризувати як випадковий.

Лінійна кореляція

Зазвичай при кореляційному аналізі досліджуються тільки лінійні зв'язки між величинами, а статистичні критерії свідчать про наявність або відсутність передбачуваного лінійного зв'язку. Тому негативна відповідь при перевірці гіпотези про кореляцію може означати не тільки відсутність зв'язку, але і можливу наявність нелінійної залежності між досліджуваними величинами.

Для кількісної оцінки лінійної кореляції користуються вибірковим коефіцієнтом парної кореляції r_{xy} – безрозмірною величиною до значень середніх квадратичних відхилень досліджуваних величин:

$$r_{xy} = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sqrt{\left(n\sum_{i=1}^{n} x_i^2 - \left[n\sum_{i=1}^{n} x_i\right]^2\right)\left(n\sum_{i=1}^{n} y_i^2 - \left[n\sum_{i=1}^{n} y_i\right]^2\right)}}. \qquad (2.36)$$

Коефіцієнт кореляції за абсолютною величиною не перевершує одиниці ($|r_{xy}| \le 1$) і може набувати таких значень:

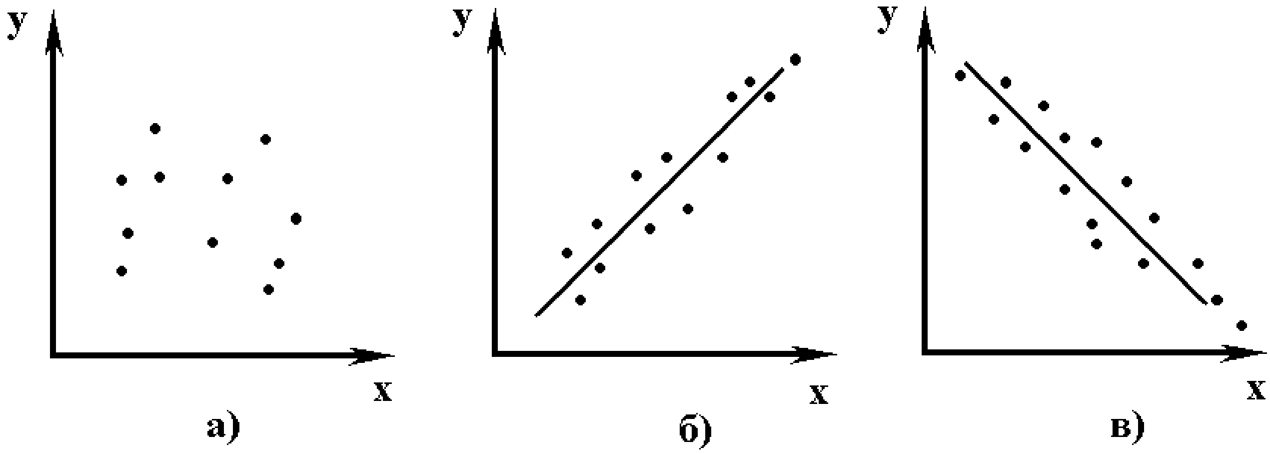
1) $r_{xy} = 0$ – цей випадок відповідає відсутності зв'язку між x і y (рис. 2.1, а);
2) $r_{xy} = +1$ – між x і y існує чіткий позитивний лінійний зв'язок (рис. 2.1, б);
3) $r_{xy} = -1$ – між x і y існує чіткий негативний зв'язок (рис. 2.1, в);
4) $1 < r_{xy} < +1$ – це випадок, що найчастіше зустрічається, і тут про

кореляцію судять уже лише з точки зору більшої або меншої імовірності.

Залежність коефіцієнта кореляції $|r_{xy}|$ слугує тільки для оцінки щільності лінійного зв'язку між величинами x і y: чим ближче абсолютна величина коефіцієнта до 1, тим зв'язок сильніше; чим ближче $|r_{xy}|$ до нуля, тим зв'язок менше. Якщо випадкові величини x і y пов'язані точною лінійною функціональною залежністю

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

то $r_{xy} = \pm 1$. Знак «+» або «–» потрібно використати залежно від знака коефіцієнта a ($a > 0$ або $a < 0$).



Рисуное 2.1 – Кореляційна залежність між випадковими величинами x і y

Залежність коефіцієнта кореляції перевіряється шляхом порівняння абсолютної величини емпіричного коефіцієнта кореляції, помноженої на $\sqrt{n-1}$, із його критичними значеннями при заданому ступені надійності (рівня довірчості) P. Критичні значення $H = |r_{xy}|\sqrt{n-1}$ при числі вимірів ів n до 10 для різноманітних значень надійності становлять:

при $P = 0{,}9$ $H = 1{,}65$;
при $P = 0{,}95$ $H = 1{,}90$;
при $P = 0{,}99$ $H = 2{,}29$.

Якщо для емпіричного коефіцієнта кореляції r_{xy} $H = |r_{xy}|\sqrt{n-1}$ виявиться більше критичного значення H, то з надійністю P варто відкинути гіпотезу про некорельованість аналізованих величин.

При інженерних розрахунках рівень довірчості $P = 0{,}95$ достатній.

Регресійний аналіз

Дослідження й оптимізація складних, неорганізованих систем можливі лише за допомогою статистичних, імовірнісних методів. Вихідною точкою для таких досліджень є аналог фізичної формули – *математичної моделі* системи, що носить назву *моделі експерименту* або *рівняння регресії*. Проте не завжди експериментальний матеріал дає змогу знайти зручний і точний вид моделі. У більш загальному випадку математична модель створюється на підставі статистичного методу – регресійного аналізу.

Регресійний аналіз – це дослідження регресійного рівняння – лінійної моделі плану. Дослідження проводиться в рамках методу найменших квадратів (МНК). Регресійний аналіз базується на трьох твердженнях.

Перше твердження.

Параметр оптимізації *y* є *випадкова величина* з нормальним законом розподілу. Для перевірки нормальності необхідно провести багато паралельних дослідів. Далі користуються одним із критеріїв згоди.

Найбільшого поширення в практиці набув критерій Пірсона. Ідея цього методу полягає в контролі відхилень гістограми експериментальних даних від гістограми з таким же кількісним інтервалів, побудованої на підставі розподілу, збіг з яким визначається, в цьому випадку з нормальним розподілом.

При великій кількості паралельних дослідів розмах відгуку наближається до інтервалу трьох сигм.

Хай, наприклад, у результаті *n* більше 50-ти паралельних дослідів середнє значення відгуку, мінімальне й максимальні значення $\bar{y}$, $y_{\min}$, $y_{\max}$, обчислили оцінку дисперсії s^2. Вибрали *m* інтервалів, обчислили величину інтервалу Δ, кванта, і підрахували частотності попадання k_i, $i = 1, 2,\ldots, m$ значень відгуку в *i*-тий інтервал, тим самим підготували дані для побудови гістограми.

Для визначення теоретичних частотностей, наприклад, в MathCad

1) задаємо ранжировану змінну $i := 0 \ldots m - 1$;
2) визначаємо значення середин інтервалів $z_i = z_{min} + (1/2 + i)\Delta$;
3) обчислюємо вектор теоретичних частотностей за формулою $K_i = d_{norm}(z_i, \bar{y}, s^2)\cdot n$.

Використання критерію Пірсона полягає в підрахунку величини x_i^2

$$\chi_i^{\,2} = \sum_{i=1}^{m} \frac{(k_i - K_i)^2}{K_i} = \sum_{i=1}^{m} \frac{(k_i - nP_i)^2}{nP_i}, \tag{2.37}$$

де k_i, K_i – експериментальні і теоретичні значення частот в *i-тому* інтервалі розбиття, *m* – кількість інтервалів розбиття, P_i – значення ймовірності в тому ж інтервалі розбиття, $n = \sum_{i=1}^{m} k_i$.

Кількість ступенів свободи змінної x_i^2 в критерії Персона обчислюється в цьому випадку за формулою $\upsilon = m - 3$, а не за формулою $n - 1$, як в критерії Бартлета. Два ступеня свободи задіяні на обчислення двох сум в чисельнику і знаменнику відношення критерію Пірсона.

Отже, задавшись рівнем значущості, зазвичай 0,05, звертаємося до функції квантилів x_i^2 розподілу, що має в MathCad ім’я dchisq, з параметрами dshisq(1-0.05, m-3). Якщо обчислене за дослідними даними значення критерію Пірсона χ^2 менше ніж обчислене за функцією квантилів, то гіпотеза нормальності приймається, в іншому випадку відкидається.

При порушенні нормальності ми позбавляємося можливості встановлення ймовірності, з якою справедливі ті або інші вислови щодо результатів експерименту.

Друге твердження.

Дисперсія відгуку не залежить від його абсолютної величини. Перевірка здійснюється за допомогою перевірки дисперсії на однорідність із застосуванням критерія Фішера або Бартлета. Порушувати це *твердження* неприпустимо. При порушенні однорідності дисперсії вдаються до функціонального перетворення відгуку. Перетворення шукається методом проб і помилок. Зазвичай починають з логарифма.

Сенс нелінійного перетворення в тому, що він підкреслює або виділяє один інтервал зміни змінної перед іншим. Наприклад, квадратичне перетворення підкреслює великі величини, а малі (менші 1) ще більше зменшує. Логарифм, навпаки, «притримує» швидкість росту монотонно зростаючих величин, і тому логарифми дисперсій можуть вже виявитися однорідними, тоді як самі дисперсії неоднорідні.

Третє твердження. Значення факторів є «невипадковими» величинами. Це несподіване твердження означає, що встановлення кожного фактора на заданий рівень і його підтримку на цьому рівні набагато точніше, ніж помилка відтворюваності.

Четверте твердження. Фактори не корельовані. Для матриці планування ця властивість виконується автоматично через властивість ортогональності.

Перевірка значущості коефіцієнтів

Перевірка здійснюється або побудовою довірчого інтервалу, або за допомогою критерію Ст'юдента:

А) *побудова довірчого інтервалу.*

Приймемо без доказу такий факт, що при використанні повного експерименту фактору або регулярних дробових реплік довірчі інтервали для всіх коефіцієнтів (у тому числі і ефектів взаємодії) рівні один одному.

Дисперсія будь-якого коефіцієнта регресії обчислюється за формулою

$$s^2_{\{b\}} = \frac{s^2_{\{y\}}}{f}, \quad f = \sum_{i=1}^{N}(n_i - 1), \tag{2.38}$$

де $s^2_{\{y\}}$ – дисперсії відтворюваності, n_i – кількість паралельних дослідів в *i-тому* експерименті, N – кількість рядків матриці планування: f_i – кількість ступенів свободи при обчисленні дисперсії відтворюваності в *i-тому* рядку плану.

Хай перший постулат регресійного аналізу про те, що відгуки суть випадкові величини з нормальним розподілом, виконаний в результаті перевірки за критерієм Пірсона. Тоді коефіцієнти регресії – також випадкові величини з нормальним розподілом. Розглянемо величину

$$\tau = \frac{1}{f}\frac{\sum_{i}^{N}\sum_{j}^{ni}(b_{i,j} - \beta_{i,j})}{S_{\{b\}}} = \frac{\Delta b}{S_{\{b\}}} \tag{2.39}$$

Величина τ має розподіл Ст'юдента.

Нагадаємо, що середнє арифметичне центрованих відносно середнього і нормованих на дисперсію випадкових величин з нормальної генеральної сукупності має розподіл Ст'юдента.

У цьому виразі відома тільки одна величина – СКО коефіцієнта регресії. Другу величину – значення випадкової величини f визначимо, задавшись рівнем значущості α. Для цього ми звертаємося до функції квантилів зворотного інтегрального розподілу Ст'юдента з аргументами $P = 1 - \alpha/2$, оскільки розподіл Ст'юдента симетричний відносно нуля, і f – кількість ступенів свободи при обчисленні випадкової величини:

$$t = qt(1\text{-}\alpha/2, f).$$

Квантиль t є величина напівінтервалу. Імовірність знаходження в інтервалі $[-t, t]$ для випадкової величини τ дорівнює $P = 1 - \alpha$, тому довірчий інтервал Δb_j визначається, як

$$\Delta b = \pm t \cdot s_{\{b\}},$$

де P – довірча імовірність.

Формулу для довірчого інтервалу можна записати в такій еквівалентній формі:

$$\Delta b = \pm \frac{t \cdot s_{\{y\}}}{\sqrt{f}}$$

Коефіцієнт значущий, якщо його абсолютна величина більша за довірчий інтервал.

Б) *використання критерію Ст'юдента.*

Коефіцієнт Ст'юдента можна використовувати і безпосередньо. Розділимо останню нерівність на стандарт коефіцієнта регресії:

$$t = \frac{|\Delta b|}{s_{\{b\}}} < \frac{|b_j|}{s_{\{b\}}} \quad \Rightarrow \quad t < \frac{|b_j|}{s_{\{b\}}}.$$

У лівій частині нерівності міститься табличне значення коефіцієнта Ст'юдента у правій – експериментально обчислювані значення.

Якщо критерій Ст'юдента менше відношення коефіцієнт-СКО – похибки коефіцієнтів, то гіпотеза про те, що коефіцієнт значущій приймається з імовірністю P.

Приклад. Нехай у результаті повного експерименту фактора 2^3 помилка з двома паралельними дослідами в кожному рядку плану похибка відтворюваності $s_{\{y\}} = 1$. Обчислені коефіцієнти: $b_0 = 10$, $b_1 = 2$, $b_2 = 1$, $b_3 = -0.5$. Побудувати довірчий інтервал і перевірити значущість коефіцієнтів.

Дисперсія коефіцієнтів регресії $s^2_{\{bj\}} = 1/8, \quad s_{\{bj\}} = \sqrt{1/8} = 0{,}354$. Табличне значення критерію Ст'юдента при рівні значущості 0,05 і кількість ступенів свободи $f = 8$ $qt(0.975, 8) = 2{,}3$.

Таким чином, довірчий інтервал коефіцієнтів регресії $\Delta b_j = \pm 2{,}3 \cdot 0{,}354 = \pm 0{,}814$.

Доведено, що всі коефіцієнти регресії окрім b_3 значущі, оскільки їх абсолютна величина більша за довірчий інтервал, а для b_3 вона менша.

Приклад 2.17. Потрібно провести регресійний аналіз залежності часу досягнення граничної втрати тиску у фільтрі ($t_н$) від Залежність у зерен завантаження (d). Вихідні результати і розрахунки наведені в таблиці 2.7.

Розв'язання. Відомо, що ступенів ою лінійної залежності двох змінних є кореляційне співвідношення і коефіцієнт кореляції. Парний коефіцієнт кореляції можна оцінити за допомогою виразів:

– для малих вибірок

$$r = \sum\left[(x_i - \bar{x}) \cdot (\bar{y}_i - \bar{\bar{y}})\right] / \sqrt{\sum(x_i - \bar{x})^2 \cdot \sum(\bar{y}_i - \bar{\bar{y}})^2}\,; \qquad (2.40)$$

– для великих вибірок

$$r = \frac{\sum xy - (\sum x) \cdot (\sum y)/n}{\sqrt{\left[\sum x^2 - (\sum x)^2/n\right] \cdot \left[\sum y^2 - (\sum y)^2/n\right]}}. \qquad (2.41)$$

Таблиця 2.7 – Вихідні дані і проміжні розрахунки

(d, см) (x_i)	τ_1 (y_1)	τ_2 (y_2)	$\bar{\tau}$ ($\bar{y}_i$)	$x_i - \bar{x}$	$\bar{y}_i - \bar{\bar{y}}$	$(x_i - \bar{x})^2 \cdot 10^{-4}$	$(\bar{y}_i - \bar{\bar{y}})^2$	$(x_i - \bar{x}) \cdot (\bar{y}_i - \bar{\bar{y}})$
0,120	16,7	16,6	16,65	-0,015	-3,84	2,25	14,74	0,0576
0,125	17,8	18,1	17,95	-0,010	-2,54	1,00	6,45	0,0254
0,130	19,3	19,4	19,35	-0,005	-1,14	0,25	1,30	0,0057
0,135	20,4	20,5	20,45	0,000	-0,04	0,00	0,00	0,0000
0,140	21,6	21,7	21,65	0,005	1,16	0,25	1,34	0,0058
0,145	23,0	23,1	23,05	0,010	2,56	1,00	6,55	0,0256
0,150	24,0	24,3	24,15	0,015	3,86	2,25	14,90	0,0579
$\bar{x}$ = 0,135			$\bar{\bar{y}}$ = 20,49			Σ = 7,00	Σ = 45,28	Σ = 0,178

Вважають [3, с. 121], що оптимальну оцінку коефіцієнта кореляції p отримують за формулою

$$r^* = r \cdot \left[1 + \frac{1 - r^2}{2 \cdot (n-3)}\right]. \qquad (2.42)$$

За формулою (2.36), за таблицею 2.7маємо:

$$r = 0{,}178 / \sqrt{0{,}0007 \cdot 45{,}28} = 0{,}9998$$

або

$$r^* = 0{,}9998 \cdot (1 + (1 - 0{,}9998^2)/\left[2 \cdot (7-3)\right]) = 0{,}99985.$$

Коефіцієнт детермінації $B = r^2$ = 0,9996, а $(r^*)^2$ = 0,9997.

Таким чином, залежність між часом досягнення граничних втрат тиску у фільтрі і Залежність от зерен завантаження добре корелюється.

Після знаходження довірчих меж для m_z (при довірчій ймовірності p = 0,95)

$$z_1 = z - 1{,}96/\sqrt{n-3} \quad \text{і} \quad z_2 = z + 1{,}96/\sqrt{n-3} \qquad (2.43)$$

можна знайти довірчі межі для генерального коефіцієнта кореляції, підставляючи z_1 і z_2 у формулу

$$r = thz = (e^{2z} - 1)/(e^{2z} + 1) \qquad (2.44)$$

Тоді для цього прикладу маємо

- стандарт розподілу

$$\sigma_r \approx (1 - r^2)/\sqrt{n} \approx (1 - 0{,}9998^2)/\sqrt{7} = 0{,}00015; \qquad (2.45)$$

- довірчі межі p

$$p = r \pm \frac{1{,}96 \cdot (1 - r^2)}{\sqrt{n}} = 0{,}9998 \pm \frac{1{,}96 \cdot (1 - 0{,}9998^2)}{\sqrt{7}} = 0{,}9998 \pm 0{,}0003; \qquad (2.46)$$

$$z = \frac{1}{2}\ln\frac{1 + r}{1 - r} = \frac{1}{2}\ln\frac{1 + 0{,}9998}{1 - 0{,}9998} \approx 4{,}6; \qquad (2.47)$$

$$\sigma_z = 1/\sqrt{n - 3} = 1/\sqrt{7 - 3} = 0{,}5. \qquad (2.48)$$

Отже, з довірчою ймовірністю p = 0,95 значення m_z перебуває в межах:

$$m_z = z \pm 1{,}96/\sqrt{n - 3} = 4{,}6 \pm 1{,}96/\sqrt{7 - 3} = 4{,}6 \pm 0{,}98;$$

$$z_1 = 4{,}6 - 0{,}98 = 3{,}62 \quad \text{и} \quad z_2 = 4{,}6 + 0{,}98 = 5{,}58;$$

$$r_{x1} = thz_1 = (e^{2 \cdot 3{,}62} - 1)/(e^{2 \cdot 3{,}62} + 1) = 0{,}99857;$$

$$r_{x2} = thz_2 = (e^{2 \cdot 5{,}58} - 1)/(e^{2 \cdot 5{,}58} + 1) = 0{,}99997,$$

тобто при рівні значущості q = 0,05 величина вибіркового коефіцієнта кореляції міститься в межах 0,998 57 < p < 0,999 97.

Вихідні дані та завдання для індивідуального розв'язання видаються викладачем.

Контрольні питання

1. Як виконується перевірка гіпотези про рівність середніх?
2. Для чого можуть бути використані похибки коефіцієнтів регресії?
3. Для чого використовують кореляційний аналіз?
4. Поясніть як можна здійснювати перевірку значущості коефіцієнтів?
5. Для чого потрібно будувати довірчій інтервал?
6. Для чого використовують критерій Стьюдента?
9. Що таке матриця суміжностей? Укажіть особливості її побудови?

2 ТЕМИ ДЛЯ САМОСТІЙНОГО ВИВЧЕННЯ

Тема 1 Поняття науки. Класифікація і структура науково-дослідної роботи

Методологія науки. Система наукових знань. Етапи наукового дослідження. Формулювання проблеми та завдань дослідження. Формулювання питань для пізнання окремих аспектів проблеми. Формулювання мети й завдань Дослідження. Робоча гіпотеза, її місце в дослідженнях. Елементи структури наукового дослідження. Теорія як результат науково-дослідної роботи.

Питання для самоперевірки

1 Що є основою пізнання?
2. Мета фундаментальних досліджень.
3. Сформулювати мету прикладного дослідження.
4. На що спрямовані прикладні дослідження?
5. Що є основою НДР?
6. Яку роль відіграє постановка проблеми у наукових дослідженнях?
7. Що являє собою робоча гіпотеза?
8. Яка послідовність проведення науково-дослідної роботи?

Тема 2 Інформаційний пошук. Аналіз інформації і формулювання завдань наукового дослідження

Місце науково-технічної інформації у розвитку суспільства. Мета пошуку. Етапи пошуку. Монографія. Вивчення патентів. Послідовність підбору літературних джерел. Науково-технічна патентна інформація. Обробка науково-технічної інформації. Критичний аналіз інформації. Висновки після обробки інформації. Актуальність і новизна теми. Останні досягнення в досліджуваній галузі.

Питання для самоперевірки

1. Охарактеризувати два етапи пошуку.
2. Що можна з'ясувати при вивченні монографій. Для чого вивчаються монографії?
3. План аналізу обробленого матеріалу.
4. Які загальні аспекти мають бути висвітлені у висновках після обробки матеріалів.
5. Як допомагають висновки при формулюванні мети й встановлення конкретного завдання наукового дослідження.

Тема 3 Основи наукової мови

Науковий текст і його основні категорії. Зв’язність, структурність, цілісність, модальність, функціонально-значеннєвий тип. Розташування значеннєвих блоків у тексті. Аналіз логіко-композиційної структури наукового тексту. Стильові типи наукових текстів, класифікація.

Академічні тексти, навчальні тексти, науково-публіцистичні тексти, науково-інформаційні тексти, навчально-наукові тексти.

Термінологія і номенклатура. Основні поняття. Системність термінології. Робота з текстом, формування механізму поверхневого розуміння тексту.

Питання для самоперевірки

1.Основні категорії наукового тексту? Дайте пояснення тематичній і логіко-композиційній структурі наукового тексту, в чому відмінність.

2.Побудова наукової статті, розташування значеннєвих блоків наукових публікацій.

3.Охарактеризуйте п'ять типів текстових структур.

4.У чому полягає основна відмінність академічних текстів від навчальних?

5. Яка специфіка науково-публіцистичних текстів?

6. Чим відрізняються науково-інформаційні тексти?

Тема 4Теоретичні дослідження

Основні завдання і методи теоретичного дослідження. Мета теоретичного дослідження. Системність вивчення об'єкта. Умови й вимоги. Стадія постановки завдання. Аналітична стадія. Використання математичних методів у дослідженнях. Етапи проведення теоретичних досліджень. Основа аналізу теоретико-експериментальних досліджень. Формулювання висновків і пропозицій. Основні вимоги до складання та оформлення звітів з науково-досліджуваної роботи.

Питання для самоперевірки

1. Назвіть завдання теоретичних досліджень.
2. Порівняйте методи розчленування і об'єднання елементів досліджуваної системи при проведенні теоретичних досліджень.
3. Загальна теорія систем, метод об'єднання.
4. Основний принцип загальної теорії систем.
5. Основний принцип вивчення об'єкта.
6. Охарактеризуйте етапи проведення теоретичних досліджень.

Тема 5 Загальні поняття теорії систем і системного аналізу

Основи системного аналізу. Елементи системи. Об'єкти в системі і зв'язок між ними. Сутність і принципи системного підходу. Проблеми узгодження цілей. Проблеми оцінки зв'язків у системі. Моделювання як метод системного аналізу.

Питання для самоперевірки

1. Що таке системний аналіз? Що входить у предметну область системного аналізу?
2. Які основні системні методи й процедури?
3. Етапи формулювання цілей і завдань системи.
4. Розв'язання проблеми проведення експерименту в системі або над системою.
5. Головні принципи теорії систем і системного аналізу.
6. Які основні системні ресурси суспільства?
7. Поясніть поняття «елемент системи».

Тема 6 Основні поняття теорії моделювання

Поняття моделі. Класифікація моделей. Засоби моделювання. Мети і завдання моделювання. Вимоги до моделей. Способи складання математичних моделей. Емпіричний і аналітичний методи моделювання. Машинні моделі. Основні принципи створення таких моделей. Приклади математичних моделей.

Питання для самоперевірки

1. Порівняти фізичне й математичне моделювання.
2. Проаналізувати різні засоби моделювання.
3. Принципи розробки математичної моделі.
4. Можливості машинних моделей.
5. Перевірка адекватності й ідентифікація моделі.

Тема 7 Експериментальні дослідження

Класифікація, типи і завдання експерименту. Формування цілей і завдань експерименту. Розходження експерименту за: способом формування умов; цілями Дослідження; організацією проведення; структурою досліджуваних об'єктів і явищ; за характером зовнішніх впливів на об'єкт дослідження; за характером взаємодії засобів експериментального дослідження з об'єктом дослідження; за типом моделей, досліджуваних в експерименті; контрольованими величинами; кількісними варіативних факторів; за характером досліджуваних об'єктів або явищ тощо. Основа аналізу теоретико-експериментальних досліджень. Формулювання висновків і пропозицій. Основні вимоги до складання та оформлення звітів з науково-дослідної роботи.

Питання для самоперевірки

1. Загальна оцінка експериментальних методів: достоїнства й недоліки.
2. Що є метою проведення експерименту.
3. Порівняти лабораторний і натуральний експеримент.
4. Відмінності пасивного й активного експериментів.
5. Порівняйте однофакторний і багатофакторний експеримент.
6. Розробка методики експерименту.
7.Основний принцип проведення будь-якого експерименту. Що для цього необхідно?
8. Що є основою для аналізу теоретико-експериментальних досліджень?
9. Принципи формулювання висновків і пропозицій.
10. Вимоги до складання наукових звітів.
11. Вимоги до оформлення звітів з науково-дослідної роботи.

Тема 8 Метрологічне забезпечення експериментальних досліджень

Метрологія. Основні проблеми метрології. Методи виміру у, абсолютні й відносні, сукупні й спільні виміри. Засоби вимірів, вимірювальні прилади. Класифікація приладів за точністю вимірів, класи точності. Вимірювальна установка (стенд). Абсолютна і відносна похибки. Чутливість приладу. Перевірка приладів. Вплив психологічних факторів на хід та якість експерименту.

Питання для самоперевірки

1. Порівняти прямі й непрямі методи виміру
2. Відмінність сукупних і спільних вимірів.
3. Дати характеристику вимірювальним засобам.
4. Вимірювальний прилад і вимірювальна установка. У чому основна відмінність?
5. Поняття ефективності.
6. Дати характеристику абсолютної і відносної похибки.
7. Обґрунтувати необхідність проведення перевірки вимірювальних приладів. Як часто необхідно це робити?
8. Принцип порівняння при проведенні перевірки приладів.
9. Чим обумовлені систематичні і випадкові похибки

Тема 9 Обробка результатів експериментальних досліджень

Основи теорії випадкових помилок і методів оцінки випадкових похибок у вимірах. Кореляційний аналіз. Види кореляції. Коефіцієнт кореляції. Тіснота зв'язку. Лінійна кореляція. Нелінійна регресія. Лінійний регресійний аналіз. Метод найменших квадратів. Теорія випадкових помилок. Нормальний закон розподілу. Генеральна і вибіркова сукупності чисел. Середні значення та їх оцінки. Оцінки довірчих границь для істинного значення вимірюваної величини. Порівняння дисперсій та середніх значень.

Питання для самоперевірки

1. На чому базується аналіз випадкових похибок.
2. Основний принцип визначення широкого значення обумовленої величини при нескінченно великої кількості вимірів.
3. Визначення мінімальної кількості вимірів у експериментальних дослідженнях.
4. Оцінки довірчих границь для істинного значення вимірюваної величини.
5. Порівняння дисперсій.
6. Порівняння середніх.
7. Перевірка гіпотези про рівність середніх.
8. Перевірка гіпотези нормальності закону розподілення випадкової величини.
9. Визначення теоретичного закону розподілення.
10. Що таке *кореляція*?
11. Що таке *регресія*?
12. Типи кореляції.
13. Суть кореляційного та регресійного аналізу.
14. Лінійна кореляція.
15. Оцінювання коефіцієнту кореляції.
16. Лінійний регресійний аналіз.
17. Оцінювання прямої регресії.
18. Що таке критерій Ст'юдента?
19. Що таке критерій Фішера?
20. Метод найменших квадратів.

3 КОНТРОЛЬНА РОБОТА

Вибір емпіричної залежності

Мета контрольної роботи – засвоєння методів математичної обробки експериментальних досліджень; встановлення емпіричних залежностей і визначення параметрів обраного рівняння з використанням методу найменших квадратів (далі – МНК). Визначення ступеня вірогідності обраної залежності й виконання кореляційного аналізу.

У процесі експериментальних вимірів одержують статистичний ряд вимірів двох величин, поєднуваних функцією

$$y = f(x). \qquad (3.1)$$

Кожному значенню функції y_1, …, y_n відповідає певне значення аргументу x_1, x_2……, x_n.

Для експериментальних даних можна підібрати алгебраїчні вирази, які називаються емпіричними формулами. У математиці емпіричними називають формули, які виходять з досвіду за допомогою спостереження і експерименту. Графічне зображення емпіричної формули складається з двох етапів:

1) вибір загального виду формули з деяким кількісним довільних (вільних) параметрів,

2) визначення її параметрів (параметрів апроксимації) з умов найкращої апроксимації даних.

Для прикладу розглянемо ряд експериментальних вимірів, що наведених у таблиці 3.1.

Таблиця 3.1 – Вихідні дані

Точки на ізотермі	Номери експериментів і значення x і y; $\frac{x}{y}$					
	1	**2**	**3**	**4**	**5**	**6**
A_1	$\frac{39{,}9}{4{,}9}$	$\frac{39{,}6}{4{,}7}$	$\frac{40{,}2}{4{,}7}$	$\frac{39{,}3}{4{,}75}$	$\frac{39{,}12}{4{,}8}$	$\frac{40{,}3}{4{,}75}$
A_2	$\frac{33{,}7}{4{,}6}$	$\frac{33{,}9}{4{,}5}$	$\frac{33{,}5}{4{,}5}$	$\frac{33{,}9}{4{,}4}$	$\frac{34{,}6}{4{,}5}$	$\frac{33{,}7}{4{,}5}$
A_3	$\frac{27{,}7}{4{,}2}$	$\frac{27{,}66}{3{,}95}$	$\frac{27{,}36}{3{,}97}$	$\frac{27{,}48}{4{,}0}$	$\frac{27{,}7}{4{,}0}$	$\frac{27{,}9}{3{,}9}$
A_4	$\frac{23{,}7}{3{,}8}$	$\frac{24{,}18}{3{,}6}$	$\frac{23{,}9}{3{,}7}$	$\frac{24{,}36}{3{,}76}$	$\frac{24{,}72}{3{,}8}$	$\frac{24{,}6}{3{,}86}$
A_5	$\frac{18{,}9}{3{,}3}$	$\frac{19{,}11}{3{,}3}$	$\frac{19{,}14}{3{,}4}$	$\frac{18{,}72}{3{,}35}$	$\frac{19{,}3}{3{,}37}$	$\frac{18{,}96}{3{,}24}$
A_6	$\frac{5{,}2}{1{,}3}$	$\frac{5{,}0}{1{,}3}$	$\frac{5{,}1}{1{,}4}$	$\frac{4{,}7}{1{,}37}$	$\frac{5{,}3}{1{,}42}$	$\frac{5{,}0}{1{,}36}$

Точки таблиці 3.1 отримані внаслідок шести експериментів (номери 1–6 по горизонталі). Досліджували адсорбційну здатність вугілля в рівноважних умовах. Визначали концентрації у воді й адсорбенті ***c*** і ***a*** (чисельник / знаменник кожного осередку таблиці 3.1 в шести пробах для кожного експерименту (точки $A_{1}...A_{6}$).

Процес сорбції зображується у вигляді залежності між поглиненою кількістю речовини в сорбенті і відповідною кількістю у розчині, та й описується функціональною залежністю $a = f(C)$. Для більш детальнішого опису необхідно:

– підібрати конкретний характер залежності;

– встановити чисельні значення параметрів (коефіцієнтів) для обраної залежності.

Для вибору конкретного різновиду аналізованої залежності необхідно скористатися найбільш частими формами опису ізотерм. Рекомендується досліджувати такі залежності:

$$y = ax + b;\; y = \frac{a}{x} + b;\; y = (ax+b)^{\frac{1}{2}};\; y = (ax+b)^{-1};$$

$$y = be^{ax};\; y = ba^{x};\; y = bx^{a};\; y = a\ln x + b;\; y = x(ax+b)^{-1}.$$

Критерієм визначення емпіричної залежності за кількісними даними, що найкраще описує досліджувану залежність, є максимальна величина емпіричного коефіцієнта кореляції r.

Обчислення коефіцієнта кореляції і параметрів a і b виконуємо за допомогою методу найменших квадратів [12, 13]. Відповідно до цього методу розглянуті залежності зводимо до лінійного вигляду:

$$Y = AX + B. \tag{3.2}$$

Наприклад, візьмемо функцію виду:

$$y = \frac{ax}{1 + bx}. \tag{3.3}$$

Для наведеної залежності характерна наявність двох параметрів (коефіцієнти a, b). Чисельні значення коефіцієнтів для розглянутої моделі виду (3.3) знайдемо методом найменших квадратів (МНК) [12].

Спочатку вихідну форму (3.3) перетворимо в лінійний вигляд. Для цього проведемо певні перетворення:

$$X = \frac{1}{x}; \qquad Y = \frac{1}{y} \tag{3.4}$$

Далі перепишемо (3.3) з урахуванням (3.4):

$$Y = \frac{1}{y} = \frac{1+bx}{ax} = \frac{1}{a} \times \frac{1}{x} + \frac{b}{a} = \frac{1}{a} \times X + \frac{b}{a} = AX + B, \tag{3.5}$$

$$A = \frac{1}{a}, \qquad B = \frac{b}{a} \tag{3.6}$$

Таким чином, вихідна функція перетворена в лінійний вигляд. Далі введемо ряд додаткових позначень:

$$X_i = \frac{1}{x_i}; \quad Y_i = \frac{1}{y_i}, \quad i = 1,...,n \tag{3.7}$$

Одержимо значення відповідних математичних очікувань (середніх значень):

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i , \qquad \overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i . \tag{3.8}$$

Після цього обчислимо ряд проміжних величин, які необхідні для одержання остаточних результатів:

$$\overline{XY} = \frac{1}{n}\sum_{i=1}^{n} x_i y_i ; \tag{3.9}$$

$$\overline{X}^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 ; \tag{3.10}$$

$$\Delta = \overline{X}^2 - (\overline{X})^2 ; \tag{3.11}$$

$$\Delta_A = \overline{XY} - \overline{X} \times \overline{Y} ; \tag{3.12}$$

$$\Delta_B = \overline{X}^2 \times \overline{Y} - \overline{XY} \times \overline{X} . \tag{3.13}$$

Тепер коефіцієнти лінійної форми A, B визначається в такий спосіб:

$$A = \frac{\Delta_A}{\Delta}, \qquad B = \frac{\Delta_B}{\Delta}. \tag{3.14}$$

Застосування МНК, як, втім, і будь-якого іншого чисельного методу моделювання, припускає оцінку адекватності отриманої моделі. Під адекватністю тут розуміється ступенів а наближення моделі й модельних значень до відповідних фактичних моделей і значень. МНК припускає для цього обчислення і оцінювання коефіцієнта кореляції r як ступенів и згаданого наближення. У розглянутому випадку значення r обчислюємо згідно з [13]:

$$r = \frac{\overline{XY} - \overline{X} \times \overline{Y}}{\sqrt{\overline{X}^2 - (\overline{X})^2} \times \sqrt{\overline{Y}^2 - (\overline{Y})^2}} ; \tag{3.15}$$

$$\overline{Y}^2 = \frac{1}{n}\sum_{i=1}^{n} y_i^2 . \tag{3.16}$$

Оцінюючи значення r для конкретної моделі, можна зробити певні висновки щодо її адекватності. Значення r, що наближаються до нуля, будуть свідчити про фактичну незалежність (відсутності статистичного зв'язку) між досліджуваними величинами (у цьому випадку C й a). Навпаки, значення r, близькі до одиниці, фактично будуть підтверджувати наявність істотного статистичного зв’язку між досліджуваними величинами: що більше значення r, то цей зв’язок тісніший, а отже, адекватніша конкретна модель.

Статистичну значущість отриманого коефіцієнта кореляції доцільно перевірити за допомогою одного з розповсюджених критеріїв. Оберемо для цієї мети критерій Стьюдента [12]. Для цього обчислюємо параметр

$$\hat{t} = \frac{r\sqrt{\nu}}{\sqrt{1 - r^2}}, \quad \nu = n - 2 . \tag{3.17}$$

Умовою позитивного результату даної перевірки є виконання умови:

$$\hat{t} \le t_{(v,p)} \qquad , \tag{3.18}$$

де $v = n - 2$, *n*- кількість пар експериментальних точок;

p – довірча імовірність, з якої приймаємо гіпотезу;

t – критерій при заданих v і p (визначаємо за табл. [12]).

Зазначимо, що для цього дослідження рівень $p = 0{,}95$ є задовільним.

У результаті обчислень для наведеної функції (3.3) у таблиці 3.2 подані отримані значення коефіцієнта кореляції і коефіцієнти *a* і *b*.

Таблиця 3.2 –Результат розрахунку МНК коефіцієнта кореляції *r* й коефіцієнтів *a* і *b* для досліджуваної функції

За наведеними в таблиці 3.2 значеннями коефіцієнта кореляції *r* і

Вид функції	Значення коефіцієнта кореляції, *r*	Коефіцієнти	
		a	*b*
$y = \dfrac{ax}{1+bx}$	0,997017114	3,021	0,138

коефіцієнтів *a* і *b* , можна зробити висновок про те, що для математичного опису групи точок залежність типу (3.3) є найоптимальнішою, а її коефіцієнти мають значення констант, що залежать від структури сорбенту і фізико-хімічних властивостей поглинаючої речовини.

Далі перевіримо значущість коефіцієнта кореляції для моделі обраного виду. Для цього за довідковою літературою [12,13] знаходимо значення t, що відповідає умовам $v = n - 2 = 34$, $p = 0{,}95$. Це значення дорівнює 2,03. Обчислюємо значення критерію:

$$\hat{t} = \frac{0{,}996\sqrt{36-2}}{\sqrt{1-0{,}996^2}} = 65\,. \tag{3.19}$$

Оскільки отримане значення більше, ніж величина 2,03, можна зробити висновок про те, що з імовірністю 0,95 отриманий коефіцієнт кореляції є значущим. Це твердження додатково підтверджує адекватність отриманої моделі.

Уточнимо параметри в (3.3). Тут y відповідає значенню концентрації в сорбенті – a, а x – концентрація в рідкій фазі С, тоді

$$a = \frac{a_\infty k \cdot C}{1 + k \cdot C} = \frac{a \cdot C}{1 + b \cdot C}\,, \qquad 9 \tag{3.20}$$

де a_∞ – гранична сорбційна ємкість, мг/г;

k – константа, що залежить від структури сорбенту й фізико-хімічних властивостей води;

C – концентрація в рідкій фазі, мг/дм^3.

У результаті обчислень отримуємо значення коефіцієнтів:

$$a_\infty = 5.233\,, \quad k = 0.0458$$

Для перевірки правильності обраної функціональної залежності (3.3) (розрахованих коефіцієнтів) було обрано на кривій точку з координатами А (35; 4,5). Підставляючи значення с у формулу (3.19), одержали a:

$$\frac{5.233 \times 0.0458 \times 35}{1 + 0.0458 \times 35} = 4.45,$$

що відповідає дійсності.

На рисунку 3.1 представлені експериментальні точки й обрана емпірична залежність.

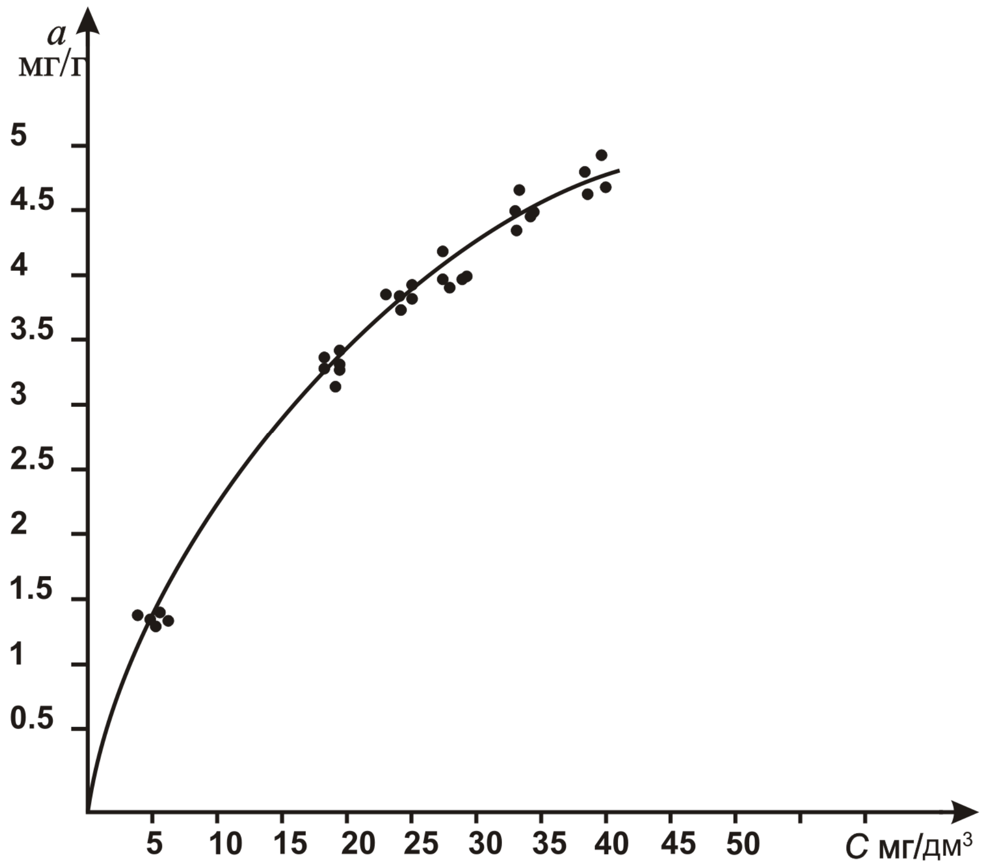


Рисунок 3.1– Графічна побудова обраної залежності $y = \dfrac{ax}{1 + bx}$.

Контрольна робота оформляється у вигляді пояснювальної записки, що повинна включати наступне: титульний аркуш, завдання, зміст, вступ, основну частину, висновок, список використаних джерел; додатки (при необхідності).

У тексті основної частини, що розбивається на розділи і підрозділи, приводяться: стисла характеристика використовуваних методів, алгоритми розрахунків, блок-схеми, тексти розроблених або опис використовуваних прикладних програм, вихідні дані, результати розрахунків і аналіз отриманих даних.

Вступ повинен містити обґрунтування проблеми, якій присвячено контрольну роботу. У висновку необхідно підкреслити все, що було зроблено при виконанні контрольної роботи, та оцінити ступінь виконання завдання.

Орієнтовний обсяг пояснювальної записки 5–7 сторінок друкованого тексту з полуторним міжрядковим інтервалом.

Завдання до контрольної роботи подане в додатку А.

СПИСОК РЕКОМЕНДОВАНИХ ДЖЕРЕЛ

1. Білуха М. Т. Методологія наукових досліджень : підручник / М. Т. Білуха – Київ : АБУ, 2002. – 480 с.

2. Рузавин Г. И. Методология научного исследования: учеб. пособие для вузов / Г. И. Рузавин – М. : Высшая школа, 1999. – 317 с.

3. Наринян А. Р. Основы научных исследований. учеб. пособие для вузов/ А. Р. Наринян. – Киев : Вища школа, 2002. – 112 с.

4. Фрумкин Р. А. Основы научных исследований : учеб. пособие для вузов / Р. А. Фрумкин. – Алчевск : АБУ, 2001. – 201 с.

5. Сиденко В. М. Основы научных исследований / В. М. Сиденко – Харьков: Вища школа, 2002. – 200 с.

6. Філіпенко А. С. Основи наукових досліджень. Конспект лекцій : посібник

7. Білущак Г. І. Теорія ймовірностей і математична статистика : практикум / Г. І. Білущак, Я. М. Чабанюк.– Львів, 2001. – 418 с.

8. Бірта Г. О. Методологія і організація наукових довартожень. [текст] : навч. посібник / Г. О. Бірта, Ю.Г. Бургу. – Київ : Центр учбової літератури, 2014. – 142 с.

9. Айвазян С. А. Прикладная статистика: исследование зависимостей: / [С. А. Айвазян и др.], справ. изд. – М. : Финансы и статистика, 1985. – 487с.

10. Румшиский Л. З. Математическая обработка результатов эксперимента. / Л. З. Румшиский. – М. : Наука, 1971. – 192 с.

11. Основи метрології та вимірювальної техніки : навчальний посібник / [В. О. Поджаренко., П. І. Кулаков., О. Г Ігнатенко, О. П. Войтович] – Вінниця: ВНТУ, 2006. – 151 с.

12. Статистична обробка експериментальних даних: навчальний посібник / О. П. Мельниченко, І. Л. Якименко, Р. Л. Шевченко − Біла Церква, 2006.− 34 с.

13. Супрун А. Н. Вычислительная математика для инженеров-экологов: методическое пособие / А. Н. Супрун, В. В. Найденко. – М.: Изд-во АСВ, 1996. – 391с.

ДОДАТОК А

Завдання до контрольної роботи

Таблиця А.1 – Вихідні дані для самостійного виконання

Номер варіанта		Експериментальні точки									
		1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
1	*x*	5	6	7	8	9	10	11	12	13	14
	y	10,2	6,7	5,3	4,8	4,15	3,6	2,9	2,7	1,8	0,6
2	*x*	2	3	4	5	6	7	8	9	10	11
	y	9,3	8,5	6,4	5,1	3,3	2,7	2,05	1,4	0,8	0,1
3	*x*	3	6	9	12	15	18	21	24	27	30
	y	57,6	41,9	31,0	22,7	16,6	12,2	8,9	6,8	4,3	2,0
4	*x*	2	3	4	5	6	7	8	9	10	11
	y	9,2	8,1	6,8	5,3	3,9	2,7	2,1	1,3	0,5	0,1
5	*x*	2	5	8	11	14	17	20	23	26	29
	y	57,9	42,5	31,3	22,9	16,7	12,4	8,9	6,9	4,5	1,8
6	*x*	2	6	10	14	18	22	26	30	34	38
	y	12,3	8,4	7,3	6,5	6,0	5,2	4,7	3,0	1,5	0,5
7	*x*	3	6	9	12	15	18	21	24	27	30
	y	57,6	41,9	31,0	22,7	16,6	12,2	8,9	6,8	4,3	2,0
8	*x*	3	6	9	12	15	18	21	24	27	30
	y	66,3	40,5	31,4	22,5	17,7	13,4	11,6	8,3	6,1	2,7
9	*x*	2	3	4	5	6	7	8	9	10	11
	y	9,8	7,5	6,0	4,9	3,1	2,2	1,5	0,7	0,1	0
10	*x*	3	6	9	12	15	18	21	24	27	30
	y	66,6	55,2	43,3	34,6	23,8	15,7	12,2	8,4	6,1	1,5
11	*x*	1	2	3	4	5	6	7	8	9	10
	y	11,2	9,8	8,3	7,1	6,4	5,2	4,1	3,3	2,1	0,5

Продовження таблиці А.1

№ варіанта		Експериментальні точки									
		1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
12	*x*	2	4	6	7	8	9	10	11	12	13
	y	10,3	6,8	5,4	4,7	4,0	3,4	2,6	2,0	1,3	0,1
13	*x*	1	3	5	7	9	11	13	15	17	19
	y	9,1	7,5	6,4	5,6	4,5	3,3	2,6	1,2	0,5	0,0
14	*x*	3	6	9	12	15	18	21	24	27	30
	y	58,6	43,9	32,0	23,7	17,6	13,2	9,9	7,8	5,3	3,0
15	*x*	2	3	4	5	6	7	8	9	10	11
	y	9,2	8,1	6,8	5,3	3,9	2,7	2,1	1,3	0,5	0,1
16	*x*	2	5	8	11	14	17	20	23	26	29
	y	57,9	42,5	31,3	22,9	16,7	12,4	8,9	6,9	4,5	1,8
17	*x*	2	6	10	14	18	22	26	30	34	38
	y	12,3	8,4	7,3	6,5	6,0	5,2	4,7	3,0	1,5	0,5
18	*x*	3	6	9	12	15	18	21	24	27	30
	y	58,1	41,7	31,0	23,7	17,2	14,1	9,4	6,8	4,1	0
19	*x*	3	6	9	12	15	18	21	24	27	30
	y	60,3	42,5	33,4	23,5	18,7	14,4	13,6	11,3	8,1	4,7
20	*x*	2	3	4	5	6	7	8	9	10	11
	y	11,8	9,5	7,0	5,4	4,1	2,9	1,7	1,0	0,5	0
21	*x*	3	6	9	12	15	18	21	24	27	30
	y	67,1	56,5	44,6	31,9	24,8	16,1	13,5	8,4	3,2	0,6
22	*x*	1	2	4	6	8	10	12	14	16	18
	y	12,1	10,3	9,7	7,3	6,2	4,2	2,7	1,5	0,8	0,1

Необхідно підібрати емпіричну залежність, визначити її коефіцієнти і побудувати графік.

Виробничо-практичне видання

Методичні рекомендації

до проведення практичних занять, самостійного вивчення
та виконання контрольної роботи
з навчальної дисципліни

«ОСНОВИ НАУКОВИХ ДОСЛІДЖЕНЬ»

(для здобувачів першого бакалаврського рівня вищої освіти всіх форм навчання зі спеціальності 194 – Гідротехнічне будівництво, водна інженерія та водні технології)

Укладачі: **ЧУБ** Ірина Миколаївна,
ТКАЧОВ В’ячеслав Олександрович

Відповідальний за випуск *Г. І. Благодарна*

За авторською редакцією

Комп’ютерне верстання *І. М. Чуб*

План 2019, поз. 128М

Підп. до друку28.12.2021. Формат 60×84/16
Електронне видання Ум. друк. арк. 3,3.

Видавець і виготовлювач:
Харківський національний університет
міського господарства імені О. М. Бекетова,
вул. Маршала Бажанова, 17, Харків, 61002
Електронна адреса: rectorat@kname.edu.ua
Свідоцтво суб’єкта видавничої справи:
ДК № 5328 від 11.04.2017.