

**Горват А.А.,
Молнар О.О., Мінькович В.В.**

**МЕТОДИ ОБРОБКИ
ЕКСПЕРИМЕНТАЛЬНИХ ДАНИХ
З ВИКОРИСТАННЯМ MS EXCEL**

Навчальний посібник

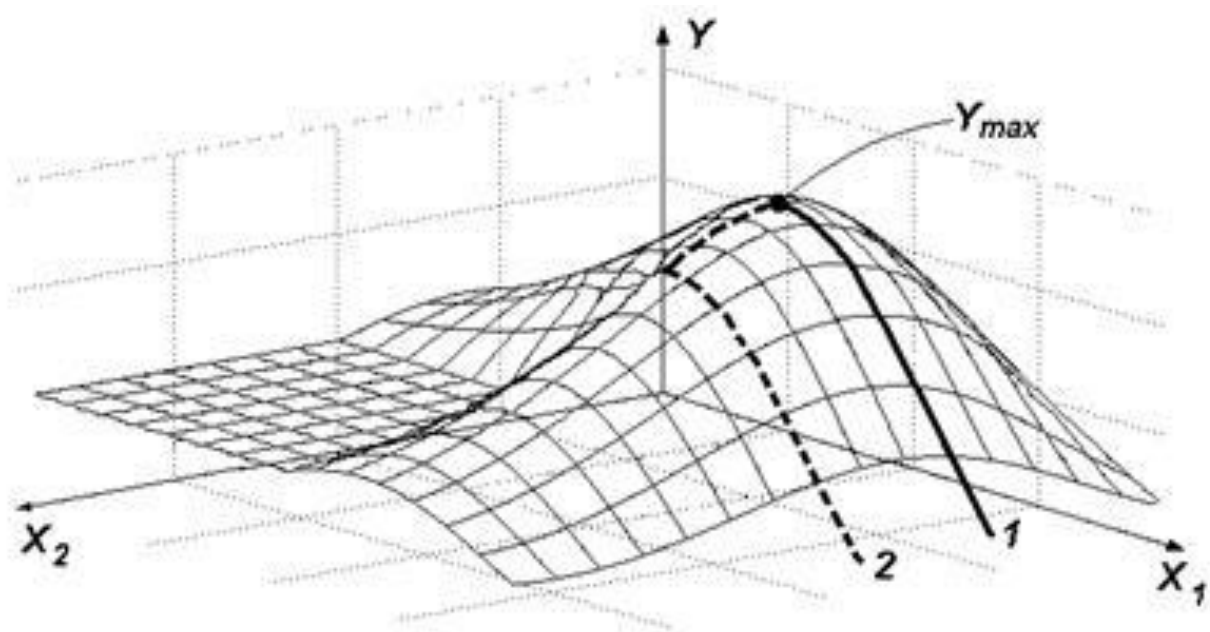


МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ДЕРЖАВНИЙ ВИЩИЙ НАВЧАЛЬНИЙ ЗАКЛАД
«УЖГОРОДСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ»
Кафедра фізики напівпровідників

Горват А.А., Молнар О.О., Мінькович В.В.

МЕТОДИ ОБРОБКИ ЕКСПЕРИМЕНТАЛЬНИХ ДАНИХ З ВИКОРИСТАННЯМ MS EXCEL

Навчальний посібник



УЖГОРОД - 2019 р.

УДК 519.22:502/504

ББК 28.081

Г67

Горват А.А., Молнар О.О., Мінькович В.В.

Методи обробки експериментальних даних з використанням MS Excel: Навчальний посібник. Ужгород: Видавництво УжНУ “Говерла”, 2019. – 160 с.: іл.

Розглянуті питання статистичної обробки результатів експерименту – визначення характеристик емпіричних розподілів, відсів грубих похибок спостережень, перевірка закону розподілу дослідних даних, основи регресійного і кореляційного аналізів та методи планування експерименту.

Навчальний посібник призначено для студентів, аспірантів, наукових працівників, які навчаються або ведуть дослідження за інженерними та природничими напрямками.

Рецензенти:

Жигуц Ю.Ю., завідувач кафедри машинобудування УжНУ,
доктор техн. наук, професор

Небола І.І., завідувач кафедри прикладної фізики УжНУ,
доктор фіз.-мат. наук, професор

Рекомендовано до друку Вченою радою УжНУ

Протокол № 12 від 19 грудня 2019 року.

ISBN 978-617-7825-00-4

©Ужгородський національний
університет, 2019 рік

ЗМІСТ

	стр.
ПЕРЕДМОВА	6
1. ЕКСПЕРИМЕНТ ЯК ПРЕДМЕТ ДОСЛІДЖЕННЯ	8
1.1. Поняття експерименту	8
1.2. Класифікація видів експериментальних досліджень	9
1.3. Класифікація вимірювань, методів і засобів вимірювань	14
Контрольні питання	19
2. КОРОТКІ ВІДОМОСТІ З ТЕОРІЇ ЙМОВІРНОСТІ ТА МАТЕМАТИЧНОЇ СТАТИСТИКИ	20
2.1. Випадкові величини і параметри їх функцій розподілу	20
2.2. Нормальний закон розподілу	30
Контрольні питання	35
3. ПОПЕРЕДНЯ ОБРОБКА ЕКСПЕРИМЕНТАЛЬНИХ ДАНИХ	36
3.1. Обчислення параметрів емпіричних розподілів.	
Точкове оцінювання	36
3.2. Оцінювання за допомогою довірчого інтервалу	44
3.2.1. Розрахунок довірчого інтервалу для математичного очікування. Розподіл Стьюдента	45
3.2.2. Побудова довірчого інтервалу для дисперсії. Розподіл Пірсона	54
3.2.3. Визначення необхідної кількості дослідів при побудові інтервальної оцінки для математичного очікування	59
3.3. Статистичні гіпотези	60
3.4. Відсів грубих похибок. Критерій Смирнова-Граббса	64
3.5. Порівняння двох рядів спостережень.	
Дисперсійний аналіз	69
3.5.1. Порівняння середніх значень. Критерій Стьюдента (t -критерій)	70

3.5.2. Порівняння двох дисперсій.	
Критерій Фішера	71
3.5.3. Перевірка однорідності декількох дисперсій.	
Критерій Кохрена	75
3.6. Перевірка гіпотез про вид функції розподілу.	
Критерій відповідності Пірсона	78
3.7. Перетворення розподілів до нормального	82
Контрольні питання	83
4. АНАЛІЗ ЕМПІРИЧНИХ ЗАЛЕЖНОСТЕЙ	84
4.1. Характеристика видів зв'язків між рядами спостережень	84
4.2. Визначення коефіцієнтів рівняння регресії	89
4.3. Визначення тісноти зв'язку між випадковими величинами.	
Коефіцієнт кореляції. Критерій Фішера	94
4.4. Лінійна регресія від одного фактора	97
4.5. Нелінійна регресія	100
4.6. Лінійна множинна регресія	101
Контрольні питання	104
5. ОСНОВИ ПЛАНУВАННЯ ЕКСПЕРИМЕНТУ	105
5.1. Виникнення та становлення планування експерименту	105
5.2. Основні поняття планування експерименту	108
5.3. Повний факторний експеримент	116
5.4. Дробовий факторний експеримент	119
5.5. Дробовий факторний експеримент	123
5.6. Планування експерименту з пошуку екстремуму відгуку	126
5.6.1. Класичні методи визначення екстремуму.	
Метод Гауса-Зейделя	126
5.6.2. Елементи симплекс планування екстремальних експериментів	129
5.6.3. Факторний метод визначення екстремуму	
Бокса-Вільсона	132
Контрольні питання	134

6. ПРАКТИЧНІ ЗАВДАННЯ ОБРОБКИ ДАНИХ З ВИКОРИСТАННЯМ MS Excel	135
6.1. Первинна обробка експериментальних даних з використання електронних таблиць MS Excel	135
6.1.1. Визначення статистичних характеристик випадкових величин	135
6.1.2. Відсів грубих похибок	139
6.1.3. Корекція статистичних характеристик	140
6.2. Оцінка виду функції розподілу випадкової величини	142
6.2.1. Оцінка виду функції розподілу випадкової величини за допомогою побудови діаграм та графічним методом	142
6.2.2. Перевірка гіпотези про вид розподілу з використанням критерію відповідності Пірсона	146
6.2.3. Визначення основних характеристик вибірки за допомогою вбудованих функцій Excel	149
6.3. Апроксимація експериментальних даних лінійною парною регресією	151
6.3.1. Обчислення коефіцієнтів рівняння лінійної регресії	151
6.3.2. Обчислення вибіркового коефіцієнта кореляції	154
6.3.3. Використання статистичних функцій табличного процесора MS Excel для визначення параметрів лінійної парної регресії	155
РЕКОМЕНДОВАНА ЛІТЕРАТУРА	158

ПЕРЕДМОВА

У всіх напрямках сучасної науки, виробництва часто і навіть суспільного життя доводиться вирішувати завдання з планування, проведення експериментальних досліджень, обробки та інтерпретації їх результатів. Після отримання перших експериментальних даних процедура експерименту не припиняється. По-перше, як правило, разовий експеримент не дає остаточної відповіді на поставлене питання. По-друге, отримані експериментальні результати потребують логічного доопрацювання, перетворення їх у науковий факт, тобто у те, в істинності чого не виникає сумнівів. Окремі експериментальні дані, отримані на початковій стадії емпіричного дослідження, самі по собі не стають фактами. Вони можуть містити помилки, пов'язані з некоректною постановкою експерименту, неправильними показаннями вимірювальних приладів, відхиленнями у функціонуванні органів чуття і т. п. Тому, як правило, проводиться не один, а серія дослідів. Уточнюються і перевіряються результати експерименту, збираються відсутні відомості, проводяться додаткові експерименти. Потім отримані у серії експериментів дані піддаються математичній обробці. При уявній простоті отримання та обробки первинних експериментальних даних, тобто результатів спостережень і вимірювань, математична обробка проводиться в рамках теорії похибок, на підставі якої кількісно визначається достовірність остаточної результату. Наскільки б точними не були спостереження та вимірювання, похибки неминучі, і завдання полягає в тому, щоб наблизити експериментальні дані до об'єктивних значень визначених величин, тобто зменшити інтервал неточності. Сучасна теорія похибок озброює експериментаторів надійними засобами коригування експериментальних даних.

Автори ставили перед собою за мету викласти питання планування і обробки результатів експерименту в максимально доступній та зрозумілій формі. Місце даного навчального посібника серед інших книг і посібників з методів планування і обробки результатів експерименту визначається бажанням дати доступний виклад елементів теорії, не переобтяжений строгими математичними викладками. Іноді для отримання кількісних результатів, корисних для практичного застосування, доводиться жертвувати строгістю

теоретичних принципів. Однак такі компромісні рішення приймалися тільки в тих випадках, коли їх доцільність підтверджувалася практикою. Більш вимогливий читач може звернутися до спеціальної літератури, зазначеної в кінці посібника.

У даний час обробка результатів експериментів немислима без використання комп'ютерів і пакетів прикладних програм. При тій вражаючій різноманітності статистичних пакетів, яким характеризується сучасний світовий та вітчизняний ринок (число різних найменувань розповсюджуваних на ринку статистичних пакетів наближається до тисячі!), важливо вміти правильно зорієнтуватися на цьому ринку, правильно розуміти область застосування статистичних методів вирішення того чи іншого класу задач. Глибоке освоєння можливостей спеціалізованих пакетів вимагає значного проміжку часу, при цьому у користувача може виникнути ілюзія освоєння і самої теорії планування та обробки експериментальних даних. Однак слід розуміти, що інструмент не замінює компетентність і професіоналізм. Комп'ютер являє собою не заміну людського інтелекту, а є лише його підсилювачем. Ніякі яскраві можливості сучасних інтерфейсів (вікна, що розкриваються, контекстне меню, кнопки і т.д.) не звільняють користувача комп'ютера від необхідності вивчення і розуміння суті методів, реалізованих в таких системах.

У кожній розглядуваній темі цього посібника по можливості при викладі теоретичного матеріалу розглянуті типові приклади розв'язання задач. Наявність достатньої кількості прикладів та приведені у шостому розділі посібника практичні завдання дозволяють використовувати представлений матеріал не тільки як короткий конспект лекцій, а і як інструмент при проведенні семінарських та практичних занять.

Зручним універсальним обчислювальним середовищем для розв'язку задач обробки експериментальних даних є табличний процесор MS Excel. Тому в методичних вказівках виклад чисельних методів розв'язку основних задач первинної обробки експериментальних даних (знаходження основних характеристик статистичних величин, відсів грубих помилок, оцінка функції розподілу) регресійного аналізу (визначення коефіцієнтів лінійної регресії, обчислення коефіцієнта кореляції) зроблений з використанням обчислювального середовища табличного процесора MS Excel.

1. ЕКСПЕРИМЕНТ ЯК ПРЕДМЕТ ДОСЛІДЖЕННЯ

1.1 Поняття експерименту

У багатьох областях наукової та практичної діяльності сучасної людини значне місце займають теоретичні методи вивчення різних об'єктів і процесів навколишнього нас світу. Однак, незважаючи на високу ефективність теоретичних методів, найчастіше доводиться зустрічатися з завданнями, вирішення яких практично неможливо без організації і проведення того чи іншого експериментального дослідження.

Із загальнофілософської точки зору експеримент - це чуттєво-предметна діяльність в науці; в більш вузькому сенсі - дослід, відтворення об'єкта пізнання, перевірка гіпотез і т.д.

У технічній літературі терміну експеримент встановлюється наступне визначення - система операцій, дій і (або) спостережень, спрямованих на отримання інформації про об'єкт дослідження.

Будучи джерелом пізнання і критерієм істинності теорій і гіпотез, експеримент грає дуже важливу роль в дослідницьких лабораторіях, на діючому виробництві, в медичних клініках, на дослідних сільськогосподарських полях, в космосі і в глибинах океану, а також в сучасному навчальному процесі як у загальноосвітніх школах, так і у вищих навчальних закладах. Питання обробки статистичних даних виникають також при проведенні соціологічних досліджень, політичних прогнозів.

Хоча об'єкти досліджень дуже різноманітні, методи експериментальних досліджень мають багато спільного:

- яким би простим не був експеримент, спочатку вибирають план його проведення;
- прагнуть скоротити число розглядуваних змінних для того щоб зменшити об'єм експерименту;
- намагаються контролювати хід експерименту;
- намагаються виключити вплив випадкових зовнішніх впливів;
- оцінюють точність вимірювальних приладів і точність отримання даних;
- і нарешті, в процесі будь-якого експерименту аналізують отримані результати і прагнуть дати їх інтерпретацію, оскільки без цього вирішального етапу весь процес експериментального дослідження не має сенсу.

На жаль, часто робота експериментатора настільки хаотична і неорганізована, а її ефективність настільки мала, що отримані результати не в змозі виправдати навіть тих зусиль і коштів, які були витрачені на проведення дослідів. Тому питання організації експерименту, зниження витрат на його проведення і обробку отриманих результатів є актуальними.

Сучасні методи планування експерименту і обробки його результатів, розроблені на основі теорії ймовірностей і математичної статистики, дозволяють істотно (найчастіше в кілька разів) скоротити число необхідних дослідів. Знання та використання цих методів робить роботу експериментатора більш цілеспрямованою і організованою, істотно підвищує як продуктивність його праці, так і надійність одержуваних ним результатів.

Як і будь-яка інша наукова дисципліна, організація і планування експерименту мають свою певну строгую термінологію, для початкового знайомства з якою ми розглянемо класифікацію видів експерименту.

1.2. Класифікація видів експериментальних досліджень

Перш за все зазначимо, що будь-який експеримент передбачає проведення тих чи інших *дослідів*.

Дослід - відтворення досліджуваного явища в певних умовах при можливості реєстрації його результатів.

За метою проведення та форми подання отриманих результатів експеримент ділять на *якісний* і *кількісний*.

Якісний експеримент встановлює тільки сам факт існування будь-якого явища, але при цьому не дає ніяких кількісних характеристик об'єкта дослідження. Будь-який експеримент, яким би складним він не був, завжди закінчується поданням його результатів, формулюванням висновків, видачею рекомендацій. Ця інформація може бути виражена в виді графіків, креслень, таблиць, формул, статистичних даних або словесних описів. Якісний експеримент якраз і передбачає саме словесний опис його результатів.

Приклад 1.1. Якщо взяти два шматки сталевого дроту, один з яких не мав термічної обробки, а другий пройшов операції відпалу, і піддати їх кільком перегинам, то легко переконатися, що термічно не оброблений метал руйнується раніше (при меншій кількості перегинів), тобто має меншу пластичність;

Приклад 1.2. Пропускання електричного струму через провідник призводить до виникнення його нагрівання.

Однак словесний опис - не найефективніший і інформативний спосіб представлення результатів експерименту, оскільки він не дозволяє дати кількісних рекомендацій. Тому в практиці основний зміст експерименту представляється числом і кількісними залежностями.

Кількісний експеримент не тільки фіксує факт існування того чи іншого явища, але, крім того, дозволяє встановити співвідношення між кількісними характеристиками явища і кількісними характеристиками способів зовнішнього впливу на об'єкт дослідження.

В умовах прикладу 1.1 а, для того щоб перевести експеримент з розряду "якісний" в "кількісний", необхідно:

- визначити і кількісно описати ті параметри процесу відпалу і ті властивості матеріалу, які за припущенням можуть вплинути на пластичність сталі (наприклад: температура відпалу, швидкість охолодження, фактичний хімічний склад сталі, з якої виготовлений дріт і т.д.);

- вибрати ту чи іншу кількісну характеристику пластичності металу;

- у результаті експерименту необхідно встановити кількісну залежність між пластичністю дроту і параметрами процесу термообробки, з урахуванням тих можливих коливань хімічного складу, які допустимі для даної марки сталі.

Отже, кількісний експеримент насамперед передбачає кількісне визначення всіх тих *способів зовнішнього впливу* на об'єкт дослідження, від яких залежить його поведінка - кількісний опис всіх *факторів*.

Фактор - змінна величина, яка за припущенням впливає на результати експерименту.

Наприклад, в якості факторів розглянутого ілюстративного експерименту можна вибрати температуру відпалу, швидкість охолодження, процентний вміст вуглецю в сталі.

В окремому конкретному досліді кожен фактор може приймати одне з можливих своїх значень - який отримав назву рівня фактора.

Рівень фактора - фіксоване значення фактора відносно початку відліку. Наприклад, одним рівнем такого фактора, як швидкість охолодження, може бути 50 °С/год, іншим - 100 °С/год і т.д.

Фіксований набір рівнів всіх факторів в кожному конкретному досліді визначає одне з можливих станів об'єкта дослідження.

При проведенні дослідів дуже багато залежить від того, наскільки активно експериментатор може "втручатися" в досліджуване явище, має він чи ні можливість встановлювати ті рівні факторів, які представляють для нього інтерес.

З цієї точки зору всі фактори можна розбити на три групи:

- *контрольовані і керовані* - це фактори, для яких можна не тільки зареєструвати їх рівень, але ще і задати в кожному конкретному досліді будь-яке його можливе значення;

- *контрольовані, але некеровані* чинники - це чинники, рівні яких можна тільки реєструвати, а ось поставити в кожному досліді їх певне значення практично неможливо;

- *неконтрольовані* - це фактори, рівні яких не реєструються експериментатором і про існування яких він навіть може і не підозрювати.

У прикладі 1.1 в якості контрольованих і керованих чинників можна розглядати температуру відпалу і швидкість охолодження. А ось фактичний процентний вміст різних хімічних елементів у сталі, очевидно, потрапить в групу контрольованих, але некерованих факторів. Справа тут в тому, що хімічний склад ще може і вдасться зареєструвати, але ось задати для кожного досліді чітко визначений процентний вміст, наприклад, вуглецю - завдання практично нездійсненне. Нарешті, до групи неконтрольованих чинників у цьому прикладі можна віднести масу причин, за якими може змінитися пластичність металу (нерівномірність деформації металу, несприятливі умови зберігання, що призводять до його підвищеної корозії, і так далі і тому подібне, на скільки в даному випадку вистачить фантазії дослідника).

У кількісному експерименті необхідно не тільки реєструвати рівні всіх контрольованих факторів, а й мати можливість встановлювати кількісний опис тої *властивості (відгуку)* досліджуваного явища, яке вивчає (спостерігає) експериментатор. Причому, оскільки на об'єкт дослідження в процесі експерименту завжди впливає величезна кількість неконтрольованих факторів, що вносить в одержувані результати певний елемент невизначеності, значення відгуку, в кожному конкретному досліді, неможливо передбачити заздалегідь. Тому відтворення досліджуваного явища при

одному і тому ж фіксованому наборі рівнів всіх контрольованих факторів завжди буде призводити до різних значень відгуку, тобто відгук - це завжди випадкова величина.

Відгук - спостережувана випадкова змінна, яка за припущенням залежить від зовнішніх чинників.

Відгуком в умовах прикладу 1.1 є пластичність сталевого дроту. Причому навіть якщо взяти шматки дроту із одного і того ж мотка (тобто метал однієї плавки - однакового хімічного складу, який має один і той же режим термообробки при однаковій температурі відпалу і швидкості охолодження), то і при цьому для кожного шматка дроту ми отримуємо різні (хоча і дуже близькі один до одного) значення пластичності металу.

І нарешті, в результаті кількісного експерименту необхідно знайти залежність між відгуком і факторами - *функцію відгуку*. Причому оскільки відгук - це випадкова величина, то, з точки зору теорії ймовірностей, його можна задати одним з параметрів свого розподілу, наприклад математичним очікуванням.

Функція відгуку - залежність математичного очікування відповіді від факторів.

У прикладі з дротом - це залежність математичного очікування величини пластичності сталі від температури відпалу, швидкості охолодження і хімічного складу металу.

З урахуванням наведеного вище поділу факторів на три групи, функцію відгуку в найзагальнішому випадку можна записати у виді

$$M_y = M_y(p, h) + \varepsilon, \quad (1.1)$$

де M_y - математичне очікування відгуку; p - контрольовані і керовані чинники; h - контрольовані, але некеровані фактори; ε - помилка експерименту, що враховує вплив неконтрольованих факторів.

По тому, якою групою факторів має справу дослідник, кількісний експеримент в свою чергу можна розділити ще на два види. Якщо в розпорядженні експериментатора немає керованих факторів, то такий експеримент носить назву *пасивного*.

Пасивний експеримент - експеримент, при якому рівні факторів в кожному досліді реєструються дослідником, але не задаються (наприклад, астрофізичні спостереження).

Оскільки при пасивному експерименті дослідник не має можливість задати рівень жодного з чинників, то при проведенні дослідів йому залишається лише "пасивно" спостерігати за явищем і реєструвати результати. Планування пасивного експерименту зводиться до визначення числа дослідів, які необхідно провести досліднику для вирішення поставленого перед ним завдання, а кінцевою метою пасивного експерименту в більшості випадків є отримання функції відгуку у виді

$$M_y = M_y(h) + \varepsilon. \quad (1.2)$$

Якщо ж експериментатор має можливість не тільки контролювати чинники, а й управляти ними, то такий експеримент носить назву *активного*.

Активний експеримент - експеримент, в якому рівні факторів у кожному досліді задаються експериментатором.

Оскільки в цьому випадку експериментатор має можливість "активно" втручатися в досліджуване явище, то природно, що активний експеримент завжди передбачає *план* його проведення.

План експерименту - сукупність даних, що визначають число, умови і порядок реалізації дослідів.

Тому активний експеримент завжди повинен починатися з планування.

До вимог, що пред'являються при плануванні активного експерименту, можна віднести ступінь точності і надійності результатів, отриманих після проведення експерименту, терміни і засоби, наявні в розпорядженні дослідника, і т.д.

Метою активного експерименту може бути або визначення функції відгуку, або пошук такого поєднання рівнів керованих факторів p , при яких досягається оптимальне (екстремальне - мінімальне або максимальне) значення функції відгуку. У цьому останньому випадку експеримент носить ще назву пошукового (екстремального) експерименту.

Наприклад, якщо у випадку з руйнуванням дроту ми б поставили перед собою мету знайти таке поєднання температури відпалу і швидкості охолодження, при яких пластичність металу була б максимальною, то наш експеримент став би пошуковим.

І нарешті, за умовами проведення розрізняють *лабораторний* і *виробничий експерименти*.

Лабораторний експеримент. У лабораторії менший вплив випадкових похибок, забезпечується велика "стерильність" умов проведення дослідів, в більшості випадків здійснюється і більш ретельна підготовка, одним словом, вища "культура експерименту". Як правило, в лабораторних умовах експериментатор може відтворити дослід "однаково" значно краще, ніж у виробництві. Це означає, що за інших рівних умов для встановлення деякого факту на заводі буде потрібно виконати значно більше дослідів, ніж у лабораторії. Інша важлива відмінність - це велика можливість варіювати (змінювати) рівні факторів. Коли в лабораторії досліджується хімічна реакція, температуру за бажанням можна змінювати в широких межах, а у *виробничих технологічних процесах*, навпаки, якщо її і можна змінювати, то в значно вужчому діапазоні і з більшою обережністю.

1.3. Класифікація вимірювань, методів і засобів вимірювань

Результатом проведення довільного кількісного фізичного експерименту є вимірювання фізичних величин.

Фізична величина – це властивість, в якісному відношенні загальна для багатьох фізичних об'єктів (фізичних систем, їхніх станів і процесів, що у них відбуваються), але в кількісному відношенні індивідуальна для кожного об'єкта. Якісна сторона визначає "вид" величини (наприклад, електричний опір), а кількісна – її "розмір" (наприклад, опір конкретного резистора). Фізичними величинами є також довжина, температура, сила струму, напруженість електричного поля, період коливань і т.д.

Конкретні реалізації однієї й тієї ж самої величини називають *однорідними* величинами. Наприклад, відстань між дзеркалом гальванометра та шкалою і висота телевежі - це конкретні реалізації однієї фізичної величини – довжини, і через це вони є однорідними фізичними величинами. Однорідні фізичні величини відрізняються одна від одної розміром.

Розмір фізичної величини – це кількісний вміст у даному об'єкті властивості, яка відповідає поняттю "фізична величина". Розміри однорідних фізичних величин різних об'єктів порівнюють між собою, якщо відомі значення цих величин.

Значенням фізичної величини називається оцінка фізичної величини за допомогою певного числа прийнятих для неї одиниць.

Одиниця фізичної величини – це фізична величина, якій за означенням надано числове значення, що дорівнює 1.

Абстрактне число, яке виражає відношення значення величини до відповідної одиниці цієї фізичної величини, називається *числовим значенням величини*. Потрібно розрізняти істинне та дійсне значення фізичної величини.

Істинне значення фізичної величини – це таке її значення, яке ідеальним чином відобразило б в якісному та кількісному відношеннях відповідну властивість об'єкта.

Дійсне значення фізичної величини – це таке значення величини, яке знайдене експериментально і настільки наближається до істинного значення, що для певної мети може бути використане замість нього.

Знаходження значень фізичної величини дослідним шляхом за допомогою спеціальних технічних засобів називають вимірюванням.

Вимірювання – це фізичний експеримент порівняння даної фізичної величини з деяким її значенням, прийнятим за одиницю порівняння. Розмір одиниці величини може бути довільним. Проте, вимірювання повинні виконуватись у загальноприйнятих одиницях і, насамперед, у міжнародній системі одиниць СІ, до якої як основні одиниці входять одиниця маси - кілограм (кг), одиниця довжини - метр (м), одиниця часу - секунда (с), одиниця температури - Кельвін (К), одиниця кількості речовини – моль, одиниця сили струму - ампер (А), одиниця сили світла – кандела (кд). Одиниця фізичної величини повинна бути матеріалізована, тобто відтворена, причому таким чином, щоб її розмір був постійним у часі і не залежав від зовнішніх дій. Наявність точно відтворюваних одиниць фізичних величин дозволяє забезпечити єдність вимірювань, виконаних у різних місцях, у різний час і різними засобами.

Вимірювання, як експериментальні процедури визначення фізичних величин, досить різноманітні, що пояснюється великою кількістю вимірюваних величин, різним характером їх зміни з часом, різними вимогами до точності вимірювання тощо.

Вимірювання в залежності від способу обробки експериментальних даних для знаходження результату поділяють на *прямі, посередні, сукупні та сумісні*.

Пряме вимірювання – вимірювання, при якому шукане значення фізичної величини знаходять безпосередньо з дослідних даних.

Приклад прямого вимірювання – вимірювання вольтметром спаду напруги на резисторі.

Посереднє вимірювання – вимірювання, при якому шукане значення фізичної величини знаходять на основі відомої залежності між цією величиною і величинами, які піддаються прямим вимірюванням. При посередньому вимірюванні значення вимірюваної величини одержують шляхом розв’язання рівняння:

$$Y = F (X_1, X_2, X_3, \dots, X_n), \quad (1.3)$$

де $X_1, X_2, X_3, \dots, X_n$ – значення величин, одержаних за допомогою прямих вимірювань. Приклад посереднього вимірювання: знаходження об’єму тіла за його геометричними розмірами; визначення опору резистора R із рівняння $R = U/I$, в яке підставляють виміряні значення спаду напруги U на резисторі та сили струму I , який через нього протікає.

Сукупні вимірювання – одночасні вимірювання кількох однойменних величин, за якими шукані значення величин знаходять розв’язуванням системи рівнянь, здобутих при прямим вимірюваннях різних сполучень цих величин або при зміні умов вимірювань. Приклад сукупного вимірювання: вимірювання опорів резисторів, з’єднаних трикутником, шляхом вимірювань опорів між різними вершинами трикутника; за результатами трьох вимірювань визначають опори резисторів.

Сумісні вимірювання – одночасне вимірювання двох або кількох неоднойменних величин для знаходження залежності між ними (наприклад, температурний коефіцієнт опору резистора визначають заданими прямим вимірювань його опору при різних температурах).

Усі вимірювання фізичних величин здійснюють за допомогою засобів вимірювання. *Засіб вимірювання* – це технічний пристрій, який використовується при вимірюваннях і має нормовані метрологічні властивості.

Розрізняють такі види засобів електричних вимірювань:

- міри та еталони;
- вимірювальні перетворювачі;
- вимірювальні прилади;
- вимірювальні установки;
- вимірювальні системи.

Еталоном одиниці фізичної величини називають засіб вимірювання (або комплекс засобів вимірювання), який забезпечує відтворення і (або) зберігання одиниці з метою передавання її розміру нижчим за перевіркою схемою засобам вимірювання.

Взірцевими засобами вимірювання називаються міра, вимірювальний прилад або вимірювальний перетворювач, які призначені для перевірки за ними інших засобів і утверджені як взірцеві. Взірцеві засоби вимірювання можуть бути також безпосередньо використані для точних вимірювань.

Мірою називається засіб вимірювання, призначений для відтворення фізичної величини заданого розміру. Розрізняють однозначну міру, багатозначну міру і набір мір. Однозначна міра відтворює фізичну величину одного розміру. Багатозначна міра відтворює ряд значень однойменних фізичних величин різного розміру. *Набір мір* це підібраний комплект мір для відтворення ряду значень однойменних величин різного розміру, причому міри можуть використовуватись як окремо, так і в різних комбінаціях. Прикладом набору мір є шальки терезів, магазин опорів, ємностей тощо. *Магазин мір* – це набір мір, конструктивно об'єднаних в одне ціле.

Вимірювальним приладом називається засіб вимірювання, який служить для вироблення сигналу вимірюваної інформації у формі, доступній для безпосереднього сприймання спостерігачем. Вимірювальні прилади класифікуються за різними ознаками. Прилади, покази яких є неперервними функціями змін вимірюваних величин, називаються *аналоговими* приладами. Вимірювальні прилади, які автоматично виробляють дискретні сигнали вимірювальної інформації і покази яких подані в цифровій формі, називаються *цифровими* приладами.

Вимірювальні прилади поділяються на *показуючі*, які допускають тільки зчитування показів оператором, та *реєструючі*, в яких передбачена реєстрація і зберігання інформації. Реєструючі прилади поділяються на *самописні* – з записом показів у формі діаграми, на якій може бути відтворена неперервна функція вимірюваної величини, і *цифрові*, в яких передбачено друк показів у цифровій формі. У приладах з вмонтованими міні-ЕОМ можлива реєстрація і зберігання інформації в оперативній пам'яті міні-ЕОМ, або, наприклад, на магнітних носіях інформації.

Для одержання результату вимірювань фізичної величини в прийнятих одиницях у процесі вимірювань обов'язково повинна брати участь міра. Вимірювальний *прилад прямої дії* може бути попередньо проградуйований в одиницях вимірюваної величини, тобто міра попередньо використовується в процесі виготовлення приладу. Існують прилади, які призначені для безпосереднього порівняння вимірюваної величини з величиною, значення якої відоме. Такі вимірювальні прилади називаються *приладами порівняння*.

За характером використання прилади поділяються на стаціонарні (щитові), корпуси яких пристосовані для жорсткого кріплення на місці встановлення, та переносні. В залежності від ступеня захищеності прилади бувають звичайними, пило-, водозахисними, герметичними тощо.

Вимірювальними перетворювачами називають засоби вимірювань, призначені для вироблення сигналів вимірної інформації у формі, зручній для передачі, подальшого перетворення, обробки і (або) зберігання, але які не піддаються безпосередньому сприйняттю спостерігачем. Залежність між вимірюваною величиною і вихідним сигналом вимірювального перетворювача називається *функцією перетворення*. В залежності від призначення вимірювальні перетворювачі поділяються на масштабні, які змінюють вхідну величину у задане число разів, і перетворювачі роду величини. До масштабних відносяться шунти, подільники напруги, вимірювальні трансформатори, електронні підсилювачі.

Слід відмітити окремо групу аналого-цифрових перетворювачів (АЦП), які перетворюють вимірювані величини у код і які широко використовуються у цифрових вимірювальних приладах.

Перетворювачі роду величин, наприклад, неелектричних в електричні, складають різноманітну групу пристроїв, які використовують у різних галузях вимірювальної техніки, наприклад, при вимірюваннях температури, тиску і т.д. Деякі види вимірювальних перетворювачів інколи називають датчиками (сенсорами, первинними перетворювачами), під якими розуміють сукупність одного або кількох вимірювальних перетворювачів і супутніх їм конструктивних елементів, розташованих безпосередньо на об'єкті вимірювання і віддалених від місця відображення, реєстрації або опрацювання вимірюваної інформації.

Вимірювальною установкою називають сукупність функціонально об'єднаних засобів вимірювання (мір, вимірювальних приладів і перетворювачів) та допоміжних пристроїв, призначених для вироблення сигналів вимірюваної інформації у формі, зручній для безпосереднього сприймання спостерігачем або її автоматичної реєстрації.

Інформаційно-вимірювальна система – це сукупність функціонально об'єднаних вимірювальних, обчислювальних та інших допоміжних технічних засобів, об'єднаних каналами зв'язку і призначених для отримання вимірюваної інформації, її перетворення, опрацювання з метою подання споживачу (в тому числі автоматизованій системі управління (АСУ)) у необхідному виді або автоматичного здійснення логічних функцій контролю, діагностування, ідентифікації.

Контрольні питання

1. Дайте означення поняття експерименту.
2. Назвіть спільні риси експериментальних досліджень.
3. Наведіть приклад якісного експерименту.
4. Наведіть приклад кількісного експерименту.
5. Охарактеризуйте фактори і функцію відгуку наведеного вами експерименту.
6. Проаналізуйте керовані, контрольовані та неконтрольовані фактори наведеного вами експерименту.
7. Наведіть приклади класифікації експериментів.
8. Що ви розумієте під поняттям вимірювання?
9. Назвіть і охарактеризуйте основні типи вимірювань.
10. Охарактеризуйте види засобів вимірювань.

2. КОРОТКІ ВІДОМОСТІ З ТЕОРІЇ ЙМОВІРНОСТІ ТА МАТЕМАТИЧНОЇ СТАТИСТИКИ

2.1. Випадкові величини і параметри їх функцій розподілу

Оскільки через вплив неконтрольованих факторів відгук - це завжди випадкова величина, при обробці результатів експерименту широко використовується апарат теорії ймовірностей і математичної статистики, тому нагадаємо деякі основні поняття і визначення цього розділу математики.

Випадкова подія - подія, реалізацію якої при певному комплексі умов неможливо заздалегідь передбачити.

Випадкова величина - величина, яка може приймати будь-яке значення з встановленого числа і характеризує випадкову подію. Випадкова величина може бути *дискретною* або *неперервною*.

Дискретна випадкова величина – величина, яка може приймати значення тільки із скінченної множини значень обмеженої або нескінченної множини дійсних чисел.

Неперервна випадкова величина - величина, яка може приймати будь-які значення з обмеженого або нескінченного інтервалу.

Якщо при фіксованому наборі рівнів всіх контрольованих факторів провести n вимірювань відгуку x , то в результаті буде отримано ряд хоча і близьких значень, але які відрізняються один від одного:

$$x_i, (i = 1, 2, \dots, n), \quad (2.1)$$

де x_i - i -те вимірювання величини x ; x_1, x_2, \dots, x_n - реалізація випадкової величини X .

Кожному значенню x_i дискретної випадкової величини X (будь-якій події A , коли випадкова величина X приймає якесь строго певне значення x_i), можна поставити у відповідність наступне відношення:

$W_i = \frac{m_i}{n}$, де m - число спостережень, в яких дискретна випадкова величина x виявилась рівною x_i ; n - загальна кількість спостережень. Величину W_i називають частотою реалізації події A .

Межа, до якого прямує відношення $W_i = \frac{m_i}{n}$ при необмеженому зростанні числа дослідів n , називається *ймовірністю випадкової події*:

$$P(A) = \lim_{n \rightarrow \infty} \frac{m_i}{n}. \quad (2.2)$$

Ймовірність $P(A)$ події A - число від нуля до одиниці, яке представляє собою межа частоти реалізації події A при необмеженій кількості повторень одного і того ж комплексу умов.

Для дискретної випадкової величини можна вказати ймовірність, з якою вона приймає кожне зі своїх можливих значень обмеженої множини дійсних чисел. Для неперервної випадкової величини задають ймовірність її попадання в один із заданих інтервалів області її визначення (оскільки ймовірність того, що вона прийме якесь конкретне своє значення, прямує до нуля).

Повністю властивості випадкової величини описуються *законом її розподілу*, під яким розуміють зв'язок між можливими значеннями випадкової величини і відповідними їм ймовірностями появи.

Розподіл випадкової величини - функція, яка однозначно визначає ймовірність того, що випадкова величина приймає задане значення або належить до деякого заданого інтервалу.

У математиці використовують два способи опису розподілів випадкових величин: *інтегральний (функція розподілу)* і *диференціальний (густина розподілу)*.

Інтегральна функція розподілу $F(x)$ - функція визначає ймовірність того, що випадкова величина X приймає значення не більш, ніж x :

$$F(x)=P(X<x). \quad (2.3)$$

Функція розподілу $F(x)$ має такі властивості (рис.2.1, а):

1. Її ордината, відповідна довільній точці x_1 , являє собою ймовірність того, що випадкова величина X буде менше, ніж x_1 , тобто $F(x_1)=P(X<x_1)$.

2. Функція розподілу приймає значення, що лежать між нулем і одиницею:

$$0 < F(x) < 1. \quad (2.4)$$

3. Функція розподілу прагне до нуля при необмеженому зменшенні x і прагне до одиниці при необмеженому зростанні x , тобто

$$\lim_{x \rightarrow -\infty} F(x) = 0, \quad \lim_{x \rightarrow +\infty} F(x) = 1. \quad (2.5)$$

4. Функція розподілу є монотонно зростаючою кривою, тобто

$$F(x_2) > F(x_1), \quad \text{якщо} \quad x_2 > x_1. \quad (2.5a)$$

5. Приріст інтегральної функції розподілу на довільному відрізку $(x_1; x_2)$ дорівнює ймовірності того що випадкова величина попаде в цей інтервал

$$F(x_2) - F(x_1) = P(X \leq x_2) - P(X \leq x_1) = P(x_1 \leq X \leq x_2). \quad (2.6)$$

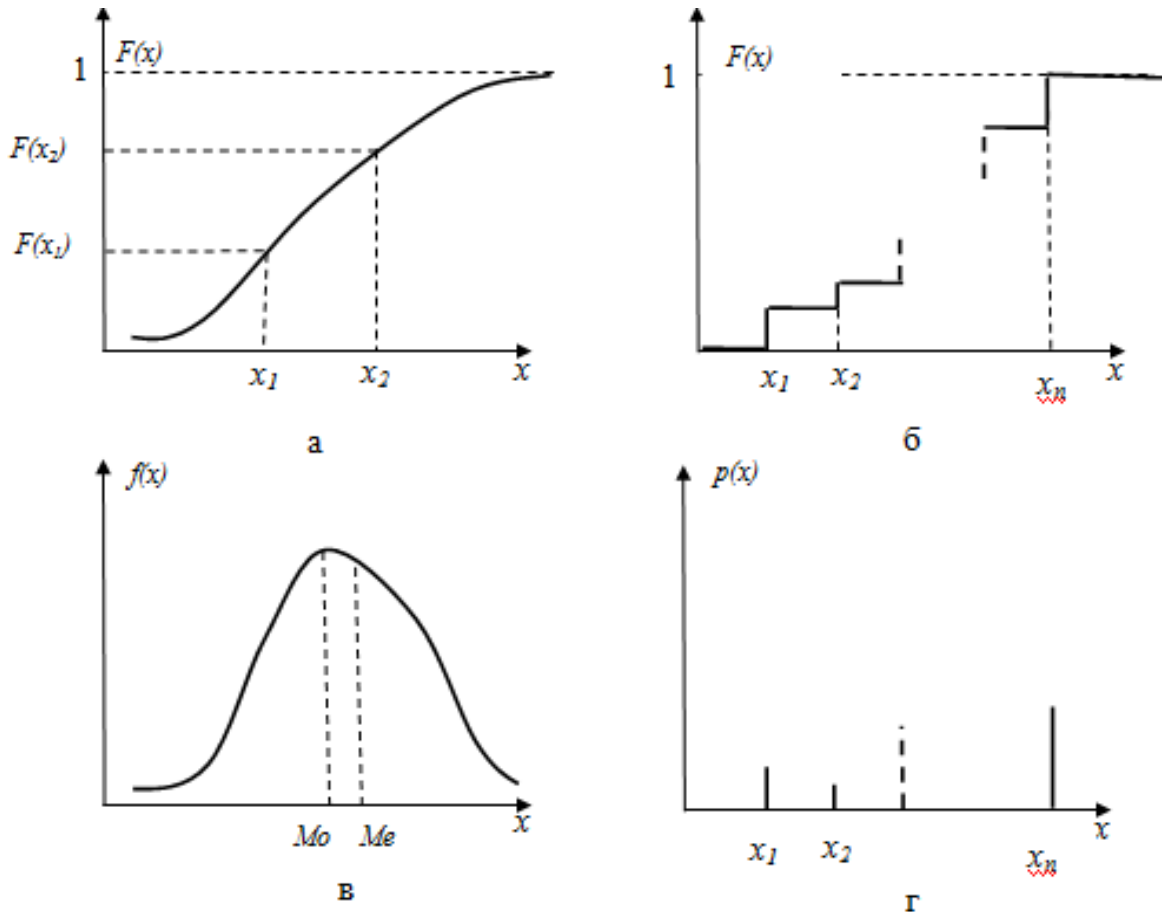


Рис.2.1. Інтегральна (а, б) та диференціальна (в) функції розподілу ймовірності для неперервної (а, в) і дискретної (б) випадкових величин. На графіку (г) ймовірності появи дискретних величин x_i .

Розглянемо, які особливості мають функції розподілу дискретних випадкових величин. Нехай X - дискретна випадкова величина, яка може приймати значення x_1, x_2, \dots, x_n з вірогідністю p_1, p_2, \dots, p_n . Функція розподілу ймовірностей цієї випадкової величини X дорівнює

$$F(x) = P(X \leq x) = \sum_i p_i,$$

де проводиться сумування ймовірностей всіх можливих значень випадкової величини X , менших ніж x . Така функція завжди розривна, східчаста (рис.2.1, б): від $-\infty$ до x_1 включно функція дорівнює нулю, в точці x_1 відбувається стрибок на величину p_1 , і функція залишається

постійною до x_2 включно і т.д., тобто можливим значенням випадкової величини відповідають скачки функції, рівні ймовірностям цих значень. Останній стрибок на p_n відбувається в точці x_n , і функція дорівнює одиниці від x_n до $+\infty$. Таким чином, сума всіх стрибків дорівнює одиниці.

Густина розподілу - перша похідна (якщо вона існує) від інтегральної функції розподілу:

$$f(x) = \frac{dF(x)}{dx}. \quad (2.7)$$

Густина функції розподілу $f(x)$ має такі властивості (рис.2.1 в):

1. Густина розподілу ймовірностей є невід'ємною функцією, тобто

$$f(x) > 0. \quad (2.8)$$

Це властивість справедливо, так як $F(x)$ є всюди неспадною функцією.

2. Функція розподілу випадкової величини X дорівнює визначеному інтегралу від густини розподілу ймовірностей в межах $(-\infty, x)$:

$$F(x) = \int_{-\infty}^x f(x) dx. \quad (2.9)$$

3. Імовірність події, яка полягає в тому що випадкова величина X прийме значення в інтервалі $[x_1, x_2]$, дорівнює визначеному інтегралу від густини розподілу ймовірностей на цьому інтервалі:

$$P(x_1 \leq X \leq x_2) = F(x_2) - F(x_1) = \int_{x_1}^{x_2} f(x) dx. \quad (2.10)$$

4. Інтеграл від густини ймовірності у нескінченно великому інтервалі $(-\infty, +\infty)$ дорівнює одиниці:

$$\int_{-\infty}^{+\infty} f(x) dx = P(-\infty < X < +\infty) = 1, \quad (2.11)$$

так як попадання випадкової величини в інтервал $-\infty < X < +\infty$ є достовірною подією.

У більшості випадків при обробці експериментальних даних, ґрунтуючись на тих чи інших припущеннях (гіпотезах) щодо властивостей досліджуваної випадкової величини, вдається записати функцію її розподілу (а отже, і густину розподілу як першу похідну від функції розподілу) з точністю до деяких невідомих параметрів.

Наприклад, для випадкової величини, яка задовольняє закону нормального розподілу (закону розподілу Гауса), функцію розподілу густини ймовірності можна записати у виді

$$f(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp\left(-\frac{(x - M_x)^2}{2\sigma_x^2}\right). \quad (2.12)$$

У (2.12) константи M_x і σ_x є параметрами розподілу, причому закон розподілу є двопараметричним.



Йоган Гаус у 1828 р.

Йоганн Карл Фрідріх Гаус (нім. Johann Carl Friedrich Gauß, (30 квітня 1777, Брауншвейг - 23 лютий 1855, Геттінген) - німецький математик, механік, фізик, астроном і геодезист Народився в німецькому герцогстві Брауншвейг. Дід Гауса був бідним селянином, батько - садівником, каменярем, доглядачем каналів. Уже в дворічному віці хлопчик показав себе вундеркіндом. У три роки він умів читати і писати, навіть виправляв помилки батька при рахуванні. Згідно з легендою, шкільний

учитель математики, щоб зайняти дітей на довгий час, запропонував їм порахувати суму чисел від 1 до 100. Юний Гаус зауважив, що попарні суми з протилежних кінців однакові: $1 + 100 = 101$, $2 + 99 = 101$ і т. д., і миттєво отримав результат. До самої старості він звик більшу частину обчислень робити усно.

З учителем йому пощастило: М. Бартельс (згодом учитель Лобачевського) оцінив винятковий талант юного Гауса і зумів виклопотати йому стипендію від герцога Брауншвейзького. Це допомогло Гауса закінчити коледж Collegium Carolinum в Брауншвейгу (1792-1795). З 1795 по 1798 рік Гаус навчався в Геттінгенському університеті, У 1798 році Гаус повернувся в Брауншвейг і жив там до 1807 року.

З 1799 року Гаус - приват-доцент Брауншвейзького університету. З ім'ям Гауса пов'язані фундаментальні дослідження майже у всіх основних областях математики: Однак після 1801 року Гаус, не пориваючи з теорією чисел, розширив коло своїх інтересів, включивши в нього і природничі науки, в першу чергу астрономію. Приводом послужило відкриття малої планети Церера (1801), втраченої незабаром після виявлення. 24-річний Гаус виконав (за декілька годин) складні обчислення, користуючись розробленим ним же новим обчислювальним методом, і з великою точністю вказав місце, де шукати «втікачку»; там вона, до загального захоплення, і була незабаром виявлена. Слава Гауса стає загальноєвропейською. Багато наукових товариств Європи обирають Гауса своїм членом, а інтерес Гауса до астрономії ще більш зростає. За рекомендацією Олександра фон Гумбольдта Гауса призначають професором в Геттінгені і директором Геттінгенської обсерваторії. Цю посаду він обіймав до самої смерті.

У 1805 році Гаус женився на Йоганні Остгоф. У них було троє дітей, вижили двоє - син Йозеф і дочка Мінна. Якраз в четверту річницю весілля померла Йоганна, незабаром після народження третьої дитини. Це найважчі роки для Гауса. У наступному 1810 році він одружився знову - на Вільгельміні («Мінне») Вальдек, подрузі Йоганни. Число дітей Гауса незабаром збільшилася до п'яти.

Гаус був найвидатнішим математиком свого часу, однак тут згадаємо його вклад в астрономію і фізику. У 1811 році з'явилася нова комета. Гаус швидко і дуже точно розрахував її орбіту. Знамениту комету усюди спостерігали, користуючись обчисленнями Гауса. В 1820 році Гаусу доручають зробити геодезичну зйомку Ганновера. Для цього він розробив відповідні обчислювальні методи (в тому числі методу практичного застосування свого методу найменших квадратів) і організував зйомку місцевості і складання карт.

У чудовій роботі «Про один новий загальний закон механіки» (1829 рік), що складається всього з чотирьох сторінок, Гаус обґрунтовує новий варіаційний принцип механіки - принцип найменшої дії. В 1831 році померла друга дружина, у Гауса почалося важке безсоння. У Геттінген приїхав запрошений з ініціативи Гауса 27-річний талановитий фізик Вільгельм Вебер, з яким Гаус познайомився в 1828 році, в гостях у Гумбольдта. Обидва ентузіасти науки здружилися, незважаючи на різницю у віці, і починають цикл досліджень електромагнетизму.

Гаус винаходить електричний телеграф (1833 рік) і (разом з Вебером) будує його діючу модель. У 1839 році Гаус в творі «Загальна теорія сил тяжіння і відштовхування, що діють обернено пропорційно квадрату відстані» виклав основи теорії потенціалу, включаючи ряд основних положень і теорем - наприклад, основну теорему електростатики (теорема Гауса). У роботі «Діоптричні дослідження» (1840 рік) Гаус розробив теорію побудови зображень в складних оптичних системах. У фізиці Гаус розвинув теорію капілярності, теорію системи лінз. Заклав основи математичної теорії електромагнетизму і при цьому першим ввів поняття потенціалу електричного поля, а в 1845 р прийшов до думки про кінцеву швидкість поширення електромагнітних взаємодій. У 1832 р створив абсолютну систему одиниць, ввівши три основні одиниці: одиницю довжини, одиницю часу, одиницю маси; ця система стала прообразом системи одиниць СГС, у якій одиниця виміру магнітної індукції дістала назву Гаус (позначення Гс, міжнародне - G).

Для мінімізації впливу помилок вимірювання Гаус використовував свій метод найменших квадратів, який зараз повсюдно застосовується у статистиці. Хоча Гаус не перший відкрив поширений в природі нормальний закон розподілу, але він настільки ретельно його досліджував, що графік розподілу з тих пір часто називають гаусівським.

Помер Гаус 23 лютого 1855 року в Геттінгені. Король Ганновера Георг V наказав викарбувати на честь Гауса медаль, на якій були викарбувані портрет Гауса і почесний титул, його портрет і винайдений ним вимірювальний інструмент «геліотрон» зображені на банкноті в 10 марок, що вийшла з обігу (після введення євро), але представляє інтерес для боністів.



Параметри розподілу - постійні, від яких залежить функція розподілу.

Отже, якщо відомий вид функції розподілу, то для того, щоб однозначно охарактеризувати випадкову величину, досить задати лише параметри її розподілу.

Найважливішими параметрами розподілу, які задають випадкову величину X , є її *математичне очікування* M_x (характеризує центр розсіювання) і *дисперсія* σ_x (характеризує ступінь розсіювання).

Математичне очікування M_x - середнє зважене по ймовірностям значення випадкової величини .

Для дискретної випадкової величини математичне очікування визначається виразом

$$M_x = \sum_{i=1}^n x_i p_i , \quad (2.13)$$

де x_i - значення дискретної випадкової величини, а $p_i = P(X=x_i)$.

Для неперервної випадкової величини математичне сподівання визначається інтегралом

$$M_x = \int_{-\infty}^{+\infty} x f(x) dx , \quad (2.14)$$

де $f(x)$ - густина розподілу неперервної випадкової величини.

Можна відзначити, що геометричний зміст математичного очікування випадкової величини - це абсциса центру ваги фігури під кривою густини розподілу x . Сказане проілюструємо на рис. 2.2, де видно, що добуток $f(x)dx$ є площа елементарної ділянки під кривою $f(x)$, а x - абсциса цієї ділянки, тобто відстань від початку координат.

Отже, інтеграл (2.14) дає абсцису центру ваги всієї площі фігури під кривою $f(x)$.

Крім математичного очікування центр розсіювання випадкової величини можна ще охарактеризувати такими параметрами її розподілу, як *мода* і *медіана*.

Мода M_o - значення випадкової величини, відповідне локальному максимуму густини ймовірностей для неперервної випадкової величини або локальному максимуму ймовірності для дискретної випадкової величини.

Медіана M_e - значення випадкової величини, для якого інтегральна функція розподілу приймає значення 0,5, або має місце «стрибок» зі значення, меншого ніж 0,5, до значення, більшого ніж 0,5.

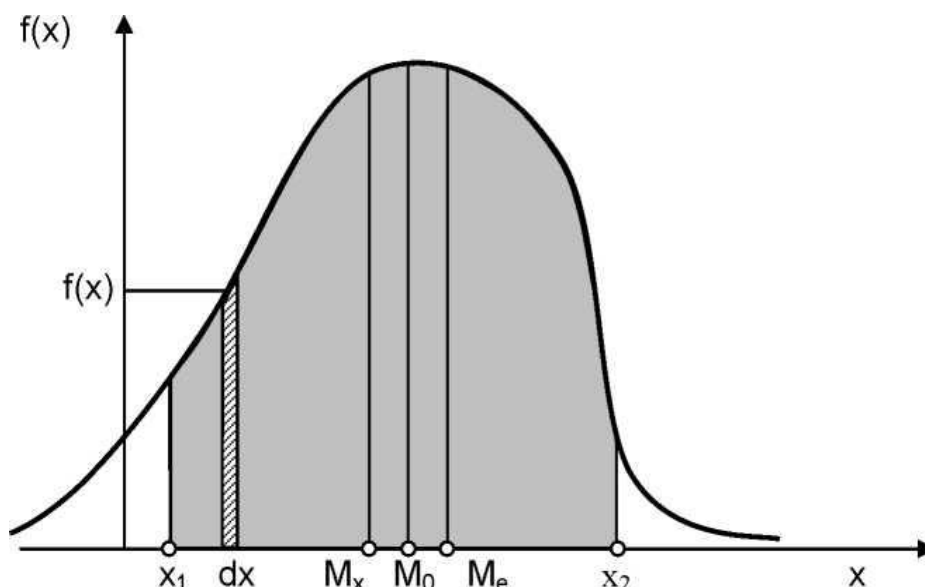


Рис. 2.2. Диференціальна функція розподілу – густина розподілу $f(x)$.

Таким чином, для диференціального закону розподілу медіана є таке значення неперервної випадкової величини X , яка ділить пополам площу під кривою густини розподілу $f(x)$.

Дисперсія випадкової величини $D = \sigma_x^2$ - це математичне очікування випадкової величини $(x - M_x)^2$ (другий центральний момент).

Для дискретної випадкової величини дисперсія визначається таким математичним виразом:

$$D = \sigma_x^2 = \sum_{i=1}^n (x_i - M_x)^2 p_i. \quad (2.15)$$

Для неперервної випадкової величини дисперсія визначається виразом

$$D = \sigma_x^2 = \int_{-\infty}^{+\infty} (x - M_x)^2 f(x) dx \quad (2.15)$$

де x - значення неперервної випадкової величини X ; $f(x)$ - густина розподілу; M_x - математичне очікування.

Дисперсія має розмірність квадрата одиниці вимірювання випадкової величини, а позитивне значення квадратного кореня з дисперсії називається середньоквадратичним відхиленням S .

Середньоквадратичне відхилення σ_x - невід'ємний квадратний корінь з дисперсії D .

На закінчення цього розділу дамо визначення ще одного параметра розподілу випадкової величини, який носить назву *квантиль*.

Квантиль порядку P - значення x_p випадкової величини, для якого функція розподілу приймає значення P або має місце «стрибок» зі значення, меншого ніж P , до значення, більшого ніж P :

$$P(x_p) = P. \quad (2.16)$$

З цього визначення квантиля слід, що медіана Me - це квантиль порядку 0,5, тобто $Me = x_{0,5}$.

Ймовірність попадання випадкової величини X в інтервал $[X_{P1}, X_{P2}]$ дорівнює

$$P(x_{P1} < X < x_{P2}) = P(X < x_{P2}) - P(X < x_{P1}) = F(x_{P2}) - F(x_{P1}) = P_2 - P_1. \quad (2.17)$$

Для опису розподілу випадкових величин навколо математичного очікування зручно використовувати такі характеристики їх диференціальної функції розподілу як *асиметрія* та *ексцес*. Обчислення асиметрії дозволяє встановити симетричність розподілу випадкової величини. Для цього знаходять третій центральний момент, що характеризує асиметрію закону розподілу випадкової величини. Якщо він дорівнює нулю, то випадкова величина симетрично розподілена відносно математичного очікування. Оскільки має розмірність випадкової величини в кубі, то вводять безрозмірну величину - *коефіцієнт асиметрії*:

$$As = \frac{\mu_3}{\sigma^3} = \frac{1}{\sigma^3} \int_{-\infty}^{+\infty} (x - M_x)^3 f(x) dx. \quad (2.18)$$

Центральний момент четвертого порядку використовується для визначення *ексцесу* і характеризує плосковершинність або гостровершинність функції густини розподілу ймовірності. *Ексцес* обчислюється за формулою:

$$Es = \frac{\mu_4}{\sigma^4} - 3 = \frac{1}{\sigma^4} \int_{-\infty}^{+\infty} (x - M_x)^4 f(x) dx - 3. \quad (2.19)$$

Число 3 віднімається для порівняння відхилення від центрального закону розподілу (нормального закону), для якого виконується рівність $\mu_4/\sigma^4 = 3$.

Отже, для нормального закону розподілу $E_s=0$. Якщо ексцес позитивний $E_s \geq 0$ то на графіку функції розподілу виділяється гостра вершина, а для негативних значень $E_s \leq 0$ максимум більш пологий. Таким чином можна встановити відхилення заданого закону від нормального. Для наочності при різних значеннях асиметрії і ексцесу графіки розподілу густини ймовірностей наведені на рис 2.3.

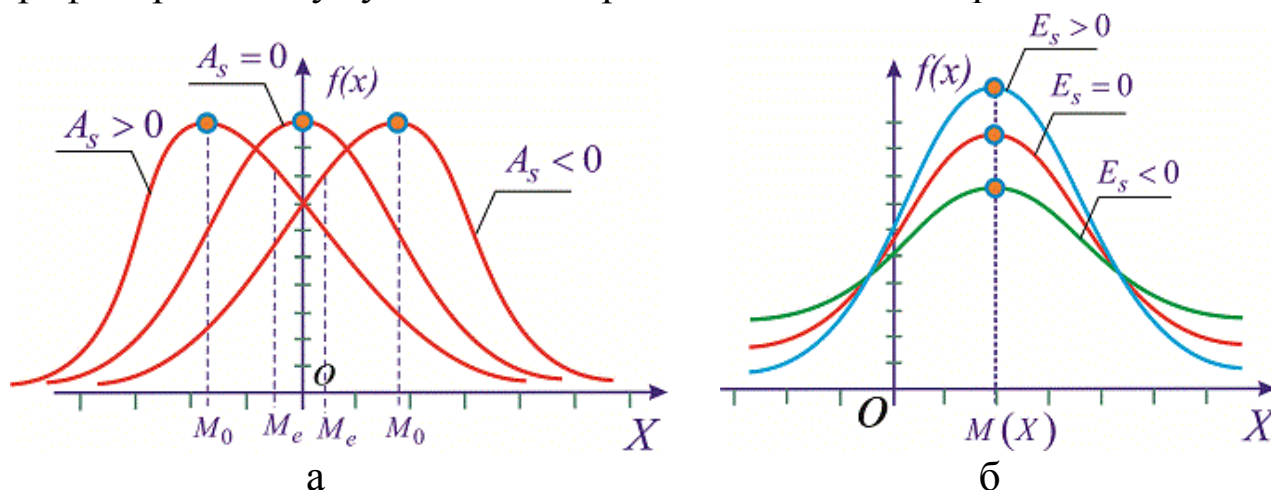


Рис. 2.3. Графіки функцій розподілу густини ймовірності із різними коефіцієнтами асиметрії (а) і ексцесу (б).

2.2. Нормальний закон розподілу

Інтегральна функція розподілу $F(x)$ і відповідна їй густина розподілу $f(x)$ представляють собою деяку математичну модель властивостей досліджуваної випадкової величини (відгуку), значення якої реєструються в ході експерименту. Тому одним з основних завдань статистичної обробки дослідних даних є знаходження таких функцій розподілу, які, з одного боку, досить добре описували б спостережувані значення випадкової величини, а з іншого - були б зручні для подальшого статистичного аналізу. При цьому вид функції розподілу бажано вибирати на основі уявлень про фізичну природу даного явища, тому що в цьому випадку виключаються можливі похибки при поширенні знайдених закономірностей за межі досліджуваного в експерименті інтервалу варіювання (зміни) випадкової величини (відгуку).

При обробці експериментальних даних дослідники найчастіше оперують з випадковими величинами, які мають так званий нормальний (Гаусовий) розподіл (рис.2.4). Не вдаючись в детальні математичні викладки, відзначимо, що, згідно з центральною граничною теоремою математичної статистики, «за певних умов

розподіл нормованої суми n незалежних випадкових величин, розподілених за довільним законом, прагне до нормального, коли n прямує до нескінченності». Необхідні умови, при яких ця теорема виявляється справедливою, полягають у тому, що різні випадкові величини повинні мати обмежені дисперсії і дисперсія будь-якої випадкової величини не повинна бути занадто великою у порівнянні з дисперсіями інших.

При обробці експериментальних даних ця теорема має дуже велике значення, оскільки відгук стає випадковою величиною в результаті впливу неконтрольованих факторів, число яких швидше за все є дуже великим. Крім того, якщо при проведенні дослідів всі найважливіші чинники контролюються, то вплив на відгук кожного з неконтрольованих факторів не повинен бути великим у порівнянні з іншими неконтрольованими факторами. Іншими словами, та дисперсія (розсіювання) відгуку, яку викликає будь-який з неконтрольованих факторів, не повинен сильно відрізнятися від дисперсій, пов'язаних з впливом інших неконтрольованих чинників. В іншому випадку фактор, дисперсія від якого істотно відрізняється від інших, обов'язково повинен бути переведений в розряд контрольованих.

Отже, якщо при плануванні експерименту враховані всі найбільш істотні фактори і потім, при проведенні дослідів, вони контролюються, то при обробці експериментальних даних можна припускати, що відгук не повинен суперечити нормальному розподілу.

Більшість інших розподілів, які використовуються в математичних статистиках (Стюдента, Фішера, Пірсона, Кохрена, а також розподіли, за якими складені різні критеріальні таблиці), отримані на основі нормального розподілу.

Не можна, однак, абсолютизувати значення нормального розподілу. Не всі випадкові величини розподілені за нормальним законом. Проте на практиці, якщо явище схильне до дії багатьох випадкових факторів, їх сумарний вплив цілком виправдано можна описати за допомогою нормального.

Як вже було зазначено, для випадкової величини, що не суперечить нормальному закону, інтегральна функція розподілу (2.12) і відповідна їй густина розподілу

$$f(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp\left(-\frac{(x - M_x)^2}{2\sigma_x^2}\right). \quad (2.12)$$

визначаються двома параметрами: M_x - математичним очікуванням і σ_x^2 - дисперсією.

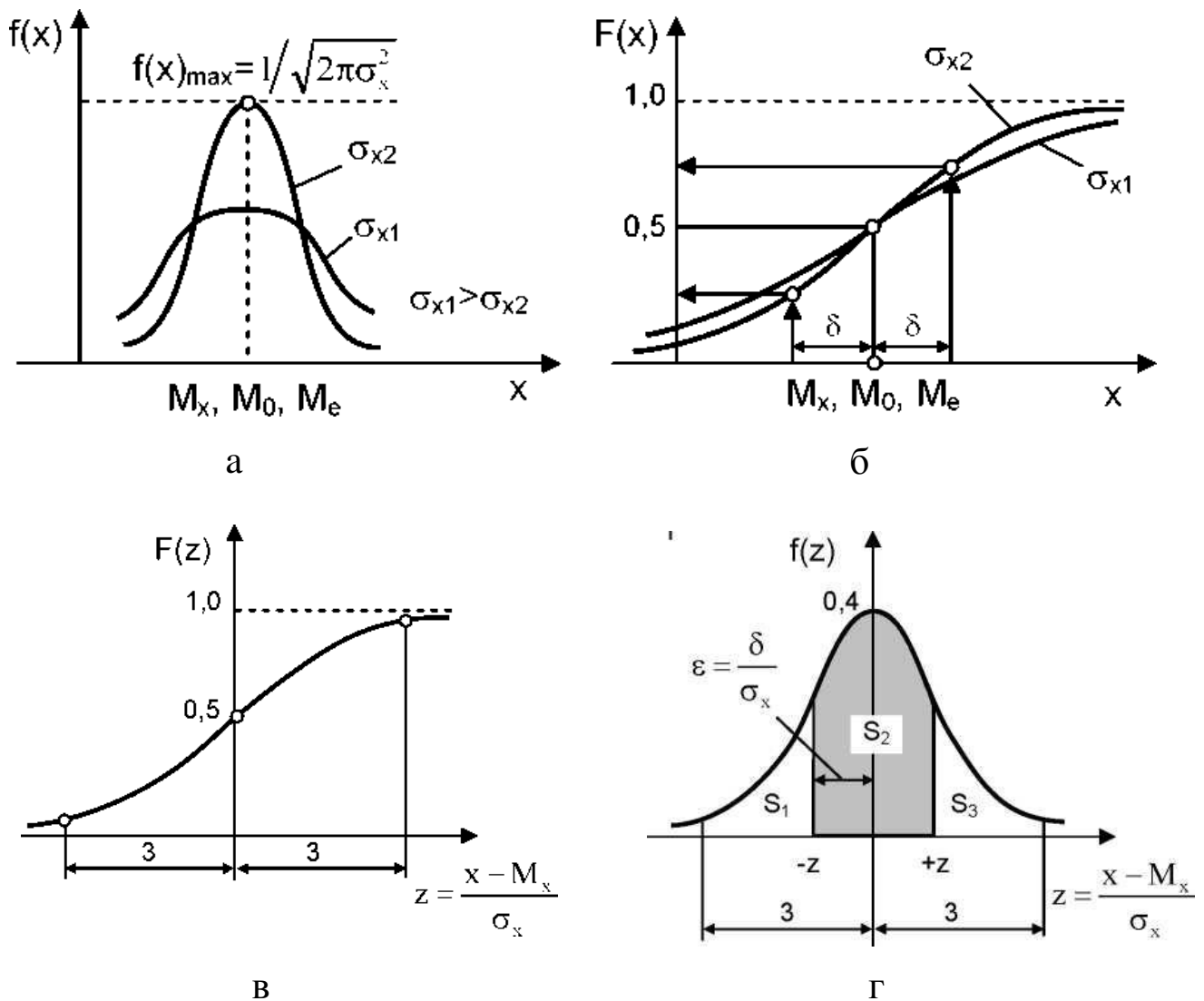


Рис.2.4. Густина розподілу (а, г) і інтегральна функція розподілу (б,в) при нормальному законі розподілу випадкових величин.

Відзначимо деякі властивості нормального закону розподілу (рис.2.4):

1. Крива густини розподілу симетрична відносно значення M_x , яке іноді називають центром розподілу.

2. При великих значеннях σ_x^2 крива $f(x)$ більш полого, тобто σ_x^2 є мірою величини розсіювання значення випадкової величини навколо значень M_x . При зменшенні параметра σ_x^2 крива нормального розподілу стискується уздовж осі Ox і витягується уздовж $f(x)$.

3. Максимум ординати кривої густини розподілу визначається виразом

$$f_{\max}(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}}, \quad (2.20)$$

що при $\sigma_x^2=1$ відповідає значенню приблизно 0,4.

Для нормального розподілу математичне сподівання, мода і медіана збігаються:

$$M_x = M_0 = M_e. \quad (2.21)$$

У ряді випадків розглядається не сама випадкова величина x , а її відхилення від математичного очікування $y = x - M_x$. Така випадкова величина y називається *центрованою*.

Відношення випадкової величини x до її середньоквадратичного відхилення називається *нормованою* випадковою величиною.

Таким чином, *центрована* випадкова величина рівна різниці між даною випадковою величиною і її математичним очікуванням, а *нормована* випадкова величина - відношення даної випадкової величини до її середнього квадратичного відхилення.

Очевидно, що математичне очікування центрованої випадкової величини дорівнює нулю, $M_y = 0$, а дисперсія нормованої випадкової величини дорівнює одиниці, $\sigma_y^2 = 1$.

Приведена випадкова величина - центрована і нормована випадкова величина

$$z = \frac{x - M_y}{\sigma_y}. \quad (2.22)$$

Математичне очікування і дисперсія приведеної випадкової величини дорівнюють відповідно нулю, $M_z = 0$, і одиниці, $\sigma_z^2 = 1$.

Нормальний *розподіл* з параметрами $M_z = 0$ і $\sigma_z^2 = 1$ називається *стандартним* (нормованим).

Для приведеної випадкової величини *нормальний стандартний інтегральний розподіл* набуває вид

$$F(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z \exp\left(-\frac{z^2}{2}\right) dz. \quad (2.23)$$

Графіки цих функцій показані на рис. 2.4 в, г, причому очевидно, що

$$F(-z) = 1 - F(z). \quad (2.24)$$

Покажемо справедливність співвідношення (2.24). Розглянемо графік густини стандартного нормального розподілу (див. рис. 2.4, г).

Позначимо площу під ним лівіше точки $-z$ через S_1 ; площу між $-z$ і $+z$ - через S_2 , площа що залишилася (правіше $+z$) - через S_3 . Тоді, поперше, з симетричності графіка густини розподілу слід, що $S_1=S_3$. По-друге, $S_1+S_2+S_3=1$ або $S_1+(S_1+S_2)=1$ (вся площа під графіком густини ймовірності дорівнює одиниці). За змістом функції розподілу $S_1=F(-z)$, $S_1+S_2=F(+z)$. Отже, $F(-z)+F(z)=1$, звідки і випливає рівність (2.24).

Значення нормованої функції (2.23) нормального розподілу (функції Лапласа) і значення густини нормованого нормального розподілу (2.22) табульовані і приведені в різних довідниках з математичної статистики. У списку статистичних функцій електронних таблиць Microsoft Excel їм відповідають НОРМРАСП (X; 0; 1; ИСТИНА) або НОРМСТРАСП (z) - для (2.22) і НОРМРАСП (x; 0; 1; ЛОЖЬ) - для (2.23).

Геометрично функція Лапласа представляє собою площу під кривою $f(z)$ в інтервалі від $-\infty$ до деякої конкретної величини z .

Відмітимо, що іноді замість функції $F(z)$ табулюється функція $F_0(z)$

$$F_0(z) = \frac{1}{\sqrt{2\pi}} \int_0^z \exp\left(-\frac{z^2}{2}\right) dz. \quad (2.24)$$

рівна площі під графіком стандартного нормального розподілу від 0 до z (див. рис.2.4, г).

В силу симетрії функції розподілу густини ймовірності відносно осі ординат

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^0 \exp\left(-\frac{z^2}{2}\right) dz = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} \exp\left(-\frac{z^2}{2}\right) dz = \frac{1}{2}, \quad (2.24)$$

тому $F(z)=0,5+F_0(z)$. Функція $F_0(z)$ непарна: $F_0(-z)=-F_0(z)$

У ряді випадків необхідно знати яка ймовірність того, що значення випадкової величини буде *відрізнятись* від математичного очікування *не більше ніж на одне середньоквадратичне відхилення*. Для цього з таблиць для стандартної функції розподілу (функції Лапласа) знаходимо $F(1)=0,84$, отже $F_0(1)=F(1)-0,5=0,84-0,5=0,34$, тобто ймовірність попадання випадкової величини в інтервал від

математичного очікування M_x до $M_x + \sigma$ становить 0,34. Тоді, враховуючи симетричність функції розподілу густини ймовірності, ймовірність попадання випадкової величини в інтервал від $(M_x - \sigma)$ до $(M_x + \sigma)$ $P(M_x - \sigma \leq x \leq M_x + \sigma)$ буде становити $0,34 + 0,34 = 0,68$.

Аналогічно для інтервалу $\pm 2\sigma$ знаходимо

$$P(M_x - 2\sigma \leq x \leq M_x + 2\sigma) = 0,95.$$

Для інтервалу $\pm 3\sigma$

$$P(M_x - 3\sigma \leq x \leq M_x + 3\sigma) = 0,997.$$

Останній вираз дістав у математичній статистиці назву *правила трьох сігм*, згідно якого ймовірність відхилення випадкової величини, яка розподілена за законом Гауса, від математичного очікування більше ніж на $\pm 3\sigma$ складає не більше ніж 0,3%.

Контрольні питання

1. Охарактеризуйте основні поняття теорії ймовірностей і математичної статистики.
2. Проаналізуйте для чого потрібні знання елементів теорії ймовірностей і математичної статистики для експериментальних досліджень.
3. Зробіть аналіз взаємозв'язку диференціальної та інтегральної функції розподілу.
4. Який зміст мають центр, мода і медіана розсіювання випадкової величини?
5. Який зміст мають математичне очікування і дисперсія розсіювання випадкової величини?
6. Що характеризують асиметрія та ексцес диференціальної функції розподілу випадкових величин?
7. Охарактеризуйте основні властивості нормального розподілу.
8. Для чого вводяться поняття нормованої, центрованої та приведеної випадкової величини?
9. Виконайте аналіз ймовірності відхилення значення випадкової величини від математичного очікування на одне, два, три середньоквадратичне відхилення.

3. ПОПЕРЕДНЯ ОБРОБКА ЕКСПЕРИМЕНТАЛЬНИХ ДАНИХ

Попередня обробка результатів вимірювань і спостережень необхідна для того, щоб в подальшому, при побудові емпіричних залежностей (функцій відгуку), з найбільшою ефективністю використовувати статистичні методи і коректно аналізувати отримані результати.

Зміст попередньої обробки полягає в оцінці дійсного значення шуканої величини і її дисперсії, відсіювання грубих похибок і, таким чином, оцінки достовірності результатів вимірювань. Іншими важливими моментами попередньої обробки даних є перевірка відповідності результатів вимірювання нормальному закону розподілу і, отже, визначення параметрів цього розподілу. Якщо гіпотеза про те, що відгук не суперечить нормальному розподілу, виявиться неприйнятною, то слід визначити, яким законом розподілу підкоряються дослідні дані або, якщо це можливо, перетворити досліджуваний розподіл до нормального.

3.1. Обчислення параметрів емпіричних розподілів.

Точкове оцінювання

Розгляд питань обробки експериментальних даних почнемо з найпростішої ситуації, коли відгук реєструється при фіксованих рівнях всіх контрольованих факторів і при проведенні дослідів (в результаті впливу неконтрольованих факторів) дослідник отримує хоча і близькі, але відмінні один від одного результати.

Приклад 3.1. При дослідженні твердих домішок у воді після її фільтрації на очисних спорудах підприємства були отримані наступні три значення: 351, 370 і 365 мг/л.

Спробуємо знайти відповідь на питання - чому дорівнює концентрація твердих домішок у воді після її очищення?

На перший погляд відповідь на поставлене запитання не викликає ніяких особливих проблем, і початківці дослідники, не дуже досвідчені в області теорії ймовірностей і математичної статистики, швидше за все дадуть відповідь, що концентрація домішок (X_1 - перший варіант відповіді) становить

$$X_1 = (351+370+365)/3 = 362,00 \text{ (мг/л) ,}$$

тобто буде знайдено середнє арифметичне (вибіркове середнє арифметичне) з трьох отриманих значень відгуку.

Однак дослідні дані можна усереднювати і іншими способами. Наприклад, можна підрахувати середнє геометричне (X2 - другий варіант відповіді):

$$X2=(351 \cdot 370 \cdot 365)^{1/3}=361,91 \text{ (мг/л)},$$

або знайти середнє, тільки між мінімальним (351) і максимальним (370) значеннями - так звану середину розмаху (X3 - третій варіант відповіді):

$$X3=(351+370)/2=360,50 \text{ (мг/л)},$$

або, розташувавши всі значення в зростаючій послідовності 351, 365, 370, взяти середній член отриманого ряду - середній член варіаційного ряду (X4 - четвертий варіант відповіді):

$$X4=365,00 \text{ (мг/л)}.$$

Можна придумати і будь-які інші способи (наприклад, дуже «оригінальною» може бути ідея ще раз усереднити всі чотири отриманих значення), проте зупинимося поки тільки на цих чотирьох варіантах відповіді на поставлене перед нами питання. Ми бачимо, що, не застосовуючи ніяких додаткових міркувань, нам поки досить важко обґрунтувати той чи інший варіант, на якому було б бажано зупинитися. Так, якщо вибирати ту відповідь, яка зажадає від нас найменшу кількість обчислень (зусиль), то тоді краще всього віддати перевагу значенню $X4=365,00$ мг/л (взагалі не вимагає ніяких розрахунків). Однак подібне обґрунтування навряд чи можна вважати досить надійним і переконливим. Тому давайте зупинимося і задумаємося про те, чому взагалі ми зіткнулися з подібною ситуацією. Адже якби, наприклад, нам потрібно було знайти відповідь на питання, яка кількість фільтрів використовується при очищенні ми при довільному числі спостережень за технологічним процесом очистки отримали б абсолютно однакові значення (припустимо, п'ять фільтрів). У подібній ситуації немає необхідності обчислювати ні вибіркоче середнє, ні середнє геометричне, ні середину розмаху, ні знаходити середній член варіаційного ряду і т. д., оскільки можна відразу вказати ту кількість фільтрів, які використовуються у процесі очищення.

Отже, між такими величинами, як число фільтрів і концентрація твердих домішок, є принципова різниця, яка полягає в тому, що перша з двох названих величин є детермінованою, а друга - випадковою. І якщо для того, щоб описати детерміновану величину, досить вказати одне її значення (наприклад, число фільтрів дорівнює п'яти), то для опису випадкової величини потрібно знати її розподіл. Іншими словами, для випадкової величини недостатньо вказати лише якесь її

значення (або комбінацію її значень, як, наприклад, вибіркове середнє арифметичне), а потрібно записати функцію, яка однозначно визначає ймовірність того, що випадкова величина приймає задане значення або належить до деякого заданого інтервалу.

Тому відповідь на питання прикладу 3.1 треба починати не з пошуку будь-яких варіантів усереднення дослідних даних, а перш за все з констатації того факту, що концентрація домішок - це випадкова величина!

Далі потрібно зазначити, що концентрація домішок - це безперервна випадкова величина, оскільки вона може приймати будь-які значення з обмеженого інтервалу і згідно нормативним документам гранично допустиме значення становить $X_{\text{МАКС}} \leq 380$ мг/л.

Після цього можна висунути гіпотезу (припущення), що така випадкова величина, як концентрація домішок не повинна суперечити нормальному закону розподілу. Згідно центральної граничної теореми математичної статистики, цю гіпотезу швидше за все можна буде прийняти в якості робочої, оскільки дослідні дані в прикладі 3.1 отримані при вимірюванні забруднення в різних точках басейна очищеної води, а найбільш суттєві фактори, які визначають забруднення зафіксовані при одних і тих же рівнях зовнішніх факторів. Крім того, відгук (концентрація домішок) стає випадковою величиною тільки в результаті впливу малозначимих неконтрольованих факторів, число яких на різних етапах виробничого циклу, очевидно, прямує до нескінченності.

Отже, в якості відповіді на питання прикладу 3.1 ми можемо сказати, що концентрація домішок - це безперервна випадкова величина, функцію розподілу якої швидше за все можна записати у виді розподілу Гауса:

$$f(X) = \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp\left(-\frac{(x-M)^2}{\sigma_x^2}\right). \quad (3.1)$$

Здавалося б, у нас з'явилася можливість визначати ймовірність того, що концентрація домішок належить до деякого заданого інтервалу (наприклад, $X \leq 380$), тобто відповідає першому етапу очищення води. Однак для того, щоб записати функцію нормального розподілу, необхідно визначити математичне очікування M і дисперсію σ_x^2 , які є характеристиками *генеральної сукупності*, тобто множини всіх можливих значень забруднення.

Іншими словами, *генеральна сукупність* - це таке уявне, нескінченно велике число результатів спостережень при необмеженому числі повторень одного і того ж комплексу зовнішніх умов (факторів).

У прикладі 3.1 під генеральною сукупністю можна розуміти, всі ділянки басейну, в яких в принципі можна було б заміряти забруднення, або взагалі забруднення води за весь час експлуатації очисних споруд.

У розпорядженні дослідника, звичайно ж, ніколи немає генеральної сукупності, і він може вивчати тільки її частину - *вибірку*, причому завжди обмеженого об'єму.

Вибірка - будь-яка обмежена підмножина генеральної сукупності, призначена для безпосередніх досліджень.

Об'єм вибірки - кількість одиниць у вибірці.

За вибіркою неможливо однозначно визначити ні інтегральну функцію розподілу, ні густину розподілу, ні параметри розподілу (наприклад, математичне очікування або дисперсію) випадкової величини, оскільки для цього потрібно необмежена (нескінченно велика) кількість результатів спостережень, тобто необхідно досліджувати всю генеральну сукупність.

Отже, маючи обмежену підмножину генеральної сукупності (*вибірку*), ми повинні або взагалі відмовитися від пошуку розподілу випадкової величини, або задовольнятися лише деякими наближеними значеннями невідомих параметрів її розподілу, тобто провести *оцінювання* випадкової величини.

Оцінювання - визначення наближеного значення невідомого параметра генеральної сукупності за результатами обмеженої кількості спостережень.

Ідея оцінювання повинна бути цілком зрозуміла з міркувань звичайної життєвої практики. Адже для того, щоб, наприклад, купити пару кілограм яблук, у нас ніколи не виникає бажання поспробувати всі наявні у даного продавця фрукти (вивчити всю генеральну сукупність), ми пробуємо часточку тільки одного яблука (досліджуємо *вибірку*), визначаємо її смак (оцінюємо) і приймаємо рішення, варто нам чи ні купувати саме ці яблука.

Вихідними даними при оцінюванні, як і при перевірці будь-яких припущень (статистичних гіпотез), що стосуються невідомого розподілу випадкової величини, звичайно ж, можуть бути лише тільки ті результати спостережень, які були отримані в ході проведення дослідів (вбірка обмеженого об'єму).

Для одного і того ж параметра розподілу може бути запропоновано декілька оцінок. У прикладі 3.1 розглядалося чотири різних оцінки для такого параметра розподілу концентрації домішок, як математичне очікування даної випадкової величини (вбіркове середнє арифметичне, вбіркове середнє геометричне, середина розмаху і середній член варіаційного ряду). Тому при оцінюванні завжди виникає проблема вибору найкращої оцінки з усіх можливих оцінок даного параметра. Причому, коли формулюються ті чи інші вимоги, за якими оцінку доцільно вважати найкращою, перш за все враховується той факт, що будь-яка оцінка - це також випадкова величина. Адже якби в умовах прикладу 3.1 було б знайдено, припустимо, вбіркове середнє арифметичне концентрації домішок в інший день, звичайно ж, зовсім не обов'язково, що воно знову виявилось б одно саме 362,00 мг/л.

З тих міркувань, що будь-яка оцінка будь-якого параметра η розподілу випадкової величини теж є випадкова величина, до оцінок застосовуються вимоги *спроможності, незміщеності і ефективності*.

Спроможність оцінки означає, що вона сходиться по ймовірності до значення оцінюваного параметра η при безмежному зростанні об'єму вбірки.

Іншими словами, для спроможної оцінки відхилення її від η на малу величину стає малоімовірним при великому об'ємі вбірки. Цілком природно, що дослідників в першу чергу цікавлять ті оцінки, які хоча б в границі (при проведенні нескінченно великої кількості спостережень) давали їм можливість визначити параметр розподілу η . Однак слід зазначити, що на практиці доводиться оцінювати невідомі параметри і при малих об'ємах вбірки. Природною є вимога, при виконанні якої оцінка не дає систематичної похибки в бік завищення (або заниження) істинного значення параметра η .

Незміщена оцінка - оцінка, математичне очікування якої дорівнює значенню оцінюваного параметра. Вимога незміщеності дозволяє усунути систематичну похибку оцінки параметра, яка залежить від об'єму вбірки і в разі спроможності оцінки прагне до нуля при збільшенні об'єму вбірки.

Ефективна оцінка – незміщена оцінка, що має найменшу дисперсію з усіх можливих незміщених оцінок даного параметра.

Після того як дослідник вибрав і підрахував спроможну, незміщену і ефективну оцінку даного параметра розподілу досліджуваної випадкової величини, перше і найбільш просте, що він може зробити, так це прийняти значення оцінки як невідоме значення параметра розподілу, тобто виконати *точкове оцінювання*.

Точкове оцінювання - спосіб оцінювання, що полягає у тому, що значення оцінки приймають за невідоме істинне значення параметра розподілу.

Розглянемо деякі точкові оцінки основних параметрів розподілу для неперервної випадкової величини, що не суперечить нормальному закону розподілу.

Вибіркове середнє арифметичне \bar{x} - сума значень випадкової величини, отриманих за результатами випробування вибірки, поділена на її об'єм:

$$\bar{x} = \frac{1}{n}(x_1 + x_2 + \dots + x_n) = \frac{1}{n} \sum_{i=1}^n x_i, \quad (3.2)$$

де n – об'єм вибірки; x_i - результат вимірювання i -тої одиниці.

У математичній статистиці доведено, що вибіркове середнє арифметичне є найкращою (спроможною, незміщеною і ефективною) оцінкою математичного очікування випадкової величини, що підкоряється нормальному закону розподілу.

У прикладі 3.1, навіть якщо припустити, що концентрація домішок суперечить нормальному закону розподілу, з чотирьох отриманих оцінок перевагу слід віддати значенням $X_1 = (351 + 370 + 365) / 3 = 362,00$ (мг/л) (вибірковому середньому арифметичному) як найкращій оцінці для математичного очікування даної випадкової величини. Три інші розглянуті в цьому прикладі оцінки також є спроможні для математичного очікування. Однак середнє геометричне $X_2 = (351 \cdot 370 \cdot 365)^{1/3} = 361,91$ - це зміщена оцінка (вона буде найкращою тільки тоді, коли випадкова величина підпорядковується так званому логарифмічному нормальному розподілу, тобто коли законом Гауса описується не випадкова величина, а її логарифм). Середина розмаху $X_3 = (351 + 370) / 2 = 360,50$ і середній член варіаційного ряду $X_4 = 365,00$ - це хоча і незміщені оцінки для математичного сподівання, але їх ефективність, як показано в математичній статистиці, менша, ніж у вибіркового середнього арифметичного (менше одиниці).

Вибіркова дисперсія S^2 - сума квадратів відхилень результатів спостережень від їх вибіркового середнього арифметичного у вибірці, поділена на $n-1$ або на n

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2. \quad (3.3)$$

Оцінки S^2 є спроможними, незміщеними і, у випадку нормального розподілу, асимптотично ефективними оцінками дисперсії σ^2 .

Для практичних розрахунків вираз (3.3) можна перетворити до виду

$$\begin{aligned} S^2 &= \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i^2 - 2x_i\bar{x} + \bar{x}^2) = \frac{1}{n-1} \left(\sum_{i=1}^n x_i^2 - 2\bar{x} \sum_{i=1}^n x_i + n\bar{x}^2 \right) = \\ &= \frac{1}{n-1} \left(\sum_{i=1}^n x_i^2 - \frac{2n\bar{x}}{n} \sum_{i=1}^n x_i + n\bar{x}^2 \right) = \frac{1}{n-1} \left(\sum_{i=1}^n x_i^2 - 2n\bar{x} \cdot \bar{x} + n\bar{x}^2 \right) = \\ &= \frac{1}{n-1} \left(\sum_{i=1}^n x_i^2 - 2n\bar{x}^2 + n\bar{x}^2 \right) = \frac{1}{n-1} \left(\sum_{i=1}^n x_i^2 - n\bar{x}^2 \right) = \\ &= \frac{1}{n-1} \left(\sum_{i=1}^n x_i^2 - n \left(\frac{1}{n} \sum_{i=1}^n x_i \right)^2 \right) = \frac{1}{n-1} \left(\sum_{i=1}^n x_i^2 - \frac{1}{n} \left(\sum_{i=1}^n x_i \right)^2 \right) \end{aligned} \quad (3.4)$$

Приклад 3.2. Для даних прикладу 3.1 вибіркова дисперсія концентрації твердих частинок в очищеній воді

$$S^2 = \frac{1}{3-1} \cdot \left[(351)^2 + (370)^2 + (365)^2 - \frac{1}{3} (351+370+365)^2 \right] = 97,00 \text{ (мг}^2\text{/л}^2\text{)},$$

а вибіркове середньоквадратичне відхилення $S=(S^2)^{1/2}=9,85$ (мг/л).

Знаючи вибіркове середнє арифметичне \bar{x} і вибіркове середнє квадратичне відхилення S^2 можна підрахувати міру відносної мінливості випадкової величини x – вибірковий коефіцієнт варіації за формулою

$$\mu = \frac{S}{\bar{x}}. \quad (3.5)$$

Для нашого прикладу 3.1 $\mu=9,85/362=0,027$, або у відсотках 2,7%.

Таблиця 3.1. Нормована інтегральна функція нормального розподілу

$$\text{(Функція Лапласа } \Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{u^2}{2}} du \text{)}$$

Z	Сотые доли Z									
	0	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5598	0,5638	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990

Через вибіркове середнє арифметичне \bar{x} і вибіркове середнє квадратичне відхилення S^2 можуть бути зроблені точкові оцінки для довільних значень інтегральної функції розподілу, а також для ймовірності потрапляння випадкової величини в будь-який з заданих інтервалів.

Так, для будь якого значення x інтегральна функція нормального розподілу

$$F(x) = P(X \leq x) = P\left(\frac{X - M}{\sigma} \leq \frac{x - M}{\sigma}\right) = P\left(z \leq \frac{x - M}{\sigma}\right) = \Phi\left(\frac{x - M}{\sigma}\right). \quad (3.6)$$

У якості точкової оцінки функції $F(x)$ можна прийняти

$$\tilde{F}(x) = \Phi\left(\frac{x - \bar{x}}{S}\right). \quad (3.7)$$

Приклад 3.3. Для оцінки ймовірності того, що концентрація твердих нерозчинених частинок у воді не перевищує гранично допустимі значення обчислимо центровану нормовану (приведену) величину $z = (380 - 362) / 9,85 = 1,827$, для якої значення нормованої функції нормального розподілу (функції Лапласа) (див. табл. 3.1.) становить 0,9664.

Таким чином в 96,5 % очищена вода відповідає встановленим вимогам однак ймовірно, що 3-4 випадках із 100 концентрація домішок перевищуватиме допустимі норми. Якщо ж необхідно оцінити квантиль порядку 0,99, то за формулою $x_p = \bar{x} + z_p S$ можна одержати таке значення

$$X(0,99) = 362 + z(0,99) \cdot 9,85 = 362 + 2,33 \cdot 9,85 = 385$$

де $z(0,99)$ - квантиль нормованого нормального розподілу порядку 0,99 - можна знайти за таблицями (див табл. 3.2), або в Microsoft Excel з використанням функції $\text{НОРМСТОБР}(0,99) = 2,326342$, а також $\text{НОРМОБР}(0,99; 0; 1) = 2,326342$.

Таким чином, якщо при вимірюванні концентрації твердих нерозчинених частинок в очищеній воді одержані значення 351, 370 і 365 мг/л, то скоріше всього тільки в одному випадку із 100 концентрація домішок перевищує 385 мг/л.

3.2. Оцінювання за допомогою довірчого інтервалу.

За значенням точкової оцінки не представляється можливим визначити, хоча б в якому діапазоні знаходиться оцінюваний нею параметр. Цей істотний недолік точкового оцінювання може бути компенсований оцінюванням за допомогою так званого довірчого інтервалу. На відміну від точкової оцінки, інтервальна оцінка дозволяє отримати вірогідну характеристику точності оцінювання невідомого параметра.

Таблиця 3.2. Квантилі z_P нормованого нормального розподілу

$$\text{порядку } P = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{u^2}{2}} du$$

1	2	3	4	5	6	7	8	9	10	11
0,80	0,842	0,845	0,849	0,852	0,856	0,860	0,863	0,867	0,871	0,874
0,81	0,878	0,882	0,885	0,889	0,893	0,896	0,900	0,904	0,908	0,912
0,82	0,915	0,919	0,923	0,927	0,931	0,935	0,938	0,942	0,946	0,950
0,83	0,954	0,958	0,962	0,966	0,970	0,974	0,978	0,982	0,986	0,990
0,84	0,994	0,999	1,003	1,007	1,011	1,015	1,019	1,024	1,028	1,032
0,85	1,036	1,041	1,045	1,049	1,054	1,058	1,063	1,067	1,071	1,076
0,86	1,080	1,085	1,089	1,094	1,098	1,103	1,108	1,112	1,117	1,122
0,87	1,126	1,131	1,136	1,141	1,146	1,150	1,155	1,160	1,165	1,170
0,88	1,175	1,180	1,185	1,190	1,195	1,200	1,206	1,211	1,216	1,221
0,89	1,227	1,232	1,237	1,243	1,248	1,254	1,259	1,265	1,270	1,276
0,90	1,282	1,287	1,293	1,299	1,305	1,311	1,317	1,323	1,329	1,335
0,91	1,341	1,347	1,353	1,359	1,366	1,372	1,379	1,385	1,392	1,398
0,92	1,405	1,412	1,419	1,426	1,433	1,440	1,447	1,454	1,461	1,468
0,93	1,476	1,483	1,491	1,499	1,506	1,514	1,522	1,530	1,538	1,546
0,94	1,555	1,563	1,572	1,580	1,589	1,598	1,607	1,616	1,626	1,635
0,95	1,645	1,655	1,665	1,675	1,685	1,695	1,706	1,717	1,728	1,739
0,96	1,751	1,762	1,774	1,787	1,799	1,812	1,825	1,838	1,852	1,866
0,97	1,881	1,896	1,911	1,927	1,943	1,960	1,977	1,995	2,014	2,034
0,98	2,054	2,075	2,097	2,120	2,144	2,170	2,197	2,226	2,257	2,290
0,99	2,326	2,366	2,409	2,457	2,512	2,576	2,652	2,748	2,878	3,090

Ідея оцінювання за допомогою довірчого інтервалу полягає в тому, щоб в околиці точкової оцінки спробувати побудувати такий інтервал (довірчий інтервал), який з деякою, відмінною від нуля, вірогідністю (довірчою ймовірністю) накрив би оцінюваний параметр розподілу.

3.2.1. Розрахунок довірчого інтервалу для математичного очікування. Розподіл Стюдента

Величина \bar{x} , знайдена за вибіркою, представляє цінність остільки, оскільки по ній можна судити про істинне математичне сподівання M_x . Цікавим є відшукування величини максимальної помилки Δ , яку ми допускаємо, припускаючи M_x рівним \bar{x} . Потрібно, отже, знайти величину, при якій

$$\bar{x} - \Delta \leq M_x \leq \bar{x} + \Delta. \quad (3.8)$$

Нерівністю (3.8) задається інтервал, в якому знаходиться значення математичного очікування M_x . Цей інтервал називається довірчим інтервалом для математичного очікування. Величина Δ залежить, очевидно, від об'єму вибірки n . Чим більше n , тим менше максимальна помилка Δ . Однак навіть при заданому n не можна абсолютно достовірно вказати величину Δ , так як розрахунок цієї величини, як і будь-який статистичний висновок, роблять на основі результатів експерименту, а вони свідомо містять помилки. Висновки, які роблять на основі неточних даних, принципово не можуть бути абсолютно достовірними, тому говорять про надійність статистичного висновку, який оцінюють величиною довірчої ймовірності p , де $0 < p < 1$. Наприклад, статистичний висновок, зроблений з довірчою ймовірністю $p=0,95$, буде справедливий в 95 випадках зі 100. Будемо користуватися частіше величиною $\alpha=1-p$, названої рівнем значимості. Рівень значимості задається заздалегідь до проведення розрахунків. Типові значення для α : 0,01; 0,05 і 0,1 або у відсотках: 1, 5, 10.

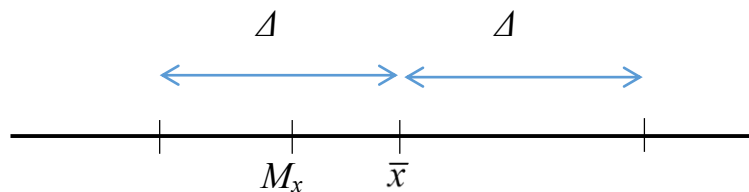


Рис. 3.1. Побудова довірчого інтервалу для математичного очікування.

Завдання отримання інтервальної оцінки в цьому випадку полягає в пошуку меж інтервалу (3.8), який із заданою довірчою ймовірністю p накриє дійсне значення математичного очікування M_x (рис.3.1).

При побудові будь-якої інтервальної оцінки, в тому числі і для математичного очікування, необхідно знати розподіл тієї точкової оцінки (випадкової величини), яка береться за основу для побудови довірчого інтервалу.

У математичній статистиці доводиться, що середні арифметичні значення \bar{x}_i з n незалежних результатів спостережень випадкової величини, розподіленої нормально з параметрами M_x і σ_x^2 також підпорядковується нормальному закону розподілу, причому з параметрами:

$$M(\bar{x}) = M_x, \quad (3.9)$$

$$\sigma^2(\bar{x}) = \frac{1}{n} \sigma^2(x) \quad (3.10)$$

Співвідношення (3.9), (3.10) не повинні інтуїтивно викликати ніяких серйозних заперечень: адже якщо підрахувати вибіркове середнє арифметичне за кількома вибірками одного і того ж об'єму, а потім знайти дисперсію одержаних значень, то найімовірніше припустити, що розкид (дисперсія) вибіркових середніх арифметичних буде меншою, ніж розкид (дисперсія) самих дослідних даних.

Прокоментуємо це положення таким ілюстративним числовим матеріалом (в продовження прикладу 3.1).

Приклад 3.4. Якщо провести серію 3 дослідів у кожному із з них з трьома вимірюваннями рівня забруднення і отримати такі результати

перша серія – 351, 370, 365 ($\bar{x}_1 = 362$, $s_1^2 = 97$)

друга серія – 375, 369, 345 ($\bar{x}_2 = 363$, $s_2^2 = 252$)

третя серія – 348, 363, 369 ($\bar{x}_3 = 360$, $s_3^2 = 117$)

і тепер оцінити дисперсію \bar{x}_1 , \bar{x}_2 і \bar{x}_3 за формулою(3.4), одержимо

$$s^2 = \frac{1}{3-1} \left(362^2 + 363^2 + 360^2 - \frac{1}{3} (362 + 363 + 360)^2 \right) = 2,33$$

Як видно з цього числового прикладу, вибіркова дисперсія середніх арифметичних $S^2 = 2,33$ за трьома вибірками (об'ємом 3) більше ніж на порядок менша тих вибіркових дисперсій (97, 252 і 117), які мають самі дослідні дані.

Для більш строгого обґрунтування співвідношення (3.10) нагадаємо, що якщо випадкова величина $Y = X1 \pm X2$, яка є сумою або різницею двох незалежних випадкових величин $X1$ і $X2$, то справедлива рівність

$$\sigma_Y^2 = \sigma_{X1}^2 + \sigma_{X2}^2. \quad (3.11)$$

Крім того, дисперсія добутку випадкової змінної X і постійної величини (константи) C дорівнює

$$\sigma^2(CX) = C^2 \sigma^2(X). \quad (3.12)$$

Закон додавання дисперсій справедливий при будь-якому числі доданків. З огляду на те, що $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ і σ_x^2 - дисперсія випадкової величини x , а також співвідношення (3.11) і (3.12), отримуємо:

$$\sigma^2(\bar{x}) = \sigma^2\left(\frac{1}{n} \sum_{i=1}^n x_i\right) = \frac{1}{n^2} \sigma^2\left(\sum_{i=1}^n x_i\right) = \frac{1}{n^2} n \sigma^2(x_i) = \frac{\sigma^2(x_i)}{n},$$

що і треба було доказати.

Повернемося до відшукування довірчого інтервалу для математичного очікування. Будемо припускати, що дисперсія вимірюваної величини x заздалегідь невідома, а її оцінка може бути знайдена з вибірки за допомогою формул (3.3) або (3.4). У цьому

випадку величину Δ , здавалося б можна замінити на $\sqrt{\left(\frac{\sigma^2}{n}\right)} = \frac{1}{\sqrt{n}} S$.

Однак таке припущення буде невірним, оскільки σ^2 є параметром нормального розподілу, який характеризує генеральну сукупність, а S^2 характеризує тільки вибірку.

На практиці, як правило, число вимірювань (наприклад, відбору проб при аналізі забруднення) звичайно не перевищує 10 ... 30. При такому малому числі спостережень фактична дисперсія σ^2 невідома, тому при побудові довірчого інтервалу для математичного очікування M використовують вибірккову дисперсію S^2 .

Функція розподілу густини ймовірності випадкової обмеженої множини центрованої і нормованої (на дисперсію), тобто приведеної величини t має розподіл, відмінний від нормального вид, а саме

$$f(t) = \frac{\Gamma\left(\frac{m+1}{2}\right)}{\sqrt{\pi m} \Gamma\left(\frac{m}{2}\right)} \left(1 + \frac{t^2}{m}\right)^{-\frac{m+1}{2}}, \quad (3.13)$$

де $\Gamma(u)$ - гамма-функція, яка є узагальненням поняття факторіала і має рекурентні властивості: $\Gamma(u+1) = u\Gamma(u)$ (для цілих чисел u справедливо $\Gamma(u+1) = u!$), m - число ступенів вільності, яке визначається різницею між об'ємом вибірки n і числом параметрів, які оцінюються за вибіркою. Функція $f(t)$, задана виразом (3.13), дістала назву *розподілу Стьюдента*.

Число ступенів вільності m обчислюється як різниця між числом експериментальних точок n і числом зв'язків, які обмежують свободу зміни випадкової величини. Так, при обчисленні вибірккової дисперсії за формулою (3.4), спостерігається один зв'язок, обумовлений рівнем середнього значення вибірки, тому число ступеней вільності дисперсії вибірки $m = n - 1$.



Вільям Сілі Госсет 1908 р.

В англomовній літературі розподіл Стьюдента бере назву зі статті Вільяма Госсета (William Sealy Gosset, 1876 -1937) в журналі «Біометрика» (1908 р.), опублікованій під псевдонімом «Стьюдент».

Поява Госсета на світ у родині полковника Королівського інженерного корпусу (Corps of Royal Engineers) гарантувало майбутньому великому статистику значні привілеї, недоступні більшості сучасників. Молодий Вільям мав намір піти по стопах батька, але виявився непридатний до військової

служби через поганий зір. Змирившись з думкою про неможливість військової кар'єри, Госсет закінчує престижний Винчестерський коледж (Winchester College), а потім навчається в Оксфорді (Oxford), де вивчає математику і хімію. Після закінчення університетського курсу в 1899 році Госсет поступає на роботу на пивоварний завод компанії «Гіннес» (Arthur Guinness Son & Co) в Дубліні. Коли Госсет починав працювати в «Гіннес», ця пивоварна компанія була вже найбільшою в світі. Навіть в порівнянні з роботою сучасних корпорацій тодішня виробнича практика Guinness була радикально орієнтованою на використання новітніх досягнень науки для поліпшення продукції, що випускається. Топ-менеджери ірландської компанії наймали на роботу найяскравіших молодих вчених, яких тільки могли знайти, і давали їм повну свободу в застосуванні інновацій у виробничих процесах і проведенні найсміливіших експериментів. Інноваційна політика Guinness створювала чудові умови для самореалізації допитливого і практично мислячого Госсета - рівнозначні шанси на втілення своїх ідей пізніше отримували невизнаний комп'ютерний геній, влаштувавшись на роботу в дослідницький центр корпорації Bell Labs в 1970-х, або вчений, який

працює в галузі досліджень штучного інтелекту, запрошений в команду Google сьогодні.

На рубежі дев'ятнадцятого і двадцятого століть Guinness приділяла свою основну увагу підтримці високої якості продукції при одночасному збільшенні масштабів і зниженні собівартості виробництва пива. За період між 1887 і 1914 випуск продукції зріс в 2 рази, досягнувши майже мільярда пінт (568 261 250 літрів). Резонним було питання: як компанія може збільшити виробництво пива, зберігши очікуване споживачами звична гарна якість пінного напою? Госсет був призначений до складу команди дослідників, які шукали відповідь на це питання. Команда «вчених-пивоварів», частиною якої став Госсет, зуміла вдосконалити процес відбору сировини для пивоваріння. Залучення Госсета до вирішення завдання пояснюється тим, що він вивчав математику в Оксфорді, а значить, «менше боявся цифр», ніж інші пивовари. Отже, Госсет приступив до роботи. Його мета - зрозуміти, наскільки знижується репрезентативність отриманих результатів при зменшенні розміру вибірки. У трохи більше академічних термінах завдання формулюється так: наскільки збільшується похибка вимірювання в разі, коли у вас є маленька вибірка з 2 або 10 зразків в порівнянні з вибіркою в 1000 зразків?

Начальство Госсета було в захваті від отриманих ним результатів: вони дозволяли приймати рішення про вибір сировини для пивоваріння, спираючись на числові дані, отримані в ході тестів, а не на інтуїцію і «свідчення» органів чуття. Жоден з конкурентів Guinness в той час нічого подібного робити не вмів. Однак самого Госсета власний метод апроксимації (наближення) не влаштовував в повному обсязі: він хотів знати, як з точки зору математики можна обґрунтувати достовірність висновків, отриманих при дослідженні малої вибірки, і правомірність застосування цих висновків до великої вибірки. Госсет повідомив керівництву «Гіннесса» про своє бажання проконсультуватися з даного питання з «якимось професійним математиком». Зобов'язана Госсету компанія «Гіннесс» оплатила йому творчу відпустку і відправила його в лабораторію Карла Пірсона в Університетському коледжі Лондона (University College London). Пірсон був одним з провідних наукових діячів свого часу і, як стали вважати пізніше, «батьком-засновником» сучасної статистики. За рік, проведений в лабораторії Пірсона, Госсет розробив математичне обґрунтування «закону помилок» («law of errors») для малих статистичних вибірок. Сьогодні його відкриття відомо як «t-розподіл

Стьюдента» (*Student's t-distribution*). Розподіл Стьюдента є основним способом визначення оцінки ймовірної помилки в залежності від розміру вибірки, і до цього дня широко застосовується в науці і промисловості. *Student's t-distribution* є одним із стовпів сучасної статистики і фундаментом концепції статистичної значущості.

Отже Гіннес був передовим підприємством харчової промисловості, і Госсет міг застосувати свої знання в області статистики як при варінні пива, так і на полях - для виведення самого урожайного сорту ячменю. Госсет і Пірсон були в хороших відносинах, і Пірсон допомагав Госсету в математичній частині його досліджень. Так, Пірсон був причетний до публікацій у 1908 році (що принесли славу Стьюденту-Госсету), але не надавав достатнього значення цьому відкриттю. Дослідження були направлені на потреби пивоварної компанії і проводилися на малій кількості спостережень. Біометристи ж, до яких належав Пірсон, зазвичай мали справу з сотнями спостережень і не відчували потреби в розвитку методів, заснованих на малій кількості дослідів.

Раніше інший дослідник, який працював на компанію Гіннес, опублікував у своїх матеріалах відомості, які становили комерційну таємницю цієї пивоварної компанії. Щоб запобігти подальшому розкриттю конфіденційної інформації, Гіннес заборонив своїм працівникам публікацію будь-яких матеріалів, незалежно від того, яка містилася в них інформація. Це означало, що Госсет не міг опублікувати свої роботи під своїм ім'ям. Він обрав собі псевдонім Стьюдент, щоб приховати своє місце роботи. Тому його найважливіше відкриття отримало назву розподілу Стьюдента, інакше б воно могло називатися розподілом Госсета.

Отже, Госсет практично всі свої роботи, включаючи роботу «Ймовірна похибка середнього» (англ. *The probable error of a mean*) опублікував в журналі Пірсона «Біометрика» під псевдонімом Стьюдент. Першим, хто зрозумів значення робіт Госсета по оцінці параметрів малої вибірки, був біолог Рональд Фішер. Госсет написав йому: «Я посилаю вам копію таблиць Стьюдента, оскільки ви, схоже, єдина людина, яка коли-небудь стане користуватися ними!». За іронією долі, статистика Стьюдента, завдяки якій став знаменитий Госсет, була фактично винаходом Фішера, тому що вона вкладалася в його теорію ступеней вільності. Фішер також застосував розподіл Стьюдента в регресійному аналізі.



Вільям Госсет 1936 р.

Інтерес Госсета до вирощування ячменю привів його до думки, що дослід треба планувати з тією метою, щоб не просто підвищити середню врожайність, але щоб вивести такі сорти ячменю, врожайність яких була б стійка до коливань складу ґрунту або клімату. Цей принцип зустрічається тільки пізніше у Р.Фішера.

У 1935 році В.Госсет покинув Дублін, щоб зайняти посаду головного пивовара, відповідального за наукову сторону виробничого процесу, в новій пивоварні Гіннеса в Парк Ройял у північно-західній

частині Лондона. У 1937 році В.Госсет помер від серцевого нападу в місті Беконсфільд в Англії.

Госсет був другом Пірсона і Фішера і був досить скромною людиною. Відомий випадок, коли він обірвав промову свого шанувальника словами «Фішер все одно б зумів відкрити все це сам».

Промислова революція початку 20-го століття і сучасні фабричні методи виробництва привели до створення продукції в раніше небачених масштабах. До настання епохи великомасштабних виробництв існувала можливість перевірки товарів з використанням якісних методів контролю. Кравці, пекарі, кораблебудівники і пивовари виробляли так мало продукції, що не представляло труднощі перевірити якість взагалі кожної одиниці товару.

Промислове виробництво принесло безліч переваг і виробникам, і споживачам, але його масштаби поставили завдання контролю якості: як брендам вберегти свою репутацію не завдаючи шкоди при попаданні у торгівлю бракованих товарів або продукції низької якості? На допомогу промисловцям прийшов Госсет, який продемонстрував, наскільки їм необхідні випадкові вибірки одиниць продукції, перевірка яких може дати уявлення про якість в цілому. Його методи в даний час стали стандартною частиною промислових протоколів контролю якості.

Основна ідея розподілу Стюдента - виправлення (adjusting) вибіркового стандартного відхилення S .

Приведена випадкова величина t у розподілі Стюдента

$$t = \frac{\bar{x} - M}{S / \sqrt{n}} \quad (3.14)$$

аналогічна приведеній (центрованій, нормованій) величині z , уведеній у попередньому розділі для стандартного нормованого розподілу Гауса, (нормального розподілу).

Найчастіше використовуються невеликі за об'ємом вибірки (менш 30), тому робота Стюдента має велике практичне значення. При $n > 30$, розподіл Стюдента практично переходить в стандартний нормальний розподіл з одиничною дисперсією. Розділ математичної статистики, присвячений обробці малочисельних вибірок ($2 < n < 20$), умовно називають мікростатистикою. В основі мікростатистичних оцінок нормально розподілених випадкових величин лежить розподіл Стюдента, який пов'язує між собою три основні характеристики вибіркової сукупності - ширину довірчого інтервалу, відповідну йому довірчу ймовірність і об'єм вибірки або число ступенів вільності вибірки $m = n - 1$. На рис.3.2. наведені залежності густини ймовірності від ширини довірчого інтервалу t у розподілі Стюдента при різному числі ступенів вільності, а також для порівняння нормальний розподіл.

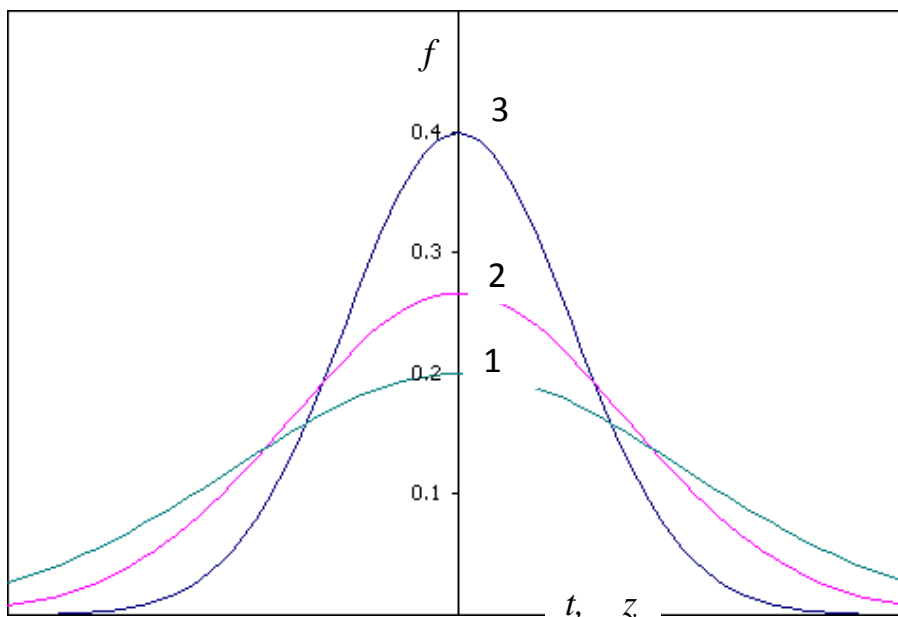


Рис. 3.2. Схематичні залежності густини ймовірності розподілу Стюдента при різному числі ступенів вільності m), а також для порівняння нормального розподілу – 3 (1 - m_1 , 2 - m_2 , $m_1 \leq m_2$).

При m , що дорівнює нескінченності, крива розподілу Стюдента збігається з кривою нормованого стандартного розподілу. Але, чим менше число ступенів вільності (об'єм вибірки), тим більш пологий вид має графік при великих значеннях аргументу t . Звідси випливає, що при однаковій ширині довірчого інтервалу довірна ймовірність, оцінена за Стюдентом, завжди менша довірчої ймовірності нормального розподілу Гауса-Лапласа. При цьому, чим менш представницька вибірка, тим більша різниця в оцінках двох типів. Або для отримання однакої довірчої ймовірності (однакових площ під кривими розподілу) для розподілу Стюдента необхідно вибрати більший довірчий інтервал.

Аналітичний вид функції $f(t)$ досить складний і громіздкий, тому зазвичай користуються таблицею заздалегідь вирахованих коефіцієнтів Стюдента $t_{pm} = t_{\alpha m}$, де $\alpha = 1 - p$ рівень значимості. Для цього потрібно знати тільки число ступеней вільності $m = n - 1$ і необхідну довірчу ймовірність p , або рівень значимості α . Істинне значення вимірюваної величини x (математичне очікування) буде знаходитись у

інтервалі $\pm \Delta = \frac{t_{pm} S}{\sqrt{n}}$ навколо середньоарифметичного значення \bar{x} , визначеного з n вимірювань з довірчою ймовірністю p :

$$\bar{x} - \frac{t_{pm} S}{\sqrt{n}} \leq M \leq \bar{x} + \frac{t_{pm} S}{\sqrt{n}}. \quad (3.15)$$

Приклад 3.5. Таким чином, для розглядуваного нами прикладу 3.1, із врахуванням того, що $t_{0.95, 2} = 4,3$ (див. Табл. 3.3), істинне значення вимірюваної величини x лежить в інтервалі $\pm \frac{4,3 \cdot 9,85}{\sqrt{3}} \approx 25$ навколо середнього значення 362 з довірчою ймовірністю 0,95.

3.2.2. Побудова довірчого інтервалу для дисперсії.

Розподіл Пірсона

При побудові довірчого інтервалу для дисперсії використовується випадкова величина χ^2 (читається: "хі-квадрат"),

$$\chi^2 = \sum_{i=1}^n \frac{(x_i - \bar{x})^2}{\sigma_x^2} = \frac{(n-1)S_x^2}{\sigma_x^2}, \quad (3.16)$$

розподіл якої має назву розподілу Пірсона (по імені англійського математика і біолога К. Пірсона), або χ^2 -розподіл ("хі-квадрат-розподіл").

Густина розподілу випадкової величини χ^2 описується рівнянням

$$f(\chi^2) = \frac{1}{2^{\frac{m}{2}} \Gamma\left(\frac{m}{2}\right)} (\chi^2)^{\frac{m-2}{2}} e^{-\frac{\chi^2}{2}}, \quad 0 \leq \chi^2 < \infty, \quad (3.17)$$

де $\Gamma\left(\frac{m}{2}\right)$ - гама функція, $m=n-1$ – число ступеней вільності.

Таблиця 3.3. Коефіцієнти розподілу Стюдента $t_{\alpha m}$ у залежності від рівня значимості α і числа ступеней вільності m

m	α				
	0,1	0,05	0,025	0,01	0,005
1	6,314	12,706	25,452	63,656	127,321
2	2,920	4,303	6,205	9,925	14,089
3	2,353	3,182	4,177	5,841	7,453
4	2,132	2,776	3,495	4,604	5,598
5	2,015	2,571	3,163	4,032	4,773
6	1,943	2,447	2,969	3,707	4,317
7	1,895	2,385	2,841	3,499	4,029
8	1,860	2,336	2,752	3,355	3,833
9	1,833	2,292	2,685	3,250	3,690
10	1,812	2,262	2,634	3,169	3,581
11	1,796	2,237	2,593	3,106	3,497
12	1,782	2,216	2,560	3,055	3,428
13	1,771	2,197	2,533	3,012	3,372
14	1,761	2,180	2,510	2,977	3,326
15	1,753	2,165	2,490	2,947	3,286
16	1,746	2,151	2,473	2,921	3,252
17	1,740	2,139	2,458	2,898	3,222
18	1,734	2,128	2,445	2,878	3,197
19	1,729	2,118	2,433	2,861	3,174
20	1,725	2,109	2,423	2,845	3,153
21	1,721	2,101	2,414	2,831	3,135
22	1,717	2,093	2,405	2,819	3,119
23	1,714	2,086	2,398	2,807	3,104
24	1,711	2,080	2,391	2,797	3,091
25	1,708	2,074	2,385	2,787	3,078
26	1,706	2,068	2,379	2,779	3,067
27	1,703	2,062	2,373	2,771	3,057
28	1,701	2,057	2,368	2,763	3,047
29	1,699	2,052	2,364	2,756	3,038
30	1,697	2,042	2,360	2,750	3,030



Карл Пірсон (Karl (Carl) Pearson, 1857, - 1936,) - англійський математик, статистик, біолог та філософ; один із засновників математичної статистики. Автор більш ніж 650 опублікованих наукових робіт. Народився в сім'ї адвоката. Закінчив Кембриджський університет в 1875 році. Згодом вивчав фізику у Гейдельберзькому та Берлінському університетах, тоді ж змінив ім'я - з "Carl" на "Karl". Деякий час він користувався обома

написаннями, після чого все ж зупинився на другому. Дехто схильний вважати, що вибрав більш "німецьку" форму свого імені Пірсон в честь Карла Маркса (Karl Marx); втім, точних доказів цієї версії не існує.

З 1884 по 1911 роки - професор прикладної математики і механіки Лондонського університету. В 1896 році був вибраний членом Королівського наукового товариства, в 1898 році був нагороджений Медаллю Дарвіна. В 1900 заснував журнал «Biometrika», який був присвячений застосуванню статичних методів в біології.

Пірсон опублікував ряд фундаментальних праць з математичної статистики (більш 400 робіт). Розробив теорію кореляції, критерії узгодженості, алгоритм прийняття рішень з оцінки параметрів. З його іменем пов'язані такі широко використовувані терміни та методи, як: криві Пірсона, розподіл Пірсона, рангова кореляція, множинна регресія та багато інших. Пірсон багато зусиль приклав для застосування своїх відкриттів в прикладних областях, передусім в біології, медицині. Ряд робіт відносяться до філософії та до історії науки. Відомим продовжувачем його робіт з прикладної математики став Рональд Фішер.

На рис.3.3 наведені криві $f(\chi^2)$ для різних значень m . Ці криві асиметричні, причому асиметрія особливо різко виражена при малих значеннях параметра m . Так, при $m=1$ крива йде в нескінченність, а при $m=2$ вона досягає максимального значення, рівного 0,5. При $m>2$ криві мають максимум при $\chi^2_{max}=m-2$. При великих значеннях m ($m>30$) χ^2 розподіл переходить в нормальний із середнім значенням $\bar{f}(\chi^2) = \sqrt{2m-1}$ і дисперсією $\sigma^2=1$.

Для побудови довірчого інтервалу для дисперсії розглянемо співвідношення

$$P(\chi_{P1}^2 \leq \chi^2 \leq \chi_{P2}^2) = P2 - P1 \quad (3.18)$$

і з урахуванням (3.16) для σ_x^2 :

$$P\left(\frac{n-1}{\chi_{P1}^2} S_x^2 \leq \sigma_x^2 \leq \frac{n-1}{\chi_{P2}^2} S_x^2\right) = P2 - P1, \quad (3.19)$$

де $\left(\frac{n-1}{\chi_{P1}^2} S_x^2 \leq \sigma_x^2 \leq \frac{n-1}{\chi_{P2}^2} S_x^2\right)$ довірчий інтервал для дисперсії з довірчою ймовірністю $p=1-\alpha$.

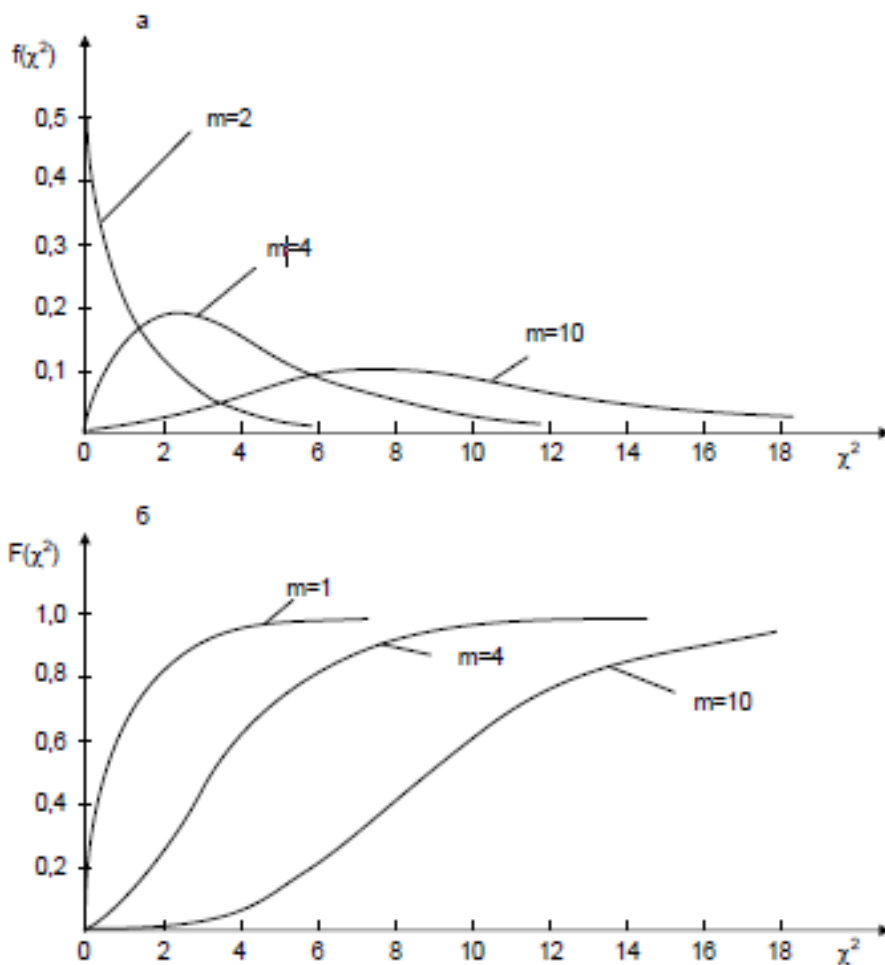


Рис.3.3. Густина розподілу ймовірності (а) та інтегральна функція розподілу (б) χ^2 .

В технічних задачах оцінку довірчого інтервалу для дисперсії проводять для такої ж довірчої ймовірності p як і для математичного очікування 0,9; 0,95; рідше 0,99, і відповідно рівень значимості α буде рівним 0,1; 0,05, рідше 0,01, причому ймовірність непопадання у довірчий інтервал (рівний α) ділиться порівну між більшими і меншими значеннями. Це значить, що довірчий інтервал починається з значень, ймовірність появи яких $\alpha/2$ і закінчується значеннями з ймовірністю появи $1 - \alpha/2$. За рахунок сильної асиметрії χ^2 -квадрат розподілу при малих значеннях ступеней вільності (малій кількості дослідів n) довірчий інтервал для дисперсії буде несиметричним. Як і при оцінці довірчого інтервалу для математичного очікування квантилі (коефіцієнти) розподілу Пірсона знаходять за таблицями (див. табл.3.4), а в Microsoft Excel для цього використовується функція ХІ2ОБР. Межі довірчого інтервалу для середнього квадратичного відхилення знаходяться як квадратний корінь з значень довірчих меж для дисперсії.

Таблиця 3.4. Квантилі χ^2 розподілу Пірсона у залежності від рівня значимості α і числа ступенів вільності m

m	α										
	0,995	0,99	0,975	0,95	0,9	0,5	0,1	0,05	0,025	0,01	0,005
1	0,0000	0,0001	0,0009	0,0039	0,016	0,455	2,706	3,841	5,024	6,635	7,879
2	0,010	0,020	0,051	0,103	0,211	1,386	4,605	5,991	7,378	9,210	10,60
3	0,072	0,115	0,216	0,352	0,584	2,366	6,251	7,815	9,348	11,34	12,84
4	0,207	0,297	0,484	0,711	1,064	3,357	7,779	9,488	11,14	13,28	14,86
5	0,412	0,554	0,831	1,145	1,610	4,351	9,236	11,07	12,83	15,09	16,75
6	0,676	0,872	1,237	1,635	2,204	5,348	10,64	12,59	14,45	16,81	18,55
7	0,989	1,239	1,690	2,167	2,833	6,346	12,02	14,07	16,01	18,48	20,28
8	1,344	1,647	2,180	2,733	3,490	7,344	13,36	15,51	17,53	20,09	21,95
9	1,735	2,088	2,700	3,325	4,168	8,343	14,68	16,92	19,02	21,67	23,59
10	2,156	2,558	3,247	3,940	4,865	9,342	15,99	18,31	20,48	23,21	25,19
11	2,603	3,053	3,816	4,575	5,578	10,34	17,28	19,68	21,92	24,73	26,76
12	3,074	3,571	4,404	5,226	6,304	11,34	18,55	21,03	23,34	26,22	28,30
13	3,565	4,107	5,009	5,892	7,041	12,34	19,81	22,36	24,74	27,69	29,82
14	4,075	4,660	5,629	6,571	7,790	13,34	21,06	23,68	26,12	29,14	31,32
15	4,601	5,229	6,262	7,261	8,547	14,34	22,31	25,00	27,49	30,58	32,80
16	5,142	5,812	6,908	7,962	9,312	15,34	23,54	26,30	28,85	32,00	34,27
17	5,697	6,408	7,564	8,672	10,09	16,34	24,77	27,59	30,19	33,41	35,72
18	6,265	7,015	8,231	9,390	10,86	17,34	25,99	28,87	31,53	34,81	37,16
19	6,844	7,633	8,907	10,12	11,65	18,34	27,20	30,14	32,85	36,19	38,58
20	7,434	8,260	9,591	10,85	12,44	19,34	28,41	31,41	34,17	37,57	40,00
21	8,034	8,897	10,28	11,59	13,24	20,34	29,62	32,67	35,48	38,93	41,40
22	8,643	9,542	10,98	12,34	14,04	21,34	30,81	33,92	36,78	40,29	42,80
23	9,260	10,196	11,69	13,09	14,85	22,34	32,01	35,17	38,08	41,64	44,18
24	9,886	10,856	12,40	13,85	15,66	23,34	33,20	36,42	39,36	42,98	45,56
25	10,520	11,524	13,12	14,61	16,47	24,34	34,38	37,65	40,65	44,31	46,93

Приклад 3.6. Використаємо дані прикладу 3.1 з трьома вибірковими значеннями 351, 370 і 365 та $S^2=97$ ($S \approx 10$). При $\alpha = 0,05$ для $P_1=\alpha/2=0,05/2=0,025$ і $m=n-1=3-1=2$ з таблиці 3.4 знаходимо $\chi_{P_1}^2 = 7,377779$ (або в Microsoft Excel для цього використовується функція ХІ2ОБР (0,025; 2)). Аналогічно для $P_2=1-0,05/2=0,975$ і $m=n-1=3-1=2$ знаходимо $\chi_{P_2}^2=0,050636$ (або в Microsoft Excel для цього використовується функція ХІ2ОБР (0,975; 2)).

Тоді довірчий інтервал для дисперсії згідно формули (3.16) складе

$$\frac{3-1}{7,38} \cdot 97 \leq \sigma^2 \leq \frac{3-1}{0,05} 97,$$

або після обчислень $26 \leq S \approx 10^2 \leq 3880$, довірчий інтервал для середнього квадратичного відхилення дорівнюватиме $5 < \sigma < 62$. Нагадаємо, що значення $S \approx 10$, яке, зрозуміло, попадає у визначений інтервал для σ .

3.2.3. Визначення необхідної кількості дослідів

при побудові інтервальної оцінки для математичного очікування

Інтуїтивно зрозуміло, що збільшення кількості вимірювань (числа проб, зразків і т.п.), може збільшити довірчу ймовірність P або звужити довірчий інтервал $\pm \Delta$ для визначення дійсного значення вимірюваної величини (математичного очікування). З виразу (3.16)

$$\Delta = \frac{t_{pm} S_x}{\sqrt{n}}$$

слідуює, що

$$n \geq \frac{t_{p(n-1)}^2 S_x^2}{\Delta^2}. \quad (3.20)$$

При розрахунках за цим рівнянням слід мати на увазі, що значення критерію Стюдента залежить не тільки від α , але і від числа ступенів вільності, останні ж визначаються числом вимірів. У зв'язку з цим рівняння (3.20) слід вирішувати методом послідовних наближень.

За перше наближення можна прийняти $n \geq \frac{t_{pm}^2 S_x^2}{\Delta^2} = t_{p\infty}^2 / (\varepsilon)^2$, де $\varepsilon = \frac{\Delta}{S_x}$.

Приклад 3.7. У прикладі 3.1, якщо необхідна точність (довірчий інтервал), наприклад, $\Delta=S_x$, ($\varepsilon=1$), то при довірчій ймовірності $p=0,95$ ($\alpha = 0,05$) $t_{0,95,\infty}=1,96$ і перша ітерація $n_1 \geq (1,96)^2 (1)^2 = 4$. Для другої ітерації $t_{0,95,4-1} = 3,18$, $n_2=3,18^2 \approx 9$, на наступній ітерації $t_{0,95,9-1}=2,26$,

$n_3=2,26^2\approx 5$; потім $t_{0,95,5-1} = 2,78$, $n_4=2,78^2\approx 7$; $n_5=2,45^2\approx 6$; $n_6=2,57^2\approx 7$. Таким чином для одержання заданої точності ($\Delta=S_x$) з довірчою ймовірністю $p=0,95$ (рівнем значимості $\alpha=0,05$) необхідно провести 7 вимірювань.

Кількість дослідів, необхідних для побудови довірчих інтервалів для математичного очікування при деяких інших $\varepsilon = \frac{\Delta}{S_x}$ і p , приведена в табл. 3.5

Таблиця 3.5. Необхідна кількість вимірювань при побудові довірчого інтервалу для математичного очікування

$\varepsilon = \frac{\Delta}{S_x}$	$\alpha=1-p$					
	0,5	0,3	0,1	0,05	0,01	0,001
1,0	2	3	5	7	11	17
0,5	3	6	13	18	31	50
0,4	4	8	19	27	46	74
0,3	6	13	32	46	78	127
0,2	13	29	70	99	171	277
0,1	41	169	273	387	668	1089
0,05	183	431	1084	1540	2659	4338
0,01	4543	10732	27161	38416	66358	108307

Принагідно відмітимо, що вимірювання проводяться за допомогою технічних пристроїв – засобів вимірювання, які мають свою систематичну похибку. Оскільки у результат вимірювання входить як систематична так і випадкова похибки, то задавати довірчий інтервал Δ для випадкової похибки менший ніж половина систематичної похибки немає змісту. Саме значення систематичної похибки у більшості випадків і визначає необхідну кількість вимірювань при заданій довірчій ймовірності.

3.3. Статистичні гіпотези

Як вже видно з викладеного вище матеріалу, при статистичному оцінюванні (тобто при наближеному визначенні випадкової величини) для обґрунтування (спроможності, незсуненості і ефективності) вибору тієї чи іншої оцінки невідомого параметра розподілу доводиться висловлювати припущення, що, наприклад, випадкова величина не суперечить нормальному закону розподілу. Крім того, використання всіх наявних значень вибірки при розрахунку оцінок, так чи інакше, передбачає, що серед них немає грубих помилок (значень, які різко виділяються). Із ще більшою кількістю різних припущень

(гіпотез) доводиться мати справу, коли необхідно не тільки визначати випадкові величини, але і порівнювати їх між собою, і тим більше, коли за результатами експерименту будується функція відгуку.

Статистична гіпотеза - будь-яке припущення, що стосується невідомого розподілу випадкової величини.

Наприклад, фахівця цікавить, чи вдалося у вже розглядуваному нами випадку домогтися підвищення пластичності дроту при його термічній обробці. Він може сформулювати наступну гіпотезу: "Пластичність збільшилася". Ця гіпотеза (нульова гіпотеза) підлягатиме перевірці в ході проведення дослідів. Крім того, можна сформулювати і будь-яку іншу (альтернативну) гіпотезу, наприклад: "Зміни пластичності не відбулося" або "Пластичність, навпаки, навіть зменшилася".

Процес прийняття рішення називається перевіркою статистичної гіпотези. Оскільки ми висували гіпотезу, спираючись тільки на випадкові вибіркові значення, наші висновки будуть носити імовірнісний характер, тобто ми не можемо дати точної відповіді: так чи ні. Можна буде лише з деякою часткою впевненості (з певною ймовірністю) стверджувати, що дані не протирічать (або суперечать) припущенням.

Статистичні гіпотези можна розділити на наступні групи.

Гіпотези про параметри розподілу. Ці гіпотези представляють собою припущення про значення деяких параметрів розподілу генеральної сукупності. Нехай, наприклад, висловлюється гіпотеза про те, що параметри (математичне очікування, дисперсії) в двох вибірках рівні між собою. Зазвичай гіпотези про параметри розподілу можна висувати, маючи в своєму розпорядженні достатньо інформації про генеральну сукупність або маючи вагомні підстави вважати відомим її закон розподілу.

Гіпотези про вид розподілу. Це більш загальні гіпотези, вони висувуються в умовах недостатньої інформації про генеральну сукупність. За вибіркою висувається гіпотеза про те, чи відповідають дані, наприклад, нормальному закону розподілу. Зауважимо, що перевірка гіпотези про нормальність розподілу може допомогти при подальшій обробці вибірки: якщо випадкову величину досить впевнено можна вважати нормально розподіленою, то до неї застосовні всі теореми про нормальні розподілені величини, зокрема є можливість побудувати довірчі інтервали для параметрів.

Нульова гіпотеза H_0 - це гіпотеза, що має найбільш важливе значення в проведеному дослідженні. Нульову гіпотезу висувають і потім перевіряють за допомогою статистичних критеріїв з метою виявлення підстав для її підтвердження або відхилення та прийняття альтернативної гіпотези.

Альтернативна гіпотеза H - кожна допустима гіпотеза, відмінна від нульової. Зазвичай в якості альтернативної гіпотези приймають гіпотезу другу за значимістю після основної.

Припущення, яке стосується невідомого параметра розподілу, коли вид розподілу відомий (наприклад, закон Гауса), називається параметричною гіпотезою, а припущення, при якому вид розподілу невідомий, називається непараметричною гіпотезою.

Завдання дослідника полягає в тому, щоб на основі аналізу експериментальних даних, отриманих з вибірки, прийняти ту чи іншу гіпотезу щодо властивостей генеральної сукупності, використовуючи при цьому будь-який спосіб (критерій) перевірки висловленого припущення.

Статистичний критерій - певний спосіб перевірки статистичних гіпотез.

Критерії для перевірки параметричних гіпотез називаються параметричними, а для перевірки непараметричних гіпотез - відповідно непараметричними.

Природно, що перш ніж використовувати той чи інший параметричний критерій, експериментатор повинен знайти спосіб переконатися в тому, узгоджується чи ні розподіл досліджуваної ним випадкової величини з тим чи іншим теоретичним (наприклад, нормальним) розподілом.

Критерій відповідності - це статистичний критерій для перевірки гіпотези про відповідність (рівність) розподілу випадкової величини досліджуваної сукупності з теоретичним розподілом або гіпотези про відповідність розподілів у двох і більше сукупностях.

Як і при статистичному оцінюванні, будь критерій може бути побудований тільки на основі тих результатів спостережень, які є в розпорядженні дослідника, тобто шляхом обчислення тієї чи іншої статистики. А як вже вище було відзначено, будь-яка статистика як деяка функція випадкової величини (функція від результатів спостережень) також є випадковою величиною.

Таким чином, статистичні гіпотези завжди носять імовірнісний характер. Це говорить про те, що, ґрунтуючись на тій чи іншій

статистиці і приймаючи нульову гіпотезу як робочу (або відкидаючи цю гіпотезу і в якості робочої і приймаючи альтернативну), дослідник може помилитися. Ситуації, що виникають під час перевірки статистичних гіпотез, представлені в табл.3.6.

Таблиця 3.6. Можливі наслідки перевірки статистичних гіпотез.

Фактична ситуація	H_0 - приймається	H_0 - відкидається
H_0 - вірна	Правильне рішення	Помилка першого роду
H_0 - не вірна	Помилка другого роду	Правильне рішення

1. Гіпотеза H_0 вірна, і вона не відкидається, тобто прийняте рішення відображає дійсний стан і приймається вірна гіпотеза.

2. Гіпотеза H_0 вірна, але вона відкидається, тобто в цьому випадку допущена *помилка першого роду*. Помилка першого роду - помилка, яка полягає в тому, що відкидають нульову гіпотезу, в той час як в дійсності ця гіпотеза вірна. Ймовірність цієї події за визначенням дорівнює рівню значимості α . Рівень значимості α - ймовірність помилки першого роду. Так як рівень значимості задається довільно, можна знизити ймовірність помилки першого роду до як завгодно низького рівня.

3. Гіпотеза H_0 невірна, і вона відкидається. Знову прийняте рішення відображає дійсний стан і відкидається невірна гіпотеза.

4. Гіпотеза H_0 невірна, але вона не відкидається. В цьому випадку допущена *помилка другого роду*. Помилка другого роду - помилка, яка полягає в тому, що приймають нульову гіпотезу, в той час як в дійсності ця гіпотеза неправильна. Якщо ймовірність помилки другого роду позначити як β , то величина $1-\beta$ називається потужність критерію. Потужність критерію - ймовірність того, що якщо вірна альтернативна гіпотеза, то нульова гіпотеза буде відкинута.

Значення застосовуваної для даного критерію статистики, при яких для обраного рівня значимості відкидається нульова гіпотеза, утворюють так звану *критичну область*.

Критична область - область з наступними властивостями: якщо значення застосовуваної статистики належать даній області, то відкидають нульову гіпотезу; в іншому випадку її приймають.

Наведені визначення намічають найпростішу форму перевірки статистичних гіпотез.

Очевидно, що при будь-якому заданому об'ємі вибірки ймовірність припуститися помилки першого роду можна скоротити, зменшивши рівень значимості α . Однак при цьому збільшується ймовірність допущення помилки другого роду (знижується потужність критерію). Таким чином, в більшості випадків не можна домогтися мінімального значення ймовірностей α і β водночас. Діють звичайно у такий спосіб: фіксують ймовірність α помилки першого роду, а потім домагаються мінімуму ймовірності β помилки другого роду. За рахунок чого можна зменшити β при фіксованому значенні α ? За рахунок правильного вибору критичної області: при заданій альтернативі H_1 критичну область вибирають таким чином, щоб значення β (ймовірність прийняти невірну гіпотезу) було найменшим з можливих. Таким чином, завдання полягає в побудові найбільш потужного критерію $(1-\beta)$ при заданому рівні значимості α .

Підводячи підсумок всьому вищесказаному, алгоритм перевірки будь-якої статистичної гіпотези в найзагальнішому випадку полягає в наступному:

- формулювання нульової гіпотези H_0 ;
- вибір однієї з альтернативних гіпотез H_1 (1), H_1 (2), H_1 (3);
- пошук критерію, за яким може бути перевірена сформульована нульова гіпотеза H_0 ;
- розрахунок значення статистики, застосовуваної для даного критерію;
- вибір рівня значимості α ;
- визначення критичної області при обраному рівні значимості α ;
- прийняття рішення: якщо значення статистики потрапило в критичну область - нульова гіпотеза відкидається, при цьому ймовірність помилки (першого роду) не перевищує обраний рівень значимості; в іншому випадку - нульова гіпотеза приймається.

3.4. Відсів грубих похибок. Критерій Смирнова-Грabbса

Часто навіть ретельно поставлені експерименти можуть давати неоднорідні дані, оскільки в процесі експерименту можуть змінитися умови проведення дослідів. Якщо експериментатор з яких-небудь причин не вловив цих змін, спостереження, що відповідають різним рівням факторів, будуть належати до різних генеральних сукупностей. Дані, які відповідають новим умовам, називають *грубими похибками*

(помилками) або *аномальними значеннями*, які різко виділяються. Грубі похибки з'являються також при неправильному записі показань приладів.

З практики проведення досліджень відомо, що експериментальні дані можуть містити $\sim 10\%$ аномальних значень. Однак ці 10% можуть дати сильний зсув при оцінці параметрів розподілу, особливо для дисперсії, так як помилки помітно відхиляються від основної групи значень, а на дисперсію особливо сильно впливають крайні члени *варіаційного ряду* (варіаційний ряд - результати спостережень, розташовані в зростаючій послідовності $x_1 < x_2 < x_3 \dots \dots < x_n$).

У разі відсіву грубих похибок (помилки) нульова гіпотеза формулюється наступним чином:

H_0 : "Серед результатів спостережень (вибіркових, дослідних даних) немає аномальних значень, які різко виділяються".

Альтернативною гіпотезою може бути або H_1 (1): "Серед результатів спостережень є тільки одна груба помилка", або H_1 (2): "Серед результатів спостережень є дві або більше грубих помилок".

У літературі можна зустріти велику кількість різних критеріїв для відсіву грубих похибок спостережень. Зазвичай експериментатори мають справу з вибірками невеликого об'єму (тобто коли генеральна дисперсія σ^2 невідома і оцінюється за дослідними даними через вибіркову дисперсію S_x^2 , причому саме в цьому випадку аномальні дані мають велику вагу. Найбільш поширеними і теоретично обґрунтованими в цьому випадку є критерій Смирнова-Граббса (використовується при H_1 (1)) і критерій Діксона (застосовується як при H_1 (1), так і при H_1 (2)). Тут розглянемо тільки *критерій Смирнова-Граббса*.

Смирнов Н.В. (Николай Васильевич Смирнов 1900 - 1966) радянський математик ХХ століття, член-кореспондент АН СРСР (1960). Лауреат Сталінської премії народився в Москві в родині дрібного церковного службовця який одночасно працював письменником в канцелярії Великого театру. З Москвою пов'язане все життя і наукова діяльність Н. В. Смирнова. Завершення його гімназійної освіти співпало за часом із завершенням Першої світової війни, під час якої він служив в санітарних частинах. Після демобілізації з Червоної армії в 1921 році вступив до фізико-математичного факультету Московського університету і зосередив свою головну увагу на вивченні математики, яка з часом витіснила всі інші його наукові інтереси і стала справою цілого життя. Смирнов -



Смирнов Н.В.

один з творців непараметричних методів математичної статистики і теорії граничних розподілів порядкових статистик. Його основні роботи лежали в галузі математичної статистики і теорії ймовірностей і присвячені вивченню граничних розподілів за допомогою асимптотичної поведінки кратних інтегралів при необмеженому збільшенні кратності. Вперше правильний розв'язок задачі щодо оцінки аномальності результатів спостережень було дано Н.В. Смирновим у 1941 р. У 1950 р.

Ф.Е. Граббс (Frank E. Grubbs) повторив ці результати Н.В. Смирнова без посилання на нього. У даний час для виявлення промахів рекомендується користуватися положеннями стандарту ДСТУ ГОСТ ISO 5725-2: 2005, в якому для виявлення промахів також пропонується використовувати критерій Граббса-Смирнова.

Якщо відомо, що є тільки одне аномальне значення (альтернативна гіпотеза H_1 (1)), то воно буде крайнім членом варіаційного ряду. Тому перевіряти вибірку на наявність однієї грубої помилки природно за допомогою статистики величини

$$u_1 = \frac{\bar{x} - x_1}{S_x} \quad (3.21)$$

якщо сумнів викликає перший член варіаційного ряду $x_1 = \min x_i$, або

$$u_m = \frac{x_n - \bar{x}}{S_x} \quad (3.22)$$

якщо сумнівний максимальний член варіаційного ряду $x_n = \max x_i$

Цей критерій вперше був запропонований Н.В. Смирновим. Він досліджував розподіл u_1 (3.21), і u_m (3.22) і склав таблиці процентних точок $u_{\alpha n}$ (квантилі порядку $p=1-\alpha$) для $\alpha=0,1; 0,05; 0,01$ при $3 < n < 20$. Якщо u_n більший розрахованого (табличного) значення (див. табл.3.5.), тобто u_n потрапляє в критичну область), то нульова гіпотеза

відхиляється, тобто викид x_n не випадковий і не характерний для даної сукупності даних, а визначається умовами що змінилися, чи грубими помилками при проведенні дослідів. У цьому випадку значення x_n виключають з розгляду, а знайдені раніше оцінки піддаються коректуванню з урахуванням відкинутого результату.

Як видно з приведених критичних значень критерію Смирнова, якщо значення вимірної величини відрізняється від середнього більше ніж на 3 середньоквадратичні відхилення, воно однозначно може бути віднесено до промахів, адже ймовірність його появи складає менше 0,01. Інші критерії відсіву промахів (грубих помилок), зокрема критерій Діксона для двох промахів, допитливий читач зможе знайти у спеціальній літературі.

Приклад 3.8. Пірометром вимірюється температура поверхні нагрітого тіла. Було проведено шість вимірювань температури $T^{\circ}\text{C}$, і отримані наступні значення: 925, 930, 950, 975, 990, 1080 ($n = 6$, причому, як видно, всі значення наведені в зростаючій послідовності, тобто у виді варіаційного ряду $T_1 = 925 < T_2 = 930 < T_3 = 950 \dots < T_6 = 1080$). Чи можна значення $T_6 = 1080$ вважати грубою похибкою, отриманою, припустимо, в результаті неправильної реєстрації показань пірометра?

Для відповіді на поставлене в цьому прикладі питання попередньо вирахуємо оцінки параметрів розподілу досліджуваної випадкової величини T (припускаючи, що вона не суперечить нормальному закону розподілу): вибіркове середнє арифметичне $T=975$ і вибіркове середнє квадратичне відхилення $S_T=57,27$

В електронних таблицях Microsoft Excel для цих розрахунків можна було б використовувати дві статистичні функції СРЗНАЧ (925; 930; 950; 975; 990; 1080) = 975 і СТАНДОТКЛОН (925; 930; 950; 975; 990; 1080) = 57,27128 .

Тепер скористаємося запропонованим вище алгоритмом перевірки статистичних гіпотез.

1. Формулюємо нульову гіпотезу H_0 : "Серед значень 925; 930; 950; 975; 990; 1080 немає грубих похибок".
2. Вибираємо наступну альтернативну гіпотезу $H_1(1)$: "Значення 1080 є (однією) грубою похибкою".
3. Сформульована нульова гіпотеза H_0 може бути перевірена за допомогою одного з критеріїв, наприклад критерієм Смирнова.
4. Значення статистики критерію Н.В. Смирнова (див. (3.22))

$$u_n = \frac{T_n - \bar{T}}{S_x} = \frac{1080 - 975}{57} = 1,83$$

5. Рівень значимості α приймемо рівним 0,05.
6. За табл. 3.7 при $\alpha = 0,05$ і $n = 6$ знаходимо $u_{0,05;6} = 1,82$, і з використанням (3.22) будемо критичну область $u \geq u_{0,05;6} = 1,82$, тобто і $u_6 > 1,82$.
7. Приймаємо рішення: оскільки значення статистики u_6 потрапило в критичну область ($1,83 > 1,82$) - нульова гіпотеза відкидається, і в якості робочої приймається альтернативна гіпотеза, тобто значення 1080°C з ймовірністю 0,95 (рівень значимості, не перевищує 0,05) за критерієм Н.В. Смирнова можна вважати грубою помилкою.
8. Цікаво відзначити, що якби на етапі 5 ми прийняли $\alpha = 0,01$, по таблиці критерію Н.В. Смирнова $u_{0,01;6} = 1,94$ і підраховане значення статистики 1,83, не потрапило б в критичну область ($1,83 < 1,94$). Отже, при $\alpha=0,01$ ми не можемо відкинути нульову гіпотезу, тобто за критерієм Н.В. Смирнова з ймовірністю 0,99 (надійністю, достовірністю) ми не можемо сказати, що значення 1080°C є грубою похибкою.

Інші критерії відсіву грубих помилок і промахів (Діксона, Ірвіна, Шовене, Романовського) можна знайти у спеціальній літературі.

Таблиця 3.7. Критичні значення критерію Смирнова $u_{\alpha n}$ у залежності від об'єму вибірки n і рівня значимості α

n	$u_{\alpha n}$		
	$\alpha=0,10$	$\alpha=0,05$	$\alpha=0,01$
3	1,15	1,15	1,15
4	1,42	1,46	1,49
5	1,60	1,67	1,75
6	1,73	1,82	1,94
7	1,83	1,94	2,10
8	1,91	2,03	2,22
9	1,98	2,11	2,32
10	2,03	2,18	2,41
11	2,09	2,23	2,48
12	2,13	2,29	2,55
13	2,17	2,33	2,61
14	2,21	2,37	2,66
15	2,25	2,41	2,70
16	2,28	2,44	2,75
17	2,31	2,48	2,78
18	2,34	2,50	2,82
19	2,36	2,53	2,85
20	2,38	2,53	2,88

3.5. Порівняння двох рядів спостережень.

Дисперсійний аналіз

При проведенні і аналізі результатів експериментальних досліджень часто доводиться порівнювати дві партії виробів, покази двох або декількох приладів, аналізувати результати роботи однотипних агрегатів, порівнювати результати досліджень двох проб матеріалів і т.д. При цьому у багатьох практичних завданнях вплив деяких величин на вихідну величину об'єкта неможливо оцінити кількісно. Треба визначити, в якій мірі на вихідну величину (на фон впливу випадкових величин) впливають дані якісні фактори, Наприклад:

- необхідно порівняти покази двох приладів, що вимірюють одну і ту ж величину, коли цими робочими засобами вимірювань отримано два ряди спостережень даної величини. Чи однакова точність вимірювання одного і того ж технологічного параметра різними приладами?

- потрібно перевірити робочий засіб вимірювання (тобто визначити, чи не виходять похибки його вимірювань за межі регламентованих значень) за допомогою зразкового засобу вимірювання. Чи математичне очікування показань даного приладу рівне дійсному значенню вимірюваного параметра?

- два агрегати випускають одну і ту ж продукцію. Необхідно зробити висновок про те, який з них кращий або гірший в будь-якому сенсі. Аналогічне завдання виникає при дослідженні партій виробів, які одержані від різних постачальників при з'ясуванні впливу різних якостей сировини на якість продукції.

Вирішення подібних завдань здійснюється також з використанням апарата перевірки статистичних гіпотез. Адже якщо нам необхідно було б порівняти дві випадкові величини X і Y , які мають нормальний розподіл, при відомих їх математичних очікуваннях і дисперсіях M_x , σ_x і M_y ; σ_y , то питання, очевидно, вирішувалося б досить просто.

Однак, як це вже неодноразово зазначалося раніше, будь-який з параметрів розподілу η випадкової величини може бути знайдений лише по всій генеральній сукупності, тобто тільки теоретично при проведенні безмежно великої кількості дослідів. Практично, за вибіркою обмеженого об'єму, дослідник може визначити тільки наближене значення параметра η^* - його оцінку. При цьому

ймовірність того, що оцінка η^* точно співпаде із істинним значення η оцінюваного параметра, дуже мала. Отже, навіть якщо рівні між собою параметри розподілів двох випадкових величин (наприклад M_x , і M_y), то їх оцінки середніми значеннями вибірок \bar{X} та \bar{Y} швидше за все не будуть однаковими.

Тому при порівнянні двох випадкових величин зазвичай доводиться перевіряти нульову гіпотезу $H_0: M_x = M_y$, при альтернативних гіпотезах типу $H_1(1): M_x < M_y$ або $H_1(2): M_x > M_y$.

3.5.1. Порівняння середніх значень. Критерій Стьюдента

Припустимо, що проведено дві серії вимірювань однієї і тієї ж величини. В одній серії з m_1 вимірювань отримано середнє значення \bar{x}_1 а в другій – з m_2 вимірювань – \bar{x}_2 . В якому випадку розбіжність між \bar{x}_1 і \bar{x}_2 є значущим, а в якому випадковим, природними для нормального розподілу випадкової величини?

Формулюємо нульову гіпотезу H_0 : «розбіжність між \bar{x}_1 і \bar{x}_2 є природними для нормального розподілу випадкової величини»

Вибираємо наступну альтернативну гіпотезу $H_1(1)$: "Різниця значення \bar{x}_1 і \bar{x}_2 є суттєвою.

Для перевірки нульової гіпотези про рівність двох вибірових середніх (\bar{x}_1 і \bar{x}_2) використовується критерій Стьюдента (t-критерій).

Для застосування даного критерію підраховують дисперсії

$$S_1^2 = \frac{1}{m_1 - 1} \sum_{i=1}^{m_1} (x_{1i} - \bar{x}_1)^2 \quad \text{та} \quad S_2^2 = \frac{1}{m_2 - 1} \sum_{i=1}^{m_2} (x_{2i} - \bar{x}_2)^2 .$$

Далі підраховують величину загальної дисперсії різниці ($\bar{x}_1 - \bar{x}_2$)

за формулою (див п.3. 5.2. , вираз (3.25)): $S = \sqrt{\frac{(m_1 - 1)S_1^2 + (m_2 - 1)S_2^2}{m_1 + m_2 - 2}}$.

Можна довести, що величина

$$t = \frac{|\bar{x}_1 - \bar{x}_2|}{S} \sqrt{\frac{m_1 m_2}{m_1 + m_2}}$$

– це той же коефіцієнт Стьюдента, який використовується для визначення довірчого інтервалу при невеликому числі вимірювань.

Розглянемо використання t-критерію на прикладі. Нехай проведено дві серії вимірювань опору R :

- у першій серії з $m_1=9$ вимірювань одержано середнє значення опору $\bar{R}_1 = 56,9 \text{ Ом}$, дисперсія $S_1^2 = 1,02$;

- у другій серії з $m_2=6$ вимірювань – середнє значення опору $\bar{R}_2 = 57,5 \text{ Ом}$, дисперсія $S_2^2 = 1,4$.

$$\text{Тоді } S = \sqrt{\frac{(m_1 - 1)S_1^2 + (m_2 - 1)S_2^2}{m_1 + m_2 - 2}} = \sqrt{\frac{8 \cdot 1,02 + 5 \cdot 1,4}{9 + 6 - 2}} = 1,08$$

$$\text{та } t = \frac{|\bar{x}_1 - \bar{x}_2|}{S} \sqrt{\frac{m_1 m_2}{m_1 + m_2}} = \frac{|56,9 - 57,5|}{1,08} \sqrt{\frac{9 \cdot 6}{6 + 9}} = 1,054.$$

Отримане значення t потрібно порівняти з табличним $t_{\alpha, m}$. При цьому m слід взяти рівним $m_1 + m_2 - 1 = 9 + 6 - 1 = 14$; при $\alpha = 0,95$, $t_{0,95; 14} = 2,2$. Бачимо, що знайдене нами значення $t < t_{\alpha, m}$ (з таблиці 3.3), отже, результати двох серій вимірювань значуще не відрізняються. Якщо ж виявиться, що $t > t_{\alpha, m}$, то результати вимірювань не можна вважати рівнозначними.

3.5.2. Порівняння двох дисперсій. Критерій Фішера

При виконанні вимірювань в різних умовах часто виникає задача порівняння ступеня розкиду (дисперсії) досліджуваних параметрів (випадкових величин).

Перевірка гіпотези про рівність дисперсій має велике значення, так як дисперсією вимірюється величина розсіювання і характеризує такі виключно важливі показники, як точність машин, приладів, стабільність технологічних процесів, якість готової продукції і т.д. Тому, наприклад, про переваги тієї чи іншої технології або про якість продукції, що випускається можна часто зробити висновок у результаті порівняння дисперсій тих параметрів, які їх характеризують.

Таким чином, потрібно встановити, чи вибіркові дисперсії $S_1^2 \neq S_2^2$ зі ступенями вільності m_1 і m_2 значно відрізняються або ж вони характеризують вибірки, взяті з однієї і тієї ж генеральної сукупності або з генеральних сукупностей з рівними дисперсіями ($\sigma_1^2 = \sigma_2^2 = \sigma^2$). В цьому випадку нульова гіпотеза формулюється у виді $H_0: \sigma_1^2 = \sigma_2^2 = \sigma^2$, тобто між двома генеральними дисперсіями відмінності немає при заданому рівні значимості α . Для перевірки цієї гіпотези використовується критерій, заснований на розподілі Фішера, який залежить тільки від числа ступенів вільності m_1 і m_2 .



Рональд Фішер

Сер Рональд Ейлмер Фішер (англ. *Sir Ronald Aylmer Fisher*, народився 17 лютого 1890, Іст-Фінчлі, Лондон, Велика Британія - помер 29 липня 1962, Аделаїда, Австралія)- англійський статистик, теоретик-еволюціоніст, генетик та євгенік, народився в сім'ї Джорджа і Кеті Фішер. Його батько був успішним торговцем предметами витонченого мистецтва. Дитинство його було щасливим, його обожнювали три старші сестри, старший брат і матір,

яка померла, коли Рональду було 14. Його батько через 18 місяців збанкрутів, провівши декілька невдалих операцій. Фішер мав поганий зір, але був не по роках розвиненим учнем і у віці 16 років виграв «*Neeld Medal*» (конкурс з математики) в школі Харроу (*Harrow School*). Унаслідок все того ж поганого зору, навчався математиці без використання «паперу і пера», що розвинуло здатність уявляти завдання в термінах геометрії. Уславився умінням отримувати відповідь, опускаючи проміжні етапи. Також виявляв сильну цікавість до біології, особливо, до еволюційного вчення.

Після закінчення Кембриджського університету неодноразово намагався записатися до лав Британської армії, але кожен раз не міг пройти медкомісію. Він став шкільним учителем і викладав фізику і математику в різних школах. Саме у цей період він почав розробляти важливі проблеми статистики та еволюційної теорії, включаючи революційну статтю «Кореляція між родичами у припущенні мендельської спадковості» («*The Correlation Between Relatives on the Supposition of Mendelian Inheritance*»). З 1919 до 1933 працював статистиком на дослідній сільськогосподарській станції. Потім, по 1943 рік, обіймав посаду професора в Лондонському університеті, а з 1943 року по 1957 завідував кафедрою генетики в Кембриджі.

Фішер дома розробляв різні програми, включаючи генетичні. Він здійснив ряд важливих робіт з генетики, євгеніці і еволюційній теорії, що знайшли завершення в його великій роботі «Генетична теорія природного відбору», опублікованій в 1930 р. У цій роботі Фішер звів

воєдино предмет генетики і еволюції шляхом природного відбору і виклав свою фундаментальну теорему: швидкість зростання пристосованості будь-якого організму в будь-який час дорівнює генотипній різноманітності його пристосованості. У той же час. Фішер стверджував, що, оскільки різноманітність у популяції підтримується мутаціями, швидкість появи мутацій визначає швидкість еволюції, тоді як природний відбір визначає її напрямом.

Багато формулювань Фішера не були повністю пояснені ним і деякі не були цілком зрозумілі його колегами-генетиками. Деякі були незалежно перевідкриті пізніше. Про Фішера кажуть, що він був генетиком з таким даром передбачення, що геніальність його висновків не розкрита досі.

Запропонував методологію планування експерименту в своїй інноваційній книзі «Планування експериментів» (1935 р.). Колосальний внесок був внесений ним в розвиток сучасної прикладної статистики.

Вираз для критерію Фішера має вид

$$F = \frac{\frac{S_1^2}{\sigma_1^2}}{\frac{S_2^2}{\sigma_2^2}} = \frac{\frac{S_1^2}{\sigma_1^2}}{\frac{S_2^2}{\sigma_2^2}} \quad (3.23)$$

Відмітимо, F може приймати тільки додатні значення. Густина розподілу величини F_{m_1, m_2} , має аналітичний вид

$$f(F) = \frac{\Gamma\left(\frac{m_1 + m_2}{2}\right) \left(\frac{m_1}{m_2}\right)^{\frac{m_1}{2}} F^{\left(\frac{m_1}{2}-1\right)}}{\Gamma\left(\frac{m_1}{2}\right) \Gamma\left(\frac{m_2}{2}\right) \left(1 + \frac{m_1}{m_2} F\right)^{\frac{m_1+m_2}{2}}}, \quad (3.24)$$

і графічно приведена на рис.3.4 а, а відповідна інтегральна функція розподілу на рис. 3.4 б. Треба мати на увазі, що швидкість зростання і спадання функції, а також величина і положення максимуму залежать від параметрів m_1 і m_2 .

Існують статистичні таблиці значень функції розподілу Фішера $f(F)$ для прийнятого рівня значимості, так і значень квантилів F_{α, m_1, m_2} цього розподілу (див. табл. 3.8).

Таблиця 3.8. Квантилі F_{α, m_1, m_2} розподілу Фішера (F -розподілу) для рівня значимості $\alpha=0.05$ у залежності від числа ступеней вільності m_1 і m_2 .

m_2	m_1															
	1	2	3	4	5	6	7	8	9	10	12	15	30	40	60	120
1	161,4	198,5	215,7	224,6	230,2	234,0	236,8	238,9	240,5	241,9	243,9	245,9	250,1	251,1	252,2	253,3
2	18,5	19,0	19,1	19,2	19,3	19,3	19,3	19,3	19,3	19,4	19,4	19,4	19,4	19,4	19,4	19,4
3	10,1	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,74	8,70	8,62	8,59	8,57	8,55
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,91	5,86	5,75	5,72	5,69	5,66
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,68	4,62	4,50	4,46	4,43	4,40
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,00	3,94	3,81	3,77	3,74	3,70
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,57	3,51	3,38	3,34	3,30	3,27
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,28	3,22	3,08	3,04	3,01	2,97
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,07	3,01	2,86	2,83	2,79	2,75
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,91	2,85	2,70	2,66	2,62	2,58
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,79	2,72	2,57	2,53	2,49	2,45
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,69	2,62	2,47	2,43	2,38	2,34
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,60	2,53	2,38	2,34	2,30	2,25
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,53	2,46	2,31	2,27	2,22	2,18
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,48	2,40	2,25	2,20	2,16	2,11
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,42	2,35	2,19	2,15	2,11	2,06
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,38	2,31	2,15	2,10	2,06	2,01
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,34	2,27	2,11	2,06	2,02	1,97
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,31	2,23	2,07	2,03	1,98	1,93
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,28	2,20	2,04	1,99	1,95	1,90
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,25	2,18	2,01	1,96	1,92	1,87
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30	2,23	2,15	1,98	1,94	1,89	1,84
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27	2,20	2,13	1,96	1,91	1,86	1,81
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	2,18	2,11	1,94	1,89	1,84	1,79
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24	2,16	2,09	1,92	1,87	1,82	1,77
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	2,15	2,07	1,90	1,85	1,80	1,75
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25	2,20	2,13	2,06	1,88	1,84	1,79	1,73
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19	2,12	2,04	1,87	1,82	1,77	1,71
29	4,18	3,33	2,93	2,70	2,55	2,43	2,35	2,28	2,22	2,18	2,10	2,03	1,85	1,81	1,75	1,70
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	2,09	2,01	1,84	1,79	1,74	1,68
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	2,00	1,92	1,74	1,69	1,64	1,58
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99	1,92	1,84	1,65	1,59	1,53	1,47
120	3,92	3,07	2,68	2,45	2,29	2,18	2,09	2,02	1,96	1,91	1,83	1,75	1,55	1,50	1,43	1,35

У разі підтвердження нульової гіпотези за двома вибірковим дисперсіями оцінка загальної генеральної дисперсії σ^2 проводиться за формулою:

$$S^2 = \frac{(m_1 - 1)S_1^2 + (m_2 - 1)S_2^2}{m_1 + m_2 - 2} \quad (3.25)$$

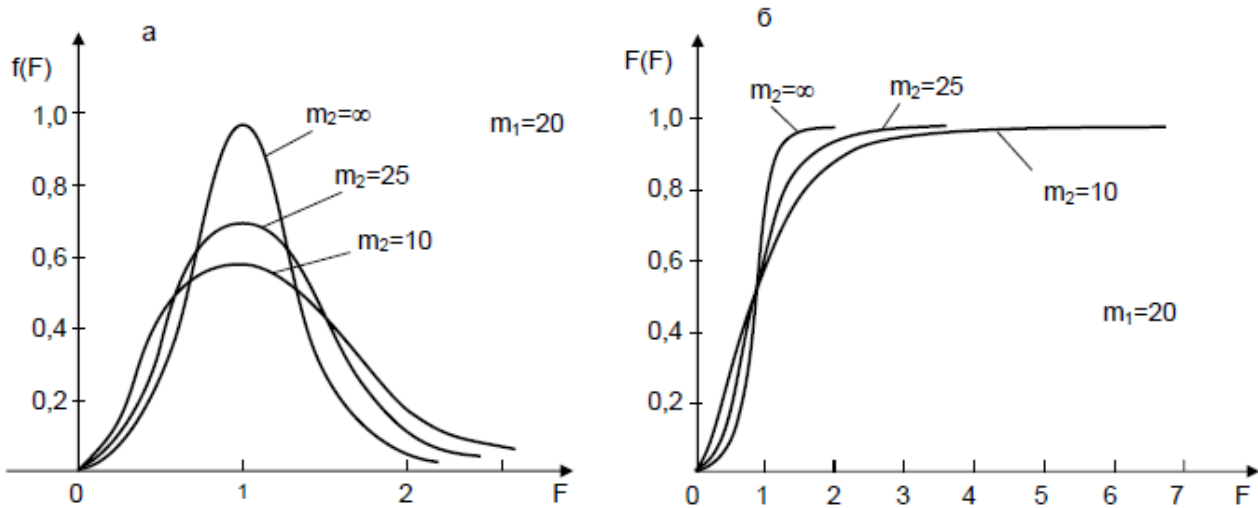


Рис.3.4 Густина розподілу Фішера (а) та інтегральна функція розподілу Фішера (б) для часткового випадку $m_1=20$.

Приклад 3.9. По двох незалежних вибірках $n_1 = 10$ і $n_2 = 15$ із генеральних сукупностей X_1 та X_2 знайдені вибіркові дисперсії $S_{x1}=0,86$ та $S_{x2}=0,43$. При $\alpha=0,05$ перевірити нульову гіпотезу H_0 : дисперсії $\sigma_1=\sigma_2$ при конкуруючій гіпотезі H_1 : $\sigma_1 > \sigma_2$. Розв'язок: спостережуваний критерій Фішера $F=0,86/0,43=2$. За таблицею 3.8. знаходимо значення $F_{кр}$ при $\alpha=1-p = 0,05$ та $m_1=10-1=9$; $m_2=15-1=14$, яке становить 2,65. Оскільки $F < F_{кр}$, то вибіркові дисперсії розрізняються незначущо, тобто вони характеризують вибірки, взяті з однієї і тієї ж генеральної сукупності або з генеральних сукупностей з рівними дисперсіями ($\sigma_1^2=\sigma_2^2=\sigma^2$). При цьому загальна дисперсія визначена за формулою (3.25) становить $S^2=0,45$.

3.5.3. Перевірка однорідності декількох дисперсій.

Критерій Кохрена

Критерій Фішера використовується для порівняння тільки двох дисперсій, однак на практиці доводиться порівнювати між собою три і більше дисперсій.

При зіставленні дисперсій ряду сукупностей нульова гіпотеза полягає у тому, що всі k сукупностей, з яких взяті вибірки, мають рівні дисперсії.



Вільям Кохрен
(William G. Cochran),
професор статистики
Гарвардського університету
(15.07.1909 - 29.03.1980)

1. $H_0: \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \dots = \sigma_k^2 = \sigma^2$, тобто перевірці підлягає припущення, що всі емпіричні дисперсії $S_1^2, S_2^2, \dots, S_k^2$ відносяться до вибірок з однієї і тієї ж генеральної сукупності з дисперсією σ^2 .

Нехай серед кількох серій вимірювань виявлена така, вибіркова дисперсія якої S_{\max}^2 помітно більше за всіх інших. Завдання полягає в тому, щоб з'ясувати, чи можна вважати відмінність виділеної дисперсії S_{\max}^2 істотною. Іншими словами, альтернативна гіпотеза може бути вибрана як

2. $H_1: \sigma_{\max}^2 > \sigma^2$.

3. При рівному об'ємі $n_1 = n_2 = n_3 = \dots = n_k = n$ всіх k вибірок

може бути використаний так званий критерій Кохрена.

4. Статистика критерію Кохрена G розраховується як відношення S_{\max}^2 до суми всіх вибірових дисперсій

$$G = \frac{S_{\max}^2}{\sum_{i=1}^k S_i^2} \quad (3.26)$$

5. Надалі для обраного рівня значимості α визначається табличне значення цього критерію, яке залежить від числа ступенів вільності $m = n - 1$ і числа порівнюваних дисперсій k - $G_{\alpha; m; k}$ (див. табл. 3.9).

6. Критична область будується як $G \geq G_{\alpha; m; k}$.

7. При $G \leq G_{\alpha; m; k}$ гіпотеза $H_0: \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \dots = \sigma_k^2 = \sigma^2$ приймається в якості робочої, тобто відмінність виділеної дисперсії S_{\max}^2 вважається несуттєвою.

Оцінка узагальненої дисперсії у цьому випадку проводиться за формулою:

$$S^2 = \frac{\sum_{i=1}^k S_i^2}{k} \quad (3.27)$$

Таблиця 3.9. Квантилі $G_{\alpha;m;k}$ розподілу Кохрена для рівня значимості $\alpha=0,05$ у залежності від числа ступеней вільності $n-1=m_1$ та об'єму вибірки $k=m_2$.

m_2	m_1												
	1	2	3	4	5	6	7	8	9	10	16	36	144
2	0,9985	0,9750	0,9392	0,9057	0,8772	0,8534	0,8333	0,8159	0,8010	0,7880	0,7341	0,6602	0,5813
3	0,9669	0,8709	0,7977	0,7457	0,7071	0,6771	0,6530	0,6333	0,6167	0,6025	0,5466	0,4748	0,4031
4	0,9065	0,7679	0,6841	0,6287	0,5895	0,5598	0,5365	0,5175	0,5017	0,4884	0,4366	0,3720	0,3093
5	0,8412	0,6838	0,5981	0,5441	0,5065	0,4783	0,4564	0,4387	0,4241	0,4118	0,3645	0,3066	0,2513
6	0,7808	0,6161	0,5321	0,4803	0,4447	0,4184	0,3980	0,3817	0,3682	0,3568	0,3135	0,2612	0,2919
7	0,7271	0,5612	0,4800	0,4307	0,3974	0,3726	0,3535	0,3384	0,3259	0,3154	0,2756	0,2278	0,1833
8	0,6798	0,5157	0,4377	0,3910	0,3595	0,3362	0,3185	0,3043	0,2926	0,2829	0,2462	0,2022	0,1616
9	0,6385	0,4775	0,4027	0,3584	0,3286	0,3067	0,2901	0,2768	0,2659	0,2568	0,2226	0,1820	0,1446
10	0,6020	0,4450	0,3733	0,3311	0,3029	0,2823	0,2666	0,2541	0,2439	0,2355	0,2032	0,1655	0,1308
12	0,5410	0,3924	0,3264	0,2880	0,2626	0,2439	0,2299	0,2187	0,2098	0,2020	0,1737	0,1403	0,1100
15	0,4709	0,3346	0,2758	0,2419	0,2195	0,2034	0,1911	0,1815	0,1736	0,1671	0,1429	0,1144	0,0889
20	0,3894	0,2705	0,2205	0,1921	0,1735	0,1602	0,1501	0,1422	0,1357	0,1303	0,1108	0,0879	0,0675
24	0,3434	0,2354	0,1907	0,1656	0,1493	0,1374	0,1286	0,1216	0,1160	0,1113	0,0942	0,0743	0,0567
30	0,2929	0,1980	0,1593	0,1377	0,1237	0,1137	0,1061	0,1002	0,0958	0,0921	0,0771	0,0604	0,0457
40	0,2370	0,1576	0,1259	0,1082	0,0968	0,0887	0,0827	0,0780	0,0745	0,0713	0,0595	0,0462	0,0347
60	0,1737	0,1131	0,0895	0,0765	0,0682	0,0623	0,0583	0,0552	0,0520	0,0497	0,0411	0,0316	0,0245
120	0,0998	0,0632	0,0495	0,0419	0,0371	0,0337	0,0312	0,0292	0,0279	0,0266	0,0218	0,0168	0,0120

Приклад 3.10. Для даних, приведених у прикладі 3.4, $S_1^2 = 97$, $S_2^2 = 252$, $S_3^2 = 117$ формулюємо гіпотезу $H_0: \sigma_1^2 = \sigma_2^2 = \sigma_3^2$ (вибірки, мають рівні дисперсії). Альтернативна гіпотеза $H_1: \sigma_2^2 = \sigma_{\max}^2 > \sigma_1^2$ та σ_3^2 .

Знаходимо критерій Кохрена
$$G = \frac{S_{\max}^2}{\sum_{i=1}^k S_i^2} = \frac{257}{97 + 257 + 117} = 0,55.$$

Надалі для обраного рівня значимості $\alpha=0,05$, числа ступенів вільності $m=n-1=3-1=2$ і числа порівнюваних дисперсій $k=m_2=3$ по таблиці 3.9 знаходимо критичне значення критерію Кохрена – $G_{\text{крит}} = G_{0,05;2;3} = 0,8709$. Оскільки експериментальне $G=0,55$ менше як $G_{\text{крит}} = G_{0,05;2;3} = 0,87$ і не попадає у критичну область, тобто відмінність виділеної дисперсії $S_2^2 = S_{\max}^2$ вважається несуттєвою у нас нема причин вважати, що дисперсії $S_1^2, S_2^2, \dots, S_3^2$ відносяться до вибірок з різних генеральних сукупностей.

Критерій Кохрена можна використовувати тільки в тих випадках, коли всі порівнювані дисперсії мають однакове число ступенів вільності $m=n-1$ (однакові об'єми вибірок $n_1=n_2=n_3=\dots=n_k=n$). Якщо ж число вимірювань n в різних серіях неоднакове, то для перевірки однорідності дисперсій можна вибрати, наприклад, критерій Бартлета.

3.6. Перевірка гіпотез про вид функції розподілу.

Критерій відповідності Пірсона

Розглянуті раніше методи оцінювання параметрів розподілу випадкової величини і критерії для перевірки статистичних гіпотез припускали, що відома функція розподілу (нормальний закон - розподіл Гауса). Однак у більшості випадків вид закону розподілу є гіпотетичним і сам по собі вимагає статистичного підтвердження.

Найбільш простим, але досить наближеним методом перевірки відповідності результатів експерименту тому чи іншому закону розподілу є графічний метод. Він полягає в оцінці емпіричної функції розподілу і зіставленні її з функцією передбачуваного теоретичного закону. Якщо побудовані експериментальні точки лежать поблизу теоретичного графіка, то можна вважати, що отримані у досліді дані не суперечать обраному теоретичному закону розподілу. Графічний метод є у значній мірі суб'єктивним і використовується на практиці у якості першого наближення при вирішенні подібних завдань.

Більш об'єктивні методи встановлення виду розподілу випадкової величини будуються на апараті перевірки статистичних гіпотез – критерію відповідності.

Нульова гіпотеза в даному випадку полягає в тому, що H_0 : - досліджувана генеральна сукупність не суперечить передбачуваному теоретичному закону розподілу. При цьому альтернативна гіпотеза зазвичай формулюється як H_1 : випадкова величина має будь-яке інший розподіл, відмінний від передбачуваного.

Розроблено досить багато критеріїв відповідності, що відрізняються як своєю потужністю, так і обсягом дослідних даних, необхідних для їх використання. Розглянемо деякі з них, і в першу чергу зупинимося на критеріях відповідності, які можуть бути використані при відносно великих об'ємах вибірки.

Коли експериментатор має досить велику кількість експериментальних даних x_j ($j=1, 2, \dots, n > 100$), то їх попередня обробка починається з групування і проводиться в наступній послідовності:

1. Знаходять найбільше (x_{\max}) і найменше (x_{\min}) значення вибірки випадкової величини і обчислюють її розмах $R=x_{\max}-x_{\min}$.

2. Розмах випадкової величини розбивають на k рівних інтервалів. Кількість інтервалів k вибирають у залежності від об'єму

вибірки. Наприклад, при $n > 100$ його значення рекомендується приймати рівним $k = 9 \div 15$ (якщо $n < 100$, то $k = 7$). Число інтервалів k можна визначити і за формулою Штюргеса $k = 1 + 3,32 \lg(n)$ з округленням отриманого значення до найближчої цілої величини.

3. Визначають ширину інтервалу $h = R/k$, для спрощення розрахунків отримані значення округлюють в будь-яку сторону, трохи збільшуючи або зменшуючи при цьому розмах варіювання R .

4. Встановлюють межі інтервалів і підраховують число попадань m_i випадкової величини в кожний з обраних інтервалів, $1 \leq i \leq k$.

Графічною формою подання неперервної випадкової величини є гістограма (рис.3.5). Послідовність побудови гістограм наступна:

1. Визначається величина ординати $f_i = \frac{P_i}{h}$, де $P_i = \frac{m_i}{n}$ (n – об'єм вибірки).

2. У системі координат $f_i = f(x)$ на ширині інтервалу h відкладають величини f_i як висоти і будуються прямокутники. Очевидно, що площа елементарного прямокутника $A_i = hf_i = P_i = \frac{m_i}{n}$ дорівнює відношенню числа дослідів m_i при яких випадкова величина виявилася всередині цього інтервалу, до загальної кількості дослідів n .

3. Площа всієї гістограми $A = \sum_{i=1}^k A_i = \sum P_i = 1$. Отже, площа, обмежена гістограмою, дорівнює одиниці. Побудова гістограми інтегральної функції розподілу F здійснюється підсумовуванням ймовірностей P_i .

Надалі здійснюється порівняння експериментально отриманого розподілу випадкової величини з деяким видом теоретичного розподілу. Для цієї мети використовуються різні критерії відповідності: χ^2 (хі-квадрат) Пірсона, Колмогорова-Смирнова та ін.

Критерій Пірсона

Розглянемо методику перевірки гіпотези нормального розподілу за критерієм χ^2 Пірсона. Цей критерій крім визначення довірчого інтервалу для дисперсії нерідко використовується для перевірки відповідності розподілів, отриманих за даними вибірки з деякою теоретичної густиною розподілу.

У даному випадку застосування критерію χ^2 передбачає використання властивостей нормованого (стандартного) нормального розподілу. Нагадаємо, що рівняння кривої густини стандартного нормального розподілу має вид

$$f(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) \approx 0,4 \exp\left(-\frac{z^2}{2}\right), \text{ де } z = \frac{x - M_x}{\sigma_x}.$$

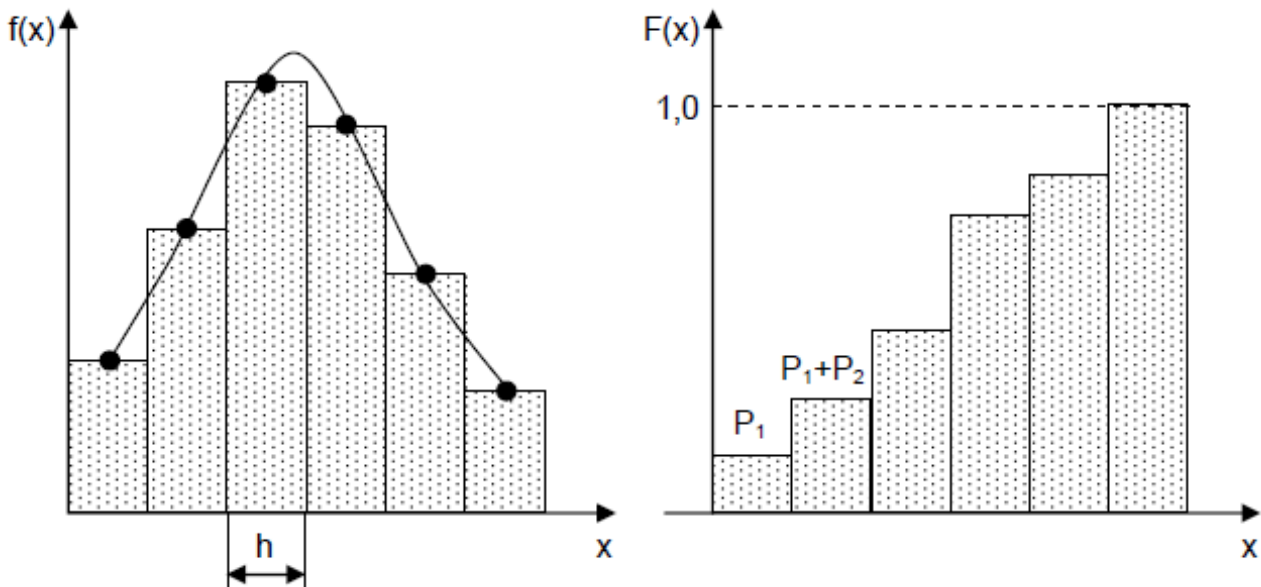


Рис.3.5. До побудови гістограми розподілу випадкової величини x .

Тоді теоретична ймовірність попадання випадкової величини в інтервал $\Delta z = z_{i+1} - z_i$ в разі нормального розподілу можна визначити за формулою

$$P_i^{теор} = F(z_{i+1}) - F(z_i) = \frac{1}{\sqrt{2\pi}} \int_{z_i}^{z_{i+1}} \exp\left(-\frac{u^2}{2}\right) du$$

Відмінність оцінки закону розподілу P від теоретичного закону розподілу $P^{теор}$ можна охарактеризувати величиною

$$\chi^2 = \sum C_i (P_i - P_i^{теор})^2, \quad (3.28)$$

де P_i і $P_i^{теор}$ - оцінка і теоретична ймовірності випадкової величини для i -го інтервалу; C_i - вагові коефіцієнти, які з більшою вагою враховують відхилення для менших P_i . Пірсон вибрав вагові коефіцієнти таким чином:

$$C_i = \frac{n}{P_i^{теор}}. \quad (3.29)$$

Пірсон показав, що при такому виборі C_i закон розподілу χ^2 слабо залежить від n і $P(x)$, а визначається в основному числом інтервалів k .

Отже,

$$\chi^2 = n \sum_{i=1}^k \frac{(P_i - P_i^{теор})^2}{P_i^{теор}} = n \sum_{i=1}^k \frac{(m_i / n - P_i^{теор})^2}{P_i^{теор}} = \sum_{i=1}^k \frac{(m_i - nP_i^{теор})^2}{nP_i^{теор}}. \quad (3.30)$$

Очевидно, що при ідеальній відповідності експериментальних даних нормальному закону розподілу значення критерію Пірсона дорівнюватиме нулю, тому що $P_i = P_i^{теор}$.

У виразі (3.28) стоїть сума квадратів k випадкових величин, проте вони не є незалежними, так як на них накладається певна кількість зв'язків. Одним з таких зв'язків є вимога, щоб площа під кривою оцінки закону розподілу дорівнювала одиниці: Іноді вимагають, щоб середнє значення x збігалось з математичним очікуванням M_x , а вибіркова дисперсія S_x^2 - з дисперсією σ_x^2 . Тому число ступенів вільності найчастіше визначається як

$$m = k - 2. \quad (3.31)$$

Теоретичне значення критерію Пірсона $\chi^2_{\alpha; m}$ визначається за довідковими даними (див. табл.3.4) або з використанням пакетів прикладних програм при заданому рівні значимості α і числі ступенів вільності m (див. Функцію ХІ2ОБР(α ; m) з електронних таблиць Microsoft Excel).

Алгоритм використання критерію Пірсона полягає в наступному:

1. Висуваються нуль-гіпотеза H_0 : "Відмінність експериментальних даних від нормального закону розподілу не суттєва" і альтернативна їй гіпотеза H_1 : "Відмінність експериментальних даних від нормального закону розподілу істотна, тобто експериментальні дані не підкоряються закону нормального розподілу".

2. За результатами експериментальних вимірювань і припущенням нормального закону їх розподілу визначається розрахункове значення критерію Пірсона χ^2 .

3. Визначають число ступенів вільності m , задаються рівнем значущості α і визначають теоретичне значення критерію Пірсона $\chi^2_{\alpha; m}$.

4. Якщо $\chi^2 < \chi^2_{\alpha; m}$, то нуль-гіпотеза H_0 про нормальний закон розподілу експериментальних даних приймається з довірчою ймовірністю $p = 1 - \alpha$. В іншому випадку нуль-гіпотеза відкидається і приймається альтернативна гіпотеза H_1 .

Відзначимо важливі рекомендації по використанню критерію χ^2 .

Якщо при деякому числі вимірювань критерій $\chi^2 > \chi^2_{\alpha; m}$, але сумніви в нормальності розподілу відсутні, то слід, якщо є можливість, збільшити число вимірювань в кілька разів і повторити аналіз за цим же критерієм.

Число ступенів вільності $m=k-2$ відноситься до такого випадку, коли обидва параметри нормального закону розподілу визначаються за результатами вимірювань, тобто коли замість точних значень M_x і σ_x застосовують їх емпіричні значення (оцінки) \bar{x} і S_x . Якщо ж значення M_x точно відомо (наприклад, при вимірюванні еталона), то число ступенів вільності дорівнює $k=n-1$; якщо відомі обидва параметри M_x і σ_x , то число ступенів вільності дорівнює $k=n$. На практиці така ситуація зустрічається відносно рідко, і тому для числа ступенів вільності не менше п'яти бажано брати число інтервалів не менше семи (іноді дев'яти).

Існують і інші критерії перевірки нормальності розподілу випадкової величини наприклад Колмогорова-Смирнова, які читач може знайти у спеціальній літературі.

3.7. Перетворення розподілів до нормального

Якщо дослідник, використавши методи, викладені в попередньому параграфі, переконався, що гіпотеза нормальності розподілу не може бути прийнята, то цілком може бути, що за допомогою існуючих методів вдасться так перетворити вихідні дані, що їх розподіл буде підкорятися нормальному закону розподілу.

Приклад 3.11. Для пояснення ідеї перетворень розглянемо якісний приклад. Нехай крива розподілу $f(x)$ має вид, представлений на рис. 3.6 а, тобто є дуже крута ліва гілка і полого права. Такий розподіл відрізняється від нормального.

Для виконання операцій перетворення кожне спостереження трансформується за допомогою логарифмічного перетворення $x'=\lg x$. При цьому ліва гілка кривої розподілу сильно розтягується і розподіл приймає наближено нормальний вид (рис.3.6 б). Якщо при перетворенні виходять значення, розташовані між 0 і 1, то всі спостережувані значення для зручності розрахунків і щоб уникнути отримання від'ємних параметрів необхідно помножити на 10 у відповідній степені, щоб всі нові отримані і перетворені значення були більші одиниці, тобто необхідно виконати перетворення $x''=\lg(x \cdot 10^a)$.

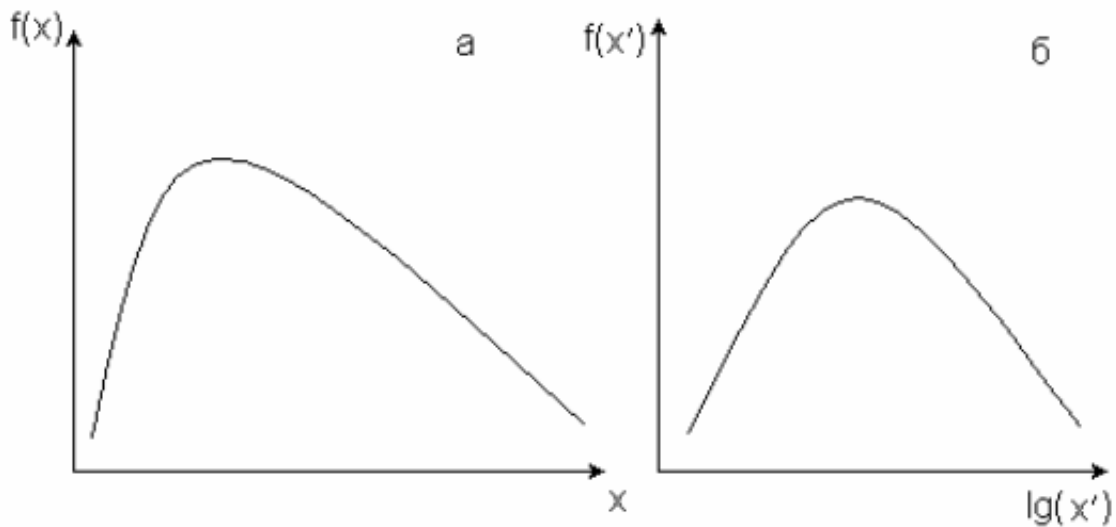


Рис. 3.6. Перетворення функції $f(x)$ до нормального розподілу.

Асиметричний розподіл з однією вершиною приводиться до нормального перетворенням $x'=\lg(x\pm a)$. В окремих випадках можна застосовувати й інші перетворення.

Контрольні питання

1. У чому полягає суть точкового оцінювання?
2. Поясніть як розраховується довірчий інтервал навколо середнього значення фізичної величини, який із заданою ймовірністю накриє дійсне значення математичного очікування M_x .
3. Поясніть, який зміст мають коефіцієнти Стьюдента.
4. Для чого використовується розподіл Пірсона (χ^2 розподіл).
5. Поясніть як вибрати необхідну кількість дослідів, необхідних для побудови заданих довірчих інтервалів для математичного очікування.
6. Опишіть процедуру відсіву грубих похибок спостережень з використанням критерію Смирнова-Граббса.
7. В якому випадку розбіжність між середніми значеннями, одержаними у двох серіях вимірювань однієї і тієї ж величини є значущим, а в якому природними для нормального розподілу випадкової величини?
8. Опишіть процедуру порівняння дисперсій з використанням критерію Фішера та критерію Кохрена.
9. Охарактеризуйте методики перевірки гіпотези нормального розподілу випадкової величини (графічний метод, критерій Пірсона).

4. АНАЛІЗ ЕМПІРИЧНИХ ЗАЛЕЖНОСТЕЙ

4.1. Характеристика видів зв'язків між рядами спостережень

На практиці сама необхідність вимірювання більшості величин обумовлена тим, що вони не залишаються постійними, а змінюються від зміни інших величин. У цьому випадку метою проведення експерименту є встановлення виду функціональної залежності $y=f(X)$. Для цього повинні одночасно визначатися як значення X , так і відповідні їм значення y , а завданням експерименту є встановлення математичної моделі досліджуваної залежності. Фактично мова йде про встановлення зв'язку між двома рядами спостережень (вимірювань).

Визначення зв'язку включає в себе вибір виду моделі і визначення її параметрів. У теорії експериментів, як вже вказувалось вище, незалежні параметри $X=(x_1, \dots, x_k)$ прийнято називати *факторами*, а залежні змінні y - *відгуками*. Координатний простір з координатами $x_1, x_2, \dots, x_i, \dots, x_k$ називається факторним простором. Експеримент з визначення виду функції

$$y=f(x), \quad (4.1)$$

де x - скаляр, називається однофакторним. Експеримент з визначення функції виду

$$y=f(X), \quad (4.1a)$$

де $X=(X_1, X_2, \dots, X_i, \dots, X_k)$ - вектор у k -мірному просторі - багатфакторним.

Геометричним представленням функції відгуку в факторному просторі є поверхня відгуку. При однофакторному експерименті ($k=1$) поверхня відгуку являє собою лінію на площині, при двофакторному ($k=2$) - поверхня в тримірному просторі.

Зв'язки в загальному випадку є досить різноманітними і складними. Зазвичай виділяють наступні види зв'язків.

Функціональні зв'язки (або залежності) - це такі зв'язки, коли при зміні величини X інша величина y змінюється так, що кожному значенню x_i відповідає цілком певне (однозначне) значення y_i (рис.4.1, а). Таким чином, якщо вибрати всі умови експерименту абсолютно

однаковими, то, повторюючи випробування, отримаємо одну і ту ж залежність, тобто криві ідеально співпадають для всіх випробувань.

Такі умови в реальності не зустрічаються. На практиці не вдається підтримувати сталість умов. При цьому вплив кожного випадкового фактору окремо може бути малим, проте у сукупності вони істотно можуть вплинути на результати експерименту. У цьому випадку говорять про стохастичний (ймовірнісний) зв'язок між змінними.

Стохастичність зв'язку полягає в тому, що одна випадкова змінна у реагує на зміну іншої X зміною свого закону розподілу, наприклад математичним очікуванням, (див. рис. 4.1, б). Таким чином, залежна змінна приймає не одне конкретне значення, а деяке із безлічі значень. Повторюючи випробування, ми будемо отримувати інші значення функції відгуку, і одному і тому ж значенню X в різних реалізаціях будуть відповідати різні значення y . Шукана залежність $y=f(X)$ може бути знайдена лише в результаті спільної обробки отриманих значень X і y .

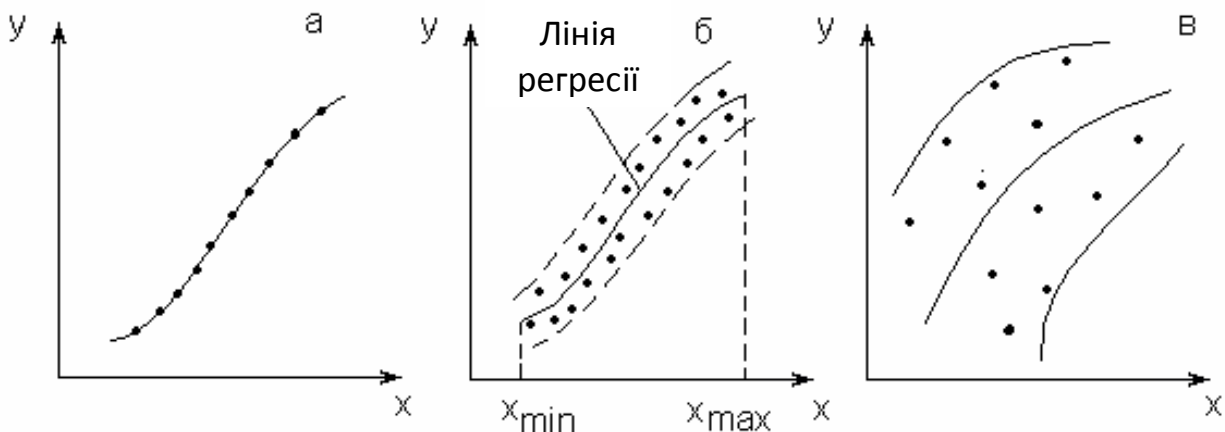


Рис.4.1. Види зв'язків: а - функціональна зв'язок, всі точки лежать на лінії; б - зв'язок досить тісний, точки групуються біля лінії регресії, але не всі вони лежать на ній; в - зв'язок слабкий.

На рис. 4.1, б - це крива, що проходить по центру смуги експериментальних точок (значень математичного сподівання), які можуть і не лежати на шуканій кривій $y=f(X)$, а займають деяку смугу навколо неї. Ці відхилення викликані похибками вимірювань, неповнотою моделі і неврахуванням всіх факторів, випадковим характером самих досліджуваних процесів і іншими причинами.

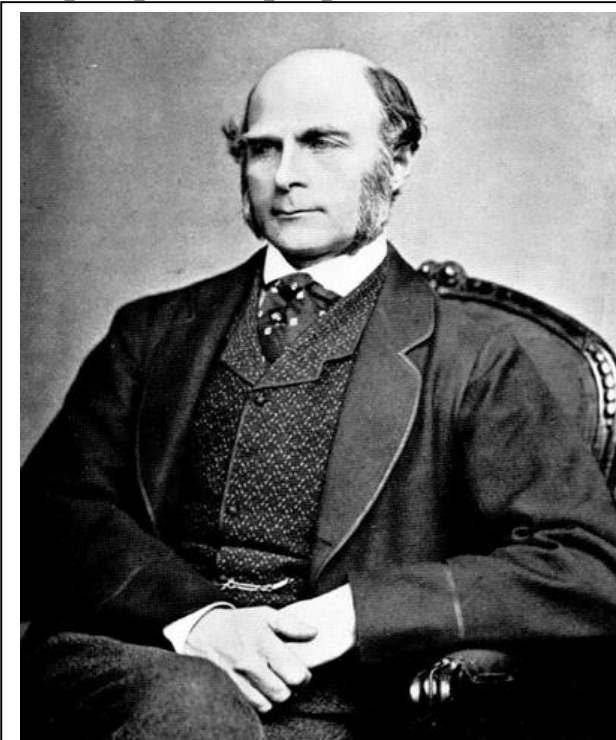
Аналіз стохастичних зв'язків призводить до різних постановок задач статистичного дослідження залежностей, які спрощено можна класифікувати наступним чином:

1) завдання кореляційного аналізу - дослідження наявності взаємозв'язків між окремими групами змінних;

2) завдання регресійного аналізу пов'язані з встановленням аналітичних залежностей між змінною y і однією або декількома змінними $x_1, x_2, \dots, x_i, \dots, x_k$, які носять кількісний характер. Регресійний аналіз (англ. Regression analysis) – це метод визначення відокремленого і спільного впливу факторів на результативну ознаку та кількісної оцінки цього впливу шляхом використання відповідних критеріїв.

3) завдання дисперсійного аналізу - завдання, в яких змінні $x_1, x_2, \dots, x_i, \dots, x_k$ мають якісний характер, а досліджується і встановлюється ступінь їх впливу на змінну y .

Під терміном "регресія" розуміють рух назад, повернення до попереднього стану. Названий термін ввів у кінці XIX ст. Френсіс Гальтон. В результаті проведеного ним дослідження зв'язку між зростом батьків і дітей, виявилось, що наявна обернена залежність. Так, у батьків з дуже високим зростом діти мають менший зріст порівняно з середнім зростом батьків. І, навпаки, у дуже низьких батьків середній зріст дітей вищий. В одному і другому випадку середній зріст дітей прямує (повертається) до середнього зросту населення певної місцевості. Саме такою залежністю і пояснюють вибір терміна "регресія".



Френсіс Гальтон

Регресійний аналіз проводиться на основі побудованого рівняння регресії і визначає внесок кожної незалежної змінної у варіацію досліджуваної (прогнозованої) залежної змінної величини.

Френсіс Гальтон (англ. Francis Galton; нар. 16 лютого 1822 - пом. 17 січня 1911) - англійський дослідник, географ, антрополог, психолог і статистик. Народився в Бірмінгемі у Англії. Гальтон був двоюрідним братом Чарльза

Дарвіна. Сім'я Гальтонів була відомою і досить успішною в сфері виготовлення зброї та банкірській справі. Рано виявив обдарованість: у півтора року знав усі букви алфавіту, самотійно читав з двох з половиною років, писав з трьох років.

З 1838 року навчається медицині: Бірмінгемський госпіталь, медична школа Лондона, в 1839 році «Кінгс-коледж» — медичне відділення. У 1840 році вступив до Кембриджського університету (Триніті-коледж) для занять математикою і природничими науками.

Френсіс Гальтон через смерть батька не завершив медичну освіту, але все життя присвятив науковій діяльності. У 1849 році публікує перше наукове повідомлення, присвячене розробці друкованого телеграфу - «телетайпа». З кінця 1850-х років займається кліматологією і метеорологією. Публікує роботу про клімат Занзібару. Першим починає випускати метеорологічні карти Європи. Відкриває феномен антициклону.

Після виходу книги «Походження видів» свого двоюрідного брата Чарльза Дарвіна став біологом. У 1860-і розробляє проблему успадкування різних ознак у людини і тварин. У 1869 році вийшла з друку його книга «Спадковий геній» — вінець наукової роботи одного з періодів його творчості. Наприкінці 1870-х розробляє методологію психометричних досліджень. Публікує безліч статей, винаходить перші прилади для психометричних дослідів. У 1884 році в Кензінгтоні відкриває першу в світі антропометричну лабораторію. Розробляє методику складових портретів.

У 1892 році його монографія про відбитки пальців «Finger prints» підводить підсумок дослідженням у цій галузі і закладає основні принципи дерматогліфіки. Вказав у класифікації шкірних візерунків на три типи - дуги, петлі, завитки. Проаналізував велику кількість відбитків пальців, отриманих від добровольців через свою лабораторію. Довів неможливість збігу відбитків пальців у людей. Займається біологічної статистикою, першим запропонував те, як обчислити коефіцієнт кореляції. Запропонував закон регресії спадкових ознак.

Стохастичні залежності характеризуються формою, тісністю зв'язку та чисельними значеннями коефіцієнтів рівняння регресії.

Форма зв'язку встановлює вид функціональної залежності $y=f(X)$ і характеризується рівнянням регресії. Якщо рівняння зв'язку лінійне,

то маємо лінійну багатовимірну регресію, в цьому випадку залежність у від X описується лінійною залежністю в k -вимірному просторі:

$$y = b_0 + \sum_{j=1}^k b_j x_j, \quad (4.2)$$

де $b_0, \dots, b_j, \dots, b_k$ - коефіцієнти рівняння. У загальному випадку види функціональних залежностей можуть бути досить різноманітними: показниковими, логарифмічними тригонометричними і т. д. Для пояснення суті використовуваних методів обмежимося спочатку випадком, коли x - скаляр.

Зауважимо, що завдання вибору виду функціональної залежності не формалізується, так як одна і та ж крива на даному інтервалі $[x_{min}; x_{max}]$ приблизно з однаковою точністю може бути описана різними аналітичними виразами. Звідси впливає важливий практичний висновок. Навіть у наш вік комп'ютерів прийняття рішення про вибір тієї чи іншої математичної моделі залишається за дослідником. Тільки експериментатор знає, для чого буде надалі використовуватися ця модель, на основі яких понять будуть інтерпретуватися її параметри.

Вкрай бажано при обробці результатів експерименту вид функції $y=f(X)$ вибирати, виходячи з умови її відповідності фізичній природі досліджуваних явищ або наявним уявленням про особливості поведінки досліджуваної величини. На жаль, не завжди є така можливість, так як експерименти найчастіше проводяться недостатньо точно або для неповністю вивчених об'єктів або явищ.

При вивченні залежності $y=f(x)$ від одного фактора при заздалегідь невідомому виді функції відгуку для приблизного визначення виду рівняння регресії корисно попередньо побудувати емпіричну лінію регресії (рис.4.2).

Для цього весь діапазон зміни x розбивають на рівні інтервали Δx . Всі точки, що потрапили в даний інтервал Δx_j , відносять до його середини - x_j . Підраховують часткові середні для кожного інтервалу:

$$\bar{y}_j = \frac{\sum_{i=1}^{n_j} y_{ji}}{n_j}, \quad (4.3)$$

тут n_j - число точок в інтервалі Δx_j , причому, $\sum_{j=1}^k n_j = n$, де k - число інтервалів розбиття; n - об'єм вибірки. Потім послідовно з'єднують

точки \bar{y}_j ; x_j відрізками прямої. Одержана ламана лінія називається емпіричною лінією регресії. За видом емпіричної лінії регресії можна у першому наближенні підібрати вид рівняння регресії $y=f(x)$.

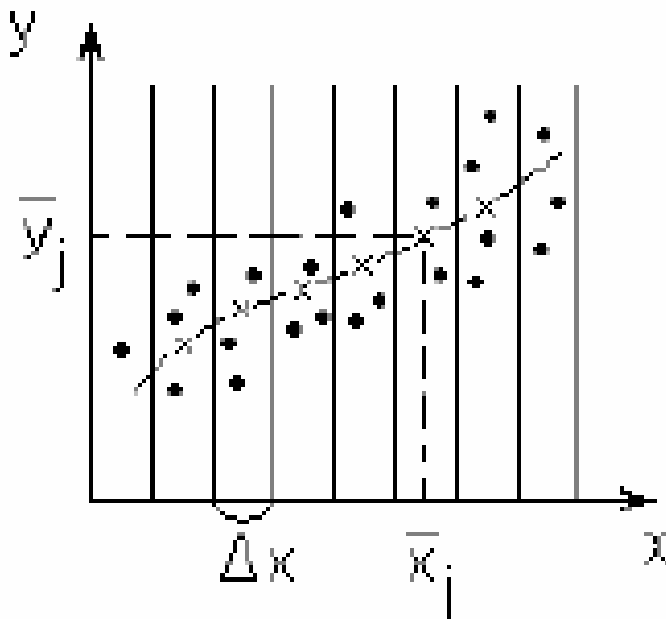


Рис.4.2. Емпірична лінія регресії.

Під тісністю зв'язку розуміється ступінь близькості стохастичної залежності до функціональної, тобто показник тісноти групування експериментальних даних навколо прийнятого рівняння моделі (див. рис. 4.1, б, в). Надалі уточнимо це положення.

4.2. Визначення коефіцієнтів рівняння регресії

Будемо вважати, що вид рівняння регресії $f(X,b)$ вже вибраний і потрібно визначити тільки конкретні чисельні значення коефіцієнтів цього рівняння $b=(b_0, b_1, b_2 \dots b_k)$. Лінія регресії дозволяє з деякою ймовірністю передбачити значення функції y при будь яких значеннях фактора X . Відзначимо попередньо, що якщо вибір виду рівняння регресії, як це вже зазначалося, - процес неформальний і не може бути повністю переданий комп'ютеру, то розрахунок коефіцієнтів вибраного рівняння регресії - операція досить формальна і її слід вирішувати з використанням комп'ютера. Це важкий і виснажливий розрахунок, в якому людина не застрахована від помилок, а комп'ютер виконає його значно швидше і якісніше.

Існує два основні підходи до знаходження коефіцієнтів b_j . Вибір того чи іншого з них визначається цілями і завданнями, що стоять перед дослідником, точністю отриманих результатів, їх кількістю і тому подібне.

Перший підхід - інтерполювання. Базується на задоволенні умови, щоб функція $y=f(X,b)$ якісно відображала експериментальні результати і співпадала з деякими точками, вибраними в якості опорних (основних, головних). У випадку однофакторного експерименту у якості такої функції може бути вибраний, наприклад, поліном k -того степеню $y=b_0+b_1x+b_2x^2+\dots+b_kx^k$. Тоді для визначення $k+1$ невідомих значень параметрів b_j використовується система рівнянь

$$f(x_i, b_0, \dots, b_j, \dots, b_k) = y_i, \quad 1 \leq i \leq m. \quad (4.4)$$

Число незалежних рівнянь системи m дорівнює числу опорних точок, максимальне число яких рівне кількості проведених дослідів n . З іншого боку, для обчислення $k+1$ коефіцієнтів необхідно не менше $k+1$ незалежних рівнянь. Якщо число n проведених дослідів і число незалежних рівнянь дорівнює числу шуканих коефіцієнтів $k+1$, то розв'язок системи може бути однозначним, а отже, точно відповідає випадковим значенням вихідних даних. Таким чином, в граничному випадку, коли число коефіцієнтів рівняння регресії дорівнює числу експериментальних точок $n=k+1$, всі експериментальні точки будуть збігатися з їх розрахунковими значеннями. Слід зауважити, що домагатися такого точного збігу шляхом значного збільшення числа коефіцієнтів рівняння регресії часто просто нерозумно, оскільки експериментальні результати отримані з більшою або меншою похибкою, і така функція може просто не відображати дійсного характеру зміни досліджуваної величини в силу впливу похибок (завад) (рис.4.3).

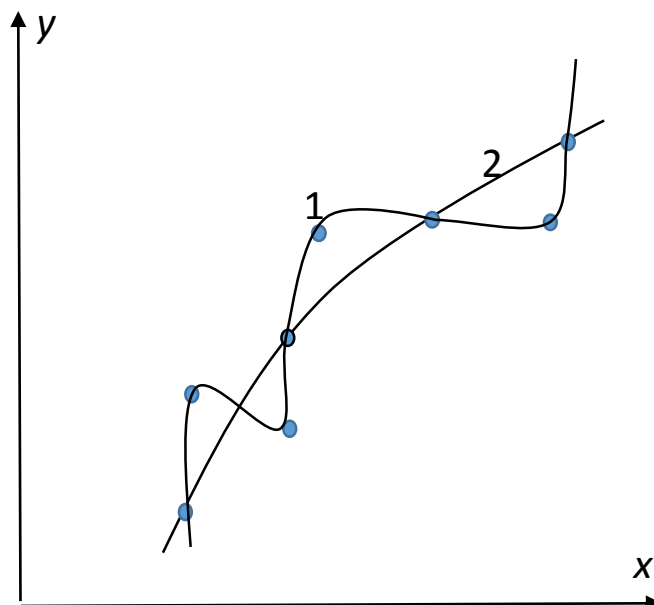


Рис.4.3. Апроксимація функції з великим (1) і невеликим (2) числом коефіцієнтів b_i .

Таким чином, завдання в кінцевому рахунку зводиться до розв'язку системи $k+1$ рівнянь з $k+1$ невідомими. Основна складність такого розв'язку пов'язана з нелінійністю системи, хоча в принципі при використанні комп'ютера вона може бути реалізована.

При числі дослідів n більшому, ніж $k+1$ шуканих коефіцієнтів, число незалежних рівнянь системи $m \leq n$ може бути надлишковим. Надлишковість інформації можна використовувати по-різному.

Після визначення чисельних значень $k+1$ параметрів перевіряється якість апроксимації шляхом зіставлення значень функції і експериментальних даних у невикористаних точках. Якщо виявлені між ними розбіжності перевищують допустимі за умовою точності, то процедуру визначення коефіцієнтів b_j можна повторити, прийнявши в якості опорних (основних) інші точки.

Таким чином, з цих рівнянь в різних комбінаціях можна скласти кілька систем рівнянь, кожна з яких окремо дасть свій розв'язок. Але між собою вони будуть несумісними. Кожний розв'язок буде відповідати своїм значенням коефіцієнтів b_j . Якщо всіх їх побудувати на графіку, то отримаємо цілий пучок апроксимуючих кривих. Це відкриває при $n > k+1$ абсолютно нові можливості. По-перше, цей пучок кривих показує форму і ширину області невизначеності проведеного експерименту. По-друге, може бути здійснене усереднення всіх знайдених кривих і отримана усереднена крива, яка буде набагато точніше і достовірніше описувати досліджуване явище, так як вона в значній степені вільна від випадкових похибок, що приводили до розкиду експериментальних точок. Пояснимо суть цього підходу на прикладі двох методів.

Метод вибраних точок (рис. 4.4). На підставі аналізу даних висувують гіпотезу про вид (форму) залежності $f(X)$. Припустимо, що вона лінійна, тобто статистичний зв'язок - це лінійна одномірна регресія

$$y = b_0 + b_1 x. \quad (4.5)$$

Вибирають дві найбільш характерні на думку дослідника, точки, через які і проходить лінія регресії (рис. 4.4). Завдання обчислення коефіцієнтів b_0 і b_1 у цьому випадку тривіальне. Якщо вибирається рівняння регресії більш високого порядку, то відповідно збільшують число обраних точок. Недоліки такого підходу очевидні, так як обрані точки вибираються суб'єктивно, а переважна частина експериментального матеріалу не використовується для визначення параметрів (коефіцієнтів) рівняння регресії, хоча його можна використати для оцінки надійності отриманого рівняння.

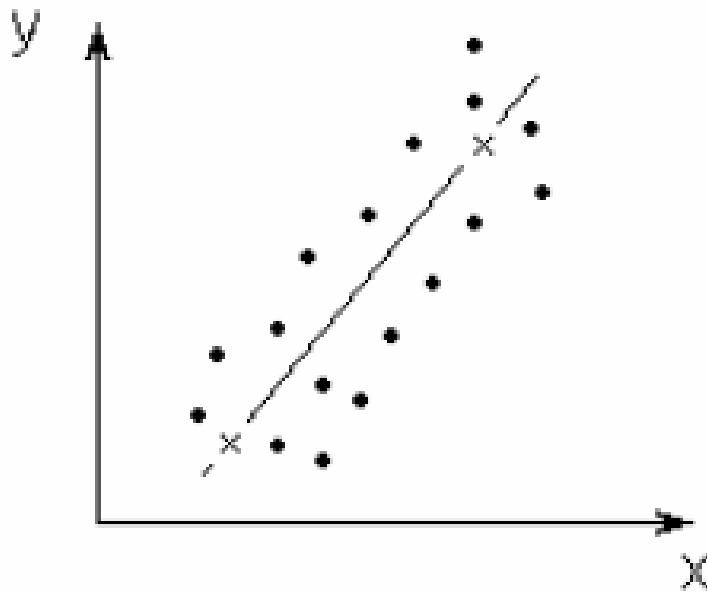


Рис.4.4. Метод вибраних точок: \times - вибрані точки.

Метод медіанних центрів. Суть цього методу пояснює рис. 4.5. Обведене контуром поле точок ділять на кілька частин, число яких дорівнює числу коефіцієнтів рівняння регресії. У кожній з цих частин знаходять медіанний центр, тобто перетин вертикалі і горизонталі зліва і справа, вище і нижче яких виявляється рівне число точок. Потім через ці медіанні центри проводять плавну криву і з розв'язку системи рівнянь визначають коефіцієнти регресії b_j . Так, у разі лінійної залежності (4.5) поле ділиться на дві частини

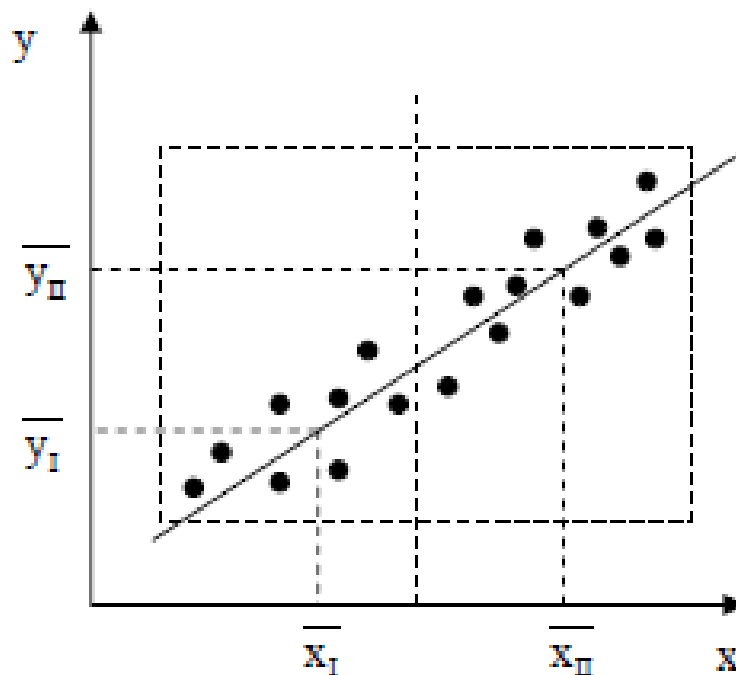


Рис.4.5. Метод медіанних точок.

Визначають середні значення $\bar{y}_I, \bar{x}_I, \bar{y}_{II}, \bar{x}_{II}$ для кожної частини, а невідомі коефіцієнти b_0, b_1 визначають з розв'язку системи рівнянь:

$$\begin{aligned}\bar{y}_I &= b_0 + b_1 \bar{x}_I \\ \bar{y}_{II} &= b_0 + b_1 \bar{x}_{II}\end{aligned}\quad (4.6)$$

Якщо при виборі виду рівняння регресії число його коефіцієнтів виявиться більше числа рівнянь (наявних результатів вимірювань) $k+1 > n$, то система (4.4) не матиме однозначного розв'язку. В цьому випадку необхідно або зменшити число $k+1$ коефіцієнтів b_j , або збільшити число дослідів n .

Другий підхід – аналітичний найбільш простими представниками якого є метод середніх значень і метод найменших квадратів.

Метод середніх значень. Метод вибраних точок включає в себе геометричні побудови з певною довільністю. Тому цей спосіб відшукування параметрів функції досить неточний і його застосовують тільки для попередніх оцінок результатів досліджень.

Ідея методу середніх значень полягає в наступному. Знайдемо для всіх експериментальних точок відхилення

$$f(x_j, b_0, b_1, b_2, \dots, b_k) - y_j = \Delta y_j. \quad (4.7)$$

Відхилення Δy_j представляють собою відстані по вертикалі між експериментальними точками і графіком вибраної функції. При цьому величини Δy_j будуть як із знаком (+), так із знаком (-).

Відповідно до методу середніх значень найкращим буде розміщення лінії вибраної функції при тій умові, коли алгебраїчна сума всіх відхилень Δy_i дорівнює нулю, тобто

$$\sum_i \Delta y_i = 0. \quad (4.8)$$

Для найбільш простого випадку лінійної апроксимації виду $y = bx$

$$\begin{aligned}bx_1 - y_1 &= \Delta y_1, \\ bx_2 - y_2 &= \Delta y_2, \\ &+ \dots \dots \dots \\ bx_n - y_n &= \Delta y_n.\end{aligned}$$

$$b(x_1 + x_2 + \dots + x_n) - y_1 - y_2 - \dots - y_n = 0,$$

звідки

$$b = \frac{y_1 + y_2 + \dots + y_n}{x_1 + x_2 + \dots + x_n}. \quad (4.9)$$

Замітимо, що керуватися сумою відхилень потрібно обережно оскільки вона може виявитись рівною нулю, наприклад, при всіх додатних відхиленнях крім одного великого відхилення від'ємного знаку.

Метод найменших квадратів. Усереднення несумісних рішень надлишкової системи рівнянь $n > k + 1$ може бути подолане методом найменших квадратів, який був розроблений ще Лежандром і Гаусом. Таким чином, метод найменших квадратів - це «новинка» майже 200-річної давності. Сьогодні, завдяки можливостям комп'ютерів, цей метод вступив, по суті, у період свого «ренесансу». Він ґрунтується на тому, що значення параметрів b_0, b_1, \dots, b_k функції $f(x_j, b_0, b_1, \dots, b_k)$ будуть найбільш ймовірними тоді, коли сума квадратів відхилень Δy_j буде мінімальною. При цьому значення відхилень повинні підлягати закону нормального розподілу.

Мірою відхилень є дисперсія σ^2 або її наближений вираз – середній квадрат відхилень

$$\Delta S_n^2 = \frac{1}{n} \sum [y_i - f(x_i, b_1, b_2, \dots, b_k)]^2 = \frac{1}{n} \sum_{i=1}^n (\Delta y_i)^2. \quad (4.10)$$

У цьому виразі параметри b_0, b_1, \dots, b_k вважаємо незалежними змінними. Умовою мінімуму функції багатьох змінних буде рівність нулю частинних похідних першого порядку. З умови мінімуму дістанемо систему k рівнянь з такою ж самою кількістю невідомих

$$-2 \sum_{i=0}^k [y_i - f(x_i, b_0, b_1, \dots, b_k)] \frac{\partial f(x_i, b_0, b_1, \dots, b_k)}{\partial b_j} = 0. \quad (4.11)$$

Система рівнянь (4.11) містить стільки ж рівнянь, скільки невідомих коефіцієнтів b_0, b_1, \dots, b_k входить в рівняння регресії, і називається в математичній статистиці системою нормальних рівнянь.

Розв'язавши цю систему знаходять параметри b_0, b_1, \dots, b_k вибраної функції $f(x, b_0, b_1, \dots, b_k)$.

Зауважимо, що, в принципі, можна оперувати і сумою інших парних степенів відхилень, але тоді обчислення будуть складнішими.

4.3. Визначення тісноти зв'язку між випадковими величинами.

Коефіцієнт кореляції. Критерій Фішера.

Визначивши рівняння теоретичної функції регресії, необхідно дати кількісну оцінку тісноти зв'язку між двома рядами спостережень. Лінії регресії, проведені на рис. 4.1, б, в, однакові, проте на рис. 4.1, б точки значно ближче (тісніше) розташовані до лінії регресії, ніж на рис. 4.1, в.

Тісноту зв'язку між випадковими величинами характеризують кореляційним коефіцієнтом ρ_{xy} . Зупинимось докладніше на фізичному змісті даного показника. Для цього введемо нові поняття.

Залишкова дисперсія $S^2_{зал}$ характеризує розкид експериментально спостережуваних точок відносно лінії регресії і є показником помилки передбачення параметра y за рівнянням регресії (рис. 4.6):

$$S^2_{зал} = \frac{1}{n-l} \sum_{i=1}^n (y_{іексп} - y_{іпрозр})^2 = \frac{1}{n-k-1} \sum_{i=1}^n (y_{іексп} - f(x_i, b_0, \dots, b_k))^2 = \frac{1}{n-k-1} \sum_{i=1}^n (\Delta y_{зал})^2, \quad (4.12a)$$

де $l=k+1$ – число коефіцієнтів b рівняння моделі.

Середній квадрат відхилення лінії регресії S^2_y від середнього значення \bar{y} (лінії $y=C$, тобто для випадку незалежності y від x):

$$S^2_y = \frac{1}{n-1} \sum_{i=1}^n [y_{іпроз} - \bar{y}]^2 = \frac{1}{n-1} \sum_{i=1}^n [f(x_i, b_0, b_1, \dots, b_k) - \bar{y}]^2 \quad (4.12б)$$

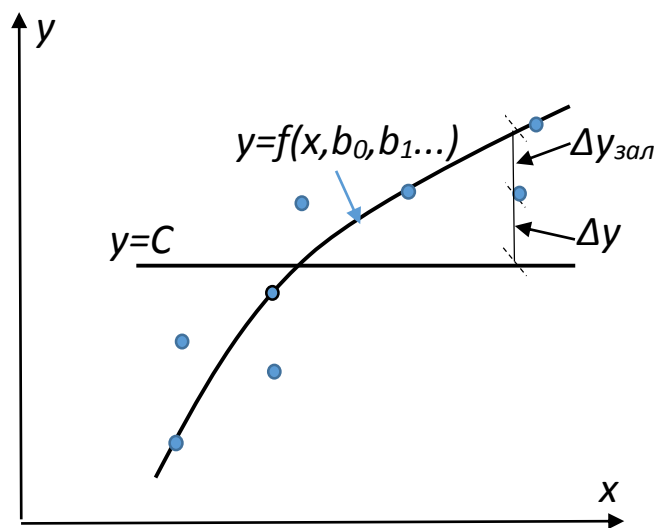


Рис. 4.6. До пояснення коефіцієнта кореляції

Загальна дисперсія S^2 (дисперсія вихідного параметра) характеризує розкид експериментально спостережуваних точок відносно середнього значення \bar{y} , тобто лінії $y = C = \frac{1}{n} \sum_{i=1}^n y_i$ (див. рис. 4.6):

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2. \quad (4.12в)$$

Прийmemo до уваги (без строгого математичного обґрунтування) що загальна дисперсія S^2 (сума квадратів відхилень відносно середнього значення y) дорівнює залишковій дисперсії $S^2_{зал}$ (сумі квадратів відхилень від лінії регресії) плюс середній квадрат відхилення лінії регресії S^2_y

$$S^2 = S^2_y + S^2_{зал} \quad (4.13)$$

Розкид експериментально спостережуваних точок щодо лінії регресії характеризується безрозмірною величиною вибіркоким коефіцієнтом кореляції

$$\rho_{xy} = \frac{\sqrt{S^2 - S^2_{зал}}}{S} = \sqrt{1 - \frac{S^2_{зал}}{S^2}} = \frac{S_y}{S} \quad (4.14)$$

Проаналізуємо властивості цього показника.

У тому випадку, коли зв'язок є не стохастичним, а функціональним, коефіцієнт кореляції дорівнює 1, так як всі точки кореляційного поля виявляються на лінії регресії, залишкова дисперсія дорівнює нулю (рис 4.1а).

Рівність нулю коефіцієнта кореляції вказує на відсутність будь-якої тісноти зв'язку між величинами x і y для даного рівняння регресії, оскільки розкид експериментальних точок щодо середнього значення і лінії регресії однаковий (рис. 4.1в).

Чим ближче розташовані експериментальні дані до лінії регресії, тим тісніший зв'язок, тим менша залишкова дисперсія і тим більший коефіцієнт кореляції. Отже, коефіцієнт кореляції може змінюватися в межах від 0 до 1.

Для функціонального зв'язку поняття кореляції практично не має сенсу (коефіцієнт кореляції завжди дорівнює 1). для стохастичною зв'язку обчислення коефіцієнта парної кореляції ρ між y і x і його статистична оцінка - важлива процедура, результати проведення якої дозволяють судити про тісноту зв'язку.

Якщо змінні y і x представляють двовимірну нормально розподілену випадкову величину, то існує дві регресії: одна визначає залежність y від x , а інша - x від y . Прямі регресії перетинаються в центрі тяжіння і утворюють «ножиці». Чим вужчі «ножиці», тим ближче стохастичний зв'язок до функціонального. При функціональному зв'язку обидві лінії співпадають.

Оцінити тісноту зв'язку за коефіцієнтом кореляції можна скориставшись такою шкалою (Шедока): $\rho < 0,2$ - зв'язку немає; $0,2 \leq \rho < 0,5$ - слабкий зв'язок; $0,5 \leq \rho < 0,75$ - середній зв'язок; $0,75 \leq \rho < 0,95$ - тісний зв'язок; $0,95 \leq \rho < 1$ - дуже тісний.

Для кількісної оцінки достовірності кореляції між змінними x та y використовується вже згадуваний критерій Фішера. Спочатку визначають параметр $F = S^2 / S_{зал}^2$ який в даному випадку показує, у скільки разів рівняння регресії передбачає результати дослідів краще, ніж середнє значення y , тобто паралельна осі x пряма лінія

$$y = \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i = C = const.$$

Якщо $F > F_{\alpha, m_1, m_2}$, то рівняння регресії адекватно і підтверджує нульову гіпотезу H_0 про правильність вибраного виду функціональної залежності $y=f(X)$. Чим більше значення F перевищує F_{α, m_1, m_2} для обраного α і числа ступенів вільності $m_1 = n-1$, $m_2 = n-1$, тим ефективніше рівняння регресії.

З огляду на те, що для комп'ютерів є пакети програм для статистичної обробки результатів досліджень, розглянемо методологію цього підходу на прикладі найпростіших лінійних і одновимірних задач. Ідеологія вирішення складніших завдань принципово не відрізняється. Більш того, як ми побачимо надалі, багато нелінійних залежностей можна звести до лінійних.

4.4. Лінійна регресія від одного фактора

Рівняння лінії регресії на площині в декартових координатах для лінійної залежності має вид виразу

$$y = b_0 + b_1 x. \quad (4.15)$$

Завдання методу найменших квадратів аналітично можна виразити таким чином:

$$\varphi(b_0, b_1) = \sum_{i=1}^n [y_i - (b_0 + b_1 x_i)]^2 \Rightarrow \min \quad (4.16)$$

Для розв'язку цього завдання, як відомо з математичного аналізу, необхідно обчислити частинні похідні функції φ за коефіцієнтами b_0 , b_1 і прирівняти їх до нуля:

$$\frac{\partial \varphi(b_0, b_1)}{\partial b_0} = 0 \quad \text{і} \quad \frac{\partial \varphi(b_0, b_1)}{\partial b_1} = 0. \quad (4.17)$$

Система нормальних рівнянь (4.11) в цьому випадку набуде вид

$$-2 \sum_{i=1}^n [y_i - (b_0 + b_1 x_i)] = 0, \quad \sum_{i=1}^n y_i - n b_0 - b_1 \sum_{i=1}^n x_i = 0; \quad (4.18a)$$

$$-2 \sum_{i=1}^n [y_i - (b_0 + b_1 x_i)] x_i = 0, \quad \sum_{i=1}^n y_i x_i - b_0 \sum_{i=1}^n x_i - b_1 \sum_{i=1}^n x_i^2 = 0. \quad (4.18б)$$

Або

$$\sum_{i=1}^n y_i - n b_0 - b_1 \sum_{i=1}^n x_i = 0; \quad (4.18a)$$

$$\sum_{i=1}^n y_i x_i - b_0 \sum_{i=1}^n x_i - b_1 \sum_{i=1}^n x_i^2 = 0. \quad (4.18a)$$

Розв'язок цієї системи відносно b_0 і b_1 має вид

$$b_0 = \bar{y} - b_1 \bar{x} \quad (4.19a)$$

$$b_1 = \frac{(\overline{xy}) - \bar{x} \cdot \bar{y}}{(\overline{x^2}) - \bar{x}^2} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}, \quad (4.19б)$$

тобто для розрахунку b_0 і b_1 необхідно визначити

$$\sum_{i=1}^n x_i, \quad \sum_{i=1}^n y_i, \quad \sum_{i=1}^n x_i y_i, \quad \sum_{i=1}^n (x_i)^2.$$

Коефіцієнт b_0 (вільний член рівняння регресії) геометрично являє собою відстань від початку координат до точки перетину лінії регресії з віссю ординат, а коефіцієнт b_1 характеризує тангенс кута нахилу лінії регресії до осі Ox .

Якщо ж визначають рівняння регресії у виді квадратного тричлену

$$y = b_0 + b_1 x + b_2 x^2, \quad (4.20)$$

то система рівнянь для знаходження b_0 , b_1 , b_2 матиме такий вид:

$$\sum_{i=1}^n y_i = b_0 n + b_1 \sum_{i=1}^n x_i + b_2 \sum_{i=1}^n x_i^2, \quad (4.21a)$$

$$\sum_{i=1}^n x_i y_i = b_0 \sum_{i=1}^n x_i + b_1 \sum_{i=1}^n x_i^2 + b_2 \sum_{i=1}^n x_i^3, \quad (4.21б)$$

$$\sum_{i=1}^n x_i^2 y_i = b_0 \sum_{i=1}^n x_i^2 + b_1 \sum_{i=1}^n x_i^3 + b_2 \sum_{i=1}^n x_i^4. \quad (4.21в)$$

З рівнянь (4.18) і (4.21) випливає правило запису будь-яких систем нормальних рівнянь: необхідно записати стільки рівнянь в системі, скільки невідомих коефіцієнтів міститься в шуканому рівнянні, кожен раз підсумовуючи добутки членів вихідного рівняння на змінну при шуканому коефіцієнті.

Оцінку сили лінійного зв'язку здійснюють за вибіркоvim (емпіричним) коефіцієнтом парної кореляції ρ_{xy} . Вибірковий коефіцієнт кореляції може бути обчислений двома способами:

1. Як окремий випадок кореляційного відношення для лінійного рівняння регресії.

Враховуючи, що $\bar{y} = b_0 + b_1\bar{x}$

$$\begin{aligned} S_y^2 &= \frac{1}{n-1} \sum_{i=1}^n [y_{i\text{ппоз}} - \bar{y}]^2 =, \\ &= \frac{1}{n-1} \sum_{i=1}^n [b_0 + b_1x_i - b_0 - b_1\bar{x}]^2 = \frac{1}{n-1} b_1^2 \sum_{i=1}^n (x_i - \bar{x})^2 = b_1^2 S_x^2 \end{aligned} \quad (4.22)$$

відношення

$$\rho_{xy} = S_y/S = b_1 S_x/S, \quad (4.23)$$

де S_x та S_y вибіркoві середньоквадратичні відхилення.

2. Як середнє значення добутку центрованих випадкових величин, віднесених до добутку їх середньоквадратичних відхилень:

$$\rho_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{(n-1)S_x S_y} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\left[\sum_{i=1}^n (x_i - \bar{x})^2 \right]} \cdot \sqrt{\left[\sum_{i=1}^n (y_i - \bar{y})^2 \right]}}. \quad (4.24)$$

Покажемо, що дві останні формули еквівалентні. Для цього перетворимо вираз (4.24) до виду

$$\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \rho_{xy} (n-1) S_x S_y. \quad (4.25)$$

Підставляючи (4.25) у вираз (4.19б) одержимо

$$b_1 = \frac{\rho_{xy} (n-1) S_x S_y}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad \text{звідки} \quad \rho_{xy} = \frac{b_1 S_x}{S_y}.$$

Як правило за результатами експерименту знаходять $\bar{x}, \bar{y}, S_x, S_y$ та за формулою (4.24) розраховують ρ_{xy} , а потім, використовуючи ці величини визначають коефіцієнти рівняння регресії

$$b_1 = \rho_{xy} \frac{S_y}{S_x}, \quad b_0 = \bar{y} - b_1 \bar{x}. \quad (4.26)$$

Коефіцієнт кореляції ρ_{xy} змінюється в межах $-1 \leq \rho_{xy} \leq +1$.

Додатна кореляція між випадковими величинами характеризує таку стохастичну залежність між величинами, коли із зростанням однієї з них інша в середньому також буде зростати. При від'ємній кореляції зі зростанням однієї випадкової величини інша в середньому буде зменшуватися. Чим ближче значення ρ_{xy} до одиниці, тим тісніший статистичний зв'язок між x та y .

4.5. Нелінійна регресія

Використовуючи підходи, викладені раніше, можна побудувати практично будь-які форми нелінійної зв'язку. З цією метою в інженерній практиці дуже часто використовують лінеаризуючі перетворення.

У табл. 4.1 наведені залежності, які досить часто зустрічаються, і лінеаризуючі перетворення змінних. Якість перетворення результатів перевіряють за допомогою рівняння $\tilde{y} = b'_0 + b'_1 x'$.

Таблиця 4.1. Функції та лінеаризуючі перетворення

№ п/п	Функція	Лінеаризующие преобразования			
		Преобразование переменных		Выражения для величин b_0 и b_1	
		y'	x'	b'_0	b'_1
1	$y = b_0 + b_1/x$	y	$1/x$	b_0	b_1
2	$y = 1/(b_0 + b_1 x)$	$1/y$	x	b_0	b_1
3	$y = x/(b_0 + b_1 x)$	x/y	x	b_0	b_1
4	$y = b_0 b_1^x$	$\lg(y)$	x	$\lg(b_0)$	$\lg(b_1)$
5	$y = b_0 \cdot e^{b_1 x}$	$\ln(y)$	x	$\ln(b_0)$	b_1
6	$y = 1/(b_0 + b_1 e^{-x})$	$1/y$	e^{-x}	b_0	b_1
7	$y = b_0 x^{b_1}$	$\lg(y)$	$\lg(x)$	$\lg(b_0)$	b_1
8	$y = b_0 + b_1 \lg(x)$	y	$\lg(x)$	b_0	b_1
9	$y = b_0 / (b_1 + x)$	$1/y$	x	b_1/b_0	$1/b_0$
10	$y = b_0 x / (b_1 + x)$	$1/y$	$1/x$	b_1/b_0	$1/b_0$
11	$y = b_0 e^{b_1/x}$	$\ln(y)$	$1/x$	$\ln(b_0)$	b_1
12	$y = b_0 + b_1 x^n$	y	x^n	b_0	b_1

Після обчислення коефіцієнтів b'_0 і b'_1 виконують зворотні перетворення, тобто по b'_0 і b'_1 визначають b_0 і b_1 . Аналогічний підхід зазвичай використовують і при множинному нелінійному регресійному аналізі, розглянутому нижче.

4.6. Лінійна множинна регресія

При вивченні множинної регресії не існує простої графічної інтерпретації багатофакторного простору. При проведенні експериментів у такій ситуації дослідник записує показання приладів про стан функції відгуку y і всіх факторів x_i , від яких вона залежить. Результат досліджень - це матриця спостережень

$$\begin{pmatrix} y_1 & x_{11} & \dots & x_{k1} \\ y_2 & x_{12} & \dots & x_{k2} \\ \dots & \dots & \dots & \dots \\ y_n & x_{1n} & \dots & x_{kn} \end{pmatrix}, \quad (4.27)$$

де n - число дослідів; k - число факторів; x_{ij} - значення j -го фактора в i -тому досліді; y_i - значення вихідного параметра для i -го досліді.

Завдання лінійної регресії полягає в побудові гіперплощини в $(k + 1)$ мірному просторі, відхилення результатів спостережень y_i від якої були б мінімальними при використанні методу найменших квадратів. Або, іншими словами, слід визначити значення коефіцієнтів $b_0, \dots, b_j, \dots, b_k$ в лінійному поліномі

$$y = b_0 + \sum_{j=1}^k b_j x_j, \quad (4.28)$$

які мінімізують вираз

$$\sum_{i=1}^n (y_i - y_{i\text{ппоз}})^2 = \sum_{i=1}^n [y_i - (b_0 + b_1 x_{1i} + \dots + b_k x_{ki})]^2 \Rightarrow \min. \quad (4.29)$$

Процедура визначення коефіцієнтів $b_0, \dots, b_j, \dots, b_k$ в принципі не відрізняється від одновимірного випадку, розглянутого раніше, і тому тут не наводиться.

Для оцінки тісноти зв'язку між функцією відгуку y і декількома факторами $x_1, x_2, \dots, x_j, \dots, x_k$ використовують коефіцієнт множинної кореляції R , який завжди позитивний і змінюється в межах від 0 до 1. Чим більше R , тим якісніше передбачення даною моделлю дослідних даних з точки зору близькості її до функціональної. При функціональній лінійній залежності $R = 1$.

Розрахунки зазвичай починають з обчислення парних коефіцієнтів кореляції, при цьому обчислюються два типи парних коефіцієнтів кореляції:

1) $\rho(y, x_j)$ - коефіцієнти, що визначають тісноту зв'язку між функцією відгуку y і одним з факторів x_j ;

2) $\rho(x_i, x_j)$ - коефіцієнти, що показують тісноту зв'язку між одним з факторів x_i і фактором x_j ($i, j = 1 \div k$).

Якщо один з коефіцієнтів $\rho(x_i, x_j)$ виявиться рівним 1, то це означає, що фактори x_i і x_j функціонально пов'язані між собою. Тоді доцільно один з них виключити з розгляду, причому залишають той фактор, у якого коефіцієнт $\rho(y, x_j)$ більший.

Після обчислення всіх парних коефіцієнтів кореляції можна побудувати матрицю коефіцієнтів кореляції такого виду:

$$\begin{vmatrix} 1 & \rho(y, x_1) & \dots & \rho(y, x_k) \\ \rho(x_1, y) & 1 & \dots & \rho(x_1, x_k) \\ \dots & \dots & \dots & \dots \\ \rho(x_k, y) & \rho(x_k, x_1) & \dots & 1 \end{vmatrix} \quad (4.30)$$

Однак парні коефіцієнти кореляції не характеризують тісноту зв'язку, так як вони обчислюються при значеннях інших факторів, які змінюються випадково. Дійсно, при розгляді трьох і більше випадкових величин коефіцієнти кореляції будь-якої пари з цих випадкових величин можуть не дати правильного уявлення про ступінь зв'язку між усіма випадковими величинами. Це пояснюється тим, що на закон розподілу ймовірностей досліджуваної пари випадкових величин можуть впливати і інші випадкові величини. Ця обставина робить необхідним введення показників стохастичності зв'язку між парою випадкових величин за умови, що значення інших випадкових величин зафіксовані. У цьому випадку говорять про статистичний аналіз часткових зв'язків. Використовуючи матрицю

(4.30), можна обчислити часткові коефіцієнти кореляції, які показують ступінь впливу одного з факторів x_j на функцію відгуку y при умові, що інші фактори залишаються на постійному рівні. Формула для обчислення часткових коефіцієнтів кореляції

$$\rho(y, x_j | x_2, \dots, x_k) = \frac{D_{1j}}{\sqrt{D_{11} \cdot D_{jj}}} \quad (4.31)$$

Тут D_{1j} - визначник матриці, утвореної з матриці (4.30) викреслюванням 1-го рядка і j -го стовпця. Визначники D_{11} і D_{jj} обчислюють аналогічно. Як і парні коефіцієнти, частинні коефіцієнти кореляції змінюються від -1 до +1.

Для обчислення коефіцієнта множинної кореляції використовують матрицю (4.30):

$$R(y, x_1, \dots, x_k) = \sqrt{1 - \frac{D}{D_{11}}}, \quad (4.32)$$

де D – визначник матриці (4.30).

Значимість коефіцієнта множинної кореляції перевіряється за критерієм Стьюдента:

$$t = \frac{R}{S_R} = \frac{R}{(1 - R^2)\sqrt{n - k - 1}} \geq t_{\alpha, m}, \quad m = n - k - 1. \quad (4.33)$$

Значимість R можна перевірити також за критерієм Фішера

$$F = \frac{R^2(n - k - 1)}{(1 - R^2)k}. \quad (4.34)$$

Якщо розрахункове значення F перевищує теоретичне $F_{\alpha; m1; m2}$, то гіпотезу про рівність коефіцієнта множинної кореляції нулю відкидають і зв'язок вважають статистично значущим. Теоретичне (табличне) значення критерію Фішера визначається для вибраного рівня значущості α і числа ступенів вільності $m1 = n - k - 1$ і $m2 = k$.

Якщо коефіцієнт множинної кореляції виявився несподівано малим, хоча апріорно відомо, що між виходом y і входами x_1, \dots, x_k повинен існувати досить тісний кореляційний зв'язок, то можливими причинами такого явища можуть бути наступні:

а) ряд істотних факторів не враховано, і слід включити в розгляд додатково ці суттєві вхідні параметри;

б) лінійне рівняння погано апроксимує насправді нелінійну залежність $y=f(x_1, x_2, \dots, x_k)$. Слід визначити коефіцієнти вже нелінійного рівняння регресії методами регресійного аналізу;

в) робочий діапазон розглянутих факторів не знаходиться в області екстремуму функції відгуку - в цьому випадку слід розширити діапазон змін вхідних змінних, а також перейти до нелінійної математичної моделі об'єкта.

Контрольні питання

1. Проаналізуйте види можливих зв'язків між факторами та відгуком в експериментальному дослідженні.
2. Поясніть термін «регресійний аналіз».
3. Опишіть побудову графічних залежностей методами вибраних точок, медіанних центрів.
4. Поясніть принципи побудови аналітичних залежностей з використанням методів середніх значень та найменших квадратів.
5. Пояснення зміст (значення) коефіцієнта кореляції.
6. Опишіть процедуру обчислення коефіцієнтів лінійної регресії від одного фактора.
7. Опишіть процедуру обчислення вибіркового коефіцієнту кореляції для лінійної регресії від одного фактора.
8. Опишіть побудову лінійної множинної регресії.
9. Проаналізуйте труднощі побудови нелінійної регресії та способи їх вирішення.

5. ОСНОВИ ПЛАНУВАННЯ ЕКСПЕРИМЕНТУ

5.1. Виникнення та становлення планування експерименту

Під *плануванням експерименту* розуміють оптимальне, найбільш ефективне керування ходом експерименту з метою одержання максимально можливої інформації на основі мінімально припустимої кількості досліджуваних даних.

Статистичні методи планування експерименту виникли ще на початку 20 століття. Честь відкриття цієї ідеї належить англійському статистику Рональду Фішеру який з 1918 року почав свою відому серію робіт на Рочемстедській агробіологічній станції в Англії. В 1935 році появилася його монографія “Design of Experiments” (“Конструювання (проектування, планування, дизайн) експериментів”), у який Фішер вперше показав доцільність одночасного варіювання всіх чинників, що впливають на об’єкт експерименту, на противагу поширеному в той час підходу, що полягав у вивченні впливу одного чинника (фактора). Доречі Фішеру належить і термін дисперсія.

Планування експерименту використовують для вивчення та математичного опису процесів і явищ шляхом побудови математичних рівнянь регресії - співвідношень, що пов’язують за допомогою ряду параметрів значення *факторів* і результатів експерименту, які називають, як вже вказувалось вище, *відгуками*.

Сутністю теорії факторного планування є побудова планів проведення експериментів, які дають змогу за результатами вимірювань у певних точках дуже просто реалізувати процедури побудови статистичних висновків про невідомі параметри у більшості випадків поліноміальних функцій регресії. Основна вимога до планів факторного експерименту (на відміну від пасивного, тобто не спланованого, експерименту) - це мінімізація кількості дослідів. У цьому разі завдання обробки результатів факторного експерименту полягає у визначенні числових значень розглядуваних параметрів. Використовуючи *активний* (спланований) експеримент, можна досягти істотно більшого - оптимізувати і стадію постановки експерименту. Увесь експеримент звичайно розбивається на кілька етапів. Інформація, отримана після кожного етапу, використовується для планування досліджень на наступному етапі.

Знадобилося ще кілька десятиріч після ідеї Фішера, щоб на початку 1950-х років з'явився новий напрям у плануванні експерименту, пов'язаний з оптимізацією процесів - *планування екстремального експерименту*. Першу працю в цій галузі 1951 року опублікували в Англії Бокс і Уїлсон (*Box, G.E. and Wilson, K.G. (1951) On the Experimental Attainment of Optimum Conditions. Journal of the Royal Statistical Society, 13, 1-45*). Ідея методу Бокса-Уїлсона вкрай проста. Експериментаторові пропонується ставити послідовні невеликі серії дослідів, які організуються таким чином, щоб після математичної обробки попередньої можна було вибрати умови проведення (тобто спланувати) наступну серію дослідів. Так послідовно, крок за кроком досягається область оптимуму.

Щоб побудувати сучасну теорію планування експерименту, бракувало однієї ланки - формалізації об'єкта дослідження. Ця ланка з'явилася 1947 року після створення Н. Вінером теорії кібернетики. Кібернетичне поняття «чорна скринька», відіграє у плануванні важливу роль. Окрім того, кібернетика стимулювала розвиток обчислювальної техніки, давши в руки дослідників інструмент, що дає змогу швидко розв'язувати математичні задачі, які постають при плануванні.



Норберт Вінер

Норберт Вінер (англ. Norbert Wiener; нар. 26 листопада 1894, Колумбія, Міссурі - пом. 19 березня 1964, Стокгольм) – американський математик-теоретик і прикладний математик. Норберт Вінер народився в єврейській сім'ї. В 4 роки Вінер вже був допущений до батьківської бібліотеки, а в 7 років написав свій перший науковий трактат по дарвінізму. Отримав домашню освіту й не навчався в загальній середній школі. Проте в 11 років від народження він вступив до престижного Тафт-коледжу,

який закінчив з відзнакою вже через три роки отримавши ступінь бакалавра мистецтв.

У 18 років Норберта Вінера зараховано доктором наук за фахом «математична логіка» в корнельському і гарвардському університетах. У дев'ятнадцятирічному віці доктор Вінер був запрошений на кафедру математики масачусетського технологічного інституту.

У 1913 році молодий Вінер починає свою подорож і навчання у Європі. Після початку війни він повертається до Америки.

У 1915 році він намагався потрапити на фронт, але не пройшов медкомісію через поганий зір. З 1919 року Вінер стає викладачем кафедри математики Массачусетського технологічного інституту. В 20-30 роках він знов відвідує Європу. В теорії радіаційної рівноваги зірок з'являється рівняння Вінера-Хопфа. Він читає курс лекцій в пекінському університеті Цинхуа. Серед його знайомих — Н. Бор, М. Борн, Ж. Адамар та інші відомі науковці. В 1926 році одружився на Маргарет Енгерман.

Перед другою світовою війною Вінер став професором Гарвардського, Корнельського, Колумбійського, Браунівського, Геттінгенського університетів, отримав у власне неподільне володіння кафедру в Массачусетському інституті, написав сотні статей по теорії ймовірностей і статистиці,

Під час Другої світової війни він працює над математичним апаратом для систем наведення зенітного вогню силами протиповітряної оборони. У 1945-1947 Вінер зацікавився системами зі зворотним зв'язком і проблемами передачі, зберігання і переробки інформації. Нову науку - загальну теорію управління і зв'язку - він назвав кібернетикою. Творець основ кібернетики, пов'язаною із теорією інформації та теорією керування («батько кібернетики»). Написав дві книги про своє життя: «Колишній вундеркінд» (*Ex-Prodigy: My Childhood and Youth*) і «Я - математик» (*I am a Mathematician*), які містять цікаві відомості про становлення Вінера як ученого.

Сьогодні методи планування та оптимізації експерименту широко застосовуються при вивченні процесів у лабораторних і напівзаводських умовах, а також (щоправда, і трохи рідше) у промисловості.

Мета планування експерименту — відшукування таких умов і правил проведення дослідів, за яких із найменшими витратами праці

вдається дістати надійну й вірогідну інформацію про об'єкт, подавши її в компактній і зручній формі з належною оцінкою точності.

Розв'язання більшості проблем у сфері створення нової технології пов'язана з виконанням складних і високовартісних експериментів. Для створення нової технології, оптимізації нових і вже відомих процесів доводиться ставити величезну кількість дослідів, а тому планування експерименту є важливою складовою проведення наукових досліджень.

Планування експерименту стрімко розвивається і викликає дедалі більший інтерес у дослідників. Інтерес цілком зрозумілий, адже дуже важливо скоротити кількість дослідів, знайти оптимальні умови перебігу процесів та дістати кількісні оцінки впливу факторів при найменших затратах.

5.2. Основні поняття планування експерименту

Проведення експерименту вважається мистецтвом, якому можна навчитися, але якому не можна навчити. Це твердження справедливе, якщо під експериментом мається на увазі загальний процес наукового дослідження і отримання нових даних, при якому на першій стадії розумової діяльності виникають ідеї, задуми, будуються гіпотези, зважуються різні варіанти втілення задуманого. На другій стадії здійснюється експериментальна перевірка, якій передуює власне планування експерименту. Планування експерименту включає наступні пункти:

1. збір і аналіз апріорної інформації, обґрунтування, розуміння факту необхідності експерименту, його матеріальне забезпечення;
 2. вибір факторів і відгуку або відгуків;
 3. вибір плану (побудова матриці планування, число реплік, спосіб пошуку екстремуму або рандомізації);
 4. власне експеримент;
- І, нарешті, на третій стадії відбувається оцінка одержаних результатів і їх осмислення, а з погляду планування експерименту -
5. вибір методики обробки результатів та аналіз даних експерименту;
 6. формулювання висновків і рекомендацій.

Основна мета планування експерименту – це пошук найкращого, оптимального в деякому розумінні рішення.

Після аналізу відомих фактів про досліджуване явище або процес складається план-програма, яка включає найменування теми дослідження, мету і задачі експерименту, робочу, гіпотезу, методику експерименту, перелік необхідних матеріалів, приладів, установок, список виконавців експерименту, календарний план робіт і кошторис на виконання експерименту. У ряді випадків включають роботи з конструювання і виготовлення нових приладів, апаратів, пристосувань, матеріалів, їх методичне і метрологічне обстеження.

Визначення мети і задачі експерименту - один з найбільш важливих етапів. На основі аналізу інформації, гіпотези і теоретичних розробок обґрунтовують мету і задачі експерименту. Вся наукова інформація дозволяє в тій чи іншій мірі судити про очікувані закономірності процесу, що вивчається, а, отже, і визначити задачі експерименту. Чітко, конкретно обґрунтовані задачі - це великий вклад в їх вирішення. Кількість задач не повинна бути дуже великою (3-4 задачі), тільки у великому дослідженні їх може бути 8-10.

Вибір засобів вимірювань - це обґрунтування необхідності використання для спостережень і вимірювань приладів, устаткування, машин, апаратів і ін. Експериментатор повинен бути добре ознайомлений з вимірювальною апаратурою, яку можна знайти у каталогах на засоби вимірювання, замовити ті, що випускаються серійно приладобудівною промисловістю. При цьому слід звернути увагу на можливі діапазони змін факторів і відгуку, необхідну точність їх визначення. У першу чергу використовують стандартні машини і прилади, що серійно випускаються, робота на яких регламентується інструкціями і іншими офіційними документами. У цих документах прийнято вказувати межу допустимої похибки γ , що означає максимально можливу похибку при рекомендованих умовах роботи приладу. Якщо інструментальна похибка розподілена за нормальним законом, то з такого визначення γ слід, що середньоквадратичне відхилення $\sigma_{\text{інстр}}$ пропорціональне $\gamma/3$ (див. вище правило 3σ). Для аналогових електровимірювальних стрілочних приладів прийнято вказувати клас точності, що записується у виді числа, наприклад, 0,05 або 4,0. Це число дає максимально можливу похибку приладу, виражену у відсотках від максимального значення $X_{\text{макс}}$ величини, вимірюваної в даному діапазоні роботи приладу. Так, для вольтметра, що працює в діапазоні вимірювань 0-30 В, клас точності $\gamma = 1,0$ визначає, що зазначена похибка у положенні стрілки на шкалі не перевищує 0,3 В. Відповідно середнє квадратичне відхилення $\sigma_{\text{інстр}}$

становить 0,1В. Реальна похибка приладу істотно залежить від умов навколишнього середовища, де встановлений прилад. Наприклад, похибка приладів залежить від температури приміщення і відрізняється від паспортної похибки, яка зазвичай наводиться для 20°C. Іншою причиною похибок може бути електромагнітне випромінювання іншого лабораторного обладнання, вібрація установки і т.д. Зазвичай ціна найменшої поділки шкали стрілочного приладу погоджена з похибкою самого приладу. Якщо клас точності використовуваного приладу невідомий, за похибку $\sigma_{\text{інстр}}$ завжди приймають половину ціни його найменшої поділки. Зрозуміло, що при зчитуванні показів недоцільно намагатися визначити частки поділки, так як результат вимірювання від цього не стане точнішим. Межу допустимої похибки цифрового вимірювального приладу розраховують за паспортними даними, що містить формулу для розрахунку похибки саме даного приладу.

Оскільки випадкова похибка зменшується при збільшенні числа вимірювань, доцільно зробити таку кількість вимірювань, щоб вона була меншою інструментальної, тобто випадкової похибкою можна було знехтувати в порівнянні з приладовою похибкою. Однак не завжди можливо здійснити необхідне число вимірювань. В результаті часто доводиться миритися з положенням, коли систематична і випадкова похибки близькі і вони обидві однаковою мірою визначають точність результату вимірювань. На жаль, суворого правила складання систематичної і випадкової похибок немає. Одне з можливих правил оцінки сумарної похибки полягає в тому, що ми умовно вважаємо систематичну похибку розподіленої також за нормальним законом і вважаємо, що зазначена величина цієї похибки відповідає значенню середньоквадратичної похибки, помноженої на коефіцієнт Стюдента при нескінченному числі вимірювань, тобто

$$\Delta_{\text{інстр}} = t_{\alpha, \infty} \sigma_{\text{інст}} = \frac{t_{\alpha, \infty} \gamma X_{\text{макс}}}{3} .$$

У цьому випадку сумарна похибка оцінюється як

$$\Delta = \sqrt{\Delta_{\text{вип}}^2 + \Delta_{\text{інстр}}^2} .$$

Вибір факторів - це встановлення основних і другорядних параметрів, що впливають на досліджуваний процес. *Відгуки* - це

реакції на дію - факторів (вхідних змінних) вихідних характеристик досліджуваного явища або процесу.

Спочатку аналізують розрахункові (теоретичні) схеми процесу. Більшість сучасних процесів характеризується наявністю значної кількості різноманітних відгуків і факторів, що впливають на процес. Необхідно прагнути розглядати випадки з одним відгуком - вихідною характеристикою. При виникненні завдань з багатьма відгуками доцільно для кожного з них побудувати свій план експерименту. Відгук (вихідний параметр) є кількісною характеристикою мети дослідження. При вирішенні екстремальних завдань, пов'язаних з визначенням оптимальних умов протікання процесу, вихідні змінні називають параметрами оптимізації. Бажано, щоб параметр оптимізації мав фізичний зміст, був простим і легко обчислюваним. Представивши процес у виді «чорної скриньки», все різноманіття діючих на його вході параметрів (факторів), як вказано у п.1.2. можна розбити на групи (рис. 5.1).

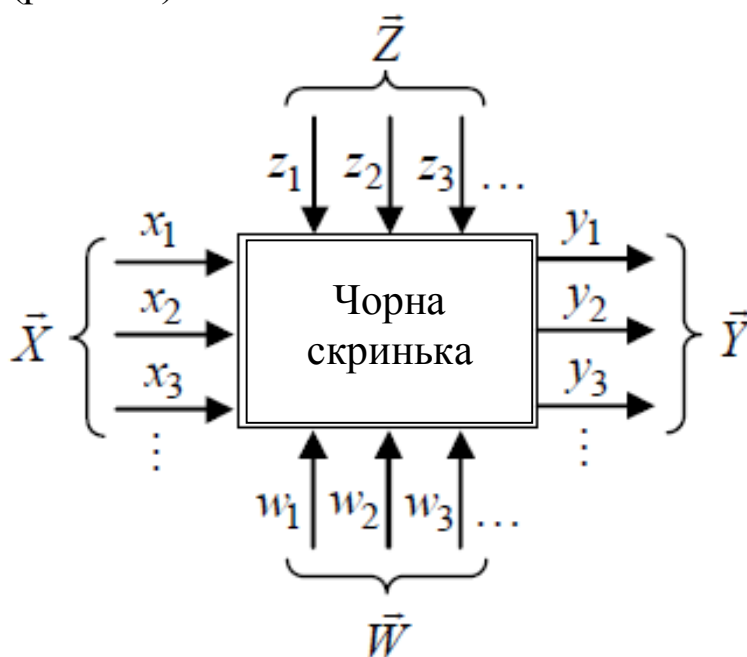


Рис. 5.1. Схема складного процесу.

Вектори відгуку Y є функцією вхідних параметрів. Перша група становить k -мірний вектор X керованих параметрів, тобто таких, які можна вимірювати і цілеспрямовано змінювати. Друга група утворює вектор W контрольованих, але некерованих параметрів, що характеризуються станом функцій відгуку на операціях, що передують досліджуваному процесу (наприклад, чистота вихідного кремнію, алюмінію, які використовуються при виготовленні ІС). Третя група

вхідних параметрів становить вектор Z неконтрольованих, а, отже, і некерованих вхідних параметрів. Сюди відносяться параметри, які надають випадкові впливи на процес.

На основі цього класифікують всі фактори і складають з них ряд по важливості для даного експерименту. Правильний вибір основних і другорядних факторів грає важливу роль в ефективності експерименту, оскільки експеримент зводиться до знаходження залежностей між відгуками і факторами. В окремих випадках важко відразу виявити роль основних і другорядних факторів. При цьому необхідно виконати невеликий за об'ємом попередній пошуковий дослід. Основним принципом встановлення ступеня важливості факторів є їх роль в досліджуваному процесі. Для цього вивчають процес залежно від якоїсь однієї змінної при решті постійних. Такий принцип проведення експерименту виправдовує себе тільки в тих випадках, коли змінних характеристик мало (1-3). Якщо ж змінних величин багато, доцільно застосовувати принципи багатофакторного аналізу.

При *проведення експерименту* припускають, що фактори вибрані і експеримент має за мету визначення функції відгуку. Основу програми експерименту складає її методика. *Методика* є системою прийомів або способів для послідовного найбільш ефективного експериментального дослідження і включає: опис проведення експерименту, обґрунтування засобів і потрібної кількості вимірювань, вибір рівнів факторів.

У методиці детально проектують *процес проведення експерименту*. На початку складають послідовність (черговість) проведення операцій вимірювань і спостережень. Потім ретельно описують кожну операцію окремо з урахуванням вибраних засобів для проведення експерименту. Велику увагу приділяють методам контролю якості операцій, що забезпечують при мінімальній кількості вимірювань високу надійність і задану точність.

При експериментальному дослідженні одного і того ж процесу (спостереження і вимірювання) повторні відліки на приладах, як правило, неоднакові. Відхилення пояснюються різними причинами – неоднорідністю властивостей тіла (матеріалу, конструкції і т.д.), що вивчається, недосконалістю приладів і класом їх точності, суб'єктивними особливостями експериментатора та інше. Чим більше випадкових факторів, що впливають на дослід, тим більше відхилення окремих вимірювань від середнього значення. Це вимагає повторних

вимірювань, отже, необхідно знати їх потрібну мінімальну кількість. Під потрібною мінімальною кількістю вимірювань розуміють таку їх кількість, яка в даному досліді забезпечує стійке середнє значення вимірюваної величини, що задовольняє заданому ступеню точності (що вже було розглянуто у п. 3.2.3). Встановлення потрібної мінімальної кількості вимірювань має велике значення, оскільки забезпечує отримання найбільш об'єктивних результатів при мінімальних витратах часу і засобів.

Вибір рівнів факторів. Фактор може мати безперервний або дискретний характер зміни. Проте, зважаючи на обмежену точність представлення безперервного фактору, він може бути описаний за допомогою дискретного набору рівнів. Це істотно полегшує побудову експерименту і спрощує оцінку його складності.

Планування експерименту починають з вибору центру плану, тобто точки, відповідної початковому значенню всіх використовуваних в експерименті факторів ($X_{10}, X_{20}, \dots, X_{k0}$), в околі якої в подальшому ставиться серія планованих дослідів. Зазвичай за початковий рівень приймають такий стан процесу, який вважається якнайкращим за умовами стійкості, оптимізації і інше. Область експериментування можна розглядати як область зміни факторів в умовах експлуатації. Даний етап неформалізованих рішень виконується на основі інтуїції і досвіду дослідника. Можна привести тільки деякі рекомендації що до вибору рівнів факторів. Очевидно, початковим значенням факторів буде відповідати початкове значення функції відгуку Y_0 . Центр плану зазвичай вибирається на основі апріорних відомостей про процес. Якщо ж їх немає, то зазвичай в якості центру плану приймається центр досліджуваної області. Значення факторів в кожному досліді відрізняється від початкового їх значення на величину інтервалу ΔX . Одним з найважливіших передумов успішного експерименту з метою розробки математичної моделі, є вибір оптимальної величини ΔX . Припустимо, досліджувана функція $Y = f(X_1)$ має вид, наведений на малюнку 5.2 (крива 1).

Якщо вибрати $\Delta X_1'$ невеликим, то при аналізі можна прийти до висновку, що Y незалежний від X_1 . Якщо ж значення $\Delta X_1''$ велике, то можна прийти до неадекватної моделі процесу (крива 2). Заздалегідь передбачити оптимальну величину інтервалу варіювання досить важко. Це залежить від рівня знань експериментатором досліджуваного процесу. Зазвичай інтервал варіювання вибирають в межах 0,05 ... 0,3 від діапазону варіювання досліджуваного фактора.

При цьому величина інтервалу варіювання повинна бути більшою подвоєної квадратичної помилки фіксації факторів. З іншого боку інтервал варіювання не повинен бути дуже великим для правильної апроксимації.

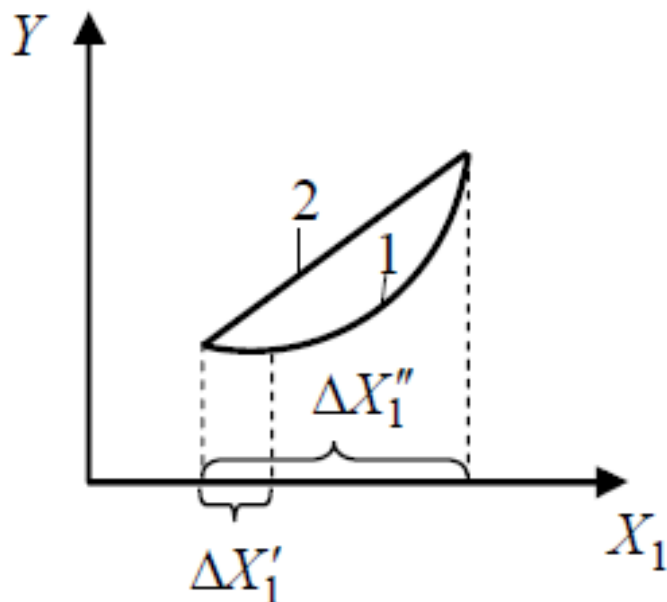


Рис. 5.2. Вид досліджуваної функції (1) і два варіанти кроку експерименту: $\Delta X_1'$ - занижена величина; $\Delta X_1''$ - завищена величина.

Для зручності обробки результатів дослідів, проводиться перетворення значень керованих змінних (тобто факторів X_i , які враховуються в експерименті) до безрозмірних величин ($x_{ij\delta}$)

$$x_{ij\delta} = \frac{X_{ij} - X_{i0}}{\Delta X_i}, \quad (5.1)$$

де X_{i0} — базове або початкове значення i -го фактора в центрі плану;
 ΔX_i — значення інтервалу варіювання i -го фактора,
 X_{ij} — поточний j -ве значення i -го фактора.

Припустимо, що базове значення температури підкладки - одного з факторів досліджуваного процесу отримання резистивних плівок (X_{20}), рівне $X_{20}=400$ °С. При цьому крок варіювання $\Delta X_2=50$ °С, а діапазон зміни фактора X_2 складає 300 - 500 °С. Переходячи від абсолютних значень до безрозмірним, отримуємо відповідно до (5.1) для верхнього рівня $(X_{2\max}-X_{20})/\Delta X_2=(500-400)/50 =+2$, а для нижнього $(X_{2\min}-X_{20})/\Delta X_2=(300-400)/50 =-2$. У безрозмірній системі координат верхній рівень фактора дорівнює +2, а нижній дорівнює -2. Координати центру плану дорівнюють нулю і збігаються з початком координат.

Некеровані фактори в традиційному експерименті або утримуються на одному рівні, або ними нехтують, вважаючи їх вплив незначним. Частку їх впливу на результати дослідів в даному випадку оцінити неможливо.

Складність експерименту визначається числом можливих станів об'єкту дослідження, який визначається кількістю і можливими рівнями факторів. (Наприклад, якщо для всіх k -факторів існує p рівнів, то число станів буде p^k . Так система з 5 факторами на 5 рівнях має 3125 станів, а 10 факторів на 4-х рівнях - $4^{10} = 1049000$.) Прямий перебір зважаючи на величезне число станів нераціональний, навіть іноді неможливий, тому вдаються до процедури планування послідовності вимірювань. Формалізація мети планування виражається у виді деякої функції, яку називають *цільовою функцією*. Побудова цільової функції найбільш відповідальний і найбільш важкий момент всього процесу планування. Коли вона побудована, то діє строгий математичний алгоритм пошуку оптимальних умов реалізації, наприклад, екстремуму відгуку.

Особлива увага в плануванні повинна бути приділена математичним *методам обробки і аналізу* дослідних даних, яка зводиться до систематизації результатів, встановленню емпіричних залежностей, апроксимації зв'язків між варійованими характеристиками, знаходженню критеріїв і довірчих інтервалів та ін. Модель для опису досліджуваного процесу залежить від знань про об'єкт, цілей дослідження і математичного апарату. Наприклад, зміна температури нагріву в часі підкоряється лінійному логарифмічному закону, коливальні процеси описуються тригонометричними функціями і так далі. Якщо вид функції невідомий, то можливе її представлення у виді розкладання в степеневі ряди. За певних умов таке розкладання в багаточлен корисно для всіх безперервних функцій.

Розробку моделі процесу слід проводити за принципом «від простого до складного». Відповідно до цього принципу починають з пропозиції, що модель досліджуваного процесу є лінійно. і має вид полінома 1-го порядку. При цьому враховується вплив на функцію відгуку не тільки кожного розглянутого в експерименті фактора окремо, але і їх взаємодію

$$Y = b_0 + \sum_{i=1}^k b_i X_i + \sum_{i \neq j} b_{ij} X_i X_j \dots \quad (5.2)$$

У цьому рівнянні враховані лише парні взаємодії $X_i X_j$.

Приклад впливу взаємодій факторів представлений на рис. 5.3.

Якщо після обробки результатів експерименту з'ясується, що зроблене припущення про лінійну залежність помилкове, переходять до пропозиції, що модель може бути представлена поліномом 2-го порядку (або іншою функцією, що впливає з фізичної суті, явища або процесу) і т.д., поки не буде розроблена адекватна процесу математична модель.

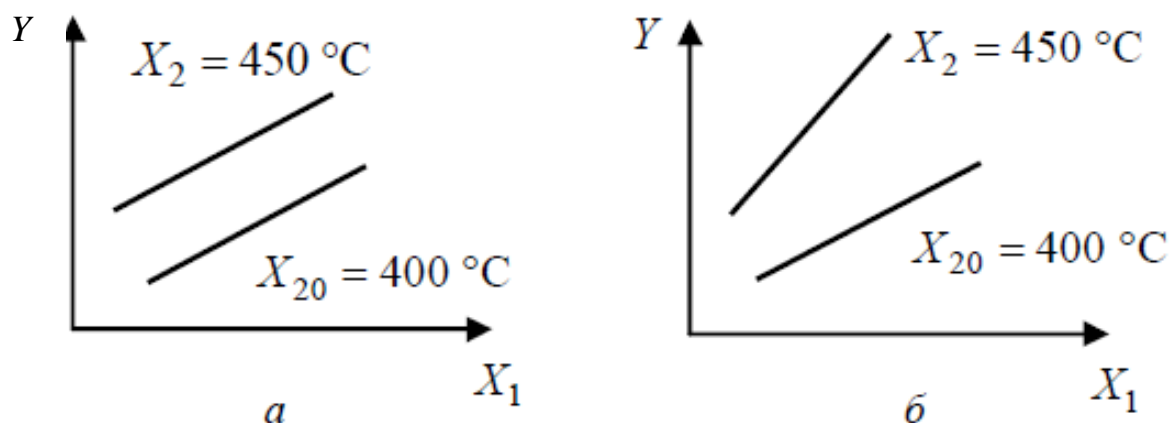


Рис.5.3. Приклад відсутності взаємодії факторів X_1 і X_2 (а) і наявності взаємодії факторів X_1 і X_2 (б).

При інтерпретації результатів дослідження, зрозуміло, що найчастіше працюють саме з 1-ю групою розглянутих вище вхідних параметрів, однак не потрібно забувати і про інших вхідних параметрах. Результати експериментів повинні бути зведені в легкі для читання форми запису - таблиці, графіки, формули, номограми, що дозволяють швидко аналізувати одержані результати.

5.3. Повний факторний експеримент

Під повним факторним експериментом (ПФЕ) розуміють такий експеримент, у якому реалізуються всі можливі комбінації рівнів факторів. У факторних експериментах (на відміну від пасивного, тобто не спланованого, експерименту) відбуваються одночасно варіювання всіма незалежними змінними. Експеримент, у результаті якого всі незалежні змінні варіюються на всіх вибраних рівнях, називається повним факторним експериментом. Як вже вказувалось вище, якщо число факторів k , а число рівнів кожного з них p , то кількість незалежних дослідів відповідає числу комбінацій $N = p^k$, а план

експерименту носить назву типу p^k . Повний факторний експеримент – це перша ланка в ланцюзі планів, послідовне створення і вдосконалення яких привело до розробки математичних методів моделювання складних процесів.

Необхідне число рівнів для факторів вибирають перш за все залежно від порядку математичної моделі, тобто виду рівняння $Y=f(X)$, воно повинне бути, принаймні, на одиницю більше, ніж порядок рівняння. Наприклад, для однофакторної задачі і лінійного рівняння регресії $Y=b_0+b_1X$ для визначення коефіцієнтів b_0 та b_1 число рівнів фактора x має бути не менше 2. (порядок математичного рівняння $Y=f(X)$ становить 1, отже число рівнів $p=1+1=2$). Тому такі плани, (типу 2^k) називають планами першого порядку. Загальне число коефіцієнтів рівняння моделі лінійної регресії (див. р-ня 5.2) при k факторах з урахуванням тільки парних взаємодій буде рівне

$$r = 1 + k + \frac{k^2 - k}{2} = 1 + \frac{k(k + 1)}{2}. \quad (5.3)$$

Наприклад, для двофакторного процесу (X_1, X_2) модель лінійної регресії з врахуванням парних взаємодій має вид $Y=b_0+b_1X_1+b_2X_2+b_{12}X_1X_2$ і очевидно, чотири невідомі коефіцієнти b_0, b_1, b_2, b_{12} , що співпадає вирахованим згідно виразу (5.3)

$$r = 1 + \frac{k(k + 1)}{2} = 1 + \frac{2 \cdot (2 + 1)}{2} = 4.$$

Факторні експерименти типу 2^k володіють властивостями оптимальності при побудові лінійної моделі. Тому за інших рівних умов слід віддавати перевагу лінійним планам. Якщо адекватність лінійної моделі при цьому не буде доведена, тільки потім слід приступати до планів вищого порядку, наприклад з квадратичними членами у функції відгуку, що однак сильно ускладнює як і методику обробки результатів, так і складання плану експерименту. Тому у подальшому основна увага звернута на складанні планів першого порядку.

Для планів першого порядку, як вже зазначалось вище, достатньо мати два рівні факторів. Оскільки фактори різні за фізичною природою і змінюються в різних динамічних діапазонах, для подальшої формалізації процесу аналізу і незалежності одержаних результатів від

зміни масштабу вхідних величин фактори попередньо кодують по аналогії з розглянутим вище методом переведу натуральних факторів у безрозмірні величини за формулою (5.1). При цьому за базове, тобто початкове, значення i -го фактора вибирається середнє значення в діапазоні його зміни $X_{i0} = \frac{X_{i\max} + X_{i\min}}{2}$, а значення інтервалу варіювання i -го фактора становить половину діапазону його зміни $X_{i0} = \frac{X_{i\max} - X_{i\min}}{2}$. Тоді максимальному $X_{i\max}$ і мінімальному $X_{i\min}$ рівням фактору X_i відповідатимуть так звані кодовані (безрозмірні) рівні $x_{i1} = \frac{X_{i\min} - X_{i0}}{\Delta X_i} = -1$ та $x_{i2} = \frac{X_{i\max} - X_{i0}}{\Delta X_i} + 1$. У подальшому будуть використовуватися кодовані змінні x_i .

Побудова матриці планування. Плани ПФЕ 2^k є найпростішими планами першого порядку і необхідні на початковому етапі планування експерименту для отримання лінійної моделі, яка ґрунтується на варіюванні факторів на двох рівнях. За формулою $N=2^k$ легко підрахувати необхідну кількість дослідів при різній кількості факторів. За наявності лише однієї змінної величини (фактору) лінійна математична модель досліджуваного процесу має вид $Y=b_0+b_1X$ і для визначення коефіцієнтів b_0 та b_1 моделі достатньо провести $2^1=2$ експерименти.

Для плану ПФЕ 2^2 кількість факторів дорівнює двом ($k=2$) і кількість рівнів фіксування факторів також 2. При реалізації ПФЕ 2^2 потрібно виконати $N=2^2=4$ дослідів. Плани ПФЕ 2^1 та 2^2 (без взаємодії, тобто добутку) можна записати у виді табл. 5.1. Такі таблиці часто називають *матрицею планування*. Кожний стовпець такої таблиці називають *вектором-стовпцем*, а кожний рядок - *вектором-рядком*. Таким чином, у табл. 5.1 один для 2^1 і два для 2^2 вектори-стовпці незалежних змінних і один вектор-стовпець функції відгуку.

Дані, наведені в табл. 5.1, можна зобразити графічно (рис. 5.4).

Завдяки такому перетворенню за результатами дослідів можна визначити коефіцієнти регресії для лінійної моделі:

$$y=b_0+b_1x_1+b_2x_2, \quad (5.4)$$

виконавши прості обчислення. Для цього підставимо в рівняння (5.4) значення x_1 та x_2 з табл. 5.1б. У результаті дістанемо систему рівнянь, що пов'язують результати дослідів y_n зі значеннями коефіцієнтів регресії b_i (n — номер дослідів; i — номер фактора):

$$\begin{aligned}
 y_1 &= b_0 + (-1)b_1 + (-1)b_2 = b_0 - b_1 - b_2; \\
 y_2 &= b_0 + (-1)b_1 + (+1)b_2 = b_0 - b_1 + b_2; \\
 y_3 &= b_0 + (+1)b_1 + (-1)b_2 = b_0 + b_1 - b_2; \\
 y_4 &= b_0 + (+1)b_1 + (+1)b_2 = b_0 + b_1 + b_2.
 \end{aligned}
 \tag{5.5}$$

Таблиця 5.1. План повнофакторного експерименту 2^1 (а) та 2^2 (б) у кодованих змінних.

а \Rightarrow

№ дослідю, n	Фактор	Функція відгуку y
	x_1	
1.	-1	y_1
2.	+1	y_2

б \Rightarrow

№ дослідю, n	Фактори		Функція відгуку y
	x_1	x_2	
1.	-1	-1	y_1
2.	+1	-1	y_2
3.	-1	+1	y_3
4.	+1	+1	y_4

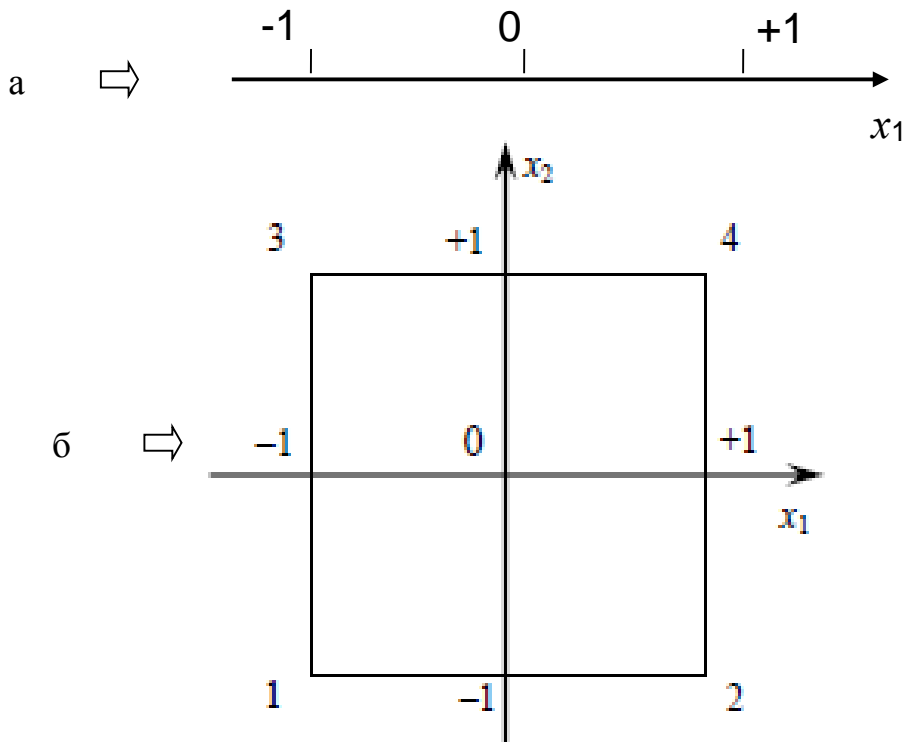


Рис. 5.4. Графічна інтерпретація ПФЕ 2^1 (а) та 2^2 (б) (змінні подано в кодованих значеннях)

Ці стовпці повністю відповідають плану ПФЕ 2^2 (стовпець для визначення b_1 відповідає стовпцю у плані для x_1 , відповідно для b_2 — стовпцю для x_2). Із цих чотирьох рівнянь неважко знайти невідомі коефіцієнти. Для цього рівняння алгебраїчно додають після множення на $+1$ або (-1) . Для обчислення коефіцієнтів регресії дістанемо такі формули:

$$\begin{aligned} b_0 &= \frac{y_1 + y_2 + y_3 + y_4}{4}; \\ b_1 &= \frac{-y_1 + y_2 - y_3 + y_4}{4}; \\ b_2 &= \frac{-y_1 - y_2 + y_3 + y_4}{4}. \end{aligned} \quad (5.6)$$

Із чотирьох рівнянь (5.5) визначено тільки три параметри (b_0, b_1, b_2). Проте в загальному випадку система з чотирьох рівнянь дає змогу визначити чотири невідомі. Отже, залишається ще один степінь вільності, яким можна скористатися для оцінювання взаємодії факторів x_1 та x_2 . Для цього необхідно, скориставшись правилами множення стовпців, дістати стовпець цих двох факторів. При обчисленні коефіцієнта, що відповідає ефекту взаємодії факторів, можна з новим вектором-стовпцем поводитись так само, як і з вектором-стовпцем будь-якого фактора.

Рівняння регресії, що враховує ефекти парних міжфакторних взаємодій, набирає виду:

$$y = b_0 + b_1 x_1 + b_2 x_2 + b_{12} x_1 x_2. \quad (5.7)$$

План ПФЕ 2^2 з урахуванням ефектів взаємодій факторів наведено в табл. 5.2. У ній є ще один вектор-стовпець x_0 . Значення його для всіх дослідів береться $+1$. Він не несе жодної інформації про змінні і введений у табл. 5.2. лише для зручності розрахунків.

Коефіцієнт b_{12} розраховується аналогічно (5.6) з урахуванням знаків стовпця $x_1 x_2$ табл. 5.2:

$$b_{12} = \frac{+y_1 - y_2 - y_3 + y_4}{4}. \quad (5.8)$$

При вивченні форми поверхні відгуку можливі експериментальні похибки $\pm \Delta y_n$ (дія в системі неврахованих факторів, похибки вимірювання), тому на практиці знаходимо лише деяку його оцінку y_n . Через це коефіцієнти регресії, розраховані за наведеними формулами, також визначаються з деякою похибкою.

Таблиця 5.2. План повнофакторного експерименту 2^2 урахуванням взаємодії факторів у кодованих змінних.

№ досліджу	x_0	Фактори		$x_1 x_2$	Функція відгуку y
		x_1	x_2		
1.	+1	-1	-1	+1	y_1
2.	+1	+1	-1	-1	y_2
3.	+1	-1	+1	-1	y_3
4.	+1	+1	+1	+1	y_4
	b_0	b_1	b_2	b_{12}	

Коефіцієнти при незалежних змінних вказують на вагомість впливу факторів на відгук. Чим більше значення коефіцієнта, тим більший вплив фактора. Величина коефіцієнта відповідає внеску даного фактора у значення відгуку під час переходу фактора з нульового рівня на верхній або нижній.

Якщо для двох факторів усі можливі комбінації рівнів легко знайти прямим перебором, то зі зростанням кількості факторів доводиться використовувати алгоритми побудови планів ПФЕ. Із приведеного розгляду випливає, що при побудові плану ПФЕ необхідно користуватись наступним правилом: при додаванні нового фактора кожна комбінація рівнів вихідного плану зустрічається двічі - у сполученні з нижнім (-1) і верхнім (+1) рівнями нового фактора. Іншими словами, матриці вихідного плану треба повторити двічі – при нижньому рівні (-1) і верхньому рівні (+1) доданого фактора. Виходячи з цього, правила можна побудувати матрицю плану трифакторного (і багатфакторного) експерименту.

Наприклад, у плані ПФЕ 2^3 наявність третього фактору у порівнянні з двофакторним приводить до появи третього стовпця x_3 і спричинює утворення чотирьох допоміжних стовпців (x_1x_2 , x_2x_3 , x_1x_3 , $x_1x_2x_3$). За наслідками поставленого за цим планом експерименту можна обчислити вісім коефіцієнтів рівняння регресії.

Властивості матриці ПФЕ типу 2^k

Симетричність матриці планування експерименту відносно центра експерименту. Алгебраїчна сума елементів вектора-стовпця кожного фактора дорівнює нулю:

$$\sum_{i=1}^n x_{i,j} = 0. \quad (5.9)$$

Умова нормування. Сума квадратів елементів кожного стовпця дорівнює кількості дослідів:

$$\sum_{i=1}^n x_{i,j}^2 = n. \quad (5.10)$$

Ортогональність матриці планування. Скалярний добуток двох векторів-стовпців дорівнює нулю для будь-яких двох різних факторів із загальної кількості P :

$$\sum_{i=1}^n x_{j,i} x_{l,i} = 0, \quad j \neq l \quad j, l = 1, 2, \dots \quad (5.11)$$

Рототабельність. Точки в матриці планування добирають так, щоб точність значень параметра оптимізації була однаковою на рівних відстанях від центра експерименту і не залежала від напрямку. План ПФЕ 2^k є також рототабельним, оскільки всі точки плану лежать на колі (або сфері, гіперсфері) центр якого збігається з центром плану.

План ПФЕ 2^k може бути насиченим, якщо кількість членів рівняння регресії дорівнює кількості дослідів, тобто $m+1=n$, або *ненасиченим* при $m+1 < n$. Таким чином, насичений план — це план ПФЕ з мінімально можливою кількістю дослідів. Такі плани мають назву *симплекс-планів*. Симплекс-плани зазвичай використовуються на стадії попереднього дослідження.

5.4. Дробовий факторний експеримент

ПФЕ 2^k є зручним з погляду проведення статистичного аналізу коефіцієнтів рівняння регресії, проте в разі великої кількості факторів ($k > 4$) проведення експерименту за планами ПФЕ стає дуже трудомістким через велику кількість дослідів. Адже кількість дослідів 2^k при значних k істотно перевищує кількість невідомих коефіцієнтів у рівнянні лінійної регресії ($k+1$), яке найчастіше використовується на практиці. Ось чому виникає природне намагання скоротити кількість необхідних дослідів. Вочевидь зі збільшенням кількості дослідів істотно підвищується точність в оцінюванні коефіцієнтів, але можна зменшити кількість дослідів без значної втрати точності. Інакше кажучи, ПФЕ передбачає проведення надлишкової кількості дослідів. Оскільки час і матеріальні ресурси в дослідника звичайно обмежені, застосовувати таку схему на практиці недоцільно. Тому за рахунок

часткової втрати інформації, яка є не суттєвою при побудові лінійної моделі при $k > 3$, кількість вимірювань можна зменшити.

У таких випадках з метою зменшення обсягу експериментальних робіт для математичного опису процесу можна використовувати певну частину повного факторного експерименту: $1/2$; $1/4$ і т.д. Такий метод називається *дробовим факторним експериментом* (ДФЕ), або *дробовою реплікою планів ПФЕ*.

Тоді замість плану ПФЕ 2^k використовують дробовий факторний план ДФЕ 2^{k-m} , виконання якого потребує проведення 2^{k-m} дослідів. Беручи до уваги, що кількість дослідів не може бути меншою за кількість коефіцієнтів рівняння, число m вибирають за умови, що $2^{k-m} \geq k+1$.

Виконуючи ДФЕ, виходять із припущення, що коефіцієнти регресії при взаємодіях вищих порядків (потрійних, четверних і т. д.) мало відрізняються від нуля. Таке припущення правомірне, оскільки експеримент ставиться в достатньо вузькій області поверхні відгуку, яку з деякою похибкою можна подати площиною (лінійне наближення) або, у крайньому випадку, поверхнею другого порядку. Тоді для оцінювання додаткових лінійних ефектів у планах ПФЕ можна використовувати стовпці, що характеризують взаємодії вищих порядків. Наприклад, у плані ПФЕ 2^3 можна використовувати стовпець, що відповідає потрійній взаємодії $x_1x_2x_3$ для включення в експеримент деякого фактора x_4 .

Утворений у такий спосіб план буде так званою *півреплікою від ПФЕ* 2^4 , оскільки замість 16 дослідів, необхідних для вивчення чотирьох факторів плану ПФЕ, потрібно поставити тільки 8. Ця піврепліка позначається звичайно як ДФЕ 2^{4-1} .

Якщо у плані ПФЕ 2^k зменшити кількість факторів на одиницю, то дістанемо план ДФЕ 2^{k-1} , називається *півреплікою*. Якщо у плані ПФЕ відкинути два фактори, то план ДФЕ 2^{k-2} являтиме собою *чвертьрепліку*, якщо відкинути три фактори — то план ДФЕ 2^{k-3} являтиме собою одну восьму репліки і т.д. Кількість дослідів ДФЕ 2^{k-m} дорівнює $N=2^{k-m}$, де 2^m — число, на яке зменшено кількість дослідів порівняно з ПФЕ.

5.5. Складання планів другого порядку

У тих випадках, коли досліджуваний процес або явище суттєво нелінійне, побудована за планом ПФЕ 2^k поверхня відгуку, збігаючись із реальною поверхнею в граничних точках, може відрізнитися в інших точках факторного простору. У таких випадках бажано перейти до планів з більшою кількістю рівнів варіювання факторів, наприклад до планів із варіюванням факторів на трьох рівнях, тобто до планів ПФЕ

3^k . Такі плани мають назву *планів другого порядку* і дають змогу сформуувати функцію відгуку у виді повного квадратичного полінома, однак кількість дослідів при реалізації такого плану різко зростає.

При нелінійному відгуку в рівнянні моделі крім лінійних членів і членів, що враховують взаємодію, необхідно, як мінімум включати й квадратичні члени, які дозволяють відобразити нелінійність яких-небудь перерізів поверхні відгуку. Можливі випадки, коли реальна поверхня відгуку визначається поліномами другого або вищого порядків.

Як і раніше, завдання планування експерименту полягає у визначенні: за вимірними значеннями відгуку y_i коефіцієнтів апроксимуючого полінома, що включає в себе квадратичні члени. Зокрема, при двофакторному експерименті така модель має вид

$$y = b_0 + b_1 x_1 + b_2 x_2 + b_{12} x_1 x_2 + b_{11} x_1^2 + b_{22} x_2^2. \quad (5.12)$$

Спроба оцінити коефіцієнти полінома за планом ПФЕ 2^2 не дає позитивних результатів. Дійсно, при побудові вектор-стовпців для $x_4 = x_1^2$ і $x_5 = x_2^2$ одержимо стовпчики, які співпадають один з одним і з стовпчиком x_0 . Оскільки вказані стовпці не відмінні, неможливо визначити, за рахунок чого формується значення b_0 ; воно залежить як власне від b_0 , так і від внесків квадратичних членів, тобто має місце змішана оцінка. Це твердження справедливе і при більшій кількості факторів. Причина полягає в тому, що для характеристики кривизни поверхні необхідно не менше трьох точок, а дворівневі плани дозволяють встановити тільки дві точки. Таким чином, з планів ПФЕ типу 2^2 неможливо одержати інформацію про коефіцієнти b_{ii} при квадратичних членах і членах більш високого порядку. Це завдання може бути вирішена при переході до плану ПФЕ з більшим числом рівнів варіювання факторів, наприклад до плану ПФЕ типу 3^k . Проте тоді число дослідів стає досить великим навіть при порівняно малому числі факторів.

Більш простим шляхом рішення є побудова плану ПФЕ 2^k (або його дробової репліки) до плану більш високого порядку. В цьому разі план ПФЕ 2^k приймають за ядро або центр плану другого порядку, а потім до нього додають симетрично розташовані допоміжні точки факторного простору, які називаються зірковими. Іншими словами, крім значень факторів на рівнях ± 1 на кожній координатній осі факторного простору вибирають дві зіркові точки $x_i = \pm 1, x_j = 0$, а також додають точку початку координат $x_i = x_j = 0$. Загальне число дослідів у

плані, побудованому таким чином при $k > 1$, складає $n = 2^k + 2k + 1$. Відмітимо, що число дослідів, визначене цим співвідношенням, суттєво менше ніж, наприклад, у плані ПФЕ 3^k при $k > 2$.

План другого порядку, який задовольняє цим умовам, прийнято називати ортогональним центральним композиційним планом (ОЦКП). Однак, побудову ПФЕ 2^k (вибір положення зіркових точок) треба робити таким чином, щоб виконувались принципи ортогональності. Щоб матриця планування була ортогональною, алгебраїчна сума елементів вектор-стовпців повинна дорівнювати нулю не тільки для кожного фактора і їх добутку, але й для вектор-стовпців відповідних квадратів факторів. Очевидно, що виконати останню умову можливо лише в тому разі, якщо квадрати факторів піддати деякому перетворенню, оскільки в протилежному разі сума квадратів будь-яких чисел не дорівнює нулю. Найпростішим перетворенням квадратів факторів є наступні: $x_i^* = x_i^2 - d$, де x_i^* – перетворений фактор; d – постійна величина, що залежить від числа факторів k . Для двофакторного експерименту $d = 2/3 = 0,67$. Більш детальний розгляд цього питання можна знайти у спеціальній літературі. Для прикладу, у таблиці 5.3 приведена матриця ОЦКП другого порядку для двофакторного експерименту.

Таблиця 5.3. Матриця ОЦКП другого порядку для двофакторного експерименту

Параметри	u	x_0	x_1	x_2	$x_3 = x_1 x_2$	$x_4^* = x_1^2 - \frac{2}{3}$	$x_5^* = x_2^2 - \frac{2}{3}$
План ПФЕ 2^2	1	+1	-1	-1	+1	1/3	1/3
	2	+1	+1	-1	-1	1/3	1/3
	3	+1	-1	+1	-1	1/3	1/3
	4	+1	+1	+1	+1	1/3	1/3
Зіркові точки	5	+1	-1	0	0	1/3	-2/3
	6	+1	+1	0	0	1/3	-2/3
	7	+1	0	-1	0	-2/3	1/3
	8	+1	0	+1	0	-2/3	1/3
Нульова точка	9	+1	0	0	0	-2/3	-2/3

5.6. Планування експерименту з пошуку екстремуму відгуку

Одним з можливих завдань планування експерименту є пошук екстремуму функції відгуку. Цей пошук відбувається шляхом дослідження поверхні відгуку. Ці дослідження виконують за допомогою вимірювання поверхні відгуку в різних точках факторного простору. Виникає запитання: якою повинна бути стратегія планування експерименту, щоб число дослідів (вимірювань), необхідних для знаходження екстремуму або близьких до нього значень, було якомога менше.

Кінцевою метою проведення експериментальних досліджень є досягнення екстремальної області й одержання екстремального значення функції відгуку. Вирішення цього завдання виконується у два етапи:

- перший етап – пошуковий рух до області екстремуму,
- другий етап – уточнення екстремальної точки або за допомогою допоміжних пошукових дослідів, або за допомогою математичної моделі області екстремуму, при цьому ця модель отримується за допомогою спеціально організованих дослідів.

Методи пошуку екстремуму при активному експерименті можна розділити на класичні і факторні, деякі з них розглянуті нижче.

5.6.1. Класичні методи визначення екстремуму.

Метод Гауса-Зейделя

За класичними методами, зокрема *методу Гауса-Зейделя*, пошукові досліді для просування до області екстремуму виконують шляхом почергового варіювання незалежних змінних (факторів). При цьому вважається, що всі інші фактори залишаються на цей час незмінними (фіксованими). Зобразимо функцію відгуку досліджуваного об'єкта топографічним способом за допомогою замкнутих ліній постійного рівня (рис. 5.5).

Суть методу Гауса-Зейделя полягає у послідовному просуванні до екстремуму, яке здійснюється шляхом почергового варіювання кожного фактору до досягнення часткового екстремуму вихідної величини. Інакше кажучи, робоча точка X -факторного простору пересувається поперемінно вздовж кожної із координатних осей x_i ; $i=1,2,\dots,k$, факторного простору, причому перехід до нової $(i+1)$ -ї координати здійснюється після досягнення часткового екстремуму

цільової функції $y=(x_1, x_1, \dots x_k)$ на попередньому напрямі, тобто в точці x_{im} , де $\frac{\partial y(x_1, x_2, \dots x_k)}{\partial x_i} = 0$.

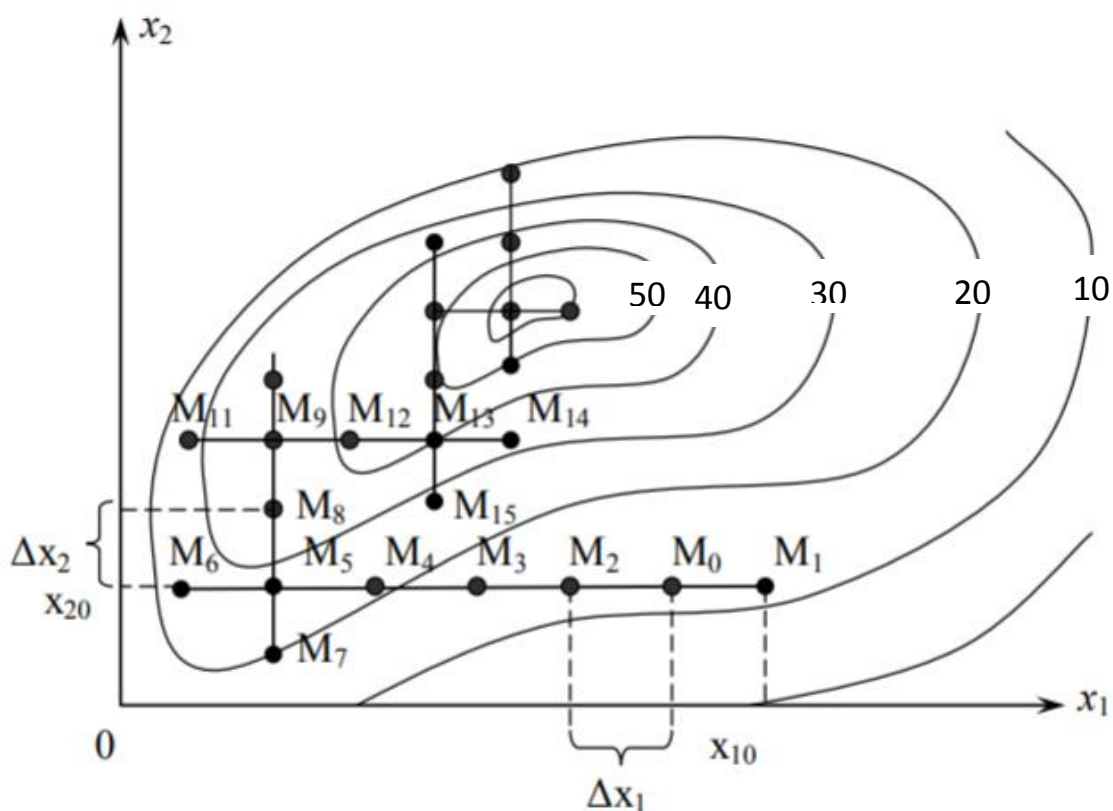


Рис.5.5. Застосування методу Гауса-Зейделя при пошуку точки екстремуму для двофакторної задачі.

Досягнувши часткового екстремуму по останній координаті x_k , переходять знову до варіювання першої і т. д. Таким чином, характерною особливістю методу є необхідність тривалої стабілізації всіх факторів (параметрів процесу), крім одного, за яким відбувається рух. Напрямок руху уздовж $(i+1)$ -ї координатної осі обирається за результатами двох пробних експериментів, які полягають у вимірюванні відклику $y(x_1, \dots x_{i+1}-\Delta x_{i+1}, \dots x_k)$ та $y(x_1, \dots x_{i+1}+\Delta x_{i+1}, \dots x_k)$ в околі базової точки x_{im} , тобто точки часткового екстремуму за попередньою i -змінною. Викладені загальні міркування ілюструються на прикладі двофакторної задачі (рис. 5.5). Тут цифрами 10, 20, 30, позначено лінії рівного рівня вихідного параметра y в деяких відносних одиницях.

При практичному використанні методу Гауса-Зейделя для оптимізації двофакторного процесу бажана така послідовність операцій:

- 1) визначається початкова точка $X=(x_{10},x_{20})$ руху до екстремуму;
- 2) задається крок варіювання Δx_i по кожній незалежній змінній x_i ($i=1,2,\dots,k$);
- 3) здійснюється пробний рух з центром у початковій точці для з'ясування напрямку руху в першому робочому циклі (вздовж осі x_1). З цією метою з базової точки X_0 варіацією параметра x_1 на $\pm\Delta x_1$ виконуються два пробних кроки. Проводиться вимірювання відклику;
- 4) здійснюється порівняння значень відклику;
- 5) здійснюється перший цикл робочого руху (з тим же кроком Δx_1) в напрямку зростання цього відклику у випадку пошуку максимуму, або зменшення відклику і випадку пошуку мінімуму.
- 6) проводиться вимірювання значень відклику після кожного робочого кроку;
- 7) припиняється перший цикл крокового руху після досягнення у деякій точці x_{1m} часткового екстремуму цільової функції по відповідній змінній;
- 8) точка x_{1m} вихідною для нових пробних експериментів по змінній x_2 (у загальному випадку по i -ій змінній). Якщо у пробному русі по i -ій змінній обидва кроки були невдалими, то переходять до варіювання наступним ($i+1$) параметром;
- 9) в подальшому процедура є аналогічною до описаних вище. Після закінчення другого циклу переходять до третього (у випадку двофакторного експерименту знову по осі x_1) і т.д.
- 10) Пошук припиняється в деякій точці X_m подальший будь-який рух від якої призводить до зменшення (якщо досягнуто мінімуму – до збільшення) значення вихідного параметра - відгуку. З точністю до кроків варіювання Δx_i це і буде точка екстремуму цільової функції.

Очевидною перевагою методу Гауса-Зейделя є його наглядність і простота. Проте шлях до головного (глобального) екстремуму звичайно буває довгим, особливо при великому числі n змінних. Крім того, при цьому важко стабілізувати на довгий час всі керовані фактори, крім одного, що викликає допоміжні похибки в знаходженні часткових екстремумів.

5.6.2. Елементи симплекс планування екстремальних експериментів.

Застосування градієнтних методів дозволяє швидше відшукати область екстремуму. Ці методи засновані на попередньому визначенні градієнта функції відгуку $y(X)=y(x_1, x_2, \dots, x_k)$. Методи градієнта відрізняються правилом вибору кроку просування до екстремуму. Суть стратегії всіх градієнтних методів полягає в тому, що на кожному етапі руху до екстремуму біля обраної (або одержаної) базової точки проводять пробні досліди, за результатами яких оцінюють напрямок градієнта у факторному просторі, після чого виконують за обраним робочим кроком просування в нову базову точку, яка буде ближче до екстремуму ніж початкова.

Пошук області оптимуму можна виконати методом, згідно з яким кількість експериментів мінімальна. Сутність такого методу полягає ось у чому. Планують експеримент таким чином, щоб координати точок, які відповідають певним значенням факторів, містились у вершинах правильного симплексу, побудованого у факторному просторі.

У розглянутих раніше планах ПФЕ 2^2 і 2^3 експериментальні точки містяться відповідно у вершинах квадрата та куба. Проте замість квадрата або куба для планування експерименту можна також використати *регулярний симплекс*. Симплексом у k -вимірному просторі називають опуклий многогранник, що має рівно $(k + 1)$ вершин, кожна з яких визначається перетином k гіперплощин даного простору. Симплекс називається *регулярним*, якщо відстані між всіма його вершинами однакові. Прикладами регулярних симплексів є правильний трикутник у двовимірному просторі та тетраедр у тривимірному.

Формально симплекс-метод є типовим пошуковим алгоритмом і заслуговує окремого розгляду. Так само, як і метод найшвидшого спуску, він широко використовується, але, що цікаво, є прикладом методу, збіжність якого теоретично обґрунтувати не вдається. Найпростішим варіантом симплекс-методу є послідовний симплекс метод, що використовує дзеркальне відображення регулярних симплексів відносно граней, протилежних вершині симплексу з найменшим значенням функції відгуку. Після відображення симплексу на кожному кроці розраховуються нові значення факторів.

На практиці планування експерименту з використанням регулярних симплексів застосовується для розв'язування задач оптимізації при русі до майже стаціонарної області.

Практичний досвід використання симплекс-методу показує, що за його допомогою доволі швидко можна визначити область екстремуму. У практичному використанні з кожним кроком розміри симплексу зменшуються, що дає змогу з достатньою точністю визначити область екстремуму.

Розглянемо симплекс-метод на прикладі пошуку оптимуму залежності функції відгуку від двох факторів. Графічно для двох факторів такий регулярний симплекс зображено на рис.5.6.

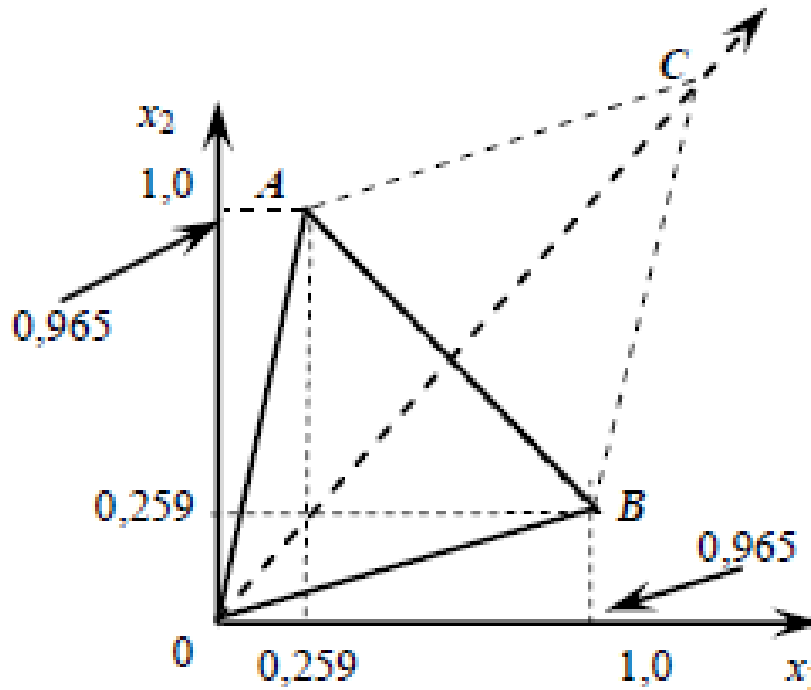


Рис. 5.6. Регулярний симплекс у двовимірному просторі.

При проведенні перших трьох дослідів доцільно вибрати координати однієї з вершин правильного трикутника так, щоб значення кодованих факторів дорівнювали нулю. Значення кодованих факторів у вершинах регулярного симплексу з довжиною ребра $L = 1$ наведено у табл. 5.4. При цьому одна вершина буде знаходитись на початку координат, а вектори, відповідні вершинам А та В регулярного симплексу складуть однакові кути з координатними осями.

Таблиця 5.4. Значення факторів при симплекс-плануванні у двовимірному факторному просторі

Номер дослідів, n	Кодовані фактори		y_n
	x_1	x_2	
1	0	0	y_1
2	0,965	0,259	y_2
3	0,259	0,965	y_3

Оптимізація методом послідовного симплекс-планування виконується в такий спосіб: вихідна серія дослідів планується так, щоб експериментальні точки утворювали регулярний симплекс у факторному просторі. Після проведення дослідів визначається вершина симплексу, що відповідає найгіршим результатам. Далі будується новий симплекс, для чого найгірша точка вихідного симплексу замінюється новою, розміщеною симетрично відносно центра грані симплексу, що перебуває навпроти найгіршої точки.

Наприклад, якщо точка 0 на рис. 5.6 найгірша, то вона замінюється новою точкою С. Нова точка разом із точками, що залишилися, утворює новий симплекс, центр мас якого зміщений у бік збільшення функції відгуку. Після проведення дослідів в наступній точці знову відшукується найгірша вершина симплексу і т.д. (рис. 5.7).

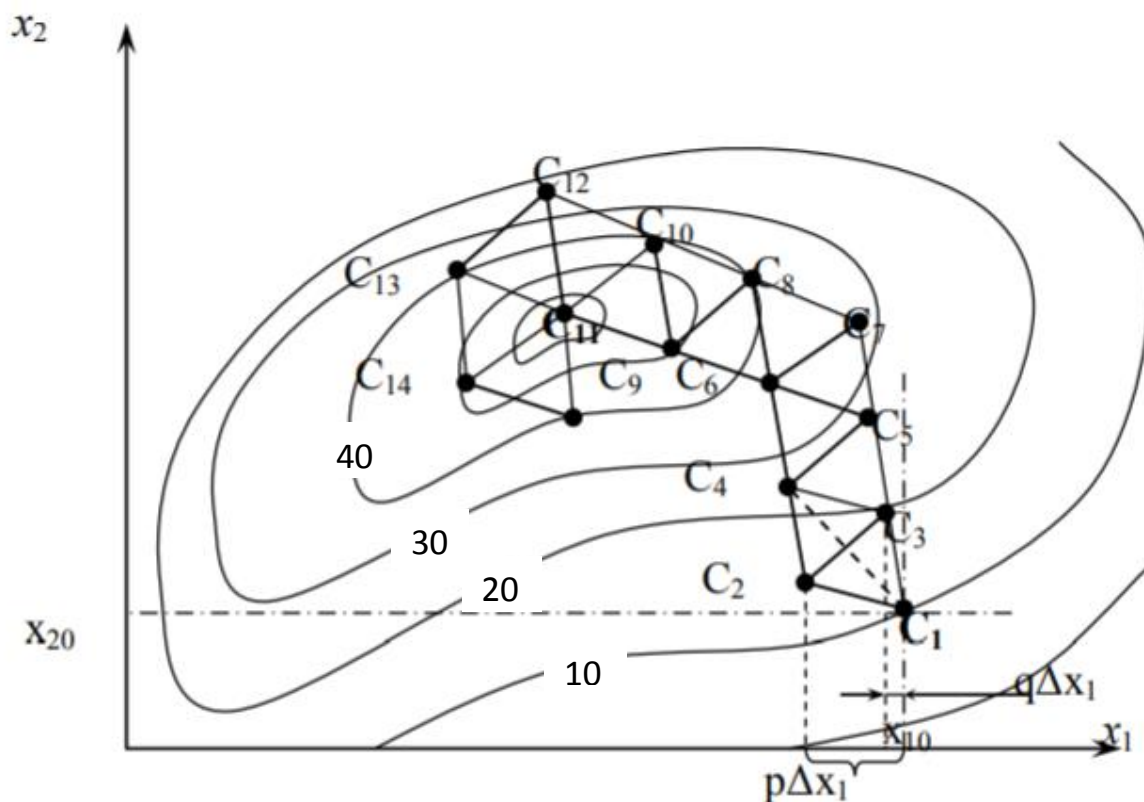


Рис. 5.7. Досягнення екстремуму поверхні відгуку методом симплекс-планування.

Таке багатократне віддзеркалення симплексу приводить до крокового руху центру симплексу до області оптимуму по траєкторії деякої ламаної лінії. При досягненні області оптимуму симплекс починає обертання навколо вершини з максимальним значенням відгуку.

5.6.3. Факторний метод визначення екстремуму Бокса-Вільсона

У 1951 році роботою американських вчених Дж.Бокса і К.Вільсона почався новий етап розвитку планування експерименту. Відомо, що найбільш короткий шлях - це рух по градієнту, тобто перпендикулярно до ліній рівного рівня, на яких функція відгуку приймає постійні значення $y(x_1, x_2, \dots, x_k) = C$.

У зв'язку з цим при оптимізації процесу рух доцільно здійснювати в напрямку найбільш швидкого зростання функції відгуку, тобто в напрямку градієнта функції y .

Існують різні модифікації градієнтного методу, одним з них є метод крутого сходження (МКС) Бокса-Вільсона. Сутність цього методу також розглянемо на прикладі двофакторної задачі. В цьому випадку кроковий рух здійснюється в напрямку найшвидшого зростання функції відгуку, тобто $\text{grad}y(x_1, x_2)$. Однак напрямок коректують НЕ після кожного наступного кроку, а при досягненні в деякій точці на даному напрямку часткового екстремуму функції відгуку (див рис. 5.8).

Серед методів, що розглянуті теорією планування експерименту, метод крутого сходження (будемо розглядати знаходження максимуму функції відгуку) є з'єднуючим між початковим і кінцевим етапами дослідження. Головна ідея методу полягає в тому, що на початковому етапі на основі ПФЕ або ДФЕ одержують самі прості лінійні моделі в якості наближеного опису деякої частини функції відгуку, далекої від екстремуму.

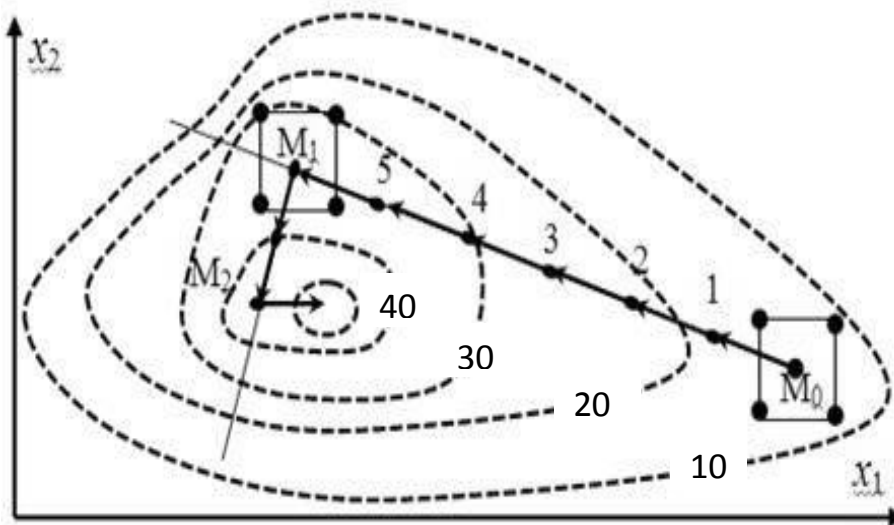


Рис.5.8. Схема планування експерименту методом Бокса- Вільсона.

Коефіцієнти моделі, одержаної в результаті ПФЕ й ДФЕ, пропорційні проекціям вектор-градієнта й дозволяють оцінити саме напрямок градієнта, тобто напрямок самого крутого схилу поверхні відгуку. Потім уздовж цього напрямку відбувається поступовий кроковий рух до області екстремуму (звідси й сама назва методу). Метод крутого сходження (МКС) поєднує кращі властивості класичних методів – градієнтного й Гауса-Зейделя. Схожість із градієнтним методом полягає в тому, що при реалізації МКС відбувається також просування до області екстремуму в напрямку градієнта, знайденого на основі пробних дослідів в околиці базової точки. Але на цьому схожість і закінчується.

За класичним градієнтним методом проводять по два пробних досліді по обидві сторони від базової точки (рис. 5.5) шляхом почергового варіювання кожної із вхідних величин при стабілізації інших $k-1$ факторів, а за МКС пробні досліді проводять відповідно до матриці плану ПФЕ або ДФЕ. За факторним експериментом в оцінці кожного коефіцієнту моделі, а значить, кожної складової градієнта беруть участь усі n точок (дослідів). Тому ці оцінки є більш точними, ніж при класичному градієнтному методі, де кожна складова градієнта обчислюється тільки за двома точками. Знайдений таким чином за МКС напрямок градієнта більш захищений і достовірний. У зв'язку з цим в знайденому напрямку градієнта можна виконати декілька, а не один, пробних кроків до досягнення часткового екстремуму. В цьому полягає схожість МКС з методом Гауса-Зейделя. Проте за класичним методом вважається фіксація $(k-1)$ незалежної змінної при просуванні до екстремуму за варійованою змінною.

Просування виконується уздовж осі - фактора, що сповільнює наближення до області екстремуму. Таким чином, МКС дозволяє значно швидше й надійніше у порівнянні з класичними методами досягти екстремальної області. Крім того, МКС дає змогу одержати інформацію про ступінь крутизни поверхні в районі базової точки – ця інформація закладена в коефіцієнтах взаємодії.

Важливою особливістю МКС є проведення статистичної оцінки результатів ПФЕ, що значно підвищує надійність інтерпретації цих результатів. Припустимо, що за результатами дослідів в області базової точки одержане лінійне рівняння регресії з якого знайдені коефіцієнти регресії можна показати, щ Вище було показано, що оцінки коефіцієнтів b_1 та b_2 пропорційні проекціям вектор-градієнта на осі - фактори.

Таким чином, за оцінками лінійних коефіцієнтів b_1 та b_2 можна оцінити й напрямок вектор-градієнта, за яким за обраним кроком можна здійснювати просування до часткового екстремуму. Про досягнення часткового екстремуму, як за методом Гауса-Зейделя, можна судити за нерівністю $y_{m-1} < y_m > y_{m+1}$. Якщо ця нерівність виконується, точка K_i є точкою часткового екстремуму, її приймають за базову і в околі реалізують ПФЕ або ДФЕ. Для випадку, наведеного на рис. 5.8, при просуванні в напрямку градієнта (промінь K_0M) такою точкою буде точка K_8 .

Перший цикл крутого сходження припиняється після досягнення часткового екстремуму (точка K_8), про що вирішують за реальними дослідями. Тому в міру наближення до часткового екстремуму необхідно частіше проводити перевірочні дослідження, а після проходження часткового екстремуму в ряді випадків поставити допоміжний перевірочний дослід у проміжках між тими двома робочими точками, після яких почалося зменшення вихідної величини і в яких досягнуті приблизно однакові й найбільш із всіх попередніх значень y .

Другий цикл починається з досягнутої точки часткового екстремуму, прийнятої за нову базову точку, в околицях якої ставиться ПФЕ або ДФЕ. У міру наближення до екстремальної області крок варіювання необхідно зменшити. Одержана нова лінійна модель дає новий напрямок градієнта вздовж якого проводяться дослідження «у думці» і перевірочні до досягнення в даному напрямку часткового екстремуму. Свідченням досягнення глобального екстремуму є неадекватність лінійної моделі, одержаної на основі ПФЕ або ДФЕ, коли коефіцієнти при парних взаємодіях різко зростають.

Контрольні питання

1. Які основні завдання процесу планування експерименту?
2. Охарактеризуйте основні етапи планування експерименту.
3. Опишіть формалізовані плани (матрицю) повного факторного експерименту.
4. Опишіть особливості матриці повнофакторного експерименту урахуванням взаємодії факторів.
5. Які основні особливості дробового факторного експерименту.
6. Поясніть класичні методи планування з пошуку екстремуму відгуку на прикладі двохфакторного експерименту. Вкажіть їх недоліки.
7. Опишіть принцип симплекс-методу пошуку оптимуму залежності функції відгуку на прикладі двохфакторного експерименту.

6. ПРАКТИЧНІ ЗАВДАННЯ ОБРОБКИ ДАНИХ З ВИКОРИСТАННЯМ MS Excel

6.1. Первинна обробка експериментальних даних з використання електронних таблиць MS Excel

У цій темі розглянуті основні завдання математичної статистики при первинній статистичній обробці даних. Як правило, статистичні експериментальні дані є результатами спостережень над деякою випадковою величиною x .

Вихідні дані: У результаті експерименту були отримані значення змінної x (наприклад діаметра в мм 10 деталей із великої їх кількості, виготовлених на автоматичному верстаті) (Табл. 6.1.).

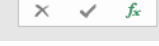
Таблиця 6.1.

i	1	2	3	4	5	6	7	8	9	10
x_i	14,85	14,80	14,84	14,81	14,63	14,81	14,80	14,85	14,84	14,80

6.1.1. Визначення статистичних характеристик випадкових величин

Мета завдання. Обчислити оцінки основних числових характеристик вибірки, одержаних в експериментальних дослідженнях, приведених у табл. 6.1.

Хід виконання завдання

1. Завантажити програму Excel, відкрити пусту книгу.
 1. У комірки A7-A16 увести нумерацію (від 1 до 10), а у комірки B7-B16 увести числові значення x_i . У комірки D7-D13 і E7-E13 увести текстові означення числових характеристик, та відповідні функції для їх визначення (див. рис.6.1.1).
 2. Відмітити курсором комірку F7. У полі  (рис.6.1.1.) вибрати **Поиск функций** наведенням курсора і вибором значка f_x . В одержаному вікні вибрати категорію **Статистические** і функцію **СЧЕТ**.

Статистична функція **СЧЕТ** використовується для визначення числа значень (рис. 6.1.2).

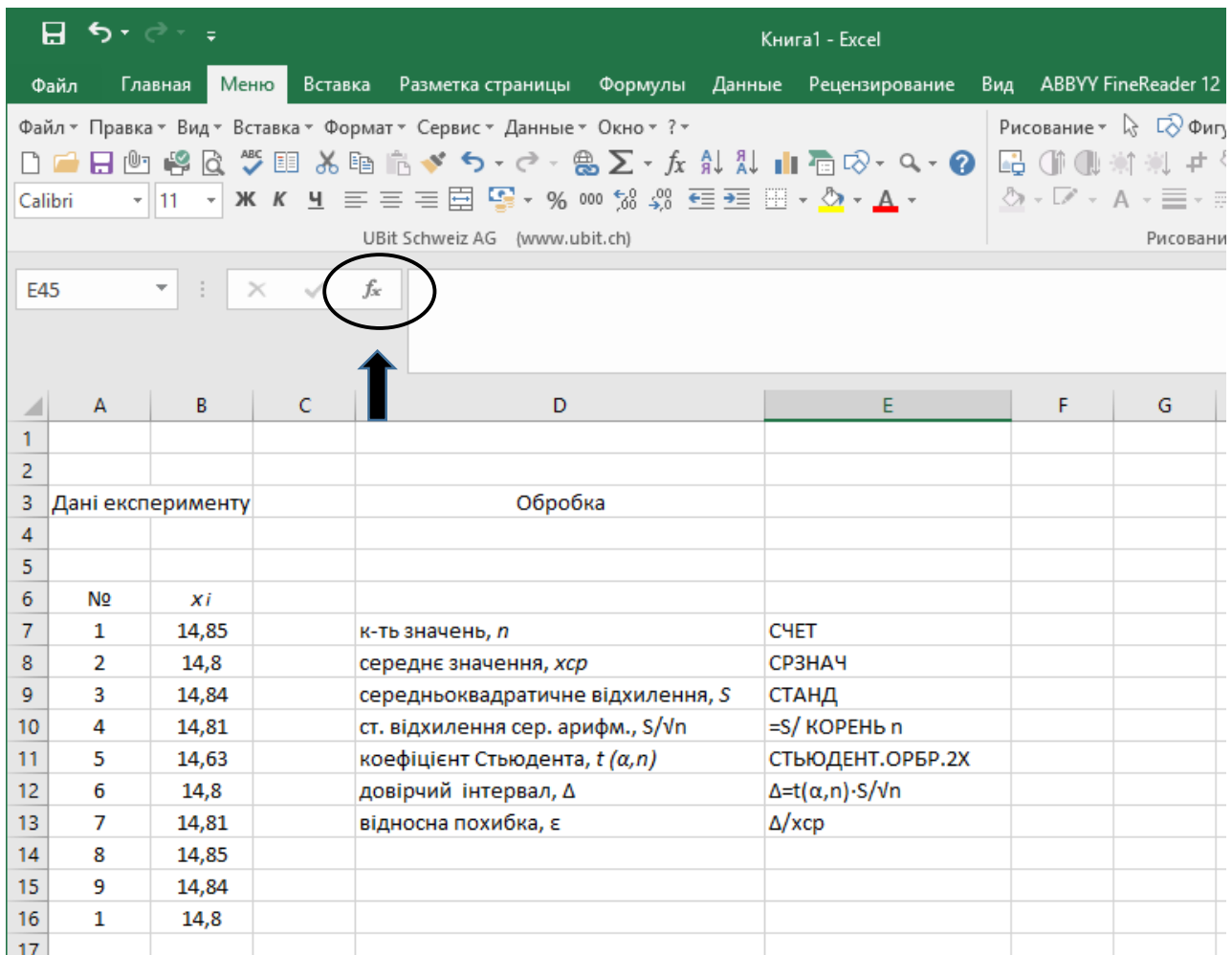


Рис.6.1.1.

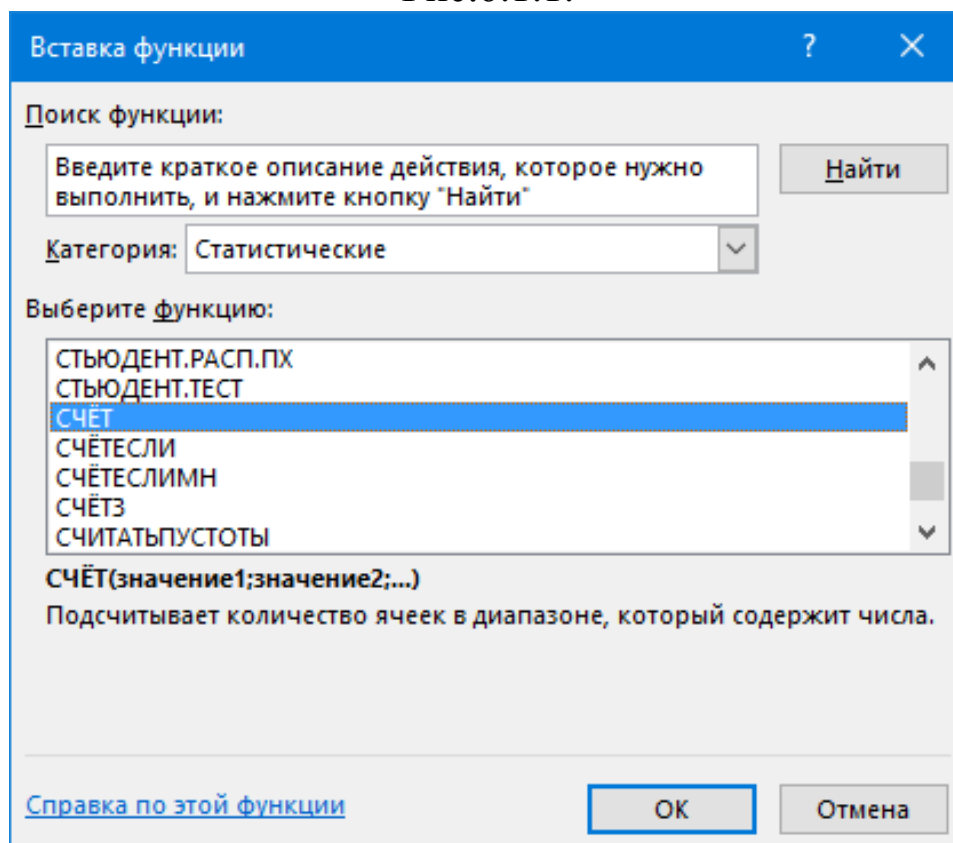


Рис. 6.1.2.

Після натискання ОК одержати вікно (рис. 6.1.3.)

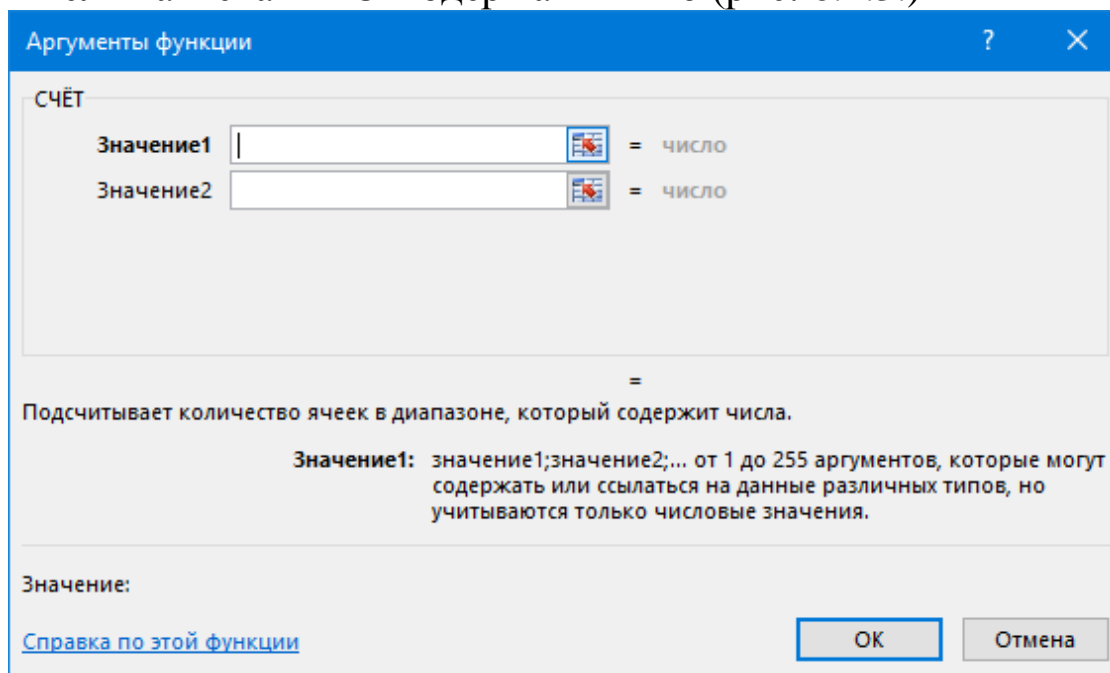


Рис. 6.1.3.

3. Навести курсор на комірку В7 і відмітити комірки В7-В16. У полі **Значение1** при цьому появиться діапазон комірок (вказуються адреси першої і останньої комірки даних, записаних через двокрапку, наприклад **В7: В16**), а також значення шуканої величини. Натисканням ОК одержане значення шуканої величини ($n=10$) появиться у комірці F7.

4. Аналогічним чином розраховуються середнє значення вибірки функцією **СРЗНАЧ**, стандартне відхилення вибірки функцією **СТАНДОТКЛОН**. Аргументами цих функцій служить все той же діапазон комірок.

5. Статистична функція **СТЮДЕНТ.РАСП.ОБР2Х** використовується для знаходження коефіцієнта Стюдента. Цей коефіцієнт залежить від числа ступенів вільності $n-1$ та ймовірності помилки (при довірчій ймовірності p рівній 0,95, яка задає надійність, критерій значимості $\alpha=0,05$, тобто ймовірність помилки становить 5%).

6. Для знаходження стандартного відхилення середнього арифметичного, довірчого інтервалу та відносної похибки використовується звичайні формули «**=F9/КОРЕНЬ(F7)**», «**=F11*F10**» та «**=F12/F8**», які записуються у відміченому на рис.6.1.4. полі. Функція **КОРЕНЬ** вибирається з категорії **Математические**.

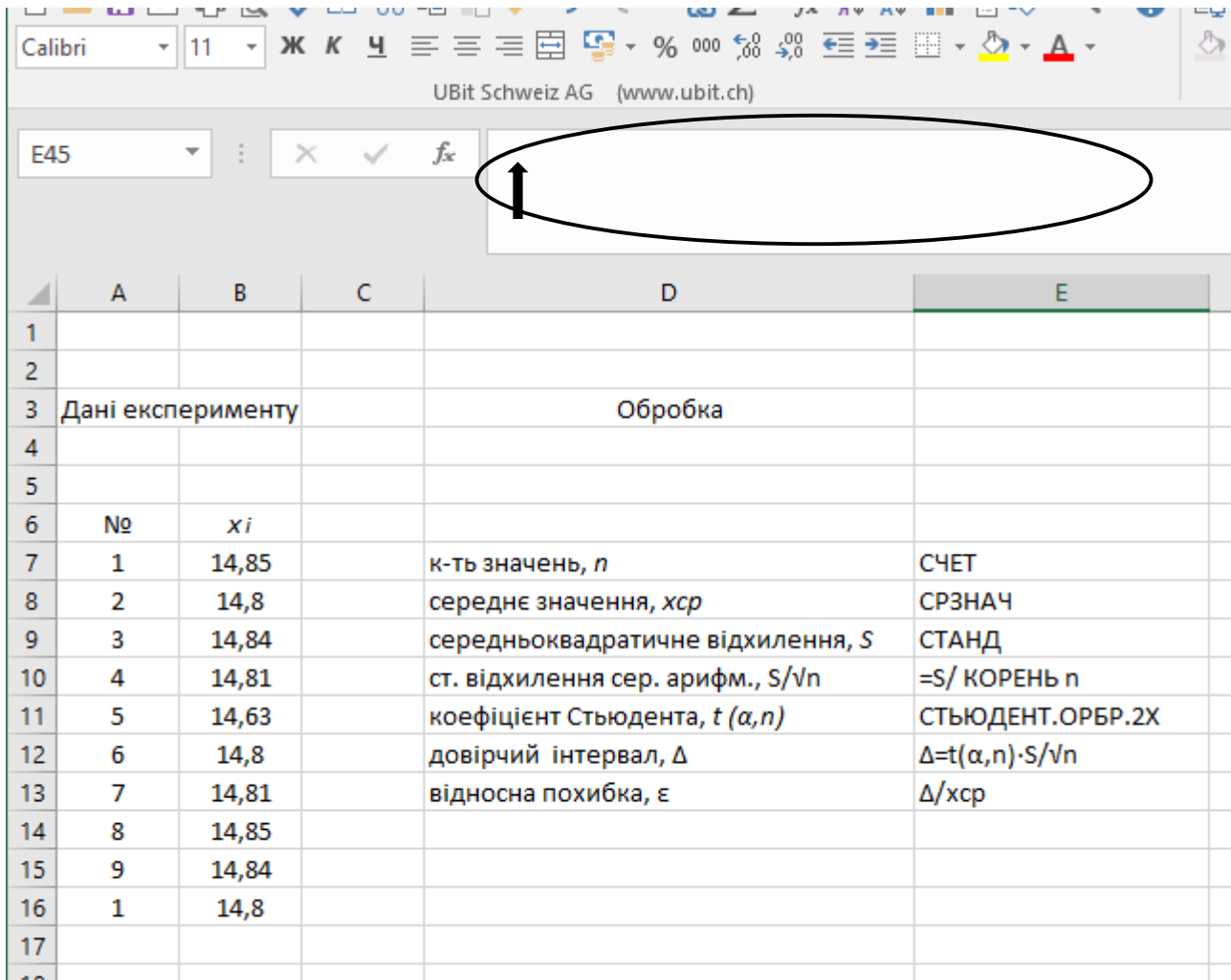


Рис. 6.1.4.

7. У результаті виконання завдання отримати вікно рис. 6.1.5.

	A	B	C	D	E	F	G
1							
2							
3	Дані експерименту		Обробка				
4							
5							
6	№	x_i					
7	1	14,85	к-ть значень, n		СЧЕТ		10
8	2	14,8	середнє значення, $x_{ср}$		СРЗНАЧ		14,803
9	3	14,84	середньоквадратичне відхилення, S		СТАНД		0,0643
10	4	14,81	ст. відхилення сер. арифм., S/\sqrt{n}		=S/КОРЕНЬ n		0,020333
11	5	14,63	коефіцієнт Стьюдента, $t(\alpha, n)$		СТЬЮДЕНТ.ОРБР.2Х		2,262157
12	6	14,8	довірчий інтервал, Δ		$\Delta=t(\alpha, n) \cdot S/\sqrt{n}$		0,05
13	7	14,81	відносна похибка, ϵ		$\Delta/x_{ср}$		0,003107
14	8	14,85					
15	9	14,84					
16	1	14,8					
17							
18							
19							

Рис. 6.1.5.

6.1.2. Відсів грубих похибок

Мета завдання. Провести первинну статистичну обробку даних, перевірити гіпотези про відсутність чи наявність промахів з допомогою критерію Смирнова-Грabbса.

Хід виконання завдання

1. У діапазон комірок **B25: B34** ввести вибіркові значення x_1, x_2, \dots, x_{10} .

2. Побудувати варіаційний ряд. Для цього скопіювати вміст комірок **B25: B34** в комірки **C25: C34**. Впорядкувати вибіркові значення, використовуючи кнопку сортування по зростанню.

3. Знайти мінімальне і максимальне значення x_i у вибірці за першим x_1' і останнім x_{10}' членами варіаційного ряду, або скориставшись функціями EXCEL «МИН», «МАКС».

4. Розрахувати приведені відхилення $u_1 = \frac{\bar{x} - x_1}{S_x}$ та $u_m = \frac{x_n - \bar{x}}{S_x}$. У результаті виконання завдання отримати вікно рис. 6.1.6.

Оскільки $u_1=2,69 \geq 0,73=u_m$, то в подальшому, наявність одного промаху буде визначатись саме для u_1 і, відповідно, першого члена варіаційного ряду $x_1'=14,63$.

№	x_i	x_i'	к-ть значень, n	середнє значення, x_{cp}	середньоквадратичне відхилення, S	мінімальне значення, $x(\min)$	максимальне значення, $x(\max)$	приведене відхилення, u_1	приведене відхилення, u_2	Табл. знач. критерію Смирнова-Грabbса
1	14,85	14,63	СЧЕТ	10						
2	14,8	14,8	SRЗНАЧ	14,803						
3	14,84	14,8	СТАНДОТКЛОН	0,0643						
4	14,81	14,8	МИН	14,63						
5	14,63	14,81	МАКС	14,85						
6	14,8	14,81						$[x(cp)-x(\min)]/S$	2,690528	$u_{1z}(кр)$
7	14,81	14,84						$[x(\max)-x(cp)]/S$	0,730953	$u_{2Su}(кр)$
8	14,85	14,84								Табл. знач. критерію Смирнова-Грabbса
9	14,84	14,85								
10	14,8	14,85								

Рис. 6.1.6

5. Для довірчої ймовірності $p=0,95$, тобто коефіцієнта значимості $\alpha=0,05$, та об'єму вибірки $n=10$ за таблицею критичних значень критерія Смирнова-Граббса $u_{\alpha n}$ знайти критичне значення $u_{0,05;10}=2,18$, тобто визначити критичну область $u \geq u_{кр}=2,18$.

Таблиця критичних значень критерію Смирнова-Граббса $u_{\alpha n}$ у залежності від об'єму вибірки n і рівня значимості α

n	$u_{\alpha n}$		
	$\alpha=0,10$	$\alpha=0,05$	$\alpha=0,01$
3	1,15	1,15	1,15
4	1,42	1,46	1,49
5	1,60	1,67	1,75
6	1,73	1,82	1,94
7	1,83	1,94	2,10
8	1,91	2,03	2,22
9	1,98	2,11	2,32
10	2,03	2,18	2,41
11	2,09	2,23	2,48
12	2,13	2,29	2,55
13	2,17	2,33	2,61
14	2,21	2,37	2,66
15	2,25	2,41	2,70
16	2,28	2,44	2,75
17	2,31	2,48	2,78
18	2,34	2,50	2,82
19	2,36	2,53	2,85
20	2,38	2,53	2,88

6. Оскільки $u_1=2,69 \geq 2,18 = u_{0,05;10}$, то можна зробити висновок, що з ймовірністю 0,95 значення 14,63 випадкової величини x є грубою помилкою або промахом.

6.1.3. Корекція статистичних характеристик

1. Враховуючи приведений вище висновок про грубу помилку у приведених в табл. 6.1.1. значеннях діаметра деталі, за схемою, приведеною у завданні 1, оцінити (перерахувати) статистичні характеристики випадкових величин, без врахування значення 14,63.

Завдання для індивідуальної роботи

При десятикратному вимірюванні напруги за допомогою цифрового вольтметра класу точності 0,1/0,05 на діапазоні 0 - 1000 мВ були отримані результати вимірювання, які наведені в таблиці (за варіантами).

Для даних завдання виключити промахи (грубі помилки вимірювань у кожному варіанті) за допомогою критерію Граббса-Смирнова.

Сумарна похибка змінюється за нормальним законом. Визначити результат вимірювання, його сумарну похибку і записати результат вимірювання. Методичною похибкою можна знехтувати.

Вар. 1	291,9	293,8	290,1	290,2	291,8	292,3	289,7	290,9	289,4	291,6
Вар. 2	394,3	396,2	392,8	399,2	392,2	393,8	394,2	392,7	393,5	393,4
Вар. 3	497,6	493,7	498,6	495,6	496,2	498,1	494,8	496,8	495,2	497,1
Вар. 4	604,5	605,6	607,2	604,8	608,2	606,4	604,3	603,6	607,4	602,8
Вар. 5	259,5	258,9	260,4	256,6	262,4	261,2	260,9	259,8	258,1	258,5
Вар. 6	689,1	690,4	690,8	688,4	690,2	689,5	692,8	688,2	691,7	689,5
Вар. 7	300,8	301,4	300,6	302,6	301,8	300,4	299,4	302,3	298,8	301,2
Вар. 8	588,9	589,7	590,1	588,4	590,7	589,4	590,5	588,3	591,2	589,5
Вар. 9	389,8	389,5	390,1	390,4	388,9	391,9	391,4	392,2	390,5	391,5
Вар. 10	625,7	626,6	628,2	624,5	622,6	623,1	624,9	627,4	628,6	627,4

6.2. Оцінка виду функції розподілу випадкової величини

Вихідні дані:. У банку протягом дня проводилось дослідження часу обслуговування клієнтів у хвилинах, результати якого наведені в табл. 6.2.

Таблиця 6.2.

№ кл.	t, хв.	№ кл.	t, хв.	№ кл.	t, хв.	№ кл.	t, хв.	№ кл.	t, хв.
1	4,8	11	13,4	21	11,0	31	11,5	41	14,9
2	13,0	12	15,1	22	14,8	32	13,0	42	12,5
3	9,1	13	8,9	23	13,3	33	14,5	43	10,7
4	11,0	14	16,5	24	7,3	34	8,5	44	13,5
5	15,0	15	12,6	25	10,6	35	13,5	45	13,0
6	6,7	16	9,0	26	15,2	36	5,2	46	17,5
7	13,2	17	13,0	27	12,7	37	11,0	47	7,0
8	10,5	18	17,0	28	11,4	38	12,5	48	15,0
9	9,2	19	9,5	29	13,0	39	15,5	49	8,8
10	12,8	20	11,5	30	11,0	40	10,5	50	11,3

6.2.1. Оцінка виду функції розподілу випадкової величини за допомогою побудови діаграм та графічним методом

Мета завдання. Провести аналіз розподілу випадкової величини (часу обслуговування клієнтів банку) різними методами.

Хід виконання завдання

Оцінка виду розподілу за допомогою побудови діаграми

Завдання вирішується за допомогою статистичних процедур і статистичних функцій бібліотеки вбудованих функцій *MS Excel*.

Алгоритм виконання завдання.

1. У чистому листку Excel в діапазон комірок **A1: AN** ввести вибіркові значення часу обслуговування клієнтів банку (50 значень з табл. 6.2.)

2. Побудувати варіаційний ряд. Для цього скопіювати комірки **A1: AN** і в комірки **B1: BN** і впорядкувати вибіркові значення від меншого до більшого, використовуючи кнопку сортування по зростанню.

3. Побудувати статистичний ряд вибірки. Для цього у комірки **C1: СК** ввести k різних вибіркових значень, які включають межі варіаційного ряду (наприклад 4, 6, 8, ...14, 16, 18; усього 8 значень i , відповідно 7 інтервалів), як це показано на рис. 6.2.1.

4. У меню **Данные** виділити підменю **Анализ данных**, виділити процедуру **Гистограмма**. У діалоговому вікні **Гистограмма** (рис. 6.2.2) у поле **Входной интервал** ввести посилання на діапазон **A1: AN**. У полі **Интервал карманов** ввести посилання на діапазон **C1: СК**. Активувати поле **Выходной интервал** і ввести в це поле посилання на комірку, в яку буде введена ліва верхня комірка таблиці результатів (наприклад **E1**). Встановити прапорець **Вывод графика**. Після натискання **ОК** отримаємо результат приведений на рис.6.2.3.

5. Оформити діаграму відповідними підписами на осях та назвою. Зробити висновок що до виду функції розподілу клієнтів за часом обслуговування.

	A	B	C	D
1	4,8	4,8	4	
2	13	5,2	6	
3	9,1	6,7	8	
4	11	7	10	
5	15	7,3	12	
6	6,7	8,5	14	
7	13,2	8,8	16	
8	10,5	8,9	18	
9	9,2	9		
10	12,8	9,1		
11	13,4	9,2		
12	15,1	9,5		
13	8,9	10,5		
14	16,5	10,5		
15	12,6	10,6		
16	9	10,7		
17	13	11		

Рис. 6.2.1.

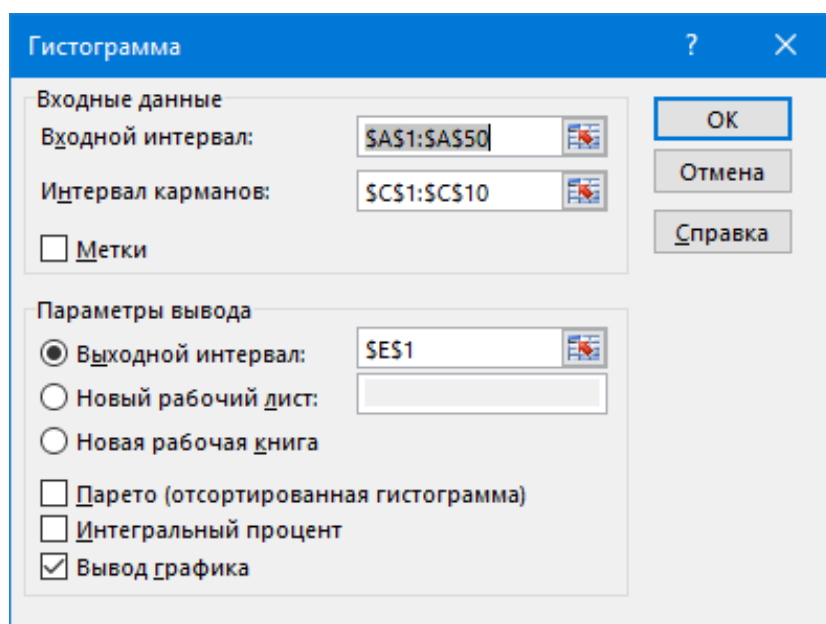


Рис. 6.2.2.

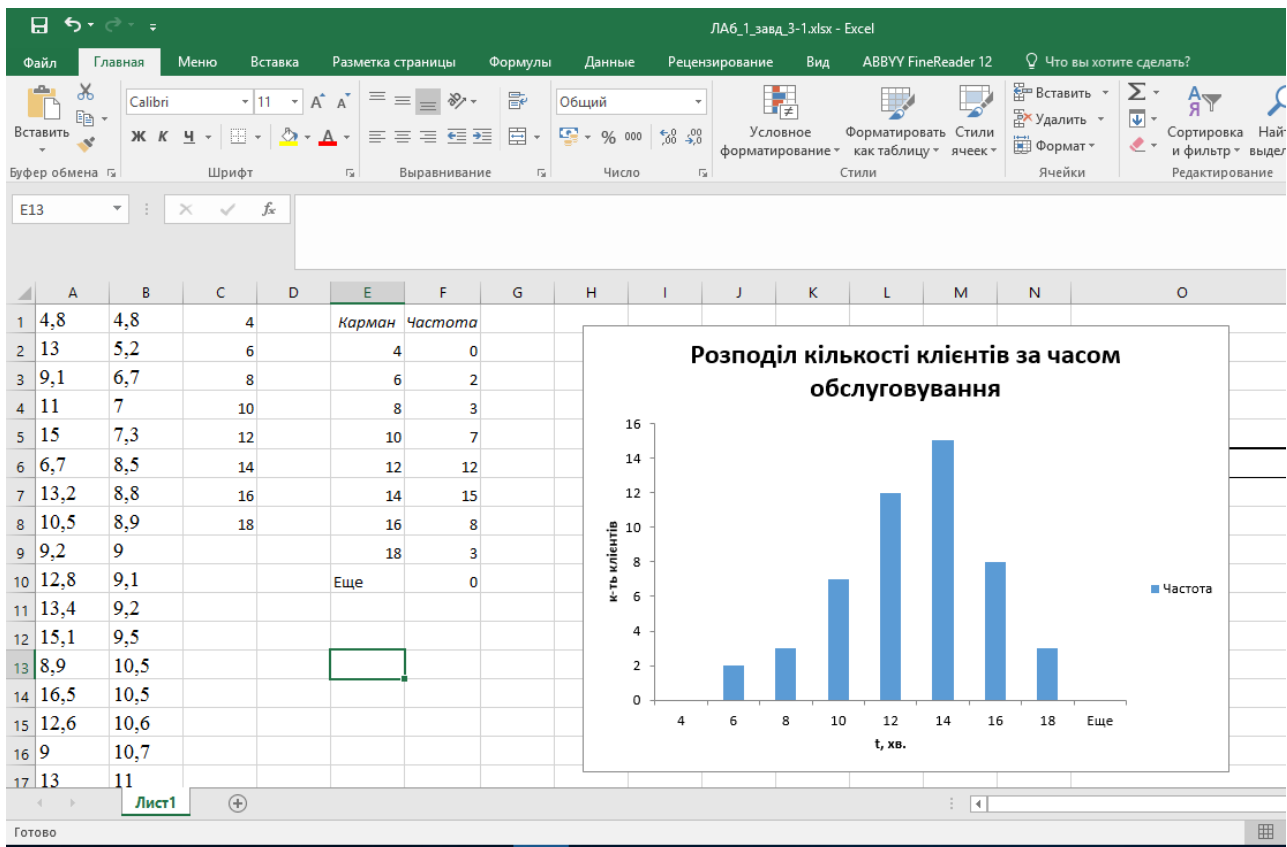


Рис. 6.2.3.

Оцінка виду розподілу графічним способом

1. У новому файлі (листу) Excel сформувати таблицю на основі одержаних вище даних. Колонка А містить нумерацію даних, у колонці В наведені інтервали часу обслуговування клієнтів, у колонці С середина відповідного інтервалу часу t , колонка D відображає кількість обслуговувань клієнтів, які попадають у відповідний інтервал часу. Колонки E та F відображають частоту попадання (ймовірність попадання) у певний інтервал та накопичену ймовірність (інтегральну функцію розподілу), тобто ймовірність того, що час обслуговування менший за t . Фрагмент цього вікна Excel показаний на рис. 6.2.4. Значення комірок двох останніх колонок розраховані: комірки колонки E одержані діленням значення комірки в колонці D на загальну суму значень комірок цієї ж колонки (загальне число клієнтів - 50.); колонка F – сума значень p_i .

2. Для побудови полігонів відносних і накопичених відносних частот скопіювати на вільне місце (наприклад у колонки H, J) колонку C та помножені на 100% значення комірок E і виділити їх. Використовуючи меню **Вставка**, застосувати до виділених чисел засіб діаграми **Точечная**. Отриманий графік є полігон відносних частот. Оформити одержаний графік підписами осей та назвою графіка (рис. 6.2.5).

	A	B	C	D	E	F	G
1							
2							
3							
4					1 день		
5	№	час обл	сер.час	К-ть кл	ймов	сум. йм	
6	1	4,0 -6,0	5	2	0,04	0,04	
7	2	6,0 -8,0	7	3	0,06	0,1	
8	3	8,0 - 10,0	9	7	0,14	0,24	
9	4	10,0 -12,0	11	12	0,24	0,48	
10	5	12,0 -14,0	13	15	0,3	0,78	
11	6	14,0-16,0	15	8	0,16	0,94	
12	7	16,0 -18,0	17	3	0,06	1	
13							
14							
15							

Рис. 6.2.4.

Якщо ці ж дії проробити з стовпцями C і F, то отримаємо полігон накопичених частот - згладжений графік емпіричної інтегральної функції розподілу.

Визначити середнє значення часу обслуговування клієнтів як суму значень у комірках **F6: F12** (11,84). Як видно з рис. 6.2.5. мода розподілу дорівнює 12 (максимум функції розподілу), а медіана 11,5 (значення x , при якому інтегральна функція розподілу приймає значення $F=0,5$).

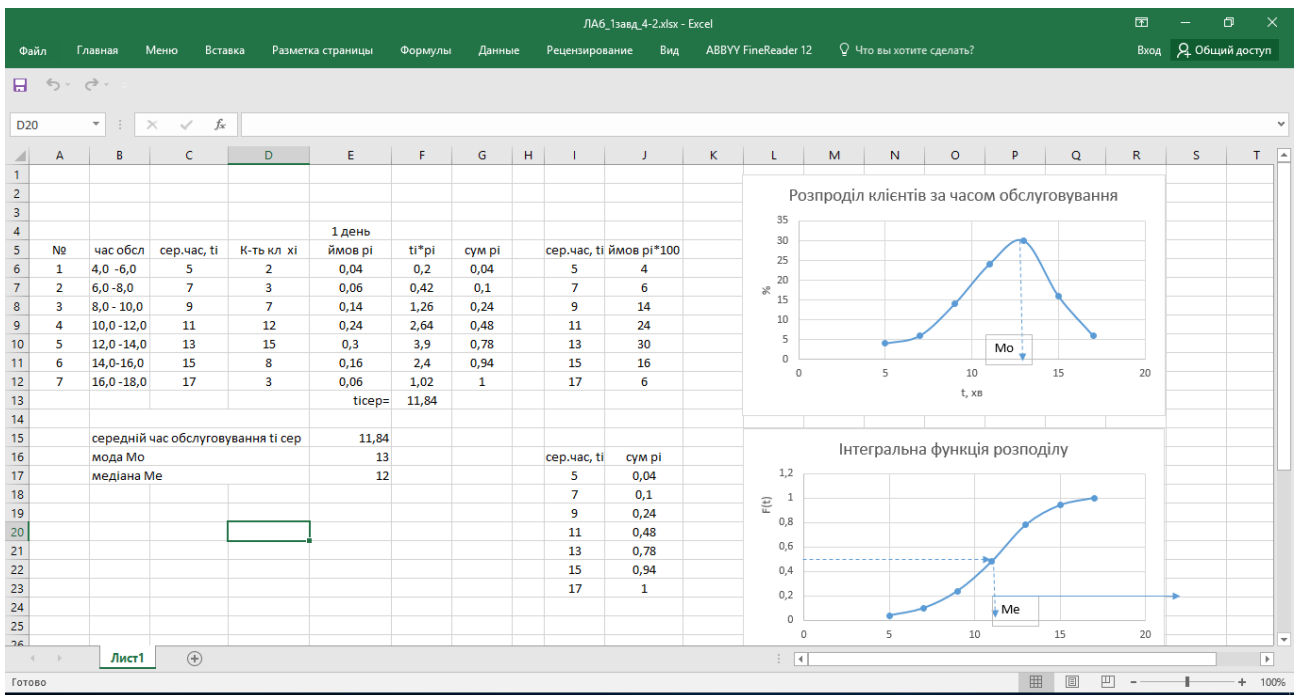


Рис. 6.2.5.

6.2.2. Перевірка гіпотези про вид розподілу з використанням критерію відповідності Пірсона

Мета завдання – перевірка гіпотези про нормальний вид розподілу часу обслуговування клієнтів банку із застосуванням критерію відповідності Пірсона (χ^2 -критерію).

Хід виконання завдання

У першому завданні було визначено оцінки статистичного розподілу часу обслуговування клієнтів банку і сформульована вихідна гіпотеза про те, що розглядуваний розподіл підпорядковується нормальному закону із середнім значенням $\bar{x}=11,84$ і середньоквадратичним відхиленням $S=2,905$.

Відмінність між експериментальним і теоретичним законами розподілу можна охарактеризувати величиною

$$\chi^2 = \sum C_i (p_i - p_i^{теор})^2, \quad (1.1)$$

де згідно припущення Пірсона $C_i = \frac{n}{p_i^{теор}}$ - вагові коефіцієнти, які з більшою вагою враховують відхилення для менших p_i .

1. Критерій відповідності Пірсона передбачає розбиття досліджуваної випадкової величини на певні інтервали значень. У

цьому завданні можна скористатися вже визначеними інтервалами у першому завданні. Межі цих інтервалів $x_{i\min}$ – $x_{i\max}$ занести у колонки С і D розрахункової таблиці (рис. 6.2.6).

	A	B	C	D	E	F	G	H	I	J
1			інтервал часу		теор інтегр ф-ція F		ймов попад	розрахунки		
2	час обсл xi	к-ть кл mi	xi(min)	xi(max)	F(x≤xi(min))	F(x≤xi(max))	в інт	n*i-pi	mi-n*i-pi	(mi-n*i-pi)^2/n*i-pi
3	5	2	4	6	0,003489271	0,022234904	0,018745633	0,956027	1,043973	1,140008249
4	7	3	6	8	0,022234904	0,093183348	0,070948443	3,618371	-0,61837	0,105678013
5	9	7	8	10	0,093183348	0,263310421	0,170127073	8,676481	-1,67648	0,323931748
6	11	12	10	12	0,263310421	0,521954071	0,25864365	13,19083	-1,19083	0,107504028
7	13	15	12	14	0,521954071	0,771346861	0,24939279	12,71903	2,280968	0,409057357
8	15	8	14	16	0,771346861	0,923859522	0,152512662	7,778146	0,221854	0,006327898
9	17	3	16	18	0,923859522	0,982986527	0,059127005	3,015477	-0,01548	7,94387E-05
10		$\Sigma m_i = 50$					$\Sigma p_i = 0.979$			$\Sigma \chi^2 = 2,09$
11		$n^* = 50 / \Sigma p_i = 51$								$\chi^2(\text{теор}) = 9,49$
12										
13		середнє значення			11,84					
14		середньоквадратичне відхилення			2,906					
15		розрахований критерій Пірсона			2,09					
16		рівень значимості α ($p=0,95$)			0,05					
17		ступені вільності r			4					
18		критичне знач. критерію Пірсона			9,487729037					
19										

Рис. 6.2.6.

2. Наступний етап розрахунків полягає у визначенні ймовірності попадання випадкової величини x з нормальним законом розподілу та передбачуваними середнім значенням 11,84 і середньоквадратичним відхиленням $S=2,905$ у задані інтервали. В Excel існує функція **НОРМ.РАСП (x; xcp; σ ; інтегральная)**. Функція **НОРМ.РАСП (x; x(cp); σ ; 1)** повертає ймовірність того, що нормально розподілена випадкова величина при середньому $x(cp)$ і середньоквадратичному відхиленні σ виявиться не більше значення x . Наприклад, для нашого випадку, ймовірність того, що час обслуговування клієнта банку не більше 8 хвилин дорівнює $F(8)=P(x<8)=\text{НОРМРАСП}(8; 11,84; 2,906; 1)=0.093$. Аналогічно для 6 хв. маємо $F(6)=P(x<6)=\text{НОРМРАСП}(6; 11,84; 2,906; 1)=0,022$. А ймовірність того, час обслуговування клієнта банку виявиться в межах 6–8 хв. дорівнює $F(8)–F(6)=0,071$.

Користуючись описаною вище функцією **НОРМРАСП (x; 11,84; 2,906; 1)** заповнити колонки Е і F розрахункової таблиці: $F(x < x_{i\min}) = \text{НОРМРАСП}(x_{i\min}; 11,84; 2,906; 1)$ - ймовірність того, що

випадкова величина виявиться менше нижньої межі заданого інтервалу, $F(x < x_{imax}) = \text{НОРМРАСП}(x_{i_max}; 11,84; 2,906; 1)$ - ймовірність того, що випадкова величина виявиться менше верхньої межі заданого інтервалу, $p_i = F(x < x_{imax}) - F(x < x_{imin})$ - ймовірність попадання випадкової величини в заданий інтервал. Для обчислення колонки **G** від значень комірок F_i необхідно відняти значення D_i .

3. За сумою комірок колонки **B** підрахувати загальну суму аналізованих спостережень $n = \sum m_i = 50$ (комірка **B10**), а по колонці **G** - загальну суму ймовірностей $\sum p_i = 0,979$ (комірка **G10**). При неухильному використанні критерію відповідності Пірсона загальна сума спостережуваних ймовірностей повинна дорівнювати 1. У цьому випадку цього не відбувається. Це обумовлено застосуванням методом побудови гістограми. У зв'язку з цим вводиться ще одна величина - припустима загальна кількість випробувань n^* . Цю величину пропонується визначити за формулою $n^* = \frac{n}{\sum p_i}$. Визначити і ввести у комірку **B11** розрахункової таблиці припустиме значення $n^* = 50/0,979 \approx 51$.

4. Визначити спостережуване значення χ^2 -статистики. Для цього у комірки колонки **H** розрахункової таблиці ввести значення n^*p_i (добуток $n^* = 51$ на значення комірок колонки **G**), у колонку **I** - $m_i - n^*p_i$ (різниця значень колонки **B** і **H**), у колонку **J** - $(m_i - n^*p_i)^2 / n^*p_i$ (квадрат значення комірок колонки **I** поділене на відповідні значення комірок колонки **H**). Сума значень комірок колонки **J** становить спостережуване значення χ^2 -статистики, яке записати у комірку **J10**.

Спостережуване значення χ^2 - статистики в розглядуваному випадку дорівнює 2,09.

5. Кількість степенів вільності дорівнює кількості спостережуваних інтервалів 7 зменшеному на кількість параметрів (2, так як раніше визначено для вибірки середнє значення і середньоквадратичне відхилення) і на 1: $r = 7 - 2 - 1 = 4$.

Для рівня значущості $\alpha = 0,05$ і числа степенів вільності $r = 4$ визначити критичне значення χ^2 - статистики. Замість таблиці можна використовувати функції Excel **ХІ2ОБР** (α ; r). Функція Excel **ХІ2ОБР** (**0,05**; **4**) дає критичне значення $\chi^2_{\text{крит}} = 9,48$.

6. Умова $\chi^2 = 2,09 \leq 9,48 = \chi^2_{\text{крит}}$ виконується, отже гіпотезу про відповідність розподілу для досліджуваної вибірки нормальному закону з ймовірністю 0,95 відкидати немає підстави.

7. Оформити результати перевірки відповідності розподілу часу обслуговування клієнтів нормальному за критерієм Пірсона (рис. 6.2.6).

6.2.3. Визначення основних характеристик вибірки за допомогою вбудованих функцій Excel

Мета завдання: за допомогою статистичних процедур і статистичних функцій бібліотеки вбудованих функцій MS Excel визначити характеристики вибірки.

Хід виконання завдання

Алгоритм виконання завдання.

Як і в попередньому завданні у діапазоні комірок **A1: AN** необхідно відобразити вибіркові значення часу обслуговування клієнтів банку (50 значень з табл. 6.2) (рис. 6.2.7). Для прискорення роботи можна скористатись готовими стовпчиками попереднього завдання копіюванням відповідних комірок на новий пустий листок Excel.

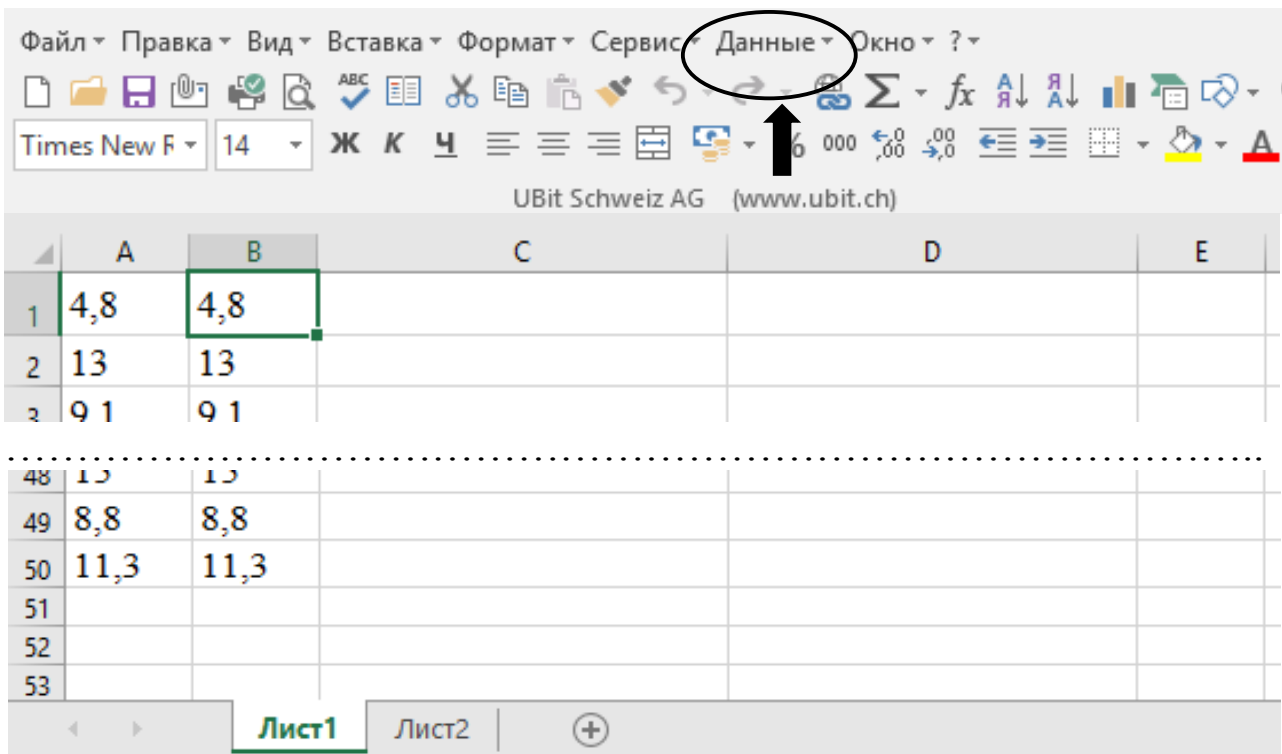


Рис. 6.2.7.

2. У меню **Данные** виділити підменю **Анализ данных**, вибрати процедуру **Описательная статистика**, в поле введення **Входной интервал** ввести посилання на діапазон комірок, що містить статистичні дані **A1: AN**. Встановити прапорець **Итоговая статистика**. Активізувати поле **Выходной интервал**, ввести в це поле посилання на комірку **C1**, яка буде лівою верхньою коміркою таблиці результатів (рис. 6.2.8).

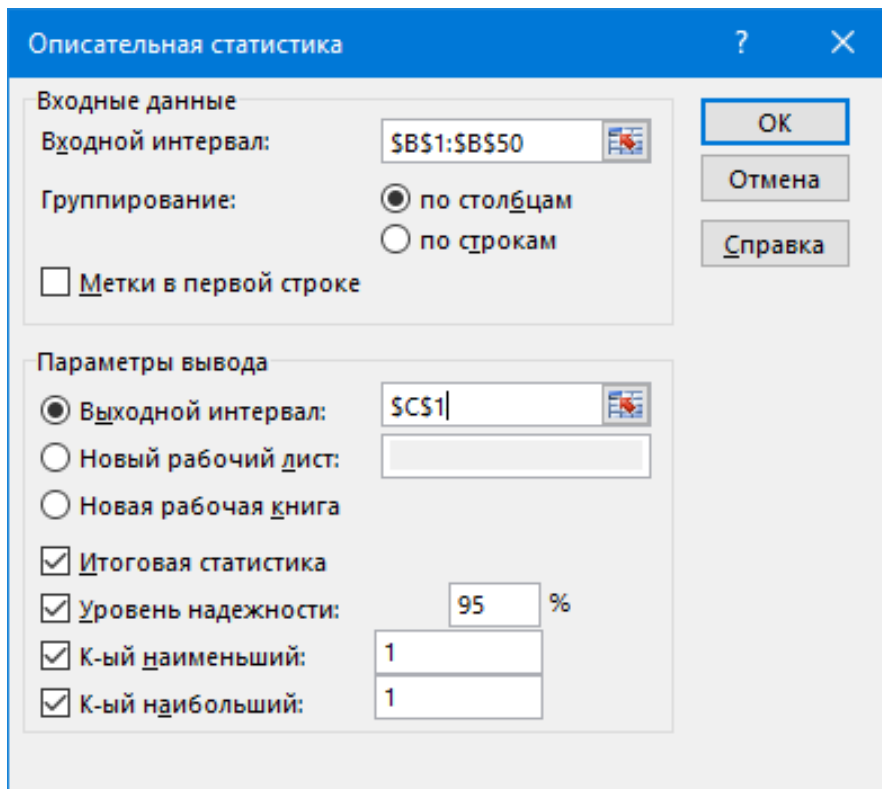


Рис. 6.2.8.

3. Після натискання ОК отримаємо результат, приведений на рис. 6.2.9.

	A	B	C	D	E	F
1	4,8	4,8				
2	5,2	5,2				
3	6,7	6,7	Среднее		11,84	
4	7	7	Стандартная ошибка		0,410952105	
5	7,3	7,3	Медиана		12,5	
6	8,5	8,5	Мода		13	
7	8,8	8,8	Стандартное отклонение		2,905870202	
8	8,9	8,9	Дисперсия выборки		8,444081633	
9	9	9	Эксцесс		-0,070922397	
10	9,1	9,1	Асимметричность		-0,393185629	
11	9,2	9,2	Интервал		12,7	
12	9,5	9,5	Минимум		4,8	
13	10,5	10,5	Максимум		17,5	
14	10,5	10,5	Сумма		592	
15	10,6	10,6	Счет		50	
16	10,7	10,7	Наибольший(1)		17,5	
17	11	11	Наименьший(1)		4,8	
18	11	11	Уровень надежности(95,0%)		0,825839174	
19	11	11				

Рис. 6.2.9.

Близькі значення середнього, моди і медіани, а також незначне значення ексцесу та симетричності свідчать про ймовірно нормальний розподіл величини часу обслуговування t одного клієнта. Ці значення добре узгоджуються із аналогічними, одержаними у попередньому пункті. Рядок *стандартная погрешность*, у цій задачі не має змісту і може бути видалений із підсумкової таблиці (пояснити чому).

6.3. Апроксимація експериментальних даних лінійною парною регресією

Ця тема включає виконання завдань, присвячених побудові і дослідженню рівняння лінійної регресії виду

$$y = f(x) = b_0 + b_1 x \quad (6.3.1)$$

Побудова лінійної регресії зводиться до оцінки її параметрів b_0 , b_1 . Класичний підхід до оцінювання параметрів лінійної регресії заснований на методі найменших квадратів (МНК).

Вихідні дані: Для визначення кореляції знань учнів з фізики (X) і математики (Y) були проведені контрольні роботи з цих предметів, результати яких представлені в табл. 6.3.

Таблиця 6.3.

Учні (i)	1	2	3	4	5	6	7	8	9	10
Оцінки КР з фізики (X)	8	11	12	9	8	8	9	9	8	12
Оцінки КР з математики (Y)	5	10	10	7	5	6	6	5	6	8

6.3.1. Обчислення коефіцієнтів рівняння лінійної регресії

Мета завдання. Обчислення коефіцієнтів рівняння лінійної регресії по вибірці, приведеній в табл. 6.3.

Хід виконання завдання

Коефіцієнти, що визначаються на основі МНК є розв'язком системи рівнянь (див. п. 4.4)

$$nb_0 + b_1 \sum_{i=1}^n x_i = \sum_{i=1}^n y_i \quad (6.3.2)$$

$$b_0 \sum_{i=1}^n x_i + b_1 \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i y_i$$

Розв'язуючи цю систему рівнянь, отримаємо

$$b_0 = \bar{y} - b_1 \bar{x}; \quad (6.3.3)$$

$$b_1 = \frac{\bar{x}\bar{y} - \bar{x} \cdot \bar{y}}{\bar{x}^2 - \bar{x}^2} = \frac{cov(x, y)}{S_x^2}; \quad (6.3.4)$$

де

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i, \quad \bar{x}\bar{y} = \frac{1}{n} \sum_{i=1}^n x_i y_i, \quad \bar{x}^2 = \frac{1}{n} \sum_{i=1}^n x_i^2$$

Для виконання завдання:

8. Завантажити програму Excel, відкрити пусту книгу.

9. У комірки A7-A16 увести нумерацію (від 1 до 10), у комірки B7-B16 увести числові значення x_i , а у комірки C7-C16 числові значення y_i .

10. Комірки D7-D16, E7-E16 заповнити відповідно значеннями $x_i y_i$ та x_i^2 з використанням функцій Excel.

11. Визначити середні значення x , y , xy та x^2 з використанням вбудованої в Excel функції **СРЗНАЧ**. Для цього відмітити курсором комірку B18 (C18, D18, E18). У полі вибрати **Поиск функций** наведенням курсора і вибором значка f_x . В одержаному вікні вибрати категорію **Статистические** і функцію **СРЗНАЧ** (рис. 6.3.1).

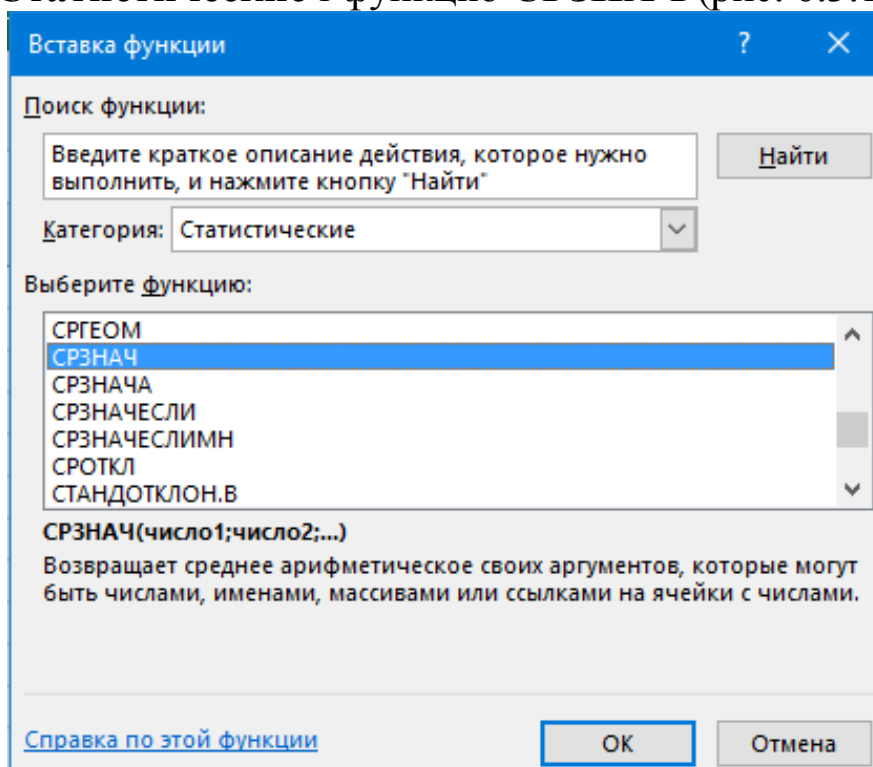


Рис. 6.3.1.

Після натискання ОК одержати вікно (рис. 6.3.2)

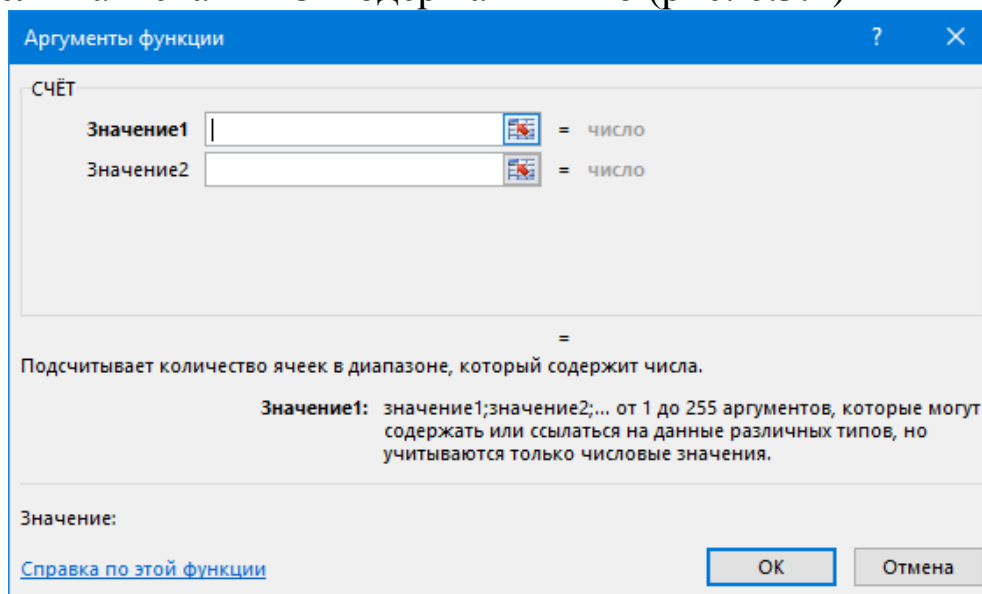


Рис. 6.3.2.

12. Навести курсор на комірку В7 і відмітити комірки В7-В16. У полі **Значение1** при цьому з'явиться діапазон комірок (вказуються адреси першої і останньої комірки даних, записаних через двокрапку, наприклад **В7: В16**), а також значення шуканої величини. Натисканням ОК одержане значення шуканої величини з'явиться у комірці В18.

13. Аналогічним чином розраховуються середні значення y , xy та x^2 .

14. За формулами (6.3.3) та (6.3.4) визначити коефіцієнти a і b , відобразити їх у комірках В20-Е20 та В21-Е21, записати рівняння регресії у комірки В22-Е22 (рис. 6.3.3).

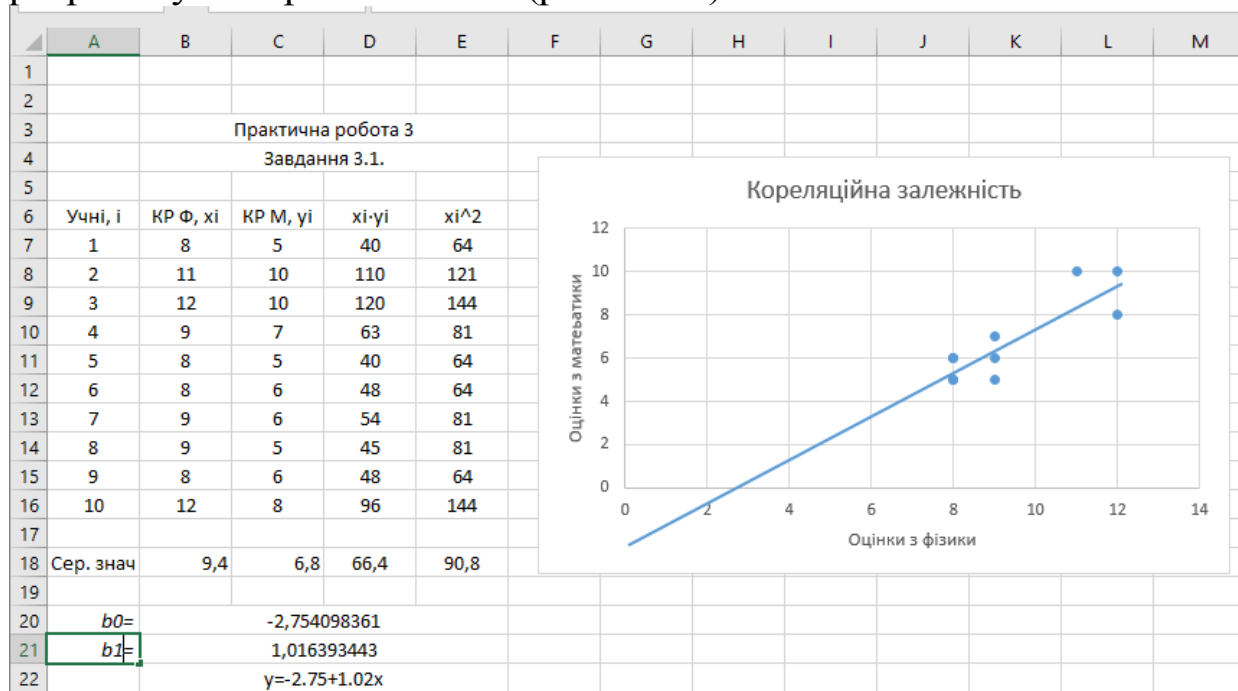


Рис. 6.3.3.

6.3.2. Обчислення вибіркового коефіцієнта кореляції

Мета завдання. Визначення тісноти зв'язку між оцінками учнів з фізики та математики.

Хід виконання завдання

Вибірковий коефіцієнт кореляції визначається співвідношенням

$$\rho_{xy} = \frac{\sqrt{S^2 - S_{зал}^2}}{S} = \sqrt{1 - \frac{S_{зал}^2}{S^2}} = \frac{S_y}{S} = \frac{b_1 S_x}{S}, \quad (6.3.5)$$

де
$$S_x^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2, \quad S^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2 \quad (6.3.6)$$

Для виконання завдання:

1. Використати результати отримані у попередньому завданні.

Доповнити таблицю стовпчиками $x_i - \bar{x}$, $(x_i - \bar{x})^2$, $y_i - \bar{y}$, $(y_i - \bar{y})^2$.

2. За формулами (6.3.6) розрахувати S_x , S та коефіцієнт кореляції ρ_{xy} .

Учні, i	КР Ф, xi	КР М, yi	xi-yi	xi^2	xi-хсер	(xi-хсер)^2	yiроз	yi-усер	(yi-усер)^2	yiроз-yi	(yiроз-yi)^2
1	8	5	40	64	-1,4	1,96	5,41	-1,8	3,24	0,41	0,1681
2	11	10	110	121	1,6	2,56	8,47	3,2	10,24	-1,53	2,3409
3	12	10	120	144	2,6	6,76	9,49	3,2	10,24	-0,51	0,2601
4	9	7	63	81	-0,4	0,16	6,43	0,2	0,04	-0,57	0,3249
5	8	5	40	64	-1,4	1,96	5,41	-1,8	3,24	0,41	0,1681
6	8	6	48	64	-1,4	1,96	5,41	-0,8	0,64	-0,59	0,3481
7	9	6	54	81	-0,4	0,16	6,43	-0,8	0,64	0,43	0,1849
8	9	5	45	81	-0,4	0,16	6,43	-1,8	3,24	1,43	2,0449
9	8	6	48	64	-1,4	1,96	5,41	-0,8	0,64	-0,59	0,3481
10	12	8	96	144	2,6	6,76	9,49	1,2	1,44	1,49	2,2201
Сума						24,4			33,6		8,4082
Сер. знач	9,4	6,8	66,4	90,8							
b0=		-2,754098361					Sx= 1,646545205		Sзал= 0,966563213		
b1=		1,016393443					S= 1,932183566				
		y=-2.75+1.02x					рху= 0,866138072				

Рис. 6.3.4.

Отримане значення коефіцієнту кореляції $\rho_{xy} = 0,87$ свідчить про сильний позитивний лінійний зв'язок (рис. 6.3.4).

6.3.3. Використання статистичних функцій табличного процесора MS Excel для визначення параметрів лінійної парної регресії


Мета завдання. Обчислення коефіцієнтів рівняння лінійної регресії.

Хід виконання завдання

В обчислювальному середовищі табличного процесора MS Excel це завдання вирішується за допомогою статистичних функцій **НАКЛОН** (нахил прямої щодо осі X, коефіцієнт b_1) і **ОТРЕЗОК** (відрізок, що відсікається прямою на осі Y, коефіцієнт b_0).

Для знайомства з цими можливостями вводяться необхідні вихідні дані у стовпчики B і C дані з табл. 6.3.

Для виконання завдання:

1. Вибрати вільну комірку. У полі  вибрати **Поиск функций** наведенням курсора і вибором значка f_x . В одержаному вікні вибрати категорію **Статистические** і функцію **НАКЛОН (ОТРЕЗОК)** - обидві ці функції мають два аргументи: діапазон комірок зі значеннями Y і діапазон комірок зі значеннями X (рис. 6.3.3). Статистична функція **КВПИРСОН** обчислює значення коефіцієнта детермінації, який рівний квадрату коефіцієнта кореляції.

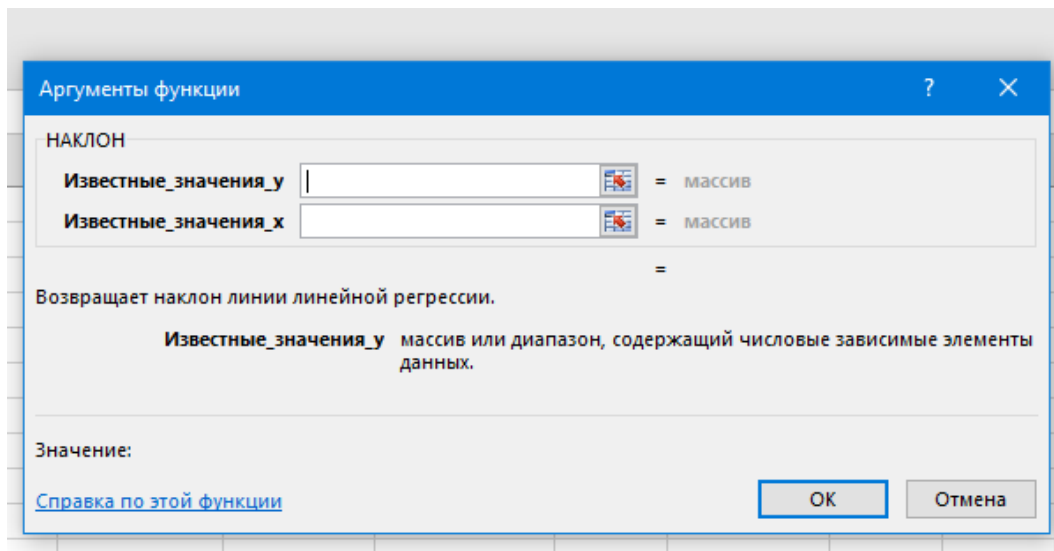


Рис. 6.3.3а.

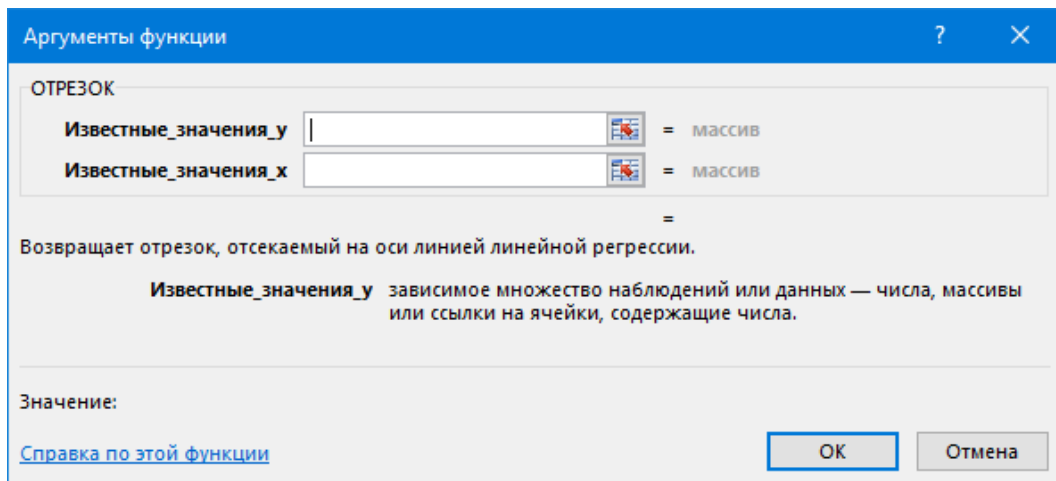


Рис. 6.3.3б.

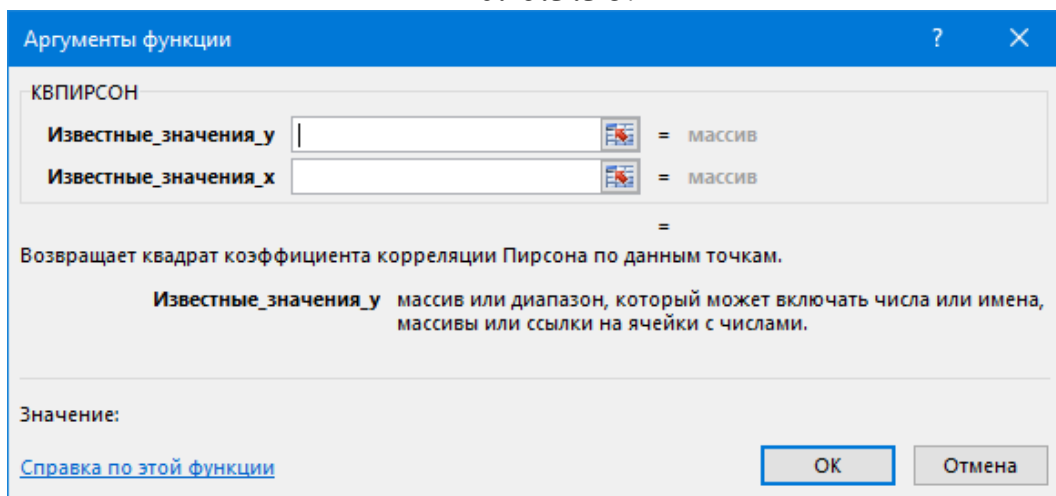


Рис. 6.3.3в.

Результат виконаних дій має вид

	A	B	C	D	E	F	G
46							
47							
48	Практична робота 3						
49	Завдання 3.3.						
50							
51	Учні, і	КР Ф, хі	КР М, уі				
52	1	8	5				
53	2	11	10		відрізок	b0=	-2,754098361
54	3	12	10		Нахил	b1=	1,016393443
55	4	9	7		КвПірсон	(рху)^2=	0,75019516
56	5	8	5				
57	6	8	6				
58	7	9	6				
59	8	9	5				
60	9	8	6				
61	10	12	8				
62							

2. Другим способом визначення параметрів лінійної парної регресії є використання функції **ЛИНЕЙН**. Функція **ЛИНЕЙН** (изв_знач_y; изв_знач_x; константа; стат) обчислює коефіцієнти

лінійної регресії, коефіцієнт детермінації $R = \rho^2$, F-статистику. В поле «изв_знач_u» вводиться діапазон значень Y (CN: CN+10); «изв_знач_x» - діапазон значень X (BN: BN+10); константа встановлюється на 0, якщо заздалегідь відомо, що вільний член b_0 дорівнює 0, і на 1 в іншому випадку; стат встановлюється на 0, якщо не потрібне виведення додаткових відомостей регресійного аналізу, та на 1 в іншому випадку.

Порядок використання функції **ЛИНЕЙН**:

1. Виділити область порожніх клітинок 5x2 (5 рядків, 2 стовпці) для виведення результатів регресійної статистики і 1x2 для виведення тільки коефіцієнтів a , b .

2. Ввести функцію ЛИНЕЙН вручну або через Майстер функцій.

3. Після коректного введення функції в лівій верхній клітинці виділеної таблиці з'явиться перший елемент таблиці. Щоб розкрити всю таблицю, слід спочатку натиснути клавішу F2, а потім одночасно натиснути клавіші [Ctrl], [Shift], [Enter].

47								
48		Практична робота 3						
49		Завдання 3.3.						
50								
51	Учні, і	КР Ф, хі	КР М, уі					
52	1	8	5					
53	2	11	10					
54	3	12	10					
55	4	9	7					
56	5	8	5					
57	6	8	6	b1=	1,016393443	-2,754098361	"=b0	
58	7	9	6	< Δb_1^2 >=	0,20736247	1,975936941	"=< Δb_1^2 >	
59	8	9	5	(ρ_{xy})^2=	0,75019516	1,024295039	"=< Δy^2 >	
60	9	8	6	F-статистика	24,025	8	"=ст. вільн	
61	10	12	8	Регр сума кв=	25,20655738	8,393442623	"=Зал сума кв	
62								

РЕКОМЕНДОВАНА ЛІТЕРАТУРА

1. Володарський Є.Т., Кошева Л.О. Статистична обробка даних: навч. посібник. – К.: НАУ, 2008. – 308 с.
2. Засименко В.М. Основи теорії планування експерименту: навч. посібник. – Львів: Видав. ДУ «ЛП», – 2000. – 205 с.
3. Гаркавий В.Г., Ярова В.В. Математична статистика. – К: Професіонал, 2004. – 484 с.
4. Джонсон Н. Лион Ф. Статистика и планирование эксперимента в технике и науке. В 2 т. Т. 1: Методы обработки данных. Пер. с англ. под. ред. Э.К. Лецкого. –М.: Мир, 1980. – 510 с.
5. Джонсон Н. Лион Ф. Статистика и планирование эксперимента в технике и науке. В 2 т. Т. 2: Методы планирования эксперимента. Пер. с англ. под. ред. Э.К. Лецкого. –М.: Мир, 1981. – 557 с.
6. Монтгомери, Д.К. Планирование эксперимента и анализ данных. Пер. с англ. В. А. Коптяева. – Л.: Судостроение, 1980. – 383 с.
7. Рубан, А. И. Методы анализа данных: уч. пособие. 2-е изд., исправл. доп. – Красноярск: ИПЦ КГТУ, 2004. – 319 с.
8. Рубан, А. И. Методы оптимизации: уч. пособие. 3-е изд., испр. и доп. – Красноярск: ИПЦ КГТУ, 2004. 528 с
9. Зажигаев Л.С., Кишьян А.А., Романиков Ю.И. Методы планирования и обработка результатов физического эксперимента. – М.: Атомиздат, 1978. – 232 с.
10. Лапач С.Н., Чубенко А.В., Бабич П.Н. Статистические методы в медико-биологических исследованиях с использованием Excel – К.: Морион, 2000. – 320 с.
11. Мінцер О.П., Вороненко Ю.В., Власов В.В. Оброблення клінічних і експериментальних даних у медицині: навч. посібник. – К.: Вища школа, 2003. – 350 с.

Навчальне видання

ГОРВАТ Андрій Андрійович
МОЛНАР Олександр Олександрович
МІНЬКОВИЧ Віктор Вікторович

**МЕТОДИ ОБРОБКИ
ЕКСПЕРИМЕНТАЛЬНИХ ДАНИХ
З ВИКОРИСТАННЯМ MS EXCEL**

Навчальний посібник

Підписано до друку 25.12.2019. Формат 60x84/16
Папір офсетний. Гарнітура Times New Roman. Друк офсетний.
Умовн. друк. арк. 9,30 Зам.№ 15. Наклад 100 прим.

Видавництво УжНУ «Говерла»
88000, м. Ужгород, вул. Капітульна, 18. E-mail: hoverla@i.ua

*Свідоцтво про внесення до державного реєстру
видавців, виготівників, і розповсюджувачів видавничої продукції
– Серія Зт № 32 від 31 травня 2006 року*

А.А. Горват , О.О. Молнар, В.В. Мінькович
Г67 **Методи обробки експериментальних даних**
з використанням MS Excel: Навчальний посібник.
Ужгород: Видавництво УжНУ “Говерла”, 2019. - 160 с.: іл.

Розглянуті питання статистичної обробки результатів експерименту – визначення характеристик емпіричних розподілів, відсів грубих похибок спостережень, перевірка закону розподілу дослідних даних, основи регресійного і кореляційного аналізів та методи планування експерименту.

Навчальний посібник призначено для студентів, аспірантів, наукових працівників, які навчаються або ведуть дослідження за інженерними та природничими напрямками.

ISBN 978-617-7825-00-4

УДК 519.22:502/504
ББК 28.081