

**НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ЦИВІЛЬНОГО ЗАХИСТУ УКРАЇНИ**  
**Факультет пожежної безпеки**  
**Кафедра пожежної профілактики в населених пунктах**

**КОНСПЕКТ ЛЕКЦІЙ**  
**З НАВЧАЛЬНОЇ ДИСЦИПЛІНИ**  
**«МЕТОДИ ОБРОБКИ КІЛЬКІСНИХ ПОКАЗНИКІВ ПОЖЕЖНОЇ**  
**НЕБЕЗПЕКИ»**

**для здобувачів третього (доктор філософії) рівня вищої освіти**  
**галузь знань 26 «Цивільна безпека», спеціальність 261 «Пожежна безпека»,**  
**3 курс, 5 семестр**

Харків 2024

Рекомендовано до друку  
кафедрою пожежної  
профілактики в населених  
пунктах НУЦЗ України  
(протокол № 19 від 18.06.2024 р.)

**Укладачі:** Ю.А. Отрош, Н.В. Рашкевич

**Рецензенти:** доктор технічних наук, професор **М. Г. Сур'янінов**, завідувач кафедри будівельної механіки Одеської державної академії будівництва та архітектури;

**Конспект** лекцій з навчальної дисципліни «Методи обробки кількісних показників пожежної небезпеки» для здобувачів третього (доктор філософії) рівня вищої освіти галузь знань 26 «Цивільна безпека», спеціальність 261 «Пожежна безпека», 3 курс, 5 семестр / Укладачі: Ю.А. Отрош, Н.В. Рашкевич. Х.: НУЦЗУ, 2024. 66 с.

## ЗМІСТ

Вступ	4
Лекція 1 Основи експериментальних досліджень	6
Лекція 2 Фізичні величини та засоби вимірювань	13
Лекція 3 Гіпотеза. Моделювання як один із методів перевірки гіпотези	21
Лекція 4 Поняття фактору. Факторний простір	27
Лекція 5 Попередня обробка даних	32
Лекція 6 Загальні відомості про методи обробки експериментальних даних	38
Лекція 7 Факторний аналіз	43
Лекція 8 Дисперсійний аналіз	50
Лекція 9 Кореляційний аналіз. Регресійний аналіз.	55

## ВСТУП

**Мета навчальної дисципліни** є формування у здобувачів вищої освіти системи базових знань і навичок для організації, проведення експериментальних досліджень та обробки їх результатів при дослідженні кількісних показників пожежної небезпеки, а також систематизація, розширення та закріплення професійних знань і формування навичок самостійного ведення наукової роботи, дослідження та експериментування.

**Завдання навчальної дисципліни** є ознайомлення з експериментальними методами проведення наукових досліджень та обробки їх результатів, розвиток практичних навичок з планування, організації та проведення експериментальних наукових досліджень, освоєння різних методів аналізу та обробки результатів експериментальних наукових досліджень, ознайомлення з методами побудови однофакторних, багатфакторних моделей та перевірки їх на адекватність, з методами пошуку та дослідження зв'язків між експериментальними даними, а також отримання практичних навичок з використання комп'ютерної техніки для обробки результатів експериментальних наукових досліджень.

У результаті вивчення навчальної дисципліни здобувач вищої освіти повинен отримати:

**знання:**

- методи планування та проведення експериментальних наукових досліджень для визначення кількісних показників пожежної небезпеки;
- методи обробки результатів експериментальних наукових досліджень;
- методи побудови однофакторних, багатфакторних моделей та перевірки їх на адекватність.

**уміння/навички:**

- застосовувати отримані навички для самостійного планування, підготовки і проведення експериментальних наукових досліджень;
- самостійно проводити обробку експериментальних даних при визначенні кількісних показників пожежної небезпеки;
- самостійно будувати однофакторні, багатфакторні моделі та перевіряти їх на адекватність за критеріями Стьюдента, Фішера, Пірсона;
- використовувати комп'ютерну техніку для обробки результатів експериментальних наукових досліджень;
- кількісних показників пожежної небезпеки об'єктів різного призначення;
- розробки і обґрунтування заходів з посилення пожежної безпеки об'єктів об'єктах різного призначення.

**комунікації:**

- донесення до фахівців і нефахівців інформації, ідей, проблем, рішень, власного досвіду та аргументації;

– вільне спілкування з питань, що стосуються сфери наукових та експертних знань, з колегами, широкою науковою спільнотою, суспільством у цілому.

**Відповідальність та автономію:**

– демонстрація значної авторитетності, інноваційність, високий ступінь самостійності, академічна та професійна доброчесність, послідовна відданість розвитку нових ідей або процесів у передових контекстах професійної та наукової діяльності;

- відповідальність за внесок до професійних знань і практики;
- здатність продовжувати навчання з високим ступенем автономії;
- здатність до безперервного саморозвитку та самовдосконалення.

## **ЛЕКЦІЯ № 1**

### **ОСНОВИ ЕКСПЕРИМЕНТАЛЬНИХ ДОСЛІДЖЕНЬ**

#### **Література:**

1. Карагодова О.О., Рожок В.Д. Дослідження операцій: навч. посіб. К.: Центр учбової літератури, 2007. 256 с.
2. Бахрушин В.Є. Методи аналізу даних: навчальний посібник для студентів. Запоріжжя: КПУ, 2011. 268 с
3. Перегуда О.В., Капустян О.А., Курилко О.Б. Статистична обробка даних: навч. посіб. Електронне видання, 2022. 103 с.
4. Рогальський Ф.Б., Курилович Я.Є., Цокурєнка А.А. Математичні методи аналізу економічних систем. Книга 1. Київ: Наукова думка, 2001. 435 с.
5. Рогальський Ф.Б., Цокурєнка А.А. Математичні методи аналізу економічних систем. Книга 2. Київ: Наукова думка, 2001. 423 с.
6. Крайчук О.В., Московська Г.К., Соколенко О.І. Теорія ймовірностей і математична статистика. Рівне, 2004.

#### **Мета лекції:**

1. Навчальна – вивчити основи організації та проведення експериментальних досліджень.
2. Виховна – виховувати професійні риси, почуття відповідальності; виховувати зацікавленість дисципліною, прагнення отримати нові знання самостійно.
3. Розвиваюча – придбання нових знань щодо основ експериментальних досліджень з метою вирішення задач у сфері пожежної безпеки.

#### **План:**

1. Експеримент. Класифікація експериментальних досліджень.
2. Підготовка до проведення експериментального дослідження.

#### **Теоретична частина:**

##### **1. Експеримент. Класифікація експериментальних досліджень.**

Як правило, будь-яке наукове дослідження включає до свого складу експеримент. Він являє собою один із способів отримати нові наукові знання.

Головною метою експерименту може бути виявлення властивостей досліджуваних об'єктів, перевірка справедливості гіпотез і на цій основі всебічне і глибоке вивчення теми наукового дослідження.

Основою експерименту є науково поставлений дослід, у якому вивчення явищ відбувається за допомогою доцільно вибраних або штучно створених умов, що забезпечують появу тих процесів, спостереження яких необхідне для встановлення закономірних зв'язків між явищами.

Саме по собі слово "експеримент" означає дію, спрямовану на створення умов, щоб вивчити певне явище.

У науковій мові і в дослідницькій діяльності терміном "експеримент" називають цілий ряд споріднених понять: дослід, цілеспрямоване спостереження, відтворення об'єкта пізнання, організацію особливих умов його здійснення, перевірка прогнозу. У це поняття вкладають наукову постановку

дослідів і спостереження досліджуваного явища у спеціально визначених умовах, які дають можливість спостерігати за ним і відтворювати його кожного разу при повторенні цих умов.

Постановка та організація експерименту визначається його призначенням. За своїм спрямування експерименти є дуже різноманітними і кожний з них має свої специфічні особливості. У першу чергу ці особливості визначаються тим, у якій галузі науки проводиться експеримент. Відповідно до цього розрізняють експерименти суто фізичні, біологічні, хімічні, психологічні, соціальні, педагогічні тощо.

Більш конкретно експерименти класифікують за конкретнішими ознаками:

- за способом формування умов - природні і штучні. Природний експеримент передбачає проведення дослідів у звичних для досліджуваного об'єкта умовах існування (найчастіше знаходить застосування у біологічних, соціальних, психологічних і педагогічних дослідженнях). Штучний експеримент передбачає створення штучних умов для його проведення (застосовується у природничих і технічних науках);

- за метою дослідження - перетворюючі, констатуючі, контролюючі, пошукові.

Перетворюючий (або ще творчий) експеримент включає активну зміну структури і функцій об'єкта дослідження відповідно до висунутої гіпотези, формування нових зв'язків і відношень між компонентами об'єкта або між досліджуваним об'єктом та іншими об'єктами. Дослідник, виходячи з виявлених тенденцій розвитку об'єкта дослідження навмисно створює умови, котрі повинні сприяти формуванню нових властивостей і якостей об'єкта.

Констатуючий експеримент застосовується для перевірки певних припущень. Під час такого експерименту констатується наявність певних зв'язків між впливом на об'єкт дослідження і досягнутим результатом, встановлюється наявність певних фактів.

Контролюючий експеримент зводиться до контролю за результатами зовнішнього впливу на об'єкт дослідження з урахуванням його стану, характеру впливу і ефекту, що очікується.

Пошуковий експеримент проводиться у тому випадку, коли важко розділити фактори, що впливають на досліджуване явище внаслідок відсутності достатніх попередніх (апріорних) даних. За результатами пошукового експерименту встановлюється значимість факторів, здійснюється відкидання тих, які мають незначний вплив;

- за організацією проведення - лабораторні, натурні, польові, виробничі тощо.

Лабораторний експеримент проводять у лабораторних умовах із застосуванням типових приладів, спеціальних моделюючих пристроїв, стендів тощо. Найчастіше у лабораторному експерименті вивчається не сам об'єкт, а його зразок. Такий експеримент дає можливість доброякісно і з потрібною повторюваністю вивчити вплив одних характеристик одночасно варіюючи інші, отримати переконливу наукову інформацію з мінімальними витратами часу і ресурсів. Але такий експеримент не завжди повністю моделює реальний хід досліджуваного процесу, і тому виникає потреба проводити натурний

експеримент.

Натурний експеримент проводять у звичайних (природних) умовах і на реальних об'єктах. Цей вид експерименту часто застосовують у процесі натурних випробувань технічних об'єктів. Залежно від місця проведення випробувань натурні експерименти можуть бути виробничими, польовими, полігонними, напівнатурними тощо. Головна наукова проблема натурального експерименту - забезпечити достатню відповідність (адекватність) умов його проведення реальній ситуації, у якій буде працювати надалі досліджуваний об'єкт;

- за структурою досліджуваних об'єктів і явищ - прості і складні.

Простий експеримент застосовується для вивчення нескладних за структурою об'єктів, що мають невелику кількість взаємопов'язаних і взаємодіючих елементів, кожний з яких виконує найпростішу функцію.

Складний експеримент вивчає явища чи об'єкти із складною структурою і великою кількістю взаємопов'язаних і взаємодіючих елементів, кожний з яких виконує складні функції. У даному разі зміна стану якого-небудь елемента або зв'язку призводить до зміни стану багатьох інших елементів системи;

- за характером зовнішнього впливу на об'єкт дослідження - речовинні, енергетичні, інформаційні.

Речовинний експеримент передбачає вивчення впливу різних речовинних факторів на стан об'єкта дослідження (наприклад, вплив різних домішок на якість певної речовини - продукту харчування, хімічного препарату, матеріалу тощо).

Енергетичний експеримент застосовують для вивчення впливу різних видів енергії (електричної, механічної, теплової і т.ін.) на об'єкт дослідження. Цей вид експерименту поширений у природничих науках.

Інформаційний експеримент має на меті вивчення впливу певної (різної за формою і змістом) інформації на об'єкт дослідження (найчастіше інформаційний експеримент застосовують у біології, психології, соціології, кібернетиці тощо). За допомогою такого експерименту вивчається зміна стану об'єкта під впливом інформації, яка надходить до нього;

- за характером взаємодії засобу експериментального дослідження з об'єктом дослідження - звичайні і змодельовані.

Звичайний (або ще класичний) експеримент включає експериментатора як суб'єкта пізнання, а також об'єкт чи предмет експериментального дослідження і ті засоби (інструменти, прилади і т. ін.), за допомогою яких здійснюється експеримент. У звичайному експерименті експериментальні засоби безпосередньо взаємодіють з досліджуваним об'єктом. Вони неначе стають посередниками між експериментатором і об'єктом дослідження.

Змодельований експеримент на відміну від звичайного має справу з моделлю досліджуваного об'єкта. Модель входить до складу експериментальної установки, замінюючи не тільки об'єкт дослідження, а іноді й умови, в яких вивчається деякий об'єкт. Змодельований експеримент, розширюючи можливості експериментального дослідження, одночасно має і деякі недоліки, пов'язані з тим, що відмінність моделі від реального об'єкта може стати джерелом помилок і, крім того, екстраполяція результатів вивчення поведінки моделі на змодельований об'єкт вимагає додаткових витрат часу і теоретичного обґрунтування правомірності такої екстраполяції;



- за типом моделей, що досліджуються у експерименті - розумовий і матеріальні.

Відмінність між знаряддями експерименту при моделюванні дає можливість виділити розумовий і матеріальний експеримент. Знаряддями розумового (його ще називають розумовим) експерименту можуть бути уявні моделі досліджуваних об'єктів чи явищ (почуттєві образи, знакові чи образно-знакові моделі). Розумовий експеримент вважається однією з форм розумової діяльності суб'єкта пізнання, у процесі якої відтворюється в уяві структура реального експерименту. В матеріальному експерименті знаходять застосування матеріальні, а не уявні об'єкти дослідження. Головна відмінність матеріального експерименту від розумового полягає у тому, що реальний експеримент являє собою форму об'єктивного матеріального зв'язку свідомості з зовнішнім світом, тоді як розумовий експеримент виступає як специфічна форма теоретичної діяльності суб'єкта. Схожість розумового експерименту з реальним у значній мірі визначається тим, що будь-який реальний експеримент, перше ніж він буде здійснений на практиці, спочатку відбувається в уяві експериментатора. Тому розумовий експеримент досить часто виступає як ідеальний план реального експерименту, відбуваючись раніше за нього. Розумовий експеримент, замінюючи собою реальний, розширює межі пізнання. Це можливе завдяки тому, що він у деяких випадках може стати єдиним джерелом одержати інформацію про досліджуваний об'єкт. Розумовий експеримент дає можливість подолати неминучу обмеженість реального дослідження шляхом абстрагування від дії небажаних факторів, уникнути впливу яких в реальному експерименті неможливо;

- за величинами, що контролюються - пасивні і активні.

Пасивний експеримент передбачає вивчення тільки обраних для цього показників (параметрів, змінних) на основі спостереження за об'єктом без штучного втручання в його функціонування. Прикладами пасивного експерименту можуть бути спостереження за якимись процесами, подіями і т.ін. Тому пасивний експеримент можна вважати по суті справою спостереженням (як методом дослідження), яке супроводжується інструментальними вимірюваннями показників стану об'єкта дослідження.

Активний експеримент пов'язаний з вибором спеціальних вхідних факторів і контролем за тим, що відбувається з об'єктом дослідження під впливом цих факторів;

- за числом факторів, що варіюються - однофакторні і багатофакторні.

Однофакторний експеримент передбачає виділення потрібних для дослідження факторів і вивчення кожного з них окремо.

Багатофакторний експеримент спрямовується на вивчення всіх виділених для дослідження факторів одночасно у їх взаємодії;

- за характером об'єктів чи явищ, що досліджуються - технологічні, соціометричні тощо.

Технологічний експеримент вивчає елементи технологічних процесів (продукцію, обладнання, діяльність працівників і т. ін.) або процеси в цілому.

Соціометричний експеримент застосовують для вивчення існуючих міжособистісних соціально-психологічних відношень у малих групах з метою наступного впливу на них.

Стверджувати, що наведена класифікація експериментальних досліджень є остаточною і повною немає підстав, адже розвиток наукового пізнання постійно приводить до розширення меж застосування експериментального методу. Крім того, залежно від завдань експерименту різні його види можуть об'єднуватись, утворюючи комплексний або комбінований експеримент (що якраз найчастіше відбувається в реальних дослідженнях).

## **2. Підготовка до проведення експериментального дослідження.**

Підготовка до проведення експериментального дослідження передбачає здійснення таких дій:

- розробку гіпотези, що підлягатиме експериментальній перевірці;
- складання програми експериментальних робіт;
- визначення способів і прийомів впливу на об'єкт дослідження;
- забезпечення умов для здійснення процедури експериментальних робіт;
- розробку прийомів фіксування проміжних і кінцевих результатів експерименту;
- підготовку експериментальних засобів (приладів, установок, моделей тощо);
- визначення і підготовку учасників експериментального дослідження.

До початку експерименту складається план (програма) його проведення. У ньому формулюють мету і завдання експерименту; визначають об'єкт, що підлягає експериментальній перевірці; обґрунтовують обсяг експерименту, кількість дослідів, послідовність їх реалізації; визначають фактори впливу на об'єкт дослідження та у якій послідовності вони будуть змінюватись у процесі експерименту; обґрунтовують засоби вимірювань; визначають способи обробки та аналізу експериментальних матеріалів.

Проводиться експеримент на основі попередньо розробленої методики експерименту.

Методика експерименту - це сукупність розумовий і реальних операцій, розташованих у певній послідовності, відповідно до якої досягається мета дослідження.

Розробка методики експерименту передбачає дотримання деяких умов:

- необхідність проведення попереднього цілеспрямованого спостереження за об'єктом дослідження чи явищем з метою визначення вихідних даних (гіпотези, факторів, що будуть вивчатись);
- створення умов, у яких можливо здійснити експеримент (добір об'єктів для експериментального впливу, усунення впливу випадкових факторів);
- визначення меж вимірювань, які будуть здійснюватись у процесі спостереження за факторами, що впливають на об'єкт дослідження;
- забезпечення можливості систематичного спостереження за процесом розвитку досліджуваного явища і точного опису одержуваних фактів;
- можливість проведення систематичної реєстрації вимірювань і оцінювання фактів різними засобами і способами;
- можливість створення повторних ситуацій або ускладнених ситуацій з метою підтвердження чи заперечення раніше одержаних даних;
- можливість здійснити перехід від емпіричного вивчення до логічних

узагальнень, до аналізу і теоретичної обробки одержаних фактичних матеріалів.

В методиці докладно розкривається процес експерименту, визначається послідовність проведення потрібних вимірювань і спостережень, описується кожна експериментальна дія з урахуванням обраних для проведення експерименту засобів, обґрунтовуються методи контролю якості експериментальних дій, які повинні забезпечувати при мінімальній (попередньо встановленій) кількості вимірювань високу надійність і задану точність. Розробляються форми протоколів (або таблиць) для фіксації результатів спостережень і вимірювань.

Важливим розділом методики є вибір методів обробки і аналізу експериментальних даних.

Обробка даних полягає у систематизації всіх кількісних показників (найчастіше це цифрові матеріали), їх класифікації та наступному аналізі. Залежно від теми наукового дослідження обсяг експерименту може бути різним. У кращому випадку щоб підтвердити гіпотезу дослідження достатньо провести тільки лабораторний експеримент. Але може бути потреба і в проведенні цілої серії експериментальних досліджень: попередніх (пошукових), лабораторних, на дослідних зразках в реальних умовах їх функціонування (на виробництві) і т.ін. Надійним помічником науковця у проведенні експериментальних досліджень є комп'ютерна техніка. Сучасний персональний комп'ютер може надати допомогу людині вже на етапі розробки програми і методики експерименту і допомогти обрати найбільш оптимальний варіант проведення дослідження. За допомогою ЕОМ можна створювати найбільш оптимальні моделі відповідно до умов експерименту, моделювати об'єкти або процеси, котрі недоцільно, а то і неможливо відтворити на практиці (наприклад, моделювання аварійних режимів і ситуацій). ЕОМ надає можливість оперативно систематизувати й аналізувати експериментальні дані та відображати їх у формі, зручній для зорового сприйняття.

#### **Питання для контролю знань:**

1. Надайте визначення: «експеримент», «дослід», «спостереження», «методика експерименту».
2. Мета експерименту.
3. Ознаки класифікації експерименту.
4. Класифікація експерименту за способом формування умов.
5. Класифікація експерименту за метою дослідження.
6. Класифікація експерименту за організацією проведення.
7. Класифікація експерименту за структурою досліджуваних об'єктів і явищ.
8. Класифікація експерименту за характером зовнішнього впливу на об'єкт дослідження.
9. Класифікація експерименту за характером взаємодії засобу експериментального дослідження з об'єктом дослідження.
10. Класифікація експерименту за типом моделей, що досліджуються у експерименті.
11. Класифікація експерименту за величинами, що контролюються.

12. Класифікація експерименту за числом факторів, що варіюються.
13. Класифікація експерименту за характером об'єктів чи явищ, що досліджуються.
14. Етапи підготовки до проведення експериментального дослідження.
15. Умови розробки методики експерименту.
16. У чому полягає обробка і аналіз експериментальних даних?

**Завдання для самостійної підготовки:**

- I. Вивчити матеріал лекції.
- II. Підготувати доповідь за темою:
  1. Сутність, мета, функції наукового експерименту.
  2. Класифікація експериментів.
  3. Методологія експериментальних досліджень.
  4. Загальні вимоги до проведення експерименту.
  5. Типові помилки в проведенні експерименту.
  6. Робоче місце експериментатора та організація експерименту.
- III. Пройти тестування за темою лекції: тест 1.1. за посиланням <https://forms.gle/zcPDNKs8vwmueb4c6>

## ЛЕКЦІЯ № 2 ФІЗИЧНІ ВЕЛИЧИНИ ТА ЗАСОБИ ВИМІРЮВАНЬ

### Література:

1. Карагодова О.О., Рожок В.Д. Дослідження операцій: навч. посіб. К.: Центр учбової літератури, 2007. 256 с.
2. Бахрушин В.Є. Методи аналізу даних: навчальний посібник для студентів. Запоріжжя: КПУ, 2011. 268 с
3. Перегуда О.В., Капустян О.А., Курилко О.Б. Статистична обробка даних: навч. посіб. Електронне видання, 2022. 103 с.
4. Рогальський Ф.Б., Курилович Я.Є., Цокурєнка А.А. Математичні методи аналізу економічних систем. Книга 1. Київ: Наукова думка, 2001. 435 с.
5. Рогальський Ф.Б., Цокурєнка А.А. Математичні методи аналізу економічних систем. Книга 2. Київ: Наукова думка, 2001. 423 с.
6. Крайчук О.В., Московська Г.К., Соколенко О.І. Теорія ймовірностей і математична статистика. Рівне, 2004.

### Мета лекції:

1. Навчальна – вивчити фізичні величини, що застосовуються у сфері пожежної безпеки, та засоби вимірювання для проведення експериментальних досліджень, обробки їх результатів.
2. Виховна – виховувати професійні риси, почуття відповідальності; виховувати зацікавленість дисципліною, прагнення отримати нові знання самостійно.
3. Розвиваюча – придбання нових знань щодо фізичних величин та засобів вимірювань для проведення експериментальних досліджень та обробки їх результатів.

### План:

1. Фізична величина. Основні відомості про одиниці фізичних величин.
2. Системи одиниць фізичних величин.
3. Міжнародна система одиниць.
4. Класифікація методів вимірювань фізичних величин.
5. Методи вимірювання фізичних величин.

### Теоретична частина:

#### **1. Фізична величина. Основні відомості про одиниці фізичних величин.**

Людина у своєму прагненні пізнати фізичні об'єкти – об'єкти пізнання – виділяє деяку відокремлену кількість властивостей, загальних у якісному відношенні для ряду об'єктів, але індивідуальних для кожного з них в кількісному відношенні. Такі властивості отримали назву *фізичних величин* (ФВ). Фізичні величини розрізняють в *якісному і кількісному* відношенні.

*Якісна сторона* – визначає «вид» величини, наприклад довжина, маса, теплоємність, вологість, тиск, температура тощо, а кількісна її розмір.

Кількісний склад властивості, що відповідає поняттю «фізична величина» в даному об'єкті – розмір фізичної величини.

Розмір фізичної величини існує об'єктивно, незалежно від того, що ми знаємо про неї.

Значення ФВ - це кількісна оцінка вимірюваної величини, повинна бути не тільки числом, а числом іменованим. Результат вимірювання повинен бути відображений у визначених одиницях, прийнятих для даної величини. Наприклад, маса 101 кг, довжина 91 м тощо.

Значення ФВ, яке ідеальним чином відображає в якісному і кількісному відношенні відповідну властивість об'єкту, повинно бути вільне від похибок вимірювань – називають **істинним значенням ФВ**. Так як усі ФВ знаходяться експериментальним або дослідним шляхом і їх значенню притаманні помилки вимірювань, то істинне значення ФВ залишається невідомим.

Значення ФВ, що знайдене експериментальним або дослідним шляхом і настільки наближене до істинного значення, що для визначеної мети може бути використано замість нього, носить назву **дійсного значення ФВ**.

При експериментальних або дослідних вимірюваннях значення ФВ, знайдене з допустимою по технічним вимогам похибкою, приймається за **дійсне значення**.

## 2. Системи одиниць фізичних величин.

Історично склалося так, що розвиток різноманітних галузей науки і техніки призвів до появи багатьох систем одиниць ФВ та великого числа позасистемних одиниць, але в міру розвитку метрології було встановлено, що доцільним є такий вибір одиниць, при якому одиниці деяких розмірів устанавлюють довільно, незалежно один від одного – такі одиниці називають основними, а одиниці інших розмірів виражають через основні – називають похідними.

Сукупність основних та похідних одиниць ФВ називають *системою одиниць фізичних величин* (метр, кілограм, секунда, ампер, кельвін, кандела, моль).

Одиниці величин, які не належать ні до основних, ні до похідних, називаються додатковими. Одиниці, що не входять у систему одиниць ФВ, називаються позасистемними (літр, калорія тощо).

Позасистемні одиниці, що визначаються із відношення двох значень величин, називають логарифмічними (бел, децибел).

Одиницю, що в ціле число разів більша системної або позасистемної одиниці, називають кратною (км, хвилина тощо).

Одиницю, що в ціле число разів менша від системної, або позасистемної одиниці, називають частковою (міліметр, мілісекунда тощо).

Величину, в розмірності якої хоча б один показник розмірності відмінний від 0, називають розмірною величиною, в розмірності якої всі показники дорівнюють нулю – безрозмірною.

Систему ФВ називають когерентною, якщо усі вхідні до неї похідні одиниці когерентні, тобто пов'язані з іншими одиницями систем рівнянням, в якому числовий коефіцієнт прийнятий рівним одиниці, як наприклад, одиниця швидкості (у СІ) метр у секунду, утворена за рівнянням зв'язку між одиницями:  $[V] = [L] \cdot [T^{-1}]$  м/с; де:  $V$  – швидкість,  $L$  – м;  $T$  – с.

### 3. Міжнародна система одиниць.

В 1960 році XI Генеральною конференцією з мір та ваг (ГКМВ), була прийнята єдина універсальна Міжнародна система одиниць (СІ).

Вона відзначається практичністю, когерентністю та раціональністю над іншими системами, що використовувались до неї.

Міжнародна система СІ побудована на семи основних та двох додаткових одиницях.

До основних одиниць СІ відносять:

*Метр* – дорівнює шляху, який проходить в вакуумі світло за  $299792458 \text{ с}^{-1}$  (17-та ГКМВ, 1983р).

Кілограм – є одиниця маси та дорівнює масі міжнародного прототипу кілограма (3-я ШКМВ, 1901р).

*Секунда* – дорівнює  $9192631770$  періодам випромінювання, яке відповідає переходу між двома рівнями основного стану атома Цезію-133 (13-та ГКМВ, 1967р).

*Ампер* – дорівнює силі незмінного струму, який під час проходження по двох безмежно довгих паралельних прямолінійних провідниках, розташованих у вакуумі на відстані  $1\text{ м}$  один від одного, викликав би на кожній ділянці провідника довжиною  $1\text{ м}$  силу взаємодії, яка дорівнює  $2 \cdot 10^{-7} \text{ Н}$  (МКВМ, 1946р)

*Кельвін* – одиниця термодинамічної температури. Дорівнює  $273,16^{-1}$  частині термодинамічної температури потрійної точки води (13-та ГКМВ, 1967р).

*Кандела* – сила світла у напрямі джерела, яке випромінює монохромне випромінювання частотою  $540 \cdot 10^{12} \text{ Гц}$ , енергетична сила якого при цьому становить  $683 \text{ Вт/см}$  (16-та ГКМВ, 1979 р).

Моль – дорівнює кількості речовини системи, яка містить стільки структурних елементів, скільки міститься атомів у нукліді  $^{12}\text{C}$  масою  $0,012 \text{ кг}$  (14-та ГКМВ, 1971р).

Додаткові одиниці СІ:

*Радіан* – дорівнює куту між двома радіусами кола, довжина дуги між якими дорівнює радіусу.

*Стерадіан* – дорівнює тілесному куту з вершиною у центрі сфери.

### 4. Класифікація методів вимірювань ФВ.

За характером знаходження значення вимірюваної величини, вимірювання класифікують на *прямі, непрямі, опосередковані, сукупні, сумісні, статичні та динамічні* вимірювання та ін.

*Пряме вимірювання* – вимірювання, в якому значення однієї величини –ни знаходять безпосередньо без перетворення її роду та використання відомих залежностей. До таких вимірювань можна віднести вимірювання маси, часу, довжини, сили електричного струму, атмосферного тиску тощо.

*Непряме вимірювання* – вимірювання, в якому значення однієї чи декількох величин, що вимірюються, знаходять після перетворення роду величини чи обчислення за відомими їхніми залежностями від декількох величин, що вимірюються прямо. До непрямих вимірювань відносять сукупне, сумісне та опосередковане.

*Опосередковане вимірювання* є різновидом непрямого вимірювання однієї величини з перетворенням її роду чи обчисленням за результатами вимірювання інших величин, з якими вимірювана величина пов'язана явною функціональною залежністю. Наприклад, при вимірюванні площі прямокутної ділянки потрібно визначити прямими вимірюваннями довжину та ширину, а значення площі обчислити за функціональною залежністю:  $S = a \cdot b$

*Сукупне вимірювання* – непряме вимірювання, в якому значення декількох одночасно вимірюваних однорідних величин отримують розв'язанням рівнянь, які пов'язують їх з іншими величинами, що вимірюються прямо чи опосередковано.

Наприклад, при визначенні струмів в розгалуженому електричному колі, за наявності значень опорів резисторів та е.р.с., складають систему рівнянь за першим і другим законами Кірхгофа.

*Сумісне вимірювання* – непряме вимірювання, в якому значення декількох одночасно вимірюваних різнорідних величин отримують розв'язанням рівнянь, які пов'язують їх з іншими величинами, що вимірюються прямо чи опосередковано.

*Статичне вимірювання* – вимірювання величини, яку можна вважати незмінною за час вимірювання.

*Динамічне вимірювання* – вимірювання величини, що змінюється за час проведення вимірювань.

Прикладом статичних вимірювань можуть бути вимірювання довжини, маси, значень постійного струму, а прикладом динамічних вимірювань можуть бути вимірювання значення температури в двигунах внутрішнього згорання, теплових потоків у швидкозмінних процесах рухомих газів, вимірювання температури в процесі нагрівання рідини тощо.

*Одноразове вимірювання* – вимірювання, яке здійснюється одноразово за один прийом (вимірювання часу по годиннику).

Багаторазові вимірювання – вимірювання однієї і тієї ж ФВ, результат яких отримують за декілька вимірювань, впорядкованих одне за одним.

Рівноточні вимірювання – ряд вимірювань будь-якої ФВ, що виконано однаковими за точністю засобами вимірювальної техніки (ЗВТ) в одних і тих самих умовах.

*Нерівноточне вимірювання* – ряд вимірювань ФВ, виконаних різними по точності ЗВТ в різних умовах.

*Метрологічне вимірювання* – вимірювання за допомогою еталонів і зразкових засобів ФВ або передачі їх розмірів до робочих ЗВТ.

## **5. Методи вимірювання ФВ**

*Методом вимірювання* називають сукупність способів використання ЗВТ та принципів вимірювань для створення вимірювальної інформації. Під принципом вимірювання розуміють сукупність фізичних явищ, на яких засновані ці вимірювання, а під ЗВТ – *технічний засіб*, який застосовується під час вимірювання і має нормовані метрологічні характеристики.

Існує велика кількість методів вимірювання. Одну і ту саму ФВ можна виміряти декількома різноманітними методами. Обрання того або іншого з них



визначається видом вимірюваної величини, необхідною точністю, швидкістю проведення вимірювань, економічними співвідношеннями.

В залежності від характеру використання ЗВТ основні методи вимірювань поділяють на:

*метод безпосереднього оцінювання вимірювання* – значення величини отримують безпосередньо по реєструючому пристрою вимірювального приладу прямої дії;

Методи порівняння засновані на порівнянні вимірювальної величини із тією, що відтворюється мірою.

Розрізняють наступні різновиди методів порівняння:

*метод порівняння з мірою* – це метод вимірювання, при якому вимірювана величина порівнюється з іншою, що відтворюється, мірою. За міру можна прийняти нормальні елементи, стандартні зразки тощо;

*метод протиставлень* – метод порівняння з мірою, при якому величина, що вимірюється, та величина, що відтворена мірою, одночасно впливають на прилад порівняння. Наприклад, вимірювання маси тіла на рівноплечих терезах;

*диференціальний метод* – метод порівняння з мірою, при якому на прилад, що вимірює, діє різниця між вимірюваною величиною та величиною, відтвореною мірою;

*метод заміщень* – метод порівняння з мірою. За цим методом величина, що вимірюється, заміщається відтвореною величиною мірою в такий спосіб, щоб показчик індикатора вимірювального приладу залишався незмінним. Прикладом може служити зважування з почерговим розміщенням вимірюваної маси та гир на одну і ту саму шальку ваг;

*метод збіжностей* – метод порівняння з мірою, при якому різниця між величиною, що вимірюється, та величиною, що відтворюється мірою, вимірюється за допомогою використання збіжності позначок шкал або періодичних сигналів. Цей метод використовується під час вимірювання геометричних розмірів тіл за допомогою штангенциркуля, інтерференцію світлових хвиль та ін.;

*нульовий метод* – метод порівняння з мірою, при якому результуючий ефект дії вимірюваної величини та відомої величини, яка відтворюється мірою, в приладі порівняння доводять до нуля. Прикладом використання нульового методу є різноманітні мостові схеми;

*контактний і неконтактний методи* – методи, за яких чутливий елемент приладу відповідно приводиться або не приводиться до контакту з об'єктом вимірювань. Наприклад, вимірювання температури тіла людини термометром здійснюється контактним методом, а температура на Сонці або в мартенівській печі здійснюється неконтактним методом.

### Основні фізичні величини

Основна величина	Одиниця фізичної величини	Умове позначення одиниці фізичної величини	Символ	Умове позначення
Довжина	метр	м	<i>L</i>	<i>l</i>
Час	секунда	с	<i>T</i>	<i>t</i>
Маса	кілограм	кг	<i>M</i>	<i>m</i>
Сила електричного струму	ампер	А	<i>I</i>	<i>I</i>
Температура	кельвін	К	⊕	<i>T</i>
Кількість речовини	моль	моль	<i>N</i>	<i>n</i>
Сила світла	кандела	кд	<i>J</i>	<i>j</i>

### Приставки, які використовуються для утворення кратних і частинних одиниць

Назва приставки	Позначення	Множник	Приклад
<b>Кратні одиниці</b>			
гіга	Г	$10^9$	Гм - гігаметр
мега	М	$10^6$	МА – мегаампер
кіло	к	$10^3$	кВ – кіловольт
гекто	г	$10^2$	гВт – гектоват
дека	да	$10^1$	даК – декакельвін
<b>Частинні одиниці</b>			
деци	д	$10^{-1}$	дм – дециметр
санти	с	$10^{-2}$	см – сантиметр
мілі	м	$10^{-3}$	мА – міліампер
мікро	мк	$10^{-6}$	мкВ – мікрвольт
нано	н	$10^{-9}$	нВт – нановат
піко	п	$10^{-12}$	пс – пікосекунда

## Одиниці та системи одиниць фізичних величин їх вимірювання

Назва	Основні одиниці	Примітка
СГС (абсолютна)	сантиметр, грам, секунда	Прийнята за пропозицією Стокса, 1862 р. Похідні: дина, ерг.
МКГСС (технічна)	метр, кілограм-сила, секунда	Увійшла в практику наприкінці XIX ст.
МКС (система одиниць механічних величин)	метр, кілограм, секунда	Запроваджена у 1901р. Дж. Джорджі.
МКСА (система одиниць електричних і магнітних величин)	метр, кілограм, секунда, ампер	"Система Джорджі" з 1958р.
МКСК (система одиниць теплових величин)	метр, кілограм, секунда, кельвін	ГОСТ 8550-57 ГОСТ 8550-61
МТС	метр, тонна, секунда	З 1919 р. у Франції

### 2. Вимірювання та їх похибки.

Теорія похибок – це вивчення і оцінка похибок у вимірюваннях.

Похибки можуть бути результатом недосконалості як методики вимірювання, так і міри. Таким чином, результат вимірювання не співпадає зі справжнім (істинним) значенням вимірювальної величини, а відображає лише її оцінку

Точність вимірювання відображає наближення результату до справжнього (істинного) значення вимірювальної величини.

Різниця між результатом вимірювання та істинним значенням називається похибкою.

Похибки можливо розділити на три класи:

- до першого класу належать систематичні похибки, які залишаються незмінними або такими, що закономірно змінюють своє значення. До них відносяться, наприклад, похибки виготовлення або градування самої міри. Зігнута стрілка вольтметра – приклад систематичної похибки.

- до другого класу похибок відносяться випадкові похибки. Причини їх виникнення невідомі. В результаті дії випадкових факторів ми не можемо, наприклад, впевнено знати результат вимірювання, знаючи попереднє значення. Дія цих випадкових факторів призводить до того, що відхилення результату від його істинного значення носить статистичний, ймовірний характер. До метрологічних характеристик точності вимірювань, зв'язаних з випадковими похибками, відносяться сходження і відтворення результатів. Сходження відображає (показує) близькість одне до одного результатів вимірювання, виконаних в ідентичних умовах. Відтворення вимірювань відображає (показує)

близькість одне до одного результатів вимірювання, виконаних в різних місцях, різними методами. Сукупність таких метрологічних характеристик, як: правильність, сходження і відтворення визначає точність вимірювання.

- третій клас похибок носить назву грубих похибок або промахів. Різке порушення умов вимірювання, неправильний запис в протоколі, поломка приладу під час проведення експерименту тощо призводять до появи цих похибок. Одна із задач статистики - визначення промахів на фоні інших випадкових відхилень.

#### **Питання для контролю знань:**

1. Що таке фізична величина?
2. На які значення поділяють фізичні величини?
3. Система одиниць вимірювання.
4. Міжнародна система одиниць.
5. Які існують вимірювання за характером знаходження значення величини, що вимірюють?
6. Класифікація методів вимірювання фізичних величин
7. Теорія похибок.
8. Три класи похибок.

#### **Завдання для самостійної підготовки:**

- I. Вивчити матеріал лекції.
- II. Підготувати доповідь за темою:
  1. Класифікація видів та методів вимірювання.
  2. Метрологічні характеристики засобів вимірювання.
  3. Основи теорії похибок вимірювань.
  4. Міжнародна система одиниць фізичних величин.
  5. Класифікація та характеристики похибок.
  6. Розрахунок похибок при прямих вимірах.
  7. Розрахунок похибок при непрямих вимірах.
  8. Систематичні похибки та їх корекція.
- III. Пройти тестування за темою лекції: тест 2.1. за посиланням <https://forms.gle/eGW6bnBrkChWnxxQ9>

## ЛЕКЦІЯ № 3

### ГІПОТЕЗА. МОДЕЛЮВАННЯ ЯК ОДИН ІЗ МЕТОДІВ ПЕРЕВІРКИ ГІПОТЕЗИ

#### Література:

1. Карагодова О.О., Рожок В.Д. Дослідження операцій: навч. посіб. К.: Центр учбової літератури, 2007. 256 с.
2. Бахрушин В.Є. Методи аналізу даних: навчальний посібник для студентів. Запоріжжя: КПУ, 2011. 268 с
3. Перегуда О.В., Капустян О.А., Курилко О.Б. Статистична обробка даних: навч. посіб. Електронне видання, 2022. 103 с.
4. Рогальський Ф. Б., Курилович Я. Є., Цокурєнко А. А. Математичні методи аналізу економічних систем. Книга 1. Київ: Наукова думка, 2001. 435 с.
5. Рогальський Ф. Б., Цокурєнка А. А. Математичні методи аналізу економічних систем. Книга 2. Київ: Наукова думка, 2001. 423 с.
6. Крайчук О.В., Московська Г.К., Соколенко О.І. Теорія ймовірностей і математична статистика. Рівне, 2004.

#### Мета лекції:

1. Навчальна – вивчити основи експериментальних досліджень.
2. Виховна – виховувати професійні риси, почуття відповідальності; виховувати зацікавленість дисципліною, прагнення отримати нові знання самостійно.
3. Розвиваюча – придбання нових знань щодо основ експериментальних досліджень з метою вирішення задач у сфері пожежної безпеки.

#### План:

1. Поняття гіпотези і її структура.
2. Види гіпотез.
3. Моделювання як один із методів перевірки гіпотези.

#### Теоретична частина:

##### 1. Поняття гіпотези і її структура.

**Гіпотезою** називається спосіб мислення, котрий полягає в побудові припущення про те, що таке досліджуване явище, та в доведенні цього припущення.

**Термін «гіпотеза»** уживається з подвійним значенням. Під гіпотезою розуміють і саме припущення, котре пояснює спостережуване явище, і спосіб мислення в цілому, який включає висування припущення, його розвиток і доведення. Гіпотеза є метод пізнання предметів і явищ навколишнього світу.

Гіпотеза створюється для того, щоб дати пояснення ще не поясненим явищам, фактам, подіям. Пізнання будь-якого явища в дійсності, як відомо, розпочинають із збирання і нагромадження окремих фактів, що відносяться до цього явища. Фактів, наявних на початку пізнання явища, завжди недостатньо, щоб повністю й відразу пояснити це явище, дати достовірний висновок про те, що воно таке, які причини його виникнення, закони розвитку тощо. Тому пізнання явищ і подій зовнішнього світу відбувається у формі гіпотези: не

очікуючи, доки накопичаться факти для кінцевого, достовірного висновку про характер і причину досліджуваного явища, роблять на початку здогадне пояснення спостережуваного явища, а потім цей здогад розвивають і доводять.

**Гіпотеза** – це форма розвитку наших знань. Мислення людини не знає інших способів логічного опрацювання емпіричного матеріалу і проникнення в сутність речей, окрім гіпотези. Побудова гіпотез у науці дає змогу переходити від окремих фактів, що стосуються явищ, до пізнання закону розвитку цього явища.

**Побудова гіпотез** – необхідний шлях до створення наукової теорії. Всяка наукова теорія висловлюється спочатку як гіпотеза. Науково доведена і підтверджена на практиці гіпотеза стає науковою теорією.

Логічна структура гіпотези складна. Гіпотеза не зводиться до якогось одного судження чи умовиводу. Вона – система суджень, понять і умовиводів. Якесь одне окремо взяте судження або умовивід ще не складає гіпотези.

Гіпотеза може складатися одночасно із різних видів умовиводів індукції, аналогії і дедукції. Наприклад, судження-припущення може бути висловлене за аналогією чи індукцією, а потім розвинуте й доведене у формі дедукції. Але припущення в гіпотезі може бути висунуте не тільки у формі індукції чи аналогії, воно висловлюється часто дедуктивно, а доводиться потім у формі індукції або дедукції тощо.

**Гіпотеза** – процес розвитку думки. Процес мислення в гіпотезі має певні стадії. Розрізняють дві такі стадії побудови і доведення гіпотези: 1) висунання гіпотези і 2) доведення гіпотези. Дехто виділяє в гіпотезі не дві, а три, чотири чи п'ять стадій: 1) вивчення обставин досліджуваного явища (збирання фактів), 2) формування гіпотези, 3) виведення з гіпотези наслідків (розвиток гіпотези), 4) перевірка цих наслідків на практиці і 5) висновок про істинність або хибність висунутої гіпотези.

Гіпотеза має відповідати **вимогам**:

- відповідність установленим у науці законам;
- узгодженість із фактичним матеріалом;
- несуперечність з погляду формальної логіки (якщо йдеться про суперечності самої об'єктивної реальності, то гіпотеза повинна містити суперечності);
- відсутність суб'єктивних, довільних припущень (що не відмінняє активності самого суб'єкта);
- можливість її підтвердження чи спростування або в ході безпосереднього спостереження, або опосередковано – шляхом виведення наслідків з гіпотези.

**Висунання гіпотези.** Гіпотеза будується не на голому місці. Щоб її висунути, необхідно мати певну сукупність фактів, що відносяться до спостережуваного явища, котрі б обґрунтовували ймовірність якогось припущення, пояснювали ймовірність невідомого. Тому побудова гіпотези завжди пов'язана зі збиранням фактів, які мають відношення до того явища, котре ми пояснюємо. На підставі зібраних фактів висловлюється припущення про те, що таке досліджуване явище, тобто формулюється гіпотеза. **Припущення** в гіпотезі в логічному відношенні є судження (або система суджень). Його висловлюють унаслідок логічного опрацювання зібраних фактів.

Факти, на підставі яких висувається гіпотеза, можуть бути логічно осмислені у формі аналогії, індукції чи дедукції. В одних випадках гіпотезу висувають за аналогією, в інших – вони є висновком індуктивного чи дедуктивного умовиводу. Наприклад, гіпотеза про існування життя на Марсі висунута за аналогією. На підставі схожості Марса і Землі в одних ознаках, а саме в тому, що Марс і Земля є планетами сонячної системи, що обидві вони обертаються навколо Сонця, мають атмосферу, воду, зміну пір року, дня й ночі і т. д., зроблено припущення про схожість цих планет і в інших ознаках, а саме, що на Марсі, як і на Землі, існує життя.

Висування припущення, тобто формулювання гіпотези, становить основний зміст гіпотези.

**Припущення** – головний елемент будь-якої гіпотези. Припущення є відповіддю на поставлене питання про сутність, причину, зв'язки спостережуваного явища. Припущення містить те знання, до якого доходять унаслідок узагальнення фактів. **Припущення – це серцевина гіпотези, навколо якої відбувається вся пізнавальна і практична діяльність.** Припущення в гіпотезі – це, з одного боку, підсумок попереднього пізнання, те головне, до чого доходять унаслідок спостереження й узагальнення фактів; з другого боку – це відправний пункт подальшого вивчення явища, визначення напрямку, яким має відбуватися все дослідження. Гіпотеза дає змогу не тільки пояснити наявні факти, а й виявити нові, на котрі не була б звернута увага, коли не була б висунута ця гіпотеза.

**Доведення гіпотези.** Висунута гіпотеза має бути доведеною. Доказ гіпотези здійснюється так. Припускаючи висунуту гіпотезу істинною, із неї дедуктивним методом виводять ряд наслідків (фактів), котрі мають існувати, якщо існує гадана причина, а потім ці наслідки перевіряють на практиці. Якщо наслідки відповідають дійсності, підтверджуються практикою, то це свідчить про те, що ця гіпотеза є правильною. Якщо ж наслідки, логічно виведені з гіпотези, не відповідають дійсності, то це означає, що висунута гіпотеза хибна.

Перевірка гіпотези відбувається завжди шляхом підтвердження наслідків, виведених із цієї гіпотези. Критерієм істинності гіпотези є практика. Гіпотеза стає достовірною теорією, коли наслідки, виведені з неї, підтверджуються практикою. Тому доведення істинності гіпотези не можна зводити до одноактної логічної дії.

Перевіряючи гіпотезу, використовують різні логічні форми. Досить часто перевірка гіпотези відбувається за схемою умовно-категоричного силогізму. Особливо широко умовно-категоричний силогізм використовується тоді, коли мають справу з умовно-виділяючими судженнями. У таких випадках наявність у дійсності всього лише одного наслідку, виведеного з гіпотези, достатня для визнання цієї гіпотези істинною. Окрім умовно-категоричних умовиводів використовуються також категоричний силогізм, розподільні умовиводи та інші логічні форми.

## **2. Види гіпотез.**

Гіпотеза може пояснити або явище (подію) в цілому, або якийсь окремий бік явища, одну його властивість, один зв'язок. Тому розрізняють гіпотези загальні й часткові.

**Загальна гіпотеза** – це припущення, котре пояснює причину явища або групи явищ у цілому.

**Часткова гіпотеза** – припущення, яке пояснює якийсь окремий бік чи окрему властивість явища чи події.

Так, гіпотеза про походження гір – це загальна гіпотеза, а гіпотеза про походження якоїсь однієї гори – часткова гіпотеза. У судовому дослідженні припущення про злочин у цілому є загальною гіпотезою, а припущення, що пояснює окремий бік злочину, наприклад, припущення про мотив злочину, про шлях проникнення злочинця до приміщення, про спосіб скоєння злочину тощо є частковою гіпотезою.

Поділ гіпотез на загальні й часткові має сенс, коли ми співвідносимо одну гіпотезу з другою. Цей поділ не є абсолютним, гіпотеза може бути частковою стосовно однієї і загальною стосовно інших гіпотез.

Окрім загальних і часткових гіпотез, існують гіпотези наукові й робочі.

**Наукова гіпотеза** – це гіпотеза, що пояснює закономірність розвитку явищ природи і суспільства. Такими є, наприклад, гіпотеза про походження сонячної системи, гіпотеза про походження життя, гіпотеза про походження людини, вулканів, нафти тощо.

**Робоча гіпотеза** – це тимчасове припущення або здогад, яким користуються, будуючи гіпотези. Робоча гіпотеза є припущенням – пробою, тимчасовим варіантом, що допомагає побудувати ту чи іншу гіпотезу. Робоча гіпотеза дає змогу перевірити, чи можна це явище якось пояснити. Висунувши робочу гіпотезу і переконавшись, що вона не може пояснити явище, котре нас цікавить, чи пояснює його неправильно, її відкидають, замінюють іншою робочою гіпотезою. Робоча гіпотеза створюється як тимчасовий здогад, тобто таке припущення, котре пояснює явище умовно. За допомогою таких робочих гіпотез тимчасово групують факти, а потім уже її формулюють. Робоча гіпотеза може стати в ході подальшого дослідження науковою гіпотезою.

### **3. Моделювання як один із методів перевірки гіпотези.**

Для систематизації, обробки і аналізу даних спостережуваних явищ з метою встановлення притаманних для них статистичних закономірностей необхідне використання специфічних методів математичної статистики та інформаційних технологій. Застосування багатого арсеналу методів статистичної обробки експериментальних даних та економіко-математичного моделювання є важливою складовою успішних досліджень та якісного аналізу в сільськогосподарській сфері.

Модель – відтворення чи відображення об'єкту, задуму (конструкцій), опису чи розрахунків, що відображає, імітує, відтворює принципи в внутрішньої організації або функціонування, певні властивості, ознаки чи/та характеристики об'єкта дослідження чи відтворення (оригіналу). Модель — це проєкт, інформаційне, натурно-матеріальне чи описово-макетне уявлення предмета. Об'єкт або явище, що є тотожною чи спрощеною версією модельованого об'єкта, проєкта чи явища (прототипу).



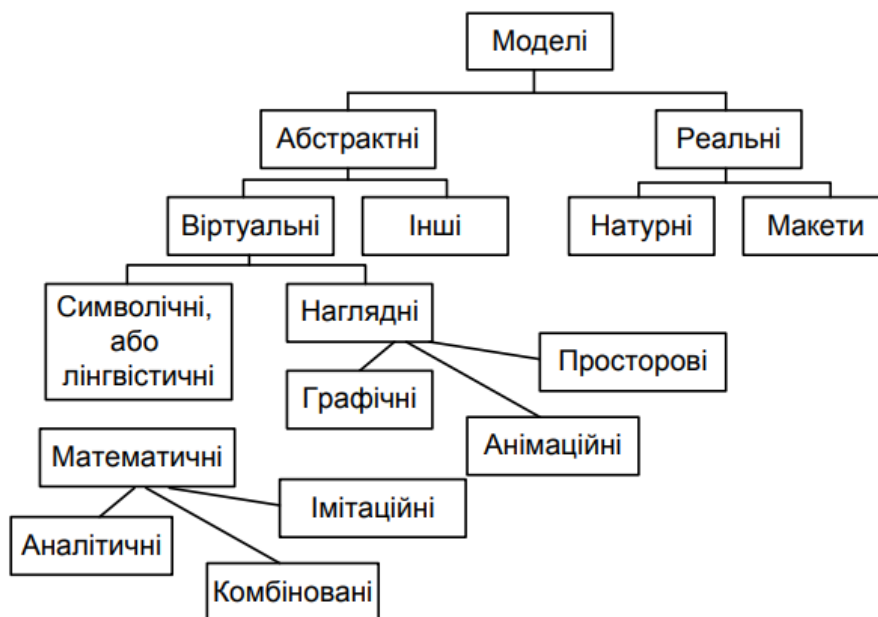


Рис. Основні типи моделей

Моделювання – це процес створення та дослідження моделі.

Моделювання – це вивчення об'єкта шляхом створення та дослідження його копії (моделі), яка за своїми властивостями відтворює властивості об'єкта, що досліджується. Моделювання використовується тоді, коли безпосереднє вивчення об'єктів з деяких причин неможливе.

Моделі, які використовуються в науці, поділяються на дві групи: До першої групи відносяться всі матеріальні предметні моделі, які імітують структуру або функції об'єкта й безпосередньо сприймаються органами відчуття. До другої групи належать обчислювальні моделі, які існують як відображення об'єктів, що не можуть бути безпосередньо сприйняті за допомогою органів відчуттів. Їх поділяють на наочно-образні і логікосимволічні. Наочно-образні моделі бувають у вигляді слів, схем, креслень або просторових конструкцій. Логіко-символічні (знакові) моделі будуються як логічні і математичні обрахунки, в них особливості реальних явищ представлені символами, і тому їх називають математичними моделями.

#### Питання для контролю знань:

1. Надайте визначення: «експеримент», «дослід», «спостереження», «гіпотеза», «припущення», «модель», «моделювання».
2. Зміст гіпотези.
3. Особливості робочої гіпотези.
4. Гіпотеза. Стадії розвитку гіпотези.
5. Вимоги до гіпотези.
6. Призначення моделювання.
7. Основні типи моделей.

#### Завдання для самостійної підготовки:

- I. Вивчити матеріал лекції.
- II. Підготувати доповідь за темою:

1. Параметричні методи класифікації без навчання.
  2. Сутність ієрархічних методів кластерного аналізу.
  3. Методи розпізнавання образів з навчанням.
  4. Методи розпізнавання образів без навчанням.
  5. Базовий алгоритм методу середнього зв'язку Кінга.
  6. Базовий алгоритм методу  $k$ -середніх Мак-Куїна.
  7. Сутність дискримінантного аналізу.
- III. Пройти тестування за темою лекції: тест 3.1. за посиланням <https://forms.gle/bvLHvRdZAWBvMehQ8>

## ЛЕКЦІЯ № 4 ПОНЯТТЯ ФАКТОРУ. ФАКТОРНИЙ ПРОСТІР

### Література:

1. Карагодова О.О., Рожок В.Д. Дослідження операцій: навч. посіб. К.: Центр учбової літератури, 2007. 256 с.
2. Бахрушин В.Є. Методи аналізу даних: навчальний посібник для студентів. Запоріжжя: КПУ, 2011. 268 с
3. Перегуда О.В., Капустян О.А., Курилко О.Б. Статистична обробка даних: навч. посіб. Електронне видання, 2022. 103 с.
4. Рогальський Ф.Б., Курилович Я.Є., Цокурєнка А.А. Математичні методи аналізу економічних систем. Книга 1. Київ: Наукова думка, 2001. 435 с.
5. Рогальський Ф.Б., Цокурєнка А.А. Математичні методи аналізу економічних систем. Книга 2. Київ: Наукова думка, 2001. 423 с.
6. Крайчук О.В., Московська Г.К., Соколенко О.І. Теорія ймовірностей і математична статистика. Рівне, 2004.

### Мета лекції:

1. Навчальна – вивчити основи експериментальних досліджень.
2. Виховна – виховувати професійні риси, почуття відповідальності; виховувати зацікавленість дисципліною, прагнення отримати нові знання самостійно.
3. Розвиваюча – придбання нових знань щодо основ експериментальних досліджень з метою вирішення задач у сфері пожежної безпеки.

### План:

1. Завдання оптимізації.
2. Оптимізація технологічних процесів з використанням планування експерименту.

### Теоретична частина:

**Експеримент** – система операцій, впливів і (або) спостережень, спрямованих на одержання інформації про об'єкт при дослідницьких випробуваннях.

**Дослід** – відтворення досліджуваного явища в певних умовах проведення експерименту при можливості реєстрації його результатів.

**План експерименту** – сукупність даних, що визначають число, умови й порядок реалізації дослідів.

**Планування експерименту** – вибір плану експерименту, що задовольняє заданим вимогам.

**Фактор** (Параметр) – змінна величина, по припущенню, що впливає на результати експерименту.

**Рівень фактору** – фіксоване значення фактору відносно початку відліку.

**Основний рівень фактору** – натуральне значення фактору, відповідне до нуля в безрозмірній шкалі.

**Нормалізація факторів** – перетворення натуральних значень факторів у безрозмірні значення.

**Апріорне ранжирування факторів** – метод вибору найбільш важливих факторів, заснований на експертній оцінці.

**Розмах варіювання фактору** – різниця між максимальним і мінімальним натуральними значеннями фактору в даному плані.

**Інтервал варіювання фактору** – половина розмаху варіювання фактору.

**Ефект взаємодії факторів** – показник залежності зміни ефекту одного фактору від рівнів інших факторів.

**Факторний простір** – простір, координатні осі якого відповідають значенням факторів.

**Область експериментування** (область планування) – область факторного простору, де можуть розміщатися крапки, що відповідають умовам проведення дослідів.

### **1. Завдання оптимізації.**

Після виявлення факторів і вибору вихідних величин формують факторний простір досліджуваної системи – встановлюють область визначення і деякі найвірогідніші (середні) значення рівня варіювання факторів, які підлягають зміні параметрів об'єкта, вихідних величин і критеріїв ефективності.

Область визначення факторів обмежується можливими крайніми їх значеннями ( $x_{\min}$ ,  $x_{\max}$ ), а рівні варіювання отримують шляхом розбиття інтервалу на деяке число рівних підінтервалів (зазвичай від 2 до 6 ... 12) залежно від розміру інтервалу і цілей дослідження.

Обмеження рівнів факторів обумовлені їх фізичною природою, техніко-економічними міркуваннями, обладнанням та приладами що застосовують, а також наявною інформацією про дослідження, які раніше виконувалися.

Основний рівень (нульова точка) являє собою центр досліджуваної області зміни даного фактора. Якщо завданням експерименту є оптимізація деякого параметра, то нульову точку розміщують якомога ближче до положення, що забезпечує оптимум параметра. Якщо завдання експерименту – отримання моделі даного процесу, то за нульову точку приймають середину інтервалу зміни даного фактору.

Рівні фактора, як правило, вибирають симетричними відносно нульової точки. Для спрощення запису умов експерименту і оброблення даних масштаби по осях вибирають так, щоб верхній рівень фактора відповідав +1, нижній –1, а основний – нулю.

Точність фіксування рівнів фактора визначають стабільністю їх у ході досліду і точністю приладів: висока точність, якщо вимірювання проводиться з похибкою не більше 1%, середня – не більше 5% і низька – більше 10%.

На вибір інтервалів варіювання рівня фактора накладають обмеження «згори» і «знизу». Інтервал варіювання не може бути менше тієї помилки, з якою експериментатор фіксує рівень фактора, інакше його верхній і нижній рівні виявляться неможливим розрізнити. З іншого боку, інтервал не може включати такі рівні фактора, при яких його верхній і нижній рівні опиняються за границями області визначення. Якщо інтервал становить не більше 10% від області визначення фактора, його вважають вузьким, не більше 30% – середнім і в решту випадках – широким.

При вирішенні задачі оптимізації для першої серії дослідів прагнуть вибрати таку підобласть зміни рівня фактора, яка давала б можливість найшвидшого руху до оптимуму. У задачах інтерполяції (описи процесу) інтервал варіювання рівнів факторів охоплює всю область експериментування.

Факторний простір може мати при цьому будь-яку конфігурацію. У разі однофакторного експерименту факторний простір є відрізком прямої на числовій осі фактора  $x_i$ . Для двофакторного експерименту воно має вигляд прямокутника, а для трифакторного – прямокутного паралелепіпеда і т.д. У загальному випадку факторний простір являє собою  $K$ -мірний паралелепіпед.

Після перенесення координат системи факторів  $x_i$ , в центр експерименту, фактори переводять із натуральних змінних, що мають зазвичай фізичну розмірність, в безрозмірні, тобто в кодованій

$$x'_i = \frac{x_i - x_{0i}}{\Delta x_i},$$

де  $\Delta x_i = 0,5 (x_{max} - x_{min})$  – півдіапазон змін  $i$ -го фактора.

де  $\Delta x_i = 0,5 (x_{max} - x_{min})$  – півдіапазон змін  $i$ -го фактора. Кодування призводить до того, що всі фактори можуть змінитися в діапазоні  $-1 \leq x'_i \leq +1$ , що перетворює  $K$ -мірний паралелепіпед в  $K$ -мірний куб (гіперкуб), а еліпсоїд розсіяння вихідних показників – у сферу. При плануванні експериментів для побудови моделей в обмеженій області визначення факторів і параметрів об'єкта фактори найчастіше варіюють на двох (+1 і -1) або на трьох (+1; 0; +1) рівнях через однакові інтервали (проміжки). При необхідності можна застосовувати багаторівневе варіювання змінних.

## 2. Оптимізація процесів з використанням планування експерименту.

Завдання оптимізації можна вирішувати двома способами графічним і аналітичним.

При графічному способі вирішення будують двомірні перетину однієї поверхні відгуку, які суміщають з двомірними січеннями іншої поверхні відгуку. Аналізуючи суміщені двомірні січення, знаходять умовні екстремуми. Для отримання двомірних перерізів в рівняння регресії підставляють значення (ймовірно близькі до оптимальних) усіх факторів, крім будь-яких двох. При заданому значенні функції відгуку отримують залежність між двома факторами, яку на площині можна представити у вигляді кривої. Задаючи різні значення параметра оптимізації, можна побудувати сукупність кривих рівного відгуку. Цим способом можна отримати наочне уявлення про вплив кожної пари факторів на параметр оптимізації.

Графічний метод досить простий і відрізняється великою наочністю, однак він зручний тільки при малому числі факторів. При числі факторів до  $k > 3$  графічний метод виявляється дуже громіздким.

Широко застосовують симплексний метод оптимізації.

Для пошуку оптимуму дуже часто використовують кроковий метод, який передбачає виконання двох етапів:

1) вивчення деякої області поверхні відгуку з метою визначення напрямку руху до оптимуму;

2) рух до оптимуму за встановленим напрямом.

Наприклад, методом стрімкого сходження необхідно знайти максимум величини  $y$ , яка залежить від двох факторів  $x_1$  і  $x_2$ . Функція відгуку  $y = f(x_1, x_2)$  являє собою рівняння поверхні, яку можна представити лініями рівного відгуку на факторній площині  $x_1x_2$  (рис.1).

Поблизу максимального значення величини  $y$  така поверхня має вигляд пагорба, вершина якого представляє собою точку, що відповідає оптимальними умовами протікання процесу. Для пошуку оптимуму проводять повний факторний експеримент  $2^2$  (досліди 1, 2, 3, 4), визначають коефіцієнти лінійного рівняння регресії і знаходять напрям градієнта. Потім проводять стрімке сходження (досліди 5, 6, 7, 8). Прийнявши точку з найбільшим значенням  $y$  (дослід 7) за центр плану, знову проводять повний факторний експеримент  $2^2$  (досліди 9, 10, 11, 12) і визначають новий напрям градієнта. Після цього, рухаючись по градієнту (досліди 13, 14, 15, 16) знаходять максимум (дослід 15).

Симплексом називають найпростішу опуклу геометричну фігуру, утворену множиною до  $k + 1$  незалежних точок у просторі. Точки, що утворюють симплекс, називають його вершинами.

Особливістю симплекс-планування є суміщення процесу вивчення поверхні відгуку й процесу руху по поверхні відгуку. Це досягається тим, що умови проведення кожного чергового дослідів встановлюють на основі оцінки попередніх дослідів, поставлених у вершинах симплекса. Всі дослідів ставлять у вершинах симплексів і рух у факторному просторі здійснюють після кожного дослідів.

Наприклад, необхідно знайти максимум величини  $y$ , яка залежить від змінних  $x_1$  і  $x_2$ . У вихідному симплексі 1, 2, 3 (рис. 2) отримані значення  $y_3 > y_2 > y_1$ . Вершина 1 з найменшим значенням функції відгуку відкинута й замінена вершиною 4. У вершинах 2, 3, 4 отримані значення  $y_4 > y_3 > y_2$ . Вершину 2 відкинули і побудували новий симплекс 3, 4, 5 і т.д. Загальне число дослідів, необхідних для досягнення області оптимуму визначають тим, що спочатку необхідно поставити  $k + 1$  дослідів, а потім кожен крок руху супроводжується проведенням тільки одного додаткового дослідів.

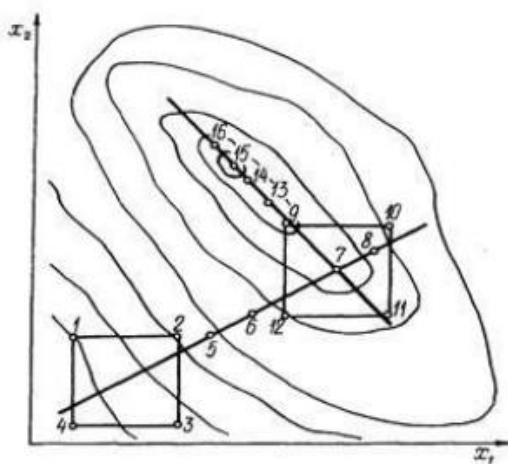


Рисунок 1. Схема руху до оптимуму при стрімкому сходженні

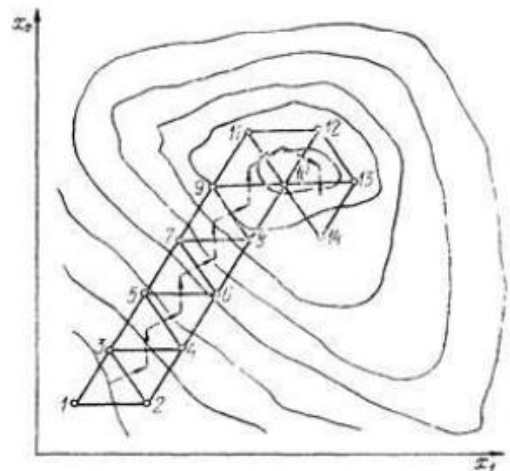


Рисунок 2. Схема руху до області оптимуму при симплексному методі

### **Питання для контролю знань:**

1. Надайте визначення: «експеримент», «дослід», «спостереження», «фактор», «факторний простір».
2. Завдання оптимізації.
3. Факторний простір досліджуваної системи.
4. Область визначення факторів.
5. Обмеження рівнів факторів.
6. Рівні фактора. Точність фіксування рівнів фактора.
7. Метод стрімкого сходження.
8. Симплексний метод оптимізації.
9. Особливість симплекс-планування.

### **Завдання для самостійної підготовки:**

- I. Вивчити матеріал лекції.
- II. Підготувати доповідь за темою:
  1. Форми звітності при науковому дослідженні.
  2. Наукова публікація: поняття, функції, основні види.
  3. Наукова монографія, наукова стаття, тези.
  4. Реферат, доповідь, виступ.
  5. Методика підготовки та оформлення публікацій.
- III. Пройти тестування за темою лекції: тест 4.1. за посиланням <https://forms.gle/5Gx5qqGGFRZHTvjL7>

## ЛЕКЦІЯ № 5 ПОПЕРЕДНЯ ОБРОБКА ДАНИХ

### Література:

1. Карагодова О.О., Рожок В.Д. Дослідження операцій: навч. посіб. К.: Центр учбової літератури, 2007. 256 с.
2. Бахрушин В.Є. Методи аналізу даних: навчальний посібник для студентів. Запоріжжя: КПУ, 2011. 268 с.
3. Перегуда О.В., Капустян О.А., Курилко О.Б. Статистична обробка даних: навч. посіб. Електронне видання, 2022. 103 с.
4. Рогальський Ф. Б., Курилович Я. Є., Цокурєнка А. А. Математичні методи аналізу економічних систем. Книга 1. Київ: Наукова думка, 2001. 435 с.
5. Рогальський Ф. Б., Цокурєнка А. А. Математичні методи аналізу економічних систем. Книга 2. Київ: Наукова думка, 2001. 423 с.
6. Крайчук О.В., Московська Г.К., Соколенко О.І. Теорія ймовірностей і математична статистика. Рівне, 2004.

### Мета лекції:

1. Навчальна – вивчити загальні відомості про методи обробки експериментальних даних.
2. Виховна – ознайомитися з методами обробки експериментальних даних.
3. Розвиваюча – прививати навички наукового підходу у вирішенні завдань застосування методів обробки експериментальних даних.

### План:

1. Аналіз даних. Розділи аналізу даних.
2. Попередня обробка даних.
  - 2.1. Зміст та завдання попередньої обробки експериментальних даних.
  - 2.2. Методи обробки експериментальних даних.
  - 2.3. Критерії відсіювання завідомо помилкових даних.

### Теоретична частина:

#### **1. Аналіз даних. Розділи аналізу даних.**

**Аналіз даних** – розділ математики, що займається розробкою методів обробки даних незалежно від їх природи.

Аналіз даних включає виконання послідовних, логічних дій з інтерпретації зібраних даних (наприклад, відповідей респондентів) та їх перетворення на статистичні форми, потрібні для ухвалення керівницьких рішень.

Можна виділити такі етапи аналізу даних: отримання даних, обробка, аналіз та інтерпретація результатів обробки.

Аналіз даних можна вважати прикладним розділом математичної статистики, проте потрібно наголосити, що аналіз даних охоплює обробку як кількісних, так і якісних даних. Причому, не обов'язково використання імовірнісних моделей в описі досліджуваних об'єктів, явищ та процесів.

Розділи аналізу даних:

1. Розвідувальний аналіз



2. Кореляційний аналіз
3. Дисперсійний аналіз
4. Регресійний аналіз
5. Коваріаційний аналіз
6. Дискримінантний аналіз
7. Кластерний аналіз
8. Аналіз часових рядів

**Розвідувальний аналіз** займається попереднім експрес-аналізом даних шляхом їх перетворення та/або представлення у зручному вигляді: графічному, табличному, схем, діаграм і т.д.

**Кореляційний аналіз** – це статистичне дослідження (стохастичної) залежності між випадковими величинами (англ. *correlation* — взаємозв'язок). У найпростішому випадку досліджують дві вибірки (набори даних), у загальному — їх багатовимірні комплекси (групи). **Мета кореляційного аналізу** – виявити чи існує істотна залежність однієї змінної від інших.

**Дисперсійний аналіз** (англ. *analysis of variance (ANOVA)*) являє собою статистичний метод аналізу результатів, які залежать від якісних ознак. Кожен фактор може бути дискретною чи неперервною випадковою змінною, яку розділяють на декілька сталих рівнів (градацій, інтервалів). Якщо кількість вимірювань (проб, даних) на всіх рівнях кожного з факторів однакова, то дисперсійний аналіз називають рівномірним, інакше – нерівномірним.

**Регресійний аналіз** – розділ математичної статистики, присвячений методам аналізу залежності однієї величини від іншої. На відміну від кореляційного аналізу не з'ясовує чи істотний зв'язок, а займається пошуком моделі цього зв'язку, вираженої у функції регресії.

**Коваріаційний аналіз** – розділ аналізу даних, що намагається визначити модель зв'язку між залежною величиною, та набором кількісних та якісних величин. Таким чином він є ніби синтезом регресійного та дисперсійного аналізів.

**Дискримінантний аналіз** – різновид багатовимірного аналізу, призначеного для вирішення задач розпізнавання образів. Використовується для прийняття рішення про те, які змінні розділяють (тобто «дискримінують») певні масиви даних (так звані «групи»).

**Кластерний аналіз** (англ. *Data clustering*) – задача розбиття заданої вибірки об'єктів (ситуацій) на підмножини, які називаються кластерами, так, щоб кожен кластер складався з схожих об'єктів, а об'єкти різних кластерів істотно відрізнялися. Завдання кластеризації відноситься до статистичної обробки, а також до широкого класу завдань навчання без вчителя.

**Аналіз часових рядів** – сукупність математико-статистичних методів аналізу, призначених для виявлення структури часових рядів і для їх прогнозування. Сюди належать, зокрема, методи регресійного аналізу. Виявлення структури часового ряду необхідно для того, щоб побудувати математичну модель того явища, яке є джерелом аналізованого часового ряду. Прогноз майбутніх значень часового ряду використовується для ефективного прийняття рішень.

## 2. Попередня обробка даних.

**Попередня обробка** - розділ аналізу даних що займається отриманням характеристик для подальшого використання у наступних розділах аналізу даних.

1. Обчислення базових характеристик (центральні моменти)
2. Перевірка основних гіпотез (симетричності, однорідності)
3. Перевірка стохастичності вибірки
4. Видалення аномальних спостережень
5. Розвідувальний аналіз

### **2.1. Зміст та завдання попередньої обробки експериментальних даних.**

Результати вимірювань – це випадкові величини, тобто в ході експерименту інформація спотворена перешкодами, і за одних і тих же умов можна отримати різні дані.

Зміст попередньої обробки даних полягає у відсіюванні грубих похибок і оцінці достовірності результатів вимірювань. Попередня обробка результатів вимірювань необхідна для того, щоб надалі, при побудові функцій відгуку, з найбільшою ефективністю використовувати статистичні методи і коректно аналізувати отримані результати.

Завданням попередньої обробки даних є перевірка відповідності результатів вимірювання нормальному закону і визначення параметрів цього розподілу. Якщо відгук суперечить нормальному розподілу, то слід визначити, якому закону розподілу підлягають дослідні дані або, якщо це можливо, перетворити досліджуваний розподіл до нормального вигляду.

### **2.2. Методи обробки експериментальних даних.**

#### **Метод найменших квадратів**

Для обробки експериментальних даних найчастіше на практиці використовують метод найменших квадратів - один з методів теорії помилок, що використовується для оцінки невідомих величин за наслідками вимірювань, що містить випадкові помилки (спричиняються різного роду випадковими причинами, які діють при кожному з окремих вимірювань непередбаченим чином).

Метод найменших квадратів запропонував К. Гаус (1794–95) і А. Лежандром (1805–06). Строге математичне обґрунтування методу було дано А.А. Марковим (старшим) і А.Н. Колмогоровим. Нині цей метод є одним з найважливіших розділів математичної статистики і широко використовується для статистичних висновків в різних областях науки і техніки.

Суть методу найменших квадратів (по Гаусу) полягає в припущенні, що «збиток» від заміни точного (невідомого) значення фізичної величини  $m$  її наближеним значенням  $X$ , обчисленим за наслідками спостережень, пропорційний квадрату помилки:

$$(X-m)^2$$

Оптимальною оцінкою визнають величину  $X$ , позбавлену систематичної помилки, для якої середнє значення «збитку» мінімальне. Задачу пошуку оптимальної оцінки звужують і як  $X$  вибирають лінійну функцію від результатів спостережень, позбавлену систематичної помилки, і таку, для якої середнє значення «збитку» мінімальне в класі всіх лінійних функцій. Якщо випадкові

помилки спостережень мають нормальний розподіл і оцінювана величина  $m$  залежить від середніх значень результатів спостережень лінійно, то рішення цієї задачі одночасно буде і рішенням загальної задачі. Оцінка  $X$ , обчислена згідно методу найменших квадратів – найвірогідніше значення невідомого параметра  $m$ .

Метод найменших квадратів дає найбільш бажаний результат тоді, коли випадкова помилка має порівняно невелику величину. В іншому разі необхідним є проведення попередньої обробки експериментальних даних, яка полягає в наступному: вихідні записи випадкових величин згладжуються певним способом, що дає змогу виявити основну тенденцію у їхній зміні.

### **Метод виключення перешкод**

Метод виключення перешкод полягає у проведенні на око середньої лінії, яка враховує тільки основні коливання змінної. Інформація, що надалі буде зніматись із цієї середньої лінії, при математичній обробці даних буде використовуватись як вихідна.

Значення випадкової величини, що не збігаються із середньою лінією, прямо не впливатимуть на подальші висновки.

### **Метод оновлюваної середньої**

Метод оновлюваної середньої полягає у використанні рекурентної формули для обчислення середнього арифметичного. Якщо випадкова величина  $x$  надходить у вигляді дискретних вимірів  $i$  для  $(N-1)$ -го виміру обчислено середнє значення, то поява нового виміру змінює попереднє середнє значення на величину

$$\frac{1}{N}(x_N - x_{N-1})$$

При згладжуванні цим методом у кожній точці на часовій осі вимірне значення замінюється на середнє, розраховане на даний момент часу. Послідовність обчислених за рекурентною формулою середніх значень є позбавленим від перешкод рядом вимірів змінної  $x$ , який використовується при подальшій обробці.

### **Метод ковзної середньої**

Метод ковзної середньої полягає в послідовному усередненні на деякому інтервалі ту значень вимірюваної величини  $x$ . Рухаючи ту уздовж осі часу для всіх точок  $t_i$ , що попали в нього, відповідні значення  $x_i$  замінимо середніми значеннями; віднесемо ці значення до середини відповідного інтервалу. Операція згладжування виконується за формулою:

$$\bar{x}_{i+\frac{l}{2}} = \frac{1}{l+1} \sum_{k=0}^l x_{i+k}; k = 0, 1, \dots, l$$

де  $i$  – номер інтервалу;  $(i+1)$  – число вимірювань у  $i$ -тому інтервалі;  $(i + \frac{l}{2})$  – номер замінюваного вимірювання.

Якщо попередня оцінка виконана невдало і після згладжування залишаються перешкоди, утворені дані знову можна піддати усередненню і робити це багаторазово, до бажаного результату.

## Метод експоненційного згладжування

Цей метод має найширші можливості. Тут використовують наступні формули:

$$\bar{x}_N = \alpha \cdot x_N + (1 - \alpha)\bar{x}_{N-1} \text{ або } \bar{x}_N = \bar{x}_{N-1} + \alpha \cdot (x_N - \bar{x}_{N-1})$$

де  $\alpha$  – параметр згладжування, який вибирається в діапазоні 0..1. Тут обчислювана середня заміняє відповідне значення вимірюваної змінної.

Від величини  $\alpha$  залежать згладжуючі властивості методу. При різних  $\alpha$  можна добувати з вихідної інформації високочастотні або низькочастотні складові. Таким чином, з'являється можливість боротися з низько – і високочастотними шумами, тобто зі швидко – і повільно змінними у часі перешкодами.

При малих, близьких до 0  $\alpha$ , сукупність обчислених середніх відобразить низькочастотні зміни вимірюваної змінної, позбавивши їх швидкозмінних перешкод. У цьому випадку говорять про *інерційне згладжування*, при якому обчислювана середня мало залежить від останнього вимірювання і значною мірою від середньої, утвореної на попередньому кроці.

### 2.3. Критерії відсіювання завідомо помилкових даних.

Дані, які відповідають умовам, що змінилися, називають грубими помилками або значеннями, що різко виділяються (аномальними). У разі відсіву грубих помилок формулюється нульова гіпотеза:  $H_0$ : "Серед результатів спостережень (вибіркових, дослідних даних) немає значень, що різко виділяються (аномальних)".

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 більш грубих помилок". Найпоширенішими і теоретично обґрунтованими в цьому випадку є

- критерій Н.В. Смірнова (використовується при  $H_1^{(1)}$ )
- критерій Діксона (застосовується як при  $H_1^{(1)}$  так і при  $H_1^{(2)}$ )

### Питання для контролю знань:

1. Надайте визначення термінам «аналіз даних», «спостереження», «експеримент», «імітація», «попередня обробка даних».
2. Зміст попередньої обробки даних.
3. Завдання попередньої обробки даних.
4. Методи обробки експериментальних даних.
5. Сутність методу найменших квадратів.
6. Сутність методу виключення перешкод.
7. Сутність методу оновлюваної середньої.
8. Сутність методу ковзної середньої.
9. Сутність методу експоненційного згладжування.
10. Критерії відсіювання завідомо помилкових даних.

### **Завдання для самостійної підготовки:**

- I. Вивчити матеріал лекції.
- II. Підготувати доповідь за темою:
  1. Статистичні методи оцінки експериментальних даних.
  2. Послідовність статистичного дослідження.
  3. Статистична оцінка імовірності досліджень. Оцінка похибок вимірювання.
  4. Статистичні критерії розходження.
  5. Критерій Н.В. Смірнова.
  6. Критерій Діксона.
- III. Пройти тестування за темою лекції: тест 5.1. за посиланням <https://forms.gle/5sHmFqXQ1wAhpoTU8>

## ЛЕКЦІЯ № 6

### ЗАГАЛЬНІ ВІДОМОСТІ ПРО МЕТОДИ ОБРОБКИ ЕКСПЕРИМЕНТАЛЬНИХ ДАНИХ

#### Література:

1. Карагодова О.О., Рожок В.Д. Дослідження операцій: навч. посіб. К.: Центр учбової літератури, 2007. 256 с.
2. Бахрушин В.Є. Методи аналізу даних: навчальний посібник для студентів. Запоріжжя: КПУ, 2011. 268 с
3. Перегуда О.В., Капустян О.А., Курилко О.Б. Статистична обробка даних: навч. посіб. Електронне видання, 2022. 103 с.
4. Рогальський Ф.Б., Курилович Я.Є., Цокурєнко А.А. Математичні методи аналізу економічних систем. Книга 1. Київ: Наукова думка, 2001. 435 с.
5. Рогальський Ф.Б., Цокурєнка А.А. Математичні методи аналізу економічних систем. Книга 2. Київ: Наукова думка, 2001. 423 с.
6. Крайчук О.В., Московська Г.К., Соколенко О.І. Теорія ймовірностей і математична статистика. Рівне, 2004.

#### Мета лекції:

1. Навчальна – вивчити загальні відомості про методи обробки експериментальних даних.
2. Виховна – виховувати професійні риси, почуття відповідальності; виховувати зацікавленість дисципліною, прагнення отримати нові знання самостійно.
3. Розвиваюча – прививати навички наукового підходу під час вирішення задач у сфері забезпечення пожежної безпеки.

#### План:

1. Загальні відомості про методи обробки експериментальних даних.
  - 1.1. Мода.
  - 1.2. Медіана.
  - 1.3. Вибіркове середнє.
  - 1.4. Розкид вибірки.
  - 1.5. Дисперсія.

#### Теоретична частина:

Методами статистичної обробки результатів експерименту називаються математичні прийоми, формули, способи кількісних розрахунків, за допомогою яких показники, отримані в ході експерименту, можна узагальнювати, приводити в систему, виявляючи приховані в них закономірності. Мова йде про такі закономірності статистичного характеру, які існують між досліджуваними в експерименті змінними величинами.

Деякі з методів математико-статистичного аналізу дозволяють обчислювати так звані елементарні математичні статистики, що характеризують вибіркоче розподіл даних, наприклад вибіркоче середнє, вибіркоче дисперсія, мода, медіана і ряд інших. Інші методи математичної статистики, наприклад дисперсійний аналіз, регресійний аналіз, дозволяють судити про динаміку зміни

окремих статистик вибірки. За допомогою третьої групи методів, скажімо, кореляційного аналізу, факторного аналізу, методів порівняння вибірових даних, можна достовірно судити про статистичні зв'язки, що існують між змінними величинами, які досліджують у даному експерименті.

## **1. Методи первинної статистичної обробки результатів експерименту**

Всі методи математико-статистичного аналізу умовно діляться на первинні і вторинні. Первинними називають методи, за допомогою яких можна отримати показники, безпосередньо відображають результати вироблених в експерименті вимірювань. Вторинними називаються методи статистичної обробки, за допомогою яких на базі первинних даних виявляють приховані в них статистичні закономірності.

До первинних методам статистичної обробки відносять, наприклад, визначення вибіркової середньої величини, вибірковою дисперсії, вибіркової моди і вибіркової медіани. У число вторинних методів звичайно включають кореляційний аналіз, регресійний аналіз, методи порівняння первинних статистик у двох або декількох вибірок.

Розглянемо методи обчислення елементарних математичних статистик.

### **1.1 Мода.**

Числовою характеристикою вибірки, як правило, не вимагає обчислень, є так звана мода. Моду називають кількісне значення досліджуваної ознаки, найбільш часто зустрічається у вибірці. Для симетричних розподілів ознак, в тому числі для нормального розподілу, значення моди збігається зі значеннями середнього і медіани. Для інших типів розподілу, несиметричних, це не характерно. Наприклад, у послідовності значень ознак 1, 2, 5, 2, 4, 2, 6, 7, 2 модою є значення 2, так як воно зустрічається частіше за інших значень - чотири рази.

Моду знаходять згідно з такими правилами:

1) У тому випадку, коли всі значення у вибірці зустрічаються однаково часто, прийнято вважати, що цей вибіровий ряд не має моди. Наприклад: 5, 5, 6, 6, 7, 7 - у цій вибірці моди немає.

2) Коли два сусідніх (суміжних) значення мають однакову частоту і їх частота більше частот будь-яких інших значень, мода обчислюється як середнє арифметичне цих двох значень. Наприклад, у вибірці 1, 2, 2, 2, 5, 5, 5, 6 частоти поруч розташованих значень 2 та 5 збігаються і дорівнюють 3. Ця частота більше, ніж частота інших значень 1 та 6 (у яких вона дорівнює 1). Отже, модою цього ряду буде величина  $= 3,5$

3) Якщо два несуміжні (не сусідніх) значення у вибірці мають рівні частоти, які більше частот будь-якого іншого значення, то виділяють дві моди. Наприклад, у ряді 10, 11, 11, 11, 12, 13, 14, 14, 14, 17 модами є значення 11 і 14. У такому випадку говорять, що вибірка є бімодальною.

Можуть існувати і так звані мультимодальні розподілу, що мають більше двох вершин (мод).

4) Якщо мода оцінюється по безлічі згрупованих даних, то для знаходження моди необхідно визначити групу з найбільшою частотою ознаки. Ця група називається модальною групою.

## 1.2. Медіана.

Медіаною називається значення досліджуваної ознаки, яке ділить вибірку, впорядковану за величиною даної ознаки, навпіл. Праворуч і ліворуч від медіани в упорядкованому ряду залишається за однаковою кількістю ознак. Наприклад, для вибірки 2, 3, 4, 4, 5, 6, 8, 7, 9 медіаною буде значення 5, так як зліва і праворуч від нього залишається по чотири показника. Якщо ряд включає в себе парне число ознак, то медіаною буде середнє, взяте як полусумма величин двох центральних значень ряду. Для наступного ряду 0, 1, 1, 2, 3, 4, 5, 5, 6, 7 медіана буде дорівнює 3,5.

Знання медіани корисно для того, щоб встановити, чи є розподіл приватних значень вивченої ознаки симетричним і наближається до так званого нормального розподілу. Середня і медіана для нормального розподілу зазвичай збігаються або дуже мало відрізняються один від одного.

Якщо вибіркоче розподіл ознак нормально, то до нього можна застосовувати методи вторинних статистичних розрахунків, засновані на нормальному розподілі даних. В іншому випадку цього робити не можна, тому що в розрахунках можуть бути серйозні помилки.

## 1.3. Вибіркове середнє.

Вибіркове середнє (середнє арифметичне) значення як статистичний показник являє собою середню оцінку досліджуваної в експерименті якості. Ця оцінка характеризує ступінь його розвитку в цілому у тій групі досліджуваних, яка була піддана обстеженню. Порівнюючи безпосередньо середні значення двох або декількох вибірок, ми можемо судити про відносну ступеня розвитку, що складають ці вибірки, оцінюється якості.

Вибіркове середнє визначається за допомогою наступної формули:

$$\bar{X} = \frac{(X_1 + X_2 + \dots + X_n)}{n} = \frac{1}{n} \cdot \left( \sum_{i=1}^n X_i \right)$$

де  $x$  – вибіркова середня величина або середнє арифметичне значення по вибірці;  $n$  – кількість випробуваних у вибірці показників, на основі яких обчислюється середня величина;  $x_k$  – приватні значення показників у окремих досліджуваних. Усього таких показників  $n$ , тому індекс  $k$  даної змінної приймає значення від 1 до  $n$ ;  $\Sigma$  – прийнятий у математиці знак підсумовування величин тих змінних, які знаходяться праворуч від цього знаку. Вираз відповідно означає суму всіх  $x$  з індексом  $k$ , від 1 до  $n$ .

## 1.4. Розкид вибірки.

Розкид (іноді цю величину називають розмахом) вибірки позначається літерою  $R$ . Це самий простий показник, який можна отримати для вибірки – різниця між максимальною і мінімальною величинами даного конкретного варіаційного ряду, тобто

$$R = X_{\max} - X_{\min}$$

Зрозуміло, що чим сильніше варіює вимірюваний ознака, тим більше величина  $R$ , і навпаки. Проте може статися так, що у двох вибіркових рядів і середні, і розмах збігаються, проте характер варіювання цих рядів буде різний. Наприклад, дано дві вибірки:



$$X = 10 \ 15 \ 20 \ 25 \ 30 \ 35 \ 40 \ 45 \ 50 \quad X = 30 \quad R = 40$$

$$Y = 10 \ 28 \ 28 \ 30 \ 30 \ 30 \ 32 \ 32 \ 50 \quad Y = 30 \quad R = 40$$

При рівності середніх і розкидів для цих двох вибірових рядів характер їх варіювання різний. Для того щоб більш чітко уявляти характер варіювання вибірок, слід звернутися до їх розподілом.

### 1.5. Дисперсія.

Дисперсія - це середнє арифметичне квадратів відхилень значень змінної від її середнього значення.

Дисперсія як статистична величина характеризує, наскільки приватні значення відхиляються від середньої величини в даній вибірці. Чим більше дисперсія, тим більше відхилення або розкид даних.

$$D = \frac{1}{n} \cdot \sum_{i=1}^n (x_i - \bar{x})^2$$

де  $D$  - вибіркова дисперсія, або просто дисперсія;

$\sum$  - вираз, що означає, що для всіх  $x$ , від першого до останнього в даній вибірці необхідно обчислити різниці між приватними і середніми значеннями, звести ці різниці в квадрат і підсумувати;

$n$  - кількість випробуваних у вибірці чи первинних значень, за якими обчислюється дисперсія. Проте сама дисперсія, як характеристика відхилення від середнього, часто незручна для інтерпретації. Для того, щоб наблизити розмірність дисперсії до розмірності вимірюваної ознаки застосовують операцію добування квадратного кореня з дисперсії. Отриману величину називають стандартним відхиленням.

Із суми квадратів, ділених на число членів ряду витягується квадратний корінь.

$$S_x = \sqrt{\frac{1}{n} \cdot \sum_{i=1}^n (x_i - \bar{x})^2}$$

Іноді вихідних приватних первинних даних, які підлягають статистичній обробці, буває досить багато, і вони вимагають проведення великої кількості елементарних арифметичних операцій. Для того щоб скоротити їх кількість і разом з тим зберегти потрібну точність розрахунків, іноді вдаються до заміни вихідної вибірки приватних емпіричних даних на інтервали. Інтервалом називається група впорядкованих за величиною значень ознаки, заміна в процесі розрахунків середнім значенням.

### Питання для контролю знань:

1. Методи статистичної обробки результатів експерименту.
2. Первинні методи математико-статистичного аналізу.
3. Вторинні методи математико-статистичного аналізу.
4. Визначення «мода».
5. Правила знаходження моди.
6. Значення понять «медіана», «вибіркове середнє», «розкид вибірки», «дисперсія».

### **Завдання для самостійної підготовки:**

- I. Вивчити матеріал лекції.
- II. Підготувати доповідь за темою:
  1. Аналіз сучасного стану на ринку статистичних пакетів.
  2. Суть та організація статистичного зведення.
  3. Візуалізація та графічне відображення даних.
  4. Статистичні критерії розходження.
  5. Застосування t-критерію Стьюдента (t-критерій).
  6. Застосування критерію Фішера (F-критерій).
  7. Застосування критерію Кохрена (G-критерій).
  8. Застосування критерію узгодженості Пірсона ( $\chi^2$ -критерій).
- III. Пройти тестування за темою лекції: тест 6.1. за посиланням <https://forms.gle/k4DnxFxCr5Ek1gfMA>

## ЛЕКЦІЯ № 7 ФАКТОРНИЙ АНАЛІЗ

**Місце проведення:** за розкладом.

**Час:** 2 акад. години.

**Матеріальне забезпечення:** 1. Мультимедійний проектор, екран.

2. Література:

1. Карагодова О.О., Рожок В.Д. Дослідження операцій: навч. посіб. К.: Центр учбової літератури, 2007. 256 с.
2. Бахрушин В.Є. Методи аналізу даних: навчальний посібник для студентів. Запоріжжя: КПУ, 2011. 268 с
3. Перегуда О.В., Капустян О.А., Курилко О.Б. Статистична обробка даних: навч. посіб. Електронне видання, 2022. 103 с.
4. Рогальський Ф. Б., Курилович Я. Є., Цокурєнко А. А. Математичні методи аналізу економічних систем. Книга 1. Київ: Наукова думка, 2001. 435 с.
5. Рогальський Ф. Б., Цокурєнка А. А. Математичні методи аналізу економічних систем. Книга 2. Київ: Наукова думка, 2001. 423 с.
6. Крайчук О.В., Московська Г.К., Соколенко О.І. Теорія ймовірностей і математична статистика. Рівне, 2004.

### Мета лекції:

1. Навчальна – вивчити основи застосування факторного аналізу.
2. Розвиваюча – прививати навички наукового підходу під час вирішення задач у сфері забезпечення пожежної безпеки.
3. Виховна – виховувати професійні риси, почуття відповідальності; виховувати зацікавленість дисципліною, прагнення отримати нові знання самостійно.

### План:

1. Сутність та завдання факторного аналізу.
2. Класифікація та сфери застосування методів факторного аналізу.
3. Загальна характеристика методів факторного аналізу.

### Теоретична частина:

Факторний аналіз дозволяє виявити об'єктивно існуючі фактори які впливають на поведінку об'єкта, що досліджується. Дуже важливим є те, що, як правило, ці фактори безпосередньо не споглядаються, а лише якимось чином впливають на досліджувані параметри.

Авторами основних концепцій факторного аналізу є американські й англійські вчені (Ч. Спірмен, Л. Л. Терстоун, Г. Х. Томсон, С. Л. Барт, Р. Б. Кеттелл).

#### 1. Сутність та завдання факторного аналізу.

Суть методу факторного аналізу полягає у виключенні опису множини характеристик, що вивчаються, і заміною їх меншою кількістю інформаційно об'ємніших змінних.

Привабливість цього методу полягає в тому, що він дозволяє вводити в обіг та корисно застосовувати великий обсяг вихідних статистичних даних. Відомо, що більшість вчених, прагнучи дослідити певну проблему цілком і повністю, спираються на величезну кількість інформації. Але проаналізувати таку кількість інформації дуже складно. Таким чином, виникає потреба в стисненні інформації до більш реальних розмірів за рахунок відбору найбільш суттєвої та відкиданням другорядної. Саме тоді, коли існує необхідність обмеження кількості змінних і застосовується факторний аналіз.

Сутність методів факторного аналізу полягає в переході від опису деякої множини досліджуваних об'єктів, заданої великим набором непрямих безпосередньо вимірюваних ознак, до їх опису меншим числом максимально інформативних глибинних змінних, що відображають найбільш істотні властивості явища. Такого роду змінні, що називаються факторами, є деякими функціями початкових ознак. В більшості випадків фактори являють собою латентні (скриті) ознаки, які не підлягають прямому вимірюванню, але здійснюють безпосередній вплив на досліджуване явище чи процес.

Факторний аналіз не вимагає апріорного підрозділу змінних на залежні і незалежні, оскільки всі змінні в ньому розглядаються як рівноправні. Завдання факторного аналізу полягає в тому, щоб визначити поняття, кількість і природу найбільш істотних і незалежних функціональних характеристик явища.

Важливою відмінною особливістю факторного аналізу є можливість одночасного дослідження скільки завгодно великої кількості взаємозалежних змінних. Він не вимагає припущення про незмінність всіх інших умов, властивого багатьом іншим методам статистичного аналізу. Завдяки цьому факторний аналіз є цінним інструментом дослідження явища у всьому різноманітті його реальних взаємозв'язків.

Фактор є розрахунковою змінною, тобто якоюсь новою характеристикою об'єктів, що вивчаються. Опис фактора в термінах його зв'язку з набором початкових ознак відшуковується у вигляді так званої факторної матриці, або матриці факторних навантажень розмірністю  $p \times k$ , де  $p$  — кількість вихідних ознак, а  $k$  — число факторів. Основою для побудови факторної матриці служить кореляційна матриця. Вона відображає ступінь взаємозв'язку між кожною парою ознак, тоді як факторна матриця характеризує ступінь зв'язку між кожною з  $p$  даних ознак і  $k$  факторів, виявлених в процесі аналізу. При цьому значення  $k$  обирається виходячи з двох умов:  $k$  повинне бути багато менше за  $p$ , а Рівень втрат в інформації достатньо малим.

Факторна матриця дозволяє виділити для кожного фактора групу параметрів, найтісніше з ним зв'язаних. Тим самим відкривається можливість зіставити фактори один з одним, дати їм змістовне тлумачення і найменування, тобто навести інтерпретацію факторів.

У зв'язку з цим важливо відзначити, що перетворення кореляційної матриці у факторну не є однозначним. Одну і ту ж кореляційну матрицю можна перетворити в різні факторні. Для вибору факторної матриці, на основі якої проводитиметься інтерпретація факторів, розробляються спеціальні прийоми. Як правило, в ролі факторної матриці вибирається така, в якій початкові параметри сильно пов'язані з одним з факторів і слабо пов'язані зі всіма іншими.

Для визначення якості подання вихідних ознак за допомогою системи факторів потрібно ввести деякий показник, який і буде виступати мірилом цієї якості. Необхідно, щоб значення такого показника-критерія можна було б визначити за коефіцієнтами матриці факторних навантажень. Найкращою матрицею факторних навантажень можна вважати ту, для якої побудований критерій приймає найбільше значення.

У факторному аналізі вихідні змінні повинні вимірюватись за допомогою шкали інтервалів. Це пов'язано з тим, що розрахунки базуються на використанні кореляційної (в загальному випадку коваріаційної) матриці. Крім того, подання змінних у вигляді лінійної комбінації скритих факторів і використання оцінок факторів через лінійні комбінації вихідних змінних для порядкових даних неможливе.

## **2. Класифікація та сфери застосування методів факторного аналізу**

Набір сучасних методів факторного аналізу достатньо великий.

Метод головних компонент. Даний метод передбачає одночасне виділення всіх головних компонент, кількість яких дорівнює кількості вихідних ознак. При цьому припускається повне пояснення дисперсії вихідних ознак через виділені латентні фактори. Детальний аналіз метода буде наведений далі.

Метод головних факторів. Дисперсія початкових ознак пояснюється не повністю: частина її залишається нерозпізнаною як результат дії характерних чинників. Фактори виділяються послідовно: спочатку найбільш впливовий, потім іншим за значущістю і т.д. Виділення факторів припиняється, коли приймається рішення про достатність частки поясненої ними дисперсії.

Однофакторний метод Ч.Спірмена. Дозволяє виділити один генеральний фактор і одну характерну ознаку.

Центроїдний метод Л. Терстоуна. Дозволяє виділити декілька латентних факторів і декілька характерностей. Принцип обчислень схожий на метод головних факторів, однак базується на іншому критеріїв оптимізації.

З другої половини 60-х років ХХ століття факторний аналіз одержав визнання як універсальний метод компактного подання великих масивів статистичних і експериментальних даних. В даний час він широко використовується для обробки даних в самих різних областях соціально-економічних досліджень.

Найбільш ефективні результати дає застосування методів факторного аналізу в тих областях теоретичних і прикладних досліджень, де сутність основних категорій визначена недостатньо чітко, або там, де існує необхідність виробити ємні комплексні надійні параметри, які можуть бути використані як шкали для порівняння множини об'єктів, як критерії класифікації об'єктів або як незалежні індикатори процесів при прогнозуванні. Широке застосування знаходять методи факторного аналізу в нових і особливо в суміжних областях досліджень, де необхідний облік множини різнопланових характеристик, вимірюваних з істотними помилками, і особливо там, де найбільш важливі параметри не піддаються безпосередньому вимірюванню.

Широке коло застосування методів факторного аналізу пов'язане з тим, що вони грають важливу роль при вирішенні такої універсальної задачі наукового дослідження, як завдання класифікації.

Перелік конкретних завдань, при рішенні яких використовуються методи факторного аналізу:

- мінімізація опису, тобто визначення основних аспектів відмінностей між об'єктами спостереження;
- формулювання та перевірка гіпотези про природу основних аспектів відмінностей між: об'єктами;
- виявлення структури взаємозв'язків в наборі ознак, перевірка гіпотези про взаємозв'язки і взаємозамінюваність ознак;
- зіставлення структури декількох наборів ознак;
- побудова узагальненого показника (або групи показників);
- ранжування об'єктів спостереження;
- здійснення типологізації об'єктів спостереження;
- перетворення вихідних даних для використання в інших моделях.

### 3. Загальна характеристика методів факторного аналізу.

У факторній моделі передбачається, що поряд зі спільними факторами на вихідні ознаки здійснюють вплив і характерності, які також є латентними ознаками. При знаходженні факторів методами факторного аналізу передбачається, що характерності вже відомі, або їх певним чином можна оцінити. Це означає, що у вихідній кореляційній матриці  $R$  діагональні елементи виражають лише спільності, а тому менші одиниці.

Сутність методів полягає в знаходженні такої матриці факторних навантажень  $W$ , яка задовольняє співвідношенню

$$WW = R - D = R'$$

де  $D$  — діагональна матриця, що виражає характерність.

$R'$  — редукована матриця кореляцій

Відмітимо, що завдання знаходження матриці  $W$  не є однозначним, оскільки матриця факторних навантажень виділяє в просторі  $Z$  підпростір  $F$  загальних факторів, але не фіксує в ньому систему координат. Тобто, матриця факторних навантажень знаходиться з точністю до ортогонального перетворення. Це дозволяє при її обчисленні накладати додаткові вимоги на неї. Однією з них є те, що загальні фактори обчислюються за впорядкованістю їх внесків у пояснення дисперсії вихідних ознак. Якщо при цьому виявляється, що серед знайдених лише декілька мають великі внески, то для практичного використання можна обмежитись лише ними.

На даний час не існує точних процедур побудови редукованої кореляційної матриці; всі вони носять евристичний характер. На практиці редукцію вихідної кореляційної матриці можна провести такими способами:

– методом найбільшої кореляції – діагональний елемент замінюється на найбільше по відповідному рядку (стовпчику) недіагональне значення коефіцієнта кореляції:

– методом Варта – для кожного рядка спочатку знаходиться середнє значення коефіцієнта кореляції; якщо воно порівняно велике, то діагональний елемент замінюється трохи більшим за найбільше по рядку значення коефіцієнта кореляції, а якщо воно порівняно мале, то діагональний елемент замінюється трохи меншим за найбільше по рядку значення коефіцієнта кореляції;

– методом тріад – діагональний елемент обчислюється за формулою:

де гікга гкс- найбільші по рядку значення коефіцієнта кореляції;

– методом середнього значення коефіцієнта кореляції:

– за методом першого центроїдного фактора. В такому випадку діагональні елементи вихідної кореляційної матриці замінюються на додатні значення найбільших за модулем по рядку (стовпчику) коефіцієнтів кореляції. Далі спільності розраховуються як відношення квадрату суми по рядку (стовпчику) модулів коефіцієнтів кореляції до суми модулів цих елементів:

за методом малого центроїдного фактора. Тоді для кожної спільності  $h$ -будується кореляційна матриця  $P$  розмірності  $(4 \times 4)$ . Вона включає відповідний діагональний елемент кореляційної матриці  $r'$ , а також коефіцієнти кореляції для трьох інших ознак, найбільш щільно пов'язаних з  $i$ -тою. За даними цієї матриці і розраховуються спільність:

де чисельник розраховується за значеннями першого стовпчика матриці  $P$ , а знаменник – за всіма значеннями матриці  $P$ ;  $kg$  за значеннями коефіцієнтів множинної кореляції:

де величина  $s_{iii}$ -тим діагональним елементом матриці, оберненої до вихідної кореляційної матриці  $R$ .

Наведені методи дають різний ступінь оцінки спільностей. Так, перший з них доцільно застосовувати у випадку великої кількості ( $p > 10$ ) вихідних ознак. Методи середнього значення та першого центроїдного фактора дають дещо занижені оцінки спільності. Найбільш розповсюдженим є останній метод, оскільки він враховує мультиколінеарність вихідних ознак, яка в свою чергу в статистичному сенсі найбільш повно відповідає поняттю спільності.

При використанні методів факторного аналізу виникає питання, наскільки щільні кореляційні зв'язки між вихідними ознаками. Тобто, наскільки значуща кореляційна матриця вихідних ознак. Адже незначущість такої матриці може означати, що для вихідної сукупності ознак взагалі не існує спільностей, і застосування методів факторного аналізу не є виправданим. В такому випадку виділені фактори фактично будуть являти собою вихідні ознаки.

Оцінка значущості кореляційної матриці може проводитись за критерієм Уїлкіса  $\chi^2$ . Спостережене значення критерію обчислюється за формулою:

Це значення порівнюється з критичним для  $\chi^2$  – розподілу, знайденим при заданому рівні значущості  $\alpha$  та кількості ступенів вільності  $v = p(p-1)/2$ . Значущість кореляційної матриці підтверджується, якщо спостережене значення перевищує критичне:

Знак "-" введений тому, що визначник матриці  $R$  як правило менший 1, а критерій повинен приймати додатні значення. Помітимо також, що чим ближче значення визначника до нуля, тим вищими будуть значення критерію. А близькість визначника до нуля свідчить про лінійну залежність його стовпчиків, а, отже, і про наявність у початкових ознак спільних рис, які підлягають виявленню методами факторного аналізу.

Основним методом факторного аналізу, який найчастіше застосовується, є метод аналізу головних компонентів (principal components analysis), що дає змогу усунути вплив мультиколінеарності зв'язків між показниками, які є в показниках податкової звітності, та зменшити розмірність даних.

Найбільшу складність при використанні факторного аналізу становить інтерпретація факторів, за відсутності якої факторний аналіз втрачає значення. Ця інтерпретація пов'язана з якісним аналізом явища та процесу, що вивчається.

Підсумовуючи, факторний аналіз як статистичний метод, то важливо згадати наступні міркування:

1. Факторний аналіз, на противагу контрольованому експерименту, опирається в основному на спостереження над природним варіюванням змінних.

2. При використанні факторного аналізу сукупність змінних, досліджуваних з погляду зв'язків між ними, не вибирається довільно: цей метод дозволяє виявити основні фактори, що роблять істотний вплив у даній області.

3. Факторний аналіз не вимагає попередніх гіпотез, на-оборот, він сам може служити методом висування гіпотез, а також виступати критерієм гіпотез, що опираються на дані, отримані іншими методами.

4. Факторний аналіз не вимагає апіорних припущень щодо того, які змінні незалежні, а які залежні, він не гіпертрофірує причинні зв'язки й вирішує питання про їхню меру в процесі подальших досліджень.

Завдяки цим властивостям факторний аналіз придбав особливе значення в психології .і суспільних науках, тому що саме в цих областях часто практично не існує можливостей для експериментального контролю окремих впливів.

Таким чином, в теоретичному плані логіка факторного аналізу зводиться по суті до з'ясування питання про те, що первинне і що вторинне. Однак, насправді все значно складніше. Це пояснюється тим, що крім причинно-наслідкової існують і інші форми зв'язку між економічними явищами, при яких розподіл показників на результативні і факторні не є однозначним. Крім того, при більш глибокому дослідженні проблеми виявляється значна кількість різних додаткових обставин, що вимагають більш широкого підходу до її розуміння. Все це передбачає істотне поглиблення теорії індексного факторного аналізу і розширення його понятійного апарату.

З практичної точки зору основними видаються два моменти. По-перше, озброєння практиків початковими науковими принципами утворення двохфакторних індексних моделей з таким розрахунком, щоб при необхідності кваліфіковано скласти саме ту модель будь-якого результативного показника, яка цікавить економіста. По-друге, врахування механізму перетворення двохфакторних індексних моделей в багатфакторні, уникаючи при цьому штучності і формалізму. Вирішення даних задач доцільно почати з розгляду видів функціональної залежності між явищами.

#### **Питання для контролю знань:**

7. У чому полягає сутність факторного аналізу?
8. Назвіть основні етапи факторного аналізу.
9. Які основні групи методів виділяють у факторному аналізі?
10. Які методи розрахунку спільностей ви знаєте?
11. Чим відрізняється редукована матриця від звичайної кореляційної матриці?
12. У чому відмінність методу головних компонент від методів факторного аналізу?



13. Як визначається необхідна та достатня кількість загальних факторів у методі головних факторів?
14. Які види обертання ви знаєте?
15. Яким чином здійснюється економічна інтерпретація отриманих чинників?
16. Який вигляд має лінійна модель методу головних компонент?
17. Назвіть основні поняття факторного аналізу.
18. Наведіть алгоритм методу головних факторів.
19. Як здійснюється розкладання дисперсії у факторному аналізі?
20. Назвіть критерії точності оцінки значень факторів.
21. Як здійснюється оцінювання рівня інформативності головних факторів?
22. Назвіть критерії вибору кількості факторів.

#### **Завдання для самостійної підготовки:**

- I. Вивчити матеріал лекції.
- II. Підготувати доповідь за темою:
  1. Поняття, типи і завдання факторного аналізу.
  2. Класифікація факторів.
  3. Систематизація факторів.
  4. Поняття, типи і завдання факторного аналізу.
  5. Метод головних факторів. Оцінювання факторів і задачі класифікації.
  6. Метод головних компонент.
  7. Приклади реалізації алгоритму методу головних компонент.
- III. Пройти тестування за темою лекції: тест 7.1. за посиланням <https://forms.gle/SUbj1CST8iyQuVDy9>

## ЛЕКЦІЯ № 8

### ДИСПЕРСІЙНИЙ АНАЛІЗ. РЕГРЕСІЙНИЙ АНАЛІЗ

#### Література:

1. Карагодова О.О., Рожок В.Д. Дослідження операцій: навч. посіб. К.: Центр учбової літератури, 2007. 256 с.
2. Бахрушин В.Є. Методи аналізу даних: навчальний посібник для студентів. Запоріжжя: КПУ, 2011. 268 с
3. Перегуда О.В., Капустян О.А., Курилко О.Б. Статистична обробка даних: навч. посіб. Електронне видання, 2022. 103 с.
4. Рогальський Ф. Б., Курилович Я. Є., Цокурєнка А. А. Математичні методи аналізу економічних систем. Книга 1. Київ: Наукова думка, 2001. 435 с.
5. Рогальський Ф. Б., Цокурєнка А. А. Математичні методи аналізу економічних систем. Книга 2. Київ: Наукова думка, 2001. 423 с.
6. Крайчук О.В., Московська Г.К., Соколенко О.І. Теорія ймовірностей і математична статистика. Рівне, 2004.

#### Мета лекції:

1. Навчальна – вивчити основи застосування дисперсійного аналізу.
2. Розвиваюча – ознайомитися з напрямками застосування дисперсійного аналізу.
3. Виховна – виховувати професійні риси, почуття відповідальності; виховувати зацікавленість дисципліною, прагнення отримати нові знання самостійно.

#### План:

1. Загальні відомості про дисперсійний аналіз.
2. Однофакторний дисперсійний аналіз.
3. Двофакторний дисперсійний аналіз.

#### Теоретична частина:

##### 1. Загальні відомості про дисперсійний аналіз.

**Дисперсійний аналіз** являє собою статистичний метод аналізу результатів, які залежать від якісних ознак.

Кожен фактор може бути дискретною чи неперервною випадковою змінною, яку розділяють на декілька сталих рівнів (градацій, інтервалів). Якщо кількість вимірювань (проб, даних) на всіх рівнях кожного з факторів однакова, то дисперсійний аналіз називають рівномірним, інакше – нерівномірним.

В основі дисперсійного аналізу є такий принцип (факт з математичної статистики): якщо на випадкову величину діють взаємно незалежні фактори А, В, ..., то загальна дисперсія дорівнює сумі дисперсій, зумовлених дією окремо кожного з факторів.

Дисперсійний аналіз був створений спочатку для статистичної обробки агрономічних дослідів. В наш час його також використовують як в економічних експериментах, так і технічних, соціальних.

Сутність цього аналізу полягає в тому, що загальну дисперсію досліджуваної ознаки розділяють на окремі компоненти, які обумовлені впливом певних конкретних чинників. Істотність їх впливу на цю ознаку здійснюється

методом дисперсійного аналізу. Відповідно до дисперсійного аналізу будь-який його результат можна подати у вигляді суми певної кількості компонент. Так, наприклад, якщо досліджується вплив певного чинника на результат експерименту, то модель, що описує структуру останнього, можна подати так:

$$x_{ij} = \bar{X} + \alpha_j + \varepsilon_{ij}$$

де  $x_{ij}$  – значення ознаки  $X$ , одержане при  $i$ -му експерименті на  $j$ -му рівні фактора.

Під рівнем фактора розуміють певну його міру.

Наприклад, якщо фактором є добрива, які вносяться в землю з метою збільшення врожайності сільськогосподарської культури, то рівнем фактора в цьому разі є кількість добрива, що вноситься в ґрунт;

$\bar{X}$  – загальна середня величина ознаки  $X$ ;

$\alpha_j$  – ефект впливу фактора на значення ознаки  $X$  на  $j$ -му рівні;

$\varepsilon_{ij}$  – випадкова компонента, що впливає на значення ознаки  $X$  в  $i$ -му експерименті на  $j$ -му рівні.

При цьому  $M(\varepsilon_{ij}) = 0$  і  $\varepsilon_{ij}$ , як випадкові величини мають закон розподілу ймовірностей  $N(0; \sigma^2)$  і між собою незалежні ( $K_{ij} = 0$ ).

Складнішою моделлю аналізу є вивчення впливу на результати експерименту кількох факторів. Зокрема при аналізі впливу двох факторів структура моделі набуває такого вигляду:

$$x_{ijk} = \bar{X} + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk}$$

де  $x_{ijk}$  – значення ознаки  $X$  в  $i$ -му експерименті на  $j$ -му рівні впливу фактора  $A$  і на  $k$ -му рівні впливу фактора  $B$ ;  $\bar{X}$  – загальна середня величина ознаки  $X$ ;  $\alpha_i$  – ефект впливу фактора  $A$  на  $i$ -му рівні,  $\beta_j$  – ефект впливу фактора  $B$  на  $j$ -му рівні;  $\gamma_{ij}$  – ефект одночасного впливу факторів  $A$  і  $B$ ;  $\varepsilon_{ijk}$  – випадкова компонента.

У разі проведення дисперсійного аналізу досліджуваний масив даних, одержаних під час експерименту, поділяють на певні групи, які різняться дією на результати експерименту певних рівнів факторів.

Вважається, що досліджувана ознака має нормальний закон розподілу, а дисперсії в кожній окремій групі здобутих значень ознаки однакові. Ці припущення необхідно перевірити.

## 2. Однофакторний дисперсійний аналіз.

Нехай потрібно дослідити вплив на ознаку  $X$  певного одного фактора. Результати експерименту ділять на певне число груп, які відрізняються між собою ступенем дії фактора.

Для зручності в проведенні необхідних обчислень результати експерименту зводять в спеціальну таблицю.

Відповідно до моделі однофакторного дисперсійного аналізу необхідно визначити дні дисперсії, а саме: міжгрупову (дисперсію групових середніх), зумовлену впливом досліджуваного фактора на ознаку  $X$ . і внутрішню групову, зумовлену впливом інших випадкових факторів.

Загальна дисперсія розглядається як сума квадратів відхилень:

$$\sum_{i=1}^{n_j} \sum_{j=1}^p (x_{ij} - \bar{x})^2$$

Тоді поділ загальної дисперсії на компоненти здійснюється так:  
Для того щоб мати виправлені дисперсії, необхідно кожен зі здобутих сум поділити на число ступенів свободи.

Так, для загальної дисперсії

$$\frac{\sum_{i=1}^{n_j} \sum_{j=1}^p (x_{ij} - \bar{x})^2}{N}$$

виправлена дисперсія дорівнюватиме

$$S^2 = \frac{\sum_{i=1}^{n_j} \sum_{j=1}^p (x_{ij} - \bar{x})^2}{N - 1}$$

Виправлена дисперсія  $S_1^2$ , що характеризує розсіювання всередині групи, зумовлене впливом випадкових факторів, обчислюється за формулою:

$$S_1^2 = \frac{\sum_{i=1}^{n_j} \sum_{j=1}^p (x_{ij} - \bar{x}_j)^2}{N - p}$$

де  $N-p=k_1$  є числом ступенів свободи для  $S_1^2$ , оскільки при цьому використовується  $p$  співвідношень при обчисленні групових середніх  $\bar{x}_j$ ,  $j=1, p$

Виправлена дисперсія  $S_2^2$ , що характеризує розсіювання групових середніх  $\bar{x}_j$  відносно загальної середньої  $\bar{x}$ , яке викликано впливом фактора на результат експерименту ознаки  $X$ , обчислюється за формулою:

$$S_2^2 = \frac{\sum_{j=1}^p n_j (\bar{x}_j - \bar{x})^2}{p - 1}$$

де  $p-1=k_2$  – це число ступенів свободи для  $S_2^2$ , оскільки групові середні варіюють відносно однієї загальної середньої  $\bar{x}$ .

Завдання виявлення впливу фактора на наслідки експерименту полягає в порівнянні виправлених дисперсій  $S_1^2$ ,  $S_2^2$ . І справді, якщо досліджуваний фактор не впливає на значення ознаки  $X$ , то в цьому разі  $S_1^2$  і  $S_2^2$  можна розглядати як незалежні оцінки загальної дисперсії  $D$ . І навпаки, якщо відношення  $S_1^2$  і  $S_2^2$  істотне, то в цьому разі вибірки слід вважати здійсненими з різних сукупностей, тобто з сукупностей з різним рівнем впливу фактора.

Порівняння двох дисперсій ґрунтується на перевірці правильності нульової гіпотези:  $H_0 : D_1 = D_2$  – про рівність дисперсій двох вибірок.

За статистичний критерій вибирається випадкова величина

$$F = \frac{S_2^2}{S_1^2} \cdot \frac{p - 1}{N - p}$$

що має розподіл Фішера-Снедекора з  $k_1=N-p$ ,  $k_2=p-1$  ступенями свободи  
За значеннями  $\alpha$ ,  $k_1=N-p$ ,  $k_2=p-1$ , знаходимо критичну точку.

Якщо  $F^* \leq F_{kp}$ , то нульова гіпотеза про вплив фактора на результати досліджень відхиляється, а коли  $F^* > F_{kp}$ , то цим самим підтверджується вплив фактора на ознаку  $X$ .

### 3. Двофакторний дисперсійний аналіз.

Нехай необхідно визначити вплив двох факторів  $A$  і  $B$  на певну ознаку  $X$ . Для цього необхідно, щоб дослід здійснювався при фіксованих рівнях факторів  $A$  і  $B$ , а також їх одночасній дії на ознаку. При цьому дослід здійснюватимемо  $n$  разів для кожного з рівнів факторів  $A$  і  $B$ .

Позначимо через  $x_{ijk}$  конкретне значення ознаки  $X$ , якого вона набуває при  $i$ -му експерименті,  $j$ -му рівні фактора  $A$  і  $k$ -му рівні фактора  $B$ .

Результат експерименту зручно подати у вигляді таблиці, яка поділена на блоки, в кожному з яких враховується на певних рівнях факторів  $A$  і  $B$  їх вплив на конкретні значення ознаки  $X=x_{ijk}$ .

$$\bar{x}_{ij} = \frac{\sum_{k=1}^p x_{ijk}}{n}$$

Виходячи з даних  $\bar{x}_{ij}$  є середнім значенням ознаки  $X$  для кожного блока;

$$\bar{z}_j = \frac{\sum_{i=1}^n \sum_{k=1}^p x_{ijk}}{np}, \quad j = \overline{1, p} \quad \text{є середнім значенням ознаки } X \text{ за стовпцями;}$$

$$\bar{y}_i = \frac{\sum_{j=1}^q \sum_{k=1}^p x_{ijk}}{np}, \quad i = \overline{1, q} \quad \text{є середнім значенням ознаки } X \text{ за рядками;}$$

$$\bar{x} = \frac{\sum_{i=1}^n \sum_{j=1}^q \sum_{k=1}^p x_{ijk}}{npq} \quad \text{є загальною середньою ознакою } X;$$

$S_1^2 = \frac{np \sum (\bar{y}_i - \bar{x})^2}{p-1} = \frac{Q_1}{p-1}$  є виправленою дисперсією, яка зумовлена впливом фактора  $A$  на ознаку  $X$ ;

$S_2^2 = \frac{nq \sum (\bar{z}_j - \bar{x})^2}{q-1} = \frac{Q_2}{q-1}$  є виправленою дисперсією, яка зумовлена впливом фактора  $B$  на ознаку  $X$ ;

$S_3^2 = \frac{\sum \sum (x_{ij} - \bar{z}_j - \bar{y}_i + \bar{x})^2}{(p-1)(q-1)} = \frac{Q_3}{(p-1)(q-1)}$  є виправленою дисперсією, яка зумовлена одночасним впливом на ознаку  $X$  факторів  $A$  і  $B$ ;

$$S_4^2 = \frac{\sum \sum \sum (x_{ijk} - \bar{x}_{jk})^2}{N - pq} = \frac{Q_4}{N - pq}$$

є виправленою дисперсією, яка зумовлена впливом на ознаку  $X$  інших, не головних факторів.

Обчислюються спостережувані значення критерію

$$F_A^* = \frac{S_\sigma^2}{S_m^2}; \quad F_B^* = \frac{S_\sigma^2}{S_m^2}; \quad F_{AB}^* = \frac{S_\sigma^2}{S_m^2}.$$

При рівні значущості  $\alpha$  визначають критичні точки:

$$F_{kp}(\alpha; k_4; k_1), \quad F_{kp}(\alpha; k_3; k_1), \quad F_{kp}(\alpha; k_2; k_2).$$

Якщо:

- 1)  $F_A^* > F_{kp}(\alpha; k_4; k_1)$ , то нульова гіпотеза про відсутність впливу фактора  $A$  відхиляється;
- 2)  $F_B^* > F_{kp}(\alpha; k_3; k_1)$ , то нульова гіпотеза про відсутність впливу фактора  $B$  відхиляється;
- 3)  $F_{AB}^* > F_{kp}(\alpha; k_2; k_1)$ , то нульова гіпотеза про відсутність спільного впливу факторів  $A$   $B$  відхиляється.

Дисперсійний аналіз створений для статистичної обробки агрономічних дослідів, а також його використовують в економічних експериментах, технічних, соціальних. Сутність цього аналізу полягає в тому, що загальну дисперсію досліджуваної ознаки розділяють на окремі компоненти, які обумовлені впливом певних конкретних чинників.

Існують однофакторний та двофакторний дисперсійний аналізи. Однофакторним досліджують вплив на ознаку певного одного фактора. Двофакторний використовують коли необхідно визначити вплив двох факторів на певну ознаку. Результати досліджень подають у вигляді таблиці.

### Питання для контролю знань:

1. Визначення понять: «аналіз», «дисперсія», «дисперсійний аналіз», «статистика», «проба», «фактор», «чинник», «гіпотеза».
2. Сутність однофакторного дисперсійного аналізу.
3. Сутність двофакторного дисперсійного аналізу.

### Завдання для самостійної підготовки:

- I. Вивчити матеріал лекції.
- II. Підготувати доповідь за темою:
  1. Теоретичні основи і принципова схема дисперсійного аналізу.
  2. Дисперсійний аналіз при групуванні даних за однією ознакою.
  3. Застосування дисперсійного аналізу для оцінки вірогідності різниці двох середніх.
  4. Дисперсійний аналіз при групуванні даних за двома ознаками.
- III. Пройти тестування за темою лекції: тест 8.1. за посиланням <https://forms.gle/MUGF8XTdWFcKz9DU8>

## ЛЕКЦІЯ № 9

### КОРЕЛЯЦІЙНИЙ АНАЛІЗ. РЕГРЕСІЙНИЙ АНАЛІЗ.

#### Література:

1. Карагодова О.О., Рожок В.Д. Дослідження операцій: навч. посіб. К.: Центр учбової літератури, 2007. 256 с.
2. Бахрушин В.Є. Методи аналізу даних: навчальний посібник для студентів. Запоріжжя: КПУ, 2011. 268 с.
3. Перегуда О.В., Капустян О.А., Курилко О.Б. Статистична обробка даних: навч. посіб. Електронне видання, 2022. 103 с.
4. Рогальський Ф. Б., Курилович Я. Є., Цокурєнка А. А. Математичні методи аналізу економічних систем. Книга 1. Київ: Наукова думка, 2001. 435 с.
5. Рогальський Ф. Б., Цокурєнка А. А. Математичні методи аналізу економічних систем. Книга 2. Київ: Наукова думка, 2001. 423 с.
6. Крайчук О.В., Московська Г.К., Соколенко О.І. Теорія ймовірностей і математична статистика. Рівне, 2004.

#### Мета лекції:

1. Навчальна – вивчити основи застосування кореляційного аналізу; вивчити основи застосування регресійного аналізу.
  2. Розвиваюча – ознайомитися з напрямками застосування кореляційного аналізу.
- Виховна – виховувати професійні риси, почуття відповідальності; виховувати зацікавленість дисципліною, прагнення отримати нові знання самостійно.

#### План:

1. Кореляційний аналіз. Загальні відомості.
2. Регресійний аналіз. Загальні відомості.

#### Теоретична частина:

##### **1. Кореляційний аналіз. Загальні відомості.**

Кореляція дослівно з латинської «correlation» – відношення, тобто це означає співвідношення, відповідність речей, понять. Кореляційним зв'язком називається такий зв'язок між ознаками суспільно-економічних явищ, за якого на величину результативної ознаки крім факторної впливають багато інших ознак, які можуть діяти в різних напрямках одночасно чи послідовно. Цей зв'язок характеризується тим, що між факторною і результативною ознаками немає повної відповідності, а лише є певне співвідношення. Особливістю кореляційного зв'язку є те, що кожному значенню факторної ознаки відповідає не одне, а ціла низка значень результативної ознаки. Кореляційний зв'язок можна виявити тільки у вигляді загальної тенденції при масовому порівнянні факторів.

**Кореляційний аналіз** (кореляційний метод) – метод дослідження взаємозалежності ознак у генеральній сукупності, які є випадковими величинами з нормальним характером розподілу.

Основними вимогами до застосування кореляційного аналізу є достатня кількість спостережень, сукупності факторних і результативних показників, а також їх кількісний вимір і відображення в інформаційних джерелах.

Застосування кореляційного аналізу тісно пов'язане з регресійним аналізом, тому його часто називають кореляційно-регресійним. Головними завданнями кореляційного аналізу є:

- визначення форми зв'язку;
- вимірювання щільності (сили) зв'язку;
- виявлення впливу факторів на результативну ознаку.

Здійснення кореляційного аналізу передбачає такі послідовні етапи:

1) встановлення причинно-наслідкових зв'язків між досліджуваними ознаками (виявлення факторів та вибір серед них тих, які найбільше впливають на результативний показник);

2) формування кореляційно-регресійної моделі (інформаційне забезпечення аналізу, вибір типу і форми зв'язку, складання моделі);

3) визначення кореляційних характеристик (показників зв'язку);

4) статистична оцінка параметрів зв'язку (економічна інтерпретація, оцінка значимості коефіцієнтів кореляції (наскільки відібрані фактори пояснюють варіацію результативного показника) та використання їх для вирішення практичних завдань, наприклад прийняття рішень, прогнозування, планування, нормування тощо.

Отже, на початковому етапі аналізу виявляються зв'язки між результативною і факторними ознаками. Ці зв'язки можуть бути різними залежно від характеру залежності, напряму дії та аналітичного виразу.



Рис. Схема кореляційно-регресійного аналізу



На другому етапі оцінюється вихідна інформація для дослідження з використанням різних статистичних критеріїв (середнє квадратичне відхилення, коефіцієнт варіації тощо), а потім формується модель стохастичного зв'язку.

Для здійснення кореляційного аналізу скористаємося прикладом, який використали раніше для демонстрації регресійного аналізу. Спочатку оцінимо типовість та однорідність даних спостереження, які визначаються відносним їх розподілом навколо середнього рівня, за допомогою таких критеріїв, як середньоквадратичне відхилення ( $\sigma$ ) та коефіцієнт варіації ( $V$ ).



Рис. Види зв'язків суспільних явищ

Формування кореляційної моделі передбачає визначення чи це буде проста (парна) кореляція (результативна ознака з одним фактором), чи множинна (результативна ознака і декілька факторів). У свою чергу за характером зв'язку кореляційні моделі можуть бути лінійними (прямолінійними, з оберненою лінійною залежністю) чи нелінійними (криволінійними).

У лінійних моделях тіснота зв'язку між досліджуваними показниками вимірюється за допомогою лінійного коефіцієнта кореляції (Пірсона)  $r$  за формулою

$$r = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sigma_x \cdot \sigma_y}$$

Крім відображення щільності зв'язку, коефіцієнт кореляції відіграє ще одну важливу роль – через коефіцієнт детермінації ( $D$ ) він характеризує розмір впливу факторів на результативну ознаку:

$$D = r^2$$

Залежності між декількома факторними ознаками складніші  $q$  виражаються рівнянням множинної лінійної кореляції

$$Y_x = a_0 + a_1x_1 + a_2x_2 + \dots + a_nx_n.$$

Коефіцієнт множинної кореляції характеризує ступінь тісноти зв'язку між залежною змінною та кількома незалежними змінними. Він не може бути

меншим, ніж абсолютна величина будь-якого коефіцієнта парної чи множинної кореляції і набуває значення від 0 до 1.

Загальний математичний вираз коефіцієнта множинної кореляції має вигляд

$$R = \sqrt{\frac{\sigma_r^2}{\sigma_y^2}},$$

де  $\sigma_r^2$  – теоретична (відтворена) варіація – дисперсія значень величини у, розрахованих шляхом підстановки значень факторів у знайдене рівняння регресії

$$Y = f(x_1, x_2, x_3, \dots, x_n),$$

де  $\sigma_y^2$  – загальна варіація – дисперсія фактичних значень величини у.

Рівняння множинної регресії і показники множинної кореляції розв'язують і визначають за допомогою спеціальних комп'ютерних програм.

Таким чином, кореляційний аналіз має велике значення в вивченні суспільних явищ і процесів. Зокрема, він допомагає вирішити такі завдання:

- встановлення характеру і тісноти зв'язку між досліджуваними явищами;
- кількісний вимір ступеня впливу окремих факторів та їх сукупності на рівень явища, яке вивчається;
- розрахунок кількісних змін аналізованого явища при прогнозуванні показників та об'єктивна оцінка господарської діяльності підприємства.

Велике значення відводиться кореляційному аналізу в дослідженні кореляційних зв'язків на виробництві, зокрема між рівнем продуктивності праці та озброєністю її основними засобами, між урожайністю і кількістю внесених добрив, між собівартістю і випуском продукції та ін.

Завдяки кореляційному аналізу є можливість глибше дослідити взаємозв'язки економічних явищ і процесів, виявити вплив факторів на результати господарської діяльності, виявити і підрахувати резерви підвищення ефективності виробництва. Все це позитивно позначається на здійсненні управлінської, маркетингової та інших видів діяльності, прийнятті економічно обґрунтованих господарських рішень.

## 2. Регресійний аналіз. Загальні відомості.

Під терміном "регресія" розуміють рух назад, повернення до попереднього стану. Названий термін ввів у кінці XIX ст. Френсіс Галтон. 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".

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.

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

Основним завданням регресійного аналізу є визначення впливу факторів на результативний показник (в абсолютних показниках). Передусім для цього необхідно підібрати та обґрунтувати рівняння зв'язку, що відповідає характеру аналітичної стохастичної залежності між досліджуваними ознаками. Рівняння регресії показує як в середньому змінюється результативна ознака ( $Y_x$ ) під впливом зміни факторних ознак ( $x_i$ ).

У загальному вигляді рівняння регресії можна представити так:

$$Y_x = f(x_1, x_2, \dots, x_n),$$

де  $Y_x$  – залежна змінна величина;

$x$  – незалежні змінні величини (фактори).

Залежно від кількості змінних величин виділяють різні види регресійного аналізу. Якщо змінна величина завжди одна, то змінних може бути як одна, так і декілька. Виходячи з цього, виділяють два види регресійного аналізу: парний (простий) регресійний аналіз і регресійний аналіз на основі множинної регресії, або багатофакторний.

Парний регресійний аналіз – вид регресійного аналізу, що включає у себе розгляд однієї незалежної змінної величини, а багатофакторний – відповідно дві величини і більше.

Зважаючи на характер зв'язку, в регресійному аналізі можуть використовуватися лінійні та нелінійні функції. Для визначення характеру залежності та, відповідно, побудови рівняння регресії доцільно застосувати графічний метод, порівняння рівнобіжних рядів вихідних даних, табличний метод.

Так, графічний метод дає найбільш наочну картину розміщення крапок на графіку, завдяки чому можна виявити напрям і вид залежності між досліджуваними показниками: прямолінійна чи криволінійна.

За допомогою порівняння рівнобіжних рядів ознак можна спостерігати за рівномірністю їх взаємних змін. Якщо зміна факторної ознаки ( $x$ ) призводить до відносно рівномірної зміни результативної ( $Y_x$ ), тоді використовується лінійна функція (наприклад, залежність між урожайністю культур і кількістю внесених добрив).

Найпростішим рівнянням парної регресії, що описує лінійну залежність між факторною і результативною ознаками, є рівняння прямої, яке має такий вигляд:

$$Y_x = a_0 + a_1 x,$$

де  $Y_x$  – залежна змінна, яка оцінюється або прогнозується (результативна ознака);

$a_0$  – вільний член рівняння;

$a_1$  – коефіцієнт регресії;

$x$  – незалежна змінна (факторна ознака), яка використовується для визначення залежної змінної.

Параметри рівняння обчислюються на основі системи нормальних рівнянь методом найменших квадратів:

$$\begin{cases} \sum y = na_0 + a_1 \sum x \\ \sum xy = a_0 \sum x + a_1 \sum x^2. \end{cases}$$

Звідси

$$a_1 = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - \sum x \sum x}$$

$$a_0 = \frac{\sum x^2 \sum y - \sum x \sum xy}{n \sum x^2 - \sum x \sum x}, \text{ або } a_0 = \bar{y} - a_1 \bar{x}$$

Основне змістовне навантаження в рівнянні регресії несе коефіцієнт регресії. Найчастіше застосовуються лінійні рівняння або приведені до лінійного вигляду. Коефіцієнт регресії – це кутовий коефіцієнт у прямолінійному рівнянні кореляційного зв'язку. У лінійній функції рівняння регресії він показує на скільки одиниць в середньому зміниться результативна ознака (y) при зміні факторної ознаки (x) на одиницю свого натурального виміру. Тобто, коефіцієнт регресії – це варіація y, яка припадає на одиницю варіації x. Коефіцієнт регресії має одиницю виміру результативної ознаки. За наявності прямого зв'язку коефіцієнт регресії є додатною величиною, а за зворотного зв'язку – від'ємною.

Параметр  $a_0$  – вільний член рівняння регресії, тобто це значення y при  $x=0$ . Цей показник має тільки розрахункове значення у випадках, коли x не має нульових значень.

У разі, коли зі зміною факторної ознаки результативна змінюється нерівномірно, використовуються нелінійні функції. Так, якщо зміна факторного показника сприяла прискореній динаміці результативного показника (наприклад, вплив обсягу грошової маси на рівень інфляції), доцільно використати степеневу функцію:

$$Y_x = ax^b$$

У випадку, коли під впливом факторної ознаки результативна змінюється нерівномірно, причому з уповільненням, використовується рівняння гіперболи:

$$Y_x = a + \frac{b}{x}$$

Прикладом такої залежності є залежність рівня продуктивності праці робітників від рівня їх заробітної плати.

Якщо зміна факторної ознаки супроводжується нерівномірною варіацією факторної ознаки із зміною напрямку зв'язку, нелінійна регресія описується рівнянням параболи:

$$Y_x = a + bx + cx^2$$

Так, за допомогою функції параболи можна виразити залежність урожайності культур від кількості опадів.

Аналіз на основі множинної регресії (анг. multiple regression analysis) – вид регресійного аналізу, який ґрунтується на використанні в рівнянні регресії

більше, ніж однієї незалежної змінної. Так, його застосовують при прогнозуванні попиту. Причому спочатку ідентифікуються фактори, що визначають попит, потім встановлюються наявні між ними взаємозв'язки та прогнозуються ймовірні майбутні їх значення. На основі отриманих даних виводиться прогнозне значення попиту.

Багатофакторне рівняння множинної регресії при лінійній залежності має такий вигляд:

$$Y_{t-x_t} = a_0 + a_1x_1 + a_2x_2 + \dots + a_nx_n,$$

де  $a_0$  – вільний член рівняння;

$a_1, a_2 \dots a_n$  – коефіцієнти регресії;

$x_1, x_2 \dots x_n$  – незалежні змінні (факторні ознаки);

$n$  – кількість незалежних змінних.

Визначення параметрів множинної регресії вимагає трудомістких розрахунків із застосуванням комп'ютерних інформаційних систем. Однак одержані результати будуть достовірними і можуть широко використовуватися в економічній та управлінській діяльності насамперед для складання довгострокових прогнозів. Відомо, що однофакторна модель придатна для короткострокових прогнозів (на 2–3 роки).

Метод регресійного аналізу вважається найдосконалішим з усіх використовуваних нині нормативно-параметричних методів. Він широко застосовується для аналізу та встановлення рівня і співвідношень вартості продукції, яка характеризується наявністю одного або декількох техніко-економічних параметрів, що характеризують головні споживчі якості. Регресивний аналіз надає можливість знайти емпіричну форму залежності ціни від техніко-економічних параметрів товарів і виробів. При цьому він виступає в ролі цільової функції параметрів.

Метод регресійного аналізу особливо ефективний за умови здійснення розрахунків за допомогою сучасних інформаційних технологій і систем.

Приклад. За даними 25 заводів дослідити залежність між рівнями озброєності праці основними засобами та її продуктивністю, використовуючи регресійний аналіз (табл.1).

Таблиця 1. Озброєність праці основними засобами та рівень її продуктивності в групі заводів України, тис. грн

№ з/п	Озброєність праці	Вироблено продукції на 1 працівника	Розрахункові величини			
			$x^2$	$y^2$	$x \cdot y$	$Y_x$
n	x	y				
1	2	3	4	5	6	7

1	9,6	12,8	92,2	163,5	122,7	11,9
2	11,0	12,6	121,0	157,9	138,2	13,3
3	11,2	15,5	125,4	241,1	173,9	13,4
4	11,4	15,7	130,0	246,2	178,9	13,6
5	11,5	9,3	132,3	87,2	107,4	13,7
6	12,1	16,4	146,4	267,5	197,9	14,3
7	12,2	17,6	148,8	308,0	214,1	14,4
8	12,3	17,6	151,3	310,6	216,8	14,5
9	12,6	14,6	158,8	213,3	184,0	14,8
10	12,8	13,7	163,8	188,2	175,6	15,0
11	13,2	14,8	174,2	219,0	195,4	15,3
12	13,2	14,8	174,2	219,0	195,4	15,3
13	13,3	16,1	176,9	258,1	213,7	15,4
14	13,3	16,1	176,9	258,1	213,7	15,4
15	14,6	23,2	213,2	539,4	339,1	16,7
16	16,0	21,5	256,0	461,1	343,6	18,0
17	16,4	25,0	269,0	625,0	410,0	18,4
18	17,3	6,9	299,3	48,0	119,9	19,2
19	17,3	6,9	299,3	48,0	119,9	19,2
20	17,6	14,9	309,8	221,8	262,1	19,5
21	17,6	14,9	309,8	221,8	262,1	19,5

22	17,7	22,2	313,3	490,7	392,1	19,6
23	17,9	18,3	320,4	335,4	327,8	19,8
24	22,8	30,1	519,8	905,5	686,1	24,5
25	24,9	31,1	620,0	970,2	775,6	26,5
Сума	369,8	422,5	5802,0	8004,8	6565,9	421,3
В середньому	14,8	16,9	232,1	320,2	262,6	16,9

Паралельне зіставлення рядів значень рівнів озброєності праці основними засобами та її продуктивності, а також крапковий графік "кореляційного поля" (рис. 1) свідчать про наявність і напрям зв'язку (прямий) між наведеними показниками. Причому зміна озброєності праці (факторної ознаки  $x$ ) приводить до відносно рівномірної зміни продуктивності праці (результативної ознаки  $y$ ), що видно із графіка.

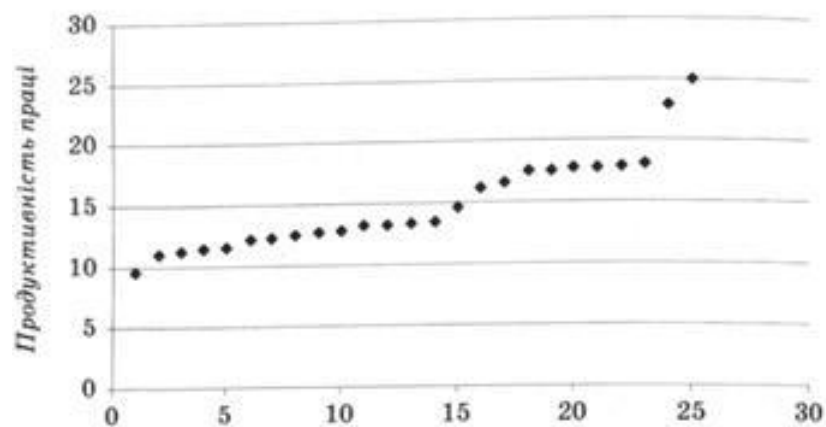


Рис. 1. Крапковий графік рівнів озброєності праці основними засобами та її продуктивності по 25 заводах України, тис. грн..

Отже, можна зробити висновок про наявність між зазначеними ознаками парного прямолінійного зв'язку, який виражається найпростішим рівнянням регресії лінійної функції

$$Y_x = a_0 + a_1 x$$

де  $a_0$  – вільний член рівняння регресії;

$a_1$  – коефіцієнт регресії.

Для зручності розрахунків скористаємося наведеними нижче формулами, в які підставляємо попередньо обчислені необхідні розрахункові дані (табл. 12.1):



$$a_0 = \frac{\sum x^2 \sum y - \sum x \sum xy}{n \sum x^2 - \sum x \sum x} = \frac{5802,0 \cdot 422,5 - 369,8 \cdot 6565,9}{25 \cdot 5802,0 - 369,8 \cdot 369,8} =$$

$$= \frac{23275,18}{8297,96} = 2,8;$$

$$a_1 = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - \sum x \sum x} = \frac{25 \cdot 6565,9 - 369,8 \cdot 422,5}{25 \cdot 5802,0 - 369,8 \cdot 369,8} =$$

$$= \frac{7907}{8297,96} = 0,95.$$

Рівняння залежності матиме такий вигляд:

$$Y_x = 2,8 + 0,95x$$

Перевірка:  $Y_x = 2,8 + 0,95 \cdot 14,8 = 16,9$ , що відповідає середньому рівню продуктивності праці (табл. 2).

Зобразимо графічно залежність між рівнями озброєності праці основними засобами та її продуктивністю, використавши рівняння регресії (рис. 2).

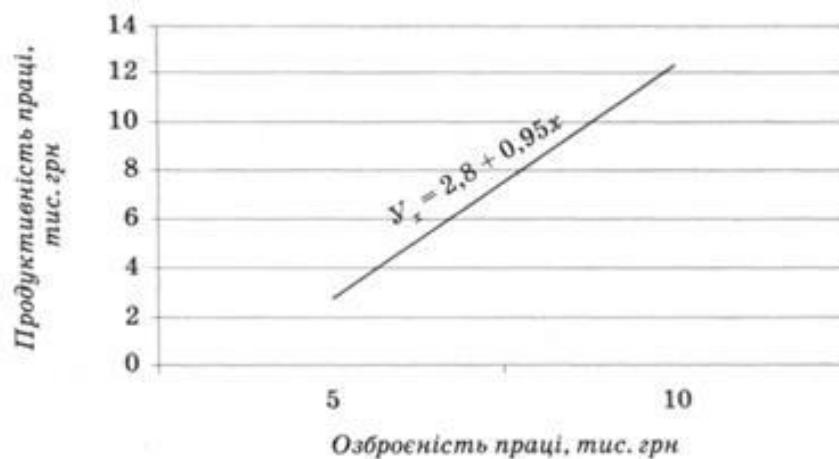


Рис. 2. Залежність між рівнями озброєності та продуктивності праці.

Якщо підставити в отримане рівняння регресії відповідні значення фактора (x) по заводах, то одержимо вирівняні значення продуктивності праці залежно від озброєності її основними засобами. Результати розрахунків наведені в табл. 12.3 (остання графа).

З рівняння регресії видно, що підвищення озброєності праці основними засобами на одну тисячу гривень забезпечує підвищення рівня продуктивності праці на 0,95 тис. грн.

Якщо б на заводі під № 1 озброєність праці була вище середньої, наприклад 17,9 тис. грн, то виробіток продукції на одного працівника становив би 19,8 тис. грн:

$$Y_x = 2,8 + 0,95x = 2,8 + 0,95 \cdot 17,9 = 19,8$$

Величина 17,9, яка взята на заводі, де озброєність праці перевищує середнє її значення – 14,8 тис. грн, називається середньою прогресивною.



Таким чином, регресійний аналіз може бути використаний ще й для визначення резервів шляхом застосування середньопрогресивного значення факторіальної ознаки. У наведеному прикладі цей резерв мають 19 заводів, що мають виробіток продукції на одного працівника менший за 19,8 тис. грн.

Також слід зазначити, що регресійний аналіз доцільно застосовувати для обґрунтування проектних, прогнозних чи очікуваних показників. Для цього необхідно підставити в одержане рівняння регресії проектне значення фактора. Зокрема, якщо б заводу № 7 поставити за мету доведення рівня озброєності праці основними засобами до 20 тис. грн, то можна зробити прогноз продуктивності праці, яка становитиме 21,8 тис. грн:

$$Y_x = 2,8 + 0,95x = 2,8 + 0,95 \cdot 20,0 = 21,8$$

Безумовно, що при застосуванні регресійного аналізу дотримуються деяких умовностей. Так, попередньо обумовлюється, що дія інших факторів, крім взятого за факторну ознаку, залишиться незмінною, а в дослідженні взято тільки один фактор.

Проте останню умовність можна усунути за допомогою застосування множинної регресії та кореляції, за яких підбирається значна кількість факторів.

#### **Питання для контролю знань:**

1. Визначення понять: «аналіз», «кореляція», «кореляційний аналіз», «коефіцієнт детермінації», «кореляційне поле».
2. Мета кореляційного аналізу.
3. Етапи кореляційного аналізу.
4. Коефіцієнт кореляції Брауе-Пірсона.
5. t-критерій Стюдента.
6. Ранговий коефіцієнт кореляції Спірмена.
7. Позитивна кореляція.
8. Негативна кореляція.
9. Слабка/нульова кореляція.
10. Визначення понять: «аналіз», «регресія», «регресійний аналіз», «коефіцієнт регресії».
11. Основне завданням регресійного аналізу.
12. Що показує рівняння регресії. Основний вид.
13. Види регресійного аналізу.
14. Парний регресійний аналіз.
15. Лінійні та нелінійні функції.
16. Переваги графічного методу в регресійному аналізі.
17. Вид рівняння прямої.
18. Рівняння параболі.
19. Сутність аналізу на основі множинної регресії.
20. Багатофакторне рівняння множинної регресії.

#### **Завдання для самостійної підготовки:**

- I. Вивчити матеріал лекції.
- II. Підготувати доповідь за темою:
  1. Кореляційний аналіз, мета і завдання.
  2. Поняття кореляції.

3. Часткова та напівчасткова кореляції.
  4. Властивості коефіцієнта кореляції.
  5. Кореляційне поле.
  6. Перевірка гіпотези про значущість коефіцієнта кореляції.
  7. Автокореляція.
  8. Кореляційна матриця.
  9. Задачі регресійного аналізу.
  10. Регресійні моделі.
  11. Базові (фундаментальні) положення класичного лінійного регресійного аналізу.
  12. Класичні оцінки параметрів регресії методом найменших квадратів і їх властивості.
  13. Рекурентний алгоритм методу найменших квадратів.
  14. Статистичний аналіз якості регресійної моделі.
  15. Регресійний аналіз найпростіших поліноміальних моделей.
  16. Регресійний аналіз при неоднорідних і корельованих збуреннях.
  17. Регресійний аналіз в умовах мультиколінеарності.
  18. Вибір найкращої структури регресійної моделі.
  19. Регресійний аналіз в умовах похибок в регресорах.
- III. Пройти тестування за темою лекції: тест 9.1. за посиланням <https://forms.gle/fjY5oZiTdRUAQYhB6>

*Навчальне видання*

**КОНСПЕКТ ЛЕКЦІЙ  
З НАВЧАЛЬНОЇ ДИСЦИПЛІНИ  
«МЕТОДИ ОБРОБКИ КІЛЬКІСНИХ ПОКАЗНИКІВ ПОЖЕЖНОЇ  
НЕБЕЗПЕКИ»**

**для здобувачів третього (доктор філософії) рівня вищої освіти  
галузь знань 26 «Цивільна безпека», спеціальність 261 «Пожежна безпека»,  
3 курс, 5 семестр**