

**Ю.Ю. Гусєва, О.В. Гребенікова**

**ТЕОРЕТИЧНІ ТА ПРАКТИЧНІ АСПЕКТИ  
ФІНАНСОВО-ЕКОНОМІЧНИХ ДОСЛІДЖЕНЬ**

**2011**

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ, МОЛОДІ ТА СПОРТУ УКРАЇНИ  
Національний аерокосмічний університет ім. М.Є. Жуковського  
„Харківський авіаційний інститут”**

**Ю.Ю. Гусєва, О.В. Гребенікова**

**ТЕОРЕТИЧНІ ТА ПРАКТИЧНІ АСПЕКТИ  
ФІНАНСОВО-ЕКОНОМІЧНИХ ДОСЛІДЖЕНЬ**

**Навчальний посібник**

**Харків „ХАІ” 2011**

УДК 336:001.89 (075.8)  
Г96

Рецензенти: д-р екон. наук, проф. О.М. Ястремська,  
д-р екон. наук, проф. П.Т. Бубенко

**Гусєва, Ю.Ю.**

Г96 Теоретичні та практичні аспекти фінансово-економічних досліджень  
[Текст]: навч. посіб. / Ю.Ю. Гусєва, О.В. Гребенікова. – Х.: Нац.  
аерокосм. ун-т ім. М.Є. Жуковського „Харк. авіац. ін-т”, 2011. – 100 с.

Розглянуто основні принципи й методи наукового дослідження, подано рекомендації щодо підготовки наукових статей, тез, доповідей. Більш докладно висвітлено питання статистичних і експертних методів в економіці та фінансах.

Аналітичні можливості та межі застосування конкретних моделей проілюстровано на фінансово-економічних прикладах. Розрахунки виконано за технологіями статистичного аналізу й оброблення даних, реалізованими в системі Statistica.

Для студентів, аспірантів і викладачів, діяльність яких пов'язана з обробленням та аналізом статистичної інформації.

Іл. 33. Табл. 33. Бібліогр.: 18 назв

**УДК 336:001.89 (075.8)**

© Гусєва Ю.Ю., Гребенікова О.В., 2011  
© Національний аерокосмічний  
університет ім. М.Є. Жуковського  
„Харківський авіаційний інститут”, 2011

## ЗМІСТ

ПЕРЕДМОВА .....	4
1 ЗАГАЛЬНА МЕТОДИКА НАУКОВОЇ ДІЯЛЬНОСТІ .....	5
1.1 Наукове вивчення як основна форма наукової роботи.....	5
1.2 Методи емпіричного дослідження .....	8
1.3 Методи теоретичних досліджень.....	9
1.4 Методи, що застосовуються на емпіричному й теоретичному рівнях дослідження.....	10
1.5 Загальна схема наукового дослідження.....	11
1.6 Наукова публікація: поняття, функції, основні види .....	12
1.7 Поняття про наукову інформацію та її роль у проведенні наукових досліджень .....	21
1.8 Організація наукової діяльності в Україні .....	23
Запитання для самоперевірки .....	25
2 МЕТОДИ АНАЛІЗУ ЕКСПЕРТНОЇ ІНФОРМАЦІЇ.....	27
2.1 Сучасна теорія вимірювань і експертне оцінювання.....	27
2.2 Методи проведення експертиз .....	29
2.3 Підбір експертів.....	31
2.4 Методи експертного оцінювання .....	33
Запитання для самоперевірки .....	36
3 СТАТИСТИЧНІ МЕТОДИ ДОСЛІДЖЕННЯ. ОСНОВИ РОБОТИ 3 СИСТЕМОЮ STATISTICA.....	37
3.1 Показники описової статистики .....	37
3.2 Формування інформаційної бази моделі .....	42
3.3 Описові статистики. Розвідувальний аналіз даних у Statistica.....	44
3.4 Однорідність і типологія .....	47
Запитання для самоперевірки .....	50
4 МЕТОДИ БАГАТОВИМІРНОЇ КЛАСИФІКАЦІЇ .....	52
4.1 Кластерні процедури класифікації .....	53
4.2 Дискримінантний аналіз і логістична регресія.....	59
4.3 Факторний аналіз .....	64
4.4 Визначення репрезентанта групи.....	73
Запитання для самоперевірки .....	76
5 ЛІНІЙНИЙ РЕГРЕСІЙНИЙ АНАЛІЗ.....	78
5.1 Лінійна регресія: побудова моделі й оцінювання якості рівняння регресії.....	78
5.2 Оцінювання якості регресійної моделі .....	86
5.3 Перевірка умов, виконання яких передбачалося під час оцінювання рівняння регресії. Статистика Дарбіна – Уотсона.....	88
Запитання для самоперевірки .....	91
БІБЛІОГРАФІЧНИЙ СПИСОК.....	93
ПРЕДМЕТНИЙ ПОКАЖЧИК.....	95
ДОДАТКИ.....	96

## ПЕРЕДМОВА

У сучасних умовах розвитку науково-технічного прогресу, інтенсивного збільшення обсягу наукової та науково-технічної інформації, швидкої змінюваності й оновлення знань особливого значення набуває підготовка у вищій школі висококваліфікованих фахівців, здатних до самостійної, творчої роботи, впровадження у виробничий процес новітніх технологій.

З цією метою до навчального плану включено дисципліну «Основи наукових досліджень», яка розкриває перед студентами зміст наукового дослідження, дає уявлення про методи й методики проведення наукового дослідження, формує потребу в здобутті нових знань та інтерес до науки. Водночас ця дисципліна розкриває великі перспективи творчих аспектів вибраної спеціальності.

Статистичний аналіз даних стає невід'ємним атрибутом системи управління на всіх її рівнях – від невеликої фірми до національної економіки в цілому. Статистичні моделі використовують для діагностики стану об'єктів управління, під час вивчення причинно-наслідкового механізму формування варіації та динаміки соціально-економічних явищ і процесів, у моніторингу економічної кон'юнктури, під час прогнозування та прийняття оптимальних управлінських рішень.

Важливим етапом професійної підготовки економіста є оволодіння багатим арсеналом методів статистичного оброблення даних з використанням комп'ютерних технологій. Відповідно до програми курсу в навчальному посібнику розглядаються моделі багатовимірного оцінювання й моделі класифікацій, експертне оцінювання, моделі лінійної регресії.

Логічна структура викладення матеріалу ілюструється на конкретних прикладах соціально-економічного змісту. Для кожного типу моделей розглядаються принципи формування інформаційної бази, вибору процедур аналізу, інтерпретації результатів.

Акцентуючи увагу студентів на параметрах моделей, таблиці з результатами аналізу й графіки наводяться у стандартному вигляді англійською мовою. Специфікація включених до моделі ознак і змістовна інтерпретація параметрів моделі розкриваються в коментарях до таблиць і графіків.

# 1 ЗАГАЛЬНА МЕТОДИКА НАУКОВОЇ ДІЯЛЬНОСТІ

## 1.1 Наукове вивчення як основна форма наукової роботи

*Наука* – це сфера людської діяльності, спрямована на отримання нових знань про природу, суспільство й мислення. Зазвичай виділяють дві її складові частини:

- установлені наукові знання, які формують уявлення про навколишню дійсність і створюють у сукупності картину світу;
- творча діяльність людей (наукові дослідження), спрямована на отримання нового знання.

Наука як система наукових знань має такі специфічні ознаки:

- 1) систематизованість накопичених наукових знань;
- 2) можливість перевірки фактів (те, що не можна перевірити, є не накопиченим науковим знанням, а гіпотезою);
- 3) відтворюваність явищ (якщо факт не було свого часу відтворено, то це означає, що отримана інформація є неправильною або необхідно ще раз урахувати всі умови й перевірити висловлені твердження новими дослідниками);
- 4) довговічність тієї або іншої системи знань (дає можливість приймати ефективні рішення, застосовувати отриману систему знань в інших галузях).

Формою здійснення й розвитку науки є наукове дослідження, тобто вивчення за допомогою наукових методів явищ і процесів, аналіз впливу на них різних чинників, а також вивчення взаємодії явищ з метою отримання переконливо доведених та корисних для науки й практики рішень з максимальним ефектом.

Мета наукового дослідження – визначення конкретного об'єкта й всебічне достовірне вивчення його структури, характеристик, зв'язків на основі розроблених у науці принципів і методів пізнання, а також отримання корисних для діяльності людини результатів, впровадження їх у виробництво з подальшим ефектом.

Будь-яке наукове дослідження – від творчого задуму до кінцевого оформлення наукової праці – здійснюється індивідуально. Проте можна визначити й деякі загальні методологічні підходи до його проведення, які прийнято називати вивченням у науковому розумінні.

Накопичення наукових фактів під час дослідження – завжди творчий процес, в основі якого лежить задум ученого, його *ідея*. У філософському розумінні ідея являє собою продукт людської думки, форму відображення дійсності. Ідея відрізняється від інших форм мислення й наукового знання тим, що в ній не лише відображено об'єкт вивчення, але й міститься усвідомлення мети, перспективи пізнання й практичного перетворення дійсності.

Ідеї народжуються із практики, спостереження навколишнього світу і потреб життя (рисунок 1.1). У їхній основі лежать реальні факти й події. Життя висуває конкретні завдання, але не завжди відразу знаходяться продуктивні ідеї для їх вирішення. Тоді дослідник має пропонувати новий, зовсім незвичний аспект розгляду завдання, яке довго не могли вирішити при звичайних підходах до справи.

Нова ідея – не просто зміна уявлення про об'єкт дослідження шляхом суворого обґрунтування, це якісний стрибок думки за межі сприйнятих почуттями даних і перевірених рішень. Нові ідеї можуть виникати під впливом парадоксальних ситуацій, коли виявляється незвичний, несподіваний результат, який надто розбіжний із загальноприйнятими положеннями науки – парадигмами. При цьому отримання нових знань відбувається за такою схемою: парадигма – парадокс – нова парадигма. Можна стверджувати, що розвиток науки – це зміна відмінних парадигм, методів, стереотипів мислення. Перехід від однієї парадигми до іншої не піддається логічному опису, бо кожна з них відкидає попередню і несе принципово новий результат дослідження, котрий не можна логічно вивести із відомих теорій. Особливу роль тут відіграють інтуїтивні механізми наукового пошуку, які не ґрунтуються на формальній логіці.

Розвиток ідеї до стадії вирішення завдання зазвичай здійснюється як плановий процес наукового дослідження. Хоча в науці й відомі випадкові відкриття, проте тільки планове, добре обладнане сучасними засобами наукове дослідження дає змогу розкрити й глибоко пізнати об'єктивні закономірності в природі. Згодом іде процес продовження цільового оброблення первинного задуму, уточнення, зміни, доповнення й розвитку накресленої схеми дослідження з використанням різних методів пізнання.

*Метод* – це сукупність прийомів чи операцій практичного або теоретичного вивчення дійсності, підпорядкованих вирішенню конкретного завдання. Фактично різниця між методом і теорією має функціональний характер: формуючись як теоретичний результат попереднього дослідження, метод є вихідним пунктом та умовою майбутніх досліджень.

У кожному науковому дослідженні можна виділити два рівні: емпіричний, на якому відбувається процес накопичення фактів; теоретичний – досягнення синтезу знань (у формі наукової теорії).

Згідно із наведеними рівнями загальні методи пізнання можна поділити на три групи, грані між якими визначено приблизно:

- методи емпіричного дослідження;
- методи теоретичного дослідження;
- методи, що використовуються на емпіричному й теоретичному рівнях.



Рисунок 1.1 – Основні структурні елементи теорії пізнання



## 1.2 Методи емпіричного дослідження

*Спостереження* – це систематичне цілеспрямоване вивчення об'єкта. Аби бути плідним, спостереження мусить відповідати:

- завчасному задуму (спостереження здійснюється для певного, чітко поставленого завдання);
- плановірності (виконується за планом, складеним відповідно до завдання спостереження);
- цілеспрямованості (спостерігаються лише певні сторони явища, які становлять інтерес під час дослідження);
- активності (спостерігач активно шукає потрібні об'єкти, риси явища);
- систематичності (спостереження ведеться безперервно або за певною системою).

*Порівняння* – це процес визначення подібності або відмінності предметів та явищ дійсності, а також знаходження спільного, притаманного двом або кільком об'єктам.

Застосування методу порівняння має результат, якщо:

- порівнюються тільки такі явища, між якими є деяка об'єктивна спільність;
- порівняння здійснюється за найбільш важливими, суттєвими (у плані конкретного завдання) рисами.

Різні об'єкти чи явища можуть порівнюватися безпосередньо або опосередковано через їх порівняння з будь-яким іншим об'єктом (еталоном). У першому випадку зазвичай отримують якісні результати (більше – менше, вище – нижче). Порівняння ж об'єктів з еталоном дає можливість отримати кількісні характеристики. Таке порівняння називають вимірюванням.

*Вимірювання* – це визначення числового значення певної величини за допомогою одиниці виміру. Вимірювання передбачає наявність таких основних елементів: об'єкт вимірювання, еталон, вимірювальні прилади, метод вимірювання.

*Експеримент* – це такий метод вивчення об'єкта, за яким дослідник активно й цілеспрямовано впливає на нього завдяки створенню штучних умов або використанню природних умов, необхідних для виявлення відповідної властивості. Переваги експериментального вивчення об'єкта порівняно зі спостереженням такі:

- під час експерименту можна вивчати явище «у чистому вигляді», звільнившись від побічних факторів, які затінюють основний процес;
- в експериментальних умовах можна дослідити властивості об'єктів;
- повторюваність експерименту: можна проводити дослідження стільки разів, скільки це потрібно.

У науковому дослідженні експеримент і теорія найтісніше взаємозв'язані. Усіляке ігнорування експерименту неодмінно призводить до помилок, тому всебічне поширення експериментальних досліджень являє собою один із найважливіших шляхів розвитку сучасної науки.

### 1.3 Методи теоретичних досліджень

*Ідеалізація* – це конструювання подумки об'єктів, які не існують насправді або майже нездійсненні (наприклад, абсолютно тверде тіло, абсолютно чорне тіло, лінія, площина).

*Формалізація* – метод вивчення різноманітних об'єктів шляхом відображення їхньої структури у знаковій формі за допомогою штучних мов, наприклад мовою математики. Переваги формалізації:

- забезпечує узагальненість підходу до вирішення проблем;
- надає стислості й чіткості фіксації значень;
- забезпечує однозначність символіки (немає багатозначності звичайної мови);
- дає змогу формувати знакові моделі об'єктів та замінювати вивчення реальних речей і процесів вивченням цих моделей.

*Аксиоматичний метод* – метод побудови наукової теорії, при якому деякі твердження приймаються без доведень, а всі інші знання виводяться з них відповідно до певних логічних правил.

*Гіпотеза й припущення.* У становленні теорій як системи наукового знання найважливішу роль відіграє гіпотеза. Гіпотеза є формою осмислення фактичного матеріалу, переходу від фактів до законів.

Розвиток гіпотези відбувається за трьома стадіями:

- накопичення фактичного матеріалу й висловлення на його основі припущень;
- формування гіпотези, тобто висновків із зробленого припущення, і на його основі – прийнятної теорії;
- перевірка отриманих результатів на практиці і на її основі уточнення гіпотези; якщо під час перевірки висновок відповідає дійсності, то гіпотеза перетворюється на наукову теорію.

Гіпотези (як і ідеї) мають імовірнісний характер. На їхній основі відбувається систематизація накопичених раніше знань і здійснюється пошук нових наукових результатів – у цьому суть і призначення гіпотези як форми розвитку науки.

*Історичний метод* дає змогу дослідити виникнення, формування й розвиток процесів і подій у хронологічній послідовності з метою виявлення внутрішніх і зовнішніх зв'язків, закономірностей і суперечностей. Цей метод дослідження використовується переважно у суспільних науках. У прикладних науках він застосовується, наприклад, під час вивчення розвитку й формування тих чи інших галузей науки і техніки.

*Системний підхід* полягає у комплексному дослідженні великих і складних об'єктів (систем), дослідженні їх як єдиного цілого з узгодженим функціонуванням усіх елементів і частин. З огляду на цей принцип треба вивчити кожен елемент системи у його зв'язку та взаємодії з іншими елементами, виявити вплив властивостей окремих частин системи на її поведінку в цілому, визначити емерджентні властивості системи і оптимальний режим її функціонування.

Ускладнення завдань та об'єктів дослідження спричиняє необхідність розподілення (декомпозиції) системи на системи нижчого рівня (підсистеми), які досліджуються автономно, причому з обов'язковим урахуванням подальшого узгодження цілей кожної підсистеми із загальною метою системи. Таким чином, декомпозиція наперед визначає створення ієрархії системи. Застосування декомпозиції зумовлено не тільки неможливістю охопити неосяжне, але й різномірністю елементів складної системи і, як наслідок, необхідністю залучення фахівців різного профілю.

#### **1.4 Методи, що застосовуються на емпіричному й теоретичному рівнях дослідження**

*Абстрагування* – відхід у думці від несуттєвих властивостей, зв'язків, відношень предметів і виділення декількох рис, котрі цікавлять дослідника.

*Аналіз і синтез.* Аналіз – метод пізнання, який дає змогу поділяти предмети дослідження на складові частини (природні елементи об'єкта або його властивості й відношення). Синтез, навпаки, припускає поєднання окремих частин або рис предмета в єдине ціле. Аналіз і синтез взаємозв'язані, вони являють собою єдність протилежностей.

Аналіз і синтез можуть бути:

- прямими, або емпіричними (використовуються для виокремлення частин об'єкта, виявлення його властивостей, найпростіших вимірювань і т. ін.);
- зворотними, або елементарно-теоретичними (базуються на деяких теоретичних міркуваннях стосовно причинно-наслідкового зв'язку різних явищ або дії будь-якої закономірності; при цьому виділяються та поєднуються явища, які здаються суттєвими, а другорядні – ігноруються);
- структурно-генетичними (потребують виокремлення у складному явищі таких елементів, які вирішально впливають на всі інші сторони об'єкта).

*Індукція та дедукція.* Дедуктивною називають таку розумову конструкцію, у якій висновок щодо якогось елемента множини робиться на основі знання загальних властивостей усієї множини. Змістом дедукції як методу пізнання є використання загальних наукових положень під час дослідженні конкретних явищ.

Під індукцією розуміють перехід від часткового до загального, коли на основі знання про частину предметів класу робиться висновок стосовно класу в цілому. Дедукція та індукція – взаємопротилежні методи пізнання.

*Моделювання* – метод, який ґрунтується на використанні моделі як засобу дослідження явищ і процесів природи. Під моделями розуміють системи, що замінюють об'єкт пізнання і є джерелом інформації стосовно нього. Моделі – це такі аналоги, подібність яких до оригіналу є суттєвою, а розбіжність – несуттєвою. Моделі поділяють на два види: матеріальні й ідеальні. Матеріальні моделі втілюються у певному матеріалі – дереві, металі, склі й т. ін. Ідеальні моделі фіксуються в таких наочних елементах, як креслення, рисунок, схема, комп'ютерна програма тощо.

Метод моделювання має таку структуру: постановка завдання; створення або вибір моделі; дослідження моделі; перенесення знань із моделі на оригінал.

### **1.5 Загальна схема наукового дослідження**

Увесь процес наукового дослідження можна зобразити у вигляді такої логічної схеми:

- обґрунтування актуальності вибраної теми;
- формулювання мети й конкретних завдань дослідження;
- визначення об'єкта й предмета дослідження;
- вибір методів (методики) проведення дослідження;
- опис процесу дослідження;
- обговорення результатів дослідження;
- формулювання висновків і оцінювання одержаних результатів.

*Обґрунтування актуальності вибраної теми* – початковий етап будь-якого дослідження.

Висвітлення актуальності не має бути багатослівним. Достатньо кількома реченнями висловити головне – суть проблеми, з чого й впливе актуальність теми. Проблема завжди виникає тоді, коли колишнє знання вже виявило свою неспроможність, а нове ще не набуло розвиненої форми. Таким чином, проблема в науці – це суперечлива ситуація, яка потребує свого вирішення. Така ситуація найчастіше виникає внаслідок відкриття нових фактів, які явно не вкладаються у рамки колишніх теоретичних уявлень, тобто коли жодна з теорій не може пояснити щойно виявлені факти.

Правильна постановка та ясне формулювання нових проблем часом мають не менше значення, ніж вирішення цих проблем. По суті, саме вибір проблеми якщо не цілком, то дуже великою мірою визначає як стратегію дослідження взагалі, так і напрямок наукового пошуку зокрема. Невипадково вважається, що сформулювати наукову проблему – означає

показати вміння відокремити головне від другорядного, виявити відоме й невідоме науці про предмет дослідження.

Від доведення актуальності вибраної теми слід перейти до *формулювання мети й конкретних завдань дослідження*, які мають бути вирішені відповідно до цієї мети. Це зазвичай робиться у формі перелічення (вивчити..., описати ..., встановити..., з'ясувати..., вивести формулу... й т. ін.).

*Мета дослідження* – це запланований результат, який має бути конструктивним, тобто спрямованим на вироблення суспільно корисного продукту з ліпшими, ніж були раніше, показниками якості або процесу її досягнення.

Надалі визначаються *об'єкт і предмет дослідження*. Об'єкт – це процес або явище, які породжують проблемну ситуацію і вибрано для вивчення. Предмет – це те, що міститься в межах об'єкта.

Об'єкт і предмет дослідження як категорії наукового процесу співвідносяться як загальне й часткове. В об'єкті виділяється та його частина, яка є предметом дослідження. Саме на нього й спрямована основна увага дослідника, саме предмет дослідження визначає тему наукової роботи, що вноситься на титульний аркуш як заголовок.

Дуже важливим етапом наукової праці є *вибір методів проведення дослідження* – інструменту отримання фактичного матеріалу й необхідної умови досягнення поставленої в роботі мети.

*Опис процесу дослідження* – основна частина наукової праці, де висвітлюються методика й техніка дослідження з використанням логічних законів і правил.

Інший важливий етап наукового дослідження – обговорення його результатів (на конференціях, засіданнях профілюючих кафедр, наукових семінарах, учених і науково-технічних радах).

Заключним етапом наукового дослідження є формулювання висновків, які містять те нове й суттєве, що становить наукові та практичні результати виконаної наукової роботи.

Наукові результати можна поділити на два види: теоретико-методологічні (для теоретичних досліджень), зокрема концепція, гіпотеза, класифікація, закон, метод; інструментальні (для прикладних та емпіричних досліджень), зокрема спосіб, технологія, методика, алгоритм.

## **1.6 Наукова публікація: поняття, функції, основні види**

Наукова публікація (в перекладі з латинської – *publicato* – оголошую всенародно, оприлюднюю) – це доведення інформації до громадськості за допомогою преси, радіомовлення, телебачення; розміщення її в різних виданнях (газетах, журналах, збірках наукових праць, монографіях, книгах, підручниках).

Головні функції публікацій:

- оприлюднення результатів наукової роботи;
- сприяння встановленню пріоритету автора за наявності аналогічних за змістом наукових статей;
- свідчення про особистий внесок дослідника в розроблення наукової проблеми;
- підтвердження достовірності основних результатів і висновків наукової роботи, її новизни та наукового рівня, оскільки після видання публікація стає об'єктом вивчення й оцінювання широкою науковою громадськістю;
- відображення основного змісту, наукового рівня й новизни дослідження;
- забезпечення первинною науковою інформацією суспільства, повідомлення про появу нового наукового знання, передача його у загальне користування.

Наукові видання можна поділити на дві групи:

- науково-дослідні;
- джерелознавчі.

До науково-дослідних видань належать:

- монографія (наукова праця, присвячена дослідженню однієї теми);
- реферат – коротке письмове викладення однієї або декількох наукових праць, яке стисло висвітлює їхній зміст;
- методичні рекомендації;
- тези доповідей, а також матеріали наукової конференції (неперіодичний збірник підсумків конференції, доповідей, рекомендацій та рішень);
- збірники наукових праць (збірники матеріалів досліджень (наукових статей), виконаних у наукових установах, навчальних закладах).

До другої групи наукових видань належать джерелознавчі, або наукові документальні видання, які містять пам'ятки культури та історичні документи, що пройшли текстологічне опрацювання, мають коментарі, вступи, статті, допоміжні покажчики тощо.

Серед наукових неперіодичних видань можна виділити:

- книги (книжкове видання обсягом понад 48 сторінок);
- брошури (книжкове видання обсягом від 4 до 48 сторінок).

**Монографія** – це книга, в якій у систематизованому вигляді висловлюються основні дані наукових досліджень самими авторами наукових розробок. Монографії можуть бути написані як одноособово, так і колективом авторів. Під час підготовки монографії слід пам'ятати, що в ній найглибше, систематизовано й аргументовано висловлюються загальні результати наукових досліджень.

**Реферат** (від лат. *referre* – доповідати, повідомляти) – коротке викладення змісту одного або декількох документів з певної теми. Під час

індивідуальної роботи з літературою рефератом є короткий запис ідей з декількох джерел. Часто реферат готують для того, щоб передати ці ідеї аудиторії. Обсяг реферату залежить від вибраної теми, змісту документів, їхньої наукової цінності або практичного значення. Найчастіше маємо справу з інформативними й розширеними (зведеними) рефератами.

Інформативний реферат повністю розкриває зміст теми, містить основні фактичні й теоретичні повідомлення. У такому рефераті має бути визначено предмет дослідження й метод роботи, наведено основні результати, викладено дані про умови дослідження, сформульовано власні пропозиції автора щодо практичного значення змісту і уміння довести його до аудиторії слухачів. Реферат може бути посібником для усного виступу з елементами імпровізації.

Інформаційні реферати можна розміщувати у первинних документах (книгах, журналах, збірниках праць, звітах про науково-дослідну роботу) і у вторинних (реферативних журналах і збірниках, інформаційних картках та ін.).

Розширений, або зведений (оглядовий), реферат містить відомості про певну кількість опублікованих і неопублікованих документів з однієї теми, зміст яких викладено у вигляді однорідного тексту.

Структура реферату:

- вступ;
- історія й теорія питання;
- вирішення проблеми в сучасних умовах;
- висновки;
- література;
- додатки (якщо потрібно).

У вступі обґрунтовується актуальність теми, її особливість і значущість у конкретній галузі науки або практики. Далі наводяться основні теоретичні, експериментальні дослідження з теми. Подається перелік основних змістовних аспектів проблеми, які раніше розглядалися вченими, визначаються недостатньо досліджені питання, з'ясовуються причини їх слабкої розробленості. У наступному розділі подається поглиблений аналіз сучасного стану процесу або явища, тлумачення основних поглядів на проблему і позицій щодо неї. Особлива увага приділяється виявленню нових ідей і гіпотез, експериментальним даним, новим методикам вивчення проблеми, практичного досвіду та висловленню власної думки щодо перспектив розвитку досліджуваної проблеми. У висновках подаються узагальнені умовиводи, ідеї, думки, оцінки, пропозиції науковця. До списку літератури включаються публікації переважно останніх 5 – 10 років. Особливо цінуються праці останнього року видання. Обсяг розширеного реферату становить 20 – 24 сторінки.

Призначення **автореферату** – широке ознайомлення наукових працівників з методикою дослідження, фактичними результатами й

основними висновками наукового дослідження (дисертації). Публікація автореферату дає змогу отримати відгуки спеціалістів певної галузі. Автореферат – своєрідна візитна картка дисертації. У ньому викладаються основні ідеї та висновки дисертації, внесок автора у проведене дослідження, ступінь новизни і практична значущість результатів дослідження.

**Методичні рекомендації** повинні відповідати таким вимогам:

- містити оригінальні способи організації відповідної діяльності;
- розкривати нові методичні прийоми, форми й методи організації навчальної, виробничої або будь-якої іншої діяльності;
- підтверджувати ефективність пропонованих методик матеріалами експериментальної роботи, ілюстраціями, таблицями, графіками.

**Тези** є найпоширенішими видами публікації результатів наукових досліджень. Поняття «теза» (від гр. thesis) – це положення, твердження. У словниках зміст цього поняття трактується по-різному. У Великій Радянській Енциклопедії воно дається таким чином: «Теза в широкому розумінні – це будь-яке твердження в спорі або викладення деяких теорій; у вузькому розумінні – основоположне твердження, принципи». У Словнику російської мови С. І. Ожегова: «Теза – положення, що стисло висловлює яку-небудь ідею, а також одну з основних думок твору, доповіді». Виходячи з цих означень, тези доповідей і виступів, опубліковані у пресі, – це коротке викладення якихось ідей або думок, відображених у наукових дослідженнях. Таким чином, у тезах повторюється і стисло формулюється результат наукового дослідження. У них завжди містяться докази, виявляється суть змісту дослідження, узагальнюється матеріал. Вони є цінними для критичного аналізу статті, доповіді, дисертації і т.д.

Обсяг тез зазвичай становить одну-три сторінки машинописного тексту через полуторний інтервал. Теза має таку структуру: теза – обґрунтування – доведення – аргументування – результат – перспектива.

Проте найбільш значущі результати наукового дослідження зазвичай прийнято відображати в наукових статтях.

**Наукова стаття** є одним із видів публікацій, в якій подаються проміжні або кінцеві результати, висвітлюються конкретні окремі питання за темою дослідження, фіксується науковий пріоритет автора. Стаття – це самостійний науковий твір, що являє собою висловлення своїх думок з актуальної наукової проблеми. У статті можуть розкриватися конкретні питання теоретичної й прикладної роботи дослідника.

Основні етапи роботи над статтею:

- визначення теми, її аналіз, складання плану (тез);
- робота над першим варіантом статті на основі плану (тез);
- роздум, обдумування, вирішення проблеми;



- послідовне викладення думки, доказове міркування, точний вибір слів;
- критичне оцінювання результатів роботи, пошук недоліків, помилок і їх виправлення.

Відповідно до Постанови президії ВАК (Вищої атестаційної комісії) України від 15.01.2003 р. № 7-05/1 “Про підвищення вимог до фахових видань, внесених до переліків ВАК України” (Бюлетень ВАК України. – 2003. – № 1) наукова стаття має містити такі обов’язкові елементи:

- постановка проблеми у загальному вигляді та її зв’язок з важливими науковими або практичними завданнями;
- аналіз останніх досліджень і публікацій, в яких започатковано вирішення проблеми і на які спирається автор;
- виділення не вирішених раніше частин загальної проблеми, яким присвячується певна стаття;
- формулювання цілей статті (постановка завдання);
- викладення основного матеріалу дослідження з повним обґрунтуванням отриманих наукових результатів;
- висновки щодо результатів і перспективи подальших досліджень у цьому напрямку.

Наукова стаття подається до редакції завершеною відповідно до вимог, які публікуються в окремих номерах журналів або збірниках у вигляді пам’ятки автору. Обсяг наукової статті, який дорівнює 0,5 – 0,7 авторського аркуша, вважається оптимальним. Рукопис статті повинен містити повну назву роботи, прізвище та ініціали автора, анотацію, список використаної літератури.

Анотація – це стисла характеристика першоджерела. Головна функція анотації – офіційно повідомляти про нові видання; у таких випадках анотація зазвичай поміщається на початку книги, на зворотному боці титульного аркуша, або в кінці книги, трохи вище за технічні дані видавництва. Анотації наукових статей наводяться або на початку (в кінці) статті, або в кінці збірки наукових праць, що полегшує пошук необхідної інформації.

Анотація зазвичай має стандартний вигляд. Вона складається із заголовної частини, до якої входять такі дані: прізвище автора книги (статті), назва книги (статті), назва видання (для статті), вихідні дані видання (місце, час виходу з друку), обсяг публікації (кількість сторінок) і описової частини, що включає вступ у тематику питань, які розглядаються в публікації, безпосереднє викладення теми з переліченням основних проблем, вказівку на актуальність роботи, її соціальну значущість і прийняту в роботі методологію дослідження.

Анотація відрізняється найсильнішим порівняно з іншими вторинними письмовими документами стисненням тексту, для неї характерні максимальна лаконічність і компактність. Під час складання анотації слід використовувати прості речення, що містять пасивні конструкції безособових речень. Обсяг анотації для статті, рекомендований ВАКом, становить 500 символів. Анотація має давати загальне уявлення про зміст статті, інформувати про наявність роботи за темою, але не повинна розкривати подробиць змісту. Якщо інформація, викладена в анотації, становить інтерес, необхідно ознайомитися з роботою в цілому.

Далі наведено зразки кліше, що використовуються в анотації:

„Досліджено (що) ...”;

„Показано (що) ...”;

„Надано статистичний матеріал (про що) ...”;

„Надано характеристику (чого) ...”;

„У систематизованому вигляді наведено аналіз (чого) ...”;

„Проаналізовано (що) ...”;

„Наголошено (на чому) ...”;

„Підкреслено (що) ...”;

„Розкрито основні методологічні положення (чого) ...”;

„Описано деякі методи (чого) ...”;

„Особливу увагу приділено питанням (чого) ...”;

„Досліджено (мало розроблені в літературі) проблеми (чого) ...”;

„Розглянуто значення (чого) ...”;

„Установлено критерії (чого) ...”;

„Розглянуто (ключову) проблему (чого) ...”;

„На основі аналізу (чого) показано (що) ...”;

„Розглянуто чинники, які сприяють (чому) ...”;

„Наведено дані, що наочно показують ...”.

Складаючи анотацію, слід уникати повторень одних і тих самих конструкцій. Вживання слів „вищевикладене”, „згаданий”, „перераховані” та інших робить виклад коротшим і концентрованим.

Зазвичай анотацію статті наводять трьома мовами – українською, російською та англійською. В таблиці 1.1 подано деякі вирази, рекомендовані для написання анотацій англійською мовою.

Розглядаючи основні вимоги до публікацій, зазначимо, що досить часто наукові розробки виконують два або більше дослідників. Таку форму наукової роботи називають співавторством. Співавторство може виражатися в різних формах. Якщо наукова робота написана в рівній мірі двома і більше авторами, вони мають право використовувати всі матеріали, опубліковані в співавторстві, в подальшій своїй науковій роботі, проте при цьому необхідно робити посилання на те, що той або інший матеріал розроблено спільно. Якщо співавторство полягає в тому, що

співавтор розробив або додаток, або список літератури і джерел, або таблиці та графіки для певної наукової публікації, то це відбивається або в передмові до публікації, або в змісті наукової роботи.

Таблиця 1.1 – Типові вирази для написання анотацій англійською мовою

Український варіант	Англійський варіант
Стисло описано...	It is described in short...
Запропоновано...	... is introduced
Показано, що ...	It is shown that...
Надано (запропоновано)...	... is given
Розглянуто ...	It is dealt with...
Забезпечено ...	... is provided for
Призначено для ...	... is designed for
Досліджено ...	... is examined, is investigated
Проаналізовано ...	... is analyzed
Сформульовано...	... is formulated
Підкреслено необхідність використання...	The need is stressed to employ...
Звернено увагу на ...	Attention is drawn to...
Наведено дані про ...	Data are given about...
Зроблено спроби проаналізувати, сформулювати ...	Attempts are made to analyze, to formulate...
Зроблено висновки ...	Conclusions are drawn...
Надано рекомендації ...	Recommendation are given...

Практика показує, що під час визначення обсягу статей, навчальних посібників, монографій та інших творів, підготовлених до друку, у авторів виникають утруднення. Тому наведемо короткі відомості про одиниці виміру, необхідні для визначення обсягу публікацій.

Авторський аркуш – одиниця обсягу авторського твору. Один авторський аркуш дорівнює 40 000 друкарських знаків, що включають усі букви, розділові знаки, цифри, а також пропуски між словами, або 700 рядкам віршованого тексту, або 3000 см<sup>2</sup> ілюстративного матеріалу. У рукопису авторський аркуш зазвичай відповідає 20 – 24 (в середньому 22) сторінкам машинописного тексту.

Друкарський аркуш – одиниця виміру натурального обсягу видання, що дорівнює площі однієї сторони паперового аркуша стандартного формату (60x90 см). Друкарський аркуш містить різну кількість друкарських знаків залежно від формату смуги, набору й кегля шрифту.

Умовним друкарським аркушем називають одиницю обсягу видання, що дорівнює друкарському аркушу формату 60x90 см. Він використовується для перерахунку й зіставлення друкарського обсягу видань різних форматів.

Результати наукових досліджень можна подати не тільки письмово, але й у вигляді доповіді перед аудиторією.

**Доповідь** може повністю збігатися, наприклад, з науковою статтею. Проте переваги усної мови полягають у тому, що є можливість урізноманітнити свою мову, зробити її живішою, відходячи від строгості наукового викладення.

До критеріїв культури мови належать: правильність, доцільність, точність, логічність, ясність, доступність, чистота, виразність, різноманітність засобів вираження, естетичність, доцільність. Існує декілька рівнів, що характеризують культуру мовної діяльності й культуру мови.

Репродуктивний рівень – проста передача інформації слухачам (відсутність зворотного зв'язку; концентрація доповідача на змісті й правильності його викладення; процес розуміння не регулюється).

Адаптаційний рівень – вільна орієнтація в матеріалі (підтримка контакту з аудиторією, урахування її особливостей, застосування різних способів роз'яснення і доведення; керування увагою слухачів, вплив на засвоєння ними матеріалу).

Творчий рівень – найвищий рівень володіння мовою (повне оволодіння аудиторією, активізація творчого мислення слухачів, спонукання до власних роздумів і висновків).

Науковий стиль в усній мові може застосовуватися:

- у діалозі (дискусії, дебатах, бесіді тощо);
- у монологі (повіді, повідомленні, під час захисту дипломної роботи (дисертації), звіті та ін.).

У наукових діалозі й монологі використовуються п'ять типів комунікативних цілей: повідомлення (повіді, інформувати), переконання (довести, обґрунтувати), схвалення (рекомендувати, підтвердити, підтримати), обговорення (спростувати, розкритикувати, заперечити, оспорювати), пояснення (конкретизувати, показати, уточнити, описати, виділити, акцентувати, прокоментувати). Монологічне висловлювання характеризується: цілеспрямованістю, зв'язністю, логічністю, самостійністю, виразністю, завершеністю, безперервністю.

Робота над доповіддю має складатися з таких частин:

- мотиваційна (повідомлення нової наукової інформації, пояснення причини явищ, опис наслідків, встановлення взаємозв'язку між фактами);
- аналітико-синтетична (формування основних думок, положень, написання тексту);
- виконавча (виступ).

Реалізація виконавчої частини потребує знання відповіді на запитання, по-перше, „як почати виступ?"; по-друге, „як зацікавити слухачів під час основного викладення?"; по-третє, „як завершити виступ?".

Починаючи виступ, можна привернути до себе увагу такими способами:

- поставити проблемне або оригінальне запитання;
- почати з цікавої цитати за темою виступу;
- навести конкретний приклад з життя;
- образно порівняти предмет виступу з конкретним явищем.

На початку доповіді обґрунтовуються тема, її актуальність й наукове значення. При цьому зацікавити слухачів під час виступу можна, виклавши незвичайні факти, розповівши те, що безпосередньо стосується всіх слухачів доповіді.

Закінчення виступу необхідно ретельно продумати, склавши декілька варіантів його завершення. Так, закінчуючи виступ, потрібно або стисло викласти основні думки, яких торкнулися в доповіді, або процитувати щонебудь за темою доповіді, або створити кульмінацію, залишивши слухачів у роздумах над поставленою проблемою. Для вибору найбільш відповідного варіанта слід підготувати кінцівки-резюме до матеріалів доповіді. Резюмування – це підведення підсумків. Форма наведення резюме може бути такою: „Таким чином...”, „Отже, можна стверджувати...”, „Основна ідея, отже, зводиться до такого...”, „Підсумуємо сказане...”.

До типових помилок під час виступу слід віднести:

- зловживання іноземною термінологією і поняттями, що ускладнюють сприйняття головної думки;
- наявність слів-паразитів („ось”, „значить”, „так би мовити” і т.д.);
- надмірну гучність голосу (слухачі через 8-10 хвилин не сприймають таку мову);
- побудову складних речень, в яких є понад 14 слів (такі фрази сприймаються дуже важко);
- монотонна інтонація, без акцентів на значущих моментах доповіді й т.д.

Якщо доповідь має відбутися перед аудиторією, більшість якої – наукові працівники-теоретики, увагу доцільніше зосередити на обґрунтуванні методу дослідження й прийнятих допущеннях, звернувши увагу на те, як ці допущення справджуються на практиці та який отримано ефект від впровадження результатів дослідження.

Коли ж аудиторію становлять головним чином практичні працівники, представники виробництва, корисно основну увагу приділити практичній цінності дослідження, наочно показати, що воно дає виробництву, шляхи його впровадження й перспективи подальших робіт.

Для того щоб зробити доповідь зрозумілішою й цікавішою для аудиторії та мати самому логічну канву свого виступу, корисно підготувати до доповіді демонстраційні матеріали, наприклад плакати, графіки, слайди і т. ін.

## 1.7 Поняття про наукову інформацію та її роль у проведенні наукових досліджень

На сучасному етапі розвитку ринкових відносин, коли темпи накопичення й передачі інформації зростають, виникло протиріччя між виробництвом інформації та можливостями її споживання, перероблення й використання. Необхідні відповідні методики орієнтації наукових працівників на найбільш продуктивний пошук і використання відповідних інформаційних матеріалів.

Наукова інформація – це логічна інформація, яка отримується в процесі пізнання, адекватно відображає закономірності об'єктивного світу і використовується в суспільно-історичній практиці.

Основні ознаки наукової інформації:

- отримується в процесі пізнання закономірностей об'єктивної дійсності, підґрунтям якої є практика, і подається у відповідній формі;
- це документовані або оприлюднені відомості про вітчизняні та зарубіжні досягнення науки, техніки, виробництва, отримані в процесі під час науково-дослідної, дослідно-конструкторської, виробничої та громадської діяльності.

Основні джерела науково-технічної інформації можна згрупувати таким чином:

- 1 Монографія.
- 2 Збірник – це видання, яке складається з окремих робіт різних авторів, присвячених одному напрямку. У збірнику публікуються закінчені праці з рекомендаціями щодо їх використання.
- 3 Періодичні видання – журнали, бюлетені та інші видання з різних галузей науки і техніки. У періодичних виданнях можуть друкуватись наукові праці та їхні результати. Матеріал викладається в популярній, доступній формі.
- 4 Спеціальні випуски технічних видань – документи інформаційного, рекламного плану, аналітичні, статистичні дані.
- 5 Патентно-ліцензійні видання (патентні бюлетені).
- 6 Стандарти – нормативно-технічні документи щодо єдиних вимог до продукції, її розроблення, виробництва й застосування.
- 7 Навчальна література – підручники, навчальні посібники, навчально-методична література.
- 8 Надруковані документи – дисертації, звіти про науково-дослідну роботу, окремі праці. Це документи для студентів, аспірантів, які займаються науково-дослідною роботою: планові, звітні документи, статистичні й опубліковані доповіді, методичні й інструкційні матеріали.
- 9 Науково-інформаційна діяльність – сукупність дій, спрямованих на задоволення потреб громадян, юридичних осіб і держави, що полягають в

її збиранні, аналітико-синтетичному обробленні, фіксації, зберіганні, пошуку й поширенні.

10 Інформаційні ресурси науково-технічної інформації – це систематизований збір науково-технічної літератури й документації, зафіксованої на паперових та інших носіях.

11 Довідково-інформаційний фонд – сукупність упорядкованих первинних документів і довідково-пошукового апарата, призначених для задоволення інформаційних потреб.

12 Довідково-пошуковий апарат – сукупність упорядкованих вторинних документів, що створюються для пошуку першоджерел.

13 Інформаційні ресурси спільного користування – сукупність інформаційних ресурсів державних органів науково-технічної інформації (бібліотеки, фірми, організації).

14 Аналітико-статистичне оброблення науково-технічної та практичної інформації.

15 Інформаційний ринок – це система економічних, організаційних і правових відносин щодо продажу й купівлі інформаційних ресурсів, технологій, продукції та послуг.

**Універсальна десяткова класифікація (УДК)** – система класифікації інформації, яка широко використовується у всьому світі для систематизації творів науки, літератури й мистецтва, періодичного друку, різних видів документів та організації картотек.

Десятковою УДК називають тому, що для її побудови використано десятковий принцип.

Десятковий принцип структури УДК практично дає можливість майже необмежено розширювати її шляхом додавання нових цифр до тих, що вже є, не впливаючи на всю систему в цілому. Індокси УДК можуть бути пов'язані один з одним у всіляких комбінаціях, завдяки чому в схемі класифікації можна відобразити незліченну кількість понять.

Розробленням таблиць УДК займається Міжнародна федерація з інформації і документації (МФД), яка веде роботу з їх вдосконалення відповідно до розвитку науки і техніки.

Основні розділи УДК:

0 Загальний відділ.

1 Філософські науки. Філософія.

2 Релігія. Атеїзм.

3 Суспільні науки (зокрема економіка).

4 (Вільний з 1961 р.).

5 Математика. Природничі науки.

6 Прикладні науки. Медицина. Техніка.

7 Мистецтво. Декоративно-прикладне мистецтво. Фотографія. Музика. Ігри. Спорт.

8 Мовознавство. Філологія. Художня література. Літературознавство.

9 Географія. Біологія. Історія.

## 1.8 Організація наукової діяльності в Україні

Науково-дослідна діяльність в Україні регламентується Законом України „Про наукову і науково-технічну діяльність”. Закон містить певні терміни.

*Наукова діяльність* – інтелектуальна творча діяльність, спрямована на одержання й використання нових знань. Основними її формами є фундаментальні та прикладні наукові дослідження.

*Науково-технічна діяльність* – інтелектуальна творча діяльність, спрямована на одержання й використання нових знань у всіх галузях техніки і технологій. Її основними формами (видами) є науково-дослідні, дослідно-конструкторські, проектно-конструкторські, технологічні, пошукові та проектно-пошукові роботи, виготовлення дослідних зразків або партій науково-технічної продукції, а також інші роботи, пов'язані з доведенням наукових і науково-технічних знань до стадії практичного їх використання.

*Науково-педагогічна діяльність* – педагогічна діяльність у вищих навчальних закладах і закладах післядипломної освіти III-IV рівнів акредитації, пов'язана з науковою та (або) науково-технічною діяльністю.

*Науково-організаційна діяльність* – діяльність, спрямована на методичне, організаційне забезпечення й координацію наукової, науково-технічної та науково-педагогічної діяльності.

*Фундаментальні наукові дослідження* – наукова теоретична та (або) експериментальна діяльність, спрямована на одержання нових знань про закономірності розвитку природи, суспільства, людини, їхні взаємозв'язки.

*Прикладні наукові дослідження* – наукова і науково-технічна діяльність, спрямована на одержання і використання знань для практичних цілей.

*Науково-дослідна (науково-технічна) установа* (далі – наукова установа) – юридична особа незалежно від форми власності, створена в установленому законодавством порядку, для якої наукова або науково-технічна діяльність є основною і становить понад 70 відсотків загального річного обсягу виконаних робіт.

*Наукова робота* – дослідження з метою одержання наукового результату.

*Науковий результат* – нове знання, одержане в процесі фундаментальних або прикладних наукових досліджень та зафіксоване на носіях наукової інформації у формі звіту, наукової праці, наукової доповіді, наукового повідомлення про науково-дослідну роботу, монографічного дослідження, наукового відкриття тощо.



*Науково-прикладний результат* – нове конструктивне або технологічне рішення, експериментальний зразок, закінчене випробування, розробка, яку впроваджено або може бути впроваджено у суспільну практику. Науково-прикладний результат може бути у формі звіту, ескізного проекту, конструкторської або технологічної документації на науково-технічну продукцію, натурального зразка тощо.

Суб'єктами наукової та науково-технічної діяльності є вчені, наукові працівники, науково-педагогічні працівники, а також наукові установи, наукові організації, вищі навчальні заклади III-IV рівнів акредитації, громадські організації у науковій та науково-технічній діяльності (далі – громадські наукові організації).

Сьогодні науково-дослідну роботу ведуть:

- науково-дослідні та проектні установи й центри Академії наук України;
- науково-виробничі, науково-дослідні, проектні установи, системи галузевих академії;
- науково-дослідні, проектні установи і центри міністерств і відомств;
- науково-дослідні установи і кафедри вищих навчальних закладів;
- науково-виробничі, проектні установи і центри при промислових підприємствах, об'єднаннях;
- Державний комітет України з питань науки і технологій, який забезпечує єдину державну політику в галузі науки та її використання на практиці.

Вищим державним науковим центром є Національна академія наук України (НАН). Вона разом з Державним комітетом у справах науки та технологій України очолює і координує фундаментальні й прикладні дослідження в різних галузях науки. НАН є державною науковою установою, яка об'єднує всі напрями науки та підтримує міжнародні зв'язки з науковими центрами інших країн. При Національній академії наук України створено міжвідомчу раду з координації фундаментальних досліджень. Очолює НАН України Президент, який обирається загальними зборами вчених. Вони ж обирають трьох віце-президентів, вченого секретаря, Президію і ревізійну комісію. НАН України має в своєму складі відділення з відповідних галузей науки, зокрема, математики, інформатики, механіки, фізики і астрономії, наук про землю, хімії, загальної біології, економіки, історії, філософії, літератури, мови та мистецтва тощо.

До складу НАН входять наукові інститути з відповідних галузей, є територіальні відділення (Донецьке, Західне, Південне та ін.) і територіальні філіали.

Відділення НАН об'єднують науково-дослідні інститути (НДІ), які є провідними у певних галузях знань.

Крім НАН в Україні функціонують галузеві академії, наприклад: Академія педагогічних наук України, Українська академія аграрних наук, у

складі якої є НДІ з економіки, Академія медичних наук України, Академія правових наук України, Академія мистецтв України.

У галузевих НДІ окремі підрозділи здійснюють наукові дослідження за темами профілю, переважно прикладного характеру.

Науково-дослідну діяльність прикладного характеру на нижчих рівнях здійснюють у НДІ відділи, лабораторії, сектори, а також вищі навчальні заклади (університети, академії, інститути). Останні мають спеціальні підрозділи, які виконують науково-дослідні роботи за рахунок державних бюджетних і госпрозрахункових коштів. Проводять дослідження науково-педагогічні працівники із залученням студентів, а також молодих учених, здобувачів кандидатських і докторських дисертацій за науковою тематикою вищих навчальних закладів.

У системі державної організації науки важливе місце належить Вищій атестаційній комісії (ВАК) України. Вона являє собою державний орган, який здійснює керівництво атестацією наукових кадрів вищої кваліфікації, забезпечує єдність вимог до здобувачів наукових ступенів кандидата та доктора наук, контролює якість дисертаційних робіт, їхню наукову й практичну цінність, беручи тим самим участь у формуванні наукового потенціалу держави.

### **Запитання для самоперевірки**

- 1 Назвіть специфічні ознаки науки як системи наукових знань.
- 2 Дайте означення поняття „метод”.
- 3 Дайте відповіді на тестові запитання:
  - А Означення терміна „спостереження”:
    - а)вивчення об'єкта, основане на активній, цілеспрямованій дії;
    - б)планомірне, систематичне й цілеспрямоване сприйняття об'єктів у цілому або окремих його сторін, при якому дослідник не втручається в поведінку об'єкта;
    - в)визначення числового значення деякої величини.
  - Б Означення терміна „порівняння”:
    - а)вивчення об'єкта, основане на активній, цілеспрямованій дії;
    - б)планомірне, систематичне й цілеспрямоване сприйняття об'єктів у цілому або окремих його сторін, при якому дослідник не втручається в поведінку об'єкта;
    - в)визначення числового значення деякої величини.
  - В Означення терміна „експеримент”:
    - а)вивчення об'єкта, основане на активній, цілеспрямованій дії;
    - б)планомірне, систематичне й цілеспрямоване сприйняття об'єктів у цілому або окремих його сторін, при якому дослідник не втручається в поведінку об'єкта;
    - в)визначення числового значення деякої величини.

Г Означення терміна „абстрагування”:

а) відвернення від деяких характеристик (властивостей) об'єктів, що вивчаються;

б) метод наукового пізнання, за допомогою якого отримують знання про одні предмети й явища на основі їх схожості з іншими;

в) визначення (ідентифікація) загального поняття, в якому знаходять віддзеркалення головне, основне, таке, що характеризує об'єкти певного класу;

г) уявне конструювання й вивчення неіснуючих об'єктів, але таких, що зберігають деякі риси реальних об'єктів.

Д Означення терміна „аналогія”:

а) відвернення від деяких характеристик (властивостей) об'єктів, що вивчаються;

б) метод наукового пізнання, за допомогою якого отримують знання про одні предмети й явища на основі їх схожості з іншими;

в) визначення (ідентифікація) загального поняття, в якому знаходять віддзеркалення головне, основне, таке, що характеризує об'єкти певного класу;

г) уявне конструювання й вивчення неіснуючих об'єктів, але таких, що зберігають деякі риси реальних об'єктів.

Е Означення терміна „узагальнення”:

а) відвернення від деяких характеристик (властивостей) об'єктів, що вивчаються;

б) метод наукового пізнання, за допомогою якого отримують знання про одні предмети й явища на основі їх схожості з іншими;

в) визначення (ідентифікація) загального поняття, в якому знаходять віддзеркалення головне, основне, таке, що характеризує об'єкти певного класу;

г) уявне конструювання й вивчення неіснуючих об'єктів, але таких, що зберігають деякі риси реальних об'єктів.

Ж Означення терміна „ідеалізація”:

а) відвернення від деяких характеристик (властивостей) об'єктів, що вивчаються;

б) метод наукового пізнання, за допомогою якого отримують знання про одні предмети й явища на основі їх схожості з іншими;

в) визначення (ідентифікація) загального поняття, в якому знаходять віддзеркалення головне, основне, таке, що характеризує об'єкти певного класу;

г) уявне конструювання й вивчення неіснуючих об'єктів, але таких, що зберігають деякі риси реальних об'єктів.

З Визначіть, індукцією чи дедукцією є висновок „Якщо йде дощ, то земля мокра. Йде дощ – земля мокра”:

а) індукцією, правильний;

- б) індукцією, неправильний;
- в) дедукцією, правильний;
- г) дедукцією, неправильний.

К Визначіть, індукцією чи дедукцією є висновок „Італія – республіка, Португалія – республіка, Фінляндія - республіка, Франція – республіка. Ці країни – західноєвропейські. Всі західноєвропейські країни – республіки”:

- а) індукцією, правильний;
- б) індукцією, неправильний;
- в) дедукцією, правильний;
- г) дедукцією, неправильний.

4 Сформулюйте й охарактеризуйте основні етапи наукового дослідження.

5 Перелічіть основні вимоги до наукових статей.

6 До якого розділу УДК належить підрозділ „Фінанси. Банківська справа. Гроші та грошовий обіг”? Для визначення УДК можна скористатися веб-ресурсом <http://www.teacode.com/online/udc/index.html>.

7 Які організації ведуть наукову роботу в Україні?

## **2 МЕТОДИ АНАЛІЗУ ЕКСПЕРТНОЇ ІНФОРМАЦІЇ**

### **2.1 Сучасна теорія вимірювань і експертне оцінювання**

Характерною особливістю моделювання й прогнозування соціально-економічних процесів є багатоваріантність, тобто можливість використання різних методів, моделей, інформаційного забезпечення, критеріїв оцінювання адекватності моделі тощо. Вибір варіанта серед конкуруючих базується на певній системі правил, що забезпечують надання обґрунтованих оцінок кожному варіанту.

Уважається, що експерт (від лат. *expertus* – досвідчений) знає цю систему правил і може порівняти варіанти, присвоючи кожному з них числа. Найчастіше перевага або відносна значущість варіантів встановлюється за допомогою методів ранжування, попарних порівнянь або безпосереднього оцінювання.

Думки експертів часто висловлюються за допомогою порядкової шкали (докладніше ці шкали розглянемо нижче), тобто експерт може сказати (і обґрунтувати), що один показник якості продукції важливіший, ніж інший, перший технологічний об'єкт більш небезпечний, ніж другий, і т.д. Але він не в змозі сказати, в скільки разів, наскільки важливіший або, відповідно, більш небезпечний. Експертів часто просять зробити ранжування (впорядкування) об'єктів експертизи, тобто розташувати їх у порядку зростання (або убування) інтенсивності експертизи характеристики, що цікавить організаторів. Ранг – це номер (об'єкта

експертизи) у впорядкованому ряду. Формально ранги виражаються числами 1, 2, 3 ..., але з цими числами не можна робити звичні арифметичні операції. Наприклад, хоча  $1 + 2 = 3$ , але не можна стверджувати, що для об'єкта, який стоїть на третьому місці у впорядкованому ряду, інтенсивність характеристики, що вивчається, дорівнює сумі інтенсивностей об'єктів з рангами 1 і 2. Так, один з видів експертного оцінювання – оцінки учнів. Навряд чи хто-небудь стверджуватиме, що знання відмінника дорівнюють сумі знань тих, хто вчиться на двійки і трійки (хоча  $5 = 2 + 3$ ). Тому очевидно, що для аналізу такого роду якісних даних необхідна не всім відома арифметика, а інша теорія, що дає базу для розроблення, вивчення й застосування конкретних методів розрахунку. Це є теорія вимірювань, основні тези якої розглянуто нижче.

### **Типи шкал**

*Шкала найменувань, або класифікації.* Використовується для опису належності об'єктів до певних класів. Усім об'єктам одного й того ж класу відповідає одне й те саме число, об'єктам різних класів – різні.

*Шкала порядку* застосовується для вимірювання впорядкування об'єктів поодиноці або сукупності ознак. Прикладом є шкала твердості мінералів. Числа шкали порядку відображають тільки порядок проходження об'єктів і не дають можливості сказати, наскільки один об'єкт переважає інший.

*Шкала інтервалів* застосовується для відображення величини відмінності між властивостями об'єктів (вимірювання температури за Фаренгейтом і Цельсієм). Шкала може мати довільні точки відліку й масштаб.

*Шкала відносин* використовується, наприклад, для вимірювання маси, довжини, ваги. Числа цієї шкали відображають відношення властивостей об'єктів, тобто в скільки разів властивість одного об'єкта перевершує властивість іншого.

*Шкала різниць* використовується для вимірювання властивостей об'єктів за необхідності вираження того, наскільки один об'єкт перевершує інший за однією або декількома ознаками. Шкала різниць є окремим випадком шкали інтервалів.

*Абсолютна шкала* – окремий вид шкали інтервалів. На шкалі позначаються нульова точка відліку й одиничний масштаб. Застосовується для вимірювання кількості об'єктів.

### **Методи вимірювань**

*Ранжування.* Під час ранжування експерт розташовує об'єкти в порядку їх переваги, керуючись одним або декількома показниками порівняння.

*Парне оцінювання, або метод парних порівнянь,* є процедурою встановлення переваг об'єктів при порівнянні всіх можливих пар.

*Безпосереднє оцінювання* є процедурою присвоєння об'єктам числових значень за шкалою інтервалів. Еквівалентним об'єктам присвоюється одне й те саме число. Це може бути здійснено тільки у разі повної інформованості експертів про властивості об'єктів. Замість числової осі можна використовувати бальне оцінювання.

*Послідовне порівняння* включає ранжування й безпосереднє оцінювання.

## **2.2 Методи проведення експертиз**

Методи експертного оцінювання – це методи організації роботи з фахівцями-експертами і аналізу думок експертів.

Методи проведення групових експертиз поділяють на такі:

**1 Очні та заочні.** При очному методі проведення експертизи експерт працює у присутності організатора експертизи. Ця необхідність може виникнути, якщо завдання сформульовано недостатньо чітко або якщо воно дуже складне і його треба уточнити. Експерт може звернутися до організатора за роз'ясненням. Для проведення очного опитування потрібно більше часу, оскільки організатор експертизи працює з кожним учасником особисто, але при складності поставленого завдання це компенсується більшою точністю отриманих результатів.

**2 Індивідуальні й колективні.** При колективному методі проведення експертизи поставлена проблема вирішується спільно, за круглим столом. При індивідуальному – кожен експерт оцінює проблему, виходячи з особистого досвіду й переконань. Під час проведення експертизи методом експертних комісій група фахівців колективно оцінює досліджувану проблему. Слід зазначити, що при такій організації на рішення групи експертів може впливати один з авторитетних її членів. Але в цьому випадку ймовірність отримання рішення поставленого завдання є більшою. Цей метод рекомендується застосовувати за необхідності знайти рішення в найкоротші терміни.

**3 Зі зворотним зв'язком і без зворотного зв'язку.** Проведення експертизи методом зі зворотним зв'язком пов'язано з великими витратами часу, оскільки в цьому випадку необхідно провести декілька турів. Але повідомлення результатів попереднього туру й подальше опитування дають можливість зменшити діапазон розкиду в індивідуальних відповідях і зблизити точки зору експертів. Експертиза закінчується, коли досягнуто достатню збіжність відповідей експертів. Досвід показує, що найчастіше достатньо чотирьох турів. Зазвичай цей метод застосовується в прогнозуванні, коли є великий ступінь невизначеності. Експертизу без зворотного зв'язку можна проводити при хорошій інформованості експертів у галузі поставленого завдання. Метод без зворотного зв'язку передбачає один тур опитування.

Прикладом методу проведення експертизи зі зворотним зв'язком є метод Дельфі, який передбачає проведення декількох турів опитування й анонімне анкетування. Назву методу надали за асоціацією зі звичаєм стародавніх греків, а саме: для отримання порад і підтримки під час прийняття рішень звертатися в Дельфійський храм. У США з кінця 50-х років минулого століття методом Дельфі стали називати експертну процедуру прогнозування науково-технічного розвитку. Так, у першому турі експерти називали ймовірні дати тих або інших майбутніх науково-технічних звершень. У другому турі кожен експерт ознайомлювався з прогнозами всіх інших (без указання прізвищ авторів прогнозів). Якщо прогноз одного експерта сильно відрізнявся від прогнозів основної маси експертів, його просили пояснити свою позицію. І такий експерт часто змінював свої оцінки, наближаючись до середніх значень. Ці середні значення і видавалися замовникові як групова думка. Експертиза за методом Дельфі найчастіше проводиться в чотири тури. Таким чином, анонімність думок (оцінок), обґрунтування точок зору експертів, що дали крайні оцінки, зворотний зв'язок, що реалізується за допомогою багатотурової процедури, – основні особливості методу Дельфі. Метод Дельфі використовується головним чином під час прогнозування і перспективного планування. Характерною особливістю цього методу є те, що розкид оцінок експертів зменшується від туру до туру, а їхня узгодженість зростає. Основною причиною цього є закладений у методі „тиск” на кожного експерта з метою наближення його оцінки до середньої оцінки групи експертів. Дійсно, кожному експертові кілька разів повідомляють середню оцінку; зрозуміло, що у нього немає підстав відхилитися від неї далі, але є підстави наблизитися до неї.

Однак явище „проходження думки більшості” експертів більшою мірою властиве іншому методу експертного оцінювання, який відомий як „мозковий штурм”. Іноді метод Дельфі позиціонується як різновид „мозкового штурму”. Він також застосовується з 50 – 60-х років минулого століття, і основною його метою є виявлення нових ідей. Організовується „мозковий штурм” як збори експертів, на виступ яких накладено одне обмеження – не можна критикувати пропозиції інших. Під час засідання експерти, „заряджаючись” один від одного, висловлюють все більш екстравагантні міркування. Години через дві записане на магнітофон або відеокамеру засідання закінчується і починається другий етап „мозкового штурму” – аналіз висловлених ідей. Зазвичай зі 100 ідей 30 заслуговують на подальше опрацювання, з них 5 – 6 дають можливість сформулювати прикладні проекти, а 2 – 3 виявляються такими, що у підсумку дають корисний ефект. „Мозковий штурм” можна вважати лише вельми специфічним методом опитування, придатним в особливих випадках, що потребують творчої атмосфери й максимальної різноманітності думок. Давно також існує і широко використовується на

практиці метод проведення експертиз, який полягає у відкритій дискусії за круглим столом з розглядуваної проблеми для вироблення єдиної думки експертів. Іноді цей метод називають методом комісії. Під час дискусії кожен експерт має можливість неодноразово висловлювати думки, змінювати й уточнювати свою точку зору з урахуванням думок інших експертів. Переваги методу: оперативний зворотний зв'язок, наявність вербальних контактів, швидке подолання нерозуміння між експертами. Недоліки дискусії: сильний вплив думок авторитетних експертів, схильність окремих експертів до думки більшості, небажання відкрито переглянути вже висловлену думку.

### 2.3 Підбір експертів

Незважаючи на різноманітність типів експертної інформації і методів проведення експертиз, можна виділити такі загальні проблеми підготовки й організації експертизи: виявлення кола фахівців з цього питання; оцінювання рівня їхньої компетентності; формування експертних груп із компетентних осіб; організація збору інформації у членів групи.

Проблема підбору експертів – одна з найбільш складних у теорії і практиці експертних досліджень. Можна виділити дві складові цієї проблеми: складання переліку можливих експертів; формування експертної комісії відповідно до компетентності кандидатів.

Якщо експертизи проводяться багато разів, то зазвичай ведеться реєстр можливих експертів (наприклад, у області державної екологічної експертизи або суддівства фігурного катання).

Якщо експертиза проводиться вперше, для формування списку експертів застосовується метод „сніжної грудки”, при якому від кожного фахівця, що залучається як експерт, отримують декілька (наприклад, п'ять) прізвищ тих, хто може бути експертом з цієї тематики. Кожного нового експерта опитують за тією ж схемою. Процес формування списку експертів припиняється, коли список розширюється до потрібних розмірів або коли нові прізвища перестають мати місце.

На практиці для визначення необхідної й достатньої кількості експертів можна використовувати таку формулу:

$$N = \frac{t_{\alpha}}{\varepsilon_1},$$

де  $t_{\alpha}$  – показник достовірності для заданої довірчої ймовірності отриманого результату;

$$\varepsilon_1 \text{ – гранична похибка в частках } \sigma, \varepsilon_1 = \frac{\varepsilon}{\sigma} \text{ (} \varepsilon \text{ – абсолютна похибка).}$$



Вибравши  $\varepsilon_1 = 0,5$  (імовірна похибка знаходиться в межах  $\pm 0,5\sigma$ ), отримаємо такі результати: для рівня довіри  $\alpha = 0,85$  (імовірність похибки становить 0,15) необхідна і достатня кількість експертів – 7 осіб, для рівня довіри  $\alpha = 0,95$  (імовірність похибки становить 0,05) – 15 осіб (таблиця 2.1).

Таблиця 2.1 – Визначення необхідної й достатньої кількості експертів

$\varepsilon_1$	$\alpha$	N
0,5	0,85	7
0,5	0,95	15

Не всі фахівці дотримуються статистичного підходу до формування експертної групи. Часто акцент робиться на якості експертів.

Питання про оцінювання компетентності експертів не менш складне. Успішність участі в попередніх експертизах – хороший критерій для діяльності дегустатора, лікаря, судді в спортивних змаганнях, тобто таких експертів, які беруть участь у тривалих серіях однотипних експертиз. Проте найцікавішими й важливішими є унікальні експертизи великих проектів, що не мають аналогів. Використання формальних показників експертів (посада, науковий ступінь і вчене звання, стаж, кількість публікацій та ін.) може мати лише допоміжний характер, хоча такі показники найпростіше застосовувати під час формування експертної групи. Часто пропонують використовувати методи самооцінювання й взаємооцінювання компетентності експертів. Проте під час самооцінювання компетентності експерт оцінює ступінь своєї самовпевненості, а не реальну компетентність. Саме поняття „компетентність” не має строгого означення. Його визначають як виражену здатність застосовувати свої знання й уміння. Іноді компетентність подають як комбінацію теоретичного знання, практичного досвіду й персональних якостей.

Інтерес також становить тестовий метод оцінювання компетентності експертів – за допомогою тестових випробувань фахівців.

Самооцінювання експертів проводиться за допомогою визначення коефіцієнта компетентності експертів

$$k_{ki} = \frac{k_{ai} + k_{оби}}{2},$$

де  $k_{ai}$  – коефіцієнт аргументації;

$k_{оби}$  – коефіцієнт обізнаності (експерт визначає його самостійно у межах від 0 до 1).

Коефіцієнт аргументації визначається за таблицею 2.2. Експерт заповнює таблицю, позначаючи знаком „+” чи будь-яким іншим знаком

ступінь вагомості для нього різних джерел аргументації (висока, середня, низька).

Таблиця 2.2 – Розрахунок коефіцієнта аргументації

Джерело аргументації	Вагомість джерел		
	Висока	Середня	Низька
Теоретичний аналіз, проведений дослідником	+		
Практичний досвід		+	
Узагальнення робіт вітчизняних авторів	+		
Узагальнення робіт зарубіжних авторів		+	
Особисте ознайомлення зі станом справ за рубежом			+
Інтуїція		+	

Особа, що проводить експертне опитування, розшифровує дані відповідно до таблиці 2.3 (коефіцієнт аргументації розраховується як сума клітинок таблиці, позначених експертом).

Для прикладу, наведеного в таблиці 2.2,  $k_{ai} = 0,85$ .

Таблиця 2.3 – Дані для розрахунку коефіцієнта аргументації

Джерело аргументації	Вагомість джерел		
	Висока	Середня	Низька
Теоретичний аналіз, проведений дослідником	0,3	0,2	0,1
Практичний досвід	0,5	0,4	0,2
Узагальнення робіт вітчизняних авторів	0,05	0,05	0,05
Узагальнення робіт зарубіжних авторів	0,05	0,05	0,05
Особисте ознайомлення зі станом справ за рубежом	0,05	0,05	0,05
Інтуїція	0,05	0,05	0,05

Репрезентативність (якість) експертної групи в цілому обчислюється за формулою

$$W = \frac{1}{n} \sum_{i=1}^n k_{ki}.$$

Умовою репрезентативності групи експертів є  $0,67 \leq W \leq 1$ .

## 2.4 Методи експертного оцінювання

При **ранжуванні** експерт повинен розмістити варіанти (фактори, моделі, об'єкти тощо) у порядку, який вважає раціональним, і присвоїти кожному з них числа натурального ряду – ранги 1, 2, ...,  $n$ . Кількість рангів дорівнює кількості варіантів. Якщо експерт надає двом і більше варіантам однакові ранги, то кожному з цих варіантів присвоюється

середній ранг, розрахований за відповідними числами натурального ряду.

Під час обґрунтування складних управлінських рішень в умовах невизначеності, довгострокового прогнозування розвитку науки, техніки, економіки використовують групові експертизи. Надійність групових оцінок залежить від узгодженості думок експертів, що потребує відповідного статистичного оброблення інформації.

При **груповій експертизі** ( $n$  експертів) для кожного  $i$ -го варіанта визначається сума рангів  $\Sigma R_i$ , за якою упорядковуються варіанти. Скажімо, перший – найвищий – ранг надається варіанту, який набирає найменшу суму рангів, а останній – варіанту з найбільшою сумою рангів. Результати опитування експертів оформляються у вигляді матриці.

Наприклад, за даними ранжування трьох варіантів п'ятьма експертами (таблиця 2.4) перший ранг надається варіанту А, для якого  $\Sigma R_i = 6$ , другий – варіанту В, третій – варіанту С. Слід зазначити, що ранги визначають лише місця варіантів поміж іншими, не враховуючи існуючих між ними відстаней.

Таблиця 2.4 – Групова експертиза трьох об'єктів

Варіант	Експерт					Сума рангів	$d$	$d^2$
	1	2	3	4	5			
А	2	1	1	1	1	6	-4	16
В	1	2	3	2	2	10	0	0
С	3	3	2	3	3	14	4	16
Разом	X	X	X	X	X	30	X	32

Статистичне оброблення результатів ранжування передбачає оцінювання ступеня узгодженості думок експертів. Мірою узгодженості є коефіцієнт конкордації  $W$ , в основу розрахунку якого покладено відхилення  $d$  сум рангів за окремими варіантами  $\Sigma R_i$  від середньої суми рангів, яка становить  $\frac{1}{2} n (m + 1)$ . Коефіцієнт конкордації – це відношення суми квадратів названих відхилень  $S = \Sigma d^2$  до максимально можливої суми квадратів відхилень  $S_{max} = n^2 (m^3 - m) / 12$ . Якщо ранги не повторюються, то

$$W = \frac{12S}{n^2(m^3 - m)},$$

де  $m$  – кількість варіантів;

$n$  – кількість експертів.

При неузгодженості думок експертів  $W = 0$ . Чим вищий ступінь узгодженості, тим більше значення  $W$  наближається до одиниці. За даними таблиці 2.4 середня сума рангів становить  $30 : 3 = 10$ , сума квадратів відхилень  $S = 32$ , а коефіцієнт конкордації

$$W = (12 \cdot 32) / 5^2 (3^3 - 3) = 0,64,$$

що свідчить про певні розбіжності в оцінках експертів щодо значущості варіантів.

Перевірка істотності коефіцієнта конкордації  $W$  здійснюється за допомогою критерію  $\chi^2$  з  $(m - 1)$  числом ступенів вільності (свободи). Статистична характеристика критерію розраховується за формулою  $\chi^2 = W \cdot n \cdot (m - 1)$ . Для наведеного прикладу  $\chi^2 = 0,64 \cdot 5 \cdot (3 - 1) = 6,4$ , що перевищує критичне значення  $\chi^2(2) = 5,99$  (таблиця 2.5). Це дає підстави стверджувати з імовірністю 0,95, що значення  $W = 0,64$  не є випадковим.

Таблиця 2.5 – Критичні значення  $\chi_{1-\alpha}^2(k)$

$k$	1	2	3	4	5	6	7	8	9	10	11
$\alpha = 0,10$	2,71	4,61	6,25	7,78	9,24	10,64	12,02	13,36	14,68	15,99	17,28
$\alpha = 0,05$	3,84	5,99	7,81	9,49	11,07	12,59	14,07	15,51	16,92	18,31	19,67

При **попарних порівняннях** експерти використовують дві оцінки: 0 і 1. Більш вагомому варіанту надається оцінка 1, менш вагомому – 0. Результати попарних порівнянь оформляються у вигляді матриці, елементами якої є кількості наданих переваг  $a_{ij}$ . Діагональні елементи такої матриці зображено нулями. Однією із властивостей матриці є те, що  $a_{ij} + a_{ji} = n$ , де  $n$  – кількість експертів. За результатами опитування (див. таблицю 2.4) матриця кількості переваг набуває такого вигляду (таблиця 2.6).

Таблиця 2.6 – Матриця кількості переваг

Варіант	A	B	C	Разом	$\omega_j$
A	0	4	5	9	0,60
B	1	0	4	5	0,33
C	0	1	0	1	0,07
Разом	1	5	9	15	1,00

Відношення кількості наданих відповідному варіанту переваг до загальної суми елементів матриці характеризує його вагомість. За даними таблиці 2.6 найвагомим виявився варіант А, для якого  $\omega = 9 : 15 = 0,60$ .

Часто завданням експерта є не ранжування варіантів, а **безпосереднє оцінювання** рівнів певного явища або окремих його властивостей, наприклад, якості продукції, конкурентоспроможності фірм тощо. У таких ситуаціях спочатку визначається шкала (діапазон) оцінок, у межах якої експерт і оцінює явище (властивість) певним балом  $z_{ij}$ , де  $i$  – властивість,  $j$  – елемент сукупності.

Для певної множини  $m$  властивостей одного явища визначається середній бал  $G_j = \sum z_{ij} / m$ .

На таких методичних основах ґрунтується більшість рейтингових систем. Так, всесвітньо відома рейтингова система CAMEL, якою користуються органи нагляду за банківською діяльністю, має п'ятибальну шкалу оцінок: від 1 (добре) до 5 (незадовільно). Для кожного банку оцінюється достатність капіталу, якість активів, ефективність менеджменту, прибутковість і ліквідність балансу. Середній бал  $G_j$  є рейтингом фінансового стану  $j$ -го банку. Від його значення залежать ступінь втручання органів банківського нагляду і комплекс заходів щодо усунення недоліків.

Якщо властивості  $z_i$  нерівновагомі, то рейтинг визначається як середня арифметична зважена  $G_j = \sum z_{ij} \omega_i$ , де  $\omega_i$  – вага  $i$ -ї властивості. Саме так оцінюються комерційні, політичні ризики тощо. Наприклад, комерційний ризик, пов'язаний з інтернаціоналізацією банківської діяльності, оцінюється індексом Бері. Ознакова множина цього індексу містить 15 різновагомих показників, які характеризують політичну та економічну ситуацію в країні-партнерові, зокрема, політичну стабільність (вага 12 %), стан платіжного балансу (вага 6 %), темп економічного розвитку (вага 10 %) та ін. Сума ваг становить 100 %.

### Запитання для самоперевірки

- 1 Перелічіть основні типи й методи експертного оцінювання.
- 2 Опишіть процедуру виконання експертного опитування за методом Дельфі. Які переваги й недоліки має цей метод?
- 3 Яким чином забезпечується якість експертного опитування?
- 4 Групи респондентів здійснили ранжування дестабілізуючих факторів економіки (таблиця 2.7).

Таблиця 2.7 – Дестабілізуючі фактори економіки

Фактор	Ранги, надані		
	промисловцями	аграріями	гуманітаріями
Податки	1	1	1
Тіньова економіка	4	2	3
Законодавство	3	4	2
Державний борг	2	3	4
Вимоги Світового банку	5	5	5

Оцініть ступінь узгодженості думок респондентів, висновок зробіть з імовірністю 0,95.

5 Мета проекту – отримати максимальний прибуток. Можливі три стратегії досягнення цілі мети (таблиця 2.8). За даними матриці переваг, наданих експертами кожній стратегії, виконайте ранжування їх за вагомістю. Зробіть висновки.

Таблиця 2.8 – Стратегії досягнення мети

Стратегія	A	B	C
A	–	3	4
B	2	–	3
C	1	2	–

6 За наведеними нижче даними (таблиця 2.9) оцініть ризик підприємця, який планує вийти на ринок з новим товаром. Фактори ризику оцінювались експертами в діапазоні від 0 до 10 балів, вагові коефіцієнти – від 0 до 100 %. Межа мінімального ризику – 2,5 бала.

Таблиця 2.9 – Вихідні дані

Фактор ризику	Бал	Вага, %
Місткість ринку	2	20
Сталість попиту	5	20
Конкурентоспроможність товару	2	25
Фінансовий стан і кредитоспроможність	4	16
Якість роботи маркетингової служби	3	12
Імідж фірми	2	7

7 Проведіть експертне опитування групи експертів (5 – 7 осіб):

- сформулюйте запитання для експертів;
- сформулюйте експертну групу;
- розробіть анкету для проведення експертного оцінювання;
- проведіть опитування експертів;
- оцініть ступень узгодженості думок експертів;
- проаналізуйте узагальнений результат опитування експертів.

Як варіант опитування можна запропонувати експертам вибрати улюблені: 1) країну для проведення відпустки; 2) професію; 3) тварину; 4) колір; 5) марку автомобілю; 6) фільм; 7) породу собак; 8) стиль музики; 9) футбольну команду. Анкета повинна містити 5 – 6 варіантів об'єктів для порівняння.

### **3 СТАТИСТИЧНІ МЕТОДИ ДОСЛІДЖЕННЯ. ОСНОВИ РОБОТИ З СИСТЕМОЮ STATISTICA**

#### **3.1 Показники описової статистики**

Описова (дескриптивна) статистика є інструментом статистичного опису даних, що становлять всю сукупність у цілому. Мета описової статистики – отримання зведених (узагальнених) показників, що

характеризують початкову сукупність даних як генеральну (а не як вибірку з деякої іншої сукупності більшого обсягу).

Для числового оцінювання узагальнених показників сукупності використовуються так звані описові статистики, які є однозначними функціями на множині спостережуваних даних і визначають значення узагальнених показників сукупності, що оцінюються.

Показники, які обчислюються за допомогою описових статистик, можна розбити на три групи: показники розташування варіантів значень ознаки, варіації ознаки й особливостей форми її розподілу.

**1 Показники розташування** описують положення в первинному ряду даних тих або інших варіантів значень ознаки, що характеризують ряд. До них належать:

- максимальне  $x_{max}$  і мінімальне  $x_{min}$  значення ознаки;
- середня арифметична величина  $\bar{x}$  (статистична оцінка математичного сподівання  $M(x)$ );
- мода  $M_o$  – варіант значення ознаки, що найчастіше має місце, або той варіант, який відповідає максимальній ординаті емпіричної кривої розподілу;
- медіана  $M_e$  – серединне значення ранжованого ряду варіантів значень ознаки;
- нижні й верхні квартилі  $Q_1$  і  $Q_3$ , які обмежують центральну зону ранжованого ряду, до якої потрапляють 50 % варіантів значень ознаки: 25 % варіантів значень, менших за серединне значення  $M_e$ , і 25 % варіантів значень, більших за  $M_e$  (рисунок 3.1).

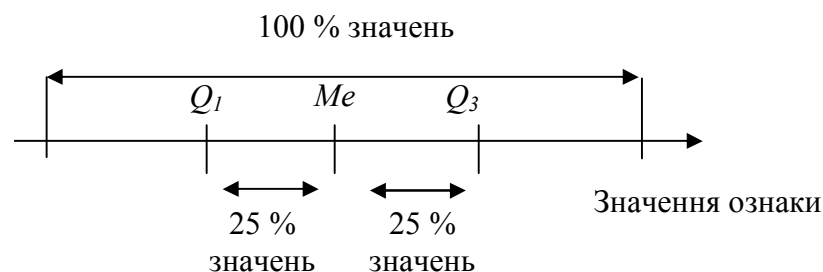


Рисунок 3.1 – Квартилі  $Q_1$  і  $Q_3$

Серед показників цієї групи найчастіше використовуються показники центра розподілу –  $\bar{x}$ ,  $M_o$  і  $M_e$ . При цьому  $\bar{x}$  розраховується для первинного ряду спостережуваних даних,  $M_o$  і  $M_e$  – для ранжованого (впорядкованого) ряду.

**2 Показники варіації** ознаки описують ступінь розкиду варіантів значень ознаки щодо свого центра  $\bar{x}$  (або  $M_e$ ). Розрізняють показники розміру та інтенсивності варіації. До показників розміру варіації відносять:

- розмах варіації  $R = x_{max} - x_{min}$ , що встановлює граничне значення амплітуди коливань ознаки;

- міжквартильний розмах  $Q_3 - Q_1$ , який визначає максимальну амплітуду коливань у центральній зоні ряду (обмеженій квартилями  $Q_1$  і  $Q_3$ );

- середнє лінійне відхилення  $\bar{d}$ , що обчислюється як середнє арифметичне значення абсолютних відхилень  $|x - x_i|$ :

$$\bar{d} = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{n};$$

- дисперсія  $\sigma^2$  (або  $D$ ), що розраховується як середнє арифметичне значення квадратів відхилень  $|x - x_i|$ :

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

- середнє квадратичне (стандартне) відхилення  $\sigma$ , що обчислюється як корінь квадратний з дисперсії  $\sigma^2$ :

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

Інтенсивність варіації ознаки вимірюється відносними показниками:

$$V_\sigma = \frac{\sigma}{x}, \quad V_d = \frac{d}{x}, \quad V_R = \frac{R}{x}, \quad V_{Me} = \frac{Me}{x}.$$

Показники  $R$ ,  $\bar{d}$  і  $\sigma$  є величинами іменованими і виражаються в тих самих одиницях, що й ознака, яка вивчається. Дисперсія  $\sigma^2$  вважається безрозмірною величиною. Відносні показники інтенсивності варіації зазвичай вимірюються у відсотках.

У статистичній практиці для оцінювання варіації найчастіше застосовуються показники розміру варіації  $\sigma^2$ ,  $\sigma$ .

Інтенсивність варіації зазвичай вимірюють коефіцієнтом варіації  $V_\sigma$ , який виражається у відсотках і обчислюється за формулою

$$V_\sigma = \frac{\sigma}{x} \cdot 100.$$

Величина  $V_\sigma$  оцінює інтенсивність коливань варіантів щодо їхньої середньої величини. Прийнято таку шкалу оцінювання інтенсивності коливань:

$0\% < V_\sigma \leq 40\%$  – інтенсивність незначна;

$40\% < V_\sigma \leq 60\%$  – інтенсивність середня (помірна);

$V_\sigma > 60\%$  – інтенсивність істотна.

Для нормальних і близьких до нормального розподілів показник  $V_\sigma$  є індикатором однорідності сукупності: прийнято вважати, що при виконанні нерівності  $V_\sigma < 33\%$  сукупність – кількісно однорідна за цією ознакою.



**3 Показники особливостей форми розподілу.** Для визначення типу закономірності емпіричного розподілу він приблизно описується відповідним теоретичним (імовірнісним) розподілом, форму кривої якого називають форматом розподілу. У тих випадках, коли форма розподілу аналізується на її близькість до нормальної форми, розбіжність між ними оцінюється показниками асиметрії і ексцесу.

За показниками асиметрії оцінюють зміщення ряду розподілу вліво або управо відносно осі симетрії нормального розподілу.

При симетричному розподілі максимальна ордината прямої розташовується точно в середині кривої (рисунок 3.2), а характеристики центра розподілу збігаються:

$$\bar{x} = Mo = Me.$$

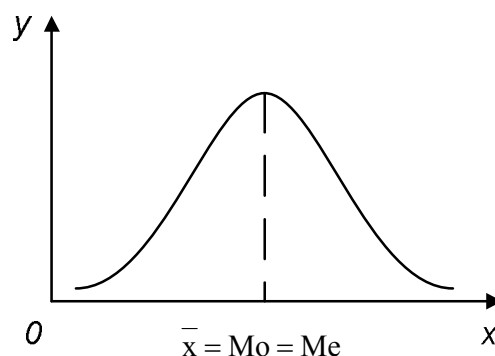


Рисунок 3.2 – Симетричний розподіл

У разі асиметричного розподілу вершина кривої знаходиться не в середині, а зміщена або вліво, або вправо.

Якщо вершина зміщена вліво, то права частина кривої є довшою від лівої (рисунок 3.3), тобто має місце правобічна асиметрія, яка характеризується нерівністю  $\bar{x} > Me > Mo$ , що означає переважну появу в розподілі більших значень ознаки.

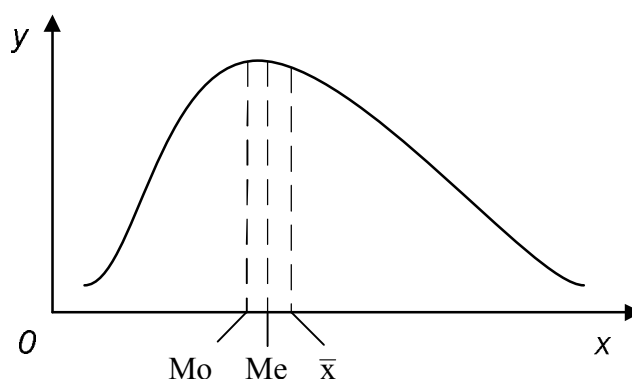


Рисунок 3.3 – Асиметричний розподіл – лівобічна асиметрія

Якщо ж вершина кривої зміщена вправо і ліва частина кривої є довшою за праву, то асиметрія – лівобічна і для неї справджується нерівність  $\bar{x} < Me < Mo$ . Це означає, що в розподілі частіше мають місце менші значення ознаки. На рисунку 3.4 наведено розподіли з лівобічною та правобічною асиметріями. Пунктиром показано форму кривої нормального розподілу.

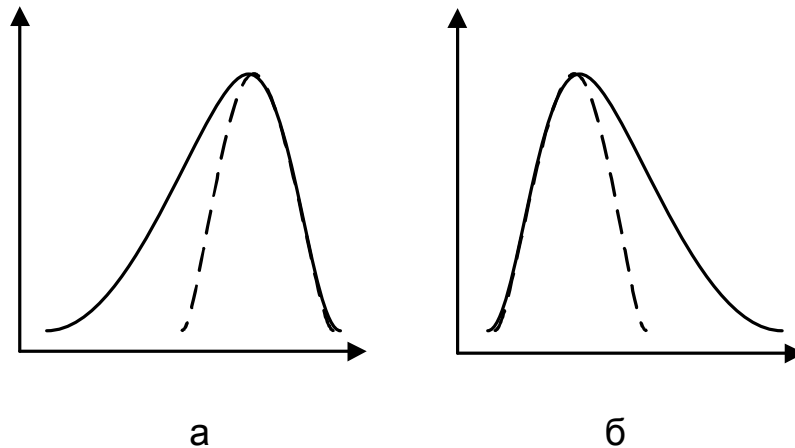


Рисунок 3.4 – Розподіли з лівобічною (а) та правобічною (б) асиметріями

Чим більша величина розбіжності між  $\bar{x}$ ,  $Me$ ,  $Mo$ , тим більш асиметричним є ряд.

Найточнішим показником асиметрії розподілу є коефіцієнт асиметрії  $As$ , що обчислюється за формулою

$$As = \frac{\sum_{i=1}^n (x_i - \bar{x})^3}{\sigma^3 \cdot n},$$

де  $n$  – кількість одиниць сукупності. Чим більша величина  $|As|$ , тим більш асиметричним є розподіл. Установлено таку оцінну шкалу асиметричності:

$|As| \leq 0,25$  – асиметрія незначна;

$0,25 < |As| \leq 0,5$  – асиметрія помітна (помірна);

$|As| > 0,5$  – асиметрія істотна.

Коефіцієнт  $|As|$  є відносною безрозмірною величиною.

Показник ексцесу характеризує крутість кривої розподілу – її загостреність або пологість порівняно з нормальною кривою (рисунок 3.5).

Для оцінювання розбіжностей ступеня крутості кривих (при однаковій силі варіації) застосовується коефіцієнт ексцесу

$$E = \frac{\sum_{i=1}^n (x_i - \bar{x})^4}{\sigma^4 \cdot n} - 3.$$

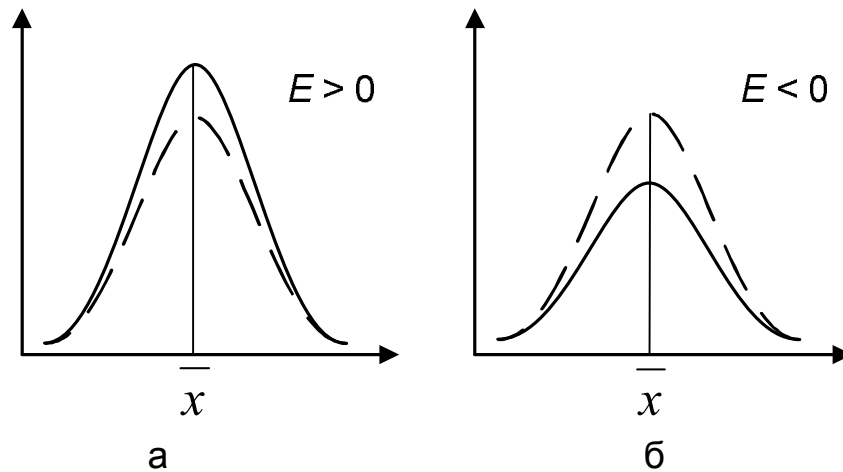


Рисунок 3.5 – Розподіли з ненульовим ексцесом (пунктиром показано форму нормального розподілу)

Зазвичай коефіцієнт ексцесу обчислюється тільки для симетричних або близьких до них розподілів. Це пояснюється тим, що за базу порівняння прийнято криву нормального розподілу, що є симетричною. При цьому:

- якщо  $E > 0$ , то вершина кривої розподілу розташовується вище за вершину нормальної кривої (рисунок 3.5, а). Це свідчить про скупчення значень ознаки в центральній зоні ряду розподілу, тобто про переважну появу в даних значень, близьких до середніх;
- якщо  $E < 0$ , то вершина кривої розподілу знаходиться нижче за вершини нормальної кривої (рисунок 3.5, б). Це означає, що значення ознаки не концентруються в центральній частині ряду, а достатньо рівномірно розсіяні в діапазоні від  $x_{max}$  до  $x_{min}$ .

Для нормального розподілу  $E = 0$ , тому чим більша абсолютна величина  $|E|$ , тим істотніше розподіл відрізняється від нормального. Зокрема, велика від'ємна величина  $E$  означає переважання в ознаці крайніх значень, причому одночасно і нижчих, і вищих. При цьому в центральній частині розподілу може утворитися „западина”, що перетворює розподіл на розподіл з двома вершинами ( $U$ -подібної форми), що є індикатором неоднорідності сукупності.

### 3.2 Формування інформаційної бази моделі

Масив первинних даних у системі *Statistica* організується й зберігається у вигляді електронної таблиці *Spreadsheet* з рядками *Cases* і стовпцями *Variables*, тобто по рядках електронної таблиці розміщуються елементи статистичної сукупності ( $j = 1, 2, \dots, n$ ), по стовпцях – ознаки ( $i = 1, 2, \dots, m$ ). Клітинки таблиці призначено для введення числової або текстової інформації. В системі реалізовано так званий механізм

„подвійного запису”, за яким встановлюється еквівалент: число = текстове значення. Наприклад, для ознаки „стать” 1 = чол., 2 = жін. Під час опрацювання інформації можна переключатися з одного типу даних на інший.

Уведення первинних даних у таблицю здійснюється різними способами, а саме:

- безпосередньо з клавіатури;
- перетворенням існуючого масиву даних за допомогою певних операцій (ранжування, стандартизації), математичних або статистичних функцій (ln, sin тощо);
- імпортуванням даних з інших додатків Windows, наприклад з Excel.

При формуванні нового файлу через команду *New* автоматично створюється таблиця розміром 10x10: по стовпцях – 10 ознак (VAR1, VAR2 ...) і по рядках – 10 спостережень. За допомогою кнопок *Vars* і *Cases* (або пункту меню *Insert*) можна додати або видалити стовпці й рядки (команди *Delete/Add Variables* і *Delete/Add Cases*).

Необхідним етапом формування інформаційної бази є специфікація ознак, тобто визначення основних параметрів кожної з них: імені, формату, коду для пропущених даних, формули.

Найпростіший спосіб надання специфікації ознакам – через команду *Specs* – *Специфікації* (кнопка *VARs*). Наведемо основні з них.

*Name* – ім'я, не більше восьми символів. У противному разі необхідно використати мітку *Label* у полі *Long Name*, наприклад, *Name: ВВП; Label: Валовий внутрішній продукт 2010 р. у поточних цінах, млрд грн*.

*MD code* – код, який присвоюється пропущеним даним. Можна вилучити дані з розрахунків, замінити їх середніми значеннями або інтерполювати. За замовчуванням *MD code* становить 9999.

*Decimals* – кількість розрядів після коми.

*Display Format* – спосіб подання ознак різних типів (числа, дати, час, науковий формат, грошовий формат, проценти).

Якщо значення певної ознаки необхідно розрахувати, то в текстовому полі *Long Name* задається формула, за якою ведеться розрахунок. Формула починається із знака «=». Якщо потрібно, у формулі використовуються математичні або статистичні функції – *Functions*.

Інші параметри електронної таблиці встановлюються за допомогою функціональних кнопок на панелі інструментів (імена спостережень (рядків) – за допомогою кнопки *Cases*). Після того як усі параметри таблиці задано, можна вводити дані.

Для роботи з даними в системі *Statistica* реалізовано всі стандартні операції методу *Drag-and-Drop*, зокрема:

- копіювання, переміщення, вставка;
- автозаповнення блоків з регулярною структурою;
- стандартизація даних;

- транспонування;
- ранжування даних;
- перекодування значень ознаки, перехід від однієї шкали вимірювання до іншої.

### 3.3 Описові статистики. Розвідувальний аналіз даних у Statistica

Статистичне оброблення даних у будь-якому модулі генерує велику кількість вихідної інформації у вигляді електронних таблиць *Scrollsheet* і графіків. Таблиці *Scrollsheet* підтримують усі стандартні операції з виділеними блоками значень (копіювання, переміщення, вставка, екстраполяція, стандартизація даних тощо). Таблицю *Scrollsheet* можна відредагувати, зберегти як файл результатів (з розширенням *scr*) або конвертувати в первинні дані (файл з розширенням *sta*), а також експортувати в інші *Windows*-додатки.

Діалогове вікно процедури (рисунок 3.6) пропонує вибрати тип аналізу.

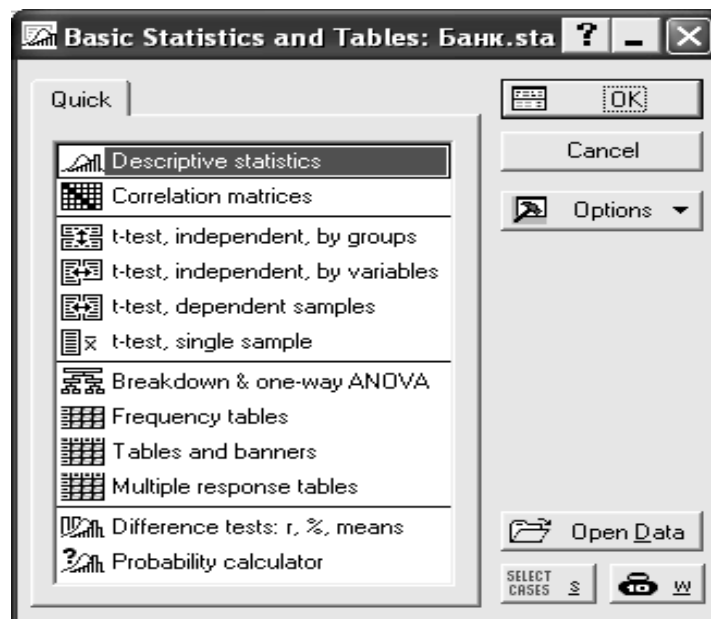


Рисунок 3.6 – Діалогове вікно *Basic Statistics/Tables*

**Описові статистики і графіки.** Програма обчислює описові статистики загального характеру: медіану, моду, квартилі, задані користувачем процентилі, середнє значення і стандартне відхилення, квартильний розмах, довірчі інтервали для середнього значення, асиметрію і ексцес (і їхні стандартні похибки), гармонійне й геометричне середні значення.

У таблиці 3.1 наведено результати обчислень для 25 комерційних банків (*Statistics – Basic Statistics/Tables – Descriptive Statistics –*

*Advanced*): *Mean* – середня величина, *Median* – медіана, *Minimum*, *Maximum* – мінімальне та максимальне значення, *Standard Deviation* – стандартне відхилення, *Skewness* – коефіцієнт асиметрії та *Kurtosis* – коефіцієнт ексцесу.

Таблиця 3.1 – Описові статистики

Показник	Mean	Median	Minimum	Maximum	Std.Dev.	Skewness	Kurtosis
Активи	26332613	17756642	7729003	79508914	19799383	1,186	0,609
Капітал	3140563	2246906	-2677417	16422065	3735416	2,059	5,983
Зобов'язання	23192050	15300592	6910870	70936104	16968193	1,260	1,086
Прибуток	-444748	-187364	-4603184	753673	991543	-3,272	13,398

**Кореляції** (*Statistics – Basic Statistics/Tables – Descriptive Statistics – Correlation matrice*). У системі є великий набір методів для дослідження кореляцій між змінними. У вікні *Product-Moment and Partial Correlation* вибираємо тип матриці – *One variable list* – *Один список ознак*, а у вікні вибору ознак – ознаки для аналізу. За командою на виконання процедури створюється нова таблиця, елементами якої є парні коефіцієнти кореляції (таблиця 3.2).

Коефіцієнти кореляції відображають наявність, напрям і силу лінійного зв'язку. Вони варіюються від -1 до +1.

Якщо коефіцієнт дорівнює -1, то це відповідає абсолютно різноспрямованій залежності (зі збільшенням однієї змінної інша зменшується) між аналізованими змінними; якщо +1 – відображає повну відповідність між змінними (тобто вони, по суті, є одним і тим же); якщо 0 – показує повну відсутність всякого зв'язку.

Таблиця 3.2 – Кореляції

Показник	Активи	Капітал	Зобов'язання	Прибуток
Активи	1,00	0,80	0,99	0,24
Капітал	0,80	1,00	0,71	0,50
Зобов'язання	0,99	0,71	1,00	0,17
Прибуток	0,24	0,50	0,17	1,00

Для зручності інтерпретації кореляцій застосовуються семантичні інтервали (таблиця 3.3).

Таблиця 3.3 – Градації коефіцієнтів кореляції

Значення коефіцієнта кореляції	Характеристика сили лінійного зв'язку
Від $\pm 0,81$ до $\pm 1,00$	Сильний
Від $\pm 0,61$ до $\pm 0,80$	Помірний
Від $\pm 0,41$ до $\pm 0,60$	Слабкий
Від $\pm 0,21$ до $\pm 0,40$	Дуже слабкий
Від $\pm 0,00$ до $\pm 0,20$	Немає кореляції

Існують два основних типи коефіцієнтів кореляції, які розраховуються залежно від виду шкали змінних, що беруть участь в аналізі:

1 Для змінних з інтервальною шкалою застосовується коефіцієнт кореляції Пірсона, який дає можливість охарактеризувати лінійний зв'язок між двома змінними

$$\rho = \frac{\sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y})}{(n-1)\sigma_x \sigma_y}.$$

2 Якщо хоча б одна з пари досліджуваних змінних має порядкову або дихотомічну шкалу, використовуються рангові коефіцієнти кореляції Спірмана або Кендала. В системі *Statistica* ці коефіцієнти розраховуються в модулі *Nonparametric Statistics* (рисунок 3.7).

**Т-критерії (та інші критерії групових відмінностей).** У цьому вікні можна обчислити Т-критерії для залежних і незалежних виборок.

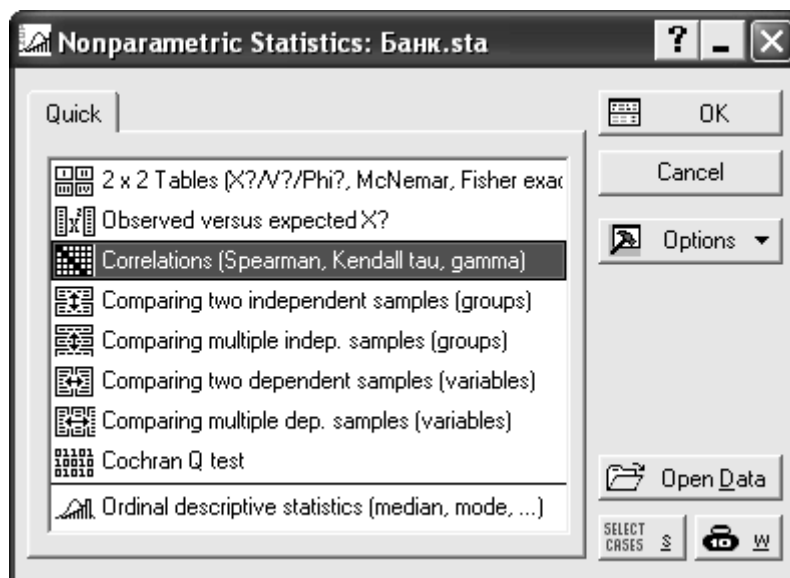


Рисунок 3.7 – Діалогове вікно *Nonparametric Statistics*

**Калькулятор імовірнісних розподілів.** Він підтримує стандартні розподіли: бета, Коші, хі-квадрат, експоненціальний, екстремальний (Гумбеля), F, гамма, Лапласа, лог-нормальний, логістичний, Парето, Релея, *t* (Стьюдента), Вейбулла і Z (нормальний).

Статистичне оброблення даних у будь-якому модулі генерує велику кількість вихідної інформації у вигляді електронних таблиць *Scrollsheet* і графіків. Таблиці *Scrollsheet* підтримують усі стандартні операції з виділеними блоками значень (копіювання, переміщення, вставка, екстраполяція, стандартизація даних тощо). Відредаговану числову, текстову та графічну інформацію можна експортувати в інші *Windows*-додатки, на принтер, у текстовий файл.

### 3.4 Однорідність і типологія

Однією з умов статистичного моделювання є однорідність сукупності. Лише в однорідній сукупності виявлені закономірності є сталими і їх можна застосувати до всіх одиниць сукупності.

Поняття *однорідності* пов'язують з наявністю в усіх одиницях сукупності таких спільних властивостей і рис, які визначають їх належність до одного й того ж типу. Оцінювання ступеня однорідності здійснюється за допомогою критеріїв математичної статистики, більшість з яких орієнтовано на аналіз форми одновершинних розподілів.

Однорідними вважаються сукупності, яким властивий симетричний, нормальний розподіл. Звісно, в соціально-економічних явищах нормальний розподіл у чистому вигляді не існує. Але він близький до інших одновершинних розподілів, його часто використовують як перше наближення під час моделювання. Деякі одновершинні розподіли зводяться до нормального вигляду перетворенням значень ознак, скажімо, заміною їх логарифмами. Логонормальною кривою можна описати низку асиметричних розподілів, передусім з правобічною асиметрією.

Основні властивості нормального розподілу:

- крива розподілу симетрична відносно максимальної ординати, яка відповідає середньому арифметичному значенню  $\bar{x}$ ;
- у межах  $\bar{x} \pm \sigma$  міститься 68,3 % усіх частот ряду розподілу, в межах  $\bar{x} \pm 2\sigma$  – 95,4 % частот, у межах  $\bar{x} \pm 3\sigma$  – 99,7 % частот;
- співвідношення стандартного відхилення  $\sigma$  і середнього модуля

відхилень  $\bar{l}$  становить  $\frac{\bar{l}}{\sigma} = \sqrt{\frac{2}{\pi}} = 0,8$ , або  $\frac{\sigma}{\bar{l}} = 1,25$ . Його значення

залежить від наявності в сукупності нетипових аномальних спостережень і може бути індикатором її «засміченості»;

- коефіцієнт асиметрії розподілу дорівнює нулю, коефіцієнт ексцесу – трьом.

Завдяки цим властивостям нормальна крива застосовується як *стандарт* і відіграє значну роль під час використання методів вибіркового, регресійного, факторного аналізу. Оцінювання ступеня наближення до цього стандарту ґрунтується на порівнянні емпіричних  $f_i$  і теоретичних  $\tilde{f}_i$  частот розподілу, де  $i$  – номер інтервалу. Теоретичні частоти визначають за формулою

$$\tilde{f}_i = n |F_{x_i} - F_{x_{i-1}}|,$$

де  $F_x = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{z^2}{2}} dz$  – інтегральна функція нормального розподілу;



$$z = \frac{x - \bar{x}}{\sigma} - \text{стандартизоване значення ознаки } x;$$

$$n = \sum f_i - \text{обсяг сукупності.}$$

З-поміж критеріїв, що застосовуються для оцінювання узгодженості емпіричного розподілу з нормальним, найпоширенішими є критерії Пірсона  $\chi^2$  та Колмогорова – Смирнова  $d$ .

Фактичне значення критерію  $\chi^2$  порівнюється з критичним значенням  $\chi^2_{1-\alpha}(k)$  (для ймовірності  $1 - \alpha$  і відповідної кількості ступенів свободи  $k = m - q - 1$ , де  $m$  – кількість інтервалів групування,  $q$  – кількість параметрів функції (для нормального розподілу  $q = 2$ )). Якщо  $\chi^2 < \chi^2_{1-\alpha}(k)$ , відхилення емпіричного розподілу від нормального є неістотним. У разі, коли фактичне значення критерію перевищує критичне, гіпотеза про узгодженість розподілів відхиляється. Критичні значення критерію  $\chi^2$  наведено в таблиці 3.4.

Таблиця 3.4 – Критичні значення  $\chi^2_{1-\alpha}(k)$

$k$	1	2	3	4	5	6	7	8	9	10	11
$\alpha = 0,10$	2,71	4,61	6,25	7,78	9,24	10,64	12,02	13,36	14,68	15,99	17,28
$\alpha = 0,05$	3,84	5,99	7,81	9,49	11,07	12,59	14,07	15,51	16,92	18,31	19,67

Висновки за критерієм  $\chi^2$  значною мірою залежать від кількості груп, частота яких має бути не менше п'яти. Крім того,  $\chi^2$  не враховує послідовність знаків відхилень частот, за наявності серій знаків надійність висновку зменшується. Цих вад позбавлений критерій Колмогорова – Смирнова  $d$ . Критичні значення  $d$  наведено в таблиці 3.5.

Таблиця 3.5 – Критичні значення критерію Колмогорова – Смирнова

Обсяг вибірки $n$	$\alpha = 0,10$	$\alpha = 0,05$	Обсяг вибірки $n$	$\alpha = 0,10$	$\alpha = 0,05$
5	0,510	0,563	15	0,304	0,338
6	470	521	16	295	328
7	438	486	17	286	318
8	411	457	18	278	309
9	388	432	19	272	301
10	368	409	20	264	294
11	352	391	25	240	264
12	338	375	30	220	242
13	325	361	$n > 30$	$\frac{1,22}{\sqrt{n}}$	$\frac{1,36}{\sqrt{n}}$

У системі *Statistica* закономірність одновимірного розподілу можна аналізувати за допомогою процедур *Distribution* стартової панелі *Descriptive statistics* (модуль *Basic Statistics and Tables*). Розподіл сукупності за варіаційною ознакою подається у вигляді таблиці *Frequency Table*, де вказуються інтервали групувань, частоти розподілу за інтервалами, а також кумулятивні частоти. Опція *Normal expected frequencies* додає в таблицю теоретичні частоти (групові й кумулятивні). Перевірка на нормальність розподілу здійснюється за опцією *K-S and Lilliefors test for normality*.

У таблиці 3.6 наведено розподіл 25 комерційних банків за розміром власного капіталу.

У верхньому лівому куті таблиці подано значення критерію  $d = 0,24636$ . Значення  $d$  менше за критичне ( $0,264$ ) при  $\alpha = 0,05$ , отже, гіпотеза про нормальний розподіл комерційних банків за рівнем власного капіталу не відхиляється.

Таблиця 3.6 – Характеристики розподілу (власний капітал)

Frequency table: Капітал: Власний капітал усього, 1-й квартал 2009 р. (Банк.sta) K-S $d=,24636$ , $p<,10$ ; Lilliefors $p<,01$						
Category	Count	Cumulative	Percent	Cumul %	% of all	Cumulative %
-5E6 <x<=0,000000	2	2	8,00	8,00	8,00	8,00
0,000000<x<=5000000,	17	19	68,00	76,00	68,00	76,00
5000000,<x<=1E7	5	24	20,00	96,00	20,00	96,00
1E7 <x<=15E6	0	24	0,000	96,00	0,00	96,00
15E6 <x<=2E7	1	25	4,00	100,00	4,00	100,00
Missing	0	25	0,00		0,00	100,00

Узгодженість емпіричного розподілу комерційних банків за рівнем власного капіталу з нормальним розподілом показано на рисунку 3.8.

Якщо необхідно перевірити гіпотезу про узгодженість даних з іншими розподілами (лог-нормальним, експоненціальним тощо), використовується модуль *Nonparametrics / Distribution – Непараметричні статистики / Розподіли*.

**Стандартизація.** Безпосереднє використання абсолютних значень ознак під час аналізу може призвести до того, що класифікацію визначатимуть ознаки, що мають найбільший розкид значень. Тому застосовуються такі види стандартизації:

1 Z-шкали (Z-Scores). Із значень змінних віднімається їхнє середнє значення, і ці значення діляться на стандартне відхилення.

2 Розкид від -1 до 1. Лінійним перетворенням змінних добиваються розкиду значень від -1 до 1.

3 Розкид від 0 до 1. Лінійним перетворенням змінних добиваються розкиду значень від 0 до 1.

4 Максимум 1. Значення змінних діляться на їхній максимум.

- 5 Середнє 1. Значення змінних діляться на їхнє середнє значення.  
 6 Стандартне відхилення 1. Значення змінних діляться на стандартне відхилення.

Histogram: Капітал: Власний капітал усього, 4й квартал 2009 р.

K-S d=,24636, p<,10 ; Lilliefors p<,01

— Expected Normal

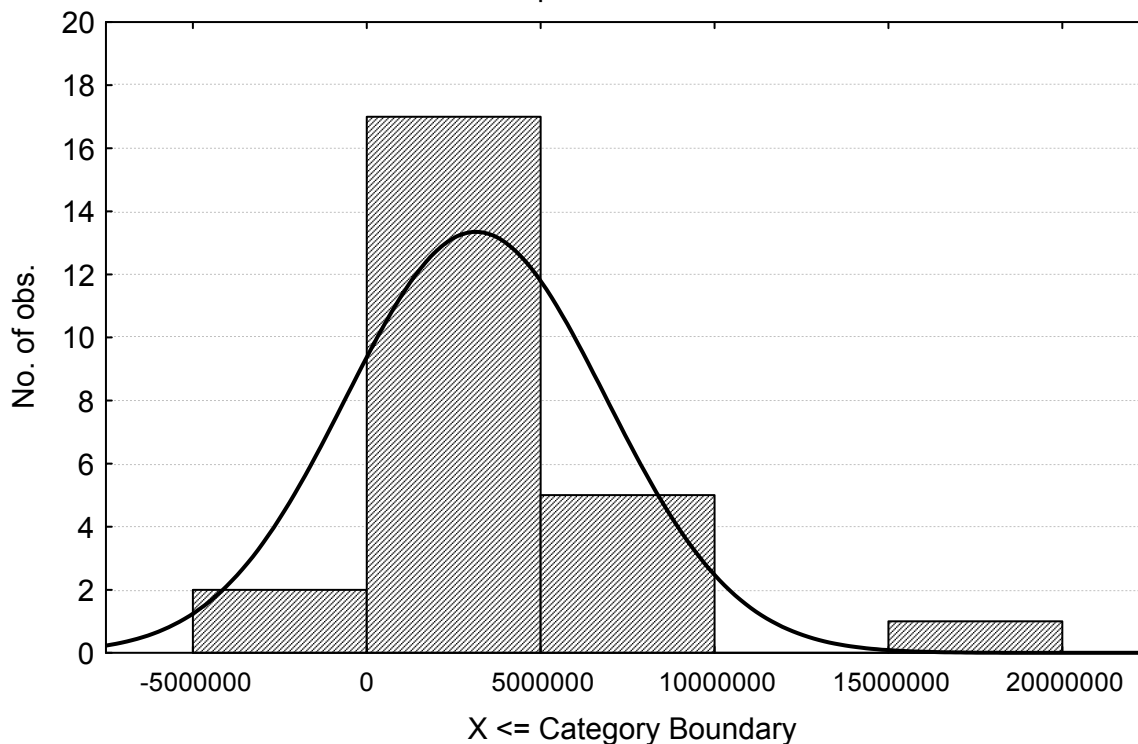


Рисунок 3.8 – Емпіричний і теоретичний розподіли комерційних банків за рівнем власного капіталу

### Запитання для самоперевірки

- 1 Перелічіть і охарактеризуйте описові статистики розташування варіантів значень ознаки.
- 2 Перелічіть і охарактеризуйте описові статистики варіації ознаки.
- 3 Перелічіть і охарактеризуйте описові статистики особливостей форми розподілу.
- 4 Назвіть основні властивості нормального розподілу.
- 5 Виробничі потужності 12 металургійних комбінатів характеризуються такими даними, млн т/рік:

Продукція	Максимальний рівень	Мінімальний рівень	Середній рівень	Середнє квадратичне відхилення
Чавун	10,4	3,2	8,5	1,3
Сталь	17,6	4,4	9,8	2,2
Прокат	12,4	3,7	7,6	1,9

За допомогою критерію Граббса перевірте однорідність сукупності металургійних комбінатів за виробничими потужностями. Висновок зробіть з імовірністю 0,95.

Істотність відхилень від основної маси можна оцінити за допомогою критерію Граббса  $T_n$ , статистичною характеристикою якого є стандартизоване граничне відхилення аномального значення  $x_n$  ( $x_1$ ) від середнього  $\bar{x}$ . Якщо йдеться про максимальне значення  $x_n$ , то

$$T_n = \frac{x_n - \bar{x}}{\sigma},$$

де  $\bar{x}$  і  $\sigma$  визначаються для сукупності в цілому.

Критичні значення  $T_{1-\alpha}(n)$ , де  $n$  – обсяг сукупності, наведено в Додатку А. Якщо фактичне значення  $T_n$  менше за критичне, то відхилення певного значення від основної маси з імовірністю  $(1 - \alpha)$  є випадковим, неістотним, а якщо  $T_n$  перевищує критичне, то відхилення є істотним, а отже, значення – аномальним і нетиповим для сукупності в цілому. В такому разі це значення вилучається і критерій застосовується до  $x_{n-1}$  і т. д., доти, доки не буде визнано, що сукупність однорідна.

6 Розподіл 400 домогосподарств за рівнем середнього доходу на душу населення характеризується такими даними:

Номер групи	Частка емпіричного розподілу, %	Імовірність теоретичного розподілу, %	
		Нормального	Лог-нормального
1	5,5	6,7	5,3
2	9,4	8,0	11,6
3	17,2	12,7	16,3
4	15,3	16,6	17,6
5	15,0	17,8	15,0
6	13,0	15,5	12,7
7	10,2	11,2	9,0
8	8,5	6,6	6,5
9	3,8	3,1	4,0
10	2,0	1,8	2,0
Разом	100	100	100

За допомогою критеріїв  $\chi^2$  та Колмогорова – Смирнова  $d$  перевірте, чи узгоджується розподіл домогосподарств за рівнем середнього доходу на душу населення з нормальним або лог-нормальним розподілом. Висновок зробіть з імовірністю 0,95.

7 Для аналізу фінансового стану банків самостійно визначіть сукупність банків (не менше 25) і для кожного з них випишіть інформацію щодо розміру активів, капіталу, зобов'язань і прибутку (за даними офіційного сайту Національного банку України [www.bank.gov.ua](http://www.bank.gov.ua)). Для

кожного банку визначить: прибутковість активів і капіталу, співвідношення капіталу і зобов'язань. У системі *Statistica* сформуєте файл первинних даних, здійсніть специфікацію показників.

8 Використовуючи процедури модуля *Basic Statistics and Tables*, здійсніть розвідувальний аналіз даних завдання 7: а) для показників прибутковості активів і капіталу визначить характеристики розподілу: *Mean* – середню величину, *Median* – медіану, *Standard Deviation* – стандартне відхилення, *Skewness* – коефіцієнт асиметрії та *Kurtosis* – коефіцієнт ексцесу; б) складіть ряд розподілу банків за рівнем прибутковості активів (*Frequency tables*), здійсніть частотний аналіз розподілу; в) для візуалізації розподілу банків за всіма показниками скористайтеся графіками.

9 За процедурою *Correlation matrices* оцініть взаємозв'язки між показниками прибутковості й співвідношенням капіталу і зобов'язань (завдання 7), подайте їх у матричному вигляді та графічно. Зробіть висновки.

## 4 МЕТОДИ БАГАТОВИМІРНОЇ КЛАСИФІКАЦІЇ

Соціальні та економічні об'єкти зазвичай характеризуються чималою кількістю параметрів, які утворюють багатовимірні вектори, і особливого значення в економічних і соціальних дослідженнях набувають завдання вивчення взаємозв'язків між компонентами цих векторів, причому ці взаємозв'язки необхідно виявляти на основі обмеженої кількості багатовимірних спостережень.

Поділ сукупності на однорідні класи (групи, кластери) називають *класифікацією*. Ідея класифікації ґрунтується на поняттях подібності й відмінності. Методологічний принцип класифікації містить два фундаментальних положення:

- в один клас об'єднуються подібні, схожі між собою одиниці сукупності;
- ступінь подібності, схожості одиниць, які належать до одного класу, вища, ніж ступінь подібності одиниць, віднесених до різних класів.

Оцінювання подібності здійснюється на основі однієї або декількох ознак, які, на думку експертів, формують „образ класу”. В традиційній схемі класифікації ці ознаки ієрархічно впорядковуються за своєю вагомістю, наприклад, класифікація шахт – за гірничо-геологічними умовами виробництва: потужністю пласта, нахилом його залягання, глибиною розробки лав, загазованістю лав і т. д. Саме так будується більшість комбінаційних групувань. На кожному кроці поділу сукупності до уваги береться лише одна ознака, тобто відбуваються послідовне формування, покрокове уточнення, детальніший опис класів. У невеликих за обсягом сукупностях можливості використання такої схеми класифікації обмежені.

У другій схемі класифікації використовується множина

класифікаційних ознак одночасно. Будь-яка одиниця сукупності, описана множиною ознак, геометрично інтерпретується як точка у багатовимірному просторі, а близькість двох точок розглядається як їх подібність, однорідність. Існують різні варіанти реалізації багатовимірної схеми класифікації. Їх можна об'єднати в два блоки:

- конструювання багатовимірних інтегральних оцінок (індексів, рейтингів), на основі яких проводиться класифікація за традиційною схемою;
- автоматична багатовимірна класифікація методами кластерного аналізу, коли поняття однорідності задається певними метриками.

Слід зазначити, що класифікація за будь-якою схемою є певною мірою суб'єктивною, оскільки її результати визначаються передусім множиною класифікаційних ознак та їхніми розмежувальними властивостями.

Багатовимірним статистичним аналізом називають розділ математичної статистики, що вивчає методи збору й оброблення багатовимірних статистичних даних, їх систематизації з метою виявлення характеру й структури взаємозв'язків між компонентами досліджуваної багатовимірної ознаки, отримання практичних висновків.

Методи зниження розмірності багатомірного простору дають можливість без істотної втрати інформації перейти від початкової системи великої кількості досліджуваних взаємозв'язаних чинників до системи істотно меншої кількості прихованих чинників, що визначають варіацію первинних ознак.

Методи багатовимірної класифікації призначено для поділення сукупностей об'єктів (що характеризуються великою кількістю ознак) на класи, до кожного з яких мають входити об'єкти, в певному значенні однорідні або близькі. Таку класифікацію можна провести методами кластерного й дискримінантного аналізу.

#### 4.1 Кластерні процедури класифікації

*Кластер* – це група, клас однорідних одиниць сукупності. Основне завдання кластерного аналізу – формування таких груп у багатовимірному просторі. Однорідність сукупності задається правилом обчислення певної метрики, що характеризує ступінь подібності (схожості)  $j$ -ї та  $k$ -ї одиниць сукупності. Такою метрикою може бути відстань між ними  $c_{jk}$  або коефіцієнт подібності  $r_{jk}$ . Близькі (схожі) за вибраними метриками одиниці вважаються такими, що належать до одного типу, тобто однорідними. Вибір метрики є важливим моментом кластерного аналізу, від якого залежить кінцевий варіант поділу сукупності на класи.

**Приклад 4.1** Двадцять банків, акції яких котируються на ринку, надали інформацію, наведену в таблиці 4.1, де  $X$  – витрати за минулий

період,  $Y$  – прибуток за минулий період. Необхідно з'ясувати, акції яких банків має сенс придбати (Buy), яких – притримати (Hold), а від яких – позбавитися (Sell).

Таблиця 4.1 – Вихідні дані (витрати й прибуток комерційних банків)

Номер банку	$X$	$Y$	Номер банку	$X$	$Y$
1	4	2	11	5	4
2	6	10	12	3	8
3	5	7	13	13	5
4	12	3	14	15	3
5	17	4	15	5	9
6	3	10	16	8	3
7	6	1	17	6	4
8	6	3	18	14	3
9	15	1	19	4	8
10	15	4	20	4	3

Графічне подання вихідних даних показано на рисунку 4.1 (графік побудовано за допомогою команди *Scatterplots* меню *Graphs*).

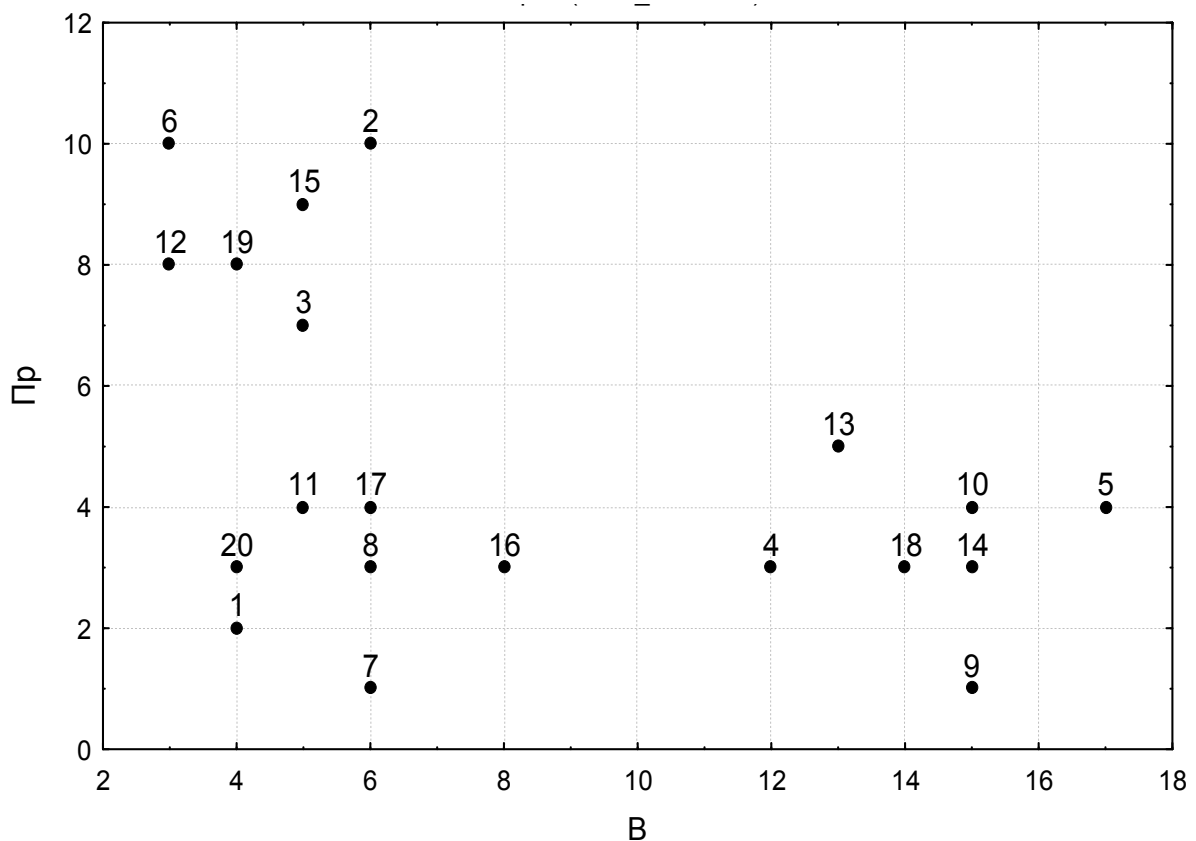


Рисунок 4.1 – Графічне подання початкових даних

Кластерний аналіз у системі *Statistica* здійснюється за допомогою пункту меню *Статистика – Методи багатовимірного аналізу – Кластерний аналіз* (рисунок 4.2).

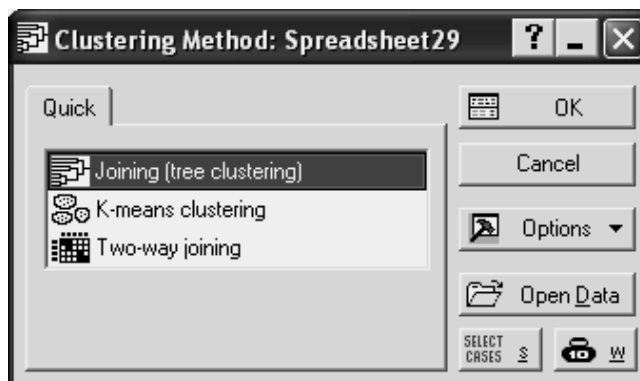


Рисунок 4.2 – Вибір типу кластерного аналізу

**Побудова деревоподібної діаграми (*Joining (tree clustering)*).** У системі *Statistica* ієрархічну процедуру класифікації – *Joining (Tree clustering)* – реалізовано в модулі *Cluster Analysis*. Діалогове вікно процедури (рисунок 4.3) пропонує вибрати:

- ознакову множину;
- тип вихідних даних: *Raw Data* – дані типу «об’єкт – ознака» або *Distance Matrix* – матриця відстаней;
- варіант класифікації: по стовпцях (*columns*) – класифікацію ознак або по рядках (*rows*) – класифікацію об’єктів;
- алгоритм об’єднання – *Amalgamation (linkage) Rules*; за замовчуванням – алгоритм одиничного зв’язку – *Single linkage (nearest neighbor)*;
- метрику відстаней – *Distance measure: Euclidean distances, City-block (Manhattan) distance* тощо.

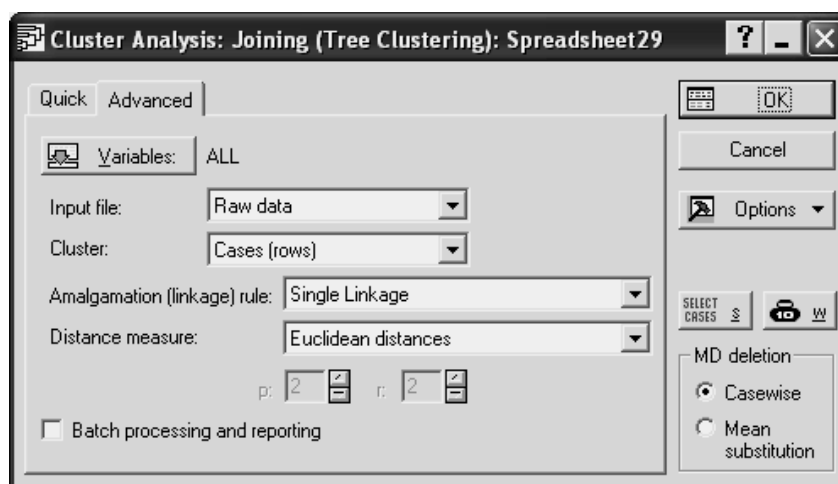


Рисунок 4.3 – Діалогове вікно процедури *Joining (tree clustering)*



За командою на виконання система видає *Joining Results* з опціями вибору типу дендрограми – горизонтальної або вертикальної (рисунки 4.4, 4.5).

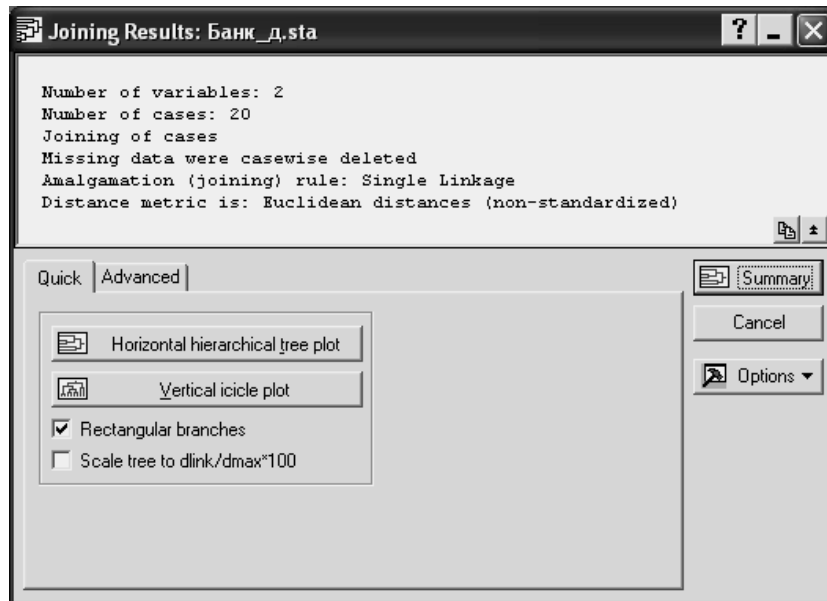


Рисунок 4.4 – Діалогове вікно *Joining Results*

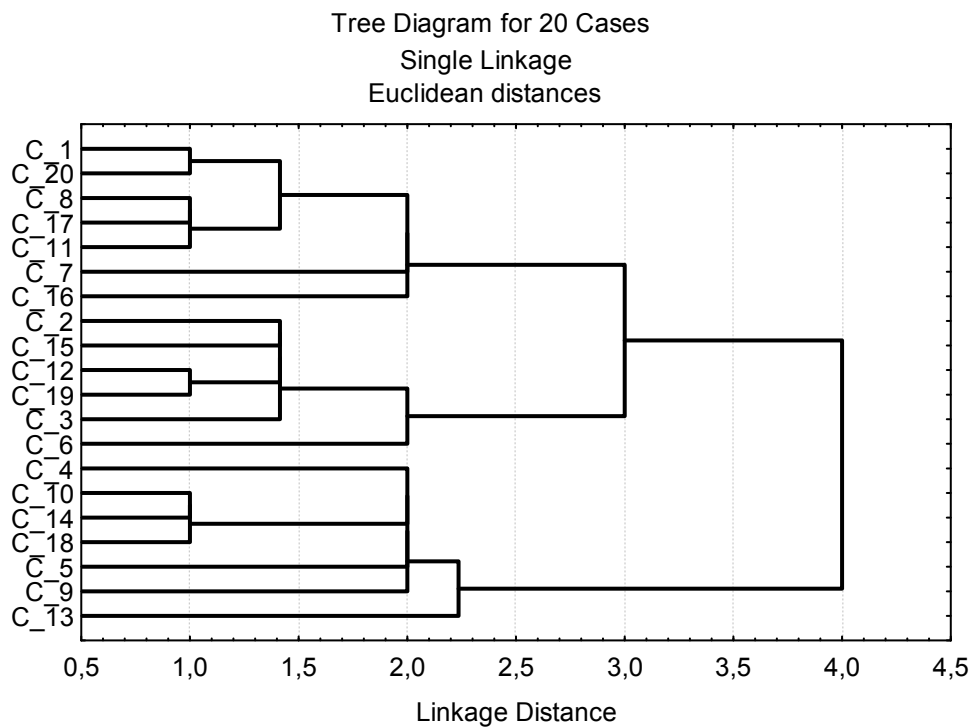


Рисунок 4.5 – Результати *Joining (tree clustering)*

На рисунку 4.5 показано результати кластерного аналізу. Відокремлено три кластери (до першого ввійшли банки 1, 20, 8, 17, 11, 7, 16; до другого – банки 2, 15, 12, 19, 3, 6; до третього – банки 4, 10, 14, 18, 5, 9, 13).

Узагалі ієрархічна кластер-процедура досить проста й прийнятна для інтерпретації. Проте для численної сукупності вона виявляється громіздкою. У таких випадках перевагу віддають ітераційним процедурам.

**Метод  $k$ -середніх ( $k$ -means clustering).** На відміну від ієрархічної процедури, яка потребує розрахунку й збереження матриці подібності, ітераційна процедура оперує безпосередньо первинними даними; формуються кластери одного рангу, ієрархічно не підпорядковані. Основні риси ітераційних кластер-процедур розглянемо на прикладі алгоритму  $k$ -середніх (рисунок 4.6), який реалізує ідею утворення груп за принципом «найближчого центра».

На першому кроці ітераційного процесу здійснюється орієнтовний поділ сукупності на класи і визначаються центри ваги (багатовимірні середні) цих класів.

На другому кроці визначаються евклідові відстані одиниць сукупності до центрів ваги виділених кластерів, і кожна з них належить до того кластера, центр ваги якого найближчий.

На третьому кроці розраховуються нові центри ваги кластерів.

Кроки 2 і 3 повторюються доти, доки склад кластерів не стабілізується. Ітерації за принципом  $k$ -середніх у явному вигляді не використовують критеріїв якості класифікації, проте вони неявно мінімізують внутрішньогрупові дисперсії, забезпечуючи тим самим однорідність сформованих кластерів.

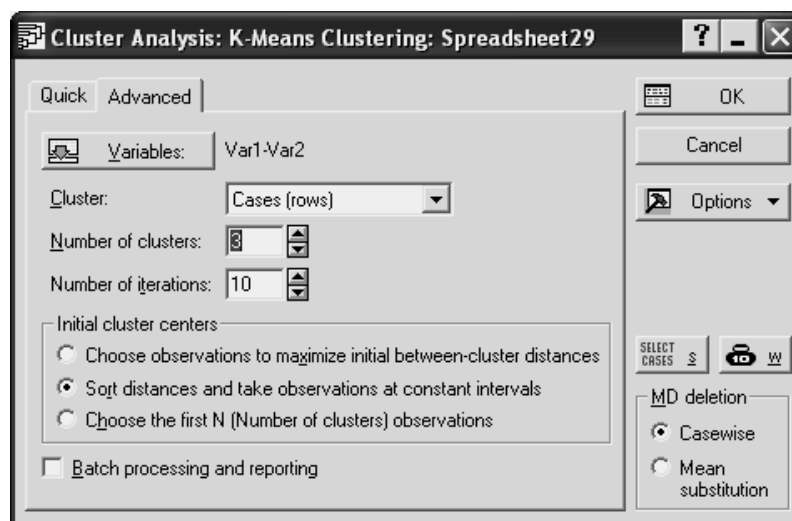


Рисунок 4.6 – Діалогове вікно процедури  $k$ -means clustering

Діалогове вікно процедури (див. рисунок 4.6) пропонує вибрати:

- ознакову множину;
- варіант класифікації: по стовпцях (*columns*) – класифікацію ознак або по рядках (*rows*) – класифікацію об'єктів;
- кількість кластерів;
- кількість ітерацій;
- метод визначення центра кластерів та ін.

За командою на виконання система видає вікно *k-means clustering Results* (рисунок 4.7).

Вкладка *Members of each cluster & distances* демонструє інформацію щодо складу кожного з кластерів і відстаней членів кластера від кластерного центра (рисунок 4.8).

За результатами аналізу отримано перший (банки 4, 5, 9, 10, 13, 14, 15); другий (банки 1, 7, 8, 11, 16, 17, 20) і третій (банки 2, 3, 6, 12, 15, 19) кластери.

За допомогою процедури класифікації отримано три групи банків: для першої характерним є високий прибуток при малих витратах (акції цих банків доцільно купувати (*Buy*)), для другої – середні величини прибутку та витрат (рекомендація *Hold*), для третьої – високі витрати та невеликий прибуток (рекомендація *Sell*).

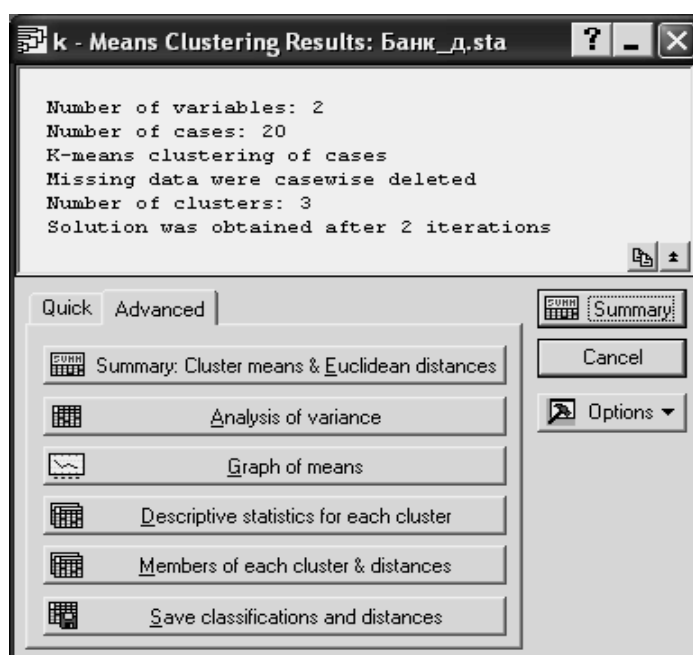


Рисунок 4.7 – Результати k-means clustering

Case No.	Case No.	Case No.	Case No.	Case No.	Case No.	Case No.	
C 4	C 5	C 9	C 10	C 13	C 14	C 18	
Distance	1,729103	1,887121	1,665986	0,646813	1,577909	0,451754	0,364216

Рисунок 4.8 – Склад кластерів

Рекомендації *Buy*, *Hold*, *Sell* наведено в таблиці 4.2.

Таблиця 4.2 – Рекомендації

Номер банку	X	Y	Рекомендація	Номер банку	X	Y	Рекомендація
1	4	2	Hold	11	5	4	Hold
2	6	10	Buy	12	3	8	Buy
3	5	7	Buy	13	13	5	Sell
4	12	3	Sell	14	15	3	Sell
5	17	4	Sell	15	5	9	Sell
6	3	10	Buy	16	8	3	Hold
7	6	1	Hold	17	6	4	Hold
8	6	3	Hold	18	14	3	Sell
9	15	1	Sell	19	4	8	Buy
10	15	4	Sell	20	4	3	Hold

Слід пам'ятати, що у кластерному аналізі поділ на кластери істотно залежить від абсолютних значень вихідних даних. Цю проблему вирішують за допомогою нормування (стандартизації).

Кластерний аналіз не потребує апріорних припущень про набір даних, не накладає обмежень на форму подання досліджуваних об'єктів, дає можливість аналізувати показники різних типів даних (інтервальні дані, частоти, бінарні дані). При цьому необхідно пам'ятати, що змінні повинні вимірюватися за порівняними шкалами.

#### 4.2 Дискримінантний аналіз і логістична регресія

Логістична регресія і дискримінантний аналіз застосовуються у тому випадку, коли необхідно класифікувати (сегментувати) об'єкти за цільовими групами, які, у свою чергу, подано рівнями однієї одноваріантної змінної.

Прикладом завдання, що вирішується за допомогою цих статистичних методів, може бути завдання класифікувати підприємства за двома групами – банкрути і небанкрути – на основі їхніх фінансових характеристик (ліквідність, прибутковість тощо). Як видно, в процедурах логістичної регресії і дискримінантного аналізу використовуються змінні – критерії сегментації і одна змінна, що кодує цільові групи, на які слід поділити об'єкти на основі критеріїв сегментації.

Необхідно зазначити, що спектр можливостей застосування логістичної регресії вузьчий, ніж для дискримінантного аналізу, тому використання дискримінантного аналізу як універсального методу є доцільнішим. Більш того, рекомендується завжди починати класифікаційне дослідження саме з дискримінантного аналізу, а не з логістичної регресії, і застосовувати останню у разі невпевненості в результатах дискримінантного аналізу.

3-поміж методів розпізнавання образів особливе місце посідає дискримінантний аналіз. На відміну від кластерного аналізу дискримінантний не утворює нових класів, а допомагає виявити різницю між існуючими класами і віднести новий (нерозпізнаний) об'єкт до одного з них за принципом максимальної схожості. Наприклад, банк, спираючись на певну систему характеристик фінансового стану клієнтів, які звертаються за позиками, поділяє їх на дві категорії: надійні та ненадійні. Дискримінантний аналіз використовується в медичній діагностиці, під час визначення ризику відмови приладів у технічних системах тощо. Основна проблема – звести помилку класифікації до мінімуму.

*Дискримінантна функція* – це лінійна комбінація певної множини ознак, які називають класифікаційними і на основі яких ідентифікують класи. Особливість дискримінантної функції полягає в тому, що класи подаються шкалою найменувань, а класифікаційні ознаки  $x_i$ , де  $i = 1, 2, \dots, m$ , вимірюються метричною шкалою. Кількість останніх не може перевищувати  $(n - 2)$ , де  $n$  — обсяг сукупності. Функціонально зв'язані й висококорельовані ознаки до ознакового простору моделі не включаються.

Дискримінантна функція  $f_j$  визначається для кожного  $j$ -го класу ( $j = 1, 2, \dots, p$ ):

$$f_j = a_{0j} + a_{1j}\bar{x}_{1j} + a_{2j}\bar{x}_{2j} + \dots + a_{mj}\bar{x}_{mj},$$

де  $a_{ij}$  – коефіцієнт функції (змістовної інтерпретації не має);

$\bar{x}_{ij}$  – середнє значення  $i$ -ї ознаки в  $j$ -му класі.

Для оцінювання спроможності дискримінантної функції розпізнавати класи у багатовимірному ознаковому просторі використовують також  $\lambda$ -статистику Вілкса (*Wilks lambda*)

$$\lambda = \prod_{j=1}^p \frac{1}{1 + \lambda_j},$$

де  $\lambda_j$  – власні значення матриці коваріацій.

$\lambda$ -статистика враховує як відмінності між класами, так і однорідність кожного класу. Оскільки  $\lambda$  розраховується як обернена величина, то чим більше різняться центроїди, тим менше її значення, і, навпаки, якщо центроїди збігаються, то  $\lambda$  прямує до 1. Отже, близькі до 0 значення  $\lambda$  свідчать про високу розпізнавальну спроможність дискримінантної функції. Істотність різниці значень центроїдів перевіряється також за допомогою критерію  $\chi^2$  або дисперсійного  $F$ -критерію, які функціонально зв'язані з  $\lambda$ -статистикою.

У системі *Statistica* процедури дискримінантного аналізу об'єднані в модулі *Discriminant Analysis* – *Дискримінантний аналіз*. Порядок використання модуля розглянемо на прикладі.

**Приклад 4.2** За даними прикладу 4.1 сформулювати процедуру класифікації акції комерційного банку (віднесення її до класу *Sell*, *Buy* або *Hold*).

Вихідні дані для аналізу наведено на рисунку 4.9.

	1 B	2 Pr	3 P
1	4	2	Hold
2	6	10	Buy
3	5	7	Buy
4	12	3	Sell
5	17	4	Sell
6	3	10	Buy
7	6	1	Hold
8	6	3	Hold
9	15	1	Sell
10	15	4	Sell
11	5	4	Hold
12	3	8	Buy
13	13	5	Sell
14	15	3	Sell
15	5	9	Sell
16	8	3	Hold
17	6	4	Hold
18	14	3	Sell

Рисунок 4.9 – Дані для проведення дискримінантного аналізу

Для запуску модуля *Дискримінантний аналіз* вибираємо пункт меню *Statistics – Multivariate Exploratory Techniques – Discriminant Analysis*.

Діалогове вікно процедури (рисунок 4.10) пропонує вибрати ознаки (*Variables*) для аналізу. В цьому випадку незалежними ознаками є витрати (*B*) і прибуток (*Pr*). Ознакою, за якою здійснюється групування, є рекомендація (*P*).

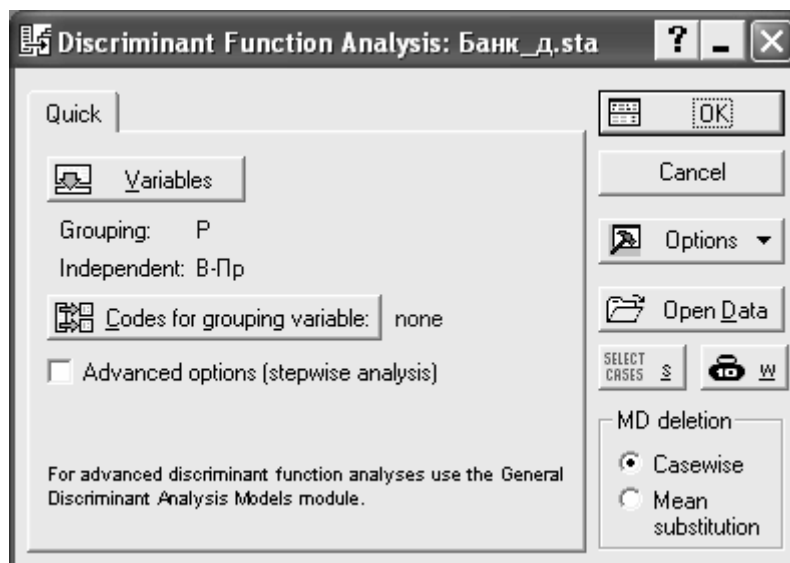


Рисунок 4.10 – Діалогове вікно процедури

За командою на виконання система видає вікно *Discriminant Function Analysis Results* (рисунок 4.11). Інформаційна панель містить дані про кількість змінних – 2, про значення лямбди Уїлкса, значення *F*-статистики, що зв'язане з лямбдою Уїлкса. Значення лямбди Уїлкса 0,066 свідчить про досить високу точність аналізу.

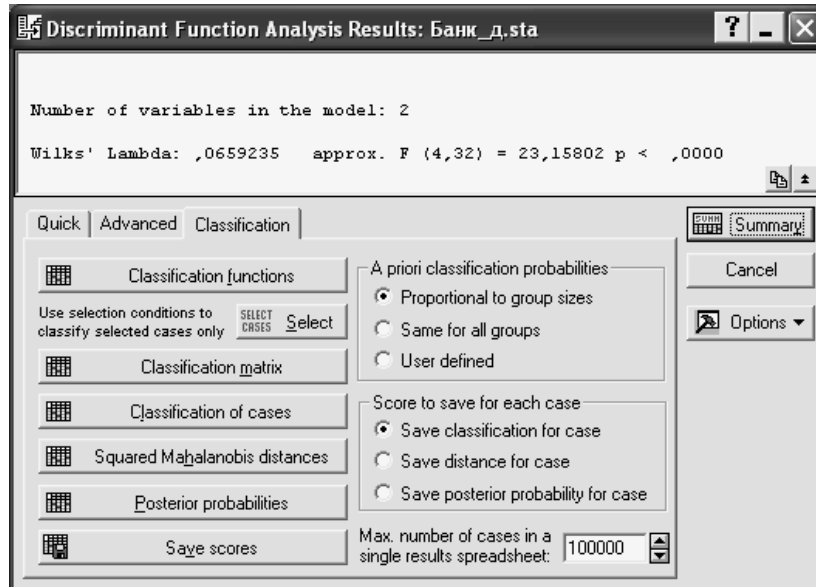


Рисунок 4.11 – Вікно *Discriminant Function Analysis Results*

Вкладка *Classification functions* дає можливість отримати функції класифікації для кожного класу (рисунок 4.12).

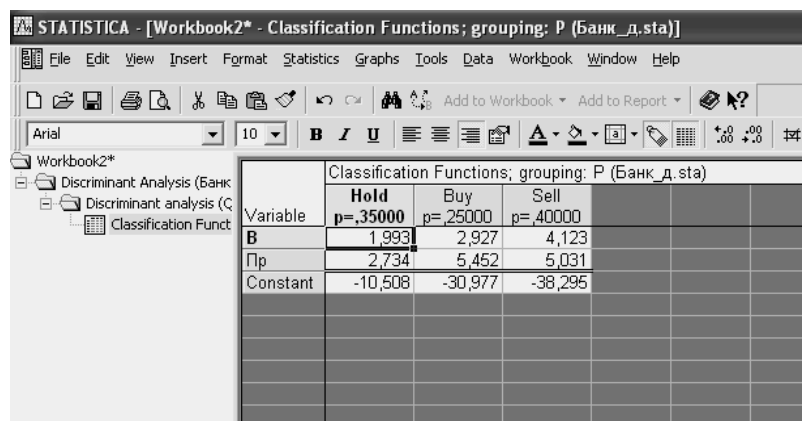


Рисунок 4.12 – Вкладка *Classification functions*

Отже, отримано такі функції класифікації:

$$P_{hold} = -10,508 + 1,993B + 2,734Пр ;$$

$$P_{buy} = -30,977 + 2,927B + 5,452Пр ;$$

$$P_{sell} = -38,295 + 4,123B + 5,031Пр .$$

Для класифікації нового об'єкта (акції) у ці формули слід підставити значення ознак (витрати, прибуток). Новий об'єкт буде віднесено до того

класу, класифікаційне значення якого є найбільшим.

Припустимо, що необхідно класифікувати акцію банку з такими характеристиками: витрати дорівнюють 6 г.о., прибуток – 8 г.о.

Підставляємо ці значення в класифікаційні функції і отримуємо

$$P_{hold} = 23,322,$$

$$P_{buy} = 30,20,$$

$$P_{sell} = 26,691.$$

Отже, акцію треба віднести до групи *Buy*.

Дані таблиці 4.3 (вкладка *Classification matrix*) показують точність класифікації. Загальна точність моделі – 95 % (помилка – в 5 % випадків). При цьому всі акції типу *Hold* і *Buy* було класифіковано правильно (точність – 100 %). При класифікації акцій типу *Sell* допущено одну помилку (одну акцію було помилково віднесено до типу *Buy*, точність – 87,5 %).

Таблиця 4.3 – Матриця класифікації

Рекомендація	Percent	Hold	Buy	Sell
Hold	100,0	7	0	0
Buy	100,0	0	5	0
Sell	87,5	0	1	7
Total	95,0	7	6	7

Вкладка *Апостеріорні ймовірності (Posterior Probabilities)* (рисунок 4.13) демонструє ймовірності, за якими той чи інший об'єкт належить до певного класу, а також показує неправильно класифіковані об'єкти (позначено знаком «\*»).

Case	Observed Classif.	Hold p=.35000	Buy p=.25000	Sell p=.40000
1	Hold	1,000	0,000	0,000
2	Buy	0,000	0,987	0,013
3	Buy	0,038	0,949	0,013
4	Sell	0,009	0,003	0,988
5	Sell	0,000	0,000	1,000
6	Buy	0,000	1,000	0,000
7	Hold	1,000	0,000	0,000
8	Hold	0,998	0,001	0,000
9	Sell	0,002	0,000	0,998
10	Sell	0,000	0,000	1,000
11	Hold	0,992	0,007	0,000
12	Buy	0,017	0,983	0,001
13	Sell	0,000	0,002	0,998
14	Sell	0,000	0,000	1,000
* 15	Sell	0,000	0,994	0,006
16	Hold	0,972	0,008	0,021
17	Hold	0,979	0,018	0,003

Рисунок 4.13 – Апостеріорні ймовірності (*Posterior Probabilities*)



Дискримінантний аналіз є більш універсальною статистичною процедурою порівняно з методами логістичної регресії. Основним результатом проведення дискримінантного аналізу є розрахована ймовірність потрапляння кожного об'єкта до тієї або іншої групи. Разом з цією інформацією за висновками аналізу можна скласти рівняння дискримінантної функції.

Слід зазначити, що дискримінантний аналіз проводиться при таких основних припущеннях:

- множину об'єктів  $M$  для аналізу поділено на дві підмножини або більше;
- у кожній підмножині знаходяться, принаймні, два об'єкти (причому всі об'єкти множини  $M$  повинні належати до однієї з підмножин);
- кількість об'єктів  $N$  перевищує кількість дискримінантних змінних  $p$  на дві одиниці ( $0 < p < N - 2$ );
- лінійна незалежність ознак, тобто жодна з ознак не є лінійною комбінацією інших ознак (інакше вона не несе нової інформації);
- закон розподілу дискримінантних змінних (ознак об'єкта) нормальний.

### 4.3 Факторний аналіз

Факторний аналіз – це сукупність методів багатовимірного статистичного аналізу, що використовуються для вивчення взаємозв'язків між значеннями змінних. Факторний аналіз дає можливість поділити масив змінних на декілька груп, які називають факторами. Класифікація здійснюється на основі критерію кореляції між змінними. В один фактор об'єднуються декілька змінних, які тісно корелюють між собою і не корелюють або слабо корелюють з іншими змінними, які входять до інших факторів.

Таким чином, унаслідок факторного аналізу отримуємо з несистематизованого масиву даних декілька макрозмінних, що описують різні характеристики досліджуваного об'єкта.

Прикладом використання факторного аналізу є сегментація ринку. Факторний аналіз застосовується для виявлення агрегатних змінних, що є основою для сегментації споживачів. Наприклад, споживачі можуть характеризуватися різним ступенем значущості, яку вони бачать у характеристиках продукту. Тут факторний аналіз дасть можливість виявити цільові сегменти споживачів на основі значущості для них різних груп чинників:

- покупці, що орієнтуються під час вибору переважно на цінові чинники (вартість, знижки);
- покупці, які орієнтуються на якість досліджуваного продукту;
- покупці, що вибирають в основному за виглядом (дизайн упаковки).

Основна складність під час проведення факторного аналізу полягає в необхідності раціонально інтерпретувати отримані макрокатегорії з погляду здорового глузду (стосовно цілей і специфіки конкретного дослідження). Ця проблема не має універсального рішення і підлягає окремому аналізу у кожному конкретному випадку. Саме складність інтерпретації результатів є істотним обмеженням цієї статистичної методики, оскільки через неможливість логічного опису отриманих категорій іноді доводиться взагалі відмовитися від використання факторного аналізу.

Ще одним обмеженням застосування факторного аналізу є ситуація, коли одна й та сама змінна належить відразу до двох або більше факторів, тобто змінну не можна однозначно класифікувати. У такому разі слід або відмовитися від використання факторного аналізу і спробувати застосувати інші статистичні методики (наприклад, кластерний аналіз), або наново перерахувати факторну модель без цієї змінної, а потім уручну віднести неоднозначну змінну до того або іншого фактора на основі логічних міркувань.

Акцент у факторному аналізі робиться на дослідженні внутрішніх причин, які формують специфіку явища, що вивчається, на виявленні узагальнених чинників, які стоять за відповідними конкретними показниками. Факторний аналіз не потребує апріорного поділення ознак на залежні й незалежні, оскільки всі ознаки розглядаються як рівноправні.

Мета факторного аналізу – сконцентрувати початкову інформацію, виражаючи велику кількість ознак через меншу кількість більш ємних внутрішніх характеристик явища, які, проте, не піддаються безпосередньому вимірюванню (наприклад, рівень інвестиційної привабливості).

Фактором називають гіпотетичну безпосередньо не вимірювану латентну (приховану) змінну, яка має лінійні кореляційні зв'язки з початковими вимірюваними змінними. Зазначимо, що за визначенням кожний з факторів безпосередньо для вимірювання недоступний – він гіпотетичний і є сумою вимірюваних кількісних ознак з різними ваговими коефіцієнтами:

$$F_i = \sum_{i=1}^k a_i \cdot x_{ij} .$$

Отже, головними задачами факторного аналізу є:

- скорочення кількості змінних (редукція даних);
- визначення структури взаємозв'язків між змінними, тобто класифікація змінних.

Основні результати факторного аналізу виражаються в наборах факторних навантажень і факторних ваг.

Факторні навантаження – це значення коефіцієнтів кореляції кожної з початкових ознак з кожним з виявлених факторів. Чим тісніше зв'язок цієї

ознаки з певним фактором, тим вище значення навантаження ознаки. Додатний знак навантаження ознаки вказує на прямий (а від'ємний знак – на зворотний) зв'язок цієї ознаки з фактором.

Факторними вагами називають кількісні значення виділених факторів для кожного з наявних об'єктів. Об'єкт з великим значенням факторної ваги має значний ступінь прояву властивостей, що описуються фактором.

Слід зазначити ще декілька обмежень при факторному аналізі:

- використовуються лише кількісні дані, які мають нормальний розподіл;

- взаємозв'язок між окремими ознаками має бути лінійним;
- необхідним є використання коефіцієнта кореляції Пірсона.

**Приклад 4.3** Провести статистичний факторний аналіз інвестиційної привабливості 50 машинобудівних підприємств України за такими показниками:

- 1 Коефіцієнт оборотності активів (трансформації).
- 2 Частина основних засобів у активі.
- 3 Фондовіддача.
- 4 Частина виробничих запасів у запасах.
- 5 Забезпеченість запасів власними оборотними коштами.
- 6 Співвідношення дебіторської та кредиторської заборгованостей.
- 7 Коефіцієнт фінансової автономії.
- 8 Коефіцієнт фінансової стійкості.
- 9 Коефіцієнт маневреності власного капіталу.
- 10 Коефіцієнт структури покриття довгострокової заборгованості.
- 11 Коефіцієнт покриття запасів нормальними джерелами фінансування.
- 12 Коефіцієнт покриття.
- 13 Коефіцієнт швидкої ліквідності.
- 14 Коефіцієнт абсолютної ліквідності.
- 15 Коефіцієнт ліквідної платоспроможності.
- 16 Коефіцієнт забезпеченості поточної діяльності власними коштами.
- 17 Коефіцієнт оборотності іммобільних активів.
- 18 Коефіцієнт оборотності мобільних активів.
- 19 Рентабельність активів.
- 20 Рентабельність власного капіталу.
- 21 Рентабельність необоротного капіталу.
- 22 Рентабельність оборотного капіталу.
- 23 Рентабельність операційної діяльності.
- 24 Рентабельність виробництва.
- 25 Рентабельність продажів за чистим прибутком.

У системі *Statistica* факторний аналіз здійснюється в модулі *Statistics – Multivariate Exploratory Techniques – Factor Analysis* (рисунок 4.14).

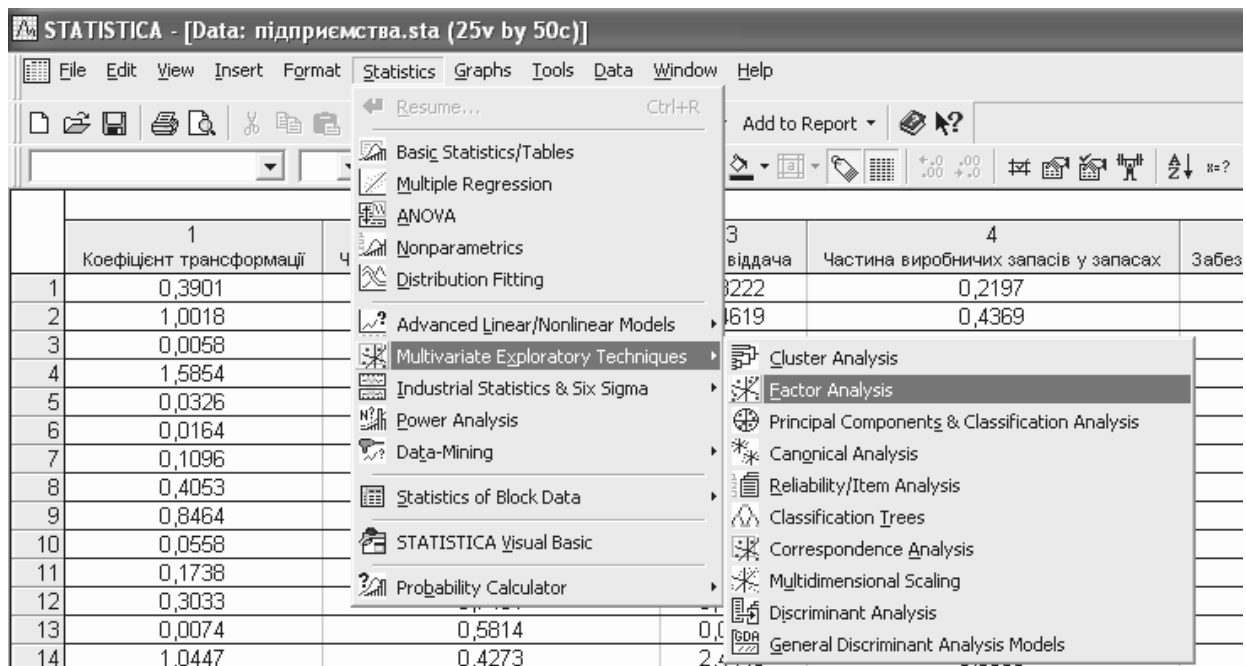


Рисунок 4.14 – Факторний аналіз у системі *Statistica*

Діалогове вікно модуля факторного аналізу (рисунок 4.15) пропонує вибрати змінні для аналізу (*Variables*) і визначити тип вихідних даних – необроблені дані (*Raw Data*) або кореляційна матриця (*Correlation matrix*).

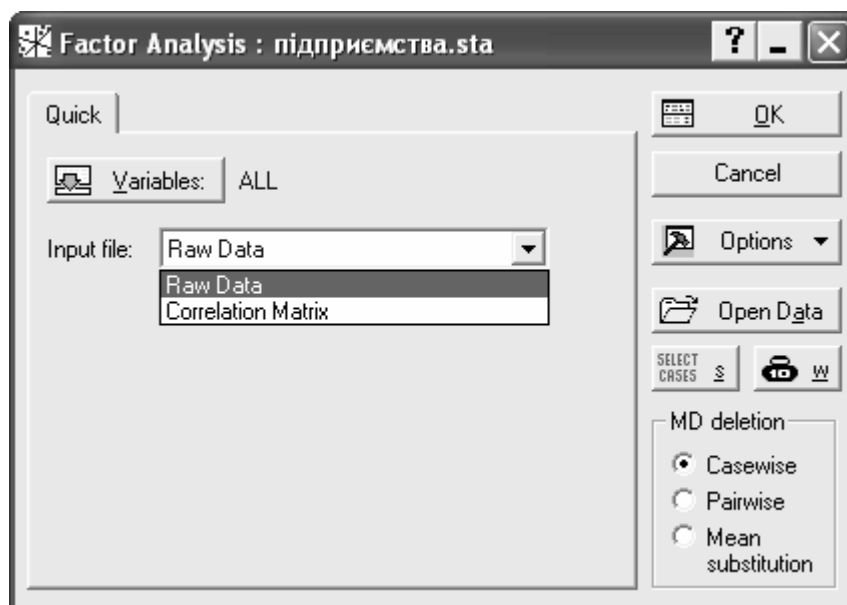


Рисунок 4.15 – Діалогове вікно модуля *Factor Analysis*

Другим кроком є вибір кількості факторів і способу факторного аналізу (рисунок 4.16).

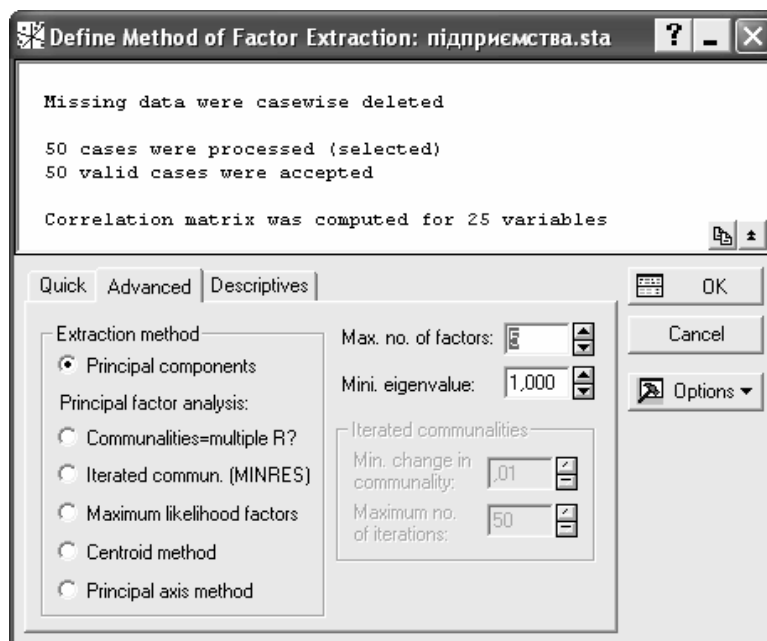


Рисунок 4.16 – Вибір кількості факторів і способу факторного аналізу

Результати проведеного аналізу показали, що на фінансово-економічну складову інвестиційної привабливості підприємства впливають п'ять факторів, які пояснюють близько 80 % її мінливості (таблиця 4.4, рисунок 4.17, вкладка *Eigenvalues*, відсоток загальної накопиченої дисперсії, *Cumulative*).

Таблиця 4.4 – Характеристики факторів (*Eigenvalues*)

Номер фактора	Власне значення фактора	Відсоток загальної дисперсії	Відсоток накопиченої дисперсії
1	10,26944	41,07778	41,07778
2	3,47075	13,88299	54,96076
3	2,86354	11,45415	66,41492
4	2,03998	8,15990	74,57482
5	1,29548	5,18193	79,75675

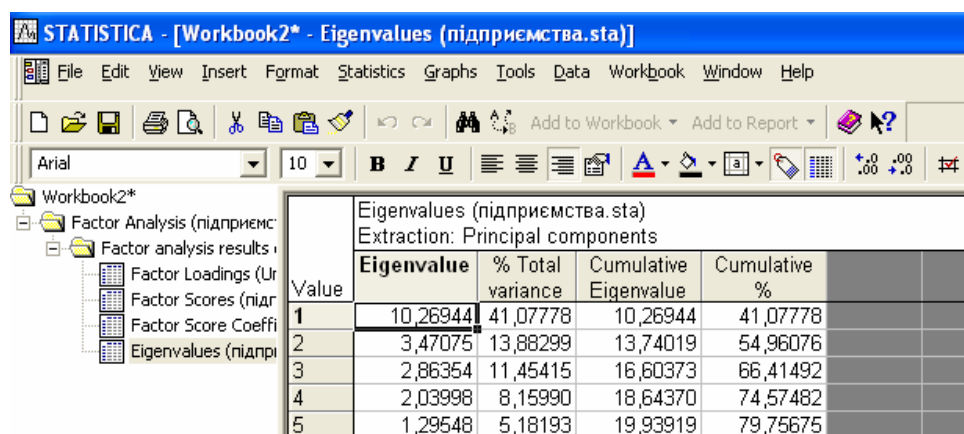


Рисунок 4.17 – Результати факторного аналізу: характеристики факторів

Результатами аналізу є набір факторних навантажень (рисунок 4.18). Значущі коефіцієнти для кожного фактора віділено.

STATISTICA - [Workbook1* - Factor Loadings (Unrotated) (підприємства)]					
Factor Loadings (Unrotated) (підприємства) Extraction: Principal components (Marked loadings are > ,700000)					
Variable	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
<b>Коефіцієнт трансформації</b>	<b>-0,50332</b>	-0,627766	-0,343993	-0,071590	0,408963
Частина основних засобів у активі	0,56918	-0,291291	0,629030	-0,070572	-0,032647
Фондовіддача	-0,40340	0,272564	-0,545065	-0,474907	-0,158795
Частина виробничих запасів у запасах	-0,21262	0,214524	-0,012790	0,561355	-0,238985
Забезпеченість запасів ВОЗ	-0,64007	0,474185	-0,241635	-0,472779	-0,121855
Співвідношення ДЗ і КЗ	-0,60652	0,063914	0,153101	0,113159	0,329416
Коефіцієнт фінансової автономії	-0,59984	0,157992	<b>0,703512</b>	-0,244855	0,116074
Коефіцієнт фінансової стійкості	-0,61135	0,159716	0,598791	-0,244678	0,141458
Коефіцієнт маневреності власного капіталу	-0,60704	0,039186	0,448959	-0,313423	0,097735
Коефіцієнт структури покриття довгострокової заборгованості	0,06478	-0,031020	-0,644995	0,091273	-0,026299
Коефіцієнт покриття запасів нормальними джерелами фінансування	-0,56112	0,489743	-0,331756	-0,494612	-0,144676
Коефіцієнт покриття	<b>-0,81749</b>	0,509088	0,003664	0,110921	0,097058
Коефіцієнт швидкої ліквідності	<b>-0,78726</b>	0,478494	-0,126102	0,208477	0,011055
Коефіцієнт абсолютної ліквідності	-0,65331	0,330872	0,064736	0,587169	0,077678
Коефіцієнт ліквідної платоспроможності	<b>-0,76887</b>	0,400633	0,156210	0,272210	0,190081
Коефіцієнт забезпеченості поточної діяльності власними коштами	<b>-0,75812</b>	0,137042	-0,091247	-0,157383	0,293735
Коефіцієнт оборотності іммобільних активів	-0,65157	-0,361426	-0,344383	0,166226	0,336319
Коефіцієнт оборотності мобільних активів	-0,29482	<b>-0,731249</b>	-0,099970	-0,234206	0,372888
Рентабельність активів	<b>-0,79143</b>	-0,339193	0,121494	0,135230	-0,301398
Рентабельність власного капіталу	-0,66478	-0,339938	0,364441	-0,183362	-0,323609
Рентабельність необоротного капіталу	<b>-0,77504</b>	-0,202965	0,109531	0,323633	-0,200357
Рентабельність оборотного капіталу	<b>-0,79535</b>	-0,372556	-0,081842	0,053425	-0,278968
Рентабельність операційної діяльності	<b>-0,79930</b>	-0,420255	-0,110069	0,103688	-0,112026
Рентабельність виробництва	<b>-0,81705</b>	-0,357999	-0,166663	0,040164	-0,113558
Рентабельність продажів за ЧП	-0,53412	-0,438311	0,041534	-0,167663	-0,357347
Expl.Var	10,26944	3,470747	2,863538	2,039976	1,295483
Prp.Totl	0,41078	0,138830	0,114542	0,081599	0,051819

Рисунок 4.18 – Результати факторного аналізу: факторні навантаження

Необхідно, щоб кожний з факторів містив значущі коефіцієнти. Цього можна досягти, застосовуючи різні методи обертання. Використання методу ортогонального обертання *Varimax* з нормалізацією Кайзера дало можливість розрахувати навантаження кожного показника на кожен з п'яти факторів (рисунок 4.19, таблиця 4.5).

Таблиця 4.5 – Факторні навантаження показників фінансової складової інвестиційної привабливості

Код показника	Фактор 1	Фактор 2	Фактор 3	Фактор 4	Фактор 5
Коефіцієнт трансформації	0,070878	0,361161	-0,067274	-0,009500	<b>0,893339</b>
Частина основних засобів у активі	-0,646410	-0,124554	0,279275	-0,486507	-0,249999
Фондовіддача	<b>0,863434</b>	0,133536	-0,119688	-0,030455	0,079165
Частина виробничих запасів у запасах	-0,096596	0,178242	-0,163852	0,576769	-0,251025
Забезпеченість запасів ВОЗ	<b>0,902144</b>	0,151141	0,249730	0,179255	-0,009413
Співвідношення ДЗ і КЗ	0,102428	0,124949	0,389378	0,461343	0,355177
Коефіцієнт фінансової автономії	0,102043	0,207258	<b>0,927557</b>	0,198289	0,008872
Коефіцієнт фінансової стійкості	0,156770	0,192818	<b>0,853027</b>	0,210768	0,062093

## Продовження таблиці 4.5

Код показника	Фактор 1	Фактор 2	Фактор 3	Фактор 4	Фактор 5
Коефіцієнт маневреності власного капіталу	0,216379	0,272379	<b>0,726842</b>	0,102333	0,139999
Коефіцієнт структури покриття довгострокової заборгованості	0,230340	-0,028484	-0,590573	0,029330	0,163387
Коефіцієнт покриття запасів нормальними джерелами фінансування	<b>0,937036</b>	0,106244	0,153049	0,128890	-0,031255
Коефіцієнт покриття	0,508272	0,139760	0,356241	<b>0,734051</b>	0,074781
Коефіцієнт швидкої ліквідності	0,499711	0,190433	0,179455	<b>0,766563</b>	0,049251
Коефіцієнт абсолютної ліквідності	0,051635	0,179738	0,141120	<b>0,913545</b>	0,042733
Коефіцієнт ліквідної платоспроможності	0,255733	0,128482	0,409307	<b>0,788395</b>	0,123692
Коефіцієнт забезпеченості поточної діяльності власними коштами	0,482659	0,170168	0,347396	0,374323	0,436107
Коефіцієнт оборотності іммобільних активів	0,132838	0,357733	-0,078729	0,345889	<b>0,736926</b>
Коефіцієнт оборотності мобільних активів	-0,089516	0,322462	0,090851	-0,289221	<b>0,788441</b>
Рентабельність активів	0,063077	<b>0,834114</b>	0,212718	0,304738	0,165701
Рентабельність власного капіталу	0,074980	<b>0,770649</b>	0,475162	0,000771	0,057930
Рентабельність необоротного капіталу	0,007390	0,697831	0,172412	0,507874	0,155140
Рентабельність оборотного капіталу	0,194476	<b>0,827792</b>	0,079574	0,239981	0,267892
Рентабельність операційної діяльності	0,138403	<b>0,755278</b>	0,072264	0,281457	0,420560
Рентабельність виробництва	0,240790	<b>0,727062</b>	0,064924	0,272854	0,415631
Рентабельність продажів за ЧП	0,129210	<b>0,755986</b>	0,137810	-0,095695	0,136269
Expl. Var	3,991017	4,898390	3,643401	4,424231	2,982148
Prp. Totl	0,159641	0,195936	0,145736	0,176969	0,119286

STATISTICA - [Workbook1\* - Factor Loadings (Varimax normalized) (підприємства)]

Factor Loadings (Varimax normalized) (підприємства)  
Extraction: Principal components  
(Marked loadings are > ,700000)

Variable	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
<b>Коефіцієнт трансформації</b>	<b>0,070878</b>	0,361161	-0,067274	-0,009500	<b>0,893339</b>
Частина основних засобів у активі	-0,646410	-0,124554	0,279275	-0,486507	-0,249999
Фондовіддача	<b>0,863434</b>	0,133536	-0,119688	-0,030455	0,079165
Частина виробничих запасів у запасах	-0,096596	0,178242	-0,163852	0,576769	-0,251025
Забезпеченість запасів ВОЗ	<b>0,902144</b>	0,151141	0,249730	0,179255	-0,009413
Співвідношення ДЗ і КЗ	0,102428	0,124949	0,389378	0,461343	0,355177
Коефіцієнт фінансової автономії	0,102043	0,207258	<b>0,927557</b>	0,198289	0,008872
Коефіцієнт фінансової стійкості	0,156770	0,192818	<b>0,853027</b>	0,210768	0,062093
Коефіцієнт маневреності власного капіталу	0,216379	0,272379	<b>0,726842</b>	0,102333	0,139999
Коефіцієнт структури покриття довгострокової заборгованості	0,230340	-0,028484	-0,590573	0,029330	0,163387
Коефіцієнт покриття запасів нормальними джерелами фінансування	<b>0,937036</b>	0,106244	0,153049	0,128890	-0,031255
Коефіцієнт покриття	0,508272	0,139760	0,356241	<b>0,734051</b>	0,074781
Коефіцієнт швидкої ліквідності	0,499711	0,190433	0,179455	<b>0,766563</b>	0,049251
Коефіцієнт абсолютної ліквідності	0,051635	0,179738	0,141120	<b>0,913545</b>	0,042733
Коефіцієнт ліквідної платоспроможності	0,255733	0,128482	0,409307	<b>0,788395</b>	0,123692
Коефіцієнт забезпеченості поточної діяльності власними коштами	0,482659	0,170168	0,347396	0,374323	0,436107
Коефіцієнт оборотності іммобільних активів	0,132838	0,357733	-0,078729	0,345889	<b>0,736926</b>
Коефіцієнт оборотності мобільних активів	-0,089516	0,322462	0,090851	-0,289221	<b>0,788441</b>
Рентабельність активів	0,063077	<b>0,834114</b>	0,212718	0,304738	0,165701
Рентабельність власного капіталу	0,074980	<b>0,770649</b>	0,475162	0,000771	0,057930
Рентабельність необоротного капіталу	0,007390	0,697831	0,172412	0,507874	0,155140
Рентабельність оборотного капіталу	0,194476	<b>0,827792</b>	0,079574	0,239981	0,267892
Рентабельність операційної діяльності	0,138403	<b>0,755278</b>	0,072264	0,281457	0,420560
Рентабельність виробництва	0,240790	<b>0,727062</b>	0,064924	0,272854	0,415631
Рентабельність продажів за ЧП	0,129210	<b>0,755986</b>	0,137810	-0,095695	0,136269
Expl. Var	3,991017	4,898390	3,643401	4,424231	2,982148
Prp. Totl	0,159641	0,195936	0,145736	0,176969	0,119286

Рисунок 4.19 – Результати факторного аналізу: факторні навантаження після обертання

Перший фактор пояснює 15,96 % загальної дисперсії (*Prp.Totl*). Найбільший вплив на нього роблять такі коефіцієнти: фондвіддача, забезпеченість запасів власними оборотними коштами, коефіцієнт покриття запасів нормальними джерелами фінансування. Отже, перший фактор можна трактувати як фактор результатів використання внутрішніх ресурсів.

Другий фактор є найбільш важливим, оскільки описує 19,6 % варіабельності. Він включає такі показники: рентабельність активів, власного капіталу, оборотного капіталу, операційної діяльності, виробництва і рентабельність продажів за чистим прибутком. Аналізуючи їхній склад, можна дійти висновку, що на величину інвестиційної привабливості впливають результати ефективного використання власних і позикових коштів підприємства. Таким чином, цей фактор доцільно інтерпретувати як фактор загальної рентабельності. Він відображає ефективність стратегічного і тактичного використання ресурсів.

Третій фактор пояснює 14,57 % загальної дисперсії. Найбільший вплив на нього роблять коефіцієнти фінансової автономії, маневреності власного капіталу й фінансової стійкості. Ці показники характеризують незалежність підприємства від зовнішніх джерел фінансування.

Четвертий фактор описує 17,7 % варіабельності. До його складу входять такі значущі показники: коефіцієнти покриття, швидкої і абсолютної ліквідності, а також ліквідної платоспроможності. Цей фактор отримав назву фактора ліквідності.

П'ятий фактор, який найменше впливає на інвестиційну привабливість підприємства (пояснює 11,92 % її мінливості), включає три значущі показники: коефіцієнти оборотності активів, мобільних і необоротних активів. Отже, можна дійти висновку, що п'ятий фактор – це фактор оборотності ресурсів. Він характеризує оборотність як внутрішніх, так і зовнішніх фінансових джерел.

Таким чином, на основі аналізу факторних навантажень за кожним фактором було визначено склад показників, загальна значущість яких перевищує 79 %. Виділено п'ять факторів: результатів використання внутрішніх ресурсів (містить три фінансових показники); ефективності стратегічного й тактичного використання ресурсів (містить шість фінансових показників); незалежності підприємства від зовнішніх джерел фінансування (містить три фінансових показники); ліквідності (містить чотири фінансових показники); оборотності як внутрішніх, так і зовнішніх фінансових джерел (містить три фінансових показники).

**Проблеми, які виникають при факторному аналізі.** Найскладнішим завданням під час проведення факторного аналізу є інтерпретація отриманих факторів. Не існує універсального рішення: у кожному конкретному випадку аналітик використовує наявний практичний досвід для того, щоб зрозуміти, чому факторна модель відносить ту або



іншу змінну до певного конкретного фактора. Існують випадки (особливо при малій кількості добре формалізованих змінних), коли фактори є очевидними і відмінності між змінними видно неозброєним оком. У такій ситуації можна обійтися без факторного аналізу і розбити змінні на групи вручну. Проте ефективність і потужність факторного аналізу виявляються в складних і нетривіальних випадках, коли змінні не можна наперед класифікувати.

Наведемо рекомендації, які допоможуть вирішити проблеми при інтерпретації результатів факторного аналізу.

Коли це можливо і прийнятно для цілей дослідження, слід формалізувати змінні до проведення факторного аналізу. Це дасть можливість аналітику наперед зробити припущення про поділення сукупності наявних змінних на групи. Завдання дослідника при інтерпретації результатів у цьому випадку спроститься, оскільки він вже не починатиме «з чистого аркуша». Його завдання зведеться до перевірки раніше висунутих гіпотез про належність тієї або іншої змінної до конкретної групи.

Іноді виникають випадки, коли змінна, віднесена до конкретного фактора, логічно ніяк не зв'язана з рештою змінних, які складають той же самий фактор. Можна перерахувати факторну модель без відсікання незначущих коефіцієнтів і подивитися, з яким ще фактором ця нелогічна змінна корелює майже з тією ж силою, як з фактором, до якого її було віднесено автоматично. У такій ситуації (при невеликій різниці в коефіцієнтах кореляції) можна спробувати віднести змінну до іншого фактора і, якщо це виявиться логічним, розглядати її в групі змінних другого фактора.

Можна вручну скоротити кількість факторів, що полегшить завдання дослідника при інтерпретації результатів чинника факторного аналізу. Проте необхідно мати на увазі, що таке скорочення знизить гнучкість факторної моделі й навіть може призвести до ситуації, коли змінні будуть помилково поділені на неправильні, з практичної точки зору, групи. Зниження кількості факторів неминуче знизить і частку однозначно класифікованих факторів.

Як варіант попереднього рішення можна запропонувати об'єднати два або більше факторів з невеликими кількостями змінних. Таке групування, з одного боку, дасть можливість знизити кількість факторів, що інтерпретуються, а з іншого – полегшить розуміння факторів.

Якщо дослідник потрапив у безвихідь і ніякі засоби не допомагають пояснити належність тієї або іншої змінної до конкретного фактора, залишається застосувати іншу статистичну процедуру (наприклад, кластерний аналіз).

Кластерний і факторний аналіз мають на меті одне й те саме: класифікувати об'єкти за однорідними групами (сегментами, кластерами).

Проте між цими методами й дискримінантним аналізом існує одна головна відмінність. При дискримінантному аналізі й логістичній регресії завжди є якась залежна (результуюча) змінна з двома або більше варіантами, рівнями (категоріями). Завдання аналізу в цьому випадку полягає в класифікації об'єктів за цими рівнями результуючої змінної. Ці два статистичних методи дають можливість сегментувати вибірку на наперед відомі цільові групи. При кластерному і факторному аналізі ситуація інша: кластери (сегменти, категорії), на які слід поділити вибірку, наперед не відомі. Завданням статистичного аналізу в цьому випадку буде не тільки формування максимально однорідних сегментів, але й виділення кластерів, за якими проводитиметься сегментація.

#### 4.4 Визначення репрезентанта групи

Іноді необхідно визначити репрезентанта („типового представника”) серед групи показників. Для досягнення цієї мети використовують метод центра ваги. Основою методу є розрахунок коефіцієнтів кореляції. Репрезентанти, що виділяються з груп показників, знаходяться поблизу „центра ваги,” груп і тому задовольняють всі основні вимоги, які ставляться до діагностичних ознак. Ці показники не корелюють між собою і одночасно сильно корелюють з іншими показниками, які не ввійшли до складу діагностичних.

Алгоритм методу центра ваги для декількох груп такий:

1 Вибираються репрезентанти для груп, до яких входить лише один показник, – він і буде репрезентантом.

2 Вибираються репрезентанти для груп, до яких входять три показники й більше. Розраховуються коефіцієнти кореляції для показників, що входять до групи, і обчислюється величина  $(1 - \rho_{ij})$  (таблиця 4.6).

Таблиця 4.6 – Обчислення величини  $(1 - \rho_{ij})$  для груп, до яких входять три показники й більше

Показники	Показник 1	Показник 2	Показник 3	...	Показник n
Показник 1	$(1 - \rho_{11})$	$(1 - \rho_{12})$	$(1 - \rho_{13})$	...	$(1 - \rho_{1n})$
Показник 2	$(1 - \rho_{21})$	$(1 - \rho_{22})$	$(1 - \rho_{23})$	...	$(1 - \rho_{2n})$
Показник 3	$(1 - \rho_{31})$	$(1 - \rho_{32})$	$(1 - \rho_{33})$	...	$(1 - \rho_{3n})$
...	...	...	...	...	...
Показник n	$(1 - \rho_{n1})$	$(1 - \rho_{n2})$	$(1 - \rho_{n3})$	...	$(1 - \rho_{nn})$
Усього	$\sum_{i=1}^n (1 - \rho_{i1})$	$\sum_{i=1}^n (1 - \rho_{i2})$	$\sum_{i=1}^n (1 - \rho_{i3})$		$\sum_{i=1}^n (1 - \rho_{in})$

Репрезентантом групи стане показник  $j$ , для якого сума  $\sum_{i=1}^n (1 - \rho_{ij})$  буде мінімальною.

3 Вибираються репрезентанти для груп, до яких входять два показники. Розраховуються коефіцієнти кореляції для показників, що входять до групи, а також для репрезентантів інших груп (таблиця 4.7).

Репрезентантом групи стане показник  $j$  групи, для якого сума  $\sum_{i=1}^n (1 - \rho_{ij})$  буде максимальною. Таким чином, вибираємо максимальне зі значень:  $\sum_{i=1}^n (1 - \rho_{i1})$ ,  $\sum_{i=1}^n (1 - \rho_{i2})$ .

Таблиця 4.7 – Обчислення величини  $(1 - \rho_{ij})$  для груп, до яких входять два показники

Показники	Показник 1	Показник 2	Показник 3 (репрезентант групи 1)	Показник 4 (репрезентант групи 2)	...	Показник n (репрезентант групи k)
Показник 1	$(1 - \rho_{11})$	$(1 - \rho_{12})$	$(1 - \rho_{13})$	$(1 - \rho_{14})$	...	$(1 - \rho_{1n})$
Показник 2	$(1 - \rho_{21})$	$(1 - \rho_{22})$	$(1 - \rho_{23})$	$(1 - \rho_{24})$	...	$(1 - \rho_{2n})$
Показник 3 (репрезентант групи 1)	$(1 - \rho_{31})$	$(1 - \rho_{32})$	$(1 - \rho_{33})$	$(1 - \rho_{34})$	...	$(1 - \rho_{3n})$
Показник 4 (репрезентант групи 2)	$(1 - \rho_{41})$	$(1 - \rho_{42})$	$(1 - \rho_{43})$	$(1 - \rho_{44})$	...	$(1 - \rho_{4n})$
...	...	...	...	...	...	...
Показник n (репрезентант групи k)	$(1 - \rho_{n1})$	$(1 - \rho_{n2})$	$(1 - \rho_{n3})$	$(1 - \rho_{n4})$	...	$(1 - \rho_{nn})$
Усього	$\sum_{i=1}^n (1 - \rho_{i1})$	$\sum_{i=1}^n (1 - \rho_{i2})$	$\sum_{i=1}^n (1 - \rho_{i3})$	$\sum_{i=1}^n (1 - \rho_{i4})$	...	$\sum_{i=1}^n (1 - \rho_{in})$

**Приклад 4.4** Визначити репрезентантів факторів, отриманих під час факторного аналізу (див. приклад 4.3).

Усі фактори містять три показники або більше. Визначимо „типового представника” для першого фактора. Для цього розрахуємо коефіцієнти кореляції для показників, які ввійшли до першого кластера (таблиця 4.8).

Далі визначаємо величину  $(1 - \rho_{ij})$  для кожного з фінансових показників, що складають перший фактор. Репрезентантом фактора став коефіцієнт покриття запасів нормальними джерелами фінансування, для якого величина  $\sum_{i=1}^n (1 - \rho_{ij})$  є мінімальною (таблиця 4.9).

Аналогічно визначаються репрезентанти факторів 2 – 5 прикладу 4.3. Для другого фактора – це рентабельність операційної діяльності, для третього – коефіцієнт автономії, для четвертого – коефіцієнт покриття і для п'ятого – коефіцієнт оборотності активів.

Таблиця 4.8 – Обчислення коефіцієнтів кореляції  $\rho_{ij}$ 

Показник	Фондовіддача	Забезпеченість запасів ВОЗ	Коефіцієнт покриття запасів нормальними джерелами фінансування
Фондовіддача	1,00	0,72	0,76
Забезпеченість запасів ВОЗ	0,72	1,00	0,97
Коефіцієнт покриття запасів нормальними джерелами фінансування	0,76	0,97	1,00

Таблиця 4.9 – Обчислення величини  $(1 - \rho_{ij})$ 

Показник	Фондовіддача	Забезпеченість запасів ВОЗ	Коефіцієнт покриття запасів нормальними джерелами фінансування
Фондовіддача	0,00	0,28	0,24
Забезпеченість запасів ВОЗ	0,28	0,00	0,03
Коефіцієнт покриття запасів нормальними джерелами фінансування	0,24	0,03	0,00
Усього	0,52	0,31	<b>0,27</b>

Таким чином, система показників для оцінювання інвестиційної привабливості машинобудівних підприємств складається з таких показників:

- коефіцієнт оборотності активів, який характеризує майновий стан і ділову активність суб'єкта господарювання, показує, скільки одиниць виручки від реалізації припадає на одиницю майна підприємства;
- коефіцієнт автономії, що відображає фінансову незалежність підприємства і характеризує його можливості виконати зовнішні зобов'язання за рахунок власних активів;
- коефіцієнт покриття запасів нормальними джерелами фінансування, який відображає рівень платоспроможності суб'єкта господарювання і показує, які джерела засобів і в якому обсязі використовуються для покриття виробничих запасів;
- коефіцієнт покриття, що є показником ліквідності й показує, скільки грошових одиниць оборотних коштів припадає на кожну одиницю короткострокових зобов'язань;

▪ рентабельність операційної діяльності, яка характеризує ефективність основної діяльності суб'єкта господарювання й показує, скільки грошових одиниць чистого прибутку припадає на одиницю витрат.

### Запитання для самоперевірки

- 1 Охарактеризуйте суть кластерного аналізу.
- 2 Які обмеження має дискримінантний аналіз?
- 3 Що є результатом дискримінантного аналізу?
- 4 Що є основною проблемою застосування факторного аналізу?
- 5 Що є метою методів класифікації?
- 6 В яких випадках застосування факторного аналізу є недоцільним?
- 7 Сформулюйте масив вихідних даних відповідно до завдання 7 розділу 3. Зробіть кластерний аналіз і виконайте класифікацію комерційних банків за допомогою:
  - деревоподібної діаграми;
  - методу k-середніх (k-means); кількість кластерів становить 3 – 4.
 Проаналізуйте отримані результати.
- 8 Дані для проведення кластерного аналізу наведено в таблицях 4.10 – 4.12.

Таблиця 4.10 – Варіанти завдання

Номер варіанта	Змінні для аналізу	Номер варіанта	Змінні для аналізу	Номер варіанта	Змінні для аналізу
1	X1,X2,X4,X5	6	X1,X2,X6,X10	11	X2,X3,X10,X11
2	X1,X2,X6,X7	7	X1,X2,X4,X7	12	X2,X3,X4,X10
3	X1,X2,X8,X9	8	X2,X3,X4,X5	13	X2,X3,X6,X11
4	X1,X2,X10,X11	9	X2,X3,X6,X7	14	X2,X3,X6,X8
5	X1,X2,X3,X8	10	X2,X3,X8,X9	15	X1,X3,X5,X9

Таблиця 4.11 – Споживання продуктів на душу населення

Країна	Продукти					
	М'ясо, кг	Масло тваринне, кг	Цукор, кг	Алкоголь, л	Фрукти, кг	Хлібопродукти, кг
	X 1	X 2	X 3	X 4	X 5	X 6
Росія	55	3,9	30	5	28	124
Австралія	100	2,6	47	8, 2	121	87
Австрія	93	5,3	37	12	146	74
Азербайджан	20	4,1	12,4	7,9	52	141
Вірменія	20	3,7	4,3	6,5	72	134
Білорусь	72	3,6	28	5,4	38	120
Бельгія	85	6,9	48	11	83	72

Продовження таблиці 4.11

Країна	Продукти					
	М'ясо, кг	Масло тваринне, кг	Цукор, кг	Алкоголь, л	Фрукти, кг	Хлібопродукти, кг
	X 1	X 2	X 3	X 4	X 5	X 6
Болгарія	65	3	18	9,5	92	156
Великобританія	67	3,5	39	8,8	91	91
Угорщина	73	1,7	40	10,9	73	106
Німеччина	88	6,8	35	8,1	138	73
Греція	83	1	24	8,8	99	108
Грузія	21	3,8	36	9,8	55	140
Данія	98	5	38	10,3	89	77
Ірландія	99	3,3	31	9,6	87	102
Іспанія	89	0,4	26	8,95	103	72
Італія	84	2,2	27	9,6	169	118
Казахстан	61	4,2	19,2	7,2	10	191
Канада	98	3,1	44	7,4	123	77
Киргизія	46	4,1	23,5	6,7	20	134

Таблиця 4.12 – Показники рівня життя населення

Країна	Кількість лікарів на 10 000 населення	Смертність на 100 000 населення	ВВП за ПКС*, %	Витрати на охорону здоров'я, в відсотках до ВВП	Урожайність зернових і зернобобових, ц / га
	X 7	X 8	X 8	X 10	X 11
Росія	44,5	84,98	20,4	3,2	14,4
Австралія	32,5	30,58	71,4	8,5	11,6
Австрія	33,9	38,42	78,7	9,2	56,1
Азербайджан	38,8	60,34	12,1	3,3	16,4
Вірменія	34,4	60,22	10,9	3,2	13,5
Білорусь	43,6	60,79	20,4	5,4	22,4
Бельгія	41	29,82	79,7	8,3	65,5
Болгарія	36,4	70,57	17,3	5,4	27,8
Великобританія	17,9	34,51	69,7	7,1	62,3
Угорщина	32,1	64,73	24,5	6	39,8
Німеччина	38,1	36,63	76,2	8,6	56,9
Греція	41,5	32,84	44,4	5,7	37,4
Грузія	55	62,64	11,3	3,5	18,6
Данія	36,7	34,07	79,2	6,7	54,4
Ірландія	15,8	39,27	57	6,7	64,2
Іспанія	40,9	28,46	54,8	7,3	22,6
Італія	49,4	30,27	72,1	8,5	46
Казахстан	38,1	69,04	13,4	3,3	7,9
Канада	27,6	25,42	79,9	10,2	25,4
Киргизія	33,2	53,13	11,2	3,4	17

\*ПКС – паритет купівельної спроможності.

Використовуючи відповідні показники (див. таблицю 4.10), виконайте кластерний аналіз і класифікацію країн за допомогою:

- деревоподібної діаграми;
- методу k-середніх (k-means); кількість кластерів становить 3.

Проаналізуйте отримані результати. Дайте назви одержаним кластерам.

## **5 ЛІНІЙНИЙ РЕГРЕСІЙНИЙ АНАЛІЗ**

### **5.1 Лінійна регресія: побудова моделі й оцінювання якості рівняння регресії**

Проблема вивчення взаємозв'язків економічних показників є однією з найважливіших проблем економічного аналізу. Будь-яка економічна політика полягає в регулюванні економічних змінних і повинна ґрунтуватися на знанні того, як ці змінні впливають на інші змінні, які є ключовими для того, хто приймає рішення. Так, у ринковій економіці не можна безпосередньо регулювати темп інфляції, але на нього можна впливати засобами бюджетно-податкової та кредитно-грошової політики. Тому, зокрема, має бути вивчена залежність між пропозицією грошей і рівнем цін. Неможливо будувати, перевіряти або покращувати економічні моделі без статистичного аналізу їх змінних з використанням реальних статистичних даних. Уся сфера економічних досліджень може бути в певному значенні охарактеризована як вивчення взаємозв'язків економічних змінних. Інструментарієм їх базового аналізу є методи статистики і економетрії.

Регресійна модель описує кореляційні зв'язки, які об'єктивно існують між явищами. За своїм характером кореляційні зв'язки надзвичайно складні й різноманітні. В одних випадках результат  $y$  зі зміною фактора  $x$  збільшується або зменшується рівномірно, в інших – нерівномірно. Іноді результат може спочатку збільшуватися, а потім – зменшуватися, і навпаки. Простежити всі ці взаємозв'язки і встановити точний функціональний вигляд майже неможливо. А тому під час вибору типу функції йдеться лише про апроксимацію відносно простими функціями незрівнянно більш складних за своєю природою взаємозв'язків. На практиці перевагу віддають моделям, які є лінійними або зводяться до лінійного вигляду шляхом перетворення змінних, наприклад логарифмуванням. Такий підхід, безперечно, містить у собі певну умовність, оскільки передбачає однаковий характер зв'язку з усіма факторами. Проте використання надто складних функцій неминуче веде до збільшення кількості параметрів, а отже, зменшує точність вимірювання й ускладнює інтерпретацію результатів.

Під час обґрунтування типу функції слід ураховувати й той факт, що межі варіації корельованих ознак у конкретних умовах простору й часу, в

конкретній сукупності значно вужчі за їхні можливі значення, і в цих межах варіації навіть лінійна функція може задовільно апроксимувати зв'язок.

У лінійному щодо параметрів рівнянні регресії індивідуальне значення результативного показника  $y_j$  (де  $j$  – порядковий номер одиниці сукупності) можна записати так:

$$y_j = b_0 + \sum_{i=1}^m b_i x_i + e_j,$$

де  $b_0$  – вільний член рівняння, який зазвичай не має економічного змісту і лише окреслює область існування моделі;

$b_i$  – коефіцієнт регресії, що показує, як у середньому змінюється  $y$  зі зміною  $x_i$  на одиницю шкали вимірювання за незмінності інших включених до моделі факторів і за інших однакових умов;

$e_j = y_j - Y_j$  – залишкова величина.

У регресійній моделі основне навантаження припадає на коефіцієнт регресії  $b_i$ , його розглядають як своєрідну міру «очищеного» впливу  $x_i$  на  $y$  і називають ефектом впливу.

Процедура оцінювання параметрів регресійної моделі ґрунтується на методі найменших квадратів (МНК).

Для оцінювання адекватності регресійної моделі використовують:

- стандартне відхилення;
- множинні коефіцієнти детермінації та кореляції;
- часткові коефіцієнти детермінації та кореляції;
- коефіцієнти окремої детермінації;
- критерії перевірки істотності зв'язку.

Вивчення залежностей економічних змінних почнемо з випадку двох змінних. Цей випадок є найбільш простим і його можна розглянути графічно.

**Приклад 5.1** Фірма вирішила використати модель лінійної регресії для визначення залежності вигляду  $y = a + bx$  між річним обсягом продажів і річними витратами на рекламу. За попередні роки було зібрано такі дані:

Обсяг продажів $y_i$ , млн дол.	Витрати на рекламу $x_i$ , тис. дол.
28	71
14	31
19	50
21	60
16	35

Необхідно побудувати лінійну регресійну модель і оцінити її якість.

На рисунку 5.1 (графік побудовано за допомогою модуля *Graphs – Scatterplots*) наведено графік залежності обсягу продажів від витрат на рекламу. Очевидно, що між цими параметрами існує лінійна залежність. Модуль *Graphs – Scatterplots* дає можливість оцінити параметри такого



зв'язку й отримати рівняння лінійної регресії, але лише для найпростіших однофакторних залежностей. Крім того, оцінити якість моделі можна лише приблизно, аналізуючи графік.

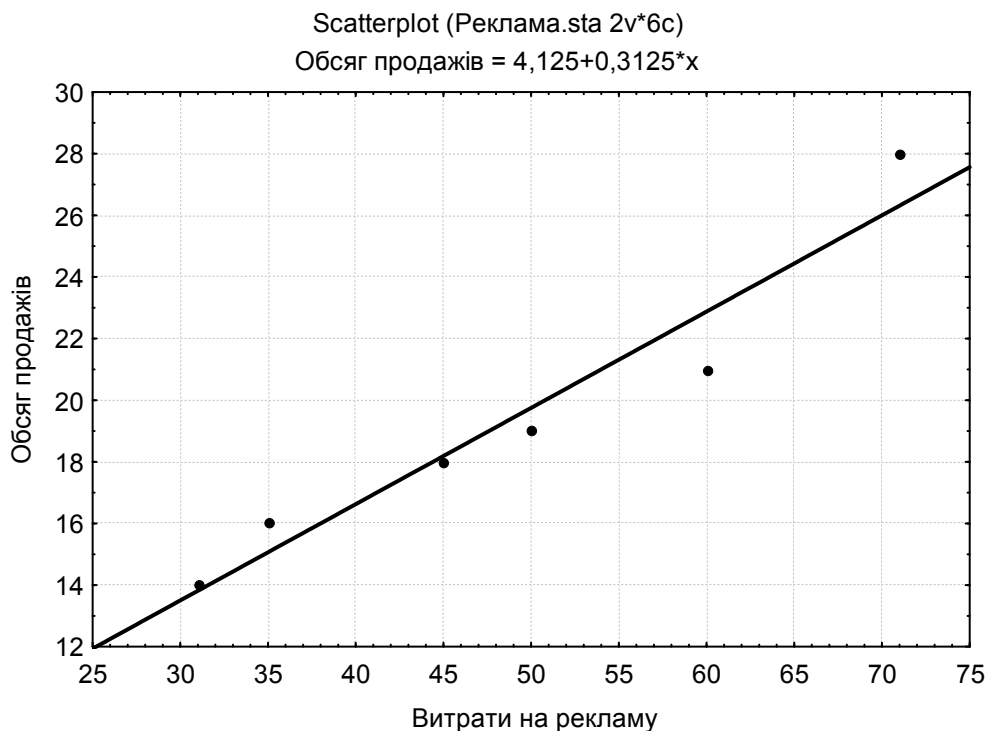


Рисунок 5.1 – Графік залежності між обсягом продажів і витратами на рекламу

У системі *Statistica* регресійний аналіз проводиться в модулі *Множинна регресія – Statistics – Multiply Regression*.

У вікні вибору змінних моделі вибираються залежна (*Dependent var.*) і незалежні (*Independent variable list*) змінні (рисунок 5.2).

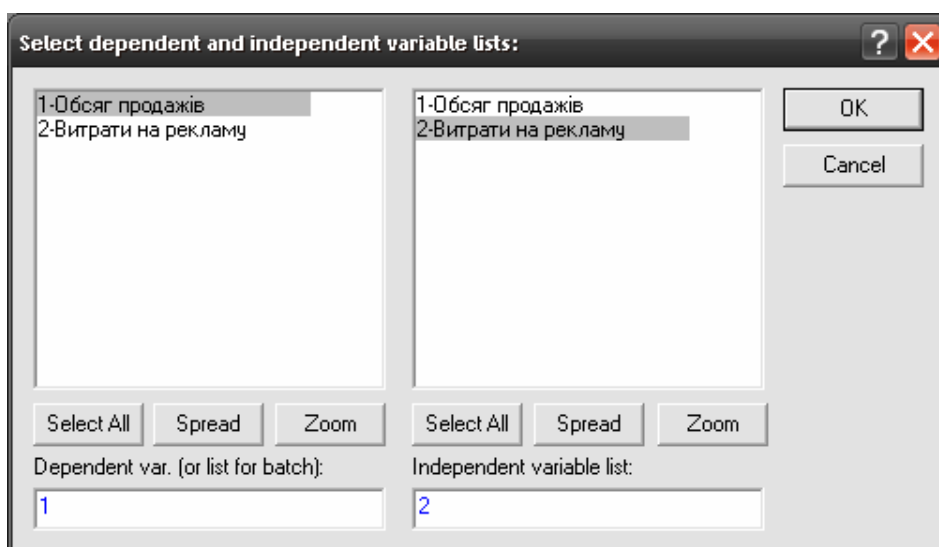


Рисунок 5.2 – Вибір змінних у модулі *Multiply Regression*

Вікно результатів аналізу (рисунок 5.3) має таку структуру: інформаційна панель, яка містить основну інформацію щодо результатів оцінювання, та функціональна панель, яка дає можливість всебічно переглянути результати аналізу.

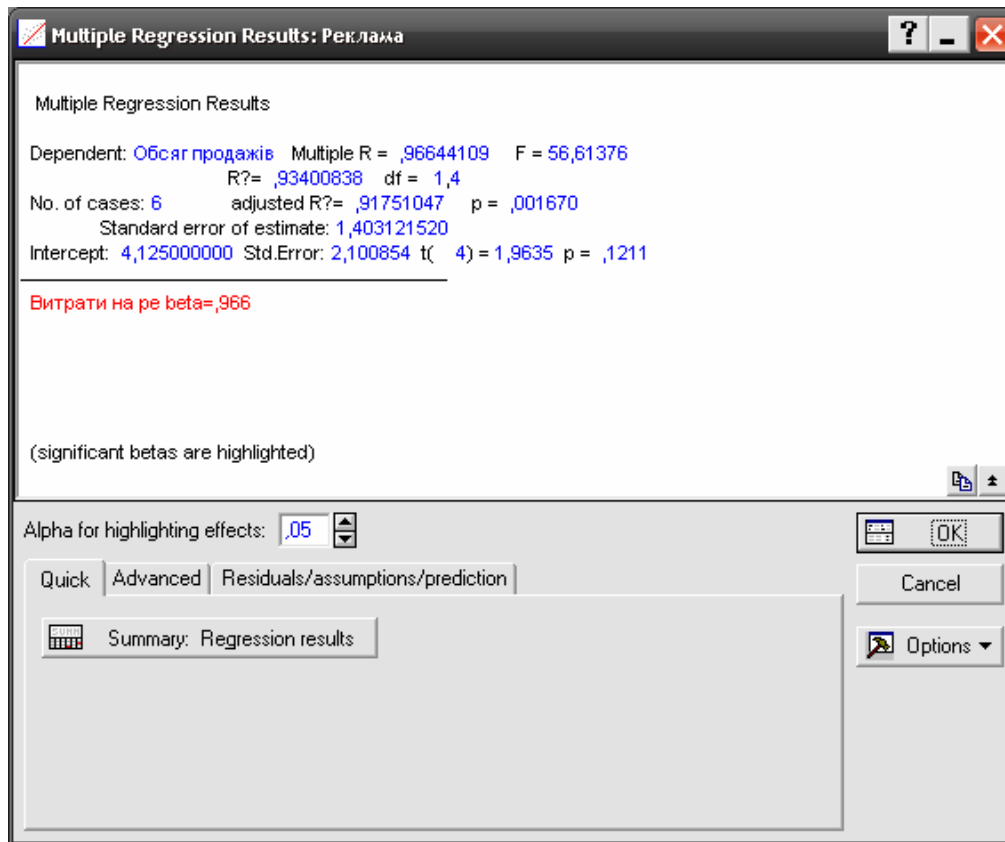


Рисунок 5.3 – Результати аналізу в модулі *Multiple Regression*

На інформаційній панелі наведено такі параметри:

- 1 **Dependent** – ім'я залежної змінної (*Обсяг продажів*).
  - 2 **No. of Cases** – кількість спостережень (6).
  - 3 **Multiple R** – коефіцієнт множинної кореляції (0, 96644109).
  - 4 **R<sup>2</sup>** – квадрат коефіцієнта множинної кореляції (коефіцієнт детермінації) (0,934008383).
  - 5 **Adjusted R<sup>2</sup>** – скоректований коефіцієнт детермінації (0,91751047).
- Відношення факторної суми квадратів до загальної характеризує частку варіації у, пов'язану з варіацією включених до моделі факторів, і має назву множинного коефіцієнта детермінації

$$R^2 = \frac{S_Y^2}{S_y^2} = \frac{S_y^2 - S_e^2}{S_y^2} = 1 - \frac{S_e^2}{S_y^2}.$$

За відсутності зв'язку  $R^2 = 0$ . Якщо зв'язок функціональний, то  $R^2 = 1$ . Очевидно, що  $R^2$  пов'язаний зі стандартним відхиленням  $s_e$ . При зменшенні  $s_e$  значення  $R^2$  збільшуватиметься і навпаки. Корінь квадратний

із коефіцієнта детермінації називають коефіцієнтом кореляції  $R = \sqrt{R^2}$ . У цьому випадку рівняння регресії пояснює 91,75 % варіації змінної *Обсяг продажів*.

Крім названих множинних коефіцієнтів щільності зв'язку у комп'ютерних програмах передбачено обчислення  $R^2$  з урахуванням числа ступенів свободи:

$$R_k^2 = 1 - \frac{s_e^2}{s_y^2} = 1 - (1 - R^2) \frac{n-1}{n-(m+1)},$$

де  $s_y^2 = \frac{\Sigma(y - \bar{y})^2}{n-1}$  – оцінка дисперсії результативної ознаки у;

$s_e^2$  – оцінка залишкової дисперсії.

Скоригований коефіцієнт множинної детермінації  $R_k^2$  відрізняється від  $R^2$  співвідношенням чисел ступенів свободи залишкової ( $n - m + 1$ ) і загальної ( $n - 1$ ) дисперсій.

6 **F** – значення F-критерію (56,61376).

7 **df** – кількість ступенів свободи F-критерію (1,4).

8 **p** – рівень значущості (0,00167).

Перевірка істотності зв'язку статистично формулюється як перевірка нульових гіпотез:  $H_0 : R^2 = 0$ ;  $H_0 : b_i = 0$ . Гіпотеза  $H_0$  відхиляється або вважається допустимою на основі статистичних критеріїв, зокрема дисперсійного F-критерію, статистична характеристика якого розраховується за співвідношенням оцінок факторної і залишкової дисперсій:

$$F = \frac{S_Y^2}{S_e^2} : \frac{m-1}{n-m-1}, \text{ або } F = \frac{R^2}{1-R^2} : \frac{m-1}{n-m-1}.$$

Критичні значення  $F_{1-\alpha}(k_1, k_2)$ , де  $\alpha$  – рівень істотності,  $k_1 = m$ ,  $k_2 = n - m - 1$  – числа ступенів свободи чисельника й знаменника,  $m$  – кількість пояснювальних (незалежних) змінних,  $n$  – кількість спостережень (див. Додаток Б).

Для наведеного рівняння  $\alpha = 0,05$ ,  $m = 1$ ,  $n = 6$ ,  $k_1 = 1$ ,  $k_2 = 4$ ,  $F_{табл} = 7,71$ ,  $F = 56,61376 > 7,71$ , гіпотеза не приймається і можна говорити про істотність зв'язку.

9 **Standard error of estimate** – стандартна похибка оцінки (1,40312). Ця статистика є мірою розкиду значень, що спостерігаються, відносно регресійної прямої.

Стандартне відхилення характеризує варіацію залишкових величин

$$s_e = \sqrt{\frac{\sum_1^n e^2}{n - (m + 1)}},$$

де  $n$  – обсяг сукупності,  $m$  – кількість коефіцієнтів регресії.

Розрахунок характеристик щільності зв'язку ґрунтується на декомпозиції (розкладанні) варіації  $y$  за джерелами формування:

$$S_y^2 = S_X^2 + S_e^2,$$

де  $S_y^2 = \sum_1^n (y - \bar{y})^2$  – загальна сума квадратів відхилень, зумовлена впливом усіх можливих факторів;

$S_Y^2 = \sum_1^n (Y - \bar{y})^2$  – факторна сума квадратів відхилень, зумовлена впливом включених до моделі факторних ознак  $x_i$ ;

$S_e^2 = \sum_1^n (y - Y)^2$  – залишкова сума квадратів відхилень, розмір якої залежить від потужності впливу не включених до моделі факторів.

10 **Intercept** – оцінка вільного члена регресії (4,125), або значення коефіцієнта  $b_0$  у рівнянні регресії.

11 **Std. Error** – стандартна похибка оцінки вільного члена регресії (2,100854).

12 **t(df), p** – **t**-критерій, що використовується для перевірки гіпотези щодо рівності нулю вільного члена регресії; **p** – рівень значущості  $t(4) = 1,9635$ ,  $p = 0,1211$ .

У таблиці 5.1 наведено характеристики регресійної моделі зв'язку між обсягом продажів і витратами на рекламу.

Таблиця 5.1 – Характеристики регресійної моделі

N = 6	Regression Summary for Dependent Variable: Обсяг продажів (Реклама.sta) R= ,96644109 R2= ,93400838 Adjusted R2= ,91751047 F(1,4)=56,614 p<,00167 Std.Error of estimate: 1,4031					
	Beta	Std.Err.	B	Std.Err.	t(4)	p-level
Intercept			4,125000	2,100854	1,963487	0,121070
Витрати на рекламу	0,966441	0,128444	0,312500	0,041533	7,524211	0,001670

Згідно з даними таблиці 5.1 рівняння регресії має такий вигляд:

$$Y = 4,125 + 0,3125x.$$

Зі збільшенням витрат на рекламу на 1 % за умови незмінності інших факторів обсяг продажів збільшується в середньому на 0,3125 %. Включені до моделі фактори пояснюють 91,7 % варіації обсягу продажів.

Паралельно з оцінюванням адекватності моделі проводиться перевірка істотності впливу окремих факторів  $x_i$  на  $y$  за допомогою **t**-критерію:

$$t = \frac{b_i}{\mu_{b_i}},$$

де  $\mu_{b_i} = \sqrt{s_e^2 c_{ii}}$  — стандартна похибка коефіцієнта регресії;

$s_e^2$  — оцінка залишкової дисперсії;

$c_{ii}$  — діагональний елемент оберненої матриці  $C$ .

Критичні значення  $t_{1-\alpha}(k)$ , де  $k = n - 1$ , наведено в Додатку В. Ефект впливу  $i$ -го фактора вважається істотним, якщо  $t_i > t_{табл}(k)$ . Так, при  $\alpha = 0,05$  і  $k = 5$   $t_{табл} = 2,57$ ,  $t(4) = 7,524 > 2,5$  вплив фактора *Витрати на рекламу* є істотним.

**Приклад 5.2** За статистичними даними (таблиця 5.2) про результати господарської діяльності підприємства побудувати множинну лінійну модель залежності показника  $Y$  (чистий прибуток підприємства, млн грн) від трьох факторів:  $X_1$  (основні фонди, млн грн);  $X_2$  (оборотні фонди, млн грн);  $X_3$  (вартість робочої сили, млн грн).

Таблиця 5.2 – Результати господарської діяльності підприємства

Період	Чистий прибуток підприємства, млн грн	Основні фонди, млн грн	Оборотні фонди, млн грн	Вартість робочої сили, млн грн
1-й	13,2	21,4	16,7	12,02
2-й	8,16	19,27	7,77	11,57
3-й	5,14	19,21	11,64	10,99
4-й	7,45	15,34	9,76	10,62
5-й	18,79	24,45	15,84	10,01
6-й	21,52	23,21	17,33	9,55
7-й	25,33	21,12	19,78	9,06
8-й	26,78	24,4	21,88	8,42
9-й	28,72	25,44	19,62	7,95
10-й	30,57	26,87	22,55	7,5
11-й	33,1	25,98	21,58	7,1
12-й	38,65	29,5	20,71	6,67

Результати аналізу наведено на рисунку 5.4 і в таблиці 5.3. Величини коефіцієнта детермінації  $R^2 = 0,959$  і  $F$ -критерію  $= 62,79$  ( $F_{табл} = 4,07$  ( $\alpha = 0,05$ ,  $k_1 = m = 3$ ,  $k_2 = n - m - 1 = 12 - 3 - 1 = 8$ )) свідчать про істотність зв'язку. Згідно зі значенням  $t$ -критерію істотною є лише одна змінна – *Вартість робочої сили*.

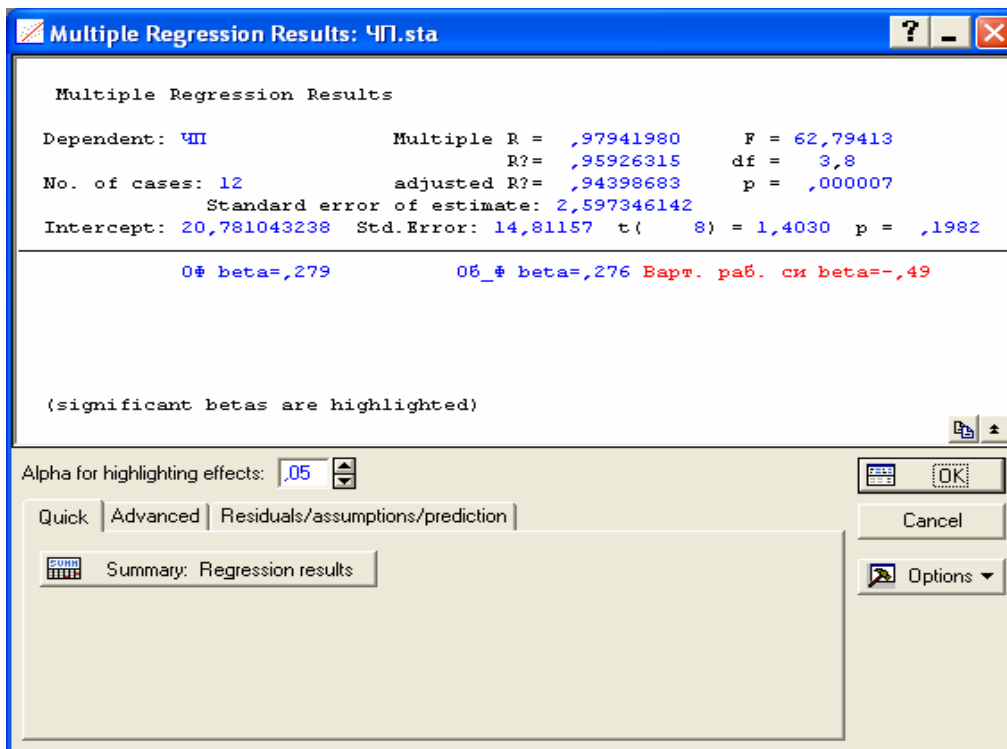


Рисунок 5.4 – Результати аналізу (три фактори)

Таблиця 5.3 – Результати аналізу (три фактори)

N=12	Regression Summary for Dependent Variable: ЧП (ЧП.sta)					
	R= ,97941980 R <sup>2</sup> = ,95926315 Adjusted R <sup>2</sup> = ,94398683 F(3,8)=62,794 p<,00001 Std.Error of estimate: 2,5973					
	Beta	Std.Err.	B	Std.Err.	t(8)	p-level
Intercept			20,78104	14,81157	1,40303	0,198207
ОФ	0,279311	0,139746	0,78154	0,39102	1,99871	0,080678
Об_Ф	0,276084	0,138632	0,60962	0,30611	1,99149	0,081585
Варт. роб. сили	-0,487967	0,135752	-2,98660	0,83087	-3,59456	0,007038

Якщо виключити з аналізу змінну *Основні фонди*, то отримаємо модель зображену на рисунку 5.5.

Модель залишилася такою, що відображає істотний зв'язок, але всі незалежні змінні (вартість оборотних фондів і вартість робочої сили) є істотними. Результати аналізу наведено в таблиці 5.4.

Рівняння регресії має такий вигляд:

$$\text{ЧП} = 40,587 + 0,902 \cdot \text{Об}_\Phi - 3,721 \cdot \text{Варт. роб. сили}.$$

Збільшення вартості оборотних фондів на 1 % у середньому приводить до збільшення чистого прибутку на 0,902 %, а збільшення вартості робочої сили на 1 % – до зменшення чистого прибутку на 3,721 %. Рівняння пояснює 93,89 % варіації чистого прибутку. Всі коефіцієнти рівняння є істотними.

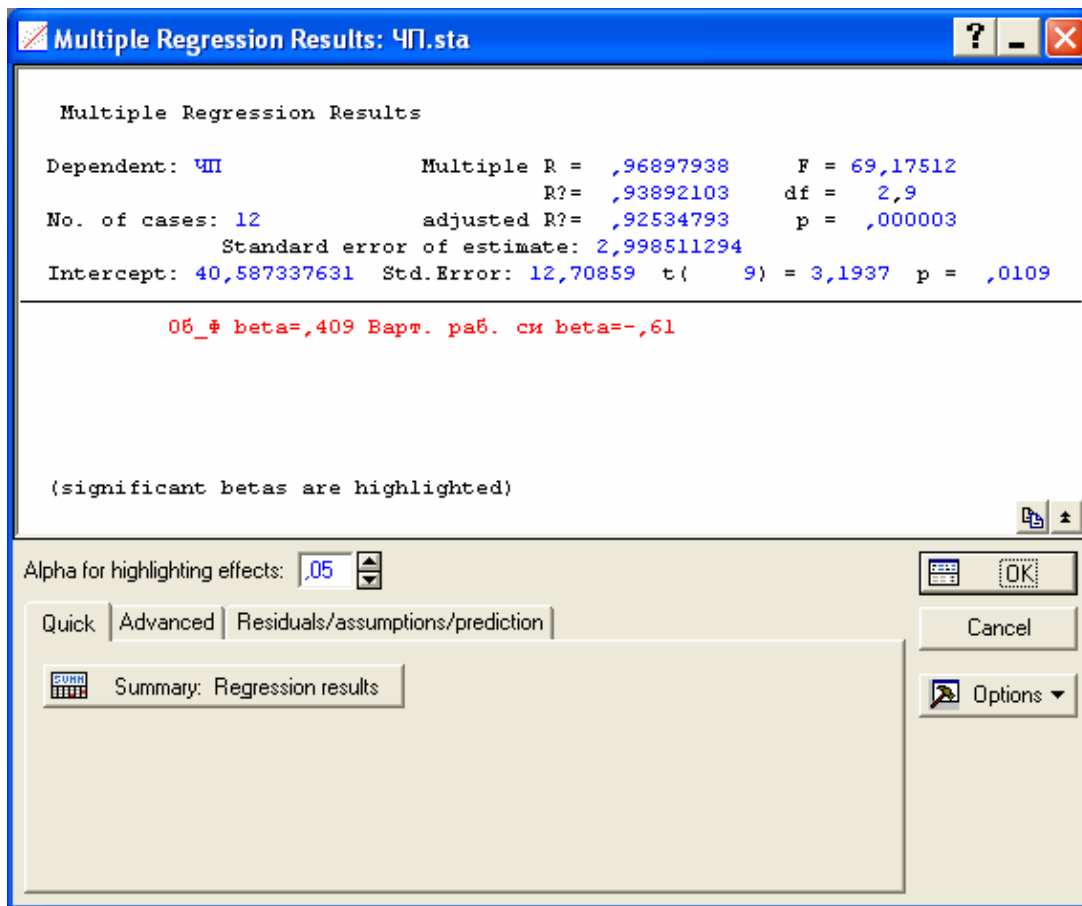


Рисунок 5.5 – Результати аналізу (два фактори)

Таблиця 5.4 – Результати аналізу (два фактори)

N=12	Regression Summary for Dependent Variable: ЧП (ЧП.sta) R= ,96897938 R²= ,93892103 Adjusted R²= ,92534793 F(2,9)=69,175 p<,00000 Std.Error of estimate: 2,9985					
	Beta	Std.Err.	B	Std.Err.	t(8)	p-level
Intercept			40,58734	12,70859	3,19369	0,010942
Об_ф	0,408532	0,140576	0,90208	0,31040	2,90614	0,017420
Варт. роб. сили	-0,607907	0,140576	-3,72069	0,86039	-4,32441	0,001921

## 5.2 Оцінювання якості регресійної моделі

Адекватність регресійної моделі означає її здатність правильно описати реальну структуру взаємозв'язків між ознаками  $x_i$  та  $y$ . Методологічною основою вирішення проблеми адекватності є теоретичний змістовний аналіз матеріальної природи процесу (явища) й обґрунтування типу і структури моделі, яка описує механізм його формування. Практично з метою забезпечення адекватності моделі змістовний аналіз поєднується з формальними процедурами перевірки гіпотез щодо дотримання логіко-

статистичних умов використання МНК.

Мірою адекватності моделі, як уже зазначалося, є відхилення фактичних значень від теоретичних  $e_j = y_j - Y_j$ . На величину цих відхилень впливає весь комплекс умов, зокрема:

- обсяг й однорідність сукупності;
- незалежність спостережень;
- інформативність включених до моделі факторів;
- стабільність не включених до моделі факторів;
- тип моделі.

Репрезентативність оцінок регресійного аналізу прямо пропорційна обсягу й однорідності сукупності. Саме недостатній обсяг сукупності та її неоднорідність вважаються найвагомими факторами неадекватності моделей. Тому під час формування ознакової множини моделі слід урахувати співвідношення між обсягом вибірки й кількістю включених до моделі факторів (воно має бути приблизно 8 : 1). Деякі джерела рекомендують співвідношення 3:1.

Оцінювання однорідності сукупності здійснюється на етапі розвідувального аналізу даних. Так, наявність аномальних значень, які не узгоджуються з розподілом основної маси даних, може бути наслідком помилок спостереження або результатом незвичайної комбінації причин і умов, у яких функціонує одиниця сукупності. Ідентифікація таких спостережень дає можливість усунути помилки, а якщо це неможливо, то вилучити аномальний об'єкт з подальшого аналізу.

Інформативність включених до моделі факторних ознак залежить як від соціально-економічного змісту, так і від шкали вимірювання ознаки. Якщо ознака за змістом не є інформативною, то ніякий спосіб моделювання не забезпечить необхідних результатів. Так само результати аналізу будуть суттєво різнитися залежно від того, за якою шкалою визначено одну й ту саму ознаку (за метричною, ранговою або номінальною).

Важливою умовою регресійного аналізу є відсутність мультиколінеарності, яка приводить до зміщення оцінок параметрів моделі та унеможлиблює коректну інтерпретацію результатів.

**Мультиколінеарність** – це корельованість двох або декількох пояснювальних змінних у рівнянні регресії. Проблема мультиколінеарності виникає тільки для випадку множинної регресії, оскільки в парній регресії лише одна пояснювальна змінна. Оцінка коефіцієнта регресії може виявитися незначущою не тільки через неістотність цього фактора, але й через те, що важко розмежувати дію на залежну змінну двох або декількох факторів. Це виникає у тому випадку, коли деякі фактори лінійно зв'язані між собою (корельовані) і змінюються синхронно. Зв'язок залежної змінної зі змінами кожної з них можна визначити, тільки якщо до пояснювальних змінних включається лише один з цих факторів.



Мультиколінеарність змінних  $x_1$  і  $x_2$  характеризується величиною коефіцієнта кореляції між ними. Чим ближче за абсолютною величиною значення коефіцієнта кореляції до одиниці, тим ближче мультиколінеарність до досконалої, тим важче поділити впливи пояснювальних змінних  $x_1$  і  $x_2$  на поведінку змінної  $y$  і тим менш надійними будуть оцінки коефіцієнтів регресії при цих змінних.

У загальному випадку якщо під час оцінювання рівняння регресії декілька факторів виявилися незначущими, то потрібно з'ясувати, чи немає серед них сильно корельованих між собою. Для цього розраховується кореляційна матриця і перевіряється статистична значущість коефіцієнтів парної кореляції. За наявності кореляції один з пари зв'язаних між собою факторів виключається або як пояснювальний фактор береться якась їхня функція. Якщо ж незначущим виявився тільки один фактор, то можна його виключити або замінити іншим (хоча, можливо, на якомусь коротшому проміжку часу цей фактор виявився б значущим).

Так, для прикладу 5.2 обґрунтуємо виключення з моделі однієї з пояснювальних змінних. Коефіцієнти кореляції, розраховані для змінних моделі, наведено в таблиці 5.5.

Таблиця 5.5 – Коефіцієнти кореляції для змінних регресійної моделі

Показник	ОФ	Об_Ф	Варт. роб. сили
ОФ	1,00	0,82	-0,81
Об_Ф	0,82	1,00	-0,81
Варт. роб. сили	-0,81	-0,81	1,00

Аналіз кореляцій підтверджує, що існує тісний лінійний зв'язок між окремими змінними моделі.

### **5.3 Перевірка умов, виконання яких передбачалося під час оцінювання рівняння регресії. Статистика Дарбіна – Уотсона**

Близьке до одиниці значення коефіцієнта детерміації  $R^2$  ще не є підтвердженням високої якості рівняння регресії. Тому наступний етап перевірки якості рівняння регресії – перевірка деяких важливих умов, виконання яких передбачалося під час оцінювання рівняння регресії.

Перед оцінюванням лінійного рівняння регресії припускалося, що реальний взаємозв'язок змінних лінійний, а відхилення від регресійної прямої випадкові, незалежні між собою і мають нульове середнє значення і стали дисперсію. Чи так це насправді? Якщо ні, то аналіз статистичної значущості коефіцієнтів регресії неточний і оцінки цих коефіцієнтів не

мають таких властивостей, як незміщеність, обґрунтованість і ефективність.

Спробуємо відповісти на запитання, в яких випадках відхилення не мають властивостей, що передбачалися. По-перше, це має статися, якщо насправді досліджуваний взаємозв'язок нелінійний. Це відбувається, якщо відхилення від лінії регресії не випадково розподілені навколо неї, а мають певну закономірність. Ця закономірність, зокрема, виражається в однаковому знаку кожних двох сусідніх відхилень. Це може бути наслідком нелінійного характеру зв'язку змінних або дією якогось фактора, що не ввійшов до рівняння регресії. Величина такого неврахованого фактора може змінювати свою динаміку в певний період, відхиляючись у досить тривалі проміжки часу в той чи інший бік від свого середнього значення. Це, очевидно, може бути причиною тривалих стійких відхилень залежної змінної від лінії регресії. Обидві вказані причини свідчать про те, що існує можливість поліпшити рівняння регресії шляхом оцінювання якоїсь нової нелінійної формули або включення деякої нової пояснювальної змінної.

Для оцінювання ступеня такої невідповідності необхідно ввести кількісну міру. Отже, однією з основних передбачуваних властивостей відхилень  $e_i$  значень  $y_i$  від регресійної формули  $y = a + bx$  є їхня статистична незалежність одна від одної. При цьому перевіряється зазвичай їхня некорельованість (що є необхідним, але недостатнім атрибутом незалежності), причому некорельованість не будь-яких, а сусідніх величин  $e_i$ . Сусідніми можна вважати сусідні в часі (у разі часових рядів) або за збільшенням змінної  $x$  (у разі перехресних вибірок) значення  $e$ . Для цих величин можна розрахувати, наприклад, коефіцієнт кореляції (так званий коефіцієнт автокореляції першого порядку):

$$r_{i,i-1} = \frac{\sum_i e_i e_{i-1}}{\sqrt{\sum_i e_i^2 \sum_i e_{i-1}^2}} \quad (\text{вважаємо, що } M|e_i| = 0).$$

На практиці, проте, використовують тісно зв'язану з  $r_{i,i-1}$  статистику Дарбіна – Уотсона  $DW$ , що розраховується за формулою

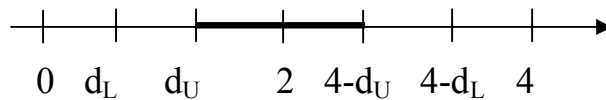
$$DW = \frac{\sum_i (e_i - e_{i-1})^2}{\sum_i e_i^2}, \quad 0 < DW < 4.$$

У разі, коли кожне відхилення  $e$  приблизно збігається з попереднім відхиленням  $e_i$ , статистика Дарбіна – Уотсона виявиться близькою до нуля. Це випадок додатної автокореляції залишків першого порядку. Інший крайній випадок виникає, коли точки спостережень по черзі відхиляються в різні боки від лінії регресії і кожне наступне відхилення  $e$  має протилежний знак, ніж попереднє відхилення  $e_{i-1}$ . Це випадок від'ємної автокореляції, оскільки при цьому  $DW = 4$ . Остання досить рідко має місце в

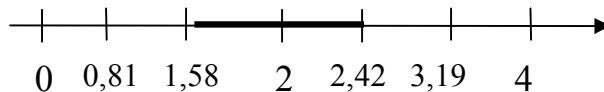
економічному аналізі. Якщо розглядаються часові ряди з річними даними, то таку закономірність поведінки послідовних відхилень досить важко проінтерпретувати. Проте вона може мати місце під час роботи, наприклад, з піврічними даними показників з сезонним характером зміни. Нарешті, якщо характер поведінки відхилень випадковий, можна припустити, що в половині випадків знаки послідовних відхилень збігаються, а в половині – ні. Оскільки їхня абсолютна величина в середньому передбачається однаковою, можна вважати, що тут в половині випадків  $e_i$  дорівнює  $e_{i-1}$ , а в іншій половині –  $-e_{i-1}$ . При цьому  $DW = 2$ . Це показує, що близькість статистики Дарбіна – Уотсона до двох є необхідною умовою випадкового характеру відхилень від лінії регресії. Потрібно, проте, мати на увазі, що під час визначення показника  $DW$  порівнюються тільки сусідні відхилення від регресії, водночас цикли зміни економічних змінних можуть бути тривалішими, ніж одна одиниця часу. Наприклад, якщо розглядаються поквартальні дані сільськогосподарського виробництва (що має річний цикл) і оцінюється їхня лінійна регресія від часу, статистика Дарбіна – Уотсона може бути близькою до двох при регулярному відхиленні залежної змінної від лінії регресії.

Якщо статистика Дарбіна – Уотсона близька до двох, то відхилення від регресії вважаються випадковими (хоча насправді вони можуть і не бути такими). Це означає, що лінійна функція, ймовірно, відображає реальний взаємозв'язок; найімовірніше, не залишилося істотних неврахованих факторів, які впливають на залежну змінну, і яка-небудь інша нелінійна формула не перевершує за статистичними характеристиками цю лінійну. Навіть якщо частка дисперсії залежної змінної, поясненої за допомогою регресії, при цьому мала, то можна чекати, що інша частина цієї дисперсії, яка залишилася непоясненою, породжена дією безлічі різних малих факторів і може бути описана як випадкова нормальна похибка.

Але як визначити, чи достатньо близька величина статистики  $DW$  до двох? Для цього є спеціальні таблиці (див. Додаток Д), що дають можливість при певній кількості спостережень, пояснювальних змінних і заданому рівні значущості знайти критичні значення статистики Дарбіна – Уотсона. Для цього за таблицями знаходяться довірчі інтервали, в межах яких нульова гіпотеза приймається, відкидається або не може бути прийнятою або знехтуваною. Важливо, що для статистики Дарбіна – Уотсона існують два критичних значення, менших за два: нижнє  $d_L$  як межа для визнання додатної автокореляції залишків і верхнє  $d_U$  як межа визнання її відсутності. Для перевірки гіпотези про від'ємну автокореляцію залишків ці критичні значення розташовуються симетрично відносно числа 2:



Оцінимо статистику Дарбіна – Уотсона для регресійної моделі з прикладу 5.2. У системі *Statistica* значення  $DW$  розраховується в модулі *Multiply Regression (Residual Analysis)*. У цьому випадку  $DW = 1,755$ . За таблицею визначимо, що для двох змінних моделі й 12 спостережень  $d_L = 0,81$ ,  $d_U = 1,58$ :



Гіпотеза про відсутність автокореляції залишків приймається, якщо  $DW$  знаходиться в межах  $d_U < DW < 4 - d_U$  ( $1,58 < DW < 2,42$ ), і відкидається, якщо  $DW < d_L$  або  $DW > 4 - d_L$  ( $DW < 0,81$ ,  $DW > 3,19$ ). В інших випадках гіпотезу не можна прийняти або відкинути. Для прикладу 5.2 значення  $DW$  становить  $1,58 < 1,75 < 2,42$ , отже, гіпотеза про відсутність автокореляції залишків приймається.

Як у загальному випадку визначити приблизно критичні величини статистики  $DW$ ? Дуже грубо в першому наближенні можна сказати, що при достатньої кількості спостережень (не менше 12 – 15) і при 1 – 3 пояснювальних змінних значення  $DW$  має бути не меншим за одиницю (і не більшим за три). Інакше слід визнати існування автокореляції залишків і поліпшити формулу. Якщо статистика знаходиться приблизно між 1,2-1,3 і 2,7-2,8, можна вважати, що статистично значущої автокореляції залишків немає. В іншому випадку достатньо надійний висновок зробити не можна.

Отже, якщо статистика Дарбіна – Уотсона становить 1,5-2,0-2,5, це означає, що відхилення від лінії регресії взаємно незалежні.

У разі наявності автокореляції залишків отримана формула регресії вважається зазвичай незадовільною. Можна пошукати іншу (нелінійну) формулу, включити невраховані до цього змінні, уточнити період проведення розрахунків або поділити його на декілька частин, або застосувати до даних перетворення, які зменшують автокореляцію залишків (наприклад, автокореляційне перетворення або метод ковзних середніх).

### Запитання для самоперевірки

- 1 Що описує регресійна модель? Що є метою регресійного аналізу?
- 2 Які показники використовуються для аналізу якості моделі лінійної регресії?
- 3 З яких етапів складається перевірка якості рівняння регресії?
- 4 Як розраховується і що показує коефіцієнт детермінації  $R^2$ ?

5 Коефіцієнт детермінації  $R^2 = 0,5$ . Що можна сказати про якість рівняння регресії в цілому? Яка потрібна додаткова інформація?

6 Поясніть значення терміна „адекватність регресійної моделі”. Які фактори впливають на адекватність моделі? Які рекомендації щодо обсягу сукупності даних для аналізу існують?

7 Що таке мультиколінеарність? Яким чином можна вирішити проблему мультиколінеарності змінних лінійної регресійної моделі?

8 Яким чином можна вирішити проблему наявності автокореляції залишків?

9 Які показники характеризують незалежність відхилень залежної змінної від лінії регресії? Як здійснюється перевірка цієї незалежності?

10 Що показує статистика Дарбіна – Уотсона? У яких межах має знаходитись статистика для того, щоб отримана формула регресії була задовільною?

11 У яких випадках спостерігається додатна автокореляція відхилень  $e$ ? Наведіть приклади з економіки.

12 Статистика Дарбіна – Уотсона виявилася близькою до чотирьох. Що це означає?

13 Регресійна модель описує залежність продуктивності праці робітників (тонн на одного робітника за зміну) від потужності вугільного пласта ( $x_1$ ) і типу вугільного комбайна. Параметри моделі становлять:

$R^2$	$a_0$	$b_1$	$a_1$	$c_{11}$
0,76	3,872	1,875	1,290	2,754

Поясніть зміст параметрів моделі.

14 За наведеними нижче даними щодо залежності обсягу продажів  $y$  (тис. од.) від витрат на рекламу  $x_1$  (тис. дол.), кількості мереж роздрібної торгівлі, які продають товар у регіоні,  $x_2$  (од.) і рівня безробіття в регіоні  $x_3$  (%) побудуйте регресійну модель і оцініть її якість:

Регіони	$y$	$x_1$	$x_2$	$x_3$	Регіони	$y$	$x_1$	$x_2$	$x_3$
1-й	16	22	2	4	6-й	21	50	3	10
2-й	17	26	2	8	7-й	32	56	6	8
3-й	26	45	3	7	8-й	18	34	3	8
4-й	24	37	4	0	9-й	30	60	5	2
5-й	22	28	4	2	10-й	20	40	3	6

## БІБЛІОГРАФІЧНИЙ СПИСОК

Анализ статистической совокупности в программе MS Excel [Текст] / сост. А. Н. Акжигитова, Н. С. Циндин, Н. Ф. Разуваева. – Пенза : Информ.-изд. центр ПГУ, 2007. – 52 с.

Андреев, Г.И. В помощь написания диссертации и рефератов: основы научной работы и оформление результатов научной деятельности [Текст]: учеб. пособие / Г.И. Андреев, С.А. Смирнов, В.А. Тихомиров. – М.: Финансы и статистика, 2004. – 272 с.

Боровиков, В.П. Популярное введение в программу STATISTICA [Текст] / В.П. Боровиков. – М.: КомпьютерПресс, 1998. – 267 с.

Боровиков, В. П. Statistica® – статистический анализ и обработка данных в среде Windows® [Текст] / В. П. Боровиков, И. П. Боровиков. – М.: Информ.-изд. дом «Филинь», 1998. – 608 с.

Голикова, Ж.А. Книга для чтения по английскому языку для студентов технических вузов [Текст]: учеб. пособие / Ж.А. Голикова, Л.В. Гусарова. – Минск: Вышэйш. шк., 1990. – 130 с.

Єріна, А. М. Статистичне моделювання та прогнозування [Текст]: навч. посіб. / А. М. Єріна. – К.: КНЕУ, 2001. – 170 с.

Замков, О.О. Математические методы в экономике [Текст]: учебник / О.О. Замков, А.В. Толстопятенко, Ю.Н. Черемных. – 2-е изд. – М.: Дело и Сервис, 1999. – 368 с.

Ковалев, А. И. Прологомены к методам научных исследований [Текст]: учеб. пособие / А. И. Ковалев. – Х.: Инжэк, 2005. – 312 с.

Орлов, А.И. Экспертные оценки [Текст]: учеб. пособие / А.И. Орлов. – М.: ИВСТЭ, 2002. – 31 с.

Основы научной речи [Текст]: учеб. пособие для студ. нефилол. высш. учеб. заведений / Н.А. Буре, М.В. Быстрых, С.А. Вишнякова и др.; под ред. В.В. Химика, Л.Б. Волковой. – СПб.: филол. ф-т СПбГУ; М.: Изд. центр «Академия», 2003. – 272 с.

Основы научных исследований [Текст]: учеб. пособие / под ред. А.А. Лудченко. – 2-е изд. – К.: Знання, 2001. – 113 с.

Плюта, В. Сравнительный многомерный анализ в эконометрическом моделировании [Текст]: пер. с польск. / В. Плюта. – М.: Финансы и статистика, 1989. – 175 с.

Практикум з теорії ймовірностей та математичної статистики [Текст]: навч. посіб. / Р.К. Чорней, О.Ю. Дюженкова, О.Б. Жильцов та ін.; за ред. Р.К. Чорнея. – К.: МАУП, 2003. – 328 с.

Словарь-справочник автора [Текст] / Л.А. Гильберг. – М.: Книга, 1979. – 301 с.

Таганов, Д.Н. SPSS: Статистический анализ в маркетинговых исследованиях [Текст] / Д.Н. Таганов. – СПб.: Питер, 2005. – 192 с.

Факторный, дискриминантный и кластерный анализ [Текст]: пер. с англ. / Дж.-О. Ким, Ч. У. Мьюллер, У. Р. Клекка и др.; под ред. И. С. Енюкова. – М.: Финансы и статистика, 1989. – 215 с.

Цехмістрова, Г.С. Основи наукових досліджень [Текст]: навч. посіб. / Г.С. Цехмістрова. – К.: Вид. дім «Слово», 2003. – 240 с.

Як підготувати і захистити дисертацію на здобуття наукового ступеня [Текст]: метод. поради / упоряд. Л.А. Пономаренко. – К.: Толока, 2001. – 80 с.

## ПРЕДМЕТНИЙ ПОКАЖЧИК

- F*-критерій, 82  
*t*-критерій, 83  
Автореферат, 14  
Анотація, 16  
Безпосереднє оцінювання, 35  
Групова експертиза, 34  
Дискримінантний аналіз, 59  
Дискусія, 31  
дисперсія, 39  
Доповідь, 19  
Експерт, 27  
Коефіцієнт аргументації, 32  
Коефіцієнт варіації, 39  
Коефіцієнт детермінації, 81  
Коефіцієнт ексцесу, 41  
Коефіцієнт кореляції, 45  
Коефіцієнт обізнаності, 32  
Компетентність експертів, 32  
Критерій Граббса, 51  
Логістична регресія, 59  
Мета наукового дослідження, 5  
Метод Дельфі, 30  
Метод центру тяжіння, 73  
Методи вимірювань, 28  
Методи експертного оцінювання, 29  
Методичні рекомендації, 15  
Мозковий штурм, 30  
Монографія, 13  
Мультіколінеарність, 87  
Наука, 5  
Наукова інформація, 21  
Наукова стаття, 15  
Описова статистика, 37  
Показники варіації, 38  
Показники особливостей форми розподілу, 40  
Попарні порівняння, 35  
Ранжування, 33  
Регресійна модель, 78  
Реферат, 13  
Середнє квадратичне (стандартне) відхилення, 39  
Тези, 15  
Типи шкал, 28  
Універсальна десяткова класифікація, 22  
Фактор, 65  
Факторний аналіз, 64  
Факторні ваги, 66  
Факторні навантаження, 65



## Додаток А

### Критичні значення критерію Граббса $T_n$

Номер спостереження	$\alpha = 0,05$	$\alpha = 0,10$	Номер спостереження	$\alpha = 0,05$	$\alpha = 0,10$
3	1,41	1,41	15	2,49	2,33
4	1,69	1,65	16	2,52	2,35
5	1,87	1,79	17	2,55	2,38
6	2,00	1,89	18	2,58	2,40
7	2,09	1,97	19	2,60	2,43
8	2,17	2,04	20	2,62	2,45
9	2,24	2,10	21	2,64	2,47
10	2,29	2,15	22	2,66	2,49
11	2,34	2,19	23	2,68	2,50
12	2,39	2,23	24	2,70	2,52
13	2,43	2,26	25	2,72	2,54
14	2,46	2,30	26	2,73	2,55

## Додаток Б

### Квантилі $F$ -розподілу ( $\alpha = 0,05$ )

$k_2$	$k_1$									
	1	2	3	4	5	6	7	8	9	10
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64
8	5,32	4,46	4,07	3,84	3,69	3,59	3,50	3,44	3,39	3,35
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99
120	3,92	3,07	2,68	2,45	2,29	2,17	2,09	2,02	1,96	1,91
$\infty$	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,88	1,83

## Додаток В

Квантилі  $t$ -розподілу Стьюдента  $t_{1-0,05}(k)$ :  $|t|$  – двобічний критерій;  $t$  –  
однобічний критерій

$k$	$ t $	$t$	$k$	$ t $	$t$
5	2,57	3,04	18	2,10	2,17
6	2,45	2,78	20	2,09	2,15
7	2,37	2,62	25	2,06	2,11
8	2,31	2,51	30	2,05	2,08
9	2,26	2,43	40	2,02	2,05
10	2,23	2,37	50	2,01	2,03
11	2,20	2,33	60	2,00	2,02
12	2,18	2,29	100	1,98	1,99
14	2,15	2,24	$\infty$	1,96	1,96
16	2,12	2,20			

## Додаток Д

### Критичні значення критерію Дарбіна – Уотсона

<i>n</i>	Кількість факторів											
	<i>m</i> = 1		<i>m</i> = 2		<i>m</i> = 3		<i>m</i> = 4		<i>m</i> = 5		<i>m</i> = 6	
	<i>d<sub>L</sub></i>	<i>d<sub>U</sub></i>	<i>d<sub>L</sub></i>	<i>d<sub>U</sub></i>	<i>d<sub>L</sub></i>	<i>d<sub>U</sub></i>	<i>d<sub>L</sub></i>	<i>d<sub>U</sub></i>	<i>d<sub>L</sub></i>	<i>d<sub>U</sub></i>	<i>d<sub>L</sub></i>	<i>d<sub>U</sub></i>
6	0,61	1,40										
7	0,70	1,36	0,47	1,90								
8	0,76	1,33	0,56	1,78	0,37	2,29						
9	0,82	1,32	0,63	1,70	0,46	2,13	0,30	2,59				
10	0,88	1,32	0,70	1,64	0,53	2,02	0,38	2,41	0,24	2,82		
11	0,93	1,32	0,76	1,60	0,60	1,93	0,44	2,28	0,32	2,65	0,12	2,89
12	0,97	1,33	0,81	1,58	0,66	1,86	0,51	2,18	0,38	2,51	0,16	2,67
13	1,01	1,34	0,86	1,56	0,72	1,82	0,57	2,09	0,45	2,39	0,21	2,49
14	1,05	1,35	0,91	1,55	0,77	1,78	0,63	2,03	0,51	2,30	0,26	2,35
15	1,08	1,36	0,95	1,54	0,81	1,75	0,69	1,98	0,56	2,22	0,30	2,24
16	1,11	1,37	0,98	1,54	0,86	1,73	0,73	1,94	0,62	2,16	0,35	2,15
18	1,16	1,39	1,05	1,54	0,93	1,70	0,82	1,87	0,71	2,06	0,44	2,02
20	1,20	1,41	1,10	1,54	1,00	1,68	0,89	1,83	0,79	1,99	0,52	1,92
22	1,24	1,43	1,15	1,54	1,05	1,66	0,96	1,80	0,86	1,94	1,57	1,85
25	1,29	1,45	1,21	1,55	1,12	1,65	1,04	1,77	0,95	1,89	1,68	1,78
28	1,33	1,48	1,26	1,56	1,18	1,65	1,10	1,75	1,03	1,85	0,76	1,73
32	1,37	1,50	1,31	1,57	1,24	1,65	1,18	1,73	1,11	1,82	0,86	1,69
35	1,40	1,52	1,34	1,58	1,28	1,65	1,22	1,73	1,16	1,80	0,91	1,67
40	1,44	1,54	1,39	1,60	1,34	1,66	1,29	1,72	1,23	1,79	1,00	1,64
45	1,48	1,57	1,43	1,62	1,38	1,67	1,34	1,72	1,29	1,78	1,07	1,64
50	1,50	1,59	1,46	1,63	1,42	1,67	1,38	1,72	1,34	1,77	1,12	1,64
60	1,55	1,62	1,51	1,65	1,48	1,69	1,44	1,73	1,41	1,77	1,21	1,64
70	1,58	1,64	1,55	1,67	1,53	1,70	1,49	1,74	1,46	1,77	1,28	1,65
80	1,61	1,66	1,59	1,69	1,56	1,72	1,53	1,74	1,51	1,77	1,34	1,65
90	1,64	1,68	1,61	1,70	1,59	1,73	1,57	1,75	1,54	1,78	1,38	1,66
100	1,65	1,69	1,63	1,72	1,61	1,74	1,59	1,76	1,57	1,78	1,42	1,67
150	1,72	1,75	1,71	1,76	1,69	1,77	1,68	1,79	1,67	1,80	1,54	1,71
200	1,76	1,78	1,75	1,79	1,74	1,80	1,73	1,81	1,72	1,82	1,61	1,74

Навчальне видання

**Гусєва Юлія Юріївна  
Гребенікова Олена Володимирівна**

**ТЕОРЕТИЧНІ ТА ПРАКТИЧНІ АСПЕКТИ  
ФІНАНСОВО-ЕКОНОМІЧНИХ ДОСЛІДЖЕНЬ**

Редактор А.М. Ємленінова

Зв. план, 2011

Підписано до видання 20.10.2011

Ум. друк. арк. 5,6. Обл.-вид. арк. 6,25. Електронне видання

---

Національний аерокосмічний університет ім. М.Є. Жуковського

«Харківський авіаційний інститут»

61070, Харків-70, вул. Чкалова, 17

<http://www.khai.edu>

Видавничий центр «ХАІ»

61070, Харків-70, вул. Чкалова, 17

[izdat@khai.edu](mailto:izdat@khai.edu)

Свідоцтво про внесення суб'єкта видавничої справи до Державного реєстру видавців, виготовлювачів і розповсюджувачів видавничої продукції, серія ДК № 391, видане Державним комітетом інформаційної політики, телебачення та радіомовлення України від 30.03.2001 р.