

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ**

**ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ ЕКОНОМІЧНИЙ УНІВЕРСИТЕТ  
ІМЕНІ СЕМЕНА КУЗНЕЦЯ**

**Методичні рекомендації  
до виконання лабораторних робіт  
з навчальної дисципліни**

**"МАТЕМАТИЧНА СТАТИСТИКА  
ТА ОПРАЦЮВАННЯ ДАНИХ"**

**для студентів галузі знань  
0515 "Видавничо-поліграфічна справа"  
денної форми навчання**

**Харків. ХНЕУ ім. С. Кузнеця, 2015**

Затверджено на засіданні кафедри економічної кібернетики.  
Протокол № 7 від 22.12.2014 р.

**Укладачі:** Клебанова Т. С.  
Гур'янова Л. С.  
Чаговець Л. О.

М 54        Методичні рекомендації до виконання лабораторних робіт з навчальної дисципліни "Математична статистика та опрацювання даних" для студентів галузі знань 0515 "Видавничо-поліграфічна справа" денної форми навчання / уклад. Т. С. Клебанова, Л. С. Гур'янова, Л. О. Чаговець. – Х. : ХНЕУ ім. С. Кузенця, 2015. – 64 с. (Укр. мов.)

Подано методичні рекомендації до виконання лабораторних робіт, метою яких є закріплення теоретичних знань та набуття навичок побудови моделей математичної статистики та порівняльного аналізу багатовимірних об'єктів.

Рекомендовано для студентів спеціальностей 8.05150102 "Технології електронних мультимедійних видань", 8.05150103 "Комп'ютерні технології та системи видавничо-поліграфічних виробництв".

© Харківський національний економічний  
університет імені Семена Кузнеця, 2015

## Вступ

У мінливих умовах розвитку вітчизняної економіки значно зросла роль фактору невизначеності в системах прийняття рішень. Однією з основних причин чого є невизначеність соціально-економічних процесів. Це приводить до необхідності обліку імовірнісного характеру впливів при формуванні управлінських рішень.

Невизначеність зовнішнього середовища істотно впливає на збільшення ризиків і втрат під час прийняття управлінських рішень, обертається кризовим станом соціально-економічних систем. Таким чином, можливість прийняття рішень при управлінні економічними об'єктами в умовах неповної інформації має ґрунтуватися на використанні нового інструментарію економіко-математичного моделювання, що припускає зменшення первісної невизначеності завдяки підвищенню якості інформації. Так, наявність множини вихідних ознак, що характеризують економічні системи як багатомірні об'єкти, викликає необхідність відбирати найбільш істотні і вивчати менший набір показників. При цьому кожний з об'єктів характеризується великою кількістю різних і стохастично зв'язаних ознак.

**Мета** методичної розробки – формування практичних навичок формалізації завдань порівняльного аналізу багатомірних об'єктів з використанням спеціалізованих методів та моделей математичної статистики.

**Об'єктом** вивчення навчальної дисципліни "Математична статистика і опрацювання даних" є соціально-економічні системи різного рівня ієрархії.

**Предметом** навчальної дисципліни є сукупність методів і моделей математичної статистики та порівняльного аналізу багатомірних об'єктів.

Дана робота містить методичні рекомендації, приклади розв'язання завдань та варіанти для індивідуального виконання за темами лабораторних робіт "Оцінка параметрів розподілу випадкових величин", "Побудова моделей класифікації агломеративними методами", "Побудова моделей класифікації ітеративними методами", "Методи дискримінантного аналізу. Класифікація з навчанням", "Методи скорочення простору ознак", "Моделі і методи редукції даних", "Моделі і методи факторного аналізу".

# Модуль 1. Методи статистичного аналізу складних систем

## Лабораторна робота №1. Оцінка параметрів розподілу випадкових величин

**Мета** – закріплення теоретичного й практичного матеріалу з оцінки параметрів розподілу випадкових величин, придбання навичок роботи в модулі *Basic Statistics/Tables*.

**Завдання** – необхідно провести аналіз варіаційного ряду для вибірових даних у модулі *Basic Statistics/Tables* ППП *Statistica*:

1. Розрахувати статистичні характеристики ряду (середнє, дисперсію, середнє квадратичне відхилення, моду, медіану, розмах варіації, коефіцієнти асиметрії та ексцесу).
2. Побудувати гістограму та полігон розподілу випадкової величини, зробити висновки щодо характеру закону розподілу.
3. За допомогою критеріїв Пірсона та Колмогорова – Смірнова перевірити гіпотезу про нормальний закон розподілу.
4. Зробити висновки щодо угруповання об'єктів за величиною відповідного показника.

### Методичні рекомендації

Для рішення й аналізу задач даного типу в ППП *Statistica* передбачений модуль *Basic Statistics/Tables* (*Основні статистики й таблиці*). Розглянемо порядок роботи в даному модулі. У меню програм вибрати програму *Statistica*, після її запуску виберіть у меню пункт *File / New* для підготовки власних даних. Перед вами з'явиться діалогове вікно, у якому необхідно вказати кількість змінних (*Number of variables*) і кількість випадків (*Number of Cases*). Після введення натисніть кнопку *OK* (рис. 1.1). Перед вами з'явиться поле, що становить таблицю розміром 25 x 2 : 25 спостережень, 2 змінні (рис. 1.2). Кожен елемент даних, тобто значення показника, займає одну комірку поля даних. Після заповнення всіх комірок поля даних ви одержите таблицю, наведену на рис. 1.3.

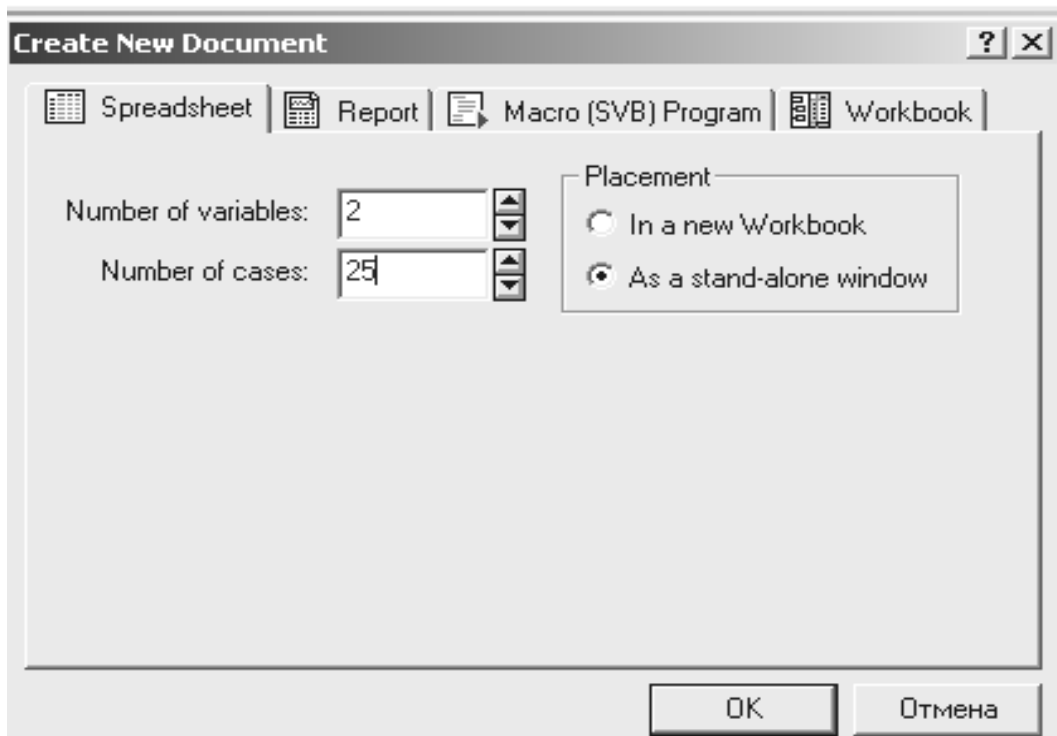


Рис. 1.1. Визначення кількості змінних і спостережень

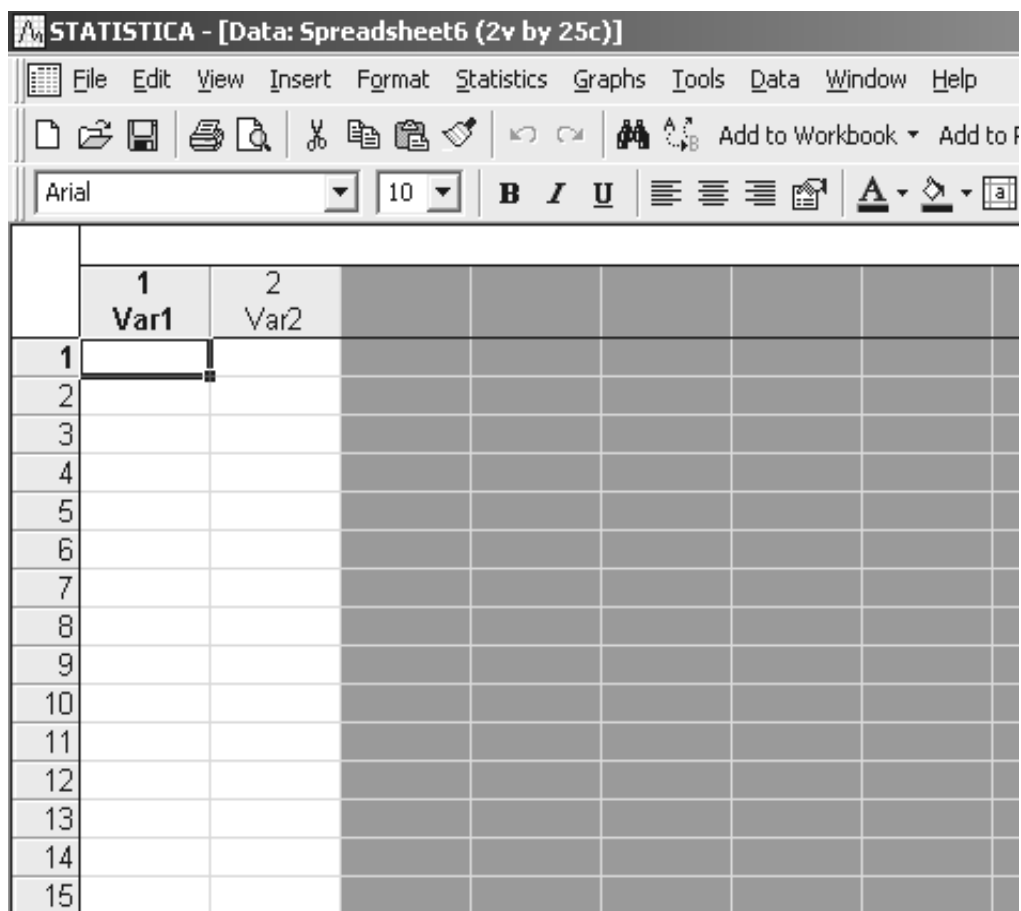


Рис. 1.2. Поле даних

	1 Bank	2 % dohod		
1	Приватбанк	1187477		
2	Проминвест	793821		
3	Аваль	876148		
4	Ощадбанк	389719		
5	Укрсоцбанк	459234		
6	Укрсиббанк	451074		
7	Укрэксим	328131		
8	Райффайзенбанк	209010		
9	Надра	273945		
10	Брокбизнесбанк	167741		
11	Укрпромбанк	232158		
12	Финансы и кредит	175292		
13	Первый укр. международный банк	111185		
14	Хрещатик	70674		
15	Форум	145468		
16	Південний	132243		
17	Правексбанк	120243		
18	Кредитпромбанк	100261		
19	Укргазбанк	104326		
20	Кредитбанк	114054		
21	Ситибанк	42602		
22	Ингбанк Украина	35241		
23	Вабанк	71296		
24	КредитДнепро	91436		
25	Донгорбанк	86384		

Рис. 1.3. Вихідні дані

Розрахуємо основні статистичні характеристики ряду (середнє, дисперсію, середнє квадратичне відхилення, моду, медіану, розмах, коефіцієнти асиметрії й ексцесу).

Щоб почати обчислювальні процедури, необхідно ввійти в позицію меню *Statistics / Basic Statistics/Tables* (рис. 1.4). Після підтвердження вибору модуля перед вами з'явиться діалогове вікно, що дозволяє задати напрямок аналізу *Descriptive statistics (описові статистики)*, подане на рис. 1.5.

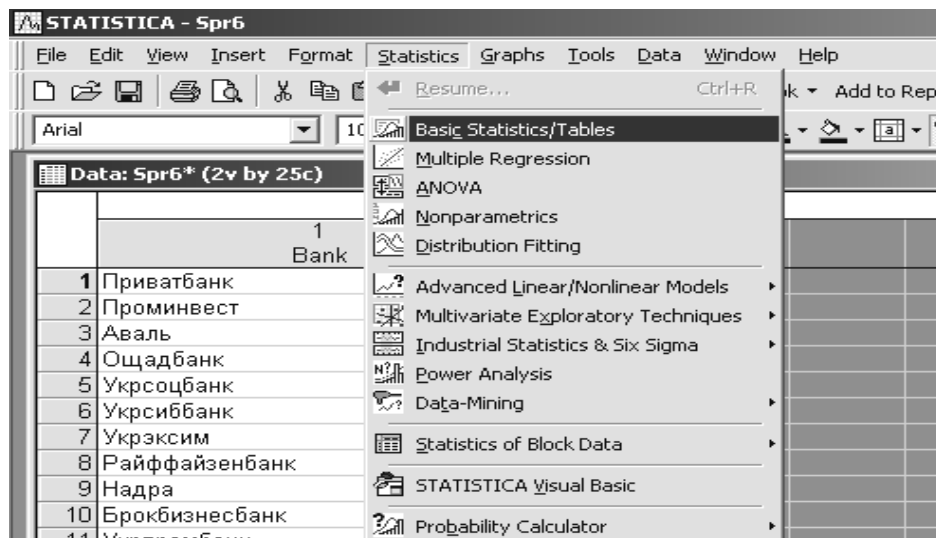


Рис. 1.4. Вибір модуля

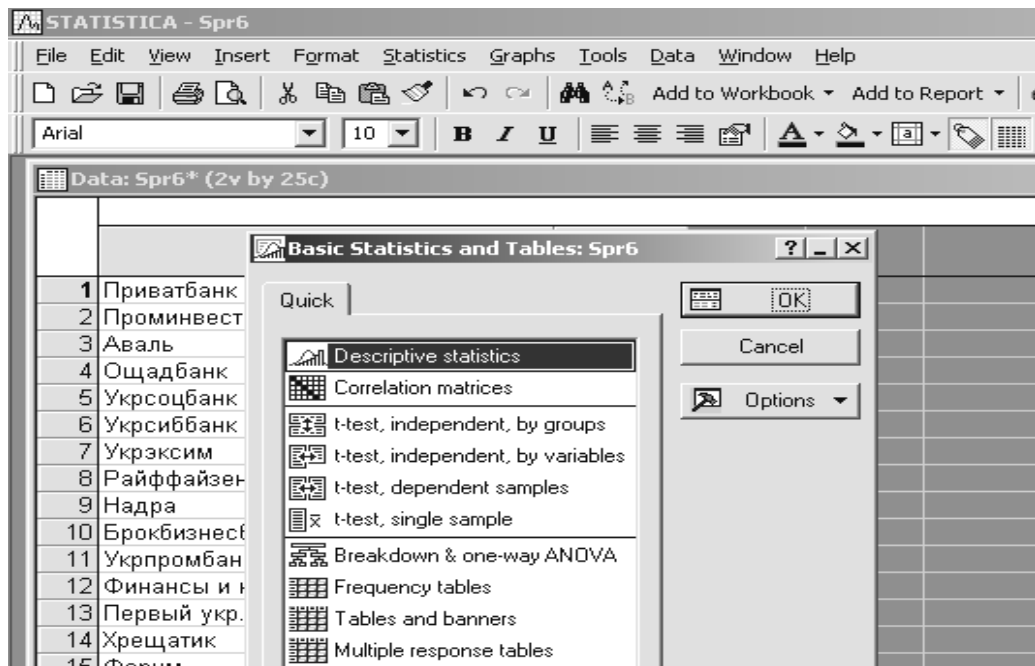


Рис. 1.5. Вибір напрямку аналізу

Після вибору напрямку аналізу з'явиться стартова панель модуля, де необхідно задати вихідні параметри: *Variable* (змінні) і відповідний набір процедур для подальшого аналізу (рис. 1.6).

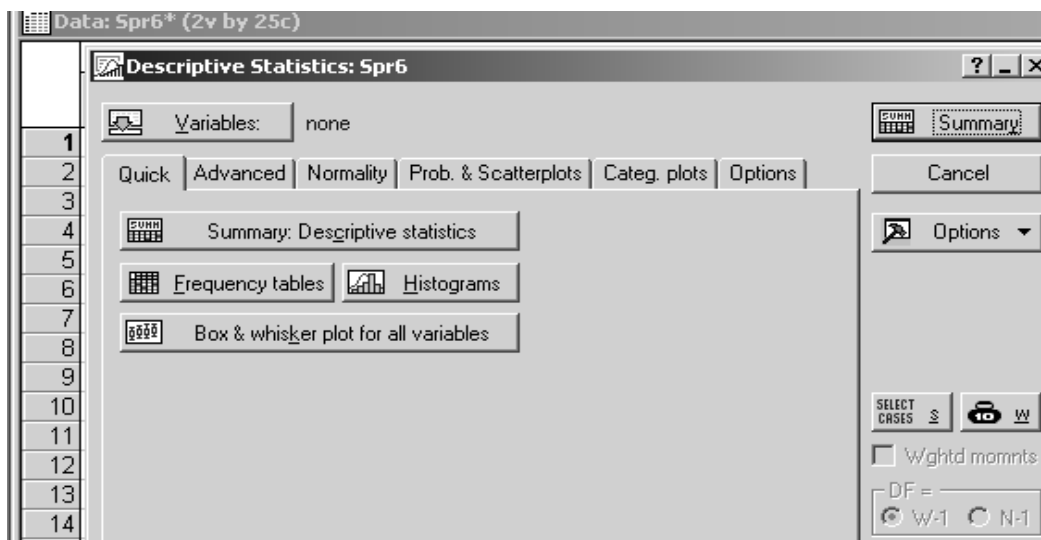


Рис. 1.6. Стартова панель модуля

Ініціюйте кнопку *Variable* (змінні) і у вікні, що з'явилося, укажіть показники, за якими здійснюється аналіз. Після зазначення змінних підтвердіть свій вибір натисканням кнопки *OK*. Далі, ініціювавши вкладиш *Advanced*, необхідно виділити основні статистики для розрахунку (рис. 1.7).

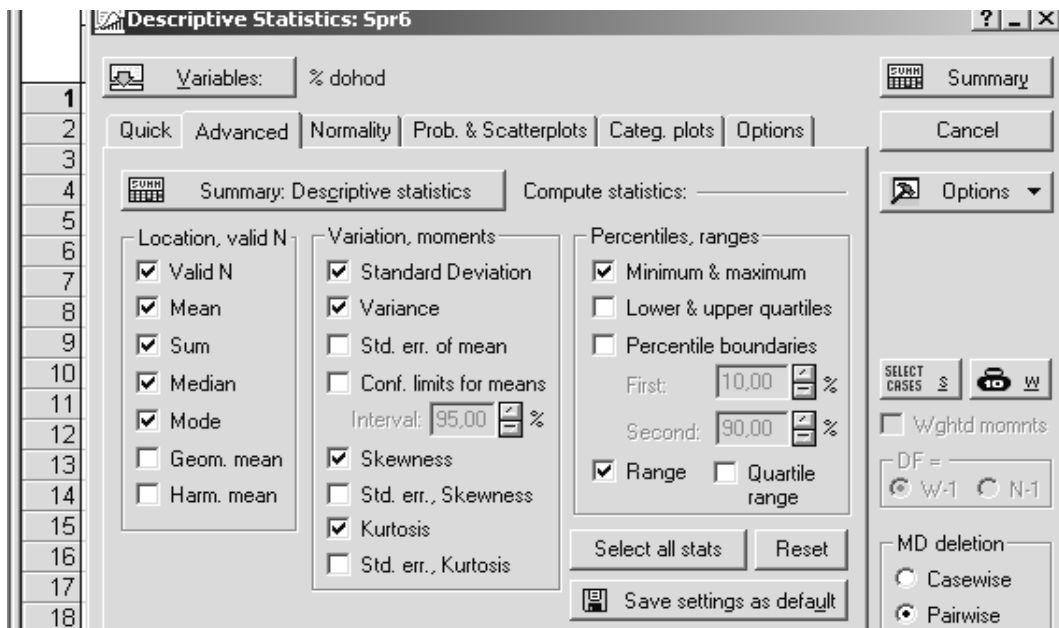


Рис. 1.7. Вибір описових статистик

Такими є: *Valid N* (число спостережень), *Mean* (середнє), *Sum* (сума значень), *Median* (медіана), *Mode* (мода), *Standard Deviation* (середнє квадратичне відхилення), *Variance* (дисперсія), *Skewness* (коефіцієнт асиметрії), *Kurtosis* (коефіцієнт ексцесу), *Min & Max* (мінімум і максимум), *Range* (розмах вибірки). Результати розрахунку описових статистик для даної вибірки одержуємо натисканням клавіші *Summary*, їх наведено на рис. 1.8.

Descriptive Statistics (Spr6)							
Variable	Valid N	Mean	Median	Sum	Minimum	Maximum	Range
% dohod	25	270766,5	145468,0	6769163	35241,00	1187477	1152236
		Variance	Std.Dev.	Skewness	Kurtosis		
		8,370020E+10	289309,9	1,999460	3,715184		

Рис. 1.8. Описові статистики

Побудуємо гістограму й полігон розподілу випадкової величини та проведемо угруповання вибірки. Для наочності подання досліджуваної сукупності побудуємо полігон розподілу. Для цього необхідно зайти в меню *Graphs / 2D Graphs / Scatterplots* (рис. 1.9), вибрати змінні (рис. 1.10) і побудувати полігон розподілу випадкової величини (рис. 1.11).



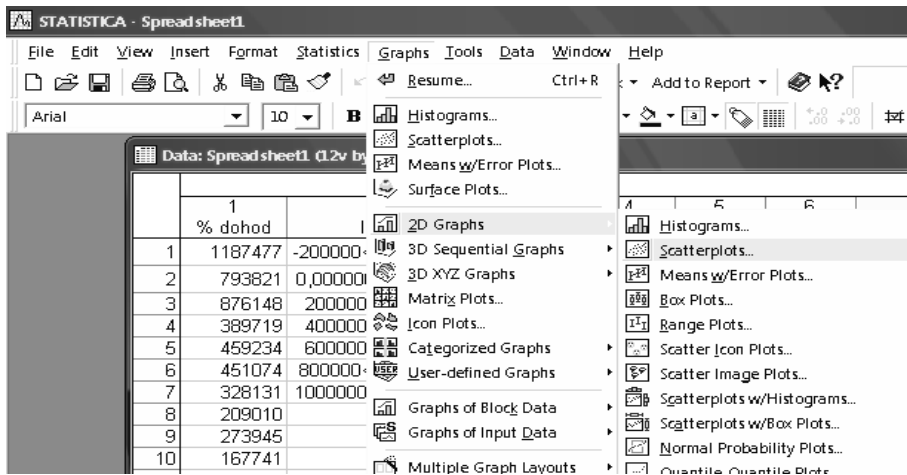


Рис. 1.9. Вибір типу графіку

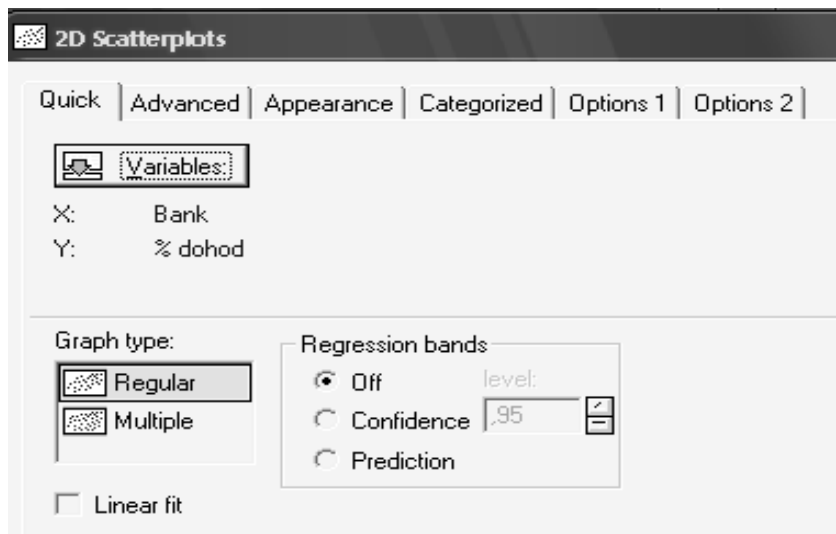


Рис. 1.10. Вибір параметрів графіку

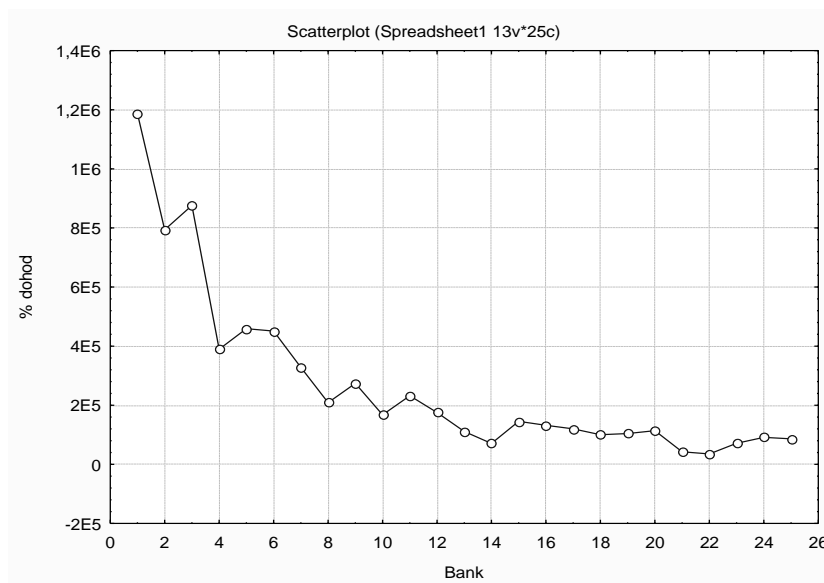


Рис. 1.11. Полігон розподілу випадкової величини

Подальший аналіз здійснюється в рамках перевірки вибірки на нормальний закон розподілу. Для проведення угруповання вибірки в стартовій панелі модуля вибираємо вкладку *Normality*, де можна задавати бажану кількість інтервалів і критерій Колмогорова – Смірнова для тестування вибірки (рис. 1.12). Ініціювавши клавішу *Frequency tables* (таблиці частот), одержуємо таку таблицю (рис. 1.13).

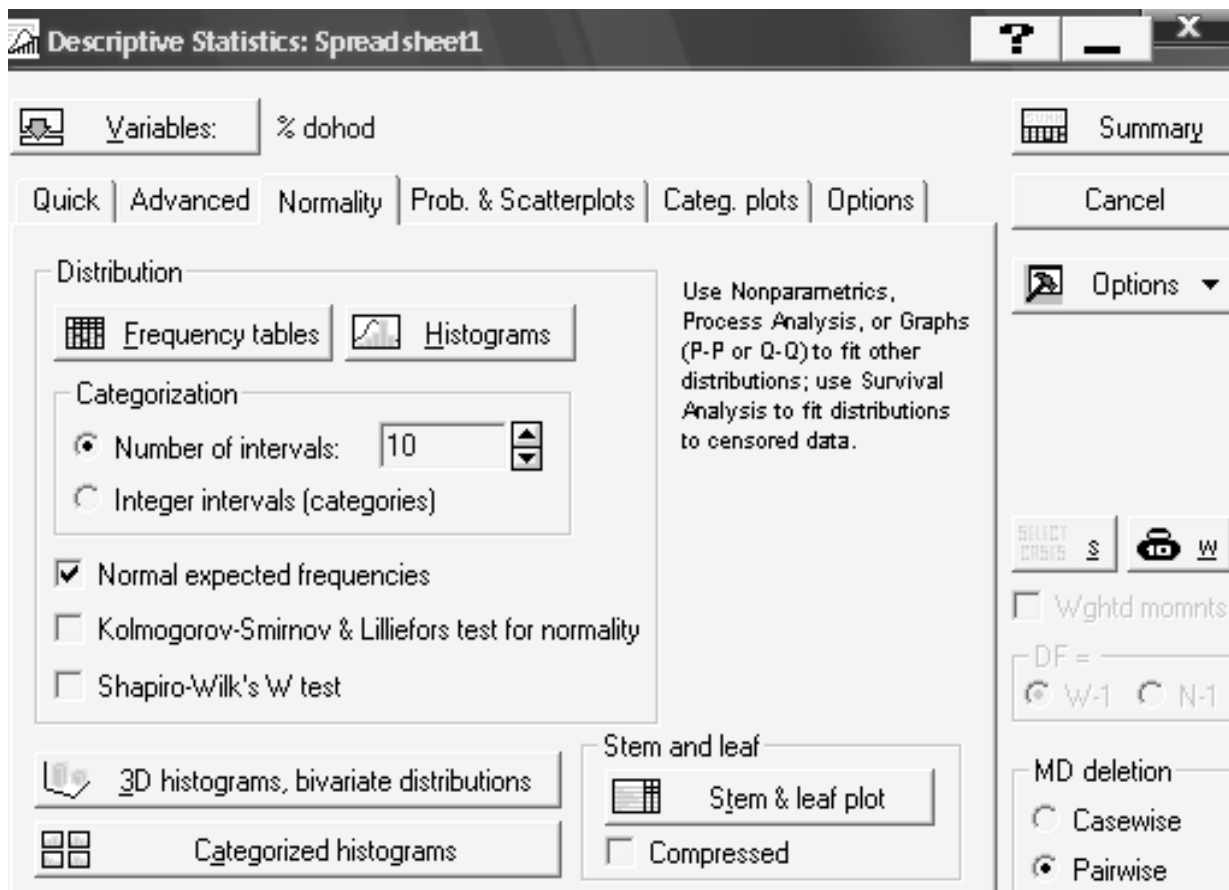


Рис. 1.12. Вибір параметрів угруповання випадкової величини

Як видно, вихідна сукупність із 25 банків розбита на 7 інтервалів, у кожному інтервалі розраховані такі характеристики: Count (частота), Cumulative Count (накопичена частота), Percent of Valid (відсоток від загальної частоти), Cumul % of Valid (накопичений відсоток від загальної частоти), % of all Cases (відсоток від загального числа спостережень), Cumulative % of all Cases (накопичений відсоток від загального числа спостережень), Expected Count (теоретична частота), Cumulative Expected (накопичена теоретична частота), % Expected (відсоток від загальної теоретичної частоти), Cumulative % Expected (накопичений відсоток від загальної теоретичної частоти).

Frequency table: % dohod (Spr6)						
K-S d=,23308, p<,15 ; Lilliefors p<,01						
Category	Count	Cumulative Count	Percent of Valid	Cumul % of Valid	% of all Cases	Cumulative % of All
-200000,<x<=0,000000	0	0	0,00000	0,0000	0,00000	0,0000
0,000000<x<=200000,0	15	15	60,00000	60,0000	60,00000	60,0000
200000,0<x<=400000,0	5	20	20,00000	80,0000	20,00000	80,0000
400000,0<x<=600000,0	2	22	8,00000	88,0000	8,00000	88,0000
600000,0<x<=800000,0	1	23	4,00000	92,0000	4,00000	92,0000
800000,0<x<=1000000,	1	24	4,00000	96,0000	4,00000	96,0000
1000000,<x<=1200000,	1	25	4,00000	100,0000	4,00000	100,0000
Missing	0	25	0,00000		0,00000	100,0000

%	Expected Count	Cumulative Expected	Percent Expected	Cumulative % Expected
00	4,366527	4,36653	17,46611	17,46611
00	5,718006	10,08453	22,87202	40,33813
00	6,726783	16,81132	26,90713	67,24526
00	4,999658	21,81097	19,99863	87,24389
00	2,347086	24,15806	9,38835	96,63224
00	0,695495	24,85355	2,78198	99,41422
00	0,129962	24,98352	0,51985	99,93407
00				

Рис. 1.13. Результат угруповання вибірки

Ініціювавши клавішу *Histograms* (вкладиш *Normality*), одержимо гістограму розподілу з накладеною кривою нормального закону розподілу (рис. 1.14).

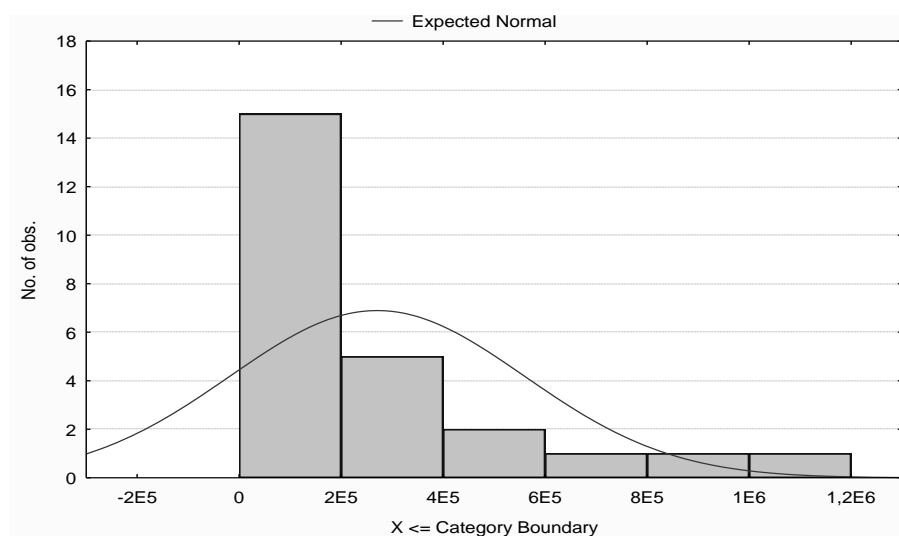


Рис. 1.14. Гістограма розподілу

Подальший аналіз вибірки передбачає розрахунок критерію Пірсона та Колмогорова – Смірнова для формування відповідних висновків про характер закону розподілу. Для визначення характеру закону розподілу та його відповідності нормальному закону дослідимо за допомогою графіків порівняння емпіричних та теоретичних частот і накопичених частот. Вихідні дані для побудови графіків та розраховані значення критерію Пірсона та Колмогорова – Смірнова наведено на рис. 1.15.

	1	2	3	4	5	6	7	8
	% dohod	Interval	m	M	m'	M	X2	M-M
1	1187477	-200000<x<=0,000000	0	0	4,366527	4,36653	4,36653	-4,36653
2	793821	0,000000<x<=200000	15	15	5,718006	10,08453	15,06739	4,91547
3	876148	200000<x<=400000	5	20	6,726783	16,81132	0,44327	3,18868
4	389719	400000<x<=600000	2	22	4,999658	21,81097	1,79971	0,18903
5	459234	600000<x<=800000	1	23	2,347086	24,15806	0,77315	-1,15806
6	451074	800000<x<=1000000	1	24	0,695495	24,85355	0,13332	-0,85355
7	328131	1000000<x<=1200000	1	25	0,129962	24,98352	5,82450	0,01648
8	209010					<b>χ<sup>2</sup></b>	<b>28,4078616</b>	
9	273945					<b>Kolmogorov</b>	<b>0,98309</b>	

Рис. 1.15. Аналіз закону розподілу випадкової величини

Для побудови графіків інтервальних значень частоти розподілу досліджуваної сукупності необхідно зайти в меню *Graphs / 2D Graphs / Scatterplots* вибрати змінні та задати параметри графіку (рис. 1.16). На рис. 1.17 та 1.18 наведені графіки порівняння емпіричних та теоретичних частот і накопичених частот, які дозволяють дійти висновків про відповідність нормальному закону розподілу й визначити розбіжність частот у кожному з досліджуваних інтервалів.

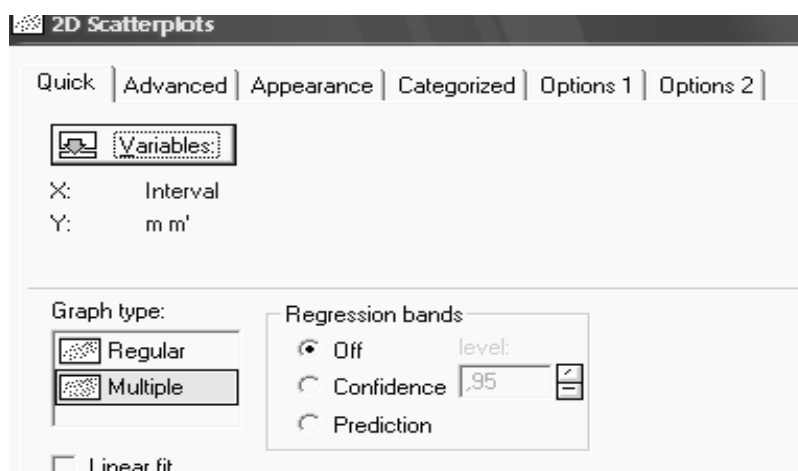


Рис. 1.16. Вибір змінних для побудови графіків

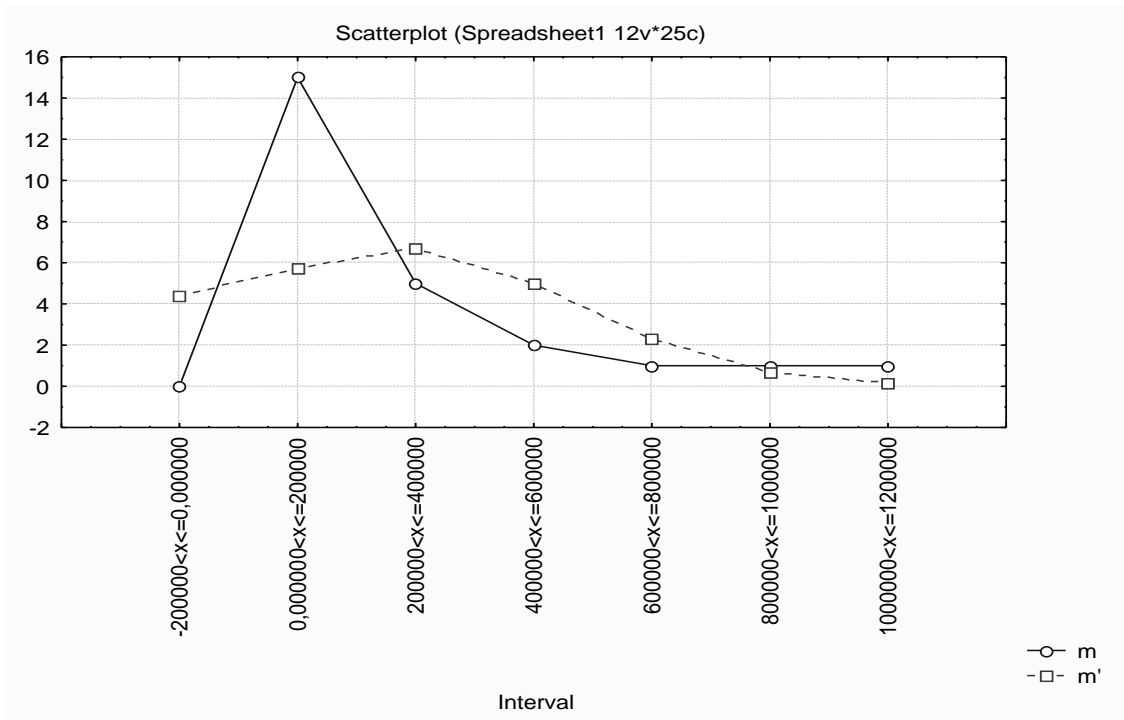


Рис. 1.17. Графік порівняння емпіричних та теоретичних частот

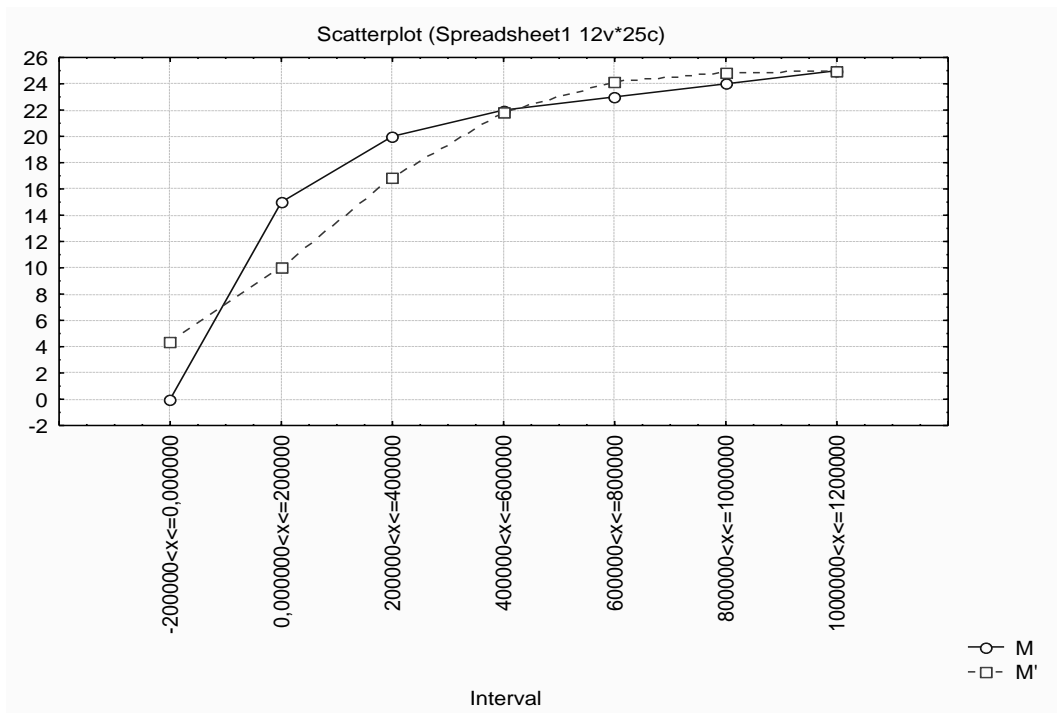


Рис. 1.18. Графік порівняння накопичених емпіричних та теоретичних частот

Далі формулюємо висновки про угруповання об'єктів. Порівнюємо обчислені значення з табличними за відповідними критеріями й формулюємо висновки про характер закону розподілу.

## Лабораторна робота № 2. Побудова моделей класифікації агломеративними методами

**Мета** – закріплення теоретичного й практичного матеріалу за темою "Побудова моделей класифікації агломеративними методами", придбання навичок роботи в модулі *Cluster Analysis*.

**Завдання** – необхідно побудувати моделі класифікації підприємств, використовуючи агломеративні методи класифікації для вибірових даних в модулі *Cluster Analysis* системи *Statistica*:

1. Побудувати моделі кластерного аналізу, використовуючи ієрархічні (деревовидні) методи кластерного аналізу.

2. Порівняти результати дослідження за різними правилами об'єднання та використовуючи різні метрики. Побудувати різні типи дендрограм класифікації. Сформулювати висновки.

### Методичні рекомендації

Для рішення задач класифікації об'єктів у багатовимірному просторі та вивчення їх особливостей в ППП *Statistica* передбачено модуль *Cluster Analysis* (кластерний аналіз). Розглянемо порядок роботи в цьому модулі. Таблицю вихідних даних для рішення задачі класифікації подано на рис. 2.1. Таким чином, задача дослідження полягає у виокремленні класів однорідних об'єктів (підприємств) за такими показниками:  $x_1$  – продуктивність праці,  $x_2$  – коефіцієнт рентабельності капіталу,  $x_3$  – коефіцієнт фондівіддачі.

Щоб перейти до обчислювальних процедур необхідно увійти в меню *Statistics / Multivariate Exploratory Techniques / Cluster Analysis* (рис. 2.2). Після підтвердження вибору модуля перед Вами з'явиться стартова панель модуля (рис. 2.3), де необхідно вибрати напрям аналізу, тобто метод класифікації: *Joining tree clustering* (деревовидна кластеризація). Після підтвердження вибору методу (вибрано метод *Joining tree clustering*) необхідно задати вихідні параметри для проведення кластеризації (рис. 2.4): *Variable* (змінні), *Cluster* (об'єкти кластеризації), *Amalgamation rule* (правила кластеризації), *Distance measure* (міру подібності).

	1 X1	2 X2	3 X3
1	9,26	13,26	1,45
2	9,38	10,16	1,3
3	12,11	13,72	1,37
4	10,81	12,85	1,65
5	9,35	10,63	1,91
6	9,87	9,13	1,68
7	8,17	25,83	1,94
8	9,12	23,39	1,89
9	5,88	14,68	1,94
10	6,3	10,05	2,06
11	6,22	13,99	1,96
12	5,49	9,68	1,02
13	6,5	10,03	1,85
14	6,61	9,13	0,88
15	4,32	5,37	0,62
16	7,37	9,86	1,09
17	7,02	12,62	1,6
18	8,25	5,02	1,53
19	8,15	21,18	1,4
20	5,72	25,17	2,22
21	6,64	19,4	1,32
22	8,1	21	1,48
23	5,52	6,57	0,68
24	9,37	14,19	2,3
25	13,17	15,81	1,37

Рис. 2.1. Вихідні дані

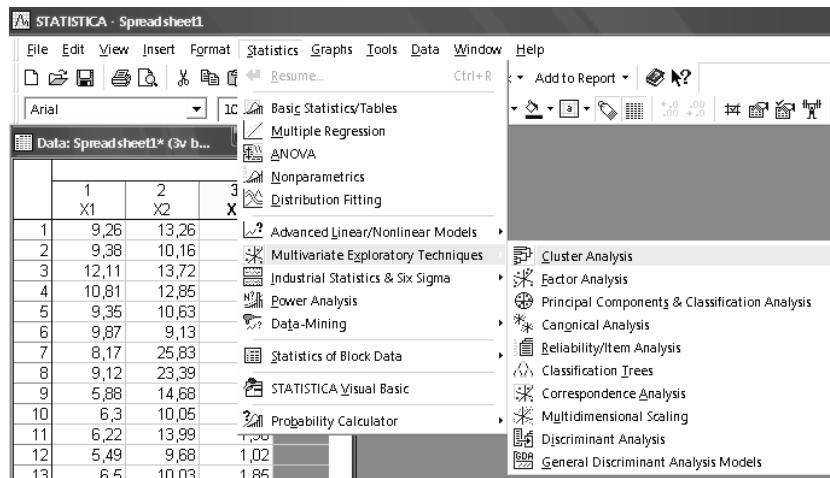


Рис. 2.2. Вибір модуля

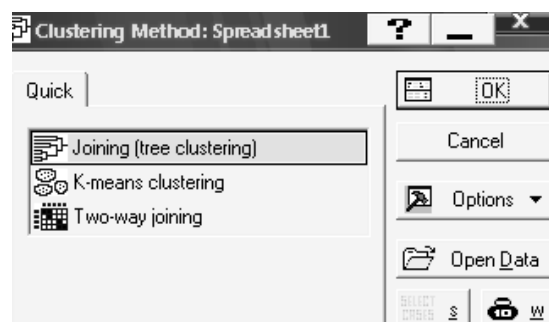


Рис. 2.3. Стартова панель модуля

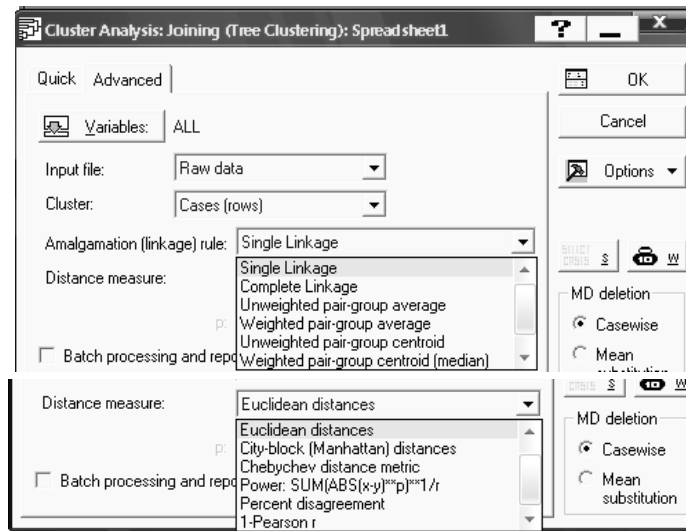


Рис. 2.4. Визначення вихідних параметрів

У даному модулі подано такі правила ієрархічного об'єднання:

- 1) Single linkage (одиначного зв'язку);
- 2) Complete linkage (повних зв'язків);
- 3) Unweighted pair-group average (незваженого попарного середнього);
- 4) Weighted pair-group average (зваженого попарного середнього);
- 5) Unweighted pair-group centroid (незваженого центроїдного);
- 6) Weighted pair-group centroid (зваженого центроїдного);
- 7) Ward's method (метод Уорда).

Як міру подібності подано такі метрики:

- 1) Euclidean distances (евклідова метрика);
- 2) Square Euclidean distances (квадрат евклідової метрики);
- 3) City-block (Manhattan) distances (манхеттенська відстань);
- 4) Chebychev distances metric (відстань Чебишева);
- 5) Power metric (степенева відстань Мінковського);
- 6) Percent disagreement (відсоток незгоди (для категоріальних даних));
- 7) 1 – Personal r (1 – коефіцієнт кореляції Пірсона).

Для дослідження виберемо процедуру *Single linkage* (одиначного зв'язку) та звичайну *Евклідову метрику* (*Euclidean distances*). Підтверджуючи вибір, ініціюйте клавішу (ОК). З'явиться вікно результатів, де у верхній частині подана основна інформація та вибрані процедури дослідження. Опції в нижній частині вікна на вкладці *Advanced* призначено для аналізу результатів кластеризації (рис. 2.5):



- *Horizontal hierarchical tree plot* (горизонтальна деревовидна діаграма);
- *Vertical icicle plot* (вертикальна деревовидна діаграма – дендрограма);
- *Amalgamation schedule* (правило об'єднання в кластери);
- *Graf of amalgamation schedule* (графік порядку об'єднання);
- *Distance matrix* (матриця відстаней);
- *Descriptive statistics* (описові статистики).

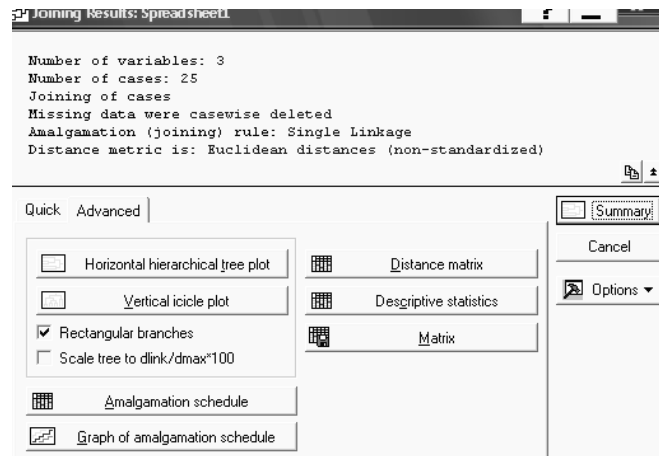


Рис. 2.5. Опції аналізу результатів

Ініціювавши клавішу *Vertical icicle plot*, будемо дендрограму класифікації (рис. 2.6), де за віссю абсцис подано об'єкти дослідження, а за віссю ординат – відстані між ними.

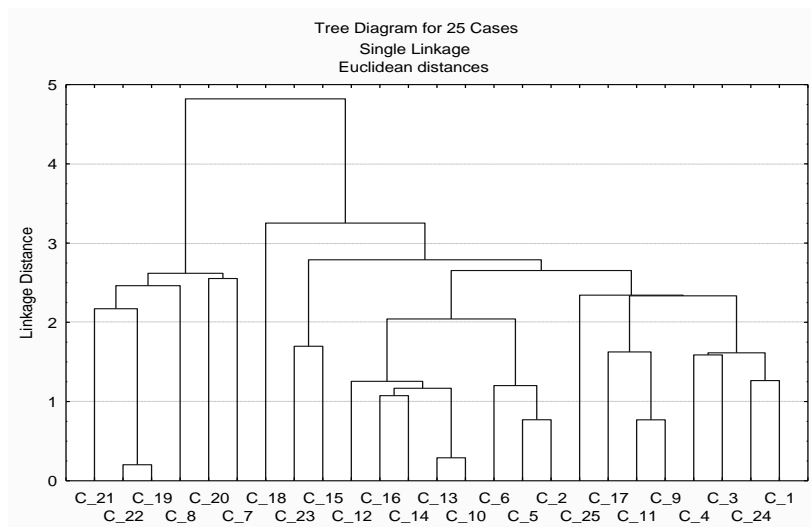


Рис. 2.6. Дендрограма класифікації

Ініціювавши клавішу *Distance matrix*, обчислюємо матрицю відстаней, фрагмент якої наведено на рис. 2.7. На рис. 2.8. наведено фрагмент матриці об'єднання (*Amalgamation schedule*).

Case No	Euclidean distances (Spreadsheet1)												
	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_10	C_11	C_12	C_13
C_1	0,0	3,1	2,9	1,6	2,7	4,2	12,6	10,1	3,7	4,4	3,2	5,2	4,3
C_2	3,1	0,0	4,5	3,1	0,8	1,2	15,7	13,2	5,8	3,2	5,0	3,9	2,9
C_3	2,9	4,5	0,0	1,6	4,2	5,1	12,7	10,1	6,3	6,9	5,9	7,8	6,7
C_4	1,6	3,1	1,6	0,0	2,7	3,8	13,2	10,7	5,3	5,3	4,7	6,2	5,2
C_5	2,7	0,8	4,2	2,7	0,0	1,6	15,2	12,8	5,3	3,1	4,6	4,1	2,9
C_6	4,2	1,2	5,1	3,8	1,6	0,0	16,8	14,3	6,8	3,7	6,1	4,5	3,5
C_7	12,6	15,7	12,7	13,2	15,2	16,8	0,0	2,6	11,4	15,9	12,0	16,4	15,9
C_8	10,1	13,2	10,1	10,7	12,8	14,3	2,6	0,0	9,3	13,6	9,8	14,2	13,6
C_9	3,7	5,8	6,3	5,3	5,3	6,8	11,4	9,3	0,0	4,7	0,8	5,1	4,7
C_10	4,4	3,2	6,9	5,3	3,1	3,7	15,9	13,6	4,7	0,0	3,9	1,4	0,3
C_11	3,2	5,0	5,9	4,7	4,6	6,1	12,0	9,8	0,8	3,9	0,0	4,5	4,0

Рис. 2.7. Матриця класифікації

linkage distance	Amalgamation Schedule (Spreadsheet1)							
	Obj. No. 1	Obj. No. 2	Obj. No. 3	Obj. No. 4	Obj. No. 5	Obj. No. 6	Obj. No. 7	Obj. No. 8
,2032240	C_19	C_22						
,2906888	C_10	C_13						
,7694803	C_9	C_11						
,7706491	C_2	C_5						
1,074523	C_14	C_16						
1,167647	C_10	C_13	C_14	C_16				
1,202248	C_2	C_5	C_6					
1,255587	C_10	C_13	C_14	C_16	C_12			
1,264713	C_1	C_24						
1,589119	C_3	C_4						
1,615735	C_1	C_24	C_3	C_4				
1,626807	C_9	C_11	C_17					
1,698117	C_15	C_23						

Рис. 2.8. Матриця об'єднання

Дендрограму класифікації за методом Уорда наведено на рис. 2.9.

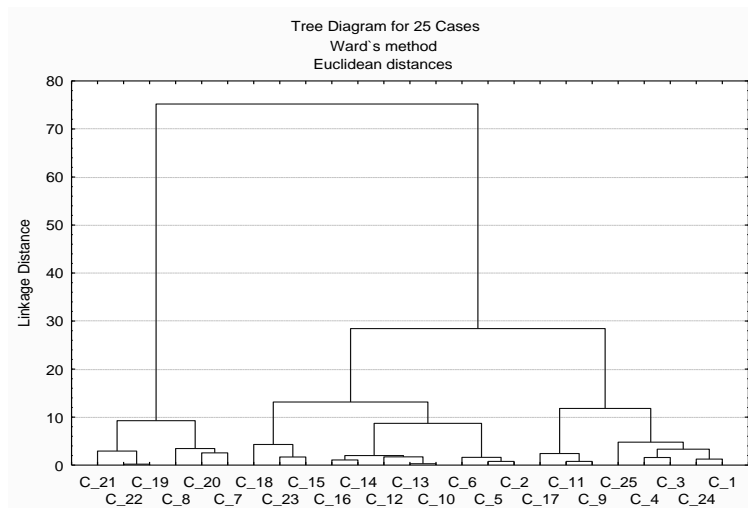


Рис. 2.9. Дендрограма класифікації за методом Уорда

Аналіз даної дендрограми дозволяє розпізнати три групи (кластери) однорідних станів у спостережуваній сукупності даних. Так, кластер № 1

має найвищий рівень рентабельності. Кластер № 2 має найвищі показники продуктивності праці та фондівіддачі. Кластер № 3 має найнижчі значення за всіма досліджуваними показникам.

### **Лабораторна робота № 3. Побудова моделей класифікації ітеративними методами**

**Мета** – закріплення теоретичного й практичного матеріалу за темою "Побудова моделей класифікації ітеративними методами", придбання навичок роботи в модулі *Cluster Analysis*.

**Завдання** – необхідно побудувати моделі класифікації підприємств, використовуючи ітеративні методи та стратегії класифікації для вибіркового даних у модулі *Cluster Analysis* ППП *Statistica*:

1. Побудувати моделі кластерного аналізу ітеративним методом кластерного аналізу.

2. Провести класифікацію об'єктів за методом К-середніх, визначити характеристики моделі.

3. Проаналізувати результати класифікації за різного значення сформованих кластерів, побудувати графіки, навести основні статистики та оцінку змінних за отриманими моделями.

4. Зробити висновки та подати економічну інтерпретацію отриманим результатам кластерних утворень.

#### **Методичні рекомендації**

Для рішення задач класифікації об'єктів у багатовимірному просторі та вивчення їх особливостей у ППП *Statistica* передбачено модуль *Cluster Analysis* (кластерний аналіз). Розглянемо порядок роботи в даному модулі.

Таблицю вихідних даних для рішення задачі класифікації подано на рис. 3.1. Таким чином, задача дослідження полягає у формуванні класів однорідних об'єктів (підприємств) за такими показниками:  $x_1$  – продуктивність праці,  $x_2$  – коефіцієнт рентабельності капіталу,  $x_3$  – коефіцієнт фондівіддачі.

Щоб почати обчислення, необхідно вибрати пункт меню *Statistics / Multivariate Exploratory Techniques / Cluster Analysis* (рис. 3.2).

	1 X1	2 X2	3 X3
1	9,26	13,26	1,45
2	9,38	10,16	1,3
3	12,11	13,72	1,37
4	10,81	12,85	1,65
5	9,35	10,63	1,91
6	9,87	9,13	1,68
7	8,17	25,83	1,94
8	9,12	23,39	1,89
9	5,88	14,68	1,94
10	6,3	10,05	2,06
11	6,22	13,99	1,96
12	5,49	9,68	1,02
13	6,5	10,03	1,85
14	6,61	9,13	0,88
15	4,32	5,37	0,62
16	7,37	9,86	1,09
17	7,02	12,62	1,6
18	8,25	5,02	1,53
19	8,15	21,18	1,4
20	5,72	25,17	2,22
21	6,64	19,4	1,32
22	8,1	21	1,48
23	5,52	6,57	0,68
24	9,37	14,19	2,3
25	13,17	15,81	1,37

Рис. 3.1. Вихідні дані

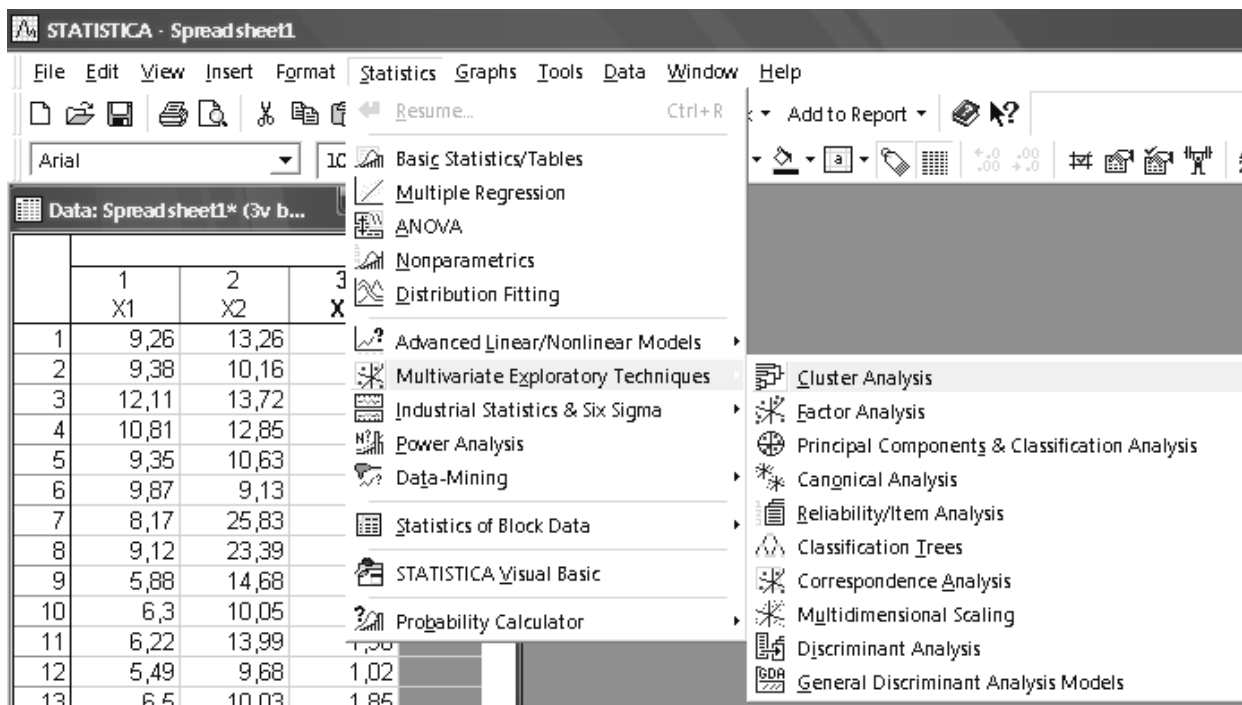


Рис. 3.2. Вибір модуля

Після підтвердження вибору модуля перед вами з'явиться стартова панель модуля (рис. 3.3), де необхідно вибрати напрям аналізу, тобто метод класифікації: *K-means clustering* (метод *K*-середніх);

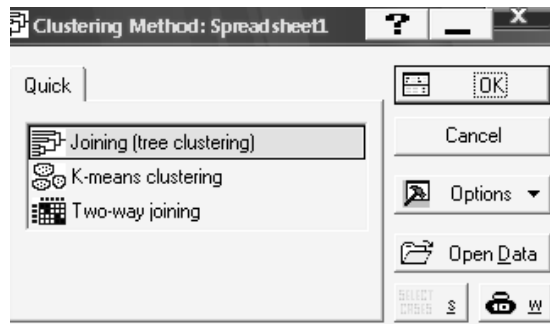


Рис. 3.3. Стартова панель модуля

Ініціювавши дану опцію, відкриваємо діалогове вікно реалізації методу, де на вкладці *Advanced* необхідно задати змінні для аналізу (*Variables*), об'єкти кластеризації (*Cluster*), число кластерів (*Number of clusters*), число ітерацій (*Number of iterations*), та початкові центри кластерів – опції (*Initial cluster centers*) (рис. 3.4)

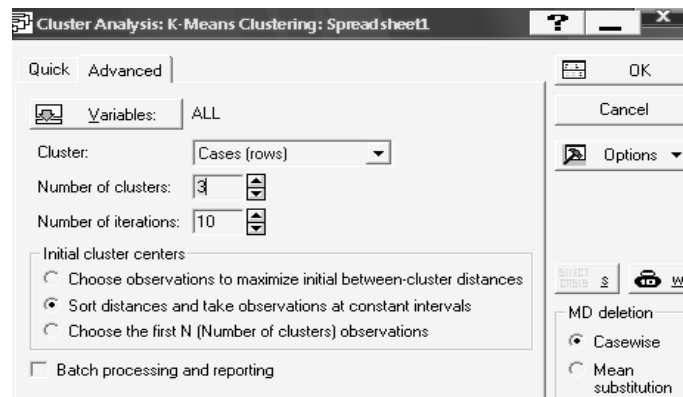


Рис. 3.4. Діалогове вікно методу *K-середніх*

Вікно аналізу результатів методу *K-середніх* наведено на рис. 3.5. Верхня частина є інформаційною, а нижня (вкладка *Advanced*) дозволяє сформулювати повну інформацію результатів аналізу. Розглянемо функціональне значення опцій даного вікна:

- *Summary: Cluster means & Euclidean distances* (евклідові відстані та середні значення станів кластерів);
- *Analysis of variance* (дисперсійний аналіз);
- *Graf of means* (графік середніх значень);
- *Descriptive statistics for each cluster* (описові статистики для кластерів);
- *Members of each cluster & distances* (члени кластерів та їх відстані до центру кластера);
- *Save classifications and distances* (збереження результатів).

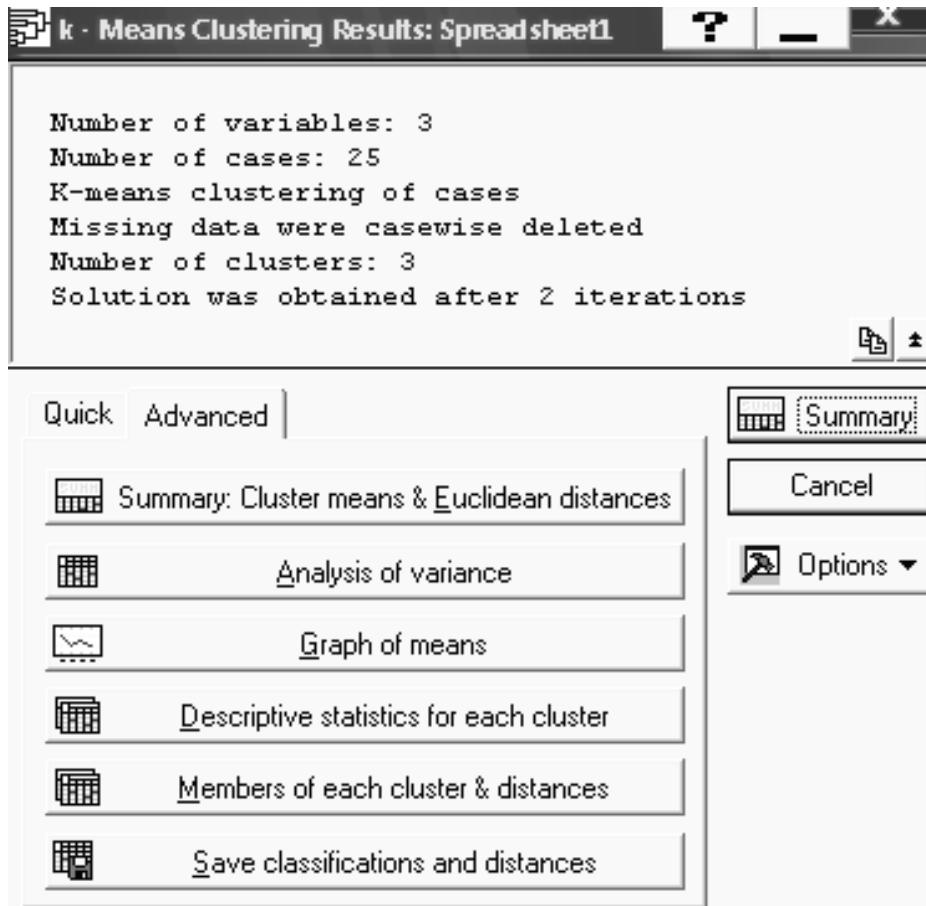


Рис. 3.5. Опції аналізу результатів методу *K-середніх*

Евклідові відстані між отриманими кластерами та середні значення для кожного досліджуваного показника подано на рис. 3.6, причому значення евклідових відстаней знаходиться під головною діагоналлю, а над – квадрат евклідових відстаней.

Cluster Number	Euclidean Distances between Clusters			Cluster Means (Spreadsheet)			
	No. 1	No. 2	No. 3	Cluster No. 1	Cluster No. 2	Cluster No. 3	Cluster No. 3
No. 1	0,000000	10,23208	28,65567	9,24333	6,96100	7,65000	7,65000
No. 2	3,198762	0,000000	67,07293	13,52778	8,50000	22,66167	22,66167
No. 3	5,353099	8,18981	0,000000	1,72778	1,27100	1,70833	1,70833

Рис. 3.6. Евклідові відстані та середні значення станів кластерів

Графік середніх значень для кластерів станів наведено на рис. 3.7. Як видно, найбільш кластери різняться за показником  $x_2$ , потім  $x_1$  і дуже малі відмінності в середніх для показника  $x_3$ .

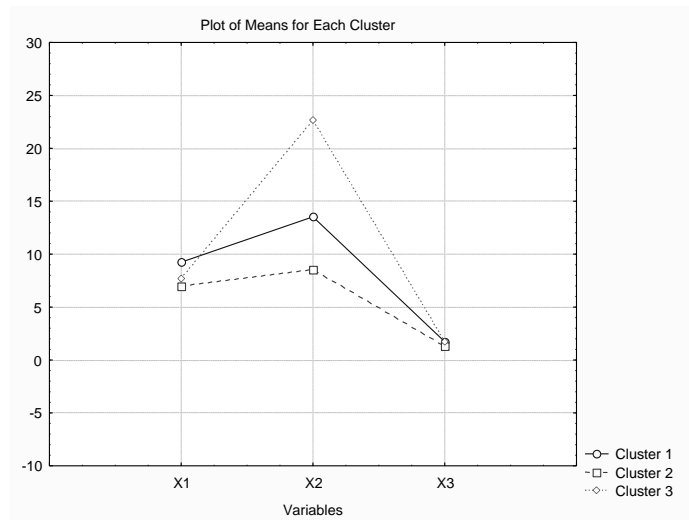


Рис. 3.7. Графік середніх значень для кластерів станів

Результати дисперсійного аналізу оцінки якості показників наведено на рис. 3.8. У таблиці наведено значення міжгрупових та внутрішньогрупових дисперсій ознак. Чим менше значення внутрішньогрупової дисперсії і більше значення міжгрупової, тим краще ознака характеризує приналежність об'єктів до кластеру. Параметри  $F$  та  $p$  визначають вклад ознаки в класифікацію.

Variable	Analysis of Variance (Spreadsheet1)					
	Between SS	df	Within SS	df	F	signif. p
X1	25,375	2	87,2736	22	3,1983	0,06035
X2	752,533	2	86,6000	22	95,5873	0,00000
X3	1,211	2	3,6763	22	3,6233	0,04363

Рис. 3.8. Таблиця дисперсійного аналізу

На рис. 3.9. подано описові статистики для виділених кластерів, а саме: середнє, середньоквадратичне відхилення та дисперсія.

Variable	Descriptive Statistics for Cluster 1 Cluster contains 9 cases			Variable	Descriptive Statistics for Cluster 3 Cluster contains 6 cases			Variable	Descriptive Statistics for Cluster 2 Cluster contains 10 cases		
	Mean	Standard Deviation	Variance		Mean	Standard Deviation	Variance		Mean	Standard Deviation	Variance
X1	9,2433	2,53757	6,43930	X1	7,65000	1,235217	1,525760	X1	6,961000	1,767939	3,125610
X2	13,5277	1,45677	2,12219	X2	22,66167	2,547763	6,491096	X2	8,500000	2,032158	4,129667
X3	1,7277	0,31920	0,10189	X3	1,70833	0,359579	0,129297	X3	1,271000	0,496061	0,246077

Рис. 3.9. Описові статистики для кластерів

Члени кластерів та їх відстані до центру відповідного кластеру наведено на рис. 3.10. Таким чином, дані таблиці дозволяють визначити склад кожного кластеру.

Members of Cluster Number 1 (Spreadsheet1) and Distances from Respective Cluster Center Cluster contains 9 cases									
Case No.	Case No.	Case No.	Case No.	Case No.	Case No.	Case No.	Case No.	Case No.	Case No.
C_1	C_3	C_4	C_5	C_9	C_11	C_17	C_24	C_25	
Distance	0,22296	1,67159	0,98655	1,67746	2,05626	1,77088	1,38847	0,51056	2,63028

Members of Cluster Number 2 (Spreadsheet1) and Distances from Respective Cluster Center Cluster contains 10 cases										
Case No.	Case No.	Case No.	Case No.	Case No.	Case No.	Case No.	Case No.	Case No.	Case No.	
C_2	C_6	C_10	C_12	C_13	C_14	C_15	C_16	C_18	C_23	
Distance	1,693911	1,734595	1,074235	1,098369	0,981268	0,473632	2,394130	0,826568	2,147790	1,431859

Members of Cluster Number 3 (Spreadsheet1) and Distances from Respective Cluster Center Cluster contains 6 cases						
Case No.	Case No.	Case No.	Case No.	Case No.	Case No.	Case No.
C_7	C_8	C_19	C_20	C_21	C_22	
Distance	1,85853	0,95295	0,92021	1,85098	1,98405	1,00262

Рис. 3.10. Члени кластерів та їх відстані до центру кластера

Так, кластер № 1 (дев'ять підприємств) має найвищі показники продуктивності праці та фондівіддачі, при середньому значенні рентабельності капіталу. Кластер № 2 (десять підприємств) має найнижчі значення за всіма досліджуваними показниками. Кластер № 3 (шість підприємств) має найвищий рівень рентабельності, а показники продуктивності праці та фондівіддача знаходяться на середньому рівні.

#### Лабораторна робота № 4. Методи дискримінантного аналізу. Класифікація з навчанням

**Мета** – закріплення теоретичного й практичного матеріалу за темою, методи і моделі дискримінантного аналізу, придбання навичок роботи в модулі *Discriminant Analysis*.

**Завдання** – необхідно побудувати модель класифікації підприємств і провести розпізнавання для вибіркового даних у модулі *Discriminant Analysis* ППП *Statistica*:

1. Побудувати модель дискримінантного аналізу на основі вибіркового даних.
2. Оцінити якість моделей розпізнавання, значимість змінних та провести канонічний аналіз функцій.
3. Побудувати моделі, використовуючи методи покрокового аналізу включення та виключення факторних змінних, оцінити якість побудованих моделей та дискримінацію змінних.



4. Проаналізувати результати розпізнавання (матриця класифікацій), подати класифікацію за дискримінантною моделлю.

5. Сформулювати висновки та прогнози (розпізнавання) за побудованою моделлю. Дати економічну інтерпретацію здобутим результатам.

### Методичні рекомендації

Для рішення й аналізу задач класифікації багатомірних сукупностей за наявності навчальних вибірок (класифікація з навчанням) в ППП *Statistica* передбачено модуль *Discriminant Analysis* (Дискримінантний аналіз). Даний модуль має широкий набір засобів, які забезпечують проведення дискримінантного аналізу даних, візуалізації та інтерпретації результатів. Розглянемо порядок роботи в даному модулі.

Таблиця вихідних даних для рішення задачі подано на рис. 4.1.

	1	2	3	4	5	6
	X1	X2	X3	X4	X5	Class
1	9,26	0,78	1,37	0,23	1,45	A
2	9,38	0,75	1,49	0,39	1,3	A
3	10,81	0,7	1,42	0,18	1,65	A
4	9,35	0,62	1,35	0,15	1,91	A
5	9,87	0,76	1,39	0,34	1,68	A
6	9,12	0,71	1,27	0,09	1,89	A
7	6,61	0,72	1,23	0,48	0,88	B
8	4,32	0,68	1,39	0,41	0,62	B
9	7,37	0,77	1,38	0,62	1,09	B
10	6,64	0,77	1,35	0,5	1,32	B
11	5,52	0,72	1,48	1,2	0,68	B
12	9,37	0,79	1,24	0,21	2,3	A
13	5,68	0,71	1,28	0,66	1,43	B
14	5,22	0,79	1,33	0,74	1,82	B
15	10,02	0,76	1,22	0,32	2,62	A
16	6,7	0,79	0,79	0,39	1,24	B
17	9,42	0,7	0,7	0,72	2,03	A

Рис. 4.1. Вихідні дані

Як чинники впливу на рівень інвестиційної привабливості, розглядаються такі коефіцієнти: продуктивність праці ( $x_1$ ), питома вага робітників у складі промислово-виробничого персоналу ( $x_2$ ), коефіцієнт змінності ( $x_3$ ), коефіцієнт браку ( $x_4$ ), коефіцієнт фондівіддачі ОВФ ( $x_5$ ). Значення даних показників та навчальні вибірки (група А – інвестиційно привабливі підприємства, група В – підприємства аутсайтери) подано на рис. 4.1.

Щоб почати обчислювальні процедури слід вибрати пункт меню *Statistics / Multivariate Exploratory Techniques / Discriminant Analysis* (рис. 4.2).

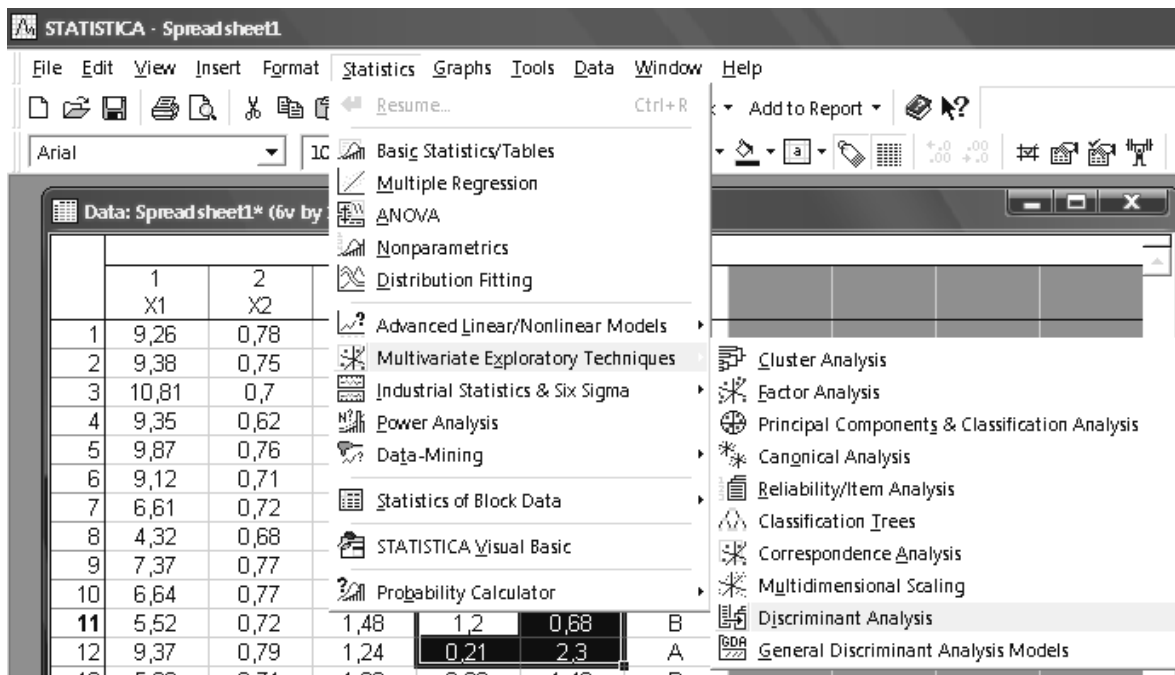


Рис. 4.2. Вибір модуля

Після підтвердження вибору модуля перед вами з'явиться стартова панель модуля, де необхідно задати вихідні параметри моделювання (рис. 4.3).

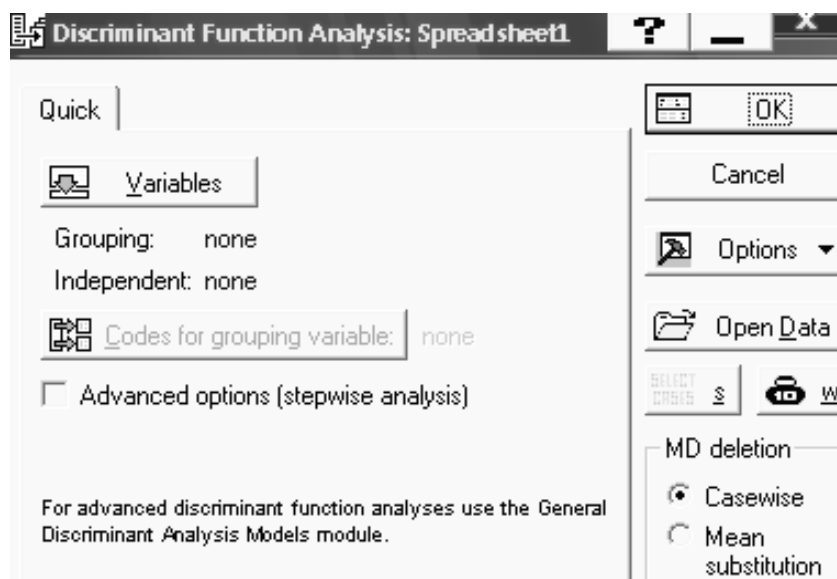


Рис. 4.3. Стартова панель модуля

Ініціюйте кнопку *Variables (змінні)* і у вікні, що з'явилось, вкажіть показники, за якими будете здійснювати аналіз (рис. 4.4). Після вибору змінних (групуєча – залежна змінна, та незалежних – коефіцієнти), підтвердіть свій вибір натисканням кнопки *OK*. Також у даному вікні можна задати коди для значень групуєчої змінної.

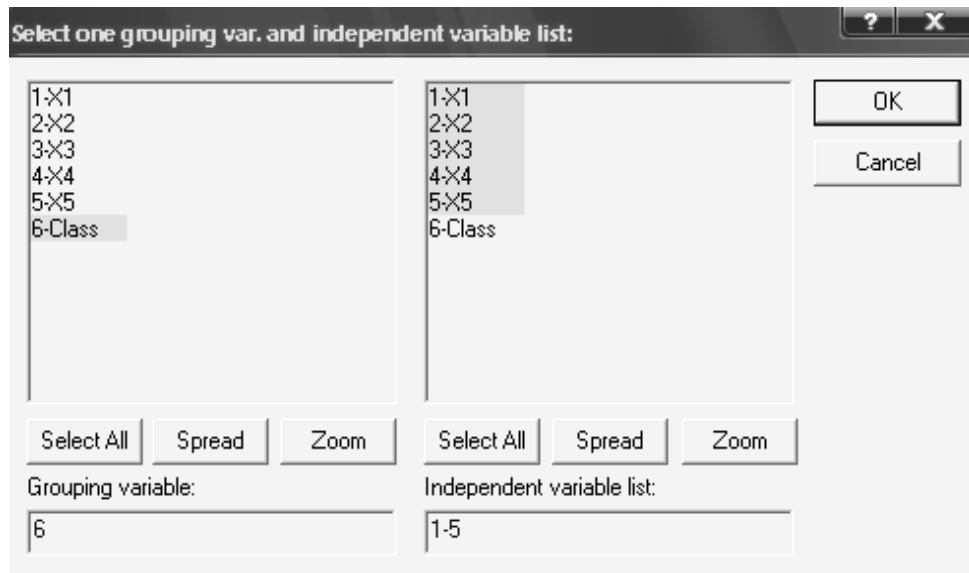


Рис. 4.4. Вибір змінних для аналізу

Стартову панель модуля *Discriminant Analysis* з опціями для аналізу наведено на рис. 4.5.

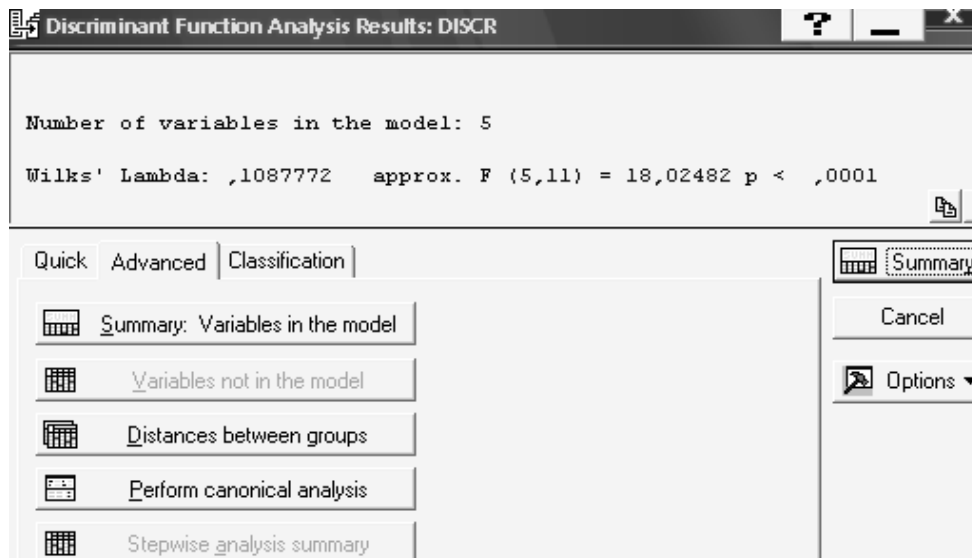


Рис. 4.5. Результати дискримінантного аналізу

У верхній інформаційній частині вікна наведено значення лямбди Уїлкса ( $\lambda$ ), яка характеризує якість дискримінації та змінюється в межах [0,1]. Значення близькі до нуля визначають гарну якість дискримінації. Значення критерія Фішера ( $F$ ) для оцінювання адекватності моделі для порівняння з табличними значеннями.

У нижній частині вікна подано різноманітні опції для поглибленого аналізу моделі та побудови графіків. Ініціювавши опцію *Summary: Variables in the model (аналіз змінних у моделі)* на вкладці *Advanced*, отримаємо результати дискримінантного аналізу, наведені на рис. 4.6.

Discriminant Function Analysis Summary (Spreadsheet1)						
No. of vars in model: 5; Grouping: Class (2 grps)						
Wilks' Lambda: ,10878 approx. F (5, 11)=18,025 p< ,0001						
N=17	Wilks' Lambda	Partial Lambda	F-remove (1, 11)	p-level	Toler.	1-Toler. (R-Sqr.)
X1	0,357719	0,304086	25,17404	0,000392	0,890477	0,109523
X2	0,125683	0,865487	1,70961	0,217705	0,847674	0,152326
X3	0,113186	0,961049	0,44583	0,518085	0,847547	0,152453
X4	0,109932	0,989494	0,11680	0,738972	0,968503	0,031497
X5	0,125269	0,868346	1,66776	0,223036	0,788661	0,211339

Рис. 4.6. Результати дискримінантного аналізу

Розглянемо стовпці даної таблиці:

*Wilks' Lambda* (лямбда Уїлкса) – значення лямбди Уїлкса, яка є результатом виключення відповідної змінної з моделі. Чим більше значення лямбди Уїлкса, тим вагоміша змінна в процедурі дискримінації.

*Partial Lambda* (часткова лямбда) – характеризує одиничний вклад відповідної змінної в дискримінацію моделі. Чим менша дана статистика, тим більший її вклад у загальну дискримінацію.

*F-remove (F-виключення)* – значення  $F$ -критерію для відповідної змінної.

*p-level (рівень значущості)*  $F$ -критерію для відповідної змінної.

*Toler (толерантність)* – визначається  $(1 - R^2)$ , де  $R^2$  – коефіцієнт множинної кореляції даної змінної зі всіма іншими змінними в моделі. Толерантність є мірою збитковості змінних у моделі.

Таким чином, можна дійти висновку, що найбільш значимими змінними для дискримінації є змінні  $x_1$ ,  $x_2$ ,  $x_5$ . Ініціювавши клавішу *Distances between groups (відстані між групами)*, будуюмо таблицю

відстаней (рис. 4.7), яка характеризує якість дискримінації спостережень і ступінь відмінностей (неоднорідність) груп.

Class	Squared Mahalanobis	
	A	B
A	0,00000	32,88619
B	32,88619	0,00000

Рис. 4.7. Матриця відстаней між групами

Ініціювавши опцію *Perform canonical analysis (канонічний аналіз)*, відкриється меню напрямів канонічного аналізу (рис. 4.8):

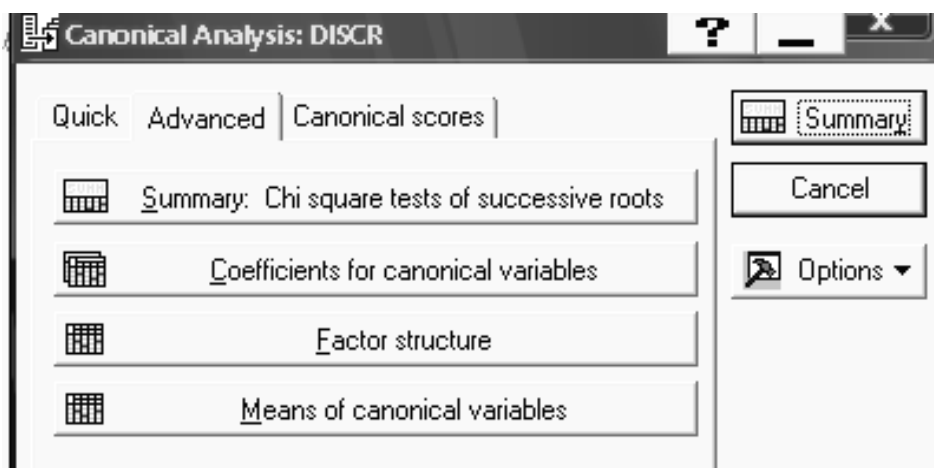


Рис. 4.8. Опції канонічного аналізу функцій

Ініціювавши опцію *Summary: Chi square tests of successive roots* (Підсумки:  $\chi^2$ -критерій послідовності коренів), обчислюємо власні значення та канонічний коефіцієнт кореляції (рис. 4.9).

Roots Removed	Chi-Square Tests with Successive Roots Removed					
	Eigen-value	Canonical R	Wilks' Lambda	Chi-Sqr.	df	p-level
0	8,19309	0,94404	0,10877	27,7306	5	0,00004

Рис. 4.9.  $\chi^2$ -критерій канонічних коренів

Дані результати дозволяють оцінити кількість значимих коренів для інтерпретації та статистичну значимість дискримінантних функцій. *Coefficients for canonical variables (Коефіцієнти канонічних змінних)*. Ініціювання даної опції дозволяє побудувати таблиці нестандартизованих та стандартизованих коефіцієнтів дискримінантних функцій (рис. 4.10).

Variable	Raw Coeff f for Canoni	Variable	Standardize f for Canonic
	Root 1		Root 1
X1	-1,19813	X1	-0,936426
X2	8,78740	X2	0,421963
X3	-1,02098	X3	-0,227084
X4	0,48913	X4	0,110328
X5	-1,05978	X5	-0,432791
Constant	5,71665	Eigenval	8,193099
Eigenval	8,19310	Cum.Prop	1,000000
Cum.Prop	1,00000		

Рис. 4.10. Коефіцієнти канонічних змінних

Дані результати використовуються для визначення значень канонічних змінних для кожного спостереження та визначення ступеню і напряму впливу змінних у кожному дискримінантну функцію. *Factor structure (факторна структура)*. У даній таблиці (рис.4.11) подано об'єднані внутрішньогрупові кореляції змінних з відповідними дискримінантними функціями, які використовуються для змістовної інтерпретації функцій. *Means of canonical variables (середні значення канонічних змінних)*. У таблиці (рис. 4.11) подано середні значення для дискримінантних функцій, які дозволяють визначити групи, які найкраще ідентифікуються конкретною дискримінантною функцією.

Variable	Factor Struc Correlations (Pooled-w it	Group	Means of
	Root 1		Root 1
X1	-0,858571	A	-2,53495
X2	0,053157	B	2,85182
X3	0,005448		
X4	0,273883		
X5	-0,334115		

Рис. 4.11. Факторна структура та середні канонічних змінних

Для побудови дискримінантних функцій необхідно ініціювати опцію *Classification functions (функції класифікації)* на вкладці *Classification* (рис. 4.12).

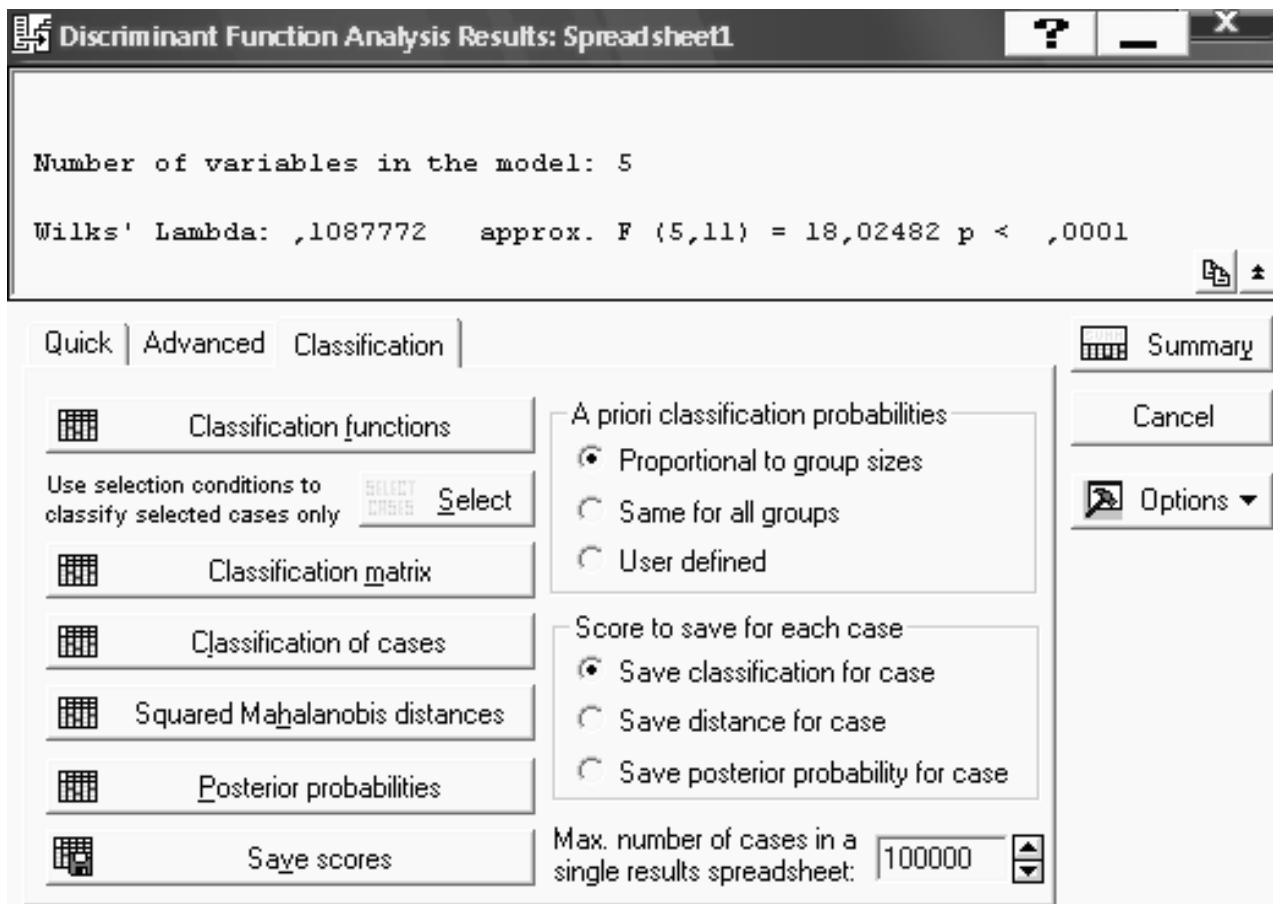


Рис. 4.12. Опції результатів класифікації

Дискримінантні функції для виділених класів станів підприємств (А, В) подано на рис. 4.13. Таким чином, лінійні дискримінантні функції мають такий вигляд:

Variable	Classification Functi	
	A p=,52941	B p=,47059
X1	13,045	6,591
X2	252,878	300,214
X3	42,081	36,582
X4	15,184	17,819
X5	8,301	2,592
Constant	-192,444	-162,621

Рис. 4.13. Дискримінантні функції

$$y_1 = -192,44 + 13,05 \cdot x_1 + 252,88 \cdot x_2 + 42,08 \cdot x_3 + 15,18 \cdot x_4 + 8,3 \cdot x_5$$

$$y_2 = -162,62 + 6,59 \cdot x_1 + 300,21 \cdot x_2 + 36,58 \cdot x_3 + 17,82 \cdot x_4 + 2,59 \cdot x_5$$

Класифікаційна матриця (*Classification matrix*) містить інформацію про кількість і відсоток коректно класифікованих спостережень у кожній з груп. Строки матриці – вихідні класи, стовпці – розпізнані класи за моделлю (рис. 4.14).

Classification Matrix (Spreadsheet)			
Rows: Observed classifications			
Columns: Predicted classifications			
Group	Percent Correct	A p=,52941	B p=,47059
A	100,0000	9	0
B	100,0000	0	8
Total	100,0000	9	8

Рис. 4.14. Матриця класифікацій

Для визначення приналежності об'єкта до одного з виділених класів на основі побудованої дискримінантної моделі необхідно скористатися такими опціями:

1) *Classification of cases (класифікація спостережень)* – таблиця класифікації для кожного спостереження;

2) *Squared Mahalanobis distances (квадрати відстаней Махаланобіса)* – таблиця квадратів відстаней Махаланобіса для кожного спостереження до центру відповідної групи;

3) *Posterior probabilities (апостеріорні ймовірності)* – таблиця апостеріорних ймовірностей приналежності кожного спостереження до відповідної групи.

Реалізацію даного напряму аналізу подано на рис. 4.15. Об'єкти, які неправильно класифіковані буде відмічено (\*).

Більш детальний аналіз змінних передбачає побудову моделей за методами покрокового включення й виключення незалежних змінних.

У даному модулі реалізовано метод покрокового включення змінних (*Forward stepwise*) і метод покрокового виключення (*Backward stepwise*). Вибір методів здійснюється на стартовій панелі модуля ініціюванням опції *Advanced options (stepwise analysis)* (рис. 4.16).



Classification of Cases (Spreadsheet)				Squared Mahalanobis Distances from				Posterior Probabilities (Spreadsheet)			
Case	Observed	1	2	Case	Observed	A	B	Case	Observed	A	B
	Classif.	p=,52941	p=,47059		Classif.	p=,52941	p=,47059		Classif.	p=,52941	p=,47059
1	A	A	B	1	A	3,77821	19,99433	1	A	0,999732	0,000268
2	A	A	B	2	A	3,17580	22,54519	2	A	0,999945	0,000055
3	A	A	B	3	A	4,18658	51,08083	3	A	1,000000	0,000000
4	A	A	B	4	A	6,53688	44,51579	4	A	1,000000	0,000000
5	A	A	B	5	A	1,09086	29,34067	5	A	0,999999	0,000001
6	A	A	B	6	A	1,50518	27,20268	6	A	0,999998	0,000002
7	B	B	A	7	B	23,55153	1,87625	7	B	0,000022	0,999978
8	B	B	A	8	B	55,92411	7,63650	8	B	0,000000	1,000000
9	B	B	A	9	B	16,87427	3,58540	9	B	0,001462	0,998538
10	B	B	A	10	B	21,12348	1,34014	10	B	0,000057	0,999943
11	B	B	A	11	B	47,18958	8,11675	11	B	0,000000	1,000000
12	A	A	B	12	A	2,81582	27,88547	12	A	0,999997	0,000003
13	B	B	A	13	B	28,48811	1,63615	13	B	0,000002	0,999998
14	B	B	A	14	B	41,50803	5,72591	14	B	0,000000	1,000000
15	A	A	B	15	A	3,79399	42,94798	15	A	1,000000	0,000000
16	B	B	A	16	B	35,09802	7,70224	16	B	0,000001	0,999999
17	A	A	B	17	A	10,49733	33,02277	17	A	0,999989	0,000011

Рис. 4.15. Матриці розпізнавання стану підприємств

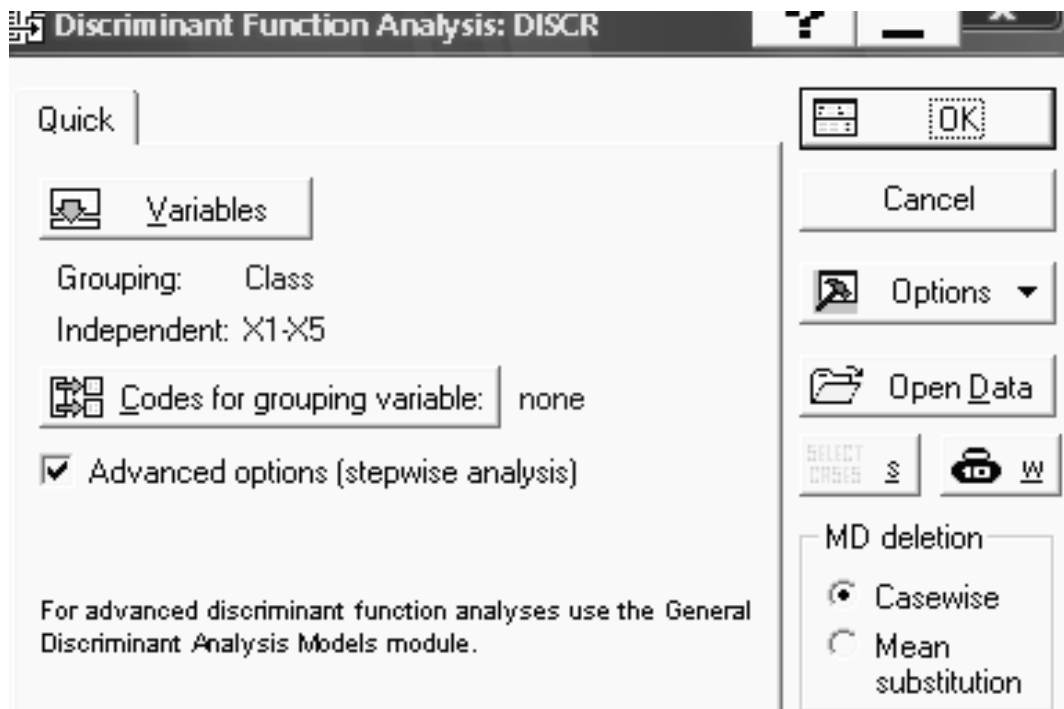


Рис. 4.16. Вибір опцій покрокового аналізу

Вибір методу оцінювання, порогові значення F-критерію включення або виключення, послідовність подання результатів вибирається у вкладці *Advanced* (рис. 4.17).

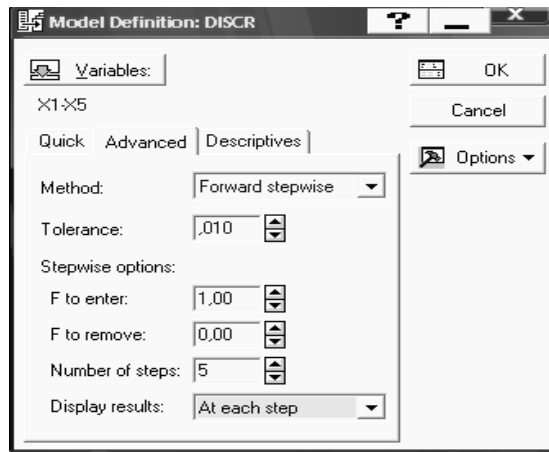


Рис. 4.17. Вибір методу покрокового дискримінантного аналізу

Послідовність етапів реалізації алгоритму покрокового включення (*Forward stepwise*) подано на рис. 4.18.

Stepwise Analysis - Step 0

Number of variables in the model: 0

Wilks' Lambda: 1,000000

Stepwise Analysis - Step 1

Number of variables in the model: 1

Last variable entered: X1  $F(1,16) = 90,59234$   $p < ,0000$

Wilks' Lambda: ,1420558 approx.  $F(1,15) = 90,59234$   $p < ,0000$

Stepwise Analysis - Step 2

Number of variables in the model: 2

Last variable entered: X2  $F(1,15) = 1,488011$   $p < ,2414$

Wilks' Lambda: ,1284077 approx.  $F(2,14) = 47,51384$   $p < ,0000$

Stepwise Analysis - Step 3

Number of variables in the model: 3

Last variable entered: X5  $F(1,14) = 1,528165$   $p < ,2367$

Wilks' Lambda: ,1149010 approx.  $F(3,13) = 33,38029$   $p < ,0000$

Stepwise Analysis - Step 3 (Final Step)

Number of variables in the model: 3

Last variable entered: X5  $F(1,13) = ,5423793$   $p < ,4745$

Wilks' Lambda: ,1149010 approx.  $F(3,13) = 33,38029$   $p < ,0000$

Рис. 4.18. Реалізація моделі покрокового включення змінних

Опції аналізу змінних які ввійшли в модель (*Variables in the model*), та виключені з моделі (*Variables not in the model*) подано на рис. 4.19. Результати аналізу змінних за методом покрокового включення змінних наведено на рис. 4.20, 4.21.

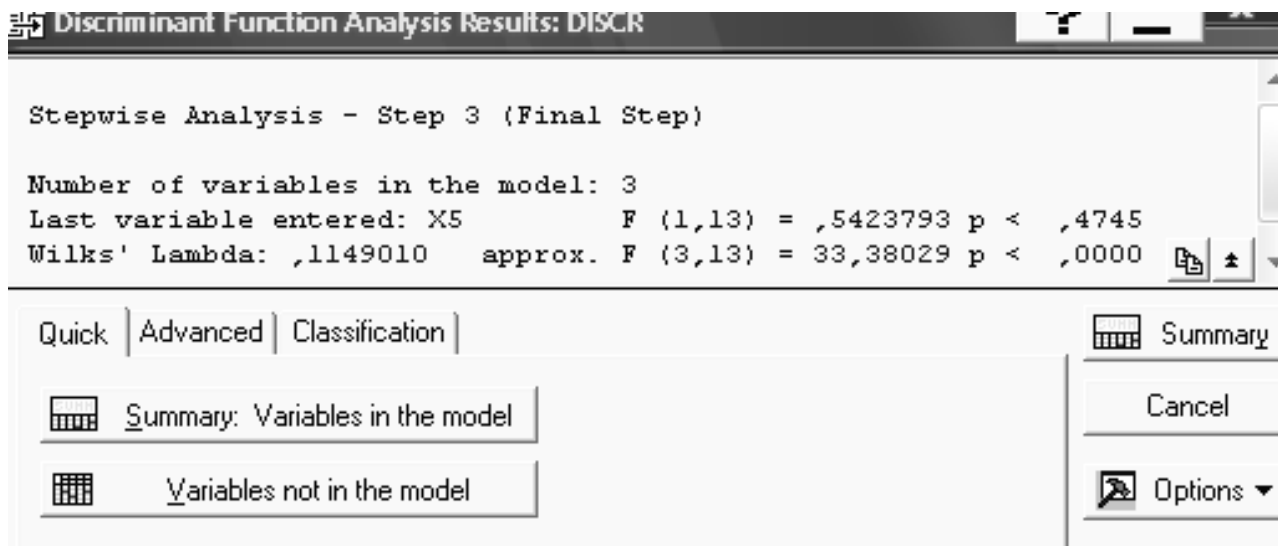


Рис. 4.19. Опції аналізу змінних

Discriminant Function Analysis Summary (Spreadsheet1) Step 3, N of vars in model: 3; Grouping: Class (2 grps) Wilks' Lambda: ,11490 approx. F (3,13)=33,380 p< ,0000						
N=17	Wilks' Lambda	Partial Lambda	F-remove (1,13)	p-level	Toler.	1-Toler. (R-Sqr.)
X1	0,471135	0,243881	40,30468	0,000025	0,922465	0,077535
X2	0,133311	0,861902	2,08293	0,172616	0,852662	0,147338
X5	0,128408	0,894814	1,52816	0,238259	0,911070	0,088930

Рис. 4.20. Аналіз змінних у моделі

Variables currently not in the model (Spreadsheet1) Df for all F-tests: 1,12						
N=17	Wilks' Lambda	Partial Lambda	F to enter	p-level	Toler.	1-Toler. (R-Sqr.)
X3	0,109932	0,956756	0,542379	0,475600	0,853756	0,146244
X4	0,113186	0,985074	0,181825	0,677359	0,975598	0,024402

Рис. 4.21. Аналіз змінних, виключених з моделі

Аналіз змінних у моделі методом покрокового виключення наведено на рис. 4.22.

Discriminant Function Analysis Summary (DISCR)						
Step 4, N of vars in model: 1; Grouping: Class (2 grps)						
Wilks' Lambda: ,14206 approx. F (1,15)=90,592 p< ,0000						
N=17	Wilks' Lambda	Partial Lambda	F-remove (1,15)	p-level	Toler.	1-Toler. (R-Sqr.)
X1	1,000000	0,142056	90,59234	0,000000	1,000000	0,00

Рис. 4.22. Аналіз змінних у моделі методом покрокового виключення

Для розпізнавання нових об'єктів (підприємств) на основі відомих значень коефіцієнтів необхідно додати в початкову таблицю вихідних даних нові спостереження (рис. 4.23).

16	6,7	0,79	0,79	0,39	1,24	B
17	9,42	0,7	0,7	0,72	2,03	A
18	12,11	0,68	1,44	0,43	1,37	
19	5,49	0,74	1,1	0,05	1,02	

Рис. 4.23. Вихідні дані для розпізнавання

Для визначення приналежності нових об'єктів до виділених класів необхідно скористатися такими опціями:

*Classification of cases (класифікація спостережень),*

*Squared Mahalanobis distances (квадрати відстаней Махаланобіса),*

*Posterior probabilities (апостеріорні ймовірності).*

Результати розпізнавання наведено на рис. 4.24.

Таким чином, можна дійти висновку, що підприємство № 18 з зазначеними характеристиками належить класу А (інвестиційно привабливі підприємства), а підприємство № 19 – класу В (підприємства аутсайтери).

Classification of Cases (Spreadsheet)				Squared Mahalanobis Distance				Posterior Probabilities (Spreadsheet)			
Case	Observed	1	2	Case	Observed	A	B	Case	Observed	A	B
	Classif.	p=,52941	p=,47059		Classif.	p=,52941	p=,47059		Classif.	p=,52941	p=,47059
1	A	A	B	1	A	3,77821	19,9943	1	A	0,99973	0,00026
2	A	A	B	2	A	3,17580	22,5451	2	A	0,99994	0,00005
3	A	A	B	3	A	4,18658	51,0808	3	A	1,00000	0,00000
4	A	A	B	4	A	6,53688	44,5157	4	A	1,00000	0,00000
5	A	A	B	5	A	1,09080	29,3406	5	A	0,99999	0,00000
6	A	A	B	6	A	1,50518	27,2026	6	A	0,99999	0,00000
7	B	B	A	7	B	23,5515	1,8762	7	B	0,00002	0,99997
8	B	B	A	8	B	55,9241	7,6365	8	B	0,00000	1,00000
9	B	B	A	9	B	16,8742	3,5854	9	B	0,00146	0,99853
10	B	B	A	10	B	21,1234	1,3401	10	B	0,00005	0,99994
11	B	B	A	11	B	47,1895	8,1167	11	B	0,00000	1,00000
12	A	A	B	12	A	2,8158	27,8854	12	A	0,99999	0,00000
13	B	B	A	13	B	28,4881	1,6361	13	B	0,00000	0,99999
14	B	B	A	14	B	41,5080	5,7259	14	B	0,00000	1,00000
15	A	A	B	15	A	3,7939	42,9479	15	A	1,00000	0,00000
16	B	B	A	16	B	35,0980	7,7022	16	B	0,00000	0,99999
17	A	A	B	17	A	10,4973	33,0227	17	A	0,99998	0,00001
18	---	A	B	18	---	17,2723	78,5461	18	---	1,00000	0,00000
19	---	B	A	19	---	44,9851	9,3938	19	---	0,00000	1,00000

Рис. 4.24. Результати розпізнавання

Діаграму розсіву об'єктів у просторі канонічних коренів можна побудувати ініціюванням клавіші *Scatterplot of canonical scores* в опції аналізу *Perform canonical analysis (канонічний аналіз)* на вкладці *Canonical scores* (рис. 4.25). Діаграму розсіву наведено на рис. 4.26.

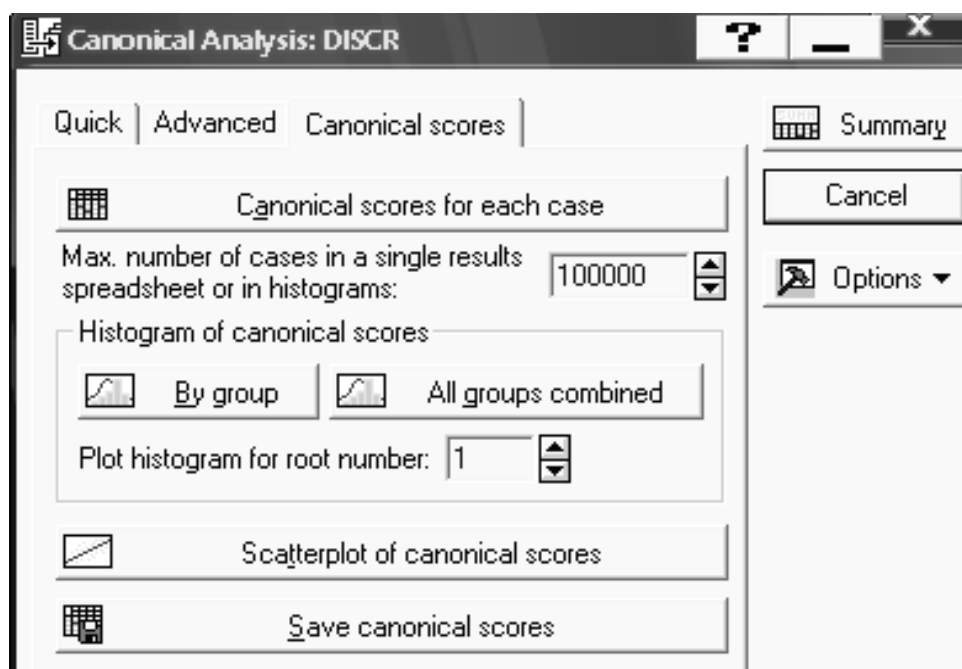


Рис. 4.25. Опції побудови діаграми

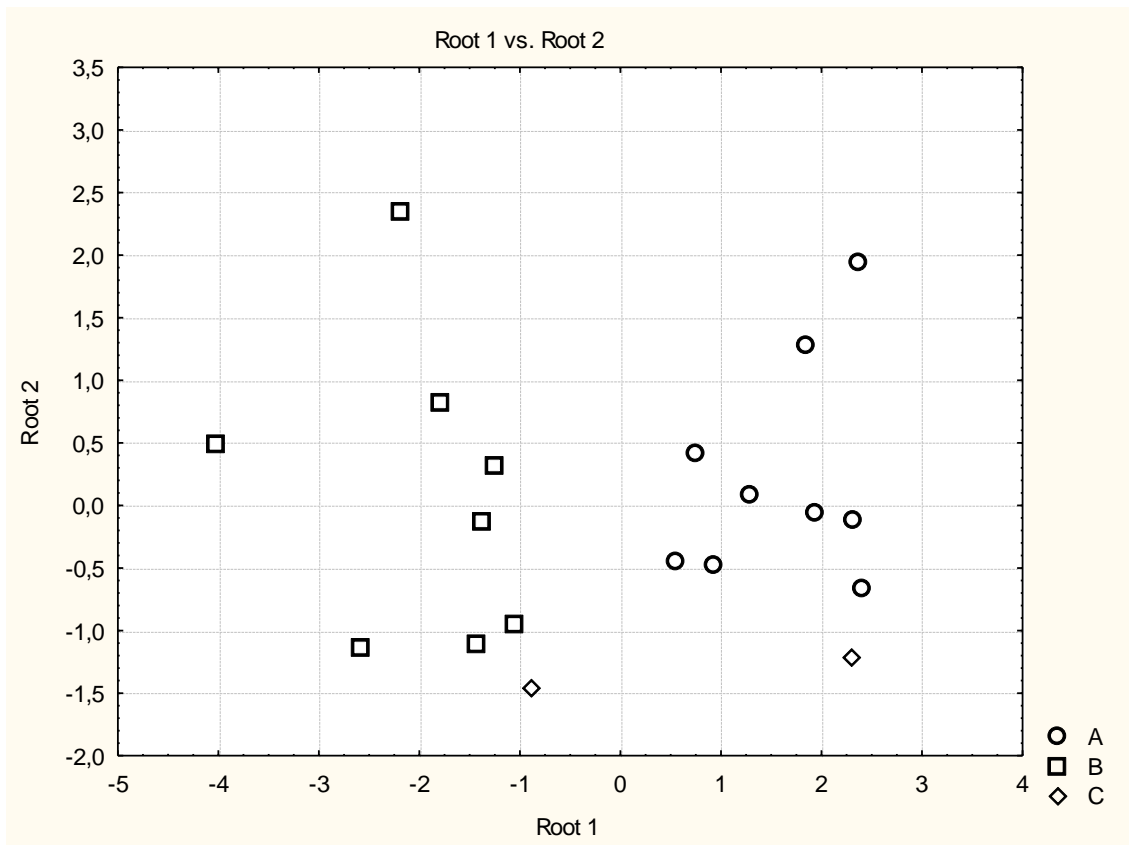


Рис. 4.26. Діаграма розсіву об'єктів у просторі канонічних коренів

Таким чином, результати діаграми підтверджують якість проведеної класифікації (виділені класи в просторі не перетинаються і мають значно відмінні значення канонічних коренів, які визначаються за дискримінантними функціями). Два нові підприємства, які було задано як клас С за результатами діаграми відносяться відповідно до класу А та класу В.

## Модуль 2. Методи опрацювання багатомірних об'єктів

### Лабораторна робота № 5. Методи скорочення простору ознак

**Мета** – закріплення теоретичного й практичного матеріалу за темою "Методи скорочення простору ознак", придбання навичок вибору показників-репрезентантів та побудови таксономічного показника рівня розвитку в середовищі *Microsoft Excel*.

**Завдання** – вибрати найбільш значимі для оцінювання стабільності фінансового стану досліджуваних підприємств показники (показники-репрезентанти) за допомогою методу "центра ваги", дати економічну інтерпретацію отриманим результатам:

1. Стандартизувати вихідні дані.
2. Розрахувати матриці Евклідових відстаней для кожної групи фінансових показників.
3. Вибрати на основі суми відстаней показник-репрезентант у кожній групі фінансових показників, дати економічну інтерпретацію отриманим результатам.

### **Методичні рекомендації**

У багатьох економічних дослідженнях виникає потреба в зменшенні числа ознак, що описують досліджувану галузь дійсності. Однак, зі скороченням кількості змінних мають дотримуватися деякі вимоги, для того щоб опис, що виходить, не спотворював дійсності. Існує досить великий спектр методів багатомірного аналізу, що дозволяє вирішувати завдання скорочення розмірності простору ознак. Такі методи підрозділяються на дві групи: методи побудови узагальнених показників, методи зменшення числа ознак.

Перша група методів спрямована на одержання інтегральної оцінки об'єктів, що мають багатоознакову природу, у вигляді деякої функції  $f(y_1, y_2, \dots, y_q)$ , що відбиває вплив усіх ознак, і у такий спосіб дозволяє впорядкувати досліджувані об'єкти.

Сутність роботи другої групи методів у заміні первісного набору  $q$  ознак  $y = (y_1, y_2, \dots, y_q)$  набором  $s$  діагностичних ознак  $x = (x_1, x_2, \dots, x_m)$ , ( $s < q$ ), які мають такі властивості:

ознаки не корельовані або слабо корельовані між собою;  
сильно корельовані з ознаками, що не входять у діагностичний набір.

Таким чином, друга група методів дозволяє виключити з первісної системи ознак ті, які дублюють інформацію, а також забезпечує вибір ознак, що найбільш повно відображають стан досліджуваних процесів.

Одним з методів другої групи є метод "центра ваги". Розглянемо порядок розрахункових процедур для вибору показників-репрезентантів на основі алгоритму даного методу в середовищі Microsoft Excel.

Алгоритм методу "центра ваги" містить такі основні кроки:

*Крок 1.* На першому кроці алгоритму формуються матриці вихідних даних за кожною групою показників стану об'єкта дослідження  $Y_1, Y_2, \dots, Y_q$ , де  $q$  – кількість груп показників. Для  $k$ -ої групи показників структуру цієї матриці може бути визначено таким чином:  $y_q = (y_{ij})$ ,  $i = [1; m]$ ,  $j = [1, n]$ , де  $y_{ij}$  – значення  $i$ -го показника в  $j$ -ому періоді;  $m$  – кількість показників, що входять у  $k$ -у групу;  $n$  – кількість періодів. Вихідні дані для вибору показників-репрезентантів наведено на рис. 5.1.

	A	B	C	D	E	F	G	H	I	J	K	L
1	№ підприємства/група показників	Показники ліквідності			Показники фінансової стабільності			Показники ділової активності			Показники рентабельності	
2		X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11
3	1	2,713	0,984	0,001	0,855	0,024	0,93	0,052	0,287	0,998	0,005	0,006
4	2	3,079	1,006	0,002	0,979	0,026	0,926	0,032	0,189	0,837	0,023	0,025
5	3	3,553	1,091	0,02	1,174	0,024	0,634	0,024	0,158	0,735	0,058	0,064
6	4	2,304	1,244	0,015	0,885	0,002	0,885	0,164	0,897	1,775	0,044	0,05
7	5	1,572	0,877	0,002	0,46	0,022	0,91	0,087	0,447	1,184	0,084	0,093
8	6	1,659	0,751	0,001	0,41	0,032	0,912	0,043	0,204	0,91	0,075	0,082

Рис. 5.1. Вихідні дані

$x_1$  – коефіцієнт поточної ліквідності,

$x_2$  – коефіцієнт швидкої ліквідності,

$x_3$  – коефіцієнт абсолютної ліквідності,

$x_4$  – коефіцієнт забезпеченості власними оборотними коштами,

$x_5$  – коефіцієнт маневреності власного капіталу,

$x_6$  – коефіцієнт заборгованості,

$x_7$  – коефіцієнт оборотності активів,

$x_8$  – коефіцієнт оборотності оборотних коштів,

$x_9$  – коефіцієнт оборотності запасів,

$x_{10}$  – коефіцієнт рентабельності активів,

$x_{11}$  – коефіцієнт рентабельності власного капіталу.

*Крок 2.* Оскільки показники може бути виражено в абсолютних і відносних величинах, а також мати різні одиниці виміру, то на другому кроці здійснюється процедура їхньої стандартизації за такою формулою:

$$z_{ij} = \frac{y_{ij} - \bar{y}_i}{S_i}$$



де  $z_{ij}$  – стандартизоване значення  $i$ -го показника в  $j$ -ому досліджуваному періоді;

$\bar{y}_i$  – середнє арифметичне значення  $i$ -го показника;

$s_i$  – стандартне відхилення  $i$ -го показника.

Результатом цього кроку є набір матриць стандартизованих значень показників кожної групи  $z_1, z_2, \dots, z_q$ . Для стандартизації вихідних даних спочатку необхідно розрахувати середнє арифметичне значення  $i$ -го показника та стандартне відхилення  $i$ -го показника в середовищі Microsoft Excel за допомогою функцій СРЗНАЧ і СТАНДОТКЛОН як це показано на рис. 5.2.

	A	B	C	D	E	F	G	H	I	J	K	L
№ підприємства/група показників		Показники ліквідності			Показники фінансової стабільності			Показники ділової активності			Показники рентабельності	
	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	
1	2,713	0,984	0,001	0,855	0,024	0,93	0,052	0,287	0,998	0,005	0,006	
2	3,079	1,006	0,002	0,979	0,026	0,926	0,032	0,189	0,837	0,023	0,025	
3	3,553	1,091	0,02	1,174	0,024	0,634	0,024	0,158	0,735	0,058	0,064	
4	2,304	1,244	0,015	0,885	0,002	0,885	0,164	0,897	1,775	0,044	0,05	
5	1,572	0,877	0,002	0,46	0,022	0,91	0,087	0,447	1,184	0,084	0,093	
6	1,659	0,751	0,001	0,41	0,032	0,912	0,043	0,204	0,91	0,075	0,082	
середнє значення	=СРЗНАЧ(В3:В8)											
стандартне відхилення	=СТАНДОТКЛОН(В3:В8)											

Рис. 5.2. Розрахунок середнього арифметичного та стандартного відхилення значення  $i$ -го показника

Результати розрахунку показників наведено на рис. 5.3

	A	B	C	D	E	F	G	H	I	J	K	L
№ підприємства/група показників		Показники ліквідності			Показники фінансової стабільності			Показники ділової активності			Показники рентабельності	
	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	
1	2,713	0,984	0,001	0,855	0,024	0,93	0,052	0,287	0,998	0,005	0,006	
2	3,079	1,006	0,002	0,979	0,026	0,926	0,032	0,189	0,837	0,023	0,025	
3	3,553	1,091	0,02	1,174	0,024	0,634	0,024	0,158	0,735	0,058	0,064	
4	2,304	1,244	0,015	0,885	0,002	0,885	0,164	0,897	1,775	0,044	0,05	
5	1,572	0,877	0,002	0,46	0,022	0,91	0,087	0,447	1,184	0,084	0,093	
6	1,659	0,751	0,001	0,41	0,032	0,912	0,043	0,204	0,91	0,075	0,082	
середнє значення	<b>2,4800</b>	<b>0,9922</b>	<b>0,0068</b>	<b>0,7938</b>	<b>0,0217</b>	<b>0,8662</b>	<b>0,0670</b>	<b>0,3637</b>	<b>1,0732</b>	<b>0,0482</b>	<b>0,0533</b>	
стандартне відхилення	<b>0,7866</b>	<b>0,1702</b>	<b>0,0084</b>	<b>0,2999</b>	<b>0,0102</b>	<b>0,1148</b>	<b>0,0523</b>	<b>0,2813</b>	<b>0,3761</b>	<b>0,0304</b>	<b>0,0334</b>	

Рис. 5.3. Результати розрахунку середнього арифметичного та стандартного відхилення значення  $i$ -го показника

Для стандартизації вихідних даних необхідно створити таблицю такої ж розмірності, що й таблиця вихідних даних. Стандартизацію вихідних даних в середовищі Microsoft Excel можна здійснити за допомогою функції НОРМАЛИЗАЦИЯ, аргументами якої є вихідні дані, середнє значення та стандартне відхилення показників рис.5.4.

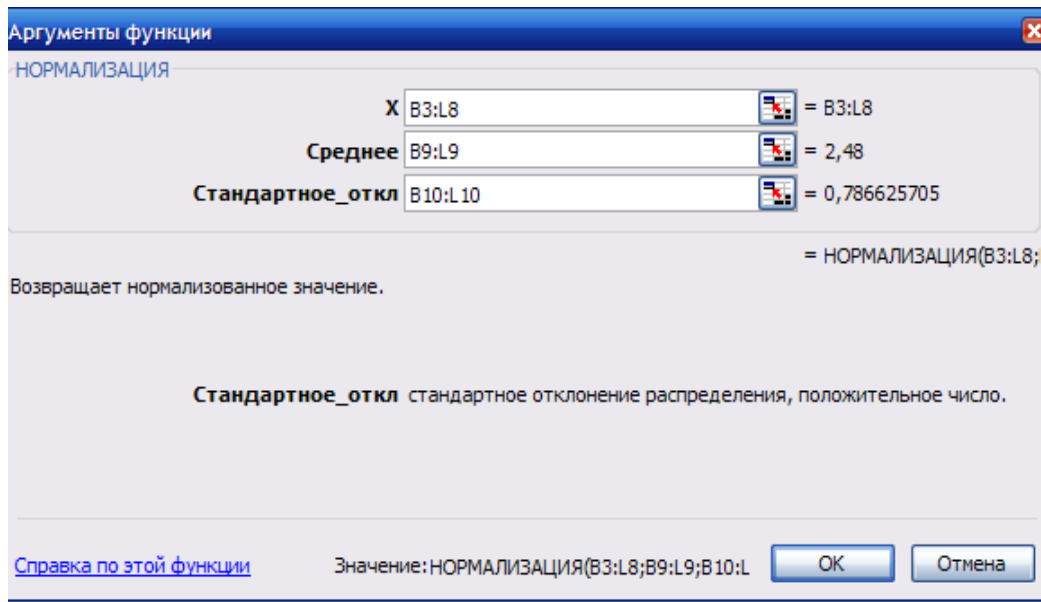


Рис. 5.4. Аргументи функції НОРМАЛИЗАЦИЯ

Для стандартизації вихідних даних необхідно в першій комірці створеної таблиці задати аргументи функції НОРМАЛИЗАЦИЯ, потім, починаючи з першої комірки, виділити діапазон, у якому будуть розраховані стандартизовані дані, натиснути клавішу F2, а потім одночасно натиснути клавіші Ctrl+Shift+Enter (Рис. 5.5)

№ підприємства/ група показників	Показники ліквідності			Показники фінансової стабільності			Показники ділової активності			Показники рентабельності	
	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11
1	=										
2	НОРМА										
3	ЛИЗАЦ										
4	ИЯ(B3:										
5	L8;B9:										
6	L9;B10:										
	L10)										

Рис. 5.5. Розрахунок стандартизованих значень показників

Результати стандартизації наведено на рис. 5.6.

№ підприємства/ група показників	Показники ліквідності			Показники фінансової стабільності			Показники ділової активності			Показники рентабельності	
	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11
1	0,2962	-0,0480	-0,6925	0,2040	0,2281	0,5559	-0,2867	-0,2726	-0,1999	-1,4209	-1,4190
2	0,7615	0,0813	-0,5737	0,6175	0,4236	0,5210	-0,6690	-0,6210	-0,6279	-0,8284	-0,8494
3	1,3641	0,5806	1,5630	1,2677	0,2281	-2,0218	-0,8220	-0,7312	-0,8991	0,3237	0,3198
4	-0,2237	1,4795	0,9694	0,3040	-1,9223	0,1640	1,8542	1,8962	1,8661	-0,1371	-0,0999
5	-1,1543	-0,6766	-0,5737	-1,1132	0,0326	0,3817	0,3823	0,2963	0,2947	1,1795	1,1892
6	-1,0437	-1,4168	-0,6925	-1,2800	1,0100	0,3991	-0,4588	-0,5677	-0,4338	0,8832	0,8594

Рис. 5.6. Таблиця стандартизованих значень показників

Крок 3. Описані вище обчислювальні процедури є основою для розрахунку матриць відстаней  $P_1, P_2, \dots, P_q$ , елементи яких відображають ступінь близькості показників усередині кожної групи. Як міра відстані використовується Евклідова відстань, що визначається за такою формулою:

$$\rho_E(z_i, z_j) = \sqrt{\sum_{l=1}^n (z_{il} - z_{jl})^2},$$

де  $\rho(z_i, z_j)$  – відстань між  $i$ -им і  $j$ -им показником групи;

$z_i, z_j$  – стандартизовані значення  $i$ -ого й  $j$ -ого показників групи в періоді  $l$ .

Матриці відстаней розраховуються для груп з кількістю показників більше двох. Це групи показників ліквідності, фінансової стабільності та ділової активності. Розрахунок відстаней для показників ліквідності наведено на рис. 5.7.

№ підприємства/ група показників	Показники ліквідності			Розрахунок відстаней		
	X1	X2	X3	$(x1-x2)^2$	$(x1-x3)^2$	$(x2-x3)^2$
1	0,2962	-0,0480	-0,6925	0,11846	0,97744	0,415347
2	0,7615	0,0813	-0,5737	0,46269	1,78283	0,429044
3	1,3641	0,5806	1,5630	0,61376	0,03956	0,964982
4	-0,2237	1,4795	0,9694	2,90094	1,42366	0,260144
5	-1,1543	-0,6766	-0,5737	0,22821	0,33704	0,010576
6	-1,0437	-1,4168	-0,6925	0,13921	0,12337	0,524696
			<b>Сума</b>	<b>4,46327</b>	<b>4,68391</b>	<b>2,604787</b>
			<b>Відстань</b>	<b>2,11265</b>	<b>2,16423</b>	<b>=КОРЕНЬ(G29)</b>
						<b>КОРЕНЬ(число)</b>

Рис. 5.7. Розрахунок відстаней для показників ліквідності

Матриця відстаней показників є симетричною з нульовими діагональними елементами (оскільки відстань від показника до самого себе дорівнює нулю). Матрицю відстаней для групи показників ліквідності наведено на рис. 5.8.

		Показники ліквідності		
		X1	X2	X3
X1		0	2,1126	2,1642
X2		2,1126	0	1,6139
X3		2,1642	1,6139	0

Рис. 5.8. Матриця відстаней для групи показників ліквідності

Аналогічно розраховуються матриці відстаней для груп показників фінансової стабільності та ділової активності (рис. 5.9).

	Показники фінансової стабільності				Показники ділової активності		
	X4	X5	X6		X7	X8	X9
X4	0	3,5542	4,0035	X7	0	0,1782	0,1537
X5	3,5542	0	3,1665	X8	0,178229	0	0,2288
X6	4,0035	3,1665	0	X9	0,153727	0,2288	0

Рис. 5.9. Матриці відстаней для груп показників фінансової стабільності та ділової активності

*Крок 4.* На четвертому кроці здійснюється вибір так званих показників-репрезентантів груп, які несуть у собі найбільш значиму інформацію, властиву групі за такими правилами:

у групах з одного елемента показники мають властивості, які сильно відрізняються від показників інших груп, тому вони належать до показників-еталонів (репрезентантів);

у групах, де кількість показників більше двох, розраховується сума відстаней кожного показника до інших показників групи:

$$\rho_i = \sum_{\substack{j=1 \\ j \neq i}}^m \rho(z_i, z_j),$$

де  $m$  – кількість показників групи.

До складу показників-репрезентантів входить показник з найменшою сумою відстаней:  $\rho_s = \min_i \rho_i$ . У групах, де число показників дорівнює двом, обчислюється сума відстаней показників, що належать до групи, від показників-репрезентантів, обраних за описаними вище правилами:

$$\rho_i = \sum_{\substack{j=1 \\ j \neq i}}^k \rho(z_i, z_j),$$

де  $k$  – кількість груп показників.

До репрезентантів належить той показник, у якого сума відстаней від відособлених елементів і елементів-репрезентантів, виділених із груп елементів із числом більше двох максимальна:  $\rho_s = \max_i \rho_i$ .

Таким чином, результатом 4-го кроку є набір показників-репрезентантів  $x = (x_1, x_2, \dots, x_n)$ , що описують найбільш важливі аспекти стану об'єкта дослідження. У таблиці вихідних даних відсутні групи з одним елементом, отже пропустимо цей крок і розглянемо групи з числом показників більше двох. Розрахуємо суми відстаней та виберемо показники-репрезентанти у цих групах на основі найменшої суми відстаней (рис. 5.10).

	Показники ліквідності			Сума
	X1	X2	X3	
X1	0	2,1126	2,1642	4,27688
X2	2,1126	0	1,6139	3,72658
X3	2,1642	1,6139	0	3,77817
	Показники фінансової стабільності			Сума
	X4	X5	X6	
X4	0	3,5542	4,0035	7,55771
X5	3,5542	0	3,1665	6,72073
X6	4,0035	3,1665	0	7,16997
	Показники ділової активності			Сума
	X7	X8	X9	
X7	0	0,1782	0,1537	0,33196
X8	0,1782	0	0,2288	0,40704
X9	0,1537	0,2288	0	0,38254

Рис. 5.10. Вибір показників-репрезентантів груп з кількістю показників більше двох

Таким чином, як показник-репрезентант в групі показників ліквідності був виділений показник швидкої ліквідності ( $x_2$ ), у групі показників фінансової стабільності – коефіцієнт маневреності власного капіталу ( $x_5$ ), у групі показників ділової активності – коефіцієнт оборотності активів ( $x_7$ ).

Оскільки група показників рентабельності містить тільки два показники, то для вибору показника-репрезентанта цієї групи знайдемо відстані кожного показника до раніше виділених показників-репрезентантів (рис. 5.11). Як показник-репрезентант вибирається той показник, у якого сума відстаней від показників-репрезентантів, виділених із груп елементів із кількістю більше двох, максимальна  $\rho_s = \max_i \rho_i$ .

	Показники рентабельності			Сума
	X2	X5	X7	
X10	3,7586	2,9688	3,0043	9,73175
X11	3,7377	3,0039	2,9707	9,71233

Рис. 5.11. Вибір показника-репрезентанта групи, де кількість показників дорівнює двом

Максимальною є сума відстаней від коефіцієнту рентабельності активів ( $x_{10}$ ) до вибраних раніше репрезентантів, отже цей показник і буде репрезентантом групи показників рентабельності.

Таким чином, до показників-репрезентантів належать: показник швидкої ліквідності ( $x_2$ ), коефіцієнт маневреності власного капіталу ( $x_5$ ), коефіцієнт оборотності активів ( $x_7$ ), коефіцієнт рентабельності активів ( $x_{10}$ ).

## Лабораторна робота № 6. Моделі і методи редукції даних

**Мета** – закріплення теоретичного й практичного матеріалу за темою "Моделі і методи редукції даних", придбання навичок побудови таксономічного показника рівня розвитку в середовищі *Microsoft Excel*.

**Завдання** – здійснити впорядкування поліграфічних підприємств рівнем стабільності фінансового стану на основі методу рівня розвитку, дати економічну інтерпретацію отриманим результатам:

1. На основі стандартизованих вихідних даних визначити еталон розвитку для наданих підприємств.

2. Розрахувати евклідові відстані між окремими підприємствами та еталоном.

3. Розрахувати таксономічний показник рівня розвитку на основі евклідових відстаней для кожного досліджуваного підприємства.

4. Здійснити впорядкування підприємств за стабільністю фінансового стану на основі таксономічного показника рівня розвитку, дати економічну інтерпретацію отриманим результатам.

### Методичні рекомендації

Для зіставлення об'єктів, які характеризуються більшим числом ознак, найчастіше застосовуються таксономічні процедури. Одним з методів дослідження багатомірних об'єктів є таксономічний показник рівня розвитку, запропонований З. Хельвігом. Цей показник являє собою синтетичну величину, "рівнодіючу" всіх ознак, що характеризують об'єкти, і дозволяє лінійно впорядкувати елементи досліджуваної сукупності.

Першим кроком процесу побудови таксономічного показника рівня розвитку є побудова елементів матриці спостережень, яку може бути подано таким чином:

$$X = \begin{bmatrix} X_{11} & X_{12} & \dots & X_{1j} & \dots & X_{1m} \\ X_{21} & X_{22} & \dots & X_{2j} & \dots & X_{2m} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ X_{i1} & X_{i2} & \dots & X_{ij} & \dots & X_{im} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ X_{\omega 1} & X_{\omega 2} & \dots & X_{\omega j} & \dots & X_{\omega m} \end{bmatrix},$$

де  $\omega$  – кількість досліджуваних об'єктів;

$m$  – кількість ознак;

$x_{ij}$  – значення  $j$ -ої ознаки для  $i$ -го об'єкта.

Оскільки ознаки, що включаються в матрицю спостережень, неоднорідні, проводиться стандартизація їхніх значень. Матрицю стандартизованих вихідних даних наведено на рис. 6.1. Наступний крок у розглянутій процедурі полягає в диференціації ознак матриці спостережень.

Усі змінні діляться на стимулятори й дестимулятори. Підставою поділу ознак на дві групи є характер впливу кожного з них на рівень розвитку досліджуваних об'єктів. Ознаки, що здійснюють позитивний, стимулюючий вплив на рівень розвитку об'єктів, називаються стимуляторами, на відміну від ознак-дестимуляторів (табл. 6.1).

Таблиця 6.1

**Класифікація ознак на стимулятори та дестимулятори**

Умовне позначення показника	Назва показника	Характер впливу на рівень стабільності фінансового стану	Група
X1	Коефіцієнт поточної ліквідності	Позитивний	Стимулятор
X2	Коефіцієнт швидкої ліквідності	Позитивний	Стимулятор
X3	Коефіцієнт абсолютної ліквідності	Позитивний	Стимулятор
X4	Коефіцієнт забезпеченості власними оборотними коштами	Позитивний	Стимулятор
X5	Коефіцієнт маневреності власного капіталу	Позитивний	Стимулятор
X6	Коефіцієнт заборгованості	Негативний	Дестимулятор
X7	Коефіцієнт оборотності активів	Позитивний	Стимулятор
X8	Коефіцієнт оборотності оборотних коштів	Позитивний	Стимулятор
X9	Коефіцієнт оборотності запасів	Позитивний	Стимулятор
X10	Коефіцієнт рентабельності активів	Позитивний	Стимулятор
X11	Коефіцієнт рентабельності власного капіталу	Позитивний	Стимулятор

Поділ ознак на стимулятори й дестимулятори є основою для побудови так званого еталона розвитку, що є точкою з координатами:

$$P_0(z_{01}, z_{02}, \dots, z_{0m})$$

$$\text{де } z_{0s} = \max_r z_{rs}, \text{ якщо } s \in I,$$

$$z_{0s} = \min_r z_{rs}, \text{ якщо } s \notin I, (s = 1, \dots, m),$$

$I$  – множина стимуляторів,

$z_s$  – стандартизоване значення ознаки  $s$  для об'єкта  $r$ .

Обчислення еталону розвитку в середовищі *Microsoft Excel* здійснюється за допомогою функцій МАКС та МИН (рис. 6.1).



№ підприємства/ група показників	Показники ліквідності			Показники фінансової стабільності			Показники ділової активності			Показники рентабельності	
	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11
1	0,2962	-0,0480	-0,6925	0,2040	0,2281	0,5559	-0,2867	-0,2726	-0,1999	-1,4209	-1,4190
2	0,7615	0,0813	-0,5737	0,6175	0,4236	0,5210	-0,6690	-0,6210	-0,6279	-0,8284	-0,8494
3	1,3641	0,5806	1,5630	1,2677	0,2281	-2,0218	-0,8220	-0,7312	-0,8991	0,3237	0,3198
4	-0,2237	1,4795	0,9694	0,3040	-1,9223	0,1640	1,8542	1,8962	1,8661	-0,1371	-0,0999
5	-1,1543	-0,6766	-0,5737	-1,1132	0,0326	0,3817	0,3823	0,2963	0,2947	1,1795	1,1892
6	-1,0437	-1,4168	-0,6925	-1,2800	1,0100	0,3991	-0,4588	-0,5677	-0,4338	0,8832	0,8594
Об'єкт-еталон (P0)	1,3641	1,4795	1,5630	1,2677	1,0100	-2,0218	1,8542	1,8962	1,8661	1,1795	1,1892
	max	max	max	max	max	min	max	max	max	max	max

**Рис. 6.1. Обчислення еталону розвитку**

Відстань між окремими точками-єдиничами й точкою  $P_0$ , що є еталоном розвитку, позначається  $C_{i0}$  й розраховується на основі евклідової відстані за формулою:

$$C_{i0} = \sqrt{\sum_{j=1}^m (z_{ij} - Z_{0j})^2}$$

Розрахунок відстаней між окремими підприємствами й об'єктом-еталоном наведено на рис. 6.2.

№ підприємства/ група показників	Показники ліквідності			Показники фінансової стабільності			Показники ділової активності			Показники рентабельності		Сума	Відстань
	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11		
Z1	0,2962	-0,0480	-0,6925	0,2040	0,2281	0,5559	-0,2867	-0,2726	-0,1999	-1,4209	-1,4190		
Z2	0,7615	0,0813	-0,5737	0,6175	0,4236	0,5210	-0,6690	-0,6210	-0,6279	-0,8284	-0,8494		
Z3	1,3641	0,5806	1,5630	1,2677	0,2281	-2,0218	-0,8220	-0,7312	-0,8991	0,3237	0,3198		
Z4	-0,2237	1,4795	0,9694	0,3040	-1,9223	0,1640	1,8542	1,8962	1,8661	-0,1371	-0,0999		
Z5	-1,1543	-0,6766	-0,5737	-1,1132	0,0326	0,3817	0,3823	0,2963	0,2947	1,1795	1,1892		
Z6	-1,0437	-1,4168	-0,6925	-1,2800	1,0100	0,3991	-0,4588	-0,5677	-0,4338	0,8832	0,8594		
Об'єкт-еталон (Z0)	1,3641	1,4795	1,5630	1,2677	1,0100	-2,0218	1,8542	1,8962	1,8661	1,1795	1,1892		
Z1-Z0													
(Z1-Z0) <sup>2</sup>	1,1403	2,3331	5,0869	1,1316	0,6115	6,6443	4,5835	4,7034	4,2681	6,7619	6,8026	44,0671	6,6383
(Z2-Z0) <sup>2</sup>	0,3631	1,9550	4,5655	0,4228	0,3439	6,4660	6,3666	6,3360	6,2201	4,0316	4,1558	41,2264	6,4208
(Z3-Z0) <sup>2</sup>	0,0000	0,8079	0,0000	0,0000	0,6115	0,0000	7,1616	6,9030	7,6464	0,7324	0,7558	24,6187	4,9617
(Z4-Z0) <sup>2</sup>	2,5211	0,0000	0,3523	0,9288	8,5987	4,7777	0,0000	0,0000	0,0000	1,7335	1,6618	20,5738	4,5358
(Z5-Z0) <sup>2</sup>	6,3421	4,6486	4,5655	5,6689	0,9554	5,7768	2,1664	2,5596	2,4693	0,0000	0,0000	35,1526	5,9290
(Z6-Z0) <sup>2</sup>	5,7973	8,3885	5,0869	6,4907	0,0000	5,8608	5,3497	6,0704	5,2896	0,0878	0,1087	48,5304	=КОРЕНЬ(M27)

**Рис. 6.2. Розрахунок відстаней між окремими підприємствами й об'єктом-еталоном**

Результати розрахунку евклідових відстаней від кожного підприємства до еталону наведено на рис. 6.3

	№ підприємства	Евклідова відстань до еталону (с <sub>і0</sub> )
31		
32	1	6,6383
33	2	6,4208
34	3	4,9617
35	4	4,5358
36	5	5,9290
37	6	6,9664

Рис. 6.3. Евклідові відстані від кожного підприємства до еталону

Обчислені відстані є вихідними величинами, використовуваними при розрахунку показника рівня розвитку:

$$d_i^* = 1 - \frac{c_{i0}}{c_0}; \quad c_0 = \bar{c}_0 + 2S_0;$$

$$\bar{c}_0 = \frac{1}{W} \sum_{i=1}^W c_{i0}; \quad S_0 = \sqrt{\frac{1}{W} \sum_{i=1}^W (c_{i0} - \bar{c}_0)^2}.$$

Розрахунок показника рівня розвитку наведено на рис. 6.4.

30	№ підприємства	Евклідова відстань до еталону (с <sub>і0</sub> )	$(c_{i0} - \bar{c}_0)^2$	$d_i^*$
31				
32	1	6,6383	=(B32-\$B\$38)^2	=1-B32/\$B\$40
33	2	6,4208	=(B33-\$B\$38)^2	=1-B33/\$B\$40
34	3	4,9617	=(B34-\$B\$38)^2	=1-B34/\$B\$40
35	4	4,5358	=(B35-\$B\$38)^2	=1-B35/\$B\$40
36	5	5,929	=(B36-\$B\$38)^2	=1-B36/\$B\$40
37	6	6,9664	=(B37-\$B\$38)^2	=1-B37/\$B\$40
38	$\bar{c}_0$	=CPЗНАЧ(B32:B37)	=CPЗНАЧ(C32:C37)	
39	S <sub>0</sub>	=КОРЕНЬ(C38)		
40	c <sub>0</sub>	=B38+2*B39		
41				

Рис. 6.4. Розрахунок таксономічного показника рівня розвитку

Отримані результати розрахунку наведено на рис. 6.5.

	№ підприємства	Евклідова відстань до еталону ( $c_{i0}$ )	$(c_{i0} - \bar{c}_0)^2$	$d_i^*$
31				
32	1	6,6383	0,53238	<b>0,1354</b>
33	2	6,4208	0,26226	<b>0,1637</b>
34	3	4,9617	0,89669	<b>0,3538</b>
35	4	4,5358	1,88465	<b>0,4092</b>
36	5	5,9290	0,00041	<b>0,2278</b>
37	6	6,9664	1,11875	<b>0,0927</b>
38	$\bar{c}_0$	5,9087	0,7825	
39	S0	0,8846		
40	C0	7,6779		
41				

Рис. 6.5. **Значення таксономічного показника рівня розвитку підприємств**

Інтерпретація показника рівня розвитку така: чим ближче значення показника рівня розвитку відносно одиниці, тим на більш високому рівні розвитку перебуває об'єкт. Впорядкуємо підприємства за рівнем розвитку (рис. 6.6)

	$d_i^*$	№ підприємства
43		
44	0,4092	4
45	0,3538	3
46	0,2278	5
47	0,1637	2
48	0,1354	1
49	0,0927	6
50		

Рис. 6.6. **Підприємства, впорядковані за рівнем розвитку**

Таким чином, найбільш стійким є четверте підприємство, а найгірше фінансове становище характерне для шостого підприємства.

## Лабораторна робота № 7. Моделі і методи факторного аналізу

**Мета** – закріплення теоретичного й практичного матеріалу за темою "Моделі факторного аналізу", придбання навичок роботи в модулі *Factor Analysis*.

**Завдання** – необхідно провести факторний аналіз вихідних даних для рішення задачі зниження розмірності для вибірових даних у модулі *Factor Analysis* ППП *Statistica*:

1. Побудувати модель факторного аналізу, визначити її характеристики і особливості.

2. Визначити мінімально необхідне для відображення всіх істотних кореляційних зв'язків число головних компонент, використовуючи критерії Кайзера, кумулятивної дисперсії та "кам'янистого осипу".

3. Провести аналіз значень факторних навантажень, побудувати графіки розсіву, проаналізувати значення при застосуванні процедури обертання.

4. За допомогою методу головних компонентів виявити латентні фактори, що здійснюють вплив на рівень досліджуваного процесу.

5. Зробити висновки, дати економічну інтерпретацію отриманим головним компонентам та розрахувати коефіцієнт їх інформативності.

### Методичні рекомендації

Для рішення задач зниження розмірності признакового простору в ППП *Statistica* існує набір підпрограм, об'єднаних у групу методів *Factor Analysis* (*Факторний аналіз*). Розглянемо порядок роботи в даному модулі.

Таблиця вихідних даних для рішення задачі подана на рис. 7.1. У якості чинників впливу на рівень інвестиційної привабливості розглянемо наступні коефіцієнти: коефіцієнт рентабельності капіталу, коефіцієнт автономії, коефіцієнт відновлення основних фондів, коефіцієнт фондівіддачі, коефіцієнт середнього виробітку та питома вага зайнятих повний робочий день. Щоб розпочати обчислювальні процедури необхідно ввійти в позицію меню *Statistics / Multivariate Exploratory Techniques / Factor Analysis* (рис. 7.2).

	1	2	3	4	5	6
	RENT	AVT	OBN_OF	FOND	VUR	ZAN
1	0	0,699	0	0,2108	3,7747	0,6224
2	0	0,7553	0,0371	0,2558	9,2943	0,6962
3	0,1439	0,8697	0,0379	1,818	42,4257	0,8764
4	0,0774	0,6514	0,0046	1,476	14,344	0,9009
5	0,0136	0,8184	0,0019	0,261	5,6978	0,8704
6	0,014	0,7408	0	0,1469	2,9477	0,5315
7	0,008	0,8807	0,002	0,0792	2,1941	0,5816
8	0,0068	0,9603	0,0007	0,1541	6,7626	0,9538
9	0	0,8232	0	0,1123	0,66	0,24
10	0	0,6785	0,0012	0,5611	0,6343	0,883
11	0	0,7373	0,0005	0,1871	4,7804	0,8017
12	0	0,7505	0	0,1651	5,0292	0,6663
13	0	0,7503	0,0005	0,0459	2,4366	0,6456
14	0	0,8965	0,0002	0,1742	5,2886	0,5134
15	0	0,868	0,0114	0,169	10,1006	0,8089
16	0,1346	0,9513	0,0178	1,697	14,523	0,9054
17	0,043	0,7991	0	0,9078	23,554	0,8538
18	0,1849	0,9751	0	8,223	23,727	0,6459
19	0	0,976	0	0,0395	4,1623	0,7048
20	0,1042	0,8205	0	1,094	12,8439	0,7404
21	0	0,8919	0	0,0821	1,9723	0,3239
22	0	0,8295	0	0,1164	2,1318	0,6084
23	0,043	0,7723	0	0,4023	3,4007	0,75
24	0	0,826	0	0,1654	7,6899	0,062
25	0,0087	0,6935	0,0006	0,535	14,6248	0,6913

Рис. 7.1. Вихідні дані

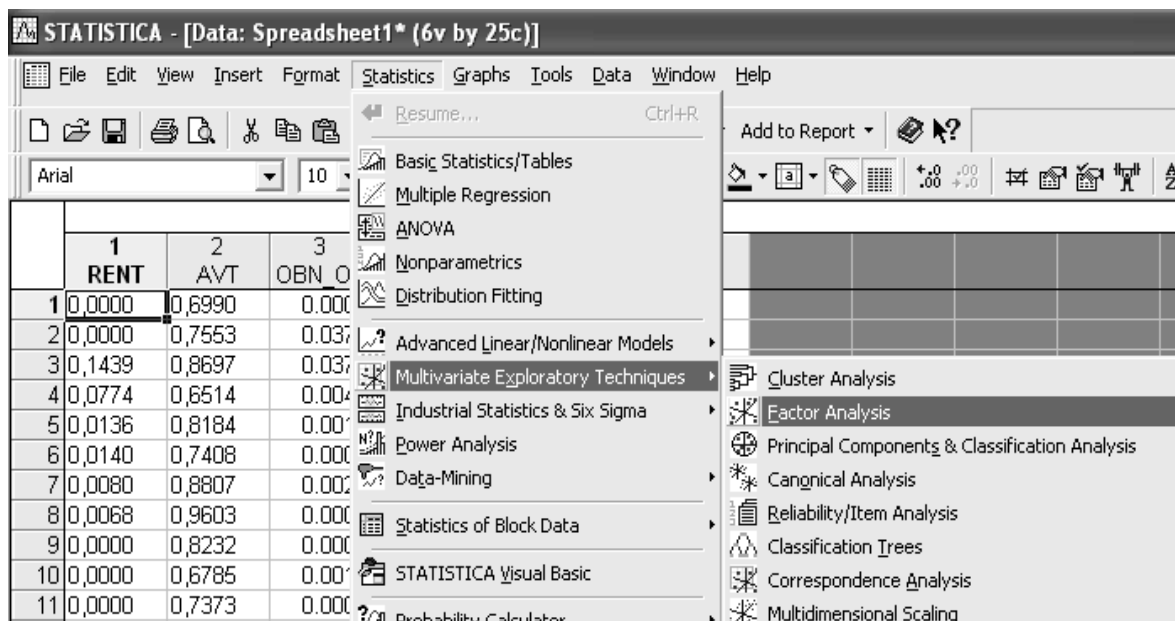


Рис. 7.2. Вибір модуля

Після підтвердження вибору модуля перед вами з'явиться стартова панель модуля, де необхідно задати вихідні параметри моделювання: *Variable* (змінні), *Input* (тип вихідних даних), *MD deletion* (спосіб обробки пропущених даних) (рис. 7.3).

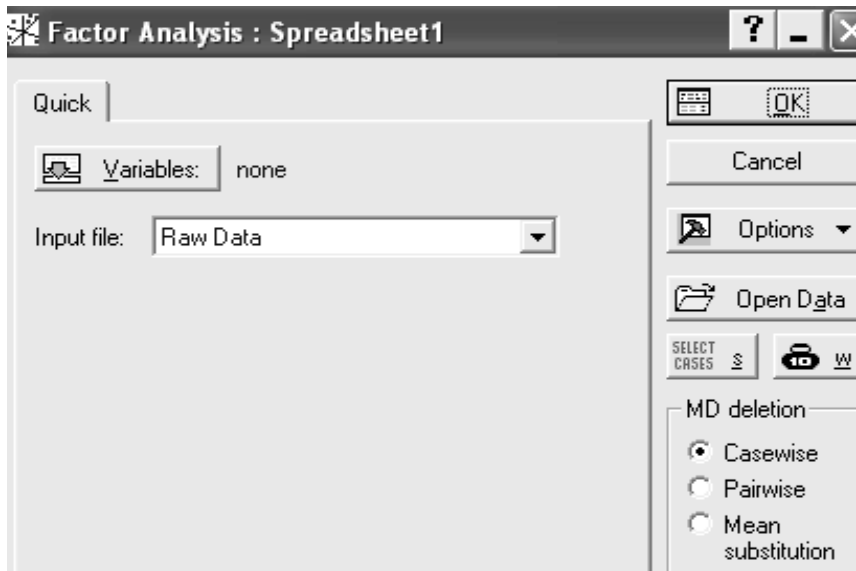


Рис. 7.3. Стартова панель модуля

Ініціюйте кнопку *Variable* (змінні), і у вікні, що з'явилося, вкажіть показники, за якими будете здійснювати аналіз (рис. 7.4).

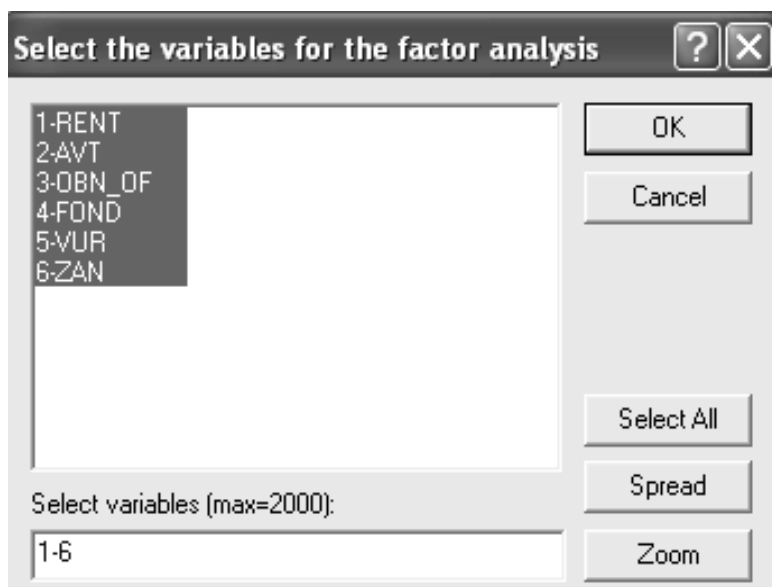


Рис. 7.4. Вибір змінних для аналізу

Після вибору змінних, підтвердіть свій вибір натисканням кнопки **OK**. Ініціюйте кнопку **OK**, і перед Вами з'явиться діалогове вікно, що дозволяє визначити: *Extraction method* (методи обробки даних) – головних компонент та аналіз головних факторів (метод спільностей, ітеративних спільностей, максимальної правдоподібності, центроїдний метод та метод головних шкал); *Max of factors* (максимальне число факторів) – може дорівнювати числу аналізованих показників; *Mini eigenvalue* (мінімальне власне число) – якщо власні значення будуть меншими за встановлений мінімум, то вони будуть ігноруватися (рис. 7.5).

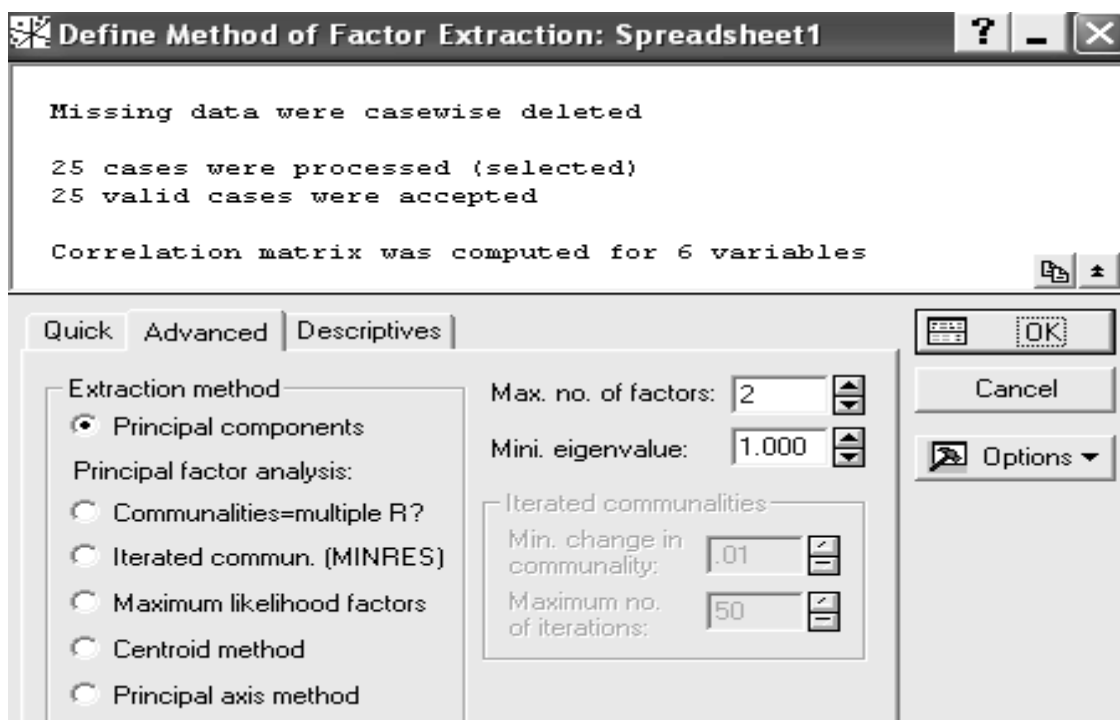


Рис. 7.5. Вибір методу обробки даних

Після підтвердження вибору перед Вами з'явиться діалогове вікно, що дозволяє переглянути результати факторного аналізу (рис. 7.6). У верхній частині вікна представлена основна інформація побудови моделі факторного аналізу: *Number of variables* (кількість змінних), *Method* (метод обробки даних), *log (10) determinant of correlation matrix* (логарифм детермінанта кореляційної матриці), *Number of factors* (число виділених факторів), *Eigenvalues* (власні числа). У нижній частині знаходяться функціональні кнопки, кожна з яких пов'язана з певним напрямком аналізу.

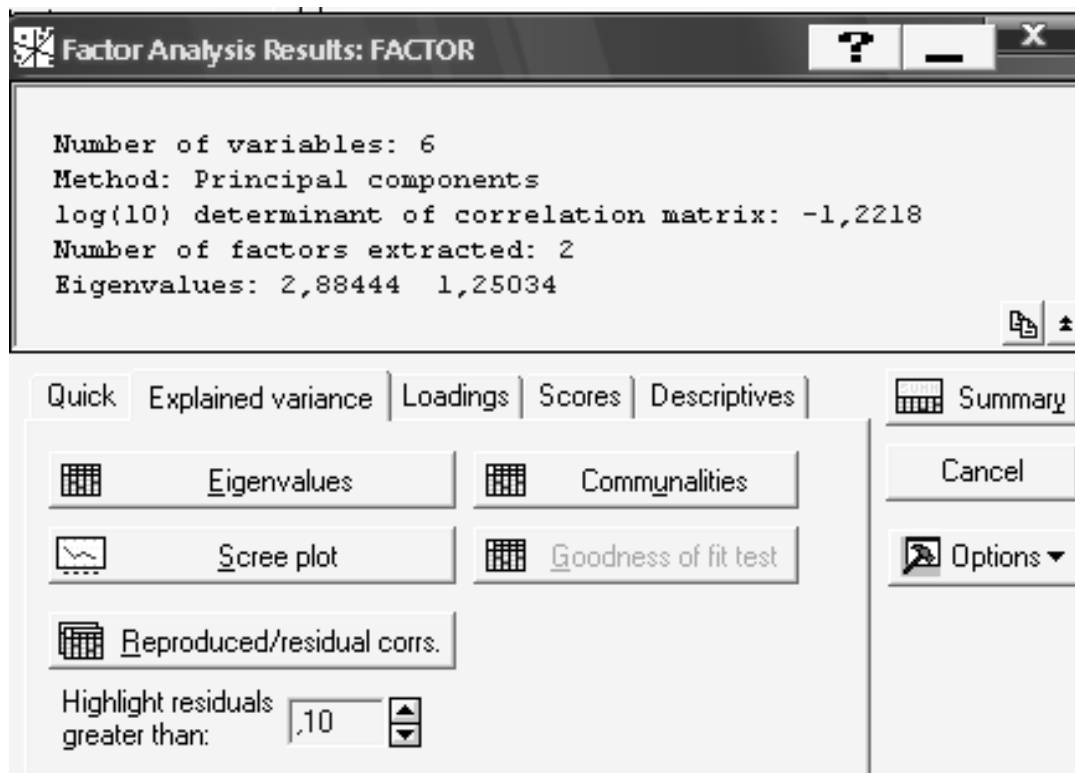


Рис. 7.6. Вибір напрямку аналізу

Ініціювання кнопки *Eigenvalues* дозволяє оцінити рівень інформативності головних компонент (рис. 7.7).

Eigenvalues (FACTOR)				
Extraction: Principal components				
Value	Eigenvalue	% Total variance	Cumulative Eigenvalue	Cumulative %
1	2,88443	48,0739	2,88443	48,0739
2	1,25033	20,8389	4,13477	68,9129

Рис. 7.7. Оцінювання рівня інформативності головних компонент

Подана на рис. 7.7 таблиця містить такі характеристики для оцінювання значимості компонентів: *Eigenvalues* – власні числа, *% Total variance* – відсоток загальної пояснювальної дисперсії, *Cumulative Eigenvalue* – кумулятивні власні числа, *Cumulative %* – кумулятивна дисперсія. Отримані головні компоненти пояснюють 69 % всієї варіації вихідної системи показників, причому перша головна компонента пояснює 48 % загальної дисперсії, а друга – 21 %.



Одним із критеріїв відбору числа головних факторів є критерій "кам'янистого осипу" (*Scree plot*), графік якого подано на рис. 7.8.

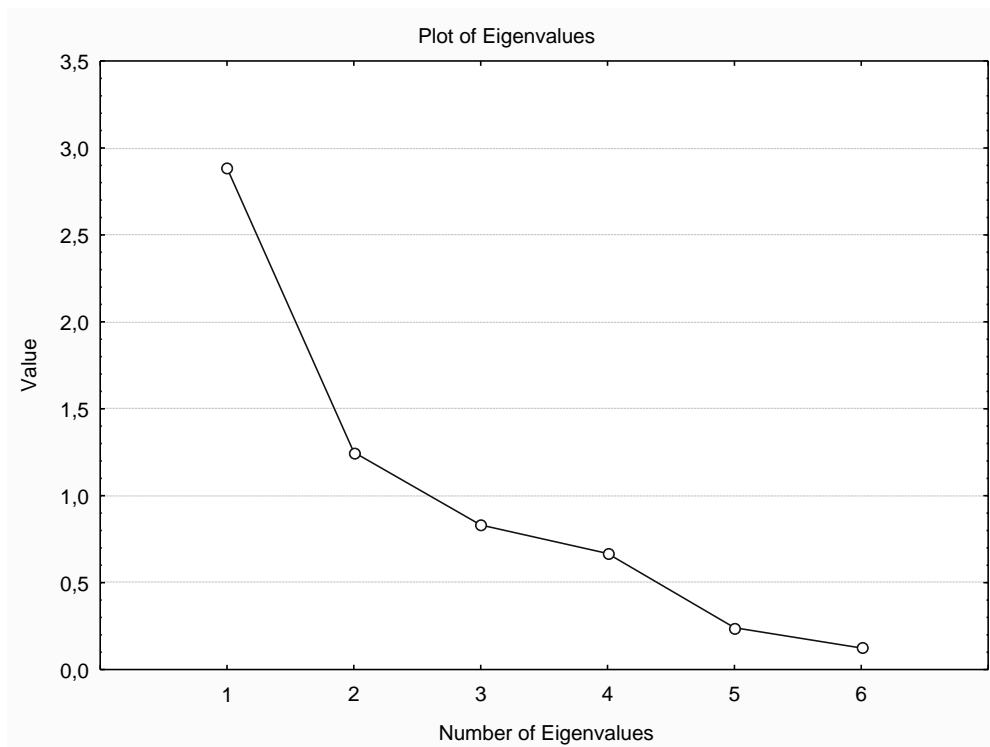


Рис. 7.8. Графік "кам'янистого осипу"

Висновок про кількість головних факторів можна зробити, знайшовши таку абсцису на графіку, в якій графік власних чисел починає візуально згладжуватися в напрямку правої, пологої його частини, і, таким чином, зменшення значень максимально сповільнюється. Відповідно до даного критерію, варто вибрати для дослідження два головних фактора.

Значення спільностей для виділених факторів подано на рис. 7.9.

Variable	Communalities (FACTOR) Extraction: Principal components Rotation: Unrotated		
	From 1 Factor	From 2 Factors	Multiple R-Square
RENT	0,859548	0,884037	0,811207
AVT	0,149911	0,519718	0,146115
OBN_OF	0,294399	0,585867	0,410079
FOND	0,594990	0,783226	0,705217
VUR	0,782680	0,814528	0,691765
ZAN	0,202910	0,547400	0,198513

Рис. 7.9. Значення спільностей

Для інтерпретації значень сформованих головних компонентів необхідно зайти у вкладку *Loadings (навантаження)* і ініціювати клавішу *Summary (Factor Loading)* (рис. 7.10). На рис. 7.11. подано значення факторних навантажень які є коефіцієнтами кореляції між змінними та виділеними факторами. Виявлення та інтерпретація закономірностей у таблицях факторних навантажень є достатньо трудомістким процесом, процедура спрощується при графічному поданні факторних навантажень.

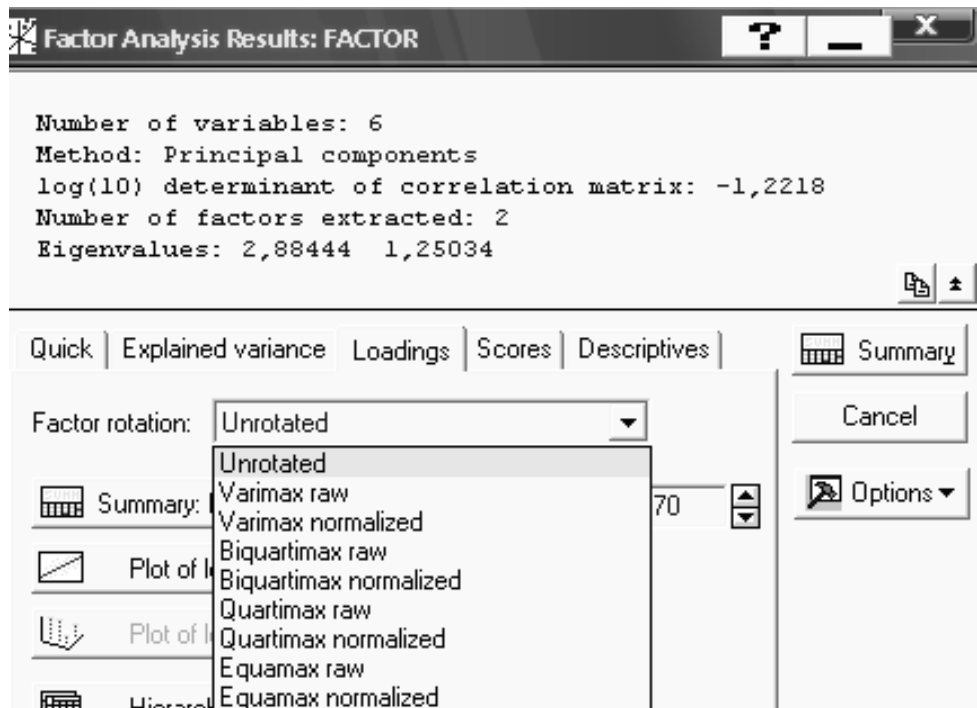


Рис. 7.10. Вибір опцій для інтерпретації компонент

Variable	Factor Loadings (Unrotated)	
	Factor 1	Factor 2
RENT	0,927118	-0,156490
AVT	0,387184	-0,608117
OBN_OF	0,542585	0,539878
FOND	0,771356	-0,433861
VUR	0,884692	0,178460
ZAN	0,450455	0,586933
Expl.Var	2,884438	1,250337
Prp.Totl	0,480740	0,208390

Рис. 7.11. Факторні навантаження

Ініціювавши клавішу *Plot of factor loadings (графік навантажень)*, отримаємо графік (рис. 7.12), який ілюструє співвідношення між факторами і групами змінних.

Для інтерпретації отриманих головних компонентів значення факторних навантажень мають виявляти закономірності. Факторні навантаження мають об'єднувати змінні в групи, для яких коефіцієнти кореляції з факторами набувають більші значення за однією групою і менші за іншою.

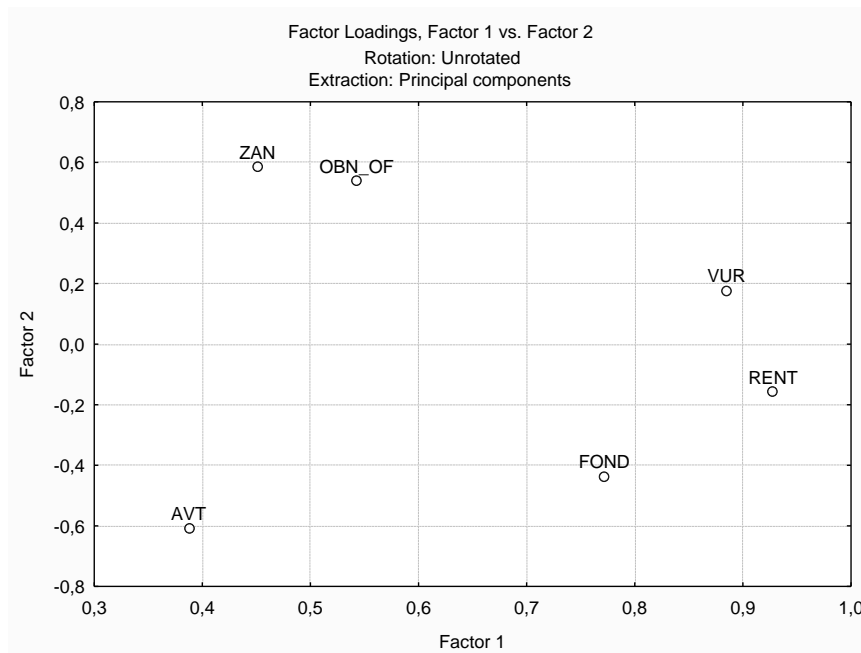


Рис. 7.12. Графічне подання факторних навантажень

Для отримання інтерпретуємих рішень в модулі реалізовано процедури повернення факторів (див. рис. 7.10):

*Varimax raw* – варімакс,

*Varimax normalized* – варімакс нормалізований;

*Biquartimax raw* – біквартімакс,

*Biquartimax normalized* – біквартімакс нормалізований;

*Quartimax raw* – квартімакс,

*Quartimax normalized* – квартімакс нормалізований;

*Equamax raw* – еквімакс,

*Equamax normalized* – еквімакс нормалізований.

На рис. 5.13. наведено факторні навантаження за методом *Equamax raw* (еквімакс), аналіз значень яких дозволяє дійти висновку, що перша головна компонента характеризує ефективність виробничо-господарської

діяльності підприємств (коефіцієнти рентабельності капіталу, фондівіддачі, середнього виробітку), друга – їх виробничий потенціал (відновлення основних фондів та питома вага зайнятих повний робочий день).

Variable	Factor Loadings (Equation 1)	
	Factor 1	Factor 2
RENT	0,918352	0,201659
AVT	0,586544	-0,419147
OBN_OF	0,301257	0,703641
FOND	0,877648	-0,113841
VUR	0,753717	0,496425
ZAN	0,198214	0,712819
Expl. Var	2,655804	1,478970
Prp. Totl	0,442634	0,246495

Рис. 7.13. Факторні навантаження після процедури повернення

На рис. 7.14. наведено групи опцій для отримання значень головних компонент. Ініціювавши клавішу *Factor Score coefficients* (рис. 7.15) отримаємо таблицю коефіцієнтів лінійних рівнянь регресії, за якими розраховуються значення факторів для спостережень.

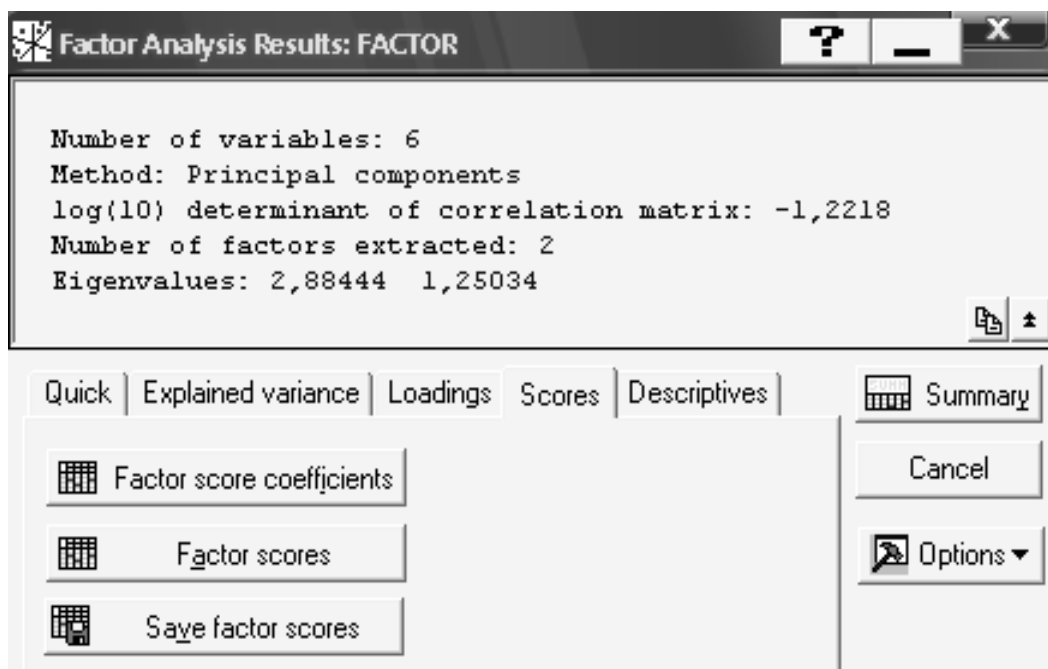


Рис. 7.14. Вибір опцій значень головних компонент

Variable	Factor Score Coefficient	
	Factor 1	Factor 2
RENT	0,344904	0,004155
AVT	0,306412	-0,400847
OBN_OF	0,012943	0,470803
FOND	0,377801	-0,221778
VUR	0,231059	0,247095
ZAN	-0,030756	0,493758

Рис. 7.15. Таблиця коефіцієнтів лінійних рівнянь регресії

На рис. 7.16 (*Factor Score*) наведено значення факторів для кожного спостереження. За цими значеннями визначають відношення підприємств до відповідних факторів. Додатне значення фактора відбиває додатне значення показників підприємства, а від'ємне – від'ємне значення.

Case	Factor Scores (FACTO)	
	Factor 1	Factor 2
1	-0,834959	0,11153
2	-0,472868	1,80944
3	1,961823	2,43020
4	0,020669	1,25895
5	-0,334352	0,29525
6	-0,631499	-0,28444
7	-0,253952	-0,68825
8	0,071234	-0,12960
9	-0,475809	-1,35280
10	-0,932788	0,71430
11	-0,716252	0,40739
12	-0,653817	0,03118
13	-0,741496	-0,04438
14	-0,149148	-0,92240
15	-0,154065	0,48862
16	1,431867	0,55427
17	0,374843	0,63573
18	3,573053	-1,55788
19	0,023406	-0,84547
20	0,634356	-0,01210
21	-0,239466	-1,41599
22	-0,470689	-0,50499
23	-0,304573	0,05750
24	-0,258156	-1,59319
25	-0,467362	0,55714

Рис. 7.16. Значення факторів для кожного підприємства

Величина позитивного фактора відповідає силі надання переваги даного фактора. Таким чином, проведення редукції даних дозволило зменшити інформаційний простір показників і виділити два значущих фактора.

## Рекомендована література

### Основна

Айвазян С. А. Прикладная статистика. Основы эконометрики : учебник для вузов в 2-х томах / С. А. Айвазян, С. В. Мхитарян. – М. : ЮНИТИ-ДАНА, 2001. – 656 с.

Дубров А. М. Многомерные статистические методы / А. М. Дубров, В. С. Мхитарян, Л. И. Трошин. – М. : Финансы и статистика, 1998. – 350 с.

Клебанова Т. С. Эконометрия / Т. С. Клебанова, Н. А. Дубровина, Е. В. Раевнева. – Х. : ИД "ИНЖЭК", 2003. – 128 с.

Орлов А. И. Прикладная статистика / А. И. Орлов. – М. : Экзамен, 2004. – 484 с.

Донченко В. С. Теорія ймовірностей та математична статистика : навч. посіб. / В. С. Донченко, М. В. Сидоров, М. М. Шарапов. – К. : ВЦ Академія, 2009. – 288 с.

Плюта В. Сравнительный многомерный анализ в экономических исследованиях / В. Плюта В. – М. : Статистика, 1980. – 151 с.

### Додаткова

Айвазян С. А. Прикладная статистика. Классификация и снижение размерности / С. А. Айвазян, В. М. Бухштабер, И. С. Енюков и др. – М. : Финансы и статистика, 1989. – 587 с.

Єгоршин О. О. Методи багатовимірного статистичного аналізу : навч. посіб. / О. О. Єгоршин, А. М. Зосімов, В. С. Пономаренко. – К. : ІЗМН, 1998. – 208 с.

Калинина В. Н. Введение в многомерный статистический анализ : учеб. пособ. / В. Н. Калинина, В. И. Соловьев. – М. : Изд-во ГУУ, 2003. – 66 с.

Шпинковський О. А. Теорія ймовірностей та математична . Конспект лекцій з базової підготовки за напрямом 0915 "Комп'ютерна інженерія" / О. А. Шпинковський, М. І. Шпинковська. – Одеса. : Наука і техніка, 2006. – 120 с.

### Ресурси мережі Інтернет

Сервер Державного комітету статистики України. – Режим доступу : [www.ukrstat.gov.ua](http://www.ukrstat.gov.ua).

Украинская инвестиционная газета. – Режим доступу : [www.investgazeta.net](http://www.investgazeta.net).

НАВЧАЛЬНЕ ВИДАННЯ

**Методичні рекомендації  
до виконання лабораторних робіт  
з навчальної дисципліни**

**"МАТЕМАТИЧНА СТАТИСТИКА  
ТА ОПРАЦЮВАННЯ ДАНИХ"**

**для студентів галузі знань  
0515 "Видавничо-поліграфічна справа"  
денної форми навчання**

Укладачі: **Клебанова Тамара Семенівна**  
**Гур'янова Лідія Семенівна**  
**Чаговець Любов Олексіївна**

Відповідальний за випуск *Клебанова Т. С.*

Редактор *Булгакова Г. К.*

Коректор *Маркова Т. А.*

План 2015 р. Поз. № 66.

Підп. до друку 27.07.2015 р. Формат 60x90 1/16. Папір офсетний. Друк цифровий.  
Ум. друк. арк. 4,0. Обл.-вид. арк. 5,0. Тираж 30 пр. Зам. № 114.

---

Видавець і виготівник – ХНЕУ ім. С. Кузнеця, 61166, м. Харків, просп. Леніна, 9-А

*Свідоцтво про внесення суб'єкта видавничої справи до Державного реєстру  
ДК № 4853 від 20.02.2015 р.*