

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ,  
МОЛОДІ ТА СПОРТУ УКРАЇНИ**

**ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ ЕКОНОМІЧНИЙ УНІВЕРСИТЕТ**

**Лабораторний практикум  
з навчальної дисципліни  
"СТАТИСТИЧНЕ МОДЕЛЮВАННЯ  
ТА ПРОГНОЗУВАННЯ"**

**для студентів напряму підготовки 6.030506**

**"Прикладна статистика"**

**денної форми навчання**

**Харків. Вид. ХНЕУ, 2013**

Затверджено на засіданні кафедри статистики та економічного прогнозування.

Протокол № 1 від 29.08.2012 р.

**Укладачі:** Раєвнєва О. В.

Стрижиченко К. А.

Чанкіна І. В.

Гольцяєва Л. А.

Л12            Лабораторний практикум з навчальної дисципліни "Статистичне моделювання та прогнозування" для студентів напряму підготовки 6.030506 "Прикладна статистика" денної форми навчання / укл. О. В. Раєвнєва, К. А. Стрижиченко, І. В. Чанкіна та ін. – Х. : Вид. ХНЕУ, 2013. – 60 с. (Укр. мов.)

Розроблено лабораторні роботи, метою яких є закріплення й поглиблення знань теоретичного та практичного матеріалу, набуття навичок аналізу різних типів даних за допомогою пакета STATISTICA.

Рекомендовано для студентів економічних спеціальностей.

## Вступ

Лабораторні роботи призначені для закріплення студентами теоретичних та практичних знань з дисципліни "Статистичне моделювання та прогнозування", набуття навичок роботи з пакетами прикладних програм Statistica 8.0 та MS Excel, які дозволяють автоматизувати розрахунок економіко-математичних методів і моделей аналізу та прогнозування розвитку соціально-економічних процесів чи явищ, що досліджуються на різних рівнях деталізації.

Важливою характеристикою системи STATISTICA 8.0 є те, що вона є повноцінною прикладною програмою Windows та містить комплекс статистичних методів, що підтримують рішення різних статистичних та економіко-математичних задач. При розробці лабораторних робіт передбачається, що студент знайомий з основними принципами й прийомами роботи в середовищі STATISTICA 8.0.

MS Excel – це універсальний та досить потужний табличний процесор, що включає в себе надбудову "Пакет аналізу" і бібліотеку з комплексом статистичних та математичних функцій. Крім того, він є найдоступнішим сучасним програмним продуктом, що дозволяє проводити статистичні розрахунки. Excel надає можливість відчувати всі деталі і тонкощі досліджуваних методів, що природним чином підвищує рівень засвоюваності навчального матеріалу.

Кожна лабораторна робота розглянута на прикладі вирішення конкретної задачі й забезпечена докладними коментарями та елементами графічного аналізу. Дисципліна складається з чотирьох змістовних модулів: методологічні основи статистичного моделювання і прогнозування, моделювання взаємозв'язків економічних процесів, моделювання часових процесів та моделювання і прогнозування багатовимірних процесів. Лабораторний практикум присвячений розробці тем першого модулю та другого модулю. Кожна робота містить мету й завдання та методичні рекомендації щодо їх виконання.

Оцінка за виконання роботи ставиться за результатами виконання й захисту лабораторної роботи. Певна увага приділяється знанням студентів теоретичного матеріалу, правильності висновків і повноті економічної інтерпретації отриманих результатів.

Компетентності, які повинен здобути студент, що вивчає навчальну дисципліну, наведені в табл. 1.

**Компетентності, які повинен здобути студент,  
що вивчає навчальну дисципліну**

Назва компетентностей	Зміст компетентностей	Вміння студента щодо даної компетентності
1. Організаційно-методологічна	1.1. Здатність впроваджувати методологію статистичного моделювання та прогнозування при оцінці статистичних явищ та процесів	1.1.1. Володіти методами аналізу рядів розподілу. 1.1.2. Володіти методами експертних оцінок. 1.1.3. Володіти методами побудови багатофакторних регресійних моделей
2. Обліково-статистична	2.1. Здатність до здійснення статистичних спостережень	2.1.1. Збирати та обробляти інформацію відповідно до наукових методів її первинної обробки. 2.1.2. Володіти сучасними інформаційними технологіями збору, обробки та поширення даних
3. Аналітична	3.1. Здатність щодо проведення статистичного аналізу	3.1.1. Аналізувати інформацію всіх видів, виконувати необхідні розрахунки й оформляти результати. 3.1.2. Визначати чинники, що впливають на зміну показників. 3.1.3. На основі аналізу інформації виробляти пропозиції щодо покращення ситуації
4. Модельно-прогнозна	4.1. Здатність до побудови моделей	4.1.1. Визначати залежні та незалежні змінні при побудові моделі. 4.1.2. Досліджувати адекватність побудованих моделей. 4.1.3. Будувати різні класи моделей залежно від потреби суспільства
	4.2. Здатність до прогнозування соціально-економічних явищ та процесів	4.2.1. Прогнозувати найімовірніший перебіг процесів на перспективу та аналізувати й оцінювати можливі наслідки зміни явища
5. Інформаційна	5.1. Здатність використовувати сучасні пакети прикладних програм для моделювання та прогнозування соціально-економічних явищ та процесів та візуалізації отриманих результатів	5.1.1. Використовувати Internet для пошуку статистичної інформації щодо економічного розвитку країн. 5.1.2. Використовувати пакет Statistica для побудов статистичних моделей, графіків і отримання прогнозів

## **Лабораторна робота № 1. Знайомство з пакетом Statistica 8.0. Дослідження статистичних характеристик варіаційного ряду**

**Мета** – закріплення теоретичного та практичного матеріалу дослідження статистичних характеристик варіаційного ряду в пакеті Statistica 8.0.

**Завдання.** Необхідно провести аналіз варіаційного ряду за допомогою описових статистик.

### **Методичні рекомендації**

Роботу потрібно починати з запуску пакету Statistica, який здійснюється аналогічно запуску інших прикладних програм – через меню ПУСК або використовуючи ярлик.

Вікно системи Statistica складається з таких основних елементів: рядок заголовка, рядок меню, панель інструментів, робоча область і рядок стану.

**Рядок заголовка** містить піктограму, назву програми Statistica і три кнопки керування розмірами основного вікна: кнопка мінімізації розмірів вікна; кнопка відновлення розмірів вікна; кнопка закриття вікна.

**Рядок меню** займає другий рядок основного вікна модуля і при наявності відкритого файлу з даними в робочій області містить так зване випадаюче меню: File – Файл, Edit – Правка, View – Вид, Insert – Вставка, Format – Формат, Statistics – Статистики, Data mining – процедури Data mining, Graphs – Графіки, Tools – Інструменти, Data – Дані, Window – Вікно, Help – Довідка.

**Панель інструментів** містить кнопки для швидкого доступу до найчастіше всього використовуваних команд меню.

**Робоча область**, в якій виводяться різні документи займає більшу частину основного вікна:

1. Електронна таблиця із вихідними даними. При першому відкритті Statistica в робочій області автоматично відкривається новий файл розмірності 10×10 з назвою Spreadsheet.sta.

2. Стартове вікно модуля статистичного аналізу, що використовується.

3. Електронні таблиці з результатами аналізу.

4. Інструменти графічного аналізу.

5. Вікно автозвіту.

**Рядок стану** розташований у найнижчій частині вікна системи Statistica. Залежно від стану, в якому знаходиться система, рядок стану містить кнопку швидкого доступу до основних статистичних модулів і пунктів меню, а також відображає різну інформацію і дозволяє управляти функціонуванням системи.

При обробці даних і побудові графіків рядок стану містить шкалу прогресу, яка відображає ступінь завершеності процесу обробки даних та таймер, що відображає час, що минув з початку обробки.

В якості прикладу, розглянемо значення капіталу 40 банків України станом на 01.09.2012 р. у млн грн (джерело інформації – офіційний сайт Асоціації українських банків).

Для цього в пункті меню **File** системи Statistica вибираємо вкладку **New**. В результаті відкриється вікно створення нового файлу (рис. 1).

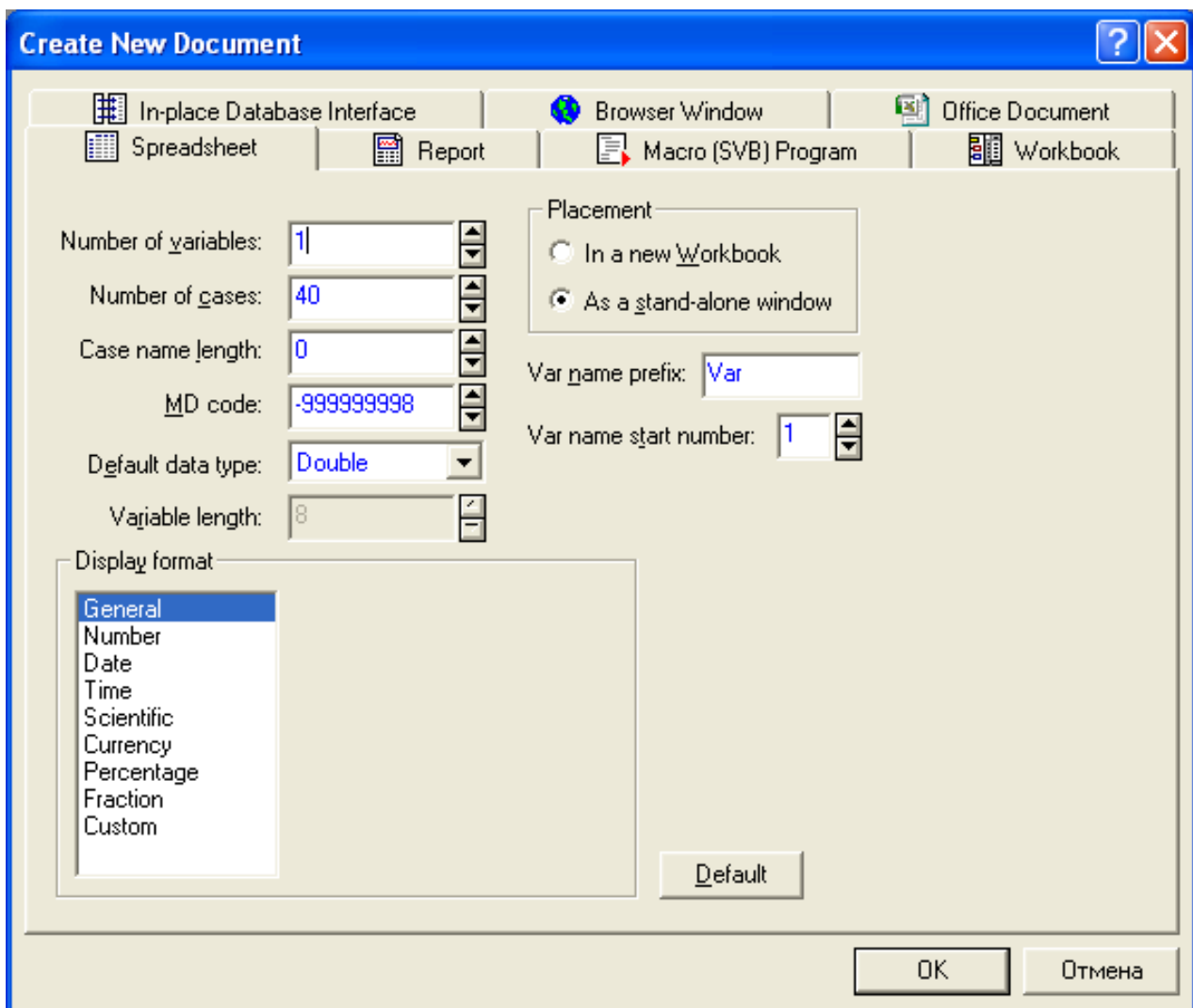


Рис.1. Стартове вікно створення нового файлу

Необхідно створити електронну таблицю з однією змінною (1 стовпець) і 40 спостереженнями (рядків) і ввести в неї числові значення. Після цього електронна таблиця виглядатиме наступним чином (рис. 2).

	1
	Капітал
ОЩАДБАНК	17816,26
ПРИВАТБАНК	17750,61
УКРСОЦБАНК	9233,952
ПРОМІНВЕСТБАНК	5366,856
АЛЬФА-БАНК	5236,242
ВТБ БАНК	4726,628
ПУМБ	4458,588
ОТП БАНК	4053,314
РОДОВІД БАНК	3619,326
СБЕРБАНК РОСІЇ	3178,131
ДЕЛЬТА БАНК	2947,366
БРОКБІЗНЕСБАНК	2938,624
ФІНАНСИ ТА КРЕДИТ	2864,028
КРЕДИТПРОМБАНК	2385,225
ФІНАНСОВА ІНІЦІАТИВА	1910,922
ІНГ БАНК Україна	1865,866
ПІВДЕННИЙ	1592,481
БТА БАНК	1520,666
ПРАВЕКС-БАНК	1252,849
ІМЕКСБАНК	1230,859
АСТРА БАНК	1113,366
КРЕДІ АГРІКОЛЬ БАНК	1016,28
УНІВЕРСАЛ БАНК	1001,577
ПІРЕУС БАНК МКБ	956,701
УНІКРЕДИТ БАНК	943,772
БАНК КРЕДИТ-ДНІПРО	885,865
МЕГАБАНК	833,342
КИЇВ	786,192
КИЇВСЬКА РУСЬ	729,389
ДІВІ БАНК	714,956
КРЕДОБАНК	700,689
ІНДУСТРІАЛБАНК	700,483
КІБ Креді Агріколь	645,15
КЛІРИНГОВИЙ ДІМ	633,611
УКРІНБАНК	629,101
СІТІБАНК УКРАЇНА	623,186
ТЕРРА БАНК	622,778
ЗЛАТОБАНК	606,734
МАРФІН БАНК	588,358
БМ Банк	573,934

Рис.2. Початкові дані

У пункті меню **File** виберіть команду **Save** і збережіть файл у створену папку під будь-яким ім'ям (наприклад, Bank.sta).

Структуру електронної таблиці можливо змінювати за допомогою меню **Data**, вибираючи різноманітні основні функції спостережень і змінних:

- додавати та видаляти (**Add** та **Delete**);
- вирізати та вставляти (**Cut** та **Paste**);
- сортувати та стандартизувати (**Sort** та **Standartize**);
- транспонувати та переносити (**Transpose** та **Move**) та ін.

Розрахунок основних числових характеристик досліджуємого варіаційного ряду можливо провести за допомогою описових статистик. У меню **Statistics** виберіть пункт **Basic statistics / Tables**. У вікні (рис. 3) виберіть пункт **Descriptive statistics**, в результаті чого відкриється вікно розрахунку комплексу описових статистик (рис. 3).

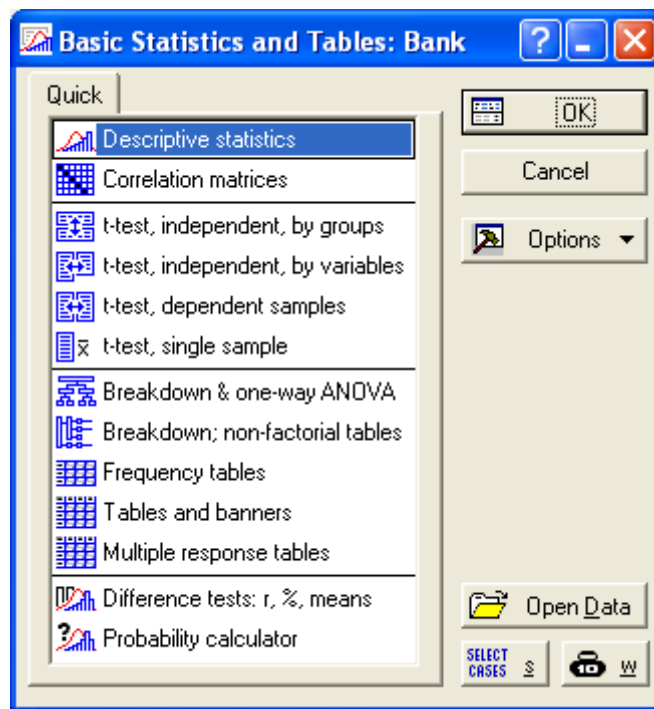


Рис. 3. Діалогове вікно розрахунку описових статистик

Далі необхідно перейти на вкладку **Advanced** та вибрати показники, які необхідно розрахувати, встановивши прапорці біля них, як показано на рисунку 4.

Таким чином, для аналізу було обрано такі показники: **Mean** – середнє арифметичне; **Median** – медіана; **Mode** – мода; **Standart Deviation** – середнє квадратичне відхилення; **Variance** – дисперсія; **Skewness** – асиметрія; **Kurtosis** – ексцес; **Minimum & Maximum** –



мінімальне та максимальне значення; **Range** – розмах варіації; **Coefficient of variation** – коефіцієнт варіації.

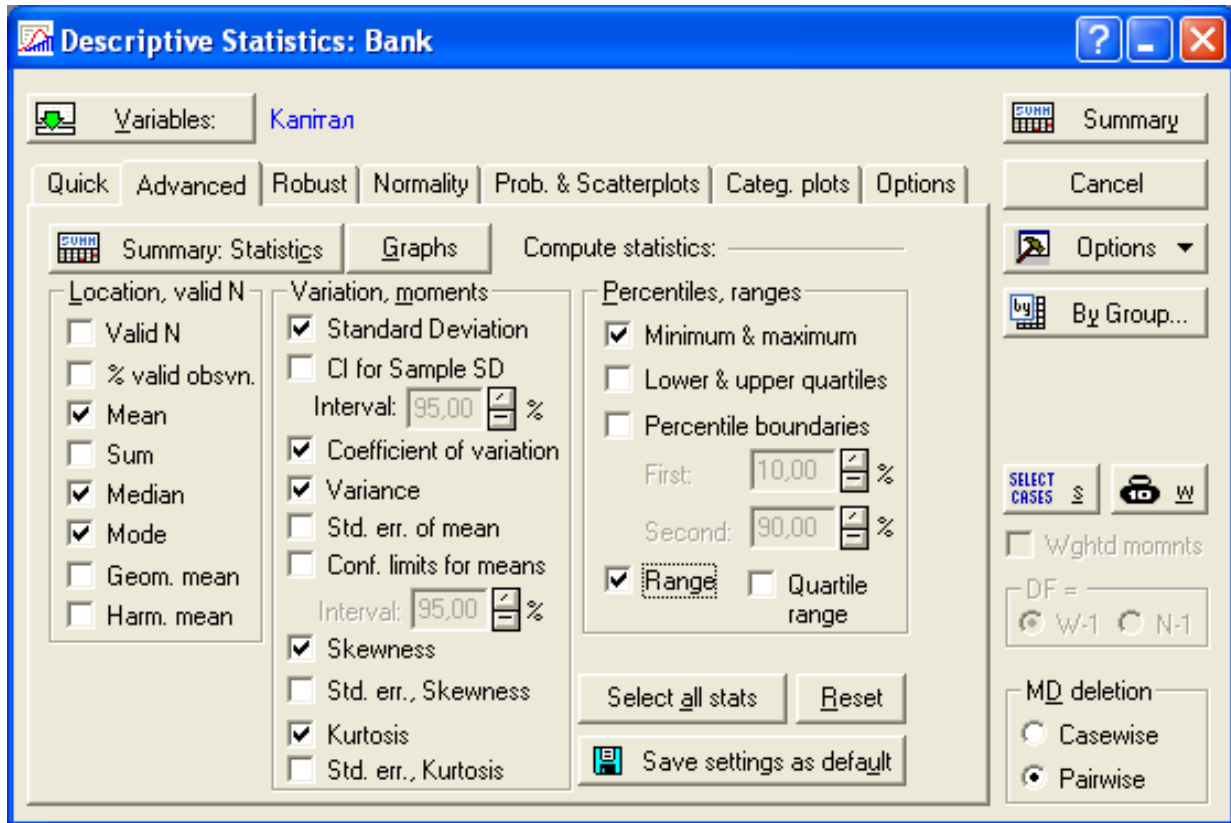


Рис. 4. Вікно вибору описових статистик

Система здійснить розрахунок зазначених показників і представить результати у вигляді таблиці (рис. 5) після натискання кнопки



Variable	Mean	Median	Mode	Frequency of Mode	Minimum	Maximum	Range	Variance	Std.Dev.	Coef.Var.	Skewness	Kurtosis
Kaniran	2781,356	1172,113	Multiple	1	573,9340	17816,26	17242,32	15510760	3938,370	141,5989	3,043798	9,622438

Рис.5. Вікно результатів розрахунку описових статистик

У системі Statistica досить потужний графічний інструментарій: гістограми, точкові і лінійні графіки, двовимірні та тривимірні і т. д. Розглянемо деякі види статистичних 2D гістограм:

**2D Histogramms** є графічними уявленнями розподілу частот вибраних змінних. Для кожного інтервалу (класу) малюється стовпець,

висота якого пропорційна частоті класу. Гістограма наочно показує, які значення або діапазони значень досліджуваної змінної є найчастішими, наскільки сильно вони розрізняються, як сконцентровано більшість спостережень навколо середнього, є розподіл симетричним чи ні, чи має воно моду або кілька мод. Розрізняють декілька видів гістограм.

**2D Histogramms Regular** (прості) становлять стовбчасту діаграму розподілу частот для обраної змінної (якщо вибрано більше однієї змінної, то для кожної з них буде побудований окремий графік).

**2D Histogramms Multiple** (складові) зображують розподіл частот для декількох змінних на одному графіку. Частоти для всіх змінних відкладаються по лівій осі  $Y$ . Значення всіх досліджуваних змінних відкладаються по одній осі  $X$ , що полегшує порівняння аналізованих змінних.

**2D Histogramms Double-Y** (з подвійною віссю  $Y$ ). Гістограму з подвійною віссю  $Y$  можна вважати комбінацією двох по-різному масштабованих складових гістограм. Для цієї гістограми можна вибрати дві різні групи змінних. Для кожної з обраних змінних буде зображено розподіл частот, але частоти змінних з першого списку, званого Left  $Y$  (ліва вісь  $Y$ ), відкладаються по лівій осі  $Y$ , а частоти змінних з другого списку, так званого Right  $Y$  (права вісь  $Y$ ), відкладатимуться по правій осі  $Y$ . Цей графік корисний для візуального порівняння розподілів змінних з різними частотами.

**2D Histogramms Hanging Bars** (висячі стовпці). Гістограма висячих стовпців є "наочним критерієм перевірки на нормальність розподілу", який допомагає визначити області розподілу, де виникають розбіжності між спостережуваними та очікуваними нормальними частотами. Вважається, що стовпчики, що представляють частоти, які спостерігаються, для послідовних діапазонів значень, "підвішуються" до найбільш придатної нормальної кривої. Якщо досліджуваний розподіл добре наближається до нормальної кривої, то нижні ребра всіх стовпців повинні утворити пряму горизонтальну лінію.

Наприклад, побудуємо гістограму розподілу банків за капіталом. Для цього в меню Graphs необхідно вибрати Histograms (Гістограми), після чого потрібно вибрати змінну "Капітал". Результат аналізу зображений на рис. 6.

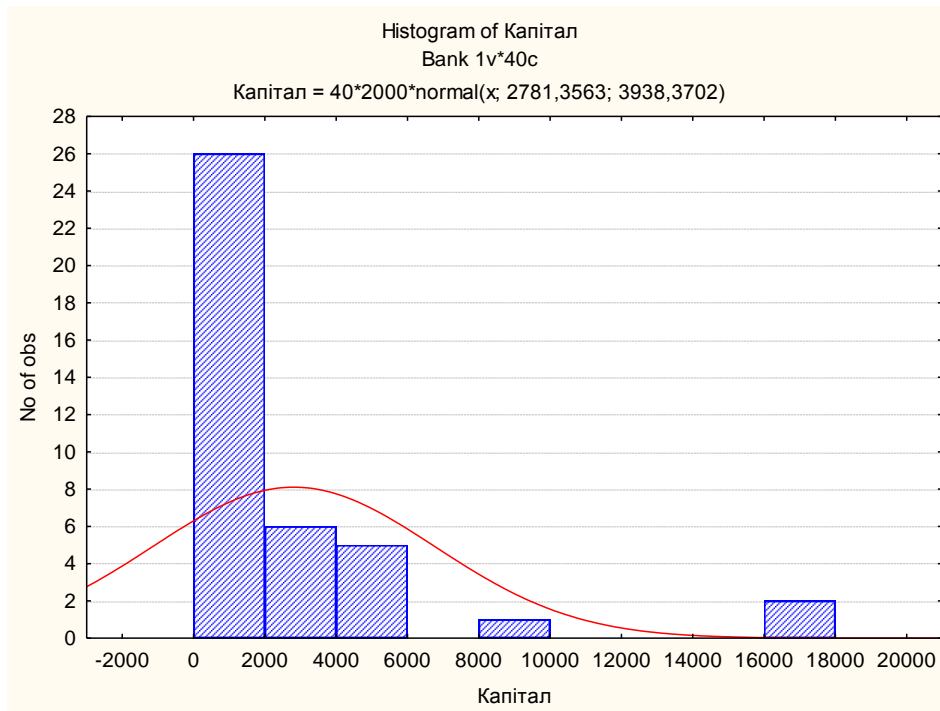


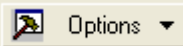
Рис. 6. Гістограма розподілу банків за капіталом

Результати всіх процедур у системі Statistica: створення електронної таблиці з даними статистичного аналізу, побудови графіків та іншого можуть бути перенесені в Автозвіт. Для створення файлу звіту необхідно на панелі інструментів знайти пункт **Add to Report**, вибрати **New Report** (Новий звіт). У результаті відкриється вікно з файлом звіту, в який автоматично буде внесено електронну таблицю з вихідними даними (рис. 7).

Bank	Капітал
ОЩАДБАНК	17816,26
ПРИВАТБАНК	17750,61
УКРСОЦБАНК	9233,952
ПРОМІВЕСТБАНК	5366,856
АЛЬФА-БАНК	5236,242
ВТБ БАНК	4726,628
ПУМБ	4458,588
ОТП БАНК	4053,314
РОДОВІД БАНК	3619,326
СБЕРБАНК РОСІІ	3178,131
ДЕЛЬТА БАНК	2947,366
БРОКБІЗНЕСБАНК	2938,624
ФІНАНСИ ТА КРЕДИТ	2864,028
КРЕДИТПРОМБАНК	2365,225
ФІНАНСОВА ІНІЦІАТИВА	1910,922
ІНГ БАНК Україна	1865,866
ПІВДЕННИЙ	1592,481
БТА БАНК	1520,666
ПРАВЕКС-БАНК	1252,849
ІМЕКСБАНК	1230,859
АСТРА БАНК	1113,366
КРЕДІ АГРІКОЛЬ БАНК	1016,28
УНІВЕРСАЛ БАНК	1001,577
ПІРЕУС БАНК МКБ	956,701
УНКРЕДИТ БАНК	943,772
БАНК КРЕДИТ-ДНІПРО	885,865

Рис. 7. Діалогове вікно автозвіту

Файл звіту є звичайним текстовим файлом, тому він може редагуватися тими ж способами, що й інші текстові файли. Після створення необхідно зберегти файл звіту під будь-якою зручною назвою.

Для автоматичного додавання результатів роботи у файл звіту необхідно виконувати наступну послідовність дій. При відкритті початкового вікна статистичного аналізу або вікна побудови графіків необхідно знайти і натиснути на цьому вікні кнопку  і вибрати Output (Вивід). У результаті буде відкрито вікно опцій виведення результатів (рис.8).

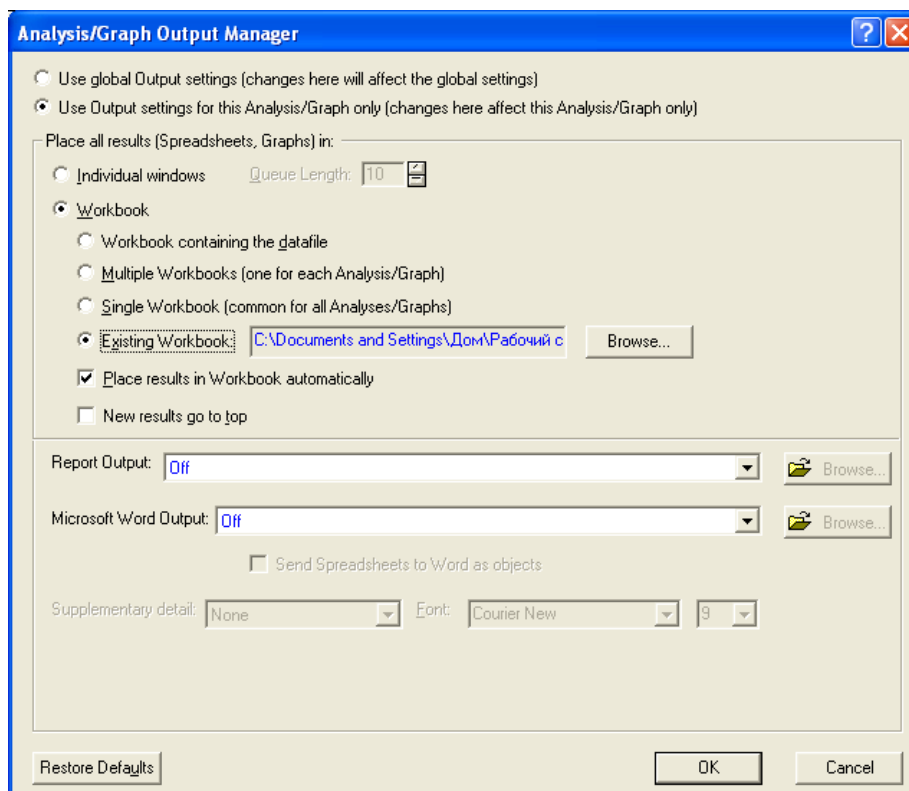


Рис. 8. Вікно встановлення параметрів внесення результатів до автозвіту

Встановіть опції виводу так, як показано на рисунку 8. Після встановлення вказаних опцій всі результати, автоматично додаватимуться у файл звіту.

**Завдання для самостійної роботи.** Необхідно знайти просторові одновимірні дані (щонайменше 30 спостережень та виконати статистичний аналіз за допомогою описових статистик та графічних процедур. Зробити висновки та навести економічну інтерпретацію результатів.

## Лабораторна робота № 2. Стратифікація вибірових даних

**Мета** – опанування студентами навичок дослідження варіаційного ряду. Навчитися стратифікувати статистичні дані за допомогою інструментів пакету Statistica.

**Завдання** – необхідно сформулювати стратифіковану вибірку та надати економічну інтерпретацію отриманим розрахункам.

### Методичні рекомендації

Під стратифікацією розуміється визначення верств (страт) у багат шаровому явищі, тобто залежностей особливого виду.

Стратифікація статистичних даних у пакеті Statistica проводиться за допомогою модуля **Data / Create a Subset / Random Sampling** (Дані/ Підмножина/Випадковий вибір). Цей модуль має три вкладки **Simple Sampling** (Простий вибір), **Stratified Sample** (Стратифікований вибір) та **Options** (Опції) (рис.9).

За допомогою вкладки Simple Sampling можливо виконувати:

**Case selection condition** (Використовувати умови вибору спостережень). Користувач самостійно вибирає змінні і визначає умови вибору спостережень у файлі даних.

**Simple random sampling** (Проста випадкова вибірка). При виборі цього правила, дані вибиратимуться випадковим чином.

**Systematic random sampling** (Систематичний випадковий вибір). Використовуючи даний метод, підмножина буде складатиметься за допомогою систематичного випадкового вибору.

Наприклад, якщо ввести число 5 у полі  $K =$ , то серед перших п'яти спостережень буде випадковим чином обрано одне, а потім система вибиратиме кожне п'яте спостереження із вихідних даних.

**Split node random sampling** (Розділений випадковий вибір). При виборі даного методу, всі спостереження будуть випадковим чином розділені на два файли даних. Необхідно вказати відсоток спостережень або приблизне число спостережень.

Розглянемо процедуру формування стратифікованої вибірки за допомогою інструментів пакету Statistica на прикладі рейтингу 100 страхових компаній України, які розподілено за змінною рівень стабільності, що приймає 2 значення: "стабільна/нестабільна".

При цьому 60 % компаній є стабільними, 40 % – нестабільними. Дані також характеризуються ще п'ятьма змінними, такими як: премії, питома вага в портфелі, виплати, рівень виплат, рівень перестраховування (рис. 9).

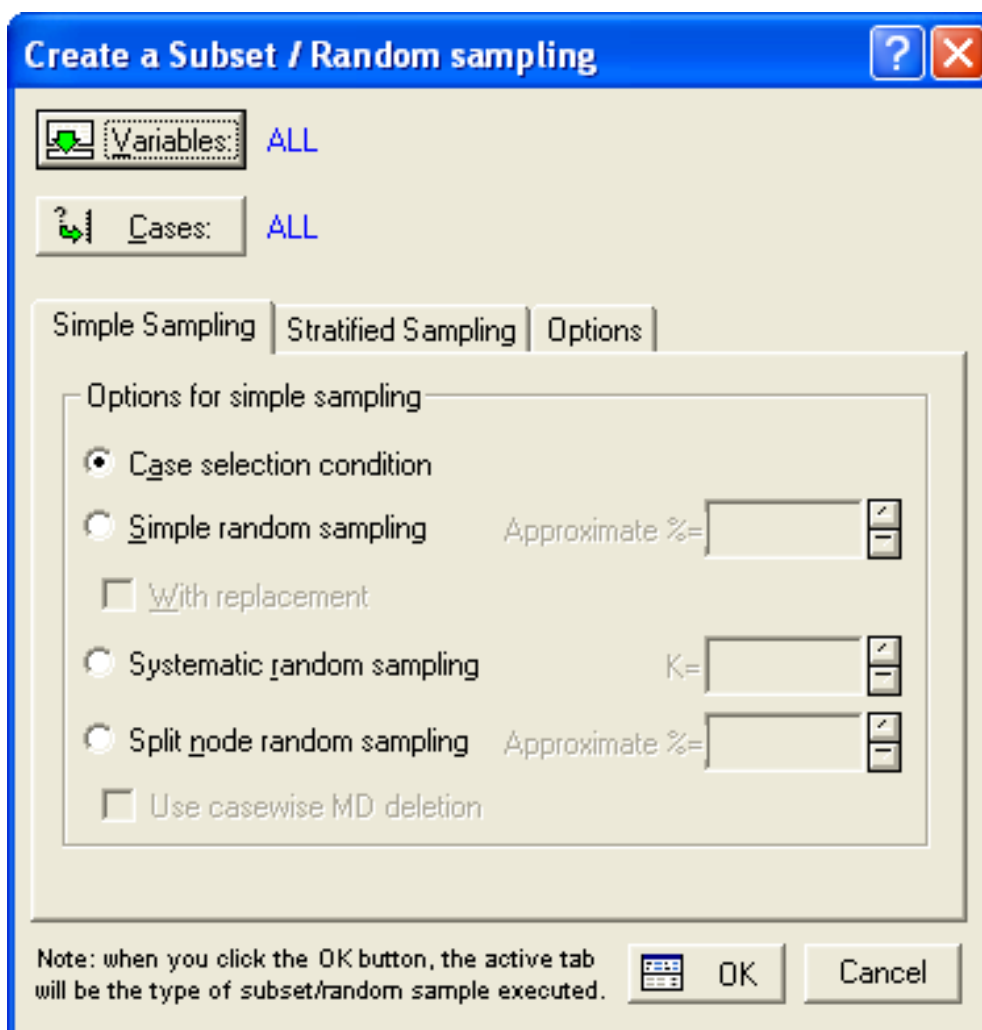


Рис. 9. Діалогове вікно модуля Підмножина / Випадковий вибір

Необхідно створити пропорційну вибірку з однаковим розміром для обох страт (по 15 спостережень для кожної страти), яка дозволить більш точно представити рідкісні події нестабільних страхових компаній та отримати більш точний прогноз про нестабільні компанії (рис. 10).

Для того, щоб побачити нерівномірність даних побудуємо гістограму початкових даних (рис. 11), за допомогою вкладки Graphs/Histograms. На гістограмі видно, що дані розподілені нерівномірно, тобто припущення про те, що всі страхові компанії є стабільними і правильними у 60 % випадків.

Це досить неточний прогноз. Для того, щоб отримати більш якісний прогноз дані, які було стратифіковано випадковою вибіркою, необхідно вирівняти. Виберемо вкладку Data/Create a Subset/Random sampling.

	1	2	3	4	5	6	7
	Страхові компанії	Премії, тис. грн	Питома вага в портфелі, %	Виплати, тис. грн	Рівень виплат, %	Рівень перестраховання, %	Рівень стабільності
1	ALLIANZ УКРАИНА	327,2	0,45	146,4	44,74	-22	нестабільна
2	PZU УКРАИНА	3663,7	2,29	573,3	15,65	10,62	стабільна
3	QBE УКРАИНА	695,7	1,89	10	1,44	7,59	нестабільна
4	АКСОР	6059,9	58,32	2,4	0,04	79,33	стабільна
5	АЛЬЯНС УКРАИНА	2277,5	15,79	0	0	48,56	стабільна
6	АКТИВ-СТРАХОВА	32,5	0,01	0	0	52,57	стабільна
7	АРМА	65810,9	32,14	412,7	0,63	0,35	стабільна
8	АЛЬФА СТРАХОВА	1560,5	1,68	505,6	32,4	1,92	стабільна
9	АХА	823	4,63	18,8	2,28	17,76	стабільна
10	АЛЬФА-ГАРАНТ	320,9	0,09	1536	478,65	7,53	стабільна
11	АЛЬЯНС	434	0,25	8	1,84	1,15	нестабільна
12	АРСЕНАЛ СТРАХО	58529	30,51	28,6	0,05	18,36	стабільна
13	АСКА	3918,7	1,01	342,6	8,74	1,21	стабільна
14	АСКО-ДОНБАСС СЕ	5982,3	14,29	1517,5	25,37	19,07	стабільна
15	АСКО-МЕДСЕРВИС	8	0,01	0	0	0	нестабільна
16	АХА СТРАХОВАНИЕ	2997	0,78	898	29,96	5,57	стабільна
17	БРОКБИЗНЕС	275,2	0,43	130	47,24	0	нестабільна
18	БУСИН	1976,8	6,07	0	0	28,13	стабільна
19	ВиДи - СТРАХОВАН	10,1	0,02	0	0	16,98	стабільна
20	ВОЕННО-СТРАХОВ.	458,6	1,39	50,6	11,03	1,98	нестабільна
21	ВУСО	2908	2,72	0,5	0,02	38,07	стабільна
22	ГАРАНТ-АВТО	598,1	0,62	200,5	33,52	93,83	нестабільна
23	ГАРАНТИЯ СО	373,8	1,58	34,9	9,34	0,83	нестабільна
24	ГОРОДСКАЯ СТРАХ	1221,3	0,6	0	0	57,25	стабільна
25	ГАРАНТИЯ СОСДО	2125,5	7,54	157,8	7,42	23,54	стабільна
26	ДЖЕНЕРАЛИ ГАРА	199,2	0,31	237,1	119,03	12,46	стабільна
27	ГАРАНТ-СИСТЕМА	11,7	0,01	0	0	44,23	стабільна
28	ГЛОБУС	485,8	1,05	0,5	0,1	0,35	нестабільна
29	ГРАВЕ УКРАИНА	489,8	2,9	3,4	0,69	13,37	нестабільна
30	ДИНАСТИЯ	2846,2	6,21	63,9	1,75	39,12	стабільна
31	ДОБРОБУТ	1291,2	1,39	26,6	2,06	0	стабільна
32	ДОБРОБУТ ТА ЗАХ	1338,2	1,44	3	0,22	34,16	стабільна
33	ДОВЕРИЕ И ГАРАН	4075,1	42,09	17	0,42	0	стабільна
34	ЕВРОПЕЙСКИЙ МУ	3005,4	4,71	187,1	6,23	32,19	стабільна
35	ЕВРОПЕЙСКИЙ СТИ	625	1,43	201,2	32,19	0,69	нестабільна
36	ЕВРОПЕЙСКИЙ СТИ	3516,2	8,06	448,9	12,77	33,71	стабільна
37	ЕВРОПЕЙСКОЕ ТУ	2655	14,44	422	15,89	18,83	стабільна
38	ИЛЬИЧЕВСКОЕ	1238,1	1,94	65,3	5,27	21,06	нестабільна
39	ИНГО УКРАИНА	2708,3	1,06	1076,3	39,74	3,01	стабільна
40	ИНДИГО	1286,2	3,66	17,1	1,33	10,36	стабільна
41	ИНТЕР	1219,4	11,73	0	0	52,06	стабільна
42	ПЗУ УКРАИНА	509,8	3,81	0	0	47,4	стабільна
43	ИНТЕР ЭКСПРЕСС	3	0,03	0	0	57,25	нестабільна

Рис.10. Фрагмент таблиці з початковими даними

Далі з'явиться діалогове вікно для створення вибірок (рис. 12).

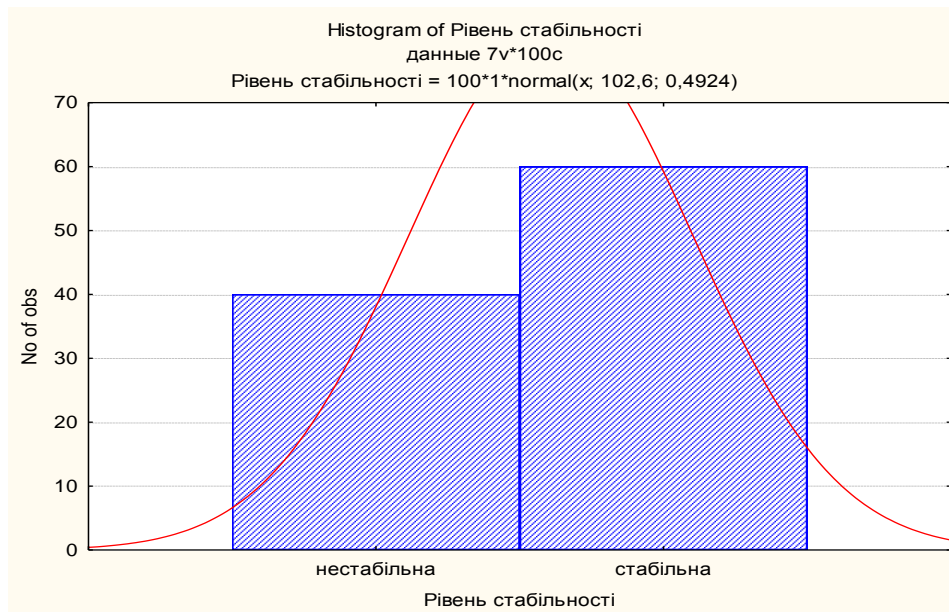


Рис. 11. Гістограма початкових даних

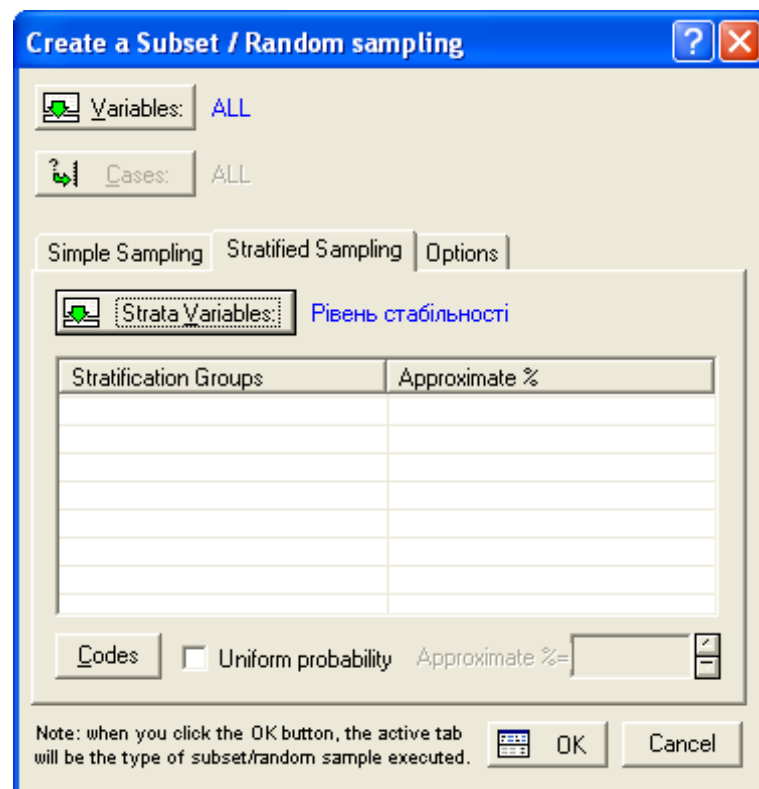


Рис. 12. Діалогове вікно випадкового формування підмножин

Вибираємо вкладку Stratified Sample, в якій є кнопка Strata Variables, що визначає змінну страти. Змінна страти – змінна, для якої вибираються пропорції вибірок, для нашого приклада це рівень стабільності.



Потім натискаємо кнопку Codes, яка дозволяє вибрати категорії цієї змінної страти. Оскільки, за умовою завдання потрібно використовувати рівень стабільності і стабільний і нестабільний, вибираємо кнопку all (рис. 13).

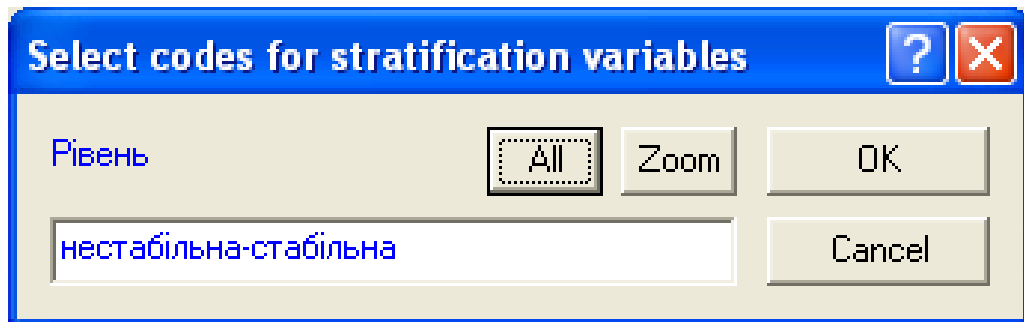


Рис. 13. Вікно вибору категорій змінної страти

Масштаб зображення Zoom дозволяє вибрати категорії і побачити сітки страти (рис. 14).

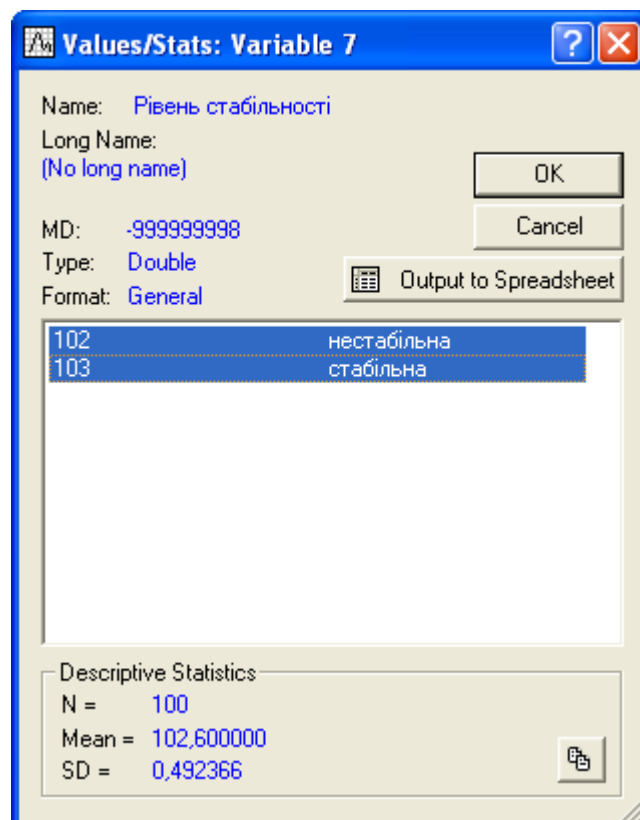


Рис. 14. Вікно сітки страт

Для того, щоб вибрати розмір вибірки необхідно внести зміни (рис. 15) у вкладці options.

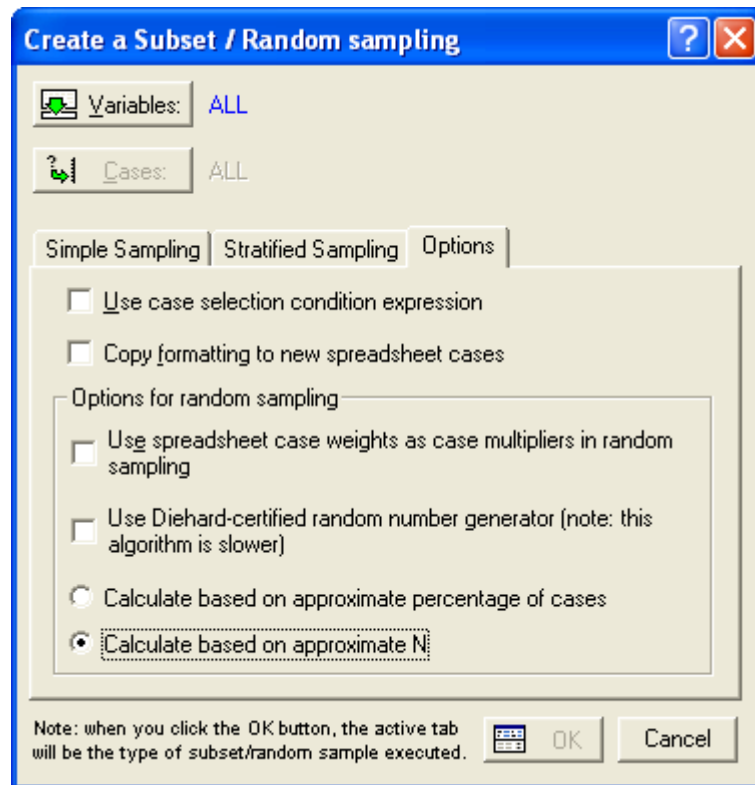


Рис. 15. Вікно вибору опцій

Потім змінити число спостережень приблизно до 15 на кожну категорію змінної страти (рис. 16).

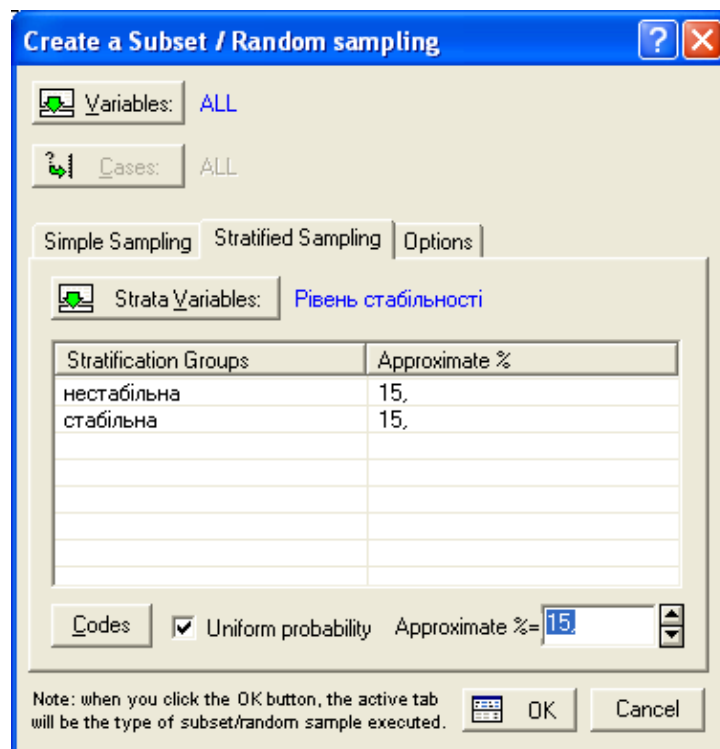
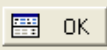


Рис. 16. Вікно вибору кількості спостережень у кожній страті

Далі натискаємо кнопку  та отримуємо таблицю зі стратифікованою вибіркою (рис. 17).

	1	2	3	4	5	6	7
	Страхові компанії	Премії, тис. грн	Питома вага в портфелі, %	Виплати, тис. грн	Рівень виплат, %	Рівень перестраховування, %	Рівень стабільності
1	АЛЬФА-ГАРАНТ	320,9	0,09	1536	478,65	7,53	стабільна
2	АЛЬЯНС	434	0,25	8	1,84	1,15	нестабільна
3	АСКО-МЕДСЕРВИС	8	0,01	0	0	0	нестабільна
4	Види - СТРАХОВАН	10,1	0,02	0	0	16,98	стабільна
5	ГАРАНТИЯ СО	373,8	1,58	34,9	9,34	0,83	нестабільна
6	ГАРАНТИЯ СОсДО	2125,5	7,54	157,8	7,42	23,54	стабільна
7	ЕВРОПЕЙСКИЙ МИ	3005,4	4,71	187,1	6,23	32,19	стабільна
8	ЕВРОПЕЙСКИЙ СТИ	3516,2	8,06	448,9	12,77	33,71	стабільна
9	КНЯЖА	1634,2	1,59	48,6	2,97	5,78	стабільна
10	КРЕДО	3558,5	19,87	3499,5	98,34	35,74	стабільна
11	КРЕМЕНЬ	8226,9	2,86	3,8	0,05	33,71	стабільна
12	КРОНА	902,7	2,57	9,6	1,06	60,69	стабільна
13	ЛЕММА	378,2	0,17	0	0	0	нестабільна
14	ЛИДЕР РЕ	2137,1	9,02	163,3	7,64	74,44	стабільна
15	МИР СК	25,6	0,09	0	0	24,97	стабільна
16	НАСТА	1311,8	3,47	22,5	1,72	44,99	стабільна
17	НЕФТЕГАЗСТРАХ	893,6	3,47	1,3	0,15	56,19	стабільна
18	ОМЕГА	792,9	2,81	140,8	17,76	4,25	нестабільна
19	ПРОМЫШЛЕННО С	935,1	5,17	0	0	18,07	стабільна
20	ПРОСТО-СТРАХОВ	1665	2,15	122	7,33	1,02	нестабільна
21	РАРИТЕТ	1,92	0	0	0	0	нестабільна
22	САЛАМАНДРА-УКР	136,3	0,07	0	0	0	нестабільна
23	ФКС	528,5	1,15	14,2	2,69	3,52	нестабільна
24	УКРАИНСКАЯ ОХРА	1549	4,35	8,3	0,54	9,22	нестабільна
25	УСК	93	0,02	0	0	59,12	нестабільна
26	УКРАИНСКИЙ СТРА	1099,5	3,38	9	0,82	68,21	нестабільна
27	УОСК	1179,3	2,7	111,2	9,43	94,81	нестабільна
28	ХАРЬКОВСКАЯ МУИ	35,1	0,04	0	0	0,65	нестабільна
29	УДИ СТРАХОВАНИЕ	2094	2,17	760	36,29	34,01	стабільна
30	ЭКСПРЕСС СТРАХ	636,7	1,08	0	0	0	нестабільна
31	ЭНЕРГОПОЛИС	1302,4	0,68	0	0	15,64	стабільна

Рис.17. Таблица стратифікованих даних

За отриманими даними на рис. 17 створюємо гістограму стратифікованої вибірки (рис. 18).

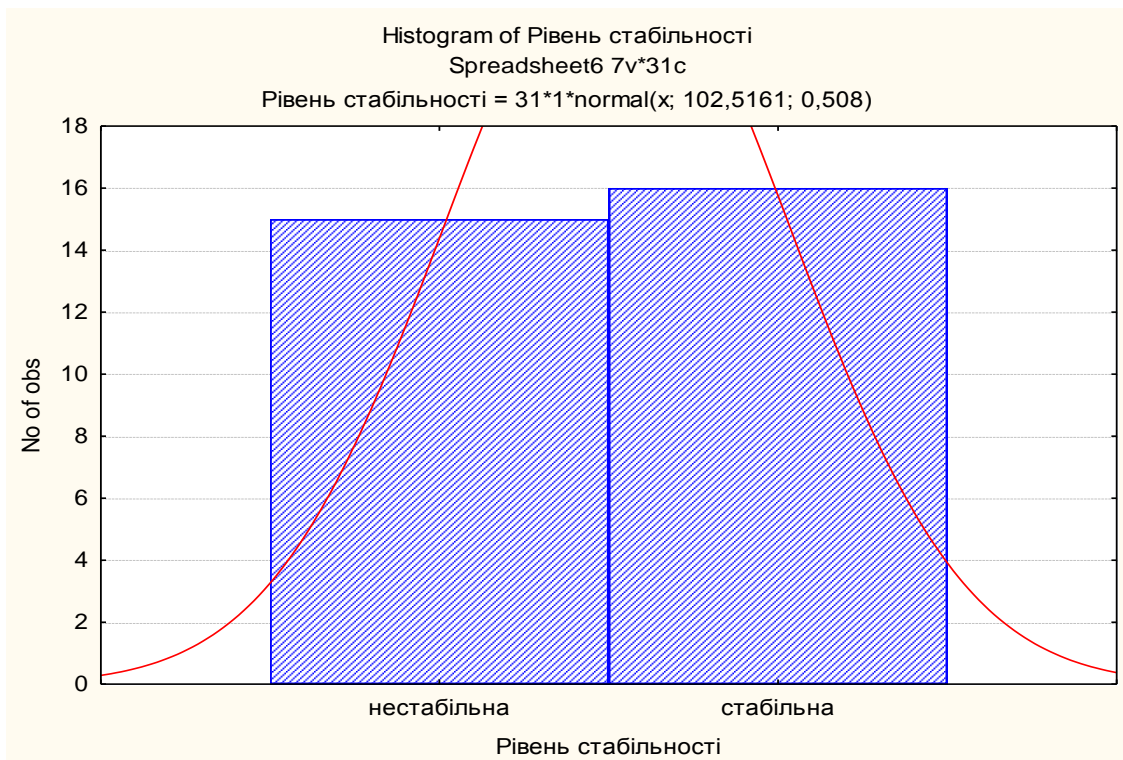


Рис. 18. Гістограма стратифікованої вибірки

З гістограми видно, що пропорція нестабільних та стабільних страхових компаній вже майже однакова, чого і треба було досягти за умовою завдання.

**Завдання для самостійної роботи.** Необхідно знайти статистичні дані (щонайменше 100 спостережень), які містять булевську змінну та створити стратифіковану 10 % вибірку. Навести економічну інтерпретацію.

### Лабораторна робота № 3. Багатофакторне ранжування

**Мета** – опанування студентами навичок проведення ранжування економічних об'єктів різними методами.

**Завдання.** Необхідно провести оцінку та аналіз ЕБК України. В якості об'єктів дослідження виступають 15 країн світу – Австрія, Бельгія, Болгарія, Фінляндія, Франція, Німеччина, Італія, Польща, Іспанія, Швеція, Швейцарія, Великобританія, Білорусія, Україна і Росія. Вхідні дані

дослідження показані в табл. 1. З таблиці видно, що для 15 країн притаманні різнопланові значення обраних для аналізу п'яти показників енергетичної безпеки

1. Частка власних джерел у балансі паливно-енергетичних ресурсів держави, % (ЧВД).

2. Частка домінуючого паливного ресурсу у споживанні паливно-енергетичних ресурсів, % (ЧДП).

3. Енергоємність ВВП, кг умовного палива/грн (ЕЄВВП).

4. Обсяг видобутку вугілля, млн тонн (ОВВ).

5. Ступінь забезпечення паливно-енергетичними ресурсами (СЗР).

Таблиця 1

### Значення показників енергетичної безпеки

	ЧДП	ЕЄВВП	ОВВ	ЧВД	СЗР
Австрія	0,512517	8,139704	28,08643	0,657682	1,229883
Бельгія	0,441011	11,18419	46,6101	0,423147	1,198028
Болгарія	0,390566	67,68311	473,7965	0,728101	0,629035
Фінляндія	0,659931	13,74058	0	0,644842	1,181073
Франція	0,598438	7,578963	6,85112	0,746463	0,569163
Німеччина	0,425501	9,358731	2922,118	0,671615	0,998937
Італія	0,422215	9,180896	4,465834	0,526936	1,333184
Польща	0,447059	27,10571	6933,373	1,057638	1,030381
Іспанія	0,493978	10,29191	4295,369	0,61782	0,923633
Швеція	0,828147	8,865035	0	0,909419	0,948364
Швейцарія	0,694735	4,621731	0	0,75509	0,952061
Великобританія	0,389766	7,828034	1005,633	1,954056	2,160322
Білорусія	0,603121	80,72323	0	0,521061	1,342974
Росія	0,562118	121,3625	3960,709	2,899916	2,942098
Україна	0,541134	179,2357	2740,338	0,676241	1,076723

### Методичні рекомендації

1. Формування матриці спостережень. З цією метою у новоствореній книзі MSExcel 2007 будуємо таблицю відповідно до наведеної вище (табл. 1) та вводимо до неї вихідні данні (рис. 19).

лаб 4 - Microsoft Excel

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1		ЧДП	ЕЄВВП	ОВВ	ЧВД	СЗР												
2	Австрія	0,512517	8,139704	2,295999	0,657682	1,229883												
3	Бельгія	0,441011	11,18419	4,823341	0,423147	1,198028												
4		0,390566102100274	67,68311	182,7928	0,728101	0,629035												
5	Фінляндія	0,659931	13,74058	0	0,644842	1,181073												
6	Франція	0,598438	7,578963	4,139735	0,746463	0,569163												
7	Германія	0,425501	9,358731	2408,544	0,671615	0,998937												
8	Італія	0,422215	9,180896	2,592751	0,526936	1,333184												
9	Польща	0,447059	27,10571	2792,817	1,057638	1,030381												
10	Іспанія	0,493978	10,29191	385,999	0,61782	0,923633												
11	Швеція	0,828147	8,865035	0	0,909419	0,948364												
12	Швейцарія	0,694735	4,621731	0	0,75509	0,952061												
13	Нікоритя	0,389766	7,828034	606,1022	1,954056	2,160322												
14	Білорусія	0,603121	80,72323	0	0,521061	1,342974												
15	Росія	0,562118	121,3625	5683,912	2,899916	2,942098												
16	Україна	0,4057	130,269	1396,38	0,676241	1,21												
17																		
18																		
19																		

Рис. 19. Формування матриці вихідних даних

2. Визначити рейтинг країн за показником енергоємності ВВП, кг умовного палива/грн (ЕЄВВП) та побудувати їх ренкінг. Для цього проводиться сортування показників за стовпцем ЕЄВВП за зростанням ознаки (рис. 20). Виходячи з того що обраний показник є дестимулятор, то рейтингове місце 1 віддається країні з найменшим значенням показника.

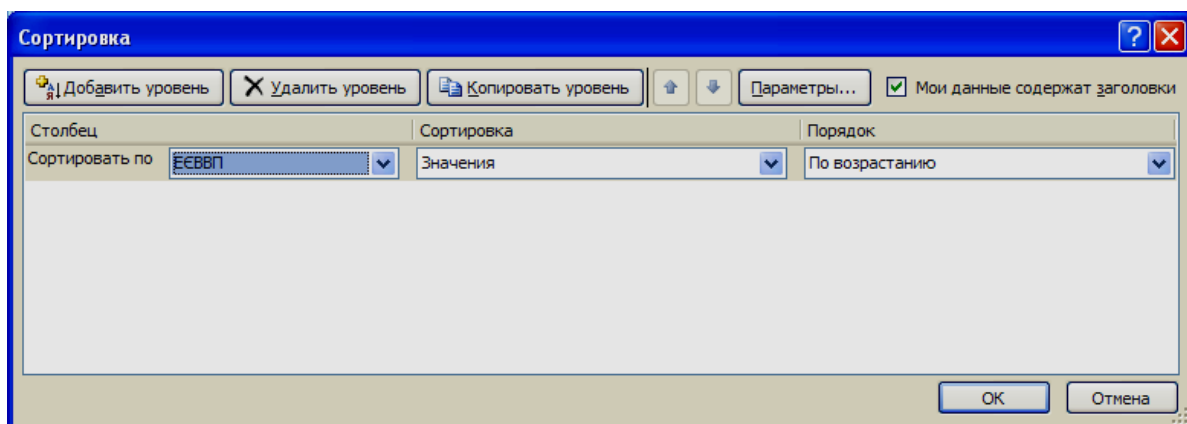


Рис. 20. Сортування за обраною ознакою

Ренкінг країн світу за показником енергоємність ВВП має вид приведений у табл. 2.

Таблиця 2

**Ренкінг країн світу за показником енергоємність ВВП**

№ п/п	Країна	ЧДП	ЕЄВВП	ОВВ	ЧВД	СЗР
1	Швейцарія	0,694735	<b>4,621731</b>	0	0,75509	0,952061
2	Франція	0,598438	<b>7,578963</b>	4,139735	0,746463	0,569163
3	Великобританія	0,389766	<b>7,828034</b>	606,1022	1,954056	2,160322
4	Австрія	0,512517	<b>8,139704</b>	2,295999	0,657682	1,229883
5	Швеція	0,828147	<b>8,865035</b>	0	0,909419	0,948364
6	Італія	0,422215	<b>9,180896</b>	2,592751	0,526936	1,333184
7	Німеччина	0,425501	<b>9,358731</b>	2408,544	0,671615	0,998937
8	Іспанія	0,493978	<b>10,29191</b>	385,999	0,61782	0,923633
9	Бельгія	0,441011	<b>11,18419</b>	4,823341	0,423147	1,198028
10	Фінляндія	0,659931	<b>13,74058</b>	0	0,644842	1,181073
11	Польща	0,447059	<b>27,10571</b>	2792,817	1,057638	1,030381
12	Болгарія	0,390566	<b>67,68311</b>	182,7928	0,728101	0,629035
13	Білорусія	0,603121	<b>80,72323</b>	0	0,521061	1,342974
14	Росія	0,562118	<b>121,3625</b>	5683,912	2,899916	2,942098
15	Україна	0,4057	<b>130,269</b>	1396,38	0,676241	1,21

3. Розрахунок узагальнюючого показника рівня розвитку. Стандартизація ознак. Стандартизації матриці вихідних даних проводиться за формулами:

$$z_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j},$$

де  $\bar{x}_j = \frac{1}{m} \sum_{t=1}^m x_{ij},$

$$s_j = \left[ \frac{1}{m} \sum_{i=1}^m (x_{ij} - \bar{x}_j)^2 \right]^{1/2},$$

$\bar{x}_j$  – середнє арифметичне значення показника  $j$ ;  $s_j$  – стандартне відхилення показника  $j$ .

Попередні розрахунки проводяться з використанням стандартних функцій СРЗНАЧ та СТАНДАРТОТКЛ, як показано на рис. 21, 22.

The screenshot shows an Excel spreadsheet with the following data:

	A	B	C	D	E	F	G
1		ЧДП	ЕЄВВП	ОВВ	ЧВД	СЗР	
2	Австрія	0,512517	8,139704	2,295999	0,657682	1,229883	
3	Бельгія	0,441011	11,18419	4,823341	0,423147	1,198028	
4	Болгарія	0,390566	67,68311	182,7928	0,728101	0,629035	
5	Фінляндія	0,659931	13,74058	0	0,644842	1,181073	
6	Франція	0,598438	7,578963	4,139735	0,746463	0,569163	
7	Германія	0,425501	9,358731	2408,544	0,671615	0,998937	
8	Італія	0,422215	9,180896	2,592751	0,526936	1,333184	
9	Польща	0,447059	27,10571	2792,817	1,057638	1,030381	
10	Іспанія	0,493978	10,29191	385,999	0,61782	0,923633	
11	Швеція	0,828147	8,865035	0	0,909419	0,948364	
12	Швейцарія	0,694735	4,621731	0	0,75509	0,952061	
13	Велика Британія	0,389766	7,828034	606,1022	1,954056	2,160322	
14	Білорусія	0,603121	80,72323	0	0,521061	1,342974	
15	Росія	0,562118	121,3625	5683,912	2,899916	2,942098	
16	Україна	0,4057	130,269	1396,38	0,676241	1,21	
17	хсередне	=СРЗНАЧ(В2:В16	898,0266	0,919335	1,243276		
18	s	СРЗНАЧ(число1; [число2]; ...)	0	0	0		

Рис. 21. Розрахунок середніх значень за кожною з ознак

The screenshot shows an Excel spreadsheet with the following data:

	A	B	C	D	E	F	G
1		ЧДП	ЕЄВВП	ОВВ	ЧВД	СЗР	
2	Австрія	0,512517	8,139704	2,295999	0,657682	1,229883	
3	Бельгія	0,441011	11,18419	4,823341	0,423147	1,198028	
4	Болгарія	0,390566	67,68311	182,7928	0,728101	0,629035	
5	Фінляндія	0,659931	13,74058	0	0,644842	1,181073	
6	Франція	0,598438	7,578963	4,139735	0,746463	0,569163	
7	Германія	0,425501	9,358731	2408,544	0,671615	0,998937	
8	Італія	0,422215	9,180896	2,592751	0,526936	1,333184	
9	Польща	0,447059	27,10571	2792,817	1,057638	1,030381	
10	Іспанія	0,493978	10,29191	385,999	0,61782	0,923633	
11	Швеція	0,828147	8,865035	0	0,909419	0,948364	
12	Швейцарія	0,694735	4,621731	0	0,75509	0,952061	
13	Велика Британія	0,389766	7,828034	606,1022	1,954056	2,160322	
14	Білорусія	0,603121	80,72323	0	0,521061	1,342974	
15	Росія	0,562118	121,3625	5683,912	2,899916	2,942098	
16	Україна	0,4057	130,269	1396,38	0,676241	1,21	
17	хсередне	0,524987	34,52889	898,0266	0,919335	1,243276	
18	s	=СТАНДОТКЛОН(В2:В16			0	0	

Рис. 22. Розрахунок стандартного відхилення

Після розрахунку середніх значень та стандартного відхилення за кожною з ознак, проводиться розрахунок стандартизованої матриці (рис. 23).



(x-xsr) <sup>2</sup>					
	ЧДП	ЄЄВВП	ОВВ	ЧВД	СЗР
Австрія	0,00016	696,389	802333,3	0,06846	0,000179375
Бельгія	0,00705	544,975	797812,1	0,2462	0,00204737
Болгарія	0,01807	1099,2	511559,4	0,03657	0,37729176
Фінляндія	0,01821	432,154	806451,8	0,07535	0,003869152
Франція	0,0054	726,298	799033,7	0,02988	0,454427204
Германія	0,0099	633,537	2281664	0,06137	0,059701577
Італія	0,01056	642,521	801801,8	0,15398	0,008083548
Польща	0,00607	55,1036	3590230	0,01913	0,045324324
Іспанія	0,00096	587,431	262172,2	0,09091	0,102171573
Швеція	0,09191	658,633	806451,8	9,8E-05	0,086972938
Швейцарія	0,02881	894,438	806451,8	0,02698	0,084806058
Великобританія	0,01828	712,936	85219,83	1,07065	0,840974048
Білорусія	0,0061	2133,92	806451,8	0,15862	0,009939715
Росія	0,00138	7540,08	22904699	3,9227	2,88599825
Україна	0,01423	9166,17	248356,1	0,05909	0,001107266
	0,23709	26523,8	36310688	6,01999	4,962894158
Standart					
	ЧДП	ЄЄВВП	ОВВ	ЧВД	СЗР
Австрія	-0,09918	-0,62756	-0,57571	-0,41302	-0,023284107
Бельгія	-0,66794	-0,55516	-0,57409	-0,78324	-0,078664043
Болгарія	-1,06918	0,78844	-0,4597	-0,30186	-1,067866059
Фінляндія	1,07335	-0,49436	-0,57719	-0,43329	-0,108139914
Франція	0,58423	-0,64089	-0,57453	-0,27288	-1,17195324
Германія	-0,79131	-0,59857	0,970855	-0,39103	-0,424786811
Італія	-0,81745	-0,6028	-0,57552	-0,61941	0,156307252
Польща	-0,61984	-0,17653	1,217838	0,21831	-0,370121053
Іспанія	-0,24664	-0,57638	-0,3291	-0,47595	-0,55570354
Швеція	2,41134	-0,61031	-0,57719	-0,01565	-0,512708117
Швейцарія	1,35018	-0,71122	-0,57719	-0,25926	-0,506280921
Великобританія	-1,07555	-0,63497	-0,18763	1,63332	1,594297701
Білорусія	0,62148	1,09854	-0,57719	-0,62868	0,173326549
Росія	0,29534	2,06498	3,076031	3,12637	2,953426413
Україна	-0,94881	2,27678	0,320307	-0,38373	-0,057850111
etalon	2,41134	2,27678	3,076031	-0,78324	-1,17195324

Рис. 23. Розрахунок значень стандартизованої матриці

4. Побудова еталону розвитку передбачає виділення підмножини стимуляторів та дестимуляторів. Серед наведених показників стимуляторами є : частка власних джерел у балансі паливно-енергетичних ресурсів держави (ЧВД), обсяг видобутку вугілля (ОВВ), Ступінь забезпечення паливно-енергетичними ресурсами (СЗР). До дестимуляторів належать: частка домінуючого паливного ресурсу у споживанні паливно-енергетичних ресурсів (ЧДП), енергоємність ВВП (ЄЄВВП). Поділ показників на дві групи проводиться відповідно до їх економічного змісту.

Еталон розвитку має такий вигляд

$$Z_0 = [z_{01}, z_{02}, \dots, z_{0n}],$$

де  $z_{0s} = \max_i z_{is}$ , як що  $s \in I$ ,

$z_{0s} = \min_t z_{it}$ , як що  $s \notin I (s = 1, \dots, n)$ ,

$I$  – множина стимуляторів.

Відповідно до поділу показників на стимулятори і дестимулятори з застосуванням стандартних функцій MAX і MIN було розраховано значення еталону розвитку (рис. 24).

	Standart				
	ЧДП	ЄСВВП	ОВВ	ЧВД	СЗР
Австрія	-0,09918	-0,62756	-0,575713	-0,41302	-0,023284107
Бельгія	-0,66794	-0,55516	-0,574088	-0,78324	-0,078664043
Болгарія	-1,06918	0,788437	-0,459702	-0,30186	-1,067866059
Фінляндія	1,073348	-0,49436	-0,577188	-0,43329	-0,108139914
Франція	0,584228	-0,64089	-0,574528	-0,27288	-1,17195324
Германія	-0,79131	-0,59857	0,9708547	-0,39103	-0,424786811
Італія	-0,81745	-0,6028	-0,575522	-0,61941	0,156307252
Польща	-0,61984	-0,17653	1,217838	0,218313	-0,370121053
Іспанія	-0,24664	-0,57638	-0,329095	-0,47595	-0,55570354
Швеція	2,411335	-0,61031	-0,577188	-0,01565	-0,512708117
Швейцарія	1,350178	-0,71122	-0,577188	-0,25926	-0,506280921
Великобританія	-1,07555	-0,63497	-0,187628	1,633319	1,594297701
Білорусія	0,621482	1,098542	-0,577188	-0,62868	0,173326549
Росія	0,295343	2,06498	3,0760308	3,12637	2,953426413
Україна	-0,94881	2,276784	0,3203065	-0,38373	-0,057850111
Z0	-1,07555	-0,71122	3,0760308	3,12637	2,953426413

Рис. 24. Розрахунок еталону розвитку

5. Розрахунок відстаней між точками-спостереженнями й точкою-еталоном  $Z_0$  проводиться за такою формулою:

$$c_{i0} = \left[ \sum_{s=1}^n (z_{ij} - z_{0j})^2 \right]^{1/2}, i = 1, 2, \dots, m.$$

Проведений розрахунок дозволяє сформулювати матрицю відстаней (рис. 25).

6. Розрахунок таксономічного показника рівня розвитку:

$$d^* = (d_1^*, d_2^*, \dots, d_m^*), \quad d_i^* = \frac{c_{i0}}{c_0},$$

$$c_0 = \bar{c}_0 + 2 * S_0, \quad \bar{c}_0 = \frac{\sum_{i=1}^m c_{i0}}{m}, \quad S_0 = \sqrt{\frac{\sum_{i=1}^m (c_{i0} - \bar{c}_0)^2}{m}}$$

На рис. 25 наведено проведені розрахунки для отримання вектору значень  $d^*$ .

	(z-z0)*2					
	ЧДП	ЄСВВП	ОВВ	ЧВД	СЗР	
Австрія	0,95329	0,006999	13,335231	12,5273	8,860805522	35,68362
Бельгія	0,16614	0,024355	13,32337	15,28505	9,193572536	37,99249
Болгарія	4,05E-05	2,248966	12,501407	11,7528	16,17079315	42,674
Фінляндія	4,617751	0,047026	13,346011	12,67118	9,373188377	40,05516
Франція	2,754856	0,004946	13,326577	11,55491	17,01875728	44,66004
Германія	0,08079	0,01269	4,4317665	12,3721	11,41232459	28,30967
Італія	0,066616	0,011755	13,333838	14,03085	7,823875603	35,26693
Польща	0,207668	0,285892	3,4528807	8,456799	11,04596776	23,44921
Іспанія	0,687081	0,018182	11,594884	12,97668	12,31399303	37,59082
Швеція	12,15835	0,010183	13,346011	9,872312	12,01408858	47,40095
Швейцарія	5,884147	0	13,346011	11,46251	11,96957484	42,66224
Великобританія	0	0,005814	10,651472	2,229203	1,847230855	14,73372
Білорусія	2,879911	3,275233	13,346011	14,10041	7,728955258	41,33052
Росія	1,879342	7,707276	0	0	0	9,586618
Україна	0,016063	8,92816	7,5940162	12,32078	9,067786304	37,92681
		еталон розвитку				
		c0i	(c0i-c0 cp)*	d*		
Австрія	35,68362	5,973577	0,033627	0,757766		
Бельгія	37,99249	6,163805	0,1395802	0,781897		
Болгарія	42,674	6,532534	0,5510595	0,828671		
Фінляндія	40,05516	6,328914	0,290213	0,802842		
Франція	44,66004	6,682817	0,7967644	0,847735		
Германія	28,30967	5,320683	0,2204462	0,674945		
Італія	35,26693	5,938597	0,0220216	0,753329		
Польща	23,44921	4,842438	0,8982531	0,614278		
Іспанія	37,59082	6,131135	0,1162366	0,777753		
Швеція	47,40095	6,884835	1,1982241	0,873362		
Швейцарія	42,66224	6,531634	0,5497236	0,828557		
Великобританія	14,73372	3,838453	3,8093182	0,486919		
Білорусія	41,33052	6,428882	0,4079137	0,815523		
Росія	9,586618	3,096226	7,2574951	0,392766		
Україна	37,92681	6,158475	0,1356259	0,781221		

Рис. 25. Розрахунки вектору значень  $d^*$ .

Для того щоб показник  $d^*$  приймав високі значення при більших значеннях стимуляторів і низькі значення при малих значеннях стимуляторів, його перетворюють до виду:

$$d_i = 1 - \frac{c_{i0}}{c_0}$$

У табл. 3 наведені результати розрахунку.

Як видно з табл. 3, Україна має не найгірший показник енергетичної безпеки серед обраних країн, але є аутсайдером. Країни, що мають нижчий показник енергетичної безпеки не видобувають вугілля, чи видобувають його в невеликій кількості. Тобто можна зробити висновок, що Україна при достатньо великих власних запасах енергоресурсів нерационально використовує їх, це підтверджується тим, що Україна має найгірший показник енергоємності ВВП серед європейських країн.

**Значення узагальнюючого показника ЕБК**

	Країна	Узагальнюючий показник енергетичної безпеки
1	Росія	0,602159791
2	Великобританія	0,548855061
3	Польща	0,420904309
4	Німеччина	0,3189488
5	Іспанія	0,317985185
6	Італія	0,255845839
7	Австрія	0,251346831
8	Бельгія	0,226923613
9	Фінляндія	0,200271334
10	Білорусія	0,198356135
11	Болгарія	0,189788944
12	Україна	0,17500943
13	Швейцарія	0,172319901
14	Франція	0,154621963
15	Швеція	0,121230012

3. **Застосування адитивного методу**, коли підсумковий рейтинг об'єкта виставляється за підсумковою сумою всіх елементів рядка. При цьому можливе використання вагових коефіцієнтів для кожного стовпця, що характеризує суттєвість ранжування за цією ознакою. Підсумкове рейтингове число розраховується в даному випадку як

$$R_i = \sum_{j=1}^n \alpha_j \times R_{ij}$$

де  $\alpha_j$  – ваговий коефіцієнт, що визначається експертним методом.

3.1. За кожним з показників необхідно виставити рейтинг (табл. 4). Процедура відповідна до розрахунків проведених у пункті 1.

3.2. Наступним етапом є визначення вагових коефіцієнтів, що виставляються експертним методом. Для визначення вагових коефіцієнтів використовується метод ранжувань. Даний метод полягає в тому, що кожного експерта просять розташувати ознаки в порядку переваги. Цифрою один позначається найбільш важлива ознака, цифрою два – наступний після нього по важливості і т.д.

Таблиця 4.

**Рейтингова оцінка за кожним з показників ЕБК**

	ЧДП	ЕЄВВП	ОВВ	ЧВД	СЗР
Австрія	8	4	1	8	4
Бельгія	5	9	2	13	5
Білорусія	10	13	3	12	2
Болгарія	2	12	4	6	12
Великобританія	1	3	5	1	1
Німеччина	4	7	6	7	8
Іспанія	7	8	7	10	11
Італія	3	6	8	11	3
Польща	6	11	9	2	7
Фінляндія	11	10	10	9	6
Франція	9	2	11	5	13
Швейцарія	12	1	12	4	9
Швеція	13	5	13	3	10

Отримані дані, щодо суттєвості ознаки для оцінки рівня енергетичної безпеки заносять у таблицю 5.

Таблиця 5

**Експертні оцінки**

	Э <sub>1</sub>	Э <sub>2</sub>	Э <sub>3</sub>	Э <sub>4</sub>	Э <sub>5</sub>
ЧДП	1	2	1,5	1	2
ЕЄВВП	2,5	2	1,5	2,5	1
ОВВ	2,5	2	3	2,5	3
ЧВД	4	5	4,5	4,5	4
СЗР	5	4	4,5	4,5	5,5

Перед тим як проводити узагальнення думок експертів, перевіримо їх на узгодженість, що є передумовою якості отриманих оцінок. Узгодженість думок експертів перевіримо за допомогою коефіцієнта рангової кореляції Кендела, Спірмена та коефіцієнта конкордації в пакеті Statistica 8.0.

З метою розрахунку коефіцієнти в парної кореляції Кендела та Спірмена необхідно представити дані в форматі, наведеному на рис. 26.

	1 E1	2 E2	3 E3	4 E4	5 E5	6 Var6	7 Var7	8 Var8	9 Var9	10 Var10
1	1	2	1,5	1	2					
2	2,5	2	1,5	2,5	1					
3	2,5	2	3	2,5	3					
4	4	5	4,5	4,5	4					
5	5	4	4,5	4,5	5,5					
6										
7										
8										
9										
10										

Рис. 26. Вихідні дані

У модулі Nonparametric Statistics обираємо меню Correlation, що дозволяє розрахувати коефіцієнти рангової кореляції. На рис. 27-30 наведено порядок дій для розрахунку коефіцієнта рангової кореляції Спірмена.

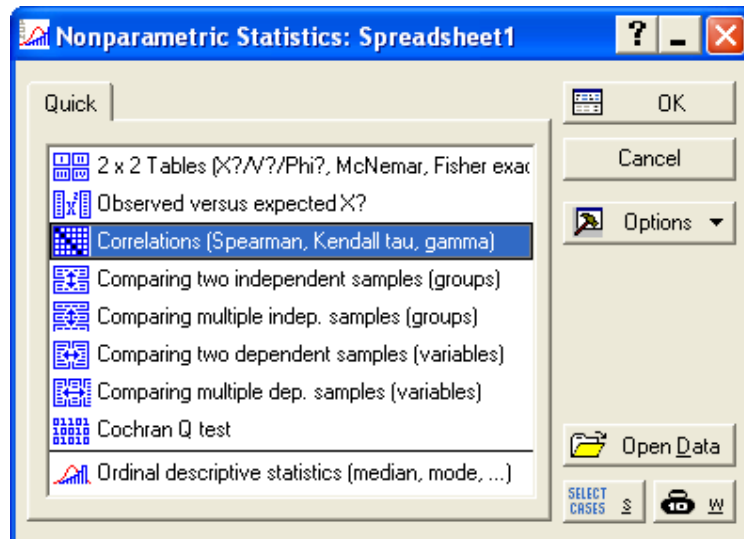


Рис. 27. Діалогове вікно вибору виду непараметричних критеріїв

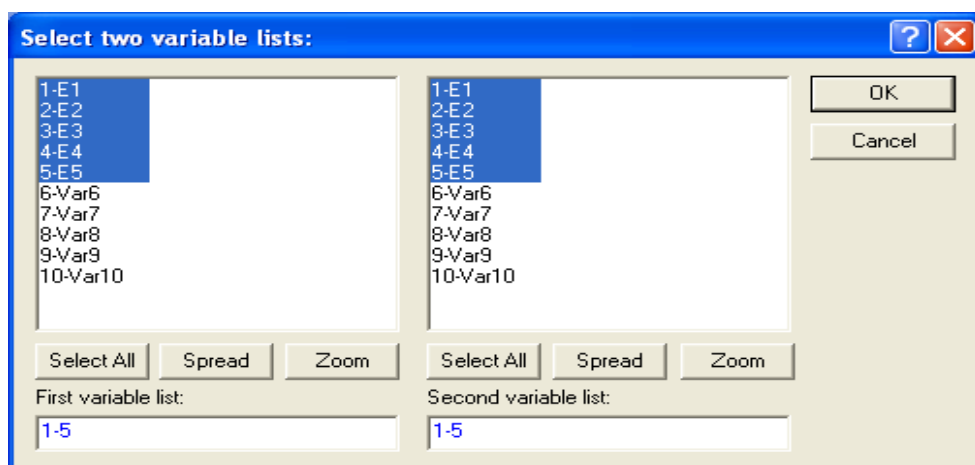


Рис. 28. Вибір вихідних даних для розрахунку

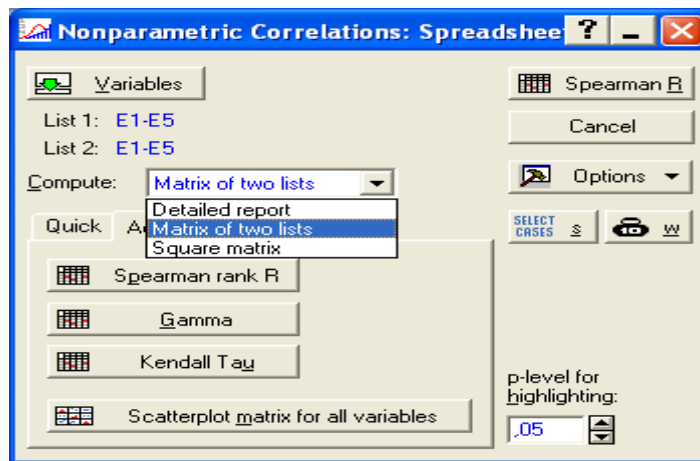


Рис. 29. Діалогове вікно модулю Непараметрична кореляція

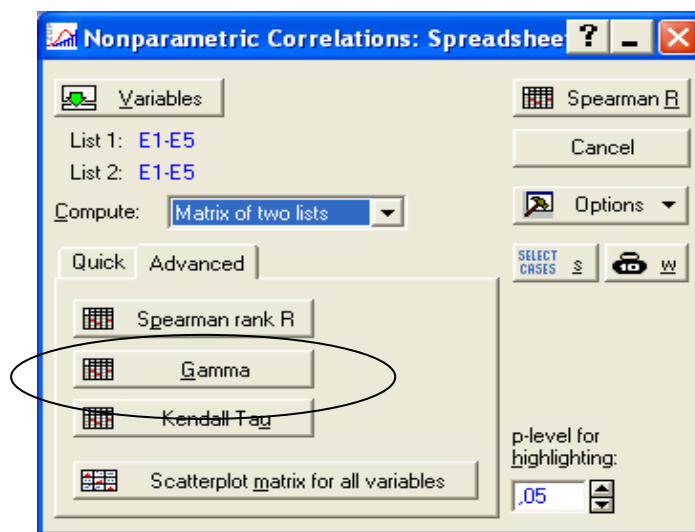


Рис. 30. Вибір коефіцієнта рангової кореляції Спірмана для розрахунку

На рис. 31 наведено матрицю рангової кореляції Спірмена, з якої видно що в більшості своїй експерти мають узгоджену думку, щодо питань дослідження.

Kendall Tau Correlations (Spreadsheet1)					
MD pairwise deleted					
Marked correlations are significant at $p < .05000$					
Variable	E1	E2	E3	E4	E5
E1	1	0,629940788	0,824957911	0,942809042	0,737864787
E2	0,629940788	1	0,801783726	0,801783726	0,597614305
E3	0,824957911	0,801783726	1	0,875	0,894427191
E4	0,942809042	0,801783726	0,875	1	0,670820393
E5	0,737864787	0,597614305	0,894427191	0,670820393	1

Рис. 31. Матриця коефіцієнта рангової кореляції Кендела

Для розрахунку коефіцієнта конкордації необхідно представити дані у вигляді, наведеному на рис. 32.

	1 X1	2 X2	3 X3	4 X4	5 X5	6 Var6
E1	1	2,5	2,5	4	5	
E2	2	2	2	5	4	
E3	1,5	1,5	3	4,5	4,5	
E4	1	2,5	2,5	4,5	4,5	
E5	2	1	3	4	5,5	

Рис. 32. Вихідні дані для розрахунку коефіцієнта конкордації

За допомогою модулю Nonparametric Statistics обираємо меню Comparing multiple dep. Sample (рис. 33) переходимо до розрахунку рангового дисперсного аналізу Фрідмена та розрахунку коефіцієнта конкордації Кендела (рис. 34).

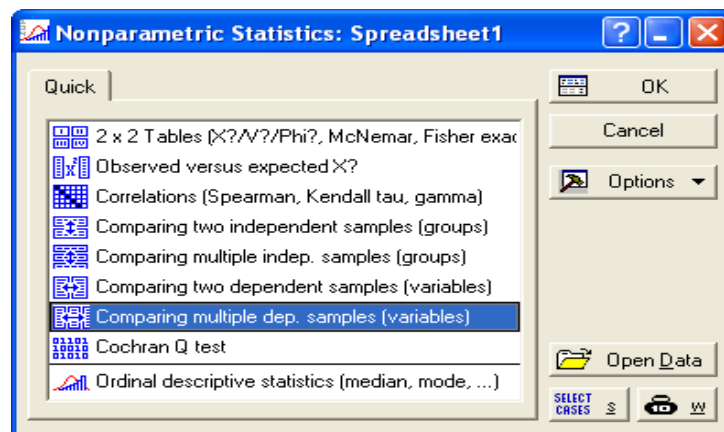


Рис. 33. Діалогове вікно модулю "Непараметричні статистики"



Рис. 34. Діалогове вікно розрахунку рангового дисперсійного аналізу та коефіцієнта конкордації



На рис. 35 наведено результати розрахунку коефіцієнту конкордації Кендела та оцінка його значимості за критерієм  $\chi^2$ . З ймовірністю 0,13 % коефіцієнт конкордації є значним, тобто експерти мають узгоджену думку

Friedman ANOVA and Kendall Coeff. of Concordance				
ANOVA Chi Sqr. (N = 5, df = 4) = 17,89011 p = ,001				
Coeff. of Concordance = ,89451 Aver. rank r = ,8681				
Variable	Average Rank	Sum of Ranks	Mean	Std.Dev.
X1	1,50000	7,50000	1,50000	0,50000
X2	1,90000	9,50000	1,90000	0,65192
X3	2,60000	13,00000	2,60000	0,41833
X4	4,40000	22,00000	4,40000	0,41833
X5	4,60000	23,00000	4,70000	0,57008

Рис. 35. Результат розрахунку рангового дисперсійного аналізу та коефіцієнту конкордації

Оскільки думка експертів є узгодженою, то результати експертизи можуть бути використані для подальших розрахунків.

3.3. За допомогою методів математичної статистики необхідно отримати узагальнену думку експертів. Визначається середній ранг, середнє статистичне значення  $S_j$   $j$ -го ознаки:

$$S_j = \frac{\sum_{i=1}^m \alpha_{ij}}{m_{kj}}$$

де  $m_{kj}$  – кількість експертів, що оцінюють  $j$ -ту ознаку ( $m_k \leq m$ );  $i$  – номер експерту;  $i = 1, \dots, m$ ;  $j$  – номер ознаки,  $j = 1, 2, \dots, n$ .

Результати розрахунку наведено в табл. 6.

Таблиця 6

### Розрахунок середнього рангу

	$\alpha_1$	$\alpha_2$	$\alpha_3$	$\alpha_4$	$\alpha_5$	S	$\alpha_i$
ЧДП	1	2	1,5	1	2	1,5	0,099338
ЕЄВВП	2,5	2	1,5	2,5	1	1,9	0,125828
ОВВ	2,5	2	3	2,5	3	2,6	0,172185
ЧВД	4	5	4,5	4,5	4	4,4	0,291391
СЗР	5	4	4,5	4,5	5,5	4,7	0,311258

Використовуючи формулу розрахунку загального рейтингу в табл. 7 розрахувати рейтинг за кожною з країн.

Таблиця 7

**Розрахунок рейтингу країн світу за рівнем ЕБК з використанням оцінок експертів**

Країни	ЧДП	ЕЄВВП	ОВВ	ЧВД	СЗР	R
Австрія	8	4	1	8	4	5,0464
Бельгія	5	9	2	13	5	7,3179
Білорусія	10	13	3	12	2	7,2649
Болгарія	2	12	4	6	12	7,8808
Великобританія	1	3	5	1	1	1,9404
Німеччина	4	7	6	7	8	6,8411
Іспанія	7	8	7	10	11	9,245
Італія	3	6	8	11	3	6,5695
Польща	6	11	9	2	7	6,2914
Фінляндія	11	10	10	9	6	8,5629
Франція	9	2	11	5	13	8,543
Швейцарія	12	1	12	4	9	7,351
Швеція	13	5	13	3	10	8,1457
$\alpha_j$	0,0993	0,1258	0,1722	0,2914	0,3113	

Аналіз розрахунку рейтингу за трьома методами довів, що введення вагових коефіцієнтів відповідних до переваг експертів, дозволило подолати недоліки перших двох методів і отримати досить адекватну оцінку ситуації.

**Лабораторна робота №4. Побудова однофакторної лінійної регресії**

**Мета** – закріплення теоретичного та практичного матеріалу, набуття навичок побудови та аналізу простих економетричних моделей у модулі.

**Завдання.** Необхідно перевірити наявність лінійного зв'язку між макроекономічними показниками (табл. 8) у модулі Multiple Regression ППП Statistica:

1. Побудувати лінійну економетричну модель і визначити всі її характеристики (параметри моделі, середньоквадратичне відхилення

параметрів моделі, дисперсію і середньоквадратичне відхилення помилок моделі, коефіцієнти кореляції і детермінації).

2. Перевірити статистичну значущість параметрів моделі та коефіцієнта кореляції за критерієм Стюдента. Перевірити адекватність моделі за критерієм Фішера.

3. Розрахувати теоретичні значення залежної змінної і помилки моделі. Побудувати графік лінійної функції з довірчими інтервалами. Побудувати гістограму та графік розподілу помилок на нормальному ймовірнісному папері.

4. Розрахувати прогнозне значення залежної змінної та довірчі інтервали зміни, якщо відомо значення незалежного показника.

5. Зробити висновки щодо адекватності побудованої моделі, дати економічну інтерпретацію даної залежності і можливості її використання.

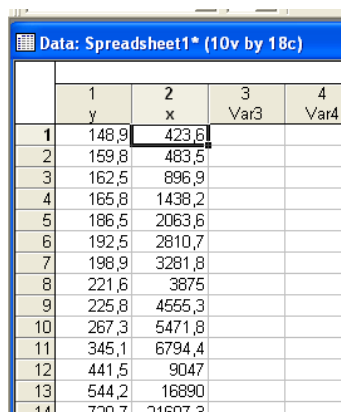
Таблиця 8.

**Рівень ВВП та прямих інвестицій в економіку України  
за 1994 – 2011 рр.**

Роки	ВВП (млрд грн)	Прямі інвестиції (млн дол.)
1994	148,9	423,6
1995	159,8	483,5
1996	162,5	896,9
1997	165,8	1 438,2
1998	186,5	2 063,6
1999	192,5	2 810,7
2000	198,9	3 281,8
2001	221,6	3875
2002	225,8	4 555,3
2003	267,3	5 471,8
2004	345,1	6 794,4
2005	441,5	9047
2006	544,2	1 6890
2007	720,7	21 607,3
2008	948,1	29 542,7
2009	913,3	35 616,4
2010	1 082,6	40 053
2011	1 316,6	44 806

## Методичні рекомендації

1. У пакеті Statistica 8.0 вносимо вихідні данні. Рівень ВВП (y) та прямі іноземні інвестиції в економіку України (x):



	1	2	3	4
	y	x	Var3	Var4
1	148,9	423,6		
2	159,8	483,5		
3	162,5	896,9		
4	165,8	1438,2		
5	186,5	2063,6		
6	192,5	2810,7		
7	198,9	3281,8		
8	221,6	3875		
9	225,8	4555,3		
10	267,3	5471,8		
11	345,1	6794,4		
12	441,5	9047		
13	544,2	16890		
14	720,7	21607,3		

Рис. 36. Фрагмент вихідних даних однофакторної моделі

2. З початку проведемо аналіз впливу факторної ознаки на результуючу. З цією метою в пакеті Statistica 8.0 обираємо Basic statistics та пункт Correlation matrices . На рис. 37 показано основні кроки побудови кореляційної матриці:

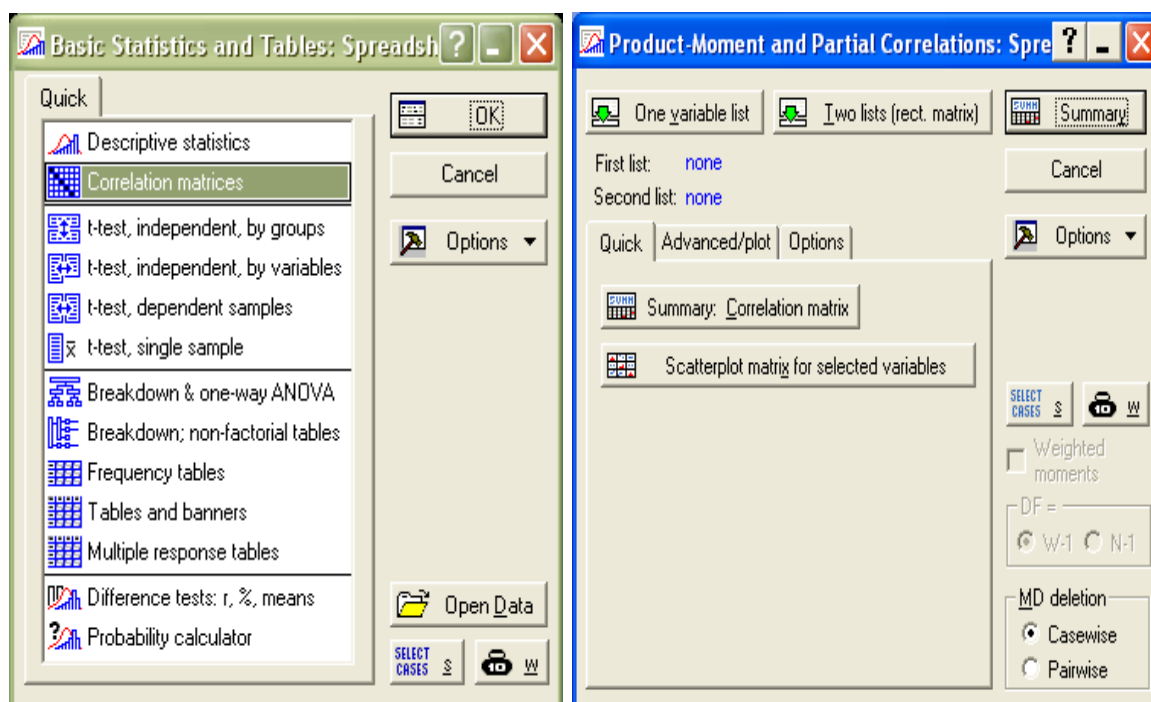
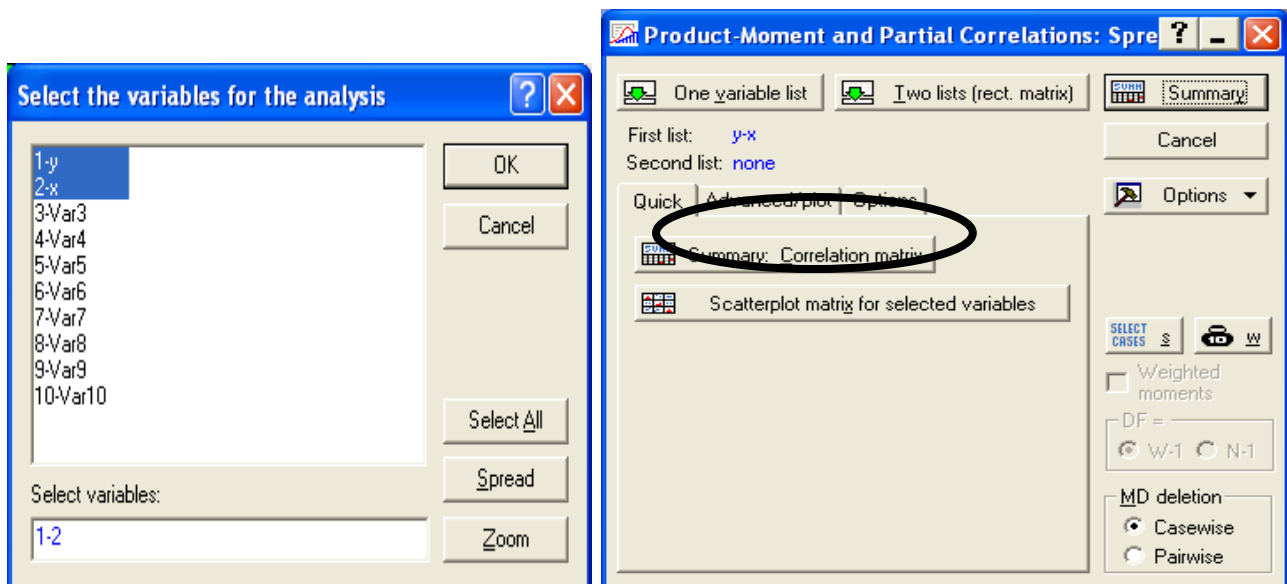


Рис. 37. Побудова кореляційної матриці



Закінчення рис. 37. Побудова кореляційної матриці

Підсумкова таблиця має вигляд наведений на рис. 38.

		Correlation Matrix Marked cells: 1 N=18 (Cases)	
Variable		y	x
y		1,00	0,99
x		0,99	1,00

Рис. 38. Підсумкова кореляційна матриця

Значення коефіцієнта парної кореляції 0,99 свідчить про високий рівень залежності між рівнем ВВП та прямими інвестиціями.

3. Далі переходимо до побудови однофакторної лінійної регресії. Щоб приступити до обчислювальних процедур, необхідно увійти в меню Statistics/Multiple Regression. Після підтвердження вибору модуля з'являється стартова панель модуля, де необхідно задати змінні для аналізу. На якій обираємо кнопку Variables у вікні, що з'явилося вказуємо Dependent (залежну) та Independent (незалежну) змінні для побудови простої одно факторної моделі. На рис. 39 наведено етапи побудови однофакторної моделі.



Характеристики моделі і ступінь їх адекватності можна отримати натиснувши кнопку Summary: Regression results (Результати регресійного аналізу). Результати побудови однофакторної економетричної моделі наведено на рис. 41.

Regression Summary for Dependent Variable: y (Spreadsheet1)						
R= ,99238117 R <sup>2</sup> = ,98482039 Adjusted R <sup>2</sup> = ,98387166						
F(1,16)=1038,0 p<,00000 Std.Error of estimate: 47,508						
N=18	Beta	Std.Err. of Beta	B	Std.Err. of B	t(16)	p-level
Intercept			140,8653	14,90628	9,45006	0,000000
x	0,992381	0,030801	0,0248	0,00077	32,21871	0,000000

Рис. 41. Результати побудови однофакторної економетричної моделі.

- Отримані результати свідчать про таке :
- Коефіцієнт множинної кореляції дорівнює 0,9923 (R).
- Коефіцієнт детермінації моделі дорівнює 0,9848 (R<sup>2</sup>).
- Скоригований коефіцієнт детермінації на число спостережень і число параметрів дорівнює 0,9838 (Adjusted R<sup>2</sup>).
- Критерій адекватності Фішера F (1,16) = 1038.
- B (a1, a2) = (140,87; 0,0248) – параметри моделі.
- Середнє квадратичне відхилення параметрів моделі становить (14,09; 0,0248).
- t (28) = (9,45; 32,22) – значимість параметрів за критерієм Стьюдента.

Аналіз наведених результатів свідчить, що модель є адекватною та має такий загальний вигляд:

$$Y = 140,87 + 0,0248 X.$$

Побудуємо графік лінійної функції з довірчими інтервалами. Для цього в меню Graphs/Scatterplots необхідно вказати змінні, лінію рівня і довірчі інтервали (рис. 42).

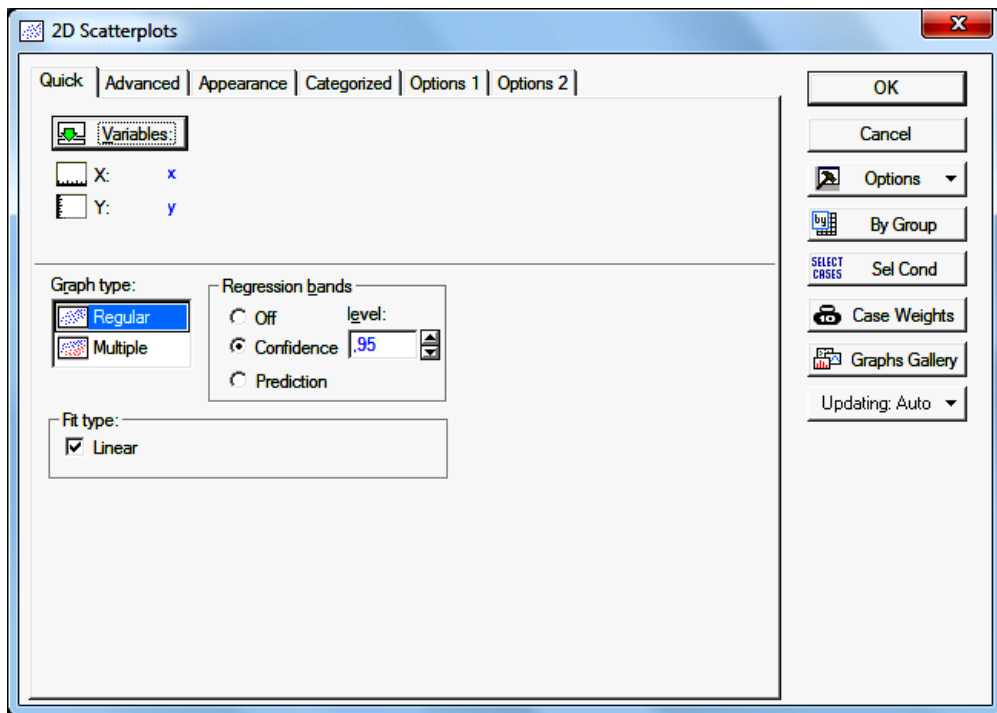


Рис. 42. Діалогове вікно побудови графіку лінійної функції

Аналіз графіка (рис. 43) доводить високу якість побудованої моделі та відповідність модельних значень фактичним.

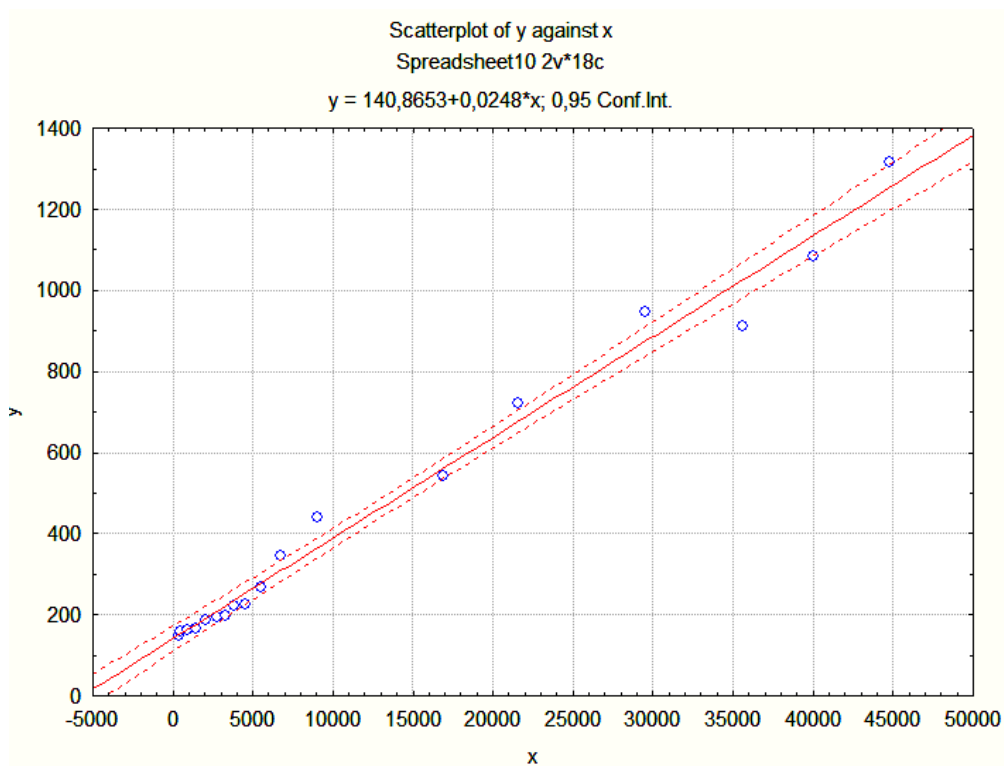


Рис. 43. Графічне представлення результатів побудованої моделі



Для розрахунку та аналізу залишків моделі, в нижній частині вікна результатів регресійної моделі див. (рис. 39) є опція *Perform residual analysis* (Всебічний аналіз залишків) (рис. 44).

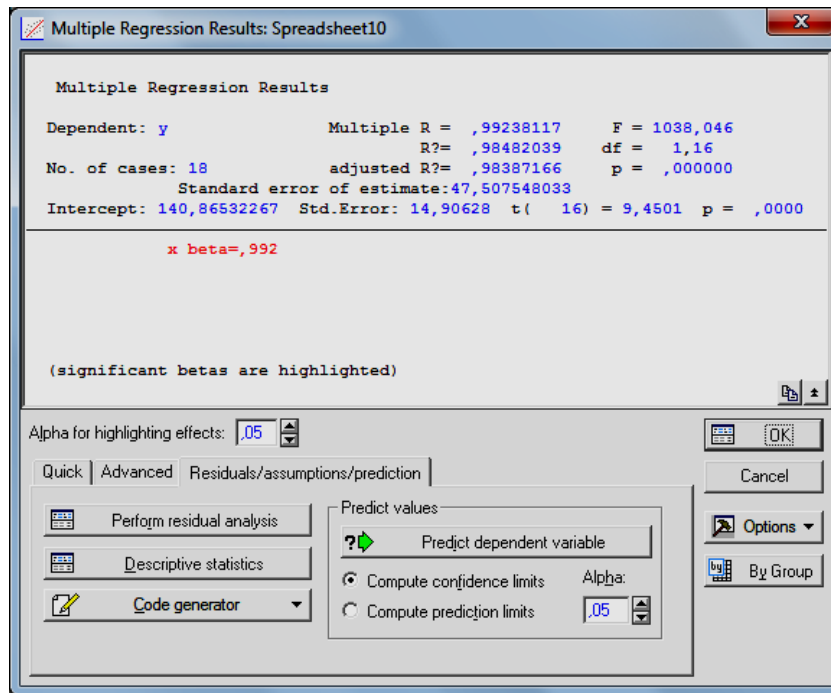


Рис. 44. Опція *Perform residual analysis* (Всебічний аналіз залишків)

Скориставшись даною опцією, отримаємо меню для аналізу помилок моделі (рис. 45).

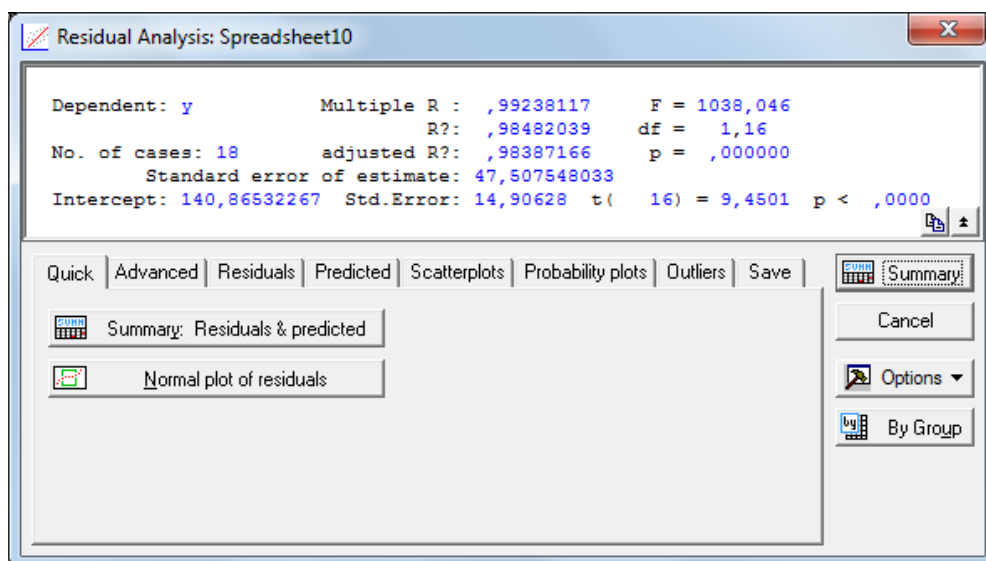


Рис. 45. Діалогове вікно модулю аналізу помилок

Кнопка аналізу помилок Summary: Residuals & Predicted дозволяє отримати таблицю, що містить фактичні значення залежної змінної (Observed value), її теоретичні значення (Predicted value) і помилки моделі (Residual) (рис. 46).

Predicted & Residual Values (Spreadsheet10)									
Dependent variable: y									
Case No.	Observed Value	Predicted Value	Residual	Standard Pred. v.	Standard Residual	Std.Err. Pred.Val	Mahalanobis Distance	Deleted Residual	Cook's Distance
1	148,900	151,390	-2,490	-0,825579	-0,05242	14,69271	0,681581	-2,754	0,000161
2	159,800	152,879	6,922	-0,821570	0,14569	14,66284	0,674977	7,650	0,001236
3	162,500	163,150	-0,650	-0,793902	-0,01368	14,45907	0,630280	-0,716	0,000011
4	165,800	176,599	-10,799	-0,757673	-0,22732	14,19866	0,574068	-11,858	0,002783
5	186,500	192,138	-5,638	-0,715815	-0,11868	13,90732	0,512392	-6,167	0,000722
6	192,500	210,701	-18,201	-0,665813	-0,38311	13,57356	0,443306	-19,819	0,007103
7	198,900	222,406	-23,506	-0,634282	-0,49478	13,37159	0,402314	-25,528	0,011437
8	221,600	237,145	-15,545	-0,594580	-0,32720	13,12715	0,353525	-16,830	0,004791
9	225,800	254,048	-28,248	-0,549048	-0,59459	12,86114	0,301454	-30,481	0,015086
10	267,300	276,819	-9,519	-0,487708	-0,20037	12,52860	0,237859	-10,231	0,001613
11	345,100	309,681	35,419	-0,399187	0,74555	12,10549	0,159351	37,879	0,020638
12	441,500	365,649	75,850	-0,248423	1,59660	11,55770	0,061714	80,622	0,085226
13	544,200	560,519	-16,319	0,276502	-0,34350	11,64205	0,076453	-17,361	0,004010
14	720,700	677,726	42,974	0,592226	0,90457	13,11302	0,350732	46,518	0,036523
15	948,100	874,891	73,209	1,123335	1,54099	17,11485	1,261882	84,127	0,203488
16	913,300	1025,800	-112,500	1,529842	-2,36804	20,88318	2,340418	-139,444	0,832367
17	1082,600	1136,033	-53,433	1,826780	-1,12472	23,84183	3,337125	-71,421	0,284608
18	1316,600	1254,127	62,473	2,144894	1,31501	27,13248	4,600569	92,714	0,621139
Minimum	148,900	151,390	-112,500	-0,825579	-2,36804	11,55770	0,061714	-139,444	0,000011
Maximum	1316,600	1254,127	75,850	2,144894	1,59660	27,13248	4,600569	92,714	0,832367
Mean	457,872	457,872	-0,000	-0,000000	-0,00000	15,26518	0,944444	-0,172	0,118497
Median	246,550	265,433	-7,579	-0,518378	-0,15952	13,74044	0,477849	-8,199	0,009270

Рис. 46. Аналіз помилок моделі

Оскільки основна гіпотеза відносно випадкової змінної свідчить, що помилки повинні бути розподілені за нормальним, представимо графік помилок моделі на нормально імовірнісному папері (Residuals/Normal plot of residuals) рис. 47.

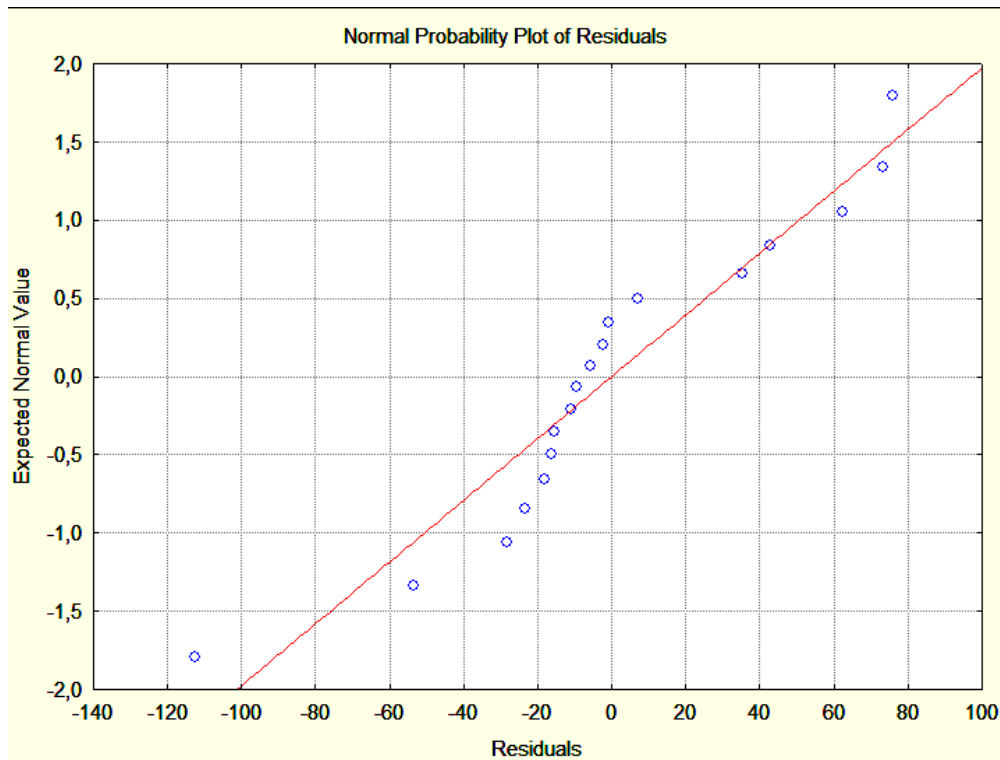


Рис. 47. Графік помилок моделі на нормально імовірнісному папері

За графіком (рис. 47) складно судити про закон розподілу помилок моделі, тому для більш детального аналізу побудуємо гістограму їх значень з нанесенням графіка нормального закону розподілу (рис. 48).

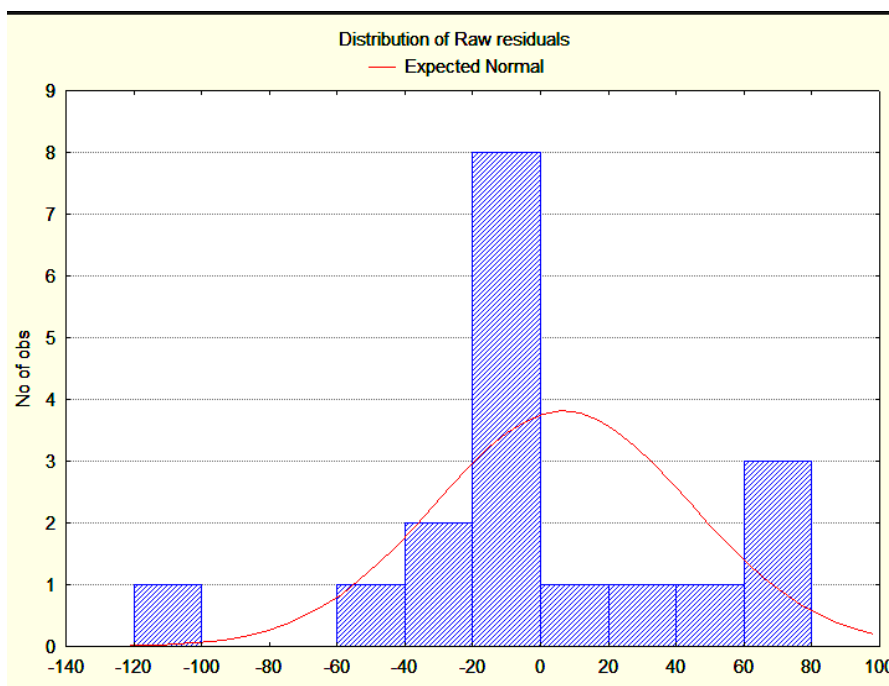


Рис. 48. Гістограма розподілу помилок моделі

Оскільки модель є адекватною, а її параметри значимі, то за моделлю можна побудувати прогноз. Щоб розрахувати прогнозні значення залежної змінної, в нижній частині вікна результатів регресійного аналізу (рис. 39) є опція Predict dependent variable (Прогнозування залежної змінної) (рис. 49).

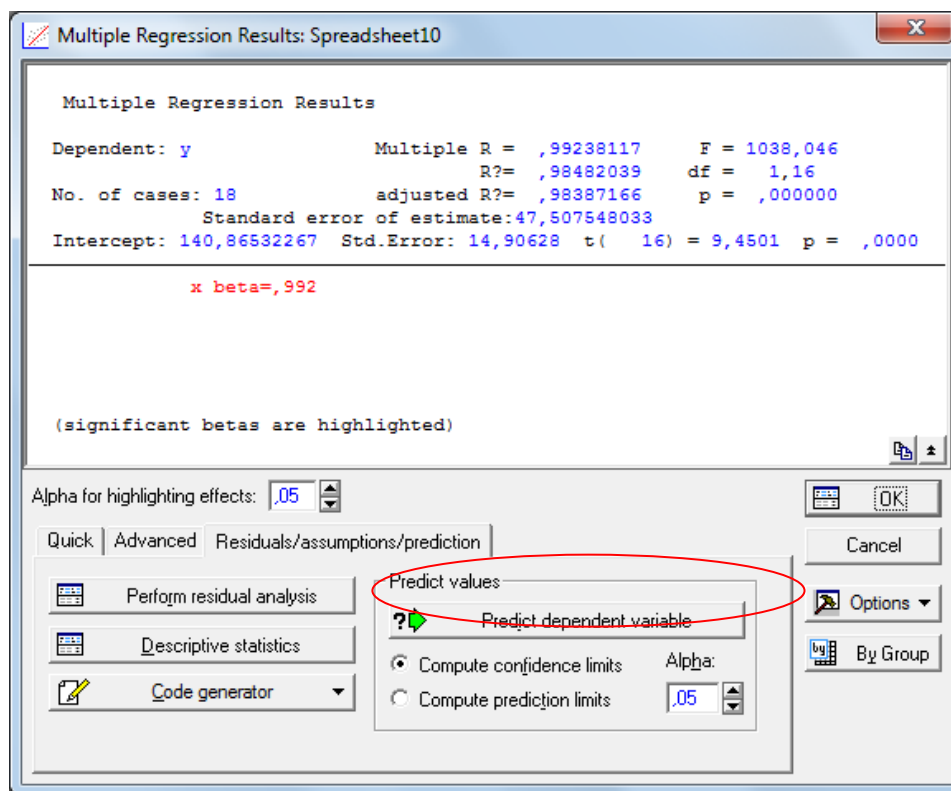


Рис. 49. Опція Predict dependent variable

Ініціювавши відповідну опцію, необхідно вказати значення обсягу прямих інвестицій на 2012 рік (рис. 50).

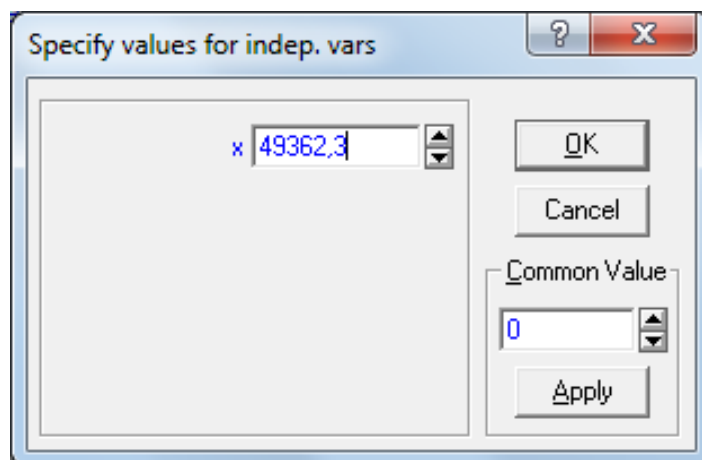


Рис. 50. Прогнозне значення обсягу прямих інвестицій

Результати прогнозування представлені у вигляді таблиці в якій додатково подано параметри моделі та значення інтервалу прогнозу (рис. 51).

Predicting Values for (Spreadsheet10) variable: y			
Variable	B-Weight	Value	B-Weight * Value
x	0,024846	49362,30	1226,469
Intercept			140,865
Predicted			1367,334
-95,0%CL			1302,958
+95,0%CL			1431,711

Рис. 51. Прогноз ВВП України на 2012 рік

Прогнозне значення ВВП (*Predicted*) = 1367,334; довірчий інтервал прогнозних значень: 1302,958 < y < 1431,711.

Висновок. У роботі був проведений аналіз однофакторної лінійної економетричної моделі залежності ВВП від прямих іноземних інвестицій в Україну. Побудована модель є адекватною і може бути використана для побудови прогнозу та оцінки впливу екзогенної змінної.

## Лабораторна робота № 5. Побудова множинної регресії

**Мета** – опанування студентами навичок регресійної моделі.

**Завдання.** Необхідно перевірити наявність лінійного множинного зв'язку між ВВП та соціально-економічних показниками (в табл. 9) наведено значення показників для України за 1998 – 2011 рр.). З цією метою необхідно провести такі розрахунки:

- Побудувати лінійну багатофакторну економетричну модель впливу соціально-економічних показників на ВВП і визначити всі її характеристики (параметри моделі, середнє квадратичне відхилення параметрів моделі, дисперсію і середнє квадратичне відхилення помилок моделі, коефіцієнти множинної кореляції і детермінації).

- Перевірити статистичну значущість параметрів моделі, коефіцієнта множинної кореляції. Перевірити адекватність моделі за критерієм Фішера.

- Розрахувати теоретичні значення залежної змінної і помилки моделі. Побудувати графік лінійної функції з довірчими інтервалами. Розрахувати прогнозне значення залежної змінної і довірчі інтервали зміни, якщо відомо значення незалежних показників.

- Зробити висновки щодо адекватності побудованої багатofакторної моделі, дати економічну інтерпретацію моделі в цілому.

Таблиця 9.

**Вихідні данні для побудови багатofакторної  
економетричної моделі**

Роки	Випуск продукції тис. грн (X1)	Обсяг роздрібногo товарообороту підприємств (юридичних осіб), млн грн (X2)	Сукупні витрати в середньому за місяць у розрахунку на одне домогосподарс- тво, грн (X3)	Прямі інвести- ції (млн дол.) (X4)	ВВП (млрд грн) (Y)
1998	22 6358	19 317	395,6	2 063,6	186,5
1999	35 6842	22 151	426,5	2 810,7	192,5
2000	37 3893	28 757	541,3	3 281,8	198,9
2001	46 0520	34 417	607	3 875	221,6
2002	50 4008	39 691	658,3	4 555,3	225,8
2003	60 3704	49 994	736,8	5 471,8	267,3
2004	80 9988	67 556	903,5	6 794,4	345,1
2005	99 5630	94 332	1 229,4	9 047	441,5
2006	1 182 179	129 952	1 442,8	16 890	544,2
2007	1 565 055	178 233	1 722	21 607,3	720,7
2008	2 072 172	246 903	2 590,4	29 542,7	948,1
2009	1 955 685	230 955	2 754,1	35 616,4	913,3
2010	2 388 289	280 890	3 072,7	40 053	1 082,6
2011	2 496 365	350 059	3 456	44 806	1 316,6

**Методичні рекомендації**

1. Відповідно до алгоритму побудови однофакторної моделі (рис. 40), були проведені розрахунки і для багатofакторної моделі (рис. 53).

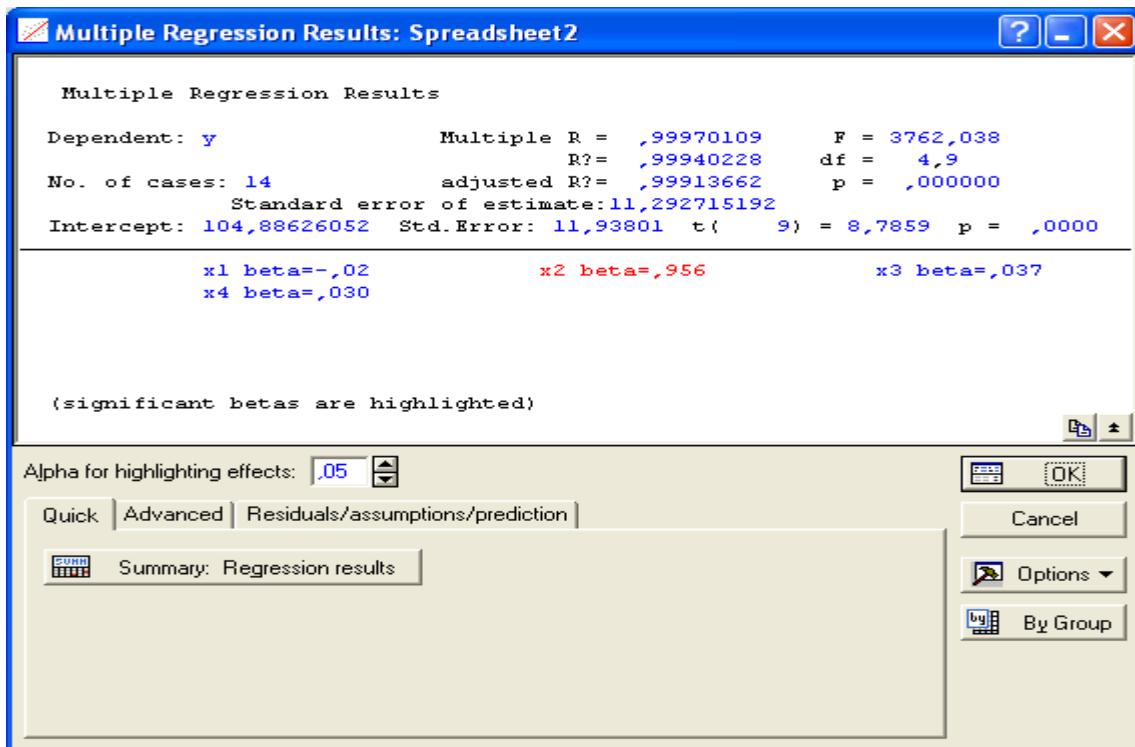


Рис. 52. Діалогове вікно модулю Multiple Regression

2.3 метою визначення параметрів та якості моделі необхідно ініціювати кнопку Summary: Regression results (Результати регресійного аналізу). Результати розрахунків наведено на рис. 53.

Regression Summary for Dependent Variable(Spreadsheet2)						
R= ,99970109 R <sup>2</sup> = ,99940228 Adjusted R <sup>2</sup> = ,99913662						
F(4,9)=3762,0 p<,00000 Std.Error of estimate: 11,293						
N=14	Beta	Std.Err. of Beta	B	Std.Err. of B	t(9)	p-level
Intercept			104,886	11,9380	8,7859	0,00001
x1	-0,02260	0,07362	-0,000	0,0000	-0,3069	0,76583
x2	0,95560	0,08281	0,003	0,0002	11,5391	0,00000
x3	0,03676	0,11198	0,013	0,0401	0,3283	0,75018
x4	0,03026	0,08187	0,000	0,0020	0,3696	0,72020

Рис. 53. Результат побудови багатофакторної економетричної моделі

Отримані результати можна інтерпретувати таким чином:

коєфіцієнт множинної кореляції дорівнює 0,997 (R). Вимірюється коєфіцієнт від -1 до +1. Оскільки значення коєфіцієнта сильно наближене до 1, то можемо говорити про адекватність моделі;

коєфіцієнт детермінації моделі дорівнює 0,999 (R<sup>2</sup>) Даний коєфіцієнт показує яка частка даних, побудованих за допомогою моделі

відповідає реальним даним. Оскільки коефіцієнт наближений до 1, то підтверджується адекватність моделі;

скоригований коефіцієнт детермінації на число спостережень і число параметрів дорівнює 0,999 (Adjusted R<sup>2</sup>);

критерій адекватності Фішера F (4,9) = 3762 отримане значення більше табличного, що підтверджує адекватність моделі;

середнє квадратичне відхилення помилок моделі складає 11,93;

вектор параметрів моделі має такий вигляд  $V(a_0, a_1, a_2, a_3, a_4) = (104,88; -0,001; 0,033; 0,00132; 0,0008)$ . Таким чином можна сформулювати загальний вид моделі:

$$Y = 104,88 - 0,001X_1 + 0,033X_2 + 0,00132X_3 + 0,0008X_4;$$

вектор значень критерію Стюдента  $t(9) = (8,79; -0,3; 11,54; 0,37)$ , що визначає значимість параметрів моделі.

Виходячи з аналізу отриманих результатів дана модель в цілому адекватна і якісна, але параметри моделі при змінних  $X_1, X_3, X_4$  є не значимі.

Для визначення середнього та середньоквадратичного відхилення вибірок всіх змінних в меню аналізу помилок ініціюємо Descriptive statistics/Means & Standard deviations (описові статистики/середнє і середньоквадратичне відхилення) (рис. 54 – 55).

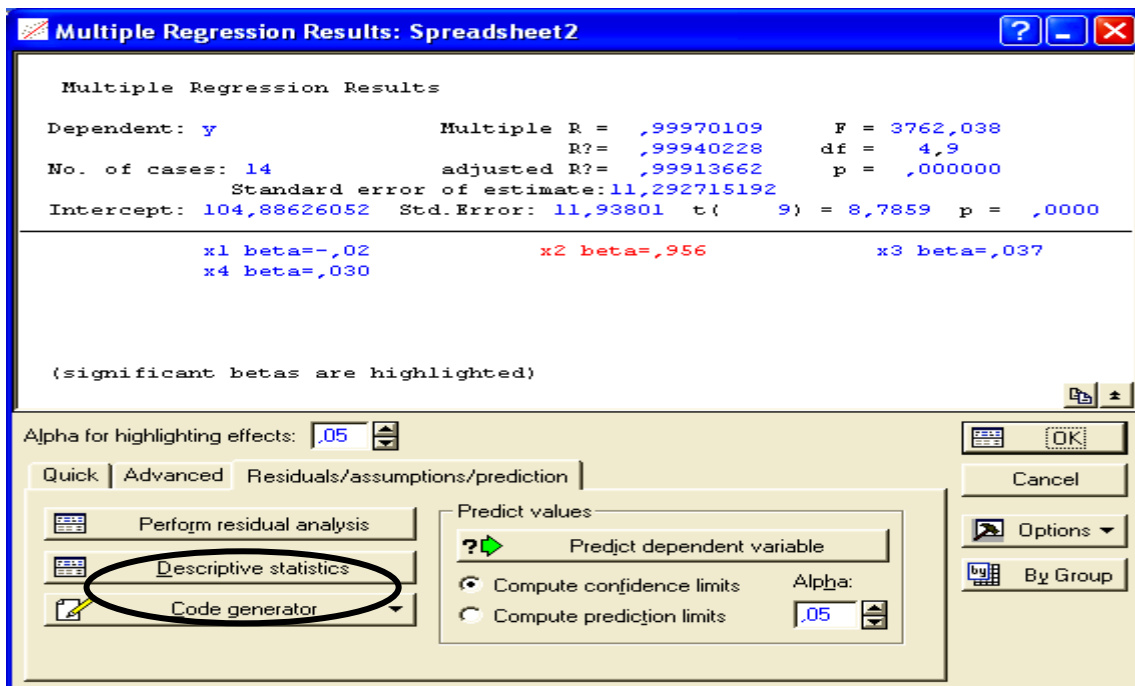


Рис. 54. Діалогове вікно аналізу результатів побудови багатофакторної моделі



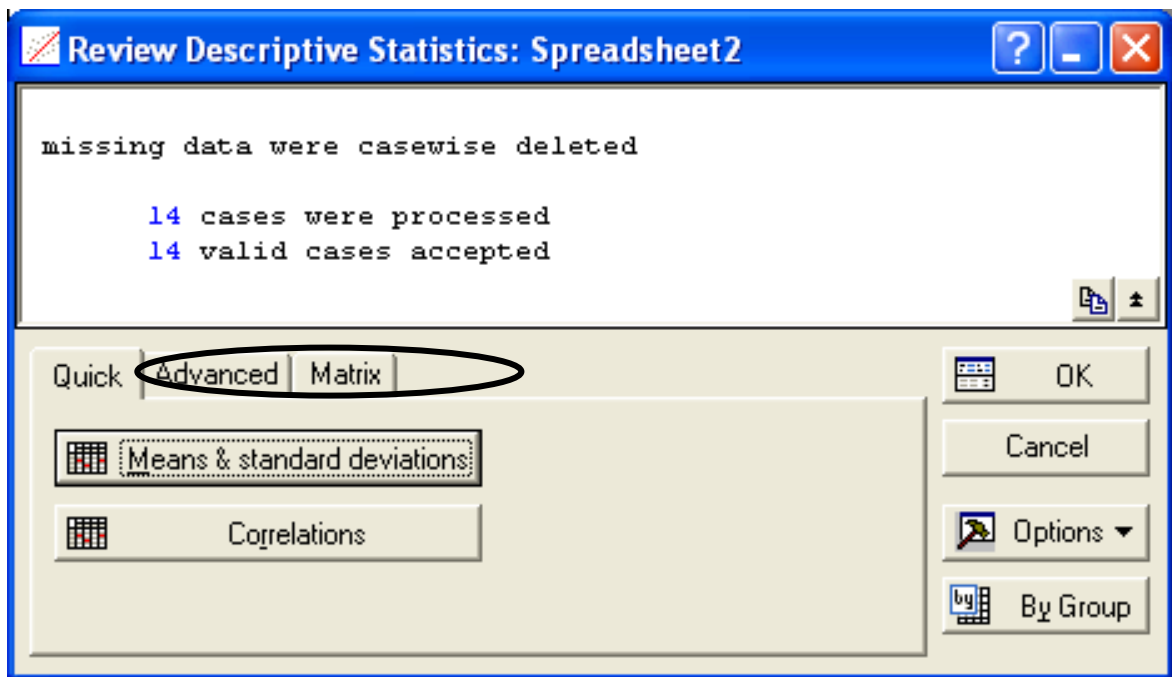


Рис. 55. **Діалогове вікно розрахунку описових статистик**

У результаті проведення розрахунків було оцінено середні значення, середнє квадратичне відхилення за екзогенними та ендогенними змінними (рис. 56).

Means and Standard Deviations (Spreadshe			
Variable	Means	Std.Dev.	N
x1	114219,	805422,	14
x2	126658,	111015,	14
x3	1467,	1070,	14
x4	16170,	15342,	14
y	543,	384,	14

Рис. 56. **Результати розрахунку описових статистик за ендогенною та екзогенними змінними**

3. За багатофакторною моделлю аналіз помилок проводиться за тим же алгоритмом, що і для однофакторної моделі (див. рис. 44, 45). На рис. 57 наведено результати розрахунку теоретичних значень за моделлю та помилок моделі.

Case No.	Predicted & Residual Values (Spreadsheet2)								
	Observed Value	Predicted Value	Residual	Standard Pred. v.	Standard Residual	Std.Err. Pred.Val	Mahalanobis Distance	Deleted Residual	Cook's Distance
1	186,50	173,13	13,366	-0,96316	1,1836	6,2985	3,1155	19,401	0,18365
2	192,50	182,07	10,423	-0,93989	0,9230	5,1640	1,7899	13,180	0,05697
3	198,90	205,61	-6,718	-0,87862	-0,5949	4,7208	1,3432	-8,140	0,01816
4	221,60	224,72	-3,125	-0,82889	-0,2767	4,1381	0,8170	-3,609	0,00274
5	225,80	242,89	-17,096	-0,78159	-1,5139	4,0156	0,7152	-19,570	0,07595
6	267,30	277,63	-10,335	-0,69117	-0,9152	3,7624	0,5145	-11,626	0,02353
7	345,10	336,71	8,387	-0,53741	0,7427	5,0407	1,6616	10,475	0,03428
8	441,50	429,29	12,201	-0,29643	1,0804	8,0138	5,6181	24,580	0,47717
9	544,20	553,88	-9,686	0,02783	-0,8577	4,2238	0,8901	-11,261	0,02782
10	720,70	716,74	3,959	0,45170	0,3506	9,3110	7,9093	12,367	0,16307
11	948,10	955,92	-7,819	1,07422	-0,6924	7,2237	4,3908	-13,235	0,11242
12	913,30	911,18	2,118	0,95778	0,1876	9,1526	7,6110	6,175	0,03929
13	1082,60	1079,27	3,320	1,39529	0,2941	7,6635	5,0583	6,158	0,02738
14	1316,60	1315,59	1,001	2,01037	0,0886	10,6181	10,5647	8,639	0,10350
Minimum	186,50	173,13	-17,096	-0,96316	-1,5139	3,7624	0,5145	-19,570	0,00274
Maximum	1316,60	1315,59	13,366	2,01037	1,1836	10,6181	10,5647	24,580	0,47717
Mean	543,19	543,19	-0,000	-0,00000	-0,0000	6,3819	3,7142	2,395	0,09614
Median	393,30	383,00	1,560	-0,41692	0,1381	5,7313	2,4527	6,166	0,04813

Рис. 57. Результати розрахунку помилок моделі

З рис. 57 видно, що найбільше значення помилки моделі спостерігається у 2002 році. Можна зробити висновок що в цей період розвиток економіки країни істотно відрізнялося від усього аналізованого періоду.

На рис. 58. наведено полігон розподілу помилок моделі.

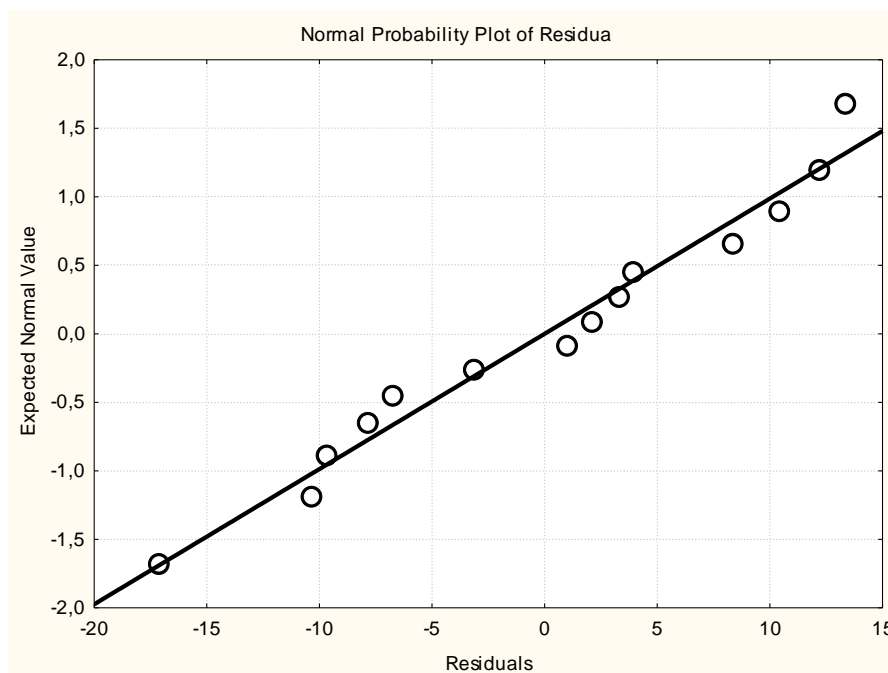


Рис. 58. Полігон розподілу помилок

На рис. 59 наведено гістограму розподілу помилок моделі.

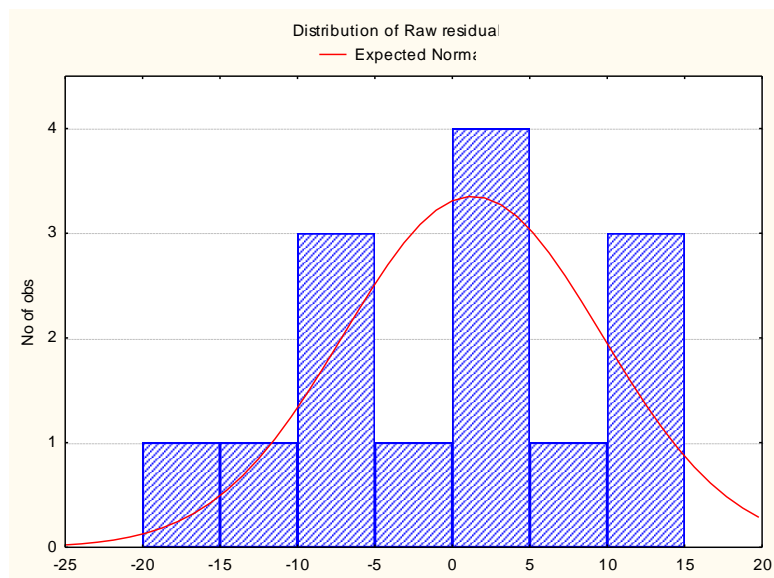


Рис. 59. Гістограма розподілу помилок

Розрахунки наведені рис. 58 та 59 доводять гіпотезу про нормальний закон розподілу помилок моделі.

Висновок. Проведені розрахунки показали, що модель є якісною, але всі параметри моделі не значимі. Таким чином можна зробити припущення про наявність мультиколінеарності в моделі і про доцільність побудови прогнозу лише після її усунення.

### Лабораторна робота № 6. Перевірка моделі на наявність мультиколінеарності та її усунення

**Мета** – опанування навичок перевірки моделі на мультиколінеарність та усунення її наслідків.

**Завдання.** Перевірити модель розроблену в лабораторній роботі № 5 (рис. 51, 52) на мультиколінеарність та усунути її у разі необхідності.

1. Для всебічної перевірки наявності мультиколінеарності в моделі доцільно використовувати алгоритм Ферара – Глобера. Всі розрахунки за алгоритмом доцільно проводити в пакеті MSExcel.

1.1. Першим кроком алгоритму є нормалізація вихідних даних за формулою:

$$z_i = \frac{x_i - \bar{x}_i}{\delta}$$

У результаті отримаємо матрицю нормалізованих даних (табл. 10).

Таблиця 10.

### Нормовані значення вихідних показників моделі

	Zi1	Zi2	Zi3	Zi4	Yi
1	-1,13709	-0,9669	-1,00042	-0,91958	-0,9281
2	-0,97508	-0,94137	-0,97157	-0,87088	-0,91249
3	-0,95391	-0,88186	-0,86436	-0,84018	-0,89584
4	-0,84635	-0,83088	-0,80301	-0,80152	-0,83677
5	-0,79236	-0,78337	-0,7551	-0,75718	-0,82585
6	-0,66858	-0,69057	-0,68179	-0,69744	-0,71786
7	-0,41246	-0,53237	-0,52612	-0,61124	-0,51543
8	-0,18197	-0,29118	-0,22178	-0,46442	-0,2646
9	0,049647	0,029675	-0,02249	0,046765	0,002621
10	0,52502	0,464577	0,238239	0,354225	0,461868
11	1,154649	1,083138	1,049196	0,871432	1,053555
12	1,010021	0,939482	1,202068	1,267298	0,963007
13	1,547135	1,389283	1,499593	1,556463	1,40352

3.1. За допомогою вбудованої функції КОРРЕЛ необхідно розрахувати матрицю парних кореляцій за нормованими даними. Розрахована матриця має такий вигляд.

	1,000000	0,997501	0,991250	0,981528	0,996972
	0,997501	1,000000	0,992652	0,987279	0,999509
r=	0,991250	0,992652	1,000000	0,992396	0,993151
	0,981528	0,987279	0,992396	1,000000	0,988197
	0,996972	0,999509	0,993151	0,988197	1,000000

З матрицею коефіцієнтів парної кореляції знаходимо його визначник як наведено на рис. 60.

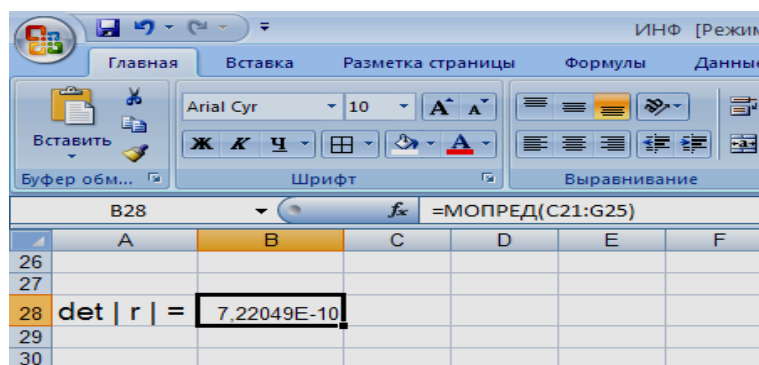


Рис. 60. Розрахунок визначника матриці коефіцієнтів парних кореляцій

1.2. Дослідження загальною мультиколінеарності моделі проводиться з використанням критерію  $\chi^2$

$$\chi^2 = -\left[ n - 1 - \frac{1}{6}(2m + 5) \ln|r| \right] = 57,6$$

$\chi^2(\alpha=0,05; k=4)=7,8 \Rightarrow |\chi^2_{\text{расч}}| > \chi^2(\alpha=0,05; k=4)$  – у моделі присутня загальна мультиколінеарність.

1.3. Необхідно розрахувати коефіцієнт Фішера за формулою

$$F_k = (c_{kk} - 1) * \frac{n - m}{m - 1}$$

Розрахунок елементів матриці с, як оберненої до матриці r наведено на рис. 61.

	J	K	L	M	N
63		268,6290252	-245,8955647	-83,4925941	76,573787
64		-245,8955647	1256,543174	53,33051964	-40,276342
65	$c=r^{(-1)}$	-83,4925941	53,33051964	143,7614884	-79,284927
66		76,57378652	-40,27634178	-79,28492728	90,922116
67					
68					

Рис. 61. Розрахунок елементів матриці С

$$F_1 = (c_{11} - 1) \times \frac{n - m}{m - 1} = (268,6 - 1) \times \frac{13 - 4}{4 - 1} = 802,1.$$

$$F_1 = (c_{22} - 1) \times \frac{n - m}{m - 1} = (1256,5 - 1) \times \frac{13 - 4}{4 - 1} = 3766.$$

$$F_1 = (c_{33} - 1) \times \frac{n - m}{m - 1} = (143,7 - 1) \times \frac{13 - 4}{4 - 1} = 428,1.$$

$$F_1 = (c_{44} - 1) \times \frac{n - m}{m - 1} = (90,92 - 1) \times \frac{13 - 4}{4 - 1} = 272.$$

Кожен з отриманих коефіцієнтів Фішера порівнюємо з табличним значенням для  $\alpha = 0,05$ ,  $k_1 = 9$ ,  $k_2 = 3$ , який становить  $F = 19,4$ . Якщо розрахункове значення перевищує табличне, то  $k$  змінна викликає мультиколінеарність, таким чином бачимо, що змінні моделі викликають мультиколінеарність.

1. Для визначення попарної мультиколінеарності використовується коефіцієнт Стюдента.

Далі визначимо приватні коефіцієнти кореляції, вони характеризують тісноту зв'язку між двома змінними за умови, що інші змінні не впливають.

$$r_{kj} = \frac{-c_{kj}}{\sqrt{c_{kk} * c_{jj}}}$$

$$r_{12} = 0,42$$

$$r_{13} = 0,43$$

$$r_{14} = -0,49$$

$$r_{23} = -0,13$$

$$r_{24} = 0,12$$

$$r_{34} = 0,69$$

Розрахуємо значимість коефіцієнтів приватної кореляції за критерієм Стюдента

$$t_{kj} = r_{kj} * \frac{\sqrt{n-m}}{\sqrt{1-r^2}}$$

Для того щоб зробити висновок про наявність мультиколінеарності необхідно порівняти отримані значення з табличними. Таким чином, проаналізувавши модель на мультиколінеарності різними методами, можна зробити висновок про те, що в моделі присутня мультиколінеарність. Це обумовлено наявністю зв'язку різного ступеня між різними ознаками.

2. Для позбавлення від мультиколінеарності використаємо методи покроково включення та виключення змінних.

У модулі Multiple Regression реалізовані методи покрокового включення параметрів (Forward stepwise) і виключення параметрів (Backward stepwise). Вибір методів здійснюється на стартовій панелі в меню Advanced (рис. 62).

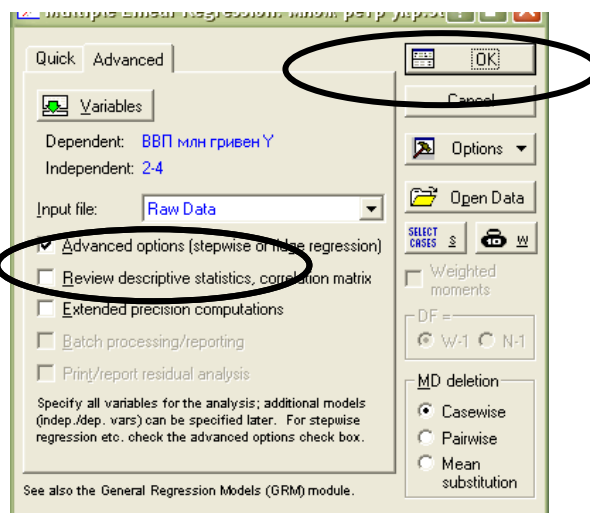


Рис. 62. Діалогове вікно модулю Multiple Regression

Першим етапом обираємо метод покрокового виключення параметрів (Backward stepwise) (рис. 63).

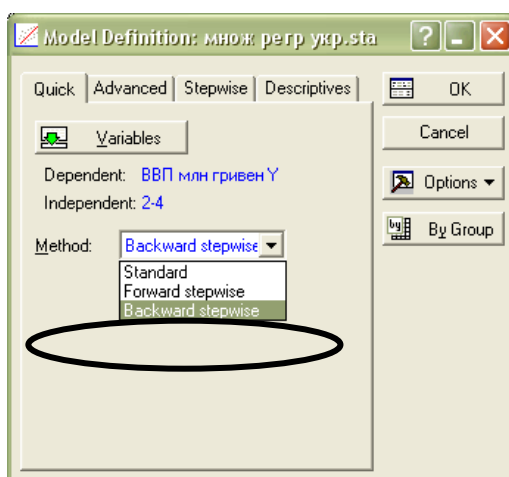


Рис. 63. Діалогове вікно вибору методу побудови багатофакторної моделі

На рис. 64 показано результати побудови багатофакторної моделі методом покрокового виключення.

Regression Summary for Dependent Variable(Spreadsheet2)						
R= ,99967377 R?= ,99934765 Adjusted R?= ,99929329						
F(1,12)=18383, p<0,0000 Std.Error of estimate: 10,217						
N=14	Beta	Std.Err. of Beta	B	Std.Err. of B	t(12)	p-level
Intercept			104,860	4,23176	24,779	0,00000
x2	0,99967	0,00737	0,003	0,00002	135,584	0,00000

Рис. 64. Багатофакторна модель побудована методом покрокового виключення

Реалізація методу покрокового включення проводиться за рахунок вибору відповідного пункту меню (рис. 65). У результаті реалізації розрахункових алгоритмів отримуємо комплекс економетричних моделей.

Regression Summary for Dependent Variable(Spreadsheet2)						
R= ,99967377 R <sup>2</sup> = ,99934765 Adjusted R <sup>2</sup> = ,99929329 F(1,12)=18383, p<0,0000 Std.Error of estimate: 10,217						
N=14	Beta	Std.Err. of Beta	B	Std.Err. of B	t(12)	p-level
Intercept			104,860	4,23176	24,779	0,00000
x2	0,99967	0,00737	0,003	0,00002	135,584	0,00000

Рис. 65. Економетрична модель побудована методом покрокового включення

Проведений аналіз усіх побудованих моделей дозволив зробити висновок, що наявність мультиколінеарності спровокована помилками специфікації, тому доцільно провести додатковий аналіз моделі чи використати методи згортання ознакового простору.

## Лабораторна робота № 7. Побудова регресійної моделі з фіктивними змінними зсуву та нахилу

**Мета** – отримання навичок побудови моделей з фіктивними змінними з використанням пакетів Statistica 8.0.

**Завдання.** Побудувати модель, що враховує вплив соціально-економічних показників розвитку регіонів України на обсяг реалізованої промислової продукції з урахуванням типу регіону. Вихідні дані наведені в табл. 11.

Таблица 11

### Показники соціально-економічного розвитку регіону

	Тип регіону	Обсяг реалізованої промислової продукції (робіт, послуг) на одну особу (Y)	Рівень середньомісячної заробітної плати (X1)	Рівень зареєстрованого безробіття (X2)
1	2	3	4	5
Республіка Крим	сільск.	7 268,5	1 707	2,1
Вінницька	сільск.	8 711,3	1 511	4,3
Волинська	сільск.	6 319,4	1 427	4,1



1	2	3	4	5
Дніпропетровська	промисл.	33 022,9	1 963	2,5
Донецька	промисл.	31 306,6	2 116	2
Житомирська	промисл.	7 790,2	1 493	4,3
Закарпатська	сільск.	4 593,1	1 562	3
Запорізька	промисл.	29 183,5	1 843	3,2
Івано-Франківська	сільск.	8 081,9	1 627	3,7
Київська	промисл.	17 640,6	1 987	2,6
Кіровоградська	сільск.	7 557,8	1 537	4,2
Луганська	промисл.	24 969,1	1 873	2,4
Львівська	сільск.	8 492,9	1 667	2,7
Миколаївська	сільск.	14 115,5	1 806	3,5
Одеська	сільск.	10 983,9	1 787	2
Полтавська	промисл.	25 876,5	1 733	4,7
Рівненська	сільск.	7 806,5	1 614	4,7
Сумська	промисл.	10 444,8	1 593	4,2
Тернопільська	сільск.	4 718,5	1 412	5,1
Харківська	промисл.	16 033,4	1 804	2,7
Херсонська	сільск.	7 348,9	1 482	3,1
Хмельницька	сільск.	7 535,4	1 521	4,3
Черкаська	сільск.	13 777,2	1 532	4,9
Чернівецька	сільск.	3 324,6	1 523	3,4
Чернігівська	сільск.	9 343,3	1 465	4,7
м. Київ	промисл.	41 384,6	3 161	0,6
м. Севастополь	промисл.	7 865,8	1 882	0,9

### Методичні рекомендації

1. Побудова моделі без урахування якісних змінних (рис. 66).

Regression Summary for Dependent Variable (Spreadsheet)						
R= ,83078285 R <sup>2</sup> = ,69020015 Adjusted R <sup>2</sup> = ,66438350 F(2,24)=26,735 p<,00000 Std.Error of estimate: 5895,0						
N=27	Beta	Std.Err. of Beta	B	Std.Err. of B	t(24)	p-level
Intercept			-46608,9	12430,3	-3,7496	0,00099
x1	1,02021	0,16654	30,3	4,94	6,1259	0,00000
x2	0,29230	0,16654	2466,9	1405,5	1,7551	0,09199

Рис. 66. Багатофакторна регресійна модель

З рис. 66. видно, що розроблена модель має задовільну якість – лише параметр при першій змінній є значимим. Тому доцільно в моделі урахувати якісну змінну, що відображала б тип регіону. Для цього доцільно використати фіктивні змінні.

## 2. Побудова моделі з урахуванням коефіцієнта нахилу.

Модель з урахування коефіцієнта нахилу має вигляд:

$$Y = a_0 + a_1X_1 + a_2X_2 + a_3d_1,$$

де  $d_1$  – фіктивна змінна нахилу, що приймає значення 0 для сільськогосподарських регіонів та 1 для промислових.

Використовуючи модуль Multiple Regression пакету Statistica 7.0, були розраховані параметри даної моделі рис. 67.

Regression Summary for Dependent Variable(Spreadsheet)						
R= ,88701946 R <sup>2</sup> = ,78680352 Adjusted R <sup>2</sup> = ,75899528						
F(3,23)=28,294 p<,00000 Std.Error of estimate: 4995,5						
N=27	Beta	Std.Err. of Beta	B	Std.Err. of B	t(23)	p-level
Intercept			-39951,0	10733,5	-3,7220	0,00111
x1	0,82407	0,15365	24,5	4,56	5,3632	0,00001
x2	0,30358	0,14117	2562,7	1191,4	2,1504	0,04226
d1	0,37207	0,11525	7561,4	2342,2	3,2282	0,00371

Рис. 67. Модель з урахуванням коефіцієнта нахилу

Наведена на рис. 67 модель є адекватною, всі параметри моделі значимі, що свідчить про наявність нахилу у вибірці даних.

Висновок. У роботі був розглянутий один з випадків, коли використання фіктивної змінної дозволяє урахувати певні якісні особливості досліджуваного процесу. А значить дозволяє розробити якісну економетричну модель, що відображає наявні зв'язки між соціально-економічними показниками.

## Рекомендована література

1. Лабораторний практикум з навчальної дисципліни "Економетрика II" для студентів напряму підготовки "Прикладна статистика" денної форми навчання / укл. О. В. Раєнєва, О. І. Горохова, І. В. Чанкіна. – Х. : Вид. ХНЕУ, 2011. – 52 с.
2. Моделі та методи соціально-економічного прогнозування : підручник / Геєць В. М., Клебанова Т. С., Черняк О. І. та ін. – Х. : ВД "ІНЖЕК", 2005. – 396 с.
3. Шашок А. В. Лабораторный практикум в системе "STATISTICA" : Методические указания к выполнению лабораторных работ по дисциплине "Статистика" для студентов всех форм обучения специальности 080507 "Менеджмент организации". – Рубцовск, 2010. – 81 с.
4. Экономическое прогнозирование : учебник. – Курск, 2009. – 92 с.
5. Офіційний сайт асоціації Українських банків. – Режим доступу : [www.aub.org.ua](http://www.aub.org.ua).
6. Офіційний сайт компанії Statsoft. – Режим доступу : [www.statsoft.ru](http://www.statsoft.ru).
7. Портал Статосфера. – Режим доступа : [www.statosphere.ru/videos/62-dmsess11.html](http://www.statosphere.ru/videos/62-dmsess11.html).

НАВЧАЛЬНЕ ВИДАННЯ

Лабораторний практикум  
з навчальної дисципліни  
**"СТАТИСТИЧНЕ МОДЕЛЮВАННЯ  
ТА ПРОГНОЗУВАННЯ"**

для студентів напряму підготовки  
6.030506 "Прикладна статистика"  
денної форми навчання

Укладачі: **Раєвська** Олена Валентинівна  
**Стрижченко** Костянтин Анатолійович  
**Чанкіна** Ірина Володимирівна та ін.

Відповідальний за випуск **Раєвська О. В.**

Редактор **Пушкар І. П.**

Коректор **Бриль В. О.**

План 2013 р. Поз. № 116.

Підп. до друку Формат 60 × 90 1/16. Папір MultiCopy. Друк Riso.

Ум.-друк. арк. 3,75. Обл.-вид. арк. 4,69. Тираж прим. Зам. №

---

Видавець і виготівник – видавництво ХНЕУ, 61166, м. Харків, пр. Леніна, 9а

---

*Свідоцтво про внесення до Державного реєстру суб'єктів видавничої справи  
Дк № 481 від 13.06.2001 р.*