

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

**ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ ЕКОНОМІЧНИЙ УНІВЕРСИТЕТ
ІМЕНІ СЕМЕНА КУЗНЕЦЯ**

**Лабораторний практикум
з навчальної дисципліни
"АНАЛІЗ ТА ПРОГНОЗУВАННЯ
РЯДІВ ДИНАМІКИ"**

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

Харків. Вид. ХНЕУ ім. С. Кузнеця, 2014

Затверджено на засіданні кафедри статистики та економічного прогнозування.

Протокол № 9 від 21.12.2013 р.

Укладачі: Раєвнєва О. В.

Чанкіна І. В.

Бровко О. І.

Л12 Лабораторний практикум з навчальної дисципліни "Аналіз та прогнозування рядів динаміки" для студентів напряму підготовки 6.030506 "Прикладна статистика" денної форми навчання / укл. О. В. Раєвнєва, І. В. Чанкіна, О. І. Бровко. – Х. : Вид. ХНЕУ ім. С. Кузнеця, 2014. – 60 с. (Укр. мов.)

Подано методичні рекомендації до виконання лабораторних робіт, метою яких є закріплення й поглиблення знань теоретичного та практичного матеріалу, набуття навичок прогнозування динамічних рядів за допомогою пакета STATISTICA та MS Excel. Рекомендовано для студентів напряму підготовки 6.030506 "Прикладна статистика".

Вступ

Лабораторний практикум призначений для засвоєння студентами теоретичних та практичних знань із дисципліни "Аналіз та прогнозування рядів динаміки", набуття навичок роботи за допомогою використання пакетів прикладних програм Statistica 8.0, MS Excel. Пакети містять множини статистичних методів, що підтримують вирішення різних економіко-статистичних задач та дозволяють реалізувати спектр економіко-математичних методів і моделей аналізу та прогнозування динамічних рядів. MS Excel – надає можливість відчувати всі деталі і тонкощі досліджуваних методів, що підвищує рівень засвоєння навчального матеріалу. ППП STATISTICA 8.0 містить комплекс статистичних методів, що підтримують рішення різних статистичних та економіко-математичних задач. Під час розробки лабораторних робіт передбачається, що студент вміє працювати в запропонованих пакетах прикладних програм.

Навчальна дисципліна складається з трьох змістових модулів: методологічні основи статистичного моделювання і прогнозування, моделювання взаємозв'язків економічних процесів, моделювання часових процесів та моделювання і прогнозування багатовимірних процесів. Кожна лабораторна робота розглянута на прикладі рішення конкретної задачі й забезпечена докладними коментарями й рисунками. Лабораторні роботи стосуються основних тем навчальної дисципліни й ґрунтуються на теоретичному матеріалі відповідної теми, а також попередніх тем. Кожна робота містить мету й завдання та методичні рекомендації щодо їх виконання.

За виконання й захист лабораторної роботи виставляється відведена кількість балів, відповідно до технологічної карти з навчальної дисципліни. Певна увага приділяється оформленню якісного звіту з лабораторної роботи, знанню теоретичного матеріалу, володінню пакетами прикладних програм, а також повноті економічної інтерпретації отриманих результатів.

**Компетентності,
що отримують студенти в процесі вивчення навчальної
дисципліни "Аналіз та прогнозування часових рядів"
(згідно з Національною рамкою кваліфікацій)**

№ п/п	Зміст компетентності	Теми дисципліни, що формують компетентність
1	2	3
1. Здатність набувати теоретичні знання з аналізу часових рядів, що описують поведінку соціально-економічних процесів		
Знання	Знання основних категорій аналізу часових рядів, вимог до статистичної інформації, що формує часовий ряд, методів визначення та аналізу аномальних спостережень, основних складових часових рядів, методів визначення присутності складових у часовому ряду	Тема 1. Основні категорії аналізу часових рядів Тема 2. Методи дослідження та виміру стійкості рівнів часового ряду Тема 3. Компонентний аналіз часового ряду Тема 4. Методи згладжування часових рядів
Вміння	Проводити декомпозицію часового ряду; визначати наявність еволюторної компоненти в ряду; здійснювати згладжування часового ряду; здійснювати усунення аномальних спостережень у часовому ряду; визначати стійкість часового ряду за допомогою критеріїв стійкості тенденції	
Комунікація	Розвиток креативного мислення під час вирішення поставлених завдань. Навички толерантного ставлення до іншої думки під час вирішенні завдання. Здатність до критики та самокритики	
Автономність і відповідальність	Здібність виділяти серед різноманітних пропозицій щодо вирішення проблеми інформацію, яка дозволяє це здійснити. Здатність до розвитку креативного мислення під час вирішення поставлених завдань. Прагнення до неперервного особистісного та професійного вдосконалення	
2. Здатність до моделювання компонентів часового ряду		
Знання	Поняття еволюторної складової часового ряду; знання методів визначення наявності тренда в ряду; знання лінійних та нелінійних видів трендових моделей; знання емпіричних методів вибору форми тренда	Тема 5. Моделювання трендової компоненти

1	2	3
Вміння	Здійснювати вибір форми тренда на підставі теоретичного аналізу та емпіричних методів. Використовувати дисперсійний аналіз для обґрунтування форми тренда	
Комунікація	Розвиток креативного мислення під час вирішення поставлених завдань. Навички толерантного ставлення до іншої думки під час вирішення завдання. Здатність до критики та самокритики	
Автономність і відповідальність	Здібність виділяти серед різноманітних пропозицій щодо вирішення проблеми інформацію, яка дозволяє це здійснити. Здатність до розвитку креативного мислення під час вирішення поставлених завдань. Прагнення до неперервного особистісного та професійного вдосконалення	
Знання	Знання розкладання часового ряду; знання методів визначення сезонних коливань; поняття циклічної складової часового ряду; знання спектрального аналізу та його основних категорій	Тема 6. Моделювання періодичних компонент часового ряду
Вміння	Виявляти сезонні коливання в часових рядах за допомогою методу абсолютних різниць, методів відношення середніх, методу абсолютних та відносних величин на підставі медіани. Моделювати сезонність соціально-економічних процесів. Визначати циклічний компонент часового ряду. Моделювати циклічний компонент за допомогою розкладання в ряд Фур'є	
Комунікація	Розвиток креативного мислення під час вирішення поставлених завдань. Навички толерантного ставлення до іншої думки під час вирішення завдання. Здатність до критики та самокритики	
Автономність і відповідальність	Здібність виділяти серед різноманітних пропозицій щодо вирішення проблеми інформацію, яка дозволяє це здійснити. Здатність до розвитку креативного мислення під час вирішення поставлених завдань. Прагнення до неперервного особистісного та професійного вдосконалення	

1	2	3
Знання	Знання завдань дослідження випадкового компонента часового ряду; знання методів перевірки випадковості залишків моделі; поняття стаціонарного процесу та методів приведення нестаціонарного процесу до стаціонарного	Тема 7. Моделювання випадкового компонента
Вміння	Використовувати методи інтеграції для побудови стаціонарного процесу. Оцінювати гіпотезу про стаціонарність залишків на підставі коефіцієнта кореляції, критерію серій, критерію мінімуму та максимуму. Оцінювати нормальність розподілу залишків на підставі коефіцієнтів асиметрії, ексцесу та RS-критерію	
Комунікація	Навички командної роботи та розробки колективного рішення. Здатність до критики та самокритики	
Автономність і відповідальність	Здатність до самостійного пошуку необхідної інформації для дослідження соціально-економічних процесів, що віддзеркалюються у вигляді часових рядів. Прагнення до неперервного особистісного та професійного вдосконалення	
3. Здатність до моделювання зв'язаних часових рядів		
Знання	Знання передумов наявності автокореляції в часовому ряду; класифікації та наслідків автокореляції; знання методів визначення автокореляції в ряду; знання методів усунення автокореляції; поняття автокореляційної та частково автокореляційної функцій; поняття авто регресійних моделей та моделей із розподіленими лагами	Тема 8. Моделі зв'язаних часових рядів (автокореляція) Тема 9. Моделі часових рядів із лаговими змінними
Вміння	Оцінювати моделі з лагами в незалежних змінних. Визначати довжину лагів змінних моделі за допомогою методу послідовного збільшення та перетворення Койка. Будувати моделі адаптивних очікувань та часткового корегування. Будувати моделі з поліноміально розподіленими лагами Алмона. Здійснювати оцінку авторегресійних моделей	

1	2	3
Комунікація	Розвиток креативного мислення під час вирішення поставлених завдань. Навички толерантного ставлення до іншої думки при вирішенні завдання. Здатність до критики та самокритики	
Автономність і відповідальність	Здібність виділяти серед різноманітних пропозицій щодо вирішення проблеми інформацію, яка дозволяє це здійснити. Здатність до розвитку креативного мислення під час вирішення поставлених завдань. Прагнення до неперервного особистісного та зростання	
4. Здатність до прогнозування часових рядів за допомогою простих методів прогнозування та сучасних інтелектуальних технологій		
Знання	Знання основних категорій прогнозування часових рядів; знання методів екстраполяції рівнів часових рядів; знання методів прогнозування на основі темпів зростання; знання методів прогнозування на підставі кривих зростання	<p data-bbox="1134 1335 1417 1458">Тема 10. Прості методи прогнозування екстраполяції</p> <p data-bbox="1134 1570 1417 1693">Тема 11. Прогнозування на підставі кривих зростання</p>
Вміння	Прогнозувати соціально-економічні процеси, що представлені у вигляді часових рядів за допомогою методу середнього рівня, середнього абсолютного приросту, методу середнього темпу зростання. Здійснювати підбір тренда та прогнозувати часові ряди на підставі його екстраполяції. Прогнозувати часові ряди на підставі кривих зростання. Оцінювати якість та адекватність прогностичних моделей. Оцінювати точність прогнозу	
Комунікація	Розвиток креативного мислення під час вирішення поставлених завдань. Навички толерантного ставлення до іншої думки під час вирішення завдання. Здатність до критики та самокритики	

1	2	3
Автономність і відповідальність	Здібність виділяти серед різноманітних пропозицій щодо вирішення проблеми інформацію, яка дозволяє це здійснити. Прагнення до неперервного особистісного та професійного вдосконалення. Навички толерантного ставлення до іншої думки. Здатність до критики та самокритики	
Знання	Знання класифікації сучасних інтелектуальних технологій у прогнозуванні часових рядів; парадигми та засад побудови штучних нейронних мереж; класифікації штучних нейронних мереж та принципів їх відбору для прогнозування часових рядів	Тема 12. Прогнозування за допомогою штучних нейронних мереж
Вміння	Обирати адекватний вид штучної нейронної мережі відповідно до специфіки часового ряду, що досліджується; будувати штучну нейронну мережу відповідно до парадигми її навчання; прогнозувати соціально-економічний процес за допомогою штучних нейронних мереж	
Комунікація	Розвиток креативного мислення під час вирішення поставлених завдань. Навички толерантного ставлення до іншої думки під час вирішення завдання	

Лабораторна робота № 1

Перевірка динамічних рядів на нормальність розподілу величин. Показники асиметрії та ексцесу. Розрахунок параметричних та непараметричних критеріїв для рядів розподілу величин

Мета – закріплення теоретичних та практичних навичок перевірки закону розподілу динамічного ряду та перевірки гіпотез щодо наявності закономірностей у них у пакеті Statistica 8.0.

Завдання. З метою порівняння закономірностей функціонування переробної та добувної галузі промисловості, необхідно дослідити індекс промислової продукції за період з 2010 по 2013 рр. (табл. 2). Перевірити закон розподілу показників, а також гіпотезу про те, що розбіжність в індексі промислової продукції для галузей є випадковою.

Індекс промислової продукції з січня 2010 по листопад 2013 рр.

Період	Добувна промисловість	Переробна промисловість	Період	Добувна промисловість	Переробна промисловість
1	106,8	114,8	25	103,4	102,5
2	102,2	110,1	26	103,5	101,5
3	103,3	113,3	27	103,6	100,6
4	105,3	115,1	28	103,4	100,9
5	106,4	114,9	29	102,8	101,5
6	106,1	114,2	30	103	100,8
7	105	113,5	31	103,6	100,3
8	104,2	113,3	32	103,4	99,6
9	103,8	113,5	33	103	98,8
10	103,7	113,4	34	102,7	98,6
11	103,5	113,6	35	102,5	98,6
12	103,7	113,9	36	101,9	98
13	106,4	116,1	37	98,1	95,5
14	106	115,6	38	99,4	94,6
15	105,2	113,9	39	99,8	93,1
16	104,8	111,6	40	99,5	94,1
17	105,3	111,4	41	98,8	92,7
18	105,6	110,9	42	98,7	92,6
19	105,8	110,8	43	98,8	92,5
20	106,2	111,1	44	99,1	92,3
21	106,4	111	45	99,3	92,1
22	106,6	110,8	46	99,5	92
23	106,6	110,2	47	99,8	92
24	106,8	109,6			

Методичні рекомендації**1. Розрахунок основних статистичних показників**

Для того щоб виконувати обчислювальні процедури необхідно увійти в меню Statistics / Basic Statistics / Tables. У вікні необхідно вибрати пункт Descriptive Statistics (Описові статистики). У стартовій панелі модуля використовую кнопку Variable (змінні) необхідно задати вихідні параметри моделі. У даному випадку це стовпець Index (рис. 1).

Далі виділяємо основні статистики для розрахунків, як зображено на рис. 1. Такими є: Valid N (число спостережень), Mean (середнє), Sum (сума значень), Median (медіана), Mode (мода), Standard Deviation (середньоквадратичне відхилення), Variance (дисперсія), Std. err. of mean (помилка середнього), Skewness (коефіцієнт асиметрії), Std. err. of skewness (помилка коефіцієнта асиметрії), Kurtosis (коефіцієнт ексцесу), Std. err. of kurtosis, Minimum & maximum (мінімум і максимум вибірки), Range (розмах вибірки).

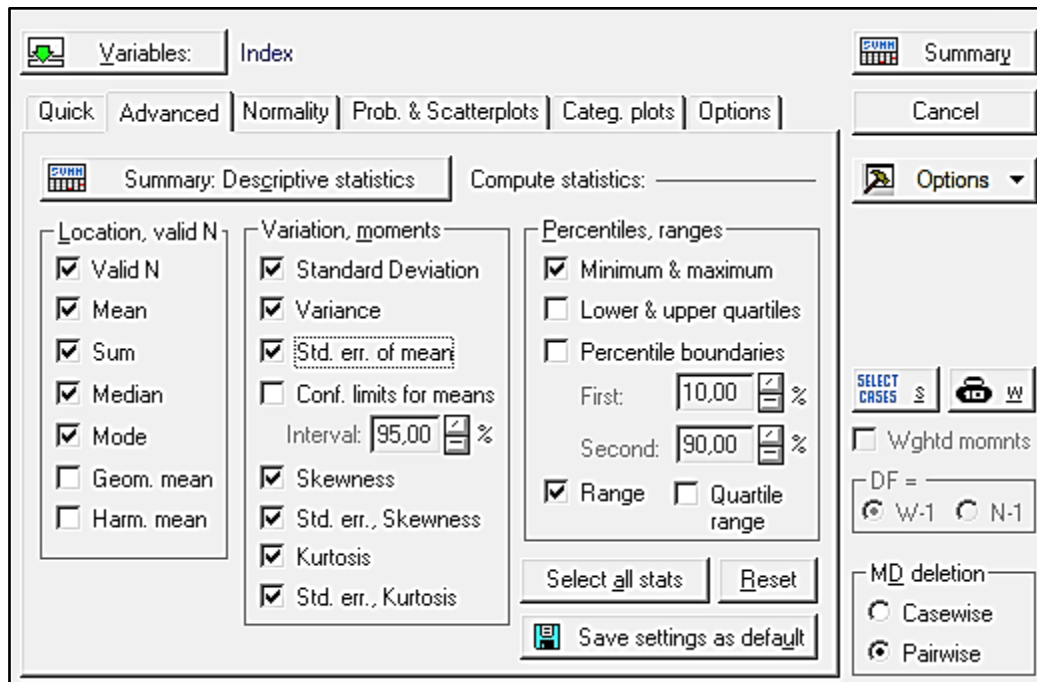


Рис. 1. Виділення основних описових статистик для розрахунку

Результат розрахунку основних описових статистик зображено на рис. 2.

Variable	Descriptive Statistics (Spreadsheet1_(Recovered))								
	Valid N	Mean	Median	Mode	Frequency of Mode	Sum	Minimum	Maximum	Variance
Добувна промисловість	47	103,261	103,500	Multipl	3	4853,30	98,1000	106,800	7,0450
Переробна промисловість	47	104,931	109,600	Multipl	2	4931,80	92,0000	116,100	74,4848

Std.Dev.	Coef.Var.	Skewness	Std.Err. Skewness	Kurtosis	Std.Err. Kurtosis
2,65424	2,57040	-0,49519	0,34657	-0,9197	0,68091
8,63045	8,22481	-0,24795	0,34657	-1,5881	0,68091

Рис. 2. Описові статистики

Виходячи з проведених розрахунків, можна зробити висновок, що середнє значення індексу промислової продукції переробної промисловості складає 104,9 %, а добувної – 103,3 %. За період із січня 2010 по листо-

пад 2013 рр. найбільше значення індексу промислової продукції для переробної галузі складав 116,1 %, а найменше – 92 %, а для добувної – 106,8 % та 98,1 % відповідно. Такі розрахунки свідчать про невелику амплітуду зміни значень показника середні що свідчить про невеликий розкид значень. Визначення форми розподілу проводиться за допомогою критеріїв:

- значення коефіцієнта асиметрії перевищує 0 у випадку правосторонньої асиметрії і менше 0 у випадку лівосторонньої. В обох вибірках спостерігається лівостороння асиметрія;
- значення коефіцієнта ексцесу перевищує 0 у випадку гостровершинного розподілу і менше 0 у випадку плосковершинного розподілу. Вибірки, що вивчаються, є плосковершинними, бо коефіцієнт ексцесу складає від'ємне значення.

2. Побудова полігону розподілу

Для цього необхідно в меню Graphs/2D Graphs/Line Plots (variables) і вибрати змінну.

З рис. 3 видно, що для обох показників спостерігається тенденція до значного зменшення значень із 1 січня 2012 р. Проте більш суттєвими є зміни для переробної промисловості.

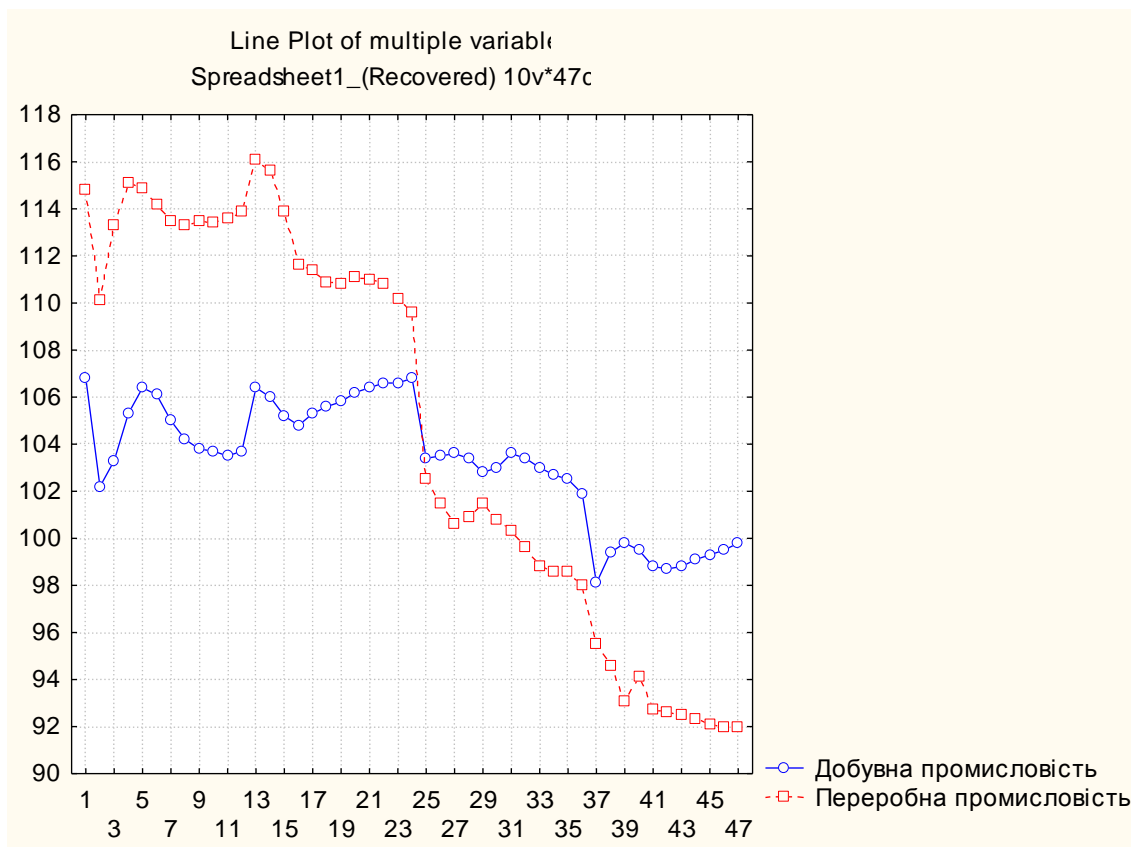


Рис. 3. Полігон розподілу

3. Перевірка вибірки на нормальний закон розподілу

3.1. Розрахунок кількості інтервалів групування. Кількість інтервалів групування розраховується за такою формулою:

$$m = 1 + 3,322 \lg n, \quad (1)$$

де n – довжина динамічного ряду.

$$m = 1 + 3,322 \lg 47 = 6,38 \approx 6.$$

3.2. Подальший аналіз здійснюється в рамках перевірки вибірки на нормальний закон розподілу. Розрахунок критерію Пірсона для формування остаточного висновку про характер розподілу величини. Для цього в головному меню Statistics вибираємо модуль Distribution Fitting (Закони розподілу). На рис. 4 показано вибір закону розподілу, який перевіряється.

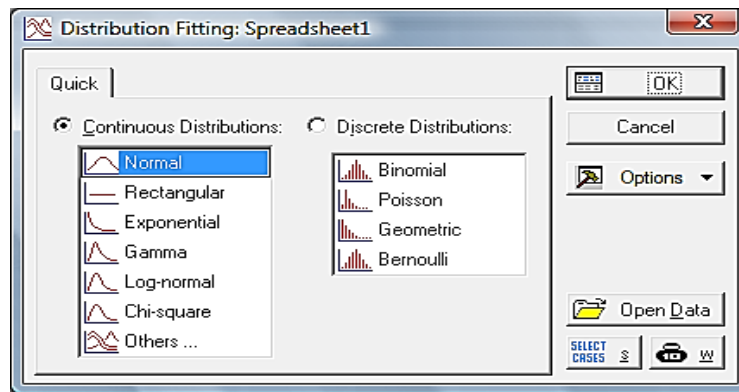


Рис. 4. Стартова панель модуля Distribution Fitting

Далі задаються параметри розрахунку, як показано на рис. 5. У вікні параметрів вказується кількість інтервалів, середнє значення індексу промисловості, мінімальне та максимальне значення показника.

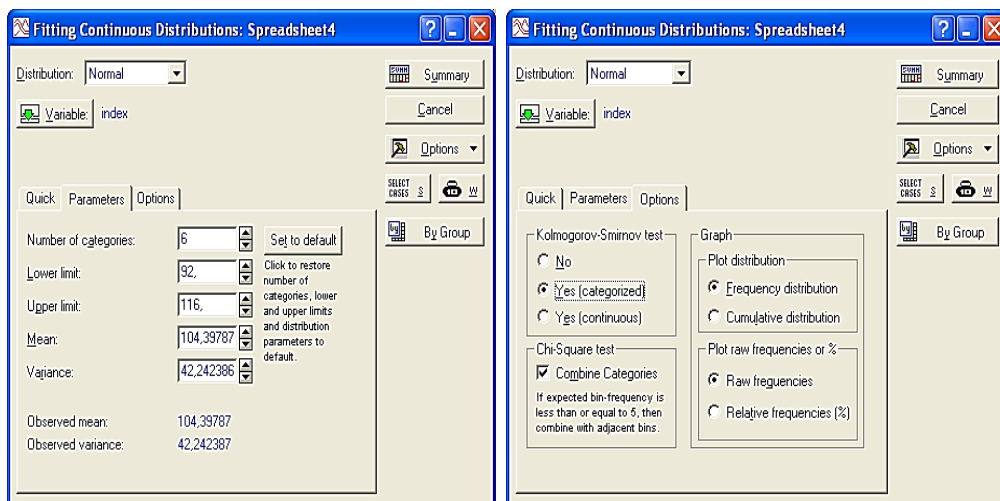


Рис. 5. Завдання параметрів розрахунку

Результати розрахунків наведено на рис. 6, 7.

Variable: Добувна промисловість, Distribution: Normal (Spreadsheet1_(Recovered)) Kolmogorov-Smirnov d = 0,08919, Chi-Square = 8,04061, df = 2 (adjusted) , p = 0,01795									
Upper Boundary	Observed Frequency	Cumulative Observed	Percent Observed	Cumul. % Observed	Expected Frequency	Cumulative Expected	Percent Expected	Cumul. % Expected	Observed-Expected
<= 98,83333	4	4	8,51064	8,51064	2,23807	2,23807	4,76171	4,76171	1,76199
100,66667	7	11	14,89362	23,40426	5,47534	7,71331	11,6496	16,4111	1,52469
102,50000	3	14	6,38298	29,78724	10,47871	18,19201	22,2952	38,7061	-7,47871
104,33333	16	30	34,04255	63,82979	12,67741	30,86951	26,9733	65,6791	3,32254
106,16667	9	39	19,1489	82,97869	9,69724	40,56674	20,6324	86,3121	-0,69724
< Infinity	8	47	17,02121	100,0000	6,43322	47,00001	13,68761	100,0000	1,56679

Variable: Переробна промисловість, Distribution: Normal (Spreadsheet1_(Recovered)) Kolmogorov-Smirnov d = 0,20561, p < 0,05 Chi-Square = 26,45456, df = 2 (adjusted) , p = 0,00000									
Upper Boundary	Observed Frequency	Cumulative Observed	Percent Observed	Cumul. % Observed	Expected Frequency	Cumulative Expected	Percent Expected	Cumul. % Expected	Observed-Expected
<= 93,33333	8	8	17,02121	17,02121	4,20591	4,20591	8,94881	8,94881	3,79409
98,66667	6	14	12,76596	29,78717	6,78901	10,99491	14,4447	23,3931	-0,78901
104,00000	9	23	19,1489	48,93607	10,48421	21,47921	22,30701	45,70001	-1,48421
109,33333	0	23	0,00000	48,93607	11,18421	32,66351	23,7963	69,49701	-11,18421
114,66667	19	42	40,42551	89,36107	8,24191	40,90551	17,53601	87,03301	10,75801
< Infinity	5	47	10,63831	100,0000	6,09441	47,00001	12,96691	100,0000	-1,09441

Рис. 6. Результати групування та розрахунку критерію Пірсона та критерію Колмогорова–Смірнова

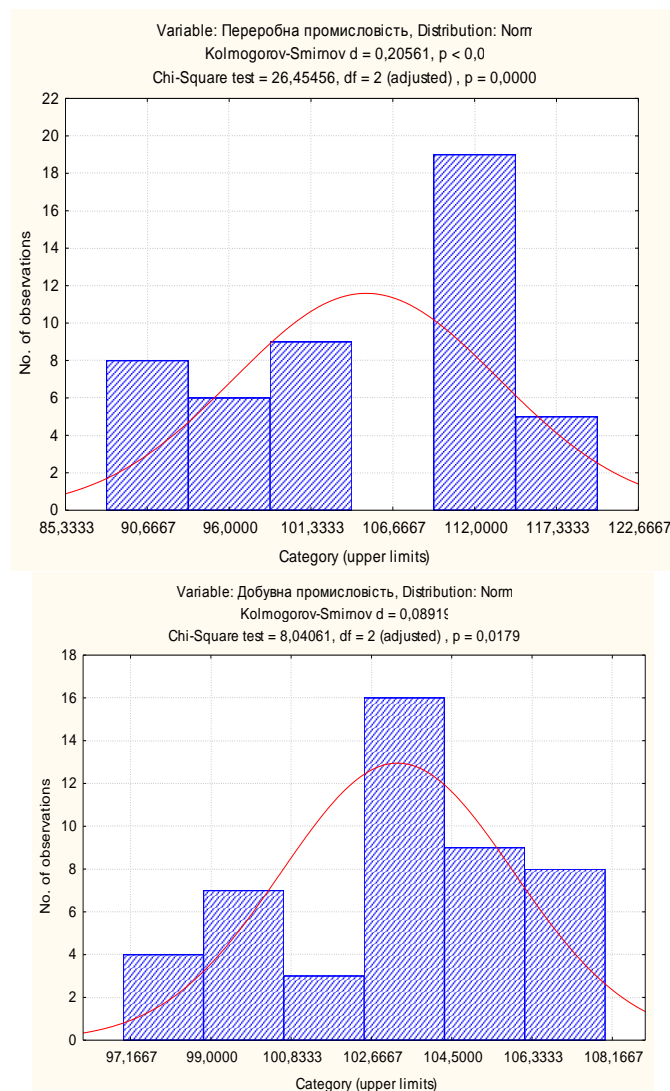


Рис. 7. Гістограма розподілу з критерієм Пірсона та Колмогорова–Смірнова

При нормальному розподілі розраховані значення критерію Пірсона та критерію Колмогорова–Смірнова менше критичного значення. Для $k - p - 1 = 4$ ступенів свободи (де k – кількість інтервалів розподілу, p – кількість параметрів що перевіряється) та рівня значимості $\alpha = 0,01$ табличне значення критерію Пірсона складає 18,47. З розрахунків видно, що критерій Пірсона складає 26,45 для індексу промислової продукції переробної промисловості та 8,06 для добувної. Критичне значення критерію Колмогорова–Смірнова при довжині вибірки 47 спостережень та $\alpha = 0,01$ складає 0,2377, а розрахункові – 0,21 та 0,09 для індексу промислової продукції переробної та добувної промисловості відповідно.

Таким чином, можна зробити висновок, що індекс промислової продукції добувної галузі розподілений за нормальним законом, для відповідного показника переробної галузі однозначного висновку не можна зробити.

4. Розрахунок t-критерію Стьюдента

Для запуску програми у верхньому меню Statistics треба вибрати команду Basic Statistic Tables (основні статистики/таблиці). Відкриється меню команди, в якому t-критерій поданий чотирма процедурами:

- t-test, independent, by variables (t-критерій для незалежних вибірок) застосовується, якщо треба порівняти середні випадкових величин, отриманих за двома різними (незалежними) вибірками;

- t-test, independent, by groups (t-критерій для незалежних вибірок із згрупованих змінних) використовується, якщо треба порівняти середні випадкових величин двох незалежних груп, отриманих з однієї вибірки за допомогою групової змінної;

- t-test, dependent samples (t-критерій для залежних вибірок) застосовується, якщо треба порівняти середні випадкових величин двох залежних груп;

- t-test, single samples (прості вибірки).

У перерахованих процедурах як і нульова гіпотеза передбачається, що середні в групах є рівними.

Рівень значущості p t-критерію дорівнює ймовірності помилково відкинутої гіпотези про рівність середніх двох вибірок, коли в дійсності ця гіпотеза має місце. Рівень значущості p – це максимально прийнятна для дослідника ймовірність помилково відхилити нульову гіпотезу, коли насправді вона правильна, тобто допускається ймовірність помилки першого роду. Величина рівня значущості встановлюється дослідником довільно, проте зазвичай приймається рівним 0,05, або 0,01, або 0,001. У програмі STATISTICA

принятною кордоном статистичної значущості прийняті значення, менші або рівні 0,05. Якщо p менше або дорівнює 0,05, то результат вважається статистично значимим, якщо менше або дорівнює 0,01, то результат вважається статистично високо значущим.

Значення витрат сировини під час виробництва продукції за старою та новою технологією отримані з різних вибірок, для порівняння середніх можна застосувати процедуру t-test, independent, by variables. Результати розрахунків наведено на рис. 8.

T-test for Independent Samples (Spreadsheet1_(Recovered))												
Note: Variables were treated as independent samples												
Group 1 vs. Group 2	Mean Group 1	Mean Group 2	t-value	df	p	t separate variances	df	p	Valid Group 1	Valid Group 2	Std. Dev. Group 1	Std. Dev. Group 2
X1 vs. X2	103,2	104,9	-1,26	92	0,207	-1,26	54,62	0,210	4	4	2,654	8,630

F-ratio Variances	p Variances	Levene F(1,df)	df Levene	p Levene	Brn-Fors F(1,df)	df Brn-Fors	p Brn-Fors
10,57269	0,000000	143,2315	92	0,000000	44,96612	92	0,000000

Рис. 8. Результати розрахунків

У таблиці з результатами аналізу, в якій будуть наведені значення таких статистик:

- Mean Group – середні значення змінних за групою;
- Std. Dev. Group – стандартні відхилення значень змінних;
- Valid N Group – число спостережень у групі;
- t – значення t-критерію;
- df – число ступенів свободи;
- p – рівень значимості t-критерію;
- F-ratio Variance – значення коефіцієнта F – тест для дисперсії;
- p Variance – рівень значимості коефіцієнта;
- F – тест для дисперсії.

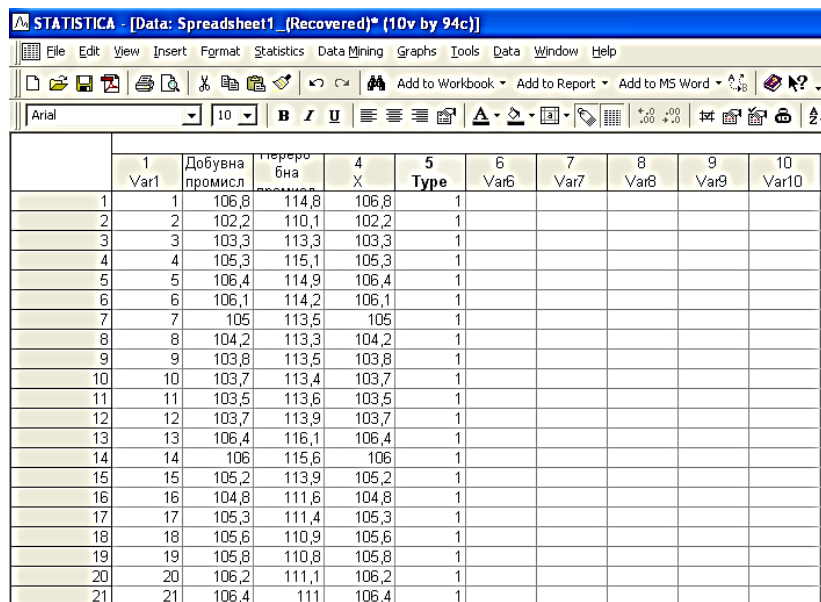
За даними рис. 8 не можна зробити висновок, що до істотності відмінності в середніх (p – рівень значимості t-критерію складає 0,22). Проте гіпотеза про рівність дисперсій вірна (p Variances значно більше, ніж 0,05).

5. Розрахунок непараметричних критеріїв порівняння середніх.

Для порівняння середніх у двох незалежних групах даних використовують критерії Wald-Wolfowitz test (Вальда-Вольфовица), Kolmogorov–Smirnov test (Колмогорова–Смирнова), Mann–Whitney test (Манна–Уїтні), що є непараметричними альтернативами t-критерію для двох незалежних вибірок і перевіряють нульову гіпотезу, що дві незалежні вибірки можуть відрізнитися не тільки середнім, а й формою розподілу.

5.1 Формування вхідного простору показників.

Для розрахунку непараметричних критеріїв, необхідно перетворити вихідну інформацію в вид наведений на рис. 9. Де X позначено стовпчик зі значенням витрат сировини, а змінна Type приймає значення 1 для добувної галузі, 2 – для переробної.



	1 Var1	2 Добувна промисл	3 переробна	4 X	5 Type	6 Var6	7 Var7	8 Var8	9 Var9	10 Var10
1	1	106,8	114,8	106,8	1					
2	2	102,2	110,1	102,2	1					
3	3	103,3	113,3	103,3	1					
4	4	105,3	115,1	105,3	1					
5	5	106,4	114,9	106,4	1					
6	6	106,1	114,2	106,1	1					
7	7	105	113,5	105	1					
8	8	104,2	113,3	104,2	1					
9	9	103,8	113,5	103,8	1					
10	10	103,7	113,4	103,7	1					
11	11	103,5	113,6	103,5	1					
12	12	103,7	113,9	103,7	1					
13	13	106,4	116,1	106,4	1					
14	14	106	115,6	106	1					
15	15	105,2	113,9	105,2	1					
16	16	104,8	111,6	104,8	1					
17	17	105,3	111,4	105,3	1					
18	18	105,6	110,9	105,6	1					
19	19	105,8	110,8	105,8	1					
20	20	106,2	111,1	106,2	1					
21	21	106,4	111	106,4	1					

Рис. 9. Вихідна інформація для розрахунку непараметричних критеріїв

5.2 Для того щоб виконувати обчислювальні процедури необхідно увійти в меню Statistics / Nonparametric Statistics. У вікні необхідно вибрати пункт Comparing two independent samples (group). Стартова панель модуля наведена на рис. 10. Модуль дозволяє розрахувати критерій серій Вальда-Вольфовица, двухвибірковий тест Колмогорова–Смирнова та U-тест Манна–Уїтні, а також побудувати діаграму розмаху та категоризовану діаграму. Для проведення розрахунків

необхідно задати незалежну змінну (X) та змінну групування (Type).

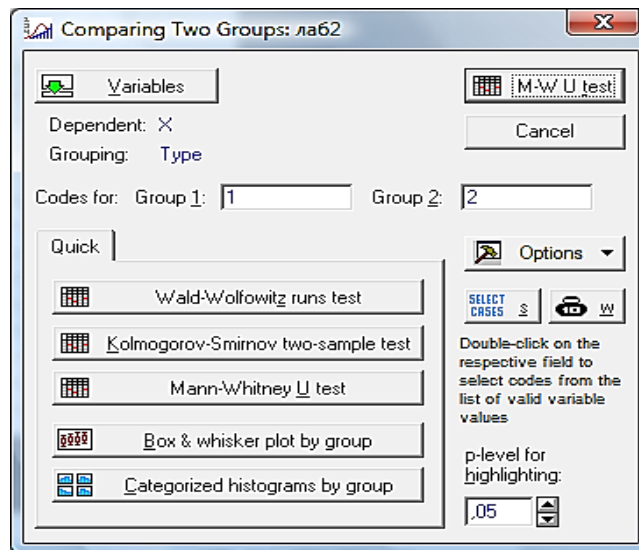


Рис. 10. Стартова панель модуля розрахунку непараметричних критеріїв

5.3 Розрахунок критерію серій Вальда-Вольфовица.

Критерій перевіряє гіпотезу про те що, дві незалежні вибірки описують два різні процеси і суттєво відрізняються не лише в значеннях середніх, а й за формою розподілу. Нульова гіпотеза полягає в тому, що дві вибірки описують один і той самий процес (рис. 11).

Wald-Wolfowitz Runs Test (Spreadsheet1_(Recovered))										
By variableType										
Marked tests are significant at p <,05000										
Variable	Valid N Group 1	Valid N Group 2	Mean Group 1	Mean Group 2	Z	p-level	Z adjstd	p-level	No. of Runs	No. of ties
X	47	47	103,261	104,931	-7,2590	0,00000	7,15538	0,00000	13	2

Рис. 11. Розрахунок критерію серій Вальда-Вольфовица

За розрахунками, можна зробити висновок що дві вибірки описують один і той самий процес.

5.4. Розрахунок двохвибіркового критерію Комогорова–Смірнова.

Двухвибірковий критерій Колмогорова–Смирнова чутливий не тільки до розбіжності в положенні двох розподілів, але також і до форми розподілу.

Фактично він чутливий до будь-якого відхилення від гіпотези однорідності, але не вказує, з яким саме відхиленням ми маємо справу (рис. 12).

Kolmogorov-Smimov Test (Spreadsheet1_(Recovered))									
By variableType									
Marked tests are significant at p <,05000									
variable	Max Neg Diffenc	Max Pos Diffenc	p-level	Mean Group 1	Mean Group 2	Std.Dev. Group 1	Std.Dev. Group 2	Valid N Group 1	Valid N Group 2
X	-0,51063	0,27659	p < .001	103,261	104,931	2,65424	8,63045	47	47

Рис. 12. Таблиця розрахунку двохвибіркового критерію Комогорова-Смірнова

Рівень значущості критерію $p < 0,001$, свідчить про значущу розбіжність між вибірками, що може бути пов'язана не лише з різницею в середніх значеннях, але й у розбіжності форм розподілу.

5.5. Розрахунок U-тесту Манна–Уїтні.

Призначений для оцінки відмінностей величин членів двох вибірок. Цей критерій заснований на підрахунку числа інверсій U (перестановок) членів у їх загальному упорядкованому ряді (рис. 13).

Mann-Whitney U Test (Spreadsheet1_(Recovered))										
By variableType										
Marked tests are significant at p <,05000										
variable	Rank Sum Group 1	Rank Sum Group 2	U	Z	p-level	Z adjusted	p-level	Valid N Group 1	Valid N Group 2	2*1 sided exact p
X	2115,50	2349,50	987,500	-0,88474	0,37629	-0,88483	0,37624	47	47	0,37800

Рис. 13. Таблиця розрахунку U-тесту Манна–Уїтні

З розрахунку видно, що критерій не є статистично значимим ($p = 0,38$), тобто можна зробити про відсутність розбіжності між даними вибірками.

5.6. Побудова діаграми розмаху.

Діаграма розмаху будується ініціалізацією кнопки Box & whisker plot by group на стартовій панелі модуля. Діаграма зображена на рис. 14.

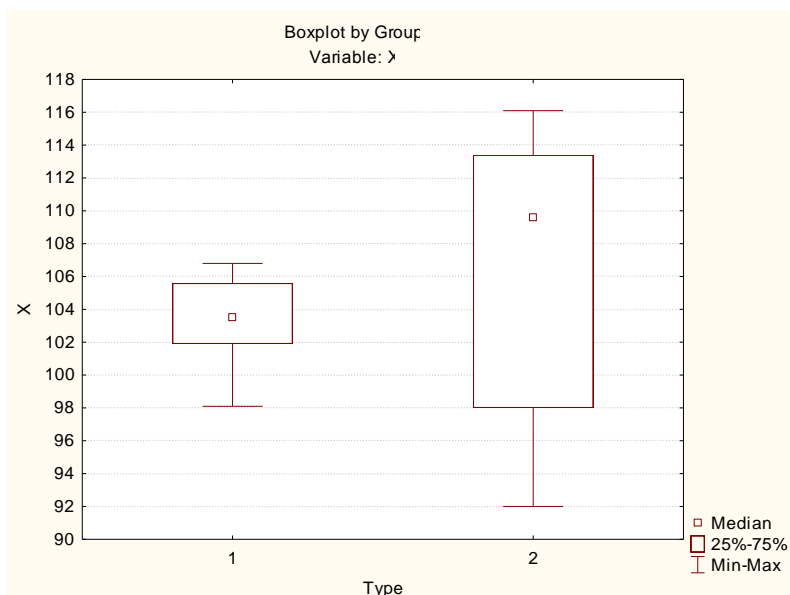


Рис. 14. Діаграма розмаху

На діаграмі розмаху для кожної змінної показується медіана, квартальний розмах і розмах.

Проведений розрахунок доводить, що наявність загальних тенденцій у зміні індексу промислової продукції добувної і переробної промисловості описують схожі процеси.

Лабораторна робота № 2

Визначення та усунення аномальних рівнів часового ряду.

Дослідження закономірностей у часовому ряду. Методи визначення стійкості часового ряду

Мета – закріплення теоретичних та практичних знань проведення попередніх етапів дослідження часового ряду в пакеті MS Excel.

Завдання. Перевірити наявність аномальних значень у динамічному ряді обсягу імпорту Вінницької області (табл. 3, стовпчик 2) та за необхідності усунути їх.

Таблиця 3

Обсяг імпорту в Вінницької області

Період	Обсяг імпорту млн \$	Період	Обсяг імпорту млн \$
1	39,967	11	60,648

2	43,725	12	59,495
3	66,066	13	57,229
4	52,42	14	57,428
5	56,811	15	55,23
6	53,137	16	62,99
7	50,151	17	39,722
8	53,069	18	47,367
9	57,908	19	59,578
10	59,004	20	54,289

Методичні рекомендації

1. Формування вихідного файлу в пакеті MS Excel.
2. За допомогою стандартної функції ДИСП розрахувати дисперсію та середнє квадратичне відхилення динамічного ряду.
3. За формулою 2 розрахувати значення λ для кожного рівня ряду:

$$\lambda = \frac{|y_t - y_{t-1}|}{S_y} \quad (2)$$

Проведені розрахунки наведено в табл. 4.

4. Відповідно до статистичних таблиць $\lambda_{табл} (0,05; 10)$, де $a = 0,05$ – рівень значущості, а $n = 20$ – кількість членів ряду. $\lambda_{табл} = 3,1$.

5. Проводити порівняння розрахованих значень λ для кожного рівня ряду з табличним. Виходячи з того що $\lambda_3 > \lambda_{табл} (3,12 > 3,1)$ та $\lambda_{17} > \lambda_{табл} (3,26 > 3,1)$ то значення y_3 та y_{17} є аномальними рівнями.

6. Проводити заміну значень аномальних рівнів на середнє арифметичне їх сусідів (табл.4).

7. Перевіряємо відповідно до зазначеного алгоритму (крок 1 – 5) наявність аномальних рівнів у новому ряду. Як видно в табл. 4, проведені розрахунки дозволили усунути аномальні рівні в ряді.

Таблиця 4

Обсяг імпорту Вінницької області та розрахунки аномальних значень ряду

Вихідний ряд				Ряд після перетворень		
t	y_t	$ y_t - y_{t-1} $	λ	y_t	$ y_t - y_{t-1} $	λ

1	2	3	4	5	6	7
1	39,9671			39,9671		
2	43,7251	3,758	0,525617	43,7251	3,758	0,637825
3	66,0657	22,3406	3,124696	48,07275	4,34765	0,737903
4	52,4204	13,6453	1,908517	52,4204	4,34765	0,737903
5	56,8112	4,3908	0,614125	56,8112	4,3908	0,745227
6	53,137	3,6742	0,513897	53,137	3,6742	0,623602
7	50,1505	2,9865	0,417711	50,1505	2,9865	0,506883
8	53,0692	2,9187	0,408228	53,0692	2,9187	0,495375
9	57,908	4,8388	0,676785	57,908	4,8388	0,821264
10	59,0041	1,0961	0,153307	59,0041	1,0961	0,186035
11	60,6483	1,6442	0,229968	60,6483	1,6442	0,279061
12	59,4952	1,1531	0,16128	59,4952	1,1531	0,195709
13	57,229	2,2662	0,316965	57,229	2,2662	0,38463
14	57,4278	0,1988	0,027805	57,4278	0,1988	0,033741

Закінчення табл. 4

1	2	3	4	5	6	7
15	55,2298	2,198	0,307426	55,2298	2,198	0,373055
16	62,9896	7,7598	1,085334	62,9896	7,7598	1,317029
17	39,7218	23,2678	3,25438	55,17815	7,81145	1,325796
18	47,3667	7,6449	1,069264	47,3667	7,81145	1,325796
19	59,5783	12,2116	1,707991	59,5783	12,2116	2,072609
20	54,2894	5,2889	0,739739	54,2894	5,2889	0,897657
	ДИСП = 51,1			ДИСП = = 34,7		
	Sy = 7,2			Sy = 5,9		

На рис. 15 наведено порівняння вихідного та згладженого динамічного ряду.



Рис. 15. Графіки вихідного та згладженого ряду динаміки

8. Проаналізувати загальну тенденцію зміни показника мінімальної заробітної плати та порівняти її зі зміною реальної заробітної плати в країні. Мінімальну та реальну заробітну плату в Україні за 2002 – 2013 рр. наведено в табл. 5.

Таблиця 5

Мінімальна заробітна плата в Україні за 2002 – 2013 рр., грн

Дата	Мінімальна заробітна плата, грн	Реальна заробітна плата, грн
1	2	3
01.01.2002 р.	140	318
01.01.2003 р.	185	395
01.01.2005 р.	262	493

Закінчення табл. 5

1	2	3
01.01.2006 р.	350	630
01.01.2007 р.	400	855
01.01.2008 р.	515	1 106
01.01.2009 р.	605	1 618
01.01.2010 р.	869	1 882
01.01.2011 р.	941	2 274
01.01.2012 р.	1 073	2 717
01.01.2013 р.	1 147	2 994

9. Побудувати графік динаміки мінімальної заробітної плати (рис. 16) з метою візуального аналізу досліджуваного динамічного ряду.

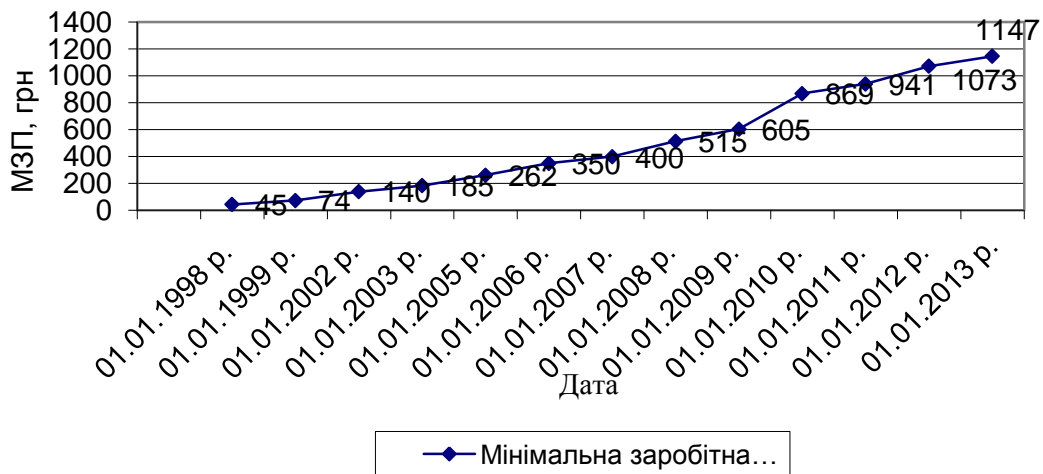


Рис. 16. Динаміка зміни рівня мінімальної заробітної плати в Україні

Як видно з графіка, рівень мінімальної заробітної плати на всьому досліджуваному періоді зростає, проте ці зміни не є однаковими. Для більш докладного дослідження загальної тенденції зміни показника, доцільно провести розрахунок показників інтенсивності динаміки.

9.1 Визначити абсолютні прирости:

у клітинку C4 ввести формулу = B4 – \$B\$3. Розтягнути цю формулу на весь стовпець;

у клітинку D5 (для 2002 р. абсолютний приріст неможливо розрахувати ланцюговим способом, тому що немає попереднього періоду – 2001 р.) ввести формулу = B5 – B4 і також розтягти цю формулу на весь стовпець.

9.2 Визначити темпи зростання:

у клітинку E4 ввести формулу = B4 / \$B\$3 X 100. Розтягти цю формулу на весь стовпець;

у клітинку F5 (для 2002 р. темп зростання, як і абсолютний приріст, неможливо розрахувати ланцюговим способом, тому що немає попереднього періоду – 2001 р.) вводимо формулу = B5 / B4 X 100 і також розтягти цю формулу на весь стовпець.

9.3 Визначити темпи приросту:

у клітинку G4 ввести формулу = E4 – 100. Розтягти цю формулу на весь стовпець;

у клітинку H5 ввести формулу = F5 – 100 і також розтягти цю формулу на весь стовпець.

9.4 Визначити абсолютне значення 1% приросту:

у клітинку I5 ввести формулу = B4 / 100. Розтягти цю формулу на весь стовпець.

Результати всіх проведених розрахунків наведено в табл. 6.

Таблиця 6

Основні характеристики динаміки мінімальної заробітної плати

Дата	Мінімальна заробітна плата, грн	Δу, грн (–)		Тр, %		Тпр, %		А
		базисний	ланцюговий	базисний	ланцюговий	базисний	ланцюговий	
01.01.2002 р.	140							
01.01.2003 р.	185	140	45	4,11	1,32	3,11	0,32	1,4
01.01.2005 р.	262	217	77	5,82	1,42	4,82	0,41	1,85
01.01.2006 р.	350	305	88	7,78	1,34	6,77	0,33	2,62
01.01.2007 р.	400	355	50	8,89	1,14	7,88	0,14	3,5
01.01.2008 р.	515	470	115	11,44	1,29	10,44	0,28	4
01.01.2009 р.	605	560	90	13,44	1,17	12,44	0,17	5,15
01.01.2010 р.	869	824	264	19,31	1,44	18,31	0,43	6,05
01.01.2011 р.	941	896	72	20,91	1,08	19,91	0,08	8,69
01.01.2012 р.	1 073	1 028	132	23,84	1,14	22,84	0,14	9,41
01.01.2013 р.	1 147	1 102	74	25,49	1,07	24,48	0,06	10,73

9.5 З метою порівняння зміни двох динамічних рядів необхідно привести ряди динаміки (див. табл. 7) до однієї основи, тобто до загальної бази порівняння. Привести порівнювані ряди до однієї основи, визначивши відносні рівні рядів: базисні темпи зростання з постійною базою порівняння за 2002 р.

у клітинку B7 ввести формулу = B2 / \$B\$2 x 100. Розтягти цю формулу до клітинки F7;

у клітинку B8 ввести формулу = B3 / \$B\$3 x 100. Розтягти цю формулу до клітинки F8.

9.6 Розрахунок коефіцієнтів випередження:

у клітинку B12 ввести формулу = B7 / B\$9. Розтягти цю формулу до клітинки F12.

Розрахунок коефіцієнта випередження/запізнення

Дата	Темп зростання		Коефіцієнт випередження /запізнення
	Мінімальна заробітна плата, грн	Реальна заробітна плата, грн	
01.01.2002 р.			
01.01.2003 р.	132,1429	124,2138	1,063834
01.01.2005 р.	187,1429	155,0314	1,207128
01.01.2006 р.	250	198,1132	1,261905
01.01.2007 р.	285,7143	268,8679	1,062657
01.01.2008 р.	367,8571	347,7987	1,057672
01.01.2009 р.	432,1429	508,805	0,849329
01.01.2010 р.	620,7143	591,8239	1,048816
01.01.2011 р.	672,1429	715,0943	0,939936
01.01.2012 р.	766,4286	854,4025	0,897035
01.01.2013 р.	819,2857	941,5094	0,870183

Як видно з розрахунків до 2009 р. приріст мінімальної заробітної плати випереджує приріст реальної заробітної плати, а з 2009 р. спостерігається зворотня тенденція.

Аналізуючи дані, наведені в табл. 6, можна зробити висновок, що протягом досліджуваного періоду мінімальна заробітна плата населення України постійно зростає, про що свідчить позитивний абсолютний приріст, також при цьому темпи зростання мінімальної заробітної плати збільшуються, про що свідчить збільшення темпу приросту. Отже, в 2011 р. зафіксоване різке скорочення темпу зміни розміру мінімальної заробітної плати – 8 %, що становить 72 грн, відносно попереднього року, а у 1999 р. спостерігається найбільший темп приросту – 89 %.

Розрахунок ланцюгових показників дає змогу провести порівняльну оцінку за роками, тоді як базисні показники дозволяють оцінити загальну динаміку за визначений період, а також провести порівняльний аналіз.

За визначений період у середньому щорічний приріст МЗП склав 41,66 %.

Лабораторна робота № 3

Згладжування часових рядів. Побудова трендових моделей часових рядів

Мета – закріплення теоретичних та практичних матеріалів із методів згладжування часових рядів та моделювання загальної тенденції соціально-економічних процесів та явищ у пакетах Statistica 8.0 та MS Excel. Проведення згладжування часових рядів. Отримання навичок дослідження наявності трендової складової ряду. Побудова трендової моделі та оцінка її адекватності.

Завдання. Провести дослідження ВВП України за період з 2006 по 2013 рік на наявність загальної тенденції (табл. 8). Перевірити вид трендової складової ряду та побудувати модель даної складової.

Таблиця 8

ВВП України за період із 2006 по 2013 р

Період	ВВП, млн грн	Період	ВВП, млн грн
1	2	3	4
1 кв. 2006	106 348	1 кв. 2010	23 860,00
2 кв. 2006	126 319	2 кв. 2010	23 986
3 кв. 2006	152 406	3 кв. 2010	24 567
4 кв. 2006	159 080	4 кв. 2010	25 322
1 кв. 2007	15 330,00	1 кв. 2011	25 689
2 кв. 2007	16 800,00	2 кв. 2011	311 022
3 кв. 2007	18 330,00	3 кв. 2011	369 818
4 кв. 2007	19 660,00	4 кв. 2011	363 557

Закінчення табл. 8

1	2	3	4
1 кв. 2008	20 530,00	1 кв. 2012	293 493
2 кв. 2008	20 800,00	2 кв. 2012	349 212
3 кв. 2008	20 610,00	3 кв. 2012	387 620
4 кв. 2008	20 210,00	4 кв. 2012	378 564
1 кв. 2009	19 860,00	1 кв. 2013	301 598
2 кв. 2009	19 800,00	2 кв. 2013	351 896
3 кв. 2009	20 290,00	3 кв. 2013	392 630
4 кв. 2009	21 560,00	–	–

Методичні рекомендації

1. Побудувати графік динамічного ряду, що досліджується (рис. 17).

Приклад побудови графіка наведено на рис. 17.

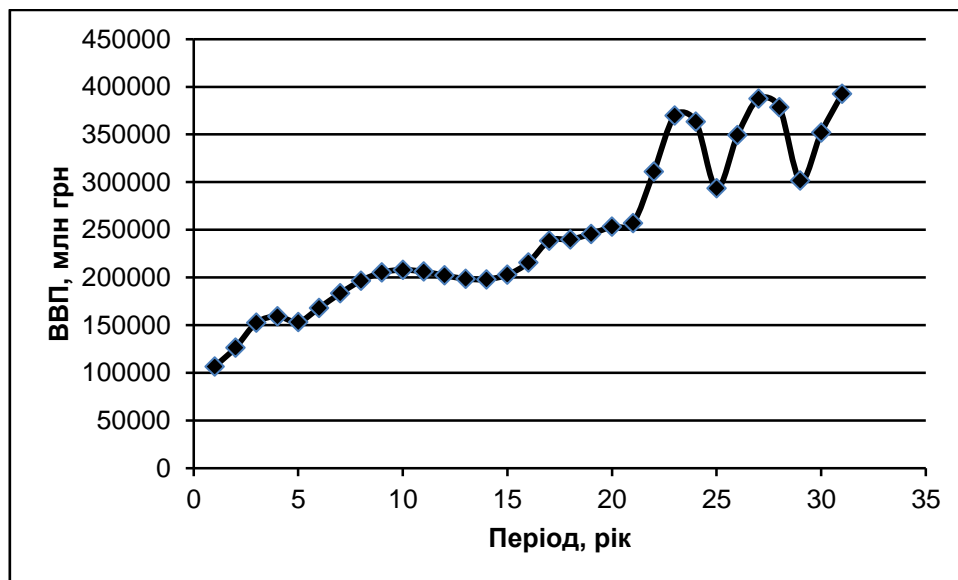


Рис. 17. Динаміка ВВП, млн грн

Виходячи з візуального аналізу динаміки ВВП за 2006 по 2013 рр., можна висунути гіпотезу про наявність трендової складової в досліджуваному ряді. Для перевірки даної гіпотези використати кумулятивний Т-критерій, критерій Аббе.

2. Розрахувати кумулятивний Т-критерій, що дозволяє визначити наявність не тільки самої тенденції, але і її математичного виразу – тренда. Висувається основна гіпотеза (H_0) про відсутність тенденції у вихідному часовому ряді. Розрахункове значення критерію визначається, як відношення накопиченої суми квадратів відхилень емпіричних значень, рівнів часового ряду від їх середнього значення ($\sum_n Z_n^2$) і самих відхилень за формулою:

$$T_p = \frac{\sum_n Z_n^2}{\sigma_n^2}, \quad (3)$$

де Z_n – накопичений підсумок відхилень емпіричних значень від середнього рівня вихідного часового ряду;

σ_n^2 – загальна сума квадратів відхилень, що визначається за формулою:

$$\sigma_n^2 = \sum_n y_t^2 - (\bar{y})^2 \cdot n;$$

y_t – вихідні значення ознаки;

\bar{y} – середні рівні вихідного ряду;

n – довжина часового ряду.

Фрагмент проведених розрахунків за формулою 3 наведено на рис. 18.

Период	yt	yt ²	yt-уср	Zt	Zt ²
1 кв 2006	106348	11309897104	-139317	-139317	1,941E+10
2 кв 2006	126319	15956489761	-119346	-258662	6,691E+10
3 кв 2006	152406	23227588836	-93259	-351921	1,238E+11
4 кв 2006	159080	25306446400	-86585	-438505	1,923E+11
1 кв 2007	153300,00	23500890000	-92365	-530870	2,818E+11
2 кв 2007	168000,00	28224000000	-77665	-608535	3,703E+11
3 кв 2007	183300,00	33598890000	-62365	-670899	4,501E+11
4 кв 2007	196600,00	38651560000	-49065	-719964	5,183E+11
1 кв 2008	205300,00	42148090000	-40365	-760329	5,781E+11
2 кв 2008	208000,00	43264000000	-37665	-797993	6,368E+11
3 кв 2008	206100,00	42477210000	-39565	-837558	7,015E+11
4 кв 2008	202100,00	40844410000	-43565	-881122	7,764E+11
1 кв 2009	198600,00	39441960000	-47065	-928187	8,615E+11
2 кв 2009	198000,00	39204000000	-47665	-975852	9,523E+11
3 кв 2009	202900,00	41168410000	-42765	-1E+06	1,038E+12
4 кв 2009	215600,00	46483360000	-30065	-1E+06	1,1E+12
1 кв 2010	238600,00	56929960000	-7064,6	-1E+06	1,115E+12
2 кв 2010	239860	57532819600	-5804,6	-1E+06	1,127E+12
3 кв 2010	245670	60353748900	5,3871	-1E+06	1,127E+12
4 кв 2010	253220	64120368400	7555,39	-1E+06	1,111E+12
1 кв 2011	256890	65992472100	11225,4	-1E+06	1,087E+12
2 кв 2011	311 022	96734684484	65357,4	-977406	9,553E+11
3 кв 2011	369 818	1,36765E+11	124153	-853253	7,28E+11
4 кв 2011	363 557	1,32174E+11	117892	-735361	5,408E+11
1 кв 2012	293 493	86138141049	47828,4	-687532	4,727E+11
2 кв 2012	349 212	1,21949E+11	103547	-583985	3,41E+11
3 кв 2012	387 620	1,50249E+11	141955	-442030	1,954E+11
4 кв 2012	378 564	1,43311E+11	132899	-309130	9,566E+10
1 кв 2013	301 598	90961353604	55933,4	-253197	6,411E+10
2 кв 2013	351896	1,23831E+11	106231	-146965	2,16E+10

Рис. 18. Фрагмент розрахунку кумулятивного T-критерію

Розрахункові значення кумулятивного T-критерію порівнюються з критичними при заданому рівні значущості $\alpha = 0,01$. Якщо розрахункове значення $T_r = 86$, перевищує критичне (табличне) значення критерію ($T_{кр} = 2,75$), то гіпотеза про відсутність тренда відкидається, отже, у вихідному часовому ряді існує тенденція, описувана трендом.

3. Критерій квадратів послідовних різниць (критерій Аббе).

Якщо

є підстави вважати, що рівні ряду динаміки щодо свого середнього значення підпорядковані нормальному (гауссову)

закону розподілу, то для з'ясування питання про наявність не випадкової складової використовують критерій послідовних різниць.

3.1. Розрахуємо різницю між рівнями ряду та між рівнем ряду та середнім таненням. Результати проведених розрахунків наведено на рис. 19.

Період	yt	yi-1 - yi	(yi+1*yi) ²	ȳ	yi - ȳ	(yi - ȳ) ²
1 кв 2006	106348	19971	398840841	-139317	19409118631	
2 кв 2006	128319	26087	680531569	-119346	14243375319	
3 кв 2006	152406	6674	44542276	-93258,6	8697168881	
4 кв 2006	199080	-5780	33408400	-86584,6	7496895192	
1 кв 2007	153300,00	14700	216090000	-92364,6	8531221717	
2 кв 2007	168000,00	15300	234090000	-77664,6	6031792097	
3 кв 2007	183300,00	13300	176890000	-62364,6	3889344943	
4 кв 2007	196600,00	8700	75690000	-49064,6	2407336239	
1 кв 2008	205300,00	2700	7290000	-40364,6	1629301975	
2 кв 2008	208000,00	-1900	3610000	-37664,6	1418623065	
3 кв 2008	205100,00	-4000	16000000	-39564,6	1565358594	
4 кв 2008	202100,00	-3500	12250000	-43564,6	1897875497	
1 кв 2009	198600,00	-600	360000	-47064,6	2215077788	
2 кв 2009	198000,00	4900	24010000	-47664,6	2271915323	
3 кв 2009	202900,00	12700	161290000	-42764,6	1828812117	
4 кв 2009	215600,00	23000	529000000	-30064,6	903880949	
1 кв 2010	238600,00	1260	1587600	-7064,61	49908755,47	
2 кв 2010	239860	5810	33756100	-5804,61	33693530,96	
3 кв 2010	245670	7550	57002500	5,387097	29,02081165	
4 кв 2010	253220	3670	13468900	7555,387	57083874,18	
1 кв 2011	256890	54132	2930273424	11225,39	126009315,5	
2 кв 2011	311 022	58796	3456969616	65357,39	4271588048	
3 кв 2011	368 818	-6261	39200121	124153,4	15414063528	
4 кв 2011	363 557	-70064	4908964096	117892,4	13898614935	
1 кв 2012	283 493	55719	3104606961	47828,39	2287554612	
2 кв 2012	349 212	38408	1475174464	103547,4	10722061375	
3 кв 2012	387 620	-9056	82011136	141955,4	20151331926	
4 кв 2012	378 564	-76966	5923765156	132899,4	17662247091	
1 кв 2013	301 598	50298	2529888804	55933,39	3128543792	

Рис. 19. Фрагмент проміжних розрахунків до критерію Аббе

3.2. Провести безпосередні розрахунки за критерієм:

$$q^2(n) = \frac{\sum_{i=1}^{n-1} (y_{i+1} - y_i)}{2n - 1} = 472620011,8;$$

$$s^2(n) = \frac{\sum_{i=1}^{n-1} (y_i - \bar{y})^2}{n - 1} = 6117496891;$$

$$\gamma(n) = \frac{q^2(n)}{s^2(n)} = 0,077.$$

3.3. Визначення y_α – величина, що розраховується за таблицею значень функції Лапласа $\Phi(x)$ для обраного рівня значущості α , виходячи зі співвідношення:

$$\Phi(y_\alpha) = \alpha, \alpha = 1 - \gamma = 1 - 0,077 = 0,92,$$

$$y_\alpha = 0,62.$$

3.4. Потім значення $\gamma(n)$ порівнюється з величиною:

$$\gamma_{\min}(n; \alpha) = 1 + \frac{y_{\alpha}}{\sqrt{n + 0,5(1 + y_{\alpha}^2)}}, = 1 + \frac{0,62}{\sqrt{31 + 0,5(1 + 0,38)}} = 1,11.$$

Правило перевірки гіпотези, якщо виконується нерівність:

$$\gamma(n) \leq \gamma_{\min}(n; \alpha),$$

$$0,077 \leq 1,11,$$

тобто з імовірністю 0,92 можна стверджувати, що в ряді ВВП присутня не випадкова складова.

4. Провести згладжування рівнів ряду за допомогою пакета Statistica 8.0. Аналіз часових рядів у пакеті проводиться в модулі Advanced Linear/Nonlinear Models/Time Series/Forecasting. Отримані результати згладжування з урахування квартальної сезонності наведено на рис. 20.

Seasonal Decomposition: Additive season (4) (Spreadsheet47)							
Y							
Case	Y	Moving Averages	Diffnrcs	Seasonal Factors	Adjusted Series	Smoothed Trend-c.	Irreg. Compon.
1	106348,0			-17883,0	124231,0	125360,0	-1129,0
2	126319,0			-1004,0	127323,0	130841,0	-3518,0
3	152406,0	136038,0	16367,0	11436,0	140969,0	141803,0	-833,0
4	159080,0	147776,0	11303,0	7450,0	151629,0	152835,0	-1206,0
5	153300,0	158196,0	-4896,0	-17883,0	171183,0	163072,0	8111,0
6	168000,0	165920,0	2080,0	-1004,0	169004,0	170431,0	-1427,0
7	183300,0	175300,0	8000,0	11436,0	171863,0	180695,0	-8832,0
8	196600,0	188300,0	8300,0	7450,0	189149,0	192838,0	-3689,0
9	205300,0	198300,0	7000,0	-17883,0	223183,0	203598,0	19585,0
10	208000,0	204000,0	4000,0	-1004,0	209004,0	205167,0	3837,0
11	206100,0	205375,0	725,0	11436,0	194663,0	203440,0	-8777,0
12	202100,0	203700,0	-1600,0	7450,0	194649,0	201583,0	-6933,0
13	198600,0	201200,0	-2600,0	-17883,0	216483,0	202542,0	13940,0
14	198000,0	200400,0	-2400,0	-1004,0	199004,0	201744,0	-2740,0
15	202900,0	203775,0	-875,0	11436,0	191463,0	206851,0	-15388,0
16	215600,0	213775,0	1825,0	7450,0	208149,0	217801,0	-9651,0
17	238600,0	224240,0	14360,0	-17883,0	256483,0	232574,0	23908,0
18	239860,0	234932,0	4927,0	-1004,0	240864,0	239771,0	1092,0
19	245670,0	244337,0	1332,0	11436,0	234233,0	245247,0	-11013,0
20	253220,0	248910,0	4310,0	7450,0	245769,0	256468,0	-10698,0

Рис. 20. Результати згладжування часового ряду
Зглажені значення ряду отримаємо в стовпчику Smoothed.

5. Провести визначення виду тренда за допомогою графічного методу. Графік зглаженого ряду наведено на рис. 21.

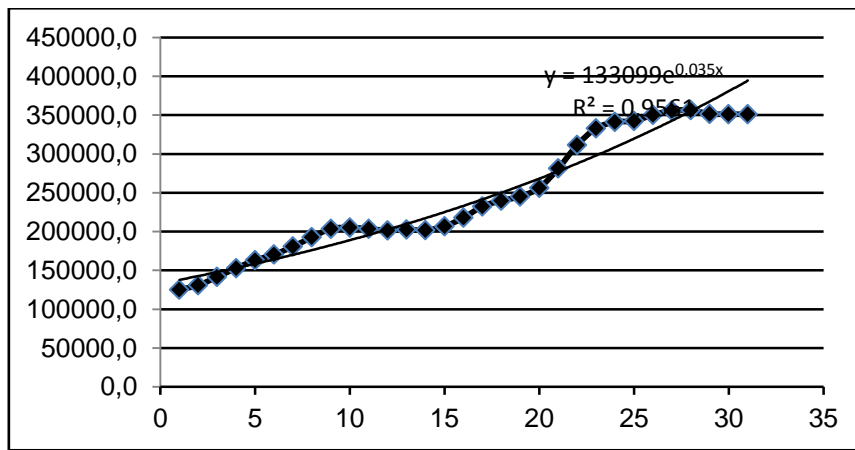


Рис. 21. Графік згладженого динамічного ряду

Як видно з графіка експоненціальна функція відображає загальну тенденцію в ряді динаміки.

6. Метод характеристик приросту. Порівнюються характеристики зміни приростів рівнів часового ряду з відповідними характеристиками обраної кривої. Розраховані характеристики наведено в табл. 9.

Таблиця 9.

Характеристики приросту

Період	y_t	$\Delta_t^{(1)}$	$\frac{\Delta_t}{y_t}$	$\log \Delta_t$	$\log \frac{\Delta_t}{y_t}$
1	2	3	4	5	6
1 кв. 2006	125 360,2				
2 кв. 2006	130 841,1	5480,9	0,041890107	3,73886	-1,377888535
3 кв. 2006	141 803,0	10961,9	0,077303693	4,03989	-1,111799757
4 кв. 2006	152 835,6	11032,6	0,072186121	4,04268	-1,141546293
1 кв. 2007	163 072,1	10236,5	0,062772814	4,01015	-1,202228403
2 кв. 2007	170 431,6	7359,4	0,043181228	3,86685	-1,364705007
3 кв. 2007	180 695,9	10 264,3	0,056804482	4,01133	-1,245617395
4 кв. 2007	192 838,8	12 142,9	0,062969389	4,08432	-1,200870523
1 кв. 2008	203 598,1	10 759,3	0,052845643	4,03178	-1,27699081

Закінчення табл. 9

1	2	3	4	5	6
---	---	---	---	---	---

2 кв. 2008	205 167,1	1 569,0	0,007647424	3,19562	-2,116484825
3 кв. 2008	203 440,3	-1 726,8	0,008487863	3,23724	-2,07120164
4 кв. 2008	201 583,3	-1 857,1	0,009212349	3,26882	-2,035629641
1 кв. 2009	202 542,6	959,3	0,004736159	2,98194	-2,324573724
2 кв. 2009	201 744,9	-797,7	0,003953838	2,90182	-2,402981123
3 кв. 2009	206 851,5	5 106,6	0,024687085	3,70813	-1,607530193
4 кв. 2009	217 801,1	10 949,6	0,050273449	4,0394	-1,298661317
1 кв. 2010	232 574,8	14 773,7	0,063522441	4,16949	-1,197072824
2 кв. 2010	239 771,6	7 196,8	0,030015143	3,85714	-1,52265959
3 кв. 2010	245 247,0	5 475,4	0,022326259	3,73842	-1,65118404
4 кв. 2010	256 468,0	11 220,9	0,043751837	4,05003	-1,359003706
1 кв. 2011	281 391,7	24 923,7	0,088573047	4,39661	-1,052698417
2 кв. 2011	311 584,8	30 193,1	0,096901749	4,47991	-1,013668384
3 кв. 2011	333 062,0	21 477,2	0,064484166	4,33198	-1,190546915
4 кв. 2011	341 119,6	8 057,6	0,023621072	3,90621	-1,626700405
1 кв. 2012	342 370,9	1 251,3	0,003654732	3,09735	-2,437144491
2 кв. 2012	350 331,9	7 961,0	0,022724165	3,90097	-1,643512062
3 кв. 2012	355 785,1	5 453,2	0,015327302	3,73665	-1,814534277
4 кв. 2012	356 420,6	635,5	0,001783006	2,80312	-2,748847275
1 кв. 2013	351 538,6	-4 882,1	0,013887692	3,6886	-1,857369908
2 кв. 2013	351 191,5	-347,1	0,000988393	2,54047	-3,005070456
3 кв. 2013	351 017,9	-173,6	0,000494441	2,23944	-3,305885772

Як видно з табл. 9, співвідношення абсолютного приросту до значення рівня ряду мають дуже невеликі значення, що свідчить про підтвердження гіпотези про експоненційний вид трендової складової.

7. Визначення параметрів трендової моделі в пакеті Statistica 8.0. Для побудови нелінійної трендової моделі необхідно перейти в Advanced Linear/Nonlinear Models/Nonlinear Estimation. У вікні модуля необхідно ввести вид функції в даному випадку експоненційної (рис. 22).

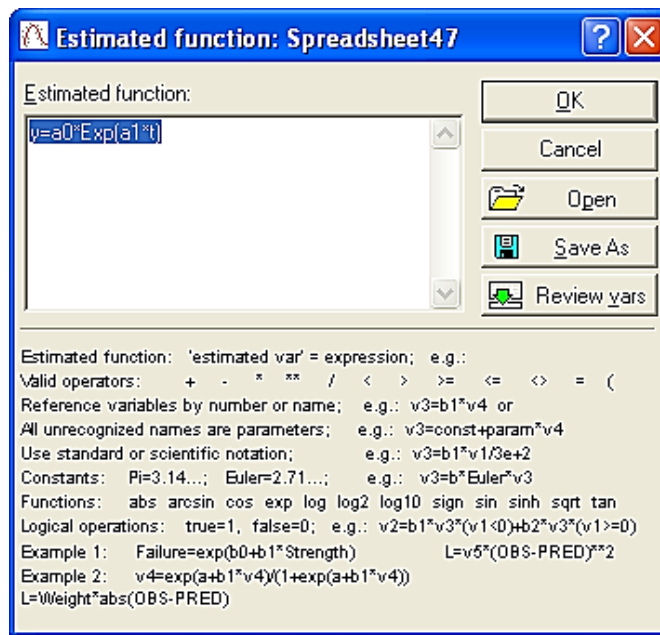


Рис. 22. Визначення виду нелінійної функції

На рис. 23 наведено результати визначення параметрів експоненційної функції.

Model is: $y=a_0 \cdot \text{Exp}(a_1 \cdot t)$
Dependent variable: y Independent variables: 1
Loss function: least squares
Final value: 22677538051,34
Proportion of variance accounted for: ,88944459 R = ,9431037

Model is: $y=a_0 \cdot \text{Exp}(a_1 \cdot t)$ (Spreadsheet47)						
Dep. Var. :y						
Level of confidence: 95.0% (alpha=0.050)						
	Estimate	Standard error	t-value df = 29	p-level	Lo. Conf Limit	Up. Conf Limit
a0	133932,9	7391,43	18,1200	0,00000	118815,;	149050,;
a1	0,0	0,002	14,3470	0,00000	0,0	0,0

Рис. 23. Результати побудови експоненційної функції

Як видно з розрахунків побудована трендова модель є якісною та адекватною, та має загальний вигляд:

$$y = 133932,9e^{0,034 \times t}$$

На рис. 24 наведено графік трендової складової та емпіричного ряду.

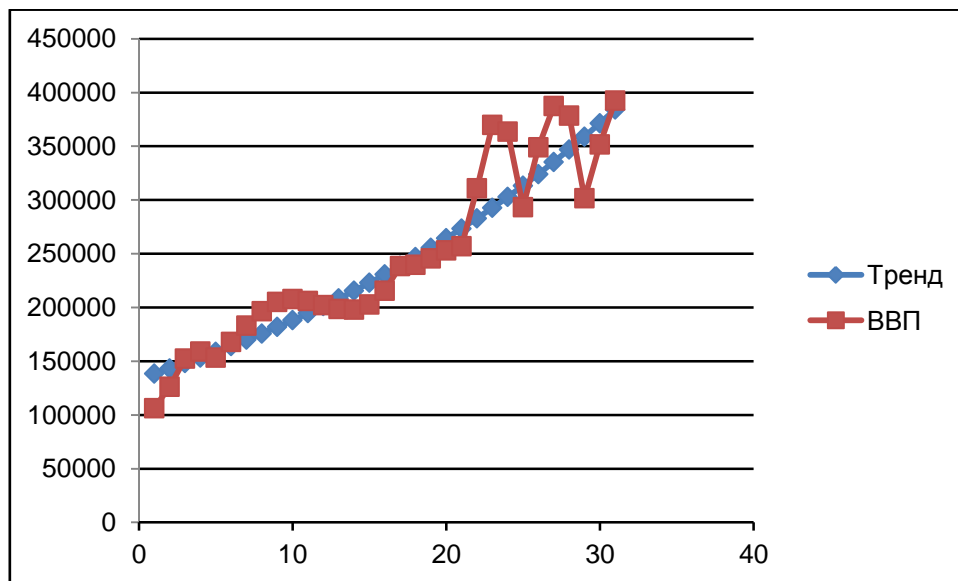


Рис. 24. Графік трендової складової та вихідного динамічного ряду

Як видно з графіка для більш точного моделювання досліджуваного процесу доцільно ураховувати крім загальної тенденції, сезонні та циклічні коливання.

Висновок. У лабораторній роботі було досліджено динамічний ряд ВВП за 7 років. Було виявлено наявність трендової складової в даному динамічному ряді та побудована модель експоненційного тренда.

Лабораторна робота № 4

Дослідження сезонної та циклічної складових часових рядів

Мета – набути навичок виділення циклічної компоненти в ряді динаміки та його моделювання засобами ПП Statistica 8.0. Освоєння навичок розкладання в ряд Фур'є. Вивчення особливостей оцінки адекватності побудованої моделі циклічної складової ряду.

Завдання. Дослідити сезонну та циклічну складову динамічного ряду ВВП України (табл. 8). Побудувати прогноз за трьома складовими та оцінити його якість.

Методичні рекомендації

1. У лабораторній роботі 3 на 4 етапі було проведено декомпозицію часового ряду на такі складові: трендово-циклічну, сезонну і випадкову, використовуючи адитивності модель часового ряду. На рис. 24 наведено результати сезонної декомпозиції обсягу продажів, які необхідно

скопювати в таблицю з вихідними даними, користуючись контекстним меню (Copy with Headers/Paste). На рис. 25 відображена динаміка сезонної складової ряду.

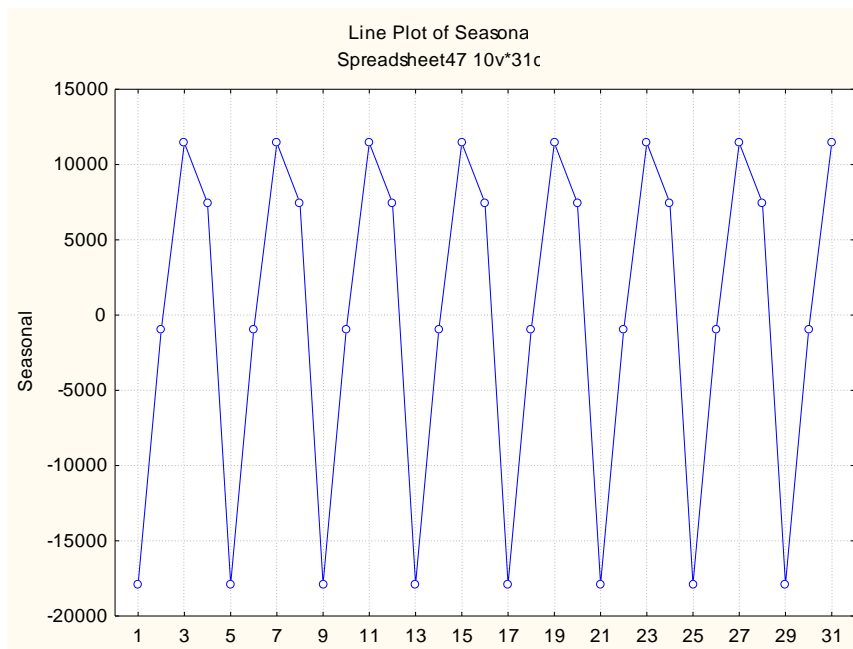


Рис. 25. Графік сезонної складової динамічного ряду

Як видно з графіка сезонна складова має від'ємне значення в першому та другому кварталі року, але найбільше в третьому кварталі, що відповідає особливостям економіки країни.

2. Розрахунок значень циклічної складової (C) можливий шляхом завдання формули розрахунку $C = \text{Smoothed} - \text{Trend}$ в області специфікації змінної Long name (рис. 26).

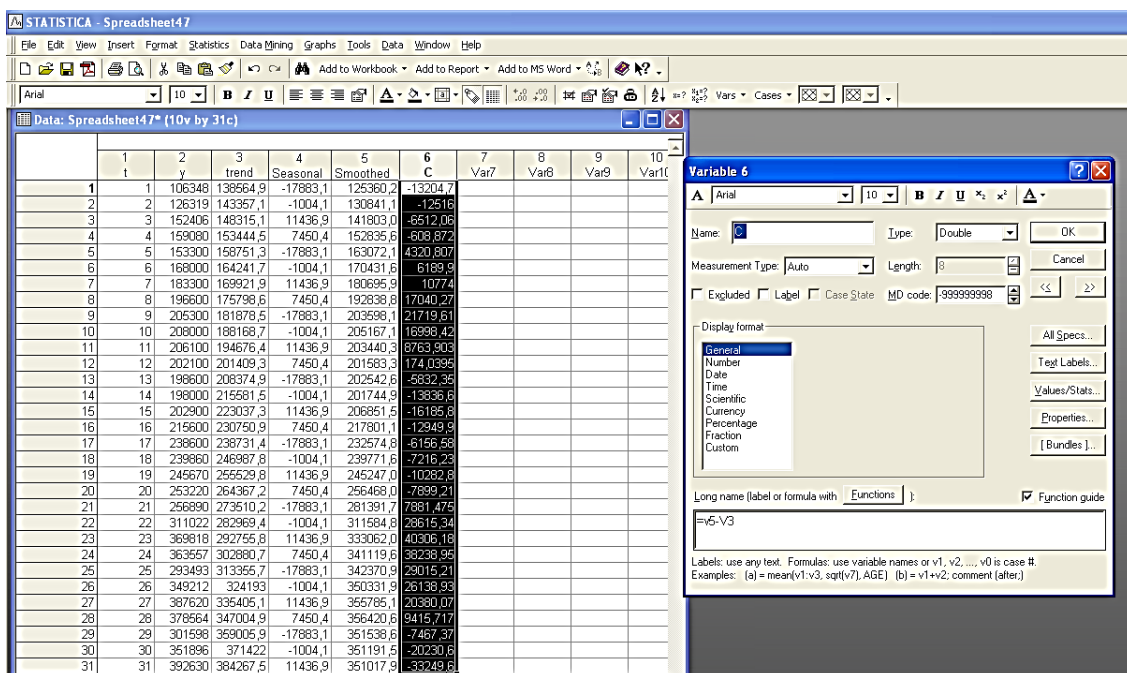


Рис. 26. Розрахунок емпіричних значень циклічної складової

Графік циклічної складової наведено на рис. 27.

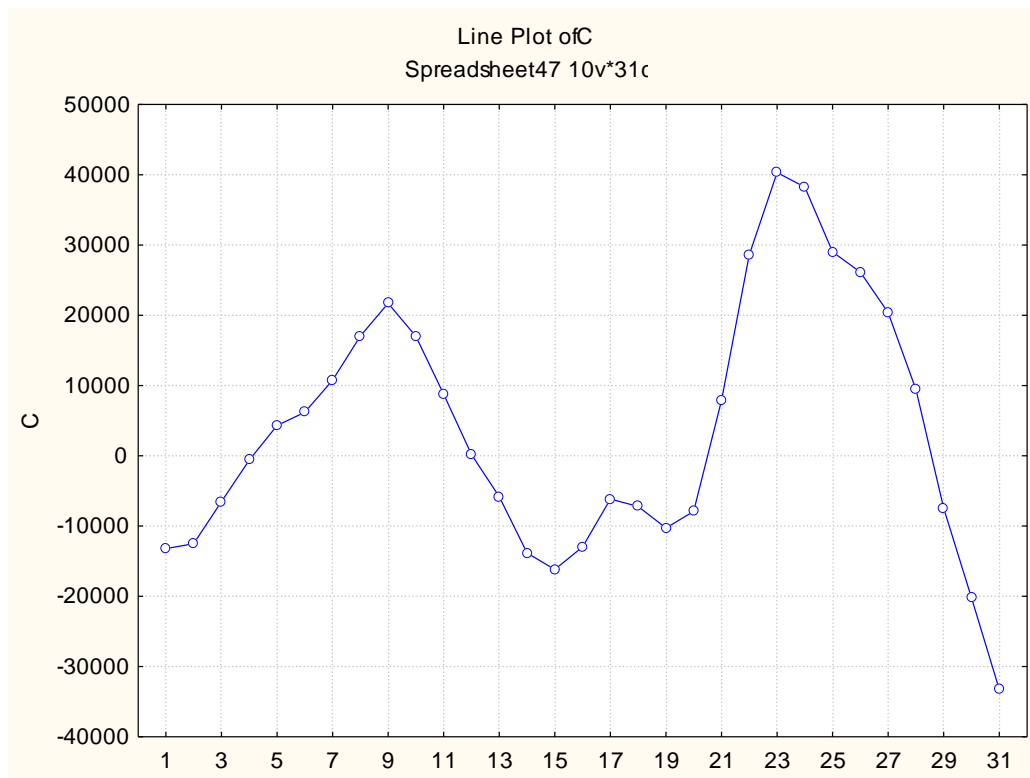


Рис. 27. Графік циклічної складової

Для аналізу циклічної компоненти необхідно побудувати ряд Фур'є. Спочатку проводимо спектральний аналіз ряду в модулі Spectral Analysis і виділяємо основні піки вершини та їх гармоніки (рис. 28).

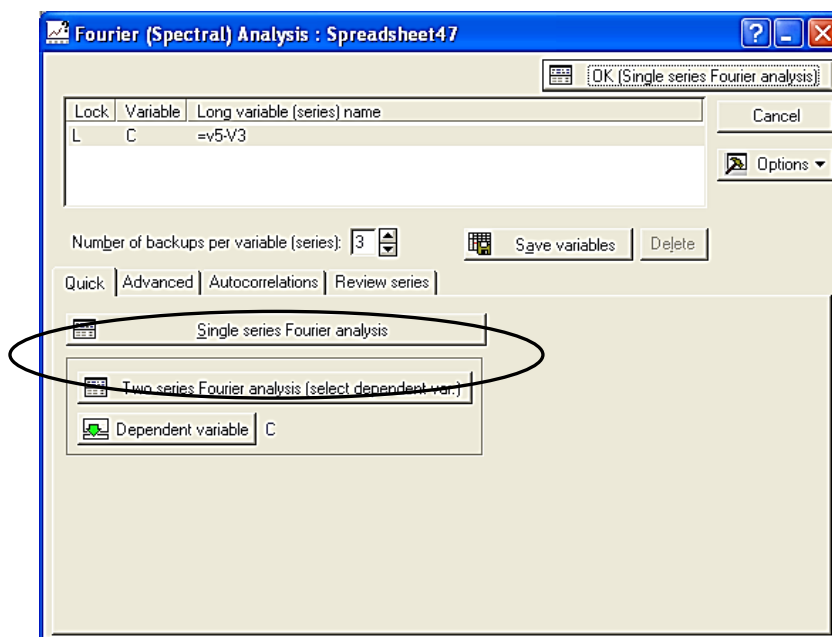


Рис. 28. Діалогове вікно модуля Spectral Analysis

Результати розкладання в ряд Фур'є наведено на рис. 29.

Spectral analysis: C : =v5-V3 (Spreadsheet47)							
No. of cases: 30							
	Frequency	Period	Cosine Coeffs	Sine Coeffs	Periodogram	Density	Hamming Weights
0	0,00000		2691,8	0,00	1,086908E+0	7,047110E+0	0,035714
1	0,03333	30,0000	3116,5	-4198,2	4,100658E+0	1,782097E+0	0,24107
2	0,06666	15,0000	-20583,7	-2055,5	6,418692E+0	3,051656E+0	0,44642
3	0,10000	10,0000	-3052,4	3298,7	3,029819E+0	1,773218E+0	0,24107
4	0,13333	7,5000	2836,7	3369,3	2,909434E+0	4,782181E+0	0,035714
5	0,16666	6,0000	-3197,6	-789,2	1,627128E+0	2,025279E+0	
6	0,20000	5,0000	-1312,9	3319,9	1,911820E+0	1,550039E+0	
7	0,23333	4,2857	2283,3	298,9	7,954125E+0	9,343531E+0	
8	0,26666	3,7500	1036,9	732,4	2,417581E+0	3,839873E+0	
9	0,30000	3,3333	533,3	291,9	5,544898E+0	1,324157E+0	
10	0,33333	3,0000	592,9	381,1	7,453219E+0	7,884593E+0	
11	0,36666	2,7272	675,9	322,8	8,415563E+0	8,333234E+0	
12	0,40000	2,5000	749,8	225,1	9,194436E+0	8,990468E+0	
13	0,43333	2,3076	824,5	-38,4	1,021946E+0	8,371324E+0	
14	0,46666	2,1428	410,8	-262,3	3,563048E+0	7,432949E+0	
15	0,50000	2,0000	899,1	0,00	1,212528E+0	7,860932E+0	

Рис. 29. Результати спектрального аналізу

Як видно з розрахунків 5 найбільш адекватних гармонік мають період 30, 15, 10, 7, 5 та 6 років.

Розрахувати теоретичні значення кожної з гармонік (C_k) за допомогою такої формули в пакеті MS Excel:

$$C_k = a_k \cos\left[\frac{2\pi}{T_k}(t-1)\right] + b_k \sin\left[\frac{2\pi}{T_k}(t-1)\right], \quad (4)$$

де a_k , b_k – це коефіцієнти тригонометричних функцій косинуса та синуса відповідно;

T_k – довжина гармоніки;

t – час.

На рис. 30 наведено розрахунки проведені в середовищі MS Excel для розрахунку теоретичних значень циклічної складової ряду.

для лабораторних работ - Microsoft Excel

Главная Вставка Разметка страницы Формулы Данные Рецензирование Вид Настройки

Вырезать Копировать Вставить Формат по образцу Буфер обмена

Verdana 10

Перенос текста

Числовой

Условное форматирование

Форматировать как таблицу

Стили

E4 =D\$43*SIN((2*PI()/E\$2*(A4-1)))+E\$43*COS((2*PI()/E\$2*(A4-1)))

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1		ВВП,	тренд	Сезонна										
2		млн. грн		складова	30	15	10	8	6	C	прогноз	e	e	e /упр
3					C30	C15	C10	C7,5	C6					
4	1	106348	138565	-17883	-4198	-2056	3299	3369	-789	-375	120307	-374.934	374.934	0.003526
5	2	126319	143357	-1004	-3459	-10250	875	4362	-3164	-11636	130717	-11635.6	11635.58	0.092113
6	3	152406	148315	11437	-2568	-16672	-1884	2468	-2375	-21030	138722	-21029.6	21029.62	0.137984
7	4	159080	153444	7450	-1565	-20211	-3922	-1059	789	-25968	134927	-25968	25968.04	0.163239
8	5	153300	158751	-17883	-493	-20256	-4463	-3885	3164	-25934	114935	-25933.7	25933.7	0.16917
9	6	168000	164242	-1004	600	-16798	-3299	-4141	2375	-21263	141974	-21263.4	21263.38	0.126568
10	7	183300	169922	11437	1667	-10436	-875	-1656	-789	-12089	169270	-12089.1	12089.12	0.065953
11	8	196600	175799	7450	2661	-2269	1884	1925	-3164	1036	184285	1035.965	1035.965	0.005269
12	9	205300	181879	-17883	3538	6290	3922	4232	-2375	15608	179603	15607.89	15607.89	0.076025
13	10	208000	188169	-1004	4261	13762	4463	3738	789	27014	214178	27013.65	27013.65	0.129873
14	11	206100	194676	11437	4798	18854	3299	771	3164	30886	236999	30885.8	30885.8	0.149858
15	12	202100	201409	7450	5125	20686	875	-2706	2375	26354	235214	26353.95	26353.95	0.130401
16	13	198600	208375	-17883	5228	18941	-1884	-4393	-789	17103	207595	17103.47	17103.47	0.08612
17	14	198000	215581	-1004	5103	13921	-3922	-3173	-3164	8765	223343	8765.141	8765.141	0.044268
18	15	202900	223037	11437	4754	6494	-4463	147	-2375	4558	239032	4558.264	4558.264	0.022466
19	16	215600	230751	7450	4198	-2056	-3299	3369	789	3003	241204	3002.67	3002.67	0.013927
20	17	238600	238731	-17883	3459	-10250	-875	4362	3164	-140	220708	-139.916	139.9162	0.000586
21	18	239860	246988	-1004	2568	-16672	1884	2468	2375	-7378	238606	-7377.76	7377.764	0.030759
22	19	245670	255630	11437	1565	-20211	3922	-1059	-789	-16573	250394	-16572.6	16572.55	0.067459
23	20	253220	264367	7450	493	-20256	4463	-3885	-3164	-22349	249468	-22349.2	22349.24	0.08826
24	21	256890	273510	-17883	-600	-16798	3299	-4141	-2375	-20615	235012	-20614.7	20614.73	0.080247
25	22	311022	282969	-1004	-1667	-10436	875	-1656	789	-12095	269871	-12094.7	12094.73	0.038887
26	23	369818	292756	11437	-2661	-2269	-1884	1925	3164	-1725	302468	-1724.86	1724.862	0.004664
27	24	363557	302881	7450	-3538	6290	-3922	4232	2375	5436	315767	5435.746	5435.746	0.014952
28	25	293493	313356	-17883	-4261	13762	-4463	3738	-789	7987	303459	7986.786	7986.786	0.027213
29	26	349212	324193	-1004	-4798	18854	-3299	771	-3164	8365	331553	8364.572	8364.572	0.023953
30	27	387620	335405	11437	-5125	20686	-875	-2706	-2375	9605	356447	9605.404	9605.404	0.02478
31	28	378564	347005	7450	-5228	18941	1884	-4393	789	11993	366448	11992.81	11992.81	0.03168
32	29	301598	359006	-17883	-5103	13921	3922	-3173	3164	12732	353855	12731.82	12731.82	0.042215
33	30	351896	371422	-1004	-4754	6494	4463	147	2375	8724	379142	8724.238	8724.238	0.024792
34	31	392630	384268	11437	-4198	-2056	3299	3369	-789	-375	395329	-374.934	374.934	0.000955
35														1.91816
36														0.061876

Рис. 30. Фрагмент розрахунку циклічної компоненти

На рис. 31. наведено графік кожної гармоніки та загальної циклічної компоненти.

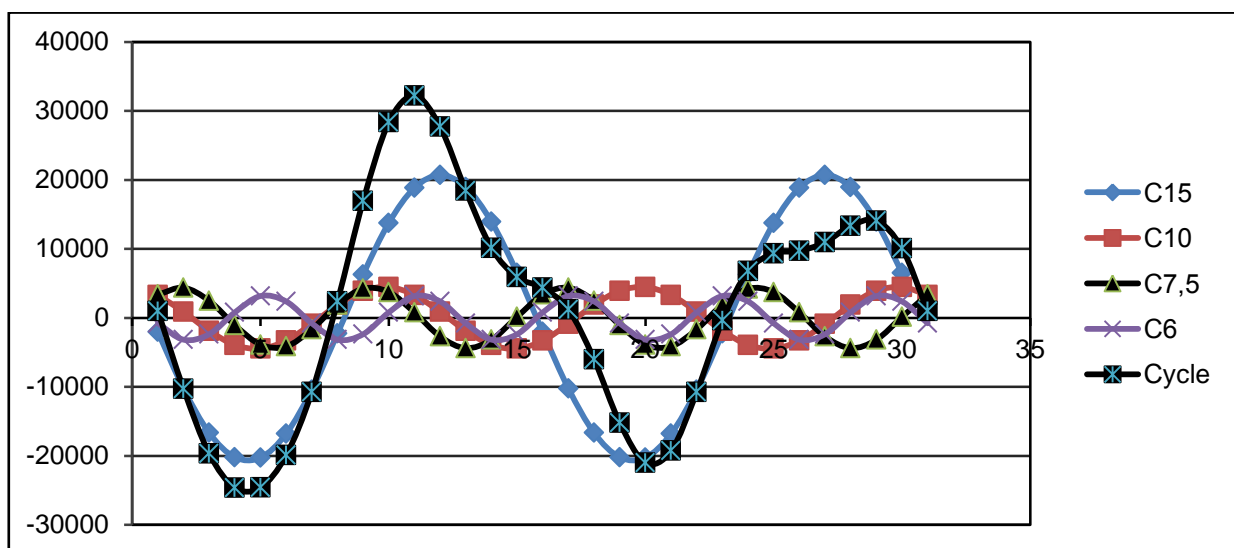


Рис. 31. Графік сезонної компоненти та її складових

Прогнозні значення ряду отримуємо, як додавання компонент ряду. Якість отриманої прогнозової моделі визначаємо за допомогою середньої абсолютної похибки у відсотках (MAPE):

$$MAPE = \sum_{i=n-m+1}^n \frac{100|y_i - \hat{y}_i|}{m|y_i|}, \quad (5)$$

де n – кількість рівнів ряду;

m – кількість пояснювальних змінних у моделі;

\hat{y}_t – оцінки рівнів ряду за моделлю;

\bar{y} – середнє арифметичне значення вибірки.

Для визначення адекватності за даним критерієм використовується така шкала (табл. 10).

Таблица 10

Залежність точності прогнозу від значення критерію

MAPE	Точність прогнозу
Менше 10 %	Висока
10 % – 20 %	Добра
20 % – 40 %	Задовільна
40 % – 50 %	Погана
Більше 50 %	Незадовільна

Розрахунок середньої абсолютної похибки у відсотках (MAPE) наведено на рис. 30. За побудованою моделлю. MAPE = 6 %, що відповідає високій точності прогнозу.

На рис. 32 наведено графік емпіричних та теоретичних значень ВВП та прогноз його на 2014 р.

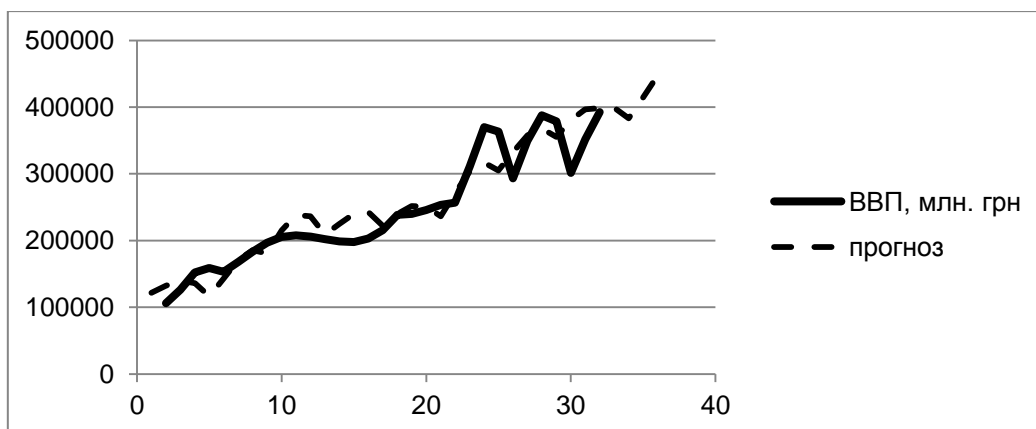


Рис. 32. Теоретичні, емпіричні значення ВВП

Висновок. У лабораторній роботі проведено дослідження циклічної та періодичної компоненти ряду. Розраховано теоретичні значення рівнів ряду за трьома компонентами, проведена оцінка якості прогнозу та побудовано прогноз ВВП на 2014 р.

Лабораторна робота № 5

Оцінка стаціонарності ряду та нормальності розподілу випадкової компоненти. Оцінка наявності автокореляції залишків. Критерій фон Неймана, нециклічний коефіцієнт автокореляції, циклічний коефіцієнт автокореляції. Методи оцінки параметрів моделі з автокорельованими залишками

Мета – закріплення теоретичного та практичного матеріалу, набуття навичок аналізу випадкової складової ряду динаміки в пакеті Statistica 8.0. Набути навичок оцінки наявності автокореляції залишків моделі в ПП Statistica 8.0 та MS Excel. Освоєння методів побудови моделі за допомогою методу Ейткіна.

Завдання. Необхідно побудувати модель впливу показників розвитку регіонів України на рівень іноземних інвестицій та перевірити модель на наявність автокореляції (табл. 11).

Таблиця 11

Показники розвитку регіонів України

Регіон	X1 Частка підприємств колективної форми власності (у % до загальної кількості в регіоні)	X2 Індекс продукції промисловості (%) 1 990 = 100	X3 Індекс продуктивності праці (%) 1 990 = 100	Y Іноземні інвестиції (млн дол США)
1	2	3	4	5
АРК	0,220	0,180	0,170	0,200
Вінницький	0,820	0,540	0,450	0,530
Волинський	0,910	0,460	0,470	0,510
Дніпропетровський	0,960	0,870	0,770	0,850
Донецький	3,490	4,490	4,810	4,050

1	2	3	4	5
Житомирський	0,310	0,840	1,210	0,960
Закарпатський	0,550	2,360	2,230	1,570
Запорізький	0,130	0,450	0,690	0,880
Івано-Франківський	0,770	0,810	0,830	0,800
Київський	0,250	0,460	0,550	0,510
Кіровоградський	0,500	1,080	1,040	0,800
Луганський	0,120	0,390	0,530	0,750
Львівський	0,220	0,180	0,170	0,200
Миколаївський	0,820	0,540	0,450	0,560
Одеський	0,910	0,460	0,470	0,570
Полтавський	0,960	0,870	0,770	0,850
Рівненський	0,160	0,710	0,880	0,860
Сумський	0,110	0,070	0,270	0,140
Тернопільський	0,020	0,140	0,250	0,220
Харківський	0,010	0,030	0,070	0,090
Херсонський	0,043	0,136	0,154	0,175
Хмельницький	0,969	0,073	0,182	0,201
Черкаський	0,172	0,105	0,178	0,289
Чернівецький	0,108	0,150	0,181	0,352
Чернігівський	0,064	0,189	0,354	0,979

Методичні рекомендації

9.6.1.1 Побудова багатofакторної економетричної моделі

Для побудови і всебічного аналізу множинних лінійних економетричних моделей у пакеті Statistica 7 передбачений модуль Multiple Regression (Множинна регресія). Щоб приступити до обчислювальних процедур, необхідно увійти в позицію меню Statistics/Multiple Regression. У стартовій панелі цього модуля необхідно задати змінні для аналізу. Результати побудови лінійної економетричної моделі наведені в діалоговому вікні (рис. 33). У верхній частині вікна міститься основна інформація про моделі, в нижній

частині знаходяться функціональні кнопки, що дозволяють всебічно розглянути результати аналізу.

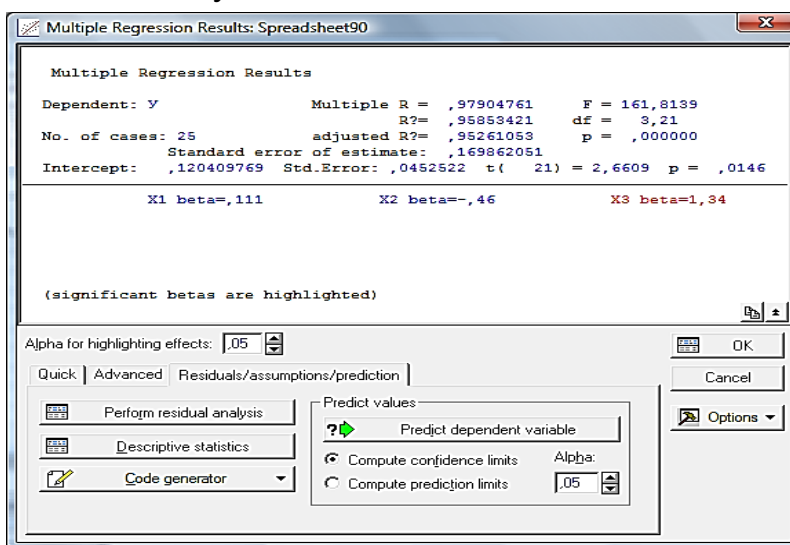


Рис. 33. Вікно результатів регресійного аналізу

Ініціювавши кнопку Summary: Regression results (Результати регресійного аналізу), визначаються найважливіші характеристики моделі і ступінь її адекватності (рис. 34).

Regression Summary for Dependent Variable: Y (Spreadsheet90)						
R= .97904761 R²= .95853421 Adjusted R²= .95261053 F(3,21)=161,81 p<,00000 Std.Error of estimate: .16986						
N=25	Beta	Std.Err. of Beta	B	Std.Err. of B	t(21)	p-level
Intercept			0,12041	0,04525	2,6609	0,01462
X1	0,11119	0,08100	0,12225	0,08906	1,3726	0,18433
X2	-0,45817	0,34634	-0,38190	0,28868	-1,3229	0,20009
X3	1,33885	0,33458	1,07875	0,26958	4,0015	0,00064

Рис. 34. Результати регресійного аналізу

Проаналізувати отримані результати моделі:

$R = 0,979$ – коефіцієнт множинної кореляції;

$R^2 = 0,9585$ – коефіцієнт детермінації моделі;

Adjusted $R^2 = 0,9526$ – скоригований коефіцієнт детермінації на число спостережень і число параметрів;

$F(3,21) = 161,81$ – критерій адекватності Фішера;

Std.Error of estimate = 0,16986 – середнє квадратичне відхилення помилок моделі;

$B(a_0, a_1, a_2, a_3) = (0,12; 0,122; -1,38; 1,0788)$ – параметри моделі, отже, модель має вигляд: $Y = 0,12 + 0,122 \times X_1 - 0,38 \times X_2 + 1,0788 \times X_3$.

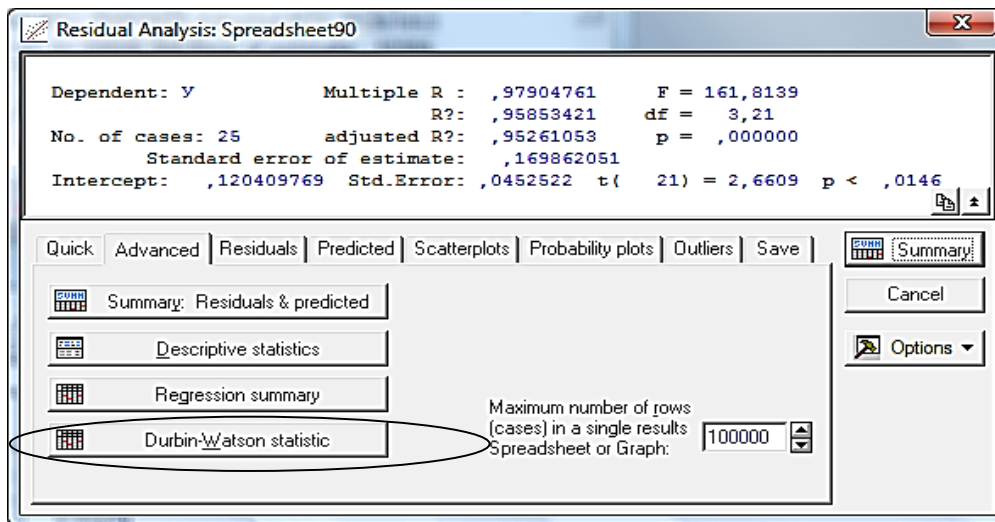


Рис. 36. Модуль аналізу помилок

Результат розрахунку коефіцієнта Дарбіна–Уотсона наведено на рис. 37.

	Durbin-Watson d (Spreadsheet9 and serial correlation of residua	
	Durbin-Watson d	Serial Corr.
Estimate	1,30236	0,20136

Рис. 37. Результат розрахунку критерію Дарбіна–Уотсона за допомогою пакета Statistica

Критичні значення критерію дорівнюють $d_l = 1,12$; $d_u = 1,65$. Тобто $d_l < d < d_u$, це свідчить що розраховане значення критерію перебуває у зоні невизначеності, та використання одного критерію недостатньо. Коефіцієнт циклічної кореляції (Serial Corr.) складає 0,2, що свідчить про наявність позитивної автокореляції.

2. Розрахунок критерію фон Неймана

Розрахунок критерію фон Неймана ведеться за такою формулою:

$$Q = \frac{\frac{\sum(u_t - u_{t-1})^2}{n-1}}{\frac{\sum u_e^2}{n}} = DW \frac{n}{n-1} = 1,3 \times \frac{25}{24} = 1,354. \quad (6)$$

Критичні значення критерію становлять 1,42 і 2,99. Значить у моделі позитивна автокореляція помилок.

3. Розрахунок нециклічного коефіцієнта автокореляції

Нециклічний коефіцієнт автокореляції показує вплив кожного наступного значення помилки на попередній і розраховується за такою формулою:

$$r = \frac{\sum_{t=2}^n (u_t \times u_{t-1}) - \frac{1}{n-1} (\sum_{t=1}^n u_{t-1}) (\sum_{t=2}^n u_t)}{\sqrt{\left[\sum_{t=1}^n u_{t-1}^2 - \frac{1}{n-1} (\sum_{t=1}^n u_{t-1})^2 \right] \left[\sum_{t=2}^n u_t^2 - \frac{1}{n-1} (\sum_{t=2}^n u_t)^2 \right]}} = 0,19. \quad (7)$$

Позитивне значення нециклічного коефіцієнта кореляції свідчить про наявність позитивної автокореляції в моделі.

У роботі була побудована багатофакторна регресійна модель, що є адекватною за коефіцієнтом Фішера. Перевірка залишків моделі на автокореляцію, довело, що в моделі існує позитивна автокореляція. Наявність автокореляції в моделі є порушенням побудови класичної регресійної моделі, тому для побудови якісної та адекватної моделі, необхідно використовувати не метод найменших квадратів, а інші методи оцінки параметрів моделі.

Метод Ейткіна відрізняється від звичайного МНК тим, що у ході оцінки параметрів моделі використовується матриця Ω , що відображає коригування вихідних даних на мінливість дисперсії. Параметри моделі визначаються за формулою:

$$a = (X^T \times \Omega^{-1} \times X)^{-1} X^T \times \Omega^{-1} \times Y; \quad (8)$$

$$\Omega^{-1} = \begin{matrix} 1 & -\rho & 0 & 0 & 0 & \dots & 0 \\ \frac{1}{1-\rho^2} & -\rho & 1+\rho^2 & -\rho & 0 & 0 & \dots & 0 \\ 0 & -\rho & 1+\rho^2 & -\rho & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & \dots & 1 \end{matrix} \quad (9)$$

На практиці для розрахунку ρ використовується те співвідношення, що й для визначення циклічного коефіцієнта кореляції:

$$\rho \approx r^0 \approx \frac{\sum_{t=1}^n u_t \times u_{t-1}}{\sum_{t=1}^n (u_t)^2}. \quad (10)$$

Як видно з розрахунків на рис. 36 значення циклічного коефіцієнта кореляції складає 0,2, тобто параметр $\rho = 0,2$.

Подальші розрахунки було проведено в пакеті MS Excel, із використанням вбудованих функцій.

1. Введення вихідних даних у робочий аркуш пакета MS Excel.

Вихідні дані для побудови багатofакторної моделі наведено на рис. 38.

	A	B	C	D
1	X1	X2	X3	y
2	0,09	0,36	0,20	0,22
3	0,12	0,74	0,53	0,82
4	0,21	0,64	0,51	0,91
5	0,84	0,91	0,85	0,96
6	11,33	1,33	4,05	3,49
7	8,10	0,35	0,96	0,31
8	0,78	0,56	1,57	0,55
9	0,30	0,08	0,88	0,13
10	1,03	0,48	0,80	0,77
11	0,98	0,26	0,51	0,25
12	0,17	0,36	0,80	0,50
13	0,25	0,08	0,75	0,12
14	0,09	0,36	0,20	0,22
15	0,12	0,74	0,56	0,82
16	0,21	0,64	0,57	0,91
17	0,84	0,91	0,85	0,96
18	0,57	0,44	0,86	0,16
19	3,75	0,08	0,14	0,11
20	0,02	0,10	0,22	0,02

Рис. 38. Вихідні дані для побудови багатofакторної моделі

2. Проведення проміжних розрахунків

Транспонування в пакеті MS Excel проводиться за допомогою функції ТРАНСП (масив) (рис. 39, 40).

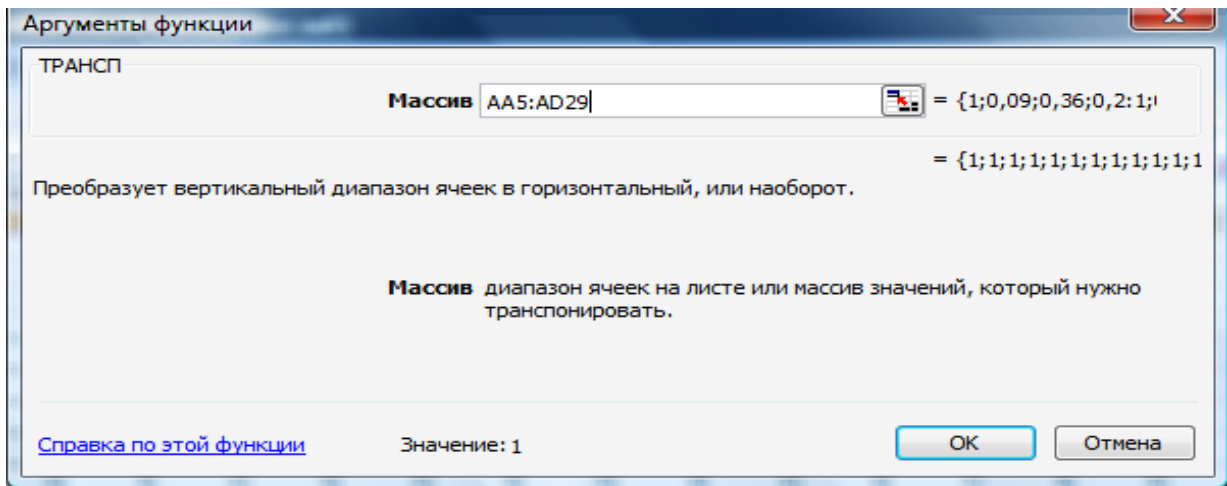


Рис. 39. Вікно функції транспонування

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
31		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
32		0,09	0,12	0,21	0,84	11,33	8,1	0,78	0,3	1,03	0,98	0,17	0,25	0,09	0,12	0,21	0,84	0,57	3,75	0,02	0,09	0,54	0,339	0,005	0,314	0,617
33	X^T	0,36	0,74	0,84	0,91	1,33	0,35	0,56	0,08	0,48	0,26	0,38	0,08	0,38	0,74	0,64	0,91	0,44	0,08	0,1	0,02	0,2	0,187	0,154	0,187	0,249
34		0,2	0,53	0,51	0,85	4,05	0,96	1,57	0,88	0,8	0,51	0,8	0,75	0,2	0,58	0,57	0,85	0,88	0,14	0,22	0,09	0,175	0,201	0,289	0,352	0,979

Рис. 40. Значення матриці X^T

З урахуванням значення циклічного коефіцієнта автокореляції та формули (9) було сформовано матрицю Ω^{-1} (рис. 41).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
5		1,0417	-0,2083	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6		-0,2083	1,0833	-0,2083	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7		0	-0,2083	1,0833	-0,2083	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8		0	0	-0,2083	1,0833	-0,2083	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9		0	0	0	-0,2083	1,0833	-0,2083	0	0	0	0	0	0	0	0	0	0	0	0	0
10		0	0	0	0	-0,2083	1,0833	-0,2083	0	0	0	0	0	0	0	0	0	0	0	0
11		0	0	0	0	0	-0,2083	1,0833	-0,2083	0	0	0	0	0	0	0	0	0	0	0
12		0	0	0	0	0	0	-0,2083	1,0833	-0,2083	0	0	0	0	0	0	0	0	0	0
13		0	0	0	0	0	0	0	-0,2083	1,0833	-0,2083	0	0	0	0	0	0	0	0	0
14		0	0	0	0	0	0	0	0	-0,2083	1,0833	-0,2083	0	0	0	0	0	0	0	0
15		0	0	0	0	0	0	0	0	0	-0,2083	1,0833	-0,2083	0	0	0	0	0	0	0
16		0	0	0	0	0	0	0	0	0	0	-0,2083	1,0833	-0,2083	0	0	0	0	0	0
17		0	0	0	0	0	0	0	0	0	0	0	-0,2083	1,0833	-0,2083	0	0	0	0	0
18	Ω^{-1}	0	0	0	0	0	0	0	0	0	0	0	0	-0,2083	1,0833	-0,2083	0	0	0	0
19		0	0	0	0	0	0	0	0	0	0	0	0	0	-0,2083	1,0833	-0,2083	0	0	0
20		0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0,2083	1,0833	-0,2083	0	0
21		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0,2083	1,0833	-0,2083	0
22		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0,2083	1,0833	-0,2083
23		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0,2083	1,0833
24		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0,2083
25		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
26		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
27		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
28		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
29		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Рис. 41. Результат розрахунку матриці Ω^{-1}

Перемноження двох матриць у пакеті MSeXcel проводиться за допомогою функції МУМНОЖ(масив1; масив2) (рис. 42).

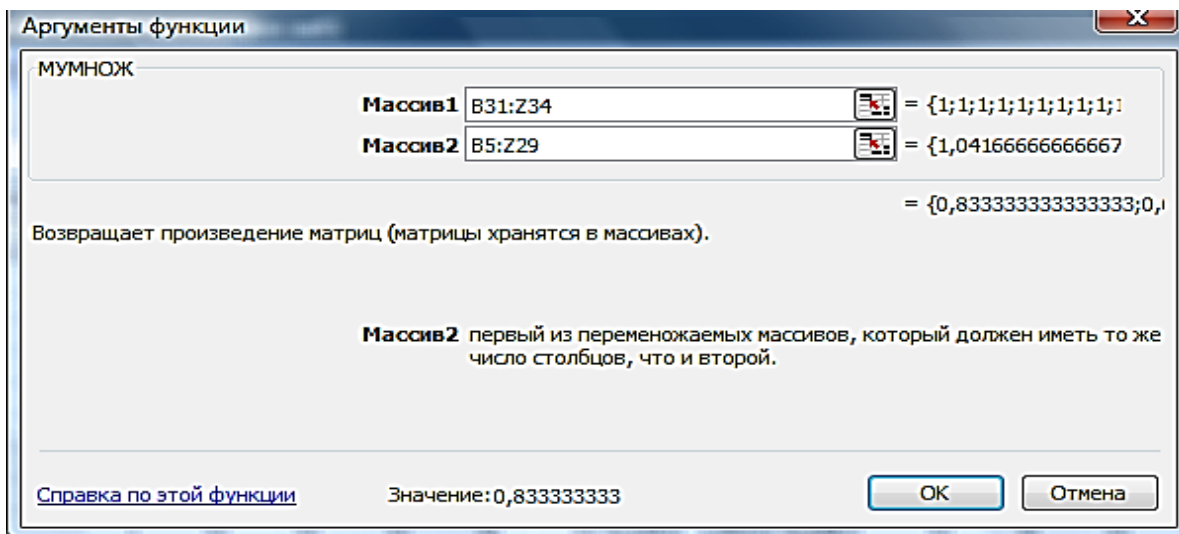


Рис. 42. Вікно реалізації функції МУМНОЖ

Для отримання матриці зворотної даній матриці, використовується функція МОБР (масив) (рис. 43).

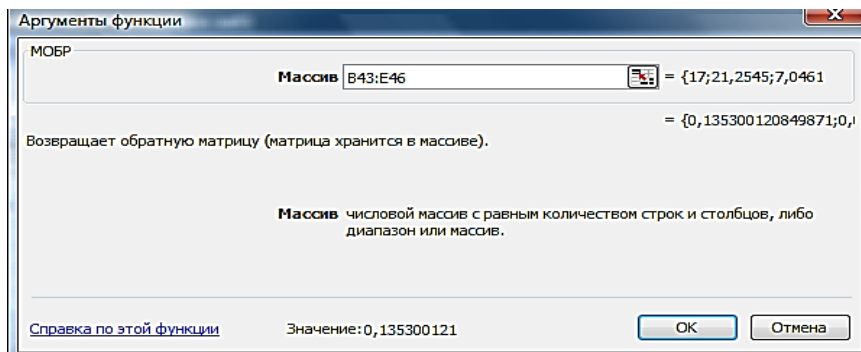


Рис. 43. Вікно реалізації функції МОБР

Використовуючи наведені функції проводяться розрахунки за формулою (8). Результати розрахунків наведено на рис. 44.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	
35																				
36																				
37		0,8333	0,6667	0,6667	0,6667	0,6667	0,6667	0,6667	0,6667	0,6667	0,6667	0,6667	0,6667	0,6667	0,6667	0,6667	0,6667	0,6667	0,6667	0
38		0,0688	0,0675	0,0275	-1,4942	10,4117	6,2521	-0,905	-0,0521	0,8492	0,8117	-0,0721	0,2167	0,0204	0,0675	0,0275	0,7475	-0,3388	3,9396	-0
39	$X^T \Omega^{-1} =$	0,2208	0,5933	0,3496	0,5754	1,17833	-0,0146	0,5171	-0,13	0,4492	0,1067	0,3192	-0,0633	0,2192	0,5933	0,3496	0,7608	0,2704	-0,0258	0
40		0,0979	0,4263	0,265	-0,0292	4,01042	-0,1308	1,3175	0,4596	0,5771	0,2192	0,6042	0,6042	-0,0563	0,4463	0,3238	0,6229	0,7254	-0,0733	0
41																				
42																				
43		17	21,255	7,0462	12,127															
44		21,255	184,34	16,02	48,214															
45	$X^T \Omega^{-1} X =$	7,0462	16,02	5,1688	8,9094															
46		12,127	48,214	8,9094	23,137															
47																				
48																				
49																				
50		0,1353	0,0009	-0,1837	-0,0021															
51	$(X^T \Omega^{-1} X)^{-1} =$	0,0009	0,0125	0,017	-0,0331															
52		-0,1837	0,017	0,8514	-0,2671															
53		-0,0021	-0,0331	-0,2671	0,2161															
54																				
55		9,1113		-0,2578																
56	$X^T \Omega^{-1} y =$	38,748	a	0,0422																
57		8,1051		1,1369																
58		17,786		0,378																

Рис. 44. Розрахунок параметрів багатофакторної регресійної моделі методом Ейткіна

3. Формування загального вигляду багатофакторної регресійної моделі

Загальний вид моделі:

$$Y = -0,2578 + 0,0422 \times X_1 + 1,1269 \times X_2 + 0,378 \times X_3.$$

4. Перевірка отриманої моделі на наявність автокореляції залишків

З метою перевірки, усунення автокореляції залишків у моделі доцільно розрахувати коефіцієнт Дарбіна–Уотсона, за формулою:

$$DW = \frac{\sum_{t=2}^T (u_t - u_{t-1})^2}{\sum_{t=1}^T u_t^2} = 1,72. \quad (11)$$

Критичні значення критерію дорівнюють $d_l = 1,12$; $d_u = 1,65$. Тобто $d_u < d < 2$, це свідчить про відсутність автокореляції моделі.

Таким чином модель побудована з використанням методу Ейткіна не має автокореляції в залишках.

Висновки. У роботі були проведені розрахунки параметрів моделі з використанням методу Ейткіна, що дозволило усунути автокореляцію в залишках моделі.

Лабораторна робота № 6

Визначення довжини лагової моделі та побудова моделі з кінченою довжиною лага. Моделі адаптивних очікувань та часткового корегування

Мета – набути навичок побудови адекватної лагової моделі в пакеті Statistica 8.0.

Завдання. Необхідно проаналізувати дані, що відображають залежність прибутку компанії (y) від витрат на маркетингові дослідження (x), що наведені в табл. 12. Визначити ефект, що дають додаткові вкладення у маркетингові дослідження від 1 тис. грн та середнього лага, що існує між вкладенням коштів у маркетингові дослідження та отриманням прибутку від цих вкладень.

Таблиця 12

Вихідні дані

Період (місяць)	Прибутки компанії тис. грн	Витрати на маркетингові дослідження, тис. грн	Період (місяць)	Прибутки компанії тис. грн	Витрати на маркетингові дослідження, тис. грн
1	988	60	11	1 281	78
2	1 035	66	12	1 253	67
3	1 089	73	13	1 302	74
4	1 082	67	14	1 382	94
5	1 073	54	15	1 426	93
6	1 126	65	16	1 468	91
7	1 177	75	17	1 513	93
8	1 234	83	18	1 593	96
9	1 265	83	19	1 612	98
10	1 258	74	20	1 628	93

Методичні рекомендації

1. Визначення оптимальної довжини лагової моделі

З метою визначення довжини максимального лага, необхідно побудувати кореляційну матрицю для змінних із різним періодом затримки. Нехай максимальний лаг затримки ефекту від маркетингових досліджень дорівнює 4.

У пакеті Statistica 7 сформувавши таблицю наведену на рис. 45.

Data: лаб7* (12v by 20c)											
	1 t	2 y	3 x	4 y1	5 x1	6 y2	7 x2	8 y3	9 x3	10 y4	11 x4
1	1	988	60	1035	66	1089	73	1082	67	1073	54
2	2	1035	66	1089	73	1082	67	1073	54	1126	65
3	3	1089	73	1082	67	1073	54	1126	65	1177	75
4	4	1082	67	1073	54	1126	65	1177	75	1234	83
5	5	1073	54	1126	65	1177	75	1234	83	1265	83
6	6	1126	65	1177	75	1234	83	1265	83	1258	74
7	7	1177	75	1234	83	1265	83	1258	74	1281	78
8	8	1234	83	1265	83	1258	74	1281	78	1253	67
9	9	1265	83	1258	74	1281	78	1253	67	1302	74
10	10	1258	74	1281	78	1253	67	1302	74	1382	94
11	11	1281	78	1253	67	1302	74	1382	94	1426	93
12	12	1253	67	1302	74	1382	94	1426	93	1468	91
13	13	1302	74	1382	94	1426	93	1468	91	1513	93
14	14	1382	94	1426	93	1468	91	1513	93	1593	96
15	15	1426	93	1468	91	1513	93	1593	96	1612	98
16	16	1468	91	1513	93	1593	96	1612	98	1628	93
17	17	1513	93	1593	96	1612	98	1628	93		
18	18	1593	96	1612	98	1628	93				
19	19	1612	98	1628	93						
20	20	1628	93								

Рис. 45. Вихідні дані для розрахунку кореляційної матриці

Для того, щоб виконувати обчислювальні процедури необхідно увійти в меню Statistics/Basic Statistics/Tables. У вікні необхідно вибрати пункт Correlation matrices (Кореляційна матриця). Отримана матриця парних кореляцій наведена на рис. 6.

Correlations (лаб6)											
Marked correlations are significant at p < ,05000											
N=16 (Casewise deletion of missing data)											
Variable	y	x	y1	x1	y2	x2	y3	x3	y4	x4	
y	1,00	0,86	0,98	0,78	0,95	0,71	0,95	0,76	0,95	0,77	
x	0,86	1,00	0,81	0,72	0,74	0,45	0,72	0,45	0,71	0,49	
y1	0,98	0,81	1,00	0,87	0,98	0,80	0,96	0,77	0,95	0,74	
x1	0,78	0,72	0,87	1,00	0,81	0,73	0,76	0,51	0,72	0,44	
y2	0,95	0,74	0,98	0,81	1,00	0,88	0,98	0,84	0,96	0,74	
x2	0,71	0,45	0,80	0,73	0,88	1,00	0,83	0,77	0,78	0,51	
y3	0,95	0,72	0,96	0,76	0,98	0,83	1,00	0,90	0,98	0,82	
x3	0,76	0,45	0,77	0,51	0,84	0,77	0,90	1,00	0,87	0,78	
y4	0,95	0,71	0,95	0,72	0,96	0,78	0,98	0,87	1,00	0,89	
x4	0,77	0,49	0,74	0,44	0,74	0,51	0,82	0,78	0,89	1,00	

Рис. 46. Матриця парних кореляцій

Найбільше значення коефіцієнта кореляції спостерігається якщо лаг дорівнює 3. Тому доцільно зробити висновок, що довжина лагової моделі складатиме 3.

2. Побудова лагової моделі

Для побудови лагової моделі в стартовому вікні модуля Time Series Analysis (рис. 47) необхідно ініціювати кнопку Distributed lags analysis.

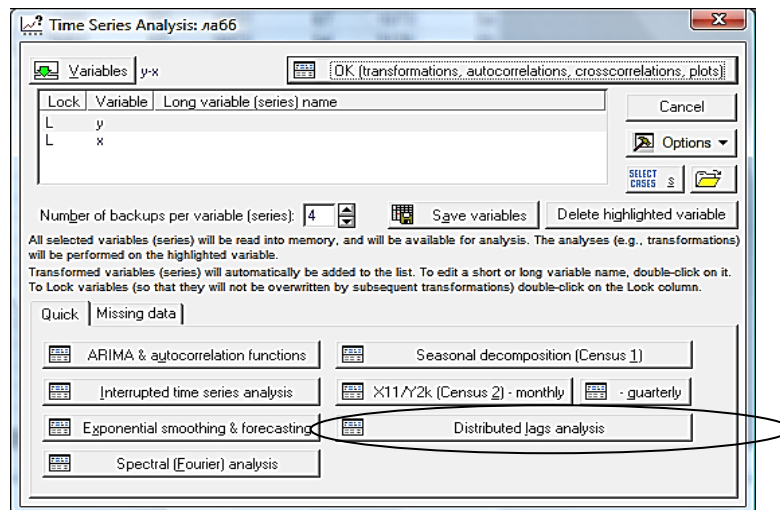


Рис. 47. Стартове вікно модуля Time Series Analysis

Для розрахунку параметрів лагової моделі необхідно задати незалежну змінну та кількість лагів у моделі (рис. 48).

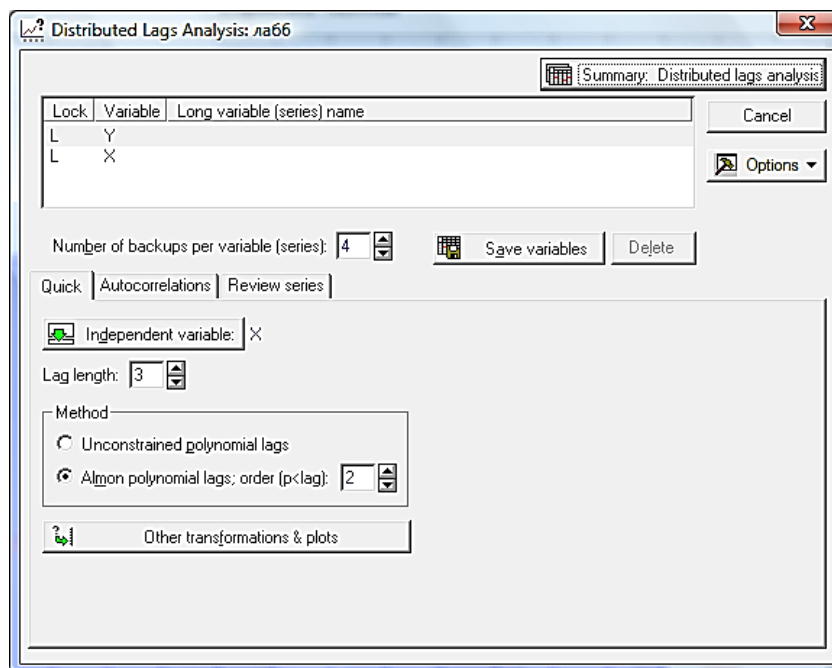


Рис. 48. Завдання параметрів побудови лагової моделі

Таким чином були розраховані параметри моделі з лагом 3. Результати розрахунків параметрів моделей наведені на рис. 49.

Almon Polyn. Distr.Lags; Regression Coefficients (пабб)				
Indep: X Dep: Y				
Lag: 3 Polyn. order: 2 R= ,9997 R-square= ,9994 N: 17				
Lag	Regressn Coeff.	Standard Error	t(13)	p
0	7,44077376427	0,87908432486	8,46423210357	0,00000119947
1	2,79036748363	0,65862210923	4,23667448224	0,00097095472
2	1,88942869082	0,66647833217	2,83494391306	0,01405679933
3	4,73795738584	0,89370391483	5,30148442589	0,00014349907

Рис. 49. Параметри моделі для лага 3

Проаналізувати зазначену модель більш детально. Модель має загальний вигляд:

$$Y_t = 7,44 \times X_t + 2,79 \times X_{t-1} + 1,88 \times X_{t-2} + 4,74 \times X_{t-3}.$$

Розрахунок суми квадратів помилок та коефіцієнта Фішера наведені на рис. 50.

Almon Polyn. Distr.Lags; Analysis of Variance (пабб)					
Indep: X Dep: Y					
Lag: 3 Polyn. order: 2 R= ,9997 R-square= ,9994 N					
Effect	Sums of Squares	df	Mean Square	F	p
Regress.	3074234	4	768558	5312,11	0,000001
Residual	18808	13	1447		
Total	3076115				

Рис. 50. Результати дисперсійного аналізу моделі

З рис. 46, 47 видно, що модель є адекватною за критерієм Фішера, всі параметри моделі значимими. Таким чином дану модель можна використовувати для дослідження ефекту від маркетингових досліджень.

3. Економічний аналіз моделі

3.1. Визначення короткострокового мультиплікатора моделі.

У лагових моделях короткостроковий мультиплікатор дорівнює параметру при X_t і в даному випадку складає 7,44. Це означає, що збільшення витрат на маркетингові дослідження на 1 тис. грн призведе до збільшення прибутку на 7,44 тис. грн у тому ж періоді.

3.2. Розрахунок відносних коефіцієнтів регресії.

Відносні коефіцієнти регресії лагової моделі розраховуються за формулою:

$$\beta_i = \frac{a_i}{\sum_{i=0}^n a_i}, \quad (12)$$

де a_i – параметри моделі з i -тим лагом затримки.

За формулою 2 можна отримати такі результати:

$$\beta_0 = \frac{7,44}{16,91} = 0,44; \quad \beta_1 = \frac{2,79}{16,91} = 0,16; \quad \beta_2 = \frac{1,89}{16,91} = 0,11; \quad \beta_3 = \frac{4,74}{16,91} = 0,28.$$

Отже, 44 % загального збільшення обсягу прибутку, викликане зростанням витрат на маркетингові дослідження в поточному періоді; 16 % у момент часу $(t + 1)$ і найбільше; 11 % в моменті часу $(t + 2)$; 28 % у момент часу $(t + 3)$.

3.3. Визначення середнього лага моделі.

У лагових моделях для визначення середньої довжини лага використовується формула:

$$L_{cp} = \sum_{i=0}^n (\beta_i \times i). \quad (13)$$

Середній лаг у цій моделі визначається як:

$$L_{cp} = 0 \times 0,44 + 1 \times 0,16 + 2 \times 0,11 + 3 \times 0,28 = 1,33 \text{ (року)}.$$

Таким чином, усередненому збільшенню витрат на маркетингові дослідження призведе до збільшення прибутку підприємства через 1 рік і 4 місяці.

Висновки. У роботі була побудована лагова модель за допомогою якої було проаналізовано ефект, що дають додаткові витрати на маркетингові дослідження.

Лабораторна робота № 7

Прості методи прогнозування часових рядів.

Прогнозування за допомогою кривих зростання

Мета – закріплення теоретичних та практичних знань із прогнозування за допомогою простих методів. Отримання навичок прогнозування за допомогою кривих зростання в пакеті MS Excel.

Завдання. Побудувати прогноз ВВП на 4 квартал 2013 р. відповідно до даних наведених в табл. 8 за допомогою простих методів прогнозування. Обрати вид кривої зростання для досліджуваного показника та розрахувати

за нею прогноз на одну точку. Порівняти результати прогнозу за різними методами.

Методичні рекомендації

1. Побудова прогнозу методом середнього абсолютного приросту.

1.1. Середній абсолютний приріст що розраховується за такою формулою:

$$\bar{\Delta} = \frac{y_n - y_1}{n - 1} = \frac{392\,630 - 106\,348}{30} = 9\,542,733 \text{ млрд грн.}$$

1.2. Розрахунок теоретичних значень ряду, наведено в табл. 13 Δ_t^2 :

Таблиця 13

Розрахунок за методом середнього абсолютного приросту

Період	y_t	Δ_t	\hat{y}_t	$(\hat{y}_t - y_t)^2$	Δ_t^2
1	2	3	4	5	6
1 кв 2006	106 348		106 348		
2 кв 2006	126 319	19 971	115 890,7	108 748 745,7	398 840 841
3 кв 2006	152 406	26 087	125 433,5	727 517 554,4	680 531 569
4 кв 2006	159 080	6 674	134 976,2	580 993 174,4	44 542 276
1 кв 2007	153 300,00	-5 780	144 518,9	77 107 131,8	33 408 400
2 кв 2007	168 000,00	14 700	154 061,7	194 277 136,1	216 090 000
3 кв 2007	183 300,00	15 300	163 604,4	387 916 659,4	234 090 000
4 кв 2007	196 600,00	13 300	173 147,1	550 036 954,9	176 890 000
1 кв 2008	205 300,00	8 700	182 689,9	511 218 129,4	75 690 000
2 кв 2008	208 000,00	2 700	192 232,6	248 610 902,8	7 290 000
3 кв 2008	206 100,00	-1 900	201 775,3	18 702 741,78	3 610 000
4 кв 2008	202 100,00	-4 000	211 318,1	84 972 753,07	16 000 000
1 кв 2009	198 600,00	-3 500	220 860,8	495 543 216,6	12 250 000
2 кв 2009	198 000,00	-600	230 403,5	10 499 889 72	360 000
3 кв 2009	202 900,00	4 900	239 946,3	1 372 425 874	24 010 000
4 кв 2009	215 600,00	12 700	249 489	1 148 464 321	161 290 000

Закінчення табл. 13

1	2	3	4	5	6
1 кв 2010	238 600,00	23 000	259 031,7	417 455 727	529 000 000
2 кв 2010	239 860	1 260	268 574,5	824 520 596	1 587 600
3 кв 2010	245 670	5 810	278 117,2	1 052 820 788	33 756 100
4 кв 2010	253 220	7 550	287 659,9	1 186 109 008	57 002 500
1 кв 2011	256 890	3 670	297 202,7	1 625 111 094	13 468 900
2 кв 2011	311 022	54 132	306 745,4	18 289 307,56	2 930 273 424
3 кв 2011	369 818	58 796	316 288,1	2 865 446 625	3 456 969 616
4 кв 2011	363 557	-6 261	325 830,9	1 423 261 136	39 200 121
1 кв 2012	293 493	-70 064	335 373,6	1 753 984 656	4 908 964 096
2 кв 2012	349 212	55 719	344 916,3	18 452 752,11	3 104 606 961
3 кв 2012	387 620	38 408	354 459,1	1 099 647 500	1 475 174 464
4 кв 2012	378 564	-9 056	364 001,8	212 057 668,8	82 011 136
1 кв 2013	301 598	-76 966	373 544,5	5 176 303 659	5 923 765 156
2 кв 2013	351896	50 298	383 087,3	972 895 116,3	2 529 888 804
3 кв 2013	392 630	40 734	392 630	1,35525E-20	1 659 258 756
Всього	7 615 603	286 282		26 202 879 901	28 829 820 720

1.3. Перевірка відповідності обраного методу. Якщо нерівність $\sigma_{\text{ост}}^2 \leq \rho^2$ виконується, то вважається що метод середнього абсолютного приросту можна використовувати для прогнозування.

$$\sigma_{\text{ост}}^2 = \frac{26202879901}{31} = 845254190,4;$$

$$\rho^2 = \frac{1}{2} \times \frac{28829820720}{31} = 464997108,4.$$

Відповідно до проведених розрахунків видно, що основна умова використання методу не виконується, тому для прогнозування ВВП доцільно використати інші методи.

2. Побудова прогнозу методом середнього темпу росту.

2.1. Розрахунок середнього темпу росту для досліджуваної сукупності:

$$\bar{T}_p = \sqrt[n-1]{\frac{y_n}{y_1}} = 1,044.$$

2.2. Розрахунок прогнозу ВВП на 4 квартал 2014 р.:

$$\hat{y}_{4 \text{ квартал}} = 392\,630 \times 1,044 = 410\,102,1 \text{ млрд. грн.}$$

Виходячи з того, що ВВП в 4 кварталі склало 404 311, то отриманий прогноз відхиляється від реальних даних на 2%, що свідчить про адекватність вибору методу прогнозування.

Лабораторна робота № 8 Прогнозування часових рядів за допомогою штучних нейронних мереж

Мета – отримання навичок проведення побудови штучної нейронної мережі в пакеті Statistica 8.0 та прогнозування її основи.

Є дані про надходження податку з доходів фізичних осіб (ПДФО) до місцевого бюджету за 2004 – 2013 рр .

Графік, що демонструє залежність надходжень податку від часу наведений на рис. 51.

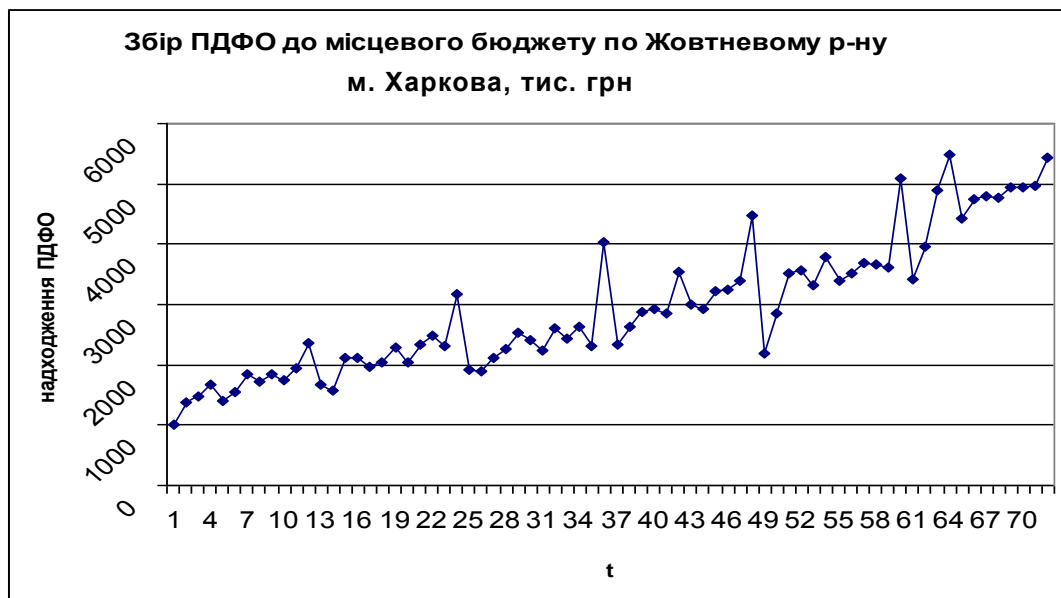


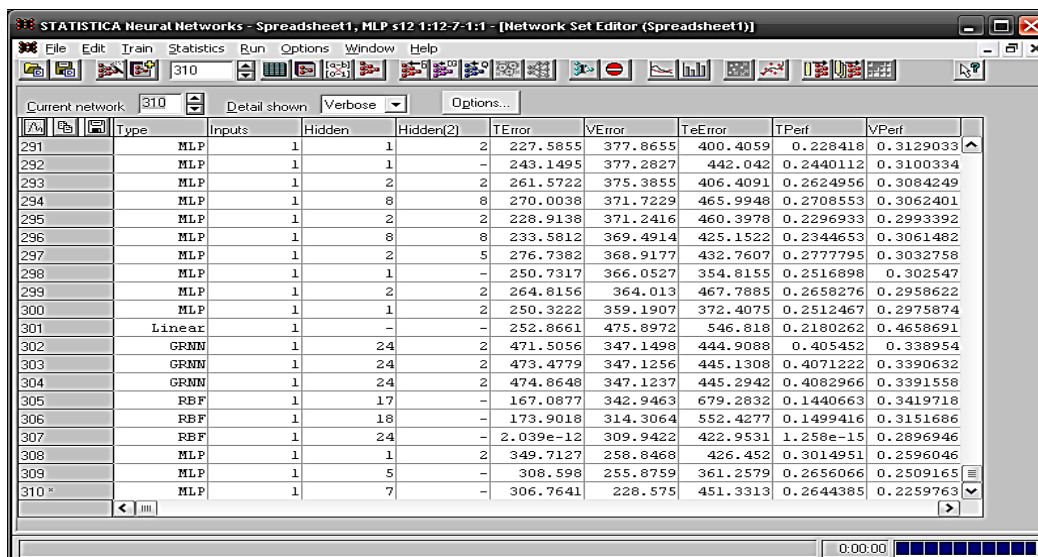
Рис. 51. Динаміка надходжень прибуткового податку до місцевого бюджету

Необхідно побудувати економіко-математичну модель, яка б з одного боку відображала залежність рівня надходжень податку від часу,

а з іншого – давала б можливість побудувати точний прогноз. Використати для цього модель побудовану на базі нейронної мережі.

Для рішення завдання прогнозування скористуємося STATISTICA Neural Networks.

Результати побудови нейронних мереж наведені на рис. 52.



№	Type	Inputs	Hidden	Hidden(2)	TError	VError	TeError	TPerf	VPerf
291	MLP	1	1	2	227.5855	377.8655	400.4059	0.228418	0.3129033
292	MLP	1	1	-	243.1495	377.2827	442.042	0.2440112	0.3100334
293	MLP	1	2	2	261.5722	375.3855	406.4091	0.2624956	0.3084249
294	MLP	1	8	8	270.0038	371.7229	465.9948	0.2708553	0.3062401
295	MLP	1	2	2	228.9138	371.2416	460.3978	0.2296933	0.2993392
296	MLP	1	8	8	233.5812	369.4914	425.1522	0.2344653	0.3061482
297	MLP	1	2	5	276.7382	368.9177	432.7607	0.2777795	0.3032758
298	MLP	1	1	-	250.7317	366.0527	354.8155	0.2516898	0.302547
299	MLP	1	2	2	264.8156	364.013	467.7885	0.2658276	0.2958622
300	MLP	1	1	2	250.3222	359.1907	372.4075	0.2512467	0.2975874
301	Linear	1	-	-	252.8661	475.8972	546.818	0.2180262	0.4658691
302	GRNN	1	24	2	471.5056	347.1498	444.9088	0.405452	0.338954
303	GRNN	1	24	2	473.4779	347.1256	445.1308	0.4071222	0.3390632
304	GRNN	1	24	2	474.8648	347.1237	445.2942	0.4082966	0.3391558
305	RBF	1	17	-	167.0877	342.9463	679.2832	0.1440663	0.3419718
306	RBF	1	18	-	173.9018	314.3064	552.4277	0.1499416	0.3151686
307	RBF	1	24	-	2.039e-12	309.9422	422.9531	1.258e-15	0.2896946
308	MLP	1	1	2	349.7127	258.8468	426.452	0.3014951	0.2596046
309	MLP	1	5	-	308.598	255.8759	361.2579	0.2656066	0.2509165
310	MLP	1	7	-	306.7641	228.575	451.3313	0.2644385	0.2259763

Рис. 52. Набір мереж

Графічне зображення найкращої нейронної мережі наведено на рис. 53.

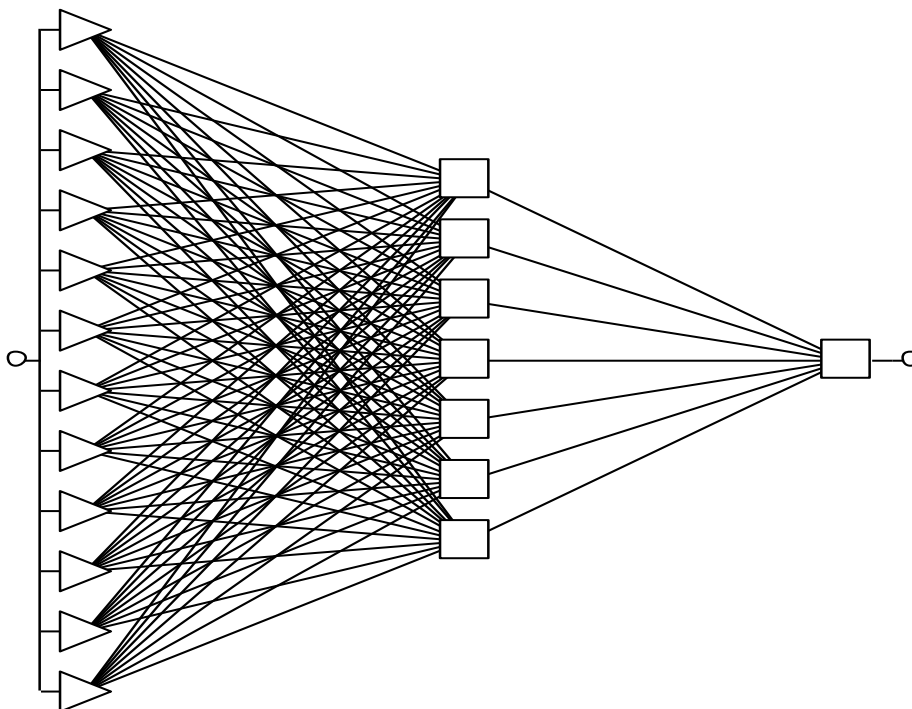


Рис. 53. Ілюстрація мережі

Для оцінки роботи нейронної мережі розглянемо статистику моделі (рис. 54).

Regression Statistics			
Variable	1		
Run			
	Tr. PDFD	Ve. PDFD	Te. PDFD
Data Mean	3350.481	3384.33	2809.268
Data S.D.	1184.742	1025.987	820.686
Error Mean	-6.510551	38.45781	95.8772
Error S.D.	313.2913	231.8487	453.8161
Abs E. Mean	230.7337	192.5209	315.255
S.D. Ratio	0.2644385	0.2259763	0.5529717
Correlation	0.9645663	0.9745855	0.8509273

Рис. 54. Статистика моделі

Виходячи з результатів класифікації добре видно, що мережа має незначні помилки. Таким чином, можна сказати, що помилки мережі незначні. Ряд помилок зображено на рис. 55.

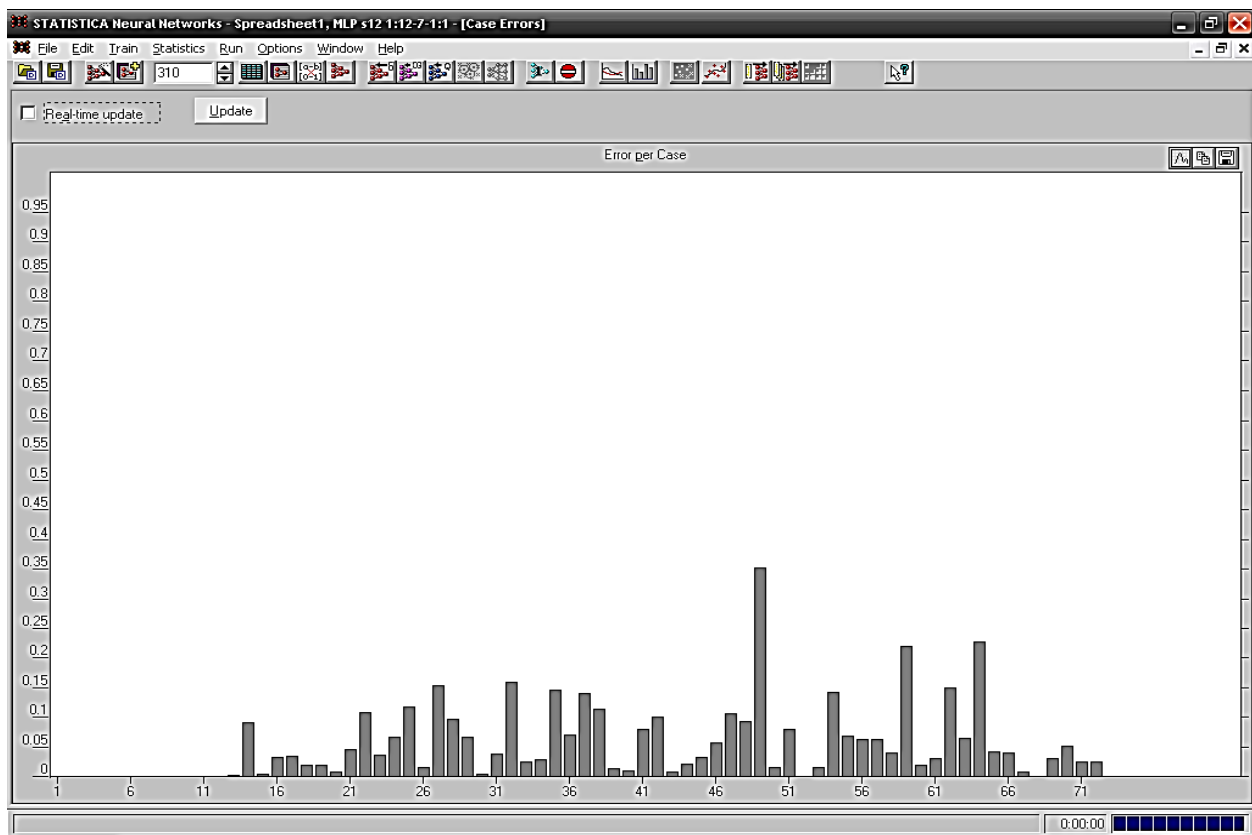


Рис. 55. Ряд помилок

На рис. 56 наведено порівняльний графік реальних значень динаміки зміни надходження податку на прибуток та прогнозні значення побудовані з використанням нейронної мережі.

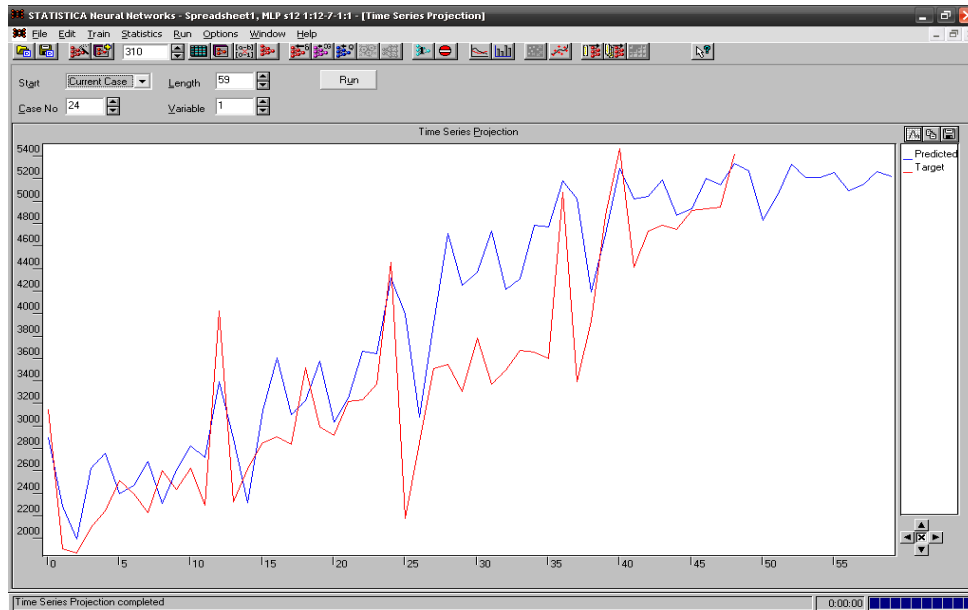


Рис. 56. Графік прогнозних значень динаміки зміни надходжень від прибуткового податку

Найважливіший показник якості моделі – це коефіцієнт кореляції (рис. 57).

```
0:00:06 An improved network has been found
0:00:06 An improved network has been found
0:00:07 An improved network has been found
0:00:15 An improved network has been found
0:00:16 An improved network has been found
0:00:17 An improved network has been found
0:00:17 An improved network has been found
0:01:01 An improved network has been found
0:01:06 An improved network has been found
0:01:43 An improved network has been found
0:02:16 An improved network has been found
0:02:24 An improved network has been found
0:02:54 An improved network has been found
0:03:24 An improved network has been found
...search has finished
718 networks were tested, 10 retained
The best network found had good performance (regression ratio 0.225976, correlation 0.974586)
```

Рис. 57. Значення коефіцієнта кореляції

Для моделі, що аналізується значення коефіцієнта кореляції становить 0,975, що вказує на досить високу якість побудованої моделі.

Таким чином модель багатоварового персептрона дозволяє отримати прогноз із коефіцієнтом кореляції – 97 %. Проте для побудови більш якіс-

ного прогнозу необхідна наявність більшої вибірки даних.

Аналіз побудованої моделі показав, що використання нейронних мереж доцільно під час прогнозування надходжень прибуткового податку.

Рекомендована література

1. Раєвнєва О. В. Лабораторний практикум з навчальної дисципліни "Економетрика II" для студентів напряму підготовки "Прикладна статистика" денної форми навчання / укл. О. В. Раєвнєва, О. І. Горохова, І. В. Чанкіна. – Х. : Вид. ХНЕУ, 2011. – 52 с.
2. Моделі та методи соціально-економічного прогнозування : підручник / В. М. Геєць, Т. С. Клебанова, О. І. Черняк, 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 / А. В. Шашок. – Рубцовск : Рубцовский 2010. – 81 с.
4. Офіційний сайт асоціації Українських банків. – Режим доступу : www.aub.org.ua.
5. Офіційний сайт компанії Statsoft. – Режим доступу : www.statsoft.ru.
6. Портал Статосфера. – Режим доступа : www.statosphere.ru/videos/62-dmsess11.html.
7. Электронный учебник StatSoft [Электронный ресурс]. – Режим доступа : // <http://www.statsoft.ru>.
8. Экономическое прогнозирование : учебник [Электронный ресурс]. – Режим доступа : //<http://www.statsoft.ru>.

