

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ ЕКОНОМІЧНИЙ УНІВЕРСИТЕТ
ІМЕНІ СЕМЕНА КУЗНЕЦЯ

СУЧАСНІ МЕТОДИ І ТЕХНОЛОГІЇ ОБРОБКИ
ІНФОРМАЦІЇ В ОСВІТІ

Методичні рекомендації
до лабораторних робіт
для здобувачів спеціальності
011 "Освітні, педагогічні науки"
освітньо-наукової програми
"Освітні, педагогічні науки"
третього (освітньо-наукового) ступеня

Харків
ХНЕУ ім. С. Кузнеця
2023

УДК [004:37](072.034)

С91

Укладачі: О. В. Раєвнєва
О. І. Бровко
К. А. Стрижиченко

Затверджено на засіданні кафедри статистики і економічного прогнозування.

Протокол № 4 від 10.11.2022 р.

Самостійне електронне текстове мережеве видання

Сучасні методи і технології обробки інформації в освіті [Електронний ресурс] : методичні рекомендації до лабораторних робіт для здобувачів спеціальності 011 "Освітні, педагогічні науки" освітньо-наукової програми "Освітні, педагогічні науки" третього (освітньо-наукового) ступеня / уклад. О. В. Раєвнєва, О. І. Бровко, К. А. Стрижиченко. – Харків : ХНЕУ ім. С. Кузнеця, 2023. – 43 с.

Подано методичні рекомендації до виконання лабораторних робіт, метою яких є закріплення й поглиблення знань теоретичного та практичного матеріалу з методології й технології обробки інформації в освіті та ефективного використання сучасних інформаційно-економічних методів і моделей дослідження освітніх процесів.

Рекомендовано для здобувачів спеціальності 011 "Освітні, педагогічні науки" освітньо-наукової програми "Освітні, педагогічні науки" третього (освітньо-наукового) ступеня.

УДК [004:37](072.034)

© Харківський національний економічний університет імені Семена Кузнеця, 2023

Вступ

Початок ХХІ століття характеризує суспільство як інформаційне, в якому здійснюється інформатизація всіх галузей науки й освіти. Інформатизація суспільства – це глобальний соціальний процес, особливість якого полягає в тому, що домінуючим видом діяльності в сфері суспільного виробництва є нагромадження, збирання, продукування, зберігання, оброблення, передавання та використання інформації. Ці процеси здійснюються на основі сучасних засобів процесорної та обчислювальної техніки, а також на базі різноманітних засобів інформаційного обміну. Перед системою вищої освіти також стоїть багато завдань, вирішення яких потребує використання сучасних методів і технологій аналізу освітніх процесів.

Сучасні методи і технології оброблення інформації в освіті – це порівняно новий напрям статистичної науки, що утворився від поєднання математики, економічної статистики та економетрики. Завдання навчальної дисципліни в межах економетрики полягає не лише в тому, щоб виявляти закони та зв'язки, які об'єктивно існують, а й описувати їх математичними методами. Економічна статистика акумулює всю інформацію про процеси, що відбуваються в реальному освітньому просторі, та уособлюють практичний досвід.

Навчальна дисципліна "Сучасні методи і технології обробки інформації в освіті" є обов'язковою навчальною дисципліною та вивчається згідно з навчальним планом підготовки здобувачів спеціальності 011 "Освітні, педагогічні науки" денної та заочної форм навчання. Вона містить теоретичні і практичні знання про якісні та кількісні властивості освітніх систем, оцінювання взаємозв'язку кількісних показників розвитку освітньої галузі, а також економетричні моделі освітніх процесів у вітчизняному та європейському просторі.

Методичні рекомендації до лабораторних робіт

Змістовий модуль 1

Засади моделювання освітніх процесів

Тема 1. Статистичне моделювання як сучасний метод обробки інформації

Методичні рекомендації до лабораторної роботи за темою 1 "Ознайомлення з пакетом *Statistica 13.0*. Первинний аналіз даних"

Мета – закріплення теоретичного й практичного матеріалу, набуття навичок роботи в модулі *Basic Statistics / Tables*.

Завдання – необхідно провести аналіз варіаційного ряду для вибіркового ряду банків України в модулі *Basic Statistics / Tables* ППП *Statistica*:

1. Розрахувати статистичні характеристики ряду (середнє, дисперсію, середньоквадратичне відхилення, моду, медіану, розмах, коефіцієнти асиметрії й ексцесу).

2. Побудувати гістограму й полігон розподілу випадкової величини, зробити висновки про характер закону розподілу.

3. За допомогою критеріїв Колмогорова – Смирнова і Пірсона перевірити гіпотезу про нормальний закон розподілу.

4. Зробити висновки щодо угруповання банків за величиною відповідного показника.

Методичні рекомендації

Для вирішення й аналізу завдань такого типу в ППП *Statistica* передбачений модуль *Basic Statistics / Tables* (*Основні статистики й таблиці*). Варто розглянути порядок роботи в цьому модулі.

1. Запуск *Statistica* і підготовка даних. У меню програм вибрати програму *Statistica*, після її запуску обрати у меню пункт *File-New* для підготовки власних даних. З'явиться діалогове вікно, у якому необхідно вказати кількість змінних (*Number of variables*) і кількість випадків (*Number of Cases*). Після введення натисніть кнопку вікна *OK* (рис. 1.1).

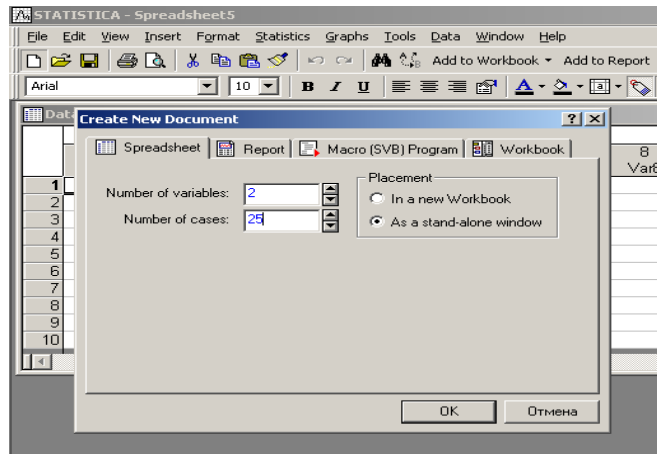


Рис. 1.1. Завдання кількості змінних і спостережень

З'явиться порожнє поле, що становить таблицю розміру 25 × 2 : 25 спостережень, 2 змінні (рис. 1.2).

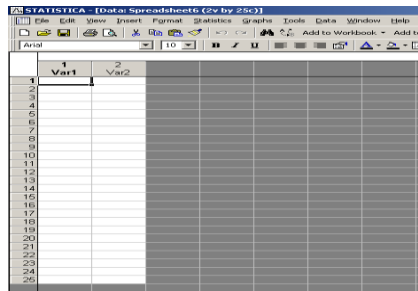


Рис. 1.2. Порожнє поле даних

Кожен елемент даних, тобто значення показника, займає одну комірку поля даних. Після заповнення всіх комірок поля даних буде одержано таблицю, аналогічну наведеній на рис. 1.3.

	1 Region	2 Kilkist zdobyvachiv
1	Vinnnytsya	1187477
2	Volyn	793821
3	Dnipropetrovsk	876148
4	Donetsk	389719
5	Zhytomyr	459234
6	Zakarpattya	451074
7	Zaporizhzhya	328131
8	Ivano-Frankivsk	209010
9	Kyiv	273945
10	Kirovohrad	167741
11	Luhansk	70674
12	Lviv	232158
13	Mykolayiv	111158
14	Odesa	175292
15	Poltava	145468
16	Rhne	132243
17	Sumy	120243
18	Temopyl	100261
19	Kharkiv	404232
20	Kherson	114054
21	Khmelnytskyi	42607
22	Cherkasy	35241
23	Chernivtsi	71296
24	Chernihiv	91436
25	city Kyiv	125384

Рис. 1.3. Вихідні дані

2. Розрахунки. Щоб приступити до обчислювальних процедур необхідно увійти в позицію меню *Statistics / Basic Statistics / Tables*. Після підтвердження вибору модуля з'явиться діалогове вікно, що дозволяє задати напрям аналізу *Descriptive statistics (Описові статистики)*, подане на рис. 1.4.

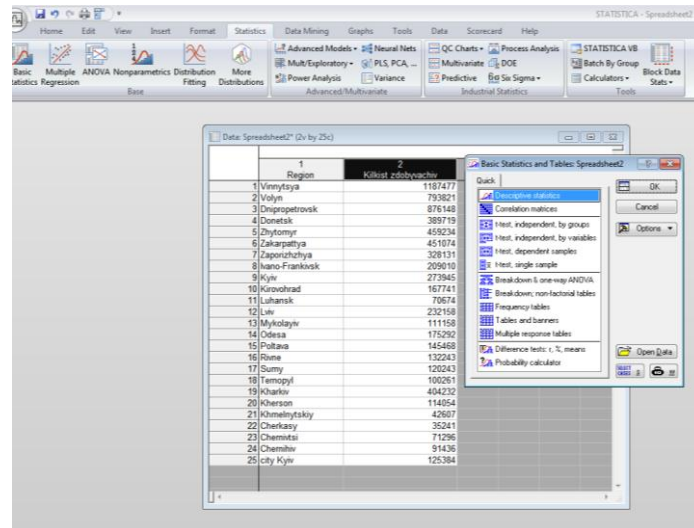


Рис. 1.4. Вибір напрямку аналізу

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

Ініціюйте кнопку *Variable (Змінні)* і у вікні, що з'явилося, укажіть показники, за якими здійснюють аналіз. Після зазначення змінних треба підтвердити вибір натисканням кнопки *OK*. Далі, ініціювавши вкладку *Advanced*, необхідно виділити основні статистики для розрахунку (рис. 1.5).

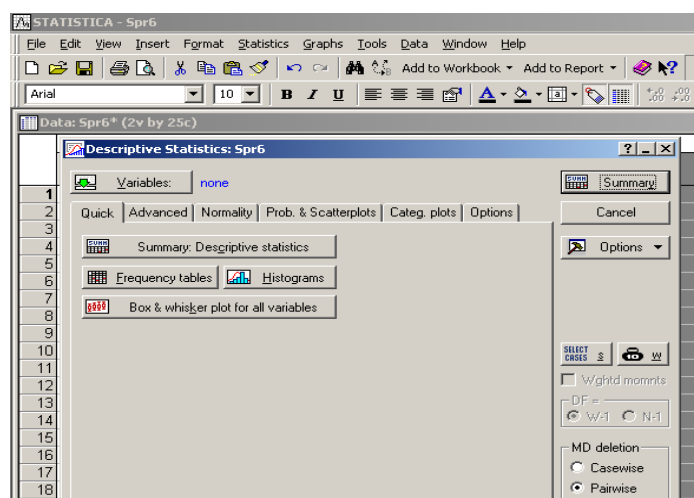


Рис. 1.5. Стартова панель модуля

Такими є: *Valid N* (число спостережень), *Mean* (середнє), *Sum* (сума значень), *Median* (медіана), *Mode* (мода), *Standard Deviation* (середньо-квадратичне відхилення), *Variance* (дисперсія), *Skewness* (коефіцієнт асиметрії), *Kurtosis* (коефіцієнт ексцесу), *Min & Max* (мінімум і максимум), *Range* (розмах вибірки) (рис. 1.6).

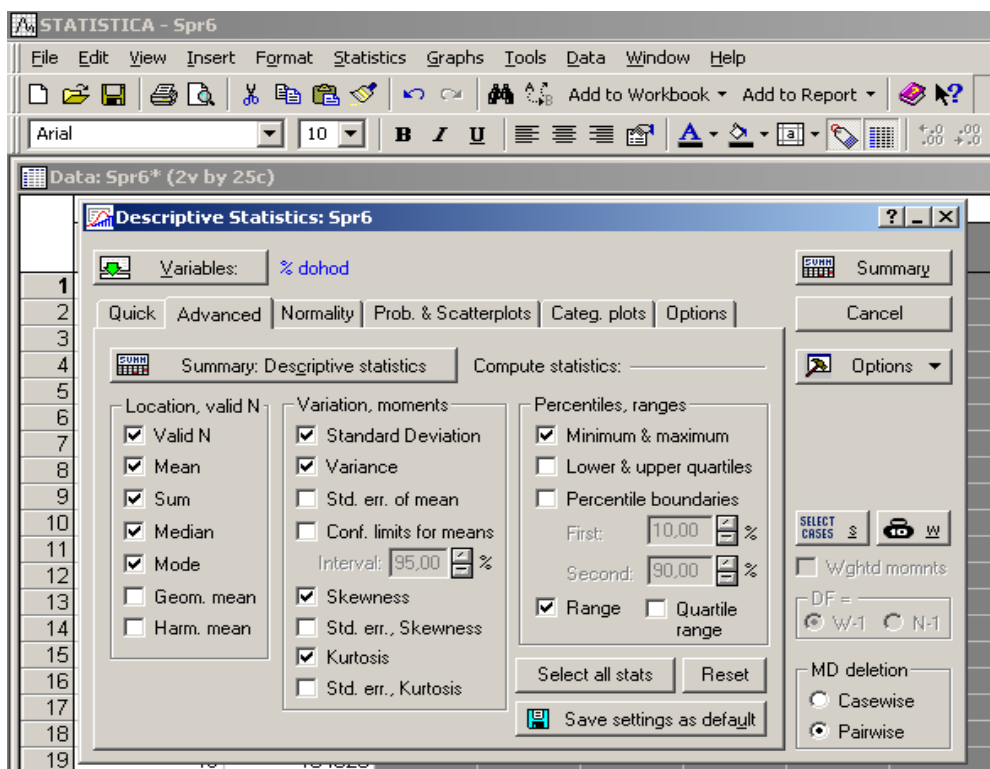


Рис. 1.6. Вибір описових статистик

Результати розрахунку описових статистик для цієї вибірки буде одержано натисканням клавіші *Summary*, і вони наведені на рис. 1.7.

Descriptive Statistics (Spreadsheet2)													
Variable	Valid N	Mean	Median	Mode	Frequency of Mode	Sum	Minimum	Maximum	Variance	Std.Dev.	Coef.Var.	Skewness	Kurtosis
Kilkist zdobyvachiv	25	284321,9	167741,0	Multiple	1	7108047	35241,00	1187477	8,256114E+10	287334,5	101,0596	1,906466	3,461855

Рис. 1.7. Описові статистики

Для наочності подання досліджуваної сукупності варто побудувати полігон розподілу. Для цього необхідно в меню *Graphs / 2D Graphs / Line Plots (variables)* обрати змінну й побудувати лінійний графік значень (рис. 1.8).

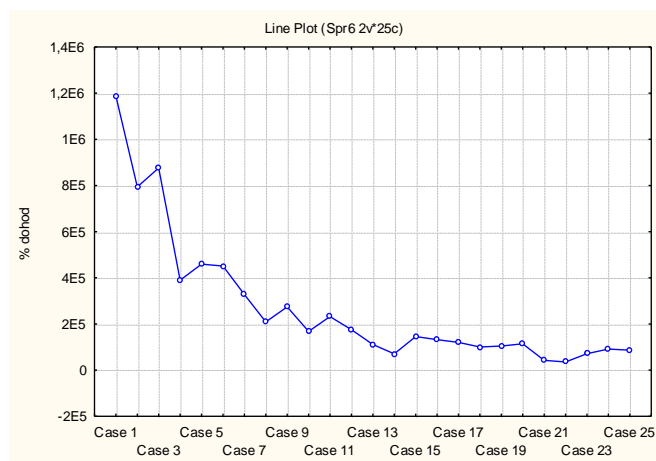


Рис. 1.8. Полігон розподілу

Подальший аналіз здійснюють в рамках перевірки вибірки на нормальний закон розподілу. Для проведення угруповання вибірки в стартовій панелі модуля обрати вкладку *Normality*, де необхідно задати кількість інтервалів і критерій Колмогорова – Смирнова для тестування вибірки. Ініціювавши клавішу *Frequency tables (Таблиці частот)*, буде одержано таку таблицю (рис. 1.9).

Frequency table: Kilkist zdobyvachiv (Spreadsheet2)						
K-S d=.21203, p<.20 ; Lilliefors p<.01						
Category	Count	Cumulative Count	Percent of Valid	Cumul % of Valid	% of all Cases	Cumulative % of All
-200000,<x<=0,000000	0	0	0,00000	0,0000	0,00000	0,0000
0,000000<x<=200000,0	14	14	56,00000	56,0000	56,00000	56,0000
200000,0<x<=400000,0	5	19	20,00000	76,0000	20,00000	76,0000
400000,0<x<=600000,0	3	22	12,00000	88,0000	12,00000	88,0000
600000,0<x<=800000,0	1	23	4,00000	92,0000	4,00000	92,0000
800000,0<x<=1000000,0	1	24	4,00000	96,0000	4,00000	96,0000
1000000,<x<=1200000,0	1	25	4,00000	100,0000	4,00000	100,0000
Missing	0	25	0,00000		0,00000	100,0000

Рис. 1.9. Результат угруповання вибірки

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

Cumulative % Expected (накопичений % від загальної теоретичної частоти).

Ініціювавши клавішу *Histograms*, буде одержано таку гістограму розподілу з накладеним нормальним законом розподілу й критерієм Колмогорова – Смирнова, значення якого порівнюють з табличним (рис. 1.10).

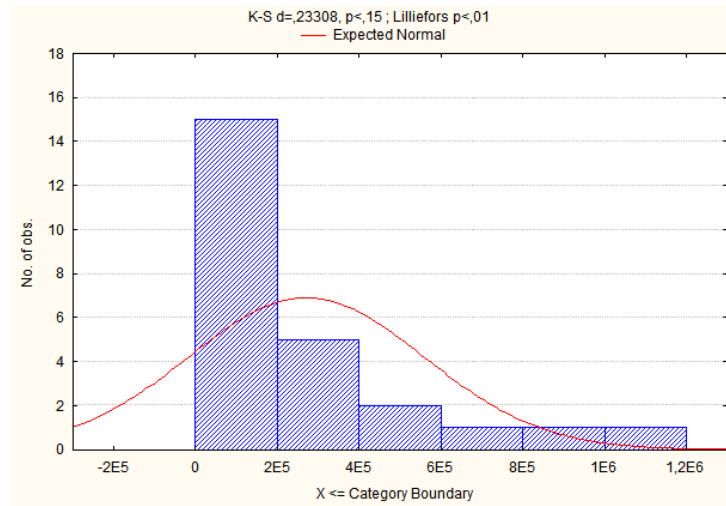


Рис. 1.10. Гістограма розподілу

Графік накопичених емпіричних частот подано на рис. 1.11.

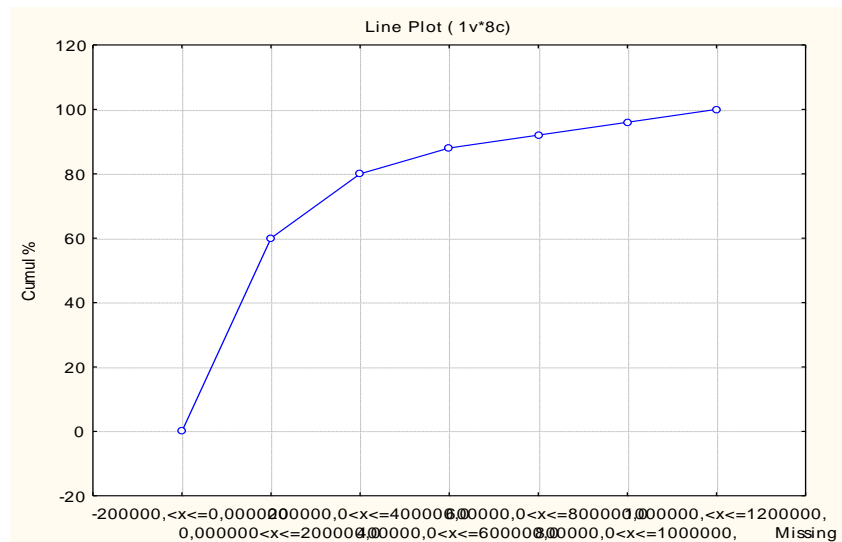


Рис. 1.11. Графік накопичених емпіричних частот

Подальший аналіз вибірки припускає розрахунок критерієм Пірсона й формування відповідних висновків про характер закону розподілу.

Для цього в головному меню *Statistics* обрати модуль *Distribution Fitting* (*Закони розподілу*). Далі обрати змінну для аналізу й нормальний закон розподілу, параметри для розрахунків й умови для критеріїв можна задати самостійно або за замовчуванням вони задані автоматично. Ініціювавши кнопку графіка буде одержано гістограму й значення відповідних критеріїв (рис. 1.12).

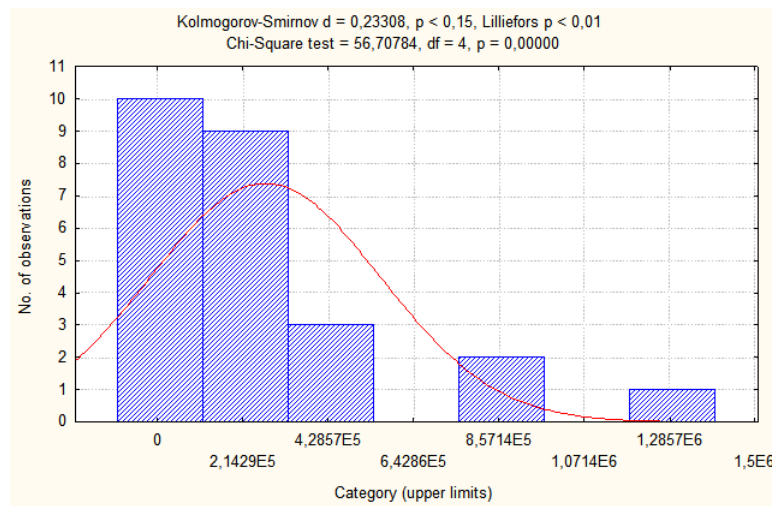


Рис. 1.12. Гістограма розподілу із критерієм Пірсона й Колмогорова – Смирнова

Далі робимо висновки про угруповання даних об'єктів за величиною показника кількість здобувачів освіти. Порівняти отримані значення з таблицями за відповідними критеріями й зробити висновки про характер закону розподілу.

Тема 2. Формування інформаційного простору наукового дослідження: попередній аналіз даних

Методичні рекомендації до виконання лабораторної роботи за темою 2 "Перевірка динамічних рядів на нормальність розподілу величин. Показники асиметрії та ексцесу", "Розрахунок параметричних та непараметричних критеріїв для рядів розподілу величин"

Мета: отримання практичних навичок щодо перевірки вихідних даних на нормальність розподілу величини з використанням пакетів *Statistica 13.0*.

А також отримання практичних навичок аналізу рядів розподілу з використанням параметричних та непараметричних критеріїв.

Завдання: вивчити основні прийоми роботи з пакетами *Statistica 13.0* для перевірки вихідної статистичної інформації. Необхідно провести аналіз ряду динаміки значень показника індекс освітніх послуг (табл. 2.1), попередньо розрахувавши показники описової статистики, а також провести перевірку ряду на нормальність закону розподілу величини.

Таблиця 2.1

Вихідні дані

№ з/п	Періоди	Індекс освітніх послуг (відсотків до відповідного періоду попереднього року)	№ з/п	Період	Індекс освітніх послуг (відсотків до відповідного періоду попереднього року)
1	2	3	4	5	6
1	січень 2007 року	112,2	25	січень 2009 року	66,2
2	лютий 2007 року	110,1	26	лютий 2009 року	67,4
3	березень 2007 року	109,3	27	березень 2009 року	68,2
4	квітень 2007 року	109,5	28	квітень 2009 року	68,2
5	травень 2007 року	108,9	29	травень 2009 року	68,3
6	червень 2007 року	108,4	30	червень 2009 року	69
7	липень 2007 року	107,9	31	липень 2009 року	69,7
8	серпень 2007 року	107,9	32	серпень 2009 року	70,5
9	вересень 2007 року	107,7	33	вересень 2009 року	71,7
10	жовтень 2007 року	108	34	жовтень 2009 року	73,6
11	листопад 2007 року	107,9	35	листопад 2009 року	76,1
12	грудень 2007 року	107,6	36	грудень 2009 року	78,1
13	січень 2008 року	105	37	січень 2010 року	111,8
14	лютий 2008 року	107	38	лютий 2010 року	108,8

1	2	3	4	5	6
15	березень 2008 року	105,7	39	березень 2010 року	110,8
16	квітень 2008 року	105,4	40	квітень 2010 року	112,6
17	травень 2008 року	105,2	41	травень 2010 року	112,6
18	червень 2008 року	104,5	42	червень 2010 року	112
19	липень 2008 року	104,4	43	липень 2010 року	111,1
20	серпень 2008 року	103,4	44	серпень 2010 року	110,9
21	вересень 2008 року	102,5	45	вересень 2010 року	110,8
22	жовтень 2008 року	100	46	жовтень 2010 року	110,7
23	листопад 2008 року	97	47	листопад 2010 року	110,6
24	грудень 2008 року	94,8			

Методичні рекомендації

1. Введення вихідних даних у робочий аркуш пакета Statistica 13.0.

Використовуючи інструменти пакета Statistica 13.0, вводять стовпець вихідних даних, що позначають як *Index* (рис. 2.1).

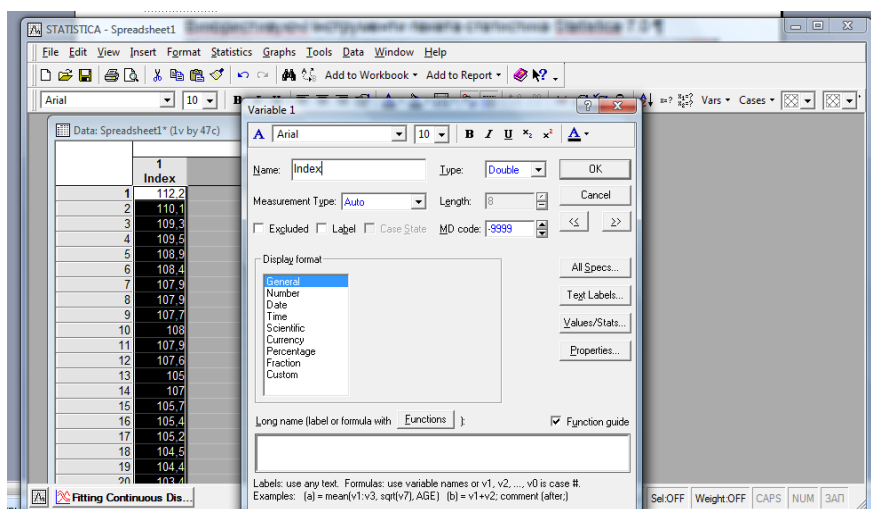


Рис. 2.1. Введення вхідної інформації

2. Розрахунок основних статистичних показників.

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 використовують кнопку *Variable* (змінні), необхідно задати вихідні параметри моделі. У цьому випадку це стовпець *Index* (рис. 2.2).

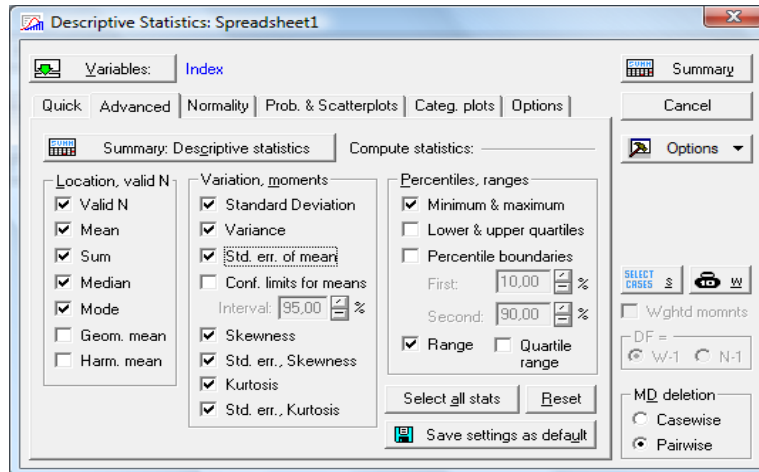


Рис. 2.2. Виділення основних описових статистик для розрахунку

Далі виділяємо основні статистики для розрахунків, як зображено на рис. 2.2. 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).

Результат розрахунку зображено на рис. 2.3.

Descriptive Statistics (Spreadsheet1)									
Variable	Valid N	Mean	Median	Mode	Frequency of Mode	Sum	Minimum	Maximum	Range
Index	47	98,0851	107,0000	107,9000	3	4610,000	66,20000	112,6000	46,40000
	Variance	Std.Dev.	Standard Error	Skewness	Std.Err. Skewness	Kurtosis	Std.Err. Kurtosis		
	281,3120	16,77230	2,446500	-1,04860	0,346570	-0,698730	0,680910		

Рис. 2.3. Описові статистики

Виходячи з проведених розрахунків, можна зробити висновок, що довжина ряду динаміки 47 значень, середнє значення індексу наданих освітніх послуг складає 98,09 %. За період з січня 2007 року до листопаду 2010 року найбільше значення індексу складало 112,6 %, а найменше – 66,2 %, що свідчить про великий розкид значень. Визначення форми розподілу проводять за допомогою критеріїв:

- значення коефіцієнта асиметрії перевищує 0 у випадку правосторонньої асиметрії і менше 0 у випадку лівосторонньої. У вибірці спостерігається лівостороння асиметрія, бо коефіцієнт асиметрії $-1,05 < 0$;
- значення коефіцієнта ексцесу перевищує 0 у випадку гостровершинного розподілу і менше 0 у випадку плосковершинного розподілу. Вибірка, що вивчається, є плосковершинною, бо коефіцієнт ексцесу складає $-0,7 < 0$.

3. Побудова полігону розподілу.

Для цього необхідно в меню *Graphs / 2D Graphs / Line Plots (variables)* вибрати змінну (рис. 2.4).

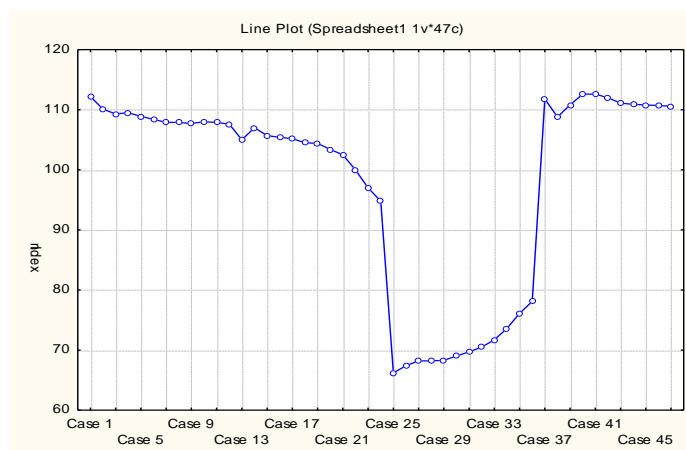


Рис. 2.4. Полігон розподілу

З рис. 2.4 видно, що тенденція зміни показника суттєво відрізняється для випадків 25 – 36, що відповідає місяцям 2009 року. У 2009 році на фоні світової кризи відбулось значне зниження обсягу промислової продукції, що і відображає полігон розподілу.

4. Перевірка вибірки на нормальний закон розподілу.

4.1. Розрахунок кількості інтервалів групування. Кількість інтервалів групування розраховують за такою формулою:

$$m = 1 + 3,322 \ln n, \quad (2.1)$$

де n – довжина динамічного ряду.

$$m = 1 + 3,322 \ln 47 = 13,78 \approx 14.$$

4.2. Подальший аналіз здійснюють у рамках перевірки вибірки на нормальний закон розподілу. Розрахунок критерію Пірсона для формування остаточного висновку про характер розподілу величини. Для цього в головному меню *Statistics* обрати модуль *Distribution Fitting* (Закони розподілу). На рис. 2.5 показано вибір закону розподілу, який перевіряється.

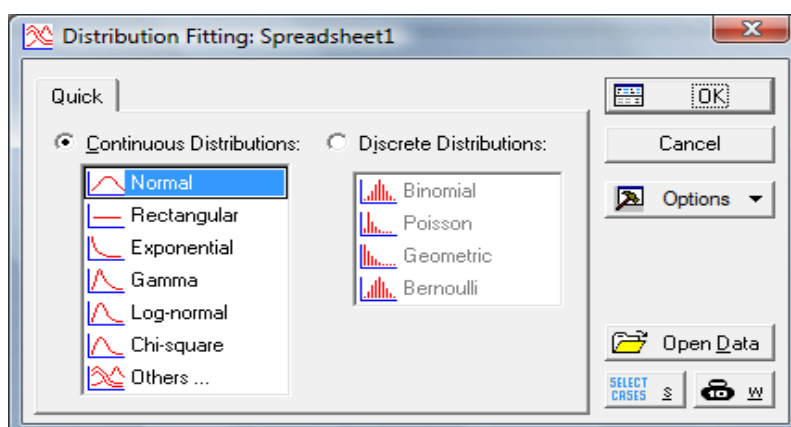


Рис. 2.5. Стартова панель модуля *Distribution Fitting*

Далі задають параметри розрахунку, як показано на рис. 2.6. У вікні параметрів вказують кількість інтервалів, середнє значення індексу промисловості, мінімальне та максимальне значення показника.

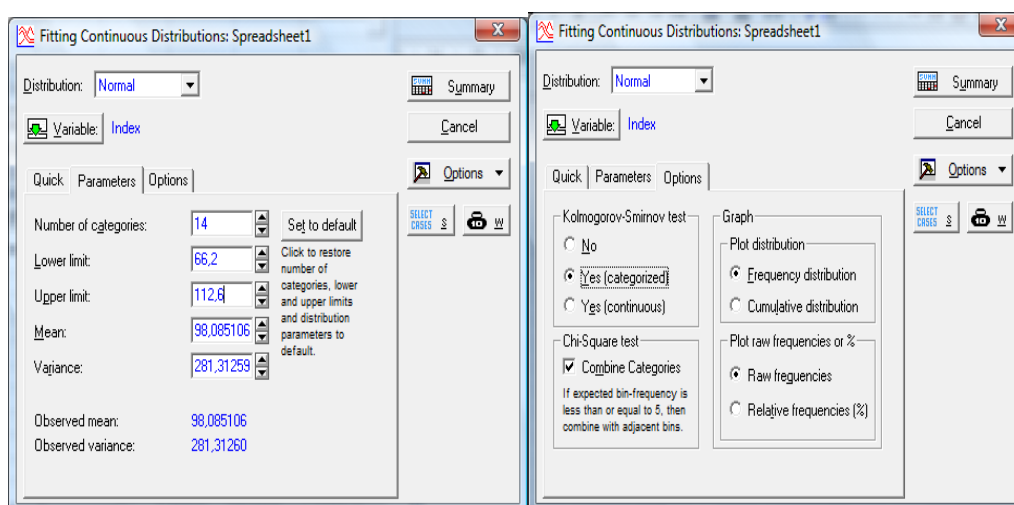


Рис. 2.6. Завдання параметрів розрахунку

Результати розрахунків наведено на рис. 2.7 і 2.8.

Variable: Index Distribution: Normal (Spreadsheet1) Kolmogorov-Smimov d = 0,26699, p < 0,01 Chi-Square = 37,66166, df = 3 (adjusted) , p = 0,00000									
Upper Boundary	Observed Frequency	Cumulative Observed	Percent Observed	Cumul. % Observed	Expected Frequency	Cumulative Expected	Percent Expected	Cumul. % Expected	Observed-Expected
<= 69,51429	6	6	12,7659%	12,7660%	2,0793%	2,0793%	4,4242%	4,4242%	3,9206%
72,82857	3	9	6,3829%	19,1489%	1,0251%	3,1045%	2,1811%	6,6054%	1,9748%
76,14286	2	11	4,2553%	23,4042%	1,3790%	4,4836%	2,9342%	9,5396%	0,6209%
79,45714	1	12	2,1276%	25,5319%	1,7844%	6,2680%	3,7966%	13,336%	-0,7844%
82,77143	0	12	0,0000%	25,5319%	2,2207%	8,4888%	4,7250%	18,061%	-2,2207%
86,08571	0	12	0,0000%	25,5319%	2,6583%	11,1471%	5,6559%	23,717%	-2,6583%
89,40000	0	12	0,0000%	25,5319%	3,0605%	14,2076%	6,5118%	30,229%	-3,0605%
92,71429	0	12	0,0000%	25,5319%	3,3891%	17,5968%	7,2110%	37,440%	-3,3891%
96,02857	1	13	2,1276%	27,659%	3,6098%	21,2066%	7,6804%	45,120%	-2,6098%
99,34286	1	14	2,1276%	29,7872%	3,6980%	24,9047%	7,8682%	52,988%	-2,6980%
102,65714	2	16	4,2553%	34,0426%	3,6438%	28,5486%	7,7528%	60,741%	-1,6438%
105,97143	7	23	14,8936%	48,9362%	3,4533%	32,0019%	7,3475%	68,089%	3,5466%
109,28571	10	33	21,2766%	70,2128%	3,1479%	35,1498%	6,6976%	74,787%	6,8520%
< Infinity	14	47	29,7872%	100,0000%	11,8501%	47,0000%	25,2130%	100,0000%	2,1498%

Рис. 2.7. Результати групування та розрахунку критерію Пірсона та критерію Колмогорова – Смірнова

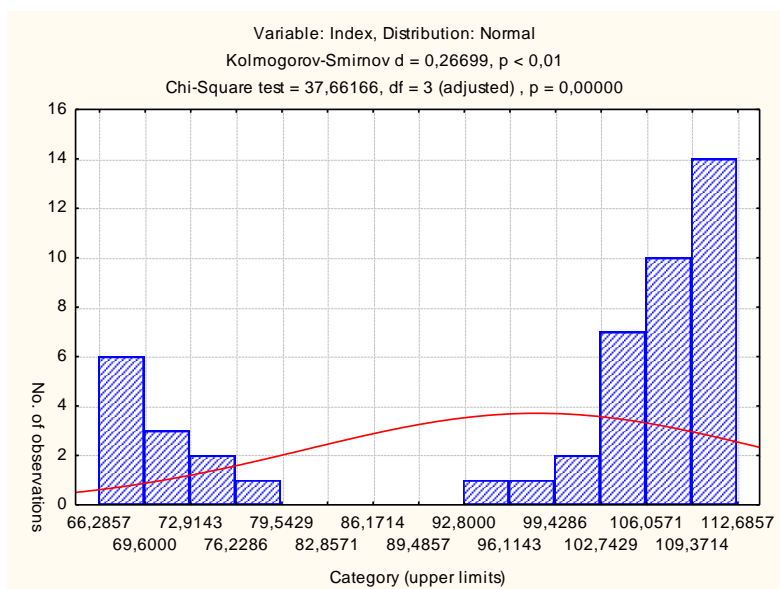


Рис. 2.8. Гістограма розподілу з критерієм Пірсона та Колмогорова – Смірнова

За умови нормального розподілу розраховані значення критерію Пірсона та критерію Колмогорова – Смірнова менше критичного значення. Для $k - p - 1 = 12$ ступенів свободи (де k – кількість інтервалів розподілу, p – кількість параметрів, що перевіряється), а також рівня значущості $\alpha = 0,01$ табличне значення критерію Пірсона складає 31,264. З розрахунків видно, що критерій Пірсона складає 37,66, тобто за критерієм Пірсона вибірка розподілена не за нормальним законом розподілу.

Критичне значення критерію Колмогорова – Смірнова за умови довжини вибірки 47 спостережень та $\alpha = 0,01$ складає 0,2377, що менше розрахункового значення 0,2669. Критерій Колмогорова – Смірнова теж спростовує гіпотезу про нормальний закон розподілу.

Таким чином, тенденція зміни показника суттєво відрізняється для випадків 25 – 36, що відповідає місяцям 2009 року. Виходячи з проведених розрахунків та проаналізувавши значення критеріїв Пірсона та Колмогорова – Смірнова, можна зробити висновок що ряд динаміки не підпорядковується нормальному закону розподілу. Тому під час подальшого аналізу доцільно нормалювати та стандартизувати значення ряду динаміки.

Змістовий модуль 2

Методи та технології кількісно-якісного аналізу та прогнозування освітньої інформації

Тема 3. Методи та технології обробки кількісної інформації

Методичні рекомендації до виконання лабораторної роботи за темою 3 "Визначення та усунення аномальних рівнів часового ряду. Дослідження закономірностей у часовому ряду. Методи визначення стійкості часового ряду"

Метою завдання є формування компетентностей дослідження закономірностей у часовому ряді та використання методів згладжування для аналізу стійкості часового ряду.

Перевірити наявність аномальних значень у динамічному ряді обсягу імпорту Вінницької області (табл. 3.1, стовпчик 2) та у разі необхідності усунути їх.

Таблиця 3.1

Обсяг наданих послуг у закладах освіти

Періоди	Обсяг наданих послуг, ум. од.	Періоди	Обсяг наданих послуг, ум. од.
1	2	3	4
1	39,967	11	60,648
2	43,725	12	59,495

1	2	3	4
3	66,066	13	57,229
4	52,420	14	57,428
5	56,811	15	55,230
6	53,137	16	62,990
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

Аналіз часових рядів, а саме визначення аномальних точок та дослідження стійкості ряду відбувається в декілька етапів:

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

$$\lambda = \frac{|y_t - y_{t-1}|}{S_y} \quad (3.1)$$

Проведені розрахунки наведено в табл. 3.2.

Таблиця 3.2

Обсяг наданих освітніх послуг 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

1	2	3	4	5	6	7
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
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		

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. Провести порівняння розрахованих значень λ для кожного рівня ряду з табличним. Виходячи з того, що $\lambda_3 > \lambda_{табл}$ (тобто $3,12 > 3,1$) та $\lambda_{17} > \lambda_{табл}$ (тобто $3,26 > 3,1$), то значення u_3 та u_{17} є аномальними рівнями.

6. Провести заміну значень аномальних рівнів на середнє арифметичне їх сусідів (табл. 3.2).

7. Перевірити відповідно до зазначеного алгоритму (етап 1 – 5) наявність аномальних рівнів у новому ряду. Як видно з розрахунків, проведених у табл. 3.2, вони дозволили усунути аномальні рівні в ряду.

На рис. 3.1 наведено порівняння вихідного та згладженого динамічного ряду.



Рис. 3.1. Графіки вихідного та згладженого рядів динаміки

Таким чином, аналізуючи дані, подані в табл. 3.2 та рис. 3.1, можна зробити висновок про те, що після використання методів згладжування досліджуваній часовий ряд має плавну динаміку розвитку. За згладженим рядом можна визначити наявність позитивних та негативних тенденцій надання обсягу освітніх послуг та побудувати прогнозні моделі.

Тема 4. Методи та технології обробки якісної інформації

Методичні рекомендації до виконання лабораторної роботи за темою 4 "Багатофакторне ранжування"

Метою завдання є опанування студентами навичок проведення ранжування економічних об'єктів різними методами.

Завдання: необхідно провести оцінювання та аналіз ЕБК України. В якості об'єктів дослідження виступають 15 країн світу: Австрія, Бельгія, Болгарія, Фінляндія, Франція, Німеччина, Італія, Польща, Іспанія, Швеція, Швейцарія, Великобританія, Білорусь, Україна і Росія. Вхідні дані дослідження показані в табл. 4.1. З таблиці видно, що для 15 країн притаманні різнопланові значення обраних для аналізу п'яти показників енергетичної безпеки.

1. Частка власних джерел у балансі паливно-енергетичних ресурсів держави, % (ЧВД).
2. Частка домінувального паливного ресурсу у споживанні паливно-енергетичних ресурсів, % (ЧДП).
3. Енергоємність ВВП, кг умовного палива/грн (ЕЄВВП).
4. Обсяг видобутку вугілля, млн тонн (ОВВ).
5. Ступінь забезпечення паливно-енергетичними ресурсами (СЗР).

Значення показників енергетичної безпеки

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

Методичні рекомендації

1. Формування матриці спостережень. З цією метою новоствореній книзі *MS Excel 2007* побудувати таблицю відповідно до наведеної таблиці (див. табл. 4.1) та ввести до неї вихідні дані (рис. 4.1).

Країни	ЧДП	ЄСВВП	ОВВ	ЧВД	СЗР
Австрія	0,5125	8,1397	28,086	0,6577	1,2299
Бельгія	0,441	11,184	46,61	0,4231	1,198
Болгарія	0,3906	67,683	473,8	0,7281	0,629
Фінляндія	0,6599	13,741	0	0,6448	1,1811
Франція	0,5984	7,579	6,8511	0,7465	0,5692
Німеччина	0,4255	9,3587	2922,1	0,6716	0,9989
Італія	0,4222	9,1809	4,4658	0,5269	1,3332
Польща	0,4471	27,106	6933,4	1,0576	1,0304
Іспанія	0,494	10,292	4295,4	0,6178	0,9236
Швеція	0,8281	8,865	0	0,9094	0,9484
Швейцарія	0,6947	4,6217	0	0,7551	0,9521
Великобританія	0,3898	7,828	1005,6	1,9541	2,1603
Білорусь	0,6031	80,723	0	0,5211	1,343
Росія	0,5621	121,36	3960,7	2,8999	2,9421
Україна	0,5411	179,24	2740,3	0,6762	1,0767

Рис. 4.1. Формування матриці вихідних даних

2. Визначити рейтинг країн за показником енергоємності ВВП, кг умовного палива/грн (ЄЄВВП) та побудувати їхній рейтинг. Для цього проводять сортування показників за стовпцем ЄЄВВП за зростанням ознаки (рис. 4.2). Виходячи з того, що обраний показник є дестимулятором, то рейтингове місце 1 віддають країні з найменшим значенням показника.

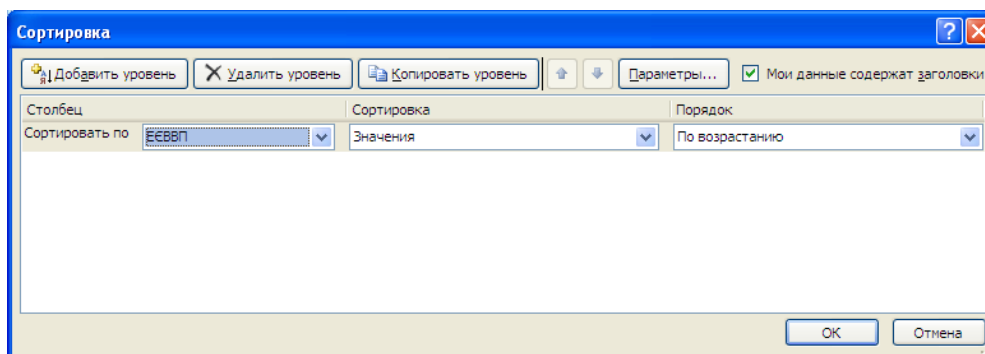


Рис. 4.2. Сортування за обраною ознакою

Рейтинг країн світу за показником енергоємності ВВП має вигляд, наведений у табл. 4.2.

Таблица 4.2

Рейтинг країн світу за показником енергоємності ВВП

№ з/п	Країни	ЧДП	ЄЄВВП	ОВВ	ЧВД	СЗР
1	Швейцарія	0,694735	4,621731	0	0,75509	0,952061
2	Франція	0,598438	7,578963	4,139735	0,746463	0,569163
3	Велика Британія	0,389766	7,828034	606,1022	1,954056	2,160322
4	Австрія	0,512517	8,139704	2,295999	0,657682	1,229883
5	Швеція	0,828147	8,865035	0	0,909419	0,948364
6	Італія	0,422215	9,180896	2,592751	0,526936	1,333184
7	Німеччина	0,425501	9,358731	2408,544	0,671615	0,998937
8	Іспанія	0,493978	10,29191	385,999	0,61782	0,923633
9	Бельгія	0,441011	11,18419	4,823341	0,423147	1,198028
10	Фінляндія	0,659931	13,74058	0	0,644842	1,181073
11	Польща	0,447059	27,10571	2792,817	1,057638	1,030381
12	Болгарія	0,390566	67,68311	182,7928	0,728101	0,629035
13	Білорусь	0,603121	80,72323	0	0,521061	1,342974
14	Росія	0,562118	121,3625	5683,912	2,899916	2,942098
15	Україна	0,405700	130,269	1396,38	0,676241	1,210000

3. Розрахунок узагальнювального показника рівня розвитку. Стандартизація ознак. Стандартизацію матриці вихідних даних проводять за формулами:

$$z_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j}, \quad (4.1)$$

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

\bar{x}_j – середнє арифметичне значення показника j ;

s_j – стандартне відхилення показника j .

Попередні розрахунки проводять з використанням стандартних функцій СРЗНАЧ та СТАНДАРТТОТКЛ, як показано на рис. 4.3 і 4.4.

	A	B	C	D	E	F	G	H	I
1	Країни	ЧДП	БЄВВП	ОВВ	ЧВД	СЗР			
2	Австрія	0,51252	8,1397	28,0864	0,65768	1,22988			
3	Бельгія	0,44101	11,1842	46,6101	0,42315	1,19803			
4	Болгарія	0,39057	67,6831	473,797	0,7281	0,62904			
5	Фінляндія	0,65993	13,7406	0	0,64484	1,18107			
6	Франція	0,59844	7,57896	6,85112	0,74646	0,56916			
7	Німеччина	0,4255	9,35873	2922,12	0,67162	0,99894			
8	Італія	0,42222	9,1809	4,46583	0,52694	1,33318			
9	Польща	0,44706	27,1057	6933,37	1,05764	1,03038			
10	Іспанія	0,49398	10,2919	4295,37	0,61782	0,92363			
11	Швеція	0,82815	8,86504	0	0,90942	0,94836			
12	Швейцарія	0,69474	4,62173	0	0,75509	0,95206			
13	Великобританія	0,38977	7,82803	1005,63	1,95406	2,16032			
14	Білорусь	0,60312	80,7232	0	0,52106	1,34297			
15	Росія	0,56212	121,363	3960,71	2,89992	2,9421			
16	Україна	0,54113	179,236	2740,34	0,67624	1,07672			
17	хсереднє	=СРЗНАЧ(B2:B16)	1494,49	0,91934	1,23439				
18		СРЗНАЧ(число1; [число2]; ...)							
19									

Рис. 4.3. Розрахунок середніх значень за кожною з ознак

Крайни	ЧДП	ЄСВВП	ОВВ	ЧВД	СЗР
Австрія	0,51252	8,1397	28,0864	0,65768	1,22988
Бельгія	0,44101	11,1842	46,6101	0,42315	1,19803
Болгарія	0,39057	67,6831	473,797	0,7281	0,62904
Фінляндія	0,65993	13,7406	0	0,64484	1,18107
Франція	0,59844	7,57896	6,85112	0,74646	0,56916
Німеччина	0,4255	9,35873	2922,12	0,67162	0,99894
Італія	0,42222	9,1809	4,46583	0,52694	1,33318
Польща	0,44706	27,1057	6933,37	1,05764	1,03038
Іспанія	0,49398	10,2919	4295,37	0,61782	0,92363
Швеція	0,82815	8,86504	0	0,90942	0,94836
Швейцарія	0,69474	4,62173	0	0,75509	0,95206
Великобританія	0,38977	7,82803	1005,63	1,95406	2,16032
Білорусь	0,60312	80,7232	0	0,52106	1,34297
Росія	0,56212	121,363	3960,71	2,89992	2,9421
Україна	0,54113	179,236	2740,34	0,67624	1,07672
хсередне	0,53402	37,7933	1494,49	0,91934	1,23439
S	=СТАНДОТКЛОН(B2:B16)			0,65574	0,59692

Рис. 4.4. Розрахунок стандартного відхилення

Після розрахунку середніх значень та стандартного відхилення за кожною з ознак проводять розрахунок стандартизованої матриці (рис. 4.5).

	ЄСВВП	ОВВ	ЧВД	СЗР	ЧДП	ЄСВВП	ОВВ	ЧВД	СЗР
2	8,1397	28,0864	0,65768	1,22988	= (B2-\$B\$17)/\$B\$18		-0,67540864	-0,3990177	-0,007551467
3	11,1842	46,6101	0,42315	1,19803	-0,738735257	-0,509820137	-0,66687685	-0,7566806	-0,06091733
4	67,6831	473,797	0,7281	0,62904	-1,139418878	0,572675644	-0,470119701	-0,2916296	-1,014136697
5	13,7406	0	0,64484	1,18107	1,000141902	-0,460840769	-0,688344927	-0,4185985	-0,089321607
6	7,57896	6,85112	0,74646	0,56916	0,51170424	-0,578894785	-0,68518938	-0,2636278	-1,114438722
7	9,35873	2922,12	0,67162	0,99894	-0,861930875	-0,544795171	0,657549059	-0,37777	-0,394449372
8	9,1809	4,46583	0,52694	1,33318	-0,888031507	-0,548202416	-0,686288015	-0,5984037	0,165506047
9	27,1057	6933,37	1,05764	1,03038	-0,690696114	-0,204770443	2,505087036	0,21091011	-0,341772046
10	10,2919	4295,37	0,61782	0,92363	-0,318019441	-0,52691585	1,290052695	-0,4598067	-0,52060423
11	8,86504	0	0,90942	0,94836	2,336278209	-0,554254181	-0,688344927	-0,015122	-0,479173021
12	4,62173	0	0,75509	0,95206	1,276589375	-0,635554119	-0,688344927	-0,2504717	-0,472979531
13	7,82803	1005,63	1,95406	2,16032	-1,145773262	-0,574122688	-0,225161918	1,57793612	1,551189106
14	80,7232	0	0,52106	1,34297	0,548901214	0,822518913	-0,688344927	-0,607363	0,181906983
15	121,363	3960,71	2,89992	2,9421	0,22321521	1,601150404	1,135912143	3,02036055	2,86087837
16	179,236	2740,34	0,67624	1,07672	0,05653972	2,709976817	0,573823279	-0,3707155	-0,264136483
17	37,7933	1494,49	0,91934	1,23439	Еталон				
18	52,1932	2171,14	0,65574	0,59692					

Рис. 4.5. Розрахунок значень стандартизованої матриці

4. Побудова еталона розвитку передбачає виокремлення підмножини стимуляторів та дестимуляторів. Серед наведених показників стимуляторами є: частка власних джерел у балансі паливно-енергетичних

ресурсів держави (ЧВД), обсяг видобутку вугілля (ОВВ), ступінь забезпечення паливно-енергетичними ресурсами (СЗР). До дестимуляторів належать: частка домінувального паливного ресурсу у споживанні паливно-енергетичних ресурсів (ЧДП), енергоємність ВВП (ЄЄВВП). Поділ показників на дві групи проводять відповідно до їх економічного змісту.

Еталон розвитку має такий вигляд:

$$z_0 = [z_{01}, z_{02}, \dots, z_{0n}],$$

де $z_{0s} = \max_i z_{is}$, якщо $s \in I$;

$z_{0s} = \min_t z_{is}$, якщо $s \notin I (s = 1, \dots, n)$;

I – множина стимуляторів.

Відповідно до поділу показників на стимулятори і дестимулятори з застосуванням стандартних функцій MAX і MIN було розраховано значення еталона розвитку (рис. 4.6).

Крайни	ЧДП	ЄЄВВП	ОВВ	ЧВД	СЗР	ЧДП	ЄЄВВП	ОВВ	ЧВД	СЗР
Австрія	0,51252	8,1397	28,0864	0,65768	1,22988	-0,170764536	-0,568151221	-0,67540864	-0,3990177	-0,007551467
Бельгія	0,44101	11,1842	46,6101	0,42315	1,19803	-0,738735257	-0,509820137	-0,66687685	-0,7566806	-0,06091733
Болгарія	0,39057	67,6831	473,797	0,7281	0,62904	-1,139418878	0,572675644	-0,470119701	-0,2916296	-1,014136697
Фінляндія	0,65993	13,7406	0	0,64484	1,18107	1,000141902	-0,460840769	-0,688344927	-0,4185985	-0,089321607
Франція	0,59844	7,57896	6,85112	0,74646	0,56916	0,51170424	-0,578894785	-0,68518938	-0,2636278	-1,114438722
Німеччина	0,4255	9,35873	2922,12	0,67162	0,99894	-0,861930875	-0,544795171	0,657549059	-0,37777	-0,394449372
Італія	0,42222	9,1809	4,46583	0,52694	1,33318	-0,888031507	-0,548202416	-0,686288015	-0,5984037	0,165506047
Польща	0,44706	27,1057	6933,37	1,05764	1,03038	-0,690696114	-0,204770443	2,505087036	0,21091011	-0,341772046
Іспанія	0,49398	10,2919	4295,37	0,61782	0,92363	-0,318019441	-0,52691585	1,290052695	-0,4598067	-0,52060423
Швеція	0,82815	8,86504	0	0,90942	0,94836	2,336278209	-0,554254181	-0,688344927	-0,015122	-0,479173021
Швейцарія	0,69474	4,62173	0	0,75509	0,95206	1,276589375	-0,635554119	-0,688344927	-0,2504717	-0,472979531
Великобританія	0,38977	7,82803	1005,63	1,95406	2,16032	-1,145773262	-0,574122688	-0,225161918	1,57793612	1,551189106
Білорусь	0,60312	80,7232	0	0,52106	1,34297	0,548901214	0,822518913	-0,688344927	-0,607363	0,181906983
Росія	0,56212	121,363	3960,71	2,89992	2,9421	0,22321521	1,601150404	1,135912143	3,02036055	2,86087837
Україна	0,54113	179,236	2740,34	0,67624	1,07672	0,05653972	2,709976817	0,573823279	-0,3707155	-0,264136483
хсередне	0,53402	37,7933	1494,49	0,91934	1,23439	Еталон	=МИН(H2:H16)	2,505087036	3,02036055	2,86087837
S	0,1259	52,1932	2171,14	0,65574	0,59692		МИН(число1; [число2]; ...)			

Рис. 4.6. Розрахунок еталона розвитку

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

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

Проведений розрахунок дозволяє сформуванати матрицю відстаней (рис. 4.7).

6. Розрахунок таксономічного показника рівня розвитку:

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

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

На рис. 4.7 наведено проведені розрахунки для отримання вектора значень d^* .

Крайни	ЧДП	ЄСВВП	ОВВ	ЧВД	СЗР	Сума	Корень суми (Co)	Co / Co	Інт. пок-к
Австрія	0,51252	8,1397	28,0864	0,65768	1,22988	30,99077521	5,56693589	0,649270821	=1-J23
Бельгія	0,44101	11,1842	46,6101	0,42315	1,19803	33,04577365	5,748545351	0,670451903	0,3295481
Болгарія	0,39057	67,6831	473,797	0,7281	0,62904	36,2967353	6,024677195	0,702657115	0,29734289
Фінляндія	0,65993	13,7406	0	0,64484	1,18107	35,3636038	5,946730513	0,693566206	0,30643379
Франція	0,59844	7,57896	6,85112	0,74646	0,56916	39,51603103	6,286177776	0,733155885	0,26684412
Німеччина	0,4255	9,35873	2922,12	0,67162	0,99894	25,64665052	5,064252217	0,590642907	0,40935709
Італія	0,42222	9,1809	4,46583	0,52694	1,33318	30,61942232	5,533481935	0,645369092	0,35463091
Польща	0,44706	27,1057	6933,37	1,05764	1,03038	18,54265128	4,30611789	0,502221824	0,49777818
Іспанія	0,49398	10,2919	4295,37	0,61782	0,92363	25,71927601	5,071417555	0,591478599	0,4085214
Швеція	0,82815	8,86504	0	0,90942	0,94836	42,69939731	6,534477585	0,762115051	0,23788495
Швейцарія	0,69474	4,62173	0	0,75509	0,95206	37,8788008	6,154575598	0,717807144	0,28219286
Великобританія	0,38977	7,82803	1005,63	1,95406	2,16032	11,25390739	3,354684395	0,391256292	0,60874371
Білорусь	0,60312	80,7232	0	0,52106	1,34297	35,53317192	5,960970719	0,695227039	0,30477296
Росія	0,56212	121,363	3960,71	2,89992	2,9421	8,751616443	2,958313108	0,345027573	0,65497243
Україна	0,54113	179,236	2740,34	0,67624	1,07672	37,63302775	6,134576412	0,715474642	0,28452536
хсередне	0,53402	37,7933	1494,49	0,91934	1,23439	37,63302775	6,134576412	0,715474642	0,28452536
с	0,1259	52,1932	2171,14	0,65574	0,59692				

Рис. 4.7. Розрахунки вектора значень d^*

Щоб показник d^* набував високих значень за умови більших значень стимуляторів і низьких значень за умови малих значень стимуляторів, то його перетворюють до вигляду:

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

У табл. 4.3 наведені результати розрахунку.

Як видно з табл. 4.3, Швеція має не найгірший показник енергетичної безпеки серед обраних країн, але є аутсайдером. Країни, що мають нижчий показник енергетичної безпеки, не видобувають вугілля чи видобувають його в невеликій кількості. Тобто можна зробити висновок, що Україна за умови достатньо великих власних запасів енергоресурсів нерационально використовує їх, це підтверджено тим, що Україна має найгірший показник енергоємності ВВП серед європейських країн.

Таблиця 4.3

Значення узагальнювального показника ЕБК

№ з/п	Країни	Узагальнювальний показник енергетичної безпеки
1	Австрія	0,35072918
2	Бельгія	0,32954810
3	Болгарія	0,29734289
4	Фінляндія	0,30643379
5	Франція	0,26684412
6	Німеччина	0,40935709
7	Італія	0,35463091
8	Польща	0,49777818
9	Іспанія	0,40852140
10	Швеція	0,23788495
11	Швейцарія	0,28219286
12	Велика Британія	0,60874371
13	Білорусь	0,30477296
14	Росія	0,65497243
15	Україна	0,28452536

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

$$R_i = \sum_{i=1}^n \alpha_i \times R_{ij}, \quad (4.2)$$

де α_i – ваговий коефіцієнт, що визначається експертним методом.

За кожним з показників необхідно виставити рейтинг (табл. 4.4). Процедура відповідна до розрахунків, проведених у пункті 1.

Таблиця 4.4

Рейтенгова оцінка за кожним з показників ЕБК

Країни	ЧДП	ЄЄВВП	ОВВ	ЧВД	СЗР
Австрія	8	4	1	8	4
Бельгія	5	9	2	13	5
Білорусь	10	13	3	12	2
Болгарія	2	12	4	6	12
Велика Британія	1	3	5	1	1
Німеччина	4	7	6	7	8
Іспанія	7	8	7	10	11
Італія	3	6	8	11	3
Польща	6	11	9	2	7
Фінляндія	11	10	10	9	6
Франція	9	2	11	5	13
Швейцарія	12	1	12	4	9
Швеція	13	5	13	3	10

Наступним етапом є визначення вагових коефіцієнтів, що виставляють експертним методом. Для визначення вагових коефіцієнтів використовують метод ранжувань. Цей метод полягає у тому, що кожного експерта просять розташувати ознаки в порядку переваги. Цифрою один позначають найбільш важливу ознаку, цифрою два – наступний за нею за важливістю і т. д.

Отримані дані щодо суттєвості ознаки для оцінювання рівня енергетичної безпеки зводять в табл. 4.5.

Таблиця 4.5

Експертні оцінки

Показники	E ₁	E ₂	E ₃	E ₄	E ₅
ЧДП	1	2	1,5	1	2
ЕЄВВП	2,5	2	1,5	2,5	1
ОВВ	2,5	2	3	2,5	3
ЧВД	4	5	4,5	4,5	4
СЗР	5	4	4,5	4,5	5,5

Перед тим, як проводити узагальнення думок експертів, треба перевірити їх на узгодженість, що є передумовою якості отриманих оцінок. Узгодженість думок експертів перевірити за допомогою коефіцієнта рангової кореляції Кендела, Спірмена та коефіцієнта конкордації у пакеті *Statistica 13.0*.

З метою розрахунку коефіцієнтів парної кореляції Кендела та Спірмена необхідно подати дані в форматі, наведеному на рис. 4.8.

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

Рис. 4.8. Вихідні дані

У модулі *Nonparametric Statistics* обрати меню *Correlation*, що дозволяє розрахувати коефіцієнти рангової кореляції. На рис. 4.9 – 4.12 наведено порядок дій для розрахунку коефіцієнта рангової кореляції Спірмена.

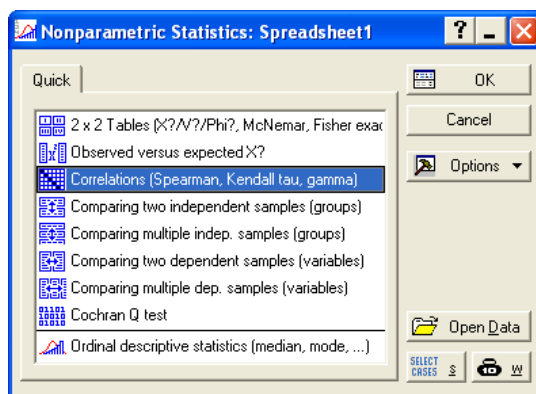


Рис. 4.9. Діалогове вікно вибору виду непараметричних критеріїв

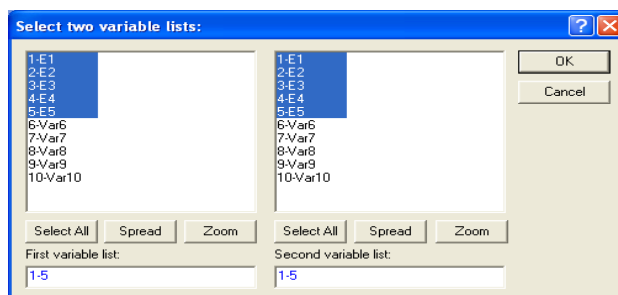


Рис. 4.10. Вибір вихідних даних для розрахунку

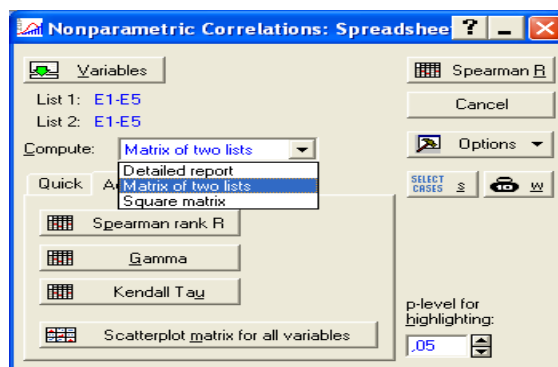


Рис. 4.11. Діалогове вікно модуля "Непараметрична кореляція"

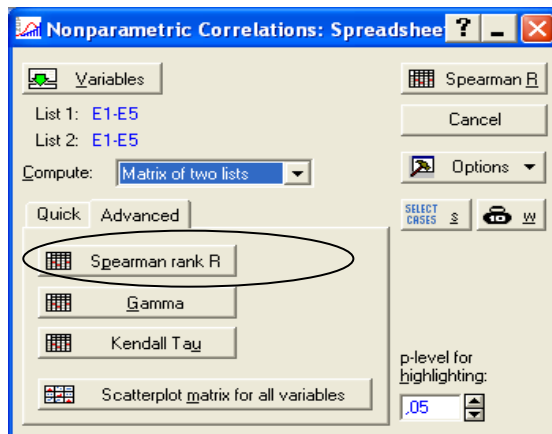


Рис. 4.12. Вибір коефіцієнта рангової кореляції Спірмена для розрахунку

На рис. 4.13 наведено матрицю рангової кореляції Спірмена, з якої видно, що в більшості своїй експерти мають узгоджену думку щодо питань дослідження.

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

Рис. 4.13. Матриця коефіцієнта рангової кореляції Кендела

Для розрахунку коефіцієнта конкордації необхідно подати дані у вигляді, наведеному на рис. 4.14.

Data: Spreadsheet1* (10v by 10c)						
	1	2	3	4	5	6
	X1	X2	X3	X4	X5	Var6
E1	1	2,5	2,5	4	5	
E2	2	2	2	5	4	
E3	1,5	1,5	3	4,5	4,5	
E4	1	2,5	2,5	4,5	4,5	
E5	2	1	3	4	5,5	

Рис. 4.14. Вихідні дані для розрахунку коефіцієнта конкордації

За допомогою модуля *Nonparametric Statistics* обрати меню *Comparing multiple dep. Sample* (рис. 4.15) та перейти до розрахунку рангового дисперсного аналізу Фрідмена та розрахунку коефіцієнта конкордації Кендела (рис. 4.16).

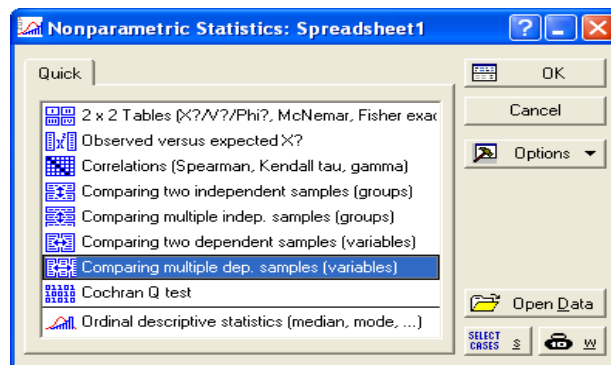


Рис. 4.15. Діалогове вікно модуля "Непараметричні статистики"

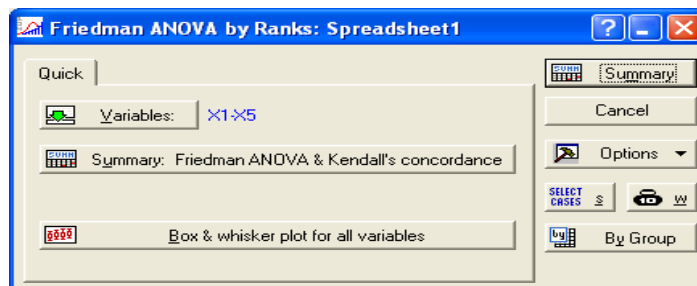


Рис. 4.16. **Діалогове вікно розрахунку рангового дисперсійного аналізу та коефіцієнта конкордації**

На рис. 4.17 наведено результати розрахунку коефіцієнта конкордації Кендела та оцінка його значущості за критерієм χ^2 . З ймовірністю 0,13 % коефіцієнт конкордації є значущим, тобто експерти мають узгоджену думку.

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

Рис. 4.17. **Результат розрахунку рангового дисперсійного аналізу та коефіцієнта конкордації**

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

За допомогою методів математичної статистики необхідно отримати узагальнену думку експертів. Визначають середній ранг, середнє статистичне значення S_j j -ї ознаки:

$$S_j = \frac{\sum_{i=1}^m \alpha_{ij}}{m_{kj}}, \quad (4.3)$$

де i – номер експерта;

$i = 1, \dots, m$; j – номер ознаки, $j = 1, 2, \dots, n$;

m_{kj} – кількість експертів, що оцінюють j -ту ознаку ($m_{kj} \leq m$).

Результати розрахунку наведено в табл. 4.6.

Таблиця 4.6

Розрахунок середнього рангу

Показники	E_1	E_2	E_3	E_4	E_5	S	α_i
ЧДП	1	2	1,5	1	2	1,5	0,099338
ЕЄВВП	2,5	2	1,5	2,5	1	1,9	0,125828
ОВВ	2,5	2	3	2,5	3	2,6	0,172185
ЧВД	4	5	4,5	4,5	4	4,4	0,291391
СЗР	5	4	4,5	4,5	5,5	4,7	0,311258

Використовуючи формулу розрахунку загального рейтингу, в табл. 4.7 розраховано рейтинг за кожною з країн.

Таблиця 4.7

Розрахунок рейтингу країн світу за рівнем ЕБК з використанням оцінок експертів

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

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

Тема 5. Класифікація, види прогнозів, оцінка якості прогнозу

Методичні рекомендації до виконання лабораторної роботи за темою 5 "Прості методи прогнозування"

Мета: отримання практичних навичок щодо моделювання тенденцій розвитку компаній з використанням пакетів *MS Excel* та *Statistica*, а також отримання практичних навичок прогнозування рядів динаміки різними методами.

Завдання: необхідно збудувати прогноз одного з показників освітньої діяльності на наступний рік, використовуючи прості методи прогнозування. Для вибору найбільш адекватного методу необхідно побудувати прогноз цього досліджуваного показника, використовуючи наведені методичні рекомендації та оцінити розбіжності фактичного та прогнозного значення.

Методичні рекомендації

1. За даними про обсяг вартості навчання в одному з регіонів країни за період січень – серпень 2002 року побудувати прогноз методом середнього абсолютного приросту на вересень – листопад 2002 року (табл. 5.1).

Середній абсолютний приріст складає:

$$\bar{\Delta} = \frac{y_n - y_1}{n - 1} = \frac{208,8 - 201,8}{8 - 1} = 1,0 \text{ ум.од.}$$

або

$$\bar{\Delta} = \frac{\sum \Delta_i}{n - 1} = \frac{0,6 + 0,7 + 0,9 + 1,2 + 1,2 + 1,2 + 1,2}{8 - 1} = 1,0 \text{ ум.од.}$$

Результати розрахунків подано в табл. 5.1.

**Розрахункова таблиця для визначення прогнозних значень
методом середнього абсолютного приросту**

Місяці	Вартість навчання, ум. од.	Абсолютні ланцюгові зростання, ($\Delta_{iy} = y_i - y_{i-1}$)	\bar{Y}_{Δ}	$(Y_i - \bar{Y}_{\Delta})$	$(Y_i - \bar{Y}_{\Delta})^2$	Δ_i^2
Січень	201,8	–	201,8	0	0	–
Лютий	202,4	0,6	202,8	–0,4	0,16	0,36
Березень	203,1	0,7	203,8	–0,7	0,49	0,49
Квітень	204,0	0,9	204,8	–0,8	0,64	0,81
Травень	205,2	1,2	205,8	–0,6	0,36	1,44
Червень	206,4	1,2	206,8	–0,4	0,16	1,44
Липень	207,6	1,2	207,8	–0,2	0,04	1,44
Серпень	208,8	1,2	208,8	0	0	1,44
Усього	1 639,3	7,0	–	–	1,85	7,42

Варто перевірити нерівність: $\sigma_{\text{ост}}^2 \leq \rho^2$.

$$\sigma_{\text{ост}}^2 = \frac{1,85}{8} = 0,231;$$

$$\rho^2 = \frac{1}{2} \times \frac{7,42}{8} = 0,464.$$

Отже, основна передумова виконується, що свідчить у тому, що даним методом можна прогнозувати вартість навчання на наступний період.

Прогнозні оцінки цього показника на вересень – листопад 2002 року становитимуть:

$$\hat{y}_{\text{вересень}} = 208,8 + 1 \times 1 = 209,8 \text{ ум. од.}$$

$$\hat{y}_{\text{жовтень}} = 208,8 + 1 \times 2 = 210,8 \text{ ум. од.}$$

$$\hat{y}_{\text{листопад}} = 208,8 + 1 \times 3 = 211,8 \text{ ум. од.}$$

2. За даними табл. 5.2 для проведення маркетингового дослідження необхідно побудувати прогноз на 2019 та 2020 роки.

Обсяг продажу продовольчих товарів

Роки	2013	2014	2015	2016	2017	2018
Обсяг освітніх послуг, тис. грн	365,00	374,00	381,00	396,00	405,00	380,00

Використовуючи метод середнього темпу зростання необхідно на початку розрахувати середній темп зростання за період 2013 – 2018 роки, використовуючи формулу:

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

Таким чином, середній абсолютний приріст за період 2013 – 2018 роки складає:

$$\bar{T}_p = \sqrt[6-1]{\frac{380,0}{365,0}} = 1,01;$$

$$\hat{y}_{2019} = 380,00 \times 1,01^1 = 383,8 \text{ тис. грн};$$

$$\hat{y}_{2020} = 380,0 \times 1,01^2 = 387,64 \text{ тис. грн.}$$

3. Відома динаміка прибутку підприємства за період 1990 – 1998 років (табл. 5.3).

Динаміки прибутку підприємства, тис. грн

Роки	1990	1991	1992	1993	1994	1995	1996	1997	1998
x_i	0	1	2	3	4	5	6	7	8
y_i	40,8	50,5	48,6	40,2	16,6	10,2	14,6	22,0	24,8

Знайти оцінки параметрів моделі.

$$a_1 = \frac{y_n - y_1}{x_n - x_1} = \frac{24,8 - 40,8}{8 - 0} = -2,0;$$

$$a_0 = y_1 - a_1 \times x_1 = 40,8 - (-2) \times 0 = 40,8.$$

Таким чином, модель має вигляд: $\hat{y} = 40,8 - 2x$.

Отримана модель може бути використана для прогнозування.

Використовуючи вихідні дані у табл. 5.3, знайти оцінки параметрів моделі a_0 та a_1 , розбивши попередньо вибірку на дві частини.

I група – дані з 1990 до 1993 р.:

$$\bar{x}_1 = \frac{0 + 1 + 2 + 3}{4} = \frac{6}{4} = 1,5;$$
$$\bar{y}_1 = \frac{40,8 + 50,5 + 48,6 + 40,2}{4} = \frac{180,1}{4} = 45,02.$$

II група – дані з 1994 до 1998 р.:

$$\bar{x}_{11} = \frac{4 + 5 + 6 + 7 + 8}{5} = 6;$$
$$\bar{y}_{11} = \frac{16,6 + 10,2 + 14,6 + 22,0 + 24,8}{5} = 17,64.$$

Знайти параметри моделі:

$$a_1 = \frac{17,64 - 45,02}{6 - 1,5} = \frac{-27,38}{4,5} = -6,084;$$
$$a_0 = 45,02 - (-6,084) \times 1,5 = 54,146.$$

Таким чином, модель має вигляд $y = 54,146 - 6,084 x$.

4. Порівнюючи параметри двох лінійних моделей, побудованих різними методами, можна переконатися у суттєвому відмінності оцінок параметрів. Отримана друга модель точніша і краще може бути використана для простого прогнозу.

Є такі дані, що однією компанією було надано освітні послуги за січень – вересень 2002 року (табл. 5.4).

Таблиця 5.4

Розрахункова таблиця для визначення знаків відхилення

Місяці	Обсяг освітніх послуг, ум. од., y_i	Знаки відхилення
1	2	3
Січень	4,5	-
Лютий	5,2	+

1	2	3
Березень	5,3	+
Квітень	6,7	+
Травень	6,1	-
Червень	6,4	+
Липень	5,8	-
Серпень	5,0	-
Вересень	4,3	-

Побудований за цими даними ряд розподілу знаків відхилень має вигляд (табл. 5.5).

Таблиця 5.5

Розподіл знаків відхилення

Види тенденцій	Довжина сприятливої тенденції, τ_i	Частота, f_i
--	0	2
--+	1	1
-++-	2	0
-+++	3	1

На основі даних таблиці визначають середню довжину сприятливої тенденції:

$$\bar{\tau} = \frac{\sum \tau f}{\sum f} = \frac{0 \times 2 + 1 \times 1 + 2 \times 0 + 3 \times 1}{2 + 1 + 0 + 1} = 1,25.$$

Інтенсивність переривань сприятливої тенденції становитиме:

$$\lambda = \frac{1}{\bar{\tau}} = \frac{1}{1,25} = 0,8.$$

Імовірність сприятливої тенденції визначають на основі такої модифікації закону розподілу Пуассона:

$$\rho = e^{-\lambda L},$$

де ρ – ймовірність здійснення сприятливої тенденції;

λ – інтенсивність переривань сприятливої тенденції;

L – період попередження (кількість років збереження сприятливої тенденції).

Таким чином, ймовірність сприятливої тенденції складе (табл. 5.6).

Таблиця 5.6

Ймовірність здійснення сприятливої тенденції

Період збереження сприятливої тенденції	t	λ	$-\lambda_t$	Ймовірність сприятливої тенденції, ρ
Жовтень	1	0,8	-0,8	0,449
Листопад	2	0,8	-1,6	0,202
Грудень	3	0,8	-2,4	0,091

Таким чином, із ймовірністю 44,9 % можна стверджувати, що обсяг освітніх послуг у жовтні зростає порівняно з вереснем.

Рекомендована література

Основна

1. Бандоріна Л. М. Моделювання економіки : навч. посіб. / Л. М. Бандоріна, Л. І. Лозовська, Л. М. Савчук. – Дніпро : УДУНТ, 2022. – 154 с.
2. Вітлінський В. В. Економіко-математичні методи та моделі: оптимізація : навч. посіб. / В. В. Вітлінський, Т. О. Терещенко, С. С. Савіна. – Київ : КНЕУ, 2016. – 303 с.
3. Галушак М. П. Прогнозування соціально-економічних процесів : навчальний посібник для економічних спеціальностей / М. П. Галушак, О. Я. Галушак, Т. І. Кужда. – Тернопіль : ФОП Паляниця, 2021. – 160 с.
4. Макроекономічне прогнозування та його принципи. Національна економіка : навч. посіб. / [В. І. Мельникова, О. П. Мельникова, Т. В. Сідлярчук та ін.]. – 2-ге вид. перероб. та доп. – Київ : Центр учбової літератури, 2012. – 248 с.

5. Прогнозування соціально-економічних процесів : навчальний посібник для студентів напряму підготовки 6.030502 "Економічна кібернетика" денної форми навчання / Т. С. Клебанова, В. А. Курзенев, В. М. Наумов та ін. – Харків : ХНЕУ ім. С. Кузнеця, 2015. – 656 с.

6. Статистичні методи та моделі оцінювання й прогнозування поведінки соціально-економічних систем в умовах інформаційної економіки : звіт кафедри статистики і економічного прогнозування за 2019 – 2020 навч. рік / викон. О. В. Раєвнєва, І. В. Аксьонова, О. І. Бровко та ін. – Харків : ХНЕУ ім. С. Кузнеця, 2020. – 107 с.

7. Моделювання управління економічними системами і процесами: навчально-методичний комплекс дисципліни [Електронний ресурс] : навч. посіб. для здобувачів ступеня доктора філософії за освітньо-науковою програмою "Менеджмент" спеціальності 073 "Менеджмент" КПІ ім. Ігоря Сікорського; уклад. : С. О. Солнцев, Ж. М. Жигалкевич. – Київ : КПІ ім. Ігоря Сікорського, 2021. – 50 с. – Режим доступу : <https://ela.kpi.ua/handle/123456789/48693>.

Додаткова

8. Демяненко А. Є. Методи і моделі первинної обробки інформації з елементами самоподібності : пояснювальна записка до атестаційної роботи здобувача вищої освіти на другому (магістерському) рівні, спеціальність 123 Комп'ютерна інженерія / А. Є. Демяненко ; М-во освіти і науки України, Харків. нац. ун-т радіоелектроніки. – Харків, 2020. – 78 с.

9. Лабораторний практикум з навчальної дисципліни "Аналіз та прогнозування рядів динаміки" для студентів напряму підготовки 6.030506 "Прикладна статистика" денної форми навчання / уклад. О. В. Раєвнєва, І. В. Чанкіна, О. І. Бровко ; Харківський національний економічний університет ім. С. Кузнеця. – Електрон. текстові дан. (9,19 МБ). – Харків : ХНЕУ ім. С. Кузнеця, 2014. – 59 с.

10. Матеріали міжнародної науково-практичної конференції "Моделювання соціально-економічних процесів: регіональні та галузеві аспекти" / за ред. Б. Ю. Кишакевича. – Дрогобич : Дрогобицький державний педагогічний університет імені Івана Франка, 2016. – 252 с.

11. Статистичне моделювання та прогнозування. Змістові модулі 3 та 4 : методичні рекомендації до лабораторних та самостійних робіт студентів спеціальності 051 "Економіка" першого (бакалаврського) рівня

/ уклад. О. В. Раєвнєва, О. І. Бровко, В. І. Дериховська. – Харків : ХНЕУ ім. С. Кузнеця, 2022. – 105 с.

12. Статистичне моделювання та прогнозування. Модуль 1 та 2. Методичні рекомендації до лабораторних та самостійних робіт для студентів спеціальності 051 "Економіка" першого (бакалаврського) рівня [Електронний ресурс] / уклад. О. В. Раєвнєва, О. І. Бровко, В. І. Дериховська. – Харків : ХНЕУ ім. С. Кузнеця, 2020. – 73 с.

13. Цибульник С. О. Сучасні методи обробки інформації. Лекції [Електронний ресурс] : навчальний посібник для здобувачів ступеня доктора філософії за освітньою програмою "Автоматизація та комп'ютерно-інтегровані технології" спеціальності 151 "Автоматизація та комп'ютерно-інтегровані технології" / С. О. Цибульник, О. М. Павловський ; КПІ ім. Ігоря Сікорського. – Електронні текстові дані (1 файл: 709 Кбайт). – Київ : КПІ ім. Ігоря Сікорського, 2021. – 111 с. – Режим доступу : <https://ela.kpi.ua/handle/123456789/45934>.

Інформаційні ресурси

14. Офіційний сайт Головного управління статистики в Харківській області. – Режим доступу : <http://uprstat.kharkov.ukrtel.net/>.

15. Офіційний сайт департаменту статистики Організації Об'єднаних Націй. – Режим доступу : <http://unstats.un.org/unsd/default.htm>.

16. Офіційний сайт Державної служби статистики України. – Режим доступу : www.ukrstat.gov.ua.

17. Сторінка курсу на платформі *Moodle* (персональна навчальна система). – Режим доступу : <https://pns.hneu.edu.ua/course/view.php?id=5004>.

Зміст

Вступ	3
Методичні рекомендації до лабораторних робіт	4
Змістовий модуль 1. Засади моделювання освітніх процесів	4
Тема 1. Статистичне моделювання як сучасний метод обробки інформації	4
Тема 2. Формування інформаційного простору наукового дослідження: попередній аналіз даних	10
Змістовий модуль 2. Методи та технології кількісно- якісного аналізу та прогнозування освітньої інформації	17
Тема 3. Методи та технології обробки кількісної інформації	17
Тема 4. Методи та технології обробки якісної інформації	20
Тема 5. Класифікація, види прогнозів, оцінка якості прогнозу	34
Рекомендована література	39
Основна	39
Додаткова	40
Інформаційні ресурси	41

НАВЧАЛЬНЕ ВИДАННЯ

СУЧАСНІ МЕТОДИ І ТЕХНОЛОГІЇ ОБРОБКИ ІНФОРМАЦІЇ В ОСВІТІ

**Методичні рекомендації
до лабораторних робіт
для здобувачів спеціальності
011 "Освітні, педагогічні науки"
освітньо-наукової програми
"Освітні, педагогічні науки"
третього (освітньо-наукового) ступеня**

Самостійне електронне текстове мережеве видання

Укладачі: **Раєвська** Олена Валентинівна
Бровко Ольга Іванівна
Стрижиченко Костянтин Анатолійович

Відповідальний за видання *О. В. Раєвська*

Редактор *В. О. Дмитрієва*

Коректор *Н. В. Завгородня*

План 2023 р. Поз. № 147 ЕВ. Обсяг 43 с.

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

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