

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

СТАТИСТИКА II

Методичні рекомендації
до лабораторних та самостійних робіт
для студентів спеціальності 051 "Економіка"
першого (бакалаврського) рівня

Харків
ХНЕУ ім. С. Кузнеця
2020

УДК 311(07.034)

C78

Укладачі: І. В. Аксьонова

О. І. Бровко

Г. І. Свидло

Затверджено на засіданні кафедри статистики і економічного прогнозування.

Протокол № 3 від 16.09.2019 р.

Самостійне електронне текстове мережеве видання

Статистика II [Електронний ресурс] : методичні рекомендації до лабораторних та самостійних робіт для студентів спеціальності 051 "Економіка" першого (бакалаврського) рівня / уклад. І. В. Аксьонова, О. І. Бровко, Г. І. Свидло. – Харків : ХНЕУ ім. С. Кузнеця, 2020. – 84 с.

Подано методичні рекомендації до виконання лабораторних та самостійних робіт, метою яких є закріплення й поглиблення знань теоретичного та практичного матеріалу з методології статистичного дослідження соціально-економічних явищ і процесів, поглиблення навичок аналізу соціально-економічних процесів за допомогою ППП Statistica та засобів Excel.

Рекомендовано для студентів спеціальності 051 "Економіка" першого (бакалаврського) рівня.

УДК 311(07.034)

© Харківський національний економічний
університет імені Семена Кузнеця, 2020

Вступ

Зростаючий інтерес до статистики та бізнес-аналітики обумовлений сучасним розвитком економіки, інтеграційних процесів, розвитком міжнародного співробітництва, впровадженням соціально-економічних реформ. Перед статистикою поставлені важливі завдання щодо подальшого вдосконалення методології розрахунку найважливіших соціально-економічних показників та її гармонізації з міжнародними стандартами, забезпечення усіх рівнів загальнодержавного управління надійною, своєчасною та вірогідною інформацією.

Статистика здійснює збирання, оброблення та аналіз даних, які характеризують усі аспекти суспільного життя, виявляє взаємозв'язки між явищами та процесами, вивчає тенденції їхнього розвитку для ухвалення ефективних управлінських рішень на мікро-, мезо- та макроекономічних рівнях.

Мету статистики визначають тими функціями, які вона виконує в системі економічних наук, що полягає в набутті та поглибленні студентами необхідних знань та вмінь під час вирішення наступних завдань: розроблення інформаційно-аналітичного забезпечення дослідження, аналітичному групуванні масових даних явищ та процесів суспільного життя й дослідженні взаємозв'язків між ними, обчисленні узагальнюючих показників структури і структурних зрушень, визначенні факторів, що формують варіацію та розвиток масових явищ й процесів, вимірюванні інтенсивності динаміки, побудові моделей і прогнозуванні тенденцій розвитку явищ та процесів, визначенні сили впливу факторних ознак на результативну в статистиці й динаміці.

Вивчення дисципліни "Статистика II" вимагає, поряд із засвоєнням теоретичних положень, практичної реалізації основних розділів навчальної дисципліни шляхом виконання відповідних лабораторних та самостійних робіт.

У результаті вивчення навчальної дисципліни "Статистика II" студент набуває наступних компетентностей, що наведені в Додатку А.

Методичні рекомендації до лабораторних та самостійних робіт

Змістовий модуль 1

Методи статистичного спостереження, побудови та аналізу рядів розподілу

Тема 1. Організація статистики в Україні. Міжнародні статистичні організації

Методичні рекомендації

Лабораторне заняття

На занятті студенти повинні виконати наступні завдання:

1. Написати есе за одним із наступних питань:

- 1) стратегія розвитку державної статистики: основні напрями;
- 2) характеристика метаописів державних статистичних спостережень;
- 3) нормативно-правова та регуляторна діяльність Держстату;
- 4) характеристика основних Класифікаторів та класифікацій, які використовують у своїй діяльності органи державної статистики;
- 5) нормативно-правова база та послідовність дій для електронного звітування;
- 6) єдиний державний реєстр підприємств та організацій України та його показники;
- 7) характеристика анкетних опитувань користувачів статистичної інформації;
- 8) статистика за кордоном: міжнародні організації та статистичні служби країн світу;
- 9) Україна в міжнародних індексах;
- 10) характеристика сайту Державної служби статистики України.

2. Зробити презентацію есе й доповідь на 5 – 7 хвилин.

В есе студент повинен освітити проблематику за обраним питанням, стан його розвитку в сучасних умовах, вказати позитивні та негативні сторони вирішення питання, навести зарубіжний досвід. Наприкінці

есе необхідно вказати компетентності, яких набув студент під час його написання й вивчення обраного питання.

Завдання для самостійного опрацювання

1. Охарактеризуйте основні завдання статистики на сучасному етапі.
2. Поясніть якими органами здійснюється статистична робота в Україні.
3. Надайте описання структурним підрозділам Державної служби статистики України.
4. Охарактеризуйте основні інформаційні джерела статистичних даних (сайти, збірники, тощо).
5. Поясніть, чим займається Статистична комісія.
6. Охарактеризуйте структуру Глобальної статистичної системи.
7. Наведіть приклади та основні задачі міжнародних органів статистики.

Тема 2. Історія статистики

Методичні рекомендації

Лабораторне заняття

На даному занятті студенти повинні виконати наступні завдання:

1. Написати есе за одним із наступних питань:
 - 1) політична арифметика та державоведення в XVI – XVII ст.;
 - 2) статистична діяльність А. Кетле;
 - 3) академічна статистика. Передумови формування стохастичної статистики;
 - 4) земська статистика в 1861 – 1917 рр.;
 - 5) зарубіжна статистика XIX ст.;
 - 6) зарубіжна статистика XX ст.;
 - 7) розвиток статистики в 20 – 30 рр. XX ст.;
 - 8) розвиток та реорганізація статистики в 50-ті рр. XX ст.;
 - 9) статистична наука в 60 – 80-ті рр. XX ст.;
 - 10) перебудова статистики в 90-ті рр. XX ст.

2. Зробити презентацію есе й доповідь на 5 – 7 хвилин.

В есе студент повинен освітити проблематику за обраним питанням, стан його розвитку в конкретний період часу, вказати позитивні та негативні сторони вирішення питання, порівняти вітчизняний та зарубіжний досвід. Наприкінці есе необхідно вказати компетентності, яких набув студент при його написанні й вивченні обраного питання.

Завдання для самостійного опрацювання

1. Поясніть, що вивчала статистика як сфера практичної діяльності у стародавні часи (Китай, Греція, Рим).

2. Охарактеризуйте основні періоди в історії статистики.

3. Поясніть на конкретних прикладах, що таке цenz.

4. Охарактеризуйте чим займалась школа політичних арифметиків.

5. Охарактеризуйте статистико-математичний напрям розвитку статистики.

6. Поясніть, у чому полягає сутність таблиці смертності і до якого напрямку розвитку статистики належить її побудова?

7. Наведіть історію розвитку кафедри статистики при Академії наук України.

Тема 3. Інформаційне забезпечення статистичного дослідження

Методичні рекомендації

Лабораторне заняття

Будь-яке статистичне спостереження проводять заздалегідь розробленим планом, який містить програмно-методологічні та організаційні питання.

Програмно-методологічні питання плану статистичного спостереження містять:

визначення мети і завдань спостереження;

вибір об'єктів та одиниць спостереження;

розроблення програми спостереження;

вибір форми, виду і способу проведення спостереження.

Мета статистичного спостереження має бути конкретною та чітко сформульованою, виходити із загальних завдань, поставлених перед статистичним дослідженням. Відповідно до принципів системного підходу завдання спостереження повинні відповідати поставленій меті, виходити з неї. Мета і завдання зумовлюють розроблення програми й вибір форм проведення спостереження, а конкретність їх формулювання забезпечують повноту отримання даних.

Залежно від мети і поставлених завдань визначають об'єкт спостереження – статистичну сукупність соціально-економічних явищ чи процесів, що підлягають обстеженню. Інколи для обмеження об'єкта спостереження використовують поняття цензу. Ценз – це низка ознак, кількісне значення яких під час проведення статистичного спостереження служить підставою для обліку (або невіднесення) одиниці до сукупності, що досліджується.

Будь-який об'єкт спостереження як статистична сукупність складається з одиниць сукупності. Одиниця сукупності – елемент статистичної сукупності, який є носієм ознак, що підлягають реєстрації.

Кількість одиниць сукупності називається обсягом сукупності.

Одиниці сукупності володіють безліччю різних ознак. Статистична ознака – це конкретна властивість, якість, відмінна риса одиниці спостереження.

Однак не кожна одиниця сукупності може надати про себе інформацію. Тому в ході обстеження виділяють одиницю спостереження (звітну одиницю). Одиниця спостереження – це первинна одиниця, від якої одержують інформацію. Одиниця сукупності та одиниця спостереження можуть співпадати, а можуть і не співпадати.

Відповідно до поставленої мети, визначеним завданням, обраним об'єктом розробляється програма спостереження.

Програма спостереження – перелік ознак, що підлягають реєстрації (в разі безпосереднього спостереження), або це перелік питань, за допомогою яких збираються відомості (при опитуваннях). Складання програми спостереження є складним і відповідальним процесом, що відбиває якість зібраної інформації. Склад та зміст програми спостереження визначаються завданнями дослідження та особливостями суспільно-економічного явища, що досліджується. Будь-яке явище має багато ознак. Збирати інформацію з усіх – недоцільно і неможливо. Тому необхідно

відібрати найбільш важливі ознаки, що відповідають поставленим завданням і відповідній меті спостереження.

Одночасно з програмою спостереження розробляється інструментарій у вигляді статистичних формулярів та інструкцій щодо їх заповнення.

Статистичний інструментарій забезпечує не тільки вхідну, але і вихідну частину інформаційної бази спостереження. Тобто, визначаючи ознаки, складаючи блоки питань, одночасно готують макети вихідних таблиць, де немає цифрової інформації. За макетами таблиць можна визначити, наскільки кожне питання (ознака) узгоджується з іншими, передбачити, як "працює" те чи інше питання (ознака), вилучити малоінформативні питання, обґрунтувати методику подальшого статистичного оброблення інших питань.

Статистичний формуляр – це документ єдиного зразка, що містить програму і результати спостереження. Він може мати різні назви: бланк обстеження, переписний лист, анкета, звіт.

Обов'язковими елементами статистичного формуляра є титульна та адресна частини. У титульній частині вказують: назву статистичного спостереження та органу, що його проводить; номер формуляра, а також, хто і коли його ухвалив. В адресній – адреса звітної одиниці, її підпорядкованість.

Інструкція передбачає порядок проведення спостереження чи заповнення формуляра. Залежно від складності програми спостереження це може бути документ у вигляді окремої брошури, або підказки у відповідях, або роз'яснення на зворотньому боці бланка.

Програма спостереження передбачає також визначення виду і способу реєстрації даних. Зазвичай вид і спосіб спостереження залежать від його мети, сутності об'єкта спостереження, обсягу і ступеня точності очікуваних результатів.

Другою складовою частиною плану спостереження є комплекс організаційних питань, до яких належать:

визначення органу (виконавця) спостереження – суб'єкт спостереження (спостереження може проводитися власними силами або організаціями, що спеціалізуються на проведенні спостереження);

визначення часу спостереження: дати початку, дати закінчення спостереження, критичної дати. Строк (період) спостереження встановлюється виходячи з обсягу роботи та чисельності персоналу, котрий буде

зайнятий збиранням даних. Критичною датою вважається конкретний день року (час дня), станом на який проводиться реєстрація ознак по кожній одиниці статистичної сукупності;

визначення місця (території) проведення спостереження.

Залежно від масштабності об'єкта спостереження, а також зацікавленості щодо його результатів суб'єктами спостереження можуть бути: центральні органи державної статистики, статистичні відділи міністерств і відомств, що здійснюють державні обстеження локального характеру за певною тематикою, спеціальні інститути, агентства, міжнародні організації, що проводять обстеження, котрі засновані на вивченні: громадської думки або мотивації, поведінки та оцінок окремих суб'єктів суспільно-економічного життя, аналітичні відділи окремих економічних структур (підприємств, організацій, фірм, банків, бірж, страхових товариств тощо), що проводять на мікрорівні обстеження маркетингового або контрольного характеру.

Обґрунтування місця обстеження передбачає визначення пункту, в якому знаходиться одиниця спостереження і реєструються дані.

Для того щоб спостереження дало ймовірні і своєчасні дані, необхідно вирішити питання часу і періоду спостереження.

Час спостереження (об'єктивний час) – це час, до якого належать дані спостереження. Коли об'єктом спостереження є процес, то вибирають інтервал часу, протягом якого накопичуються дані. Якщо об'єктом спостереження є певний стан, то вибирають критичний момент – момент часу, станом на який реєструються дані. Критичний момент використовують під час проведення перепису населення (коли потрібно точно зареєструвати його стан в один "момент"). Зрозуміло, що переписати все населення країни в одну мить неможливо, тому крім часу спостереження встановлюється період спостереження (суб'єктивний час), протягом якого реєструються дані.

Для проведення будь-якого статистичного спостереження необхідно відповідне матеріально-технічне забезпечення: друковані засоби, обчислювальна та розмножувальна техніка, транспортні засоби, статистичний інструментарій та рекламні носії.

Одночасно методологічним (як засіб запобігання помилок) і організаційним (як засіб виявлення та виправлення помилок) питанням плану спостереження є контроль даних.

Контроль означає, насамперед, перевірку даних обстежень щодо їх повноти та достовірності. Повнота даних контролюється, як правило, візуально: перевіряють наявність даних по всім одиницям і позиціям. Достовірність даних перевіряють засобами логічного та арифметичного контролю. Логічний контроль – це перевірка сумісності даних, яка полягає в порівнянні взаємозалежних ознак. Встановити розмір помилки і виправити її можна засобами арифметичного контролю: прямим або побічним перерахунком зареєстрованих даних.

На цьому лабораторному занятті студенти повинні виконати наступні завдання:

1. Обрати приклад статистичного спостереження та назвати: мету спостереження, об'єкт спостереження, одиницю спостереження, одиницю сукупності; розробити програму спостереження; передбачити контроль даних спостереження. Відобразити програмно-методологічні та організаційні питання плану даного статистичного спостереження у вигляді таблиці, в якій вказати: загальну характеристику складових програмно-методологічних та організаційних питань плану статистичного спостереження та конкретну характеристику стосовно конкретного спостереження.

Приклади статистичних спостережень:

- а) ринок праці Харківської області;
- б) процеси природного руху населення країни;
- в) міграційні процеси за регіонами країни;
- г) зовнішня торгівля товарами та послугами;
- д) інвестиційна діяльність країни;
- е) інноваційна діяльність підприємств та організацій регіону;
- є) шлюбна ситуація в регіонах країни;
- ж) тенденції розвитку макроекономічної ситуації;
- з) ринок освітніх послуг;
- и) ринок послуг системи охорони здоров'я.

2. Зробити презентацію та доклад. Презентація повинна містити основні питання за формуванням інформаційного забезпечення конкретного спостереження, а саме: законодавчу базу для аналізу обраного явища або процесу; основні джерела інформації (форми статистичної звітності,

вибіркові обстеження тощо) для аналізу обраного явища або процесу; характеристику програмно-методологічних та організаційних питань плану статистичного спостереження обраного явища або процесу; перелік питань програми спостереження й яку інформацію вони нададуть.

Завдання для самостійного опрацювання

1. Складіть проект програми статистичного вивчення бюджету часу студента вашого курсу і факультету відносно підготовки до лабораторних занять.

2. Перелічіть програмно-методологічні й організаційні питання статистичного спостереження: перепис населення; цінова політика підприємства на ринку; зайнятість випускників університету.

3. Визначте об'єктивний та суб'єктивний час спостереження за наступними даними: а) пробний перепис населення здійснюється з 17 на 18 лютого (станом на 0 годин) протягом 10 днів; б) дані обліку дебіторської заборгованості банку на початок року мають бути подані протягом 10 днів з початку наступного року; в) строк подання місячних звітів про виробничо-фінансову діяльність державних підприємств – не пізніше 7 числа наступного місяця.

4. Визначте вид статистичного спостереження за ступенем охоплення одиниць та за часом реєстрації даних: а) облік числа мігрантів; б) реєстрація даних торгів на валютних біржах; в) реєстрація рівня цін на продукти харчування, що продаються на ринках великих міст; г) обстеження бюджетів молодих сімей регіону; д) облік чисельності померлих; е) опитування домогосподарств регіону з питань їх доходів.

5. Визначте організаційну форму та спосіб спостереження: а) інвентаризація устаткування на підприємствах транспорту; б) перелік усіх зареєстрованих економічних структур із зазначенням їх реквізитів (назва, адреса, телефон, телефакс), виду та сфери діяльності; в) обстеження користувачів соціальних мереж із приводу їх інтерфейсу; г) щоденний облік обсягу реалізації продукції підприємством; д) фінансова звітність підприємств промисловості; е) маркетингове дослідження насиченості ринку айфонами.

Тема 4. Перевірка статистичних гіпотез

Методичні рекомендації

Лабораторна робота 1. Набуття навичок перевірки статистичних гіпотез за допомогою ППП Statistica 10.0

Завдання. На курсах вивчення іноземних мов використовують дві методики – традиційну та нову. Для порівняння ефективності нової методики було проведено тестування 10-ти студентів за 100-бальною шкалою (табл. 1). Оцінити, чи є випадковістю розбіжність в отриманих балах в обох випадках, дослідивши розбіжність у середніх і в дисперсії між вибірками, використовуючи параметричні та непараметричні критерії.

Таблиця 1

Оцінювання студентів за різними методиками

Номер студента	Традиційна методика	Нова методика
1	98	98
2	98	95
3	96	95
4	98	96
5	94	96
6	97	94
7	97	93
8	98	94
9	97	95
10	96	93

1. Введення вихідних даних в робочий аркуш пакету Statistica 10.0.

Використовуючи інструменти пакета статистика Statistica 10.0, вводиться два стовпця вихідних даних, що позначається як X1 та X2 (рис. 1).

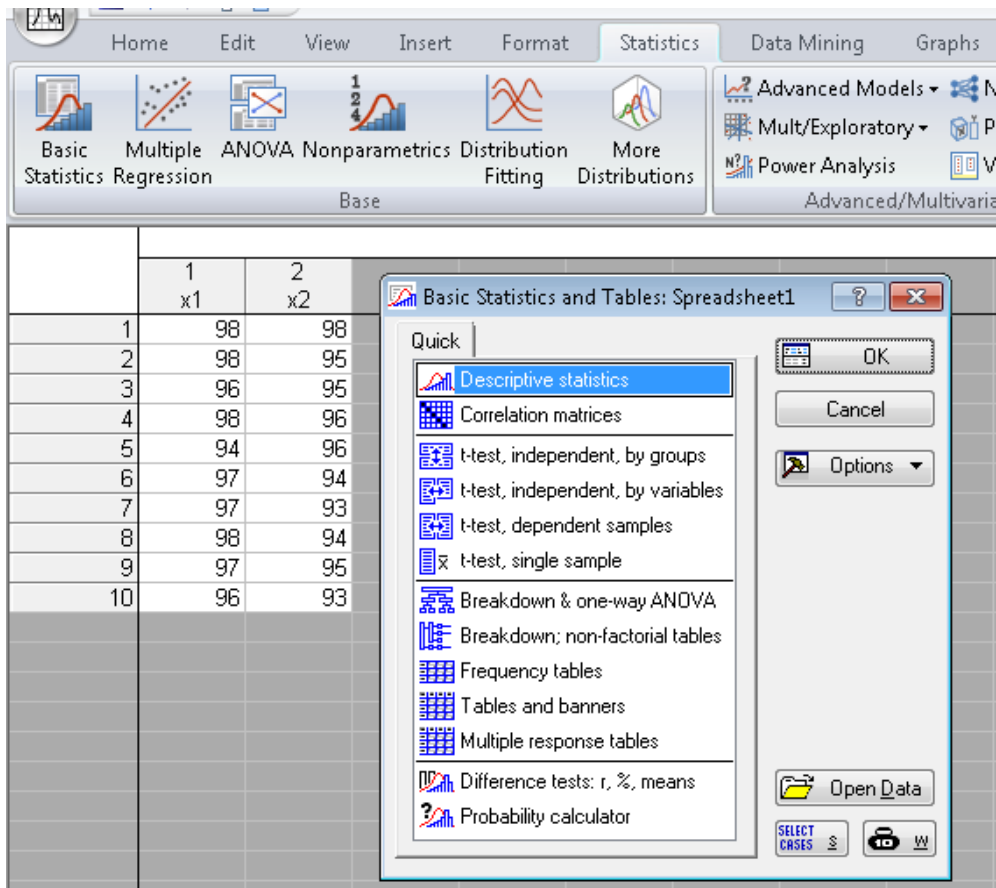


Рис. 1. Введення вхідної інформації

2. 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 представлений чотирма процедурами (рис. 1):

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 груповою змінною) 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 в якості нульової гіпотези передбачається, що середні в групах рівні.

Рівень значущості p t-критерію дорівнює ймовірності помилково відкинути гіпотезу 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 прийнятою межею статистичної значущості є 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 (тобто гіпотеза H_0 відхиляється з певним рівнем достовірності). Значення отриманих балів студентами під час використання традиційної та нової методики навчання отримані з різних вибірок, для порівняння середніх можна застосувати процедуру t-test, independent, by variables (рис. 2).

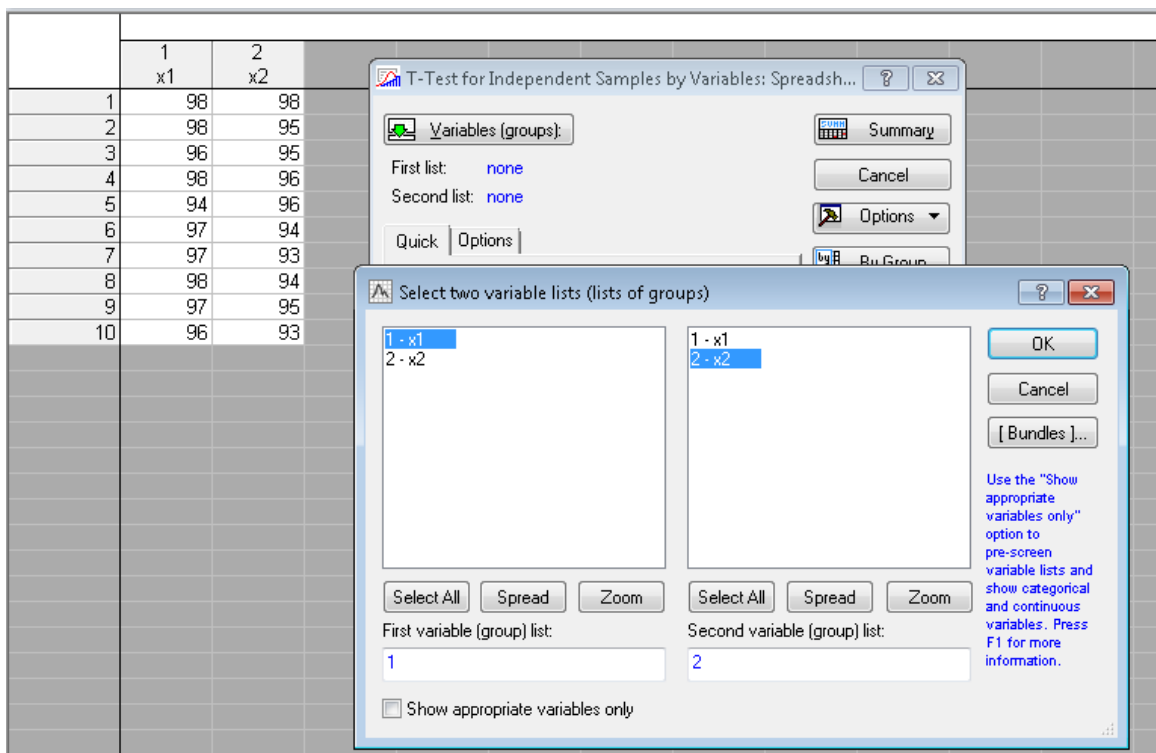


Рис. 2. Вибір вікон під час розрахунку t-критерію

Результати розрахунків наведено на рис. 3.

		T-test for Independent Samples (Spreadsheet1)									
		Note: Variables were treated as independent samples									
Group 1 vs. Group 2	Mean Group 1	Mean Group 2	t-value	df	p	Valid N Group 1	Valid N Group 2	Std.Dev. Group 1	Std.Dev. Group 2	F-ratio Variances	p Variances
x1 vs. x2	96,90000	94,90000	3,171099	18	0,005288	10	10	1,286684	1,523884	1,402685	0,622318

Рис. 3. Результати розрахунків

На рис. 3 наведені значення наступних 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. 3 можна зробити висновок, що середні бали студентів, які навчалися за різними методиками, відрізняються істотно (p – рівень значимості t-критерію менше 0,05), тобто гіпотеза H_0 , яка полягала в тому, що отримані бали студентів, які навчалися за традиційною та новою методиками, є однаковими і відхилення в середніх між ними є випадковим, відхиляється. Маємо, що методики навчання відрізняються ефективністю і відхилення в середніх балах студентів не випадково. Висновок є статистично значимим, оскільки вірна гіпотеза про рівність дисперсій (p Variances значно більше, ніж 0,05).

3. Розрахунок непараметричних критеріїв порівняння середніх.

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

3.1. Формування вхідного простору показників.

Для розрахунку непараметричних критеріїв необхідно перетворити вихідну інформацію у вигляд, наведений на рис. 4, де X – це стовпчик зі значеннями отриманих балів, а змінна Type приймає значення 1 для традиційної методики навчання, 2 – для нової.

	1 x1	2 x2	3 Type	4 X
1	98	98	1	98
2	98	95	1	98
3	96	95	1	96
4	98	96	1	98
5	94	96	1	94
6	97	94	1	97
7	97	93	1	97
8	98	94	1	98
9	97	95	1	97
10	96	93	1	96
11			2	98
12			2	95
13			2	95
14			2	96
15			2	96
16			2	94
17			2	93
18			2	94
19			2	95
20			2	93

Рис. 4. Вихідна інформація для розрахунку непараметричних критеріїв

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 модулю наведена на рис. 5. Модуль дозволяє розрахувати критерій серій Вальда – Вольфовіца, двовибірковий тест Колмогорова – Смірнова 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).

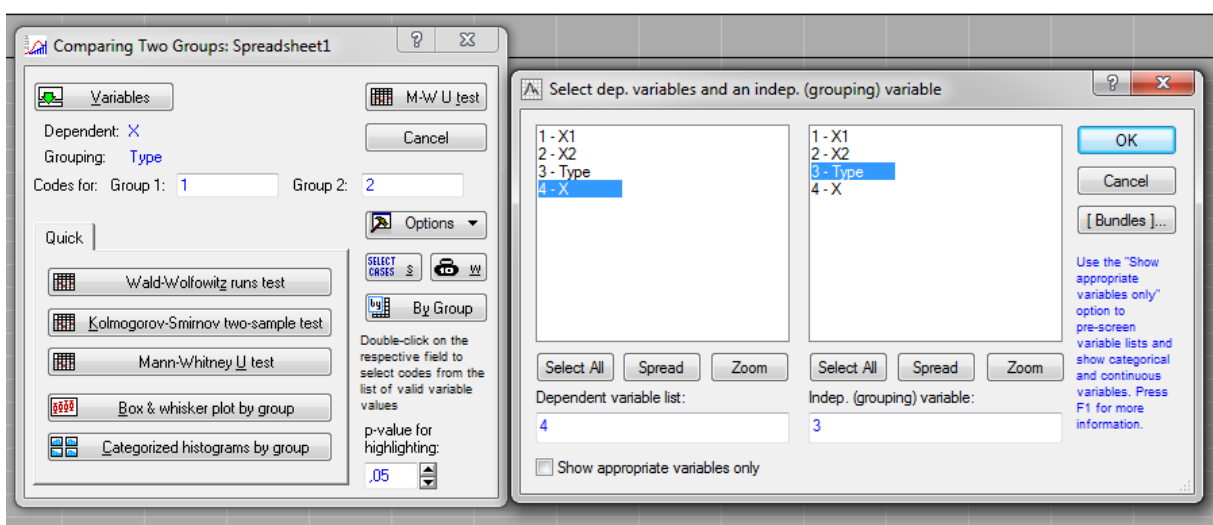


Рис. 5. Стартова панель модулю розрахунку непараметричних критеріїв

3.2. Розрахунок критерію серій Вальда – Вольфовіца.

Критерій перевіряє гіпотезу про те, що дві незалежні вибірки описують два різні процеси, і які суттєво різняться не лише в значеннях середніх, а й за формою розподілу. Нульова гіпотеза (H_0) полягає в тому, що дві вибірки описують один і той самий процес, тобто мають однакову форму розподілу балів за методиками.

За розрахунками (рис. 6) можна зробити висновок, що критерій має рівень значущості $p = 0,25$, що менше $0,05$, тобто даний тест надає аргументи для відхилення нульової гіпотези з 5 % рівнем значущості.

Wald-Wolfowitz Runs Test (Spreadsheet1)										
By variable Type										
Marked tests are significant at $p < ,05000$										
Variable	Valid N Group 1	Valid N Group 2	Mean Group 1	Mean Group 2	Z	p-value	Z adjstd	p-value	No. of Runs	No. of ties
X	10	10	96,90000	94,90000	-1,37840	0,168079	1,148671	0,250693	8	5

Рис. 6. Розрахунок критерію серій Вальда – Вольфовіца

3.3. Розрахунок двовибіркового критерію Колмогорова – Смірнова.

Двовибірковий критерій Колмогорова – Смірнова чутливий не тільки до розбіжності в положенні двох розподілів, але також і до форми розподілу (рис. 7). Фактично він чутливий до будь-якого відхилення від гіпотези однорідності, але не вказує, з яким саме відхиленням ми маємо справу.

Kolmogorov-Smirnov Test (Spreadsheet1)									
By variable Type									
Marked tests are significant at $p < ,05000$									
variable	Max Neg Differnc	Max Pos Differnc	p-value	Mean Group 1	Mean Group 2	Std.Dev. Group 1	Std.Dev. Group 2	Valid N Group 1	Valid N Group 2
X	0,00	0,600000	$p < .10$	96,90000	94,90000	1,286684	1,523884	10	10

Рис. 7. Таблиця розрахунку двовибіркового критерію Колмогорова – Смірнова

Якщо рівень значущості критерію був би менше $0,01$ (const), то це свідчило б про значиму розбіжність між вибірками, що може бути пов'язано не лише з різницею в середніх значеннях, але й з розбіжностями у формі

розподілу, тобто вибірки з різних сукупностей, отже різні методики навчання передбачають різні результати навчання, що означало б, що H_0 відхиляється. У нашому випадку $p < 0,1$ і це може бути як менше, ніж $0,01$, а може бути й більше ніж $0,01$, тобто цей критерій надає недостатньо аргументів для відхилення нульової гіпотези.

3.4. Розрахунок U-тесту Манна – Уїтні.

Цей тест призначений для оцінювання відмінності величин – членів двох вибірок. Цей критерій заснований на підрахунку числа інверсій U (перестановок) членів у їх загальному упорядкованому ряді.

З розрахунку видно, що критерій є статистично значимим ($p_{\text{теор}} = 0,011 > p_{\text{факт}} = 0,010$), тобто можна зробити висновок про наявність розбіжності в середніх двох вибірок та формі розподілу в них, отже гіпотеза H_0 відхиляється.

З розрахунків на рис. 8 також видно, що величина $Z = 2,53$ перевищує квантиль стандартного розподілу $u_{1-1/2\alpha} = u_{0,95} = 1,645$ (табличне значення), що також підтверджує те, що гіпотеза H_0 відхиляється.

Mann-Whitney U Test (Spreadsheet1)										
By variable Type										
Marked tests are significant at p < ,05000										
variable	Rank Sum Group 1	Rank Sum Group 2	U	Z	p-value	Z adjusted	p-value	Valid N Group 1	Valid N Group 2	2*1sided exact p
X	139,0000	71,00000	16,00000	2,532362	0,011330	2,574319	0,010044	10	10	0,008931

Рис. 8. Таблиця розрахунку U-тесту Манна – Уїтні

Висновок: проведений розрахунок параметричних та непараметричних критеріїв, свідчить про наявність розбіжності між середніми двох вибірок та форми їх розподілу. Отже, можна зробити висновок, що використання нової методики навчання дає можливість підвищити його якість.

Завдання для самостійного опрацювання

1. За результатами одноразового опитування мешканців міста про доцільність відкриття курсів "бізнес-леді" таку ідею підтримують 80 жінок із 125 та 55 чоловіків із 100. Перевірте гіпотезу про те, що розбіжності у ставленні чоловіків та жінок до ідеї відкриття курсів "бізнес-леді" мають

випадковий характер. Сформулюйте альтернативну гіпотезу і за допомогою t-критерію Стюдента зробіть висновок з імовірністю 0,95.

2. З метою перевірки ефективності нового пристрою для вдосконалення технології оброблення деталей робітників бригади розділили на дві групи: на верстатах з новим пристроєм працювало 4 робітники, за традиційною технологією оброблення деталей – 6. Результати експерименту такі: у I групі середній виробіток за годину становить 84 деталі при $\sigma^2 = 10$, у II – 77 деталей при $\sigma^2 = 8$.

Сформулюйте нульову та альтернативну гіпотези. За допомогою t-критерію Стюдента зробіть висновок, чи дійсно нова технологія впливає на продуктивність праці.

3. Під час опитування населення одне з питань передбачало оцінювання орієнтації респондентів на індивідуальну незалежність від держави. Відповіді розподілилися так: із 40 респондентів віком до 30 років 28 вважають, що держава має створити лише умови для плідної роботи, а людина повинна сама забезпечити свою сім'ю; із 60 респондентів старшого віку (понад 30 років) таку думку поділяють 24.

Сформулюйте нульову та альтернативну гіпотези і за допомогою t-критерію Стюдента перевірте, випадковий чи не випадковий характер мають розбіжності думок респондентів молодшого і старшого віку. Висновок зробіть з імовірністю 0,95.

4. Компанія, яка виробляє засоби для втрати ваги, стверджує, що приймання таблеток у сукупності зі спеціальною дієтою дозволить збавити в середньому 400 г ваги за тиждень. Випадково відібрані 25 осіб, які використовували цю терапію, і виявлено, що в середньому тижнева втрата ваги склала 430 г із середнім квадратичним відхиленням 110 г. Перевірте гіпотезу стосовно того, що втрата ваги складає 400 г. Рівень значущості $\alpha = 0,05$.

5. Компанія, яка випускає у продаж новий сорт розчинної кави, провела перевірку смаків покупців методом випадкової вибірки (400) осіб і з'ясувала, що 220 із них надали перевагу новому сорту. Перевірте з рівнем значущості $\alpha = 0,01$ гіпотезу стосовно того, що 52 % споживачів нададуть перевагу новому сорту кави.

6. Виробники нового типу аспірину стверджують, що він знімає головну біль за 30 хвилин. Випадкова вибірка 100 осіб, які страждають головними болями, показала, що новий тип аспірину знімає головну біль

за 28,6 хвилин за середнього квадратичного відхилення 4,2 хвилини. Перевірте з рівнем значущості $\alpha = 0,05$ ствердження виробників аспірину стосовно того, що ці ліки виліковують головний біль за 30 хвилин.

7. Виробник деякого виду продукції стверджує, що 95 % продукції, яка випускається, не має дефектів. Випадкова вибірка (100 виробів) показала, що тільки 92 з них вільні від дефектів. Перевірте ствердження виробника на рівні значущості $\alpha = 0,05$.

Тема 5. Метод аналітичних групувань

Методичні рекомендації

Лабораторна робота 2. Отримання вмінь щодо побудови аналітичних групувань та проведення аналізу взаємозв'язків між факторами на підставі використання правила додавання дисперсій за допомогою MS Excel

Завдання. За даними про відсоток порушень технологічної дисципліни та втратами від браку продукції побудувати аналітичне групування виробничих ділянок й дослідити залежність втрат від браку продукції на цих ділянках від проценту порушень технологічної дисципліни на підставі кореляційного відношення та коефіцієнту детермінації, використовуючи засоби MS Excel.

1. Для побудови інтервального варіаційного ряду необхідно в діалоговому вікні інструменту "Гістограма" задати верхні межі інтервалів. Оскільки вони невідомі, для їх знаходження можна скористатися режимом автоматичного розрахунку меж інтервалів. Однак в цьому режимі в кишенях генеруємої таблиці видаються нижні межі інтервалів. Для отримання на їх основі верхніх меж необхідно виконати дві дії:

а) виключити з кишень згенерованої таблиці нижню межу першого інтервалу x_{\min} ;

б) додати кишеню з верхньою межею x_{\max} останнього інтервалу.

Дане перетворення наведено на рис. 9.

	A	B	C	D	E	F	G	H	I
	Номер виробничої ділянки	Відсоток порушень технологічної дисципліни, %	Втрати від браку продукції, тис. грн.						
1					<i>Карман</i>	<i>Частота</i>		<i>Карман</i>	<i>Частота</i>
2	1	2,22	1,7		1,2	1			1
3	2	1,57	1,56		1,575	4		1,575	4
4	3	2,58	1,92		1,95	7		1,95	7
5	4	1,2	1,23		2,325	3		2,325	3
6	5	2,45	2,1		<u>Еще</u>	<u>7</u>		<u>2,7</u>	<u>7</u>
7	6	1,79	1,76						
8	7	1,58	1,3						
9	8	1,67	1,48						
10	9	1,75	1,54						
11	10	1,95	1,8						
12	11	1,31	1,27						
13	12	2,7	1,98						
14	13	1,23	1,26						
15	14	2,12	1,67						
16	15	2,4	2,3						
17	16	1,96	1,84						
18	17	2,33	1,9						
19	18	1,92	1,79						
20	19	2,34	1,7						
21	20	1,76	1,47						
22	21	2,55	2,45						
23	22	1,34	1,45						
24									

Рис. 9. Схема переходу від нижніх меж інтервалів до верхніх

2. Для побудови інтервального ряду застосовується режим формування ряду по заданим верхнім межах інтервалів. Для цього необхідно знову звернутися до інструменту "Гістограма", задавши в діалоговому вікні наступні параметри: Вхідний інтервал, Інтервал кишень, Вихідний інтервал, Інтегральний відсоток, Висновок графіка.

Вихідна таблиця містить: межі інтервалів – у графі з ім'ям "Карман" та частоти інтервалів – у другій графі. Для переходу від цієї таблиці до результативної таблиці необхідно провести вручну наступні перетворення: замінити назви стовпців; видалити рядок "Ще"; межі інтервалів привести до формату "нижня межа – верхня межа"; додати і заповнити підсумковий рядок (рис. 10).

Групи ділянок за порушенням технологічної дисципліни	Кількість ділянок
1,2-1,575	5
1,575-1,95	7
1,95-2,325	3
1,235-2,7	7
Усього	22

Рис. 10. Результативна таблиця групування

3. Для побудови аналітичного групування необхідно розрахувати середні значення втрат від браку продукції за кожною групою (рис. 11).

G2		fx =СУММ((C3+C5+C12+C14+C22)/5)				
A	B	C	D	E	F	G
Номер виробничої ділянки	Відсоток порушень технологічної дисципліни, %	Втрати від браку продукції, тис. грн.		Групи ділянок за порушенням технологічної дисципліни	Кількість ділянок	Середнє значення втрат продукції, тис. грн.
1						
2	1	2,22	1,7	1,2-1,575	5	1,354
3	2	1,57	1,56	1,575-1,95	7	1,59142857
4	3	2,58	1,92	1,95-2,325	2	1,77
5	4	1,2	1,23	2,325-2,7	7	2,05
6	5	2,45	2,1	Усього	21	1,89135714
7	6	1,79	1,76			
8	7	1,58	1,3			
9	8	1,67	1,48			
10	9	1,75	1,54			
11	10	1,95	1,8			
12	11	1,31	1,27			
13	12	2,7	1,98			
14	13	1,23	1,26			
15	14	2,4	2,3			
16	15	1,96	1,84			
17	16	2,33	1,9			
18	17	1,92	1,79			
19	18	2,34	1,7			
20	19	1,76	1,47			
21	20	2,55	2,45			
22	21	1,34	1,45			

Рис. 11. Аналітичне групування

4. Далі проводиться розрахунок загальної та міжгрупової дисперсії, коефіцієнта детермінації та кореляційного відношення (рис. 12).

Для розрахунку загальної дисперсії необхідно побудувати групування за результативною ознакою – втратами від браку продукції.

$=((1,23+1,535)/2)^2 \cdot 19$

D	E	F	G	H	I	J
	Групи ділянок за порушенням технологічної дисципліни	Кількість ділянок	Середнє значення втрат продукції, тис. грн.	$y-уср$	$(y-уср)^2 \cdot f$	$\bar{y}^2 \cdot f$
	1,2-1,575	5	1,354	-0,337357143	0,569049209	13,38
	1,575-1,95	7	1,59142857	-0,099928571	0,069900036	22,78
	1,95-2,325	2	1,77	0,078642857	0,012369398	15,88
	2,325-2,7	7	2,05	0,358642857	0,900372893	9,202
	Усього	21	1,89135714		1,551691536	61,24
				Групи за втратами від браку	Кількість ділянок	
	Карман	Частота				
	1,23	1		1,23-1,535	7	
	1,535	6		1,535-1,84	8	
	1,84	8		1,84-2,145	4	
	2,145	4		2,145-2,45	2	
	Еще	2		Усього	21	
	Міжгрупова дисперсія	0,07389007		Коефіцієнт детермінації	0,8591839	
	Загальна дисперсія	0,0860003		Кореляційне відношення	0,926921733	

Рис. 12. Розрахунок показників зв'язку між факторною та результативною ознаками

Висновок. Аналітичне групування показало наявність прямого зв'язку між порушенням технологічної дисципліни та втратами від браку продукції: з підвищенням факторної ознаки середнє значення результативної ознаки також підвищується (див. рис. 11).

Коефіцієнт детермінації дорівнює 0,86, тобто варіація втрат від браку продукції на 86 % пояснюється варіацією процента порушень технологічної дисципліни і на 14 % – варіацією інших факторів. Значення кореляційного відношення вказує на те, що зв'язок між ознаками дуже сильний (див. рис. 12).

Завдання для самостійного опрацювання

1. Використовуючи коефіцієнт Фехнера, виявіть напрямок та характер зв'язку між прожитковим мінімумом та середньою заробітною платою населення по 10 регіонах (табл. 2).

Таблиця 2

Розподіл населення регіонів за рівнем прожиткового мінімуму та розміром заробітної плати

Номер регіону	Середня заробітна плата	Прожитковий мінімум на душу населення, грн / місяць
1	480	287
2	463	250
3	350	200
4	540	317
5	585	385
6	605	366
7	513	392
8	421	300
9	390	297
10	620	390

2. Експерти оцінили споживчі якості дев'яти партій рибних консервів. Одночасно проведено опитування споживачів щодо оцінки якості консервів.

За даними, наведеними в табл. 3, визначте коефіцієнт рангової кореляції між об'єктивними та суб'єктивними оцінками споживчих якостей консервів, перевірте істотність зв'язку з імовірністю 0,95, зробіть висновки.

Таблиця 3

Експертні оцінки споживних якостей товару

Партія консервів	Експертні оцінки споживних якостей (max = 100 балів)	Середній бал якості за результатами опитування споживачів
1	57,6	3,8
2	45,8	3,1
3	50,7	5,2
4	44,4	2,6
5	42,9	2,5
6	53,5	5,0
7	44,8	3,6
8	45,6	2,8
9	44,2	4,0

3. За допомогою коефіцієнта рангової кореляції оцініть ступінь узгодженості оцінок двох груп експертів на конкурсі професійної майстерності модельєрів (табл. 4). Висновок зробіть з імовірністю 0,95.

Таблиця 4

Оцінки експертів професійності модельєрів

Модельєри	Ранг, визначений експертами	
	художниками	промисловиками
1	5	4
2	1	3
3	6	5
4	3	2
5	2	1
6	4	6
7	7	7

4. За даними аудиторської перевірки 24 комерційних банків виявлено залежність між рівнем ліквідності активів та їх прибутковістю (табл. 5).

Таблиця 5

Результати аудиторської перевірки банків

Коефіцієнт ліквідності активів	Кількість банків	Середній рівень прибутковості активів, %
0,2 – 0,4	6	15
0,4 – 0,7	5	8
0,7 – 1,0	5	6
1,0 і більше	8	3
У цілому	24	7,7

Загальна дисперсія прибутковості активів дорівнює 31. Визначте відносні ефекти впливу, міжгрупову дисперсію та кореляційне відношення, поясніть його зміст. Перевірте істотність зв'язку з імовірністю 0,95.

5. За даними звітів сільськогосподарських підприємств рівень рентабельності виробництва залежить від ступеня забезпеченості ресурсами.

Визначте міжгрупову дисперсію та кореляційне відношення, якщо загальна дисперсія рентабельності виробництва становить 116. Зробіть висновки щодо щільності зв'язку та перевірте його на істотність з імовірністю 0,95 (табл. 6).

Таблиця 6

Дані звітів сільхозпідприємств

Коефіцієнт забезпеченості ресурсами	Кількість підприємств	Рівень рентабельності, %
До 0,9	31	10
0,9 – 1,1	425	16
1,1 і більший	24	35
У цілому	100	18,7

6. Маємо наступні дані про вартість основних засобів і випуск товарної продукції промисловими підприємствами за рік (табл. 7).

**Розподіл підприємств за випуском
продукції та вартістю основних засобів**

Номер підприємства	Вартість основних засобів, тис. грн	Товарна продукція, тис. грн	
		за планом	фактично
1	480	560	590
2	690	1 435	1 470
3	1 750	3 535	3 742
4	3 420	7 150	8 210
5	1 146	1 124	1 217
6	1 205	1 784	1 955
7	1 308	2 966	2 994
8	364	408	409
9	2 120	3 151	3 276
10	1 332	2 150	2 166
11	1 000	2 100	2 300
12	4 073	10 967	12 099
13	1 014	2 976	2 138
14	586	978	992
15	905	1 296	1 404
16	1 746	4 010	4 157
17	3 140	6 572	6 206
18	529	1 088	1 057
19	2 861	6 068	6 412
20	407	618	724
21	1 990	3 912	4 075

Зробіть аналітичне групування підприємств за вартістю основних засобів, виділивши п'ять груп. У кожній групі підрахуйте число підприємств, вартість основних засобів, товарну продукцію за планом і фактично, відсоток виконання плану, фондівіддачу за планом і фактично.

Результати розрахунку зведіть у таблицю, зробіть короткі висновки.

За результатами групування, використовуючи інформацію щодо груп підприємств за вартістю основних засобів, числа підприємств та середній обсяг фактичної товарної продукції, що припадає на кожну групу, розрахуйте тісноту зв'язку між факторною та результативною ознаками на підставі правила додавання дисперсій, розрахувавши емпіричний коефіцієнт детермінації. Прокоментуйте проведення розрахунків та отримані результати.

Тема 6. Регресійний аналіз взаємозв'язку

Методичні рекомендації

Лабораторна робота 3. Поглиблення навичок побудови однофакторних та багатфакторних кореляційно-регресійних моделей за допомогою ППП Statistica 10.0

Завдання 1. За даними щодо оптових цін за одну пляшку на марочні вина (y) залежно від року закладки вина (x) (табл. 8), побудувати однофакторну регресійну модель та оцінити її параметри.

Таблиця 8

Оптові ціни за одну пляшку на марочні вина (y) в залежності від року закладки вина (x)

Рік (x)	Ціна (y. о.)
1890	50,00
1900	35,00
1920	25,00
1931	11,98
1934	15,00
1935	13,00
1940	6,98
1941	10,00
1944	5,99
1948	8,98
1950	6,98
1952	4,99
1955	5,98
1960	4,98

Регресійний аналіз застосовується для визначення виду зв'язку і дає можливість для прогнозування значення однієї (залежної) змінної відштовхуючись від значення іншої (незалежної) змінної.

Рівняння $y = \beta_1 x + \beta_0 + \varepsilon$ називають рівнянням простої лінійної регресії, воно має такі складові:

x – незалежна змінна;

y – залежна змінна;

β_1, β_0 – невідомі параметри, що описують пряму, яку потрібно оцінити за результатами спостережень $(x_i, y_i), i = 1, n$;

ε – незалежна випадкова помилка з нульовим середнім, яка інтерпретується як помилка спостережень.

За даними табл. 8 потрібно побудувати рівняння простої регресії за результатами спостереження (x_i, y_i) . У цьому випадку одержати точні значення параметрів β_1, β_0 неможливо, оскільки потрібна нескінченна вибірка, тому отримують їх оцінки b_1, b_0 , і рівняння регресії записується у вигляді $y = b_1x + b_0$.

Після запуску системи необхідно вибрати модуль Multiple Regression – множинна регресія, після запуску якого з'являється "Стартова панель" модуля, що містить основні операції, доступні в модулі і різні параметри аналізу.

На першому кроці необхідно створити електронну таблицю з вихідними даними, в якості змінних в розглянутому прикладі виступають "Рік" закладки вина і його "Ціна" на аукціоні.

Для проведення статистичного аналізу спочатку викликається "Стартова панель" модуля, в якому задаються різні параметри аналізу. Вибирається залежна змінна "Ціна" і незалежна змінна "Рік". Для завдання змінних використовують кнопку Variables – Змінні, під час натискання на яку відкривається вікно Select dependent and independent variable list – вибір залежної змінної і списку незалежних змінних, в якому вибираються відповідні змінні (ціна, рік) (рис. 13).

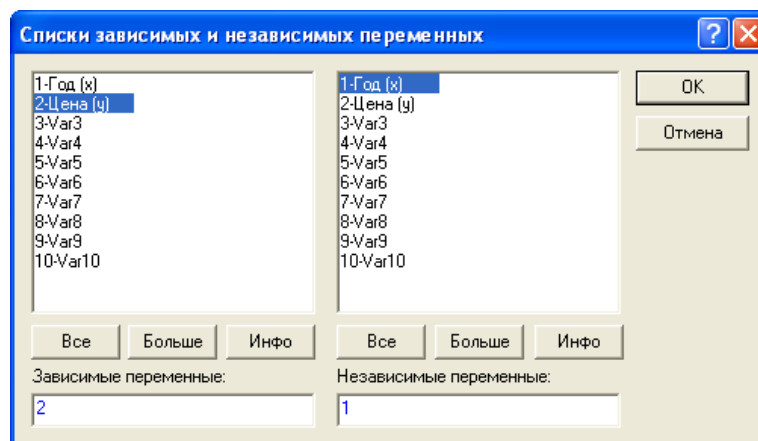


Рис. 13. Вибір залежної змінної і списку незалежних змінних

Натисканням на кнопку ОК повертаємося в "Стартову панель" модуля (рис. 14).

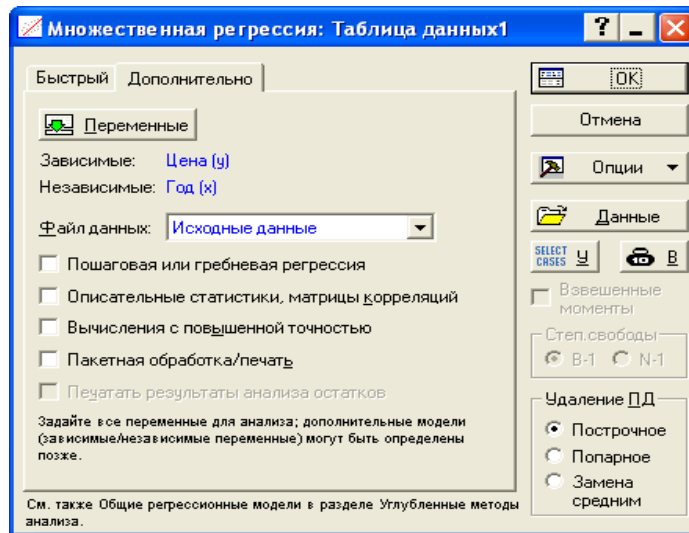


Рис. 14. Стартова панель модуля "Множинна регресія"

Для виведення результатів і їх аналізу на стартовій панелі необхідно натиснути на кнопку ОК. Система робить обчислення, і через секунду вікно результатів з'явиться на екрані. Вікно результатів аналізу має наступну просту структуру: верхня частина вікна – інформаційна, нижня містить функціональні кнопки, що дозволяють всебічно ознайомитися з результатами аналізу (рис. 15).

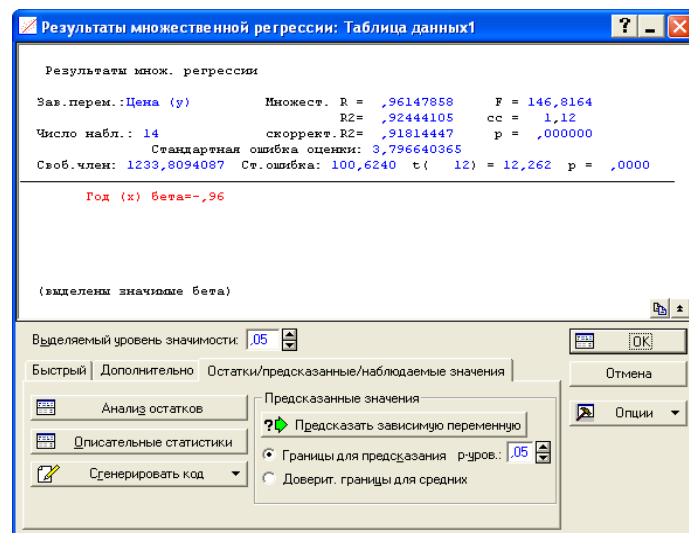


Рис. 15. Вікно результатів аналізу множинної регресії

В інформаційній частині вікна міститься коротка інформація про проведений аналіз, а саме:

Dep. Var. – ім'я залежної змінної: Вартість;

No. Of Cases – число випадків, за якими побудована регресія: 14;

Multiple R – коефіцієнт множинної кореляції;

R – square – квадрат коефіцієнта множинної кореляції;

Adjusted R – square – скоригований коефіцієнт кореляції, який визначається як $\text{Adjusted R – square} = 1 - (1 - R - \text{Square}) \times (n / (n - p))$, де n – число спостережень в моделі; p – число параметрів моделі;

Std. Error of estimate – стандартна помилка оцінки. Ця статистика є мірою розсіювання спостережуваних значень щодо прямої регресії;

Intercept – оцінка вільного члена регресії. Значення коефіцієнта b_0 в рівнянні регресії;

Std. Error – стандартна помилка оцінки вільного члена. Стандартна помилка коефіцієнта b_0 в рівнянні регресії;

$t(df)$ and p -value – значення критерію Стьюдента та рівня помилки. Критерій Стьюдента використовується для перевірки гіпотези про рівність нулю вільного члена регресії.

F – значення критерію Фішера.

df – число ступенів свободи F-критерію.

p – рівень значущості.

У розглянутому прикладі значення коефіцієнта детермінації $R^2 = 0,924$, що означає, що побудована регресія пояснює 92,4 % розкиду значень ціни щодо середнього значення.

Для перевірки значимості регресії використовують F-критерій. У такому випадку цей критерій використовується для перевірки гіпотези, яка стверджує, що між залежною змінною і незалежною змінною немає лінійної залежності, тобто $b_1 = 0$ проти альтернативи $b_1 \neq 0$. Велике значення F-критерію: $F = 146,82$ і рівень значимості $p = 0,0000$ показує, що побудована регресія високо значима.

При натисканні на кнопку Regression summary – Короткі результати регресії на екрані з'являється електронна таблиця з результатами аналізу (рис. 16).

Итоги регрессии для зависимой переменной: Цена (Таблица 1)						
R= ,96147858 R2= ,92444105 Скорректир. R2= ,91814447						
F(1,12)=146,82 p<,000000 Станд. ошибка оценки: 3,7966						
N=14	БЕТА	Стд. Ош. БЕТА	В	Стд. Ош. В	t(12)	p-уров.
Св. член			1233,809	100,6240	12,2616	0,000000
Рік	-0,961479	0,079351	-0,630	0,0520	-12,1168	0,000000

Рис. 16. Короткі результати регресії

У третьому стовпці цієї таблиці розташовані оцінки невідомих параметрів моделі:

$$b_0 = 1233,809;$$

$$b_1 = -0,630, \text{ тобто модель має вигляд:}$$

$$\text{ЦІНА} = -0,63 \times \text{ВІК} + 1233,809.$$

Параметри моделі визначають вид зв'язку. У цьому випадку зв'язок зворотний, тобто зі зростанням року виготовлення падає ціна на вино. Такий висновок підтверджується параметром $b_1 = -0,630$, а саме: знаком "-" перед значенням параметру.

Завдання 2. Побудувати багатофакторну регресійну модель за даними, наведеними в табл. 9.

Таблиця 9

Вихідні дані для побудови багатофакторної економетричної моделі

Рік	Випуск продукції тис. грн (X1)	Обсяг роздрібного товарообороту підприємств (юридичних осіб), млн грн (X2)	Сукупні витрати в середньому за місяць у розрахунку на одне домогосподарство, грн (X3)	Прямі інвестиції (млн дол.) (X4)	ВВП (млрд грн) (Y)
2002	226 358	19 317	395,6	2 063,6	186,5
2003	356 842	22 151	426,5	2 810,7	192,5
2004	373 893	28 757	541,3	3 281,8	198,9
2005	460 520	34 417	607,0	3 875,0	221,6
2006	504 008	39 691	658,3	4 555,3	225,8
2007	603 704	49 994	736,8	5 471,8	267,3
2008	809 988	67 556	903,5	6 794,4	345,1
2009	995 630	94 332	1 229,4	9 047,0	441,5
2010	1 182 179	129 952	1 442,8	16 890,0	544,2
2011	1 565 055	178 233	1 722,0	21 607,3	720,7
2012	2 072 172	246 903	2 590,4	29 542,7	948,1
2013	1 955 685	230 955	2 754,1	35 616,4	913,3
2014	2 388 289	280 890	3 072,7	40 053,0	1 082,6
2015	2 496 365	350 059	3 456,0	44 806,0	1 316,6

Відповідно до алгоритму побудови однофакторної моделі, проводять розрахунки й для багатофакторної моделі (рис. 17).

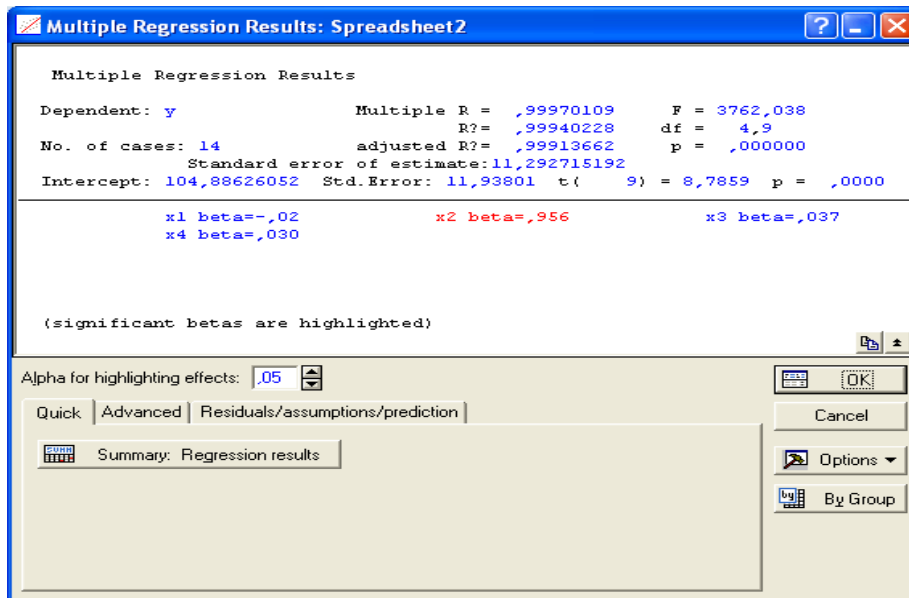


Рис. 17. Діалогове вікно модулю Multiple Regression

З метою визначення параметрів, та якості моделі необхідно ініціювати кнопку Summary: Regression results (Результати регресійного аналізу). Результати розрахунків наведено на рис. 18.

Regression Summary for Dependent Variable(Spreadsheet2)						
R= .99970109 R²= .99940228 Adjusted R²= .99913662						
F(4,9)=3762,0 p<,00000 Std.Error of estimate: 11,293						
N=14	Beta	Std.Err. of Beta	B	Std.Err. of B	t(9)	p-level
Intercept			104,886	11,9380	8,7859	0,00001
x1	-0,02260	0,07362	-0,000	0,0000	-0,3069	0,76583
x2	0,95560	0,08281	0,003	0,0002	11,5391	0,00000
x3	0,03676	0,11198	0,013	0,0401	0,3283	0,75018
x4	0,03026	0,08187	0,000	0,0020	0,3696	0,72020

Рис. 18. Результат побудови багатofакторної економетричної моделі

Отримані результати, які подані на рис. 18 можна інтерпретувати наступним чином:

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

коефіцієнт детермінації моделі дорівнює 0,999 (R²). Цей коефіцієнт показує, яка частка даних, побудованих за допомогою моделі відповідає

реальним даним. Так як коефіцієнт наближений до 1, то підтверджується адекватність моделі;

скоригований коефіцієнт детермінації на число спостережень і число параметрів дорівнює 0,999 (Adjusted R²);

критерій адекватності Фішера F (4,9) = 3 762 отримане значення більше табличного, що підтверджує адекватність моделі;

середнє квадратичне відхилення помилок моделі складає 11,93;

вектор параметрів моделі має наступний вигляд $V(a_0, a_1, a_2, a_3, a_4) = (104,88; -0,001; 0,033; 0,00132; 0,0008)$.

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

Отже, за даними, що подано на рис. 18, можна сформулювати загальний вид багатфакторної регресійної моделі:

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

Виходячи з аналізу отриманих результатів дана модель у цілому адекватна та якісна, але параметри моделі за змінними X₁, X₃, X₄ є незначимими.

Для визначення середнього значення та середньоквадратичного відхилення вибірок всіх змінних у меню аналізу помилок ініціюємо Descriptive statistics / Means & Standard deviations (описові статистики / середнє і середньоквадратичне відхилення) (рис. 19 – 20).

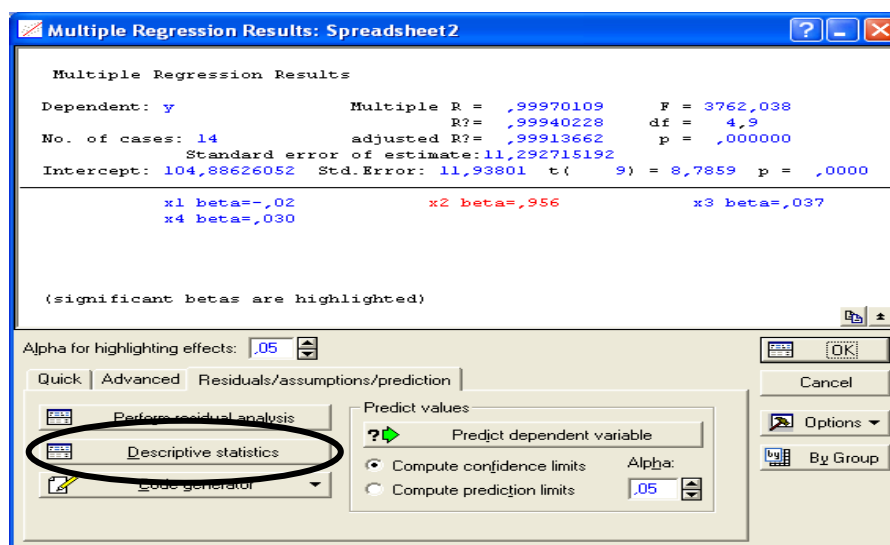


Рис. 19. Діалогове вікно аналізу результатів побудови багатфакторної моделі

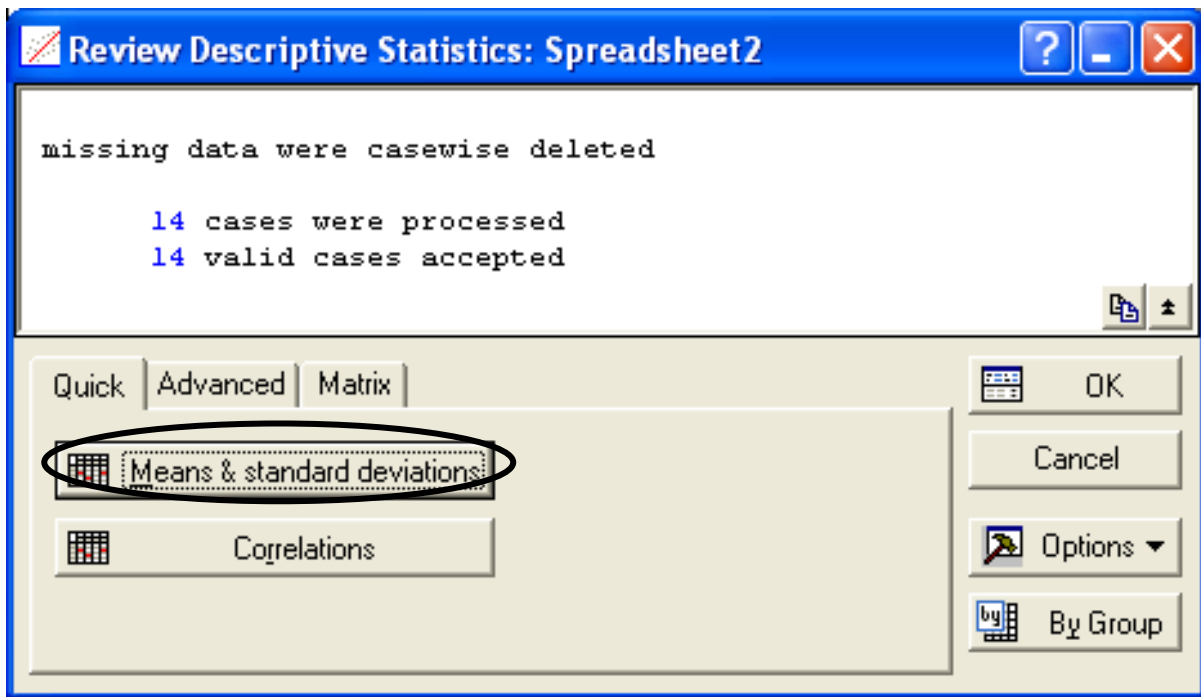


Рис. 20. **Діалогове вікно розрахунку описових статистик**

У результаті проведення розрахунків було оцінено середні значення, середнє квадратичне відхилення за екзогенними та ендogenous змінними (рис. 21).

Variable	Means and Standard Deviations (Spreadsheet2)		
	Means	Std.Dev.	N
x1	114219,8	805422,8	14
x2	126658,8	111015,8	14
x3	1467,8	1070,8	14
x4	1617,8	15342,8	14
y	543,8	384,8	14

Рис. 21. **Результати розрахунку описових статистик за ендogenousною та екзогенними змінними**

За багатофакторною моделлю аналіз помилок проводиться за тим же алгоритмом, що і для однофакторної моделі. На рис. 22 наведено результати розрахунку теоретичних значень за моделлю та помилок моделі.

Predicted & Residual Values (Spreadsheet2)									
Dependent variable:									
Case No.	Observed Value	Predicted Value	Residual	Standard Pred. v.	Standard Residual	Std.Err. Pred.Val	Mahalanobis Distance	Deleted Residual	Cook's Distance
1	186,500	173,134	13,366	-0,96316	1,18360	6,29854	3,11557	19,4017	0,18365
2	192,500	182,070	10,4230	-0,93989	0,92300	5,16400	1,78990	13,1800	0,05697
3	198,900	205,618	-6,7182	-0,87862	-0,59492	4,72080	1,34327	-8,1400	0,01816
4	221,600	224,720	-3,1250	-0,82889	-0,27672	4,13812	0,81700	-3,6097	0,00274
5	225,800	242,890	-17,096	-0,78159	-1,51392	4,01560	0,71527	-19,5700	0,07595
6	267,300	277,630	-10,3350	-0,69117	-0,91522	3,76240	0,51450	-11,6260	0,02353
7	345,100	336,712	8,3878	-0,53741	0,74272	5,04070	1,66160	10,4750	0,03428
8	441,500	429,290	12,2010	-0,29643	1,08040	8,01380	5,61810	24,5800	0,47717
9	544,200	553,880	-9,6860	0,02783	-0,85772	4,22380	0,89010	-11,2610	0,02782
10	720,700	716,740	3,9590	0,45170	0,35060	9,31100	7,90930	12,3670	0,16307
11	948,100	955,920	-7,8190	1,07422	-0,69240	7,22370	4,39080	-13,2350	0,11242
12	913,300	911,180	2,1180	0,95778	0,18760	9,15260	7,61100	6,1750	0,03929
13	1082,600	1079,270	3,3220	1,39529	0,29410	7,66350	5,05830	6,1580	0,02738
14	1316,600	1315,590	1,0010	2,01037	0,08860	10,61810	10,56470	8,6390	0,10350
Minimum	186,500	173,134	-17,096	-0,96316	-1,51392	3,76240	0,51450	-19,5700	0,00274
Maximum	1316,600	1315,590	13,366	2,01037	1,18360	10,61810	10,56470	24,5800	0,47717
Mean	543,190	543,190	-0,0000	-0,00000	-0,00000	6,38190	3,71420	2,3950	0,09614
Median	393,300	383,000	1,5600	-0,41692	0,13810	5,73130	2,45270	6,1660	0,04813

Рис. 22. Результати розрахунку помилок моделі

На рис. 22 видно, що найбільше значення помилки моделі спостерігається у 2006 році. Тобто можна зробити висновок, що в цей період розвиток економіки країни істотно відрізнявся від усього аналізованого періоду.

Для аналізу помилок моделі доцільно побудувати полігон та гістограму розподілу помилок. На рис. 23 наведено полігон розподілу помилок моделі.

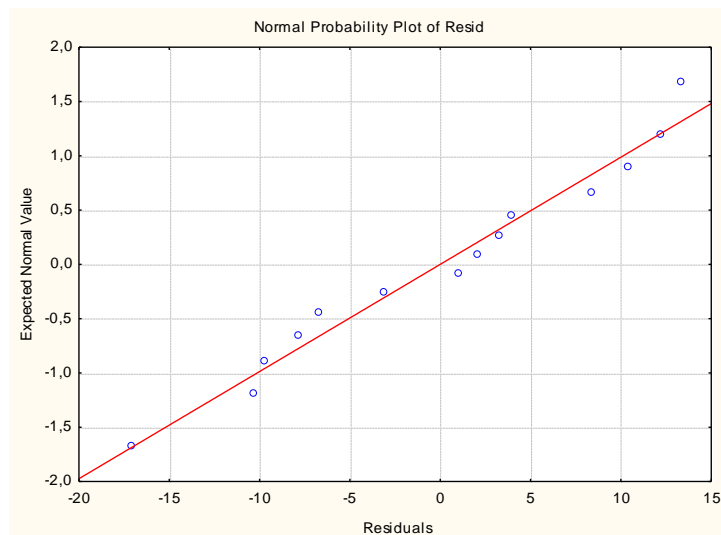


Рис. 23. Полігон розподілу помилок

На рис. 24 наведено гістограму розподілу помилок моделі.

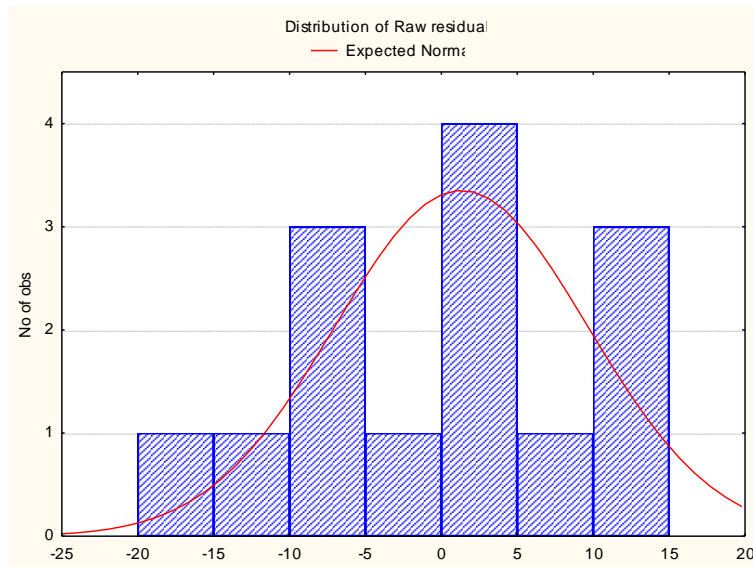


Рис. 24. Гістограма розподілу помилок

Дані, наведені на рис. 23 та рис. 24 доводять гіпотезу про нормальний закон розподілу помилок моделі.

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

Завдання для самостійного опрацювання

1. За десятьма однорідними підприємствами за звітний рік маємо наступні дані про середню виробітку продукції та електроозброєність праці (табл. 10).

Таблиця 10

Розподіл підприємств за рівнем електроозброєності та продуктивності праці

Номер підприємства	Електроозброєність праці, тис. кВт / год на одну людину	Випуск продукції на одного працівника, тис. грн
1	2	3
1	5	6,25
2	6	7,54

1	2	3
3	7	8,75
4	4	6,00
5	3	3,75
6	5	6,24
7	4	5,90
8	6	7,80
9	7	8,78
10	3	3,60

Визначте лінійне рівняння зв'язку, що характеризує залежність між електроозброєністю праці працівника та випуском продукції на одного працівника. Для виявлення тісноти зв'язку розрахуйте лінійний коефіцієнт кореляції. Поясніть зміст розрахованих показників кореляції.

2. Маємо наступні дані по 10 заводах (табл. 11).

Таблиця 11

**Розподіл заводів за вартістю основних засобів
і випуском продукції**

Номер заводу	Вартість основних засобів, млн грн, X	Випуск продукції, млн грн, Y
1	12	6,2
2	13	6,8
3	10	4,8
4	12	5,6
5	8	4,2
6	10	4,1
7	6	2,4
8	9	3,6
9	15	5,0
10	11	4,6

Для вивчення зв'язку між вартістю основних засобів і випуском продукції розрахуйте лінійне рівняння зв'язку та лінійний коефіцієнт кореляції. Поясніть зміст розрахованих показників.

Тема 7. Методи оцінювання стохастичного зв'язку

Методичні рекомендації

Лабораторне заняття

Стохастичний зв'язок – це зв'язок між величинами, при якому одна з них, випадкова величина y , реагує на зміну іншої величини x або інших величин x_1, x_2, \dots, x_n (випадкових або не випадкових) зміною закону розподілу. Це обумовлено тим, що залежна змінна (результативна ознака), окрім змінних, що розглядаються, знаходиться під впливом інших неврахованих або неконтрольованих (випадкових) факторів. Оскільки значення залежної змінної підпорядковані випадковому розсіянню, вони не можуть бути передбачені з достатньою точністю, а тільки вказані з певною імовірністю.

Характерною особливістю стохастичних зв'язків є те, що вони виявляються у всій сукупності, а не в кожній її одиниці, причому невідомий ані повний перелік факторів, що визначають результативну ознаку, ані точний механізм їх функціонування та взаємодії з результативною ознакою. Отже, завжди має місце вплив випадковості.

Виявлення стохастичних зв'язків знаходиться під дією закону великих чисел: тільки в достатньо великій кількості одиниць індивідуальні особливості згладжуються і залежність, якщо вона має місце, виявиться достатньо чітко.

При стохастичному зв'язку різним значенням однієї змінної (фактора X) відповідають різні розподіли іншої змінної (результату Y). Кореляційний зв'язок – це частковий випадок стохастичного зв'язку, за якого різним значенням фактора X відповідають різні середні значення результату Y .

Наявність стохастичного зв'язку, який проявляється зміною умовних розподілів, можна виявити на основі комбінаційного розподілу елементів сукупності.

У статистиці широко використовують непараметричні методи визначення взаємозв'язків, які ґрунтуються на кількісних значеннях ознак, і не потребують обчислення параметрів їх розподілів. Якщо в кореляційно-регресійному аналізі всі ознаки – ознаки метричної шкали, а в методі аналітичного групування це стосується результативної ознаки, то непараметричні методи застосовуються і тоді, коли є ознаки порядкової чи номінальної шкали.

Використання непараметричних методів оцінювання тісноти стохастичного зв'язку ґрунтується на порівнянні частот або часток умовних розподілів у таблицях взаємної спряженості.

Таблицею взаємної спряженості називають таблицю, яка містить зведену числову характеристику вивчаємої сукупності за двома та більше атрибутивними ознаками або комбінацією кількісних і атрибутивних ознак. Макет таблиці спряженості розмірністю $i \times j$, де $i = 1, 2, \dots, k$ – число варіантів значень однієї ознаки (A); $j = 1, 2, \dots, n$ – число варіантів значень другої ознаки (B) наведено на рис. 25.

B A	B ₁	B ₂	...	B _j	Разом
A ₁	f ₁₁	f ₁₂	...	f _{1j}	f ₁₀
A ₂	f ₂₁	f ₂₂	...	f _{2j}	f ₂₀
...
A _i	f _{i1}	f _{i2}	...	f _{ij}	f _{i0}
Разом	f ₀₁	f ₀₂	...	f _{0j}	f ₀₀

Рис. 25. Загальна схема таблиці спряженості

Для визначення тісноти зв'язку двох якісних ознак, кожна із яких складається тільки із двох груп, застосовуються коефіцієнти асоціації та контингенції:

$$K_a = \frac{f_{11} \times f_{22} - f_{12} \times f_{21}}{f_{11} \times f_{22} + f_{12} \times f_{21}}, \quad (1)$$

$$K_k = \frac{f_{11} \times f_{22} - f_{12} \times f_{21}}{\sqrt{(f_{11} + f_{12}) \times (f_{12} + f_{22}) \times (f_{11} + f_{21}) \times (f_{21} + f_{22})}}. \quad (2)$$

Значення коефіцієнта контингенції завжди менше значення коефіцієнта асоціації. Зв'язок між ознаками вважається підтвердженим, якщо $K_a \geq 0,5$ або $K_k \geq 0,3$.

Корисною мірою під час аналізу 4-клітинкових таблиць взаємної спряженості є відношення перехресних добутоків або відношення шансів:

$$W = \frac{f_{11}f_{22}}{f_{12}f_{21}}. \quad (3)$$

Відношення шансів характеризує міру відносного ризику.

Коли кожна із якісних ознак складається більше, ніж із двох груп, то для визначення тісноти зв'язку застосовують коефіцієнт взаємної спряженості Пірсона, коефіцієнт взаємної спряженості Чупрова та модифікація коефіцієнта спряженості Чупрова – коефіцієнт Крамера:

1. Коефіцієнт взаємної спряженості Пірсона:

$$K_p = \sqrt{\frac{\chi^2}{n + \chi^2}}, \quad (4)$$

де χ^2 – коефіцієнт квадратичної спряженості хі-квадрат Пірсона, який характеризує розбіжність між частотами (частками) умовного та безумовного розподілу і розраховується за формулою:

$$\chi^2 = n \left[\sum_i \sum_j \frac{f_{ij}^2}{f_{i0} \times f_{0j}} - 1 \right], \quad (5)$$

де n – кількість спостережень.

Фактичні значення хі-квадрат Пірсона порівнюються з критичними. Якщо фактичне значення перевищує критичне, то з імовірністю 0,95 істотність зв'язку між ознаками доведена;

2. Коефіцієнт взаємної спряженості Чупрова:

$$K_c = \sqrt{\frac{\chi^2}{n \sqrt{(m_x - 1) \times (m_y - 1)}}}, \quad (6)$$

де m_x – число груп за ознакою x ;

m_y – число груп за ознакою y .

Оскільки у разі незалежності ознак $\chi^2 = 0$, то $K_c = 0$. За функціонального зв'язку $K_c = 1$ за умови, що $m_x = m_y$.

3. У разі, коли $m_x \neq m_y$, зручніше користуватися формулою Крамера:

$$K_k = \sqrt{\frac{\chi^2}{n(m_{\min} - 1)}}, \quad (7)$$

де m_{\min} – мінімальне число груп (m_x або m_y).

Якщо $m_x = m_y$, то значення коефіцієнтів, обчислені за формулами Чупрова та Крамера збігаються.

Значення коефіцієнтів Чупрова, Крамера і Пірсона коливаються в межах від 0 до 1. Коефіцієнт Чупрова враховує кількість виділених груп за кожною ознакою і дає найбільш обережну оцінку зв'язку. У разі значення $K_c \geq 0,3$, можна говорити про помірний або щільний зв'язок між ознаками.

Перевірка істотності зв'язку здійснюється на основі критерію хі-квадрат Пірсона.

Методи аналізу таблиць взаємної спряженості можна використати і для кількісних ознак.

В аналізі соціально-економічних явищ та процесів використовують умовні оцінки, наприклад, ранги. Ранжування – це процедура впорядкування об'єктів вивчення, яка виконується на основі переваги. Ранг – це порядковий номер значень ознаки, розташованих в порядку зростання або спадання їх величин.

Якщо значення ознаки мають однакову кількісну оцінку, то ранг усіх цих значень приймається рівним середній арифметичній від відповідних номерів місць, які визначають. Такі ранги називаються зв'язними.

Під час вимірювання зв'язку між ознаками порядкової шкали використовують коефіцієнт кореляції рангів Спірмена. Його розрахунок ґрунтується на різниці рангів $d = R_x - R_y$, де R_x, R_y – ранги елементів сукупності відповідно за першою і другою ознаками. Даний коефіцієнт обчислюють за формулою:

$$\rho = 1 - \frac{6 \sum d^2}{n(n^2 - 1)}, \quad (8)$$

де n – число елементів сукупності.

Коефіцієнт кореляції рангів, як і лінійний коефіцієнт кореляції, може набувати значень від -1 до $+1$. Якщо два рядки рангів повністю збігаються, то $\sum d^2 = 0$. Отже, має місце повний прямий зв'язок і $\rho = 1$. У разі повного зворотного зв'язку (ранги двох рядків розташовані у зворотному порядку) $\rho = -1$.

Істотність зв'язку можна перевірити порівнявши критичне значення коефіцієнта рангової кореляції для певного рівня істотності α та числа

спостережень n з фактичним значенням. Якщо фактичне r більше від критичного, то наявність і істотність зв'язку між ознаками доводиться з імовірністю 95 %.

Слід зазначити, що для зворотних зв'язків з критичним значенням коефіцієнта рангової кореляції порівнюється абсолютна величина фактичного значення r .

На цьому лабораторному занятті студенти повинні виконати практичні завдання на придбання навичок аналізу стохастичних зв'язків між явищами та процесами, побудови та аналізу таблиць взаємної спряженості:

1. Розподіл зареєстрованих шлюбних пар за сімейним станом нареченого та нареченої характеризується даними табл. 12.

Визначте коефіцієнти контингенції і з імовірністю 0,95, зробіть висновок про поширеність однорідних за сімейним станом шлюбів (табл. 12).

Таблиця 12

Розподіл зареєстрованих шлюбних пар

Наречений	Наречена		Разом
	не перебувала у шлюбі	перебувала у шлюбі (розлучена, вдова)	
Не перебував у шлюбі	80	10	90
Перебував у шлюбі (розлучений, вдівець)	12	18	30
Разом	92	28	120

2. Розподіл пацієнтів клініки за результатами легеневих проб характеризується даними (табл. 13).

Таблиця 13

Розподіл пацієнтів клініки за результатами легеневих проб

Наявність звички паління	Результати легеневих проб		Разом
	нормальні	аномальні	
Не палить	40	20	60
Палить	10	30	40
Разом	50	50	100

За допомогою відношення шансів оцініть відносний ризик куріння цигарок. Визначте щільність зв'язку між курінням цигарок та захворюваннями легенів, перевірте істотність зв'язку з імовірністю 0,95.

3. Маємо дані про результати опитування молодих сімей щодо допомоги їм від батьків (табл. 14).

Таблиця 14

Дані опитування молодих сімей

Молода сім'я проживає	Кількість молодих сімей		Разом
	отримують допомогу від батьків	не отримують допомоги від батьків	
З батьками	60	10	70
Окремо	20	30	50
Разом	80	40	120

За допомогою коефіцієнта контингенції оцініть щільність зв'язку між ознаками, перевірте істотність зв'язку, висновки зробіть з імовірністю 0,95.

4. Маємо дані про розподіл молодих робітників за ступенем задоволеності умовами праці та професійною мобільністю.

Проаналізуйте комбінаційний розподіл робітників, оцініть щільність зв'язку між задоволеністю умовами праці та професійною мобільністю за допомогою коефіцієнта взаємної спряженості. Перевірте істотність зв'язку з імовірністю 0,95 та зробіть висновок (табл. 15).

Таблиця 15

Дані опитування молодих спеціалістів

Ступінь задоволеності умовами праці	Чи маєте намір змінити професію?			Разом
	Так, найближчим часом	Так, у перспективі	Ні	
Задоволений	–	20	26	46
Ставлюсь байдуже	7	18	9	34
Не задоволений	15	5	–	20
Разом	22	43	35	100

Завдання для самостійного опрацювання

1. За наведеними даними (табл. 16) про розподіл 100 домогосподарств за кількістю дітей та відносним рівнем доходу визначте коефіцієнт взаємної спряженості, поясніть його зміст. Зробіть висновок про наявність та істотність зв'язку з імовірністю 0,95.

Таблиця 16

Розподіл домогосподарств за показниками

Кількість дітей	Рівень доходу		Разом
	нижче прожиткового мінімуму	вище прожиткового мінімуму	
Немає дітей	1	6	7
1 дитина	10	38	48
2 дитини	15	20	35
3 і більше	6	4	10
Разом	32	68	100

2. Залежність звільнення робітників від місця роботи досліджувалось у процесі опитування 200 респондентів, результати цього опитування наведено в табл. 17.

Таблиця 17

Дані опитування респондентів

Думка респондента	Робітники		Разом
	державні підприємства	приватні підприємства	
Дуже імовірно	55	48	103
Практично виключено	45	52	97
Усього	100	100	200

Визначте коефіцієнти асоціації та контингенції. Зробіть висновки.

3. Розподіл основних категорій потенційних мігрантів за рівнем освіти характеризуються наступними даними (табл. 18).

Таблиця 18

Розподіл потенційних мігрантів за рівнем освіти

Освіта	Основні категорії потенційних мігрантів				Усього
	керівники	спеціалісти	службовці	робітники	
вища	55	48	12	7	122
неповна вища	5	3	3	5	16
середня спеціальна	36	44	51	39	170
середня загальна	4	4	33	39	80
неповна середня	0	1	1	10	12
Усього	100	100	100	100	400

Розрахуйте усі можливі модифікації коефіцієнтів взаємної спряженості. Зробіть висновки.

Змістовий модуль 2

Методи і моделі статистичного дослідження рядів динаміки та тенденцій розвитку

Тема 8. Аналіз тенденцій розвитку

Методичні рекомендації

Лабораторна робота 4. Придбання навичок проведення екстраполяції та інтерполяції в MS Excel

Завдання. Маємо дані про валовий внутрішній продукт (ВВП) за 2006 – 2018 рр. (рис. 26). Необхідно визначити за допомогою аналітичного вирівнювання прогнозне значення обсягу прибутку у 2019 р.

	А	В
1	Роки	ВВП, млн грн
2	2007	751 106
3	2008	990 819
4	2009	947 042
5	2010	1 120 585
6	2011	1 349 178
7	2012	1 459 096
8	2013	1 522 657
9	2014	1 586 915
10	2015	1 988 544
11	2016	2 385 367
12	2017	2 983 882
13	2018	3 558 706

Рис. 26. ВВП країни за роками

Екстраполяція та інтерполяція належать до статистичних методів прогнозування. Прогнозування – це визначення орієнтованих розмірів явища в майбутньому, поширення виявленої закономірності на інші періоди часу. Під час прогнозування явища вирішуються наступні проблеми:

визначення довжини базисного періоду, закономірність якого поширюють на майбутнє;

визначення довжини майбутнього періоду, на який поширюють виявлену закономірність.

Екстраполяція – це метод прогнозування, який передбачає, що закономірність розвитку, що діяла в минулому, збережеться і в прогнозованому майбутньому.

Якщо для вимірювання основної тенденції проводиться аналітичне вирівнювання ряду динаміки, то для екстраполяції використовується рівняння тренда, в яке підставляється значення t у майбутні періоди. Так як в базисному періоді рівень коливається близько тренду, то коливання ймовірні і в майбутньому. Отже, фактичний рівень в майбутньому не обов'язково точно збігається з екстрапольованим за трендом;

Інтерполяція – це визначення невідомого рівня всередині динамічного ряду. Під час інтерполяції припускають, що виявлена тенденція та її характер, істотно не змінювалися в тому проміжку часу, рівень якого невідомий. Існують наступні методи інтерполяції даних:

на основі середнього абсолютного приросту, середнього темпу зростання;

якщо, окрім кінцевого і базисного рівнів, відомі також деякі проміжні рівні, інтерполяції можна здійснити на основі рівняння тренду.

У Excel можна вибрати один з п'яти типів екстраполяції за допомогою діалогового вікна (рис. 27).

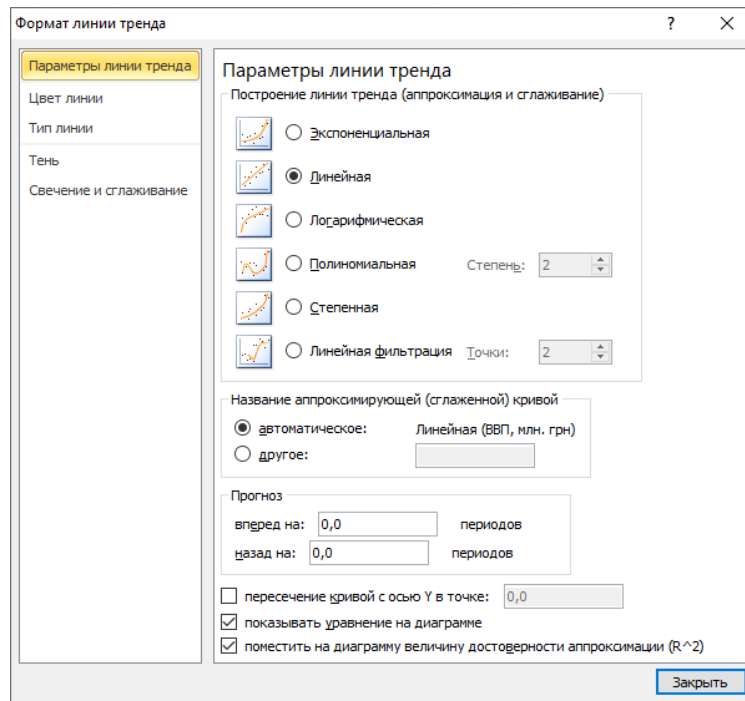


Рис. 27. Діалогове вікно: лінія тренда

Для побудови лінії тренду необхідно виділити часовий ряд та вибрати в контекстному меню команду "Добавить линию тренда". Після вибору команди відкриється діалогове вікно (рис. 28), в якому вибираємо тип, параметри побудови тренду.

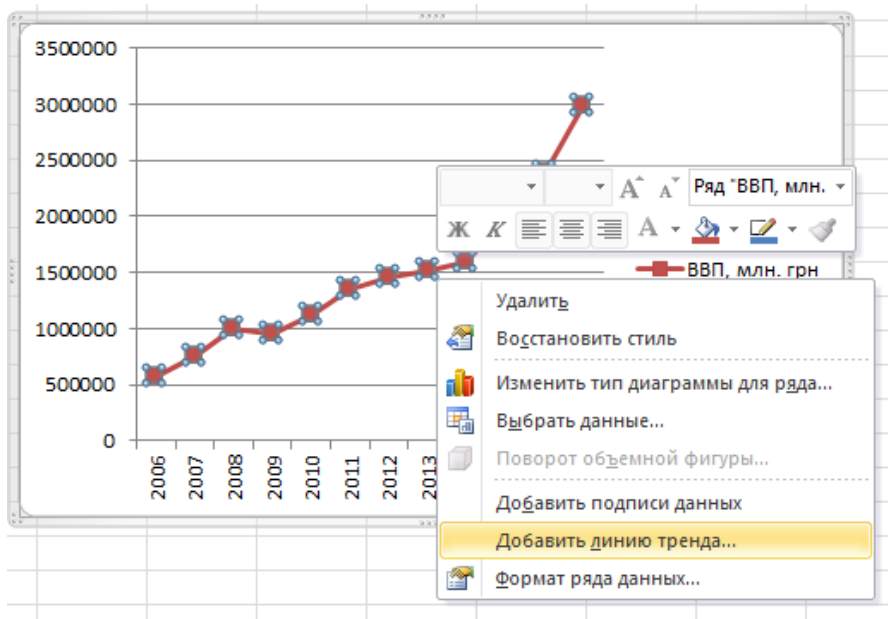


Рис. 28. Меню для побудови лінії тренду

Результат побудови лінійного тренду подано на рис. 29.

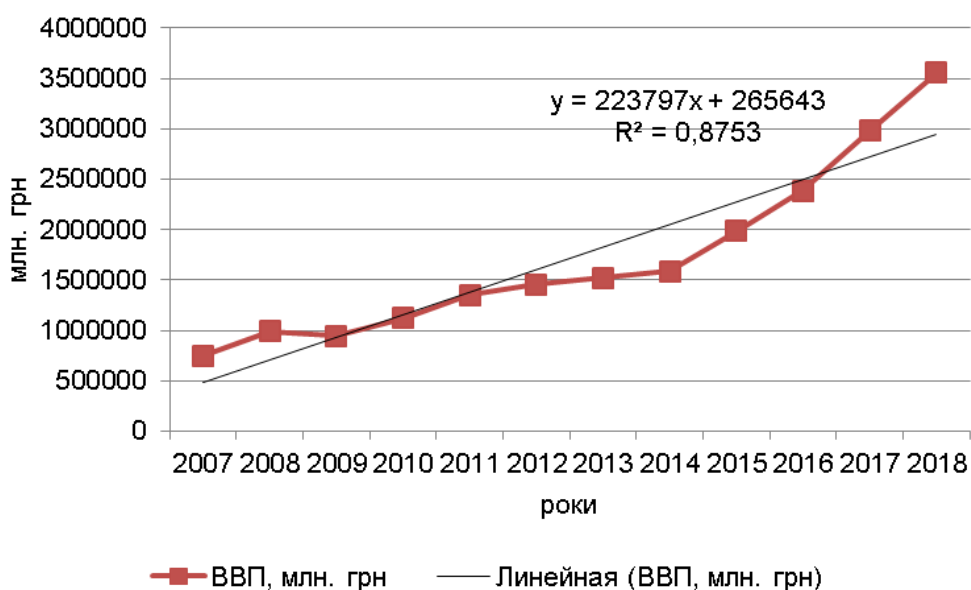


Рис. 29. Результати розрахунку параметрів тренду

Таким чином, рівняння прямої має вигляд:

$$Y = 265\,643 + 223\,797 \times t.$$

Розрахувати прогнозний обсяг ВВП на 2019 рік можна двома способами:

1. Зробити точковий прогноз, використовуючи рівняння прямої, де t – це умовний показник часу, на який буде будуватись прогноз. Для цього використовують стандартну функцію ПРЕДСКАЗ. Після вибору даної функції відкривається діалогове вікно (рис. 30), в якому задаємо параметри.

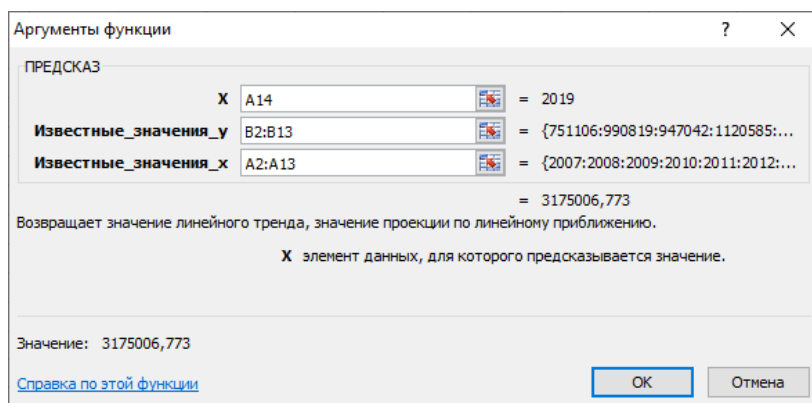


Рис. 30. Діалогове вікно: аргументи функції ПРЕДСКАЗ

Результат функції ПРЕДСКАЗ подано на рис. 31.

Буфер обмена		Шрифт		
В14		fx =ПРЕДСКАЗ(A14;B2:B13;A2:A13)		
	A	B	C	D
1	Роки	ВВП, млн. грн		
2	2007	751106		
3	2008	990819		
4	2009	947042		
5	2010	1120585		
6	2011	1349178		
7	2012	1459096		
8	2013	1522657		
9	2014	1586915		
10	2015	1988544		
11	2016	2385367		
12	2017	2983882		
13	2018	3 558 706		
14	2019	3175006,773		

Рис. 31. Результат функції ПРЕДСКАЗ

2. Зробити точковий прогноз, використовуючи стандартну функцію ТЕНДЕНЦИЯ. Після вибору даної функції відкривається діалогове вікно (рис. 32), в якому задаємо параметри.

Аргументы функции ? X

ТЕНДЕНЦИЯ

Известные_значения_y	B2:B13	=	{751106;990819;947042;1120585;1...
Известные_значения_x	A2:A13	=	{2007;2008;2009;2010;2011;2012;...
Новые_значения_x	A14	=	{2019}
Конст		=	логическое

= {3175006,77272725}

Возвращает значения в соответствии с линейной аппроксимацией по методу наименьших квадратов.

Известные_значения_y множество значений y, для которых уже известно соотношение $y = mx + b$.

Значение: 3175006,773

[Справка по этой функции](#) OK Отмена

Рис. 32. Діалогове вікно: аргумент функції ТЕНДЕНЦИЯ

Результат функції ТЕНДЕНЦИЯ подано на рис. 33.

	A	B	C	D
1	Роки	ВВП, млн. грн		
2	2007	751106		
3	2008	990819		
4	2009	947042		
5	2010	1120585		
6	2011	1349178		
7	2012	1459096		
8	2013	1522657		
9	2014	1586915		
10	2015	1988544		
11	2016	2385367		
12	2017	2983882		
13	2018	3 558 706		
14	2019	3175006,773		

Рис. 33. Результат функції ТЕНДЕНЦИЯ

Визначимо межі обсягу ВВП у 2019 р.: $y \pm t_{\alpha} S_{yt}$.

Розрахунок критерію Стьюдента за допомогою вбудованої функції СТЬЮДРАСПОБР (рис. 34) у разі довірчої імовірності 0,95.

Кількість спостережень дорівнює 12, тобто $n = 12$, кількість параметрів моделі – 2, тобто $m = 2$, звідси число ступенів свободи $12 - 2 = 10$.

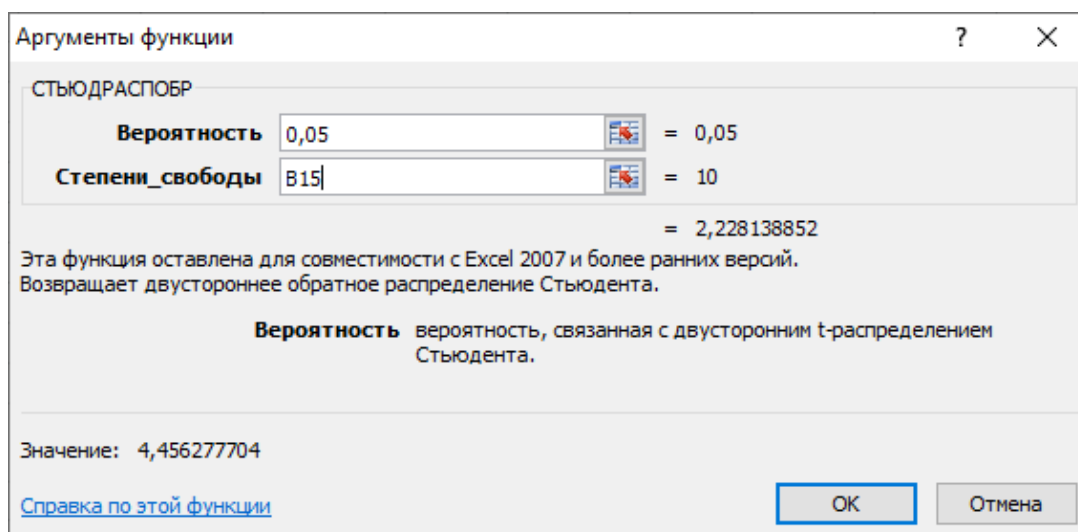


Рис. 34. Диалогове вікно: аргументи функції СТЬЮДРАСПОБР

Результат функції: $t_{\alpha} = 2,23$.

Розрахуємо S_{yt} (рис. 35) за формулою:

$$S_{yt} = \sqrt{\frac{(y_i - y_t)^2}{n - m}}, \quad (9)$$

де n – число рівнів ряду динаміки;

m – число параметрів адекватної моделі тренду (для рівняння прямої $m = 2$).

1	A	B	C				D				E				F				G				H				I			
			Розрахунок параметрів								Розрахунок F-критерію																			
2	Роки	ВВП, млн. грн (y _t)	Умовний показник часу, t	t ²	y*t	y _t	y _i - y _t	(y _i - y _t) ²	(y _i - \bar{y}) ²																					
3	2007	751106	-5,5	30,25	-4131083	489440	261666	68469115684,16	939384985351,56																					
4	2008	990819	-4,5	20,25	-4458686	713237	277581,8	77051658020,60	532178639283,06																					
5	2009	947042	-3,5	12,25	-3314647	937034	10007,57	100151455,91	597966211447,56																					
6	2010	1120585	-2,5	6,25	-2801463	1160832	-40246,7	1619793990,15	359687767730,06																					
7	2011	1349178	-1,5	2,25	-2023767	1384629	-35450,9	1256766211,65	137749910035,56																					
8	2012	1459096	-0,5	0,25	-729548	1608426	-149330	22299488582,12	68240459826,56																					
9	2013	1522657	0,5	0,25	761328,5	1832223	-309566	95831335659,84	39072539390,06																					
10	2014	1586915	1,5	2,25	2380373	2056021	-469106	220060065263,54	17798161395,06																					
11	2015	1988544	2,5	6,25	4971360	2279818	-291274	84840447342,61	71941566070,56																					
12	2016	2385367	3,5	12,25	8348785	2503615	-118248	13982606042,19	442281194285,06																					
13	2017	2983882	4,5	20,25	13427469	2727412	256469,7	65776704865,90	1596576924027,56																					
14	2018	3 558 706	5,5	30,25	19572883	2951210	607496,5	369051950781,75	3379645620351,56																					
15	Разом	20643897		143	32003005			1020340083900,40	8182523979194,25																					
16	t ₂	2,23																												
17	a ₀	1720324,75																												
18	a ₁	223797,23																												
19	y _{пр}	3175006,773																												
20	S _{yt}	319427,6262																												
21	нижня межа	2463277,67																												
22	верхня межа	3886735,877																												
23	залишкова дисперсія загальна	85028340325																												
24	дисперсія коеф.	681876998266,19																												
25	детермінації	0,88																												
26	F-розрах.	70,19408537																												
27																														

Рис. 35. Розрахунок прогнозного значення обсягу ВВП у 2019 р. за допомогою аналітичного вирівнювання

Для визначення меж прогнозного значення обсягу ВВП в 2019 р. необхідно провести розрахунки:

параметр a_0 розраховується в комірці B17 за формулою =B15/12;

параметр a_1 розраховується в комірці B18 за формулою =E15/D15;

прогнозне значення ($y_{пр}$) розраховується в комірці B19 за формулою =B17+B18*6,5;

S_{yt} розраховується в комірці B20 за формулою =КОРЕНЬ(H15/10);

нижня межа розраховується в комірці B21 за формулою =B19-B20*B16;

верхня межа розраховується в комірці B22 =B19+B20*B16 за формулою.

Прогнозне значення обсягу ВВП в 2019 р. буде знаходитися у межах:

$$3175006,773 - 2,23 \times 319427,6262 \leq y_{\text{пр}} \leq 3175006,773 + 2,23 \times 319427,6262,$$

$$2463277,67 \leq y_{\text{пр}} \leq 3886735,877.$$

Отже, прогнозне значення обсягу ВВП у 2019 р. буде знаходитись у межах від 2 463 277,67 до 3 886 735,877 млн грн. Це можна стверджувати з імовірністю 95 %.

Перевіримо адекватність тренду за допомогою критерію Фішера.

У комірці G3 вводимо формулу =B3-F3. Розтягуємо цю формулу на весь стовбець.

У комірці H3 вводимо формулу =G3^2. Розтягуємо цю формулу на весь стовбець. По стовпчику розраховуємо суму.

У комірці I3 вводимо формулу =(B3-\$B\$17)^2. Розтягуємо цю формулу на весь стовбець. За стовпцем розраховуємо суму.

Далі розраховуємо наступні показники:

залишкова дисперсія: в комірці B24 вводимо формулу =H15/I12;

загальна дисперсія: в комірці B25 вводимо формулу =I15/I12;

теоретичний коефіцієнт детермінації: в комірці B26 вводимо формулу =1-B24/B25;

критерій Фішера (F):

в комірці B27 вводимо формулу =(B26/(1-B26))*(12-2)/(2-1).

$$F_{\text{крит}} = 19,39;$$

$k_1=10, k_2=2$

$$F_{\text{факт}} > F_{\text{крит}}$$

$k_1=10, k_2=2$

Параметри моделі можна інтерпретувати в такий спосіб:

параметр тренду a_0 (вільний член рівняння регресії) характеризує вплив неврахованих в моделі факторів на зміну ВВП, а чисельно дорівнює середньому рівню для моменту або періоду часу, прийнятого за початок відліку, тобто для $t = 0$;

$a_1 = 223797,23$ – параметр рівняння тренду, який означає, що щороку обсяг ВВП зростає у середньому на 223 797,23 млн грн.

Значення коефіцієнта детермінації $\eta_T^2 = 0,88$ та дисперсійного критерію Фішера $F(2,10) = 70,19$ свідчать про адекватність побудованої моделі тренду.

Завдання для самостійного опрацювання

1. Маємо наступні дані про надходження металу на підприємство (табл. 19).

Таблиця 19

Динаміка постачання металу

Показник	Місяць											
	1	2	3	4	5	6	7	8	9	10	11	12
Обсяг постачання, тис. т	8,7	9,2	9,6	8,8	10,2	9,8	12,2	11,5	12,0	10,8	11,6	13,0

Замініть рядок щомісячних надходжень металу рядом середньоквартального надходження, рядом ковзної середньої. Зробіть висновки щодо тенденції в постачанні металу.

2. На підприємстві у 2017 році відбулися зрушення в організаційній структурі. Зімкніть динамічний ряд, обчисліть базисні абсолютні прирости обсягу виробництва, поясніть їхній зміст (табл. 20).

Таблиця 20

Вартість обладнання на підприємстві

Обсяг виробництва, млн грн	2013	2014	2015	2016	2017	2018	2019
До 2017 року	220	245	280	289	298		
Після 2017 року					280	260	240

3. Маємо дані про прибуток підприємства (табл. 21). Необхідно визначити відсутній рівень динамічного ряду та спрогнозувати обсяг прибутку на 8 років. Побудувати як точковий, так і інтервальний прогноз. Зробити висновки.

Дані про прибуток підприємства

Роки	1	2	3	4	5	6	7
Прибуток підприємства, млн грн	3,75	4,05	–	4,27	4,82	4,99	5,12

Тема 9. Аналіз коливань та сталості динамічних рядів

Методичні рекомендації

Лабораторна робота 5. Придбання навичок пошуку сезонної складової в динамічних рядах за допомогою MS Excel

Завдання 1. На основі поданих на рис. 36 даних необхідно розрахувати індекси сезонності.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1		Місяць											
2	Рік	1	2	3	4	5	6	7	8	9	10	11	12
3	2015	70	71	82	190	280	472	295	108	605	610	184	103
4	2016	71	85	84	308	383	443	261	84	630	450	177	168
5	2018	63	60	59	261	348	483	305	129	670	515	185	104
6													

Рис. 36. Вхідні дані

Для розрахунку індексів сезонності використовують різні методи, вибір яких залежить від характеру загальної тенденції ряду динаміки. Якщо ряд динаміки не має ярко вираженої тенденції розвитку, то індекси сезонності розраховують суто за емпіричними даними без їх попереднього вирівнювання. Для розрахунку індексу сезонності необхідно мати дані в місячному розрізі мінімум за три роки.

Для кожного місяця розраховується середній рівень (\bar{y}_i), а потім необхідно визначити середньомісячний рівень для всього досліджуваного ряду (\bar{y}). За цими даними знаходимо індекс сезонності (I_s) як відсоткове

відношення середніх для кожного місяця до загального середньомісячного рівня ряду:

$$I_s = \frac{\bar{y}_i}{\bar{y}} \times 100, \quad (10)$$

де \bar{y}_i – середньомісячні рівні ряду (за однойменними місяцями);

\bar{y} – загальний середній рівень ряду (постійна середня).

Використовуючи формулу середньоарифметичної простої ($\bar{y}_i = \frac{\sum y_i}{n}$), проведемо розрахунок середньомісячних рівнів за три роки (в місячному розрізі). Перед початком розрахунку середньої величини доцільно розрахувати загальну суму за місяцями, використовуючи вбудовану функцію СУММ (рис. 37).

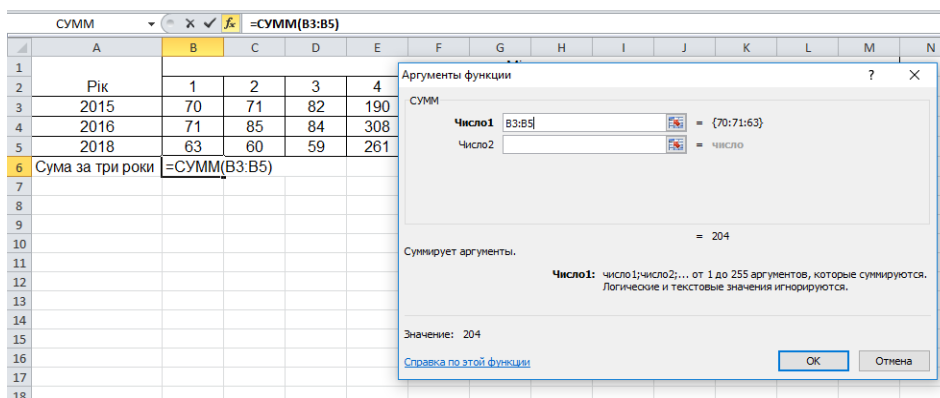


Рис. 37. Діалогове вікно "Аргументи функції СУММ"

Наступним кроком є розрахунок середньомісячної величини за три роки (\bar{y}_i). У комірці B7 прописуємо формулу =B6/3. Результати розрахунку середньомісячної величини подано на рис. 38.

		Місяць											
Рік		1	2	3	4	5	6	7	8	9	10	11	12
2015		70	71	82	190	280	472	295	108	605	610	184	103
2016		71	85	84	308	383	443	261	84	630	450	177	168
2018		63	60	59	261	348	483	305	129	670	515	185	104
Сума за три роки		204	216	225	759	1011	1398	861	321	1905	1575	546	375
\bar{y}_i		68	72	75	253	337	466	287	107	635	525	182	125

Рис. 38. Результат розрахунку середньомісячної величини

Наступним кроком є розрахунок загальної (постійної) середньої (\bar{y}). У комірці B8 прописуємо формулу =СУММ(B7:M7)/12.

Після розрахунку загальної середньої визначимо індекси сезонності (I_s) за кожним місяцем у річному розрізі, використовуючи формулу розрахунку: $B9=B7/\$B\$8*100$. Результат розрахунку індексів сезонності подано на рис. 39.

B9		=B7/\$B\$8*100											
	A	B	C	D	E	F	G	H	I	J	K	L	M
1							Місяць						
2	Рік	1	2	3	4	5	6	7	8	9	10	11	12
3	2015	70	71	82	190	280	472	295	108	605	610	184	103
4	2016	71	85	84	308	383	443	261	84	630	450	177	168
5	2018	63	60	59	261	348	483	305	129	670	515	185	104
6	Сума за три роки	204	216	225	759	1011	1398	861	321	1905	1575	546	375
7	\bar{y}_i	68	72	75	253	337	466	287	107	635	525	182	125
8	\bar{y}	261											
9	I_s	26,1	27,6	28,7	96,9	129,1	178,5	110,0	41,0	243,3	201,1	69,7	47,9
10													

Рис. 39. Результат розрахунку індексу сезонності

Для більш візуального аналізу можна побудувати графік сезонної хвилі розвитку досліджуваного явища протягом 12 місяців за три роки (рис. 40).

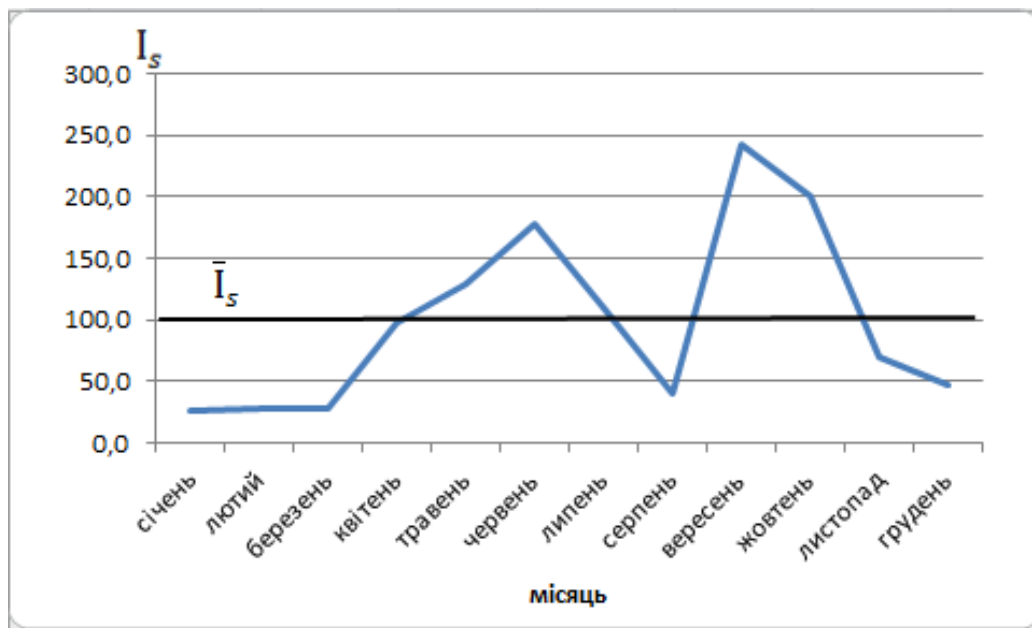


Рис. 40. Сезона хвиля (зміна індексу сезонності в середньому за три роки)

Таким чином, проведені розрахунки індексу сезонності свідчать про те, що найменший рівень розвитку явища припадає на січень-лютий, а найбільший – на вересень-жовтень.

Завдання 2. За даними про середньорічну динаміку заготовок сільськогосподарської продукції регіону в кварталному розрізі за три роки, необхідно розрахувати індекси сезонності.

Вхідна інформація для розрахунку подана на рис. 41.

	A	B	C	D
1		Заготовлено продукції, тис. грн		
2	Квартал	Перший рік	Другий рік	Третій рік
3	I	162	159	158
4	II	170	193	225
5	III	177	178	187
6	IV	151	168	172

Рис. 41. Вхідна інформація для розрахунку

Якщо рівні ряду динаміки проявляють тенденцію до зростання або зниження, то відхилення від постійного середнього рівня можуть змінити сезонні коливання. У таких випадках фактичні дані необхідно співставити з вирівняними.

Для розрахунку індексу сезонності в таких рядах динаміки використовується формула:

$$I_s = (\sum \frac{y_i}{\bar{y}_t} \times 100) / n, \quad (11)$$

де y_i – емпіричні рівні ряду;

\bar{y}_t – теоретичні рівні ряду;

n – кількість років.

Теоретичні значення \bar{y}_t знаходяться за рівнянням:

$$\bar{y}_t = a_0 + a_1 t. \quad (12)$$

Проведемо розрахунок параметру a_0 . На початку необхідно розрахувати загальну суму заготовленої продукції за три роки. У комірці B17

вводимо формулу =СУММ(B3:B6;B8:B11;B13:B16). Аналогічно розраховується загальна сума за всіма стовпчиками (рис. 42).

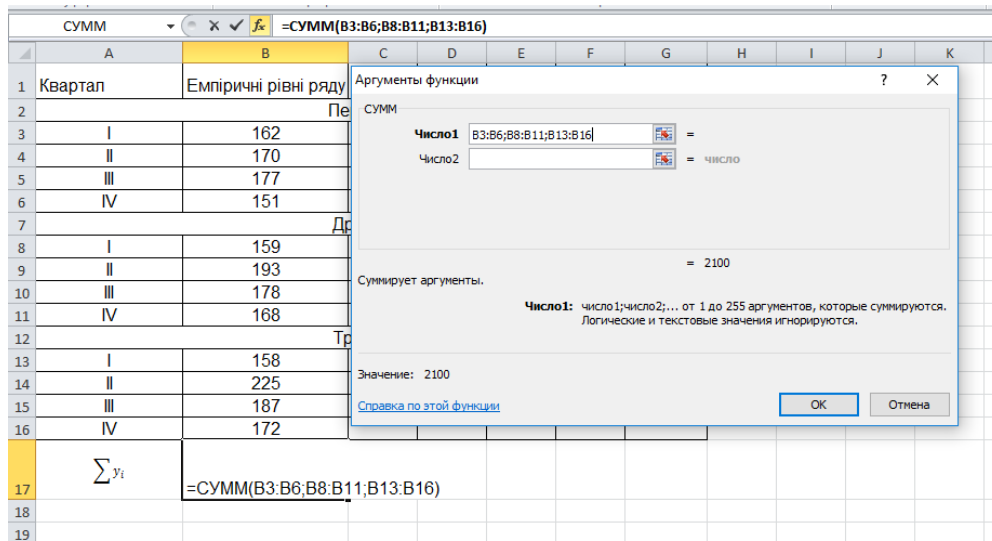


Рис. 42. Діалогове вікно визначення суми

Після розрахунку загальної суми знаходимо параметр a_0 . У комірці B18 записуємо формулу =B17/12, а для знаходження параметру a_1 в комірці B19 вписуємо формулу =E17/D17. Результат розрахунку параметрів a_0 і a_1 подано на рис. 43.

Квартал	Емпіричні рівні ряду y_i	t	t ²	y*t	\bar{y}_t	y_i / \bar{y}_t
Перший рік						
I	162	-5,5	30,25	-891		
II	170	-4,5	20,25	-765		
III	177	-3,5	12,25	-619,5		
IV	151	-2,5	6,25	-377,5		
Другий рік						
I	159	-1,5	2,25	-238,5		
II	193	-0,5	0,25	-96,5		
III	178	0,5	0,25	89		
IV	168	1,5	2,25	252		
Третій рік						
I	158	2,5	6,25	395		
II	225	3,5	12,25	787,5		
III	187	4,5	20,25	841,5		
IV	172	5,5	30,25	946		
Сума	2100	0	143	323		
a_0	175					
a_1	2,258741259					

Рис. 43. Результат розрахунку параметрів a_0 і a_1

Таким чином, рівняння прямої буде мати вигляд:

$$\bar{y}_t = 175 + 2,26 \times t.$$

Після визначення параметрів моделі необхідно в рівняння прямої підставити значення t (квартальні), в результаті чого отримаємо вирівняні значення ряду.

Наприклад, для першого року перший квартал у комірці F3 прописуємо формулу =B18+B19*C3. Аналогічно проводяться розрахунки для всіх кварталів по кожному року. Результат вирівняних значень ряду подано на рис. 44.

F4 fx =B18+B19*C4						
	A	B	C	D	E	F
1	Квартал	Емпіричні рівні ряду y_t	t	t^2	$y \cdot t$	\bar{y}_t
2	Перший рік					
3	I	162	-5,5	30,25	-891	162,58
4	II	170	-4,5	20,25	-765	164,84
5	III	177	-3,5	12,25	-619,5	167,09
6	IV	151	-2,5	6,25	-377,5	169,35
7	Другий рік					
8	I	159	-1,5	2,25	-238,5	171,61
9	II	193	-0,5	0,25	-96,5	173,87
10	III	178	0,5	0,25	89	176,13
11	IV	168	1,5	2,25	252	178,39
12	Третій рік					
13	I	158	2,5	6,25	395	180,65
14	II	225	3,5	12,25	787,5	182,91
15	III	187	4,5	20,25	841,5	185,16
16	IV	172	5,5	30,25	946	187,42
17	Сума	2100	0	143	323	
18	a₀	175				
19	a₁	2,258741259				

Рис. 44. Результат вирівнювання значень ряду динаміки

Наступним кроком необхідно знайти для кожного кварталу відсоткове відношення емпіричних рівнів ряду до теоретичних рівнів. Для першого кварталу першого року в комірці G3 прописуємо формулу =B3/F3*100. Аналогічним чином проводимо розрахунок для всіх кварталів по трьом рокам. Результат розрахунку подано на рис. 45.

G4		fx =B4/F4*100					
	A	B	C	D	E	F	G
1	Квартал	Емпіричні рівні ряду y_t	t	t ²	y*t	\bar{y}_t	y_t / \bar{y}_t
2	Перший рік						
3	I	162	-5,5	30,25	-891	162,58	99,65
4	II	170	-4,5	20,25	-765	164,84	103,13
5	III	177	-3,5	12,25	-619,5	167,09	105,93
6	IV	151	-2,5	6,25	-377,5	169,35	89,16
7	Другий рік						
8	I	159	-1,5	2,25	-238,5	171,61	92,65
9	II	193	-0,5	0,25	-96,5	173,87	111,00
10	III	178	0,5	0,25	89	176,13	101,06
11	IV	168	1,5	2,25	252	178,39	94,18
12	Третій рік						
13	I	158	2,5	6,25	395	180,65	87,46
14	II	225	3,5	12,25	787,5	182,91	123,01
15	III	187	4,5	20,25	841,5	185,16	100,99
16	IV	172	5,5	30,25	946	187,42	91,77
17	Сума	2100	0	143	323		
18	a ₀	175					
19	a ₁	2,258741259					

Рис. 45. Розрахунок відсоткового відношення емпіричних рівнів ряду до теоретичних

Після розрахунку відношення необхідно додати отримані відсоткові співвідношення емпіричних рівнів ряду до теоретичних рівнів за три роки за однойменними кварталами, використовуючи вбудовану функцію СУММ.

Тобто для першого кварталу в комірці B20 прописуємо формулу: =G3+G8+G13; для другого кварталу в комірці B21=G4+G9+G14.

Аналогічним чином проводиться розрахунок для третього та четвертого кварталу. Результат розрахунку подано на рис. 46.

B20		fx =G3+G8+G13	
	A	B	C
20	1 квартал	279,76	
21	2 квартал	337,15	
22	3 квартал	307,98	
23	4 квартал	275,11	

Рис. 46. Розрахунок суми відсоткового відношення

Після проведених розрахунків визначаємо індекси сезонності, тобто в комірці A25 вводимо формулу =B20/3. Аналогічним чином проводимо

розрахунок для наступних кварталів. Результат розрахунку індексів сезонності подано на рис. 47.

	A	B
24	Індекс сезонності, %	
25	1 квартал	93,25
26	2 квартал	112,38
27	3 квартал	102,66
28	4 квартал	91,70

Рис. 47. Результат розрахунку індексу сезонності, %

Отже, індекси сезонності характеризують розміри заготовок сільськогосподарської продукції залежно від пори року. Найбільша питома вага заготовок сільськогосподарської продукції припадає на другий квартал, а найменша – на четвертий квартал.

Завдання для самостійного опрацювання

1. Маємо дані про динаміку продаж туристичних путівок (табл. 22).

Таблиця 22

Динаміка продажу туристичних путівок в Харківському регіоні за 2017 – 2019 роки

Квартали	Продано туристичних путівок, од.		
	2017 р.	2018 р.	2019 р.
I	170	161	159
II	175	215	242
III	190	191	207
IV	171	186	195

Для аналізу динаміки продажу туристичних путівок впродовж року необхідно визначити індекс сезонності методом змінної середньої та зробити висновки.

2. Є дані про реалізацію яєць у магазинах міста по місяцях у млн шт. (табл. 23).

Таблиця 23

Реалізація яєць за місяцями

Місяць	2017 рік	2018 рік	2019 рік
1	11,9	16,2	22,9
2	11,2	17,6	13,6
3	12,1	18,2	16,4
4	42,6	77,4	89,8
5	40,2	68,2	74,6
6	58,4	65,8	50,3
7	41,8	41,3	20,8
8	22,8	38,6	14,1
9	14,6	27,9	13,6
10	15,3	25	24,2
11	16,2	27,3	31,3
12	17,8	31,1	20,9

Проаналізуйте сезонні коливання продажів за допомогою індексів сезонності, що розраховуються методом постійної середньої. Результати подайте графічно.

3. За умовами контракту поставка товару оптовому покупцю повинна виконуватися однаковими партіями з інтервалом у п'ять днів.

У табл. 24 подано дані щодо товару, поставленого в минулому місяці.

Таблиця 24

Дані щодо поставок товару

П'ятиднівка	1	2	3	4	5	6
Поставка	126	108	184	142	108	232

Визначте середній рівень та коефіцієнт нерівномірності поставок, поясніть його зміст.

4. Обсяг реалізації товару оптовому покупцю за роками наведено у табл. 25.

Таблиця 25

Дані про оптові поставки

Роки	1	2	3	4	5	6	7	8	9	10	11	12
Кількість одиниць	192	297	250	219	212	270	321	175	250	370	383	251

Здійсніть згладжування ряду динаміки за допомогою ряду Фур'є.

Тема 10. Кореляція динамічних рядів

Методичні рекомендації

Лабораторна робота 6. Придбання навичок виявлення та усунення автокореляції в рядах динаміки за допомогою ППП Statistica 10.0

Завдання. Вхідна інформація для виявлення та усунення автокореляції у динамічному ряді представлена в лабораторній роботі 3 (табл. 26).

Існує два найбільш поширених метода виявлення та усунення автокореляції залишків у динамічних рядах:

побудова графіка залежності залишків від часу і візуальне визначення наявності або відсутності автокореляції;

використання критерію Дарбіна – Уотсона і розрахунок величини.

Критерій Дарбіна – Уотсона (або DW-критерій) використовується для тестування автокореляції першого порядку елементів досліджуваної послідовності. Найбільш часто застосовується при аналізі часових рядів і залишків регресійних моделей.

Для визначення коефіцієнта Дарбіна – Уотсона, для оцінювання наявності автокореляції в залишках моделі в меню аналізу помилок ініціюємо кнопку Descriptive statistics (рис. 48).

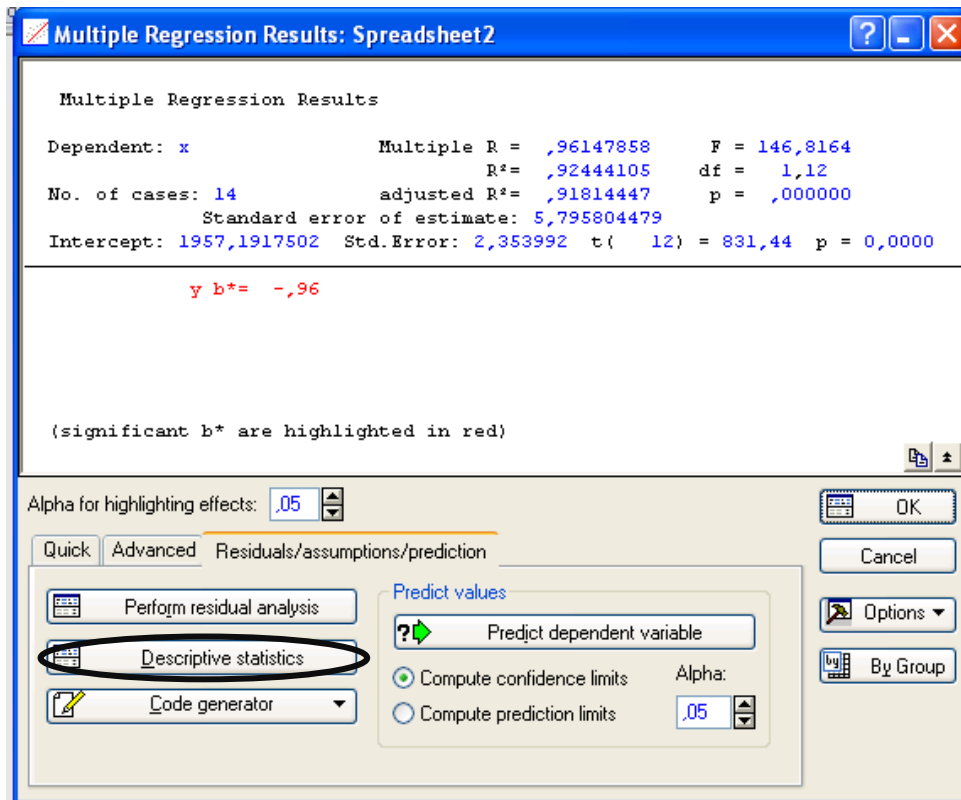


Рис. 48. Меню аналізу помилок

У меню аналізу помилок ініціювавши кнопку Durbin-Watson statistic (Статистика Дарбіна – Уотсона) (рис. 49) отримаємо значення автокореляції помилок моделі за критерієм Дарбіна – Уотсона і за значенням нециклічного коефіцієнта автокореляції (рис. 49).

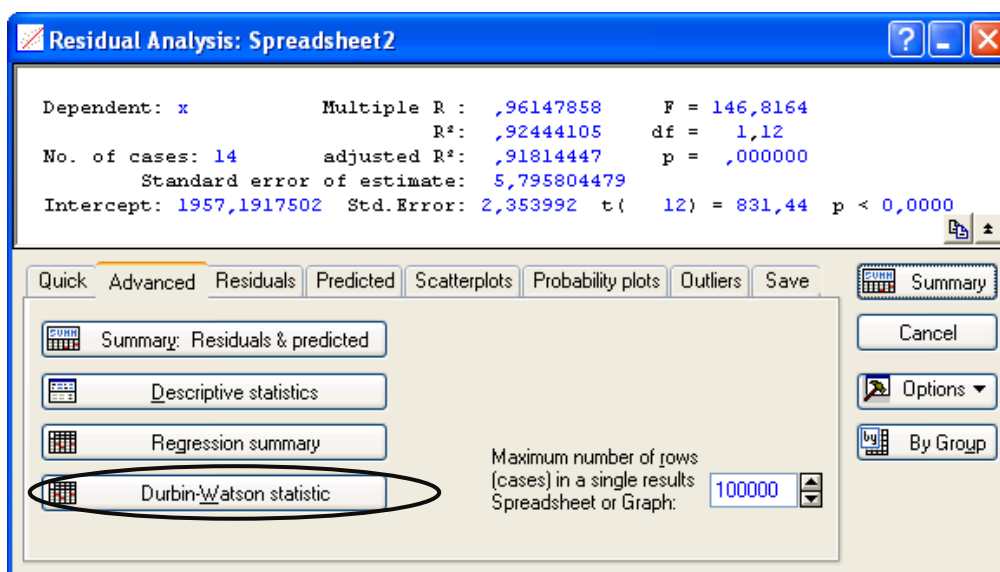


Рис. 49. Діалогове вікно вибору критерію Дарбіна – Уотсона

Ініціювавши дану кнопку отримаємо результати розрахунків, які подані на рис. 50.

Durbin-Watson d (Spreadsheet2) and serial correlation of residuals			
	Durbin-Watson d	Serial Corr.	
Estimate	1,127374	0,350804	

Рис. 50. Результат розрахунку критерію Дарбіна – Уотсона

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 0,35, що свідчить про наявність позитивної автокореляції.

Завдання для самостійного опрацювання

1. Визначте, чи існує автокореляція між послідовними значеннями серед рівнів динамічного ряду середньомісячної заробітної плати у звітному році (табл. 26). Зробіть висновки.

Таблиця 26

Вихідні дані

Місяць	Середньомісячна заробітна плата, у. о.
Січень	1 521
Лютий	1 633
Березень	1 702
Квітень	1 735
Травень	1 774
Червень	1 883
Липень	1 930
Серпень	1 872
Вересень	1 916
Жовтень	1 917
Листопад	1 823
Грудень	2 001

2. Визначте, що підвищувалося більш швидкими темпами: годинна заробітна плата або ціни (табл. 27). Зробіть висновки та поясніть проведення розрахунків.

Таблиця 27

Динаміка цін і заробітної плати

Місяць	1	2	3	4	5	6	7	8	9
Ціна, грн	120	121	123	123	124	126	134	137	145
Годинна заробітна плата, у. о.	20	21	21	23	24	24	26	27	28

Тема 11. Динамічний індексний аналіз

Методичні рекомендації

Лабораторна робота 7. Поглиблення навичок аналізу динаміки показників за допомогою індексного методу в MS Excel

Завдання 1. Маємо дані про реалізацію товарів за червень та липень поточного року (табл. 28). На основі наведених даних необхідно визначити: загальний індекс товарообігу; загальний індекс цін; загальний індекс фізичного обсягу товарообігу; абсолютну суму економії або перевитрат за рахунок зміни двох факторів, розрахувати мультиплікативну залежність між індексами. Зробити висновок.

Таблиця 28

Реалізація товару за червень та липень поточного року

Найменування товару	Ціна товару, грн за кг		Продано товару, тис. кг	
	червень	липень	червень	липень
Черешня	45	65	18	15
Персик	70	50	22	27

Розрахунки загальних індексів наведено на рис. 51.

Для визначення загального індексу товарообігу в комірці F3 вводимо формулу =C3*E3 і розтягуємо цю формулу на весь стовпчик, у комірці F5 розраховуємо суму по стовбцю: або нажавши знак автосуми Σ на панелі інструментів, або нажавши "=" та обравши формулу СУММ у рядку формул, у комірці G3 вводимо формулу =B3*D3 і розтягуємо цю формулу на весь стовпчик, у комірці G5 розраховуємо суму. Для розрахунку індексу товарообігу у комірці H5 вводимо формулу =F5/G5.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Найменування	Ціна товару, грн (р)		Прдано товару, тис. кг (q)		r1q1	p0q0	Ipq	p0q1	Ip	Iq	Δp	Ipq=Ip*Iq
2	товару	червень	липень	червень	липень								
3	Черешня	45	65	18	15	975	810		675				
4	Персик	70	50	22	27	1350	1540		1890				
5	Σ					2325	2350	0,99	2565	0,91	1,09	-240,00	0,99

Рис. 51. Розрахунок загальних індексів

Загальний індекс товарообігу дорівнює 0,99 або 99 %, тобто товарообіг у поточному періоді зменшився на 1 % порівняно з базисним періодом за рахунок зміни цін на продукцію, а також за рахунок зміни кількості реалізованої продукції.

Для розрахунку загального індексу ціни в комірці I3 вводимо формулу =B3*E3 і розтягуємо цю формулу на весь стовпчик, у комірці I5 розраховуємо суму по стовпчику. Для розрахунку індексу ціни у комірці J5 вводимо формулу =F5/I5.

Загальний індекс ціни дорівнює 0,91 або 91 %, тобто товарообіг у поточному періоді зменшився на 9 % порівняно з базисним періодом за рахунок зміни цін на продукцію при незмінній кількості реалізованої продукції.

Для розрахунку загального індексу фізичного обсягу продаж в комірці K5 вводимо формулу =I5/G5.

Загальний індекс фізичного обсягу продаж дорівнює 1,09 або 109 %, тобто товарообіг у поточному періоді збільшився на 9 % порівняно з базисним періодом за рахунок зміни кількості реалізованої продукції за незмінної ціни на продукцію.

Для розрахунку абсолютної суми економії або перевитрат від зміни двох факторів у комірці L5 вводимо формулу =F5-G5. Якщо розраховане значення зі знаком мінус – це економія, зі знаком плюс – перевитрати.

У даному випадку $\Delta q_r = -240,00$ грн, тобто абсолютна сума зниження товарообороту за рахунок двох факторів склала 240,00 грн.

Для розрахунку мультиплікативного зв'язку між індексами у комірці М5 вводимо формулу $=J5*K5$.

Завдання 2. Маємо дані про витрати на виробництво продукції та про зміни в собівартості на підприємстві (табл. 29). Необхідно визначити загальні індекси: 1) собівартості; 2) обсягу виробництва; 3) витрат на виробництво за умови, що витрати на виробництво у другому кварталі порівняно з першим кварталом збільшилися на 11 %. Зробити висновок.

Таблиця 29

Характеристика виробництва з урахуванням асортименту

Назва виробу	Витрати на виробництво у другому кварталі, тис. грн	Зміни собівартості одиниці виробу у другому кварталі порівняно з першим кварталом, %
А	22 526	+12
Б	11 920	+15

Розрахунок загальних індексів наведено на рис. 52.

	А	В	С	Д	Е	Ф	Г	Н
	Назва виробу	Витрати на виробництво у 2 кварталі, тис. грн. (z_1q_1)	Зміни собівартості одиниці виробу у 2 кварталі порівняно з 1 кварталом, %	iz	z_1q_1/iz	Iz	Izq	Iq
1								
2	А	22526	12	1,12	20113			
3	Б	11920	15	1,15	10365			
4	Σ	34446			30478	1,13	1,11	0,98

Рис. 52. Розрахунок загальних індексів

Для розрахунку індексу собівартості необхідно використовувати формулу середньозважених індексів, оскільки відомі витрати за звітний період (z_1q_1) і зміна собівартості (I_z). У комірці D2 розраховуємо індивідуальний індекс собівартості за формулою $=(C2+100)/100$ та розтягуємо по стовпчику, по другому стовпчику в комірці В4 (витрати на виробництво

у другому кварталі, тис. грн) розраховуємо суму, нажавши знак автосуми Σ на панелі інструментів, попередньо виділивши весь стовпчику. В комірці E2 вводимо формулу =B2/D2 та розтягуємо на весь стовпчику, у комірці E5 розраховуємо суму по стовпчику: або нажавши знак автосуми Σ на панелі інструментів, або нажавши "=" та обравши формулу СУММ у рядку формул. Для отримання значення загального індексу собівартості в комірці F4 вводимо формулу =B4/E4.

Загальний індекс собівартості дорівнює 1,13 або 113 %, тобто загальні витрати на виробництво продукції збільшилися на 13 % у звітному періоді порівняно з базовим за рахунок збільшення собівартості одиниці кожного виробу та за незмінного випуску продукції.

Виходячи з умови завдання, загальний індекс витрат (I_{zq}) складає 1,11 або 111 %, оскільки витрати на виробництво продукції у другому кварталі порівняно з першим збільшуються на 11 % за рахунок зміни собівартості одиниці виробу та обсягу випуску продукції.

Для розрахунку загального індексу обсягу виробництва використаємо взаємозв'язок між індексами ($I_{pq} = I_p \times I_q$). У комірці H4 вводимо формулу =G4/F4.

Загальний індекс обсягу виробництва склав 0,98 або 98 %, тобто загальні витрати на виробництво продукції зменшилися на 2 % за рахунок зміни обсягу виробництва продукції при незмінній собівартості одиниці виробу.

Завдання 3. Маємо дані про виробництво однорідної продукції та її собівартість на двох підприємствах (табл. 30). Необхідно визначити індекси собівартості: мінливого складу; постійного складу; структурних зрушень. Пояснити різницю між цими індексами.

Таблица 30

Дані виробництва продукції та її собівартість на підприємствах

Продукція	Базовий період		Звітний період	
	Кількість, шт.	Собівартість за одиницю, у. о.	Кількість продукції, шт.	Собівартість за одиницю, у. о.
А	1 050	142	920	155
Б	980	185	1 000	200

Для визначення індексів, насамперед, розрахуємо суму за стовпчиками В і D (кількість продукції).

Для розрахунку індексу собівартості мінливого складу в комірці F3 вводимо формулу =E3*D3 і розтягуємо на весь стовпчик, далі розраховуємо суму за цим стовпчиком. В комірці G5 вводимо формулу =F5/D5, в комірці H3 вводимо формулу =C3*B3 і розтягуємо на весь стовпчик, далі розраховуємо суму за цим стовпчиком. У комірці I5 вводимо формулу =H5/B5. Остаточні розрахунки індексу собівартості мінливого складу наведено у комірці J5 за формулою =G5/I5.

Розрахунок індексу собівартості мінливого складу наведено на рис. 53.

	A	B	C	D	E	F	G	H	I	J
1		Базовий період		Звітний період						
2	Продукція	Кількість продукції, шт.	Собівартість за одиницю, ум.од.	Кількість продукції, шт.	Собівартість за одиницю, ум.од.	z1q1	$\frac{\sum z_1 q_1}{\sum q_1}$	z0q0	$\frac{\sum z_0 q_0}{\sum q_0}$	І мінл.скл
3	А	1050	142	920	155	142600		149100		
4	Б	980	185	1000	200	200000		181300		
5		2030		1920		342600	178,44	330400	162,76	1,10

Рис. 53. Розрахунок індексу собівартості мінливого складу

Значення індексу собівартості мінливого складу становить 1,10 або 110 %, тобто середня собівартість одиниці продукції збільшилась на 10 % за рахунок зміни собівартості та кількості виробленої продукції.

Розрахунок індексу собівартості постійного складу наведено на рис. 54.

	A	B	C	D	E	F	G	H	I	J
1		Базовий період		Звітний період						
2	Продукція	Кількість продукції, шт.	Собівартість за одиницю, ум.од.	Кількість продукції, шт.	Собівартість за одиницю, ум.од.	z1q1	$\frac{\sum z_1 q_1}{\sum q_1}$	z0q1	$\frac{\sum z_0 q_1}{\sum q_1}$	І пост. Скл
3	А	1050	142	920	155	142600		130640		
4	Б	980	185	1000	200	200000		185000		
5		2030		1920		342600	178,44	315640	164,40	1,09

Рис. 54. Розрахунок індексу собівартості постійного складу

Для розрахунку індексу собівартості постійного складу в комірці F3 вводимо формулу =E3*D3 і розтягуємо на весь стовпчик, далі розраховуємо суму за цим стовпчиком. У комірці G5 вводимо формулу =F5/D5, в комірці H3 вводимо формулу =D3*C3 і розтягуємо на весь стовпчик, далі розраховуємо суму за цим стовпчиком. У комірці I5 вводимо формулу =H5/D5. Остаточні розрахунки індексу собівартості постійного складу наведено у комірці J5 за формулою =G5/I5.

Індекс собівартості фіксованого (постійного) складу склав 1,09 (109 %), тобто середня собівартість продукції збільшилась на 9 % за рахунок зміни собівартості одиниці продукції за незмінної кількості виробленої продукції.

Розрахунок індексу собівартості структурних зрушень наведено на рис. 55.

	A	B	C	D	E	F	G	H	I	J
1		Базовий період		Звітний період						
2	Продукція	Кількість продукції, шт.	Собівартість за одиницю, ум.од.	Кількість продукції, шт.	Собівартість за одиницю, ум.од.	z_0q_1	$\frac{\sum z_0q_1}{\sum q_1}$	z_0q_0	$\frac{\sum z_0q_0}{\sum q_0}$	І стр.зруш.
3	A	1050	142	920	155					
4	B	980	185	1000	200	130640		149100		
5		2030		1920		315640	164.40	330400	162.76	1.01

Рис. 55. Розрахунок індексу собівартості структурних зрушень

Для розрахунку індексу собівартості структурних зрушень у комірці F3 вводимо формулу =D3*C3 і розтягуємо на весь стовпчик, далі розраховуємо суму за цим стовпчиком. У комірці G5 вводимо формулу =F5/D5, у комірці H3 вводимо формулу =C3*B3 і розтягуємо на весь стовпчик, далі розраховуємо суму за цим стовпчиком. У комірці I5 вводимо формулу =H5/B5. Остаточні розрахунки індексу собівартості постійного складу наведено в комірці J5 за формулою =G5/I5.

Індекс собівартості структурних зрушень склав 1,01 (101 %), тобто середня собівартість продукції збільшилась на 1 % за рахунок зміни кількості виробленої продукції за незмінної собівартості одиниці продукції.

Завдання для самостійного опрацювання

1. Маємо дані про випуск фахівців закладами вищої освіти в країні, тис. осіб (табл. 31).

Таблиця 31

Динаміка випуску фахівців ЗВО країни

Роки	Випуск фахівців, тис. осіб
1	833
2	841
3	849
4	855
5	859
6	840
7	768

Визначте ланцюгові та базисні індекси. Перевірте взаємозв'язок розрахованих індексів та зробіть висновок.

2. Динаміка депозитних процентних ставок для юридичних та фізичних осіб характеризується даними (табл. 32).

Таблиця 32

**Динаміка депозитних процентних ставок
залежно від юридичного статусу вкладника**

Вкладники	Середня депозитна ставка, %		Сума залучених депозитів, тис. грн	
	базисний період	звітний період	базисний період	звітний період
Юридичні особи	40	26	820	950
Фізичні особи	25	20	180	450

Визначте: а) середню депозитну ставку в базисному та звітному періодах та індекс середньої депозитної ставки (змінного складу); б) індекси середньої депозитної ставки фіксованого складу та структурних зрушень. Поясніть їх економічний зміст та покажіть взаємозв'язок.

3. Маємо наступні дані про продуктивність праці робітників різних професій (табл. 33).

Таблиця 33

Розподіл працівників різних професій за рівнем продуктивності праці

Професія	Обсяг наданих послуг за період, тис. грн		Загальні витрати часу на надання послуг за період, люд.-год	
	базовий	звітний	базовий	звітний
1. Перукар	15,1	15,5	1069	1095
2. Майстер з ремонту побутової техніки	4,8	5,0	340	350
3. Майстер з ремонту взуття	11,7	12,1	670	710

За допомогою індексного аналізу визначте як змінилась продуктивність праці за кожною професією та середня продуктивність праці за всіма професіями. Зробіть висновки.

4. Є дані про обсяг попиту на банківські послуги і чисельність клієнтів банку (табл. 34).

Таблиця 34

Дані за клієнтами банку

Групи клієнтів банку	Базовий період		Звітний період	
	Сума наданих кредитів, млн грн	Середня чисельність позичальників, одиниць	Сума наданих кредитів, млн грн	Середня чисельність позичальників, одиниць
А	2,61	163	5,83	258
Б	7,02	330	13,20	350

Розрахувати:

а) індекси рівня середнього обсягу наданих позик змінного, фіксованого складу та структурних зрушень;

б) абсолютний розмір приросту загального обсягу наданих позик за рахунок зміни середнього рівня позик і чисельності позичальників. Зробити висновки.

5. За наведеними даними (табл. 35) визначте зведений індекс біржових цін акцій та абсолютний приріст обсягу продажу за рахунок динаміки цін.

Таблиця 35

Дані за емітентами

Емітент	Обсяги продажу, млн грн		Індекс цін
	Вересень	Жовтень	
А	400	450	1,1
В	200	150	1,3
С	300	420	0,9

6. Процентні ставки на кредити різного призначення характеризуються такими змінами (табл. 36).

Таблиця 36

Дані з кредитування

Цільове спрямування кредиту	Кредитна ставка, %		Сума наданих кредитів, млн грн	
	Базисний період	Поточний період	Базисний період	Поточний період
Розрахункові операції	15	20	3,5	2,0
Інвестиції	22	24	6,5	8,0

Визначте індекси змінного, постійного складу та структурних зрушень та зробіть висновок.

7. Реалізація зерна та цукрових буряків аграрними підприємствами регіону характеризується даними (табл. 37).

Таблиця 37

Дані щодо реалізації товару

Вид продукції	Загальний обсяг продажу, млн грн		Індивідуальні індекси фізичного обсягу реалізації
	1 рік	2 рік	
Зерно	316,4	284,2	0,820
Цукрові буряки	115,0	112,3	0,845
Разом	431,4	396,5	X

Визначте зведений індекс цін та обчисліть абсолютну зміну товарообігу за рахунок зміни цін. Побудуйте мультиплікативну й адитивну моделі товарообігу. Зробіть висновки.

8. Визначте, як змінилася середня виробітка одного робітника, якщо чисельність робітників зменшилася на 10 %, обсяг випуску продукції збільшився на 6%.

9. Фактичний обсяг продукції збільшився на 3,6 %. Собівартість продукції знизилася на 1,5 %. Як змінилися витрати на виробництво продукції?

10. Дані про виручку від реалізації жіночого одягу і зміну фізичного обсягу (кількості) реалізованого товару наведено в табл. 38.

Таблиця 38

Дані щодо товарообігу

Групи товару	Товарообіг у фактичних цінах (тис. грн) за період		Темп приросту цін на реалізовані товари у поточному періоді порівняно з базисним
	базисний	поточний	
Верхній одяг	400,4	619,5	+10,0
Сукня	849,6	980,5	+25,0
Разом	1 250,0	1 600,0	X

Обчисліть загальний індекс цін, абсолютну суму перевитрат грошових коштів населення за рахунок підвищення цін та загальний індекс фізичного обсягу реалізації. Зробіть висновки.

Тема 12. Просторовий індексний аналіз

Методичні рекомендації

Лабораторне заняття

Індексний метод не тільки характеризує динаміку складного явища, але й аналізує вплив на неї окремих факторів. Індексний метод дозволяє оцінити вплив окремих факторів як у відносному, так і в абсолютному вираженні.

У статистичній практиці прийнято наступне правило факторного аналізу: якщо результативний показник можна представити як добуток кількісного та якісного фактора, то, визначаючи вплив кількісного фактора на зміну результативного показника, якісний фактор фіксують на рівні

базисного періоду; якщо ж визначається вплив якісного показника, то кількісний фактор фіксують на рівні звітного періоду.

Розглянуті індексні системи являють собою двофакторні системи, але, якщо результативна ознака залежить від трьох та більше факторів, то будують багатфакторні індексні моделі.

Використовують два методи розкладання загального індексу на факторні:

- 1) метод ізольованого вивчення факторів;
- 2) метод ланцюгової підстановки (взаємозв'язаного вивчення факторів).

Оскільки в дійсності всі явища взаємозв'язані, то основним є другий метод, який потребує правильного розташування факторів під час побудови моделі результативної ознаки (наприклад, $W = a \times b \times c \times d$).

На першому місці в моделі необхідно ставити якісний фактор. Збільшення ланцюга факторів на один фактор (наприклад, $a \times b$) кожен раз повинно приводити до показника, що має реальний економічний сенс.

Під час визначення впливу першого фактора, усі інші фактори беруться на рівня звітного періоду (і у чисельнику, і у знаменнику).

Під час визначення впливу другого фактора, перший фактор зберігається на рівні базисного періоду, третій і всі наступні – на рівні звітного періоду і т. д.

Наприклад, $w = a \times b \times c \times d$. Звідси послідовно-ланцюгове розкладання факторів буде мати вигляд:

$$I_w = \frac{W_1}{W_0} = \frac{a_1 b_1 c_1 d_1}{a_0 b_0 c_0 d_0} = I_a \times I_b \times I_c \times I_d \quad (13)$$

або

$$\frac{a_1 b_1 c_1 d_1}{a_0 b_0 c_0 d_0} = \frac{a_1 b_1 c_1 d_1}{a_0 b_1 c_1 d_1} \times \frac{a_0 b_1 c_1 d_1}{a_0 b_0 c_1 d_1} \times \frac{a_0 b_0 c_1 d_1}{a_0 b_0 c_0 d_1} \times \frac{a_0 b_0 c_0 d_1}{a_0 b_0 c_0 d_0} = \frac{a_1}{a_0} \times \frac{b_1}{b_0} \times \frac{c_1}{c_0} \times \frac{d_1}{d_0}. \quad (14)$$

Індексні системи застосовуються і для визначення в абсолютному вираженні зміни складного явища за рахунок впливу окремих факторів. Розрахунки, що пов'язані з визначенням в абсолютному вираженні зміни результативного показника за рахунок окремих факторів, називаються розкладанням абсолютного приросту (скорочення) за факторами.

Так, розглянута індексна система чотирифакторного зв'язку (14) може бути представлена в абсолютних величинах наступним чином:

$$\begin{aligned} & a_1 b_1 c_1 d_1 - a_0 b_0 c_0 d_0 = \\ & = (a_1 - a_0) b_1 c_1 d_1 + (b_1 - b_0) a_0 c_1 d_1 + (c_1 - c_0) a_0 b_0 d_1 + (d_1 - d_0) a_0 b_0 c_0. \end{aligned} \quad (15)$$

Така модель називається адитивною.

Статистичні індекси використовують не тільки для дослідження змін явищ і процесів у часі, а й для характеристики змін рівнів соціально-економічних явищ у просторі. Зокрема, статистика широко застосовує метод порівняння показників у розрізі підприємств, міст, економічних районів, областей, країн. Узагальнюючі показники, які характеризують співвідношення рівнів складних економічних явищ у просторі, тобто в розрізі територій і об'єктів, називаються територіальними індексами. Побудова територіальних індексів має свої особливості порівняно з індексами, які характеризують динаміку явищ. Щодо обчислення індивідуальних територіальних індексів, то тут ніяких труднощів не виникає, адже обчислюються звичайні відносні величини порівняння.

Під час побудови загальних територіальних індексів виникає питання вибору бази порівняння і об'єкта, на рівні якого слід зафіксувати вагу індексу. У кожному конкретному випадку його вирішують, виходячи з мети самого дослідження.

Порівняння показників можна здійснювати або за двома територіями (об'єктами), або за низкою територій (об'єктів). У першому випадку базою може бути показник будь-якої з територій, а в другому – база порівняння повинна бути економічно обґрунтованою. Так, якщо порівнюється, наприклад, рентабельність по колу однотипних підприємств з приблизно однаковими техніко-економічними умовами виробництва, то за базу порівняння слід узяти підприємство, яке має найвищий рівень рентабельності, тобто найкраще в даному випадку підприємство.

Під час побудови територіальних індексів інтенсивних показників вагами можуть бути:

середня величина екстенсивного показника за сукупністю одиниць порівнюваних територій;

екстенсивний показник, що належить до території, на якій інтенсивний показник є більш економічним;

екстенсивний показник, який прийнято за стандарт.

Під час побудови територіальних індексів для екстенсивних показників як вимірники можуть виступати середній рівень інтенсивного показника: а) по території, за якою здійснюється порівняння; б) встановлений для території, прийнятої за стандарт.

Саме стандартні показники найчастіше використовують як вагу та сумірники під час побудови територіальних індексів.

Під час динамічних порівнянь, за ваги в агрегатному індексі цін беруть кількість виробленої продукції у звітному періоді. А під час територіальних порівнянь поняття "звітний період" і "базисний період" мають умовне значення. Якщо порівнювати регіон А з регіоном Б, то базою буде рівень цін в регіоні Б, а за ваги треба брати кількість продукції в регіоні А.

У цьому випадку індекс цін матиме вигляд:

$$I_p = \frac{\sum p_A q_A}{\sum p_B q_A} \quad (16)$$

Проте зовсім не обов'язково порівнювати регіон А з регіоном Б. На тій самій підставі можна порівнювати регіон Б з регіоном А. За такого зіставлення базою буде рівень цін регіону А, а вагою "звітного періоду" – кількість продукції регіону Б. Отже, індекс цін буде мати такий вигляд:

$$I_p = \frac{\sum p_B q_B}{\sum p_A q_B} \quad (17)$$

На занятті студенти повинні виконати практичні завдання щодо поглиблення навичок аналізу територіальних співставлень на підставі індексного методу:

1. Дані про виробництво молока у двох регіонах (табл. 39).

Таблиця 39

Дані про виробництво молока

Порода корів	Регіон А		Регіон Б	
	валовий надій молока, тис. т	жирність молока, %	валовий надій молока, тис. т	жирність молока, %
Лебединська	72	3,5	45	3,2
Симентальська	48	3,1	55	3,4

Визначте: а) територіальні індекси середньої жирності молока змінного та фіксованого складу (беручи за вагу загальний для двох регіонів обсяг молока); б) індекс структурних зрушень, користуючись стандартним процентом жирності молока, який становить 3,4 %.

Поясніть економічний зміст індексів.

2. Маємо дані про собівартість та кількість вироблених електродвигунів ВП-2 та ВП-3 у двох регіонах (табл. 40).

Таблиця 40

Вихідні дані

Вид електродвигуну	Вироблено, тис. од.		Собівартість одиниці продукції, грн	
	Харківський	Дніпропетровський	Харківський	Дніпропетровський
ВТ-2	8	8,3	2 500	2 400
ВТ-3	2	3,1	2 600	2 300

Визначте територіальні індекси змінного та фіксованого складу та індекс структурних зрушень. Зробіть висновки.

Завдання для самостійного опрацювання

1. Маємо дані про роботу підприємства (табл. 41).

Таблиця 41

Вихідні дані

Показники	Базисний період	Звітний період
Середня фактична тривалість робочого періоду, днів	83	84
Середня фактична тривалість робочого дня, год	7,8	7,5
Середній виробіток одного робітника, грн	400	420
Питома вага робітників, %	0,82	0,8

Визначте: вплив середньогодинного виробітку, середньої тривалості робочого дня, середньої кількості днів роботи й частки робітників у загальній чисельності працюючих на зміну середньої продуктивності одного працюючого. Зробіть висновки.

Рекомендована література

Основна

1. Гусаров В. М. Теория статистики / В. М. Гусаров. – Москва : ЮНИТИ-ДАНА, 2003. – 464 с.
2. Єріна А. М. Статистика : навч.-метод. посіб. для самост. вивч. дисц. / А. М. Єріна, Р. М. Моторін, А. В. Головач та ін. ; за заг. ред. А. М. Єріної, Р. В. Моторіна. – Київ : КНЕУ, 2001. – 448 с.
3. Єріна А. М. Статистичне моделювання та прогнозування: навч. посіб. / А. М. Єріна. – Київ : КНЕУ, 2001. – 170 с.
4. Єріна А. М. Теорія статистики : практикум / А. М. Єріна, З. О. Пальян. – Київ : Знання, 2005. – 256 с.
5. Ефимова М. Р. Общая теория статистики : учебник / М. Р. Ефимова, Е. В. Петрова, В. Н. Румянцев. – 2-е изд., испр. и доп. – Москва : ИНФРА-М, 2007. – 416 с.
6. Лугінін О. Є. Статистика : підручник / О. Є. Лугінін. – 2-е вид., перероб. та доп. – Київ : Центр учбової літератури, 2007. – 608 с.
7. Мармоза А. Т. Практикум з теорії статистики / А. Т. Мармоза. – Київ : Ельга, Ніка-Центр, 2003. – 344 с.
8. Мармоза А. Т. Теорія статистики / А. Т. Мармоза. – Київ : Ельга, Ніка-Центр, 2003. – 392 с.
9. Опря А. Т. Статистика (з програмованою формою контролю знань). Математична статистика. Теорія статистики / А. Т. Опря. – Київ : Центр навчальної літератури, 2005. – 472 с.
10. Статистика : навч. посіб. / за ред. О. В. Раєвської. – Харків : ВД "Інжек", 2011. – 504 с.
11. Фещур Р. В. Статистика: теоретичні засади і прикладні аспекти / Р. В. Фещур, А. Ф. Барвінський, В. П. Качур та ін. ; за заг. ред. Р. В. Фещура. – Львів : "Інтелект-Захід", 2003. – 576 с.
12. Щурик М. В. Статистика : навч. посіб. / М. В. Щурик – Львів : "Магнолія-2006", 2009. – 546 с.

Додаткова

13. Бабешко Л. О. Основы эконометрического моделирования: учебное пособие / Л. О. Бабешко. – изд. 3-е, стереотипное. – Москва : КомКнига, 2007. – 432 с.

14. Бек В. Л. Теорія статистики : навч. посіб. / В. Л. Бек. – Київ : Центр навчальної літератури, 2002. – 288 с.

15. Джессен Р. Методы статистических обследований / Р Джессен пер. с англ. Ю. П. Лукашина, Я. Ш. Паппэ ; под ред. Е. М. Четыркина. – Москва : Финансы и статистика, 1985. – 318 с.

16. Кендэл М. Временные ряды / М. Кендэл. – пер. с англ. и предисл. Ю. П. Лукашина. – Москва : Финансы и статистика, 1981. – 358 с.

17. Плошко Б. Г. История статистики / Б. Г. Плошко, И. И. Елисеева. – Москва : Финансы и статистика, 1990. – 296 с.

18. Про державну статистику : Закон України // Голос України. – 1992, № 43.

19. Сєрова І. А. Організація статистичних спостережень : конспект лекцій / І. А. Сєрова, І. В. Асьонова. – Харків : Вид. ХНЕУ, 2008. – 236 с.

20. Теория статистики / под ред. Р. А. Шмойловой. – 3-е изд., перераб. – Москва : Финансы и статистика, 1999. – 464 с.

Інформаційні ресурси

21. Сайт Державної служби статистики України. – Режим доступу : <http://www.ukrstat.gov.ua>.

22. Сайт ГУС в Харківській області. – Режим доступу : <http://kh.ukrstat.gov.ua>.

23. Статистика України : науковий журнал [Електронний ресурс]. – Режим доступу : <http://www.ukrstat.gov.ua>.

24. Статистична звітність емітентів України [Електронний ресурс]. – Режим доступу : <http://www.smida.gov.ua/db/emitent>.

Додатки

Додаток А

Таблиця А. 1

Професійні компетентності, яких набувають студенти після вивчення навчальної дисципліни

Назва компетентності	Складові компетентності
Володіти теоретичними знаннями у галузі статистики, як науки, що пізнає масові процеси та явища	Вивчати й узагальнювати передовий досвід у сфері статистики. Аналізувати проблеми та перспективи сучасної української статистики
Організувати та проводити статистичні спостереження й формувати інформаційну базу статистичного дослідження	Формувати програмно-методологічні та організаційні питання плану статистичного спостереження
Аналізувати ряди розподілу статистичних величин	Перевіряти статистичні гіпотези на основі різноманітних критеріїв
Формувати однорідні групи об'єктів та виявляти зв'язок між статистичними показниками	Виявляти види зв'язку між соціально-економічними явищами. Будувати кореляційно-регресійні моделі за допомогою пакетів прикладних програм. Виявляти стохастичні зв'язки між явищами та аналізувати їх за допомогою непараметричних методів
Визначати тенденції розвитку об'єктів статистичного дослідження	Визначати перспективний стан розвитку об'єкта дослідження (підприємства, організації, регіону та ін.) за допомогою пакетів прикладних програм Statistica 10.0 та MS Excel. Виявляти сезонні коливання під час дослідження соціально-економічних явищ та процесів. Виявляти та усувати автокореляцію в динамічних рядах
Аналізувати та виявляти чинники, що впливають на зміну показників за допомогою індексного методу	Аналізувати якісні та кількісні показники на підставі використання динамічних індексів. Аналізувати просторово-територіальні співставлення соціально-економічних явищ на підставі використання просторових індексів

Зміст

Вступ.....	3
Методичні рекомендації до лабораторних та самостійних робіт	4
Змістовий модуль 1. Методи статистичного спостереження, побудови та аналізу рядів розподілу	4
Тема 1. Організація статистики в Україні. Міжнародні статистичні організації.....	4
Тема 2. Історія статистики	5
Тема 3. Інформаційне забезпечення статистичного дослідження.....	6
Тема 4. Перевірка статистичних гіпотез.....	12
Тема 5. Метод аналітичних групувань.....	20
Тема 6. Регресійний аналіз взаємозв'язку	27
Тема 7. Методи оцінювання стохастичного зв'язку	38
Змістовий модуль 2. Методи і моделі статистичного дослідження рядів динаміки та тенденцій розвитку.....	45
Тема 8. Аналіз тенденцій розвитку.....	45
Тема 9. Аналіз коливань та сталості динамічних рядів	54
Тема 10. Кореляція динамічних рядів	63
Тема 11. Динамічний індексний аналіз.....	66
Тема 12. Просторовий індексний аналіз	75
Рекомендована література.....	80
Основна	80
Додаткова	81
Інформаційні ресурси	81
Додатки.....	82

НАВЧАЛЬНЕ ВИДАННЯ

СТАТИСТИКА II

**Методичні рекомендації
до лабораторних та самостійних робіт
для студентів спеціальності 051 "Економіка"
першого (бакалаврського) рівня**

Самостійне електронне текстове мережеве видання

Укладачі: **Аксьонова** Ірина Вікторівна
Бровко Ольга Іванівна
Свидло Ганна Ігорівна

Відповідальний за видання *О. В. Раєвська*

Редактор *О. В. Анацька*

Коректор *Н. В. Грінченко*

План 2020 р. Поз. № 106 ЕВ. Обсяг 84 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.*