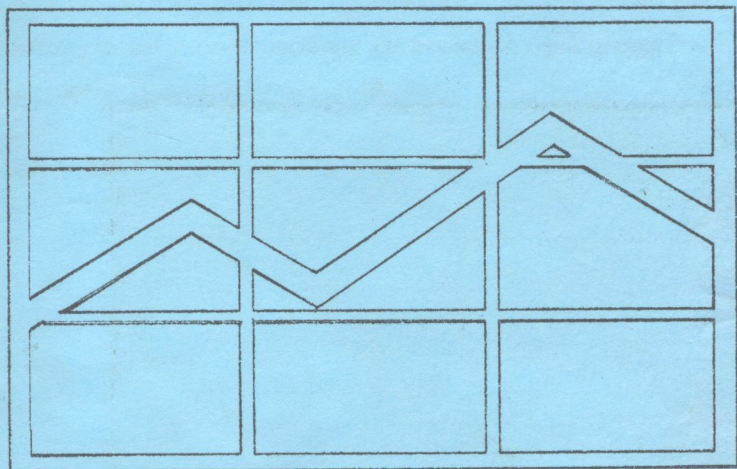


С60я73
Г75

Б.Є. Грабовецький

ЗАГАЛЬНА ТЕОРІЯ СТАТИСТИКИ



МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

Вінницький державний технічний університет

НТБ ВДТУ



3198-6

С60я73

Г 75

2001

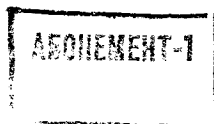
Бібліотека: 01. Кабінет на території університету

Б.Є. ГРАБОВЕЦЬКИЙ

ЗАГАЛЬНА ТЕОРІЯ СТАТИСТИКИ

Затверджено Ученою радою Вінницького державного технічного університету як навчальний посібник для студентів економічних спеціальностей. Протокол №9 від 27.04.2001р.

Вінниця ВДТУ 2001



РЕЦЕНЗЕНТИ:

О.В.Мороз, доктор економічних наук, професор
К.Д.Якимчук, доктор економічних наук, професор
В.В.Зянько, кандидат економічних наук, доцент

Рекомендовано до видання Ученою радою Вінницького державного технічного університету Міністерства освіти і науки України

Грабовецький Б.С.

Г 75 Загальна теорія статистики: Навчальний посібник. - Вінниця: ВДГУ, 2001, - 147 с.

Навчальний посібник складений у відповідності з навчальною програмою курсу "Загальна теорія статистики". Містить описання основних категорій і понять теорії статистики як складової частини комплексної науки "Статистика".

Детально подані такі методи статистики, як групування статистичних даних; абсолютні, відносні і середні статистичні величини; ряди динаміки; індекси; аналіз кореляцій і регресії; вибіркові спостереження.

Наведені приклади розв'язання типових задач, подані задачі до кожного розділу для самостійного розв'язування.

Розрахований на студентів економічних спеціальностей всіх форм навчання.

ЗМІСТ

ПЕРЕДМОВА.....	6
РОЗДІЛ 1 ПРЕДМЕТ І МЕТОД СТАТИСТИЧНОЇ НАУКИ.....	7
1.1 Предмет статистики.....	7
1.2 Метод статистики.....	8
1.3 Категорії, поняття і терміни статистичної науки.....	9
Питання для самоконтролю.....	10
РОЗДІЛ 2 СТАТИСТИЧНЕ СПОСТЕРЕЖЕННЯ.....	11
2.1 Поняття і класифікація статистичних спостережень.....	11
2.2 Планування статистичного спостереження.....	14
2.3 Помилки статистичного спостереження та методи їх усунення.....	16
Питання для самоконтролю.....	17
РОЗДІЛ 3 ЗВЕДЕННЯ ТА ГРУПУВАННЯ СТАТИСТИЧНИХ ДАНИХ.....	17
3.1 Поняття та види зведення.....	17
3.2 Поняття, види і правила побудови статистичних групувань.....	19
3.3 Вторинне групування.....	24
Питання для самоконтролю.....	26
Задачі.....	26
РОЗДІЛ 4 АБСОЛЮТНІ І ВІДНОСНІ СТАТИСТИЧНІ ВЕЛИЧИНИ.....	28
4.1 Поняття і види абсолютних величин.....	28
4.2 Поняття, види та методи обчислення відносних величин.....	31
Питання для самоконтролю.....	33
Задачі.....	33

РОЗДІЛ 5 СЕРЕДНІ ВЕЛИЧИНИ І ПОКАЗНИКИ ВАРІАЦІЇ.....	37
5.1 Статистична середня та її суть.....	37
5.2 Види середніх величин і способи їх обчислення.....	38
5.3 Вибір виду середньої.....	39
5.4 Мода і медіана	43
5.5 Поняття і показники варіації.....	45
Питання для самоконтролю.....	52
Задачі.....	53
РОЗДІЛ 6 РЯДИ ДИНАМІКИ.....	58
6.1 Поняття та види рядів динаміки.....	58
6.2 Правила формування рядів динаміки.....	60
6.3 Статистичні характеристики рядів динаміки.....	61
6.4 Основні тенденції динамічного ряду та методи їх виявлення.....	67
6.5 Статистичне вивчення сезонних коливань.....	70
Питання для самоконтролю.....	72
Задачі.....	72
РОЗДІЛ 7 ВИБІРКОВЕ СПОСТЕРЕЖЕННЯ.....	76
7.1 Поняття і теоретичні основи вибіркового спостереження.....	76
7.2 Схеми і способи відбору одиниць із генеральної сукупності.....	78
7.3 Середня і гранична помилки вибірки.....	80
7.4 Визначення необхідної чисельності вибірки.....	86
7.5 Способи розповсюдження результатів вибірових спостережень на генеральну сукупність.....	88
Питання для самоконтролю.....	88
Задачі.....	89

РОЗДІЛ 8 ІНДЕКСИ.....	92
8.1 Визначення, класифікація і правила побудови індексів.....	92
8.2 Середні індекси.....	101
8.3 Індеси середніх величин.....	104
8.4 Ланцюгові і базисні індекси.....	107
8.5 Індексний факторний аналіз.....	109
Питання для самоконтролю.....	118
Задачі.....	119
РОЗДІЛ 9 СТАТИСТИЧНІ МЕТОДИ ВИВЧЕННЯ	
ВЗАСМОЗВ'ЯЗКІВ.....	126
9.1 Задачі статистики у вивченні і вимірюванні зв'язків.	
Види зв'язків.....	126
9.2 Методи кореляції і регресії.....	128
9.3 Система параметрів і характеристик рівняння регресії.....	131
9.4 Вирівнювання рядів динаміки. Часовий тренд.....	138
9.5 Непараметричні показники зв'язку.....	140
Питання для самоконтролю.....	143
Задачі.....	143
СПИСОК ЛІТЕРАТУРИ.....	146

ПЕРЕДМОВА

Курс “Загальна теорія статистики”, як складова частина комплексної дисципліни “Статистика”, є однією з фундаментальних економічних дисциплін і вивчається студентами всіх економічних спеціальностей. Методологія цього курсу використовується майже в усіх науках як суспільних, так і природничих.

Методологія “Економічної статистики”, а також галузевих статистик, базується на методах “Загальної теорії статистики”.

Загальна теорія статистики покликана озброїти студентів знаннями, які мають не тільки пізнавальне, але і велике практичне значення. Опанування інших економічних дисциплін неможливе без знання методології загальної теорії статистики.

В інших економічних дисциплінах широко використовуються поняття, терміни, категорії, показники, методика загальної теорії статистики.

Засвоєння дисципліни дозволяє студентам на підставі інформації порівнювати явища між собою в часі і просторі, обчислювати узагальнюючі показники масових явищ (середні, відносні величини, варіації, індекси та ін.), вивчати тенденції та прогнозувати явища і процеси, будувати таблиці, графіки, робити висновки та давати пропозиції на основі результатів статистичного аналізу.

Специфічний характер статистики, її показників та методології вимагає від студентів вивчення цієї науки послідовно і систематично, засвоєння логічного змісту статистичних показників і методів.

В процесі вивчення курсу важливо не тільки засвоєння теоретичних положень, але і розв’язання задач, які подані в кінці кожного розділу, що дозволить набути вміння та практичних навиків обчислення статистичних показників, аналізу статистичних даних і, що дуже суттєво, дозволить успішно виконувати контрольні роботи.

РОЗДІЛ І ПРЕДМЕТ І МЕТОД СТАТИСТИЧНОЇ НАУКИ

1.1 Предмет статистики

В науку термін “статистика” був впроваджений німецьким ученим, професором Геттингенського університету Г.Ахенвальдом. І хоча статистичний облік існує з глибокої давнини, статистика як наука виникла тільки в XVII ст.

Розвиток статистичної науки, постійне розширення сфери її використання призвели до трансформації змісту самого поняття “статистика”.

В сучасному розумінні “статистика” використовується у трьох значеннях:

а) під статистикою розуміють галузь практичної діяльності, яка має за мету збір, обробку, аналіз і публікацію даних про різні явища суспільного життя (в цьому тлумаченні “статистика” виступає як синонім словосполучення “статистичний облік”);

б) статистикою називають цифровий матеріал, який служить для характеристики будь-якої області суспільних явищ, або територіальний розподіл певного показника;

в) статистикою називають галузь знань, наукову дисципліну, яка вивчається у вищих і середніх спеціальних навчальних закладах.

Як всяка самостійна наука статистика має свій предмет, тобто те, що вона вивчає; свій метод, тобто за допомогою яких засобів вона досліджує свій предмет, і нарешті, систему категорій, понять, термінів.

Предметом статистики є кількісна сторона масових суспільних явищ у нерозривному зв'язку з їх якісним змістом.

Статистика – комплексна, багатогалузева наука, яка складається із окремих самостійних дисциплін, що тісно пов'язані між собою, зокрема:

а) загальна теорія статистики, де розглядаються категорії статистичної науки, а також спільні для будь-яких масових явищ методи і засоби аналізу;

б) економічна статистика, яка, використовуючи методологію загальної теорії статистики, вивчає явища і процеси, що мають місце в економіці, розробляє і вдосконалює систему економічних показників, методи вивчення народного господарства чи регіону як єдиного цілого;

в) галузеві статистики (промисловості, сільського господарства, транспорту та ін.), які, опираючись на загальнотеоретичні положення статистичної науки, розробляють зміст і методи обчислення показників, що відображають особливості кожної окремої галузі;

г) соціальна статистика, предметом якої є вивчення соціальних умов і характеру праці, рівня життя, доходів, споживання матеріальних благ і послуг населення.

Як суспільна наука статистика тісно пов'язана з економічною теорією і соціологією, а також практично з усіма прикладними економічними дисциплінами.

1.2 Метод статистики

Метод статистики – це сукупність (система) спеціальних засобів, дій, способів, операцій, що базується на діалектико-логічних, формально-логічних і загальнонаукових принципах.

Арсенал статистичних методів достатньо широкий, що дозволяє здійснити групування статистичних даних, оцінити сталість і варіацію явищ, визначити динаміку та інтенсивність розвитку явищ, обчислити сукупність індексів і дослідити закономірності змін явищ тощо. Статистичні методи тісно пов'язані з математикою. У них спільні методи обробки і оцінки даних, проте різні предмети пізнання.

Математична статистика вивчає закономірності масових явищ в абстрактній формі; статистика як суспільна наука досліджує розміри і співвідношення суспільних явищ у конкретних умовах їх існування та розвитку.

2.1 Категорії, поняття і терміни статистичної науки

Для дослідження кількісних характеристик масових суспільних явищ статистика використовує систему категорій, понять, термінів, зокрема:

Статистична сукупність – це безліч реально існуючих предметів, явищ, процесів, осіб, які є однаковими по одній або декількох ознаках. Специфічною властивістю статистичної сукупності є їх масовість.

Одиниця сукупності – це первинний елемент сукупності, із яких власне створюється сукупність, і кожен із них відображає якісну однорідність певної сукупності.

Одиниці статистичної сукупності повинні бути носіями ряду властивостей, або **ознак**, тобто таких властивостей, які можуть бути встановлені і виміряні. Наприклад, чисельність працюючих, вартість основних фондів, обсяг виробництва тощо.

Ознаки бувають **кількісні і якісні (атрибутивні)**. До кількісних відносяться ознаки, які мають числове вираження. Наприклад, вага певного виробу, урожайність певної культури, вік працівника тощо.

Якісні ознаки не мають кількісного вираження, а характеризуються специфічною особливістю. Наприклад, освіта, професія, стать працівників;

істотні (головні) ознаки, які відображають змістовну сторону явищ і тому повинні підлягати статистичному дослідженню, і **неосновні (несуттєві, другорозрядні)**:

первинні, які встановлюються безпосередньо в процесі реєстрації властивостей об'єкта, і **вторинні**, які визначаються в процесі додаткової обробки первинних даних.

Окремі значення ознаки називають **варіантами**.

Ознаки – дещо окремі, розрізнені фактори. Для їх узагальнення використовуються статистичні показники, з допомогою яких формується, передається і зберігається статистична інформація.

Статистичний показник – це узагальнююча характеристика суспільних явищ і процесів, в якій поєднується їх кількісна і якісна визначеність.

В складі категорій статистичної науки особливе місце належить статистичній закономірності.

Статистичними називаються закономірності масових процесів, які виражають у середньому результат взаємодії значного числа однорідних явищ або взаємозв'язок послідовних станів системи. Це, по суті, повторювання, послідовність і порядок зміни явищ.

Природа статистичної закономірності може бути досліджена достатньо повно лише враховуючи дію **закону великих чисел**, в основі якого лежать властивості масовості і причинної обумовленості явищ.

Питання для самоконтролю

1. Визначіть предмет і метод статистики.
2. Назвіть основні самостійні дисципліни статистичної науки.
3. Наведіть приклад статистичної сукупності і одиниці сукупності.
4. В чому відмінність якісних і кількісних ознак одиниці сукупності?
5. Дайте визначення статистичного показника.

РОЗДІЛ 2 СТАТИСТИЧНЕ СПОСТЕРЕЖЕННЯ

2.1 Поняття і класифікація статистичних спостережень

Статистичне дослідження здійснюється на основі масових первинних даних, які отримують шляхом статистичного спостереження. Таким чином, статистичне спостереження – перший етап будь-якого статистичного дослідження.

Статистичне спостереження – це планомірна, науково-організована, систематична реєстрація даних про явища і процеси суспільного життя.

Основні вимоги до зібраних в результаті спостереження даних: **повного, точність, достовірність, своєчасність, порівнянність.**

В статистичній практиці застосовуються різні **форми, види і способи статистичного спостереження.**

Класифікація статистичних спостережень наведена на рис.2.1.

Звітність буває: загальнодержавна і відомча; типова і спеціалізована; централізована і децентралізована.

Спеціально організоване спостереження охоплює первинні дані, які не містяться у звітності (наприклад, перепис населення).

Реєстрове спостереження – це форма неперервного статистичного спостереження за довгостроковими процесами, які мають початкову і кінцеву стадію розвитку (наприклад, успішність студентів на протязі строку їх навчання).

Суцільне спостереження передбачає реєстрацію усіх одиниць сукупності.

Несуцільне спостереження передбачає реєстрацію не усіх одиниць сукупності, а лише їх частину, що відбирається в певному порядку.

Вибіркове спостереження – це реєстрація частини одиниць сукупності, які відбираються у випадковому порядку (наприклад, обстеження суб'єктів малого бізнесу).



Рисунок 2.1 - **Класифікація статистичних спостережень**

Спостереження основного масиву – це реєстрації даних таких одиниць сукупності, які мають визначальне значення для характеристики певного процесу (наприклад, середні ціни на ринку по країні визначаються шляхом обстеження цін в найбільш великих населених пунктах).

Монографічне спостереження – це детальне обстеження окремих одиниць сукупності з метою їх глибокого ідеального вивчення (наприклад, при обстеженні продуктивності худоби в одному або декількох господарствах детально вивчаються усі чинники, які визначають рівень продуктивності продукції).

Анкетне обстеження передбачає одержання від респондентів відомостей, які записуються в анкету, що попередньо висилається за відповідними адресами.

Моніторинг – це спеціально організоване спостереження за станом певного явища (наприклад, стан водоймищ, ціни на ринку).

Поточне спостереження – це реєстрація фактів в міру їх виникнення (наприклад, виробництво або реалізація продукції, надходження платежів).

Періодичне – це спостереження, яке проводиться, як правило, через рівні проміжки часу (наприклад, перепис основних фондів, перепис населення).

Одноразове – це спостереження, яке здійснюється за потреби дослідження певних явищ (наприклад, стадії життєвого циклу вибору).

Безпосередній облік – це реєстрація фактів, що здійснюється особою, яка проводить її шляхом підрахування, вимірювання, оцінки, огляду.

Документальний облік – це реєстрація даних, які містяться в документах первинного обліку (наприклад, складання звітностей).

Опитування – це спостереження, яке здійснюється шляхом опитування респондентів.

Експедиційне – це спостереження, при якому інформація реєструється в спеціальних бланках зі слів респондентів (наприклад, перепис населення).

Самореєстрація – це реєстрація фактів самими респондентами після проведення відповідного інструктажу (наприклад, обстеження сімейних бюджетів).

Кореспондентське опитування – це реєстрація фактів на місцях їх виникнення добровільно відібраними особами з подальшим надсиланням результатів у відповідні органи (наприклад, реєстрація побажань споживачів щодо якості товарів).

Анкетне опитування – це реєстрація думок, намірів і мотивів респондентів шляхом самостійного заповнення ними спеціально оформлених анкет (наприклад, наміри суб'єктів ринку щодо придбання товарів тривалого використання).

2.2 Планування статистичного спостереження

Проведення будь-якого статистичного спостереження починається із розробки плану статистичного спостереження, який складається із двох розділів: програмно-методологічного і організаційного (рисунок 2.2).

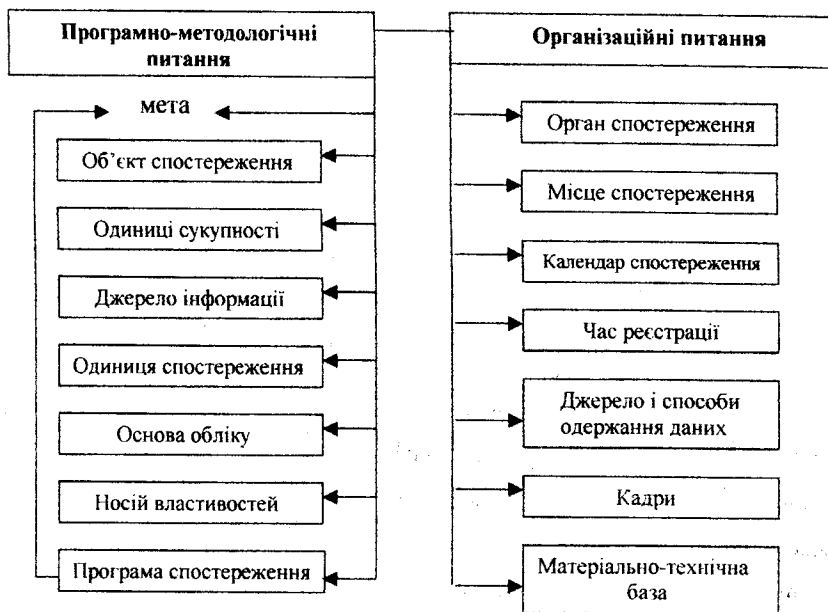


Рисунок 2.2 - Структура плану статистичного спостереження

Об'єкт спостереження – це сукупність одиниць досліджуваного явища, що підлягають обстеженню.

Одиниця сукупності – первинний елемент об'єкта, який є носієм інформації, що необхідно реєструвати.

Одиниця спостереження – це первинна одиниця, від якої повинні бути одержані необхідні дані.

Запитання програми спостереження фіксуються в **статистичних формулярах (бланках)**. Вони мають форму статистичного звіту, анкети, карти, бланку, переписного або опитувального листа. Формуляри бувають індивідуальні (карткові) і спискові.

2.3 Помилки статистичного спостереження та методи їх усунення

Одержана в результаті статистичного спостереження інформація повинна підлягати ретельному контролю з метою виявлення і усунення помилок.

Для перевірки точності і достовірності матеріалів спостереження застосовується арифметичний і логічний контроль.

Арифметичний контроль – це перевірка зібраних даних шляхом проведення перерахунків, оцінки узагальнюючих показників.

Логічний контроль полягає в порівнянні планових (нормативних) даних з фактичними даними, а також з даними за попередні періоди. В деяких випадках логічний контроль базується на здоровому глузді: чи реально можливе або неможливе певне значення показника, одержаного в результаті статистичного спостереження.

Розбіжності або похибки спостереження називаються **помилками**.

Залежно від причини виникнення розрізняють помилки реєстрації і репрезентативності.

Помилки реєстрації виникають внаслідок неправильного встановлення фактів в процесі статистичного спостереження.

Помилки репрезентативності виникають лише при вибіркових спостереженнях як результат порушення принципів випадковості відбору.

За природою виникнення розрізняють випадкові і систематичні помилки.

Випадкові помилки виникають внаслідок збігу випадкових обставин – через неуважність, недбалість, некваліфікованість реєстратора.

Оскільки такі помилки допускаються в одну і іншу сторону, то вони врівноважуються і не впливають на кінцевий результат.

Систематичні помилки, як правило, допускаються лише в одному напрямі, що призводить до збільшення або зменшення кінцевих результатів.

Систематичні помилки бувають ненавмисними або навмисними.

Ненавмисні помилки обумовлюються недосконалістю програми спостереження, некомпетентністю або неуважністю реєстратора, неосвіченістю респондентів.

Навмисні помилки виникають внаслідок свідомого викривлення фактів з метою подання дійсності в прикрашеному або погіршеному стані.

Питання для самоконтролю

1. В чому полягає суть і призначення статистичного спостереження?
2. Назвіть способи статистичного спостереження.
3. Наведіть приклад об'єкта спостереження, одиниці сукупності і одиниці спостереження.
4. Назвіть складові частини плану статистичного спостереження.
5. Якими методами здійснюється контроль матеріалів спостереження?
6. Які помилки можливі при проведенні статистичного спостереження?

РОЗДІЛ 3 ЗВЕДЕННЯ ТА ГРУПУВАННЯ СТАТИСТИЧНИХ ДАНИХ

3.1 Поняття та види зведення

Одержаний в результаті статистичного спостереження розрізнений і несистематизований матеріал потребує певної обробки, систематизації,

зведення розрізних даних воедино.

Тому другим етапом статистичної роботи є зведення і групування статистичних даних.

Зведення – це систематизація первинних даних з метою одержання узагальнюючих характеристик досліджуваного явища за рядом суттєвих для нього ознак.

Призначенням зведення є одержання на основі зведених даних узагальнюючих статистичних показників, які відображають суть і закономірності розвитку соціально-економічних явищ.

Статистичне зведення розрізняється за складністю побудови, формою обробки і технікою виконання.

За складністю побудови розрізняють прості і складні зведення.

Просте зведення передбачає одержання загальних підсумків усіх одиниць сукупності.

Складне зведення передбачає елементи групування, розрахунок групових і загальних підсумків, а також подання результатів зведення у вигляді статистичних таблиць і графіків.

За формою обробки розрізняють централізоване і децентралізоване зведення.

При **централізованому зведенні** первинний матеріал статистичного спостереження обробляють і систематизують у центральному органі державної статистики.

При **децентралізованому зведенні** усі операції з узагальнення даних спостереження виконують на місцях, а в центральний орган надсилають лише зведені матеріали.

За технікою або способом виконання зведення може бути **річним і механізованим (автоматизованим)**.

Зведення статистичних даних не дозволяє виявити важливі властивості і основні типи явищ, що досліджуються. Тому для посилення

аналітичних можливостей інформації, одержаної в результаті спостереження, використовується групування статистичних даних.

3.2. Поняття, види і правила побудови статистичних групувань

Групування – це розподіл одиниць сукупності на окремі групи за істотними, суттєвими для них ознаками. Ознаки, за якими проводиться розподіл одиниць певної сукупності на групи, називаються **групувальними ознаками**.

Ознаки бувають **кількісними і якісними; результативними і факторними; альтернативними і варіаційними**.

На групування в статистичному аналізі покладаються певні функції, зокрема:

- а) визначення структури і структурних зрушень;
- б) типологія соціально-економічних явищ;
- в) аналіз причинно-наслідкових взаємозв'язків між явищами.

Відповідно до цих функцій розрізняють структурні, типологічні і факторні групування.

Структурне групування дозволяє вивчити склад однорідної сукупності, а також проаналізувати структурні зрушення (наприклад, структура посівної площі фермерського або кооперативного господарства).

Типологічне групування дозволяє виділити найхарактерніші групи, типи явищ, з яких складається неоднорідна статистична сукупність, визначити істотні відмінності між ними, а також ознаки, які є спільними для усіх груп (наприклад, класовий склад населення).

Факторні групування відображають причинно-наслідкові взаємозв'язки між факторними і результативними ознаками (наприклад, змінний виробіток робітника-верстатника залежно від стажу роботи).

Статистичне групування починається з добору групувальної ознаки і утворення груп.

Ознаки, які приймають різні кількісні значення, називаються **кількісними**, а ознаки, які приймають різні якісні значення, наприклад, професії, суспільні групи населення, називаються **якісними (атрибутивними)**.

Тому розрізняють групування за кількісною ознакою і групування за якісною ознакою.

Групування проводяться за однією ознакою (**просте групування**) або за кількома ознаками (**складне групування**). Останнє може бути **комбінаційним**, якщо в його основі послідовно скомбіновано дві або більше ознак, або **багатомірним**, якщо воно проводиться за кількома ознаками одночасно. Наприклад, групування населення за віком і статтю буде комбінаційним, якщо в кожній віковій групі виділені підгрупи за статтю. А групування сімей за рівнем споживання продовольчих і непродовольчих товарів одночасно буде багатомірним.

Основу простого групування становить ряд розподілу, який складається із двох елементів: варіант і частот.

Варіанти – це конкретні значення групувальної ознаки; частоти – це числа, які показують, скільки разів повторюються окремі значення варіантів. Відносні частоти (% до підсумку) називаються **частками**; накопичену частоту (частку) називають **кумулятивною**.

Розподіл за ознаками може бути кількісним або якісним (атрибутивним).

Ряд розподілу одиниць сукупності за кількісною ознакою називається **варіаційним рядом розподілу**.

За способом побудови варіаційні ряди бувають дискретними і інтервальними.

Дискретні ряди – це такі ряди, де значення варіантів подані окремими числами.

Інтервальні ряди – це такі ряди, де значення варіантів подані у вигляді інтервалів.

Інтервал – це проміжок між значеннями двох кількісних ознак, в межах якого усі значення ознаки відносяться до однієї групи. Інтервал представляє собою проміжок між максимальним і мінімальним значеннями ознаки в групі.

Інтервали бувають рівні і нерівні (зростаючі і спадаючі), відкриті і закриті.

Відкритий – це інтервал, у якого є тільки одна межа: верхня – у першому, нижня – у останньому.

Закритий – це інтервал, у якого є нижня і верхня межа.

Зустрічаються також групування, де може бути відкритий тільки верхній або нижній інтервал.

Процес групування статистичної сукупності з рівними інтервалами передбачає визначення кількості груп (рядів) і величини інтервалу, для чого використовуються такі формули:

$$i = \frac{X_{\max} - X_{\min}}{k}, \quad (3.1)$$

$$i = \frac{X_{\max} - X_{\min}}{1 + 3,322 \lg n}, \quad (3.2)$$

де i – величина інтервалу;

X_{\max} , X_{\min} – відповідно максимальне і мінімальне значення ознаки в сукупності;

k – кількість груп;

n – кількість одиниць сукупності.

Правила побудови групувань

Перш ніж побудувати групування необхідно скласти таблицю з двох колонок: в першу колонку записуються варіанти групувальної ознаки, а в другу – частоти (при необхідності додатково виділяють колонки для частки і кумулятивної частоти).

Для складання дискретного групування необхідно в першій колонці записати послідовно усі значення варіант по мірі їх зростання або зменшення, а в другій колонці записати частоти, тобто скільки разів повторюється відповідне значення певної ознаки.

Приклад. За наведеними даними провести дискретне групування студентів групи за віком і зазначити, який вік є найбільш типовим:

21	20	21	23	24	23	24	24	20	25
22	24	23	21	25	20	24	25	21	20
20	22	25	23	22	20	21	23	22	20

У таблиці 3.1 за наведеними даними показані результати групування, здійсненого згідно з зазначеним вище правилом. В графі 1 показаний вік студентів по мірі зростання; в графі 2 – кількість студентів певного віку; в графі 3 – частка кількості студентів кожного віку в загальній чисельності студентів.

Таблиця 3.1 - Групування студентів за віком

Варіанта (вік студентів), x	Частота (кількість студентів), f	Частка, % f''
1	2	3
20	7	23,3
21	5	16,7
22	4	13,3
23	5	16,7
24	5	16,7
25	4	13,3
Разом	30	100,0

Найбільш типовим віком студентів групи є 20 років, про що свідчать максимальні значення частоти і частки.

Для побудови інтервального групування необхідно:

а) визначити величину інтервалу і кількість груп за формулами 3.1, 3.2;

б) в першу групу (перший рядок) за нижню межу значення варіанти прийняти мінімальне число сукупності; додавши до останнього величину інтервалу, отримуємо верхню межу значення варіанти першої групи;

в) верхня межа значення варіанти першої групи стає нижньою межею значення варіанти другої групи; додавши до останнього величину інтервалу отримуємо верхню межу значення варіанти другої групи і аналогічно для інших груп;

г) визначити кількість одиниць сукупності, які входять в кожен групу;

д) в кожному ряд включати варіанти до верхньої межі значення: якщо є дані на рівні верхньої межі значення інтервалу, то вони включаються в наступну групу (рядок).

Розглянемо приклад складання інтервального групування.

Є такі дані про виробіток продукції в розрахунку на одного працівника за рік (тис. грн.):

7,5	5,4	7,4	6,0	6,7	9,5	8,6	5,8	5,1	6,4
6,2	6,8	5,5	8,1	5,6	6,8	9,3	6,2	8,8	7,5

Необхідно побудувати інтервальний варіаційний ряд з рівними інтервалами.

Виходячи з того, що в формулах 3.1 і 3.2 співпадають результати і чисельники, то однаковими повинні бути і знаменники, тобто:

$$k = 1 + 3,322 \lg n \quad (3.3)$$

Звідси:

$$k = 1 + 3,322 \lg 20 = 1 + 3,322 \cdot 1,3 = 5$$

$$i = \frac{X_{\max} - X_{\min}}{k} = \frac{9,5 - 5,1}{5} = 0,9$$

Значення "і" приймається в тих же одиницях виміру, що і початкові дані.

У таблиці 3.2 наведено інтервальне групування даних з використанням зазначених вище етапами і правилами побудови.

Таблиця 3.2 - Групування працівників за річним виробітком продукції

Інтервал групування, тис. грн., х	Кількість працівників, f	Частка, f'	Кумулятивна частота, f''
1	2	3	4
5,1 - 6,0	5	25,0	5
6,0 - 6,9	7	35,0	12
6,9 - 7,8	3	15,0	15
7,8 - 8,7	2	10,0	17
8,7 - 9,6	3	15,0	20
Разом	20	100	—

Кумулятивну частоту отримують шляхом послідовного накопичення частот.

Формально кумулятивні частоти можуть бути визначені за такими формулами:

$$f_1'' = f_1; f_i'' = f_{i-1}'' + f_i \text{ (для } i > 1); \sum f_i = f_n'' \text{.} \quad (3.4)$$

3.3. Вторинне групування

В економічних дослідженнях інколи доводиться вдаватися до повторного групування. Така потреба виникає в тому випадку, коли первинне групування не задовольняє дослідника.

Вторинне групування проводиться, як правило, в таких напрямках:

- розширення або, навпаки, звуження величини інтервалу;

- об'єднання декількох групувань в одне, причому розміри інтервалів в них можуть бути однакові або різні за величиною.

При проведенні будь-якого групування, в тому числі і вторинного, слід дотримуватися такого правила: розмір інтервалів повинен бути постійним (рівним), або поступово збільшуватися чи зменшуватися.

За даними таблиці 3.3 розглянемо приклад проведення вторинного групування шляхом укрупнення інтервалів.

Порядок розрахунку нових частот (кількості підприємств) наведений в графі 4.

Таблиця 3.3 - Групування підприємств регіону за чисельністю працюючих

Первинне групування		Вторинне групування	
Чисельність працюючих, чол.	Кількість підприємств	Чисельність працюючих, чол.	Кількість підприємств
1	2	3	4
До 100	4	До 200	$4+0,5\cdot 16=12$
100 – 300	16	200 – 500	$0,5\cdot 16+2/7\cdot 35=18$
300 – 1000	35	500 – 1000	$5/7\cdot 35=25$
1000 – 2000	28	1000 – 3000	$28+1/3\cdot 12=32$
2000 – 5000	12	3000 і більше	$2/3\cdot 12+5=13$
5000 і більше	5	—	—
Разом	100	Разом	100

Питання для самоконтролю

- 4.2. Наведіть види статистичних групувань.
- 4.2. Які види інтервалів вам відомі?
- 4.2. В якій послідовності будуються дискретні і інтервальні групування (з різними інтервалами).

4.2. В чому полягає суть вторинного групування і в яких випадках до нього вдаються?

Задачі

3.1. За день роботи на обслуговуванні фірма продала чоловічі чоботи таких розмірів:

42	43	41	38	39	44	42	41	39	41
42	39	42	40	40	40	39	43	41	41
42	41	44	43	42	42	45	42	39	41
41	43	42	42	40	40	43	40	42	42
41	38	44	40	43	40	40	44	41	45

Побудувати за розміром взуття дискретний варіаційний ряд розподілу.

3.2. Якість ґрунту в господарствах однієї з областей характеризується такими балами:

65	62	86	82	88	90	75	74	68	76
77	81	79	76	86	70	80	72	90	68
73	80	70	76	74	66	88	82	77	92
70	81	79	79	76	86	70	77	62	78

Побудувати інтервальний ряд розподілу господарств за якістю ґрунту з рівними інтервалами.

3.3. У районах області встановлено таку кількість телефонних апаратів:

135	175	162	191	123	182	167	142	124	178
163	170	145	164	152	198	162	156	156	126
154	138	162	131	130	175	132	160	160	170

Здійснити групування районів за кількістю встановлених телефонних апаратів, утворивши такі групи: до 130, 130-150, 150-170, 170-190, 190 і більше.

3.4. Провести вторинне групування підприємств галузі А відповідно з групуванням підприємств галузі Б.

Галузь А		Галузь Б	
Група підприємств за чисельністю робітників, чол.	Питома вага, %	Група підприємств за чисельністю робітників, чол.	Питома вага, %
До 50	1,0	До 100	4,3
50 – 70	1,0	100 – 200	18,4
70 – 100	2,0	200 – 300	19,5
100 – 150	10,0	300 – 500	28,0
150 – 250	20,0	понад 500	29,8
250 – 400	20,0		
400 – 500	22,0		
понад 500	24,0		
Разом	100,0	Разом	100,0

3.5. Користуючись методом вторинного групування, провести перегрупування заводів за обсягом товарної продукції, створивши для цього три групи: дрібні – до 2,5 млн. грн., середні – 2,5-12 млн. грн., великі – від 12 млн. грн. і більше за такими даними:

Обсяг товарної продукції, млн. грн.	До 1,0	1,0-2,0	2,0-3,0	3,0-5,0	5,0-10,0	10,0-20,0	20,0 і більше	Разом
Кількість заводів у % до підсумку	7,0	10,0	24,0	21,0	15,0	52,3	13,6	100,0

3.6. Є такі дані про розподіл заводів за вартістю основних фондів:

Групи заводів за вартістю основних фондів, млн. грн.	1-3	3-5	5-10	10-30	30-50	Разом
Кількість заводів у % до підсумку	4,6	13,6	15,9	52,3	13,6	100,0

Користуючись методом вторинного групування, створити групи заводів за такою вартістю основних фондів, млн. грн.: 1-3, 3-5, 5-10, 10-15, 15-20, 20-30, 30-50.

3.7. Нижче наведені дані про виконання плану товарообігу магазинами району:

Групи магазинів за рівнем виконання плану, %	90-95	95-100	100-102	102-105	105-110	110-115	115-120	120-125	125-130	Разом
Кількість магазинів	2	3	2	4	2	3	5	4	2	27

Провести укрупнення інтервалів цього групування, створивши не більше п'яти груп.

РОЗДІЛ 4 АБСОЛЮТНІ І ВІДНОСНІ СТАТИСТИЧНІ ВЕЛИЧИНИ

4.1 Поняття і види абсолютних величин

В результаті статистичного спостереження економічних явищ і систематизації даних одержують абсолютні статистичні величини.

Абсолютні величини відображають кількісні сторони економічних явищ і процесів.

За способом вираження розмірів явищ і процесів розрізняють індивідуальні, групові і загальні абсолютні величини.

Індивідуальні – це абсолютні величини, які відображають розміри кількісних ознак у окремих одиниць сукупності.

Групові і підсумкові абсолютні величини відображають розмір ознак відповідно у окремих груп або в усіх одиницях сукупності.

Абсолютні величини – це іменовані числа; вони завжди мають одиницю виміру.

Розрізняють такі одиниці виміру абсолютних величин:

- **натуральні** – тони, метри, літри та інші фізичні одиниці виміру;

- **грудові** – людино-години, людино-дні;

- **вартісні** – коп., грн., тис. грн. і т.д.;

- **комплексні** – т-км.

До складу натуральних включають також **умовно-натуральні вимірники**, які використовуються у разі потреби звести воедино декілька різновидів однієї споживчої вартості, наприклад, умовне паливо, консерви в тисячах умовних банках (туб) тощо.

Перерахунок в умрвні одиниці здійснюється за допомогою спеціальних коефіцієнтів-сумірників. Роль загальної міри, еталону для розрахунків і порівняння виконус один різновид, коефіцієнт-сумірник для якого дорівнює одиниці ($k = 1$).

Обсяг продукції в умовно-натуральному вимірі визначається за формулою:

$$y = \sum k \cdot x \quad (4.1.)$$

де y – обсяг продукції в умовно-натуральному вимірі;

k – коефіцієнт-сумірник;

x – виробництво продукції певного виду у натуральному вимірі.

Коефіцієнти-сумірники або задаються в умові задачі, або розраховуються як співвідношення значення величини заданого параметра (характеристики) певного виду продукції до значення того ж параметра (характеристики) еталонного зразка.

Розглянемо методику визначення обсягу продукції, в умовно-натуральному вимірі на прикладах.

Приклад 1. В таблиці 4.1 наведені дані про надходження молочної продукції в торговельну мережу.

Таблиця 4.1 - Поставки молочної продукції в торговельну мережу

Види продукції	Кількість, т (x)	Коефіцієнт-сумірник (к)
Молоко 3,2%	144,0	1,0
Молоко 6,0%	107,0	2,0
Ряжанка	6,2	2,0
Сметана	113,0	8,5

Необхідно визначити обсяг надходження молочної продукції в торговельну мережу в умовно-натуральному вимірі.

Згідно з формулою 4.1:

$$y = \sum kx = 1 \cdot 144,0 + 2 \cdot 107,0 + 2 \cdot 6,2 + 8,5 \cdot 113,0 = 1330,9 \text{ умовних т.}$$

Приклад 2. Відомі такі дані про виробництво електродвигунів (табл.4.2.):

Таблиця 4.2 - Виробництво електродвигунів за звітний період

Потужність двигуна, кВт (N)	Кількість, шт (X)
120	800
150	940
180	1700
300	2300

Необхідно визначити загальний обсяг виготовлених електродвигунів в умовних одиницях.

В першу чергу необхідно визначити коефіцієнт-сумірник, для цього в якості технічного параметра використаєм потужність двигунів, а за еталонний вид продукції – двигун потужністю $N=120$ кВт. Тоді для: $N=120$ $K=1$; $N=150$ $K=150:120=1,25$; $N=180$ $K=180:120=1,5$; $N=300$ $K=300:120=2,5$.

Загальний обсяг виробництва умовних двигунів за звітний період становить:

$$y = \sum kx = 1 \cdot 800 + 1,25 \cdot 940 + 1,5 \cdot 1700 + 2,5 \cdot 2300 = 10275 \text{ умовних шт., або } 10275 \text{ умовних електродвигунів потужністю } 120 \text{ кВт.}$$

4.2 Поняття, види та методи обчислення відносних величин

Відносна величина – це співвідношення величини даного явища з величиною будь-якого іншого явища, або з величиною того ж самого явища, взятого за інший період або з іншої території.

Іншими словами, відносна величина в алгебраїчній формі – це частка від ділення двох різноіменних або одноіменних величин. При використанні одноіменних величин знаменник відношення розглядається як база порівняння, або як основа відносної величини.

В статистичних дослідженнях використовуються такі види відносних величин:

- 1) **відносні величини планового завдання; виконання плану і динаміки.**

Введемо такі умовні позначення:

а – значення показника у попередньому періоді;

б – значення показника у поточному періоді згідно з планом (контрактом, держзамовленням, договором);

в – фактичне (звітне) значення показника за поточний період.

Тоді:

$$K_1 = \frac{b}{a} \cdot 100 \text{ -- планове завдання;} \quad (4.2.)$$

$$K_2 = \frac{a}{b} \cdot 100 \text{ -- виконання плану;} \quad (4.3.)$$

$$K_3 = \frac{v}{a} \cdot 100 \text{ -- динаміка.} \quad (4.4.)$$

Між зазначеними показниками існує така взаємозалежність:

$$K_3 = K_1 \cdot K_2 \quad (4.5)$$

Дійсно:

$$\frac{a}{a} = \frac{b}{a} \cdot \frac{a}{b} = K_3 ;$$

- 2) **відносні величини просторового порівняння** – одержують в результаті порівняння одноіменних показників, що відносяться до різних об'єктів і які взяті за один і той же період (наприклад, порівняння врожайності пшениці двох господарств. двох країн);
- 3) **відносні величини динаміки** – найбільш поширені відносні величини – одержують шляхом порівняння абсолютних або середніх величин, взятих за різні періоди. Розрізняють базисні відносні величини динаміки, коли порівнюється рівень показника за будь-який період з рівнем показника за базисний період, і ланцюгові відносні величини динаміки, коли порівнюється рівень показника за певний період з рівнем показника за попередній період;
- 4) **відносні величини структури** – це відношення складових частин до цілого, або частка (питома вага) частин в обсязі явища (наприклад, структура посівної площі господарства показує частку (питому вагу) площі під кожен культуру до загальної посівної площі, %);
- 5) **відносні величини координації** – це відношення однієї частини цілого до іншої частини (наприклад, співвідношення площі посіву під пшеницю і кукурудзу; чисельності міського і сільського населення);
- 6) **відносні величини інтенсивності** – характеризують ступінь поширення чи розвитку явища у певному середовищі. Такими показниками є демографічні коефіцієнти (народжуваності, смертності, шлюбності та ін.); коефіцієнт плінності робочої сили; чисельність лікарів, вчителів на 100000 населення тощо.

Важливим питанням, що виникає при розрахунках відносних величин, є правильний вибір бази (знаменника відносної величини при порівнянні одноіменних показників). База повинна бути типовою, а не специфічною, не випадковою.

Економічний зміст має різницю між чисельником і знаменником одноіменних величин – це збільшення (зменшення) рівня показника у звітному періоді в порівнянні з базисним періодом (або з планом), або на скільки значення показника на одному об'єкті більше (менше) значення того ж показника, взятого на іншому об'єкті.

Питання для самоконтролю

1. Які величини в статистиці називають абсолютними?
2. Перерахуйте одиниці виміру абсолютних величин і наведіть приклади їх використання.
3. В яких випадках використовуються умовно-натуральні одиниці виміру абсолютних величин? Наведіть приклади.
4. Розкрийте зміст відносних величин.
5. Назвіть основні види відносних величин та методи їх обчислення.
6. Які основні вимоги до вибору бази відносних величин?

Задачі

4.1. Є такі дані про витрати палива на виробничі потреби:

Вид палива	Витрачено	Коефіцієнт-сумірник
Вугілля донецьке, т	460	0,9
Мазут топковий, т	800	1,37
Газ природний, тис.м ³	940	1,2

Визначити загальний об'єм спожитого палива в умовних одиницях.

4.2. Відомі такі дані про виробництво продукції прокатним цехом металургійного заводу за звітний період:

Сталь кругла, діаметр, мм	Виготовлено, т		Коефіцієнт- сумірник
	За планом	Фактично	
30	7000	7100	1,0
35	5000	5050	0,9
40	4000	4300	0,8
45	1900	1800	0,7

Визначити виконання плану (в %) в цілому по всій виготовленій продукції у натуральному і умовно-натуральному вимірі.

4.3. Поставки молока на молокозавод характеризуються такими даними:

Номер господарства	Здано молока. ц	Жирність молока, %
1	2600	3,3
2	700	2,8
3	1500	3,6
4	1050	4,0

Визначити загальну кількість молока, що надійшло на молокозавод в перерахунку на умовне за жирністю 3,2%, а також ступінь виконання плану, якщо мало надійти за планом 5500ц. Молока жирністю 3,2%.

4.4. Є такі дані по підприємству:

Показник	Звітний період		Темп росту в порівнянні з попереднім періодом, %
	за планом	фактично	
Товарна продукція, млн. грн.	18920	19294	109,4

Визначити:

- 1) відносні величини планового завдання і виконання плану;
- 2) абсолютну величину товарної продукції в попередньому періоді.

4.5. Виробництво продукції цехами підприємства характеризується такими даними:

Номер цеху	Виробництво продукції, тис. грн.		Виконання плану, %
	за планом	фактично	
1	220,4	220,4	103,5
2	?	356,2	102,8
3	418,6	436,2	?

Внести в таблицю відсутні показники і визначити по трьох цехах разом рівень виконання плану.

4.6. Є такі дані по механічному цеху за місяць:

Показник	Звітний період		Попередній період
	за планом	фактично	
1. Товарна продукція, тис. грн.	208,6	?	201,5
2. Середньоспівомовна чисельність працюючих, чол.	135	132	?
3. Середній виробіток на 1-го працюючого, грн. (ряд1: ряд2)1000	?	1673	1504

Визначити:

а) відсутні показники в таблиці;

б) відносні величини планового зведення, виконання плану і динаміки по товарній продукції;

в) перевірити взаємозв'язок визначених відносних величин.

4.7. Річним планом передбачається зростання обсягу товарної продукції на 5,5%. План перевиконаний на 3,8%. Визначити, на скільки процентів виросла товарна продукція у звітному році в порівнянні з попереднім.

4.8. План виробництва продукції А виконаний на 102,5%. Обсяг продукції у звітному році виріс в порівнянні з попереднім на 3,4%. Визначити планове завдання по зростанню обсягу продукції А.

4.9. Планом на звітний період передбачено збільшити продуктивність праці на 5,0%. Фактичне зростання продукції праці становило 5,8%. Визначити процент виконання плану зростання продуктивності праці.

4.10. Підприємство у попередньому році виготовило 2700 одиниць продукції А. Планом на поточний рік передбачено збільшити обсяг виробництва продукції А на 5,5%. Фактично виготовлено в поточному році 3000 одиниць продукції А. Визначити процент виконання плану зростання продуктивності праці.

4.11. Відомі такі дані про використання колективним господарством посівної площі під зернові культури:

Назва зернових культур	Посівна площа, га	
	Попередній рік	Звітний рік
Пшениця	480	500
Жито	80	70
Ячмінь	120	95
Овес	70	65
Кукурудза	100	120
Разом:	850	850

Визначити:

- відносні величини структури посівних площ під окремі зернові культури;
- динаміку площі посіву під пшеницю, жито, кукурудзу.

Зробити висновки.

РОЗДІЛ 5 СЕРЕДНІ ВЕЛИЧИНИ І ПОКАЗНИКИ ВАРІАЦІЇ

5.1 Статистична середня та її суть

Середнім показником в статистиці називається узагальнююча або типова характеристика суспільних явищ по одній кількісній ознаці.

Статистичні середні показники відрізняються від математичних; якщо останні є величинами абстрактними, то перші завжди характеризують конкретні суспільні процеси.

Середня відображає типовий рівень ознаки лише в тому випадку, коли статистична сукупність, яку вона характеризує, якісно однорідна.

За допомогою середньої можна здійснити порівняльний аналіз кількох сукупностей, дати характеристику закономірностям розвитку соціально-економічних явищ і процесів.

Середній показник, як і будь-яке інше поняття, являє собою наукову абстракцію, якій нарівні з позитивним, що характеризує кожну абстракцію, притаманні і недоліки: в ній нівелюються індивідуальні відмінності досліджуваних явищ. Це на практиці може привести до того, що при некритичному використанні середніх за високими показниками не помічають незадовільні результати. Тому, якщо навіть сукупність є більш-менш однорідною, загальні середні рекомендується доповнити груповими середніми.

Розраховуючи середні, потрібно мати на увазі, що вони дають узагальнюючу характеристику явищ лише по одній ознаці. Кожне ж явище характеризується значною кількістю ознак. Тому з метою дослідження суспільних явищ рекомендується обчислювати не одну ізольовану середню, а систему середніх, що дозволяє оцінити явище з різних сторін.

5.2. Види середніх величин і способи їх обчислення

В статистиці є значна кількість середніх.

В табл.5.1 наведені середні, які найчастіше застосовуються на практиці.

Таблиця 5.1 - Види середніх величин та формули для їх розрахунку

Види середніх	Розрахункові формули	
	Прості (незгруповані дані)	Зважені (згруповані дані)
Арифметична	$\bar{x} = \frac{\sum x}{n} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n}$	$\bar{x} = \frac{\sum xf}{\sum f} = \frac{x_1f_1 + x_2f_2 + x_3f_3 + \dots + x_nf_n}{f_1 + f_2 + f_3 + \dots + f_n}$
Гармонічна	$\bar{x} = \frac{n}{\sum \frac{1}{x}} = \frac{1+1+1+\dots+1}{\frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3} + \dots + \frac{1}{x_n}}$	$\bar{x} = \frac{\sum w}{\sum \frac{w}{x}} = \frac{w_1 + w_2 + w_3 + \dots + w_n}{\frac{w_1}{f_1} + \frac{w_2}{f_2} + \frac{w_3}{f_3} + \dots + \frac{w_n}{f_n}}$
Квадратична	$\bar{x} = \sqrt{\frac{\sum x^2}{n}} = \sqrt{\frac{x_1^2 + x_2^2 + x_3^2 + \dots + x_n^2}{n}}$	$\bar{x} = \sqrt{\frac{\sum x^2 f}{\sum f}} = \sqrt{\frac{x_1^2 f_1 + x_2^2 f_2 + x_3^2 f_3 + \dots + x_n^2 f_n}{f_1 + f_2 + f_3 + \dots + f_n}}$
Геометрична	$\bar{x} = \sqrt[n]{x_1 \cdot x_2 \cdot x_3 \cdot \dots \cdot x_n}$	$\bar{x} = \sqrt[m]{x_1^{f_1} \cdot x_2^{f_2} \cdot x_3^{f_3} \cdot \dots \cdot x_n^{f_n}}$

Середня гармонійна застосовується в тих випадках, коли відомі не самі варіанти, а їх обернені числа.

Властивості середньої арифметичної:

а) алгебраїчна сума відхилень усіх віріант від середньої дорівнює нулю:

$$\sum (x_i - \bar{x}) = 0$$

б) якщо кожному варіанту розділити або зменшити на будь-яку постійну величину, то і середня зміниться на ту ж величину:

$$\frac{\sum (x_i \pm a)f}{\sum f} = \bar{x} + a$$

в) якщо кожному варіанту розділити або зменшити на будь-яке число, то і середня зменшиться або збільшиться в стільки ж разів:

$$\frac{\sum(x_i \cdot a)f}{\sum f} = \bar{x} \cdot a$$

г) якщо частоти усіх варіантів збільшити або зменшити в одне й те ж саме число разів, то середня при цьому не зміниться;

д) добуток середньої на суму частот дорівнює сумі добутоків варіантів на частоти:

$$\bar{x} \sum f = \sum x_i f$$

е) сума квадратів відхилення варіантів від середньої менша за будь-яку іншу величину:

$$\sum(x_i - \bar{x})^2 \rightarrow \min$$

5.3 Вибір виду середньої

Середня арифметична і середня гармонічна визначаються як частка від ділення обсягу значень ознаки на число одиниць сукупності, тобто за допомогою логічної формули, яка являє собою дріб.

Якщо є ряд даних з двох взаємозв'язаних показників, і по одному з них потрібно розрахувати середню величину, то можливі такі варіанти:

а) відомі чисельні значення знаменника логічної формули, а значення чисельника невідоме, однак може бути визначене як сума добутоків заданих величин, - середня в такому випадку визначається як середня зважена арифметична;

б) відомі чисельні значення чисельника логічної формули, а значення знаменника не відоме, однак може бути визначене як сума частки від

ділення одного показника на інший, - середня в такому випадку визначається як середня зважена гармонічна;

в) в умові задачі відомі чисельні значення як чисельника, так і знаменника – середня в такому випадку визначається за логічною формулою, тобто діленням першого на другий.

Перед вибором формули для обчислення середньої потрібно спочатку визначити, які показники потрібні для обчислення певної середньої, тобто уявити собі необхідну логічну формулу.

На підставі даних, які представлені в таблиці 5.2, розглянемо принципи вибору виду середньої для обчислення середньої заробітної плати по підприємству при наявності різних комбінацій даних.

Таблиця 5.2 - Показники оплати праці і чисельності робітників по цехах підприємства

Номер бригади	Середня заробітна плата, грн. (\bar{x}_i)	Фонд оплати праці, грн. (F_i)	Чисельність робітників, чол. (f_i)
1	2	3	4
1	190	5700	30
2	205	3690	18
3	212	2120	10
Разом	198,4	11510	58

Згідно методики обчислення величини середньої заробітної плати останнє визначається в будь-якому випадку як відношення фонду оплати праці до чисельності працівників, тобто:

$$\bar{x} = \frac{\sum F_i}{\sum f_i}$$

Варіант 1. Відомі дані граф 3 і 4. В такому випадку середня заробітна плата по підприємству визначається за наведеною вище формулою:

$$\bar{x} = \frac{\sum F_i}{\sum f_i} = \frac{11510}{58} = 198,4 \text{ грн.}$$

Варіант 2. Відомі дані граф 2 і 3, тобто відомий чисельник (фонд оплати праці), проте не відомий знаменник (чисельність робітників). З наявної інформації чисельність робітників в кожному цеху може бути визначена діленням фонду оплати праці на середню заробітну плату $\left(\frac{F_i}{\bar{x}_i}\right)$.

Середня заробітна плата по підприємству в такому випадку визначається як середня гармонічна зважена:

$$\bar{x} = \frac{\sum F_i}{\sum \frac{F_i}{\bar{x}_i}} = \frac{5700 + 3690 + 2120}{\frac{5700}{190} + \frac{3690}{205} + \frac{2120}{212}} = \frac{11510}{58} = 198,4 \text{ грн.}$$

Варіант 3. Відомі дані граф 2 і 4, тобто відомий знаменник (чисельність робітників), однак невідомий чисельник (фонд оплати праці). Останній можна визначити для кожного цеху як добуток середньої заробітної плати на чисельність робітників $(\bar{x}_i \cdot f_i)$.

Середня заробітна плата в такому випадку визначається як середня арифметична зважена:

$$\bar{x} = \frac{\sum \bar{x}_i f_i}{\sum f_i} = \frac{190 \cdot 30 + 205 \cdot 18 + 212 \cdot 10}{30 + 18 + 10} = \frac{11510}{58} = 198,4 \text{ грн.}$$

Правила вибору виду середньої дозволяють більш ґрунтовно зрозуміти зміст середньої гармонічної простої, яка застосовується для визначення витрат ресурсів (трудових, матеріальних) на одиницю продукції (робіт, послуг).

Наприклад, студент А на розв'язання однієї задачі витратив 12хв., студент Б – 15 хв., студент В – 14 хв., студент Г – 16 хв., студент Д – 15 хв.

Необхідно визначити, скільки в середньому витрачено часу на розв'язання однієї задачі.

Якщо визначити середнє значення як середню арифметичну просту, тобто:

$$(12+15+14+16+15):5=14,4 \text{ хв.},$$

то це буде невірно.

Для обґрунтування вибору виду середньої необхідно логічно уявити, яким чином слід розв'язати таку задачу. Природньо, щоб визначити, скільки часу потрібно в середньому на розв'язання однієї задачі, необхідно увесь витрачений час розділити на кількість розв'язаних задач.

Припустимо, що студенти розв'язували задачі протягом академічного часу (умовний вибір часу значення не має). Тоді кожний студент витратив по 45хв., а щоб визначити кількість розв'язаних задач протягом встановленого терміну, необхідно загальні витрати часу кожним студентом розділити на витрати часу на розв'язання однієї задачі, тобто:

$$\begin{aligned}\bar{x} &= \frac{45+45+45+45+45}{12+15+14+16+15} = \frac{45(1+1+1+1+1)}{45\left(\frac{1}{12}+\frac{1}{15}+\frac{1}{14}+\frac{1}{16}+\frac{1}{15}\right)} = \\ &= \frac{1+1+1+1+1}{\frac{1}{12}+\frac{1}{15}+\frac{1}{14}+\frac{1}{16}+\frac{1}{15}} = 13,3\text{хв.},\end{aligned}$$

що відповідає формулі середньої гармонічної простої:

$$\bar{x} = \frac{n}{\sum \frac{1}{x}}$$

яка міститься в таблиці 5.1.

5.4 Мода і медіана

Крім перерахованих раніш середніх в статистиці використовуються ще два вида середніх, які умовно називають структурними або порядковими. Це мода і медіана. Їх використовують для різних цілей, зокрема, у тих випадках, коли складно або майже неможливо визначити середню величину, наприклад, середню ціну продажу певної продукції на ринку.

Мода використовується також в тих випадках, коли необхідно дати характеристику величинам деяких ознак, які частіше всього зустрічаються в сукупності (наприклад, середня заробітна плата, середні ціни, середній розмір взуття тощо).

Модою в статистиці називається величина ознаки (варіанти), яка найчастіше зустрічається в сукупності даних. У варіаційному ряду модою є варіанта, яка має найбільшу частоту.

Медіаною в статистиці називається варіанта, яка знаходиться в середині варіаційного ряду. Медіана ділить ряд навпіл: по обидві сторони від неї знаходиться однакова кількість одиниць сукупності.

У випадках, коли у варіаційному ряду зустрічаються варіанти, де максимальне значення частоти повторюється, ряд вважається двомодальним, тримодальним і т.д.

Методи розрахунку моди і медіани залежать від виду варіаційного ряду – дискретного або інтервального.

Модою в дискретному варіаційному ряду є варіанта, якій відповідає найбільша частота (частка).

Медіані в дискретному ряду відповідає варіанта, місце якої визначається за формулою:

$$f_i'' \geq \frac{\sum f + 1}{2}. \quad (5.1)$$

Формула 5.1 визначає номер медіанної одиниці ряду.

Щоб визначити моду в інтервальному ряду, необхідно спочатку визначити модальний інтервал.

Модальним вважається інтервал з найбільшою чистотою.

Для визначення моди в інтервальному ряду використовується формула:

$$M_0 = X_{M_0} + i_{M_0} \frac{f_{M_0} - f_{M_0-1}}{(f_{M_0} - f_{M_0-1}) + (f_{M_0} - f_{M_0+1})}, \quad (5.2)$$

де M_0 – мода;

X_{M_0} – нижня межа модального інтервалу;

i_{M_0} – величина модального інтервалу;

$f_{M_0}, f_{M_0-1}, f_{M_0+1}$ – значення частот відповідно модального інтервалу, інтервалу, що передує модальному, інтервалу частоти, що слідує за модальним.

Перш ніж обчислити медіану в інтервальному ряду, необхідно визначити медіанний інтервал. **Медіанним** вважається інтервал, який знаходиться за номером медіанної одиниці ряду:

$$f_i'' \geq \frac{\sum f}{2}. \quad (5.3)$$

Медіана в інтервальному ряду визначається за формулою:

$$M_e = X_{M_e} + i_{M_e} \frac{\sum f}{2} - S_{M_e-1}, \quad (5.4)$$

де M_e – медіана;

X_{M_e} – нижня межа медіанного інтервалу;

i_{M_e} – величина медіанного інтервалу;

$\sum f$ – сума частот ряду;

S_{M_e-1} – сума накопичених частот в інтервалі, що передує медіанному;

f_{M_e} – частота медіанного інтервалу.

5.5 Поняття і показники варіації

Варіація, тобто коливання, мінливість значень будь-якої ознаки, є властивістю статистичної сукупності. Вона зумовлена дією безлічі взаємозв'язаних причин – суттєвих і несуттєвих.

Вивчення варіації має важливе значення для оцінки сталості та диференціації соціально-економічних явищ при застосуванні вибіркових та інших статистичних методів.

Варіація, по суті, - це нерівність одного і того ж показника у різних об'єктах; вона сприяє вивченню суті досліджуваних явищ.

Для виміру варіації в статистиці застосовується декілька показників (див. таблицю 5.3).

Таблиця 5.3 - Показники варіації та формули для їх обчислення

Назва показників варіації	Розрахункові формули	
	Незгруповані дані	Згруповані дані
1	2	3
1. Розмах варіації	$R = X_{\max} - X_{\min}$	$R = X_{\max} - X_{\min}$
2. Середнє лінійне відхилення	$\bar{d} = \frac{\sum x - \bar{x} }{n}$	$\bar{d} = \frac{\sum x - \bar{x} f}{\sum f}$
3. Середнє квадратичне відхилення	$\sigma = \sqrt{\frac{\sum (x - \bar{x})^2}{n}}$	$\sigma = \sqrt{\frac{\sum (x - \bar{x})^2 f}{\sum f}}$
4. Дисперсія	$\sigma^2 = \frac{\sum (x - \bar{x})^2}{n}$	$\sigma^2 = \frac{\sum (x - \bar{x})^2 f}{\sum f}$ $\sigma^2 = \bar{x}^2 - (\bar{x})^2$
5. Коефіцієнт варіації	$V_\sigma = \frac{\sigma}{x} \cdot 100$	$V_\sigma = \frac{\sigma}{x} \cdot 100$
6. Лінійний коефіцієнт варіації	$V_d = \frac{\bar{d}}{\bar{x}} \cdot 100$	$V_d = \frac{\bar{d}}{\bar{x}} \cdot 100$
7. Коефіцієнт осциляції	$V_R = \frac{R}{\bar{x}}$	$V_R = \frac{R}{\bar{x}}$

Середнє лінійне відхилення і середньоквадратичне відхилення – іменовані величини, і кожна з них має свої одиниці виміру. Тому їх неможливо використати при порівнянні різних ознак або навіть однієї і тієї ж ознаки, якщо використовуються різні розмірності. Так, врожайність культур може бути оцінена в центнерах або тоннах. Середнє лінійне відхилення і середньоквадратичне відхилення буде в такому випадку залежати від одиниці виміру.

Тому для оцінки варіації, особливо коли потрібно порівняти мінливість різних ознак, використовується відносна величина – коефіцієнт варіації.

Розглянемо на прикладах методика обчислення середніх і показників варіації для дискретного і інтервального групувань.

Усі проміжні розрахунки для зручності здійснюються безпосередньо в таблицях.

Таблиця 5.4 - Розподіл підприємств за чисельністю робітників

Чисельність робітників, чол., x	Кількість Підприємств, f	xf	$ x - \bar{x} $	$ x - \bar{x} f$	$(x - \bar{x})^2$	$(x - \bar{x})^2 f$	f'	xf'	f''
30	1	10	45,25	45,25	2047,6	2047,6	2,5	2,5	1
40	6	240	35,25	211,50	1242,6	7455,5	15,0	600	7
50	5	250	22,25	126,25	637,6	3188,0	12,5	625	12
60	4	240	15,25	61,00	232,6	930,4	10,0	600	16
70	5	350	5,25	26,25	27,6	138,0	12,5	875	21
80	4	320	4,75	19,00	22,5	90,0	10,0	800	25
90	3	270	14,75	44,25	217,5	625,5	7,5	675	28
100	4	400	24,75	99,00	612,5	2450,0	10,0	1000	32
110	5	550	34,75	173,75	1207,5	6037,5	12,5	1375	37
120	3	360	44,75	134,25	2002,5	6007,5	7,5	900	40
Σ	40	3010	-	940,5	-	28997,0	100	7525	-

Для обчислення показників варіації слід розрахувати середню арифметичну.

Для обчислення показників використовуються дані підсумкового рядку.

Середня арифметична незгрупованих даних і в дискретному групуванні, побудованому на основі цих даних. (незалежно від використаної частоти (f) або частки (f'), за значенням завжди співпадає. Останнє відноситься і до показників варіації.

Розраховані на підставі підсумкових даних таблиці 5.4 статистичні показники дорівнюють:

середня арифметична

$$\bar{x} = \frac{\sum xf}{\sum f} = \frac{3010}{40} = 75,25 \quad \text{або} \quad \bar{x} = \frac{\sum xf'}{\sum f'} = \frac{7525}{100} = 75,25$$

середнє лінійне відхилення

$$\bar{d} = \frac{\sum |x - \bar{x}| f}{\sum f} = \frac{940,5}{40} = 23,25$$

дисперсія

$$\sigma^2 = \frac{\sum (x - \bar{x})^2 f}{\sum f} = \frac{28997,0}{40} = 724,9$$

середнє квадратичне відхилення

$$\sigma = \sqrt{\sigma^2} = \sqrt{724,9} = 26,9$$

коефіцієнт варіації

$$V = \frac{\sigma}{\bar{x}} \cdot 100 = \frac{26,9}{75,25} \cdot 100 = 35,7\%$$

Мода $M_0 = 40$, так як цій ознаці відповідає найбільша частота або, іншими словами, в сукупності підприємств чисельність робітників в 40 чол. є найбільш типовою.

Для обчислення медіани необхідно в першу чергу визначити номер медіанної одиниці ряду:

$$f'' \geq \frac{\sum f + 1}{2} \geq \frac{40 + 1}{2} = 20,5 \quad f'' \geq 21$$

Цій одиниці ряду відповідає медіана $M_e = 70$, тобто певне підприємство з чисельністю 70 робітників ділить варіаційний ряд розподілу на дві рівні частини.

Середня арифметична в інтервальних групуваннях визначається аналогічно дискретному варіаційному ряду, однак спочатку потрібно визначити середнє значення варіанти в кожній групі, тобто умовно перетворити інтервальне групування в дискретне. Середнє значення варіанти визначається за формулою:

$$x' = \frac{x_{\max} + x_{\min}}{2}, \quad (5.5)$$

де x' – середнє значення інтервалу;

x_{\max}, x_{\min} – відповідно верхня і нижня межа (границя) інтервалу.

У відкритих інтервальних групуваннях у першій і останній групах (перший і останній рядки) розмір інтервалу приймається на рівні груп (рядків), що знаходяться поруч (другий і передостанній).

У такому разі умовна нижня границя першої групи визначається як різниця між заданою верхньою границею і розміром інтервалу другої групи, а умовна верхня границя останньої групи визначається як сума заданої нижньої границі і розмір інтервалу передостанньої групи (див. таблицю 5.5).

Таблиця 5.5 - Розподіл робітників за обсягом виготовленої продукції

Групи робітників за кількістю виготовленої продукції за зміну, шт., x	Кількість робітників, f	Середнє Значення Інтервалу $x' =$ $= \frac{x_{\max} + x_{\min}}{2}$	$x'f$	$ x' - \bar{x} $	$ x' - \bar{x} f$	$(x' - \bar{x})^2$	$(x' - \bar{x})^2 f$	f''
До 5	10	4	40	3,5	35,0	12,25	122,5	10
5-7	30	6	180	1,5	45,0	2,25	67,50	40
7-9	40	8	320	0,5	20,0	0,25	10,0	80
9-11	15	10	150	2,5	37,0	6,25	93,75	95
Понад 14	5	12	60	4,5	22,5	20,25	101,25	100
Σ	100	-	750	-	160,0	-	395	-

Середні та показники варіації визначаються наступним чином:

$$\bar{x} = \frac{\sum x'f}{\sum f} = \frac{750}{100} = 7,5,$$

$$\bar{d} = \frac{\sum |x' - \bar{x}|f}{\sum f} = \frac{160}{100} = 1,6,$$

$$\sigma^2 = \frac{\sum (x' - \bar{x})^2 f}{\sum f} = \frac{395}{100} = 3,95,$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{3,95} = 1,99,$$

$$v = \frac{\sigma}{\bar{x}} \cdot 100 = \frac{1,99}{7,5} \cdot 100 = 26,5\%.$$

Модальним є інтервал 7-9, тому має найбільшу частоту ($f=40$).

Згідно з формулою 5.2:

$$M_0 = X_{m_0} + i_{m_0} \left(\frac{f_{m_0} - f_{m_0-1}}{(f_{m_0} - f_{m_0-1}) + (f_{m_0} + f_{m_0+1})} \right) = 7 + 2 \left(\frac{40 - 30}{(40 - 30) + (40 - 15)} \right) = 7,6.$$

Для визначення медіального інтервалу використаєм формулу 5.3:

$$f^* \geq \frac{\sum f}{2} \geq \frac{100}{2} \geq 50.$$

Перший інтервал, який відповідає цій вимозі, є інтервал 7-9. Таким чином, медіанний і модельний інтервали співпадають.

Згідно з формулою 5.4:

$$M_0 = X_{m_0} + i_{m_0} \frac{\frac{\sum f}{2} - S_{m_0-1}}{f_{m_0}} = 7 + 2 \frac{\frac{100}{2} - 40}{40} = 7,5.$$

Після наведених прикладів продовжимо подальше вивчення варіації ознаки.

Варіація ознаки залежить від безлічі факторів, які нерідко діють в різних напрямках.

Наприклад, врожайність при інших рівних умовах залежить від кількості внесених добрив, опадів, якості ґрунту тощо.

В процесі оцінки дії факторів необхідно в першу чергу здійснити групування одиниць спостереження, а потім визначити загальну, міжгрупову і внутрішньогрупову дисперсії.

Величина загальної дисперсії (σ^2) характеризує варіацію ознак під впливом усіх умов, які викликають цю варіацію. Вона обчислюється за формулою:

$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2 f_i}{\sum f_i}. \quad (5.6)$$

Міжгрупова дисперсія (δ^2) характеризує систематичну варіацію, тобто різницю у величині досліджуваної ознаки, яка виникає під впливом однієї умови (ознаки-фактора), що закладена в основу групування. Вона обчислюється за формулою:

$$\delta^2 = \frac{\sum (\bar{x} - \bar{x}_i) \cdot n_i}{\sum n_i}, \quad (5.7)$$

де \bar{x}_i – середня окремих груп;

n_i – чисельність одиниць сукупності окремих груп.

Внутрішньогрупова дисперсія, або середня із групових дисперсій ($\bar{\sigma}^2$) відображає випадкову варіацію, тобто частину варіації, яка виникає під впливом інших, неврахованих факторів, і не залежить від умови (ознаки-фактора), що покладена в основу групування. Вона визначається за формулою:

$$\bar{\sigma}^2 = \frac{\sum \sigma_i^2 n_i}{\sum n_i}, \quad (5.8)$$

де σ_i^2 – групова дисперсія. Повинна виконуватися умова:

$$\sigma^2 = \delta^2 + \bar{\sigma}^2.$$

Методику обчислення загальної, міжгрупової і внутрішньогрупової дисперсій розглянемо на прикладі, дані яких наведені в таблиці 5.6.

Таблиця 5.6 - Стаж роботи та годинна заробітна плата робітників

Стаж роботи, роки	Годинна заробітна плата, грн.	Стаж роботи, роки	Годинна заробітна плата, грн.
2	0,62	13	0,95
4	0,68	14	1,01
7	0,88	17	0,93
10	0,84	22	1,02
11	0,82	28	1,05

Серед робітників п'ятеро пройшли навчання з годинною оплатою (в коп.): 84, 93, 95, 101, 102, решта навчання не пройшли.

Обчислимо для зазначених сукупностей

а) групові середні

$$\bar{x}_1 = \frac{84+93+95+101+102}{5} = \frac{475}{5} = 95 \text{ коп.};$$

$$\bar{x}_2 = \frac{62+68+82+88+105}{5} = \frac{405}{5} = 81 \text{ коп.};$$

б) загальну середню

$$\bar{x} = \frac{475+405}{100} = 88 \text{ коп.};$$

в) групові дисперсії

$$\sigma_1^2 = \frac{(84-95)^2 + (93-95)^2 + (95-95)^2 + (101-95)^2 + (102-95)^2}{5} = \frac{210}{5} = 42;$$

$$\sigma_2^2 = \frac{(68-81)^2 + (68-81)^2 + (82-81)^2 + (88-81)^2 + (105-81)^2}{5} = \frac{1156}{5} = 231,2;$$

г) загальну дисперсію

$$\sigma^2 = [(84-88)^2 + (93-88)^2 + (95-88)^2 + (101-88)^2 + (102-88)^2 + (62-88)^2 + (68-88)^2 + (82-88)^2 + (88-88)^2 + (105-88)^2] : 10 = 185,6.$$

На підставі розрахованих показників обчислимо:

міжгрупову дисперсію:

$$\delta^2 = \frac{\sum (\bar{x}_i - \bar{x})^2 n_i}{\sum n_i} = \frac{(95-88)^2 \cdot 5 + (81-88)^2 \cdot 5}{5+5} = \frac{490}{10} = 49,0;$$

внутрішньогрупову дисперсію

$$\bar{\sigma}^2 = \frac{\sum \sigma_i^2 n_i}{\sum n_i} = \frac{42 \cdot 5 + 231,2 \cdot 5}{5+5} = \frac{1366}{10} = 136,6;$$

$$\sigma^2 = \delta^2 + \bar{\sigma}^2 = 49,0 + 136,6 = 185,6.$$

Якщо припустити загальну дисперсію в 100%, то варіація під впливом міжгрупової дисперсії (підготовка кадрів) становила $(49:185,6) \cdot 100 = 26,4\%$ від загальної, а під впливом внутрішньогрупової дисперсії $(136,6:185,6) \cdot 100 = 73,6\%$.

Таким чином, коливання заробітної плати майже на три чверті залежить від випадкових (неврахованих) факторів.

Питання для самоконтролю

1. Дайте визначення середньої.

2. В чому відмінність середньої в математиці і статистиці?
3. Які види середніх вам відомі?
4. Сформулюйте правила вибору виду середньої.
5. Дайте визначення понять “мода” і “медіана”. В яких випадках вони використовуються?
6. Які показники варіації вам відомі?
7. В чому перевага коефіцієнта варіації в порівнянні з середнім лінійним і середнім квадратичним відхиленнями?
8. Яким чином за допомогою показників дисперсії можна оцінити вплив на варіацію ознак врахованого і випадкових (неврахованих) факторів?

Задачі

5.1. Місячна заробітна плата робітників характеризується такими даними:

Місячна заробітна плата, грн.	110	130	160	190	220	Разом
Кількість робітників	2	6	16	12	14	50

Визначити середню заробітну плату, розмах варіації, середнє лінійне відхилення, дисперсію, коефіцієнт варіації, моду і медіану.

5.2. По виконанню норм виробітку робітники розподілені таким чином:

Виконання норми виробітку, %	90-100	100-110	110-120	120-130	Понад 130	Разом
Кількість робітників	2	5	4	6	4	21

Визначити середній процент виконання норми виробітку, середнє лінійне відхилення, дисперсію, коефіцієнт варіації, моду і медіану.

5.3. Розподіл заводів цементної промисловості за розміром добової виробничої потужності такий:

Виробництво цементу, тис. т	До 100	100- 200	200- 300	300- 500	500- 700	По- нал 700	Разом
Питома вага заводів в % до підсумку	10	15	25	21	16	13	100

Визначити середню виробничу потужність, середнє лінійне відхилення, дисперсію, коефіцієнт варіації, моду і медіану.

5.4. Середньодобовий видобуток вугілля на протязі 10 днів характеризується такими даними:

Дні	1	2	3	4	5	6	7	8	9	10
Видобуток вугілля, т за добу	4,5	4,6	4,9	5,0	5,4	6,0	5,4	5,8	5,9	6,2

Визначити середній добовий видобуток вугілля, середнє лінійне відхилення, дисперсію, коефіцієнт варіації.

5.5. Врожайність зернових в господарствах характеризується такими даними:

Групи господарств за врожайністю зернових, ц/га	До 25	25-29	29-33	33-37	37 і більше	Разом
Кількість господарств	3	6	11	8	2	30

Визначити середню врожайність зернових, середнє лінійне відхилення, дисперсію, коефіцієнт варіації, моду і медіану.

5.6. Витрати часу трьома робітниками на обробку однієї деталі відповідно становить: 20, 24, 16 хв. Загальні витрати часу кожним робітником відповідно дорівнюють: 420, 480, 400 хв.

Визначити середні витрати часу на обробку однієї деталі.

5.7. Фабрикою виготовлено 62,0 тис. м тканини, в тому числі: першого сорту 92,0% по 6,4 грн. за 1м; другого сорту 6,0% по 6,2 грн. за 1м; третього сорту 2,0% по 6,0 грн. за 1м.

Визначити середню ціну 1м тканини.

5.8. Відомі такі дані по виконанню плану виробництва продукції:

Номер підприємства	Обсяг продукції за планом, тис. грн.	Виконання плану, %
1	340	95,0
2	450	100,0
3	1550	102,0

Визначити середній процент виконання плану по трьох підприємствах.

5.9. Витрати сировини на виробництво продукції характеризуються такими даними:

Номер бригади	Витрати сировини на виробництво одного виробу, кг	Загальні витрати сировини, кг
1	11	6500
2	10	5500
3	9	5000

Визначити середні витрати сировини на одиницю продукції.

5.10. Виконання плану реалізації продукції підприємствами об'єднання характеризується такими даними:

Показник	Номер підприємства			
	1	2	3	4
Обсяг реалізації за звітом, млн. грн.	715	510	850	800
Виконання плану, %	103,1	98,0	100,0	102,5

Визначити процент виконання плану в цілому по об'єднанню.

5.11. Собівартість продукції А по трьох підприємствах характеризується такими даними:

Номер підприємства	Витрати на виробництво, тис. грн.	Собівартість одиниці продукції, грн.
1	200	20
2	460	23
3	110	22

Визначити середню собівартість одиниці продукції.

5.12. Відомі такі дані про жирність молока:

Жирність молока	Надійшло молока на молокозавод, ц
3,9	1500
3,8	400
3,7	250
3,6	150

Визначити середній процент жирності молока.

5.13. Врожайність пшениці по колективному господарству характеризується такими даними:

Культура	Попередній рік		Звітний рік	
	Врожайність, ц/га	Посівна площа, тис. га	Врожайність, ц/га	Валовий збір, тис. ц
Пшениця озима	26,9	8,2	28,0	200,0
Пшениця яра	20,3	15,8	21,0	220,0

Визначити середню врожайність пшениці у попередньому і звітному роках.

5.14. Частка розфасованих продовольчих товарів в загальному обсязі виробництва характеризується такими даними:

Групи продовольчих товарів	Загальний обсяг виробництва, тис. т	Частка розфасованих товарів в загальному обсязі виробництва, %
Маргаринова продукція	290,0	40,1
Вершкове масло	1250,0	25,0

Визначити середню частку розфасованих товарів в загальному обсязі їх виробництва.

5.15. Якість продукції за звітний квартал характеризується такими даними:

Вид продукції	Процент браку	Вартість виготовленої продукції, тис. грн.
А	0,8	7,4
Б	1,2	8,2
С	2,5	6,9
Д	0,8	7,5

Визначити середній процент браку по продукції в цілому.

5.16. Впровадження інтенсивної технології для вирощування зернових культур характеризується такими даними:

Групи сільськогосподарських культур	Загальний розмір посівної площі, тис. га	Частка площі, обробленої за інтенсивною технологією, %
Озимі зернові	500	55,2
Зернобобові	130	35,0

Визначити середню частку площі, що обробляється за інтенсивною технологією.

Розділ 6. Ряди динаміки

6.1 Поняття та види рядів динаміки

Суспільні явища безперервно змінюються, розвиваються в часі і просторі. Так, протягом певного часу змінюється обсяг виготовленої продукції, споживання матеріальних благ, рівень продуктивності праці, структура суспільного виробництва тощо.

Суспільні явища слід вивчати в неперервному розвитку, для чого формуються хронологічні ряди, які також називають рядами динаміки, динамічними рядами.

Динамічний ряд – це послідовність упорядкованих в часі кількісних показників, які характеризують рівень розвитку досліджуваних явищ. Цифри (показники), із яких складається динамічний ряд, називаються **рівнями**.

Ряд динаміки складається з двох елементів: **рівнів ряду і моментів (періодів) часу**, до яких належать ці рівні.

За часом, що відображається в рядах динаміки, вони поділяються на моментні і інтервальні.

В **моментних рядах динаміки** рівні висвітлюють величину явища на певну дату (табл.6.1)

Таблиця 6.1 - Залишки готової продукції на складі підприємства на початок місяця

Місяць	1.01	1.02	1.03	1.04	1.05	1.06	1.07
Залишки готової продукції тис. тис. грн.	3,5	2,8	3,0	2,9	2,0	2,7	1,9

В інтервальних рядах рівні висвітлюють величину явищ за певний проміжок часу (таблиця 6.2).

Таблиця 6.2 - **Обсяг виготовленої продукції**

Місяць	Січень	Лютий	Березень	Квітень	Травень	Червень
Обсяг виготовленої продукції, тис. грн.	355	340	360	365	370	375

На відміну від моментних рядів динаміки, рівні явищ за окремі періоди, тобто рівні інтервальних рядів, можна підсумовувати, в результаті чого одержують нові рівні за більш тривалий період. Так, підсумовуючи дані за три місяці (див. табл.6.2) можна одержати рівні за I і II квартали.

За відстанню між датами або інтервалами часу виділяють повні і неповні ряди динаміки.

У **повних рядах** дати або періоди ідуть один за одним з рівними інтервалами.

У **неповних рядах** принцип рівних інтервалів не дотримується.

Наведені в таблиці 6.1 і 6.2 динамічні ряди є повними.

Неповний динамічний ряд можна показати на підставі наведених в таблиці 6.1 даних про залишки готової продукції (таблиця 6.3).

Таблиця 6.3 - **Залишки готової продукції на складі підприємства на початок місяця**

Місяць	1.01	1.03	1.04	1.07
Залишки готової продукції, тис. грн.	3,5	3,0	2,9	1,9

Динамічний ряд, який характеризує зміну тільки одного показника, називається **одномірним**, двох і більше показників – **багатомірним**.

Паралельні ряди динаміки характеризують зміну в часі різних показників одного об'єкта (обсяг продукції, чисельність працюючих, вартість основних виробничих фондів окремого підприємства), або одного показника різних об'єктів (продуктивність праці різних підприємств).

За формою подання рівнів ряди динаміки можуть бути рядами **абсолютних, середніх і відносних величин** (обсяг продукції, обсяг продукції на одного працюючого, темпи росту обсягу продукції).

6.2 Правила формування рядів динаміки

Щоб забезпечити об'єктивну оцінку розвитку явища (процесу) за допомогою числових рівнів, необхідно дотримуватися певних правил при формуванні рядів динаміки, зокрема:

а) періодизації розвитку, тобто розподіл рядів динаміки на однорідні етапи. Це означає, що потрібно формувати динамічні ряди за суворо однорідними періодами і етапами, що одночасно не заперечує можливість побудови і дослідження рядів динаміки за більш тривалі відрізки часу, які включають різні етапи розвитку явища;

б) однакості окремих рівнів динамічного ряду. Для цього необхідно попередньо провести типологічне або структурне групування, оцінити динаміку розвитку окремих груп, а потім явища в цілому. Так, динаміку середнього доходу необхідно досліджувати після вивчення динаміки середнього доходу окремих класів або груп населення;

в) порівнянність рівнів динамічного ряду – це означає, що рівні повинні бути в однакових одиницях виміру, визначатися по єдиній методології, включати однорідне коло об'єктів;

г) послідовність і неперервність в часі рівнів динамічного ряду – іншими словами, рівні динаміки повинні послідовно охоплювати весь етап розвитку явища від початку до кінця. Відсутність даних за деякі проміжки часу може перевернути уяву про тенденцію розвитку явища при аналізі динамічного ряду;

д) величина ряду динаміки повинна відповідати інтенсивності досліджуваних явищ. Для більш мінливих явищ використовуються менші інтервали часу; для явищ з меншою варіацією – більші інтервали часу, або більші відрізки часу. Величина інтервалу в принципі повинна бути узгоджена з тривалістю циклів досліджуваних процесів: якщо виробництво виробу продовжується на протязі місяця, то немає необхідності вимагати звіту про готову продукцію за п'ять днів або декаду.

Для поглибленого вивчення явищ слід вивчати не один динамічний ряд, який характеризує зміну тільки однієї ознаки, однієї сторони явища, а досліджувати систему динамічних рядів, що дозволить всебічно вивчити тенденції і закономірності розвитку явища.

Наприклад, динаміка валового збору цукрового буряку відображає далеко не всі процеси в системі бурякосіяння: потрібно одночасно досліджувати динаміку посівних площ, врожайності, технічного забезпечення тощо.

6.3 Статистичні характеристики рядів динаміки

На основі рівнів ряду динаміки визначають систему показників, які дозволяють всебічно оцінити напрямок та інтенсивність змін явища в часі. Ці показники називають **аналітичними (статистичними) характеристиками** рядів динаміки і обчислюють їх шляхом порівняння рівнів динамічних рядів. Оскільки розвиток явищ не має зворотного

напрямую, то порівнюють наступний рівень з попереднім, приймаючи останній за базу, а не навпаки. Якщо порівнюють лише з одним і тим же початковим рівнем, то отримують **базисні показники динаміки**; якщо ж порівнюють усі показники з найближчим попереднім рівнем, то отримують **ланцюгові показники динаміки**.

Статистичні характеристики рядів динаміки і формули для їх обчислення наведені в таблиці 6.4.

Таблиця 6.4. - Статистичні характеристики рядів динаміки
і формули для їх обчислення

Статистичні характеристики	Розрахункові формули
1	2
Абсолютний приріст:	
ланцюговий	$\Delta y' = y_i - y_{i-1}$
базисний	$\Delta y = y_i - y_0$
середній	$\Delta \bar{y} = \frac{y_n - y_0}{n - 1}$
Коефіцієнт росту:	
ланцюговий	$K'p = \frac{y_i}{y_{i-1}}$
базисний	$Kp = \frac{y_i}{y_0}$
за весь період	$Kp_n = \frac{y_n}{y_0}$
середній	$\bar{K}p = \sqrt[n-1]{\frac{y_n}{y_0}}$
Темп росту	$Tp' = Kp' \cdot 100; Tp = Kp \cdot 100$ $Tp_n = Kp_n \cdot 100; \bar{T}p = \bar{K}p \cdot 100$

Продовження таблиці 6.4.

1	2
Темп приросту	$T'_{np} = T'_p - 100$; $T_{np} = T_p - 100$ $T'_{np_n} = T'_{p_n} - 100$; $\bar{T}_{np} = \bar{T}_p - 100$
Абсолютний розмір 1% приросту:	
ланцюговий	$\Delta y_{i\%} = \frac{y_i - y_{i-1}}{y_{i-1}} \cdot 100 = \frac{y_{i-1}}{100}$
середній за весь період	$\Delta \bar{y}_{i\%} = \frac{\Delta \bar{y}}{\bar{T}_{np}}$
Коефіцієнт випередження (відставання)	$K = \left(\frac{y_i}{y_{i-1}} \right) : \left(\frac{x_i}{x_{i-1}} \right)$

Наведені в таблиці 6.4 коефіцієнти (темпи) росту за весь період і середній темп (коефіцієнт) росту можуть бути визначені тільки на підставі абсолютних даних динамічного ряду. Якщо ж такі дані відсутні, а є тільки ланцюгові темпи росту (приросту), то слід дотримуватися такого правила:

добуток ланцюгових коефіцієнтів росту дорівнює коефіцієнту росту за весь період, тобто,

$$Kp_n = \frac{y_n}{y_0} = K'p_1 \cdot K'p_2 \cdot K'p_3 \cdot \dots \cdot K'p_n = \Pi \frac{y_i}{y_{i-1}}, \quad (6.1)$$

а середній коефіцієнт росту в такому випадку визначається як

$$\bar{K}p = \sqrt[n]{\Pi \frac{y_i}{y_{i-1}}}. \quad (6.2)$$

Якщо в умові задачі подані темпи приросту, то їх спочатку потрібно перетворити в темпи росту ($T'_p = T'_{np} + 100$), а потім перевести їх в коефіцієнти росту ($K'_p = T'_p : 100$).

Приклад. Обсяг продукції збільшувався щорічно (по відношенню до попереднього року) за перші три роки п'ятирічки на 7,5%; за четвертий рік на 9,0%, а за останній рік він знизився на 0,4%. Необхідно визначити темп росту за п'ятиріччя і середньорічний темп приросту.

Перетворимо темпи приросту в коефіцієнти росту $K'_p = [100 + T'_{пр}]:100$:

$$K'_{p1} = (100 + 7,5):100 = 1,075;$$

$$K'_{p2} = (100 + 7,5):100 = 1,075;$$

$$K'_{p3} = (100 + 7,5):100 = 1,075;$$

$$K'_{p4} = (100 + 9,0):100 = 1,090; K'_{p5} = (100 - 0,4):100 = 0,996.$$

На підставі формули 6.1 визначається коефіцієнт росту за п'ятирічку:

$$\frac{Y_n}{Y_0} = K'_{p1} \cdot K'_{p2} \cdot K'_{p3} \cdot K'_{p4} \cdot K'_{p5} = 1,075 \cdot 1,075 \cdot 1,075 \cdot 1,090 \cdot 0,996 = 1,349.$$

Звідси темп росту за п'ятирічку становить $1,349 \cdot 100 = 134,9\%$, а темп приросту за п'ятиріччя дорівнює $T_{прн} = T_p - 100 = 134,9 - 100 = 34,9\%$, тобто обсяг продукції за п'ятирічку виріс на 34,9%.

У формулі середнього коефіцієнта росту під коренем знаходиться коефіцієнт росту за весь період.

$$\text{Отже, } \bar{K}_p = \sqrt[n]{\frac{Y_n}{Y_0}} = \sqrt[5]{1,349} = 1,051.$$

Звідси середньорічний темп росту становить:

$$\bar{T}_p = 1,051 \cdot 100 = 105,1\%,$$

а середньорічний темп приросту дорівнює:

$$\bar{T}_{пр} = \bar{T}_p - 100 = 105,1 - 100 = 5,1\%.$$

При аналізі динамічних рядів слід також зважити на такі правила.

Частка від ділення базисного коефіцієнта росту за поточний період на базисний коефіцієнт росту за попередній період дорівнює ланцюговому коефіцієнту росту за поточний період:

$$\frac{Y_t}{Y_0} : \frac{Y_{t-1}}{Y_0} = \frac{Y_t}{Y_{t-1}}; \quad (6.3)$$

Добуток двох послідовних ланцюгових коефіцієнтів росту дорівнює базисному коефіцієнту росту по відношенню до першого періоду:

$$\frac{Y_1}{Y_0} \cdot \frac{Y_2}{Y_1} = \frac{Y_2}{Y_0} \quad (6.4)$$

Порівняння двох послідовних ланцюгових абсолютних приростів показує **абсолютне прискорення або уповільнення динаміки**:

$$\Delta_i = \Delta y_i - \Delta y_{i-1} \quad (6.5)$$

Якщо $\Delta_i > 0$, то має місце абсолютне прискорення; якщо $\Delta_i < 0$ - то уповільнення.

Співвідношення додатних абсолютних ланцюгових приростів показує **відносне прискорення**:

$$K = \frac{\Delta y_i}{\Delta y_{i-1}} \quad (6.6)$$

На підставі даних таблиці 6.5 розрахуємо основні статистичні характеристики наведеного ряду динаміки.

Таблиця 6.5 - Виробництво товарної продукції за роки п'ятирічки

Показник	1995 - базисний рік	Роки п'ятирічки				
		1996	1997	1998	1999	2000
Обсяг товарної продукції в порівнянних цінах, млн. грн	22,6	23,4	25,1	27,0	28,9	30,5

Для обчислення статистичних характеристик динамічного ряду використані формули, які подані в таблиці 6.4, а також формули 6.5 і 6.6.

Результати розрахунків наведені в таблиці 6.6.

Таблиця 6.6 - Статистичні характеристики динамічного ряду

Рік	Обсяг продукції, млн. грн., у	Абсолютний приріст, млн. грн.		Темп росту, %		Темп приросту, %		Розмір 1-го приросту, млн. грн.	Прискорення	
		ланцюговий $\Delta y_i = y_i - y_{i-1}$	базисний $\Delta y_i = y_i - y_0$	ланцюговий $T_p' = \frac{y_i}{y_{i-1}} \cdot 100$	базисний $T_p = \frac{y_i}{y_0} \cdot 100$	ланцюговий $T_{np}' = T_p' - 100$	базисний $T_{np} = T_p - 100$		ланцюговий $\Delta y_{i-1} = \frac{y_{i-1}}{100}$	абсолютне $\Delta_i = \Delta y_i - \Delta y_{i-1}$
1995	22,6	-	-	100,0	100,0	-	-	-	-	-
1996	23,4	0,8	0,8	103,2	103,2	3,2	3,2	0,226	-	-
1997	25,1	1,7	2,5	107,3	111,1	7,3	11,1	0,234	0,9	2,13
1998	27,0	1,9	4,4	107,6	119,5	7,6	19,5	0,251	0,2	1,12
1999	28,9	1,9	6,3	107,0	127,9	7,0	27,9	0,270	0	1,00
2000	30,5	1,6	7,9	105,5	134,9	5,5	34,9	0,289	-0,3	0,84
Середнє Значення		1,58		105,1		5,1		0,31	-	

Однією із узагальнюючих характеристик динамічного ряду є середній рівень, методи обчислення якого залежать від його виду.

Середній рівень інтервального ряду визначається за допомогою середньої арифметичної простої:

$$\bar{y} = \frac{\sum y}{n} \quad (6.7)$$

Середній рівень динамічного ряду, який наведений в таблиці 6.5, дорівнює:

$$\bar{y} = \frac{22,6 + 23,4 + 25,1 + 27,0 + 28,9 + 30,5}{6} = \frac{157,5}{6} = 26,3 \text{ млн. грн.}$$

Середній рівень повного моментного ряду визначається за формулою:

$$\bar{y} = \frac{y_1 + y_2 + y_3 + \dots + y_n}{n-1} = \frac{y_1 + y_n + \sum_{i=2}^{n-1} y_i}{n-1} \quad (6.8)$$

Для даних табл.6.1 середній рівень становить:

$$\bar{y} = \frac{3,5 + 2,8 + 3,0 + 2,9 + 2,0 + 2,7 + \frac{1,9}{2}}{7-1} = \frac{16,1}{6} = 2,7 \text{ тис. грн.}$$

Середній рівень неповного моментного ряду визначається за формулою:

$$\bar{y} = \frac{(y_1 + y_2) \cdot t_1 + (y_2 + y_3) \cdot t_2 + \dots + (y_{n-1} + y_n) \cdot t_{n-1}}{2(t_1 + t_2 + \dots + t_{n-1})} \quad (6.9)$$

де t – інтервал часу між двома суміжними періодами.

Середній рівень неповного динамічного ряду, який наведений в таблиці 6.3, становить:

$$\bar{y} = \frac{(3,5 + 3,0) \cdot 2 + (3,0 + 2,9) \cdot 1 + (2,9 + 1,9) \cdot 3}{2(2 + 1 + 3)} = \frac{33,3}{12} = 2,8 \text{ тис. грн.}$$

6.4 Основна тенденція динамічного ряду та методи її виявлення

Тенденція – це певний напрямок розвитку, тривала еволюція, яка має вигляд плавної траєкторії. Статистичне дослідження тенденції ґрунтується на розкладанні динамічного ряду на дві складові:

$$y_t = f(t) + \varepsilon_t, \quad (6.10)$$

де $f(t)$ – основна тенденція, зумовлена впливом постійно діючих факторів (чинників);

ε_1 – залишкова величина, що характеризує випадкові, невраховані фактори.

Для виявлення основної тенденції застосовується ряд методів, зокрема, трендові криві (будуть детально розглядатися в 9-ому розділі), укрупнення інтервалів, плинна середня.

Використання методу **укрупнення інтервалів** пов'язано з тим, що в коротких інтервалах із-за впливу різноманітних факторів, що діють, як правило, в різних напрямках, в рядах динаміки має місце підвищення або зниження їх рівнів.

Суть цього методу полягає в тому, що окремі рівні ряду об'єднуються в групи по періодах, а потім розраховується середній показник рівня за цей період.

Метод плинних середніх має не тільки самостійне значення для вивчення тенденції, але й може бути використаний для попередньої обробки занадто варіюючих рядів динаміки.

Цей метод також передбачає укрупнення рядів динаміки, але воно проводиться шляхом послідовних зміщень на одну дату, залишаючи постійний інтервал часу.

Для методів укрупнення інтервалів і плинної середньої застосовуються, як правило, непарні інтервали ($i=3;5$ і т.д.).

В таблиці 6.7 наведені методики укрупнення інтервалів і визначення плинної середньої для виявлення тенденції зміни врожайності пшениці.

Таблиця 6.7 - Виявлення тенденції зміни врожайності пшениці

Укрупнення інтервалів				Плинна середня			
Рік	Фактична врожайність, ц/га	Сума врожайності за 5 років	Середня врожайність, ц/га	Рік	Фактична врожайність, ц/га	Сумарю плинних за 5-ти річним інтервалом	Середня плинна, ц/га
1	16,5	109,0	21,8	1	16,5	109,0	21,8
2	20,2			2	20,2		
3	27,7			3	27,7		
4	22,4			4	22,4		
5	22,2			5	22,2		
6	27,5	136,7	27,3	6	27,5	123,4	24,7
7	22,5			7	22,5	128,9	25,8
8	28,8			8	28,8	136,7	27,3
9	27,9			9	27,9	145,6	29,1
10	30,0			10	30,0	165,7	33,1
11	36,4	228,0	45,6	11	36,4	188,2	37,6
12	42,6			12	42,6	206,5	41,3
13	51,3			13	51,3	228,0	45,6
14	46,2			14	46,2		
15	51,5			15	51,5		

Аналіз результатів укрупнення інтервалів свідчить про наявність тенденції росту середньої врожайності пшениці, розрахованої за кожен п'ятиріччю.

Плинна середня показує неперервний ріст врожайності, тоді як початкові дані характеризуються постійним коливанням врожайності в сторону збільшення і зменшення.

Отже, зазначені методи дозволяють достатньо наглядно виявити основну тенденцію розвитку явища (процесу).

6.5 Статистичне вивчення сезонних коливань

Окремим соціально-економічним процесам притаманні сезонні коливання.

Сезонними коливаннями називаються більш менш сталі внутрішньорічні коливання в ряду динаміки, що зумовлені специфічними умовами виробництва і споживання певної продукції.

Прикладами виробництва, яким притаманні сезонні коливання, є будівництво, виробництво овочевих консервів, молока тощо.

Прагнення пом'якшити негативні наслідки сезонних коливань вимагає їх дослідження і оцінки.

Сезонні коливання характеризуються спеціальними показниками, які називаються **індексами сезонності**; сукупність їх створює **сезонну хвилю**.

Індекс сезонності за методом середньої визначається за формулою:

$$I_c = \frac{\bar{y}_i}{\bar{y}} \cdot 100, \quad (6.11)$$

де I_c - індекс сезонності;

\bar{y}_i - середнє значення за квартал (місяць);

\bar{y} - загальна середня (квартальна або місячна).

$$\bar{y}_i = \frac{\sum_{j=1}^n y_{ij}}{n}, \quad (6.12)$$

$$\bar{y} = \frac{\sum_{i=1}^m \bar{y}_i}{m} = \frac{\sum_{j=1}^n \sum_{i=1}^m y_{ij}}{n \cdot m}, \quad (6.13)$$

де y_{ij} - значення показника в i -му кварталі (місяці) j -го року ($i=1; 2; 3; \dots; m; j=1; 2; 3; \dots; n$).

Індекс сезонності, крім оцінки сезонних коливань, може бути використаний також для розподілу планового (прогнозного) річного

обсягу продукції (робіт, послуг) на наступний рік по кварталам (місяцям), для чого використовується така формула:

$$\hat{q}_i = \left[\left(\hat{Q} \cdot I_c \right) : m \right] : 100, \quad (6.14)$$

де \hat{q}_i – планове (прогнозне) значення показника на i -й квартал(місяць) наступного року;

\hat{Q} – планове (прогнозне) значення показника на весь наступний рік;

I_c – індекс сезонності;

m – кількість кварталів (місяців).

В таблиці 6.8 наведені дані про виробництво продукції, послідовність розрахунків показників для визначення індексу сезонності та розподіл річного плану виробництва продукції на наступний рік в обсязі 43.0 т по кварталам.

Таблиця 6.8 - Оцінка сезонних коливань обсягу продукції

Квар- тал	Обсяг продук- ції, тис. т			Разом за три роки (гр2+гр3+ +гр4)	Середнє значення $\bar{y}_i = \frac{y^5}{3}$	Індекс сезон- ності $\frac{y^6}{\bar{y}} \cdot 100$	Плановий об- сяг продукції на наступний рік, тис. т
	1-й рік	2-й рік	3-й рік				
I	11,0	11,7	11,2	33,9	11,3	108,1	11,6
II	9,3	10,5	9,8	29,6	9,9	94,7	10,2
III	8,4	9,2	9,4	27,0	9,0	86,2	9,3
IV	11,8	11,3	11,8	34,9	11,6	111,0	11,9
Разом	40,5	42,7	42,2	125,4	41,8	400,0	43,0
Середнє значення.	—	—	—	—	10,45	100,0	—

$$\bar{y} = \frac{\sum_{зр.6}}{4} = \frac{41,8}{4} = 10,45 \quad \text{або} \quad \bar{y} = \frac{\sum_{зр.5}}{3 \cdot 4} = \frac{125,4}{12} = 10,45.$$

В графі 8 наведені дані про розподіл річного плану виготовлення продукції на наступний рік по кварталам.

Питання для самоконтролю

1. В чому полягає суть і значення рядів динаміки в економічних дослідженнях?
2. З яких елементів складається ряд динаміки?
3. Які принципи і правила побудови рядів динаміки?
4. В чому відмінності моментного і інтервального рядів динаміки? Наведіть приклади кожного з них.
5. Як обчислюються середні в моментних і інтервальних рядах динаміки?
6. Назвіть основні статистичні характеристики рядів динаміки. За якими формулами вони розраховуються?
7. Якими методами вивчається основна тенденція в рядах динаміки?
8. В чому полягає суть і призначення індексу сезонності?

Задачі

6.1. Обсяг виробництва зерна в колективному господарстві характеризується такими даними:

Рік	1995	1996	1997	1998	1999	2000
Виробництво зерна, тис. ц	48	50	54	62	60	70

Визначити:

- а) абсолютні прирости – ланцюгові, базисні, середній;
- б) темп росту – ланцюговий, базисний, середній;
- в) темп приросту – ланцюговий, базисний, середній;
- г) розмір одного проценту приросту – ланцюговий, середній за весь період;
- д) середній рівень виробництва зерна.

6.2. Збір зернових по району характеризується такими даними:

Рік	Збір зерна, тис. т	Базисні показники динаміки		
		Абсолютний приріст, тис. т	Темп росту, %	Темп приросту, %
1994	120	—	100,0	—
1995	?	4,8	?	?
1996	?	?	106,8	?
1997	?	4,2	?	?
1998	?	?	?	1,6
1999	?	?	100,5	?
2000	?	7,3	?	?

Використовуючи взаємозв'язок показників динаміки, визначити:

- рівні ряду і відсутні в таблиці базисні показники динаміки;
- середнє значення рівня динаміки.

6.3. Виробництво продукції характеризується такими даними:

Рік	Вироблено продукції, тис. т	Ланцюгові показники динаміки			
		Абсолютний приріст, тис. т	Темп росту, %	Темп приросту, %	Абсолютне значення 1% приросту, тис. т
1994	81,5	—	—	—	—
1995	?	?	?	?	?
1996	?	2,5	?	?	0,836
1997	?	?	73,6	?	?
1998	?	?	?	5,7	?
1999	?	?	?	?	?
2000	?	?	108,6	?	0,720

Визначити:

- а) рівень виробництва продукції за наведені в таблиці роки і відсутні показники динаміки;
- б) середнє значення рівня виробництва продукції.

6.4. Ланцюгові темпи росту випуску продукції за роки п'ятирічки на заводі становили:

1996	1997	1998	1999	2000
102,1	99,8	102,0	101,0	103,0

Визначити:

- а) темп росту і приросту за п'ятирічку;
- б) середньорічний темп росту і приросту.

6.5. Ланцюгові темпи приросту прибутку підприємства характеризуються такими даними:

1996	1997	1998	1999	2000
6,1	5,4	7,2	7,4	7,6

Визначити:

- а) темп росту і приросту прибутку за п'ятирічку;
- б) середньорічний темп росту і приросту прибутку.

6.6. Виробництво продукції у 1995р. становило 2500 тис. грн., а в 2000р. в порівнянні з 1995р. воно збільшилось на 50%.

Визначити:

- а) середній річний абсолютний приріст;
- б) середній річний темп приросту;
- в) середнє абсолютне значення одного проценту приросту.

6.7. За період з 1995р. по 2000р. кількість робітників зросла на 20%. Абсолютне значення одного процента приросту становило 14 чол. Визначити кількість робітників в 2000р.

6.8. Абсолютний розмір одного процента приросту становив 2 млн. грн. Чому дорівнює обсяг продукції у звітному періоді, якщо в порівнянні з базисним він зріс на 2,5%.

6.9. Виробництво продукції за перші два роки п'ятирічки збільшилось щорічно в 1,2 рази, за третій рік воно підвищилось на 2,5%, а в останні два роки обсяг продукції щорічно знижувався на 2,0%.

Визначити:

- темпу росту і приросту виробництва продукції за п'ятирічку;
- середньорічний темпу росту і приросту виробництва продукції.

6.10. Є такі дані про виробничі запаси на складі:

Місяць	1.01	1.02	1.04	1.06	1.07	1.09	1.10	31.12
Виробничі запаси, Тис. грн.	118	124	128	132	136	135	141	148

Визначити середньомісячний запас товарів на складі.

6.11. Спискова чисельність робітників на початок звітної місяця становила:

Місяць	1.01	1.02	1.03	1.04	1.05	1.06	1.07	1.08	1.09	1.10	1.11	1.12	1.01
	2000р.												2001р.
Спискова чисельність, чол.	347	350	349	351	345	349	357	359	351	352	359	353	353

Визначити:

- середньоспискову чисельність робітників у першому і другому півріччях і за рік в цілому;

б) абсолютний приріст середньоспискової чисельності робітників і темп приросту у другому півріччі в порівнянні з першим півріччям.

6.12. Продуктивність праці щорічно зростала на 2,5%. На скільки підвищилась продуктивність праці за три роки?

6.13. Інфляція в поточному році в порівнянні з груднем попереднього року зросла на 25%. Визначити середньомісячний рівень інфляції у поточному році.

6.14. Темпи росту обсягу продукції по підприємству в порівнянні з 1990р. становили в 1995р. 105,1%; в 2000р.- 102,0%. Визначити середній річний темп росту і приросту продукції за 1996-2000рр.; 1991-1995рр.; 1991-2000рр.

Розділ 7 ВИБІРКОВЕ СПОСТЕРЕЖЕННЯ

7.1 Поняття і теоретичні основи вибіркового спостереження

Вибірковим називається таке несущільне спостереження, при якому спостереженню підлягають не всі одиниці сукупності, що вивчаються, а лише частина сукупності, яка відібрана за певними правилами.

Разом з тим відібрана сукупність в процесі обробки даних спостережень повинна забезпечити одержання таких результатів, які характеризували б усю сукупність в цілому. Цим забезпечується репрезентативність вибіркової сукупності, тобто властивість її відтворювати усю генеральну сукупність.

Вибіркові спостереження широко використовуються в промисловості (оцінка якості продукції; оцінка використання робочого часу); в сільському господарстві (визначення очікуваної врожайності; продуктивність худоби); в торгівлі (якість товарів, що надходять;

ефективність нових форм організації торгівлі, маркетингові дослідження); в соціології (опитування громадської думки).

Вибіркове спостереження має ряд переваг, зокрема:

а) відносно невеликі витрати сил і засобів, швидкість одержання результатів;

б) відсутність інколи альтернатив.

Щодо останньої переваги, то прикладом може бути оцінка споживчих властивостей або технічних характеристик окремих виробів, при обстеженні яких вироби повністю руйнуються або знищуються. Такі способи, природно, не можуть бути використані для оцінки якості всієї продукції.

Вибіркові спостереження дають можливість, не вдаючись до суцільного спостереження, одержати узагальнюючі показники, які правдиво характеризують усю сукупність одиниць.

Уся сукупність одиниць, з якої проводять відбір окремих одиниць, називається **генеральною (N)**, а та частина сукупності одиниць, яка відібрана і підлягає обстеженню, називається **вибірковою сукупністю (n)**.

При вибіркового спостереженні використовують дві категорії узагальнюючих показників: частку і середню величину.

Частка (p) визначається як відношення числа одиниць сукупності, яким притаманні ознаки, що цікавлять дослідника, до загальної кількості одиниць сукупності.

Завдання вибіркового спостереження полягає в тому, щоб на підставі оцінки вибіркової частки, дати вірну оцінку частки в генеральній сукупності.

Середнє значення варіюючої ознаки в усій сукупності одиниць називається **генеральною середньою (\bar{x})**, а середнє значення ознаки у

одиниць, які підлягають вибірковому спостереженню, називається **вибірковою середньою** (\bar{x}).

Завдання вибіркового спостереження в даному випадку полягає в тому, щоб на підставі вибіркової середньої одержати вірну уяву про середню генеральну.

Таким чином, першочергове завдання вибіркового спостереження полягає в тому, щоб на підставі вибірових оцінок визначити невідомі характеристики генеральної сукупності.

7.2 Схеми і способи відбору одиниць із генеральної сукупності

Кожне вибіркове спостереження проводиться за певними правилами і згідно з попередньо складеним планом. Відбір одиниць сукупності, що підлягають вибірковому обстеженню, може бути організований за різними схемами і способами.

Однак, за якими б схемами і способами не здійснювався відбір, слід додержуватися загального правила: у окремих одиниць генеральної сукупності повинна бути рівна можливість попасти в число одиниць, які підлягають обстеженню; точніше, відбір окремих одиниць сукупності повинен бути проведений у випадковому порядку.

В теорії вибіркового методу розглядаються дві схеми відбору – повторна і безповторна.

При **повторній вибірці** чисельність одиниць генеральної сукупності залишається незмінною, тому що кожна одиниця сукупності, яка попала у вибірку, після ресстрації знову повертається в генеральну сукупність, і вона, таким чином, зберігає рівні можливості з усіма іншими одиницями при новому відборі знову попасти у вибірку.

Безповторна – це така вибірка, при якій одиниця сукупності, що попала у вибірку, в подальших нових обстеженнях участі не приймає, тобто після першого обстеження вона не повертається в генеральну сукупність. Тому остання в процесі вибірових спостережень постійно скорочується після кожної нової вибірки.

Розрізняють такі види (способи) відбору:

а) простий випадковий (власне випадковий) відбір – здійснюється за допомогою жереба;

б) систематичний (механічний) відбір – із генеральної сукупності відбір проводиться в певному механічному порядку, наприклад, відбір кожної десятої (двадцятої) деталі, або для проведення опитування з телефонної книги за алфавітом відбирається кожне десяте (двадцяте і т.д.) прізвище;

в) типовий (розширений) відбір – орієнтований на забезпечення представництва у вибірці відповідних типових груп генеральної сукупності. При цьому генеральна сукупність розбивається (розширюється) на однотипні, однорідні групи. Потім з кожної групи за одним із вказаних вище методів відбирається число одиниць, пропорційних питомій вазі групи в загальній сукупності;

г) багатоступенева вибірка – це декілька ступенів (стадій) відбору. При цьому кожна стадія має свою одиницю відбору. Наприклад, галузь, підприємство, робітник.

д) багатофазна вибірка – вибірка сукупності створюється так, що одні відомості (дані) збираються з усіх одиниць сукупності, потім відбираються із вибіркової сукупності ще деякі одиниці для обстеження по більш широкій програмі;

е) серійна вибірка – проводиться випадковий відбір не окремих одиниць сукупності, а цілих серій (гнізд). В середині окремих серій проводиться обстеження усіх одиниць сукупності.

7.3 Середня і гранична помилки вибірки

Досягнути точного збігу узагальнюючих показників вибіркової і генеральної сукупності практично неможливо. Таке завдання і не ставиться в процесі досліджень. Мова може йти тільки про максимальне наближення показників вибіркової сукупності до показників генеральної сукупності, визначення можливої границі відхилення цих показників і умов, від яких залежить величина цих відхилень.

Можливі границі відхилень вибіркової частки і вибіркової середньої від частки і середньої в генеральній сукупності носять назву **помилки вибірки**.

У кожній окремій вибірці відхилення між узагальнюючими показниками у вибірковій і генеральній сукупності приймають різні значення. Тому мова може йти тільки про середню помилку вибірки.

Чим більша чисельність вибіркової сукупності, тим при інших рівних умовах, менша середня величина помилки вибірки.

При однаковій чисельності вибіркової сукупності середня помилка вибірки буде менша в тій сукупності, яка відібрана із генеральної сукупності з меншою мінливістю досліджуваної ознаки.

Звідси, помилка вибірки для середньої визначається за формулою:

$$\mu_x = \sqrt{\frac{\sigma^2}{n}}, \quad (7.1)$$

де μ_x – середня помилка вибірки для середньої;

σ^2 – дисперсія варіюючої ознаки;

n – чисельність одиниць вибіркової ознаки.

Середня помилка для частки дорівнює:

$$\mu_p = \sqrt{\frac{p(1-p)}{n}}, \quad (7.2)$$

$$p = \frac{m}{n}, \quad (7.3)$$

де μ_p – середня помилка вибірки для частки;

p – вибіркова частка ознак, що досліджуються;
 m – число одиниць у вибірковій сукупності, яким притаманні ознаки, що досліджуються;

n – число одиниць вибіркової сукупності.

Наприклад, якщо у відібраній партії із 100 деталей 5 виявилось бракованими, то $n = 100$, $m = 5$, $p = 5:100=0,05$.

Чисельник формули 7.2 є дисперсією для альтернативної ознаки (частки), що доводиться таким чином.

Ознаки, які мають одні одиниці сукупності, а інші не мають, називаються **альтернативними**.

Частка одиниць, які мають ознаку в чисельності всієї сукупності, позначають через p , а частка одиниць, що не мають цієї ознаки – через q .

Звідси:

$$p + q = 1, \text{ а } q = 1 - p.$$

Варіація альтернативної ознаки дорівнює нулю у одиниць, які не мають цієї ознаки. Звідси:

$$\bar{x} = \frac{1 \cdot p + 0 \cdot q}{p + q} = \frac{p}{p + q} = \frac{p}{1} = p. \quad (7.4)$$

Дисперсія альтернативної ознаки дорівнює:

$$\sigma^2 = \frac{(1-p)^2 p + (0-p)^2 q}{p+q} = q^2 p + p^2 q = pq(p+q) = p(1-p), \quad (7.5)$$

що і потрібно було довести.

Формули 7.1 і 7.2 відносяться до повторної вибірки. Для безповторної вибірки зазначені формули мають такий вигляд:

$$\text{для середньої } \mu_x = \sqrt{\frac{\sigma^2}{n} \left(1 - \frac{n}{N}\right)}, \quad (7.6)$$

$$\text{для частки } \mu_p = \sqrt{\frac{p(1-p)}{n} \left(1 - \frac{n}{N}\right)}. \quad (7.7)$$

Наведені формули середньої помилки характеризують тільки середнє відхилення узагальнюючих характеристик.

Однак стверджувати, що генеральна середня не вийде за певні границі, можна не з абсолютною достовірністю, а лише з певним ступенем ймовірності.

Для того, щоб визначити границі, в яких знаходиться середнє значення і частка в генеральній сукупності, визначається **гранична помилка вибірки**.

Між середньою і граничною помилкою вибірки існує таке співвідношення:

$$\Delta = t\mu, \quad (7.8)$$

де Δ - гранична помилка вибірки;

μ - середня помилка вибірки;

t - коефіцієнт довіри, який визначається в залежності від рівня ймовірності (P):

$$\text{для } P = 0,683 \quad t = 1,0,$$

$$\text{для } P = 0,954 \quad t = 2,0,$$

$$\text{для } P = 0,997 \quad t = 3,0.$$

В таблиці 7.1 наведені формули для визначення граничної помилки вибірки на основі узагальнювальної формули 7.8.

Таблиця 7.1 - Граничні помилки вибірки

Узагальнювальні характеристики	Схема відбору	
	повторна	безповторна
Середня	$\Delta\bar{x} = t\sqrt{\frac{\sigma^2}{n}}$	$\Delta\bar{x} = t\sqrt{\frac{\sigma^2}{n}\left(1 - \frac{n}{N}\right)}$
Частка	$\Delta\bar{p} = t\sqrt{\frac{p(1-p)}{n}}$	$\Delta\bar{p} = t\sqrt{\frac{p(1-p)}{n}\left(1 - \frac{n}{N}\right)}$

На підставі граничних помилок визначаються граничні значення середньої і частки в генеральній сукупності:

$$\bar{x} - \Delta x \leq \bar{x} \leq \bar{x} + \Delta \bar{x} \text{ - для середньої величини,} \quad (7.9)$$

$$\bar{p} - \Delta \bar{p} \leq \bar{p} \leq \bar{p} + \Delta \bar{p} \text{ - для частки.} \quad (7.10)$$

ПРИКЛАДИ

Приклад 1. Методом випадкової вибірки було взято для перевірки ваги 200 шт. деталей. В результаті було встановлено, що середня вага деталі 30г, при середньоквадратичному відхиленні 4г.

З ймовірністю 0,954 потрібно визначити границі, в яких знаходиться середня вага деталі в генеральній сукупності.

За умовою задачі:

$$\bar{x} = 30 \text{ г ;}$$

$$n = 200 \text{ ;}$$

$$\sigma = 4 \text{ г ;}$$

$$P = 0,954 \text{ ;}$$

$$t = 2,0 \text{ .}$$

Згідно з формулою 7.9 середня вага деталі в генеральній сукупності буде знаходитися в межах:

$$\bar{x} - \Delta \bar{x} \leq \bar{x} \leq \bar{x} + \Delta \bar{x} \text{ .}$$

Для визначення меж потрібно в першу чергу розрахувати $\Delta \bar{x}$:

$$\Delta \bar{x} = t \sqrt{\frac{\sigma^2}{n}} = 2 \sqrt{\frac{4^2}{200}} = 0,56$$

звідси:

$$30 - 0,56 \leq \bar{x} \leq 30 + 0,56 \text{ .}$$

$$29,44 \leq \bar{x} \leq 30,56 \text{ ,}$$

тобто з ймовірністю 0,954 можна стверджувати, що середня вага однієї деталі в генеральній сукупності буде знаходитися в межах 29,44 – 30,56.

Приклад 2. При дослідженні 100 зразків виробів, відібраних із партії у випадковому порядку, 20 виявились нестандартними. З ймовірністю 0,954 визначити межі, в яких знаходиться частка нестандартної продукції в генеральній сукупності.

За умовою задачі:

$$n = 100$$

$$m = 20$$

$$P = 0,954$$

$$t = 2,0$$

Згідно з формулою 7.10 частка нестандартної продукції у генеральній сукупності буде знаходитися в межах:

$$\bar{p} - \Delta\bar{p} \leq \bar{p} \leq \bar{p} + \Delta\bar{p}.$$

Частка нестандартної продукції у вибірці становить:

$$\bar{p} = \frac{m}{n} = \frac{20}{100} = 0,2.$$

Гранична помилка частки дорівнює:

$$\Delta\bar{p} = t \sqrt{\frac{p(1-p)}{n}} = 2 \sqrt{\frac{0,2(1-0,2)}{100}} = 0,08.$$

Частка нестандартної продукції в генеральній сукупності буде знаходитися в межах:

$$0,2 - 0,08 \leq \bar{p} \leq 0,2 + 0,08$$

$$0,12 \leq \bar{p} \leq 0,28,$$

тобто з ймовірністю 0,954 частка нестандартної продукції в генеральній сукупності буде знаходитися в межах 0,12 - 0,28.

Для серійної вибірки узагальнюючі характеристики визначаються за такими формулами:

середня

$$\mu_x = \sqrt{\frac{\delta^2}{s}} \text{ -- повторна вибірка,} \quad (7.11)$$

$$\mu_x = \sqrt{\frac{\delta^2}{s} \left(1 - \frac{s}{S}\right)} - \text{безповторна вибірка,} \quad (7.12)$$

частка

$$\mu_p = \sqrt{\frac{\delta_p^2}{s}} - \text{повторна вибірка,} \quad (7.13)$$

$$\mu_p = \sqrt{\frac{\delta_p^2}{s} \left(1 - \frac{s}{S}\right)} - \text{безповторна вибірка,} \quad (7.14)$$

де δ^2 – міжсерійна дисперсія;

s – число відібраних серій;

S – загальне число серій в генеральній сукупності;

δ_p^2 – міжсерійна (міжгрупова) дисперсія частки, яка визначається за формулою:

$$\delta_p^2 = \frac{\sum (p_i - \bar{p})^2}{S},$$

де p_i – частка ознаки в i -й серії;

\bar{p} – частка ознаки для всієї серії.

Середня гранична помилка для серійної вибірки визначається як добуток середньої помилки вибірки (формули 7.11, 7.12, 7.13, 7.14) на коефіцієнт довіри t .

В таблиці 7.2 наведені середні граничні помилки для серійної вибірки.

Таблиця 7.2 - Середні граничні помилки для серійної вибірки

Узагальнювальні характеристики	Схема відбору	
	повторна	безповторна
Середня	$\Delta \bar{x} = t \sqrt{\frac{\delta^2}{s}}$	$\Delta \bar{x} = t \sqrt{\frac{\delta^2}{s} \left(1 - \frac{s}{S}\right)}$
Частка	$\Delta \bar{p} = t \sqrt{\frac{\delta_p^2}{s}}$	$\Delta \bar{p} = t \sqrt{\frac{\delta_p^2}{s} \left(1 - \frac{s}{S}\right)}$

Визначення меж, в яких знаходиться середня в серійних вибірках, розглянемо на прикладі.

Із сукупності, що розбита на 100 рівних по величині серій, методом механічного відбору відібрано 10 серій. Міжсерійна дисперсія дорівнює 20, а середня величина ознаки у вибірці – 140. З ймовірністю 0,997 визначити межі, в яких знаходиться середня в генеральній сукупності:

За умовою задачі:

$$S = 100 ;$$

$$s = 10 ;$$

$$\delta^2 = 20 ;$$

$$\bar{x} = 140 ;$$

$$P = 0,997 ;$$

$$t = 3,0 .$$

Середня знаходиться в межах:

$$\bar{x} - \Delta\bar{x} \leq \bar{\bar{x}} \leq \bar{x} + \Delta\bar{x} .$$

Середня гранична помилка для серійної вибірки:

$$\Delta\bar{x} = t \sqrt{\frac{\delta^2}{s} \left(1 - \frac{s}{S}\right)} = 3,0 \sqrt{\frac{20}{10} \left(1 - \frac{10}{100}\right)} = 4 .$$

$$140 - 4 \leq \bar{\bar{x}} \leq 140 + 4 .$$

$$136 \leq \bar{\bar{x}} \leq 144 .$$

Отже, середня величина ознаки в генеральній сукупності з ймовірністю 0,997 буде в межах 136-144 одиниць.

8.4 Визначення необхідної чисельності вибірки

Важливим моментом при плануванні і організації вибіркового спостереження є визначення необхідної чисельності вибірки, тобто який обсяг вибіркової сукупності, який забезпечує необхідну точність і репрезентативність результатів.

Необхідна чисельність вибірки визначається на основі формул середньої граничної помилки (таблиця 7.3).

Таблиця 7.3 - Визначення обсягу вибіркової сукупності

Способи відбору	Схема відбору	
	повторна	безповторна
Механічний відбір: при визначенні середньої	$n = \frac{t^2 \sigma^2}{\Delta_x^2}$	$n = \frac{N t^2 \sigma^2}{N \Delta_x^2 + t^2 \sigma^2}$
при визначенні частки	$n = \frac{t^2 p(1-p)}{\Delta_p^2}$	$n = \frac{N t^2 p(1-p)}{N \Delta_p^2 + t^2 p(1-p)}$
Серійний відбір: при визначенні середньої	$s = \frac{t^2 \delta^2}{\Delta_x^2}$	$s = \frac{S t^2 \delta^2}{S \Delta_x^2 + t^2 \delta^2}$
при визначенні частки	$s = \frac{t^2 \delta_p^2}{\Delta_p^2}$	$s = \frac{S t^2 \delta_p^2}{S \Delta_p^2 + t^2 \delta_p^2}$

Визначення обсягу вибіркової сукупності розглянемо на прикладі.

Для визначення середньої довжини деталі необхідно провести вибіркове спостереження методом повторної випадкової вибірки.

Яку кількість деталей необхідно відібрати, щоб з ймовірністю 0,954 ($t=2$) помилка вибірки не перевищувала 2мм при середньоквадратичному відхиленні 8мм.

Для повторної вибірки при визначенні середньої помилки:

$$n = \frac{t^2 \sigma^2}{\Delta_x^2} = \frac{2^2 \cdot 8^2}{2^2} = 64 \text{ шт.}$$

7.5 Способи розповсюдження результатів вибіркового спостережень на генеральну сукупність

Кінцевою метою будь-якого вибіркового спостереження є поширення його узагальнюючих характеристик на генеральну сукупність.

На практиці застосовується два способи поширення результатів вибіркового спостереження:

- а) спосіб прямого перерахунку;
- б) спосіб коефіцієнтів.

Спосіб прямого перерахунку використовується в тому випадку, коли на основі вибірки обчислюють показники генеральної сукупності, використовуючи для цього вибіркові середні і вибіркові частки.

Так, визначивши середню вагу кореня цукрового буряка на спеціально виділених ділянках, розповсюджують ці дані для визначення врожайності і валового збору цукрових буряків по господарству в цілому.

Спосіб коефіцієнтів, як правило, використовують при проведенні вибіркового спостереження для перевірки і уточнення даних суцільного спостереження. В цьому випадку, порівнюючи дані вибіркового спостереження з суцільними, розраховують поправковий коефіцієнт, який і використовують для внесення поправок в матеріал суцільних спостережень.

Так, після закінчення перепису худоби проводять 10-ти процентне контрольне вибірконе спостереження домашніх господарств. При цьому інколи виявляють неповний облік худоби. За допомогою такої контрольної вибірки визначається поправковий коефіцієнт, який використовують для уточнення даних перепису.

ПИТАННЯ ДЛЯ САМОКОНТРОЛЮ.

1. В чому полягає суть і з якою метою використовується вибірконе спостереження?
2. Які схеми і методи відбору вам відомі? Наведіть приклади.

3. Чим відрізняється середня помилка вибірки від середньої граничної помилки вибірки?
4. Яким чином визначається середня помилка вибірки при різних схемах відбору?
5. В чому особливості серійної вибірки?
6. Яким чином визначається необхідний обсяг вибіркової сукупності?
7. Назвіть основні способи розповсюдження даних вибіркової сукупності на генеральну сукупність.

ЗАДАЧІ

7.1. Зі 100 вагонів вугілля, що надійшли на електростанцію, в порядку випадкової вибірки взяті проби для визначення зольності. На основі аналізу одержані такі дані:

Зольність	До 12	12-14	14-16	16-18	18-20	Понад 20
Кількість проб	5	10	35	25	15	10

З ймовірністю 0,954 ($t = 2,0$) визначити можливі межі, в яких очікується середня зольність вугілля.

7.2. Вибірковим обстеженням 10% виробів партії готової продукції (безповторна вибірка) одержані такі дані щодо вмісту вологи в зразках:

Вологість, %	До 14	14-16	16-18	18-20	Понад 20
Кількість проб	5	25	35	20	15

З ймовірністю 0,997 ($t = 3,0$) визначити можливі межі, в яких передбачається середній процент вологості виробів в усій партії.

7.4. Науково-дослідним інститутом для вивчення суспільної думки населення області щодо проведення певних заходів в порядку випадкового відбору було опитано 600 чол. З числа опитуваних 360 чол. схвалили заходи. З ймовірністю 0,997 ($t = 3,0$) визначити межі, в яких знаходиться частка осіб, що схвалили заходи.

7.5. Для вивчення думки студентів щодо проведення певних заходів із сукупності, що складалася з 10000 чол., методом випадкової безповторної вибірки опитано 600 студентів, 240 ухвалили план заходів. З ймовірністю 0,954 ($t=2,0$) визначити межі, в яких знаходиться частка студентів, що ухвалили заходи, у всій сукупності.

7.6. З 5 тис. чоловік, що здійснили правопорушення на протязі року, було обстежено 500 правопорушників методом механічного відбору. В результаті обстеження встановлено, що 300 чоловік вирости в ненормальних сімейних умовах.

З ймовірністю 0,997 ($t = 3,0$) визначити частку правопорушників, які вирости у ненормальних сімейних умовах, в генеральній сукупності.

7.7. В порядку випадкової вибірки було досліджено 900 дерев. При цьому був встановлений середній діаметр одного дерева 235 мм і коефіцієнт варіації 11,5%.

З ймовірністю 0,683 ($t = 1,0$) визначити межі, в яких знаходиться середній діаметр дерев.

7.8. Для оцінки врожайності картоплі з 800 господарських дворів було відібрано 50.

З ймовірністю 0,954 ($t = 2,0$) визначити середню граничну помилку врожайності при середньоквадратичному відхиленні 2,5.

7.9. В порядку випадкової повторної вибірки з партії взято 100 проб продукту А. В результаті дослідження встановлена середня вологість в вибірці 9,0% при середньоквадратичному відхиленні 1,5%.

З ймовірністю 0,954 ($t = 2,0$) визначити межі, в яких знаходиться середня вологість продукту А в партії.

7.10. З механічного цеху заводу на склад готової продукції надійшло 500 ящиків. Для встановлення частки бракованих деталей методом механічної вибірки було перевірено 10 ящиків. Результати перевірки показали, що частка бракованих деталей складає 10%. Дисперсія середньої вибірки дорівнює 25.

З ймовірністю 0,954 ($t = 2,0$) визначити межі, в яких знаходиться частка бракованих деталей в усій партії деталей, що надійшли на склад готової продукції.

7.11. З 5000 робітників необхідно встановити частку, що мають вищу освіту. Яка повинна бути чисельність вибірки, щоб з ймовірністю 0,954 ($t = 2,0$) помилка вибірки не перевищувала 0,05 при дисперсії, що дорівнює 0,25.

7.12. Сукупність розбита на 100 серій. Міжсерійна дисперсія дорівнює 20. Скільки серій потрібно відібрати безповторним методом, щоб з ймовірністю 0,997 ($t = 3,0$) помилка вибіркової середньої не перевищувала 4.

7.13. З метою визначення частки робітників, які попрацювали на заводі не більше 5 років, передбачається провести вибіркове дослідження. Яку кількість робітників необхідно відібрати, щоб з ймовірністю 0,683 ($t = 1,0$) помилка вибірки не перевищувала 3,0%; на основі обстеження відомо, що дисперсія типової вибірки дорівнює 2100.

7.14. З ймовірністю 0,954 ($t = 2,0$) визначити середні межі помилок вибірових середніх $\bar{x}_1 = 300$; $\bar{x}_2 = 30$, якщо коефіцієнти варіації відповідно дорівнюють 50% і 20%. Обсяг вибірки 400 одиниць.

РОЗДІЛ 8 ІНДЕКСИ

8.1 Визначення, класифікація і правила побудови індексів

Латинське слово “Індекс” означає “показник”, “показчик”.

Економічні індекси – це відносні величини, які характеризують зміну економічних явищ (процесів, об’єктів) впродовж часу (в динаміці), в просторі і в порівнянні з будь-яким еталоном (планом, нормативами, кращими зразками тощо).

Звідси найпоширенішими є індекси динаміки, просторові (територіальні); виконання плану.

Індекси передбачають співвідношення одного і того ж економічного явища.

Тому не кожна відносна величина є індексом. Не відносяться до індексів відносні величини структури, координації, інтенсивності.

Існує декілька ознак, згідно з якими проводиться класифікація індексів.

За ступенем охоплення елементів сукупності розрізняють індивідуальні і зведені індекси. Останні в свою чергу поділяються на загальні і групові.

Індивідуальні індекси (і) відображають зміну в динаміці або просторі тільки одного елемента складного явища (наприклад, обсяг виробництва певного виду продукції, ціна, собівартість певного виду продукції тощо).

Зведені індекси (І) відображають зміну сукупності елементів складного явища (наприклад, зміну обсягу різнорідних видів продукції).

Якщо індекси охоплюють зміну не всіх елементів складного явища, а тільки певну частину (їх групу), то такі індекси називають **груповими або субіндексами**.

За характером явищ, що вивчаються, індекси поділяють на індекси кількісних і індекси якісних показників.

Індекси кількісних показників відображають зміну обсягу сукупності (обсяг продукції, чисельність працюючих, посівну площу та інші показники, що відображають кількісну ознаку).

Індекси якісних показників характеризують зміну показників, які визначені в середньому на одну одиницю сукупності (ціна, собівартість одиниці продукції, урожайність певної культури, продуктивність праці, фондovіддача, матеріалоємність тощо).

Загальні і групові індекси в залежності від методології обчислення поділяють на агрегатні (підсумкові) і середні з індивідуальних індексів.

Агрегатна форма індексів розглядається як основна, а середні з індивідуальних індексів – похідні, які одержують в результаті перетворення агрегатних індексів.

За допомогою індексів вирішуються такі задачі:

- а) вимір динаміки середнього значення економічного показника;
- б) вимір співвідношення показників середніх регіонів;
- в) проведення факторного аналізу;
- г) приведення показників до порівнянних цін.

Не виникає складностей, коли потрібно визначити індекс як зміну однорідного елемента складного явища (обсяг певного виду продукції, ціни на певний вид продукції, урожайність певної культури тощо), тобто індивідуальний індекс.

Проте в економічних дослідженнях найчастіше доводиться вивчати зміни складних явищ, що формуються з різнорідних елементів, які не можна підсумовувати, наприклад, різнорідні за споживчою вартістю види продукції в натуральному вимірі.

Неприпустимо також підсумовувати ціни, собівартість одиниці окремих видів продукції, урожайність окремих культур тощо.

Для побудови індексів, які характеризують зміну сукупності різнорідних елементів складного явища або якісних показників, які розраховані у відповідності до однієї одиниці (ціна, собівартість, трудомісткість тощо) використовуються **показники-сумірники** (співмножники, ознака-вага).

Такими показниками-сумірниками для якісних величин є система натуральних показників. Так, для цін, собівартості, трудомісткості – це фізичний обсяг продукції, для урожайності культур – посівна площа і т.д.

Для узагальнення різнорідних видів продукції показниками сумірниками є ціна, собівартість, трудомісткість.

Вся теорія індексів спрямована перш за все на визначення змін різнорідних елементів складного явища в цілому.

Показник, зміна якого визначається за допомогою індексу, називається **індексовим**.

Слід звернути увагу на те, що добуток індексового показника і показника-сумірника, які пов'язані між собою, повинен відображати певний економічний зміст або, іншими словами, визначити більш складний показник.

Введемо такі умовні позначення:

g - кількість виготовленого (або проданого) товару (продукції) певного виду у натуральному вимірі;

p - ціна одиниці товару (продукції);

z - собівартість одиниці продукції;

t-трудомісткість або затрати робочого часу на виготовлення одиниці певної продукції;

w - продуктивність праці (виробіток продукції на одного працюючого);

y - урожайність певної культури;

п- розмір посівної площі під певну культуру;

"0" - порядковий знак показника базисного періоду;

"пл"- позначення показника згідно плану;

"1"- порядковий номер показника звітнього періоду.

Якщо зміна явища досліджується за декілька періодів, то кожен з них позначається відповідно «0», «1», «2», «3» і т.д.

Виходячи з цих позначень, а також змісту певних техніко-економічних показників, добуток індексовного показника на показник-сумірник характеризує:

pq - загальну вартість виготовленої або проданої продукції (товарообіг певного виду);

zq - витрати на виробництво певного виду продукції;

tq –загальні витрати робочого часу(праці) на виготовлення певного виду продукції;

уп – валовий збір певної сільськогосподарської культури.

Із огляду на приведене вище можна дати таке визначення агрегатного індексу.

Агрегатний індекс динаміки (або виконання плану) – це дріб, в чисельнику якого сума добутоків індексовного показника за звітний період на показник-сумірник, а в знаменнику сума добутоків індексовного показника за базисний період (або планове значення) на цей же показник-сумірник, причому показник-сумірник приймається за звітний період, якщо індексовна величина є якісною, і за базисний період (план), якщо індексовна величини кількісна.

Індивідуальний індекс – це по суті коефіцієнт(темп) росту або рівень виконання плану певного показника.

Грунтуючись на змісті і принципах побудови індексів, в таблицях 8.2. наведені відповідно індивідуальні і агрегатні (зведені) індекси окремих показників.

Таблиця 8.1 - Індивідуальні індекси та формули для їх обчислення

Назва індексів	Розрахункова формула
1. Фізичного обсягу виготовленої (проданої) продукції	$i_q = q_1 : q_0$
2. Ціни	$i_p = p_1 : p_0$
3. Собівартості одиниці продукції	$p_z = z_1 : z_0$
4. Трудомісткості окремого виду продукції	$i_t = t_1 : t_0$
5. Урожайності певної культури	$i_y = y_1 : y_0$
6. Продуктивності праці	$i_w = w_1 : w_0$

Таблиця 8.2 - Агрегатні індекси та формули для їх обчислення

Назва індексів	Розрахункова формула
1. Фізичного обсягу виготовленої (проданої) продукції	$I_q = \sum q_1 p_0 : \sum q_0 p_0$
2. Ціни	$I_p = \sum q_1 p_1 : \sum q_1 p_0$
3. Собівартості	$I_z = \sum q_1 z_1 : \sum q_1 z_0$
4. Трудомісткості	$I_t = \sum q_1 t_1 : \sum q_1 t_0$
5. Купівельної спроможності грошової одиниці	$I_{к.с.гр.од.} = 1 : I_p = \sum q_1 p_0 : \sum q_1 p_1$

Аналогічно, користуючись теоретико-методологічними положеннями і визначеннями, можна побудувати індивідуальні і агрегатні індекси і для інших показників.

Індекси можуть бути виражені у вигляді коефіцієнтів, якщо базисний рівень приймається за 1, або у вигляді відсотків, якщо базисний рівень приймається за 100.

Кожний індекс має певний економічний зміст. Так, індивідуальний індекс фізичного обсягу показує, у скільки разів змінюється фізичний обсяг продукції у звітному періоді у порівнянні з базисним (планом) або скільки відсотків становить ріст (зниження) фізичного обсягу продукції у звітному періоді у порівнянні з базисним (планом).

Агрегатний індекс фізичного обсягу показує, у скільки разів змінюється вартість продукції в результаті змін обсягу її виробництва, або скільки процентів становить ріст (зниження) вартості продукції через зміни її фізичного обсягу.

Враховуючи, що індивідуальні індекси інших показників інтерпретуються аналогічно індивідуальному індексу фізичного обсягу, розглянемо зміст агрегатних індексів ще деяких показників.

Агрегатний індекс ціни показує, у скільки разів змінюється вартість продукції в результаті зміни ціни або скільки процентів становить ріст (зниження) вартості продукції через зміну цін.

Агрегатний індекс собівартості показує, у скільки разів знизилась витрати виробництва в результаті зміни собівартості окремих видів продукції або скільки процентів становить ріст (зниження) витрат виробництва через зміну собівартості окремих видів продукції.

Аналогічно і для інших показників.

У процесі використання індексів у техніко-економічних розрахунках слід звернути увагу на такі правила:

1. Індекси пов'язані між собою у тому ж відношенні, що і показники, зміну яких вони відображають. Наприклад, якщо результативний показник є добуток двох інших показників (факторів), то і індекс результативного показника дорівнює добутку індексів показників-факторів.

Так, товарообіг є добутком фізичної кількості продукції на її ціну, тому і індекс товарообігу дорівнює добутку індексу фізичного обсягу на індекс цін, тобто:

$$i_{gp} = i_g \cdot i_p, \quad \text{звідси } i_{gp} = \frac{g_1}{g_0} \cdot \frac{p_1}{p_0} = \frac{g_1 p_1}{g_0 p_0}. \quad (8.1)$$

Витрати виробництва

$$i_{gz} = i_g \cdot i_z, \quad i_{gz} = \frac{g_1}{g_0} \cdot \frac{z_1}{z_0} = \frac{g_1 z_1}{g_0 z_0}. \quad (8.2)$$

Валовий збір певної культури

$$i_{ny} = i_n \cdot i_y, \quad i_{ny} = \frac{n_1}{n_0} \cdot \frac{y_1}{y_0} = \frac{n_1 y_1}{n_0 y_0}. \quad (8.3)$$

Аналогічно і для агрегатних індексів.

Товарообіг

$$I_{gp} = \frac{\sum g_1 p_0}{\sum g_0 p_0} \cdot \frac{\sum g_1 p_1}{\sum g_1 p_0} = \frac{\sum g_1 p_1}{\sum g_0 p_0}. \quad (8.4)$$

Витрати виробництва

$$I_{gz} = I_g \cdot I_z, \quad I_{gz} = \frac{\sum g_1 z_0}{\sum g_0 z_0} \cdot \frac{\sum g_1 z_1}{\sum g_1 z_0} = \frac{\sum g_1 z_1}{\sum g_0 z_0}. \quad (8.5)$$

Валовий збір культур

$$I_{ny} = I_n \cdot I_y, \quad I_{ny} = \frac{\sum n_1 y_0}{\sum n_0 y_0} \cdot \frac{\sum n_1 y_1}{\sum n_1 y_0} = \frac{\sum n_1 y_1}{\sum n_0 y_0}. \quad (8.6)$$

2. Добуток індивідуальних ланцюгових індексів дорівнює кінцевому базисному індексу:

$$\frac{q_1}{q_0} \cdot \frac{q_2}{q_1} = \frac{q_2}{q_0}. \quad (8.7.)$$

3. Частка від ділення індивідуального базисного індексу поточного періоду на індивідуальний базисний індекс попереднього періоду дорівнює ланцюговому індексу поточного періоду:

$$\frac{q_2}{q_0} : \frac{q_1}{q_0} = \frac{q_2}{q_1} \quad (8.8)$$

Методику обчислення та інтерпретацію індивідуальних та агрегатних індексів проілюструємо на основі даних таблиці 8.3.

Таблиця 8.3 - Обсяг товарообігу

Виріб	Ціна одиниці (грн.)		Продано одиниць товарів (шт.)		Проміжні розрахунки		
	базисний період	звітний період	базисний період	звітний період	$q_0 p_0$	$q_1 p_1$	$q_1 p_0$
	p_0	p_1	q_0	q_1			
А	0,30	0,25	1540	1760	462	440	528
Б	1,30	1,20	460	520	598	624	676
Разом	-	-	-	-	1060	1064	1204

Необхідно визначити:

а) індивідуальні індекси фізичного обсягу, цін і товарообігу;

б) загальні індекси фізичного обсягу, цін і товарообігу.

Індивідуальні індекси визначаються за формулами:

$$i_q = \frac{q_1}{q_0}; \quad i_p = \frac{p_1}{p_0}; \quad i_{qp} = \frac{q_1 p_1}{q_0 p_0}$$

Індивідуальні індекси за даними таблиці 8.3 становлять:

Індекси фізичного обсягу

$$i_{q_a} = 1760 : 1540 = 1,143 \quad \text{або } 114,4\%$$

$$i_{q_B} = 520 : 460 = 1,130$$

або 113,0%.

Індекси цін

$$i_{p_A} = 0,25 : 0,30 = 0,833$$

або 83,3%;

$$i_{p_B} = 1,20 : 1,30 = 0,923$$

або 92,3%.

Індекси товарообігу

$$i_{qp_A} = 440 : 462 = 0,952$$

або 95,2%;

$$i_{qp_B} = 624 : 598 = 1,043$$

або 104,3%.

Перевіримо взаємозв'язок індексів

$$i_{qp} = i_q \cdot i_p$$

$$i_{qp_A} = i_{q_A} \cdot i_{p_A} = 1,143 \cdot 0,833 = 0,952,$$

або 95,2%;

$$i_{qp_B} = i_{q_B} \cdot i_{p_B} = 1,130 \cdot 0,923 = 1,043,$$

або 104,3%.

Отже, як показує перевірка через взаємозв'язок індексів, всі розрахунки виконані вірно.

Результати розрахунків свідчать про те, що фізичний обсяг проданих товарів А і Б виріс відповідно на $114,3 - 100,0 = 14,3\%$ і $113,0 - 100 = 13,0\%$.

На відміну від фізичного обсягу, ціни відповідно знизились на $83,3 - 100 = -16,7\%$ і $92,3 - 100 = -7,7\%$.

Зниження товарообігу продукції А пов'язано з тим, що темп зростання фізичного обсягу був нижчий темпу зниження цін.

Для оцінки зміни фізичного обсягу, цін і товарообігу по всій продукції використасмо відповідні загальні індекси в формі агрегатних індексів.

Індекс фізичного обсягу:

$$I_q = \frac{\sum q_1 p_0}{\sum q_0 p_0} = 1204 : 1060 = 1,136$$

або 113,6%.

Індекс цін

$$I_p = \frac{\sum q_1 p_1}{\sum q_1 p_0} = 1064:1204 = 0,884 \quad \text{або } 88,4 \%$$

Індекс товарообігу

$$I_{qp} = \frac{\sum q_1 p_1}{\sum q_0 p_0} = 1064:1060 = 1,004 \quad \text{або } 100,4\%$$

Перевіримо взаємозв'язок індексів

$$I_{qp} = I_q \cdot I_p = 1,036 \cdot 0,884 = 1,004$$

Тобто, всі розрахунки виконані вірно.

Товарообіг, як свідчать розрахунки, у звітному періоді виріс в порівнянні з базисним всього на 0,4%, причому за рахунок змін фізичного обсягу він виріс на 13,6%, проте зниження цін призвело до зниження товарообігу на 11,6%.

8.2 Середні індекси

В практиці індексних розрахунків бувають випадки, коли відсутні абсолютні значення індексовних ознак і показників-сумірників, а є тільки їх відносні величини, які показують зміну зазначених показників у звітному періоді в порівнянні з базисним (або планом), тобто індивідуальні індекси.

В таких випадках загальну зміну індексовної ознаки визначають як середню із індивідуальних індексів, і такий індекс має назву **середнього індексу**.

Перетворення агрегатного індексу в середній із індивідуальних індексів проводиться так:

- в чисельнику або в знаменнику агрегатного індексу індексовний показник замінюється його виразом через відповідний індивідуальний індекс. Якщо таку заміну зробити в чисельнику, то агрегатний індекс

перетвориться в середньоарифметичний зважений, якщо ж у знаменнику, то у середньогармонійний зважений.

Як правило, середньоарифметичний зважений індекс використовується тоді, коли індексовна величина кількісна, а середньогармонійний зважений індекс – коли індексовна величина якісна.

Агрегатні індекси перетворюються в середні таким чином.

Агрегатний індекс фізичного обсягу

$$I_q = \frac{\sum g_1 p_0}{\sum g_0 p_0}; \text{ так як } i_x = \frac{g_1}{g_0}, \text{ то } g_1 = i_x g_0.$$

Якщо підставити $q_1 = i_q \cdot q_0$ в агрегатну форму, то одержимо середній індекс кількісного показника – фізичного обсягу

$$I_q = \frac{\sum g_1 p_0}{\sum g_0 p_0} = \frac{\sum i_x g_0 p_0}{\sum g_0 p_0} \quad (8.9)$$

Агрегатний індекс ціни

$$I_p = \frac{\sum g_1 p_1}{\sum g_1 p_0}; i_p = \frac{p_1}{p_0} \Rightarrow p_0 = \frac{p_1}{i_p} \quad (8.10)$$

Звідси середній індекс якісного показника - ціни буде мати такий вигляд:

$$I_p = \frac{\sum g_1 p_1}{\sum \frac{g_1 p_1}{i_p}} \quad (8.10)$$

Середній індекс фізичного обсягу – це середня арифметична, а середній індекс ціни – середня гармонійна.

Розрахунки середніх індексів проілюструємо на основі даних таблиці 8.4.

Таблиця 8.4 - **Обсяг товарообігу**

Продукція	Товарообіг, тис.грн		Індивідуальні індекси	
	базисний період	звітний період	фізичного обсягу	цін
	$q_0 p_0$	$q_1 p_1$	i_q	i_p
Молоко	1500	2025	1,5	0,9
Масло	1450	1740	1,2	1,0
Яйця	390	936	3,0	0,8

Необхідно визначити загальні індекси фізичного обсягу, ціни і товарообігу. Перевірити взаємозв'язок індексів.

Оскільки в умові задачі відсутні абсолютні значення проданих товарів і цін на них, то для розв'язання задачі необхідно скористатись середніми індексами.

Обчислимо індекс фізичного обсягу:

$$I_q = \frac{\sum g_1 p_0}{\sum g_0 p_0} = \frac{\sum i_q g_0 p_0}{\sum g_0 p_0} = \frac{1,5 \cdot 1500 + 1,2 \cdot 1450 + 3,0 \cdot 390}{1500 + 1450 + 390} = \frac{5160}{3340} = 1,545, \text{ або } 154,5\%.$$

тобто товарообіг у звітному періоді виріс в порівнянні з базисним за рахунок зміни фізичного обсягу на 54,5%.

Індекс ціни дорівнює :

$$I_p = \frac{\sum g_1 p_1}{\sum g_1 p_0} = \frac{2025 + 1740 + 936}{\frac{2025}{0,9} + \frac{1740}{1,0} + \frac{936}{0,8}} = \frac{4701}{5160} = 0,911, \text{ або } 91,1\%.$$

Значення індексу цін свідчить, що товарообіг у звітному періоді в порівнянні з базисним зменшився за рахунок зниження ціни на 8,9%.

Загальний індекс товарообігу становить:

$$I_{pq} = \frac{\sum g_1 p_1}{\sum g_0 p_0} = \frac{2025 + 1740 + 936}{1500 + 1450 + 390} = \frac{4701}{3340} = 1,407 \text{ або } 140,7\%.$$

Товарообіг, таким чином, виріс за досліджуваний період на 40,7%.

Взаємозв'язок індексів:

$$I_{qr} = I_q \cdot I_r = 1,545 \cdot 0,911 = 1,407,$$

що свідчить про вірність розрахунків.

Якщо в умові задачі подані не індивідуальні індекси, а зміна індексного показника у звітному періоді в порівнянні з базисним (темпом приросту), то для обчислення індивідуального індексу необхідно до 100 додати величину зміни, якщо вона додатна, або від 100 відняти, якщо вона від'ємна, а потім для зручності розрахунків перевести значення індивідуальних індексів, оцінених в процентах, в коефіцієнти шляхом ділення на 100.

8.3 Індекси середніх величин

В економічних дослідженнях нерідко доводиться аналізувати і порівнювати різні середні величини: середню ціну, середню собівартість, середню урожайність, середню продуктивність праці, середню заробітну плату тощо. Такі середні обчислюються на основі групових середніх, а показниками-сумірниками (вагою) повинні бути кількісні показники, які можна підсумовувати. Останніми, наприклад, може бути однойменна продукція, яка виготовляється на декількох підприємствах, або різнойменна, але однієї споживчої вартості продукція, яка виробляється на окремому підприємстві:

Середній рівень, який визначається для якісних показників, умовно обчислюється за формулою:

$$\bar{X} = \frac{\sum xf}{\sum f} = \sum x \cdot d, \quad (8.11)$$

де \bar{X} – середня величина якісної ознаки ;

f – частота (кількісна ознака) ;

d – частка окремих елементів кількісної ознаки.

Із формули 8.11 видно, що середня величина залежить від значення усередненого показника (x) і частоти (величини окремих варіант кількісної ознаки). Проте на середню впливає не абсолютна величина частот, а співвідношення між ними, тобто структура кількісної ознаки.

Відповідно динаміка середньої визначається зміною усередненої величини і зміною співвідношення між частотами (структурними зрушеннями).

Динаміку середньої величини певного показника характеризує **індекс змінного складу**.

Зміна усередненого показника описується **індексом фіксованого складу**.

Нарешті, зміна співвідношення частот описується **індексом структурних зрушень**.

Так, для ціни індекси змінного складу, фіксованого складу і структурних зрушень записуються відповідно такими формулами:

$$I_p^k = \frac{\sum g_1 p_1}{\sum g_1} : \frac{\sum g_0 p_0}{\sum g_0} = \bar{p}_1 : \bar{p}_0; \quad (8.12)$$

$$I_p^{fc} = \frac{\sum g_1 p_1}{\sum g_1} : \frac{\sum g_1 p_0}{\sum g_1}; \quad (8.13)$$

$$I_p^{cs} = \frac{\sum g_1 p_0}{\sum g_1} : \frac{\sum g_0 p_0}{\sum g_0}; \quad (8.14)$$

$$I_p^k = I_p^{fc} \cdot I_p^{cs}. \quad (8.15)$$

Приклад обчислення системи середніх індексів при дослідженні динаміки ціни розглянемо на основі даних таблиці 8.5.

Таблиця 8.5 - Обсяг продажу картоплі

Ринок	Ціна за 1 т., грн.		Продано картоплі, т	
	базисний	звітний	базисний	звітний
	період	період	період	період
	p_0	p_1	q_0	q_1
А	300	250	100	140
Б	200	150	120	300

Необхідно визначити індекси цін змінного складу, фіксованого складу і структурних зрушень та перевірити взаємозв'язок індексів.

Обчислимо систему індексів для оцінки динаміки середніх цін.

Індекс ціни картоплі змінного складу

$$I_z^c = \frac{\sum g_1 p_1}{\sum g_1} \cdot \frac{\sum g_0 p_0}{\sum g_0} = \frac{140 \cdot 250 + 300 \cdot 150}{140 + 300} \cdot \frac{100 \cdot 300 + 120 \cdot 200}{100 + 120} = \frac{80000}{440} \cdot \frac{5400}{220} = 1818,2455 \approx 0,74 \text{ або } 74,0\%$$

тобто, середня ціна картоплі на двох ринках у звітному періоді в порівнянні з базисним періодом знизилась на 26,0%.

Індекс ціни картоплі фіксованого складу

$$I_p^c = \frac{\sum g_1 p_1}{\sum g_1} \cdot \frac{\sum g_1 p_0}{\sum g_1} = 1818 \cdot \frac{140 \cdot 300 + 300 \cdot 200}{140 + 300} = \frac{1818}{2318} = 0,784 \text{ або } 78,4\%$$

тобто, ціна 1 т картоплі в середньому знизилась на 21,6%.

Індекс структурних зрушень

$$I_p^s = \frac{\sum g_1 p_0}{\sum g_1} \cdot \frac{\sum g_0 p_0}{\sum g_0} = 2318 \cdot \frac{2455}{2200} = 0,944 \text{ або } 94,4\%$$

тобто, за рахунок структурних зрушень середня ціна картоплі знизилась на 5,6%. Це пов'язано з тим, що на ринку Б, де ціна 1 т картоплі нижча в порівнянні з ціною на ринку А, питома вага проданої картоплі у звітному періоді виросла в порівнянні з базисним:

$300 : (140 + 300) \cdot 100 = 68,2\%$ - у звітному періоді і

$120 : (100 + 120) \cdot 100 = 54,4\%$ - у базисному періоді.

Взаємозв'язок індексів:

$$I^x = I^{bc} \cdot I^{ct} = 0,784 \cdot 0,944 = 0,740.$$

Отже, розрахунки виконані вірно.

8.4 Ланцюгові і базисні індекси

Індекси, як відносні показники динаміки, можуть бути ланцюговими і базисними.

Якщо необхідно досліджувати явища за декілька періодів, то ланцюгова система індексів передбачає порівняння з попереднім періодом – **система індексів зі змінною базою порівняння**; **базисна система індексів** передбачає порівняння явищ за різні періоди з аналогічним явищем, взятим за певний період (базою) – **система індексів з постійною базою**.

Якщо позначити чотири послідовних періоди порядковими цифрами 0; 1; 2; 3, то індивідуальні індекси певних показників, наприклад, фізичного обсягу і цін, можна записати у вигляді, як це показано в табл.8.6.

Таблиця 8.6 - Індивідуальні індекси

Індекси	Базисні	Ланцюгові
Фізичного обсягу	$\frac{q_1}{q_0} \cdot \frac{q_2}{q_0} \cdot \frac{q_3}{q_0}$	$\frac{q_1}{q_0} \cdot \frac{q_2}{q_1} \cdot \frac{q_3}{q_2}$
Цін	$\frac{p_1}{p_0} \cdot \frac{p_2}{p_0} \cdot \frac{p_3}{p_0}$	$\frac{p_1}{p_0} \cdot \frac{p_2}{p_1} \cdot \frac{p_3}{p_2}$

Для індивідуальних індексів (розглянемо на прикладі індексів фізичного обсягу) справедливі такі правила:

а) добуток проміжних за періодами ланцюгових індексів дорівнює базисному індексу за весь період:

$$\frac{q_1}{q_0} \cdot \frac{q_2}{q_1} \cdot \frac{q_3}{q_2} = \frac{q_3}{q_0} \quad \text{або} \quad I_{1/0} \cdot I_{2/1} \cdot I_{3/2} = I_{3/0}; \quad (8.16)$$

б) відношення базисного індексу звітнього періоду до базисного індексу попереднього періоду дорівнює ланцюговому індексу звітнього періоду:

$$\frac{q_3}{q_0} : \frac{q_2}{q_0} = \frac{q_3}{q_2} \quad \text{або} \quad I_{3/0} : I_{2/0} = I_{3/2}. \quad (8.17)$$

Система агрегатних індексів наведена в таблиці 8.7.

Таблиця 8.7 - Агрегатні індекси

Індекси	Базисні			Ланцюгові		
Фізичного обсягу	$\frac{\sum q_1 p_0}{\sum q_0 p_0}$	$\frac{\sum q_2 p_0}{\sum q_0 p_0}$	$\frac{\sum q_3 p_0}{\sum q_0 p_0}$	$\frac{\sum q_1 p_0}{\sum q_0 p_0}$	$\frac{\sum q_2 p_0}{\sum q_1 p_0}$	$\frac{\sum q_3 p_0}{\sum q_2 p_0}$
Цін	$\frac{\sum q_1 p_1}{\sum q_1 p_0}$	$\frac{\sum q_2 p_2}{\sum q_2 p_0}$	$\frac{\sum q_3 p_3}{\sum q_3 p_0}$	$\frac{\sum q_1 p_1}{\sum q_1 p_0}$	$\frac{\sum q_2 p_2}{\sum q_2 p_1}$	$\frac{\sum q_3 p_3}{\sum q_3 p_2}$

Для агрегатних індексів кількісних показників, зокрема, фізичного обсягу справедливі такі правила:

а) відношення базисного індексу звітнього періоду до базисного індексу попереднього періоду дорівнює ланцюговому індексу звітнього періоду

$$\frac{\sum q_3 P_0}{\sum q_0 P_0} : \frac{\sum q_2 P_0}{\sum q_0 P_0} = \frac{\sum q_3 P_0}{\sum q_2 P_0} \quad \text{або} \quad I_{3/0} : I_{2/0} = I_{3/2}; \quad (8.18)$$

б) базисний індекс звітнього періоду може бути отриманий шляхом множення відповідних ланцюгових індексів

$$\frac{\sum q_1 P_0}{\sum q_0 P_0} \cdot \frac{\sum q_2 P_0}{\sum q_1 P_0} \cdot \frac{\sum q_3 P_0}{\sum q_2 P_0} = \frac{\sum q_3 P_0}{\sum q_0 P_0} \quad \text{або} \quad I_{1/0} \cdot I_{2/1} \cdot I_{3/2} = I_{3/0}. \quad (8.19)$$

На агрегатні індекси якісних показників, зокрема ціни, зазначені правила не поширюються.

8.5 Індексний факторний аналіз

В економічних дослідженнях важливе значення має оцінка впливу окремих факторів на рівень результативного показника, тобто проведення факторного аналізу.

Індексний метод широко використовується в факторному аналізі, однак тільки в тих випадках, коли між результативним і факторними показниками є функціональний зв'язок.

Існує декілька методів проведення індексного факторного аналізу. Вибір певного методу залежить в першу чергу від наявної інформації і виду функціональної залежності.

Розглянемо два основних види функціонального зв'язку:

- факторна система являє собою взаємозв'язок агрегатних індексів;
- факторна система відображає залежність результативного показника від добутку двох або більше факторів

$$(y = a \cdot b \cdot c \cdot d)$$

Якщо факторна система зображена у вигляді взаємозалежностей агрегатних індексів, то слід звернути увагу на такі положення:

а) індекс результативного показника дорівнює добутку індексів показників-факторів;

б) різниця між чисельником і знаменником індексу показує абсолютну зміну показника у звітному періоді в порівнянні з базисним (планом);

в) зміна результативного показника дорівнює сумі змін показників-факторів.

Відносно товарообігу зазначені положення записуються відповідно так:

$$I_{pq} = I_p \cdot I_q; \quad I_{pq} = \frac{\sum q_1 p_1}{\sum q_0 p_0}; \quad I_q = \frac{\sum q_1 p_0}{\sum q_0 p_0}; \quad I_p = \frac{\sum q_1 p_1}{\sum q_1 p_0};$$

$$\Delta pq = \sum q_1 p_1 - \sum q_0 p_0; \quad \Delta p = \sum q_1 p_1 - \sum q_1 p_0; \quad \Delta q = \sum q_1 p_0 - \sum q_0 p_0;$$

$$\Delta pq = \Delta p + \Delta q = (\sum q_1 p_1 - \sum q_0 p_0) + (\sum q_1 p_0 - \sum q_0 p_0) = \sum q_1 p_1 - \sum q_0 p_0.$$

Методику проведення факторного аналізу на основі взаємозалежностей агрегатних індексів розглянемо на прикладах.

Приклад 1. Реалізація продукції (таблиця 8.8.) характеризується такими даними.

Таблиця 8.8 - **Обсяг товарообігу**

Товари	Базисний період		Звітний період		q ₀ p ₀	q ₁ p ₁	q ₁ p ₀
	Продано, т	Ціна за 1	Продано, т	Ціна за 1 т			
		т					
	q ₀	p ₀	q ₁	p ₁			
А	12	50	23	60	600	1380	1150
Б	34	60	56	80	2040	4480	3360
В	13	40	15	50	520	750	600
Разом	-	-	-	-	3160	6610	5110

Необхідно визначити:

- а) загальні індекси фізичного обсягу цін і товарообігу;
- б) наскільки змінився товарообіг у звітному періоді в порівнянні з базисним – в цілому, в тому числі в результаті змін:

- 1) фізичного обсягу;

- 2) цін.

На основі даних таблиці 8.8. обчислимо систему агрегатних індексів.

Загальний індекс фізичного обсягу:

$$I_g = \frac{\sum g_1 P_0}{\sum g_0 P_0} = \frac{5110}{3160} = 1,617 \quad \text{або } 161,7\%;$$

тобто, товарообіг у звітному періоді в порівнянні з базисним виріс за рахунок зростання фізичного обсягу на 6,7%.

Загальний індекс цін:

$$I_p = \frac{\sum g_1 P_1}{\sum g_1 P_0} = 6610 : 5110 = 1,294 \quad \text{або } 129,4\%.$$

Отже, товарообіг у звітному періоді виріс в порівнянні з базисним за рахунок зростання цін на 29,4%.

Загальний індекс товарообігу:

$$I_{gp} = \frac{\sum g_1 P_1}{\sum g_0 P_0} = 6610 : 3160 = 2,092 \quad \text{або } 209,2\%.$$

Товарообіг за звітний період в порівнянні з базисним виріс на 109,2%.

$$I_{gp} = I_q \cdot I_p = 1,617 \cdot 1,294 = 2,092.$$

В абсолютному виразі товарообіг у звітному періоді виріс в порівнянні з базисним періодом на:

$$\Delta r q = \sum q_1 P_1 - \sum q_0 P_0 = 6610 - 3160 = 3450 \text{ грн.}$$

в тому числі за рахунок змін:

а) фізичного обсягу

$$\Delta q = \sum q_1 P_0 - \sum q_0 P_0 = 5110 - 3160 = 1950 \text{ грн.}$$

б) цін

$$\Delta p = \sum q_1 P_1 - \sum q_1 P_0 = 6610 - 5110 = 1500 \text{ грн.}$$

$$\Delta r q = \Delta p + \Delta q = 1950 + 1500 = 3450 \text{ грн.}$$

Перевірка взаємозв'язків показників свідчить про вірність проведених розрахунків.

Приклад 2. Витрати на видобуток вугілля та його собівартість характеризуються такими даними.

Таблиця 8.9 - Видобуток вугілля та його собівартість

Шахта	Видобуток вугілля, т		Собівартість 1т, грн.		q ₀ z ₀	q ₁ z ₁	q ₁ z ₀
	базисний період	звітний період	базисний період	звітний період			
	q ₀	q ₁	z ₀	z ₁			
А	86	127	9,0	8,8	774,0	1117,6	1143,0
Б	131	98	12,1	12,3	1585,1	1205,4	1185,8
Разом	217	225	-	-	2359,1	2323,0	2328,8

Необхідно визначити:

а) індекси собівартості змінного складу, фіксованого складу і структурних зрушень;

б) наскільки змінилась середня собівартість вугілля у звітному періоді в порівнянні з базисним в цілому, в тому числі під впливом змін:

1) собівартість вугілля на окремих шахтах;

2) структурних зрушень.

Система індексів динаміки середньої собівартості:

Індекс собівартості змінного складу

$$I_c^z = \frac{\sum g_1 z_1}{\sum g_1} \cdot \frac{\sum g_0 z_0}{\sum g_0} = \frac{2323,0}{2250} \cdot \frac{2359,1}{2170} = 103,109 = 0,945 \text{ або } 94,5\%$$

Індекс собівартості фіксованого складу

$$I_c^g = \frac{\sum g_1 z_1}{\sum g_1} \cdot \frac{\sum g_1 z_0}{\sum g_1} = 103,104 = 0,990 \text{ або } 99,0\%$$

Індекс структурних зрушень

$$I_c^s = \frac{\sum g_1 z_0}{\sum g_1} \cdot \frac{\sum g_0 z_0}{\sum g_0} = 104,109 = 0,954 \text{ або } 95,4\%$$

$$I_c^z = I_c^g \cdot I_c^s = 0,990 \cdot 0,954 = 0,945 \text{ або } 94,5\%$$

У вартісному вимірі середня собівартість у звітному періоді знизилась в порівнянні з базисним періодом на

$$\Delta \bar{z} = \bar{z}_1 - \bar{z}_0 = \left(\frac{\sum g_1 \bar{z}_1}{\sum g_1} - \frac{\sum g_0 \bar{z}_0}{\sum g_0} \right) = 10,3 - 10,9 = -0,6 \text{ грн.},$$

в тому числі в результаті змін:

а) собівартості на окремих підприємствах

$$\Delta z_{\text{фк}} = \left(\frac{\sum g_1 \bar{z}_1}{\sum g_1} - \frac{\sum g_1 \bar{z}_0}{\sum g_1} \right) = 10,3 - 10,4 = -0,1 \text{ грн.};$$

б) структурних зрушень

$$\Delta z_{\text{ср}} = \left(\frac{\sum g_1 \bar{z}_0}{\sum g_1} - \frac{\sum g_0 \bar{z}_0}{\sum g_0} \right) = 10,4 - 10,9 = -0,5 \text{ грн.},$$

$$\Delta z_{\text{ср}} = \Delta z_{\text{фк}} + \Delta z_{\text{ср}} = (-0,1) + (-0,5) = -0,6 \text{ грн.}$$

Отже, середня собівартість 1 т вугілля знизилась у звітному періоді в порівнянні з базисним періодом головним чином, за рахунок структурних зрушень, про що свідчать і дані таблиці 8.9. Обсяг видобутку на шахті А, де собівартість вугілля нижча в порівнянні з шахтою Б, у звітному періоді виріс, а на шахті Б, де собівартість вугілля вища, видобуток, навпаки, знизився.

Якщо факторна система відображається залежністю

$$Y = a \cdot b \cdot c \cdot d,$$

то факторний аналіз проводиться за такими методами:

1. Метод ланцюгових підстановок – за цим методом, як найуніверсальнішим, система факторних індексів визначається шляхом послідовної заміни базисних (планових) значень факторів на звітні:

$$I_a = \frac{a_1 b_0 c_0 d_0}{a_0 b_0 c_0 d_0}; \quad I_b = \frac{a_1 b_1 c_0 d_0}{a_1 b_0 c_0 d_0}; \quad I_c = \frac{a_1 b_1 c_1 d_0}{a_1 b_1 c_0 d_0}; \quad I_d = \frac{a_1 b_1 c_1 d_1}{a_1 b_1 c_1 d_0}.$$

Для оцінки впливу кожного фактора на результативний показник необхідно від значення чисельника кожного індексу відняти значення знаменника, тобто:

$$\Delta y_a = a_1 b_0 c_0 d_0 - a_0 b_0 c_0 d_0;$$

$$\Delta y_b = a_1 b_1 c_0 d_0 - a_1 b_0 c_0 d_0;$$

$$\Delta y_c = a_1 b_1 c_1 d_0 - a_1 b_1 c_0 d_0;$$

$$\Delta y_d = a_1 b_1 c_1 d_1 - a_1 b_1 c_1 d_0.$$

Загальна зміна результативного показника дорівнює:

$$\Delta y = y_1 - y_0.$$

У всіх випадках при проведенні факторного аналізу загальна зміна результативної ознаки повинна дорівнювати сумі зміни під впливом окремих факторів, тобто:

$$\Delta y = \Delta y_a + \Delta y_b + \Delta y_c + \Delta y_d.$$

2. Метод різниць – цей метод є частковим випадком методу ланцюгових підстановок і впливає із вище наведених обчислень впливу окремих факторів шляхом арифметичних перетворень:

$$\Delta y_a = (a_1 - a_0) b_0 c_0 d_0;$$

$$\Delta y_b = a_1 (b_1 - b_0) c_0 d_0;$$

$$\Delta y_c = a_1 b_1 (c_1 - c_0) d_0;$$

$$\Delta y_d = a_1 b_1 c_1 (d_1 - d_0);$$

$$\Delta y = y_1 - y_0;$$

$$\Delta y = \Delta y_a + \Delta y_b + \Delta y_c + \Delta y_d.$$

Методи ланцюгових підстановок і різниць використовуються лише в тому випадку, коли відомі абсолютні значення результативного показника і всіх факторів за два періоди.

3. Метод відносних величин застосовується в тому випадку, коли відсутні абсолютні значення показників – факторів, а є тільки їх відносні зміни (індивідуальні індекси).

За цим методом вплив окремих факторів на зміну результативного показника визначається за такими формулами:

$$\Delta y_a = (i_a - 1)y_0;$$

$$\Delta y_b = i_a(i_b - 1)y_0;$$

$$\Delta y_c = i_a i_b(i_c - 1)y_0;$$

$$\Delta y_d = i_a i_b i_c(i_d - 1)y_0;$$

$$\Delta y = y_1 - y_0;$$

$$\Delta y = \Delta y_a + \Delta y_b + \Delta y_c + \Delta y_d.$$

Слід звернути увагу на методологічні принципи побудови і аналізу факторних систем.

У факторну систему спочатку включаються і оцінюються кількісні показники, а потім якісні.

Якщо у факторну систему входить декілька кількісних, або декілька якісних показників-факторів, то послідовність їх розміщення і аналізу така: в першу чергу розміщуються і аналізуються головні (первинні) фактори, а потім неголовні (вторинні).

Послідовне розміщення факторів в індексній системі – спочатку кількісні, а потім якісні – необхідне для виключення перекрученого впливу зміни кількісних показників, який вони здійснюють на кожен одиницю якісних показників.

Якщо факторна система складається тільки з двох показників – факторів ($y = a \cdot b$), то вплив кожного фактора на зміну результативного показника за методом різниць визначається в такому порядку.

Обчислюється послідовно різниця між значеннями кожного із факторів за звітний і базисний періоди, і одержане число (абсолютна зміна фактора, вплив якого досліджується) перемножується на інший фактор-співмножник, причому останній береться за базисний період, якщо фактор, вплив якого досліджується, є кількісним, і за звітний період, якщо цей фактор якісний.

Так, для факторної системи:

$$Y = a \cdot b,$$

де фактор **a** – кількісний;

b – якісний фактор.

Оцінка впливу кожного фактора на зміну результативного показника здійснюється за такою методикою:

$$\Delta y_a = (a_1 - a_0) b_0;$$

$$\Delta y_b = (b_1 - b_0) a_1;$$

$$\Delta y = y_1 - y_0 = \Delta y_a + \Delta y_b.$$

Розглянемо на прикладах методику проведення факторного аналізу за методами різниць і відносних величин.

Приклад 1. Дані про посівну площу і урожайність пшениці наведені в таблиці 8.10.

Таблиця 8.10 - Посівна площа і урожайність пшениці

Показник	Умовне позначення	Базисний період	Звітний період
Посівна площа, га	п	750	800
Урожайність пшениці, ц/га	у	35	40

Необхідно визначити, на скільки зміниться валовий збір пшениці у звітному періоді в порівнянні з базисним - в цілому, в тому числі під впливом змін:

- а) розміру посівної площі;
- б) урожайності пшениці.

Валовий збір пшениці складає:

$$\text{у базисному періоді } п \cdot у_0 = 750 \cdot 38 = 26250 \text{ ц};$$

$$\text{у звітному періоді } п \cdot у_1 = 800 \cdot 40 = 32000 \text{ ц}.$$

Приріст валового збору пшениці у звітному періоді в порівнянні з базисним дорівнює:

$$\Delta п \cdot у = п \cdot у_1 - п \cdot у_0 = 32000 - 26250 = 5750 \text{ ц},$$

в тому числі під впливом змін:

а) посівної площі

$$\Delta p_{y_n} = (p_1 - p_0) y_0 = (800 - 750) \cdot 35 = 1750 \text{ ц};$$

б) урожайності пшениці

$$\Delta p_{y_s} = (y_1 - y_0) p_1 = (40 - 35) \cdot 800 = 4000 \text{ ц};$$

$$\Delta p_{y_s} = \Delta p_{y_n} + \Delta p_{y_s} = 1750 + 4000 = 5750 \text{ ц}.$$

Приклад 2. Обсяг продукції підприємства у базисному періоді становить 120000 грн. У звітному періоді в порівнянні з базисним чисельність працівників зросла на 2,0%, а продуктивність праці виросла на 5,0%.

Необхідно визначити приріст обсягу продукції у звітному періоді в порівнянні з базисним в цілому, в тому числі під впливом змін:

а) чисельності працюючих;

б) продуктивності праці.

В умові задачі відсутні абсолютні значення показників факторної системи крім обсягу продукції за базисний період. Тому факторний аналіз можливо провести тільки за методом відносних величин.

Позначимо:

Q – обсяг продукції;

N – чисельність працюючих (середньоспискова);

W – продуктивність праці (виробіток продукції на 1-го працюючого).

В умові задачі є такі дані:

$$Q_0 = 120000 \text{ грн.}$$

$$I_w = 100 + 5,0 = 105,0 \text{ або } 1,05.$$

$$I_n = 100 + 2,0 = 102,0 \text{ або } 1,02.$$

Оскільки $W = \frac{Q}{N}$, то $Q = N \cdot W$.

Згідно з правилом про аналогічність зв'язків показників та індексів, що характеризують їх зміну, впливає:

$$i_Q = i_N \cdot i_W = 1,02 \cdot 1,05 = 1,071.$$

Оскільки $i_Q = \frac{Q_1}{Q_0}$, то $Q_1 = i_Q \cdot Q_0 = 1,071 \cdot 120000 = 128520$ грн.

Приріст продукції у звітному періоді в порівнянні з базисним становить:

$$\Delta Q = Q_1 - Q_0 = 128520 - 120000 = 8520 \text{ грн.},$$

в тому числі за рахунок змін:

а) чисельності працюючих

$$\Delta Q_N = (i_N - 1)Q_0 = (1,02 - 1) \cdot 120000 = 6120 \text{ грн.};$$

б) продуктивності праці

$$\Delta Q_W = i_N (i_W - 1)Q_0 = 1,02(1,05 - 1) \cdot 120000 = 2400 \text{ грн.}$$

$$\Delta Q = Q_1 - Q_0 = \Delta Q_N + \Delta Q_W = 2400 + 6120 = 8520 \text{ грн.}$$

Питання для самоконтролю

1. Які відносні величини називаються індексами?
2. Дайте визначення кількісного і якісного показника.
3. За якими правилами будуються індивідуальні і агрегатні індекси?
4. Розкрийте суть поняття "показник-сумірник". За якими правилами він вибирається?
5. В яких випадках використовуються середні індекси? За якими принципами вибираються види середніх?
6. В чому полягає економічна суть індексу змінного складу? Які індекси визначають його зміну?
7. Перерахуйте і дайте характеристику методу факторного аналізу. Обґрунтуйте, за якими критеріями вибирається метод індексного факторного аналізу?
8. Що означає різниця між чисельником і знаменником індексу?
9. На які складові можна поділити приріст товарообігу?

Задачі

8.1. Є такі дані по торговельному підприємству:

Вид товару	Базисний період		Звітний період	
	Ціна за одиницю, грн.	Продано товарів, шт.	Ціна за одиницю, грн.	Ціна за одиницю, шт.
А	45	2500	87	1700
Б	27	830	35	2300
В	12	610	14	1000

Визначити:

- а) індивідуальні індекси фізичного обсягу, цін і товарообігу;
- б) загальні індекси фізичного обсягу, цін і товарообігу.

Перевірити взаємозв'язок індексів:

в) на скільки змінився товарообіг у звітному періоді в порівнянні з базисним – в цілому, в тому числі під впливом змін:

- 1) фізичного обсягу;
- 2) цін.

8.2. На підставі даних, наведених в таблиці:

Продукція	План		Звіт	
	Виготовлено продукції, т.	Собівартість грн.	Виготовлено продукції, т.	Собівартість т, грн.
А	2800	29,0	2942	28,3
Б	164	21,8	171	21,0
В	48	28,3	50	27,9

Визначити:

а) індивідуальні індекси виконання плану фізичного обсягу, собівартості, витрат виробництва;

б) загальні індекси виконання плану фізичного обсягу, собівартості витрат виробництва;

в) на скільки збільшились витрати виробництва за звітом в порівнянні з планом – в цілому, в тому числі під впливом змін:

- 1) фізичного обсягу;
- 2) собівартості продукції.

8.3. Є такі дані:

Товар	Базисний період		Звітний період		Індивідуальні індекси, %	
	Ціна за 1шт, грн	Кількість, ц	Ціна за 1шт, грн	Кількість, ц	Цін	Фізичного обсягу
А	1,47	270,8	1,51	?	?	112,5
Б	0,83	131,6	0,72	?	?	105,7
В	1,37	?	?	314,6	96,8	125,9

Визначити:

- а) відсутні показники в таблиці;
- б) загальні індекси фізичного обсягу, цін і товарообігу;
- в) на скільки змінився товарообіг у звітному періоді в порівнянні з базисним – в цілому, в тому числі в результаті змін:
 - 1) фізичного обсягу;
 - 2) цін.

8.4. Є такі дані :

Товар	Собівартість одиниці, грн.		Виготовлено, тис. шт.		Індивідуальні індекси, %	
	Базисний період	Звітний період	Базисний період	Звітний період	Собівартості	Фізичного обсягу
А	1,32	?	1,27	1,31	0,92	?
Б	2,39	2,46	?	2,96	?	0,71
В	?	3,07	0,72	?	1,99	2,92

Визначити:

- а) відсутні показники в таблиці;
- б) загальні індекси фізичного обсягу, собівартості, витрат виробництва;

в) на скільки зросли витрати виробництва у звітному періоді в порівнянні з базисним – в цілому, в тому числі в результаті змін:

- 1) фізичного обсягу;
- 2) собівартості.

8.5. Виробництво зерна в колективному господарстві характеризується такими даними:

Зернові культури	Посівна площа, га		Урожайність, ц/га	
	Базисний період	Звітний період	Базисний період	Звітний період
А	540	550	24	30
Б	10	50	13	12

Визначити:

- а) індивідуальні індекси урожайності окремих культур;
- б) загальні індекси урожайності змінного, фіксованого складу і структурних зрушень;
- в) на скільки змінилась середня урожайність у звітному періоді в порівнянні з базисним в цілому, в тому числі під впливом змін:

- 1) урожайності окремих культур;
- 2) структурних зрушень.

8.6. Відомі такі дані про якість виготовленої продукції:

Сорт продукції	Виготовлено одиниць, шт.		Ціна за одиницю, грн.	
	Базисний період	Звітний період	Базисний період	Звітний період
I	400	432	10,0	10,5
II	120	128	7,5	7,6
III	25	22	6,0	6,1

Визначити:

а) загальні індекси цін змінного і фіксованого складу та структурних зрушень;

б) на скільки змінилась середня ціна одиниці продукції у звітному періоді в порівнянні з базисним в цілому, в тому числі під впливом змін:

1) цін на окремі сорти продукції;

2) структурних зрушень.

8.7. Відомі такі дані про виробництво продукції:

Продукція	Обсяг виготовленої продукції, тис. грн.		
	Базисний рік	Звітний рік	
		У фактично діючих цінах	У цінах базисного року
А	4,3	4,0	3,5
Б	4,9	5,5	2,6

Визначити:

а) загальні індекси фізичного обсягу, цін і обсягу виробництва;

б) на скільки змінився обсяг виробництва у звітному році в порівнянні з базисним – в цілому, в тому числі під впливом змін:

1) фізичного обсягу;

2) цін.

8.8. Є такі дані про товарообіг і зміну цін:

Товари	Товарообіг у фактичних цінах, тис.грн.		Зміна цін у звітному періоді в порівнянні з базисним, %
А	440	407	+ 10
Б	327	330	+ 5
В	149	115	-7

Визначити:

- а) індивідуальні індекси цін, товарообігу і фізичного обсягу;
- б) загальні індекси фізичного обсягу, цін і товарообігу;
- в) на скільки змінився товарообіг у звітному періоді в порівнянні з базисним - в цілому, в тому числі в результаті змін:
 - 1) фізичного обсягу;
 - 2) цін.

8.9. Є такі дані:

Вид продукції	Виготовлено продукції у травні, тис.грн.	Зміна фізичного обсягу виготовленої продукції у червні в порівнянні з травнем, %
А	30	+ 15,0
Б	10	- 20,0
В	25	+ 10,0

Визначити індивідуальні і загальний індекси фізичного обсягу.

8.10. Є такі дані:

Товар	Товарообіг звітного періоду, тис.грн.	Індекс цін, %
А	50	103,0
Б	10	97,0
В	40	100,0

Визначити:

- а) загальний індекс цін;
- б) економію (перевитрати) коштів населення внаслідок зміни цін.

8.11. Є такі дані:

Вид продукції	Загальні витрати на виробництво продукції, тис.грн.		Зміна фізичного обсягу %
	Базисний період	Звітний період	
А	7,4	18,5	+10,0
Б	3,6	13,2	-3,0

Визначити:

а) індивідуальні індекси фізичного обсягу, собівартості, витрат на виробництво;

б) загальні індекси фізичного обсягу, собівартості, витрат на виробництво;

в) на скільки змінились витрати на виробництво у звітному періоді в порівнянні з базисним – в цілому, в тому числі під впливом змін:

1) фізичного обсягу;

2) собівартості продукції.

8.12. Відомі такі дані по підприємству:

Товар	Обсяг реалізації у базисному періоді, тис.грн.	Зміна фізичного обсягу у звітному періоді в порівнянні з базисним, %
А	48	+2,0
Б	81	-6,0

Визначити:

а) індивідуальний індекс фізичного обсягу;

б) загальні індекси фізичного обсягу і цін, враховуючи, що товарообіг у фактичних цінах збільшився на 55,0%;

в) на скільки змінився обсяг товарообігу у звітному періоді в порівнянні з базисним – в цілому, в тому числі під впливом змін:

1) фізичного обсягу;

2) цін.

8.13. С такі дані по підприємству:

Показник	Базисний період	Звітний період
1. Собівартість одиниці продукції, грн.	250	245
2. Виготовлено продукції, шт.	200	210

Визначити, на скільки змінились загальні витрати у звітному періоді в порівнянні з базисним періодом – в цілому, в тому числі в результаті змін:

- а) собівартості одиниці продукції;
- б) кількості виготовленої продукції.

8.14. Відомі такі дані по малому підприємству:

Показник	Базисний період	Звітний період
1. Товарна продукція, грн.	120000	144000
2. Чисельність працюючих, чол.	30	32

Визначити приріст товарної продукції у звітному періоді в порівнянні з базисним – в цілому, в тому числі під впливом змін:

- а) чисельності працюючих;
- б) виробітку на одного працюючого (продуктивності праці).

8.15. Обсяг продукції у базисному періоді становив 250 тис. грн. У звітному періоді чисельність працюючих знизилась на 2,0%, а продуктивність праці виросла на 5,0%.

Визначити, на скільки змінився обсяг продукції у звітному періоді в порівнянні з базисним – в цілому, в тому числі в результаті змін:

- а) чисельності працюючих;
- б) продуктивності праці.

8.16. Відомі такі дані про зміну фізичного обсягу товарообігу:

Групи товарів	Індекс фізичного обсягу товарообігу	Структура товарообігу у базисному періоді, %
Продовольчі	108,5	55,8
Непродовольчі	97,8	44,2

Визначити загальний індекс товарообігу.

8.17. Зміна (зростання) показників у звітному періоді в порівнянні з попереднім характеризується такими даними:

- витрати виробництва – на 4,8%;
- собівартість одиниці продукції – на 0,8%;
- обсяг виробництва у вартісному вимірі – на 5,2%.

Визначити, на скільки знизились ціни на виготовлену продукцію.

РОЗДІЛ 9 СТАТИСТИЧНІ МЕТОДИ ВИВЧЕННЯ ВЗАЄМОЗВ'ЯЗКІВ

9.1 Задачі статистики у вивченні взаємозв'язків. Види зв'язків.

Суспільні явища органічно пов'язані між собою, залежать один від одного, обумовлюють одне одного і знаходяться у постійному русі і розвитку.

Тільки розкриваючи взаємозв'язок і взаємозалежності між явищами, пізнають їх суть і закони розвитку. Тому вивчення взаємозв'язків – важлива задача будь-якого статистичного аналізу і вирішується вона за допомогою спеціальних методів. Вибір певного методу для оцінки взаємозалежностей показників визначається характером інформації і метою дослідження.

Важливий етап в процесі дослідження зв'язків явищ – це з'ясування їх соціально-економічної природи, суті, яке досягається за допомогою якісного аналізу.

Якісний аналіз природи явищ – це сукупність понять, категорій і методів, за допомогою яких розкривається внутрішня сутність явищ, їх зв'язки, схожість, відмінність.

Існують різні точки зору щодо класифікації видів взаємозв'язків. Одна із них передбачає виділення трьох видів взаємозв'язків:

- факторні – досліджується методом групувань і методами кореляції і регресії;

- компонентні – вивчаються на основі індексних методів;

- балансові – вивчаються на основі побудованих балансів.

Балансові методи широко застосовуються в аналізі і особливо в плануванні.

Найпростіший баланс – це матеріальний баланс, який відображає наявність матеріальних ресурсів на початок і кінець періоду (M_n ; M_k), надходження і використання їх (H ; B) на протязі звітного періоду:

$$M_n + H = B + M_k$$

Можна побудувати баланси будь-яких ресурсів (матеріальних, трудових, енергетичних, фінансових) на рівні підприємства, галузі, регіону, країни. Щодо групувань, то для оцінки взаємозв'язків використовуються тільки факторні групування, які дозволяють вивчити наявність або відсутність взаємозалежностей між досліджуваними явищами.

На підставі факторних групувань можна встановити напрямки зв'язків між ознаками і приблизну уяву про форми цих зв'язків.

Значна роль належить групуванням у дослідженні взаємозв'язків між якісними і кількісними ознаками.

У дослідженнях для аналізу взаємозв'язків особливо ефективні багатofакторні групування, які дозволяють оцінити вплив комплексу факторів на рівень результативної ознаки. Так, за допомогою багатомірного групування можна визначити взаємозв'язок собівартості продукції з матеріало-, фондо-, зарплатоємністю.

Проте за допомогою факторних групувань можна охарактеризувати лише загальні риси зв'язку, його тенденції. Кількісно оцінити рівень взаємозв'язку групування не дозволяє.

Індексний метод дозволяє оцінити взаємозв'язок змін окремих показників, а також вплив окремих факторів на зміну досліджуваного показника.

Проте найбільш ефективним засобом встановлення і оцінки взаємозв'язків окремих явищ в статистичних дослідженнях є метод аналізу кореляції і регресії.

9.2 Методи кореляції і регресії

Існує два основних види залежностей – **функціональна і стохастична (імовірна)**.

Функціональна залежність означає, що зі зміною однієї ознаки (аргументу) друга ознака (функція) зміниться в суворо визначеному порядку. Іншими словами, кожному значенню аргументу відповідає тільки одне значення функції. Так, у формулі рівномірного руху $S = V \cdot T$ за умови постійної швидкості (V) кожному значенню часу (T) буде відповідати тільки одне значення пройденого шляху (S).

На відміну від функціональної залежності **стохастична (імовірна) залежність**, окремим випадком якої є **кореляційна залежність**, відображає зв'язки відносні, неповні.

Суть її в тому, що кожному значенню фактора (аргументу) відповідає не одне повне значення іншої ознаки (функції), а декілька їх значень. Тому зв'язок може проявитися лише в зміні результативної ознаки в середньому. Отже, ці зв'язки виявляються не в кожному окремому випадку, а лише в масі, в середньому. Це пояснюється тим, що в кореляційних зв'язках величина результативної ознаки змінюється під впливом багатьох факторних ознак, причому деякі можуть бути невідомими.

Кореляційний зв'язок можна умовно описати рівнянням:

$$Y = f(x) + \varepsilon,$$

де $f(x)$ – частина результативної ознаки, яка сформувалась під впливом врахованих, відомих факторних ознак (однієї або декількох), що знаходяться в кореляційному зв'язку з ознакою y ;

ε – частина результативної ознаки, яка виникає в результаті дії другорядних, випадкових, неврахованих факторів.

За напрямком зв'язки між явищами бувають **прямі і зворотні**.

При **прямих зв'язках** зі збільшенням (зменшенням) ознаки x відповідно збільшується (зменшується) ознака y , а при зворотніх зв'язках – навпаки.

В залежності від того, якою лінією описуються, зв'язки бувають **прямолінійні і криволінійні (нелінійні)**.

Наявність зв'язків визначається наступним чином:

по-перше, багато зв'язків встановлюються шляхом теоретичного аналізу;

по-друге, для підтвердження зв'язків будується кореляційне поле, на якому розміщуються емпіричні лінії, або іншими словами, будується графік, який дозволяє виявити наявність або відсутність зв'язків між досліджуваними явищами;

по-третє, нарешті, наявність зв'язків встановлюється шляхом факторних групувань.

В економічних дослідженнях використовуються одночасно положення як кореляційного аналізу, так і регресивного аналізу. Тому зазначений апарат дослідження називають **кореляційно-регресивний аналіз**, або **методи кореляції і регресії**.

Кореляційно-регресивний аналіз охоплює ряд етапів:

- постановка задачі дослідження;
- теоретичний аналіз економічної природи (суті) досліджуваної проблеми;
- відбір результативної та факторіальної ознак;
- збір і підготовка інформації;
- вибір виду рівняння (форми зв'язку);
- розрахунки статистичних характеристик і параметрів рівняння;
- статистична і логічна оцінка якості рівняння;
- внесення необхідних коректив і повторний розрахунок параметрів і характеристик рівняння (проводиться до тих пір, поки рівняння не буде відповідати необхідним вимогам);
- розрахунок нарівні з первинними додаткових параметрів і характеристик рівняння;
- економічний аналіз досліджуваних показників;
- побудова прогнозу (при необхідності);
- складання пояснювальної записки і передача результатів дослідження споживачам інформації.

В залежності від кількості факторів, що входять в склад рівняння регресії, розрізняють **парні рівняння** (входить тільки один фактор) і **множинні (багатофакторні, багаточинникові) рівняння**, в склад яких входить два і більше факторів.

Задача кореляційно-регресивного аналізу в першу чергу полягає в тому, щоб визначити і оцінити рівняння, яке б найбільш адекватно відображало природу взаємозв'язків досліджуваних явищ, або іншими словами, знайти алгебраїчну форму необхідного рівняння.

В економічних дослідженнях використовуються різноманітні рівняння, зокрема такі парні:

$$y = a_0 + a_1x \quad - \text{лінійне};$$

$$y = a_0 + a_1x + a_2x^2 \quad - \text{квадратичне};$$

$$y = a_0 + \frac{a_1}{x} \quad - \text{гіперболічне};$$

$$\ln y = a_0 + a_1 \ln x_1 \quad - \text{логіфімічне};$$

$$y = a_0x^{a_1} \quad - \text{степеневе}.$$

Що стосується багатofакторних рівнянь, то в економічних дослідженнях найчастіше застосовуються такі рівняння:

$$y = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + \dots + a_nx_n \quad - \text{лінійне};$$

$$\ln y = a_0 + a_1 \ln x_1 + a_2 \ln x_2 + \dots + a_n \ln x_n \quad - \text{логіфімічне}.$$

Переважа зазначених багатofакторних рівнянь полягає в тому, що їх параметри (a_i) відображають певний економічний зміст.

9.3 Система параметрів і характеристик рівняння регресії

Після теоретичних досліджень поставленої проблеми, вибору виду рівняння, збору і обробки інформації настає важлива стадія в процесі використання методів кореляції і регресії – визначення параметрів рівняння (a_0, a_i) і статистичних характеристик, за допомогою яких оцінюються якість рівняння.

Параметри будь-якого рівняння регресії визначаються методом найменших квадратів:

$$\sum (y_i - \hat{y}_i)^2 = \min, \quad (9.1)$$

де y_i – фактичне значення результативної ознаки;

\hat{y}_i – розрахункове (на підставі реалізованого рівняння) значення результативної ознаки.

На основі методу найменших квадратів формується система нормальних рівнянь, склад яких залежить від виду рівняння регресії.

Стосовно парного лінійного рівняння і квадратичного рівняння система нормальних рівнянь має такий вигляд:

$$\begin{aligned} & \underline{y = a_0 + a_1 x} \\ & \begin{cases} a_0 n + a_1 \sum x = \sum y; \\ a_0 \sum x + a_1 \sum x^2 = \sum xy. \end{cases} \end{aligned} \quad (9.2)$$

$$\begin{aligned} & \underline{y = a_0 + a_1 x + a_2 x^2} \\ & \begin{cases} \sum y = a_0 n + a_1 \sum x + a_2 \sum x^2; \\ \sum xy = a_0 \sum x + a_1 \sum x^2 + a_2 \sum x^3; \\ \sum x^2 y = a_0 \sum x^2 + a_1 \sum x^3 + a_2 \sum x^4. \end{cases} \end{aligned} \quad (9.3)$$

В рівняннях 9.2 і 9.3 x і y – відомі величини, a_0 ; a_1 ; a_2 – невідомі, які можна визначити шляхом реалізації системи рівнянь.

Для спрощення методики визначення параметрів рівняння (a_0, a_1) можна використати метод детермінантів. Для лінійного рівняння за цим методом:

$$\begin{aligned} a_0 &= \frac{\sum y \sum x^2 - \sum xy \sum x}{n \sum x^2 - (\sum x)^2}, \\ a_1 &= \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - (\sum x)^2}. \end{aligned} \quad (9.4)$$

Найбільш важливою характеристикою рівняння, що відображає його якість і цінність, є тіснота зв'язку між фактором (факторами) і результативною ознакою.

Для оцінки тісноти зв'язку використовуються:

- в парних рівняннях – парний лінійний коефіцієнт кореляції (r);
- в парних нелінійних рівняннях - парне кореляційне відношення (η);

- в багатофакторних множинних лінійних рівняннях –множинний лінійний коефіцієнт кореляції (R);

- в множинних нелінійних рівняннях - множинне кореляційне відношення (η).

Парний лінійний коефіцієнт кореляції обчислюється за формулою:

$$r = \frac{n \sum xy - \sum x \sum y}{\sqrt{[n \sum x^2 - (\sum x)^2] \cdot [n \sum y^2 - (\sum y)^2]}} \quad (9.5)$$

Лінійний множинний коефіцієнт кореляції, кореляційне відношення (парне і множинне) розраховуються за однаковою формулою:

$$R_{(\eta)} = \sqrt{1 - \frac{\sum (y - \bar{y})^2}{\sum (y - \bar{y})^2}} \quad (9.6)$$

Для оцінки рівня зв'язку між досліджуваними ознаками використовуються певні межі (таблиця 9.1):

Таблиця 9.1 - Оцінка рівня зв'язку

Величина R, r, η	До 0,5	0,5 – 0,6	0,6 – 0,7	0,7 – 0,8	0,8 – 0,9	Більше 0,9
Сила зв'язку	дуже низький	низький	помітний	середній	сильний	дуже сильний

Окрім тісноти зв'язку для оцінки якості рівнянь застосовують систему статистичних характеристик, серед яких особлива увага приділяється середній помилці апроксимації ($\bar{\varepsilon}$):

$$\bar{\varepsilon} = \frac{1}{n} \sum \frac{|y - \bar{y}|}{y} \cdot 100 \quad (9.7)$$

Зі змінюю обсягу вибіркової сукупності значення параметрів і характеристик рівнянь, як правило, коливаються. Тому з певною імовірністю потрібно бути впевненим, що значення цих показників,

по-перше, не будуть рівними нулю у генеральній сукупності (спростування, так званої, “нульової гіпотези”) і, по-друге, величина їх буде знаходитися в певних інтервалах довіри. Оцінка статистичної надійності параметрів і статистичних характеристик рівняння, відома під назвою перевірка істотності, визначається за допомогою t-критерія Стьюдента. Взагалі, t- критерій розраховується як відношення значення певного показника до його стандартної помилки.

Так наприклад, t-критерій для коефіцієнта множинної лінійної кореляції дорівнює:

$$t_R = \frac{R}{\sigma_R}; \quad \sigma_R = \frac{1 - R^2}{\sqrt{n - m - 1}}, \quad (9.8)$$

де n – кількість одиниць сукупності;

m - число факторів в рівнянні.

При заданому рівні істотності (λ) можна з імовірністю $P=1 - \lambda$ стверджувати, що коефіцієнт множинної лінійної кореляції у генеральній сукупності буде знаходитись в інтервалі :

$$\bar{R} - t\sigma_R \leq \bar{R} \leq \bar{R} + t\sigma_R. \quad (9.9)$$

Істотність рівняння перевіряється за F – критерієм Фішера:

$$F = \frac{\sigma_y^2}{\sigma_{\text{ост}}^2}; \quad \begin{cases} \sigma_y^2 = \frac{\sum (y_i - \bar{y})^2}{n - 1}; \\ \sigma_{\text{ост}}^2 = \frac{\sum (y_i - \bar{y})^2}{n - m - 1}. \end{cases} \quad (9.10)$$

Про істотність показників і рівнянь можна стверджувати в тому випадку, коли виконується умова:

$$t_p \geq t_T; \quad F_p \geq F_T. \quad (9.11)$$

де p, T – індекси слів відповідно “розрахункове” і “табличне”.

В лінійних рівняннях коефіцієнти при невідомих (а) є **коефіцієнтами регресії**. Вони показують, на скільки одиниць зміниться функція зі зміною певного фактора на одну одиницю за умови фіксованого значення решти факторів.

Для усунення відмінності в одиницях виміру показників, що входять в рівняння, визначається коефіцієнт еластичності, який для лінійного рівняння обчислюється за формулою:

$$E_i = a_i \cdot \frac{\bar{X}_i}{\bar{y}}. \quad (9.12)$$

Коефіцієнт еластичності показує, на скільки процентів зміниться функція зі зміною певного фактора на 1% за умови фіксованого значення інших факторів.

В логарифмічних рівняннях коефіцієнти (a_i) по суті є коефіцієнтами еластичності.

Для визначення ступеня впливу кожного фактора на результативний показник в прийнятих одиницях виміру обчислюється **гранична продуктивність (гранична віддача, додатковий продукт)**.

Для логарифмічного рівняння гранична продуктивність обчислюється за формулою:

$$\frac{\partial y}{\partial X_i} = a_i \frac{\bar{y}}{\bar{X}_i}. \quad (9.13)$$

Деякі фактори в рівняннях взаємозамінні. Для визначення взаємозамінності факторів обчислюється **гранична норма заміщення**, яка дорівнює співвідношенню додаткових продуктів зі знаком мінус:

$$\frac{\partial y}{\partial X_2} : \frac{\partial y}{\partial X_1} = - \frac{\partial X_1}{\partial X_2}. \quad (9.14)$$

Первинні і вторинні параметри рівняння широко використовуються в економічному аналізі і прогнозуванні.

В багатофакторних лінійних рівняннях абсолютна зміна результативного показника визначається за формулою:

$$\Delta y = a_1 \Delta X_1 + a_2 \Delta X_2 + a_3 \Delta X_3 + \dots + a_n \Delta X_n. \quad (9.15)$$

В багатофакторних рівняннях відносна зміна результативної ознаки визначається за формулою:

$$\frac{\Delta y}{y} = a_1 \frac{\Delta X_1}{X_1} + a_2 \frac{\Delta X_2}{X_2} + a_3 \frac{\Delta X_3}{X_3} + \dots + a_n \frac{\Delta X_n}{X_n}. \quad (9.16)$$

Коефіцієнт множинної лінійної кореляції і множинне кореляційне відношення в квадраті ($R^2; \eta^2$) називається **коефіцієнтом детермінації**. Останній показує вплив включених в рівнянні факторів на рівень результативної ознаки. Наприклад, якщо $R = 0,9$, то $R^2 = 0,81$. Це означає, що рівень досліджуваного показника на 81% залежить від включених в рівняння факторів і на 19% - від неврахованих факторів.

Розрахунок параметрів і характеристик лінійного рівняння, яке відображає зв'язок між вартістю основних фондів і обсягом товарної продукції, розглянемо на основі таблиці 9.2.

Таблиця 9.2 - Обсяги товарної продукції і вартості основних фондів

Номер підприємства	Обсяг товарної продукції, млн.грн. (y)	Вартість основних фондів, млн.грн. (x)	xy	x ²	y ²	$\hat{y} = a_0 + a_1 x$	y - \hat{y}	$\frac{ y - \hat{y} }{y}$
1	2	3	4	5	6	7	8	9
1	20	1	20	1	400	19,4	0,6	0,030
2	25	2	50	4	625	25,0	0	0
3	31	3	93	9	961	30,6	0,4	0,013
4	31	4	124	16	961	36,2	5,2	0,168
5	40	5	200	25	1600	41,8	1,8	0,045
6	56	6	336	36	3136	47,4	8,6	0,154
7	52	7	364	49	2704	53,0	1,0	0,019
8	60	8	480	64	3600	58,6	1,4	0,023
9	60	9	540	81	3600	64,2	4,2	0,070
10	70	10	700	100	4900	69,8	0,2	0,003
Разом	445	55	2907	385	2248	446	-	0,525

Параметри рівняння розрахуємо за формулами 9.4:

$$a_0 = \frac{\sum y \sum x^2 - \sum xy \sum x}{n \sum x^2 - (\sum x)^2} = \frac{445 \cdot 385 - 2907 \cdot 55}{10 \cdot 385 - 55^2} = 13,8;$$

$$a_1 = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - (\sum x)^2} = \frac{10 \cdot 2907 - 55 \cdot 445}{10 \cdot 385 - 55^2} = 5,6.$$

Таким чином, рівняння, яке описує зв'язок між вартістю основних фондів та обсягом товарної продукції, має такий вигляд:

$$y = 13,8 + 5,6x.$$

Розрахункове значення функції (\hat{y}) одержане шляхом послідовного використання фактичного значення фактора (x) в побудованому рівнянні:

$$\hat{y}_1 = 13,8 + 5,6 \cdot 1 = 19,4; \quad \hat{y}_2 = 13,8 + 5,6 \cdot 2 = 25,0 \text{ і т.д.}$$

Про вірність розрахунків параметрів рівняння (a_0, a_1) свідчить рівність

$$\sum y = \sum \hat{y}. \quad (9.17)$$

Незначні відхилення викликані, як правило, округленням розрахунків. Парний лінійний коефіцієнт кореляції обчислимо за формулою 9.5:

$$r = \frac{n \sum xy - \sum x \sum y}{\sqrt{[n \sum x^2 - (\sum x)^2] \cdot [n \sum y^2 - (\sum y)^2]}} = \frac{10 \cdot 2907 - 55 \cdot 445}{\sqrt{[10 \cdot 385 - 55^2] \cdot [10 \cdot 22487 - 445^2]}} = 0,976.$$

тобто, зв'язок між вартістю основних фондів і обсягом товарної продукції дуже сильний (див. таблицю 9.1)

Середню помилку апроксимації розрахуємо за формулою 9.7:

$$\bar{\varepsilon} = \frac{1}{n} \sum \frac{|y - \hat{y}|}{y} \cdot 100 = \frac{1}{10} \cdot 0,525 \cdot 100 = 5,25\%.$$

В економічних дослідженнях значення середньої помилки не повинно перевищувати 10%.

Коефіцієнт еластичності розрахуємо за формулою 9.13:

$$E = a_1 \cdot \frac{\bar{x}_y}{\bar{y}} = 5,6 \cdot \frac{55 : 10}{445 : 10} = 0,7\%$$

Згідно з інтерпретацією економічної суті коефіцієнтів регресії і еластичності зі зміною величини основних фондів на 1 млн. грн. обсяг товарної продукції зміниться на 5,6 млн. грн., а у відносному значенні – зі зміною вартості основних фондів на 1% обсяг продукції зміниться на 0,7%.

Оцінку коефіцієнтів регресії і еластичності доцільніше здійснювати в багатофакторних рівняннях.

9.4 Вирівнювання рядів динаміки. Часовий тренд

Нарівні з використанням методів укрупнення інтервалів, плинної середньої для вирівнювання динамічних рядів застосовуються також методи кореляції і регресії.

Динамічні ряди вирівнюються за допомогою залежності:

$$\hat{y}_t = f(t), \quad (9.18)$$

де \hat{y}_t – рівні динамічного ряду, які обчислюються відповідним аналітичним виразом (рівнянням) в момент часу t .

Залежність 9.18 називається **часовим трендом**. Вона відображає зміну явища (процесу) в часі.

Часовий тренд може бути описаний різними рівняннями, від яких визначається характером зміни динаміки показників.

Для реалізації рівняння тренда (визначення параметрів і характеристик) застосовуються ті ж формули, які використовуються і для парного рівняння регресії, тільки замість фактора “ x ” використовується час “ t ”.

Рівняння часового тренда застосовується для **інтерполяції і екстраполяції** рядів динаміки.

Інтерполяція дозволяє знайти рівень показника за відсутній період в межах ретроспективи: $t_1 \div t_n$; екстраполяція дозволяє прогнозувати рівень показників за межами ретроспективи: $t_{n+1} \div t_T$ ($T=1;2;3$ і т.д.).

Розглянемо приклад реалізації лінійного рівняння тренда на підставі даних таблиці 9.3.

Таблиця 9.3 - Споживання кондитерських виробів за рік на одну людину

N року	Споживання кондитерських виробів, кг.(y)	t	ty	t ²	y ²	\hat{y}	$ y - \hat{y} $	$\frac{ y - \hat{y} }{y}$
1	10,7	1	10,7	1	114,5	10,60	0,10	0,009
2	11,5	2	23,0	4	132,3	11,55	0,05	0,004
3	12,2	3	36,6	9	148,8	12,50	0,30	0,025
4	13,4	4	53,6	16	179,6	13,45	0,05	0,004
5	15,0	5	75,0	25	225,0	14,40	0,60	0,040
6	15,0	6	90,0	36	225,0	15,35	0,35	0,023
Разом:	77,8	21	288,9	91	1025,2	77,85	-	0,105

Розрахуємо параметри рівняння:

$$a_0 = \frac{\sum y \sum t^2 - \sum ty \sum t}{n \sum t^2 - (\sum t)^2} = \frac{77,8 \cdot 91 - 288,9 \cdot 21}{6 \cdot 91 - (21)^2} = 9,65;$$

$$a_1 = \frac{n \sum ty - \sum t \sum y}{n \sum t^2 - (\sum t)^2} = \frac{6 \cdot 288,9 - 21 \cdot 77,8}{6 \cdot 91 - (21)^2} = 0,95.$$

Отже, рівняння тренда має вигляд:

$$y = 9,65 + 0,95t.$$

Парний лінійний коефіцієнт кореляції дорівнює:

$$r = \frac{n \sum ty - \sum y \sum t}{\sqrt{[n \sum t^2 - (\sum t)^2] \cdot [n \sum y^2 - (\sum y)^2]}} = \frac{6 \cdot 288,9 - 77,8 \cdot 21}{\sqrt{[6 \cdot 91 - 21^2] \cdot [6 \cdot 1025,2 - 77,8^2]}} = 0,982.$$

Розрахункові значення функції на підставі побудованого рівняння:

$$\hat{y}_1 = 9,65 + 0,95 \cdot 1 = 10,60;$$

$$\hat{y}_2 = 9,65 + 0,95 \cdot 2 = 11,55;$$

$$\hat{y}_3 = 9,65 + 0,95 \cdot 3 = 12,50 ;$$

$$\hat{y}_4 = 9,65 + 0,95 \cdot 4 = 13,45 ;$$

$$\hat{y}_5 = 9,65 + 0,95 \cdot 5 = 14,40 ;$$

$$\hat{y}_6 = 9,65 + 0,95 \cdot 6 = 15,35 .$$

Результати розрахунків записані у відповідну графу таблиці 9.3.

Середня помилка апроксимації дорівнює:

$$\bar{\varepsilon} = \frac{1}{n} \sum \frac{|y - \hat{y}|}{y} \cdot 100 = \frac{1}{6} \cdot 0,105 \cdot 100 = 1,75\% .$$

Значення парного лінійного коефіцієнта кореляції і середньої помилки апроксимації свідчать, що побудоване рівняння часового тренда відрізняється високими статистичними характеристиками. Це дозволяє застосувати рівняння для побудови прогнозу шляхом використання в рівнянні тренда значення: $t = n + \tau$ ($n = 6$, $\tau = 1 \div 5$)

Прогнозні значення (\hat{y}_t) на роки наступної п'ятирічки дорівнюють:

$$\hat{y}_7 = 9,65 + 0,95 \cdot 7 = 16,3 \text{ кг};$$

$$\hat{y}_8 = 9,65 + 0,95 \cdot 8 = 17,3 \text{ кг};$$

$$\hat{y}_9 = 9,65 + 0,95 \cdot 9 = 18,2 \text{ кг};$$

$$\hat{y}_{10} = 9,65 + 0,95 \cdot 10 = 19,2 \text{ кг};$$

$$\hat{y}_{11} = 9,65 + 0,95 \cdot 11 = 21,0 \text{ кг}.$$

9.5 Непараметричні показники зв'язку

Методи кореляції і регресії орієнтовані на виявлення і оцінку тісноти зв'язку між кількісними показниками.

Статистика одночасно оцінює наявність і вимірює тісноту зв'язку і між якісними показниками, для чого застосовуються **непараметричні методи**.

Серед значної кількості непараметричних методів розглянемо метод, який дозволяє досліджувати зв'язки паралельних рядів не на основі первинних даних, а на основі рангів.

Ранжування - це упорядкування досліджуваних об'єктів на основі переваг. Згідно з ранжуванням кожній одиниці сукупності присвоюється порядковий номер ряду, який надається їй за рівнем ознаки. Таким чином, ряд значень ознаки ранжується, а номер кожної окремої одиниці буде її рангом. **Ранг** – це порядковий номер значень ознак, установлений в порядку зростання або зменшення їх величин.

Для ранжованих об'єктів показником зв'язку є **коефіцієнт рангової кореляції**.

Серед непараметричних методів для визначення тісноти зв'язку між ознаками на практиці використовуються коефіцієнти рангової кореляції Спірмена (ρ) і Кендалла (τ).

Коефіцієнт рангової кореляції Спірмена обчислюється за такою формулою:

$$\rho = 1 - \frac{6 \cdot \sum d^2}{n(n^2 - 1)}, \quad (9.19)$$

де n - кількість одиниць вибіркової сукупності ;

$d^2 = (R_x - R_y)^2$ $\left\{ \begin{array}{l} R_x - \text{ранг факторної ознаки;} \\ R_y - \text{ранг результативної ознаки.} \end{array} \right.$

Коефіцієнт кореляції рангів Кендалла **визначається за формулою:**

$$\tau = \frac{2S}{n(n-1)}. \quad (9.20)$$

Порядок розрахунку цього показника такий:

- а) значення x і y ранжуються, тобто визначаються R_x і R_y ;
- б) значення R_x записуються суворо в порядку зростання або, навпаки, зниження ($R_x = 1; 2; 3; \dots; n$);
- в) ранги результативної ознаки (R_y) розміщуються в порядку, які відповідають значенню "x" в початкових даних;
- г) для кожного значення R_y підраховують число наступних за ним рангів більш високого порядку. Загальна сума таких випадків, які враховуються зі знаком "+", позначається буквою P;

д) аналогічно для кожного значення R_y послідовно підраховується кількість наступних за ним рангів, менших за значенням. Вони враховуються зі знаком “-“, а загальна сума позначається буквою Q ;

ж) визначається загальна сума $S=P+Q$.

Величина коефіцієнтів рангової кореляції знаходиться в межах: $-1 \leq \rho, \tau \leq +1$

Приклад розрахунків коефіцієнтів рангової кореляції Спірмена і Кендалла розглянемо на підставі даних таблиці 9.4.

Таблиця 9.4 - **Обсяг йоду в воді та їжі і захворювання населення зобом**

Кількість йоду у воді та їжі, г (x)	Захворювання населення зобом, % (y)	Ранг		$d=R_x-R_y$	d^2	Підрахунок балів	
		R_x	R_y			“+”	“-”
201	0,2	1	7	-6	36	-	6
178	0,6	2	6	-4	16	-	5
155	1,1	3	4	-1	1	1	3
154	0,8	4	5	-1	1	-	3
126	2,5	5	3	+2	4	-	2
1	4,4	6	2	+4	16	-	1
71	16,9	7	1	+6	36	-	-
Разом	-	-	-		110	P=1	Q=-20

На підставі результатів розрахунків, наведених в таблиці 9.4, обчислимо:

коефіцієнт рангової кореляції Спірмена

$$\rho = 1 - \frac{6 \sum d^2}{n(n^2 - 1)} = 1 - \frac{6 \cdot 110}{7(7^2 - 1)} = -0,964;$$

коефіцієнт рангової кореляції Кендалла:

$$\tau = \frac{2S}{n(n-1)} = \frac{2[1+(-20)]}{7(7-1)} = -0,905.$$

Значення розрахованих коефіцієнтів рангової кореляції свідчать, по-перше, про значний вплив обсягу споживання йоду, який міститься у

їжі та воді, на рівень захворювання зубом і, по-друге, між досліджуваними показниками є обернений зв'язок, який дозволяє зробити такий висновок - зі збільшенням споживання йоду захворювання зубом знижується.

Питання для самоконтролю:

1. Наведіть приклади статистичних зв'язків.
2. В чому відмінність між функціональною і стохастичною залежністю?
3. Визначіть основні етапи проведення кореляційного і регресивного аналізу.
4. Які види рівнянь вам відомі ?
5. Якими методами визначаються параметри рівнянь?
6. Дайте інтерпретацію економічної суті параметрам рівнянь.
7. За допомогою яких показників визначається тіснота зв'язку між результативною ознакою і факторіальними ознаками?
8. З якою метою і в яких напрямках використовуються рівняння часового тренда?
9. В яких випадках використовуються непараметричні методи вивчення зв'язків?
10. Як оцінюється тіснота зв'язку в ранжованих ознаках?

Задачі

9.1. Є такі дані про строк збирання і урожайність кукурудзи:

Строк збирання, дні	27	23	18	20	25	30	24	34	16	20
Урожайність кукурудзи, ц/га	25	45	48	44	41	22	45	20	52	50

Визначити взаємозв'язок між наведеними показниками за допомогою парного лінійного рівняння і обчислити:

Визначити взаємозв'язок між наведеними показниками за допомогою парного лінійного рівняння і обчислити:

- а) параметри рівняння;
- б) парний лінійний коефіцієнт кореляції;
- в) коефіцієнт еластичності;
- г) середню помилку апроксимації.

Обґрунтувати напрямок зв'язку і зробити висновки.

9.2. Вихід цукру (відношення кількості одержаного цукру до обсягу перероблених буряків, %) і цукристість буряків (вміст цукру в буряках, %) характеризуються такими даними:

Вихід цукру, %	11,7	12,9	11,2	12,8	13,2	13,4	13,1	12,4	12,0	12,5
Цукристість буряків, %	15,4	13,3	14,8	16,5	16,7	16,9	16,1	16,0	15,7	16,2

Взаємозв'язок наведених показників описати парним лінійним рівнянням.

Визначити:

- а) параметри рівняння;
- б) парний лінійний коефіцієнт кореляції;
- в) середню помилку апроксимації.

Зробити висновки.

9.3 Динаміка рівня механізації робіт характеризуються такими даними:

Рівень механізації, робіт %	40	44	46	49	53	57
Роки (t)	1	2	3	4	5	6

За наведеними даними побудувати лінійне рівняння тренда, оцінити його за допомогою парного лінійного коефіцієнта кореляції і

середньої помилки апроксимації. На основі одержаного рівняння скласти прогноз рівня механізації робіт на наступні п'ять років.

9.4 Урожайність сільськогосподарської культури в залежності від кількості внесених добрив характеризується такими даними:

Внесено добрив, ц/га	1	2	3	4	5	6	7	8	9	10
Урожайність, ц/га	10	13	15	16	19	18	21	23	22	23

Взаємозв'язок показників описати за допомогою квадратичного рівняння.

Визначити:

- параметри рівняння;
- кореляційне відношення;
- середню помилку апроксимації.

Зробити висновки.

9.5 Є такі дані по магазинах району:

Товарообіг, тис.грн.	670	560	580	630	610	650	520	500	560	470
Витрати обігу, тис. грн.	35	27	30	40	36	31	28	30	24	70

Для оцінки тісноти зв'язку між показниками визначити коефіцієнти рангової кореляції Спірмена і Кендалла.

СПИСОК ЛІТЕРАТУРИ

1. Єріна А.М., Пальян З.О. Теорія статистики: Практикум. - К.: Товариство "Знання", КОО, 1997.
2. Ковтун Н.В., Столяров Г.С. Загальна теорія статистики: Курс лекцій.- К.: Четверта хвиля, 1996.
3. Кулинич О.І. Теорія статистики: Підручник. – Кіровоград: Центрально-Українське видавництво, 1996.
4. Кулинич О.І. Теорія статистики: Задачник. – Кіровоград: Центрально-Українське видавництво, 1995.
5. Статистика: Підручник / За ред. С.С. Герасименко – К.: КНЕУ, 1998.
6. Вашків П.Г. та ін. Статистика підприємництва: Навчальний посібник. / За ред. П.Г. Вашківа, В.П. Сторожука.- К.: Слобожанщина, 1999.
7. Єлисеєва И. И., Юзбашев М. М. Общая теория статистики.-М.: Финансы и статистика, 1995.
8. Ефимова М.Р. и др. Общая теория статистики: Учебник для студентов вузов. – М.: ИНФРА-М. 1996.
9. Общая теория статистики: Статистическая методология в изучении коммерческой деятельности: Учебник./ Под ред. А.А. Спирина, О.Э. Башиной. – М.: 1995.
10. Статистика: Курс лекцій./ Под ред. В.Г. Ионина. -- Новосибирск: Издательство НГАЭ и У, М.: ИНФРА, 1998.
11. Теория статистики: Учебник для студентов вузов. / Под ред. Р.Я. Шмойловой.-М.: Финансы и статистика, 1996.

Навчальне видання

Грабовецький Б.Є.

ЗАГАЛЬНА ТЕОРІЯ СТАТИСТИКИ

Навчальний посібник

Оригінал-макет підготовлено автором

Редактор Скалоцька О.Д.

Підписано до друку *27.08.2004*
Формат 29,7×42¹/₄ Гарнітура Times New Roman
Друк різнографічний Ум.друк.арк. *9.72*
Тираж 75 прим.
Зам. № *2001-177*

Віддруковано у комп'ютерному інформаційно-видавничому центрі
Вінницького державного технічного університету
21021, м. Вінниця, Хмельницьке шосе, 95, ВДТУ,ГНК, 9-й поверх
Тел. (0432) 44-01-59