МІНІСТЕРСТВО ОСВІТИ І НАУКИ, МОЛОДІ ТА СПОРТУ УКРАЇНИ
ДЕРЖАВНИЙ ВИЩИЙ НАВЧАЛЬНИЙ ЗАКЛАД
«ДОНЕЦЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ»
АВТОМОБІЛЬНО-ДОРОЖНІЙ ІНСТИТУТ

«ЗАТВЕРДЖУЮ»
Директор АДІ ДВНЗ «ДонНТУ»
М. М.Чальцев
25.01.2012

Кафедра «Економіка і фінанси»

МЕТОДИЧНІ ВКАЗІВКИ ДО ВИКОНАННЯ ПРАКТИЧНИХ РОБІТ З ДИСЦИПЛІНИ «СТАТИСТИКА» І ЧАСТИНА (ДЛЯ СТУДЕНТІВ СПЕЦІАЛЬНОСТЕЙ 6.030502 «ЕКОНОМІЧНА КІБЕРНЕТИКА» ТА 6.030601 «МЕНЕДЖМЕНТ» ВСІХ ФОРМ НАВЧАННЯ)

18/71-2012-02

«РЕКОМЕНДОВАНО»
Навчально-методична комісія
факультету «Економіка і управління»
Протокол № 3
від 21.12.2011 р.

«РЕКОМЕНДОВАНО»
Кафедра
«Економіка і фінанси»
Протокол № 6
від 19.11.2011 р.

Горлівка – 2012

УДК 31(07)

Методичні вказівки до виконання практичних робіт з дисципліни «Статистика» І частина (для студентів спеціальностей 6.030502 «Економічна кібернетика» та 6.030601 «Менеджмент» всіх форм навчання) [Електронний ресурс] / укладачі: В. П. Полуянов, Г. П. Доровських. – Електрон. дані. – Горлівка: ДВНЗ «ДонНТУ» АДІ, 2012. – 1 електрон. опт. диск (CD-R); 12 см. – Систем. вимоги: Pentium; 32 RAM; WINDOWS 98/2000/NT/XP; MS Word 2000. – Назва з титул. екрану.

В методичних вказівках наведено практичні вказівки для рішення задач з дисципліни, варіанти задач до виконання практичних робіт, а також перелік посилань.

Укладачі: Полуянов В. П., д.е.н., проф.
Доровських Г. П., асистент

Відповідальний за випуск: Полуянов В. П., д.е.н., проф.
каф. «Економіка і фінанси»

Рецензент: Вовк Л. П., д.т.н., проф.
каф. «Вища математика»

ЗМІСТ

ВСТУП 4
ТЕМА 1 СТАТИСТИЧНЕ СПОСТЕРЕЖЕННЯ 5
ТЕМА 2 ЗВЕДЕННЯ ТА УГРУПОВАННЯ СТАТИСТИЧНИХ ДАНИХ 7
ТЕМА 3 СЕРЕДНІ ВЕЛИЧИНИ У СОЦІАЛЬНО-ЕКОНОМІЧНИХ ДОСЛІДЖЕННЯХ 9
ТЕМА 4 ПОКАЗНИКИ ВАРІАЦІЇ 19
ТЕМА 5 ВИБІРКОВИЙ МЕТОД 25
ТЕМА 6 МЕТОДИ ВИВЧЕННЯ ЗВ'ЯЗКУ МІЖ ЯВИЩАМИ ТА ЇХ КОРИСТУВАННЯ ДЛЯ УПРАВЛІННЯ СОЦІАЛЬНО-ЕКОНОМІЧНИМИ ПРОЦЕСАМИ 30
ТЕМА 7 МОДЕЛЮВАННЯ ТА АНАЛІЗ ДИНАМІКИ СОЦІАЛЬНО-ЕКОНОМІЧНИХ ЯВИЩ 37
ТЕМА 8 ІНДЕКСНИЙ МЕТОД АНАЛІЗУ 42
ПЕРЕЛІК ПОСИЛАНЬ 44

ВСТУП

У сучасному суспільстві статистика стала одним з найважливіших інструментів управління народним господарством. Вона збирає інформацію, що характеризує розвиток економіки країни, культури і життєвого рівня народу. За допомогою статистичної методології вся отримана інформація узагальнюється, аналізується й у результаті дає можливість побачити систему взаємозв'язків в економіці, динаміці розвитку, дозволяє робити міжнародні зіставлення.

Статистика – це галузь науки, що поєднує принципи і методи роботи з чисельними даними, що характеризують масові явища.

Роль статистики в умовах ринкової економіки не знижується, потреба в достовірній інформації зростає як для державних підприємств, так і для підприємств інших форм власності.

Кожна група даних представляє певну інформацію; ступінь її придатності залежить частково від мети, з якою вона збирається, і частково від способу її збирання й обробки. У певних умовах вона може дати точні характеристики подій, однак найбільш корисне застосування статистики являє собою одержання такої інформації, що давала б можливість приймати рішення в умовах невизначеності. Цей підхід до статистики, як і до процесу прийняття рішень, з'явився порівняно недавно і виправдав себе.

ТЕМА 1 СТАТИСТИЧНЕ СПОСТЕРЕЖЕННЯ

Для того щоб вивчити кількісну сторону масових суспільних явищ і процесів, насамперед потрібно зібрати про них необхідну статистичну інформацію. З цією метою організовується масове статистичне спостереження, яке є першим етапом статистичного дослідження. У процесі статистичного спостереження формуються дані, від якості яких залежить успіх усього дослідження.

Статистичне спостереження – це планомірне, науково організоване збирання або одержання масових відомостей про явища і процеси суспільного життя (наприклад: переписи населення, основних фондів, заповнення анкет, бланків, форм статистичної звітності). Матеріали статистичного спостереження є основою для одержання узагальнюючих показників, які характеризують ті або інші суспільні явища і процеси.

Залежно від ступеня охоплення одиниць досліджуваної сукупності розрізняють два види статистичного спостереження: *суцільне та несуцільне*.

Суцільним називають таке спостереження, при якому обстеженню підлягають усі без винятку одиниці сукупності (наприклад, перепис населення, облік виходу продукції). Матеріали суцільного спостереження дають максимально повне уявлення про все розмаїття форм і можливість одержання точних характеристик про досліджувані соціально-економічні явища і процеси.

Несуцільним називають таке спостереження, при якому обстежується тільки частина сукупності (наприклад, вивчення використання робочого часу і устаткування, цін на ринках). Несуцільне спостереження базується на обліку деякої частини, як правило, достатньо масової частини одиниць спостереження, яка дає змогу на основі наукового відбору одиниць одержати стійкі узагальнюючі характеристики всієї сукупності. Несуцільне спостереження в свою чергу підрозділяється на *вибіркове, монографічне, анкетне і основного масиву*.

Вибірковим називають таке спостереження, при якому обстеженню підлягає певна частина сукупності явищ, яку отримали на основі ненавмисного випадкового відбору.

Монографічне спостереження – це докладний і всебічний опис окремих одиниць досліджуваної сукупності, які цікавлять дослідника. Монографічне обстеження може бути спрямоване на вивчення процесу розвитку окремого трудового колективу, узагальнення передового досвіду, опис нових технологій виробництва і форм організації праці.

Анкетний спосіб спостереження ґрунтується на заповненні анкет певного кола осіб або установ. Заповнення і повернення їх до органів, які проводять спостереження, є добровільним. Як правило, заповнених анкет повертається менше, ніж розсилається. Крім того, неможливо проконтролювати правильність відповідей на запитання анкети. Тому такий спосіб спостереження може застосовуватись у тих випадках, коли не вимагається висока точність відомостей, а потрібні приблизні характеристики.

Обстеження основного масиву являє собою спостереження за частиною найбільш крупних одиниць, питома вага яких переважає в загальному обсязі досліджуваної сукупності.

За способом реєстрації або отримання статистичних даних розрізняють три основних види спостереження: *безпосереднє, документальне та опитування.*

Безпосереднім називають спостереження, яке здійснюється шляхом реєстрації досліджуваних одиниць та їхніх ознак на основі безпосереднього огляду, підрахунку, зважування, зняття показників приладів спеціальними особами, які проводять спостереження, інакше кажучи, реєстраторами, в завдання яких входить, поряд зі встановленням і оцінкою фактів, фіксування їх у документах первинного обліку.

Документальне спостереження ґрунтується на використанні різних документів первинного обліку підприємств, організацій установ (звітності, первинних бухгалтерських документів, річних звітів).

Опитування – це спосіб спостереження, при якому відомості отримують зі слів опитуваних осіб. Опитування може бути усним та письмовим. Розрізняють три способи опитування: *експедиційний, самореєстрації і кореспондентський.*

ТЕМА 2 ЗВЕДЕННЯ ТА УГРУПУВАННЯ СТАТИСТИЧНИХ ДАНИХ

Другим етапом статистичного дослідження є **зведення,** суть якого складається в обробці первинних матеріалів спостереження з метою одержання підсумкових або упорядкованих певним чином числових характеристик тієї або іншої досліджуваної сукупності. За місцем проведення, зведення може бути централізованим і децентралізованим.

Централізованим називається зведення, при якому всі первинні статистичні матеріали зосереджуються в одному місці, де вони розробляються за єдиною програмою в потрібних розрізах і групах.

Децентралізованим називається таке зведення, при якому підсумкові дані одержують на основі їх обробки послідовними етапами.

Основним і дуже важливим моментом зведення є **групування,** тобто об'єднання статистичних даних в однорідні групи за певними ознаками. За допомогою групувань вирішують різні завдання. Найважливішими з них є: 1) виділення і всебічна характеристика різних соціально-економічних явищ; 2)характеристика структури досліджуваних явищ; 3)вивчення взаємозв'язків між окремими ознаками сукупності. Відповідно до цього розрізняють три види групувань: *типологічні, структурні й аналітичні*.

Групування, що приводять до виділення соціально-економічних типів, класів, одноякісних груп або сукупностей, називають **типологічними**. Вони широко застосовуються в економічних, соціальних і демографічних дослідженнях.

Структурні групування характеризують розподіл одиниць досліджуваної сукупності за будь-якою ознакою. За їх допомогою проводиться аналіз структури сукупності і структурних зрушень в розвитку соціально-економічних явищ і процесів.

Аналітичними називають групування, за допомогою яких вивчаються взаємозв'язки між окремими ознаками статистичної сукупності. Вони можуть бути побудовані за результативною (собівартістю, виходом продукції з одиниці земельної площі) і факторною (якість грунту, кількість добрив) ознакою.

Результати статистичного зведення і групування, як правило, оформляються у вигляді статистичних таблиць.

Статистичні таблиці – це форма систематизованого, раціонального і наочного викладення статистичних даних про явища і процеси суспільного життя. Складену таблицю, але не заповнену цифрами, прийнято називати **макетом таблиці**.

По суті, статистична таблиця являє собою статистичне речення, яке має *підмет і присудок*. **Підметом** таблиці є одиниці статистичної сукупності або їх групи, які підлягають характеристиці і вивченню. **Присудком** таблиці –

цифрові дані, що характеризують підмет.

Залежно від побудови (розробки) підмета розрізняють три види статистичних таблиць: *прості, групові та комбінаційні*.

Простими називають такі статистичні таблиці, в підметі яких міститься простий перелік будь-яких об'єктів, територіальних підрозділів або хронологічних дат.

Груповими називають статистичні таблиці, в яких статистичний підмет складається з груп, виділених за будь-якою однією суттєвою ознакою, а присудок містить ряд ознак, які характеризують зазначені групи.

Комбінаційними називають статистичні таблиці, в яких підмет являє собою комбінацію, сполучення двох або кількох ознак, а в присудку наводяться ознаки, що характеризують виділені групи і підгрупи.

ТЕМА 3 СЕРЕДНІ ВЕЛИЧИНИ У СОЦІАЛЬНО-ЕКОНОМІЧНИХ ДОСЛІДЖЕННЯХ

Середньою величиною у статистиці називають узагальнюючий показник, який характеризує типовий розмір ознаки в якісно однорідній сукупності. Обчислюється середня величина у більшості випадків шляхом ділення загального обсягу ознаки на кількість одиниць, що володіють цією ознакою. Якщо, наприклад, відомий фонд місячної заробітної плати і кількість робітників за місяць, то середню місячну заробітну плату можна визначити шляхом ділення фонду заробітної плати на кількість робітників.

Середні величини, які обчислюються як з абсолютних, так і з відносних величин, є показниками іменованими і виражаються в тих самих одиницях вимірювання, що і усереднювана ознака. Вони характеризують одним числом значення досліджуваної сукупності. В середніх величинах знаходить відображення об´єктивний і типовий рівень соціально-економічних явищ і процесів.

Залежно від характеру усереднюваної ознаки і наявної вихідної інформації в статистиці застосовуються різні види середніх величин, серед яких найбільше використовуються такі: середня арифметична, середня гармонічна, середня геометрична і середня квадратична

Поряд з переліченими видами середніх величин в статистичній практиці знаходять застосування також середня хронологічна, середня ковзна, середня прогресивна, середня багатовимірна і так звані структурні середні: мода, медіана та ін.

Кожну середню можна визначити як просту, коли значення варіант спостерігаються тільки один раз або однакову кількість разів, і як зважену, коли значення варіант повторюється різну кількість разів.

Середні величини, що застосовуються в статистиці, належать до загального типу ступеневих середніх. Відрізняються вони тільки показником ступеня. Математична статистика виводить різні середні з формули ступеневої середньої, яка являє собою корінь k-го ступеня з частки від ділення суми індивідуальних значень ознаки k-го ступеня на кількість індивідуальних значень:

проста зважена

$$\bar{x} = \sqrt[\kappa]{\frac{\sum\limits_{i=1}^{n} x_i^{\kappa}}{n}} \qquad \bar{x} = \sqrt[\kappa]{\frac{\sum\limits_{i=1}^{n} x_i^{\kappa} \cdot f_i}{\sum f}}, \tag{3.1}$$

де k – показник ступеня, який визначає тип середньої.

Підставляючи у наведену формулу замість *k* відповідні значення показника ступеня, одержимо такі середні:

	проста	зважена	
при $k = 1$ – арифметичну	$\bar{x} = \frac{\sum x_i}{n}$	$\bar{x} = \frac{\sum x_i \cdot f_i}{\sum f}$;	(3.2)
при $k = -1$ – гармонічну	$\bar{x} = \frac{n}{\sum \frac{1}{x_i}}$	$\bar{x} = \frac{\sum w_i}{\sum \frac{w_i}{x_i}}$;	(3.3)
при $k = 0$ – геометричну	$\bar{x} = \sqrt[n]{x_1 \cdot x_2 \cdot ... \cdot x_n}$	$\bar{x} = \sqrt[\sum f]{x_1^{f_1} \cdot x_2^{f_2} \cdot ... \cdot x_n^{f_n}}$	(3.4)
при $k = 2$ – квадратичну	$\bar{x} = \sqrt{\frac{\sum x_i^2}{n}}$	$\bar{x} = \sqrt{\frac{\sum x_i^2 \cdot f_i}{\sum f}}$,	(3.5)

де $\bar{x}$ – середнє значення досліджуваної ознаки;
x_i – окремі значення усереднюваної ознаки (варіанти);
n – кількість одиниць досліджуваної сукупності;
f – частота повторень (вага) варіант;
$w = x_i \cdot f_i$ – обсяг явища.

Середня арифметична проста застосовується в тих випадках, коли відомі дані про окремі значення ознаки та їх кількість в сукупності. В статистичній практиці вона застосовується, як правило, для розрахунку середніх рівнів ознак, представлених у вигляді абсолютних показників.

Наприклад, якщо є дані про посівну площу овочів у чотирьох бригадах господарства (га): 50; 61; 38; 46 і необхідно визначити середній розмір посівної площі.

Рішення:

Розрахунок середньої величини необхідно здійснювати за формулою середньої арифметичної простої, оскільки значення усереднюваної ознаки зустрічаються однакову кількість раз (по одному разу):

$$\bar{x} = \frac{\sum x_i}{n} = \frac{50 + 61 + 38 + 46}{4} = 49 \text{ га}$$

Отже, середній розмір посівної площі з розрахунку на одну бригаду становить 49 га.

Середня арифметична зважена застосовується у тих випадках, коли окремі значення ознак повторюються.

Якщо частоти виражені в процентах, то формула середньої арифметичної зваженої записана в такому вигляді:

$$\bar{x} = \frac{\sum x_i d_i}{\sum d}, \tag{3.6}$$

де $d_i = \frac{f_i}{\sum f} \cdot 100$ – питома вага кожної частини в загальному обсязі всіх частот (у процентах).

Оскільки для всієї сукупності $\sum d_i$ =100%, то формулу можна записати так:

$$\bar{x} = \frac{1}{100} \cdot \sum x_i d_i = 0{,}01 \cdot \sum x_i d_i \tag{3.7}$$

Якщо частоти виражені в коефіцієнтах (частках), $\sum w_i = 1$, тоді формула середньої спрощується:

$$\bar{x} = \sum x_i w_i \tag{3.8}$$

Наприклад, є наступні дані у таблиці 3.1.

Таблиця 3.1 – Розподіл 50 робітників за тарифним розрядом

Тарифний розряд, *xi*	2	3	4	5	6
Кількість робітників, *fi*	8	10	15	11	6

Потрібно визначити середній тарифний розряд робітників.

Рішення:

Зробимо розрахунок по формулі середньої арифметичній зваженій:

$$\bar{x} = \frac{\sum x_i \cdot f_i}{\sum f} = \frac{2 \cdot 8 + 3 \cdot 10 + 4 \cdot 15 + 5 \cdot 11 + 6 \cdot 6}{8 + 10 + 15 + 11 + 6} = 3{,}94 \text{ (≈ 4 розряд)}$$

Середня гармонічна є зворотньою до середньої арифметичної, обчисленої з обернених значень усереднюваної ознаки. Залежно від характеру наявного матеріалу її застосовують тоді, коли вагу доводиться не множити, а ділити на варіанти, або, що те ж саме, множити на обернене їх значення. Таким чином, середня гармонічна розраховується, коли відомі дані про обсяг ознаки ($W = x \cdot f$) та індивідуальні значення ознаки (x) і невідому вагу (f). Оскільки обсяги ознак являють собою добуток значень ознаки (x) на частоту (f), то частоту (f) визначають як $f = W : x$. Середня гармонічна є перетвореною формою середньої арифметичної. Замість гармонічної завжди можна розрахувати середню арифметичну, попередньо визначивши вагу окремих значень ознаки. При обчисленні середньої гармонічної вагою є обсяги ознак. Середня гармонічна проста застосовується у випадках, коли обсяги явищ по кожній ознаці рівні.

Наприклад, три комбайнера працюють на збиранні зернових культур. Перший комбайнер на збирання з 1 га протягом зміни затратив 33 хв., другий – 29 хв., третій – 30 хв. Потрібно визначити середні затрати праці на збирання 1 га зернових культур.

Рішення:

Розрахунок середніх затрат часу на збирання 1 га зернових культур можна здійснити за формулами середньої арифметичної простої або середньої гармонічної простої:

Середня арифметична проста:

$$\bar{x} = \frac{\sum x_i}{n} = \frac{33 + 29 + 30}{3} = 31 \text{ хв.}$$

Середня гармонічна проста:

$$\bar{x} = \frac{n}{\sum \frac{1}{x_i}} = \frac{3}{\frac{1}{33} + \frac{1}{29} + \frac{1}{30}} = 31 \text{ хв.}$$

Середню геометричну застосовують, коли загальний обсяг явища є не сума, а добуток значень ознаки. Ця середня використовується здебільшого для розрахунку середніх коефіцієнтів (темпів) зростання і приросту при вивченні динаміки явищ і має такий вигляд:

$$\overline{x} = \sqrt[n]{x_1 \cdot x_2 \cdot ... \cdot x_n} = \sqrt[n]{\Pi(x)}, \text{ або } \overline{x} = \sqrt[n-1]{\frac{y_n}{y_1}}, \qquad (3.9)$$

де n – кількість коефіцієнтів зростання;

y_1 і y_n – початковий і кінцевий рівні динамічного ряду.

Величина середньої геометричної залежить тільки від співвідношення кінцевого і початкового рівнів. Якщо не змінюються в цих межах інші рівні, величина середньої не зміниться.

Наприклад, у районі в 1993 р. було зареєстровано 20913 злочинів, а в 2002 р. – 31308. Визначити середній темп зростання зареєстрованих злочинів за весь період.

Рішення:

Розрахуємо середній темп зростання зареєстрованих злочинів за весь період за формулою середньої геометричної:

$$\overline{x} = \sqrt[n-1]{\frac{y_n}{y_1}} = \sqrt[10-1]{\frac{31308}{20913}} = 1,046\ \%\ ,$$

або 104,6 %, тобто з 1993 р. по 2002 р. злочинність щорічно збільшувалася на 4,6 %.

Середня квадратична використовується переважно для розрахунку показників варіації (коливання) ознаки – дисперсії і середнього квадратичного відхилення, які обчислюються на основі квадратів відхилень індивідуальних значень ознаки від їхньої середньої арифметичної. Крім того, вона застосовується для узагальнення ознак, виражених лінійними мірами якихнебудь площ (при обчисленні середніх діаметрів стовбурів дерев, кошиків, листків, бульб тощо).

Наприклад, є два квадрати зі сторонами 20 і 30 см. Визначити середній розмір сторони квадрату.

Рішення:

Знаючи площу двох квадратів, можна визначити сторону рівновеликого квадрата за формулою середньої квадратичної простої:

$$\overline{x} = \sqrt{\frac{\sum x_i^2}{n}} = \sqrt{\frac{20^2 + 30^2}{2}} = 25{,}5 \text{ см}$$

Середня хронологічна являє собою середню величину з показників, що змінюються у часі. Вона розраховується із рівнів моментного або інтервального рядів динаміки за принципом середньої арифметичної простої і зваженої.

Для інтервального ряду динаміки середня хронологічна проста обчислюється за формулою:

$$\bar{y} = \frac{\sum y_i}{n}, \qquad (3.10)$$

де y_i – рівень ряду динаміки,
n – кількість рівнів у ряду динаміки.

Для моментного ряду динаміки (при рівній відстані періодів, наприклад, місяць, квартал тощо) *середня хронологічна проста* обчислюється за формулою:

$$\bar{y} = \frac{\frac{1}{2}y_1 + y_2 + ... + \frac{1}{2}y_n}{n-1}. \qquad (3.11)$$

Середня хронологічна зважена має вигляд $\bar{y} = \frac{\sum y_i \cdot t}{t}$, якщо відомий час, протягом якого зберігалось кожне значення y. Тут t – період часу, який відокремлює один рівень від іншого.

Наприклад, є інформація про списочну чисельність працівників підприємства у 2003 р. у таблиці 3.2.

Обчислити середньорічну чисельність працівників підприємства.

Таблиця 3.2 – Списочна чисельність працівників підприємства у 2003 р.

Дата	1.01.03	1.04.03	1.07.03	1.10.03	1.01.04
Чисельність працівників, осіб	260	265	270	274	270

Рішення:

Середньорічна чисельність працівників підприємства у моментних динамічних рядах з рівними проміжками між датами обчислюється за формулою:

$$\bar{y} = \frac{\frac{1}{2}y_1 + y_2 + ... + \frac{1}{2}y_n}{n-1} = \frac{0{,}5 \cdot 260 + 265 + 270 + 274 + 0{,}5 \cdot 270}{5-1} = 269 \text{ осіб.}$$

Для виявлення тенденції зміни досліджуваного явища у часі, розраховують **середню ковзну**. Суть способу її розрахунку полягає в тому, що склад періоду безперервно і постійно змінюється – відбувається зсув на одну дату при збереженні постійного інтервалу (трирічча, п’ятиріччя тощо).

В аналізі і плануванні сільськогосподарського виробництва застосовується також **середня прогресивна**. Цей вид середньої на відміну від загальної середньої дає узагальнену характеристику не всієї сукупності, а тільки тієї її частини, яка представлена показниками, вищими за загальну середню.

Середню прогресивну обчислюють у такій послідовності: 1) з усіх варіант обчислюють загальну середню; 2) відбираються варіанти, що за величиною перевищують загальну середню; 3) за відібраними варіантами обчислюють середню. Вона й буде середньою прогресивною. Наприклад, якщо сукупність представлена рядом чисел $x_1, x_2, ..., x_8$ та їх середнім значенням $\bar{x}$, серед яких x_1, x_2 і x_8 виявляться більшими за розміром, ніж загальна середня, то середня прогресивна становитиме:

$$\bar{x}_{прог} = \frac{x_1 + x_2 + x_3}{3}. \qquad (3.12)$$

Особливим видом середніх величин є **середня багатовимірна**, яка являє собою середню величину кількох ознак для однієї одиниці сукупності. Оскільки неможливо розрахувати середню величину за абсолютними значеннями різних ознак (різноякісних, виражених у різних одиницях виміру), то багатовимірна середня визначається з відносних величин (часток, процентів тощо), як правило, з відношень абсолютних значень для одиниці сукупності до середніх значень цих ознак.

Середня багатовимірна – похідна величина, що розраховується для статистичної сукупності чисельністю N одиниць з порядковими номерами i (i = 1, 2, 3, ..., N), які володіють k ознаками (x) з порядковими номерами j (j = 1, 2, 3, ..., k) у такий спосіб. Спочатку обчислюють відношення P_{ij} значень кожної ознаки (x) у кожної величини сукупності до її середнього значення за формулою:

$$P_{ij} = \frac{x_{ij}}{\bar{x}_j}, \qquad (3.13)$$

де x_{ij} – значення j-ої ознаки у i-ої одиниці сукупності;
$\bar{x}_j$ – середнє значення j-ої ознаки.

Після цього визначають середню з цих відношень для кожної одиниці сукупності, яку і називають багатовимірною середньою:

$$\overline{P_i} = \frac{\sum P_{ij}}{k}. \qquad (3.14)$$

Багатовимірні середні дають узагальнену характеристику кожної одиниці сукупності за кількома ознаками одночасно.

До так званої **структурної (позиційної)середньої** найчастіше застосовують моду і медіану. Величина моди і медіани залежить лише від характеру частот, тобто від структури розподілу. Якщо величина середньої арифметичної залежить від усіх значень ознаки, то величина моди і медіани не залежить від крайніх значень ознаки. Це особливо важливо для рядів розподілу, в яких крайні значення ознаки мають нечітко виражені межі (до і понад).

Модою називають значення ознаки, що має найбільшу частоту в статистичному ряді розподілу. Спосіб обчислення моди залежить від того, в якому вигляді дано значення ознаки: дискретного чи інтервального ряду розподілу. В дискретних варіаційних рядах моду обчислюють без додаткових розрахунків за значенням варіанти з найбільшою частотою.

В інтервальному варіаційному ряду розподілу модою наближено вважають центральний варіант так званого модального інтервалу, тобто того інтервалу, який має найбільшу частоту. В межах інтервалу необхідно знайти те значення ознаки, яке є модою.

В інтервальних варіаційних рядах розподілу моду визначають за формулою:

$$M_0 = x_0 + h\frac{f_2 - f_1}{(f_2 - f_1) + (f_2 - f_3)} , \qquad (3.15)$$

де x_0 – нижня (мінімальна) межа модального інтервалу;
h – величина інтервалу;
f_1 – частота передмодального інтервалу;
f_2 – частота модального інтервалу;
f_3 – частота після модального інтервалу.

Медіаною називають таке значення ознаки, яке поділяє ранжирований ряд розподілу на дві рівні частини, тобто значення, яке перебуває всередині ряду розподілу. Якщо в дискретному варіаційному ряду 2 m + 1 випадків, то значення ознаки у випадку m + 1 є медіанним. Якщо в ряду парне число 2 m випадків, медіану визначають як середню арифметичну з двох серединних значень.

Наприклад, якщо 17 комбайнерів держгоспу розташувати у порядку зростання, тобто в ранжирований ряд за кількістю намолоченого ними зерна, то намолот зерна у дев’ятого комбайнера буде медіанним. Якщо ж кількість комбайнерів буде 18 осіб, то медіаною буде середнє значення намолоту зерна дев’ятого і десятого комбайнерів.

В інтервальному варіаційному ряду розподілу медіану визначають за формулою:

$$M_e = x_0 + h\frac{\frac{\sum f}{2} - S_{M_e - 1}}{f_{M_e}} \qquad (3.16)$$

де x_0 – нижня (мінімальна) межа медіанного інтервалу;

h – величина інтервалу;

$\frac{\sum f}{2}$ – половина суми нагромаджених частот інтервального ряду розподілу;

$S_{M_e - 1}$ – сума нагромаджених частот інтервалу, що передує медіанному;

f_{M_e} – частота медіанного інтервалу.

Для визначення медіани в інтервальному варіаційному ряду розподілу треба обчислити нагромаджені частоти і відшукати медіанний інтервал. Під нагромадженими частотами розуміють наростаючий підсумок частот, починаючи з першого інтервалу. Медіанним є той інтервал, на який припадає перша нагромаджена частота, що перевищує половину всього обсягу сукупності. Наприклад, розрахунок моди і медіани в інтервальному варіаційному ряду розподілу покажемо на прикладі розподілу 100 господарств за надоєм молока на корову.

Рішення:

Інтервал, в якому міститься мода, буде 36 – 38 ц, тому що цей інтервал має найбільшу частоту табл. 3.3.

Підставивши відповідні числові значення у формулу моди, одержимо:

$$M_0 = x_0 + h\frac{f_2 - f_1}{(f_2 - f_1) + (f_2 - f_3)} = 6 + 2 \cdot \frac{25 - 19}{(25 - 19) + (25 - 14)} = 36{,}71 \text{ ц.}$$

Таблиця 3.3 – Розподіл 100 господарств за надоєм молока

Група	Групи господарств за надоєм молока на корову, ц	Кількість господарств	Нагромаджені частоти
I	30 – 32	9	9
II	32 – 34	5	24 (9 + 15)
III	34 – 36	19	43 (24 + 19)
IV	36 – 38	25	68 (43 + 25)
V	38 – 40	14	82 (68 + 14)
VI	40 – 42	11	93 (82 + 11)
VII	42 – 44	7	100 (93 + 7)
Разом	–	100	–

Отже, у досліджуваній сукупності найбільше число господарств має продуктивність корів 36,71 ц.

Обчислимо медіану за даними цього ж інтервального ряду розподілу, побудуємо ряд нагромаджених частот і знайдемо медіанний інтервал. Медіанним інтервалом є інтервал 36 – 38 ц, тому що на цей інтервал припадає перша нагромаджена частота, що перевищує половину всього обсягу сукупності (68 перевищує $\sum f : 2 = 100 : 2 = 50$).

Медіанне значення продуктивності корів становитиме:

$$M_e = x_0 + h\frac{\frac{\sum f}{2} - S_{M_e - 1}}{f_{M_e}} = 36 + 2 \cdot \frac{\frac{100}{2} - 43}{25} = 36{,}56 \text{ ц.}$$

Отже, продуктивність корів, рівна 36,56 ц, і є варіантою, що поділяє варіаційний ряд розподілу 100 господарств на дві рівні частини (50 господарств має надій на корову менше 36,56 ц і 50 господарств – більше 36,56 ц).

ТЕМА 4 ПОКАЗНИКИ ВАРІАЦІЇ

Варіацією ознаки називають наявність відмінностей в чисельних значеннях ознак у одиниць сукупності. Термін «варіація» походить від латинського слова variation – зміна, коливання, відмінність. Варіація є властивістю статистичної сукупності. Вона зумовлена множиною взаємопов'язаних між собою необхідних та випадкових, внутрішніх і зовнішніх факторів, серед яких є основні та другорядні. Основні фактори формують центр розподілу, другорядні – варіацію ознак, спільна їх дія – форму розподілу. Наприклад, продуктивність тварин залежить від рівня годівлі, породності, рівня механізації та інших об'єктивних факторів. Спільна їх дія зумовлює той чи інший рівень продуктивності тварин в окремих господарствах, а також закономірність розподілу господарств за цією ознакою.

Основними завданнями вивчення варіації ознак є: 1) вивчення міри варіації, тобто кількісного її вимірювання. Це завдання вирішується за допомогою розрахунку спеціальних показників варіації; 2) вивчення причин, які викликають варіацію, причинно-наслідкова оцінка характеру розсіювання, що передбачає дослідження закономірностей випадкової варіації в статистичних сукупностях; 3) розкладання загальної варіації ознаки на варіацію, що породжується систематичними та випадковими причинами.

Всі показники варіації поділяються на дві групи: абсолютні та відносні. До абсолютних належать: розмах варіації, середнє лінійне відхилення, дисперсія, середнє квадратичне відхилення, середнє квартальне відхилення. Друга група показників розраховується як відношення абсолютних показників варіації до середньої арифметичної (або медіани). Відносними показниками варіації є: коефіцієнт осциляції, варіації, відносне лінійне відхилення та ін. Кожний з названих показників має певні аналітичні переваги при вирішенні тих чи інших статистичних задач. Вибір того чи іншого показника залежить від конкретних завдань статистичного аналізу і наявної інформації.

Найпростішим показником варіації є **розмах варіації (*R*)**, який представляє собою різницю між максимальним і мінімальним значеннями ознаки.

$$R = x_{\max} - x_{\min} \tag{4.1}$$

В інтервальних рядах розподілу розмах варіації визначають як різницю між верхньою межею останнього та нижньою межею першого або як різницю між серединами інтервалів.

Більш досконалим показником вимірювання варіації є середнє лінійне та середнє квадратичне відхилення, які усувають зазначенні вище недоліки розмаху варіації.

Середнє лінійне відхилення ($\bar{d}$ *)* являє собою середню арифметичну з абсолютних значень відхилень окремих варіант від середньої арифметичної.

$$\bar{d} = \frac{\sum\left|x_i - \bar{x}\right|}{n} \text{ – просте (для незгрупованих даних),} \qquad (4.2)$$

$$\bar{d} = \frac{\sum\left|x_i - \bar{x}\right| \cdot f_i}{\sum f} \text{ – зважене (для варіаційного ряду).} \qquad (4.3)$$

Прямі дужки означають, що абсолютні значення відхилень беруться по модулю, тобто підсумовування виконується без врахування знаків (плюс або мінус). Така умовність пояснюється тим, що оскільки сума відхилень індивідуальних значень ознаки від середньої в першому ступені дорівнює нулю (нульова властивість середньої арифметичної), то для одержання суми всіх відхилень, відмінної від нуля, кожне відхилення слід брати як додатну величину.

Намагання скласти показник варіації, який би усував недоліки розмаху варіації та середнього лінійного відхилення призводить до дисперсії та середнього квадратичного відхилення.

Дисперсією (σ^2) називають середній квадрат відхилень індивідуальних значень ознаки від середньої арифметичної. Її визначають за формулою:

$$\sigma^2 = \frac{\sum(x_i - \bar{x})^2}{n} \text{ – просте (для незгрупованих даних),} \qquad (4.4)$$

$$\sigma^2 = \frac{\sum(x_i - \bar{x})^2 \cdot fi}{\sum f} \text{ – зважене (для варіаційного ряду).} \qquad (4.5)$$

Середнє квадратичне відхилення (σ) одержують шляхом добування кореня квадратного з дисперсії:

$$\sigma = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n}} \text{ – просте (для незгрупованих даних),} \qquad (4.6)$$

$$\sigma = \sqrt{\frac{\sum(x_i - \bar{x})^2 \cdot f_i}{\sum f}} \text{ – зважене (для варіаційного ряду.} \qquad (4.7)$$

Змістовне значення середнього квадратичного відхилення таке ж саме, як і середнього лінійного відхилення. Воно показує, на скільки в середньому відхиляються індивідуальні значення варіант від їх середнього значення.

Якщо показником центру розподілу використовується медіана, то для характеристики варіації можна застосувати так зване **квартильне відхилення** (Q):

$$Q = \frac{Q_3 - Q_1}{2}, \tag{4.8}$$

де Q_1, Q_3 – відповідно перший і третій квартилі розподілу.

Цей показник також можна застосувати замість розмаху варіації, щоб запобігти недоліків, пов'язаних з використанням крайніх значень ознаки.

При порівнянні коливання сукупностей, що мають різні одиниці вимірювання та значення середніх величин, робити висновки про ступінь варіації за середнім лінійним і середнім квадратичним відхиленнями важко. Тому з метою одержання порівняних даних, необхідно від абсолютних показників варіації перейти до відносних. Ці показники розраховуються як відношення абсолютних показників варіації до середньої арифметичної (медіани). Використовуючи за абсолютні показники варіації розмах варіації, середнє лінійне відхилення, середнє квадратичне відхилення та квартальне відхилення, одержимо відносні показники коливання (найчастіше вони виражаються у відсотках):

– коефіцієнт осциляції

$$V_R = \frac{R}{\overline{x}} \cdot 100\%; \tag{4.9}$$

– відносне лінійне відхилення

$$V_{\overline{d}} = \frac{\overline{d}}{\overline{x}} \cdot 100\%; \tag{4.10}$$

– коефіцієнт варіації

$$V\sigma = \frac{\sigma}{\overline{x}} \cdot 100\%; \tag{4.11}$$

– відносне квартильне відхилення

$$V_Q = \frac{Q}{\overline{x}} \cdot 100\% \text{ або } V_Q = \frac{Q}{Me} \cdot 100\% \text{ або}$$

$$V_Q = \frac{Q_3 - Q_1}{2Q_2} \cdot 100\%, \tag{4.12}$$

де Q – квартильне відхилення;

Q_1 – перший квартиль;

Q_2 – медіана;

Q_3 – третій квартиль.

Наприклад, розрахунок перелічених показників варіації здійснимо за даними розподілу 100 господарств за надоєм молока на корову (табл.4.1)

Таблиця 4.1 Розрахунок абсолютних показників варіації

Група	Групи господарств за надоєм молока на корову, ц	Середина інтервалу, ц	Кількість господарств	Середнє лінійне відхилення		Середнє квадратичне відхилення	
		x	f	$\lvert x_i - \bar{x} \rvert$	$\lvert x_i - \bar{x} \rvert \cdot f_i$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^2 \cdot f_i$
I	30-32	31	9	6	54	36	324
II	32-34	33	15	4	60	16	240
III	34-36	35	19	2	38	4	76
IV	36-38	37	25	0	0	0	0
V	38-40	39	14	2	28	4	56
VI	40-42	41	11	4	44	16	176
VII	42-44	43	7	6	42	36	252
Разом	-	-	100	-	266	-	1124

Середня арифметична $\bar{x}$ = 37 ц, перший квартиль – Q_1 = 34,11 ц, медіана – Q_2 = 36,56 ц, третій квартиль – Q_3 = 39 ц.

Перший квартиль знаходиться в інтервалі, в який входить перша нагромаджена частота, що перевищує чверть загального обсягу сукупності (0,25·100 = 25). Отже, перший квартиль Q_1 знаходиться в третьому інтервалі (з надоем молока від 34 до 36 ц), який має суму нагромаджених частот 43.

Щоб визначити третій квартиль, знайдемо інтервал, в якому він знаходиться. На цей інтервал припадає перша нагромаджена частота, що перевищує три чверті загального обсягу сукупності (0,75 · 100 = 75). Отже, третій квартиль знаходиться в інтервалі 38 – 40 ц, який має суму нагромаджених частот, рівну 82.

В інтервальному ряді розподілу перший і третій квартиль розраховують за такими формулами:

$$Q_1 = x_0 + h\frac{0,25\sum f - S_{Q_1} - 1}{f_{Q_1}} = 34 + 2 \cdot \frac{0,25 \cdot 100 - 24}{19} = 34,11 \text{ ц},$$

$$Q_3 = x_0 + h\frac{0,75\sum f - S_{Q_3} - 1}{f_{Q_3}} = 38 + 2 \cdot \frac{0,75 \cdot 100 - 68}{14} = 39 \text{ ц}.$$

Абсолютні показники варіації.
Розмах варіації:

$$R = x_{\max} - x_{\min} = 44 - 33 = 14 \text{ ц.}$$

Середнє лінійне відхилення (розрахунок показників у табл. 4.1):

$$\overline{d} = \frac{\sum \left| x_i - \overline{x} \right| \cdot f_i}{\sum f} = \frac{266}{100} = 2,66 \text{ ц.}$$

Дисперсія:

$$\sigma^2 = \frac{\sum (x_i - \overline{x})^2 \cdot f_i}{\sum f} = \frac{1124}{100} = 11,24 \text{ ц.}$$

Середнє квадратичне відхилення:

$$\sigma = \sqrt{\sigma^2} = \sqrt{11,24} = 3,4 \text{ ц.}$$

Квартильне відхилення:

$$Q = \frac{Q_3 - Q_1}{2} = \frac{39 - 34,11}{2} = 2,4 \text{ ц.}$$

Відносні показники варіації:
– коефіцієнт осциляції

$$V_R = \frac{R}{\overline{x}} \cdot 100\% = \frac{14}{37} \cdot 100\% = 37,8\%;$$

– відносне лінійне відхилення

$$V_{\overline{d}} = \frac{\overline{d}}{\overline{x}} \cdot 100\% = \frac{2,66}{100} \cdot 100\% = 7,2\%;$$

– коефіцієнт варіації

$$V\sigma = \frac{\sigma}{\overline{x}} \cdot 100\% = \frac{3,4}{37} \cdot 100\% = 9,2\%;$$

– відносне квартильне відхилення

$$V_Q = \frac{Q}{Me} \cdot 100\% = \frac{2,4}{36,56} \cdot 100\% = 6,6\%.$$

Отже, надої молока по даній сукупності господарств коливаються в межах ± 3,4 ц (по σ), або на 9,2 % по відношенню до середнього надою.

ТЕМА 5 ВИБІРКОВИЙ МЕТОД

Сукупність методів математичної статистики, що застосовуються для обґрунтувань та висновків при проведенні вибіркового спостереження, називають вибірковим методом.

Вибірковим спостереженням в статистиці називають такий вид спостереження, який дає можливість зробити висновок про всю сукупність одиниць при обстеженні тільки її частини. Прикладом вибіркового спостереження може бути вибіркове обстеження діяльності домогосподарств, визначення втрат урожаю, якості продукції, польові досліди тощо.

Сукупність відібраних для обстеження одиниць прийнято називати вибірковою, а сукупність одиниць, з якої проводиться відбір – генеральною.

Властивість вибіркової сукупності відтворювати генеральну сукупність дістала назву репрезентативності, що означає представництво з певною точністю і вірогідністю. В зв'язку з тим, що при вибірковому спостереженні обстежується тільки частина одиниць генеральної сукупності, то характеристики вибіркової сукупності, як правило, відрізняються від характеристик вибіркової сукупності, тобто мають місце так звані помилки репрезентативності (відповідності, відображення).

Тому одним із основних завдань вибіркового методу є одержання таких вибіркових характеристик, які б якомога точніше відтворювали характеристики генеральної сукупності, тобто давали найменші помилки репрезентативності.

Результати вибіркового спостереження характеризуються середніми і відносними узагальнюючими показниками. Узагальнюючі показники генеральної сукупності (середня, частка, дисперсія та ін.) називають генеральними, а відповідні узагальнюючі показники вибіркової сукупності – вибірковими.

У вибірковому спостереженні прийняті такі позначення. Обсяг генеральної сукупності позначають через N, а вибіркової через n. Середню величину і дисперсію ознаки в генеральній сукупності називають генеральною середньою і генеральною дисперсією. Генеральну середню позначають через $\bar{x}$, а генеральну дисперсію через σ_0^2.

Середню величину і дисперсію ознаки у вибірковій сукупності називають вибірковою середньою $\tilde{x}$ і вибірковою дисперсією σ^2. Генеральну частку позначають через p, а частість через ω.

Середнє квадратичне відхилення вибіркових середніх від генеральної середньої називають **середньою помилкою вибірки (μ).**

У практичних розрахунках застосовують дві формули середньої помилки вибірки: для середньої і для частки.

При вибірковому вивченні середніх показників середня помилка має вигляд:

– при повторній вибірці:

для середньої — для частки

$$\mu_{\tilde{x}} = \sqrt{\frac{\sigma^2}{n}}, \qquad \mu_{\omega} = \sqrt{\frac{w \cdot (1-w)}{n}}; \qquad (5.1)$$

– при безповторному відбору:

для середньої — для частки

$$\mu_{\tilde{x}} = \sqrt{\frac{\sigma^2}{n} \cdot (1-\frac{n}{N})}, \qquad \mu_{\omega} = \sqrt{\frac{w \cdot (1-w)}{n} \cdot (1-\frac{n}{N})}, \qquad (5.2)$$

де σ^2 – вибіркова дисперсія;

n – чисельність вибірки;

N – чисельність генеральної сукупності;

w – вибіркова частка.

Помилка вибірки, що обчислена із заданим ступенем надійної імовірності, називається **граничною помилкою вибірки** ε_i. Величина ε_i пов’язана з нормованим відхиленням *t*, яке визначається як відношення граничної помилки вибірки ε_i до середньої помилки μ:

$$t = \frac{\varepsilon_i}{\mu}. \qquad (5.3)$$

Із виразу можна знайти можливу граничну помилку вибірки:

$$\varepsilon_i = t \cdot \mu. \qquad (5.4)$$

Гранична помилка вибірки залежить від величини середньої помилки і нормованого відхилення і дорівнює ± кратному числу середніх помилок вибірки.

Нормоване відхилення функціонально пов’язане з ймовірністю. Для знаходження значень *t* складені спеціальні таблиці, за якими можна знайти значення *t* при заданому рівні довірчої ймовірності та значення ймовірності при відомому *t*.

Наведемо значення коефіцієнта довіри або нормованого відхилення *t* та відповідні до них імовірності (табл. 5.1) для вибірок з чисельністю $n \geq 30$, що найчастіше використовується в практичних розрахунках. Середня помилка в малих вибірках розраховується за формулою:

$$\mu_{\tilde{x}} = \frac{\sigma}{\sqrt{n}}. \tag{5.5}$$

Таблиця 5.1 – Таблиця значень ймовірностей

t	1,00	1,96	2.00	2,58	3,00
P	0,6827	0,9500	0,9545	0,9901	0,9973

Середнє квадратичне відхилення в малих вибірках визначається з врахуванням числа ступенів свободи варіації (n - 1):

$$\sigma = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}. \tag{5.6}$$

В математичній статистиці доводиться, що математичне сподівання вибіркової дисперсії не дорівнює дисперсії генеральної сукупності. Тому вибіркова дисперсія є зміщеною оцінкою генеральної дисперсії.

Для отримання незміщеної оцінки дисперсії генеральної сукупності необхідно вибіркову дисперсію (σ^2) помножити на так звану поправку Бесселя $\frac{n}{n\ -\ 1}$. Тоді виправлена, або скорегована, дисперсія ($\sigma_0{}^2$) може бути визначена за формулою:

$$\sigma_0{}^2 = \sigma^2 \cdot \frac{n}{n\ -\ 1}. \tag{5.7}$$

При n > 30 (великі вибірки) практично немає різниці між оцінками σ^2 і $\sigma_0{}^2$. При малих же значеннях (n < 30; малі вибірки) поправочний коефіцієнт значно відрізняється від одиниці. Тому при малому обсязі вибірки завжди потрібно користуватися незміщеною оцінкою дисперсії $\sigma_0{}^2$.

Наприклад: визначимо граничну помилку вибірки для середньої урожайності і для частки ділянок з урожайністю 25 ц/га і більше. Довірчий рівень ймовірності приймемо дорівнюючим P = 0,9545. За табл. 5.1 знайдемо значення t = 2.

Таблиця 5.2 – Розподіл урожайності ячменю

Групи ділянок за урожайністю ячменю, ц/га (x)	Кількість ділянок	
	Всього (генеральна сукупність) N	В тому числі відібрано (вибіркова сукупність) n
24	70	6
25	150	15
26	80	9
Разом	300	30

Рішення:

1. Визначимо за цими даними середню урожайність, дисперсію і частку ділянок з урожайністю 25 ц/га і більше для генеральної і вибіркової сукупностей:

а) для генеральної сукупності:

$$\bar{x} = \frac{\sum x_i \cdot f_i}{\sum f} = \frac{24 \cdot 70 + 25 \cdot 150 + 26 \cdot 80}{70 + 150 + 80} = 25,0 \text{ ц/га},$$

– дисперсія

$$\sigma_0^2 = \frac{\sum (x_i - \bar{x})^2 \cdot f\,i}{\sum f} = \frac{(24-25)^2 \cdot 70 + (25-25)^2 \cdot 150 + (26-25)^2 \cdot 80}{70 + 150 + 80} = 0,50,$$

частка ділянок з урожайністю 25 ц/га і більше

$$p = \frac{150 + 80}{300} = 0,77 \text{ або } 77\ \%;$$

б) для вибіркової сукупності:

середня урожайність

$$\tilde{x} = \frac{\sum x_i \cdot f_i}{\sum f} = \frac{24 \cdot 6 + 25 \cdot 15 + 26 \cdot 9}{6 + 15 + 9} = 25,1 \text{ ц/га},$$

дисперсія

$$\sigma^2 = \frac{\sum (x_i - \tilde{x})^2 \cdot f\,i}{\sum f} = \frac{(24-25,1)^2 \cdot 6 + (25-25,1)^2 \cdot 15 + (26-25,1)^2 \cdot 9}{6 + 15 + 9} = 0,49,$$

частка ділянок з урожайністю 25 ц/га і більше

$$\omega = \frac{15+9}{30} = 0,80 \text{ або } 80\ \%.$$

2. Середня помилка вибірки:
а) середньої урожайності ячменю

$$\mu_{\tilde{x}} = \sqrt{\frac{\sigma^2}{n} \cdot (1 - \frac{n}{N})} = \sqrt{\frac{0,50}{30} \cdot (1 - \frac{30}{300})} = \pm 0,12 \text{ ц/га;}$$

б) частки ділянок з урожайністю 25 ц/га і більше

$$\mu_p = \sqrt{\frac{p(1-p)}{n} \cdot (1 - \frac{n}{N})} = \sqrt{\frac{0,77 - 0,23}{30} \cdot (1 - \frac{30}{300})} = \pm 0,07.$$

Середня урожайність ячменю в генеральній сукупності $\tilde{x} \pm \mu_{\tilde{x}} = 25,1 \pm 0,12$ ц/га, тобто знаходиться в межах від 24,98 до 25,22 ц/га

Частка ділянок з урожайністю 25 ц/га і більше в генеральній сукупності $p = \omega \pm \mu_p = 0,80 \pm 0,07$, тобто знаходиться в межах від 73 % до 87 %.

Гранична помилка середньої урожайності ячменю:

$$\varepsilon_{\tilde{x}} = t \ \mu_{\tilde{x}} = 2 \ \ 0,12 = \pm\, 0,24 \text{ ц/га.}$$

Отже, різниця між вибірковою середньою урожайністю і генеральною середньою буде не більше 0,24 ц/га. Межі середньої урожайності в генеральній сукупності: $\bar{x} = \tilde{x} \pm \varepsilon_{\tilde{x}} = 25,1 + 0,24$, тобто від 24,86 до 25,34 ц/га.

Гранична помилка частки ділянок з урожайністю 25 ц/га і більше:

$$\varepsilon_p = t \cdot \mu_p = 2 \cdot 0,07 = \pm 0,14.$$

Отже, гранична помилка у визначенні частки ділянок з урожайністю 25 ц/га і більше не перевищить 14 %, тобто питома вага ділянок із зазначеною урожайністю в генеральній сукупності знаходиться в межах: $p = \omega \pm \varepsilon_p = 0,80 \pm 0,14$, тобто від 66 % до 94 %.

ТЕМА 6 МЕТОДИ ВИВЧЕННЯ ЗВ'ЯЗКУ МІЖ ЯВИЩАМИ ТА ЇХ КОРИСТУВАННЯ ДЛЯ УПРАВЛІННЯ СОЦІАЛЬНО-ЕКОНОМІЧНИМИ ПРОЦЕСАМИ

Вивчення реальної дійсності показує, що практично кожне суспільне явище знаходиться в тісному зв'язку і взаємодії з іншими явищами, якими б випадковими вони не здавалися на перший погляд.

Дослідження і вимірювання взаємозв'язків і взаємозалежностей соціально-економічних явищ є одним з найважливіших завдань статистики.

Термін кореляція – співвідношення, відповідність (взаємозв'язок, взаємозалежність) між ознаками, що виявляється при масовому спостереженні зміни середньої величини однієї ознаки залежно від значення іншої.

Зв'язки між факторами досить різноманітні. При цьому одні ознаки виступають в ролі факторів, що діють на інші, зумовлюючи їх зміну, другі – в ролі дії цих факторів. Перші з них називають **факторними ознаками**, другі – **результативними**.

Досліджуючи зв'язки між ознаками, необхідно виділити насамперед два види зв'язків: функціональний (повний) і кореляційний (статистичний) зв'язок.

Для орієнтованої оцінки тісноти зв'язків користуються непараметричними показниками статистики. До них відносять: коефіцієнт кореляції знаків, коефіцієнт кореляції рангів, коефіцієнт асоціації і коефіцієнт взаємної спряженості

Коефіцієнт кореляції знаків Фехнера застосовується для оцінки тісноти зв'язку на основі порівнянь знаків відхилень значень результативної і факторної ознак від їх середніх за формулою:

$$К_{ф} = \frac{\sum a - \sum b}{\sum a + \sum b}, \qquad (6.1)$$

де $\sum a$ – сума збігів знаків;

$\sum b$ – сума незбігів знаків.

Коефіцієнт Фехнера змінюється від 0 до ± 1.

Коефіцієнт кореляції рангів – це один з найпростіших показників тісноти зв'язку (його ще називають ранговим коефіцієнтом кореляції Спірмена). Суть його розрахунку полягає в тому, що парні спостереження двох взаємопов'язаних ознак (результативної і факторної) ранжируються, а потім відповідно величині ознаки їм надається ранг від 1 до n. Тіснота зв'язку визначається на основі близькості рангів.

$$\rho = 1 - \frac{6\sum d^2}{n^3 - n}, \tag{6.2}$$

де d – різниці між величинами рангів в порівнюваних рядах;
n – кількість спостережень.

Коефіцієнт кореляції рангів, як і лінійний коефіцієнт кореляції, може приймати значення від –1 до +1.

Тісноту зв’язку між атрибутивними (якісними) ознаками можна виміряти за допомогою спеціальних коефіцієнтів асоціації і контингенції.

Для розрахунку вказаних коефіцієнтів вимірювання тісноти зв’язку між альтернативними ознаками використовується таблиця взаємної спряженості у вигляді кореляційної таблиці, яка носить назву «чотирьохклітникової таблиці»:

Таблиця 6.1 – Взаємна спряженість

α	b	α+b
c	d	c+d
α+c	b+d	α+b+c+d

При застосуванні таблиці з частотами *α, b, c, d* обчислюється:
– коефіцієнт асоціації:

$$К_а = \frac{ad - bc}{ad + bc}. \tag{6.3}$$

У випадках, коли один з показників чотирьохклітинної таблиці відсутній, величина коефіцієнта асоціації буде дорівнювати одиниці, що дає завищену оцінку тісноти зв’язку між ознаками. У цьому випадку необхідно розраховувати:
– коефіцієнт контингенції:

$$К_{конт} = \frac{ad - bc}{\sqrt{(a+b)\cdot(c+d)\cdot(a+c)\cdot(b+d)}}. \tag{6.4}$$

Коефіцієнти можуть приймати будь-які значення від –1 до +1. Коефіцієнт контингенції завжди менше коефіцієнта асоціації.

Критерій χ^2 Пірсона (згоди) використовується при перевірці гіпотез про згоду (відповідність) вибіркового і теоретичного розподілів, про незалежність розподілів, про однорідність сукупностей.

Застосування критерію χ^2 вимагає дотримання ряду умов, найважливішими серед яких є:

– обсяг вибірки повинен бути досить великим (при $n < 50$ потужність критерію χ^2 значно знижується);

– чисельність окремих інтервалів (класів) має бути не менше п'яти одиниць. Якщо ця умова не виконується, то проводиться об'єднання малочисельних інтервалів з числом одиниць менше 5 (як виняток таких інтервалів може бути не більше 20 % від загальної кількості їх);

– частоти не можна перетворювати в частки, оскільки це може призвести до збільшення величини відхилень між емпіричними і теоретичними частотами ($n_{емп.} - n_{теор.}$).·

$n1$	$n2$
$n3$	$n4$

Якщо вихідні дані подано у вигляді чотирьохклітинної таблиці розподілу частот за двома ознаками (2 × 2) з чисельностями ni, то фактичне значення $\chi 2$ може бути визначене за формулою:

$$\chi^2 = \frac{(n_1 \cdot n_4 - n_2 \cdot n_3) \cdot (n_1 + n_2 + n_3 + n_4)}{(n_1 + n_2) \cdot (n_3 + n_4) \cdot (n_1 + n_3) \cdot (n_2 + n_4)}. \quad (6.5)$$

Щоб оцінити близькість емпіричного і теоретичного розподілів, необхідно розрахувати фактичне значення χ^2 і порівняти його з табличним значенням при заданому рівні значущості (α) і відповідній кількості ступенів свободи k.

Якщо вихідні дані подано у вигляді таблиці розподілу частот і необхідно перевірити гіпотезу щодо незалежності розподілу двох ознак, то кількість ступенів свободи визначають за формулою:

$$К = (\alpha - 1) \cdot (b - 1), \quad (6.6)$$

де α – кількість рядків;

b – кількість стовпців.

Якщо отримане за вибіркою значення $\chi^2_{факт} \le \chi^2_{а}$, то нульова гіпотеза приймається. Якщо ж $\chi^2_{факт} > \chi^2_{а}$, то нульова гіпотеза відхиляється.

При парній лінійній залежності тіснота зв'язку визначається за допомогою лінійного коефіцієнта кореляції:

$$r = \frac{\overline{x \cdot y} - \overline{x} \cdot \overline{y}}{\sigma_x \cdot \sigma_y} \quad (6.7)$$

$$\text{або } r=\frac{\sum(y-\bar{y})\cdot(x-\bar{x})}{n\cdot\sigma_x\cdot\sigma_y}; \tag{6.8}$$

$$\sigma_x=\sqrt{\frac{\sum x^2}{n}-(\bar{x})^2}; \tag{6.9}$$

$$\sigma_y=\sqrt{\frac{\sum y^2}{n}-(\bar{y})^2}. \tag{6.10}$$

Коефіцієнт кореляції знаходиться в межах від 0 до ± 1.

Ш. Кендел запропонував ще одну міру зв'язку між перемінними x і y – коефіцієнт кореляції рангів Кендела – τ.

$$\tau=\frac{2\cdot S}{n\cdot(n-1)}, \tag{6.11}$$

де S – алгебраїчна сума чисел вищих рангів стосовно кожного нижчого рангу, узятому послідовно як значення y і зіставленому поруч зі значенням x у висхідному і спадному порядках.

Коефіцієнт Кендела змінюється від -1 до +1.

У зв'язку з використанням експертних оцінок у різних галузях знань, одержав поширення ранговий коефіцієнт згоди Кендела (коефіцієнт конкордації), що обчислюється для оцінки ступенів тісноти зв'язку між декількома ознаками за наступною формулою:

$$\omega=\frac{12\cdot S}{m^2\cdot(n^3-n)}, \tag{6.12}$$

де S – сума квадратів відхилень сум по рядках від їх загального середнього значення.

m – кількість рядків

n – кількість стовпців,

Змінюється даний коефіцієнт у межах від 0 до 1.

В економічних дослідженнях взаємозв'язку двох факторів серед множини функцій часто розглядається прямолінійна форма зв'язку, яка виражається рівнянням прямої лінії:

$$\bar{y}_x=a+b\cdot x \tag{6.13}$$

де $\bar{y}_x$ – вирівняне значення результативної ознаки (залежна змінна);

x – значення факторної ознаки (незалежна змінна);

a – початок відліку або значення при b = 0 (економічного змісту не має);

b – коефіцієнт регресії, який показує середню змінну залежної змінної при зміні незалежної змінної на одиницю.

Криволінійні форми зв'язку досить різноманітні. В статистичному аналізі найчастіше використовують параболу другого порядку, гіперболу і ступеневу функцію.

1. Парабола другого порядку:

$$\overline{y}_x = a + b \cdot x + c \cdot x^2. \tag{6.14}$$

2. Експонентна форма тренда:

$$\overline{y}_x = a \cdot k^x, \tag{6.15}$$

де k – темп зміни в разах; константа тренда.

3. Логарифмічна форма тренда:

$$\overline{y}_x = a + b \cdot \log x. \tag{6.16}$$

4. Гіпербола:

$$\overline{y}_x = a + \frac{b}{x}. \tag{6.17}$$

Параметри в рівняннях регресії визначаються методом найменших квадратів. Цей метод найкращим чином відповідає кореляційній таблиці і припускає знаходження таких значень параметрів рівняння регресії, при яких сума квадратів відхилень табличних (фактичних) значень результативної ознаки y від теоретичних Y за лінією регресії була б мінімальною. Використовуючи рівняння регресії можна знайти теоретичне значення Y для будь-якого значення факторної ознаки x.

Для оцінки впливу факторної ознаки на результативну може розраховуватись коефіцієнт еластичності в середньому для усієї сукупності:

$$\mathrm{K}_e = b \cdot \frac{\overline{x}}{\overline{y}}, \tag{6.18}$$

де $\overline{x}, \overline{y}$ – середні величини фактичних даних відповідно за факторною та результативною ознаками, в цілому для сукупності.

Коефіцієнт еластичності показує, на скільки процентів у середньому зміниться результативна ознака при зміні факторної ознаки на 1 %.

Наприклад, розрахувати параметри лінійного рівняння парної регресії, що характеризує залежність між тижневим роздрібним товарообігом (грн) на душу населення та доходами населення (грн), та зробити аналіз параметрів регресії за даними:

Таблиця 6.2 – Дані товарообігу на душу населення та доходу населення

Доходи населення	18	20	21	22	24	25	27	28	29	31
Роздрібний товарообіг	17	18	19	20	21	23	24	25	26	27

Рішення:

За даними факторною ознакою (x) будуть тижневі доходи населення на одну особу за відповідними групами, а результативною ознакою (y) – роздрібний товарообіг (табл. 6.2).

Лінійне рівняння парної регресії, що дозволяє встановити теоретичну залежність Y за фактичними даними таблиці, записується у вигляді $Y = f(x)$:

$$Y = a + b \cdot x, \tag{6.19}$$

де a, b – параметри теоретичної залежності.

Отримаємо розрахункові параметри за допомогою методу найменших квадратів, записавши систему нормальних рівнянь за формулою 6.18.

Таблиця 6.3 – Визначення параметрів теоретичної залежності

№ п/п	x	y	x^2	$x \cdot y$	Y
1	18	17	324	306	16,67
2	20	18	400	360	18,31
3	21	19	441	399	19,31
4	22	20	484	440	19,95
5	24	21	576	504	21,59
6	25	23	625	575	22,41
7	27	24	729	648	24,05
8	28	25	784	700	24,87
9	29	26	841	754	25,69
10	31	27	961	837	27,33
Всього	245	220	6165	5523	
В середньому	24,5	22,0	616,5	552,3	

$$\begin{cases} n \cdot a + b\sum x = \sum y \\ a\sum x + b\sum x^2 = \sum y \cdot x \end{cases}. \tag{6.20}$$

Підставляючи в цю систему значення, обчислені за даними таблиці, дістанемо систему рівнянь:

$$\begin{cases} 10a + 245b = 220 \\ 245a + 6165b = 5523 \end{cases} \rightarrow b = 0,82 \quad a = 1,91.$$

Отже, теоретична залежність роздрібного товарообігу від доходів населення має вигляд:

$$Y = 1,91 + 0,82x.$$

Зробимо економічні висновки. Параметр b = 0,82 характеризує граничний розмір витрат на купівлю товарів у роздрібній торгівлі. Тобто, коли дохід збільшується на одиницю, то обсяг роздрібного товарообігу зростає на 0,82 грн.

При цьому можна визначити коефіцієнт еластичності, який показує ефект фактора x на результат y, тобто залежності роздрібного товарообігу від доходів населення:

$$\mathrm{K}_e = b \cdot \frac{\overline{x}}{\overline{y}} = 0,82 \cdot \frac{24,5}{22,0} = 0,91.$$

На підставі коефіцієнту еластичності можна дістати висновку, що зі збільшенням доходів населення на 1 % роздрібний товарообіг зростає на 0,91 %.

ТЕМА 7 МОДЕЛЮВАННЯ ТА АНАЛІЗ ДИНАМІКИ СОЦІАЛЬНО-ЕКОНОМІЧНИХ ЯВИЩ

Динамікою в статистиці прийнято називати процес розвитку, руху соціально-економічних явищ у часі.

Рядом динаміки називають ряд статистичних показників, які характеризують зміну суспільних явищ у часі.

Рівнем ряду динаміки називають статистичний показник, який характеризує величину суспільного явища на даний момент або за певний період часу. Розрізняють моментні та інтервальні ряди динаміки.

Моментними називають ряди динаміки, які характеризують розмір явища на певний період часу.

Інтервальними називають ряди динаміки, які характеризують розмір явищ за певний період часу.

Для виявлення напрямку та інтенсивності змін досліджуваних суспільних явищ за певні періоди часу визначають систему абсолютних і відносних показників динаміки. До таких показників належать: абсолютний приріст, темп (коефіцієнт) зростання, темп приросту, абсолютне значення одного процента приросту і середні показники ряду динаміки (середній рівень ряду динаміки, середній абсолютний приріст, середній темп зростання і приросту)

При порівнянні кожного рівня з одним і тим самим рівнем, взятим за базу порівняння, одержують **базисні показники**; при порівнянні кожного рівня з безпосередньо попереднім рівнем отримують **ланцюгові показники**.

Абсолютний приріст являє собою різницю між двома рівнями, один з яких взято за базу порівняння. Він показує, на скільки одиниць кожен даний рівень відрізняється від рівня, взятого за базу порівняння.

– *ланцюговий*:

$$A = y_n - y_{n-1}, \tag{7.1}$$

де y_n – поточний рівень;

y_{n-1} – попередній рівень ряду.

– *базисний*:

$$A = y_n - y_0, \tag{7.2}$$

де y_0 – початковий рівень ряду.

Темп (коефіцієнт) зростання (К) – це відношення двох рівнів, один з яких взято за базу порівняння. Він характеризує відносну швидкість зміни явища і показує у скільки разів кожний даний рівень більший або менший рівня, який взято за базу порівняння.

– ланцюговий: – базисний:

$$К_р = \frac{y_i}{y_{i-1}}; \qquad К_р = \frac{y_i}{y_0}. \qquad (7.3)$$

Якщо коефіцієнти росту виражають у відсотках, то вони називаються **темпами росту ($Т_Р$)**:

$$Т_р = К_р \cdot 100\,\%. \qquad (7.4)$$

Поряд з темпами зростання відносна зміна явища у часі може бути також охарактеризована за допомогою **темпів приросту**, які являють собою відношення абсолютного приросту до рівня, взятого за базу порівняння (Тпр):

$$Тпр = Тр – 100\,\%. \qquad (7.5)$$

Абсолютне значення 1 % приросту обчислюється тільки з ланцюгових абсолютних приростів розподілом їх на темпи приросту або розподілом попередніх рівнів на 100:

$$А\,\% = \frac{y_i - y_{i-1}}{Т_{пр}} \text{ або } А\,\% = \frac{y_{i-1}}{100}. \qquad (7.6)$$

Ця величина показує, наприклад, скільки злочинів утримується в 1 % приросту злочинності.

Середній рівень ряду динаміки називається ще середньою характеристикою. Методика його розрахунку залежить від виду ряду динаміки (інтервальний або моментний). Середній рівень інтервального ряду обчислюється за формулою середньої арифметичної простої:

$$\overline{y} = \frac{y_0 + y_1 + y_2 + \ldots + y_n}{n} = \frac{\sum_{i=1}^{n} y_i}{n}. \qquad (7.7)$$

Для моментного ряду з рівними інтервалами розраховується середня хронологічна проста:

$$\overline{y} = \frac{\frac{y_1}{2} + y_2 + y_3 + \ldots + \frac{y_n}{2}}{n} = \frac{\sum_{i=1}^{n} y_i}{n-1} \qquad (7.8)$$

де y_1, y_2, y_n – фактичні значення рівнів моментного ряду;

n – число рівнів.

Середній абсолютний приріст обчислюється за формулою середньої арифметичної простої з абсолютних приростів, розрахованих ланцюговим способом:

$$\overline{A} = \frac{\sum A_{\text{Л.П.}}}{n} \quad \text{або} \quad \overline{\text{A}} = \frac{y_n - y_0}{n - 1}, \qquad (7.9)$$

де n – кількість періодів або ланцюгових приростів.

Середній коефіцієнт росту обчислюється за формулою середньої геометричної з ланцюгових темпів росту:

$$\overline{\text{К}}_{\text{р}} = \sqrt[n]{\text{К}_1 \cdot \text{К}_2 \cdot ... \cdot \text{К}_n}, \qquad (7.10)$$

де К_1, К_2, ..., К_n – добуток ланцюгових коефіцієнтів росту;

n – число коефіцієнтів росту;

$$\text{або } \overline{\text{К}}_{\text{р}} = \sqrt[n-1]{\frac{y_n}{y_0}}, \qquad (7.11)$$

де n – кількість періодів часу, включаючи базисний;

y_n, y_0 – відповідно рівень звітного і базисного періодів.

Наприклад, є дані про виробництво зерна в одному з господарств за 5 років.

Таблиця 7.1 – Дані про виробництво зерна в господарстві

Рік	1993	1994	1995	1996	1997
Виробництво зерна, тис. ц., y_i	50	54	62	70	80

Розрахувати: середній рівень за 5 років; щорічні темпи росту; середнорічний темп росту за 4 роки.

Рішення:

1) Якщо інтервальний ряд, то середній рівень ряду (середньорічне виробництво зерна) визначимо як середню арифметичну просту:

$$\overline{y} = \frac{\sum y_i}{n} = \frac{50 + 54 + 62 + 70 + 80}{5} = 63,2 \text{ тис.ц.}$$

2) щорічні коефіцієнти росту знаходимо як відношення рівня кожного року до попереднього:

$$К_1 = \frac{54}{50} = 1{,}08 \; 1994 \qquad К_3 = \frac{70}{62} = 1{,}129 \; 1996$$

$$К_2 = \frac{62}{54} = 1{,}148 \; 1995 \qquad К_4 = \frac{80}{70} = 1{,}143 \; 1997$$

множачи коефіцієнти на 100 %, одержуємо темпи росту.

3) середньорічний темп росту можна розрахувати як середньо геометричне з річних темпів росту:

1варіант: $\overline{\text{Т}}\text{р}=\sqrt[n]{K_1 \cdot K_2 \ldots K_n} \cdot 100\,\% = \sqrt[4]{108 \cdot 114{,}8 \cdot 112{,}9 \cdot \ldots \cdot 114{,}3} = 112{,}5\,\%$.

2 варіант:

$\overline{\text{Т}}_{\text{р}} = \sqrt[n]{\frac{y_n}{y_0}} \cdot 100\,\% = \sqrt[4]{\frac{80}{50}} \cdot 100\,\% = 112{,}5\,\%$ середньорічний темп росту за 4 роки дорівнює 112,5 %.

Наприклад, є дані в таблиці 7.2:

Таблиця 7.2 – Чисельність населення міста за 5 років (на початок року)

Рік	1993	1994	1995	1996	1997
Чисельність населення, тис. осіб	72	78	83	87	90

Знайти лінію тренда, використовуючи отримане рівняння, визначити чисельність населення в 2000 р. (прогноз).

Рішення:

Припустивши, що чисельність населення зміниться в часі по прямій $\overline{y}_x = a + b \cdot x$ для перебування параметрів *a* і *b* вирішуємо систему нормальних рівнянь, що відповідають вимозі найменших квадратів.

$$\begin{cases} n \cdot a + b\sum t = \sum y \\ a\sum t + b\sum t^2 = \sum y \cdot t \end{cases}$$ Підставляючи отриману суму:

$$\begin{cases} 5\text{a} + 15\text{b} = 410 \\ 15\text{a} + 55\text{b} = 1275 \end{cases} \quad \text{b} = 4{,}5 \;\; \text{a} = 68{,}5.$$

Звідси рівняння тренда $y_t = 68{,}5 + 4{,}5t$. Підставляючи значення y_t (табл. 7.3, гр. 6). Для 2000 р. t = 8.

Таблиця 7.3 – Розрахунок трендових значень

Рік	Чисельність населення, тис. осіб., y_i	Умовна позначка часу	t^2	y_t	$y_t = 68{,}5 + 4{,}5t$
1	2	3	4	5	6
1993	72	1	1	72	73
1994	78	2	4	156	77,5
1995	83	3	9	249	82
1996	87	4	16	348	86,5
1997	90	5	25	450	91
Σ	410	15	55	1275	410

Отже, за прогнозом чисельність населення міста в 2000 р. складе

$68{,}5 + 4{,}5 \cdot 8 = 104{,}5$ тис.осіб.

ТЕМА 8 ІНДЕКСНИЙ МЕТОД АНАЛІЗУ

Індексами у статистиці називають складні відносні показники, що характеризують середню зміну сукупності, яка складається з несумірних елементів.

У статистиці розрізняють індивідуальні і загальні індекси. До індивідуальних індексів відносяться показники, що характеризують зміну одного якого-небудь індивідуального явища або його співвідношення, що виражають, з іншими показниками. Індивідуальними індексами є відносні величини планового завдання, виконання плану, динаміки і порівняння.

Звітним (1) називається період, рівні якого порівнюються, базисним (0) – період, з рівнем якого проводиться порівняння. Індивідуальний індекс позначається через *i*. При аналізі соціально-правових явищ найбільше часто розраховують наступні індивідуальні індекси:

1) індивідуальний індекс кількості або фізичного обсягу:

$$I_q = \frac{q_1}{q_0}, \tag{8.1}$$

де q_1 і q_0 – кількість зробленої продукції в звітному і базисному періодах.

2) індивідуальний індекс ціни:

$$I_p = \frac{p_1}{p_0}, \tag{8.2}$$

де p_1 і p_0 – ціна одиниці продукції відповідно в звітному і базисному періодах.

3) індивідуальний індекс собівартості:

$$I_z = \frac{z_1}{z_0}, \tag{8.3}$$

де z_1 і z_0 – собівартість одиниці виробу відповідно в звітному і базисному періодах.

4) індивідуальний індекс трудомісткості

$$I_t = \frac{t_1}{t_0}, \tag{8.4}$$

де t_1, t_0 – витрати часу на одиницю продукції відповідно в звітному і базисному періодах;

5) індекс обороту з продажу одного виду товару

$$I_{pq} = \frac{\sum p_1 q_1}{\sum p_0 q_0}, \tag{8.5}$$

де p_1q_1, p_0q_0 – оборот із продажу товару відповідно в звітному і базисному періодах.

Загальні індекси характеризують співвідношення величин складного явища. У статистиці використовують агрегатні індекси – основну форму загальних індексів. Розглянемо їхні види:

1) агрегатний індекс цін:

$$I_p = \frac{\sum q_1 p_1}{\sum q_1 p_0}, \tag{8.6}$$

де p_1 і p_0 – ціна одиниці продукції відповідно в звітному і базисному періодах;

q_1, q_0 – кількість продукції відповідно в звітному і базисному періодах.

2) агрегатний індекс фізичного обсягу (товарообігу):

$$I_q = \frac{\sum q_1 p_0}{\sum q_0 p_0}. \tag{8.7}$$

3) агрегатний індекс витрат на виробництво (собівартості)

$$I_z = \frac{\sum q_1 z_1}{\sum q_0 z_0}. \tag{8.8}$$

4) агрегатний індекс витрат праці (трудомісткості):

$$I_t = \frac{\sum q_1 t_1}{\sum q_0 t_0}. \tag{8.9}$$

У зв'язку з тим, що середні арифметичні і середні гармонійні індекси складні для розуміння, їх обчислюють тільки тоді, коли з якої-небудь причини відсутні дані, необхідні для обчислення агрегатного індексу.

Агрегатний індекс обсягу продукції у вигляді середнього арифметичного розраховується за формулою:

$$I_q = \frac{\sum i_q q_0 p_0}{\sum q_0 p_0}. \tag{8.10}$$

Середній гармонійний індекс цін також має практичне значення. Індекси різних цін основної маси товарів і загальний індекс цін для всієї сукупності товарів розраховуються за формулою середнього гармонійного індексу з вагами звітного періоду, тобто:

$$I_p = \frac{\sum q_1 p_1}{\sum \frac{q_1 p_1}{i_p}}. \tag{8.11}$$

Індексом змінного складу називають індекс середньої величини, оскільки він відбиває не лише зміни значень ознаки *x*, а й зміни в структурі сукупності:

$$I_{з.с.} = \bar{x}_1 : \bar{x}_0 = \frac{\sum x_1 f_1}{\sum f_1} : \frac{\sum x_0 f_0}{\sum f_0}. \qquad (8.12)$$

Індекс фіксованого складу показує, як у середньому змінилися значення ознаки при незмінній, фіксованій структурі:

$$I_{ф.с.} = \frac{\sum x_1 f_1}{\sum f_1} : \frac{\sum x_0 f_1}{\sum f_1}. \qquad (8.13)$$

Індекс структурних зрушень показує, як змінилася середня за рахунок структурних зрушень; значення ознаки *x* фіксуються на постійному рівні:

$$I_{стр} = \frac{\sum x_0 f_1}{\sum f_1} : \frac{\sum x_0 f_0}{\sum f_0} \quad \text{або} \quad I_{\text{стр}} = \frac{I_{\text{п.с.}}}{I_{\text{ф.с.}}}$$

Наприклад, є наступні дані в таблиці 8.1.

Таблиця 8.1 – Випуск продукції

Найменування виробів	Зміна випуску в травні в порівнянні з квітнем, %	Випуск продукції в квітні, грн.
Столи	+ 13	30
Дивани	+ 9	45
Стільці	+ 18	25

Визначити збільшення випуску всієї продукції в травні в порівнянні з квітнем (у відсотках), розрахувати загальний індекс фізичного обсягу.

Рішення:

Загальний індекс фізичного обсягу може бути розрахований як середній арифметичний:

$$I_{ф.о} = \frac{\sum i_q \cdot q_0 p_0}{\sum q_0 p_0} = \frac{1{,}13 \cdot 30 + 1{,}09 \cdot 45 + 1{,}18 \cdot 25}{30 + 45 + 25} = 1{,}125 \ (112{,}5\ \%)$$

У середньому по підприємству випуск продукції в травні в порівнянні з квітнем збільшений на 12,5 %.

ПЕРЕЛІК ПОСИЛАНЬ

1. Елисеева И. И., Общая теория статистики: учебник / под ред. чл.-корр. РАН И. И. Елисеевой, М. М. Юзбашев – 4-е изд., перераб. и доп. – М.: Финансы и статистика, 1999. – 480 с.
2. Практикум по теории статистики: учеб. пособие / под ред. проф. Р. А. Шмойловой. – М.: Финансы и статистика, 1999. – 416 с.
3. Теорія статистики: навчальний посібник / П. Г. Вашків, П. І. Пастер, В. П. Сторожук, Є. І. Ткач. – К.: Либідь, 2001. – 320 с.
4. Годин А. М. Статистика: учебник / А. М. Годин. – М.: Издательско – торговая компания «Дашков и К°», 2003. – 472 с.
5. Статистика: підручник / С. С. Герасименко та ін. К.: – КНЕУ, 1998. – 468 с.
6. Статистика: учебное пособие / Л. П. Харченко, В. Г. Довженкова, В. Г. Ионин и др. – М.: ИНФРА – М, 2001. – 384 с.
7. Уманець Т. В. Статистика: навч. посіб. / Т. В. Уманець, Ю. Б. Пігарєв. – К.: Вікар, 2003. – 623 с.
8. Голуб Л. А. Статистика в вопросах и ответах: учебное пособие / Л. А. Голуб, П. В. Жидченко. – Донецк: ДонГУ, 1999. – 168 с.
9. Єріна А. М. Теорія статистики: практикум / А. М. Єріна, З. О. Пальян. – К.: Товариство «Знання», КОО, 1997. – 325 с.
10. Захожай В. Б., Практикум з основ статистики: навч. посіб / В. Б. Захожай, І. І. Попов, О. В. Коваленко. – К.: МАУП. – 2001. – 176 с.
11. Скоробогатова Н. В. Практикум по статистике / Н. В. Скоробогатова, Г. С. Овечко. – Донецк: ДонГУ, 1996. – 129 с.
12. Лугінін О. Є. Статистика: підручник / О. Є. Лугінін, С. В. Білоусова. – К.: Центр навчальної літератури, 2005. – 580 с.
13. Мармоза А. Т. Теорія статистики: підручник / А. Т. Мармоза. – К.: Ельга, Ніка – Центр, 2003. – 392 с.