

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
МІСЬКОГО ГОСПОДАРСТВА імені О. М. БЕКЕТОВА

Н. О. Кондратенко, Т. М. Колесник, О. О. Вороніна

СТАТИСТИКА

КОНСПЕКТ ЛЕКЦІЙ

(для студентів всіх форм навчання першого (бакалаврського) рівня вищої освіти спеціальностей 281 – Публічне управління та адміністрування, 073 – Менеджмент, 242 – Туризм)

Харків
ХНУМГ ім. О. М. Бекетова
2021

УДК 351.1

Кондратенко Н О. Статистика : конспект лекцій для студентів всіх форм навчання для першого (бакалаврського) рівня вищої освіти спеціальностей 281 – Публічне управління та адміністрування, 073 – Менеджмент, 242 – Туризм / Н. О. Кондратенко, Т. М. Колесник, О. О. Вороніна ; Харків. нац. ун-т міськ. госп-ва ім. О. М. Бекетова. – Харків : ХНУМГ ім. О. М. Бекетова, 2021. – 136 с.

Автори:

д-р екон. наук, проф. Н. О. Кондратенко,
канд. екон. наук, доц. Т. М. Колесник,
канд. екон. наук, ст. викл. О. О. Вороніна

Рецензент

С. І. Плотницька, доктор економічних наук, професор кафедри менеджменту і публічного адміністрування Харківського національного університету міського господарства імені О. М. Бекетова

*Рекомендовано кафедрою менеджменту і публічного адміністрування,
протокол №1 від 29.08.2019.*

© Н. О. Кондратенко, Т. М. Колесник, О. О. Вороніна, 2019

© ХНУМГ ім. О. М. Бекетова, 2019

ЗМІСТ

Вступ.....	5
Тема 1 Методологічні засади статистики. Організація статистики в Україні.....	6
1.1 Становлення статистики як науки.....	6
1.2 Предмет, основні поняття і категорії статистичної науки.....	10
1.3 Організація і завдання статистики.....	17
Тема 2 Статистичне спостереження.....	20
2.1 Поняття про статистичне спостереження.....	20
2.2 Програмно-методологічні та організаційні питання спостереження.....	22
2.3 Форми, види та способи спостереження.....	25
2.4 Помилки спостереження та способи їх виправлення.....	28
Тема 3 Зведення та групування статистичних даних.....	30
3.1 Суть і завдання статистичного зведення.....	30
3.2 Групування, його суть, завдання та види.....	31
3.3 Основні питання методології статистичних групувань та їх класифікація в практичній діяльності	34
3.4 Статистичні ряди розподілу.....	38
Тема 4 Узагальнюючі статистичні показники та загальні принципи їх застосування	41
4.1 Суть і види статистичних показників	42
4.2 Абсолютні статистичні величини, одиниці їх вимірювання.....	43
4.3 Характеристика відносних величин.....	45
4.4 Середні величини, умови наукового їх застосування	51
4.5 Середня арифметична, основні її властивості.....	54
4.6 Середня гармонійна, її різновиди і сфера використання	58
4.7 Структурні середні – мода і медіана.....	60
Тема 5 Ряди розподілу та їх аналіз.....	64
5.1 Поняття варіації та її основні показники.....	64
5.2 Види дисперсії та методи її обчислення розрахунку.....	69
5.3 Розрахунок дисперсії методом моментів.....	72
Тема 6 Статистичні методи вимірювання взаємозв'язків.....	75
6.1 Види взаємозв'язків між явищами.....	75
6.2 Суть і значення дисперсійного аналізу.....	77
6.3 Основи кореляційно-регресійного аналізу.....	79
6.4 Множинна регресія і кореляція.....	84

Тема 7 Аналіз тенденцій розвитку.....	86
7.1 Поняття про ряди динаміки. Види рядів динаміки та правила їх побудови.....	86
7.2 Статистичні характеристики динамічних рядів та їх взаємозв'язок.....	89
7.3 Середні показники динаміки.....	91
7.4 Зіставність в рядах динаміки.....	94
7.5 Методи аналізу тенденцій розвитку.....	96
7.6 Аналітичне вимірювання ряду.....	99
7.7 Інтерполяція і екстраполяція. Аналіз сезонних коливань.....	102
Тема 8. Індексний метод аналізу.....	104
8.1 Загальне поняття про статистичні індекси. Види індексів.....	104
8.2 Індивідуальні індекси: методика визначення і економічний зміст.....	106
8.3 Агрегатний індекс як основна форма загального індексу.....	108
8.4 Середньозважені індекси.....	112
8.5 Індекси середніх величин (змінного складу, постійного складу, структурних зрушень).....	115
8.6 Факторний індексний аналіз.....	119
Тема 9 Вибіркове спостереження.....	122
9.1 Вибіркове спостереження, причини і умови його застосування	122
9.2 Способи і методи відбору одиниць у вибіркочу сукупність.....	124
9.3 Помилки вибіркового спостереження: зміст та методика розрахунку.....	126
9.4 Визначення обсягу вибірки та способи поширення результатів вибіркового спостереження.....	129
Список використаних джерел.....	131

ВСТУП

У системі економічної освіти місце статистики, як навчальної дисципліни, обумовлено її роллю у науковій і практичній діяльності суспільства. Статистика як окрема галузь суспільної науки і спеціальна наукова дисципліна має свій предмет і метод дослідження. Вона вивчає закономірності суспільних явищ у нерозривному зв'язку з їх якісним змістом. Статистика розробляє способи кількісного аналізу, які у сукупності утворюють статистичну методологію і використовуються іншими суспільними науками як засіб соціального пізнання.

У конспекті лекцій розглядаються такі питання, як: предмет, метод і теоретичні основи статистики; завдання та організація статистики; форми, види та способи проведення статистичного спостереження; зведення і групування матеріалів спостереження; обчислення узагальнюючих статистичних показників, вивчення динаміки суспільних явищ, тенденцій і закономірностей їх розвитку; показники і методи вивчення взаємозв'язків між явищами; правила побудови та аналізу таблиць; способи графічного зображення статистичних даних. Наведені приклади розрахунків показників побудовані на основі офіційних публікацій статистичних матеріалів або на умовних даних. Для більш наочного подання теоретичних питань використовуються рисунки, таблиці та графіки.

У виданні використано законодавчі, нормативні та інструктивні матеріали стосовно статистики, а також рекомендовані літературні джерела зі статистики.

Конспект лекцій розрахований на студентів вищих навчальних закладів усіх форм навчання, слухачів інститутів та факультетів післядипломної освіти, курсів перепідготовки, працівників підприємницьких та фінансових структур для самостійного вивчення дисципліни «Статистика».

ТЕМА 1 МЕТОДОЛОГІЧНІ ЗАСАДИ СТАТИСТИКИ.

ОРГАНІЗАЦІЯ СТАТИСТИКИ В УКРАЇНІ

- 1.1 *Становлення статистики як науки.*
- 1.2 *Предмет, основні поняття і категорії статистичної науки.*
- 1.3 *Організація і завдання статистики.*

1.1 Становлення статистики як науки

Слово статистика має латинське походження: від лат. «status – «стан», що означало політичний стан держави. У науку термін «статистика» ввів німецький учений Готфрід Ахенваль в 1746 році, запропонувавши замінити назву курсу «Державознавство» на «Статистику», поклавши тим самим початок розвитку статистики як науки і навчальної дисципліни. Незважаючи на це, статистичний облік вівся набагато раніше: проводилися перепису населення в Стародавньому Китаї, здійснювалося порівняння військового потенціалу держав, вівся облік майна громадян в Стародавньому Римі тощо.

Первиною формою статистики був господарський облік, поява якого відноситься до глибокої давнини і пов'язана з утворенням держав. Для управління державою потрібна була інформація про чисельність населення, склад земель, поголів'я худоби, стан торгівлі тощо.

З розвитком суспільства все складнішими ставали взаємовідносини між окремими народами, державами, суспільними формуваннями. Первинна реєстрація фактів, накопичення масових даних про суспільні явища і процеси, необхідність їх узагальнення зумовили потребу в кількісному їх вираженні та комплексній оцінці.

Розвиток бухгалтерського обліку та первинної реєстрації фактів, накопичування масових даних про суспільні явища і необхідність їх узагальнення, підвищення попиту щодо кількісного вимірювання явищ і закономірностей суспільного життя, розвиток таких фундаментальних наук, як

філософія і математика, – ось неповний перелік умов, завдяки яким стало неминучим формування статистики.

Вважається, що статистика як самостійна наука почала формуватися і розвиватися з середини XVII століття. Біля витоків статистичної науки стояли дві школи: німецька описова і англійська школа політичних арифметиків. Представники описової школи (Герман Конрінга, Готфрід Ахенваль, Август Людвіг Шленцер) своїм завданням вважали опис пам'яток держави: території, населення, клімату, політичного устрою, віросповідання, торгівлі тощо – без аналізу закономірностей і зв'язків між явищами.

Представники школи політичних арифметиків (Вільям Петті, Джон Граунт, Едмунд Галлей) своїм головним завданням вважали виявлення на основі великого числа спостережень різних закономірностей і взаємозв'язків в досліджуваних явищах.

Кожна школа розвивалася своїм шляхом, використовуючи свої методи в дослідженнях, але предмет вивчення у них був загальний – держава, суспільство і, зокрема, масові явища і процеси, що відбуваються в ньому. Статистика сформувалася як наука в результаті синтезу державознавства і політичної арифметики, причому від останньої вона взяла більше, оскільки статистика і в даний час покликана виявляти насамперед різного роду закономірності в досліджуваних явищах.

На Русі першими і основними обліково-статистичними джерелами були літописи (IX–XI ст.). Як метод пізнання явищ, вітчизняна статистика почала формуватися у XVIII–XIX ст.

Найбільш прогресивні для цього часу теоретичні основи статистики як самостійної науки були створені Д. П. Журавським (1810–1856рр.). Вершиною його статистичного дослідження стала робота «Статистичний опис Київської губернії», у якій велику увагу він приділив питанню достовірності даних, методу групування, розкрив принцип єдності кількісного і якісного аналізу. Не можна не згадати і такий напрям статистичної практики, як земська статистика. Головним завданням земств було вивчення загальних умов життя народу,

бюджету селянських сімей за дуже розгорнутими програмами спостереження, з використанням методів групування. Представниками Чернігівського земства були П. Червинський (1849–1931 рр.) та О. Шлікевич (1849–1909 рр.), визнаний усіма винахідник особливого типу таблиць – комбінаційних. На початку ХХ ст. відбулося формування основних наукових напрямків вітчизняної статистики, проникнення ідей англо-американської статистики в середовище вітчизняних вчених, зародження й злет стохастичної статистики (від грецького «стохастика» – припускати), а також порівнювальної статистики. Статистичний метод стає універсальним. Представниками цих наукових напрямків були видатні вчені Ю. Янсон, О. І. Чупров, О. О. Чупров, М. Каблуков, А. Кауфман. Становлення статистики як науки в сучасному її розумінні пов'язане з іменами Ф. Гальтона (1822–1911 рр.) та К. Пірсона (1857–1936 рр.) – засновниками біометрії, кореляційно-регресійного аналізу; П. Чебишева (1821–1894 рр.), А. Маркова (1856–1922 рр.), О. Ляпунова (1857–1918 рр.), О. О. Чупрова (1890–1968 рр.) – математичні основи вибіркового методу; Р. Фішера (1890–1968 рр.) – теорія статистичного оцінювання та перевірки гіпотез і багатьох-багатьох інших. Отже історія статистики доводить, що статистична наука сформувалася внаслідок теоретичних узагальнень накопиченого людством досвіду обліково-статистичних робіт (практична статистика існує майже 4000 років, в той час як теоретична – лише близько 300), зумовлених потребами управління виробництвом і життям суспільства. Сучасна статистична наука є складною, багатогалузевою системою наукових дисциплін.

Сьогодні статистику можна розглядати в трьох основних значеннях:

- як вид діяльності;
- як наука або галузь знання;
- як інформація.

Як вид діяльності статистика вирішує питання збору, накопичення, обробки й аналізу даних про явища й процеси, що відбуваються в суспільстві. Організаційно-технічною базою статистики є мережа спеціалізованих

статистичних органів (Державна служба статистики, її регіональні підрозділи тощо), а також підприємства, організації й установи з їхньою матеріальною базою й кадрами, що представляють у статистичні органи за встановленою формою й у встановлений час звітність про свою діяльність.

Статистика як наука являє собою галузь знання, що вивчає соціально-економічні явища в суспільстві й утворені в них кількісні характеристики й взаємозв'язки. Це багатогалузева наука яка складається з трьох основних частин:

- загальна теорія статистики;
- соціально-економічна статистика;
- статистика галузей промисловості й усього господарства (транспорту, с/г, торгівлі тощо).

Статистика, як інформація являє собою набір даних, відомостей про ті або інші явища, що відбуваються в суспільстві, які є необхідні органам управління різних рівнів для прийняття рішень.

Значним кроком у розвитку статистичної науки послужило застосування економіко-математичних методів і широке використання комп'ютерної техніки в аналізі соціально-економічних явищ і процесів.

Між статистичною наукою і практикою існує тісний зв'язок і залежність. Статистична наука використовує дані практики, узагальнює їх і розробляє методи проведення статистичних досліджень. В свою чергу, в практичній діяльності застосовуються теоретичні положення статистичної науки для рішення конкретних управлінських завдань. Протягом тривалого і складного процесу формування статистики ці дві функції постійно взаємодіяли, і практичні потреби неминуче впливали на розвиток статистичної науки. Статистика як наука завжди обслуговувала та виконувала завдання того суспільного ладу, який був її основним фундатором і користувачем статистичної інформації.

1.2 Предмет, основні поняття і категорії статистичної науки

Статистика як наука має свій предмет дослідження. Об'єктом вивчення статистики є людське суспільство, явища і процеси суспільного життя. Специфічна риса предмета – це призначення статистики говорити за допомогою цифр. Вивчаючи кількісну сторону явищ, статистика відображає її в своїх числах – показниках, і саме цим характеризує конкретну міру явищ, встановлює загальні властивості, виявляє схожість і відмінність окремих рис, виявляє певні типи явищ.

Статистику як суспільну науку характеризують такі важливі риси:

- статистика вивчає суспільні явища;
- кількісна сторона суспільних явищ вивчається статистикою у нерозривному зв'язку з їх якісною стороною (наприклад: кількість книжок виданих у 2010 р. – 6 308, з них українською мовою – 3 140, російською – 2 364 тощо);
- кількісну сторону суспільних явищ статистика вивчає в конкретних умовах місця і часу;
- статистика вивчає масові суспільні явища;
- статистика також вивчає вплив природних і технічних чинників на кількісні зміни суспільного життя і вплив розвитку суспільства на природні умови життя суспільства.

Отже, предметом статистики як суспільної науки є кількісна сторона масових суспільних явищ і процесів у нерозривному зв'язку з їх якісною стороною, кількісне відображення закономірностей суспільного розвитку в конкретних умовах місця й часу.

Розглядаючи суспільні явища як масові і спираючись на облік всієї сукупності фактів, що відносяться до цих явищ, статистика за допомогою чисел показує ступінь їх розвитку, напрям і швидкість змін, тісноту взаємозв'язків і взаємозалежностей. Все це дає підставу стверджувати, що статистика – один із засобів пізнання суспільного життя.

Життя настільки складне і різноманітне, що майже завжди можна знайти факти, приклади, які б або підтвердили, або спростували одне і теж твердження. Тому, щоб охарактеризувати масове суспільне явище в цілому, необхідно розглянути всю чи дуже велику масу окремих явищ та процесів, що його стосуються. Таку масу називають статистичною сукупністю, що є об'єктом статистичного дослідження.

Для вивчення цього об'єкта в статистиці застосовують ряд важливих понять і категорій:

- статистична сукупність;
- одиниця статистичної сукупності;
- одиниця статистичного спостереження;
- ознака;
- варіація;
- статистична закономірність;
- статистичний показник.

Під *статистичною сукупністю* розуміють масу однорідних у певному відношенні елементів (фактів, явищ тощо), які мають єдину якісну основу, але різняться між собою певними ознаками. Статистична сукупність зовсім не обов'язково вміщує велику кількість одиниць, вона може бути і досить маленькою, наприклад, обсяг сукупності малої вибірки може складати іноді 5 – 10 одиниць. Важливо пам'ятати, що статистична сукупність складається з реально існуючих матеріальних об'єктів, які підлягають дослідженню. Наприклад: в якості такої множини об'єктів можуть виступати навчальні заклади, підприємства, сім'ї, студенти, громадяни будь-якої країни тощо.

Характерні риси статистичної сукупності наведені на рисунку 1.1.

Одиниця статистичної сукупності – це окремий елемент множини, який має всі властивості, що притаманні досліджуваній сукупності. Одиниця сукупності є носієм ознак, значення яких підлягають реєстрації. Наприклад: студенти інституту – це сукупність, кожен студент – її елемент, який має

спільні з іншими елементами ознаки: стать, вік, спеціальність, екзаменаційна оцінка тощо.

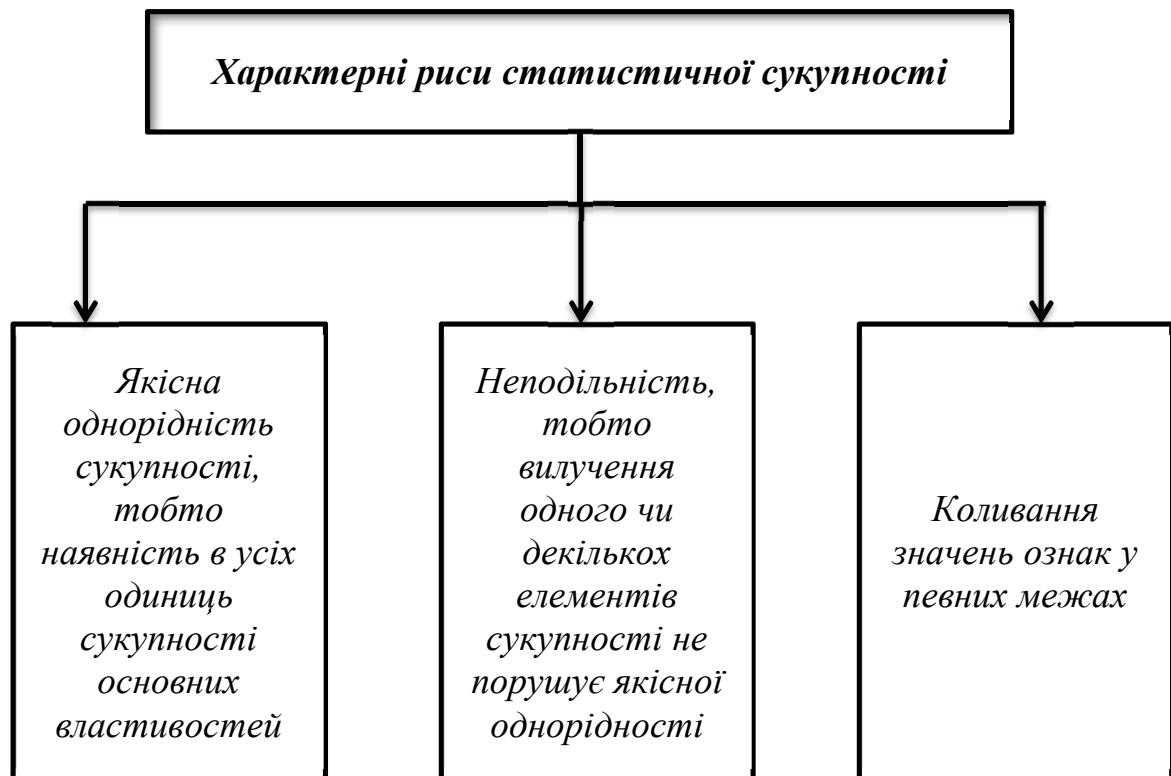


Рисунок 1.1 – Характерні риси статистичної сукупності

Одиниця статистичного спостереження – це первинна одиниця, від якої одержують інформацію. Зверніть увагу на те, що одиниці сукупності та одиниці спостереження можуть збігатися, наприклад, при вивченні суспільної думки населення це – кожний респондент – людина, яка висловлює свою думку. Але найчастіше обидві одиниці не збігаються. Так, у разі дослідження ефективності роботи друкарського обладнання одиницею спостереження є власник обладнання, а одиницею сукупності – окремий верстат.

Ознака – це статистичний еквівалент властивостей (рис, якостей), що притаманні елементам (одиницям) сукупності і відбивають їх сутність, характер та особливість. Наприклад: рівень акредитації, факультети, відділення, спеціалізації, термін та форми навчання – ознаки закладів освіти.

За характером вираження ознаки поділяються на такі:

- атрибутивні (описові, якісні) – не мають кількісного вираження, а лише словесне. Не піддаються прямому числовому визначенню; їх можна виразити тільки числом одиниць в сукупності або реєстрацією наявності чи відсутності властивостей (альтернативними), наприклад: стать, професія, ім'я тощо;
- кількісні (варіаційні) – мають числове вираження. На відміну від якісних вони можуть бути вимірюваними, тобто визначеними шляхом вимірів, зважувань і підрахунків (ціни, чисельність, вага, зріст, стаж тощо).

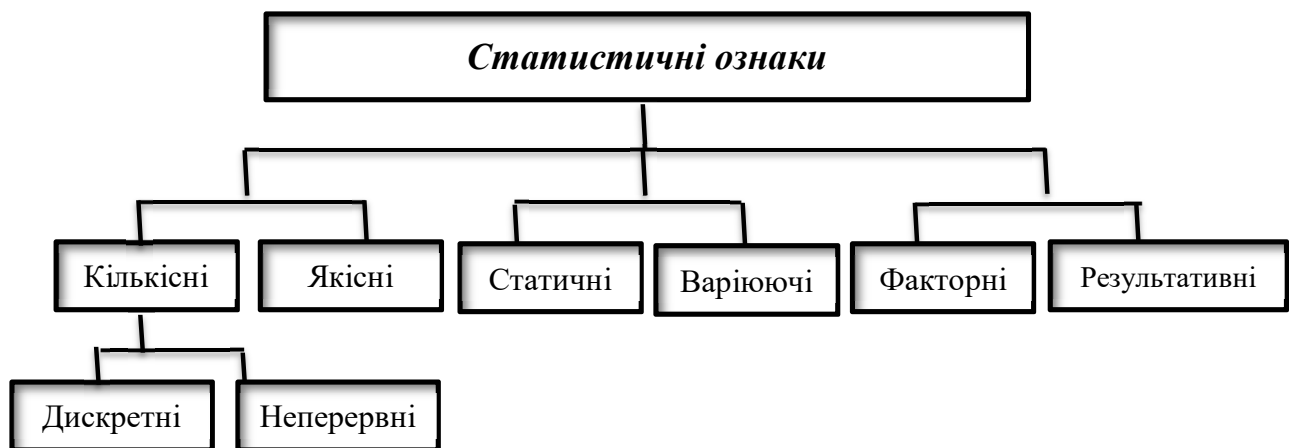


Рисунок 1.2 – Класифікація статистичних ознак

Варіація – це зміна значень конкретної ознаки при переході від однієї одиниці сукупності до іншої, інакше кажучи, це окреме значення (індивідуальний рівень) ознаки кожного елемента (одиниці) сукупності, наприклад, «стать» – чоловіча, жіноча, «оцінка» – 2, 3, 4, 5 тощо. Саме наявність варіації зумовлює необхідність статистики.

За характером варіації кількісні ознаки поділяються на дискретні та неперервні. Дискретні ознаки мають лише окремі числові значення (розряд робітника, кількість дітей в сім'ї, кількість поверхів у будинку). Неперервні ознаки мають будь-які значення в певних межах варіації (вік людини, зарплата, ціна одиниці товару тощо).

Рівень значення ознаки у окремих елементів може вимірюватися за допомогою номінальної шкали (шкали якісних величин, найменувань),

порядкової (рангової) шкали, метричної шкали (кількісних фізичних, економічних величин).

Крім того, ознаки поділяються на статичні і варіюючі. Статичні ознаки мають незмінні значення в усіх одиниць сукупності. Варіюючі ознаки приймають різні значення в межах статистичної сукупності. Ознаки, що характеризують елементи статистичної сукупності, взаємопов'язані між собою, тому розрізняють факторні та результативні ознаки. Факторні ознаки – це ознаки, що впливають на інші, пов'язані з ними ознаки (ціна, кількість реалізованих товарів). Результативними називають ознаки, розмір і динаміка яких формуються під впливом інших ознак (товарооборот).

Окремо взяті елементи будь-якої сукупності характеризуються практично необмеженим числом різних ознак. Які саме з цих ознак підлягають вимірюванню в конкретному випадку, залежить від мети дослідження.

Статистична закономірність – це кількісна закономірність змін у просторі та часі масових явищ і процесів суспільного життя, які складаються з множини елементів (одиниць сукупності). Це форма прояву причинного зв'язку, виражена у послідовності, регулярності, повторюваності подій, якщо причини (умови), що викликали ці події не змінюються або мають незначні зміни. Характерною особливістю статистичної закономірності є те, що вона властива не окремим одиницям сукупності, а усій сукупності в цілому і виявляється лише в масових явищах при значній кількості одиниць сукупності. Пояснити це можна так: формування масового процесу відбувається під впливом основних причин (загальних) і причин випадкових (індивідуальних). Якщо кількість подій достатньо велика, вплив випадкових причин усувається, взаємно урівноважується (закон великих чисел). Наприклад, успішне навчання окремих студентів залежить від загальних для усіх студентів причин (організація навчального процесу, наявність в достатній кількості підручників, комп'ютерів та спецобладнання), а також від індивідуальних (бажання, здібності, стан здоров'я, матеріальні умови). Якщо дослідити рівень успішності

багатьох студентів у нормальних (середніх) умовах навчання, то виявиться, що навчаються вони переважно добре, оскільки загальні причини є основними.

Статистичний показник – це узагальнююча кількісна характеристика властивостей сукупності загалом чи її частин зокрема щодо конкретних умов місця і часу. Статистичні показники поділяються на абсолютні, відносні та середні величини. Якісний зміст показника визначається суттю явища і знаходить своє відображення у назві. Кількісну сторону представляють число та розмірність (вимірник). Наприклад: середня заробітна плата працівників освіти України за січень 2018 року становила 6 578 грн. На відміну від значень ознак значення статистичних показників отримують розрахунковим шляхом. Це може бути простий підрахунок одиниць сукупності (кількість студентів академічної групи – 30 осіб), сума їх значень ознаки (з них: дівчат – 20 осіб, хлопців – 10 осіб), порівняння двох або декількох величин (дівчат у два рази більше) чи більш складні розрахунки (частка дівчат у групі – 67 %). Важливо правильно визначити зміст статистичних показників та методи їх побудови.

Для вивчення свого предмету статистика розробляє і використовує різноманітні методи, сукупність яких утворює статистичну методологію.

Статистична методологія – це сукупність науково обґрунтованих способів, правил і методів статистичного вивчення масових соціально-економічних явищ та процесів, які встановлюють порядок збирання, опрацювання і аналізу статистичної інформації.

Статистична методологія ґрунтується на загально-філософських (діалектична логіка) принципах і загальнонаукових (порівняння, аналіз, синтез) методах.

Згідно з принципами діалектичної логіки статистика будь-яке суспільне явище розглядає не ізольовано, а у взаємозв'язку з іншими, виявляє чинники, які спричиняють варіацію значень ознак у межах сукупності, оцінює ефект впливу чинників і щільність причинно-наслідкових зв'язків.

Суспільні явища динамічні, тому статистика вивчає їх у розвитку, оцінюючи тенденції та циклічні коливання, інтенсивність динаміки та

структурних зрушень. Статистична методологія ґрунтується на поєднанні аналізу і синтезу. Розглядаючи сукупності елементів, статистика, з одного боку, визначає в них схожі риси і відмінності, об'єднує елементи в групи, виділяючи при цьому різні типи й форми явищ, а з іншого – узагальнює інформацію як за окремими групами (типами), так і за сукупністю в цілому.

Особливості статистичної методології пов'язані, по-перше, з точним вимірюванням і кількісним описуванням масових суспільних явищ; по-друге, з аналізом їх диференціації; по-третє, з використанням узагальнюючих показників для характеристики об'єктивних статистичних закономірностей.

Всю безліч статистичних методів можна систематизувати за їх цільовим призначенням залежно від етапу статистичного дослідження

Будь-яке статистичне дослідження містить в собі такі етапи:

- розробка програми статистичного дослідження;
- проведення статистичного спостереження для збору первинної інформації;
- статистичні зведення та угруповання отриманої інформації;
- обробка отриманих результатів і виявлення закономірностей у розвитку явища.

На кожному з етапів статистичного дослідження використовують ті методи, які спроможні дати глибоку всебічну характеристику явищ, що вивчаються.

До основних прийомів і способів статистичного дослідження належать:

1. Статистичне спостереження – це збір первинного матеріалу шляхом реєстрації фактів. Масове статистичне спостереження дає інформаційну базу для статистичних узагальнень і характеристики об'єктивних закономірностей.

2. Статистичне зведення і групування матеріалів статистичного спостереження. Зібрані факти класифікують і систематизують, їх поділяють за ознаками відмінності і об'єднують за ознаками подібності. Найважливішим специфічним прийомом на цьому етапі дослідження є метод групування, за

допомогою якого досліджувані явища поділяють на характерні групи і підгрупи. Наприклад, населення можна класифікувати за статтю, за соціальним станом, за місцем проживання тощо. За допомогою групувань виокремлюються характерні риси та різноякісні типи явищ.

3. Статистичний аналіз зведеного та опрацьованого матеріалу, в процесі якого виявляються закономірності та зв'язки суспільних явищ. В арсеналі статистичних методів аналізу – методи вивчення варіації, диференціації та сталості, швидкості та інтенсивності розвитку, узагальнюючі індекси, регресійні моделі тощо. Специфічним для статистичного аналізу є застосування узагальнюючих показників – підсумкових, відносних і середніх величин, статистичних коефіцієнтів.

1.3 Організація і завдання статистики

Вивченням економічного і соціального розвитку країни, окремих її регіонів, галузей, об'єднань, фірм, підприємств займаються спеціально створені для цього органи, сукупність яких називається статистичною службою. В Україні функції статистичної служби виконують органи державної статистики і органи відомчої статистики. Централізоване керівництво веденням статистики в нашій країні здійснює спеціально уповноважений центральний орган виконавчої влади в галузі статистики, який утворюється відповідно до статті 106 Конституції України. З 1997 по 2011 роки функції державного управління у сфері статистики здійснює Державний комітет статистики України. З 2012 р. згідно з Указом Президента України слова «Державний комітет статистики України» у всіх відмінках замінити відповідно словами «Державна служба статистики України» у відповідному відмінку «Про Державну службу статистики від» від 22.02.2012 за № 396/2011) – утворилась служба статистики.

До структури Державної служби статистики України належать:

- Головне управління статистики підприємств;
- Управління статистики промисловості;

- Управління статистики інвестицій та будівництва;
- Управління структурної статистики;
- Управління зведеної інформації;
- Управління методології та планування;
- Управління національних рахунків;
- Управління статистики послуг та соціальних програм;
- Управління статистики праці;
- Управління статистики населення;
- Управління статистики сільського господарства та навколишнього середовища;
- Управління статистики торгівлі;
- Управління статистики зовнішньої торгівлі;
- Управління статистики фінансів;
- Управління статистики цін;
- Управління статистики зарубіжних країн та міжнародного співробітництва;
- Управління обстеження умов життя домогосподарств;
- Управління регіональної статистики;
- Управління інформатизації Усередині департаментів є ряд відділів.

Зазначені органи становлять єдину систему органів державної статистики України.

Державна служба статистики України виконує роботу зі збирання, опрацювання та аналізу науково обґрунтованих даних, що характеризують економічний соціальний розвиток країни, хід виконання державних і регіональних програм, пов'язаних із рішенням важливих народногосподарських проблем. Ці дані подаються у встановлені терміни до Верховної Ради, адміністрації Президента, Кабінету Міністрів України, місцевих органів державної виконавчої влади, представницьких органів регіонального самоврядування, міністерств, відомств, інших керівних органів. Організацію статистичної роботи на місцях здійснюють територіальні органи статистики.

Основними завданнями органів державної статистики є такі:

- участь у формуванні державної політики в галузі статистики та забезпечення її реалізації;
- збирання, опрацювання, аналіз, поширення, збереження, захист та використання статистичної інформації щодо масових економічних, соціальних, демографічних, екологічних явищ і процесів, які відбуваються в Україні та її регіонах;
- забезпечення надійності та об'єктивності статистичної інформації;
- розроблення, вдосконалення і впровадження статистичної методології;
- забезпечення розроблення, вдосконалення та впровадження системи державних класифікаторів техніко-економічної та соціальної інформації, які використовуються для проведення статистичних спостережень;
- створення і ведення Єдиного державного реєстру підприємств та організацій України та реєстрів респондентів статистичних спостережень;
- впровадження новітніх інформаційних технологій з опрацювання статистичної інформації;
- взаємодія інформаційної системи органів державної статистики з інформаційними системами державних органів, органів місцевого самоврядування, інших юридичних осіб, міжнародних організацій та статистичних служб інших країн шляхом взаємного обміну інформацією, проведення методологічних, програмно-технологічних та інших робіт, спрямованих на ефективне використання інформаційних ресурсів;
- координація дій державних органів, органів місцевого самоврядування та інших юридичних осіб у питаннях організації діяльності, пов'язаної із збиранням та використанням адміністративних даних;
- забезпечення доступності, гласності й відкритості статистичної інформації, її джерел та методології складання;
- збереження і захист статистичної інформації.

ТЕМА 2 СТАТИСТИЧНЕ СПОСТЕРЕЖЕННЯ

2.1 *Поняття про статистичне спостереження.*

2.2 *Програмно-методологічні та організаційні питання спостереження.*

2.3 *Форми, види та способи спостереження.*

2.4 *Помилки спостереження та способи їх виправлення.*

2.1 Поняття про статистичне спостереження

Першим етапом статистичного дослідження є збір статистичної інформації про соціально-економічні процеси, тобто статистичне спостереження.

Статистичне спостереження – планомірний, науково організований процес збирання даних щодо масових явищ та процесів, які відбуваються в економічній, соціальній та інших сферах життя України та її регіонів, шляхом їх реєстрації за спеціальною програмою, розробленою на основі статистичної методології.

Основним завданням статистичного спостереження є одержання об'єктивної, достовірної та повної інформації, яка характеризує кожную одиницю досліджуваної сукупності. Державні статистичні спостереження проводяться органами державної статистики відповідно до затвердженого Кабінету Міністрів України плану державних статистичних спостережень, або за окремими рішеннями КМ України.

Спостереження є фундаментом статистичного дослідження. В процесі спостереження формуються дані, які на наступних етапах підлягають обробці і аналізу. Від якості даних спостереження залежать результати подальшого статистичного дослідження. Тому вони мають відповідати певним вимогам.

Перша вимога – це вірогідність даних, тобто їх відповідність реальному стану.

Друга вимога – це повнота даних як за їх обсягом, так і по суті. Повнота забезпечується, по-перше, охопленням одиниць досліджуваної сукупності, по-друге, системним добором кількох взаємозалежних ознак явища, і по-третє, при вивченні явища у часі повнота передбачає одержання даних за максимально тривалі періоди.

Третя вимога – своєчасність даних. Інформація має дійти до користувача, перш ніж застаріє, інакше вона втрачає корисність.

Четверта вимога – порівнянність даних у часі або у просторі.

П'ятою вимогою є доступність даних. Доступ до статистичної інформації забезпечується шляхом: систематичної публікації її в друкованих виданнях; поширення її засобами масової інформації; безпосереднього її надання органам державної влади та органам місцевого самоврядування, іншим юридичним, а також фізичним особам.

У статистичній практиці багатьох країн, в тому числі і в Україні використовується дві основні форми організації спостереження – звітність та спеціально організовані спостереження.

Основним джерелом первинної інформації є статистична звітність, яка охоплює підприємства, організації та установи всіх форм власності, згідно з класифікатором видів економічної діяльності та кодом ЄДРПОУ (Єдиний державний реєстру підприємств та організацій України).

Статистичний звіт – це документ, який вміщує систему показників діяльності суб'єктів господарювання. Зміст звіту, форма та термін подання в органи державної статистики затверджуються Державною службою статистики України.

Крім загально державної статистичної звітності існує відомча звітність, яка розробляється міністерствами та відомствами для внутрішніх управлінських потреб.

Кожна форма статистичної звітності повинна мати такі реквізити:

- найменування форми звітності;
- номер та дата затвердження Державною службою статистики України.;

- на яку дату або за який період наводяться дані;
- юридична назва та адреса суб'єкта господарювання;
- відомча підлеглисть або форма власності;
- кому подається статистичний звіт (назва та адреса);
- підписи осіб, які відповідають за якість статистичної звітності.

Види статистичної звітності:

- типова, яка має єдину форму та зміст для всіх суб'єктів господарювання.
- спеціалізована, яка стосується окремих галузей економіки або виробництва.

За періодичність подання звітність поділяється на: щоденну, щотижневу, щомісячну, квартальну та річну. За способом подання статистична звітність буває: поштовою, телефонною, телеграфною та електронною.

Спеціально організовані статистичні спостереження проводяться згідно з планом державних спостережень для одержання первинних даних, які не охоплюються статистичною звітністю. До цих спостережень відносяться різноманітні переписи, одноразові обліки, опитування тощо.

2.2 Програмно-методологічні та організаційні питання спостереження

Статистичне спостереження проводиться відповідно до плану статистичного спостереження, що містить програмно-методологічні та організаційні питання. План статистичного спостереження включає широке коло питань методики та організації збору статистичної інформації, контролю її якості і вірогідності.

Розробка програмно-методологічних питань плану спостереження полягає в науково-практичному обґрунтуванні та визначенні суті явища, умов його формування та прояву. Крім того, добирається система ознак, що характеризують явище, ураховується можливість їх кількісної обробки та перевірки на точність.

Комплекс програмно-методологічних питань може бути поданий у послідовності їх появи та розв'язування (рис. 2.1).



Рисунок 2.1 – Програмно-методологічні питання статистичного спостереження

Мета спостереження – отримання статистичних даних, які є підставою для узагальнення характеристики стану та розвитку явища або процесу з визначенням відповідної закономірності.

Відповідно до мети визначають об'єкт і одиниці спостереження. **Об'єкт спостереження** – це сукупність явищ, що вивчаються в процесі спостереження. Для того щоб встановити об'єкт спостереження необхідно чітко визначити склад та межі статистичної сукупності. Для цього формулюються певні правила, а деяких випадках встановлюється ценз. Ценз – це обмежувальна ознака, якій повинні відповідати всі одиниці сукупності.

Одиниця спостереження – це та первинна ланка (елемент), від якої одержують необхідні відомості.

Одиниця сукупності – це первинний елемент об'єкта спостереження, ознаки якого підлягають реєстрації.

Важливим питанням плану статистичного спостереження є його програми. **Програма спостереження** – це перелік питань, на які намічають

дістати відповіді в процесі спостереження відносно кожної одиниці спостереження.

Програма статистичного спостереження оформлюється у вигляді формуляра, в якому містяться питання та відведені місця для реєстрації відповіді. У статистичній практиці використовуються два види формулярів – списковий та індивідуальний. Індивідуальний формуляр заповнюється для однієї одиниці сукупності, а списковий – для двох і більше. З метою забезпечення єдиного підходу до заповнення формулярів готується інструкція, у якій роз’яснюється суть запитань та наводяться варіанти відповідей.

Організаційні питання плану спостереження визначають місце і час його проведення, організаційну форму, вид та спосіб проведення, підготовчі роботи, порядок приймання матеріалів тощо (рис. 2.2).

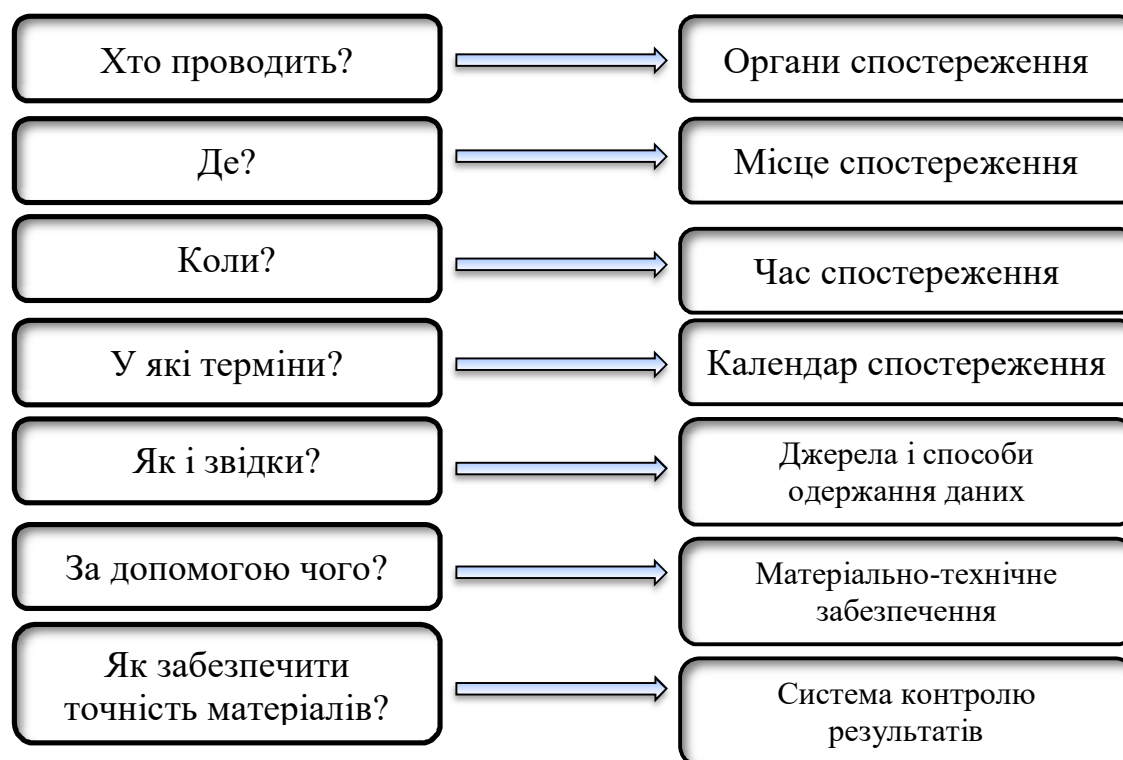


Рисунок 2.2 – Організаційні питання статистичного спостереження

Наступним питанням є обґрунтування місця спостереження. **Місцем спостереження** вважають пункт, де безпосередньо реєструються ознаки окремих одиниць сукупності. Здебільшого воно збігається з місцем знаходження одиниць спостереження.

Розрізняють об'єктивний та суб'єктивний час спостереження. **Об'єктивний час спостереження** – це той момент або період часу, станом на який реєструються ознаки. У тому випадку, коли об'єкт спостереження швидко змінюється і нараховує дуже велику кількість одиниць, встановлюється так званий критичний час. **Суб'єктивний час спостереження** – це період відведений на заповнення формулярів.

У плані статистичного спостереження вказуються органи державної статистики, які здійснюють його підготовку та проведення, а також несуть відповідальність за якість його результатів.

2.3 Форми, види та способи спостереження

З точки зору організації статистичного спостереження розрізняють три основні організаційні форми спостереження: звітність, спеціально організоване статистичне спостереження та реєстри (рис. 2.3).

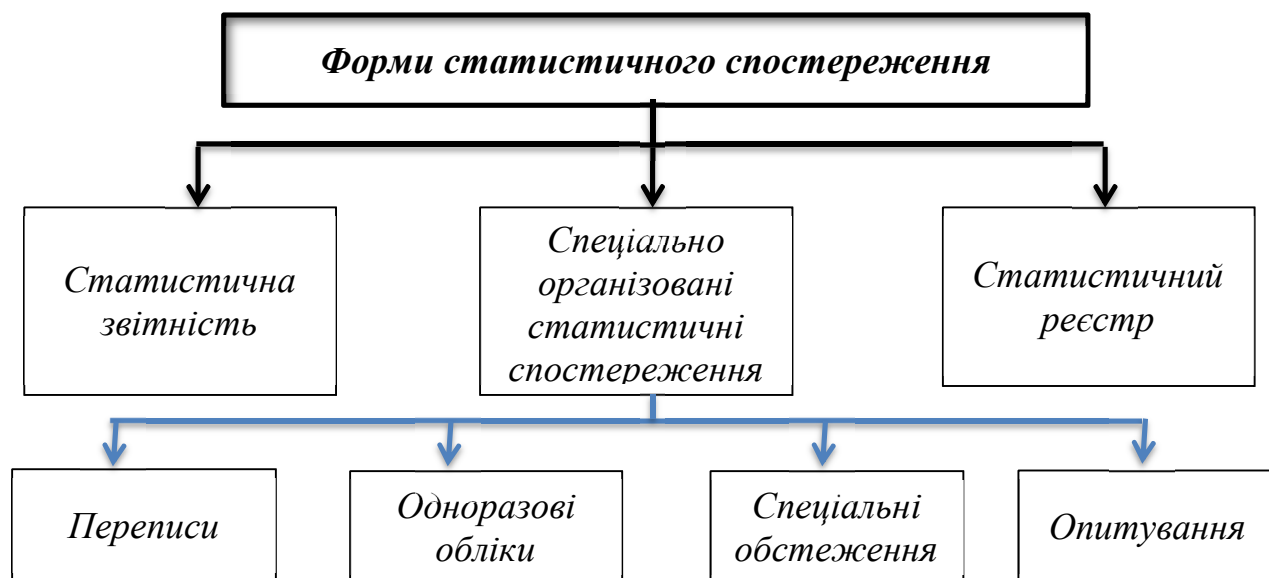


Рисунок 2.3 – Форми статистичного спостереження

Статистична звітність є основним джерелом інформації про соціально-економічні явища та процеси. Вона надається в органи статистики для

проведення державних статистичних спостережень у вигляді документів (звітів) спеціально затвердженої форми.

Спеціально організовані статистичні спостереження охоплюють ті сторони суспільного життя, які не відобразились у звітності. До них належать переписи, одноразові обліки, спеціальні обстеження, опитування. Переписи проводяться періодично або одноразово і дають повну характеристику масового явища станом на якусь дату або певний момент часу.

Обліки – суцільні спостереження масових даних, які ґрунтуються на даних огляду, опитування та документальних записів.

Спеціальні обстеження – переважно несучільне обстеження масових явищ згідно з певною тематикою, що виходить за межі звітності. Вони можуть бути періодичними або одноразовими.

Опитування – це, як правило, несучільне спостереження з метою вивчення думок, мотивів, оцінок, що реєструються зі слів респондентів. Третьою формою статистичного спостереження є статистичний реєстр – це список або перелік одиниць певного об'єкта спостереження із зазначенням необхідних ознак, який складається та оновлюється під час постійного відстежування.

У статистичній практиці використовують різноманітні види спостереження та способи його проведення. За охопленням сукупності спостереження поділяється на суцільні та несучільні.

Суцільне статистичне спостереження – спостереження щодо всіх без винятку одиниць сукупності, яка вивчається. Несучільне статистичне спостереження охоплює окремі одиниці сукупності відібрані відповідним чином.

Розрізняють такі види несучільного спостереження:

- монографічне, яке передбачає детальне та всебічне обстеження типових одиниць сукупності (наприклад, типові лікарні, школи, підприємства);
- основного масиву, коли ознаки реєструються тільки по найбільших одиницях сукупності (наприклад, найбільші підприємства, найбільші міста);

– вибіркоче, коли для обстеження одиниці відбираються випадковим чином;

– анкетне, оскільки заповненими повертаються не всі роздані анкети.

За моментом реєстрації даних спостереження поділяються на поточні, періодичні та одноразові.

Поточне статистичне спостереження полягає в реєстрації даних по мірі їх одержання (наприклад, реєстрація народжень, смертей, шлюбів, розлучень). Періодичне спостереження проводиться через певні рівні проміжки часу, а одноразові по мірі необхідності без будь-якої періодичності.

Існує три основних способи одержання первинної статистичної інформації: безпосереднє спостереження, документальне спостереження та опитування.

Безпосереднє спостереження передбачає, що реєстратори шляхом підрахунку, зважування, вимірювання тощо. встановлюють певний факт і здійснюють запис у формулярі. Документальне спостереження передбачає, що запис у формулярі здійснюється на підставі первинного документа. У свою чергу опитування поділяється на експедиційні, самореєстрацію, кореспондентське та інтерв'ю.

Експедиційне спостереження полягає у тому, що спеціально підготовлені реєстратори опитують респондентів і з їх слів здійснюють заповнення формуляру. Самореєстрація передбачає, що формуляр заповнює сам респондент. Кореспондентське опитування передбачає, що особи з якими органи державної статистики уклали відповідну угоду, надсилають у визначені терміни заповнені формуляри. Інтерв'ю передбачає, що відповіді на питання у формулярі записуються у довільній формі.

2.4 Помилки спостереження та способи їх виправлення

Проблема надання органам державного управління, окремим підприємствам та організаціям, науковцям та фахівцям надійної, достовірної та репрезентативної статистичної інформації завжди була і залишається центральною у теорії та практиці статистичної науки.

Точністю вважається міра відповідності даних спостереження дійсній їх величині, достовірністю – міра об'єктивного відображення суті явищ і процесів. Чим ближче значення показників, отриманих в результаті спостереження, їх фактичним значенням, тим вище точність статистичного спостереження.

Статистична інформація, яку одержують в процесі статистичного спостереження, підлягає перевірці та контролю. Перевіряється як повнота охоплення одиниць сукупності, так і повнота і правильність заповнення формулярів. Розбіжність між записами у формулярі та реальними значеннями ознак має назву помилки спостереження. Розрізняють помилки реєстрації та помилки репрезентативності.

Помилки реєстрації виникають в наслідок неправильного встановлення факту або помилкового запису у формулярі. Вони поділяються на випадкові та систематичні. Випадкові помилки виникають під впливом випадкових чинників і дають викривлення даних як більшу, так і в меншу сторони. Систематичні помилки призводять до викривлення інформації у певному напрямку. Вони бувають навмисними і ненавмисними.

Навмисні помилки виникають внаслідок свідомого викривлення фактів або даних. Їх часто називають приписками. Посадові особи, які допускають такого роду помилки притягаються до відповідальності. Ненавмисні помилки пов'язані з особливостями психіки людини, наприклад схильність до заокруглення цифр.

Помилки репрезентативності виникають лише при проведенні несучільних спостережень і зумовлюються тим, що обстежується тільки частина сукупності. Такі помилки є об'єктивними, їх можна кількісно оцінити та врахувати.

Помилки спостереження виявляються та усуваються після перевірки та контролю формулярів. Контроль починається із зовнішнього огляду формуляру на предмет повноти та якості його заповнення. Далі здійснюється логічний та арифметичний контроль.

Арифметичний контроль передбачає перевірку даних протягом всього процесу одержання і опрацювання інформації. Це перевірка підсумкових даних, перевірка правильності розрахунку середніх і відносних показників, балансового методу, погодження тих показників, які виводяться один з одного. Завданням лічильного контролю є виправлення підсумків і окремих кількісних показників.

Логічний контроль застосовують у взаємодії з арифметичним контролем у процесі опрацювання статистичної інформації. Він полягає в зіставленні відповідей на питання і з'ясуванні їх логічної сумісності. У процесі логічного контролю можуть бути виявлені нереальні або неправдоподібні відповіді.

Основними прийомами контролю статистичних даних є:

- співставлення відповідей на різні взаємопов'язані питання в документах;
- порівняння записів у документі, що перевіряється, з аналогічними даними в інших документах;
- співставлення звітних показників за суміжні періоди;
- застосування методу балансової ув'язки показників.

Наведені прийоми перевірки статистичних даних шляхом арифметичного і логічного контролю можуть бути використані при перевірці як матеріалів спеціальних обстежень, так і звітності. Якість контролю статистичних даних значною мірою залежить від стану первинного обліку і статистичної звітності, рівня підготовки тієї особи, що перевіряє, знання тих фактів, які висвітлені в документах.

ТЕМА 3 ЗВЕДЕННЯ ТА ГРУПУВАННЯ СТАТИСТИЧНИХ ДАНИХ

3.1 Суть і завдання статистичного зведення.

3.2 Групування, його суть, завдання та види.

3.3 Основні питання методології статистичних групувань та їх класифікація в практичній діяльності.

3.4 Статистичні ряди розподілу.

3.1 Суть і завдання статистичного зведення

В результаті статичного спостереження одержують первинну інформацію про кожен одиницю досліджуваної сукупності. Подальше завдання дослідження полягає у приведенні даних в певний порядок, їх систематизації та одержанні зведених, узагальнених характеристик сукупності в цілому. Цю роботу виконують на другому етапі статистичного дослідження, який називають статистичним зведенням.

Статистичне зведення – це процес упорядкування, систематизації і наукової обробки первинного статистичного матеріалу для виявлення типових рис і закономірностей явищ і процесів, що вивчаються. Зведення є основою для подальшого аналізу статистичної інформації. За зведеними даними розраховуються узагальнюючі показники, виконується порівняльний аналіз, а також аналіз причин групових відмінностей, вивчаються взаємозв'язки між ознаками.

Зведення – передбачає перехід від індивідуальних значень показників до характеристики сукупності в цілому або її окремих частин. Зведення у вузькому розумінні слова – це підрахунок підсумку, тобто сумування індивідуальних значень показника (так зване просте зведення).

Зведення у широкому розумінні – етап статистичного дослідження, на якому узагальнюються результати спостереження (так зване складне зведення).

Статистичне зведення проводиться згідно з розробленим планом та програмою. У плані зведення знаходять відображення його мета, форма організації та техніка проведення. Програма зведення – це перелік узагальнюючих показників, що характеризують сукупність, та перелік груп, на які вона розбивається. Програма зведення оформлюється у вигляді макетів статистичних таблиць, які мають бути заповнені в процесі зведення.

За організацією робіт і формою обробки даних розрізняють централізоване і децентралізоване зведення. При централізованому зведенні матеріали статистичного спостереження надсилають у центральний орган державної статистики, де їх обробляють за певною, заздалегідь встановленою програмою. При централізованому зведенні матеріали статистичного спостереження надходять у Державну службу статистики, де вони обробляються та систематизуються. При децентралізованому зведенні первинна інформація узагальнюється територіальними органами статистики.

У статистичній практиці виділяють два види зведення: просте і складне. Просте зведення полягає у підрахунку узагальнюючих показників шляхом сумування. Складне зведення передбачає групування одиниць сукупності з підрахунком сумарних значень як по кожній групі, так і по сукупності в цілому.

3.2 Групування, його суть, завдання та види

Одним із основних і найбільш поширених методів обробки і аналізу первинної статистичної інформації є групування. Статистичним групуванням називають розчленування, розподіл одиниць сукупності на класи, групи та підгрупи за суттєвими ознаками. Та ознака, що покладена в основу групування, тобто за якою утворюються групи, має назву групувальної.

У системі статистичних методів дослідження методу групувань займає дуже важливе місце. По-перше, він дозволяє дослідити склад та розподіл одиниць сукупності за певною ознакою. По-друге, на його основі базується

використання інших статистичних методів (середніх величин, дисперсійний аналіз тощо).

При використанні методу групувань необхідно дотримуватись трьох основних вимог:

1) всебічний логічний та економічний аналіз досліджуваних процесів та явищ;

2) визначення суттєвих ознак;

3) обґрунтоване визначення числа груп та інтервалів групування.

З допомогою групувань вирішують три типи завдань:

1) виділення соціально-економічних типів явищ та процесів;

2) вивчення складу та структури сукупності, структурних зрушень та закономірностей розподілу;

3) виявлення взаємозв'язку між явищами або показниками, що їх характеризують.

Кожному з названих завдань відповідає окремий вид групування: типологічне, структурне та аналітичне.

Групування, які дають можливість виділити із загального масиву інформації типові явища і процеси, називають *типологічними*. За допомогою типологічних групувань виділяють найхарактерніші групи, типи явищ, з яких складається неоднорідна статистична сукупність, визначають істотні відмінності між ними, а також ознаки, що є спільними для усіх груп, їх застосовують при вивченні розподілу підприємств за формами власності, при групуванні населення за суспільними групами, розподілі суспільного виробництва за економічним призначенням продукції тощо.

Структурні групування характеризують розподіл якісно однорідної сукупності на групи за певною ознакою. Потреба в таких групуваннях виникає тому, що однорідність явищ і елементів, з яких складається статистична сукупність, ще не означає їх тотожності. У межах однорідної сукупності елементи відрізняються один від одного числовими значеннями властивих їм ознак.

Структурні і типологічні групування описові, вони характеризують структуру сукупності, виділяють характерні риси та особливості, але відрізняються рівнем якісних відмінностей між групами.

Найбільш дієвим методичним інструментом для обробки економічної інформації є аналітичні групування, які застосовуються для вивчення взаємозв'язків між явищами, впливу однієї ознаки на іншу. Масштаби практичного застосування аналітичних групувань ставлять їх в число пріоритетних методів економічних досліджень.

Крім цього, групування класифікуються за видом групувальної ознаки, а також у залежності від їх числа та співвідношення.

За видом групувальної ознаки групування поділяються:

- на групування за атрибутивною ознакою;
- на групування за дискретною кількісною ознакою;
- на групування за інтервальною кількісною ознакою;
- на групування за альтернативною ознакою.

За числом групувальних ознак та співвідношенням між ними групування бувають:

- прості (групувальна ознака одна);
- багатомірні (ознак дві та більше, їх вагомість однакова);
- комбінаційні (ознак дві та більше, вони розглядаються у комбінації);
- ієрархічні (ознак дві та більше, вони мають стале співвідношення «старшості»). Наприклад, групування підприємств за галузями економіки, а кожної галузі – на підгалузі та окремі виробництва.

Використання того чи іншого виду групування обумовлюється метою дослідження, наявністю первинної інформації, можливістю використання комп'ютерної техніки тощо.

3.3 Основні питання методології статистичних групувань та їх класифікація в практичній діяльності

До основних методологічних питань групування належать:

- вибір груповальної ознаки;
- визначення числа груп;
- встановлення інтервалів групування.

Визначення груповальної ознаки (або ознак) здійснюється з врахуванням суті та законів розвитку явищ та процесів, що досліджуються. Головна вимога до груповальної ознаки – її суттєвість, тобто вагомість з точки зору мети дослідження. Особливі складності при виборі груповальних ознак виникають при топологічних групуваннях, які передбачають якісні відмінності між групами.

Визначення числа груп у багатьох випадках залежить від особливостей груповальної ознаки. В тому випадку, коли групування здійснюється за атрибутивною ознакою, число груп відповідає числу варіантів цієї ознаки (наприклад, групування населення за рівнем освіти нараховує 7 груп). Якщо груповальна ознака є альтернативною, то завжди утворюється тільки дві групи.

В тому випадку, коли груповальна ознака є кількісною дискретною з обмеженою кількістю варіантів, число груп співпадає з числом варіантів цієї ознаки. У разі групування за інтервальною ознакою кількість груп визначається шляхом логічного аналізу або шляхом випробувань. При цьому виникає потреба утворення інтервалів групування.

Інтервал групування – це проміжок між двома значеннями груповальної ознаки, в межах яких всі одиниці сукупності відносяться до даної групи. Відповідно менше та більше число мають назву нижньої та верхньої межі інтервалу групування, а різниця між ними називається величиною інтервалу.

Інтервал групування називають закритим, якщо у ньому є нижня і верхня межа. В тому випадку, коли одне число відсутнє, інтервал є відкритим зверху

або знизу. Наприклад, групування робітників за рівнем середньомісячної заробітної плати (у грн): до 450, 450–500, 500–600, понад 600.

Інтервали називаються рівними, якщо їх величина у всіх групах однакова. У протилежному випадку утворюються нерівні інтервали, котрі в свою чергу бувають зростаючими або спадаючими.

Інтервали, утворені за однією ознакою, але різні за величиною у різних сукупностях, мають назву спеціалізованих. Наприклад, групування міських поселень та сільських населених пунктів за чисельністю жителів.

Якщо виконується групування з рівними інтервалами, величина інтервалу визначається за формулою:

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

де X_{max} – найбільше значення групувальної ознаки;

X_{min} – найменше значення групувальної ознаки;

m – число груп.

Орієнтовно число груп можна визначити за формулою Стерджесса:

$$m = 1 + 3,332 \lg n. \quad (3.2)$$

Наведену формулу доцільно застосовувати для багаточисельних сукупностей з нормальним розподілом одиниць за групувальною ознакою. В деяких випадках необхідно забезпечити приблизно рівну кількість одиниць у кожній групі, але тоді групування може не мати рівні інтервали.

Іноді необхідно зробити вторинне групування. Вторинним групуванням називають групування, які виконуються не по первинних даних, а по вже згрупованих раніше, тобто по рядах розподілу. Вони здійснюються двома способами:

- збільшенням або зменшенням інтервалів групування;
- перегрупування за чисельністю або питомою вагою груп.

Розглянемо на прикладі перший спосіб вторинного групування. Маємо вихідний ряд розподілу лікарень за числом ліжок:

Кількість ліжок	Число лікарень
20–30	2
31–40	7
41–50	4
51–60	10
61–70	5
71–80	20
81–90	30
91–100	60
Разом	138

Необхідно утворити три групи лікарень з відповідним числом ліжок:

Кількість ліжок	Число лікарень
До 50	$2 + 7 + 4 = 13$
50–80	$10 + 5 + 20 = 35$
понад 80	$30 + 60 = 90$
Разом	138

Розглянемо другий спосіб вторинного групування, використовуючи наведений ряд розподілу лікарень. Необхідно утворити три групи, які нараховують по 46 лікарень у кожній, тобто по 33,3 %.

У першу групу мають увійти перші п'ять груп ($2 + 7 + 4 + 10 + 5 = 28$), а також частина шостої групи ($46 - 28 = 18$). Звідси верхня межа першого інтервалу дорівнює:

$$71 + \frac{(80-71) \cdot 18}{20} = 71 + 8,1 = 79 .$$

Таким чином, перший інтервал вторинного групування має вигляд:

До 79 включно 46 лікарень (33,3 %)

Тоді у другий інтервал входять 2 лікарні з шостої групи, 30 лікарень сьомої групи і частина лікарень з восьмої групи ($2 + 30 + 14 = 46$). Верхня межа другого інтервалу буде дорівнювати:

$$91 + \frac{(100-91) \cdot 14}{60} = 91 + 2,1 = 93 .$$

Звідси, другий інтервал має вигляд:

46 лікарень (33,3 %)

Третій інтервал отримуємо автоматично

94–100 46 (60–14) лікарень (33,3 %)

Отже, маємо нове групування (табл. 3.1).

Таблиця 3.1 – Зведені дані

Кількість ліжок	Число лікарень	Питома вага, %
До 79 включно	46	33,3
80–93	46	33,3
94–100	46	33,3
Разом	138	100

Таким чином, за допомогою вторинних групувань можна здійснювати перетворення груп, одержаних унаслідок первинного групування.

Зазвичай в економічній практиці застосовують нерівні інтервали, прогресивно зростаючі або спадаючі. При визначенні величини інтервалу і розподілі одиниць об'єкта спостереження за групами важливе значення має точне встановлення меж, які в більшості позначаються зазначенням значень ознак «від» і «до» для одиниць, включаються до цієї групи.

Отримання серединного значення інтервалу можна двома методами:

1. Цей показник можна розрахувати підсумовуванням верхніх і нижніх кордонів інтервалу і діленням суми навпіл (середньоарифметична проста).

2. Розрахунок здійснюють додатком до середнього значення другого інтервалу величини рівного інтервалу.

3.4 Статистичні ряди розподілу

Результати зведення і групування можна оформляти у вигляді рядів розподілу, статистичних таблиць і графіків.

Статистичний ряд розподілу – це ряд, який характеризує розподіл одиниць сукупності по групах за будь-якою ознакою, різновидності якої розташовані у певному порядку. Ряди розподілу складаються з двох елементів – варіантів і частот. *Варіанта* – окреме значення групувальної ознаки, а *частоти* – кількість елементів у групі з відповідним значенням ознаки. Частоти, які відповідають певній ознаці, можуть подаватись як в абсолютних значеннях, так і у відносних, виражених коефіцієнтом або відсотком (часткою). Накопичену частоту (частку) називають кумулятивною.

Залежно від статистичної природи групувальної ознаки (якісна чи кількісна) ряди розподілу поділяють на *атрибутивні* та *варіаційні*. Ряд розподілу, утворений за якісною (атрибутивною) ознакою, називається атрибутивним. Прикладом атрибутивного ряду розподілу може бути розподіл підприємств за формами власності.

Ряд розподілу сукупності за ознакою, що має кількісне вираження, називається варіаційним. Варіаційні ряди залежно від групувальної ознаки поділяють на *дискретні* і *інтервальні*. За дискретною ознакою, кількість значень якої обмежена, утворюється дискретний ряд розподілу. Прикладом такого ряду може бути розподіл студентів вузів за віком, успішністю та ін.

За дискретною ознакою, що варіює в широких межах, або за неперервною ознакою будують інтервальний ряд розподілу. При цьому варіанти об'єднуються в інтервали, а частоти (частки) відносяться не до окремого значення ознаки, як у дискретних рядах, а до всього інтервалу.

Ряди розподілу допомагають досліджувати структуру явищ. Вони мають самостійне значення при вивченні варіації групувальної ознаки.

Для унаочнення часто користуються різними способами графічного зображення варіаційних рядів. Графік дозволяє в найбільш простій і доступній

формі піддати аналізу (візуально) статистичний ряд розподілу. Для графічного зображення рядів розподілу використовують такі графіки як полігон, гістограма, кумулята, огіва, крива концентрації (Лоренца), показова крива, крива Парето, антимода.

Гістограма будується для інтервальних рядів розподілу. При цьому по осі X відкладаються інтервали групування, а по осі Y – абсолютні або відносні частоти. В тому випадку, коли виконується групування з рівними інтервалами, ширина стовпчиків однакова, а якщо інтервали групування нерівні – різні. Наведемо приклад побудови гістограми, використовуючи ряд розподілу сімей за рівнем місячного доходу (грн) (рис. 3.1).

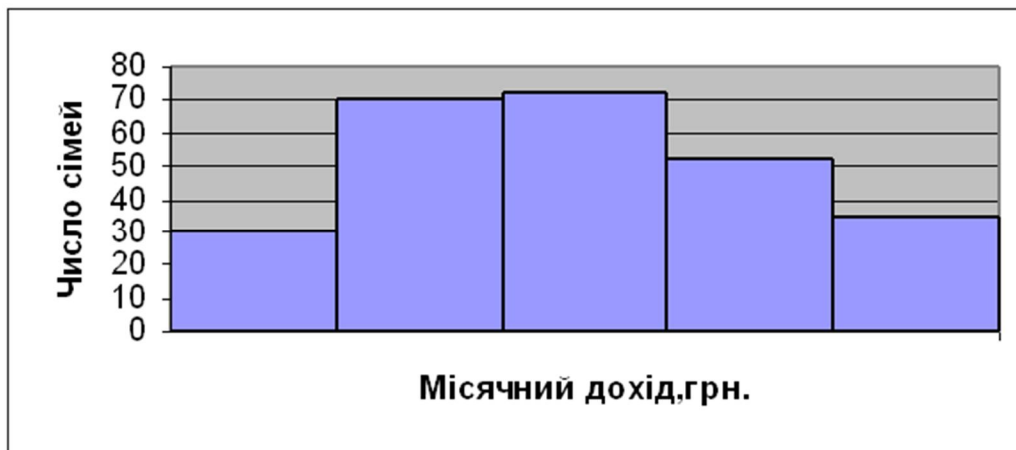


Рисунок 3.1 – Гістограма розподілу

Полігон використовується для графічного зображення дискретних та атрибутивних рядів розподілу. Це лінійний графік, при цьому по осі X відкладаються значення варіант, а по осі Y – частоти. Гістограму можна перетворити у полігон, з'єднавши відрізками прямої середини верхівок стовпчиків. Наведемо приклад побудови полігону, побудованого на основі дискретного ряду розподілу шкіл за кількістю класів (рис. 3.2).



Рисунок 3.2 – Полігон розподілу

Кумулята призначена для графічного подання рядів розподілу з нагромадженими частотами. Це може бути стовпчикова діаграма (для дискретного та атрибутивного рядів розподілу – лінійний графік). Будується вона аналогічно попереднім графікам, тільки по осі Y подаються нагромаджені частоти. Наведемо приклад побудови кумуляти (рис. 3.3).

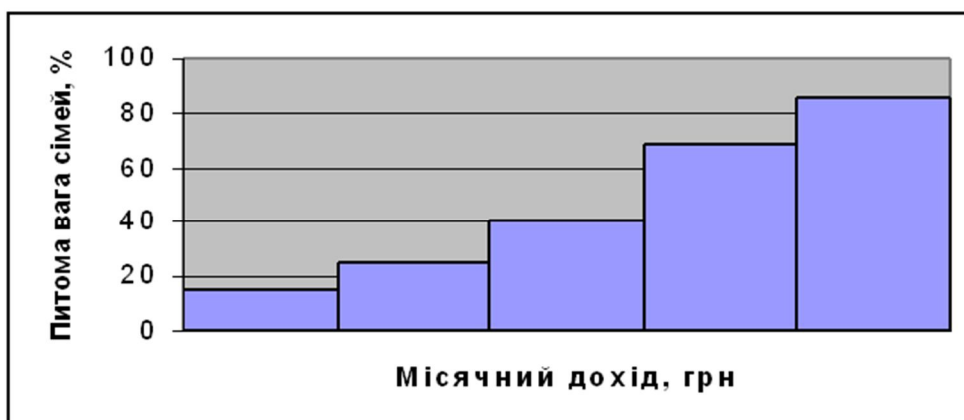


Рисунок 3.3 – Кумулята розподілу

Статистичні таблиці призначені для найбільш раціонального, наочного та систематизованого викладення результатів зведення та групування. У статистичній таблиці розрізняють підмет та присудок. Підметом таблиці є бічні заголовки (заголовки рядків), як правило, це перелік об'єктів або територій, які характеризуються числовими даними. Присудок таблиці – верхні заголовки (заголовки стовпців або граф), тобто це показники, що характеризують

статистичну сукупність. Перетин рядків та граф утворюють клітини таблиці, призначені для числових характеристик.

Таблиця обов'язково повинна мати загальний заголовок, який характеризує зміст таблиці, час та територію, до яких відносяться дані.

Правила побудови статистичної таблиці:

- таблиця повинна мати оптимальний розмір і стосуватися тільки досліджуваного явища або процесу;
- загальний та внутрішні заголовки повинні бути чіткі, короткі та змістовні;
- якщо рядків або стовпців (граф) багато, їх прийнято нумерувати;
- одиниці виміру показників обов'язково вказуються, у разі потреби для них виділяється графа або рядок;
- кількісні показники в межах однієї графи або рядка повинні наводитися з однаковою точністю, наприклад, до 0,1 тощо;
- якщо немає відомостей про розмір явища, у відповідній клітині проставляються крапки (...), відсутність або нульове значення позначають тире (–), якщо клітина не заповнюється проставляється знак «х», коли число значно менше від інших, записується 0,000.

ТЕМА 4 УЗАГАЛЬНЮЮЧІ СТАТИСТИЧНІ ПОКАЗНИКИ ТА ЗАГАЛЬНІ ПРИНЦИПИ ЇХ ЗАСТОСУВАННЯ

- 4.1 *Суть і види статистичних показників.*
- 4.2 *Абсолютні статистичні величини, одиниці їх вимірювання.*
- 4.3 *Характеристика відносних величин.*
- 4.4 *Середні величини, умови наукового їх застосування.*
- 4.5 *Середня арифметична, основні її властивості.*
- 4.6 *Середня гармонійна, її різновиди і сфера використання.*
- 4.7 *Структурні середні – мода і медіана.*

4.1 Суть і види статистичних показників

Статистична інформація, яку одержують у процесі статистичного дослідження, являє собою сукупність статистичних показників.

Статистичний показник – це одна з основних категорій статистики, яка характеризує суспільні явища та процеси у поєднанні кількісної та якісної визначеності.

Будь-який статистичний показник є поєднанням якісної та кількісної складових. Якісна складова показника обумовлюється сутністю та змістом явища і знаходить відображення у назві показника, наприклад, чисельність постійного населення, прибуток від операційної діяльності, ціна тощо. Крім цього, до якісної складової належить одиниця виміру, територія та час, до яких відноситься показник. Кількісна частина показника – це його чисельне значення у відповідності з якісним змістом. Наприклад, чисельність постійного населення Харківської області на 1.01.2019. – 2 млн 667 тис. осіб.

Статистичні показники повинні мати такі властивості:

1. Адекватність – спроможність характеризувати ті властивості, що досліджуються.

2. Достовірність – відповідність реальному стану речей.

3. Точність вимірювання – відповідність змісту показника, організації спостереження та обробки даних діючим вимогам.

Класифікація показників:

– за способом одержання показники поділяються на *первинні* та *вторинні* (похідні). Перші одержують шляхом статистичного спостереження, зведення та групування, а другі розраховуються на основі перших;

– за ознакою часу показники поділяються на *моментні* та *інтервальні*. Моментні показники дають кількісну характеристику явища на певний момент часу, а інтервальні – за період часу (доба, тиждень, декада, місяць, квартал, рік). Наприклад, чисельність студентів на 1.09.2019. – моментний показник, а число відрахованих студентів протягом навчального року – інтервальний;

– за можливістю узагальнення (сумування) показники поділяються на *адитивні* та *неадитивні*. Значення адитивних показників можна додавати, оскільки їх сума має економічний зміст. Неадитивні показники не можна додавати, оскільки така дія не має економічного змісту.

Серед статистичних показників існують пари взаємнообернених показників, які паралельно характеризують одне й те саме явище. Прямий показник X зростає з підсиленням явища, обернений $1/X$, навпаки, зменшується. Прикладом можуть бути такі показники:

1) купівельна спроможність грошової одиниці – прямий показник, ціна одиниці товару – обернений;

2) продуктивність праці за одиницю часу – прямий показник, трудомісткість одиниці продукції – обернений.

4.2 Абсолютні статистичні величини, одиниці їх вимірювання

У процесі статистичного спостереження одержують дані про значення тих чи інших ознак, що характеризують кожну одиницю досліджуваної сукупності. Для характеристики сукупності у цілому або окремих її частин дані про окремі одиниці сукупності систематизують з метою одержання узагальнюючих показників, в яких відображуються результати дослідження суспільних явищ.

Узагальнюючі показники можуть бути абсолютними, відносними і середніми величинами. Кожен вид показника має певне значення і посідає визначене місце у процесі пізнання реальної дійсності.

Абсолютні статистичні величини безпосередньо пов'язані з фізичними властивостями та соціально-економічною суттю явищ, які вивчають. **Абсолютні статистичні величини** – це кількісні показники, які характеризують розміри (рівні, обсяги) суспільних явищ у певних умовах місця і часу. Отримують їх методами статистичного спостереження і зведення

вихідної інформації. Розміри суспільних явищ можуть бути виражені у вигляді кількості одиниць, або у вигляді величини ознаки

Абсолютні статистичні величини широко використовуються у статистичному обліку та аналізі, управлінні, плануванні, менеджменті, маркетингу тощо. Вони поділяються на *індивідуальні, групові та загальні* (сумарні). Індивідуальна абсолютна величини відносяться до окремої одиниці сукупності. Групова та загальна відповідно характеризує окрему групу одиниць або всю сукупність разом, їх одержують шляхом сумування індивідуальних величин. Загальна абсолютна величина має назву «обсяг ознаки».

Абсолютні величини – це іменовані числа. Вони виражаються у натуральних, трудових та вартісних вимірниках.

Натуральними або фізичними називають такі одиниці виміру, які відтворюють фізичні властивості одиниць сукупності (вагу, площу, довжину, потужність тощо). Складна одиниця виміру утворюється з допомогою двох простих, наприклад, тоно-кілометр, кВт-год. У разі потреби використовуються умовно-натуральні одиниці виміру. Тоді рівень ознаки обчислюється у одиницях певного еталону. Перехід від натуральних до умовно-натуральних одиниць виміру здійснюється за допомогою перехідних коефіцієнтів (табл. 4.1).

Таблиця 4.1 – Перехід від натуральних до умовно-натуральних одиниць виміру

Потужність двигуна, кВт	Кількість, шт.	Перехідний коефіцієнт	Кількість в умовно-натуральних одиницях
10	100	1,0	100
15	50	1,5	75
20	200	2,0	400
Разом	350	x	575

Якщо перший вид двигунів прийнято за еталон, то перехідні коефіцієнти визначаються в такий спосіб:

$$K_1 = \frac{10}{10} = 1,0;$$

$$K_2 = \frac{15}{10} = 1,5;$$

$$K_3 = \frac{20}{10} = 2,0.$$

Кількість двигунів в умовно-натуральних одиницях становить:

$$H_1 = 100 \cdot 1,0 = 100; H_2 = 50 \cdot 1,5 = 75; H_3 = 200 \cdot 2,0 = 400.$$

Отже, загальна кількість двигунів дорівнює 350 шт., або 575 у. о. в перерахунку на двигуни потужністю 10 кВт.

Трудові одиниці виміру (людино-день та людино-година) використовуються для характеристики робочого часу або витрат праці. Вартісні вимірники характеризують розмір явища у грошовому виразі. Вони є найбільш універсальними, але їх головним недоліком є залежність від зміни цін та курсів валют.

4.3 Характеристика відносних величин

Для того, щоб повною мірою проаналізувати досліджувані соціально-економічні явища і процеси, виявити взаємозв'язки і закономірності, зробити правильні висновки, недостатньо лише абсолютних величин. Окремо абсолютні величини не дають належної уяви про досліджувані явища і процеси. В аналізі статистичної інформації поруч з абсолютними величинами важливе місце посідають відносні величини.

Відносними статистичними величинами називають показники, які виражають кількісні співвідношення між явищами суспільного життя.

Будь-який відносний показник одержують в результаті співставлення двох величин. Результатом порівняння є відносна величина, яка характеризує міру кількісного співвідношення різнойменних чи однойменних показників.

Відносна величина показує у скільки разів одна величина більша або менша за іншу, або яку частину займає одна величина по відношенню до іншої, або скільки одиниць однієї сукупності приходить на одиницю іншої сукупності.

Кожна відносна величина становить частку від ділення двох величин:

$$ВВ = \frac{А}{Б}. \quad (4.1)$$

Показник, що знаходиться у чисельнику (А), називається порівнюваною величиною або звітною, а той, що стоїть у знаменнику (Б) – базою порівняння або базисною величиною.

Залежно від бази порівняння відносні величини можуть виражатись у формі:

- 1) коефіцієнтів – якщо база порівняння приймається за одиницю;
- 2) процентів (%) – якщо база порівняння береться за 100;
- 3) проміле (‰) – якщо за базу порівняння взято 1 000;
- 4) продециміле (‱) – якщо база порівняння становить 10 000;
- 5) просантиміле (‱‱) – якщо база порівняння прийнята за 100 000.

Різноманітність співвідношень і пропорцій реального життя для свого відображення потребує різних за змістом і статистичною природою відносних величин. В залежності від змісту і пізнавального значення відносні величини, що використовуються в статистиці, поділяються на наступні основні види: відносні величини планового завдання, виконання плану, динаміки, структури, координації, порівняння і інтенсивності.

Відносна величина планового завдання ($ВВ_{ПЗ}$) – показує на скільки відсотків передбачається змінити рівень показника у звітному періоді порівняно з базисним. Розраховується у формі коефіцієнта чи відсотка:

$$ВВ_{ПЗ} = \frac{П_1}{Ф_0} \quad (4.2)$$

або

$$ВВ_{ПЗ} = \frac{\Pi_1}{\Phi_0} \cdot 100,$$

де Π_1 – плановий рівень показника у звітному періоді;

Φ_0 – фактичний рівень показника у базисному періоді.

Відносна величина виконання плану ($ВВ_{ВП}$) – показує на скільки відсотків фактично перевиконано або недовиконано планове завдання у звітному періоді. Тобто це процентне відношення фактично досягнутого рівня до запланованого за відповідний період часу (місяць, квартал тощо). Розраховується у формі коефіцієнта чи відсотка:

$$ВВ_{ВП} = \frac{\Phi_1}{\Pi_1} \quad (4.3)$$

або

$$ВВ_{ВП} = \frac{\Phi_1}{\Pi_1} \cdot 100,$$

де Φ_1 – фактичний рівень показника у звітному періоді

У тому випадку, коли планується не абсолютне значення показника, а його збільшення або зменшення у процентах, відносна величина виконання плану визначається як відношення фактичної зміни у процентах до планової. Наприклад, по підприємству планувалося у серпні збільшити обсяг реалізації продукції на 4 %, а фактично він зріс лише на 0,7 %. Тоді

$$ВВ_{ВП} = \frac{100+0,7}{100+4,0} \cdot 100 = 96,8 \%$$

Отже, план по реалізації продукції у серпні недовиконаний на 3,2 % ($96,8 - 100 = -3,2$).

Відносна величина динаміки характеризує зміну показника у часі і визначається як відношення значення у наступному періоді до величини у попередньому періоді:

$$ВВ_{д} = \frac{\Phi_1}{\Phi_0} \quad (4.4)$$

або

$$ВВ_{д} = \frac{\Phi_1}{\Phi_0} \cdot 100.$$

Наприклад, якщо у III кварталі фактичний прибуток становив 12 170 грн, а у II кварталі – 13 450 грн відносна величина динаміки становить:

$$ВВ_{д} = \frac{12\,170}{13\,450} = 0,905$$

або 90,5 %.

Отже, у III кварталі порівняно з II обсяг прибутку зменшився на 9,5 % (90,5 – 100 = –9,5).

Між названими видами відносних величин, якщо вони підраховані за одними даними, існує взаємозв'язок:

$$ВВ_{д} = ВВ_{ПЗ} \cdot ВВ_{ВП} = \frac{\Pi_1}{\Phi_0} \cdot \frac{\Phi_1}{\Pi_1} = \frac{\Phi_1}{\Phi_0}.$$

Звідси:

$$ВВ_{ПЗ} = \frac{ВВ_{д}}{ВВ_{ВП}} \quad (4.5)$$

або

$$ВВ_{ВП} = \frac{ВВ_{д}}{ВВ_{ПЗ}}.$$

Відносна величина структури ($ВВ_C$) характеризує співвідношення частини та цілого. Вона показує, яку частину або скільки процентів становить частина від загального підсумку. Якщо ця відносна величина визначається у вигляді коефіцієнту, вона називається часткою, а якщо у процентах – питомою вагою.

$$ВВ_C = \frac{\text{Частина}}{\text{Ціле (Сума частин)}} \cdot \quad (4.6)$$

Наприклад, в області за рік народилося 8 216 хлопчиків та 8 203 дівчинки.
Тоді

$$ВВ_C = \frac{8\,416}{8\,416 + 8\,203} = \frac{8\,416}{16\,619} = 0,506 \text{ або } 50,6 \% ;$$

$$ВВ_C = \frac{8\,203}{16\,619} = 0,494 \text{ або } 49,4 \% .$$

Варто мати на увазі, що сума відносних величин структури становить 1 ціле або 100 %.

Відносна величина координації показує співвідношення між окремими частинами одного цілого, при цьому одна частина приймається за базу порівняння. Вона може визначатися на 100, 1 000 або 10 000 одиниць знаменника.

$$ВВ_K = \frac{\text{Частина1}}{\text{Частина2}} \cdot 100(1\,000, 10\,000). \quad (4.7)$$

Наприклад, у демографії співвідношення між статями для новонароджених характеризується числом хлопчиків на 100 дівчаток:

$$ВВ_K = \frac{8\,416}{8\,203} \cdot 100 = 103 \text{ (осіб).}$$

Відносна величина порівняння ($ВВ_{\Pi}$) – це співвідношення однойменних показників, що обчислені по різних об'єктах або територіях за однаковий час:

$$ВВ_{\Pi} = \frac{\text{Показник по А}}{\text{Показник по Б}}. \quad (4.8)$$

Наприклад, на 1.01.2019 чисельність населення у м. Тернополі становила 1 млн 419 тис. осіб, а у м. Києві – 2 млн 960 тис. осіб. Звідси

$$ВВ_{\Pi} = \frac{2 \text{ млн } 960 \text{ тис. осіб}}{1 \text{ млн } 419 \text{ тис. осіб}} = 2,08 .$$

Отже, у м. Києві проживало у 2,08 рази більше населення, ніж у м. Харкові.

Відносна величина інтенсивності визначається як відношення двох різних показників і переважно характеризує ступінь поширення чи розвитку явища у певному середовищі. Ця відносна величина може мати одиниці виміру вихідних показників.

$$ВВ_{\text{інт}} = \frac{\text{Показник 1}}{\text{Показник 2}}. \quad (4.9)$$

Форма виразу відносної величини інтенсивності визначається суттю досліджуваних процесів та економічною логікою. Вони широко застосовуються для аналізу економічного розвитку, демографічних, соціальних та політичних процесів. Прикладами відносних величин інтенсивності можуть бути такі показники: густота населення, виробництво продукції на душу населення, випуск продукції у розрахунку на одиницю основних засобів (фондовіддача).

Відносна величина інтенсивності показують, скільки одиниць однієї сукупності припадає на одиницю іншої сукупності. Наприклад, за даними статистичного щорічника України у звітному році валовий внутрішній продукт

у фактичних цінах становив 1 454 931 млн грн, середня чисельність наявного населення – 45,4 млн осіб, територія – 603,7 тис. км². За цими даними можна обчислити такі відносні показники інтенсивності:

1) ВВП у розрахунку на душу населення
 $1\,454\,931 : 45,4 = 32046,9$ грн/осіб;

2) щільність населення $45,4 : 603,7 = 75$ осіб/км².

Ефективність використання статистичних показників залежить від специфіки і умов розвитку суспільних явищ і процесів, а також від комплексного застосування абсолютних і відносних величин у статистичному дослідженні.

4.4 Середні величини, умови наукового їх застосування

У статистиці середня величина є найрозповсюдженішою формою узагальнюючого показника, оскільки він дає кількісну характеристику масових соціально-економічних явищ та процесів.

Середня величина – це узагальнена характеристика однорідної сукупності за варіюючою ознакою, що показує типовий рівень цієї ознаки у одиниці сукупності.

Характерний, типовий рівень ознаки формується під впливом так званих систематичних (невипадкових, постійних) факторів, а відхилення індивідуальних значень від типового рівня зумовлені дією випадкових факторів, котрі впливають по-різному на окремі одиниці сукупності. Таким чином, середня величина відображає те спільне, загальне, що є характерним для усіх одиниць досліджуваної сукупності.

З допомогою середніх величин вирішуються наступні завдання статистичного дослідження:

- характеристика досягнутого рівня розвитку явища або процесу;
- порівняння показників, обчислених по різних сукупностях;
- характеристика розвитку (варіації) явища у часі та просторі;

– вивчення взаємозв'язку між показниками.

При визначенні середньої величини необхідно дотримуватись двох головних вимог:

1) сукупність повинна бути якісно однорідною;

2) достатньо велика кількість одиниць у сукупності, тобто наявність масових даних.

Середня відображає типовий рівень ознаки лише в тому випадку, коли статистична сукупність, за якою вона обчислюється, якісно однорідна. Це одна з основних умов наукового застосування середніх у статистиці. Саме тому застосування методу середніх в статистиці пов'язують з методом групування. Крім того, типовий рівень ознаки, що вивчається, проявляє себе лише у випадку узагальнення масових фактів. У разі узагальнення масових фактів випадкові відхилення індивідуальних величин від загальної тенденції взаємно погашаються в середній величині. Ця вимога стосовно обчислення середніх величин пов'язує метод середніх із законом великих чисел.

Бажано, щоб при визначенні середньої величини враховувалися значення показника по усіх одиницях сукупності.

Середні величини поділяються на загальні та групові. Загальна середня величина характеризує сукупність в цілому, а групова – окрему групу одиниць. Якщо сукупність складається з якісно різнорідних груп, загальна середня величина не буде типовою характеристикою сукупності, тому обов'язково необхідно визначати групові середні величини. Наприклад, при вивченні витрат на підготовку фахівців необхідно враховувати наявність різних форм навчання: денної, заочної та вечірньої.

За методикою розрахунку всі середні величини, які використовуються у статистиці, відносяться до класу степеневих середніх, формула якої в загальному має вигляд:

$$\bar{x} = \sqrt[m]{\frac{\sum x^m}{n}}, \quad (4.10)$$

де x – індивідуальні значення ознаки;

n – чисельність сукупності (кількість одиниць).

Залежно від значення показника степеня m розрізняють такі середні величини:

$m = -1$ – середня гармонійна;

$m = 0$ – середня геометрична;

$m = 1$ – середня арифметична;

$m = 2$ – середня квадратична.

При використанні середніх величин основним питанням є вибір виду середньої, при цьому визначальним фактором є відповідність обчислень економічній суті осереднюваного показника та характеру вихідних даних.

Зазначена формула для ступеневої середньої дозволяє визначити ступеневу середню просту, у випадку, коли значення X_i зустрічаються один раз, або коли значення X не упорядковані. У випадках, коли значення X_i повторюються, для визначення середньої ступеневої застосовується формула середньої ступеневої зваженої:

$$Z_{зв} = Z \sqrt{\frac{\sum(X^z_i \cdot f_i)}{\sum f_i}}, \quad (4.11)$$

де f_i – частоти, з якими зустрічаються ті або інші значення X_i .

Ці частоти називаються статистичними вагами, тому середню називають зваженою.

Обмежене використання у статистиці знаходять середня квадратична та середня геометрична величини. Середня квадратична (проста і зважена) обчислюються за формулами:

$$\bar{x} = \sqrt{\frac{\sum x^2}{n}} \quad \text{та} \quad \bar{x} = \sqrt{\frac{\sum x^2 f}{nf}}.$$

Вона використовується при розрахунках показників варіації (середнього квадратичного відхилення) у модифікованому вигляді.

Середня геометрична величина застосовується тоді, коли обсяг ознаки дорівнює не сумі, а добутку варіант. Її формула має вигляд:

$$\bar{x} \equiv \sqrt[n]{x_1 \cdot x_2 \cdot x_3 \cdot \dots \cdot x_n} . \quad (4.12)$$

На практиці найчастіше використовують середню арифметичну та середню гармонійну.

4.5 Середня арифметична, основні її властивості

Одним із найпоширеніших видів середніх величин є середня арифметична. Її застосовують в тих випадках, коли обсяг варіаційної ознаки для всієї сукупності формується як сума значень ознаки окремих одиниць досліджуваної сукупності.

Середня арифметична може бути простою і зваженою. Розглянемо розрахунок середньої арифметичної простої на наступному прикладі.

Припустимо, що треба обчислити середній рівень кваліфікації робітників підприємства, тарифний розряд яких становить: 6, 3, 4, 3, 5, 2, 4, 5, 4, 4. Названі числа – це індивідуальні значення ознаки або варіанти. Для обчислення середнього тарифного розряду треба суму всіх значень ознаки (суму розрядів), тобто обсяг ознаки, розділити на кількість одиниць сукупності (кількість робітників):

$$\bar{X} = (6 + 3 + 4 + 3 + 5 + 2 + 4 + 5 + 4 + 4) : 10 = 4 \text{ розряд.}$$

Позначивши варіанти $x_1, x_2, x_3, \dots, x_n$, цей розрахунок можна записати так:

$$\bar{X}_a = \frac{\sum x_i}{n}, \quad (4.13)$$

де X_i – індивідуальні значення ознаки;

n – кількість значень ознаки.

Наведена формула має назву середньої арифметичної простої і застосовується тоді, коли розрахунок здійснюють на основі первинних, не згрупованих даних. Проте в практиці аналітичної роботи нерідко виникає потреба розрахувати середні величини на основі згрупованих даних, передусім даних варіаційного ряду розподілу.

Середня арифметична зважена визначається по формулою:

$$\bar{X}_{\text{азв}} = \frac{\sum(X_i f_i)}{\sum f_i}. \quad (4.14)$$

Наприклад, відомий дискретний ряд розподілу пацієнтів за терміном їх госпіталізації у днях (табл. 4.2).

Таблиця 4.2 – Дані дискретного ряду розподілу пацієнтів

Число днів госпіталізації (x)	Число пацієнтів (f)	x·f
8	2	16
9	5	45
10	9	90
11	12	132
12	10	120
13	11	143
14	8	112
15	5	75
16	1	16
19	1	19
Разом	64	768

$$\bar{x} = \frac{\sum xf}{\sum f} = \frac{768}{64} = 12 \text{ днів}.$$

У багатьох випадках вихідні дані для визначення середньої арифметичної являють собою інтервальний ряд розподілу. Тоді спочатку інтервальний ряд розподілу перетворюється у дискретний шляхом знаходження середини кожного інтервалу, а далі розрахунок здійснюється як у попередньому випадку за формулою середньої арифметичної зваженої.

Наприклад, відомий ряд розподілу за розміром штрафу (табл. 4.3).

Таблиця 4.3 – Дані ряду розподілу за штрафом

Розмір штрафу, грн.	Число штрафів (f)	Середина інтервалу (x)	x·f
До 100	4	50	200
100–200	20	150	3 000
200–400	26	300	7 800
400–600	15	500	7 500
600–800	8	700	5 600
800–1 000	3	900	2 700
1 000–2 000	2	1 500	3 000
2 000–3 000	2	2 500	5 000
Разом	80	x	34 800

Середній розмір штрафу:

$$\bar{x} = \frac{\sum xf}{\sum f} = \frac{34\,800}{80} = 435 \text{ грн.}$$

Якщо вихідні дані являють собою результат групування і відомі середні значення показника по кожній групі (групові середні), то розрахунок загальної середньої здійснюється виключно за формулою середньої арифметичної зваженої:

$$\bar{x} = \frac{\sum \bar{x}_i f_i}{\sum f_i}, \quad (4.15)$$

де \bar{X}_i – групові середні величини;

f_i – число одиниць у і-тій групі.

Наприклад, групування вкладників за розміром вкладу наведено у таблиці 4.4.

Групи за розміром вкладу	Середній розмір вкладу, грн	Число вкладників, осіб	$\bar{X}_i \cdot f_i$
Невеликий	2 300	2 130	4 899 000
Середній	5 700	650	3 705 000
Великий	14 200	97	1 377 400
Разом	x	2 877	9 981 400

Загальна середня дорівнює:

$$\bar{x} = \frac{\sum \bar{x}_i f_i}{\sum f_i} = \frac{9981400}{2877} = 3470 \text{ грн.}$$

Середня арифметична величина має ряд властивостей, що використовуються при обчисленнях:

1. При збільшенні або зменшенні кожної частоти в k разів, середня не зміниться:

$$\bar{x} = \frac{\sum xf}{\sum f} = \frac{\sum x \frac{f}{k}}{\sum \frac{f}{k}}. \quad (4.16)$$

2. При збільшенні або зменшенні кожної варіанти в k разів середня зміниться в стільки ж разів:

$$k \cdot \bar{x} = \frac{\sum k \cdot x \cdot f}{\sum f} \quad (4.17)$$

або

$$\frac{\bar{x}}{k} = \frac{\sum \frac{x}{k} \cdot f}{\sum f}.$$

3. При збільшенні або зменшенні кожної варіанти та сталої величини A , середня зміниться на цю ж величину:

$$\bar{x} - A = \frac{\sum (x-A)f}{\sum f} \quad \text{або} \quad \bar{x} + A = \frac{\sum (x+A)f}{\sum f}. \quad (4.18)$$

4. Сума відхилень значень ознаки (варіант) від середньої арифметичної дорівнює нулю:

$$\sum(x - \bar{x}) = 0 \quad \text{або} \quad \sum(x - \bar{x})f = 0 . \quad (4.19)$$

5. Середня арифметична, що помножена на чисельність сукупності, дорівнює обсягу ознаки:

$$\bar{x} \cdot n = \sum x \quad \text{або} \quad \bar{x} \cdot \sum f = \sum xf . \quad (4.20)$$

6. Сума квадратів відхилень варіант від середньої арифметичної є мінімальною величиною із всіх можливих:

$$\sum(x - \bar{x})^2 = \min \quad \text{або} \quad \sum(x - \bar{x})^2 f = \min . \quad (4.21)$$

Властивість 1 свідчить про те, що середню арифметичну можна визначити як за абсолютними, так і за відносними частотами.

Властивості 2 та 3 використовуються для спрощення підрахунків середньої арифметичної зваженої в інтервальних рядах розподілу (метод «моментів»). Середнє значення при використанні цього методу визначається за формулою:

$$\bar{x} = m_i \cdot i + A , \quad m_i = \frac{\sum\left(\frac{x-A}{i}\right)f}{\sum f} , \quad (4.22)$$

де: m_1 – момент першого порядку;

i – величина інтервалу;

A – середина інтервалу з найбільшою частотою.

4.6 Середня гармонійна, її різновиди і сфера використання

При розрахунку середньої з обернених показників використовують середню гармонічну. Середня гармонічна – це величина, обернена середній арифметичній з обернених значень ознаки. Буває простою і зваженою.

Припустимо, що придбано товару у двох продавців на одну й ту саму суму – на 10 грн, але за різною ціною: по 30 грн за 1 кг у першого продавця і по 20 грн – у другого. Як визначити середню ціну покупки? Середня арифметична $(30 + 20) : 2 = 25$ грн за 1 кг нереальна, оскільки за такою ціною на 20 грн можна придбати $20 : 25 = 0,8$ кг товару. Насправді придбано товару у першого продавця $(10 : 30) = 0,33$ кг, у другого – $(10 : 20) = 0,50$ кг, тобто разом $0,33 + 0,50 = 0,83$ кг, а середня ціна становить $20 : 0,83 = 24$ грн. Описаний порядок розрахунку називають середньою гармонічною простою.

У випадку, коли статистична інформація не містить відомостей про частоти варіантів сукупності й представлена у вигляді добутків цих варіантів на їхню частоту $M_i = X_i \cdot f_i$, застосовується середня гармонійна величина – проста й зважена. Середня гармонійна зважена величина визначається за формулою:

$$\overline{X}_{\text{гар.зв}} = \frac{\sum M_i}{\sum \frac{M_i}{X_i}}, \quad (4.23)$$

де $M_i = X_i \cdot f_i$;

$$X_i \neq 0, \text{ а } f_i = \frac{M_i}{X_i}.$$

Підставивши значення $M_i = X_i \cdot f_i$ та $f_i = \frac{M_i}{X_i}$ у рівняння середньої гармонійної зваженої, отримаємо:

$$\overline{X}_{\text{гар.зв}} = \frac{\sum X_i \cdot f_i}{\sum f_i} = \overline{X}_{\text{а.зв}}, \quad (4.24)$$

У такий спосіб $\overline{X}_{\text{гар.зв}}$ отримана з формули для $\overline{X}_{\text{а.зв}}$

Середня гармонійна проста може бути визначена з формули $\overline{X}_{\text{гар.зв}}$

$$\bar{X}_{\text{гар.зв}} = \frac{\sum M_i}{\sum \frac{M_i}{X_i}} = \frac{n \cdot M_i}{M_i \cdot \sum \frac{1}{X_i}} = \frac{n}{\sum \frac{1}{X_i}} = \bar{X}_{\text{гар}}, \quad (4.25)$$

де n – число індивідуальних значень ознаки.

Приклад

За даними, наведеним у таблиці 4.5, розрахувати середній відсоток і середній індекс виконання плану.

Таблиця 4.5 – Дані підприємств для розрахунку

Підприємство	Зроблено продукції, M_i - одиниць	Виконання плану	
		X_i , %	Індекс виконання плану X_i'
1	50	101	1,01
2	70	98	0,98
3	30	102	1,02
Разом:		150	

Середній відсоток виконання плану дорівнює:

$$\bar{X}_{\text{гар.в}} = \frac{\sum M_i}{\sum \frac{M_i}{X_i}} = \frac{50 + 70 + 30}{50/101 + 70/98 + 30/102} = 99,8 \% .$$

Середній індекс виконання плану розрахувати аналогічно.

$$\bar{X}'_{\text{гар.в}} = \frac{50 + 70 + 30}{50/1,01 + 70/0,98 + 30/1,02} = 0,998 .$$

4.7 Структурні середні – мода і медіана

Для характеристики розподілу одиниць сукупності за певною ознакою використовується так звані порядкові або структурні середні – мода і медіана.

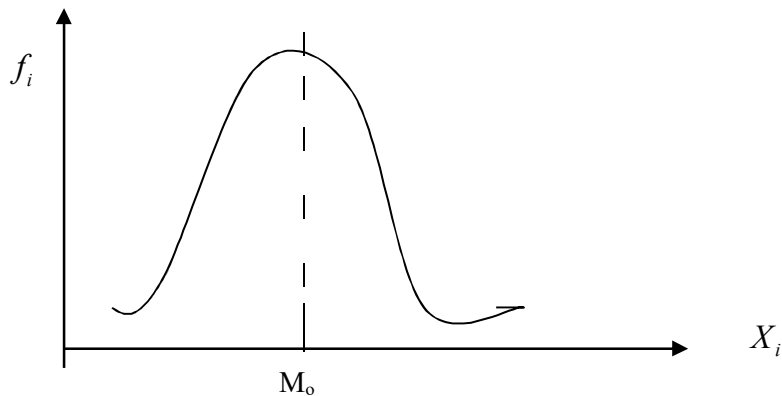
Під модою розуміють значення ознаки, що найбільш часто зустрічається

в ряді розподілу. Якщо ряд розподілу дискретний – мода, це варіант X_i із найбільшою частотою.

Наприклад:

X_i	3	5	7	$M_0 = 7$
f_i	14	10	27	

Якщо ми маємо графічне зображення ряду, то модою буде значення ознаки, що відповідає максимальній точці кривій розподілу.



Для інтервальних рядів з рівними інтервалами модальним інтервалом є інтервал з найбільшою частотою f_i .

Наприклад:

X_i	f_i
До 500	14
500–1 500	18
1 000–1 500	28
1 500–2 000	11
2 000–2 500	14
2 500–3 000	15

У нашому прикладі модальним буде інтервал 1 000 – 1 500 , у якого $f_i = 28 = \max$

Значення моди визначають за формулою:

$$M_o = X_{mo} + i_{mo} \cdot \frac{f_{mo} - f_{mo-1}}{(f_{mo} - f_{mo-i}) + (f_{mo} - f_{mo+1})} = 1\,000 + 500 \cdot \frac{28 - 18}{(28 - 18) + (28 - 11)} = 1\,185;$$

$$M_o = X^{mo} - i_{mo} \cdot \frac{f_{mo} - f_{mo+1}}{(f_{mo} - f_{mo+1}) + (f_{mo} - f_{mo-1})} = 1500 - 500 \cdot \frac{28 - 11}{(28 - 11) + (28 - 18)} = 1185.$$

де X_{mo} ; X^{mo} – нижня та верхня межі модального інтервалу;

i_{mo} – довжина модального інтервалу;

f_{mo} , f_{mo-1} , f_{mo+1} – частоти модального, попереднього та наступного інтервалів.

У формулі розрахунку моди прийняте припущення, що відстань від початку модального інтервалу до моди й від моди до кінця модального інтервалу пропорційна різниці частот між модальним і прилягаючих до нього інтервалів

Медіана – це значення ознаки, що доводиться на середину впорядкованого ряду й ділить ряд на дві рівні за обсягом частини. Для дискретних рядів розподілу медіана – це варіант, розташований у центрі ряду.

Для її розрахунку необхідно визначити накопичені частоти f'_i й тоді медіаною буде значення X_i для якого $f'_{me} \geq 0.5 \sum f_i$

X_i	2	4	6	8	10	12
f_i	3	7	1	4	5	15
f'_i	3	10	11	15	20	35

У нашому випадку $f'_{me} \geq 0,5 \cdot 35 \geq 17,5$. Тоді медіана дорівнює:

$$i_a = 10.$$

Для інтервальних рядів з рівними інтервалами для визначення медіано – утримуючого інтервалу необхідно визначити $f'_{me} \geq 0.5 \sum f_i$ (табл. 4.6).

Таблиця 4.6 – Дані для інтервальних рядів

№ п/п	Xi_{int}	f_i	f'_i
1	До 500	14	14
2	500–1 000	18	14 + 18 = 32
3	1000–1 500	28	60
4	1 500–2 000	11	71
5	2 000–2 500	14	85
6	2 500–3 000	15	100

Для нашого прикладу $f'_{me} \geq 0,5 \sum f_i \geq 0,5 \cdot 100 \geq 50$.

Тоді медіанним інтервалом є інтервал 1 000 – 1 500.

Значення медіани визначають за формулою:

$$M_e = X_{me} + i_{me} \frac{0,5 \sum f_i - f'_{me-1}}{f_{me}} = 1000 + 500 \cdot \frac{0,5 \cdot 100 - 32}{28} = 1321;$$

де f'_{me-1} – накопичена сума частот інтервалу попереднього медіанному;

f_{me} – власна частота медіанного інтервалу.

$$M_e = X^{me} - i_{me} \cdot \frac{f'_{me} - 0,5 \sum f_i}{f_{me}} = 1500 - 500 \frac{60 - 0,5 \cdot 100}{28} = 1321.$$

Для інтервальних рядів з нерівними інтервалами для знаходження медіанного інтервалу необхідно визначити абсолютну щільність розподілу D_i , після чого знайти відносну щільність розподілу D'_i , та накопичену відносну щільність розподілу P''_i (табл. 4.7).

Таблиця 4.7 – Для визначення медіанного інтервалу

№ п/п	X_i	f_i	$P_i = \frac{f_i}{i}$	$P'_i = \frac{P_i}{\sum P_i}$	P''_i
1	До 200	10	0,05	0,265	0,265
2	200 – 500	18	0,06	0,317	0,265 + 0,317 = 0,58
3	500 – 1 000	28	0,056	0,296	0,878
4	1 000 – 2 000	11	0,011	0,058	0,936
5	2 000 – 4 000	14	0,007	0,037	0,973
6	4 000 – 7 000	15	0,005	0,027	1 000
	Разом:	96	0,189	1,00	

У нашому прикладі медіанним інтервалом є інтервал 200–500, оскільки у ньому перебуває середнє значення $P''_{me} \geq 0,5 \Sigma P'_{me} \geq 0,5 \cdot 1 \geq 0,5$.

Тоді одержимо:

$$M_e = 500 - 300 \cdot \frac{0,582 - 0,5 \cdot 1}{0,317} = 422,6.$$

Основною властивістю медіани широко використовуваною на практиці є те, що сума відхилень значень X_i від медіани мінімальна

$$\Sigma(X_i - M_e) = \min.$$

ТЕМА 5 РЯДИ РОЗПОДІЛУ ТА ЇХ АНАЛІЗ

- 5.1 *Поняття варіації та її основні показники.*
- 5.2 *Види дисперсії та методи її обчислення розрахунку.*
- 5.3 *Розрахунок дисперсії методом моментів.*

5.1 Поняття варіації та її основні показники

В одних сукупностях індивідуальні значення ознаки щільно групуються навколо центру розподілу, в інших – значно відхиляються. Чим менші відхилення, тим однорідніша сукупність, а отже, тим більш надійні і типові середні характеристики розподілу.

Коливання окремих значень ознаки характеризують показники варіації. Термін «варіація» походить від латинського *variatio* – зміна, коливання, відмінність. Варіацією в статистиці називають кількісні зміни величини досліджуваної ознаки в межах однорідної сукупності, які зумовлені впливом дії різних чинників.

Неважко уявити дві сукупності, в яких середні величини варіюючої ознаки однакові, проте наближення індивідуальних значень у кожній сукупності до середньої є різним. Ось чому для характеристики сукупності велике практичне значення має вивчення відхилень досліджуваної ознаки

окремих одиниць сукупності від середньої величини. Треба брати до уваги не лише крайні відхилення, а й сукупність відхилень усіх варіантів. Від розміру і розподілу цих відхилень залежить типовість і надійність середніх величин.

Варіація зумовлена дією багатьох факторів, які поділяються на систематичні та випадкові. При вивченні варіації вирішуються три головних завдання (відповідно існує й три групи показників):

1. характеристика центру розподілу (середня, мода і медіана);
2. характеристика розміру та ступеня варіації;
3. характеристика виду та типу розподілу.

Вивчення варіації має велике значення з точки зору аналізу диференціації соціально-економічних явищ та процесів. Показники варіації покладено в основу вивчення взаємозв'язку між ознаками (дисперсійний аналіз), а також вибіркового спостереження.

Для вимірювання та оцінки варіації використовуються абсолютні й відносні характеристики. До абсолютних показників варіації належать:

- розмах варіації R ;
- середнє лінійне відхилення l ;
- середнє квадратичне відхилення σ ;
- дисперсія або квадрат середнього квадратичне відхилення σ^2 .

Відносні характеристики представлені низкою коефіцієнтів. До відносних показників варіації відносяться:

- коефіцієнт осциляції V_R ;
- лінійний коефіцієнт варіації $V_{\bar{p}}$,
- квадратичний коефіцієнт варіації V_{σ} .

Всі перелічені показники є іменованими величинами, крім коефіцієнтів варіації, які обчислюються у відсотках.

Система абсолютних показників вимірювання та оцінки розміру варіації:

1. Розмах варіації (R), що характеризує максимальну амплітуду коливань значень ознаки у сукупності:

$$R = X_{\max} - X_{\min}, \quad (5.1)$$

де: X_{\max} – найбільше та найменше значення ознаки;

X_{\min} – відповідно найбільше та найменше значення ознаки сукупності.

В інтервальних рядах розподілу розмах варіації визначається як різниця між верхньою межею останнього та нижньою межею першого інтервалу. Перевагою даного показника є простота обчислення та ясність економічної інтерпретації. Головний недолік полягає у тому, що він визначається по двох граничних величинах, які часто є випадковими.

2. Середнє лінійне відхилення (l), що характеризує середній розмір коливань значень ознаки навколо середнього рівня:

$$l = \frac{\sum |X - \bar{X}|}{n} \quad \text{просте,} \quad l = \frac{\sum |X - \bar{X}| \cdot f}{\sum f} \quad \text{зважене;} \quad (5.2)$$

де X – значення ознаки або варіанти;

\bar{X} – середнє значення;

n – кількість одиниць у сукупності;

f – частоти.

Просте середнє лінійне відхилення визначається по індивідуальних даних, а зважене – в рядах розподілу. В інтервальних рядах розподілу спочатку знаходиться середина кожного інтервалу, а далі робляться обчислення за наведеною формулою.

3. Середнє квадратичне відхилення:

$$\sigma = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n}} \quad \text{– просте;} \quad (5.3)$$

$$\sigma_{зв.} = \sqrt{\frac{\sum(x_i - \bar{x})^2 \times f_i}{\sum f_i}} - \text{зважене.} \quad (5.4)$$

Між величиною σ й l існує взаємозв'язок $\sigma \cong 1,25 l$. Між σ і R також існує взаємозв'язок $D_n = \frac{R}{\sigma}$ де значення D_n представлені в статистичних таблицях для різного числа одиниць сукупності. Дана залежність використовується для експрес розрахунку величини σ . У цьому випадку визначають значення R_i для кожної групи сукупності, потім визначають \bar{R} після чого величина середньоквадратичне відхилення буде дорівнювати:

$$\sigma = \frac{\bar{R}}{D_n}, \quad (5.5)$$

4. Дисперсія (σ^2) – це середній квадрат відхилень значень ознаки від середнього рівня:

$$\sigma^2 = \frac{\sum X^2}{n} - \left(\frac{\sum X}{n}\right)^2 \quad \text{або} \quad \sigma^2 = \frac{\sum X^2 f}{\sum f} - \left(\frac{\sum X f}{\sum f}\right)^2. \quad (5.6)$$

Вирахування дисперсії зручно робити за такими формулами:

$$\sigma^2 = \left(\overline{x^2}\right) - (\bar{x})^2 - \text{проста;} \quad (5.7)$$

$$\sigma^2 = \frac{\sum x_i^2 f_i}{\sum f_i} - \left(\frac{\sum x_i f_i}{\sum f_i}\right)^2 - \text{зважена.} \quad (5.8)$$

Середнє квадратичне відхилення є мірилом надійності середньої. Чим менше середнє квадратичне відхилення, тим об'єктивніше середня арифметична відображує всю сукупність.

Усі розглянуті показники варіації – розмах варіації, середнє лінійне відхилення та середнє квадратичне відхилення – завжди виражають у одиницях вихідних даних ряду та середньої величини. Всі вони є абсолютним виміром варіації. А це означає, що безпосередньо порівнювати абсолютні показники варіації у варіаційних рядах різних явищ не можна. Для того, щоб забезпечити їх порівняння, потрібно обчислити показники, які характеризують варіацію, виражену в стандартних величинах, наприклад, у відсотках. Відношення абсолютних характеристик варіації до середньої величини називаються коефіцієнтами варіації. Існують такі коефіцієнти варіації:

1. Лінійний коефіцієнт варіації:

$$V_l = \frac{\bar{l}}{\bar{x}} 100. \quad (5.9)$$

Квадратичний коефіцієнт варіації характеризує відносне середнє відхилення індивідуальні.

2. Квадратичний коефіцієнт варіації:

$$V_\sigma = \frac{\sigma}{\bar{x}} 100. \quad (5.10)$$

Для характеристики і порівняння варіації різних сукупностей і різних явищ найчастіше використовують квадратичний коефіцієнт варіації. Цей показник вживається для оцінки однорідності сукупності, тобто надійності і типовості середньої величини. Розрізняють такі значення відносних коливань: Вважають, що сукупність є однорідною, а середня – типовою, коли квадратичний коефіцієнт варіації не перевищує $V_\sigma = 0,33$.

3. Коефіцієнт осциляції:

$$V_R = \frac{R}{\bar{x}} \cdot 100 \%. \quad (5.11)$$

Коефіцієнт осциляції відображує відносне коливання крайніх значень ознаки навколо середньої.

Отже, можна зробити такі висновки:

1. Середня величина (середня арифметична) має важливе пізнавальне значення, однак вона не завжди об'єктивно і не завжди із однакою ступенем достовірності відображує внутрішній стан статистичної сукупності. При однаковому значенні середньої статистичні сукупності можуть бути досить нерівноцінні за рівнем коливань (варіації).

2. Чим менше відхилення, тим типовіша середня, тим більш однорідна сукупність.

3. Універсальним показником варіації є коефіцієнт варіації, цінність якого полягає в тому, що ним можна користуватись для характеристики і порівняння варіації різних сукупностей і різних явищ.

5.2 Види дисперсії та методи їх розрахунку

Дисперсія має певні математичні властивості. Сформулюємо найважливіші з них:

1. Дисперсія постійної величини (A) рівна нулю: $\sigma^2(A) = 0$.

2. Дисперсія не міняється, якщо всі варіанти збільшити або зменшити на одне й те ж число: $\sigma^2(x + A) = \sigma^2(x)$.

3. Якщо всі варіанти помножити (розділити) на число k , дисперсія збільшиться (зменшиться) в k^2 раз.

4. Якщо вагу збільшити або зменшити в одне й те ж число раз, то дисперсія не зміниться.

5. Дисперсія щодо середньої арифметичної менше ніж середній квадрат відхилення від будь-якого числа X_0 на величину $(X_0 - \bar{X})^2$. Ця властивість має назву властивості мінімальності дисперсії від середньої.

Дисперсія посідає особливе місце у статистичному аналізі соціально-економічних явищ. На відміну від інших характеристик варіації завдяки своїм

математичним властивостям вона є невіддільним і важливим елементом інших статистичних методів, зокрема дисперсійного аналізу. Дисперсія має певні математичні властивості, найважливіші з них:

1. Якщо всі значення варіант x_j зменшити на сталу величину A , то дисперсія не зміниться:

$$\sigma_{(x-A)}^2 = \sigma_x^2 . \quad (5.12)$$

2. Якщо всі значення варіант x_j змінити в A раз, то дисперсія зміниться в A^2 раз:

$$\sigma_{xA}^2 = \sigma_x^2 A^2 . \quad (5.13)$$

3. Якщо частоти замінити частками, дисперсія не зміниться.

Якщо сукупність розбита на групи за одним фактором, то для неї можна обчислити такі види дисперсій: загальну, групові (часткові), середню з групових (часткових) і міжгрупову.

Загальна дисперсія визначається як середня арифметична з квадратів відхилень кожного значення ознаки від їх загальної середньої величини. Дана дисперсія характеризує варіацію досліджуваної ознаки за рахунок впливу всіх факторів.

Внутрішньогрупова дисперсія характеризує варіацію індивідуальних значень відносно групових середніх.

Для всіх в цілому розраховується **середня з групових дисперсій**, зважених на частоти відповідних груп:

$$\bar{\sigma}^2 = \frac{\sum_1^n \sigma_i^2 f_i}{\sum_1^n f_i} , \quad (5.14)$$

Міжгрупова дисперсія розраховується за формулою:

$$\delta^2 = \frac{\sum_1^n (\bar{y}_i - \bar{y})^2 f_i}{\sum_1^n f_i}, \quad (5.15)$$

де \bar{y} – загальна середня ознаки, що варіює.

Взаємозв'язок між трьома дисперсіями отримав назву: правило складання дисперсій:

$$\sigma^2 = \delta^2 + \bar{\sigma}^2, \quad (5.16)$$

Дисперсія альтернативної ознаки

В економічному аналізі поряд з кількісною дисперсією має місце і якісна, атрибутивна дисперсія. При цьому якщо мають місце дві взаємовиключних варіанти, то дисперсія називається **альтернативною**. Якщо більше двох – **множинною**. Позначимо наявність ознаки у спостережуваного явища через 1, а його відсутність через 0, частку варіантів, які володіють ознакою через – «*P*», а які не володіють – через «*q*». Тоді дисперсія альтернативної ознаки дорівнює:

$$\delta^2 = p \cdot q, \quad (5.17)$$

а середнє квадратичне відхилення:

$$\delta = \sqrt{p \cdot q}. \quad (5.18)$$

5.3 Розрахунок дисперсії методом моментів

Метод моментів заснований на використанні властивостей дисперсії. Під моментом у статистиці розуміють середню величину варіаційного ряду.

Для дискретних рядів величину моменту визначають за формулою:

$$M^k = \overline{(x_i - A)^k}, \quad \text{де } k - \text{порядок моменту,} \quad (5.19)$$

$$M_k = \frac{\Sigma(x_i - A)^k}{n} - \text{простий,} \quad (5.20)$$

$$M_{k.зв} = \frac{\Sigma(x_i - A)^k f_i}{\Sigma f_i} - \text{зважений.} \quad (5.21)$$

З цих формул можна визначити значення моменту різного порядку:

$$M_1 = \frac{\Sigma(x_i - A)f_i}{\Sigma f_i} - \text{момент першого порядку,} \quad (5.22)$$

$$M_2 = \frac{\Sigma(x_i - A)^2 f_i}{\Sigma f_i} - \text{момент другого порядку,} \quad (5.23)$$

$$M_3 = \frac{\Sigma(x_i - A)^3 f_i}{\Sigma f_i} - \text{момент третього порядку.} \quad (5.24)$$

Якщо $A = 0$, то моменти називають початковими й тоді:

$$M_1 = \bar{x}; \quad M_2 = \overline{(x^2)}; \quad M_3 = \overline{(x^3)} \text{ тощо.}$$

Початкові моменти самостійного значення не мають і використовуються тільки для спрощення розрахунків.

Якщо $A = \bar{x}$, то момент називається центральним і тоді

$\mu_1 = 0$, $\mu_2 = \sigma^2$, $\mu_3 = 0$ для симетричного розподілу значень ознаки.

Тоді для дискретного ряду будемо мати:

$$\sigma^2 = \overline{x^2} - (\bar{x})^2 = \frac{\sum x_i^2 f_i}{\sum f_i} - \left(\frac{\sum x_i f_i}{\sum f_i} \right)^2, \quad (5.25)$$

але $\frac{\sum x_i^2 f_i}{\sum f_i} = \overline{(x^2)} = M_2$, а $\left(\frac{\sum x_i f_i}{\sum f_i} \right)^2 = (\bar{x})^2 = M_1^2$,

тоді

$$\sigma^2 = M_2 - M_1^2. \quad (5.26)$$

Якщо ряд інтервальний, то моменти в цьому випадку дорівнюють:

$$M_{\text{інт.к}} = \left(\frac{x_i - A}{i} \right)^k, \quad (5.27)$$

де i – довжина інтервалу.

Величина $M_{\text{інт.к}}$, як середнє арифметичне, визначається за формулою:

$$M_{\text{інт.к}} = \frac{\sum \left(\frac{x_i - A}{i} \right)^k f_i}{\sum f_i}, \quad (5.28)$$

Тоді

$$M_{\text{інт.1}} = \frac{\sum \left(\frac{x_i - A}{i} \right)^1 f_i}{\sum f_i}; \quad M_{\text{інт.2}} = \frac{\sum \left(\frac{x_i - A}{i} \right)^2 f_i}{\sum f_i}, \quad \text{тощо.}$$

Порівнюючи значення $M_{\text{інт.к}}$ для інтервального ряду зі значеннями M_k для дискретного ряду отримаємо

$$M_{\text{інт.к}} = \frac{M_k}{i^k}, \quad (5.29)$$

$$M_k = M_{\text{інт.к}} \cdot i^k \quad \text{і тоді}$$

$$\sigma^2 = M_2 - M_1^2 = M_{\text{інт.2}} \cdot i^2 - (M_{\text{інт.1}} \cdot i)^2 = i^2 (M_{\text{інт.2}} - M_{\text{інт.1}}^2). \quad (5.30)$$

Приклад

Розрахувати значення середньої арифметичної, дисперсії й середнього квадратичного відхилення товарообігу на підприємствах регіону методом моментів (табл. 5.1).

Таблиця 5.1 – Методів

№ з/п	Група підприємств за т/о тис. грн	Число підприємств f_i	Розрахункові показники				
			\bar{x}_{im}	$\bar{x}_{im} - A$ $A = 110$	$i = 20$ $\frac{\bar{x}_{im} - A}{i}$	$\frac{\bar{x}_{im} - A}{i} \cdot f_i$	$\left(\frac{\bar{x}_{im} - A}{i}\right)^2 \cdot f_i$
1	60–80	21	70	–40	–2	–42	84
2	80–100	27	90	–20	–1	–27	27
3	100–120	24	110	0	0	0	0
4	120–140	16	150	20	1	16	16
5	140–160	8	150	40	2	16	32
6	160–180	4	170	60	3	12	36
	Разом:	100				–25	195

Визначаємо середню арифметичну методом моментів: $\bar{X} = A + iM_{\text{інт.1}}$.

Виходячи з формули для $M_{\text{інт.1}}$ можна записати, що

$$M_{\text{інт.1}} = \frac{\sum \left(\frac{x_{im} - A}{i} \right) \cdot f_i}{\sum f_i} = \frac{-25}{100} = -0,25, \text{ тоді } \bar{X} = 110 + 20(-0,25) = 105 \text{ тис. грн.}$$

Момент $M_{\text{інт.2}}$ буде дорівнювати:

$$M_{\text{інт.2}} = \frac{\sum \left(\frac{x_{iu} - A}{i} \right)^2 \cdot f_i}{\sum f_i} = \frac{195}{100} = 1,95,$$

і тоді

$$\sigma^2 = i^2 (M_{\text{інт.2}} - M_{\text{інт.1}}^2) = 20^2 [1,95 - (-0,25)^2] = 400(1,95 - 0,0625) = 754 \text{ тис. грн,}$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{754} = 27,4 \text{ т. грн.}$$

ТЕМА 6 СТАТИСТИЧНІ МЕТОДИ ВИМІРЮВАННЯ ВЗАЄМОЗВ'ЯЗКІВ

- 6.1 Види взаємозв'язків між явищами.
- 6.2 Суть і значення дисперсійного аналізу.
- 6.3 Основи кореляційно-регресійного аналізу.
- 6.4 Множинна регресія і кореляція.

6.1 Види взаємозв'язків між явищами

Усі явища у суспільному житті існують не ізольовано, а у нерозривному зв'язку, тобто залежать одне від одного. При цьому виділяються факторні (x) і результативні (y) ознаки.

Для кількісних ознак залежності між окремими явищами можуть бути:

– функціональними, коли певному значенню однієї змінної, чиннику (x), відповідає певне значення результативної ознаки (y);

– кореляційними (статистичними), коли із зміною факторної ознаки (x) змінюються групові середні результативної ознаки (y).

Особливістю функціонального зв'язку є те, що він зберігає повністю свою силу для кожної одиниці сукупності і в усіх спостереженнях. Цю особливість називають стійкою тісністю. Така залежність притаманна фізичним, хімічним явищам тощо. У суспільних процесах це найчастіше зв'язок складових елементів розрахункових формул відповідних показників, наприклад залежність продуктивності праці від обсягу товарообороту і чисельності робітників підприємства.

При кореляційному зв'язку конкретному значенню факторної ознаки x відповідає певне середнє значення результативної ознаки y. За допомогою формули цей зв'язок можна описати неявною функцією $y = f(x)$.

Кореляційний зв'язок виявляється при великій кількості випробувань і лише через середні величини. Тому його класифікують, як неповний зв'язок.

Між ознаками суспільних явищ переважають саме кореляційні зв'язки. Наприклад, на продуктивність праці робітника впливає його кваліфікація, але не тільки. Умови праці, трудова дисципліна, технічне оснащення робочого місця, оплата праці є важливими важелями продуктивності праці.

За напрямом розрізняють зв'язки прямі і обернені.

Прямий зв'язок – це такий зв'язок, коли із зростанням факторної ознаки, результативна також зростає.

При оберненому зв'язку із збільшенням факторної ознаки результативна зменшується або, навпаки, із зменшенням факторної ознаки, результативна зростає.

За формою зв'язок ділиться на прямолінійний і криволінійний.

При прямолінійній кореляційній залежності рівним змінним середніх значень факторної ознаки відповідають приблизно рівні зміни середніх значень результативної ознаки.

При криволінійній кореляційній залежності рівним змінним середніх значень факторної ознаки відповідають нерівні зміни середніх значень результативної ознаки.

Статистичне вивчення взаємозв'язків розв'язує такі завдання:

- а) визначаються форми зв'язку;
- б) вимірюється тіснота (сила) зв'язку;
- в) виявляється вплив окремих чинників на результативну ознаку.

Зв'язки і залежності суспільних явищ вивчаються різними методами, які дають уявлення про їх наявність і характер. До цих методів відносять: балансовий метод, метод порівняння паралельних рядів, графічний метод, метод аналітичних групувань, індексний метод, кореляційно-регресійний аналіз тощо.

Наявність або відсутність зв'язків можна виявити, використовуючи:

- балансовий метод;
- метод аналітичних групувань;
- метод порівняння паралельних рядів;
- графічний метод;
- кореляційно-регресійний аналіз;
- індексний метод тощо.

6.2 Суть і значення дисперсійного аналізу

Дисперсійний аналіз – це метод статистичної оцінки залежності результативної ознаки від одного або кількох чинників.

Дисперсійний аналіз можна подати у вигляді п'яти послідовно виконуваних етапів:

- 1) визначення і розкладання варіації;
- 2) визначення числа ступенів свободи варіації;
- 3) обчислення дисперсій та їх співвідношень;
- 4) аналіз дисперсій та їх співвідношень;

5) оцінка вірогідності різниці між середніми і формулювання висновків з перевірки нульової гіпотези.

Найбільш трудомісткою частиною дисперсійного аналізу є перший етап – визначення і розкладання варіації за джерелами її утворення. Для кількісної оцінки зв'язку між явищами на базі матеріалів аналітичного групування обчислюють кореляційне відношення та індекс кореляції.

Кореляційне відношення показує питому вагу міжгрупової дисперсії у загальній дисперсії, тобто визначає, ступінь варіації ознаки від впливом чинника, покладеного в основу групування. Його обчислюють за формулою:

$$\eta = \sqrt{\frac{\sigma_m^2}{\sigma^2}} = \sqrt{\frac{\sigma_m^2}{\sigma_m^2 + \overline{\sigma_t^2}}}, \quad (6.1)$$

σ_m^2 – міжгрупова дисперсія результативної ознаки;

σ^2 – загальна дисперсія результативної ознаки;

$\overline{\sigma_t^2}$ – середня із внутрішньогрупових дисперсій результативної ознаки.

Емпіричне кореляційне відношення змінюється в межах від 0 до 1. Чим ближче його значення наближається до 1, тим сильнішим є взаємозв'язок між ознаками. При $\eta = 1$ зв'язок вважається функціональним.

Перевірка істотності відхилень групових середніх здійснюється за допомогою критеріїв математичної статистики. Вона ґрунтується на порівнянні фактичного значення η з так званим критичним. Останнє є тим максимально можливим значенням кореляційного відношення, яке може виникнути випадково при відсутності кореляційного зв'язку. Якщо фактичне значення η більше від критичного, то зв'язок між результативною і факторною ознаками вважається істотним. Якщо фактичне значення η менше критичного, то наявність кореляційного зв'язку між ознаками не доказана і зв'язок вважається неістотним.

Крім оцінки тісноти взаємозв'язку необхідно встановити, чи не є відмінність між середніми значеннями \bar{Y} по групах випадковою (несуттєвою),

тобто здійснити перевірку суттєвості зв'язку. Для цього значення η^2 порівнюються з певним критичним значенням, яке знаходиться з спеціальної таблиці, і якщо це відношення більше 1, коливання середніх можна вважати не випадковим, а зв'язок – суттєвим. Для перевірки суттєвості взаємозв'язку між X та Y часто використовують запропонований Р. Фішером показник, який має назву F-критерія та визначається за формулою:

$$F = \frac{\eta^2}{1+\eta^2} \cdot \frac{K_2}{K_1}, \quad (6.2)$$

Розподіл η залежить від числа ступенів вільності міжгрупової K_1 і середньої з групових K_2 дисперсій. В аналітичному групуванні вони обчислюються так:

$$K_1 = m - 1;$$

$$K_2 = n - m,$$

де n – кількість елементів досліджуваної сукупності;

m – число груп.

Критичні значення F-критерія для рівнів значимості 0,05 та 0,01 занесені у спеціальні таблиці. Із цих таблиць у відповідності зі значеннями K_1 та K_2 визначається так зване табличне значення F-критерія ($F_{\text{табл}}$). Якщо виконується умова $F > F_{\text{табл}}$, зв'язок між показниками можна вважати суттєвим, (невипадковим).

6.3 Основи кореляційно-регресійного аналізу

Кореляційно-регресійний аналіз є логічним продовженням дисперсійного аналізу і поглиблено досліджує та дає кількісну оцінку характеру і механізму взаємодії факторних і результативних ознак.

Кореляційно-регресійний аналіз об'єднує:

– кореляційний аналізі, який полягає у дослідженні тісноти зв'язку між

результативною і факторною (факторними) ознаками;

– регресійний аналіз, який вирішує питання побудови, розв'язання і оцінки рівнянь регресії.

Термін кореляція походить від англійського слова *correlation* – співвідношення, відповідність (взаємозв'язок, взаємозалежність) між ознаками, що виявляється при масовому спостереженні зміни середньої величини однієї ознаки залежно від значення іншої. Ознаки, що пов'язані між собою кореляційним зв'язком, називають корельованими.

Схематично кореляційно-регресійний аналіз можна поділити на п'ять етапів:

1) постановка завдання, встановлення наявності зв'язку між досліджуваними ознаками;

2) відбір найістотніших чинників для аналізу;

3) визначення характеру зв'язку, його напрямку і форми, вибір математичного рівняння для вираження існуючих зв'язків;

4) розрахунок числових характеристик кореляційного зв'язку (визначення параметрів рівняння і показників тісноти зв'язку);

5) статистична оцінка вибіркового показника зв'язку.

Науково обґрунтоване застосування кореляційно-регресійного методу потребує перед усім глибокого розуміння суті взаємозв'язків соціально економічних явищ. Сам метод не встановлює наявності і причини виникнення зв'язків між досліджуваними явищами, його призначення полягає в їх кількісному вимірюванні. На першому етапі кореляційно-регресійного аналізу здійснюється загальне ознайомлення з досліджуваним об'єктом і явищами, уточнюються мета і завдання дослідження, встановлюється теоретична можливість причинно-наслідкового зв'язку між ознаками.

Відбір чинників для включення їх в кореляційно-регресійну модель повинен базуватися передусім на теоретичних основах і практичному досвіді аналізу досліджуваного соціально-економічного явища.

Одною з головних проблем побудови регресійної моделі є визначення

форми зв'язку і на цій основі встановлення типу аналітичної функції, що відображає механізм зв'язку результативної ознаки з факторною (факторними). Під формою кореляційного зв'язку розуміють тип аналітичного рівняння, що виражає залежність між досліджуваними ознаками.

Залежно від вихідних даних теоретичною лінією регресії можуть бути різні типи кривих або пряма лінія:

Лінійна $\gamma_x = a_0 + a_1X$

Параболічна $\gamma_x = a_0 + a_1X + a_2X^2$

Кубічна $\gamma_x = a_0 + a_1X + a_2X^2 + a_3X^3$

Степенева $\gamma_x = a_0 \cdot a_1X^{a_1}$

Гіперболічна $\gamma_x = a_0 + \frac{a_1}{X}$

Розглянемо методику кореляційно-регресійного аналізу на прикладі лінійного зв'язку, який описується лінійним рівнянням регресії. На першому етапі доцільно застосувати графічний метод для формування гіпотези про наявність саме лінійного взаємозв'язку між ознаками.

На другому етапі визначаються параметри лінійного рівняння регресії:

$$\gamma_x = a_0 + a_1X . \quad (6.3)$$

Для цього найчастіше застосовують метод найменших квадратів, розроблений К. Гауссом. Суть його полягає у відшукуванні таких параметрів рівняння зв'язку, за яких залишкова варіація (сума квадратів відхилень) змінної від її значень, обчислених за рівнянням, буде мінімальною. Тобто, використовується метод найменших квадратів та розв'язується система рівнянь відносно a_0 і a_1 :

$$na_0 + a_1 \sum X = \sum Y;$$

$$a_0 \sum X + a_1 \sum X^2 = \sum XY \quad . \quad (6.4)$$

З наведеної системи параметри рівняння регресії розраховуються різними способами, в тому числі за формулами:

$$a_0 = \frac{\sum Y \sum X^2 - \sum XY \sum X}{n \sum X^2 - \sum X \sum X} \quad \text{або} \quad a_0 = \bar{Y} - a_1 \bar{X};$$

$$a_1 = \frac{n \sum XY - \sum X \sum Y}{n \sum X^2 - \sum X \sum X} \quad \text{або} \quad a_1 = \frac{\frac{1}{n} \sum XY - \bar{X} \bar{Y}}{\sigma^2_x}.$$

Параметри a_0 і a_1 мають певний економічний зміст, зокрема, a_1 називається коефіцієнтом регресії, що показує, на скільки одиниць змінюється Y при збільшенні X на одну одиницю. Якщо цей коефіцієнт додатний – зв'язок прямий, а якщо від'ємний – зв'язок між показниками обернений.

Коефіцієнт регресії a_1 показує, на скільки одиниць збільшиться результативна ознака при збільшенні факторної ознаки на одиницю. Коефіцієнт регресії відбиває міру впливу відповідного чинника на результативну ознаку.

Якщо $a_1 > 1$ – зв'язок результативної ознаки з факторною є прямим; $a_1 < 1$ – зв'язок зворотний.

Коефіцієнт регресії у невеликих за обсягом сукупностях схильний до випадкових коливань. Тому здійснюється перевірка його істотності за допомогою t-критерію Стьюдента:

На третьому етапі за одержаним рівнянням регресії розраховуються теоретичні значення результативної ознаки Y_x та показники рівня апроксимації (наближення), які показують розбіжність між фактичними і теоретичними значеннями Y :

– середня квадратична (стандартна) помилка:

$$S = \sqrt{\frac{\sum(Y - Y_x)^2}{n}}; \quad (6.5)$$

– коефіцієнт апроксимації:

$$V = \frac{S}{Y} \cdot 100. \quad (6.6)$$

Чим меншими є значення S та V , тим краще рівняння регресії описує (апроксимує) взаємозв'язок між X та Y .

На четвертому етапі оцінюється тіснота зв'язку за допомогою лінійного коефіцієнта кореляції (r):

$$r = \frac{\overline{XY} - \bar{X} \cdot \bar{Y}}{\sigma_x \sigma_y} = \frac{n \sum XY - \sum X \sum Y}{\sqrt{(n \sum X^2 - \sum X \sum X)(n \sum Y^2 - \sum Y \sum Y)}};$$

Значення коефіцієнта кореляції r знаходиться в межах від -1 до $+1$. При $r > 0$ зв'язок між показниками прямий, а при $r < 0$ – обернений. Якщо: $|r| < 0,3$ вважається, що зв'язок між X та Y практично відсутній; $0,3 < |r| < 0,5$ – зв'язок слабкий; $0,5 < |r| < 0,7$ – зв'язок середній; $0,7 < |r| < 0,9$ – зв'язок сильний; $0,9 < |r| < 1$ – зв'язок дуже сильний.

Коефіцієнт детермінації $D = r^2$ показує, на скільки відсотків варіація Y обумовлюється варіацією X .

Часто також визначається коефіцієнт еластичності (E) за формулою:

$$E = a_1 \frac{\bar{X}}{\bar{Y}}. \quad (6.7)$$

Цей коефіцієнт показує, на скільки процентів змінюється Y при

збільшенні X на 1 %.

На п'ятому етапі здійснюється перевірка суттєвості (невипадковості) взаємозв'язку між показниками за допомогою F-критерія Фішера:

$$F = \frac{r^2}{1-r^2} \cdot \frac{K_2}{K_1}, \quad (6.8)$$

де $K_1 = m-1$;

$K_2 = n-m$;

n – кількість одиниць у сукупності;

m – кількість параметрів у рівнянні регресії.

6.4 Множинна регресія і кореляція

Як відомо, більшість соціально-економічних показників формується під впливом не одного, а багатьох факторів. Метод побудови моделі такого зв'язку має назву багатофакторного кореляційно-регресійного аналізу. В цьому випадку результативна ознака (Y) пов'язується з допомогою рівняння множинної регресії з двома або більше факторними ознаками ($X_1, X_2, X_3, \dots, X_m$).

Найважливішими умовами побудови багатофакторної моделі зв'язку є достатня кількість одиниць у сукупності (як мінімум у 8 разів більше, ніж число факторів) та відсутність мультиколінеарності факторів (близького до функціонального зв'язку між ними). В тому випадку, якщо два факторних показники мультиколінеарні, один з них повинен бути виключений з моделі.

На практиці використовуються два види рівнянь множинної регресії:

– лінійне(адитивне):

$$Y_x = a_0 + a_1X_1 + a_2X_2 + \dots + a_mX_m;$$

– нелінійне (мультиплікативне):

$$Y_x = a_0 \cdot X_1^{a_1} \cdot X_2^{a_2} \cdot \dots \cdot X_m^{a_m},$$

де $a_0, a_1, a_2, \dots, a_m$ – параметри рівняння множинної регресії;

$X_1, X_2, X_3, \dots, X_m$ – факторні ознаки.

Оцінка параметрів рівняння множинної регресії здійснюється методом найменших квадратів. Параметри a_1, a_2, \dots, a_m називаються коефіцієнтами регресії та показують, на скільки одиниць змінюється Y при збільшенні X на одиницю, при умові, що інші фактори є сталими. Наприклад, рівняння залежності ціни (Y) від рівня продуктивності праці (X_1) та якості сировини (X_2):

$$Y_x = 10,2 + 12,6X_1 + 0,7 X_2.$$

Для вимірювання тісноти взаємозв'язку між двома ознаками, що включені у модель, визначають парні коефіцієнти кореляції ($r_{yx_1}, r_{yx_2}, r_{x_1x_2}$). Тісноту зв'язку між результативною ознакою (Y) та факторною (при спільному впливі всіх факторів) характеризують часткові коефіцієнти кореляції (R_{yx_1}, R_{yx_2}).

Тісноту взаємозв'язку між результативною ознакою та сукупністю всіх факторних ознак визначають на основі коефіцієнта множинної кореляції R . Величина $D = R^2$ називається коефіцієнтом детермінації, що показує, на скільки процентів варіація Y обумовлюється варіацією всіх факторних ознак, включених у модель.

ТЕМА 7 АНАЛІЗ ТЕНДЕНЦІЙ РОЗВИТКУ

7.1 Поняття про ряди динаміки. Види рядів динаміки та правила їх побудови.

7.2 Статистичні характеристики динамічних рядів та їх взаємозв'язок.

7.3 Середні показники динаміки.

7.4 Зіставність в рядах динаміки.

7.5 Методи аналізу тенденцій розвитку.

7.6 Аналітичне вимірювання ряду.

7.7 Інтерполяція і екстраполяція. Аналіз сезонних коливань.

7.1 Поняття про ряди динаміки. Види рядів динаміки та правила їх побудови

Суспільні явища безперервно змінюються. Протягом певного часу – місяць за місяцем, рік за роком змінюється чисельність населення, обсяг і структура суспільного виробництва, рівень продуктивності праці тощо. Процес розвитку суспільних явищ у часі називається динамікою. Зміна явищ і процесів у часі відбувається під впливом різних соціальних, економічних, технічних та інших чинників. Вивчення динаміки дозволяє виявити і оцінити особливості розвитку явищ протягом часу під впливом різних чинників.

Основна мета статистичного вивчення динаміки полягає у виявленні закономірностей соціально-економічних явищ. Система статистичних методів вивчення динаміки явищ дозволяє визначити як розвиваються суспільні явища, зростають чи зменшуються їх розміри, швидко чи повільно відбуваються ці зміни тощо.

Динамічний ряд – це статистичні показники, розташовані в хронологічній послідовності, які характеризують розвиток того чи іншого соціально-економічного явища у часі. Для будь-якого динамічного ряду характерні перелік хронологічних дат (моментів) або інтервалів часу і конкретні значення

відповідних статистичних показників. Окремі числові значення розмірів явищ називають рівнями ряду. Рівень ряду відображає стан явищ, досягнутий за будь-який період або на певний момент часу. Складовими динамічного ряду є ознака часу «t» (момент або інтервал) та числові значення показника – рівні. Перший показник ряду називається початковим, а останній – кінцевим.

При вивченні динаміки важливі не лише числові значення рівнів, але і послідовність їх. Як правило, часові інтервали поміж рівнями однакові (доба, декада, календарний місяць, квартал, рік). Приймаючи будь-який інтервал за одиницю, послідовність рівнів можна записати так: $y_1, y_2, y_3, \dots, y_n$ де «n» – число рівнів (довжина динамічного ряду).

Відповідно до класифікації показників за ознакою часу динамічні ряди поділяють на моментні та інтервальні. У моментних рядах рівні фіксують стан явища на певні моменти часу, в інтервальних – за певний проміжок часу. Приклади зазначених рядів динаміки наведено нижче:

– *моментний ряд:*

Дата, t	на 01.01	на 01.04	на 01.07	на 01.10	на 01.01
Кількість студентів, У	1 000	1 120	1 110	1 150	1 130

– *інтервальний ряд:*

Періоди, t	I кв.	II кв.	III кв.	IV кв.
Товарообіг, У у. о.	400	550	370	520

Характерні особливості інтервального динамічного ряду: залежність величини рівня від величини проміжку часу; показники періодичного (інтервального) ряду динаміки, як правило, можна додавати і ці показники мають реальний конкретний зміст; наприклад, в результаті додавання можна одержати новий ряд динаміки, кожний показник якого характеризує величину явища за збільшені періоди часу.

Залежно від статистичної природи показника-рівня розрізняють динамічні ряди первинні і похідні, ряди абсолютних, середніх і відносних величин.

Крім цього ряди динаміки поділяють на одно- і багатомірні. Одномірні характеризують зміну одного показника (наприклад, видобуток газу), багатомірні – двох, трьох і більше показників. У свою чергу, багатомірні динамічні ряди поділяються на два види: паралельні та ряди взаємопов'язаних показників.

Паралельні відображають динаміку або одного і того самого показника щодо різних об'єктів (національний дохід по країнах, прибуток по підприємствах тощо), або різних показників одного і того ж об'єкта (видобуток вугілля, нафти і газу в регіоні).

Ряди взаємопов'язаних показників характеризують динаміку декількох показників, взаємопов'язаних між собою. Зв'язок між показниками багатомірного динамічного ряду може бути функціональним (адитивним чи мультиплікативним) або кореляційним.

Прикладом адитивно зв'язаних рядів є динаміка цілого і його складових частин (чисельності всього населення і у тому числі міського і сільського); мультиплікативно зв'язаних – динаміка посівної площі, урожайності і валового збору певної сільськогосподарської культури; кореляційно зв'язаних – динаміка фондоозброєності і продуктивності праці.

При побудові рядів динаміки необхідно забезпечити порівнянність рівнів ряду за територією, методикою розрахунку показників, періодом або моментом часу, об'єктом і одиницею спостереження, ступенем охоплення одиниць досліджуваної сукупності, одиницями вимірювання тощо.

Насамперед, всі рівні ряду динаміки повинні характеризувати одне і теж явище. Цього можна досягти лише тоді, коли протягом всього періоду, який охоплюється динамічним рядом, будуть незмінними зміст і межі об'єкта та одиниці спостереження. Наприклад, при характеристиці динаміки чисельності студентів ВНЗ по роках не можна в одні роки враховувати тільки чисельність

студентів денної форми навчання, а в інші – чисельність студентів всіх видів навчання.

Кожен рівень динамічного ряду повинен бути визначений (розрахований) за однією методологією. Важливою умовою порівнянності рядів динаміки є вираження їх рівнів в однакових одиницях вимірювання. З різницею одиниць вимірювання приходиться зустрічатися при обліку продукції в натуральному вираженні.

Приведення до порівняння різних видів продукції досягається вираженням її у вартісних або трудових вимірниках. Необхідно забезпечити порівнянність рівнів інтервальних рядів динаміки щодо тривалості відрізків часу, а в моментних рядах щодо відношення до однієї й тієї ж дати року.

Потрібно також забезпечити територіальну порівнянність, тобто використовувати дані по території в одних і тих же межах.

Таким чином, всі вище названі обставини слід враховувати при підготовці інформації для аналізу змін явищ у часі.

7.2 Статистичні характеристики динамічних рядів та їх взаємозв'язок

При вивченні закономірностей соціально-економічного розвитку статистика вирішує низку завдань:

- характеристика інтенсивності окремих змін у рівнях ряду від періоду до періоду або від дати до дати;
- визначення середніх показників динамічного ряду за той чи інший період;
- виявлення основних закономірностей динаміки досліджуваного явища на окремих етапах або за весь період, що вивчається;
- виявлення чинників, що зумовили зміни досліджуваного об'єкту у часі;
- прогноз розвитку явищ на майбутнє.

Вивчаючи особливості розвитку соціально-економічних явищ визначають абсолютні та відносні характеристики динаміки:

- абсолютний приріст та абсолютне значення 1 % приросту;
- темп зростання;
- темп приросту;
- темп нарощування.

Розрахунок цих характеристик ґрунтується на порівнянні рівнів динамічного ряду. Якщо база порівняння постійна, характеристики динаміки називаються базисними, якщо база порівняння змінна – ланцюговими.

Абсолютний приріст (зменшення) ΔY_i – це різниця рівнів динамічного ряду:

$$\text{базисні } \Delta Y_{б.i} = Y_i - Y_o, \text{ ланцюгові } \Delta Y_{л.i} = Y_i - Y_{i-1}. \quad (7.1)$$

Зв'язок базисних і ланцюгових абсолютних приростів має такий вигляд:

$$\sum_{i=1}^n \Delta Y_{л.i} = \Delta Y_{б.n} \quad (7.2)$$

Темп зростання $T_{зр}$ розраховується як відношення рівнів ряду. Він виражається коефіцієнтом або процентом:

$$\text{базисні } T_{зр.б.i} = \frac{Y_i}{Y_o}; \text{ ланцюгові } T_{зр.л.i} = \frac{Y_i}{Y_{i-1}}. \quad (7.3)$$

Зв'язок базисних і ланцюгових темпів зростання має наступний вигляд:

$$T_{зр.б.n} = \prod_{i=1}^n T_{зр.л.i} = T_{зр.л.1} \cdot T_{зр.л.2} \cdot \dots \cdot T_{зр.л.n}. \quad (7.4)$$

Темп приросту $T_{пр}$ показує, на скільки процентів рівень Y_i більше (менше) рівня, взятого за базу порівняння. Його можна визначити як відношення абсолютного приросту до бази порівняння або безпосередньо на основі темпу зростання:

$$\text{базисні } T_{пр.б.i} = \frac{\Delta Y_{б.i}}{Y_o}; \quad \text{ланцюгові } T_{пр.л.i} = \frac{\Delta Y_{л.i}}{Y_{i-1}} \quad (7.5)$$

$$T_{пр.б.i} = T_{зр.б.i} - 1 \quad T_{пр.л.i} = T_{зр.л.i} - 1.$$

Темпи нарощування $T_{нар.}$ Визначають як відношення абсолютного ланцюгового приросту до постійної бази порівняння:

$$T_{нар.i} = \frac{\Delta Y_{л.i}}{Y_o}. \quad (7.6)$$

Зв'язок темпу нарощування і темпу приросту має вигляд:

$$T_{пр.б.n} = \sum_{i=1}^n T_{нар.i}. \quad (7.7)$$

Абсолютне значення 1 % приросту показує, чого вартий один процент, розраховується як співвідношення абсолютного приросту і темпу приросту. Алгебраїчно це співвідношення дорівнює 0,01 рівня, взятого за базу порівняння.

$$A \% = \frac{\Delta Y_i}{T_{пр.i}} \quad (7.8)$$

Для базисних темпів приросту значення A % однакові.

7.3 Середні показники динаміки

З плином часу змінюються рівні динамічних рядів і обчислені на їх основі показники аналізу ряду динаміки. Постає потреба узагальнення притаманних динамічному ряду властивостей, визначення типових характеристик розвитку.

Для узагальнюючої характеристики динаміки досліджуваного явища за низку періодів визначають різного роду середні показники. Розглянемо дві категорії таких показників:

- 1) середні рівні ряду;
- 2) середні показники змін рівнів ряду.

Середні рівні використовують, насамперед, для узагальнення коливних рядів. Порядок обчислення середнього рівня динамічного ряду залежить від виду ряду динаміки.

В інтервальному ряді абсолютних величин з рівними періодами часу використовується середня арифметична проста.

Статистичні характеристики рядів динаміки, розраховані за рівнями ряду, змінюються в часі. Вони варіюють за роками і можуть мати різні масштаби виміру. Все це вимагає їхнього узагальнення і розрахунку середніх показників:

- середнього рівня ряду;
- середніх темпів росту;
- середніх темпів приросту;
- середнього абсолютного приросту.

Середній рівень ряду називається також хронологічної середньої. Він показує, яка середня величина рівня характерна для аналізованого періоду. Середній рівень ряду обчислюється по-різному для інтервальних і моментних часових рядів.

Для рівно інтервального ряду середній рівень обчислюють за формулою:

$$\bar{Y} = \frac{\sum Y_i}{n}, \quad (7.9)$$

де n – число інтервалів ряду,

Y_i – рівень часового ряду в i -тім інтервалі.

Для моментного ряду середня величина складових його рівнів, у загальному випадку дорівнює:

$$\bar{Y} = \frac{\int_{t_1}^{t_n} f(t) dt}{t_n - t_1}; \quad (7.10)$$

де t_1, t_n – початкова й кінцева дати або моменти часу;

$f(t)$ – функція залежності рівня ряду від часу t .

Оскільки функція $Y = f(t)$ є безперервною функцією, а моментний ряд являє собою сукупність дискретних значень, то при реалізації загальної формули для розрахунку \bar{Y} виходять із припущення про рівномірну зміну рівнів часового ряду в інтервалі заданих значень t_1, t_n .

При цьому, якщо проміжки часу між датами рівні, то значення середньої рівнів можна визначити за формулою середньої хронологічної:

$$\bar{Y} = \frac{\frac{1}{2}Y_1 + Y_2 + Y_3 + \dots + Y_{n-1} + \frac{1}{2}Y_n}{n-1}. \quad (7.11)$$

У рядах динаміки з різними проміжками часу між датами середній рівень ряду визначають за формулою:

$$\bar{Y} = \frac{\sum Y_i t_i}{\sum t_i} = \frac{Y_1 \cdot t_1 + Y_2 \cdot t_2 + \dots + Y_n t_n}{t_1 + t_2 + \dots + t_n}; \quad (7.12)$$

Для рівноінтервального динамічного ряду середній рівень ряду визначається за формулою:

$$\bar{Y} = \frac{\sum_{i=1}^n Y_i}{n}. \quad (7.13)$$

Середній темп зростання $\bar{T}_{зр.}$ розраховують за формулою середньої геометричної:

$$\bar{T}_{зр.} = \sqrt[n]{\prod_{i=1}^n T_{зр.л.i}} = \sqrt[n]{T_{зр.л.1} \cdot T_{зр.л.2} \cdot \dots \cdot T_{зр.л.n}},$$

$$\bar{T}_{зр.} = \sqrt[n]{\frac{Y_n}{Y_0}}. \quad (7.14)$$

де n – число ланцюгових темпів зростання.

Середній темп приросту \bar{T}_{np} розраховують за формулою:

$$\bar{T}_{np} = \frac{\sum_{i=1}^n T_{np.л.i}}{n} \text{ або } \bar{T}_{np} = \bar{T}_{zp} - 1. \quad (7.15)$$

Середній абсолютний приріст обчислюють за формулою:

$$\bar{\Delta Y} = \frac{\sum_{i=1}^n \Delta Y_{y,i}}{n}. \quad (7.16)$$

7.4 Зіставність в рядах динаміки

Найважливішою умовою побудови рядів динаміки є співставність усіх рівнів. Неспівставність може виникнути з багатьох причин: територіальні та адміністративні зміни, перегляд методики розрахунку показників, їх одиниць виміру, зміна цін тощо. Для приведення ряду динаміки у співставний вигляд використовують метод прямого перерахунку або метод зімкнення ряду динаміки.

Метод прямого перерахунку полягає у тому, що нові рівні ряду динаміки розраховуються повторно з врахуванням тих змін, які відбулися. Метод зімкнення ряду динаміки передбачає, що нові значення рівнів ряду динаміки визначаються на основі перехідного коефіцієнту. Цей коефіцієнт розраховуються як відношення значення показника в нових умовах до значення того ж показника у старих умовах, які обчислені за однаковий період або момент часу.

Наприклад,

Місяці	1	2	3	4	5	6	7
Обсяг виробництва до зміни цін, тис. грн	300	312	320	360			
Обсяг виробництва після зміни цін, тис. грн				540	507	480	510

Визначаємо перехідний коефіцієнт (K_{Π}):

$$K_{\Pi} = \frac{540}{360} = 1,5.$$

Підрахуємо скоректовані рівні ряду динаміки до зміни цін:

$$y_1 = 300 \cdot 1,5 = 450; \quad y_2 = 312 \cdot 1,5 = 468; \quad y_3 = 320 \cdot 1,5 = 480.$$

Таким чином, якщо записати одержані числа у нижній рядок, одержимо співставний ряд динаміки обсягів виробництва продукції у нових цінах.

Ряди динаміки класифікуються за показником часу та способом виразу рівнів ряду. За показником часу розрізняють моментні та інтервальні ряди динаміки. В моментних рядах рівні ряду відносяться до певного моменту часу або дати, а в інтервальних – до періоду або інтервалу. В моментних рядах динаміки сума рівнів ряду не має економічного місту, тому середній рівень ряду визначається за формулою середньої хронологічної:

$$\bar{y} = \frac{\frac{1}{2}y_1 + y_2 + \dots + \frac{1}{2}y_n}{n-1}, \quad (7.17)$$

де n – кількість рівнів ряду динаміки.

Наприклад, маємо ряд динаміки чисельності працівників:

Дата	1.01.07	1.01.08	1.01.09	1.01.10	1.01.11	1.01.12	1.01.13
Чисельність працівників, чол.	820	834	811	807	800	804	740

Середня чисельність працівників за 2010–2017 рр. становить:

$$\bar{y} = \frac{\frac{1}{2}y_1 + y_2 + \dots + \frac{1}{2}y_n}{n-1} = \frac{\frac{1}{2}820 + 834 + 811 + 807 + 804 + 800 + \frac{1}{2}740}{7-1} = 806 \text{ чол.},$$

В інтервальних рядах динаміки сума рівнів ряду характеризує значення показника за більший інтервал часу і є самостійним показником. Середній рівень ряду в даному випадку визначається за формулою середньої арифметичної простої або зваженої:

$$\bar{y} = \frac{\sum y}{n} \quad \text{або} \quad \bar{y} = \frac{\sum y \cdot n}{\sum n}. \quad (7.18)$$

Наприклад, маємо ряд динаміки виробництва продукції протягом року:

Квартал	I	II	III	IV
Обсяг виробництва, т	200	212	195	220

Середньоквартальне виробництво продукції становить:

$$\bar{y} = \frac{\sum y}{n} = \frac{200+212+195+220}{4} = \frac{827}{4} = 207 \text{ т.}$$

Отже, за квартал обсяг виробництва продукції становить в середньому 207 тонн. За способом виразу рівнів розрізняють ряди динаміки абсолютних, відносних та середніх величин.

7.5 Методи аналізу тенденції розвитку

Застосування розглянутих нами показників динаміки є першим кроком аналізу динамічних рядів, що дозволяють виявити швидкість і інтенсивність розвитку явища, яке представлено даним рядом. Подальший аналіз рядів динаміки пов'язаний з більш складними узагальненнями, з визначенням основної тенденції розвитку, коливання рівнів та зв'язку рядів.

Щоб виявити і схарактеризувати основну тенденцію розвитку динамічного ряду застосовують механічні способи згладжування та аналітичного вирівнювання динамічних рядів. До механічних способів

згладжування відносять метод укрупнення інтервалів часу та метод ковзних середніх, які дозволяють усунути первинні коливання рівнів ряду. Метод ковзних середніх застосовують також для попередньої обробки дуже коливних динамічних рядів.

Метод укрупнення інтервалів. Якщо розглядати рівні економічних показників за короткі проміжки часу, то, у силу впливу різних факторів діючих у різному кількісному сполученні з різною інтенсивністю та у різному напрямку, у рядах динаміки спостерігаються їхнє зниження або збільшення. Тому не завжди можна побачити основну тенденцію розвитку цього явища.

Наприклад, вивчаючи дані врожайності зернових за певні проміжки часу можна побачити значні коливання рівнів ряду (табл. 7.1).

Рік, t	Валовий збір зерна Y_i млн т	Рік, t	Валовий збір зерна Y_i млн т
2000	125,5	2008	169,5
2001	130,8	2009	162,4
2002	140,2	2010	186,8
2003	107,5	2011	181,2
2004	152,1	2012	168,2
2005	121,1	2013	222,5
2006	171,2	2014	195,7
2007	147,0	2015	140,1

Представлені дані показують наявність коливання значень рівнів ряду, що не дозволяє визначити основну тенденцію розвитку. Застосовуючи метод укрупнення інтервалів (по п'ятирічках) одержимо такі значення (табл. 7.2).

Таблиця 7.2 – Метод укріплення інтервалів (по п'ятирічках)

П'ятирічки	Сумарний валовий збір	Середньорічний збір
2000–2005	651,7	130,3
2006–2010	837,8	167,6
2011–2015	907,7	181,5

Отриманий ряд динаміки дозволяє визначити основну тенденцію розвитку.

Метод ковзної середньої. При виявленні основної тенденції розвитку методом ковзної середньої інтервали часу укрупнюються по-особливому.

При розрахунку ковзних середніх кожний наступний інтервал утворюється на основі попереднього заміною одного рівня. Оскільки середня \bar{Y}_i належить до середини інтервалу, то доцільно формувати інтервали з непарного числа рівнів первинного ряду.

У разі парного числа рівнів необхідна додаткова процедура центрування усереднення кожної пари значення « \bar{Y}_i ».

Ряд ковзних середніх коротший за первинний на (n-1) рівнів, що потребує уважного ставлення до вибору ширини інтервалу «n». Якщо первинному ряду динаміки притаманна певна періодичність коливань, то інтервал згладжування має бути рівним або кратним періоду коливань.

Розглянемо порядок згладжування методом ковзної середньої на конкретному прикладі:

Приклад:

Зробити згладжування тимчасового ряду використовуючи три членну ковзну середню.

Таблиця 7.3 – Згладжування тимчасового ряду

Місяць звітного періоду	Місячний рівень ряду, Y_i	Рухливі 3-членні підсумки	3 ^x членна ковзна середня
1	2	3	4
1	75		
2	80	237	79
3	82	243	81
4	81	246	82
5	83	248	82,7
6	84	249	83
7	82	246	82
8	80	251	83,7
9	89	249	83
10	80	252	84
11	83	253	84,3
12	90		

7.6 Аналітичне вирівнювання динамічного ряду

Вивчення тренда методом ковзної середньої й іншими механічними методами є емпіричними прийомами попереднього аналізу динаміки ряду. Для вираження тренда кількісною моделлю використовують аналітичне вирівнювання ряду, яке полягає в підборі для даного ряду теоретичної кривої, що виражає основні риси фактичної динаміки.

При аналітичному вирівнюванні ряду динаміки фактичні значення O_t змінюються обчисленими на основі певної функції $Y = f(t)$, яку називають Трендовим рівнянням.

Вибір типу функції ґрунтується на теоретичному аналізі суті явища і характері його динаміки. Суттєвою підмогою при виборі функції є аналіз ланцюгових характеристик динаміки.

1. Якщо ланцюгові абсолютні прирости відносно стабільні, вирівнювання ряду виконується на основі лінійної функції $Y_t = a + b \cdot t$.

2. Якщо відносно стабільними є ланцюгові темпи приросту, та найбільш адекватною такому характеру динаміки є експонента $Y_t = a \cdot b^t$.

3. Якщо характеристики швидкості розвитку зростають (зменшуються) використовують інші функції.

Звичайно аналітичне вирівнювання ряду виконують застосовуючи метод найменших квадратів, основна вимога якого формулюється в такий спосіб:

$$\sum \left(Y_i - \hat{Y}_{it} \right)^2 = \min, \quad (7.19)$$

де Y_i – фактичні значення рівнів ряду;

\hat{Y}_{it} – розрахункові значення рівня ряду.

Параметри Трендових рівнянь визначають розв'язуванням системи нормальних рівнянь.

Для лінійної функції $Y_t = a + bt$ система рівнянь має вигляд:

$$n \cdot a + b \cdot \sum t = \sum Y \quad (7.20)$$

$$a \cdot \sum t + b \cdot \sum t^2 = \sum Yt$$

Система рівнянь спрощується, якщо початок відліку часу ($t = 0$) перенести в середину динамічного ряду. Тоді значення t , розміщені вище середини, будуть від'ємними, а нижче – додатними. При непарному числі членів ряду (наприклад $n = 5$) змінній t надаються значення з інтервалом одиниця: $-2, -1, 0, 1, 2$; при парному:

$-2,5; -1,5; -0,5; +0,5; 1,5; 2,5$. В обох випадках $\sum t = 0$ а система рівнянь набирає вигляду:

$$n \cdot a = \sum Y; \quad b \sum t^2 = \sum Yt. \quad (7.21)$$

Отже, $a = \frac{\sum Y}{n}$ а $b = \frac{\sum Yt}{\sum t^2}$, значення $\sum t^2$ можна визначити за формулою:

$$\sum t^2 = \frac{n(n^2 - 1)}{12}. \quad (7.22)$$

Порядок обчислення параметрів лінійної функції розглянемо на прикладі видобутку вугілля в регіоні (табл. 7.4).

Таблиця 7.4 – Обчислення лінійної функції

Рік	Y_i , млн т	ΔY_{in}	Зміни часу t	$Y_i \cdot t$	t_i^2	Y_{ti}
2000	63,5	–	–3	–190,5	9	63,1
2001	66,8	3,3	–2	–133,6	4	66,9
2002	71,0	4,2	–1	–71,0	1	70,7
2003	74,3	3,3	0	0	0	74,5
2004	76,9	2,6	1	76,9	1	78,3
2005	82,2	5,3	2	164,4	4	82,1
2006	86,8	4,6	3	260,4	9	85,9
Разом	521,5	–	0	106,6	28	521,5

Ланцюгові абсолютні прорости $\Delta Y_{\text{л}}$ практично стабільні, тому тенденцію можна описати лінійною функцією. Параметри трендового рівняння становлять:

$$a = \sum Y_i : n = 521,5 : 7 = 74,5;$$

$$b = \sum Y_i t : \sum t^2 = 106,6 : 28 = 3,8, \text{ де } t^2 = 2 = \sum t_i^2 \text{ або } t^2 = \frac{7(7^2 - 1)}{12} = 28$$

Лінійний тренд буде мати такий вигляд:

$$Y_t = 74,5 + 3,8t .$$

Це означає, що середній рівень видобутку вугілля становить 74,5 млн т а середньорічний приріст – 3,8 млн т.

Обчислимо теоретичні рівні ряду для кожного року за Трендовим рівнянням. Результати розрахунків наведено в останній графі таблиці. Суми фактичних рівнів $\sum Y_i$ і теоретичних $\sum Y_{ti}$ однакові і становлять 521,5 млн т.

Продовження виявленої тенденції за межі ряду динаміки називають екстраполяцією тренду. Це один із методів статистичного прогнозування. Припускаючи, що умови, в яких формувалась тенденція видобутку вугілля, найближчим часом не зміняться, то в 2005 році очікуваний видобуток вугілля досягне

$$Y_{t+1} = 85,9 + 3,8 \cdot 1 = 89,7 \text{ млн т.}$$

Метод екстраполяції дає точковий прогноз. На практиці, як правило, визначають довірчі межі прогнозного рівня

$$Y_{t+v} \pm \Delta, \text{ де } \Delta \text{ – гранична похибка прогнозу.}$$

7.7 Інтерполяція і екстраполяція. Аналіз сезонних коливань

Під час аналізу рядів динаміки доводиться стикатися з такими випадками, коли в рядах відсутні дані про їхні рівні за той або інший період. Такі дані можуть бути відсутні або всередині ряду, або спочатку чи в кінці його.

Приблизне визначення відсутніх рівнів усередині одноякісного періоду, коли відомі рівні, що лежать по обидві сторони невідомого, називають *інтерполяцією ряду динаміки*. Приблизне визначення невідомих рівнів, що лежать за його межами, тобто в майбутньому (або в минулому), називають *екстраполяцією ряду динаміки*. Відповідно екстраполювання може здійснюватися як у бік майбутнього (перспективна екстраполяція), так і у бік минулого (ретроспективна екстраполяція). По суті, екстраполяція являє собою продовження ряду динаміки на основі виявленої закономірності зміни рівнів за досліджуваний відрізок часу.

Інтерполяцію (як і екстраполяцію) здійснюють виходячи з припущення, що зміни в межах періоду, що виражають закономірність розвитку, відносно стійкі, тобто що ні виявлена тенденція, ні її характер не зазнали і не зазнають суттєвих змін у тому проміжку часу, рівні якого нам невідомі.

Щоб мати досить надійні результати обчислення відсутніх рівнів, інтерполяцію та екстраполяцію слід проводити в межах однорідних періодів, яким властива одна закономірність розвитку.

Інтерполяцію і екстраполяцію ряду динаміки можна проводити різними способами. Найпростішим способом є використання середніх характеристик досліджуваного ряду динаміки: середнього абсолютного приросту (при стабільних ланцюгових абсолютних приростах) і середнього коефіцієнту зростання (при стабільних темпах зростання). Однак визначення відсутніх рівнів ряду динаміки, і особливо при екстраполяції, найчастіше пов'язують з аналітичним вирівнюванням рядів способом найменших квадратів, який дає точніші результати. При цьому для виходу за межі періоду, для якого знайдена залежність від часу, досить продовжити значення незалежної змінної – часу.

Аналіз сезонних коливань. Фактичні рівні динамічних рядів варіюють, відхиляючись від основної тенденції розвитку. В одних рядах коливання мають систематичний, закономірний характер, повторюються через певні інтервали часу, в інших – не мають такого характеру і тому називаються випадковими. У конкретному ряді можуть поєднуватися систематичні та випадкові коливання.

Найпростішою оцінкою систематичних коливань є коефіцієнти нерівномірності, які обчислюють відношенням максимального і мінімального рівнів динамічного ряду до середнього. Чим більша різниця між цими двома коефіцієнтами, тим більша нерівномірність процесу.

Наприклад, споживання електроенергії за добу становить 9 600 кВт. год, у середньому $96\ 006 : 24 = 400$ кВт. год.

Найбільший рівень споживання електроенергії в період від 21 до 22 години – 600 кВт. год, найменший – у період від 3 до 4 годин – 200 кВт. год.

Коефіцієнти нерівномірності такі:

$$K_{\max} = 600 : 400 = 1,5;$$

$$K_{\min} = 200 : 400 = 0,5.$$

Амплітуда коливань споживання електроенергії складає 100 пунктів $[100 \cdot (1,5 - 0,5)]$, що свідчить про істотну нерівномірність споживання електроенергії.

Окремим соціально-економічним процесам притаманні сезонні піднесення і спади. Наприклад, коливання попиту на товари, нерівномірне завантаження транспорту, виробництво сільськогосподарської продукції тощо. Такі коливання виявляються на основі рядів щомісячних або щоквартальних даних.

Характер сезонних коливань описується «сезонною хвилею», яку утворюють індекси сезонності. Однак місячні рівні за один рік можуть бути нетиповими через вплив випадкових причин. Тому на практиці індекси сезонності визначають за місячними даними за кілька років (три роки і більше). В цьому разі для кожного місяця встановлюють середню величину рівня за кілька років (наприклад, три роки), далі з них розраховують середньомісячний

рівень для всього ряду. Після цього кожен середньомісячний рівень порівнюють з середньомісячним річним рівнем за кілька років, а знайдений результат перемножують на сто процентів.

Як характеристики сезонності можуть бути використані показники варіації: середнє квадратичне відхилення і коефіцієнт варіації. Чим менша величина цих показників, тим меншою є сезонність досліджуваного явища.

ТЕМА 8 ІНДЕКСНИЙ МЕТОД АНАЛІЗУ

8.1 *Загальне поняття про статистичні індекси. Види індексів.*

8.2 *Індивідуальні індекси: методика визначення і економічний зміст.*

8.3 *Агрегатний індекс як основна форма загального індексу.*

8.4 *Середньозважені індекси.*

8.5 *Індекси середніх величин (змінного складу, постійного складу, структурних зрушень).*

8.6 *Факторний індексний аналіз.*

8.1 Загальне поняття про статистичні індекси. Види індексів

Індексний метод – один із найпоширеніших статистичних прийомів дослідження соціально-економічних явищ і процесів. Основне призначення статистичних індексів – кількісно охарактеризувати відносну зміну складних економічних явищ у часі і просторі.

У статистиці під *індексом* розуміють специфічну відносну величину, яка характеризує зміну показника у часі та просторі. Індекси можуть визначатися у неоднорідній сукупності, що складається з елементів, які безпосередньо не можна сумувати.

За допомогою індексів вирішують такі завдання статистичного аналізу:

– визначають середній процент зміни показника у часі в цілому по сукупності або окремій групі;

- визначають середній процент зміни середнього значення показника;
- здійснюють порівняння показника у просторі;
- оцінюють вплив окремих факторів на зміну показника у часі або просторі.

Залежно від охоплення одиниць сукупності індекси поділяються на індивідуальні, групові (субіндекси) та загальні. За методикою визначення та економічним змістом індекси бувають агрегатними, середньозваженими, індексами середніх величин, територіальними.

В залежності від бази порівняння індекси поділяються на ланцюгові та базисні, а за економічним змістом – на індекси динаміки, планового завдання, виконання плану.

При визначенні індексу спочатку записується (будується) його формула, згідно якої робляться обчислення. Показник, який досліджується з допомогою індексу, називається індексованим.

Для запису формули індексу використовують умовні позначення індексованих показників, які поділяються на якісні, кількісні та об'ємні. Якісні показники характеризують рівень ознаки у одиниці сукупності, а саме P – ціна одиниці товару; Z – собівартість одиниці продукції; f – зарплата одного працівника; y – врожайність з 1 га.

Кількісні показники характеризують чисельність сукупності, а саме: q – кількість (фізичний обсяг) товару або продукції; T – чисельність працівників; S – посівні площі тощо.

Об'ємні показники являють собою обсяг ознаки і завжди є добутком якісного та одного кількісного показників, наприклад, $pq = p \cdot q$ – товарообороті або вартість продукції; $zq = z \cdot q$ – витрати на виробництво продукції; $fT = f \cdot T$ – фонд заробітної плати працівників; $yS = y \cdot S$ – валовий збір культур тощо.

Для позначення часу використовують під строкові символи: 0 – попередній або базисний період, 1 – наступний або звітний період. Якщо відомі значення показників більш ніж за два періоди, використовують порядкові номери 1,2,3 тощо.

8.2 Індивідуальні індекси: методика визначення і економічний зміст

Індивідуальні індекси визначаються по окремих одиницях статистичної сукупності та характеризують зміну індивідуальних значень індексованого показника. Ці індекси прийнято позначати і з підстроковим позначенням індексованого показника. Індивідуальні індекси якісних показників визначаються за формулами:

$$i_p = \frac{p_1}{p_0}; i_z = \frac{z_1}{z_0}; i_f = \frac{f_1}{f_0}; i_y = \frac{y_1}{y_0}. \quad (8.1)$$

Індивідуальні індекси кількісних показників мають вигляд:

$$i_q = \frac{q_1}{q_0}; i_T = \frac{T_1}{T_0}; i_S = \frac{S_1}{S_0}. \quad (8.2)$$

Індивідуальні індекси об'ємних показників можна записати так:

$$i_{pq} = \frac{p_1 q_1}{p_0 q_0} = i_p * i_q; i_{zq} = \frac{z_1 q_1}{z_0 q_0} = i_z * i_q; i_{fT} = \frac{f_1 T_1}{f_0 T_0} = i_f * i_T. \quad (8.3)$$

Варто мати на увазі, що взаємозв'язок між індивідуальними індексами можна використовувати для розрахунку третього індексу по двох відомих, наприклад: індекс фонду заробітної плати становить 1,756, а індекс заробітної плати визначається в такий спосіб:

$$i_f = \frac{i_{fT}}{i_T} = \frac{1,756}{0,8} = 2,197.$$

Наведемо приклад розрахунку та економічної інтерпретації індивідуальних індексів (табл. 8.1).

Таблиця 8.1 – Розрахунки індивідуальних індексів

Продукція	Собівартість, грн		Кількість, шт.		z ₀ q ₀	z ₁ q ₁
	I квартал z ₀	II квартал z ₁	I квартал q ₀	II квартал q ₁		
А	80	90	100	120	8 000	10 800
Б	60	60	70	80	4 200	4 800
В	70	65	140	100	9 800	6 500

Визначимо індивідуальні індекси собівартості:

$$i_z^A = \frac{z_1}{z_0} = \frac{90}{80} = 1,125 \quad i_z^B = \frac{60}{60} = 1 \quad i_z^B = \frac{65}{70} = 0,929.$$

Отже, по продукції А собівартість зросла на 12,5 %, по продукції Б залишилася без змін, а по продукції В – зменшилася на 7,1 %.

Індивідуальні індекси кількості (фізичного обсягу) продукції:

$$i_q^A = \frac{q_1}{q_0} = \frac{120}{100} = 1,2 \quad i_q^B = \frac{80}{70} = 1,143 \quad i_q^B = \frac{100}{140} = 0,714.$$

Таким чином, по продукції А виробництво зросло на 20 %, по продукції Б – на 14,3 %, а по продукції В – зменшилася на 28,6 %.

Індивідуальні індекси витрат на виробництво продукції:

$$i_{zq}^A = \frac{z_1 q_1}{z_0 q_0} = \frac{10800}{8000} = 1,350 \quad i_{zq}^B = \frac{4800}{4200} = 1,143 \quad i_{zq}^B = \frac{6500}{9800} = 0,663.$$

Витрати на виробництво продукції А зросли на 35 %, по продукції Б – на 14,3 %, а по продукції В – скоротилися на 33,7 %.

Взаємозв'язок індексів:

$$i_{zq} = i_z \times i_q = 1,125 \times 1,2 = 1,350 \quad i_{zq} = 0,929 \times 0,714 = 0,663$$

$$i_{zq} = 1 \times 1,143 = 1,143$$

8.3 Агрегатний індекс як основна форма загального індексу

Агрегатні індекси відносяться до загальних індексів, які характеризують середню зміну індексованого показника у часі та просторі. В агрегатних індексах у чисельнику та знаменнику знаходяться суми добутків двох взаємопов'язаних показників, один з яких – якісний, а другий – кількісний. Позначаються агрегатні індекси літерою I з підстроковим символом індексованого показника.

В залежності від правил побудови агрегатний індексів розрізняють індексні системи Ласпейреса, Пааше та Фішера. В статистиці України використовується комбінована система агрегатних індексів, яка будується за наступними правилами.

В агрегатних індексах якісних показників індексований показник у чисельнику береться за звітний період, а у знаменнику – за базисний, а співмножник (кількісний показник) у чисельнику і знаменнику фіксується на рівні звітного періоду (метод Пааше).

Наприклад:

– агрегатний індекс ціни:

$$I_p = \frac{\sum p_1 q_1}{\sum p_0 q_1}; \quad (8.4)$$

– агрегатний індекс собівартості:

$$I_x = \frac{\sum z_1 q_1}{\sum z_0 q_1}; \quad (8.5)$$

– агрегатний індекс зарплати:

$$I_f = \frac{\sum f_1 T_1}{\sum f_0 T_1}; \quad (8.6)$$

– агрегатний індекс урожайності:

$$I_y = \frac{\sum y_1 S_1}{\sum y_0 S_1}. \quad (8.7)$$

Таким чином, у чисельнику агрегатного індексу якісного показника знаходиться сума значень об'ємного показника за звітний період, а у знаменнику – розрахункові значення об'ємного показника у звітному періоді при умові збереження якісного показника на базисному рівні.

В агрегатних індексах кількісних показників індексований індексований показник у чисельнику береться за звітний період, а у знаменнику за базисний, а співмножник (якісний показник) у чисельнику і знаменнику фіксується, тобто береться однаковим, на рівні базисного періоду (метод Ласпейреса).
Наприклад:

– індекс фізичного обсягу:

$$I_q = \frac{\sum p_0 q_1}{\sum p_0 q_0} \quad (8.8)$$

або

$$I_q = \frac{\sum z_0 q_1}{\sum z_0 q_0};$$

– індекс чисельності працівників:

$$I_T = \frac{\sum f_0 T_1}{\sum f_0 T_0}; \quad (8.9)$$

– індекс посівних площ:

$$I_s = \frac{\sum y_0 S_1}{\sum y_0 S_0}. \quad (8.10)$$

Отже, у знаменнику агрегатних індексів кількісних показників знаходиться сума значень об'ємного показника за базисний період, а у

чисельнику – розрахункові значення об’ємного показника при умові збереження якісного показника на базисному рівні.

В агрегатних індексах об’ємних показників у чисельнику знаходиться сума добутків якісного і кількісного показників за звітний період, а у знаменнику – за базисний, тобто індексуються обидва показники.

Наприклад:

– індекс товарообороту або вартості продукції:

$$I_{pq} = \frac{\sum p_1 q_1}{\sum p_0 q_0}, \quad (8.11)$$

– індекс витрат на виробництво продукції:

$$I_{zq} = \frac{\sum z_1 q_1}{\sum z_0 q_0}, \quad (8.12)$$

– індекс фонду заробітної плати:

$$I_{fT} = \frac{\sum f_1 T_1}{\sum f_0 T_0}, \quad (8.13)$$

– індекс валового збору (урожаю):

$$I_{yS} = \frac{\sum y_1 S_1}{\sum y_0 S_0}. \quad (8.14)$$

Отже, в чисельнику цих індексів сумуються значення об’ємного показника за звітний період, а у знаменнику – за базисний.

Між агрегатними індексами показників існує взаємозв’язок: агрегатний індекс об’ємного показника дорівнює добутку агрегатних індексів якісного та кількісного показників. Наприклад, $I_{pq} = I_p \times I_q$, $I_{zq} = I_z \times I_q$, $I_{fT} = I_f \times I_T$, $I_{yS} = I_y \times I_S$.

Наведемо приклад розрахунку та економічної інтерпретації агрегатних індексів (табл. 8.2).

Таблиця 8.2 – Розрахунок економічної інтерпретації агрегатних індексів

Товар	Ціна, грн		Кількість, шт.		p_0q_0	p_0q_1	p_1q_1
	I кв. (p_0)	II кв. (p_1)	I кв. (q_0)	II кв. (q_1)			
А	170	210	70	80	11 900	13 600	16 800
Б	190	280	90	60	17 100	11 400	16 800
В	150	180	40	20	6 000	3 000	3 600
					35 000	28 000	37 200

Загальний агрегатний індекс ціни:

$$I_p = \frac{\sum p_1q_1}{\sum p_0q_1} = \frac{37200}{28000} = 1,329.$$

Загальний агрегатний індекс фізичного обсягу:

$$I_q = \frac{\sum p_0q_1}{\sum p_0q_0} = \frac{28000}{35000} = 0,800.$$

Загальний агрегатний індекс товарообороту:

$$I_{pq} = \frac{\sum p_1q_1}{\sum p_0q_0} = \frac{37200}{35000} = 1,063.$$

Отже, по трьох товарах ціни в середньому зросли на 32,9 %, кількість проданих одиниць (фізичний обсяг) зменшився в середньому на 20 %, а товар оборот зріс на 6,3 %.

На основі агрегатних індексів можна визначити як загальний приріст об'ємного показника в абсолютному виразі, так і прирости за рахунок зміни якісного та кількісного показників. Для цього від чисельника відповідного індексу необхідно відняти знаменник. Наприклад:

– загальний приріст товарообороту:

$$\Delta = \sum p_1q_1 - \sum p_0q_0; \quad (8.15)$$

приріст товарообороту за рахунок зміни цін:

$$\Delta_p = \sum p_1q_1 - \sum p_0q_1; \quad (8.16)$$

– приріст товарообороту за рахунок зміни фізичного обсягу:

$$\Delta_q = \sum p_0 q_1 - \sum p_0 q_0 . \quad (8.17)$$

Приклад розрахунків на основі наведених вище даних:

– загальний приріст товарообороту:

$$\Delta = \sum p_1 q_1 - \sum p_0 q_0 = 37200 - 35000 = 2200 \text{ грн.};$$

– приріст товарообороту за рахунок зміни цін:

$$\Delta_p = \sum p_1 q_1 - \sum p_0 q_1 = 37200 - 28000 = 9200 \text{ грн.};$$

– приріст товарообороту за рахунок зміни фізичного обсягу:

$$\Delta_q = \sum p_0 q_1 - \sum p_0 q_0 = 28000 - 35000 = -7000 \text{ грн.}$$

Агрегатні індекси можна визначати як ланцюгові та базисні. В ланцюгових індексах індексований показник береться за суміжні періоди часу (наступний і попередній), а у базисних – у знаменнику беруться значення індексованого показника за базисний період. Таким чином, перші індекси характеризують середню зміну індексованого показника за одиницю часу (у поточному періоді порівняно з попереднім), а другі – за певний період часу (у поточному періоді порівняно з базисний).

8.4 Середньозважені індекси

Агрегатні індекси кількісних та якісних показників можна перетворити у середньозважені індекси – середньоарифметичний або середньогармонійний відповідно. Середньозважені індекси використовуються у тих випадках, коли відомі індивідуальні індекси якісних або кількісних показників. По своїй суті ці індекси є середніми зваженими величинами, у яких варіантами виступають значення індивідуальних індексів досліджуваного показника.

Агрегатні індекси кількісних показників можна перетворити у середньоарифметичні індекси наступним чином:

$$I_q = \frac{\sum p_0 q_1}{\sum p_0 q_0} = \frac{\sum i_q * p_0 q_0}{\sum p_0 q_0}, \text{ оскільки } i_q = \frac{q_1}{q_0}, \quad q_1 = i_q * q_0;$$

$$I_T = \frac{\sum f_0 T_1}{\sum f_0 T_0} = \frac{\sum i_T * f_0 T_0}{\sum f_0 T_0}, \text{ оскільки } i_T = \frac{T_1}{T_0}, \quad T_1 = i_T * T_0. \quad (8.15)$$

Отже, середньоарифметичний індекс доцільно використовувати у тому випадку, якщо відомі індивідуальні індекси кількісного показника і значення об'ємного показника за базисний період. За своїм економічним змістом ці індекси аналогічні агрегатним.

Приклад розрахунку середньоарифметичного індексу посівних площ подано у таблиці 8.3.

Таблиця 8.3 – Середньоарифметичний індекс

Культури	Валовий збір у базисному році, ц ($y_0 S_0$)	Зміна посівних площ порівняно з попереднім роком, %	Індивідуальні індекси посівних площ (I_s)
Пшениця	127 900	+5	1,050
Жито	34 400	-10	0,900
Ячмінь	20 500	+16	1,160
Всього	182 800	x	x

Середньоарифметичний індекс посівних площ:

$$I_s = \frac{\sum i_s * y_0 S_0}{\sum y_0 S_0} = \frac{1,050 * 127900 + 0,900 * 34400 + 1,160 * 20500}{182800} = 1,034.$$

Отже, по трьох культурах посівні площі зросли в середньому на 3,4 %.

Агрегатні індекси якісних показників можна перетворити у середньогармонійні індекси наступним чином:

$$I_p = \frac{\sum p_1 q_1}{\sum p_0 q_1} = \frac{\sum p_1 q_1}{\sum \frac{p_1 q_1}{i_p}}, \text{ оскільки } i_p = \frac{p_1}{p_0}, \text{ то } p_0 = \frac{p_1}{i_p};$$

$$I_x = \frac{\sum z_1 q_1}{\sum z_0 q_1} = \frac{\sum z_1 q_1}{\sum \frac{z_1}{i_x} q_1}, \text{ оскільки } i_x = \frac{z_1}{z_0}, \text{ то } z_0 = \frac{z_1}{i_x}.$$

Середньогармонійні індекси доцільно використовувати в тих випадках, коли відомі індивідуальні індекси якісного показника і значення об'ємного показника у звітному періоді.

Приклад розрахунку середньогармонійного індексу заробітної плати (табл. 8.4).

Таблиця 8.4 – Розрахунок середньогармонійного індексу

Цехи	Фонд заробітної плати у звітному періоді, тис. грн ($f_1 T_1$)	Зміна рівня заробітної плати порівняно з базисним періодом, %	Індивідуальні індекси заробітної плати (I_f)
№1	25,4	+15	1,150
№2	17,3	+7	1,070
№3	19,6	+25	1,250
Всього	62,3	x	x

Середньогармонійний індекс заробітної плати:

$$I_f = \frac{\sum f_1 T_1}{\sum \frac{f_1 T_1}{i_f}} = \frac{62,3}{\frac{25,4}{1,150} + \frac{17,3}{1,070} + \frac{19,6}{1,250}} = 1,156.$$

Можна зробити висновок, що у середньому по трьох цехах заробітна плати у звітному періоді порівняно з базисним зросла на 15,6 %.

На основі середньозважених індексів також можна розрахувати приріст об'ємного показника за рахунок індексованого, для чого від чисельника індексу необхідно відняти його знаменник.

8.5 Індeksi середніх величин (змінного складу, постійного складу, структурних зрушень)

У статистичному аналізі часто буває необхідним дослідити зміну у часі або просторі середнього значення якісного показника, наприклад, ціни, собівартості, урожайності, заробітної плати тощо. У цьому випадку середнє значення показника розраховується як середня арифметична зважена або як відношення обсягу ознаки до чисельності сукупності. Слід мати на увазі, що середній рівень будь-якої ознаки формується під впливом двох факторів – варіацією індивідуальних значень та структури сукупності. Наприклад, середній рівень заробітної плати може зростати за рахунок зростання оплати праці кожного працівника і за рахунок збільшення питомої ваги високооплачуваних працівників.

Для вивчення динаміки середнього значення якісного показника використовується система трьох індексів: індекс змінного складу, індекс постійного складу та індекс структурних зрушень.

Індекс змінного складу характеризує зміну у процентах середнього значення якісного показника у звітному періоді порівняно з базисним під впливом двох чинників разом. Цей індекс складається з двох дробів, причому перший дріб містить значення якісного та кількісного показників у звітному періоді, а другий – у базисному, тобто індекс є відношенням звітного середнього значення показника до базисного. Наприклад:

– індекс ціни змінного складу:

$$I_{\bar{p}} = \frac{\sum p_1 q_1}{\sum q_1} : \frac{\sum p_0 q_0}{\sum q_0} = \frac{\bar{p}_1}{\bar{p}_0}; \quad (8.16)$$

– індекс собівартості змінного складу:

$$I_{\bar{z}} = \frac{\sum z_1 q_1}{\sum q_1} : \frac{\sum z_0 q_0}{\sum q_0} = \frac{\bar{z}_1}{\bar{z}_0}; \quad (8.17)$$

– індекс заробітної плати змінного складу:

$$I_{\bar{f}} = \frac{\sum f_1 T_1}{\sum T_1} : \frac{\sum f_0 T_0}{\sum T_0} = \frac{\bar{f}_1}{\bar{f}_0}. \quad (8.18)$$

Індекс постійного складу показує зміну (у %) середнього значення показника під впливом одного фактора – динаміки його індивідуальних значень. У цьому індексі індексується (змінюється) якісний показник, а кількісний фіксується на рівні звітного періоду. Наприклад:

– індекс ціни постійного складу:

$$I_p = \frac{\sum p_1 q_1}{\sum q_1} : \frac{\sum p_0 q_1}{\sum q_1} = \frac{\sum p_1 q_1}{\sum p_0 q_1}; \quad (8.19)$$

– індекс собівартості постійного складу:

$$I_z = \frac{\sum z_1 q_1}{\sum q_1} : \frac{\sum z_0 q_1}{\sum q_1} = \frac{\sum z_1 q_1}{\sum z_0 q_1}; \quad (8.20)$$

– індекс зарплати постійного складу:

$$I_f = \frac{\sum f_1 T_1}{\sum T_1} : \frac{\sum f_0 T_1}{\sum T_1} = \frac{\sum f_1 T_1}{\sum f_0 T_1}. \quad (8.21)$$

Індекс структурних зрушень показує, на скільки процентів змінилося середнє значення показника під впливом змін у структурі сукупності. У даному випадку індексується кількісний показник, а якісний фіксується на рівні базисного періоду. Наприклад:

– індекс структурних зрушень ціни:

$$I_{cz} = \frac{\sum p_0 q_1}{\sum q_1} : \frac{\sum p_0 q_0}{\sum q_0}; \quad (8.22)$$

– індекс структурних зрушень собівартості:

$$I_{cz} = \frac{\sum z_0 q_1}{\sum q_1} : \frac{\sum z_0 q_0}{\sum q_0}; \quad (8.23)$$

– індекс структурних зрушень зарплати:

$$I_{cz} = \frac{\sum f_0 T_1}{\sum T_1} : \frac{\sum f_0 T_0}{\sum T_0}. \quad (8.24)$$

Між названими трьома індексами існує взаємозв'язок: індекс змінного складу дорівнює добутку індексу постійного складу та індексу структурних зрушень. Отже,

$$I_{\bar{p}} = I_p \times I_{cz}; \quad I_{\bar{z}} = I_z \times I_{cz}; \quad I_{\bar{f}} = I_f \times I_{cz}.$$

На основі вказаних індексів можна визначити приріст середнього значення показника в абсолютному виразі загальний та за рахунок окремих факторів – якісного та кількісного (структурного). Наприклад:

– приріст середньої ціни (загальний):

$$\Delta = \frac{\sum p_1 q_1}{\sum q_1} - \frac{\sum p_0 q_0}{\sum q_0} = \bar{p}_1 - \bar{p}_0; \quad (8.25)$$

– приріст середньої ціни за рахунок зміни цін:

$$\Delta p = \frac{\sum p_1 q_1}{\sum q_1} - \frac{\sum p_0 q_1}{\sum q_1} = \bar{p}_1 - \frac{\sum p_0 q_1}{\sum q_1}; \quad (8.26)$$

– приріст середньої ціни за рахунок структурних зрушень:

$$\Delta_{сз} = \frac{\sum p_0 q_1}{\sum q_1} - \frac{\sum p_0 q_0}{\sum q_0} = \frac{\sum p_0 q_1}{\sum q_1} - \bar{p}_0. \quad (8.27)$$

Наведемо приклад обчислення та економічної інтерпретації індексів середньої собівартості (табл. 8.5).

Таблиця 8.5 – Інтерпретації індексів середньої собівартості

Продукція	Собівартість, грн		Кількість, тис шт.		z ₀ q ₀	z ₀ q ₁	z ₁ q ₁
	z ₀	z ₁	q ₀	q ₁			
А	30	37	100	135	3 000	40 550	4 995
Б	41	40	70	90	2 870	3 690	3 600
В	50	53	140	170	7 000	8 500	9 010
Разом	x	x	310	395	12 870	16 240	17 605

Індекс собівартості змінного складу:

$$I_z = \frac{\sum z_1 q_1}{\sum q_1} : \frac{\sum z_0 q_0}{\sum q_0} = \frac{17605}{395} \div \frac{12870}{310} = 44,57 \div 41,52 = 1,073.$$

Індекс собівартості постійного складу:

$$I_z = \frac{\sum z_1 q_1}{\sum q_1} : \frac{\sum z_0 q_1}{\sum q_1} = \frac{17605}{395} \div \frac{16240}{395} = 44,57 \div 41,11 = 1,084.$$

Індекс структурних зрушень:

$$I_{сз} = \frac{\sum z_0 q_1}{\sum q_1} : \frac{\sum z_0 q_0}{\sum q_0} = 41,11 \div 41,52 = 0,990.$$

Таким чином, середня собівартість зросла на 7,3 % , в тому числі під впливом зміни індивідуальних значень показника – на 8,4 %, а за рахунок структурних зрушень середнє значення зменшилося на 1 %.

Приріст середньої собівартості:

- загальний: $\Delta = \bar{Z}_1 - \bar{Z}_0 = 44,57 - 41,52 = 3,05$ грн.;
- за рахунок зміни собівартості: $\Delta z = 44,57 - 41,11 = 3,46$ грн.;
- за рахунок структурних зрушень: $\Delta cз = 41,11 - 41,52 = -0,41$ грн.

8.6 Факторний індексний аналіз

Факторний індексний аналіз використовується для вивчення впливу окремих факторних показників на результативний показник з допомогою системи взаємозв'язаних індексів. При цьому результативний показник функціонально залежить від факторних показників та дорівнює їх добутку: $y = a \cdot b \cdot c \cdot d$.

На відміну від кореляційно-регресійного аналізу, вплив кожного фактора розглядається ізольовано, тобто без врахування взаємодії факторів. Слід мати на увазі, що порядок співмножників повинен бути таким, щоб кожний добуток мав економічний зміст.

При побудові індексів використовується таке правило: індексований показник у чисельнику за звітний період, у знаменнику – за базисний, показники, які знаходяться перед індексованим показником фіксуються на рівні звітного періоду, а ті, що розташовані після індексованого показника – на рівні базисного періоду.

Отже,

$$I_y = \frac{\sum a_1 b_1 c_1 d_1}{\sum a_0 b_0 c_0 d_0}; \quad I_a = \frac{\sum a_1 b_0 c_0 d_0}{\sum a_0 b_0 c_0 d_0}; \quad I_b = \frac{\sum a_1 b_1 c_0 d_0}{\sum a_1 b_0 c_0 d_0}; \quad I_c = \frac{\sum a_1 b_1 c_1 d_0}{\sum a_1 b_1 c_0 d_0}; \quad I_d = \frac{\sum a_1 b_1 c_1 d_1}{\sum a_1 b_1 c_1 d_0}.$$

Між індексами існує взаємозв'язок: $I_y = I_a \cdot I_b \cdot I_c \cdot I_d$.

На основі обчисленої системи індексів можна визначити загальний абсолютний приріст результативного показника та факторні прирости, зумовлені впливом кожного фактора зокрема. Ці прирости визначаються як різниця між чисельником та знаменником відповідного індексу:

$$\begin{aligned}\Delta y &= \sum a_1 b_1 c_1 d_1 - \sum a_0 b_0 c_0 d_0; \\ \Delta a &= \sum a_1 b_0 c_0 d_0 - \sum a_0 b_0 c_0 d_0; \\ \Delta b &= \sum a_1 b_1 c_0 d_0 - \sum a_1 b_0 c_0 d_0; \\ \Delta c &= \sum a_1 b_1 c_1 d_0 - \sum a_1 b_1 c_0 d_0; \\ \Delta d &= \sum a_1 b_1 c_1 d_1 - \sum a_1 b_1 c_1 d_0; \\ \Delta y &= \Delta a + \Delta b + \Delta c + \Delta d.\end{aligned}$$

Приклад розрахунку та економічної інтерпретації індексів (табл. 8.6).

Таблиця 8.6 – Розрахунок економічної інтерпретації індексів

Продукція	Витрати сировини на одиницю продукції (питомі витрати), кг		Ціна 1 кг сировини, грн		Кількість виробленої продукції, шт.	
	I кв. (m ₀)	IV кв. (m ₁)	I кв. (p ₀)	IV кв. (p ₁)	I кв. (q ₀)	IV кв. (q ₁)
А	20	22	100	160	40	50
Б	17	13	220	270	100	140
В	40	45	350	360	200	170

Результативний показник – витрати на виробництво (у):

$$y = m \cdot p \cdot q.$$

Індекс витрат сировини на одиницю продукції (питомих витрат):

$$I_m = \frac{\sum m_1 p_0 q_0}{\sum m_0 p_0 q_0} = \frac{22 \times 100 \times 40 + 13 \times 220 \times 100 + 45 \times 350 \times 200}{20 \times 100 \times 40 + 17 \times 220 \times 100 + 40 \times 350 \times 200} = \frac{3524000}{3254000} = 1,083.$$

Індекс цін на сировину:

$$I_p = \frac{\sum m_1 p_1 q_0}{\sum m_1 p_0 q_0} = \frac{22 \times 160 \times 40 + 13 \times 270 \times 100 + 45 \times 360 \times 200}{3524000} = \frac{3731800}{3524000} = 1,059.$$

Індекс кількості (фізичного обсягу) продукції:

$$I_q = \frac{\sum m_1 p_1 q_1}{\sum m_1 p_1 q_0} = \frac{22 \times 100 \times 45 + 13 \times 220 \times 140 + 45 \times 360 \times 170}{3731800} = \frac{3421400}{3731800} = 0,917.$$

Індекс витрат на виробництво:

$$I_y = \frac{\sum m_1 p_1 q_1}{\sum m_0 p_0 q_0} = \frac{3421400}{3254000} = 1,051.$$

Отже, витрати на виробництво зросли на 5,1 %, в тому числі за рахунок збільшення витрат сировини на одиницю продукції – на 8,3 %, за рахунок зростання цін – на 5,9 %, а під впливом зменшення кількості виробленої продукції витрати зменшилися на 8,3 %.

Визначимо абсолютний приріст витрат на виробництво:

$$\Delta y = \sum m_1 p_1 q_1 - \sum m_0 p_0 q_0 = 3421400 - 3254000 = 167400 (\text{грн.}).$$

Прирости за рахунок зміни:

– витрати сировини на одиницю продукції:

$$\Delta m = \sum m_1 p_0 q_0 - \sum m_0 p_0 q_0 = 3524000 - 3254000 = 270000 (\text{грн.});$$

– цін на сировину:

$$\Delta p = \sum m_1 p_1 q_0 - \sum m_1 p_0 q_0 = 3731800 - 3524000 = 207800 (\text{грн.});$$

– кількості (фізичного обсягу) продукції:

$$\Delta q = \sum m_1 p_1 q_1 - \sum m_1 p_1 q_0 = 3421400 - 3731800 = -310400 (\text{грн.}).$$

При цьому

$$\Delta y = 270000 + 207800 - 310400 = 167400 (\text{грн}).$$

Отже, факторний індексний аналіз є ефективним методом дослідження мультиплікативних детермінованих зв'язків при кількості факторних ознак більше двох.

ТЕМА 9 ВИБІРКОВЕ СПОСТЕРЕЖЕННЯ

9.1 Вибіркове спостереження, причини і умови його застосування.

9.2 Способи і методи відбору одиниць у вибірккову сукупність.

9.3 Помилки вибіркового спостереження: зміст та методика розрахунку.

9.4 Визначення обсягу вибірки та способи поширення результатів вибіркового спостереження.

9.1 Вибіркове спостереження, причини і умови його застосування

Статистичне спостереження за охопленням одиниць обстеження сукупності поділяють на суцільне і несучільне. З усіх видів несучільного спостереження в практиці статистичних досліджень найбільше визнання і застосування дістало вибірккове спостереження.

Вибіркове спостереження – це такий вид несучільного спостереження, при якому обстежуються не всі елементи сукупності, що вивчається, а лише певним чином відібрана їх частина. Сукупність, з якої вибирають елементи для обстеження, називають генеральною, а сукупність, яку безпосередньо обстежують, – вибірковою. Статистичні характеристики вибіркової сукупності розглядаються як оцінка відповідних характеристик генеральної сукупності.

Безумовно, що суцільне спостереження забезпечує найбільш повну інформацію про загальну кількість одиниць, про середні і відносні показники

досліджуваної сукупності. Однак на практиці є чимало випадків, коли суцільне спостереження недоцільне або зовсім неможливе. При вивченні певного кола соціально-економічних явищ вибіркоче спостереження єдино можливе. Це стосується передусім перевірки якості продукції (жирності молока, чистоти та вологості зерна, міцності пряжі, тривалості горіння електроламп тощо).

Практика вибіркових спостережень досить різноманітна. Це обстеження домогосподарств, маркетингові дослідження, аудиторські перевірки великих фірм, вивчення громадської думки тощо. Часом вибіркоче спостереження поєднується із суцільним. Наприклад, при перепису населення певна його частина обстежується за більш широкою програмою, ніж основна маса.

У процесі вибіркового спостереження вирішуються такі завдання:

- визначається мета спостереження;
- складається план і програма спостереження;
- визначається вид та спосіб відбору, чисельність вибірки;
- проведення відбору, тобто формування вибіркової сукупності;
- реєстрація ознак;
- розраховуються вибіркочві характеристики;
- визначаються помилки репрезентативності та поширюються результати

на генеральну сукупність.

Уся сукупність одиниць, з яких виконується відбір для подальшого обстеження, називається генеральною сукупністю, а її чисельність позначається N . Частина генеральної сукупності, що попала у вибірку має назву вибіркової сукупності (її чисельність позначається n). Відношення n/N називається часткою відбору, а $100 \times n/N$ – процентом відбору.

Як вибіркова, так і генеральна сукупності характеризуються рядом показників, що відповідно називаються вибірковими та генеральними характеристиками. Розбіжність між ними, яка об'єктивно виникає внаслідок несцільності спостереження, має назву помилки репрезентативності. Помилки репрезентативності, на відміну від помилок реєстрації, можна оцінити (тобто визначити їх розмір), що дозволяє врахувати їх при поширенні результатів

вибіркового спостереження на генеральну сукупність.

Генеральні характеристики:

N – чисельність генеральної сукупності;

\bar{X} – середнє значення ознаки у генеральній сукупності (генеральна середня);

σ_2 – дисперсія;

p – генеральна частка;

σ_{2p} – дисперсія альтернативної ознаки $\sigma_{2p} = p(1 - p)$.

Вибіркові характеристики:

n – чисельність вибіркової сукупності;

\bar{x} – середнє значення ознаки у вибірковій сукупності (вибіркова середня);

σ_2 – дисперсія;

W – вибірка частка;

σ_{2w} – дисперсія альтернативної ознаки $\sigma_{2w} = W(1 - W)$.

Доведено, що для достатньо великих сукупностей генеральна та вибірка дисперсії співпадають, тому на практиці для розрахунків помилок репрезентативності використовують вибірку дисперсію.

9.2 Види та способи відбору одиниць у вибірку сукупність

В залежності від того, скільки разів відібрані для обстеження одиниці приймають участь у відборі, розрізняють повторний та неповторний відбір. При повторному відборі обстежені одиниці «повертаються» у генеральну сукупність і знову приймають участь у відборі. При без повторному відборі одиниці, що попали у вибірку, більше не приймають участі у відборі, таким чином кожна одиниця може бути відбраною лише один раз

При формуванні вибіркової сукупності використовують також наступні види відбору:

– індивідуальний, при якому у вибірку сукупність вибирають по одній

одиниці з генеральної сукупності;

- груповий або серійний , при якому вибирається група (серія) одиниць;
- комбінований, тобто сполучення перших двох видів відбору.

Розрізняють чотири основних способи формування вибіркової сукупності:

1) власне випадковий відбір (повторний чи безповторний), при якому вибірка сукупність формується виключно випадково (методом жеребкування, за таблицями чисел тощо);

2) механічний (систематичний) відбір, при якому у вибірку сукупність попадають одиниці з певними порядковими номерами. При цьому всі одиниці генеральної сукупності спочатку впорядковуються та їм присвоюються порядкові номери. Далі визначається пропорція відбору та крок. Наприклад, пропорція відбору $1/20$, отже крок (різниця між порядковими номерами) становить 20. Далі з першої групи випадковим чином визначається перший порядковий номер, а наступні – шляхом додавання кроку відбору. Наприклад, з перших 20 одиниць обрано 7-му, тоді наступні одиниці – 27, 47, 67 і т.д. Цей спосіб відбору є безповторним;

3) типовий відбір передбачає , що генеральна сукупність поділяється на однорідні групи і з кожної групи випадковим або механічним способом формується вибірка сукупність. Якщо з кожної групи відбирається однаковий процент одиниць, типовий відбір називається пропорційним, а якщо однакова кількість одиниць – непропорційним. Типовий відбір може бути повторним і безповторним;

4) серійний відбір, при якому у вибірку сукупність відбираються групи одиниць (серії), які надалі обстежуються суцільно.

У статистичній практиці застосовується відбір у часі, наприклад, моментне спостереження, що передбачає реєстрацію ознак на певний момент часу, як правило, через рівні інтервали.

9.3 Помилки вибіркового спостереження:

зміст та методика розрахунку

Помилки вибіркового спостереження виникають внаслідок обстеження частини сукупності, або при порушенні правил формування вибіркової сукупності. Вони проявляються у розбіжності між генеральними і вибірковими характеристиками. Ці помилки поділяються на випадкові та систематичні. Випадкові помилки (помилки репрезентативності) можна оцінити із заданим рівнем імовірності. Систематичні помилки, що виникають внаслідок невдалого відбору, оцінюванню не підлягають, тому їх не можна враховувати.

Випадкові помилки вибіркового спостереження залежать від двох факторів:

- чисельність вибіркової сукупності (або частки чи процента відбору);
- варіації ознаки.

Доведено, що чим більшою є чисельність вибіркової сукупності (частка відбору), тим меншою є помилка вибіркового спостереження, і навпаки, чим більшою є варіація ознаки, тим більша й помилка.

Залежність величини помилки вибіркового спостереження від названих факторів виражається через формули граничної помилки вибірки:

а) при повторному випадковому відборі гранична помилка визначається:

- для середньої

$$\Delta_x = t \sqrt{\frac{\sigma^2}{n}} ; \quad (9.1)$$

- для частки

$$\Delta_w = t \sqrt{\frac{w(1-w)}{n}} . \quad (9.2)$$

В наведених формулах t – коефіцієнт довіри, який залежить від рівня ймовірності наступним чином:

рівень ймовірності (λ)	коефіцієнт довіри (t)
0,863	1
0,954	2
0,997	3
0,950	1,96
0,990	2,58
0,999	3,28

Приклад розрахунку помилок середньої та частки при випадковому повторному відборі. Обстежено 200 одиниць продукції, з яких 150 відповідають вимогам, а 50 – не відповідають. Середня вага одиниці продукції у вибірці – 850 г, дисперсія ваги – 184.

Гранична помилка середньої ваги при рівні ймовірності 0,954:

$$\Delta_x = t \sqrt{\frac{\sigma^2}{n}} = 2 \sqrt{\frac{184}{200}} = 1,9(z);$$

при рівні ймовірності 0,990:

$$\Delta_x = t \sqrt{\frac{\sigma^2}{n}} = 2,58 \sqrt{\frac{184}{200}} = 2,5(z).$$

Гранична помилка частки:

– при рівні ймовірності 0,954:

$$W = \frac{m}{n} = \frac{150}{200} = 0,75;$$

$$\Delta_w = t \sqrt{\frac{W(1-W)}{n}} = 2 \sqrt{\frac{0,75(1-0,75)}{200}} = 0,061;$$

– при рівні ймовірності 0,990:

$$\Delta_w = t \sqrt{\frac{W(1-W)}{n}} = 2,58 \sqrt{\frac{0,75(1-0,75)}{200}} = 0,079;$$

б) при без повторному випадковому та механічному відборі гранична помилка визначається за формулами:

– для середньої

$$\Delta_x = t \sqrt{\frac{\sigma^2}{n} \left(1 - \frac{n}{N}\right)}; \quad (9.3)$$

– для частки

$$\Delta_x = t \sqrt{\frac{W(1-W)}{n} \left(1 - \frac{n}{N}\right)}. \quad (9.4)$$

Наприклад, з 2 000 одиниць продукції обстежено 200 одиниць, з яких відповідають вимогам 100 одиниць. Середня вага у вибірці – 950 г, дисперсія ваги – 190.

Гранична помилка середньої ваги:

$$\Delta_x = t \sqrt{\frac{\sigma^2}{n} \left(1 - \frac{n}{N}\right)}. \quad (9.5)$$

– при ймовірності 0,950:

$$\Delta_x = 1,96 \sqrt{\frac{190}{200} \left(1 - \frac{200}{2000}\right)} = 1,8 \text{ (г)};$$

– при ймовірності 0,997:

$$\Delta_x = 3 \sqrt{\frac{190}{200} \left(1 - \frac{200}{2000}\right)} = 2,8 \text{ (г)}.$$

Гранична помилка частки одиниць, що задовольняють вимоги:

– при ймовірності 0,950:

$$W = \frac{m}{n} = \frac{100}{200} = 0,5 \quad \Delta_w = 1,96 \sqrt{\frac{0,5(1-0,5)}{200} \left(1 - \frac{200}{2000}\right)} = 0,066;$$

при ймовірності 0,997:

$$\Delta_w = 3 \sqrt{\frac{0,5(1-0,5)}{200} \left(1 - \frac{200}{2000}\right)} = 0,101.$$

9.4 Визначення обсягу вибірки та способи поширення результатів вибіркового спостереження

При підготовці вибіркового спостереження виникає необхідність визначити мінімальну чисельність одиниць у вибірці, яка забезпечує її репрезентативність. При цьому величини σ^2 та W визначаються наближено, або беруться результати попередніх чи пробних досліджень.

Мінімальна чисельність вибіркової сукупності розраховується за формулами:

при повторному випадковому відборі:

$$n = \frac{t^2 \sigma^2}{\Delta_x^2} \quad \text{та} \quad n = \frac{t^2 W(1-W)}{\Delta_w^2}; \quad (9.6)$$

при без повторному випадковому відборі та механічному відборі:

$$n = \frac{t^2 \sigma^2 N}{\Delta_x^2 N + t^2 \sigma^2} \quad \text{та} \quad n = \frac{t^2 W(1-W)N}{\Delta_w^2 N + W(1-W)t^2}. \quad (9.7)$$

Наприклад, планується провести механічний відбір з 3 000 одиниць. При цьому з ймовірністю 0,954 помилка середньої не повинна перевищувати 3,5 г при дисперсії ваги 164. Отже,

$$n = \frac{2^2 \times 164 \times 3000}{3,5^2 \times 3000 + 2^2 \times 164} = 53 \text{ од.}$$

У статистичній практиці результати вибіркового спостереження найчастіше використовуються для одержання генеральної середньої та генеральної частки.

Якщо вибіркові характеристики мають дискретний характер, то генеральні набувають інтервального вигляду з певним рівнем ймовірності.

Межі генеральної середньої:

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

Межі генеральної частки:

$$w - \Delta_w \leq p \leq w + \Delta_w.$$

Наприклад, якщо середнє значення показника у вибірковій сукупності становило 750 кг, а помилка вибіркового спостереження з ймовірністю 0,997 дорівнює 20 кг, генеральна середня буде знаходитися в межах:

$$\begin{aligned}x - \Delta_x &\leq \bar{x} \leq x + \Delta_x; \\750 - 20 &\leq \bar{x} \leq 750 + 20; \\730 &\leq \bar{x} \leq 770.\end{aligned}$$

У деяких випадках для визначення генеральних характеристик використовується метод прямого перерахунку. При цьому визначається коефіцієнт поправки, який призначений для коректування результатів вибіркового спостереження. Цей метод найчастіше застосовується тоді, коли вибіркоче спостереження має на меті перевірку якості суцільного спостереження.

Наприклад, при проведенні перепису транспортних засобів по області методом суцільного спостереження підраховано 135 680 одиниць. За результатами вибіркового спостереження (один район) одержано 17 584 одиниці проти 17 630 одиниць при суцільному спостереженні. Отже, поправочний коефіцієнт становить:

$$K_n = \frac{17\,584}{17\,630} = 0,997.$$

Тоді скоректований результат суцільного спостереження дорівнює:

$$N = 135\,680 \cdot 0,997 = 135\,527 \text{ од.}$$

Отже, цей метод дозволяє одержати генеральні характеристики у вигляді дискретних величин.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

Основна література

1. Бек В. Л. Теорія статистики: Курс лекцій : навч. посібник / В. Л. Бек – Київ : ЦУЛ, 2002.
2. Горкавий В. К. Статистика : навчальний посібник / В. К. Горкавий – Київ : Алерта, 2012.
3. Гусаров В. М. Теория статистики : уч. пособие для вузов / В. М. Гусаров – М. : Аудит, ЮНИТИ, 1998.
4. Ефимова М. Р. Общая теория статистики : учебник / М. Р. Ефимова, Е. В. Петрова, В. Н. Румянцев – 2-е изд., испр. и доп. – М. : ИНФРА – М, 2007.
5. Єріна А. М. Теорія статистики : практикум. 6-е вид. / А. М. Єріна, З. О. Пальян. – Київ : «Знання», 2008.
6. Ковалевский Г. В. Статистика : учебник / Г. В. Ковалевский ; Харьков. нац. акад. гор. хоз-ва. / Г. В. Ковалевский. – Харьков : ХНАГХ, 2012.
7. Ковалевский Г. В. Статистика зарубежных стран / Г. В. Ковалевский, В. М. Селиванов. – Харьков : ХНУ, 2001.
8. Ковалевський Г. В. Статистика / Г. В. Ковалевський. – Харків : ХНАМГ, 2010.
9. Ковтун Н. В. Теорія статистики : підручник / Н. В. Ковтун. – Київ : Знання, 2012. – 400 с.
10. Костюк В. О. Статистика : навч. посібник / В. О. Костюк, І. В. Мількін ; Харків. нац. ун-т міськ. госп-ва ім. О. М. Бекетова. – Харків : ХНУМГ ім. О. М. Бекетова, 2015. – 166 с.
11. Лугінін О. Є. Статистика : підручник. / О. Є. Лугінін. – 2-е видання, перероблене та доповнене. – Київ : Центр учбової літератури, 2007.
12. Мармоза А. Т. Теорія статистики : підручник / А. Т. Мармоза. – 2-ге вид., переробл. і доповн. – Київ : Центр учбової літератури, 2013. – 592 с.
13. Матковський С. О. Теорія статистики : навч. посібник / С. О. Матковський, О. Р. Марець. – Київ : Знання, 2008.

14. Общая теория статистики. Статистическая методология в изучении коммерческой деятельности / Под. ред. О. Э. Башиной, А. А. Спирина. – М. : Финансы и статистика, 2000.

15. Попов І. І. Теорія статистики : навч. посібник / І. І. Попов, В. С. Федорченко. – Київ : КНЕУ, 2001.

16. Про державну статистику : Закон України // Голос України. – 1992.– 21 жовтня.

17. Про заходи щодо розвитку державної статистики : Указ Президента України від 22 листопада 1997 р. № 1299/97// Статистика України. – 1998. – № 1.

18. Статистика: навч. посібник / [О. В. Раєвська, І. В. Аксьонова, Л. В. Гриневич та ін. ; під ред. О. В. Раєвської] ; М-во освіти і науки України, Харк. нац. екон. ун-т. – Х. : ІНЖЕК, 2011. – 504 с.

19. Статистика : навч.-методичний посібник для самостійного вивчення дисципліни. / О. В. Прокопова, О. О. Горошанська, І. А. Бігдан, Н. Н. Гаркуша / Харків. держ. університет харчування та торгівлі. – Харків, 2006. – 194 с.

20. Статистика : навчально-методичний посібник. / О. В. Прокопова, О. О. Горошанська / Харків. держ. університет харчування та торгівлі. – Харків, 2013. – 136 – 234 с.

21. Статистика : підручник / А. В. Головач, А. М. Єріна, О. В. Козирєв та ін. ; За ред. А. В. Головача, А. М. Єріної, О. В. Козирєва. – Київ : Вища шк., 1993.

22. Статистика : підручник / С. С. Герасименко, А. В. Головач, А. М. Єріна та ін. ; За наук. ред. д-ра екон. наук С. С. Герасименка. – 2-е вид., перероб. і доп. – Київ : КНЕУ, 2000.

23. Статистика : практикум / О. В. Прокопова, О. О. Горошанська. ДОД ХДУХТ, 2010. – 153 с.

24. Статистика : структурно-логічні схеми та задачі : навч. посібник : РМОіНУ/ А. М. Єріна, В. Б. Захожай, І. Г. Манцуров та ін. ; За наук. ред. А. М. Єріної. – Київ : КНЕУ, 2007. – 304с.

25. Статистика : учебник для вузов / Под ред. И. И. Елисейевой. – СПб. :

Питер, 2010.

26. Стегней М. І. Статистика : кредитно-модульний курс : навч. посібник / М. І Стегней, І. О. Іртищева; М-во освіти і науки України, Мукачівський держ. ун-т. – Київ : Кондор, 2012. – 306 с.

27. Тарасенко І. О. Статистика : навчальний посібник / І. О. Тарасенко. – Київ : Центр навчальної літератури, 2006. – 344 с.

28. Теория статистики : учебник / Под. ред. Г. Л. Громько. – М. : ИНФРА – М., 2000.

29. Теория статистики : учебник / Под. ред. Р. А. Шмойловой. – 2-е изд., доп. и перераб. – М. : Финансы и статистика, 1998.

30. Теорія статистики : навч. посібник / П. Г. Вашків, П. І. Пастер, В. П. Сторожук, Є. І. Ткач. – Київ : Либідь, 2001.

31. Теорія статистики : навч. посібник / Г. І. Мостовий, А. О. Дегтяр, В. К. Горкавий, В. В. Ярова ; за заг. ред. проф. Г. І. Мостового. – Харків : Вид-во ХарРІУАДУ «Магістр», 2002.

32. Ткач Є. І. Загальна теорія статистики : підручник (для студ. вищ. навч. закл.) / Ткач Є. І., Сторожук В. П., 3-тє вид. – К.: Центр учбової літератури, 2009.

33. Ткач Є.І. Загальна теорія статистики: підручник [для студ. вищ. навч. зал.] / Є. І. Ткач, В. П. Сторожук – [3-тє вид.] – Київ : Центр навчальної літератури, 2009. – 442 с.

34. Толбатов Ю. А. Загальна теорія статистики засобами EXCEL : навч. посібник / Ю. А. Толбатов. – Київ : Четверта хвиля, 1999.

35. Уманець Т. В. Статистика : навч. посібник / Т. В. Уманець, Ю. В. Пігарев. – Київ : Вікар, 2003.

36. Чернелевський Л. М. Статистика : підручник / Л. М. Чернелевський, Л. М Соломчук, М. В. Перепяцько. – Київ : НУХТ, 2012. – 207 с.

37. Щурик М. В. Статистика : навч. посібник. – 2-ге видання, оновлене і доповнене. – Львів : Магнолія-2006, 2009.

Додаткові джерела

1. Большой экономический словарь / Под. ред. А. Н. Азриляна. – 2-е изд., доп. и перераб. – М. : Ин-т новой экономики, 1997.
2. Елисеєва І. І. Общая теория статистики : учебник / І. І. Елисеєва, М. М. Юзбашев ; Под ред. І. І. Елисеєвой. – М : Финансы и статистика, 1995.
3. Елисеєва І. І. Международная статистика : учебник / І. І. Елисеєва, Л. І. Хоменко. – Минск : Выш. шк., 1995.
4. Елисеєва І. І. Общая теория статистики : учебник / І. І. Елисеєва, М. М. Юзбашев ; Под ред. І. І. Елисеєвой. – М. : Финансы и статистика, 2004. – 544 с.
5. Єріна А. М. Організація вибіркового спостереження : навч. посібник / А. М. Єріна. – Київ : КНЕУ, 2004. – 127 с.
6. Захожай В. Б. Статистика : підручник / В. Б. Захожай, І. І. Попов. – Київ : МАУП, 2006. – 536 с.
7. Индексный метод в экономике. – М. : Финансы и статистика, 1989.
8. Кевіш П. Теория индексов и практика экономического анализа : пер. с венг. / П. Кевіш – М. : Финансы и статистика, 1990.
9. Ковалевский Г. В. Индексный метод в социальной статистике / Проблемы социальной статистики. / Г. В Ковалевский. – М. : Наука, 1986.
10. Ковалевский Г. В. Харьковская экономическая школа (1804-2004). / Г. В. Ковалевский, В. М. Селиванов. – Харьков : ХНУ, 2004.
11. Общая теория статистики : Статистическая методология в изучении коммерческой деятельности : учебник / А. И. Харламов, О. Э. Башина и др. ; Под ред. А. А. Спирина. – М. : Финансы и статистика, 2006 – 296 с.
12. Опря А. Т. Статистика : навч. посібник / А. Т. Опря. – Київ : Центр навч. л-ри, 2005. – 472 с.
13. Парфенцева Н. О. Міжнародні статистичні класифікації в Україні : впровадження й використання / Н. О. Парфенцева. – Київ : Основа, 2000. – 351 с.

14. Парфенцева Н. А. Международные статистические классификации в Украине : Внедрение и использование. / Н. А. Парфенцева. – Київ : Основа, 2000.
15. Плошко Б. Г. История статистики : учебное пособие / Б. Г. Плошко, И. И. Елисеева. – М. : Финансы и статистика, 1990.
16. Попов І. І. Теорія статистики. Практикум : навч. посібник / І. І. Попов. – Київ : КНТЕУ, 2006. – 290 с.
17. Сигел З. Практическая бизнес-статистика : пер. с англ. / З. Сигел. – М. : Вильямс, 2002. – 1021 с.
18. Статистика. Курс лекций / Харченко Л. П. и др. – Новосибирск : 1997.
19. Статистичний словник / за ред. О.Г. Осауленка ; Держ. служба статистики України, Наук.-техн. комплекс стат. досліджень. – Київ : Інформ.-аналіт. агентство, 2012. – 499 с.
20. Теория статистики : учебник / Под. ред. Р. А. Шмойловой. – М. : Финансы и статистика, 1996.

Навчальне видання

КОНДРАТЕНКО Наталія Олегівна,

КОЛЕСНИК Тетяна Миколаївна,

ВОРОНІНА Олена Олегівна

СТАТИСТИКА

КОНСПЕКТ ЛЕКЦІЙ

*(для студентів першого (бакалаврського) рівня вищої освіти
всіх форм навчання спеціальностей
281 – Публічне управління та адміністрування, 073 – Менеджмент,
242 – Туризм)*

Відповідальний за випуск *М. К. Гнатенко*

За авторською редакцією

Комп'ютерне верстання *О. Г. Ткаченко*

План 2019, поз. 183 Л

Підп. до друку 06.12.2019. Формат 60 × 84/16.

Друк на ризографі. Ум. друк. арк. 7,9.

Тираж 50 пр. Зам. №

Видавець і виготовлювач:

Харківський національний університет
міського господарства імені О. М. Бекетова,
вул. Маршала Бажанова, 17, Харків, 61002.

Електронна адреса: rektorat@kname.edu.ua

Свідоцтво суб'єкта видавничої справи:

ДК № 5328 від 11.04.2017.