МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

Національний аерокосмічний університет ім. М.Є. Жуковського
«Харківський авіаційний інститут»

В.Л. Петрик

СТАТИСТИКА

Навчальний посібник для студентів факультету заочного навчання

Харків «ХАІ» 2009

УДК 681. 3. 06 + 6196

Петрик В.Л. Статистика: навч. посібник для студентів ф-ту заочного навчання / В.Л. Петрик. – Х.: Нац. аерокосм. ун-т «Харк. авіац. ін-т», 2009. - 100 с.

Розглянуто комплекс статистичних методів збору, обробки та аналізу статистичної інформації, теоретичні й методичні основи побудови системи статистичних показників, які використовують для вивчення закономірностей суспільних явищ. Особливу увагу приділено статистичній методології, можливостям її використання в умовах суттєвих змін в економіці. Наведено контрольні запитання. Подано приклади розв'язання задач.

Для студентів та аспірантів економічних спеціальностей «Фінанси», «Економіка», «Менеджмент організацій», «Маркетинг» при виконанні домашніх завдань, курсових і дипломних робіт.

Іл. 10. Табл. 52. Бібліогр.: 5 назв

Рецензенти: д-р техн. наук, проф. І.П. Гоманюк,
канд. екон. наук Н.С. Кучмієва

Вступ

У всьому світі зростає інтерес до статистики. У нашій країні ця увага тим більше загострена у зв'язку зі здійсненням економічних реформ, що стосуються інтересів усіх людей. Перехід до ринкової економіки надає нового змісту роботі комерсантів, економістів, менеджерів. Це ставить підвищені вимоги до рівня їхньої статистичної підготовки. Оволодіння статистичною методологією – це одна з неодмінних умов пізнання кон'юнктури ринку, вивчення тенденцій і прогнозування попиту та пропозиції, ухвалення оптимальних рішень на всіх рівнях комерційної діяльності на ринку товарів і послуг.

Тема 1. ПРЕДМЕТ І МЕТОД СТАТИСТИЧНОЇ НАУКИ

Статистика (від лат. status – стан) – наука, яка вивчає розміри та кількісні співвідношення масових суспільно-економічних явищ і процесів у нерозривному зв'язку з їхнім змістом.

Статистика має багатовікову історію. Її виникнення і розвиток були обумовлені такими суспільними потребами: підрахунок населення, худоби, облік земельних угідь, майна і т.ін. Найраніші відомості про такі роботи в Китаї відносяться до V століття II тис. до н.е.

У міру розвитку суспільного виробництва, внутрішньої і зовнішньої торгівлі збільшувалася потреба в статистичній інформації. Це розширювало сферу діяльності статистики, вдосконалювало її прийоми і методи.

Вважається, що основи статистичної науки були закладені англійським економістом У. Петті (1623 -1687). Послідовники У. Петті започаткували науковий напрям, який отримав назву «політична арифметика».

Засновником іншого напряму розвитку статистичної науки був визначний німецький вчений Г. Конрінг (1606 - 1681), який розробив систему опису держави. Його послідовник професор філософії й права Г. Ахенваль (1719 - 1772) вперше в Марбурзькому університеті (1746) почав викладати нову дисципліну, названу ним статистикою. Державознавство знайшло віддзеркалення і у ряді робіт М.В. Ломоносова (1711 - 1765), в яких розгляд питань населення, природних багатств, фінансів, торгівлі Росії був проілюстрований статистичними даними. Цей напрям розвитку статистики отримав назву *описового*.

У XIX в. бельгійський статистик А. Кетле поклав основу вченню про *середні величини*.

Математичний напрям у статистиці розвивали англійці Ф. Гальтон (1822 - 1911), К. Пірсон (1857 - 1936), В. Госсет – псевдонім Стьюдент (1876 - 1937), Р. Фішер (1890 - 1962) та ін. Представники цього напряму вважали основою статистики теорію вірогідності.

У розвитку статистики значне місце належить представникам російської науки та практики. В епоху Петра I І.К. Кирилов (1689 - 1737) і В.М. Татіщев (1686 - 1750) у своїх працях статистику трактували переважно як описову науку. Але вже з другої половини XIX в. на перший план висувається пізнавальне значення статистики. Так, В.С. Порошин (1809 - 1868) в роботі «Критичне дослідження про основи статистики» підкреслював, що наука не може обмежитися лише одним описанням.

Предмет статистики – кількісна сторона масових суспільних явищ у нерозривному зв'язку з їх якісною стороною або змістом.

Предмет статистики вивчається за допомогою різноманітних методів, сукупність яких утворює ***статистичну методологію***.

Усе різномаїття статистичних методів вивчення комерційної діяльності систематизується згідно з їх цільовим вживанням у послідовно виконуваних трьох ***основних стадіях економіко-статистичного дослідження***:

- статистичне спостереження;
- статистичне зведення і групування первинної інформації;
- аналіз статистичної інформації.

Статистичне спостереження – це спланований, систематичний та науково організований збір масових даних про різноманітні соціально-економічні явища та процеси.

Для здійснення спостереження застосовують ***методи масового спостереження***, які дозволяють отримати певні значення досліджуваних ознак від кожної одиниці статистичної сукупності шляхом реєстрації їх на підставі розробленої програми.

Статистичне зведення і групування первинних даних необхідні для систематизації матеріалу статистичного спостереження. Вони полягають у перевірці даних, їх групуванні за певними ознаками: підведенні групових і загальних підсумків, розрахунку різних показників, проектуванні таблиць і внесенні до них даних. На цій стадії застосовують ***метод групувань***, який дає можливість виділити в досліджуваній сукупності соціально-економічні типи явищ, охарактеризувати їх структуру, виявити взаємозв'язки і взаємозалежності між показниками.

Аналіз статистичної інформації передбачає проведення аналізу даних на основі обчислення узагальнюючих статистичних показників: абсолютних, відносних і середніх величин, статистичних коефіцієнтів, показників варіації ознак і динаміки явищ, індексів і показників, що характеризують густину зв'язку між явищами і т.п.

Предмет статистики вивчають за допомогою певних категорій – ***понять, які відображають загальні й істотні*** властивості, ознаки, зв'язки і відносини предметів і явищ об'єктивного світу.

Основні категорії теорії статистики:

1. ***Статистична сукупність*** – це велика кількість одиниць явища, об'єднаних відповідно до задачі дослідження єдиною якісною основою, загальним зв'язком, але відмінних один від одного окремими ознаками (наприклад, сукупність промислових підприємств України, сукупність сімей, сукупність підприємств, фірм, об'єднань і т.п.).

Сукупність називається однорідною, якщо один або декілька істотних ознак її об'єктів, що вивчаються, є загальними для всіх одиниць. Сукупність, до якої входять явища різного типу, вважається різнорідною.

2. ***Одиниця статистичної сукупності*** – первинний елемент статистичної сукупності, який є носієм ознак, що підлягають реєстрації. Наприклад, при переписі населення одиницею сукупності є кожна людина; при обстеженні проданих на біржі квартир – кожна продана квартира.

3. ***Ознака*** – якісна особливість одиниці сукупності – відмінна риса, властивість, якість, що є характерною для окремих одиниць, об'єктів (явищ). Наприклад, ознаками промислового підприємства є обсяги виробництва, розмір основних виробничих фундацій, чисельність персоналу та ін.; демографічні й соціально-економічні ознаки людини: вік, рівень освіти, професія, підлога і т.п.

4. Явища та процеси в житті суспільства статистика вивчає за допомогою статистичних показників.

Статистичний показник – узагальнююча характеристика соціально-економічного явища або процесу, в якій поєднуються якісна та кількісна визначеність останнього (наприклад, чисельність населення, продукція промислового підприємства, рівень продуктивності праці, рівень рентабельності й тому подібне).

5. ***Система статистичних показників*** **–** це сукупність показників, що відображає взаємозв'язки, які об'єктивно існують між явищами.

Контрольні запитання

1. Хто і коли започаткував описовий напрям у статистиці? У чому полягає суть цього напряму?
2. З якого часу статистика почала розвиватися як наука? Яку роль у розвитку статистики відіграв У. Петті?
3. Дайте визначення предмета статистики. Чим відрізняється предмет статистики від предмета інших суспільних наук?
4. Назвіть коло суспільних явищ, що вивчає статистика.
5. Чи обов'язково входять до статистичної сукупності якісно однорідні елементи та явища і чому?
6. Яку роль відіграє статистика в соціально-економічних дослідженнях і в розвитку інших суспільних наук?

7. Назвіть основні категорії статистики.
8. Назвіть специфічні методи статистичного дослідження явищ.
9. Які галузі статистики ви знаєте?
10. Що таке статистичні показники?

Тема 2. СТАТИСТИЧНЕ СПОСТЕРЕЖЕННЯ

Статистичне спостереження – перша стадія статистичного дослідження, науково організований за єдиною програмою облік фактів, що характеризують явища та процеси суспільного життя, а також збір отриманих на основі цього обліку масових даних.

Статистичні дані – це масові системні кількісні характеристики соціально-економічних явищ і процесів. Статистичні дані повинні бути: достовірними, повними (за обсягом і суттю), своєчасними, порівнянними за часом і простором, доступними.

Мета статистичного спостереження – отримання статистичних даних, які є підставою для узагальненої характеристики стану і розвитку явища або процесу.

Об'єкт спостереження – сукупність соціально-економічних явищ і процесів, які підлягають дослідженню, або точні границі, в межах яких реєструватимуться статистичні дані.

Статистичне спостереження може бути:

- первинне – це реєстрація даних, що надходять безпосередньо від об'єкта, який їх продукує (наприклад, поточний облік);
- вторинне – збір раніше зареєстрованих і оброблених даних (наприклад, банківських звітів).

Про статистичне спостереження можна говорити лише тоді, коли:

1) забезпечується реєстрація встановлюваних фактів у спеціальних облікових документах;
2) вивчаються статистичні закономірності.

Тому статистичне спостереження повинне бути планомірним, масовим і систематичним.

Будь-яке статистичне дослідження починається зі складання ***плану спостереження*** – сукупності програмно-методологічних і організаційних питань.

Програмно-методологічні питання плану спостереження – це перелік пунктів, які відповідатимуть на питання:

1) для чого проводиться обстеження (формулювання його мети та конкретних задач);
2) що обстежується (визначення об'єкта обстеження);
3) складові частини об'єкта (одиниця сукупності);

4) джерело інформації (одиниця спостереження);

5) розробка програми спостереження, вибір виду та способу спостереження.

Одиниця статистичного спостереження – це та первинна ланка, з якої мають бути отримані необхідні статистичні дані.

Наприклад, під час перепису населення об'єкт спостереження – сукупність людей, що проживають у даній країні на час перепису. Одиниця сукупності – окрема людина, і в цьому випадку така одиниця збігається з одиницею спостереження. При визначенні обсягу роздрібного товарообігу одиницями спостереження є торговельні підприємства, одиницями сукупності – акти купівлі-продажу товарів населенню.

Програма спостереження – це перелік чітко сформульованих питань, на які необхідно отримати відповіді, або перелік ознак і показників, що підлягають реєстрації.

Для забезпечення однакового тлумачення програми в обліковому формулярі, в якому вона міститься, даються пояснення, викладені в документі, який називається ***інструкцією***.

У програму спостереження також включаються: розробка статистичного інструментарію, визначення виду і способу обстеження.

Статистичний інструментарій – це набір статистичних формулярів, інструкцій і роз'яснень щодо проведення спостереження.

Статистичний формуляр – обліковий документ єдиного зразка, який містить адресну характеристику об'єкта спостереження та статистичні дані про нього. Наприклад, звіти, переписні й опитні листки, бланки документів, анкети.

При складанні формулярів враховують не тільки зміст та інформативність ознак, але і можливість їх статистичної обробки, яка забезпечується завдяки вживанню системи шкал.

Шкала – засіб впорядкування і кількісного виразу ознак. Використовують такі види шкал:

1) номінальна – шкала найменувань, яка встановлює відношення подібності елементів, для якої порядок розташування ознак значення не має. Наприклад, класифікатор сфер економічної діяльності, перелік форм власності, видів підприємницької діяльності й т.п.;

2) порядкова (рангова) – шкала, яка встановлює послідовність інтенсивності прояву ознаки. Застосовується під час визначення ступеня економічного ризику підприємництва, рівня кваліфікації робітників або при визначенні відношення респондентів до якогось явища або процесу. Ознаки порядкової шкали оцифровують шляхом присвоєння рангів (балів) окремим значенням у міру збільшення або зменшення їх інтенсивності;

3) метрична – кількісна шкала, в основу якої покладено результати

безпосереднього вимірювання. Метричною шкалою визначаються обсяги виробництва та реалізації продукції, розміри капіталу, чисельність зайнятих у виробництві осіб, кількість і вартість приватизованих об'єктів тощо.

Організаційні питання статистичного спостереження включають визначення суб'єкта (того, хто проводить ці спостереження), місця, часу, форми та способу спостереження.

У статистичній практиці використовуються дві організаційні форми спостереження – звітність і спеціально організоване спостереження.

Статистична звітність – офіційний документ, що містить статистичні дані про роботу підприємств і фірм у вигляді обов'язкових звітів про їх роботу за встановленими формами і у відповідних термінах (у вигляді формулярів регламентованого зразка).

На підставі даних звітності формують усю поточну статистику, вивчають закономірності розвитку окремих підприємств, об'єднань, галузей народного господарства.

Спеціально організоване статистичне спостереження – форма спостереження за явищами і процесами, не охопленими статистичною звітністю. Статистичні дані одержують на підставі перепису, одноразового обліку та спеціального статистичного обстеження.

Перепис – спеціально організоване статистичне спостереження, основне завдання якого полягає в повному обліку чисельності та характеристиці складу явища шляхом запису в статистичний формуляр даних про досліджувані одиниці статистичної сукупності.

Одноразовий облік – це переписи, в яких статистичні формуляри заповнюють на основі матеріалів первинного обліку. Наприклад, переписи залишків сировини, устаткування, облік посівних площ.

Способи спеціального статистичного обстеження:

1. При безпосередньому спостереженні відомості одержують шляхом особистого обліку одиниць сукупності: перерахунку, вимірювання і т.д.

2. Документальний спосіб збору статистичної інформації базується на систематичних записах у первинних документах, підтверджуючих той або інший факт.

3. Опитування – це спостереження, яке передбачає, що відомості фіксуються зі слів опитуваного. Розрізняють три способи опитування:

- експедиційний спосіб – призначена особа опитує іншу і з його слів заповнює бланк дослідження;
- анкетний;
- самореєстрація – відповідні документи заповнюють опитувані;
- кореспондентський спосіб – до органів, які здійснюють спостереження, посилають відомості їх кореспонденти.

Залежно від задач статистичного дослідження і характеру явища, що вивчається, облік фактів можна проводити:

- систематично, постійно охоплюючи факти у міру їх виникнення, – це буде поточне спостереження (звітність);
- регулярно, але не постійно, а через певні проміжки часу, – це буде періодичне спостереження (переписи населення).

Види статистичних спостережень розділяють за такими ознаками: повнота обхвату одиниць сукупності, час реєстрації фактів, джерела відомостей, способи реєстрації, збір даних.

З погляду повноти обхвату фактів (одиниць) статистичне спостереження може бути суцільним і несуцільним.

Суцільне спостереження є повним обліком усіх одиниць сукупності, що вивчається: наприклад, перепис населення.

Несуцільне спостереження організовують як облік частини одиниць сукупності, на основі якої можна отримати узагальнюючу характеристику всієї сукупності.

Види несуцільного спостереження:

1) ***спосіб основного масиву*** – відбирають найкрупніші одиниці спостереження, в яких зосереджена значна частка всіх фактів, які підлягають вивченню;
2) ***вибіркові спостереження*** – випадково відбирають зі всієї сукупності обмежене число об'єктів;
3) ***монографічні описи*** застосовують для докладного вивчення одиничних, але типових об'єктів, наприклад, окремих підприємств.

У разі ***поточного спостереження*** облік фактів здійснюють у міру їх виконання, тобто систематично (наприклад, факти, що характеризують роботу підприємства, господарства або виручку магазинів від реалізації товарів, реєстрацію актів цивільного стану або роботу транспорту).

Періодичним вважається спостереження, повторюване через певні проміжки часу. Наприклад, щорічний облік худоби, який здійснюють станом на 1 січня, облік фахівців у народному господарстві.

Одноразові спостереження здійснюють за необхідності, час від часу, без дотримання певної періодичності або ж взагалі лише один раз (наприклад, вивчення думки читачів або глядачів з приводу того або іншого питання, думки покупців про якість товарів і тому подібне).

При ***безпосередньому обліку*** представники органів статистики одержують необхідну інформацію про досліджувану сукупність шляхом особистого підрахунку. Наприклад, під час інвентаризації основних засобів в обліковий формуляр зведення заносять на підставі особистого огляду реєстратором кожного об'єкта.

Документальний облік фактів є спостереженням, при якому джерелом відомостей є відповідні документи.

Контрольні запитання

1. Що таке статистичне спостереження? У чому полягає його суть?
2. Назвіть організаційні форми статистичного спостереження та їхні особливості.
3. У чому полягає сутність статистичної звітності?
4. Які розрізняють види звітності?
5. Що таке спеціально організоване статистичне спостереження?
6. Назвіть види спеціально організованих статистичних спостережень.
7. Які види статистичного спостереження за ступенем охоплення одиниць сукупності ви знаєте?
8. Яке значення для організації статистичною дослідження має програма спостереження?
9. Що саме містять програмно-методологічні та організаційні питання статистичного спостереження?
10. Поясніть, що таке одиниця статистичного спостереження та одиниця статистичного обліку.
11. Що таке об'єкт статистичного спостереження?

Тема 3. ЗВЕДЕННЯ І ГРУПУВАННЯ СТАТИСТИЧНИХ ДАНИХ

Статистичне зведення – упорядкування, систематизація і наукова обробка даних з метою отримання узагальненої характеристики досліджуваного об’єкта або процесу.

Статистичне зведення здійснюється за науково розробленою програмою, яка включає визначення груп і підгруп, системи показників, видів таблиць.

За складністю побудови зведення поділяють на прості й групові.

Просте підсумкове зведення не передбачає попереднього розподілу на групи отриманих відомостей, воно визначає загальний підсумок усіх одиниць сукупності або загальний обсяг досліджуваного показника. Наприклад, щоб знайти загальну чисельність студентів в Україні, достатньо скласти дані про чисельність студентів у всіх вищих навчальних закладах.

Групове зведення передбачає попередній розподіл одиниць на групи (наприклад, рентабельні й збиткові підприємства), що дає можливість підрахувати кількість одиниць сукупності й обсяг досліджуваної ознаки в кожній групі.

За способом здійснення статистичне зведення поділяють на централізоване та децентралізоване.

При ***централізованому зведенні*** матеріали спостереження зосереджуються, систематизуються і узагальнюються у центральному органі державної статистики, у Міністерстві статистики України.

Децентралізоване зведення передбачає узагальнення матеріалу від низу до верху по ієрархічних сходинках управління з відповідною обробкою на кожній з них.

Групування – це розбиття сукупності на однорідні групи на підставі розподілу всієї сукупності досліджуваного явища на окремі групи за найістотнішими ознаками.

Основні задачі, які вирішуються за допомогою статистичних групувань: виділення соціально-економічних типів явищ; вивчення структури досліджуваних явищ і структурних змін; дослідження взаємозв'язку і залежності між ознаками суспільних явищ.

Згідно з цими задачам розрізняють типологічні, структурні й аналітичні групування.

Типологічне групування призначене для виділення соціально-економічних типів явищ в якісно неоднорідній сукупності, визначення істотних відмінностей між ними і загальних ознак для всіх груп. Типологічне групування використовують при вивченні розподілу підприємств за формами власності та суспільного виробництва, за економічним призначенням продукції, при групуванні населення за соціальними групами тощо.

Структурне групування характеризує структуру якісно однорідної сукупності за якою-небудь однією варіюючою ознакою. Вона дає можливість описати складові частини сукупності та проаналізувати структурні зрушення (наприклад, вивчення підприємств за галузями виробництва, розмірами основних виробничих фондів, за рівнем механізації виробництва, кількістю працівників, обсягом продукції, для дослідження складу населення – за статтю, віком, національністю, походженням тощо).

Структурні групування, як і типологічні, можна здійснювати з урахуванням атрибутивних і кількісних ознак.

Типологічні групування відрізняються від структурних лише метою дослідження, за формою ж вони цілком збігаються (табл. 3.1).

Таблиця 3.1

Схема структурного і типологічного групувань

Межі групування за істотною ознакою	Кількість одиниць сукупності	Система показників		
Разом				

Аналітичне групування дозволяє оцінювати взаємозв'язки між двома (і більше) взаємодіючими ознаками, з яких один розглядається як наслідок, а інший – як фактор. У кожній групі факторної ознаки визначається середня величина результативної ознаки (табл. 3.2).

За наявності зв'язку між ознаками середні групові систематично збільшуються (прямий зв'язок) або зменшуються (зворотний зв'язок).

Таблиця 3.2

Схема аналітичного групування

Межі групування за факторною ознакою x_i	Кількість одиниць сукупності f_i	Середнє значення результативної ознаки $\bar{y}_i$
	f_1 f_2 …	$\bar{y}_1$ $\bar{y}_2$ …
Разом	$\sum f_i$	-

Комбінаційне групування містить групи, що виділені за однією ознакою, які підрозділяються на підгрупи за іншою ознакою.

Особливим видом групувань, який широко використовується в статистиці, є класифікація – стійке розмежування об'єктів, яке грунтується на найістотніших ознаках (наприклад, класифікація галузей народного господарства, класифікація основних фондів і т.д.).

При проведенні групування доводиться вирішувати такі задачі:

1) виділення групувальної ознаки;

2) визначення числа груп і величини інтервалів;

3) за наявності декількох групувальних ознак опис того, як вони комбінуються між собою;

4) встановлення показників, за якими характеризують групи.

Групувальна ознака – це ознака, за якою відбувається об'єднання окремих одиниць сукупності в однорідні групи.

Ознаки розрізняються за способами їх вимірювання й іншими особливостями, що впливають на прийоми статистичного вивчення. Це дає підставу для класифікації ознак (табл. 3.3).

Таблиця 3.3

Основна класифікація ознак

За формою виразу	За способом обчислення	Відносно об'єкта	За характером варіації	За ознакою часу	За роллю у взаємозв'язку явищ
Атрибутивні: - номінальні, - порядкові, варіаційні	Первинні, похідні (вторинні)	Прямі, непрямі	Альтернативні, варіаційні	Моментні, інтервальні	Факторні, результативні

Атрибутивні ознаки характеризують властивості, якості явищ і не мають кількісного виразу (стать, національність людини, освіта, професія).

Номінальні – описові ознаки, за якими не можна ранжувати дані.

Порядкові – ознаки, за якими можна ранжувати, упорядковувати дані.

Кількісні (***варіаційні***) ***ознаки*** набувають різного кількісного виразу в певних одиницях (вік людини, заробітна платня, дохід фірми, стаж роботи працівника, обсяг виробництва).

Первинні ознаки характеризують одиницю сукупності в цілому, можуть бути зміряні, злічені, зважені; існують самі по собі, незалежно від їх статистичного вивчення і надаються у формі абсолютних величин (кількість і сума внесків у банку, потужність двигунів тощо).

Похідні (вторинні) ознаки є співвідношеннями первинних ознак (продуктивність праці, собівартість одиниці продукції, рентабельність, врожайність).

Прямі ознаки – це властивості, безпосередньо притаманні тому об'єкту, який вони характеризують (обсяг продукції заводу, чисельність його робітників).

Непрямі ознаки – це властивості, що характерні не самому об'єкту, а іншим сукупностям, які відносяться до об'єкта (оплата праці робітників – непряма ознака заводу, але пряма для робітників).

Альтернативні ознаки – це протилежні за значенням варіанти ознаки, тобто ті, які можуть набувати тільки двох значень. Наприклад, продукція може бути якісною або неякісною.

Варіаційні ознаки мають багато кількісних значень (чисельність працівників).

Моментні ознаки характеризують об'єкт у якийсь момент часу, встановлений планом статистичного дослідження (чисельність населення, кількість худоби, розміри житлової площі).

Інтервальні ознаки характеризують результати процесів, явищ за певний час (рік, місяць, доба), але не на момент часу.

Факторні ознаки впливають на інші ознаки.

Результативні ознаки – це ознаки, розмір і динаміка яких формуються під впливом інших ознак.

Після визначення групувальної ознаки важливим кроком є розподіл одиниць сукупності на групи. Для цього потрібно визначити кількість утворюваних груп і розмір інтервалу.

Інтервал обкреслює кількісні межі груп. Як правило, він є проміжком між максимальними і мінімальними значеннями ознаки в групі. Інтервали бувають:

- *відкриті,* коли є тільки або верхня, або нижня межа;
- *закриті,* коли є і нижня, і верхня межі.

Закриті інтервали поділяють на:

- *рівні,* коли різниця між максимальним і мінімальним значеннями в кожному з інтервалів однакова;

- *нерівні,* коли, наприклад, ширина інтервалу поступово збільшується, а верхній інтервал часто не закривається зовсім.

Число груп ***при рівних інтервалах*** визначають за формулою Стерджеса

$$n=1+3{,}322*\lg n, \qquad (3.1)$$

де n – кількість елементів сукупності.

Користуватися цією формулою можна лише в тих випадках, коли досліджувана сукупність достатньо велика і зміна груп ознаки має відносно плавний характер.

Для рівних інтервалів їх величина може бути визначена за формулою

$$l = \frac{x_{max} - x_{min}}{n}, \qquad (3.2)$$

де x_{max}, x_{min} – відповідно максимальне і мінімальне значення ознаки; n – кількість елементів сукупності.

Нижня межа першого інтервалу a_0 може збігатися з мінімальним значенням ознаки або визначається за формулою

$$a_0 = x_{min} - \frac{l}{2}.$$

Верхню межу першого інтервалу і нижню межу другого розраховують за формулою

$$a_1 = a_0 + l.$$

Межі подальших інтервалів знаходять за формулою $a_{j+1} = a_j + l$.

Основою будь-якого групування є ряд розподілу.

Ряд розподілу – групування, в якому для характеристики груп, впорядкованих за значенням групувальної ознаки, застосовується один показник – чисельність групи.

Ряд розподілу складається з двох елементів – варіантів і частот.

Варіанти – окремі значення групувальної ознаки.

Частоти – числа, що показують, скільки разів повторюються окремі значення варіантів або кожної групи варіаційного ряду.

Замість частот може бути ***частка*** – доля частоти від загальної чисельності сукупності, виражена коефіцієнтом або відсотком.

Накопичену частоту або частку називають ***кумулятивною***.

Обсяг сукупності – сума всіх частот, що визначає число елементів всієї сукупності.

Ряди розподілу, побудовані за атрибутивною ознакою, називаються ***атрибутивними***: наприклад, розподіл населення за статтю, зайнятістю, національністю, професією тощо.

Ряди розподілу, побудовані за кількісною ознакою, називаються ***варіаційними рядами***: наприклад, розподіл населення за віком, робітників – за стажем роботи, за заробітною платнею тощо.

Варіаційні ряди, залежно від способу побудови, підрозділяються на дискретні й інтервальні.

За дискретною ознакою, кількість значень якої обмежена, утворюється ***дискретний ряд розподілу.***

Якщо ознака може приймати хоча і дискретні значення, але їх число велике, або може набувати дробових значень в області свого існування, будується ***інтервальний варіаційний ряд*** – таблицю з двох рядків або граф: інтервалів ознаки, варіація якого вивчається, і частот.

Засобом наочного виразу результатів статистичного дослідження є статистичні таблиці, які служать для подання результатів зведення і групування матеріалів спостереження.

Підмет таблиці – статистична сукупність, об'єкти або їх частини, які характеризуються рядом числових показників.

Присудок таблиці – показники, які характеризують статистичну сукупність, тобто підмет таблиці.

Складену таблицю, але не заповнену цифрами, називають ***макетом таблиці*** (табл. 3.4).

Таблиця 3.4

Макет статистичної таблиці

Загальний заголовок							
Присудок / Підмет	Заголовки граф (верхні заголовки)						
А	1	2	3	4	5	6	
Бічні заголовки							Рядки таблиці
Підсумок							
	Графи таблиці						

У комерційній роботі складаються різноманітні таблиці, які залежно від побудови підмета діляться на прості, групові, комбінаційні.

У ***простих таблицях*** підмет містить перелік об'єктів, адміністративних і територіальних одиниць або перелік періодів, дат.

У ***групових таблицях*** підмет містить групи, утворені за однією істотною ознакою або виявленим зв'язком між рядом показників.

Найпростіший вид групових таблиць – ряди розподілу, побудовані за атрибутивними або кількісними ознаками. Присудок таких таблиць має лише один показник – кількість одиниць сукупності, які входять до кожної групи.

У підметі ***комбінаційних таблиць*** групи, побудовані за однією ознакою, поділяють на підгрупи за іншими ознаками. Іноді групи за однією ознакою розміщують у підметі, а за іншими ознаками – в присудку. Такий вигляд мають комбінаційні ряди розподілу.

3.1. Розв'язання типових задач

Приклад 3.1

Є дані про робітників підприємства (табл. 3.5).

Таблиця 3.5

№ п/п	Стаж роботи, років	Місячний виробіток робітником нормативно-чистої продукції, грн (НЧП)
1	1,0	200
2	1,0	202
3	3,0	205
4	6,5	290
5	9,2	298
6	4,4	250
7	6,9	280
8	2,5	230
9	2,7	223
10	16,0	310
11	13,2	284
12	14,0	320
13	11,0	295
14	12,0	279
15	4,5	222

1. Побудувати ряд розподілу робітників за стажем, утворивши п'ять груп з однаковими інтервалами. За кожною групою визначити кількість робітників та їх питому вагу до загальної кількості робітників.
2. Для вивчення залежності між стажем роботи і місячним виробітком працівників зробіть:

а) аналітичне групування робітників за стажем, утворивши п’ять груп з рівними інтервалами. Кожну групу охарактеризуйте за числом робітників; середнім стажем роботи; місячним виробітком продукції — усього й у середньому на одного робітника;

б) комбінаційне групування за двома ознаками: стажем роботи і середнім місячним виробітком продукції на одного робітника.

Розв'язання.

1. Обчислимо величину інтервалу ознаки групування (стажу роботи) за формулою (3.2):

$$I = \frac{16 - 1}{5} = 3 \text{роки}.$$

Отже, перша група робітників має стаж 1-4 роки; друга – 4-7 і т.д. Для кожної групи потрібно підрахувати чисельність робітників і оформити результати у вигляді табл. 3.6. Для проведення подальшого аналізу чисельність робітників може бути подана у відсотках (табл. 3.6, графа 4).

Таблиця 3.6

Номер групи	Групи робітників за стажем, років	Число робітників, чол.	Число робітників, відсоток до підсумку,%
I	1 - 4	5	33,4
II	4 - 7	4	26,7
III	7 - 10	1	11,5
IV	10 - 13	2	13,4
V	13 - 16	3	15,0
Разом		15	100

Результати групування показують, що більше половини робітників, тобто 60,1%, мають стаж роботи від 1 до 7 років.

2а. Застосовуючи метод групувань для взаємозв'язку, необхідно визначити факторну ознаку, яка впливає на взаємозалежні з нею ознаки. Такою ознакою в нашому прикладі є стаж роботи, який має бути покладений в основу групування. За умовою потрібно виділити п'ять груп робітників за стажем з рівними інтервалами. За основу аналітичного групування візьмемо ці ж групи.

Для побудови й оформлення результатів групування спочатку оформимо робочу табл. 3.7, в якої для кожної групи робітників за стажем підрахуємо кількість робітників, їх загальний стаж та місячний виробіток.

Таблиця 3.7

№п/п	Групи робітників за стажем, років	Номер робітника	Стаж, років	Місячний виробіток НЧП, грн
I	1 - 4	1	1,0	200
		2	1,0	202
		3	3,0	205
		8	2,5	230
		9	2,7	223
Разом		5	10,2	1060

Закінчення табл. 3.7

№п/п	Групи робітників за стажем, років	Номер робітника	Стаж, років	Місячний виробіток НЧП, грн
II	4 - 7	*6*	4,4	250
		4	6,5	290
		7	6,9	280
		15	4,5	222
	Разом	4	22,3	1042
III	7 - 10	5	9,2	298
	Разом	1	9,2	298
IV	10 - 13	13	11,0	295
		14	12,0	279
	Разом	2	23,0	574
V	13 - 16	10	16,0	310
		11	13,2	284
		12	14,0	320
	Разом	3	43,2	914
	Усього	15	107,9	3888

Групові показники робочої таблиці й обчислені на їх основі середні показники занесемо у відповідні графи зведеної аналітичної табл. 3.8.

Таблиця 3.8

Номер групи	Групи робітників за стажем, років	Число робітників, чол.	Середній стаж роботи, років	Місячний виробіток НЧП, грн	
				усього	на одного робітника
А	Б	1	2	3	4
I	1- 4	5	2,04	1060	212,0
II	4 - 7	4	5,575	1042	260,5
III	7 - 10	1	9,2	298	298,0
IV	10 - 13	2	11,5	574	287,0
V	13 - 16	3	14,4	914	304,7
Разом		15	7,19	3888	259,2

Порівнюючи графи 2 і 4 табл. 3.8, бачимо, що зі збільшенням стажу робітників зростає місячний виробіток продукції, і, отже, між досліджуваними ознаками існує пряма залежність.

2б. Щоб зробити комбінаційне групування за двома ознаками (табл. 3.9) – стажем роботи і середнім місячним виробітком продукції– потрібно в кожній групі робітників за стажем виділити групи за другою ознакою (середнім місячним виробітком продукції на одного робітника) й охарактеризувати групи за необхідними показниками.

Таблиця 3.9

№ п/п	Групи робітників		Число робітників, чол.	Середній стаж роботи, років	Середньомісячний виробіток продукції, грн	
	За стажем, років	За середньомісячним виробітком продукції, грн			усього	на одного робітника
I	1 - 4	200 - 224	4	1,9	830	207,5
		224 - 248	1	2,5	230	230
		248 - 272	0	0	0	0
		272 - 296	0	0	0	0
		296 - 320	0	0	0	0
Разом			5	4,4	1060	218,7
II	4 - 7	200 - 224	1	4,5	222	222
		224 - 248	0	0	0	0
		248 - 272	1	4,4	250	250
		272 - 296	2	6,7	570	285
		296 - 320	0	5,5	0	0
Разом			4	15,6	1042	252
III	7 - 10	200 - 224	0	0	0	0
		224 - 248	0	0	0	0
		248 - 272	0	0	0	0
		272 - 296	0	0	0	0
		296 - 320	1	9,2	298	298
Разом			1	9,2	298	298
IV	10 - 13	200 - 224	0	0	0	0
		224 - 248	0	0	0	0
		248 - 272	0	0	0	0
		272 - 296	2	11,5	574	287
		296 - 320	0	0	0	0
Разом			2	11,5	574	287
V	13 - 16	200 - 224	0	0	0	0
		224 - 248	0	0	0	0
		248 - 272	0	0	0	0
		272 - 296	1	13,2	284	284
		296 - 320	2	15	630	315
Разом			3	28,5	914	599
Усього за підгрупами		200 - 224	5	1,2	1052	210,4
		224 - 248	1	2,5	230	230
		248 - 272	1	4,4	250	250
		272 - 296	5	9,9	1428	285,6
		296 - 320	3	13	928	309,3
Усього			15	31,1	3888	259,2

Дані таблиці свідчать, що виробіток продукції робітників має пряму залежність від стажу.

Ще один вид комбінаційного групування наведено в табл. 3.10.

Таблиця 3.10

	200 - 224	224 - 248	248 - 272	272 - 296	296 - 320	разом
1 - 4	4	1	0	0	0	5
4 - 7	1	0	1	2	0	4
7 - 10	0	0	0	0	1	1
10 - 13	0	0	0	2	0	2
13 - 16	0	0	0	1	2	3
Разом	5	1	1	5	3	15

Контрольні запитання

1. Що таке другий етап статистичного дослідження?
2. Які види зведення ви знаєте? Коротко їх охарактеризуйте.
3. Що саме називають статистичним групуванням і групувальними ознаками?
4. Які завдання вирішує статистика за допомогою методу групування?
5. Назвіть основні види статистичних групувань.
6. У чому полягають принципи вибору групувальної ознаки та утворення груп?
7. Які групування називають типологічними, структурними, аналітичними**?**
8. У чому полягає взаємозв'язок усіх зазначених видів групувань?
9. Як визначають кількість груп і межі інтервалів між ними?
10. Що таке ряди розподілу і за якими ознаками вони можуть утворюватись?
11. За якими принципами поділяються варіаційні ряди розподілу?
12. Які форми розподілу найчастіше трапляються в статистиці?
13. У чому полягає суть статистичної таблиці? З яких елементів вона складається?
14. Які види таблиць за характером підмета ви знаєте?
15. У чому перевага групових і комбінаційних таблиць порівняно з простими?

Тема 4. СТАТИСТИЧНІ ГРАФІКИ

Статистичний графік – це спосіб наочного подання і викладення статистичних даних за допомогою геометричних знаків та інших графічних засобів з метою узагальнення й їх аналізу.

Основні елементи графіка:

- ***поле графіка*** – простір, на якому розміщують графічне зображення;

- ***графічний образ*** – сукупність геометричних або графічних знаків, за допомогою яких відображаються статистичні дані;

- ***масштабні орієнтири*** – масштаб, масштабні шкали, масштабні знаки для визначення розмірів;

- ***просторові орієнтири*** – це елементи, які використовують для визначення порядку розміщення графічних знаків у полі графіка;

- ***експлікація графіка*** – пояснення, що розкривають його зміст і основні елементи (заголовок, одиниці вимірювання, умовні позначення).

Класифікація статистичних графіків:

1. За *загальним призначенням* графіки поділяють на аналітичні, ілюстративні й інформаційні.

2. За *функціонально-цільовим* призначенням виділяють графіки групувань і рядів розподілу, графіки порівняння, динаміки, графіки взаємозв'язаних показників. Можливі комбінації цих графіків, наприклад, графічне зображення варіації в динаміці або динаміки взаємозв'язаних показників і т.п.

3. За *способом побудови (за* виглядом поля) графіки поділяють на діаграми і статистичні карти (картодіаграми і картограми).

4. *За формою графічного образу* розрізняють графіки крапкові, лінійні, площинні (стовпчикові, почасові, квадратні, кругові, секторні, фігурні), просторові (об'ємні) і фігурні.

5. За *типом системи координат* виділяють графіки в прямокутній і полярній системах.

6. За *типом масштабних шкал* – графіки з рівномірними, нерівномірними (функціональними) і змішаними шкалами.

7. *За функціонально-цільовим призначенням* розрізняють графіки:

- порівняння статистичних величин (діаграми – стовпчикові, стрічкові, квадратні, кругові, прямокутні, фігурні);

- структури і структурних зрушень (структурні діаграми – стовпчикові, стрічкові, секторні);

- зображення динаміки статистичних показників (динамічні графіки, стовпчикові, стрічкові, квадратні, кругові, фігурні, діаграми темпів, лінійні діаграми (їх різновиди радіальні));

- контролю виконання плану (лінійні графіки виконання плану, обліково-планові графіки);

- розповсюдження в просторі (картограма, картодіаграма, центрограма);

- варіаційних рядів (полігон, гістограма, кумулята, огіва, крива концентрації (Лоренца), крива Парето і т.ін.);

- взаємозв'язку і взаємозалежності (графіки кореляційної залежності).

Діаграма – креслення, на якому статистична інформація зображається за допомогою геометричних фігур, ліній або символічних знаків.

Стовпчикова діаграма – графік, на якому статистичні дані зображені у вигляді стовпчиків-прямокутників однакової ширини, розташованих вертикально на осі абсцис і будь-якої висоти. Кожний стовпець характеризує окремий об'єкт.

Різновидом стовпчикової діаграми є ***стрічкова діаграма***, для якої характерні горизонтальна орієнтація стовпчиків (смуг) і вертикальне розташування базової лінії. Смугова діаграма особливо зручна в тих випадках, коли окремі об'єкти порівняння характеризуються протилежними за знаком показниками.

Квадратні та кругові діаграми використовують для порівняння декількох абсолютних значень, при цьому сторона квадрата (радіус круга) – корінь квадратний з абсолютного значення, що характеризує явище. Всередині квадратів і кругів слід проставляти величини показників, що зображаються.

В ***об'ємних діаграмах*** (наприклад, у вигляді кубів) лімітні розміри графічного образу пропорційні кореням кубічним з порівнюваних величин.

Прямокутні діаграми використовують у випадках, коли необхідно порівняти три взаємозв'язані показники, один з яких дорівнює додатку двох інших, і показати роль кожного з них у формуванні першої величини. У разі прямокутних діаграм установлюють два масштаби: один – для множника, який беруть за основу, другий – для множника, який беруть за висоту.

У ***фігурних діаграмах*** геометричні фігури замінюють рисунками.

Структурні діаграми – діаграми співвідношення питомої ваги, які характеризують співвідношення окремих частин сукупності в їх загальному об'ємі (стовпчикові, стрічкові, секторні). Основна форма структурних діаграм – ***секторні діаграми*** – графічне зображення на площі круга, працюючим геометричним параметром в якому є величина кута між радіусами: 1% на діаграмі дорівнює 3,6°, а сума всіх кутів, яка становить 360° – 100 %.

Динамічні графіки призначені для зображення економічних явищ, що протікають у часі. У динамічних діаграмах об'єктом відображення являються процеси.

Лінійні графіки характеризують зміни явищ у часі, залежність між двома показниками. У статистиці комерційної діяльності на ринку товарів і послуг вони мають найбільше поширення.

Різновидом лінійних діаграм є ***радіальні діаграми***, які відображають процеси і явища, що періодично повторюються в часі.

Картограма – це схематична контурна карта або план місцевості, на якій окремі території залежно від величини показника позначають за допомогою штрихування, забарвлення, крапок.

Картодіаграма – поєднання контурної карти місцевості з діаграмою; її використовують для зображення розподілу досліджуваного явища, що позначають на контурній карті у вигляді спеціальних знаків-символів (квадратів, кружків тощо).

Для зображення варіаційних рядів застосовуються лінійні та площинні діаграми, побудовані в прямокутній системі координат.

При дискретній варіації ознаки графіком варіаційного ряду служить ***полігон розподілу*** – графічне зображення варіаційного ряду в прямокутній системі координат, де значення, які варіюються, відкладаються на осі абсцис, а відповідні їм частоти – на осі ординат.

Гістограма – графічне зображення інтервального варіаційного ряду, в якому на осі абсцис відкладаються варіанти, а прямокутники пропорційні по висоті частотам значень ознаки для кожного інтервалу.

Кумуляти (кумулятивні діаграми) використовують для графічного порівняння двох або більше варіаційних розподілів з рівними або нерівними інтервалами. На осі абсцис відкладають відрізки інтервалів групувань, на осі ординат – накопичені частоти або частості.

4.1. Розв'язання типових задач

Приклад 4.1

Показники виробництва телевізорів кольорового зображення деякими заводами України за певний період часу наведено в табл. 4.1.

Таблиця 4.1

Завод	Загальний випуск телевізорів, тис. шт.	Середня ціна одного телевізора, грн	Загальна вартість виготовлених телелевізорів, тис. грн
Львівський	30	650	19500
Київський	25	450	11250
Сімферопольський	15	670	10050

Побудувати графіки порівняння статистичних величин для діаграм:
1) стовпчикової; 2) стрічкової; 3) прямокутної.

Розв'язання.

1. При побудові стовпчикової діаграми на осі абсцис на однаковій відстані один від одного відкладають три рівні відрізки – основи стовпчиків. Назви заводів відображають на графіку ранжирувано. Діаграма дає наочне порівняння стовпчиків за висотою відповідно до чисельності телевізорів. Внизу під стовпчиками вказують назви об’єктів порівняння – назви заводів за місцем їх розташування (рис. 4.1).

Рис. 4.1. Стовпчикова діаграма

2. Для стрічкової діаграми характерні горизонтальна орієнтація стовпчиків і вертикальне розташування базової лінії (рис. 4.2).

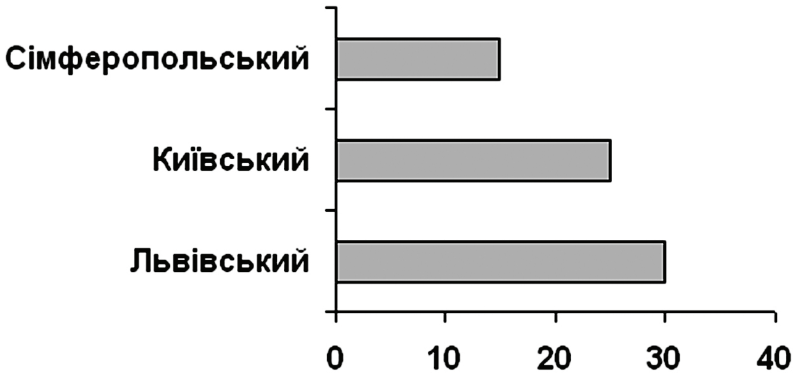

Рис. 4.2. Стрічкова діаграма

3. Для побудови прямокутної діаграми (рис. 4.3) за основу прямокутника виберемо кількість телевізорів, за висоту – середню ціну одного телевізора. Площа прямокутника – вартість усіх виготовлених телевізорів. Виберемо масштаб: для основи прямокутника (10 тис. шт. = 1 см) і висоти (200 грн = 1см).

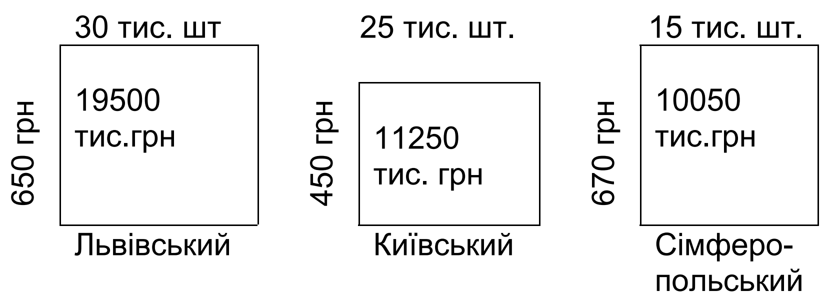

Рис. 4.3. Прямокутна діаграма: порівняльні обсяги роботи деяких заводів за кількістю виготовлених телевізорів, середньою ціною і вартістю

Приклад 4.2

Для ряду розподілу робітників за стажем (приклад 3.1, табл. 4.2) побудувати полігон і гістограму розподілу та кумулятивну гістограму.

Таблиця 4.2

Номер групи	Групи робітників за стажем, років	Число робітників, чол.	Число робітників, відсоток до підсумку
I	1- 4	5	33,4
II	4 - 7	4	26,7
III	7 - 10	1	11,5
IV	10 - 13	2	13,4
V	13 - 16	3	15,0
Разом		15	100,0

Розв'язання.

1. Для побудови полігону для інтервального ряду по осі x необхідно відкласти середини інтервалів – груп робітників за стажем, а на осі ординат – число робітників, відсоток до підсумку (рис. 4.4).

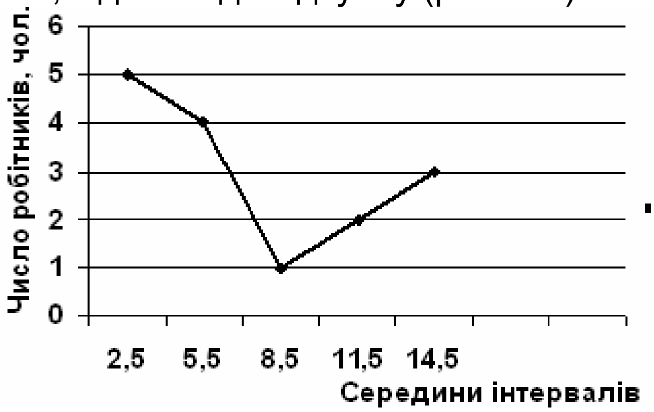

Рис. 4.4. Полігон розподілу робітників за стажем

2. Для побудови гістограми розподілу робітників за стажем (рис. 4.5) на осі абсцис відкладають межі інтервалів, на осі ординат – число робітників, відсоток до підсумку.

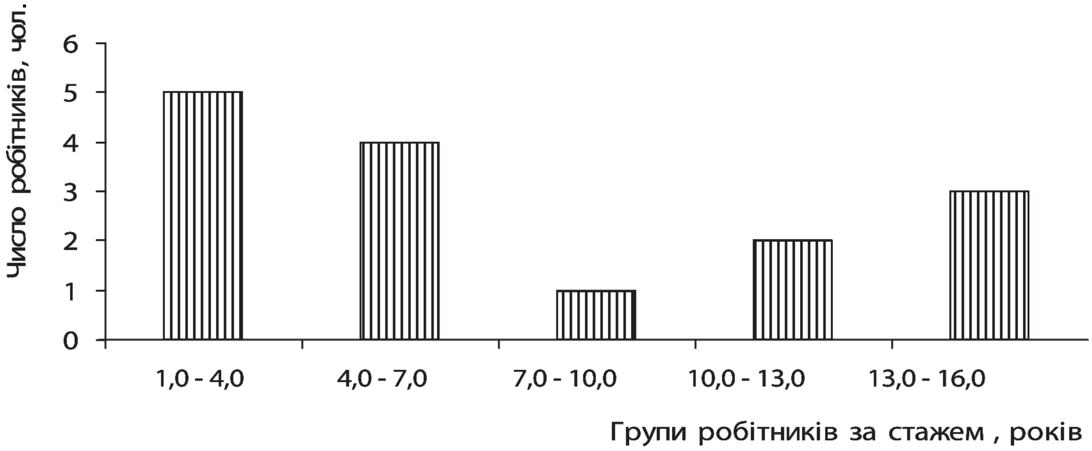

Рис. 4.5. Гістограма розподілу робітників за стажем

Утворені прямокутники пропорційні за висотою частотам значень ознаки для кожного інтервалу.

3. При побудові кумулятивної гістограми інтервальної ознаки нижній межі першого інтервалу відповідає частота, яка дорівнює нулю, а верхній – частота першого інтервалу. Верхній межі другого та наступних інтервалів відповідають їхні нагромаджені частоти, а останнього інтервалу – сума всіх частот (рис. 4.6).

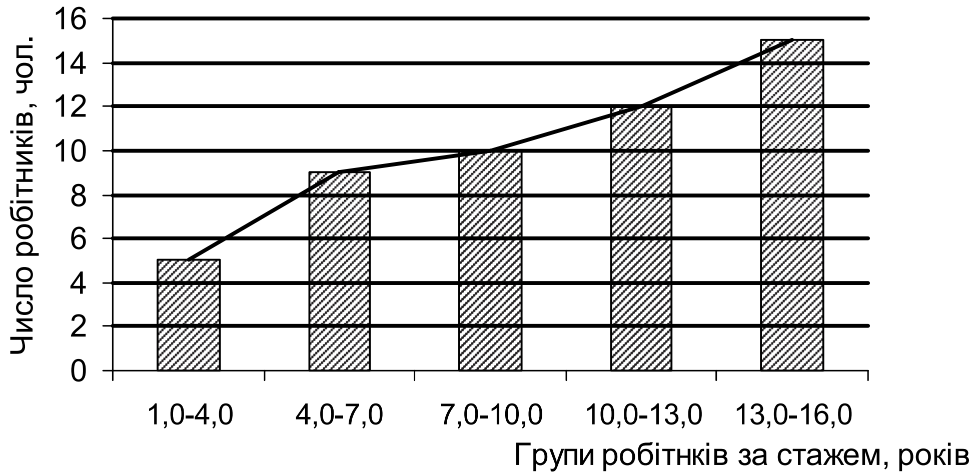

Рис. 4.6. Кумулятивна гістограма розподілу робітників за стажем

Приклад 4.3

За даними обсягів виробництва деякої продукції п'ятьма підприємствами регіону за декілька років (табл. 4.3) побудувати зображення структури явищ і структурних зрушень у вигляді: стовпчикової діаграми; стрічкової діаграми; секторної діаграми.

Таблиця 4.3

Рік	Обсяг продукції, тис. т				
	Підприємство 1	Підприємство 2	Підприємство 3	Підприємство 4	Загальний обсяг
1987	1,9	1,7	0,7	0,4	4,7
1997	3,3	3,3	1,0	1,1	8,7
2007	6,7	5,1	0,8	2,4	15,0

Розв'язання.

1. Для характеристики та ілюстрації обсягу і структури виробництва м'яса в регіоні побудуємо стовпчикову діаграму (рис. 4.7). На осі абсцис на однаковій відстані, побудуємо стовпчики, розбивши їх на частини відповідно до обсягу виробництва.

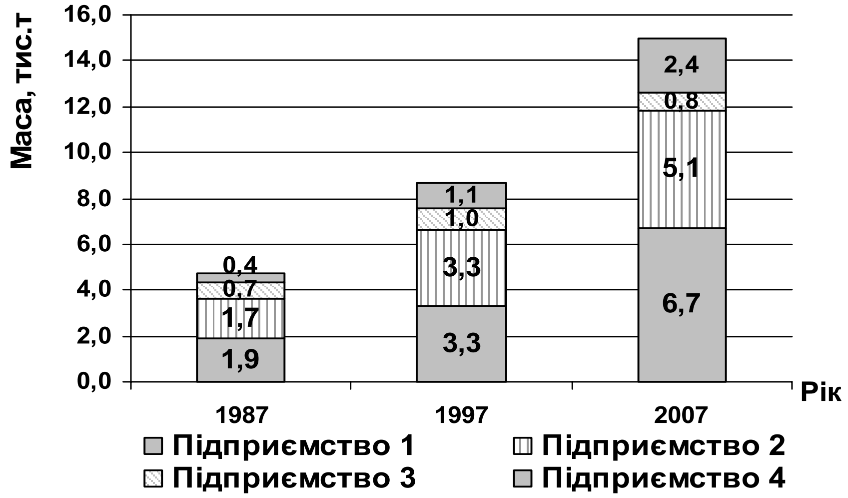

Рис. 4.7. Стовпчикова діаграма

2. Аналогічно будуємо стрічкову діаграму (рис. 4.8). Статистичні показники відображаємо перпендикулярно до осі ординат.

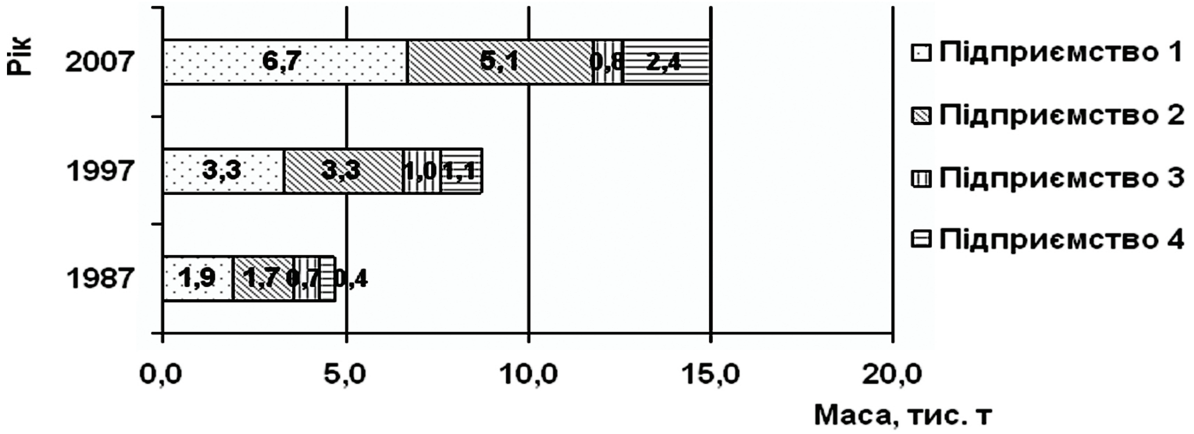

Рис. 4.8. Стрічкова діаграма

3. Для побудови секторної діаграми, яка враховує лише питому вагу частин виробництва, його загальний розмір приймають за 100%, далі розраховують частки окремих частин у процентах. Круг поділяють на сектори пропорційно до частин зображуваного цілого (рис. 4.9).

Рис. 4.9. Секторна діаграма

Контрольні запитання

1. Що таке статистичні графіки? Які вимоги до їх побудови?
2. Яке призначення графіків?
3. Які основні елементи графіків?
4. Що відображує шкала графіка і які вона має види?
5. Назвіть основні види графіків.
6. Як будують стовпчикові та стрічкові діаграми?
7. Як будують квадратні та кругові діаграми?
8. Які явища характеризують секторні діаграми і як їх будують?
9. Що таке прямокутні діаграми?
10. У чому перевага лінійних діаграм, для чого їх використовують і які правила їх побудови?

Тема 5. УЗАГАЛЬНЮЮЧІ СТАТИСТИЧНІ ПОКАЗНИКИ

5.1. Суть і класифікація статистичних показників

Статистичний показник – узагальнююча характеристика соціально-економічного явища або процесу, в якій об'єднуються якісна та кількісна визначеність останнього.

Якісний зміст показника залежить від суті явища (процесу) і знаходить своє віддзеркалення в назві (наприклад, народжуваність, прибутковість тощо).

Кількісну сторону явища відображають число і його вимірник.

Статистичні показники можна підрозділити на два основних види:

- обліково-оцінні показники (розміри, обсяги, рівні явища);
- аналітичні показники (відносні та середні величини, показники варіації тощо).

Показники розрізняються способом обчислення, ознакою часу і аналітичними функціями (табл. 5.1).

Таблиця 5.1

Класифікація показників

За якісною визначеністю	За кількісною визначеністю	За часом	За характерними властивостями
Показники властивостей конкретних об'єктів, показники статистичних властивостей будь-яких масових явищ, процесів	Первинні, вторинні	Інтервальні, моментні	Прямі, обернені

Первинні показники визначаються зведенням даних статистичного спостереження і подаються у формі абсолютних величин (наприклад, кількість і сума внесків у Ощадбанку).

Вторинні показники обчислюються на базі первинних або вторинних показників і мають форму відносних або середніх величин (наприклад, середня зарплата, індекс середньої зарплати).

Моментні показники відображають фактичну наявність або рівень явища на певний момент, дату (наприклад, наявність запасів матеріалів або оборотних коштів, величину незавершеного виробництва, чисельність проживаючих і т.д.).

Інтервальні показники – підсумковий накопичений результат за період в цілому (день, декаду, місяць, рік): наприклад, обсяг виробленої продукції за місяць або рік, приріст населення за певний період, величина валового збору зерна за рік, споживання води тощо.

Інтервальні показники залежать від довжини періоду, за який вони обчислюються. Особливістю первинних інтервальних показників є адитивність – можливість підсумовування. Вторинні показники в основному неаддитивні.

Прямий показник х зростає з посиленням явища; обернений показник *1/х,* навпаки, зменшується. Наприклад, купівельна спроможність грошової одиниці – прямий показник, ціна одиниці товару – зворотний, продуктивність роботи на одиницю часу – прямий, трудомісткість одиниці продукції – зворотний.

5.2. Абсолютні величини

Результати статистичного спостереження реєструються перш за все у формі первинних абсолютних величин.

Абсолютні статистичні величини – це показники, які виражають розміри соціально-економічних явищ і процесів у конкретних умовах місця і часу. Вони є основою для розрахунку відносних величин, аналітичних та узагальнюючих показників.

За способом виразу розмірів досліджуваних явищ розрізняють такі види абсолютних величин:

- ***індивідуальні абсолютні величини***, які характеризують кількісні ознаки у окремих одиниць (рівень виробітку окремого робітника за конкретний період);
- ***групові й загальні абсолютні величини***, які виражають значення тієї або іншої ознаки у всіх одиниць даної сукупності, разом узятих або у окремих їх груп;
- ***підсумкові абсолютні величини***, які характеризують розмір ознаки сукупності, одержаної від додавання значень ознак окремих одиниць сукупності (складання посівної площі сільськогосподарських підприємств

за даними річної звітності дає можливість знайти показник абсолютного розміру площі району).

У статистиці всі абсолютні величини є іменованими, вимірюються в конкретних одиницях і можуть бути як позитивними, так і негативними (збитки, спад, втрати і т.п.).

5.3. Види і взаємозв'язки відносних величин

Відносна величина – це узагальнюючий показник, який виражає кількісні співвідношення між соціально-економічними явищами або процесами. Оскільки багато абсолютних величин явищ взаємозв'язані, то і відносні величини одного типу в ряді випадків можуть визначатися через відносні величини іншого типу.

Загальні принципи розрахунку відносної величини:

1) наявність реальних зв'язків між порівнюваними у відносній величині показниками (наприклад, при обчисленні відносного показника, що характеризує письменність населення, зі всього населення необхідно виключити дітей до 6 років та інвалідів;

2) порівнювані початкові показники можуть розрізнятися тільки одним атрибутом: видом ознаки, часом або тільки фактичним, плановим характером показників;

3) необхідно знати можливі межі існування відносного показника.

За способом отримання відносні показники – завжди величини похідні.

Кожна відносна величина – це співвідношення, яке показує, у скільки разів порівнювана величина перевищує базисну або яку частину перша складає від другої, іноді – скільки одиниць однієї величини доводиться на 100, на 1000 і т.д. одиниць іншої – базисної величини.

Залежно від того, до якого значення прирівнюється база порівняння, частину від розподілу можна виразити у формі коефіцієнтів, відсотків, промілле, продецімілле тощо.

Якщо значення бази порівняння беруть за одиницю, то відносна величина є коефіцієнтом. Якщо значення бази порівняння беруть за 100%, то відносну величину виражають у процентах (%). Якщо базу порівняння беруть за 1000, то відносну величину виражають у промілле (‰).

Класифікація відносних величин за аналітичною функцією:

1) відношення однойменних показників:

- відносні величини виконання плану, договірних зобов'язань;
- відносні величини динаміки;
- відносні величини структури;
- відносні величини координації;

- відносні величини порівнянь;

2) відношення різнойменних показників:

- відносні величини інтенсивності.

Відносна величина виконання плану або договірних зобов'язань розраховується як відношення величини ознаки, яка спостерігається в даному періоді, до запланованої (середньої, максимальної або мінімальної) величини ознаки.

Будь-яке відхилення відносної величини від 1 або 100% (якщо обчислюють у процентах) свідчить про порушення оптимальності процесу.

На практиці розрізняють два види відносних показників виконання плану:

1) порівнюють фактичні й планові рівні;

2) у плановому завданні встановлюють абсолютну величину приросту або зниження показника, і, відповідно, перевіряють ступінь виконання плану за цією величиною.

Відносна величина динаміки і характеризує динаміку процесу та напрям й інтенсивність зміни явища в часі; визначається співвідношенням значення ознаки у визначений (поточний період або момент часу) до значення ознаки в попередній (базисний період або момент часу).

Якщо значення показника зменшується, то відносна величина буде менше одиниці.

Відносні величини структури характеризують склад, структуру, частку (відношення частини до цілого), питому вагу складових елементів сукупності за тією або іншою ознакою. Як правило, їх одержують у формі процентнів:

$$i = \frac{\text{величина частки сукупності}}{\text{сумарна величина сукупності}} 100\% = \frac{d_j}{\sum d_j} 100\% .$$

Відносні величини структури називають частками. Вони аддитивні, причому сума відносних величин структури дорівнює одиниці або 100%.

Сукупність відносних величин структури показує будову явища, що вивчається. Скільки складових величин, стільки відносних величин структури.

Різниця між відповідними частинами (частками) називається процентним пунктом (п.п.).

За допомогою відносних величин структури можна оцінити ***структурні зрушення*** – зміни у складі сукупності за певний період часу. Оцінка базується на порівнянні частин d_j за два періоди шляхом віднімання однієї з іншої.

Аналогічно можна порівнювати структури різних за обсягом сукупностей.

Відносні величини координації характеризують співвідношення окремих частин певної сукупності:

$$i = \frac{\text{величина j - й частки сукупності}}{\text{величина к - й частки сукупності}} 100\% = \frac{d_j}{d_к} 100\%.$$

Відносні величини координації показують, у скільки разів порівнювана частина сукупності є більшою чи меншою від тієї частини, яку взято за базу порівняння, або скільки одиниць однієї частини припадає на 1, 10, 100, 1000 ... одиниць іншої частини, узятої за базу порівняння.

Відносні величини порівняння – це співвідношення однойменних абсолютних величин, що відносяться до одного і того ж періоду або моменту часу, але до різних об'єктів або територій.

Частіше за все – це регіональні або міжнародні порівняння показників економічного розвитку або життєвого рівня.

Зіставляючи показники динаміки різних явищ, одержують ще один вид відносних величин порівняння – коефіцієнти випередження (відставання) за темпами зростання або приросту.

Відносні величини інтенсивності – це показники, які характеризують ступінь поширення, розвиток явища в певному середовищі. Вони завжди є відношенням двох різнойменних величин. У чисельнику – величина явища (показник), ступінь поширення якого вивчають, а в знаменнику – величина того середовища, в якому розвивається (поширюється) це явище. У кожному конкретному випадку таке співвідношення характеризує інтенсивність поширення, розвиток явища в певному середовищі. Відносні величини інтенсивності іменують одиницями вимірювань чисельника і знаменника співвідношення.

5.4. Розв’язання типових задач

Приклад 5.1

Роздрібний товарообіг державної кооперативної торгівлі області в 2007 р. становив 2260 млн грн. На 2008 р. роздрібний товарообіг планом передбачений у розмірі 2373 млн грн.

Обчислити відносну величину планового завдання області за роздрібним товарообігом на 2008 р.

Розв'язання.

Відносна величина планового завдання у відсотках розраховується як відношення планового завдання на наступний звітний період до фактичного виконання в базовому періоді:

$$i = 2373 / 2260 * 100\% = 105\%.$$

Таким чином, у 2008 р. було заплановано збільшити роздрібний товарообіг державної і кооперативної торгівлі області на 5%, порівняно з фактичним товарообігом.

Приклад 5.2

Планом на 2008 р. передбачено підвищення продуктивності праці робітників заводу на 5%. Фактично в звітному періоді вона збільшилася на 8% порівняно з 2007 р.

Визначити відносну величину виконання плану, яка характеризує зростання продуктивності праці робітників заводу.

Розв'язання.

Насамперед, заданий планом і фактичний приріст продуктивності праці в звітному періоді необхідно виразити у відсотках, включаючи базу порівняння. Припускаємо, що базисний рівень продуктивності праці в 2007 р. – 100%. Плановий відносний рівень продуктивності в 2008 році буде дорівнювати 100+5=105% (1,05), а її фактичне зростання – 100+8=108% (1,08). Відносна величина виконання плану зростання продуктивності праці обчислюється так:

$$i = 108 / 105 * 100\% = 102{,}9\% \text{ або } 1{,}08 / 1{,}05 = 102{,}9,$$

тобто план зростання продуктивності праці робітників заводу перевиконано на 2,9%.

Приклад 5.3

Дані про віковий склад населення регіону, тис. чол., наведено в табл. 5.2.

Обчислити відносні величини, які б характеризували:

1) динаміку чисельності населення;

2) його структуру за віком у кожному році та структурні зрушення;

3) співвідношення працездатного населення з чисельністю допрацездатного і старшого від працездатного віку.

Таблиця 5.2

Вікова група, років	2005	2008
0-14	192,4	211,6
15-59	469,9	586,8
60 і старше	77,7	163,6
Разом	740	962

Розв'язання.

1. Відносну величину динаміки обчислюють як відношення рівня показника в поточному році до базового. Так, загальна чисельність населення регіону в 2008 р. збільшилась порівняно з 2005 р. на 30%:

$$962 : 740 = 1{,}3, \text{ або } 130\ \%.$$

За окремими віковими групами відносні величини динаміки наведено в табл. 5.3.

Таблиця 5.3

Вікова група, років	Відносні величини					
	динаміки	структури, %		структурних зрушень, п.п.	координації	
		2005	2008		2005	2008
0-14	110,0	26,0	22,0	-4,0	40,9	36,1
15-59	124,9	63,5	61,0	-2,5	100	100
60 і старше	210,6	10,5	17,0	+6,6	16,5	27,9
Разом	130,0	100	100	0	-	-

2. Відносна величина структури – це співвідношення частини і цілого, наприклад, частка населення допрацездатного віку в 2005 р. становила 26% (192,4:740 = 0,26), а в 2007 р. – 22 % (211,6:962 = 0,22).

Отже, частка населення допрацездатного віку зменшилась на чотири процентні пункти. Результати розрахунків за іншими віковими групами наведено в табл. 5.3.

3. Співвідношення чисельності населення окремих вікових груп називають відносною величиною координації. У 2005 р. на 100 чол. працездатного віку припадало 40,9 чол. допрацездатного, в 2008 р. – 36,1, тобто навантаження населення працездатного віку дітьми зменшилось, а пенсіонерами, навпаки, збільшилось з 16,5 до 27,9.

Приклад 5.4

За даними про територію, чисельність населення і валовий внутрішній продукт двох країн (табл. 5.4) обчислити відносні величини: 1) інтенсивності; 2) порівняння.

Таблиця 5.4

Країна	Територія, тис. км^2	Чисельність населення, тис. чол.	ВВП, млн ум.од.
А	912	16390	32483
В	1285	18710	16866

Розв'язання.

1. Відносну величину інтенсивності обчислюють як співвідношення двох різнойменних показників, тобто:

$$\frac{\text{Чисельність населення, тис. чол.}}{\text{Територія, тис. км}^2} = \text{густота населення, чол./км}^2 ,$$

$$\frac{\text{ВВП, млн дол}}{\text{Чисельність населення, тис.чол.}} = \text{ВВП на душу населення, дол.}$$

Результати розрахунків наведено в табл. 5.5.

Таблиця 5.5

Країна	Густота населення, чол./км^2	ВВП на душу населення, ум. од.
А	18,0	1982
В	14,6	901

2. Відносну величину порівняння визначають як співвідношення рівнів певного показника за двома об'єктами.

Так, у країні А порівняно з країною В густота населення більше в 1,2 раза, тобто 18/14,6=1,2.

Виробництво ВВП на душу населення – в 2,2 раза, тобто

$$1982 / 901 = 2{,}2.$$

Вибір бази порівняння довільний.

Приклад 5.5

Кількість підприємств роздрібної торгівлі регіону на кінець року становила 6324. Чисельність населення цього регіону на ту ж дату становила 234,2 тис. чол. Обчислити, скільки доводиться підприємств на 10000 чоловік.

Розв'язання.

Маємо відносну величину інтенсивності:

$i = \dfrac{6324 \cdot 10000}{234200} = 270$ підприємств припадає на 10000 чол. регіону.

Контрольні запитання

1. Що таке абсолютні статистичні величини, їх значення в статистиці?
2. Які види абсолютних величин можна виділити за способом їх вираження?
3. У яких одиницях можна виражати абсолютні величини?
4. Що називають відносною величиною?
5. У якій формі можна виражати відносні величини?
6. Які види відносних величин ви знаєте?
7. Як обчислюють відносні величини планового завдання?
8. Що характеризують відносні величини динаміки?
9. Що виражають відносні величини структури та координації?
10. Для характеристики яких явищ використовують відносні величини інтенсивності та порівняння в просторі?

Тема 6. СЕРЕДНІ ВЕЛИЧИНИ

6.1. Суть і види середньої величини

Середня величина – це узагальнююча кількісна характеристика варіюючої ознаки в статистичній сукупності за конкретних умов місця і часу, яка характеризує її рівень з розрахунку на одиницю сукупності.

Середня величина відображає те загальне, що властиве всім одиницям досліджуваної сукупності. Однак вона ігнорує індивідуальні відмінності окремих одиниць сукупності, обумовлені випадковими обставинами через дію закону великих чисел.

Принципи розрахунку середніх величин:

1. Середню величину визначають для сукупностей, що складаються з якісно однорідних одиниць.
2. Середню величину визначають на підставі масових даних.
3. Середню величину слід обчислювати з урахуванням економічного змісту досліджуваного показника.
4. *Головна вимога до формули розрахунку середнього значення:* отриману середню величину слід обчислювати так, щоб при заміні кожного індивідуального значення осереднюваного показника його середньою величиною залишався без зміни деякий підсумковий зведений показник, званий визначальним, який зв'язаний тим або іншим чином з усереднюваним показником.

Види середніх величин:

1) ***степеневі середні***, до яких відносяться середня геометрична, середня арифметична, гармонійна і середня квадратична. Залежно від форми подання початкових даних середні величини можуть бути простими або зваженими;

2) ***структурні середні***, до яких відносяться мода, медіана, квартилі, децилі.

6.2. Види степеневих середніх і способи їх обчислення

Проста середня обчислюється за первинними, незгрупованими даними і має загальний вигляд

$$\overline{x}_k = \sqrt[k]{\frac{\sum_{i=1}^{n} x_i^k}{n}} = \left(\frac{\sum_{i=1}^{n} x_i^k}{n}\right)^{\frac{1}{k}},$$

де x_i – рівень (варіанта) усереднюваної ознаки; к – показник степеня середньої; n – число варіант.

Зважена середня обчислюється за згрупованими даними і має загальний вигляд

$$\bar{x}_k = \sqrt[k]{\frac{\sum_{i=1}^{m} x_i^k f_i}{\sum_{i=1}^{m} f_i}} = \left(\frac{\sum_{i=1}^{m} x_i^k f_i}{\sum_{i=1}^{m} f_i}\right)^{\frac{1}{k}},$$

де x_i – рівень (варіанта) усереднюваної ознаки або серединне значення інтервалу, в якому вимірюється варіанта; к – показник степеня середньої; n – число варіант; f_i – відповідні частоти (ваги) – кількість одиниць сукупності в різних групах, інтервалах, що показують, скільки разів зустрічається i-е значення усереднюваної ознаки, причому $\sum_{i=1}^{m} f_i = n$.

Основні формули степеневих середніх наведено в табл. 6.1.

Таблиця 6.1

Формули степеневих середніх

Ступінь	Вид середньої	Формула розрахунку	
		Проста	Зважена
-1	Гармонійна	$\bar{x} = \frac{n}{\sum_1^n \frac{1}{x_i}}$ (6.1)	$\bar{x}_{-1} = \frac{\sum_{i=1}^m z_i}{\sum_{i=1}^m \frac{z_i}{x_i}}, z_i = x_i f_i$ (6.2)
0	Геометрична	$\bar{x} = \sqrt[n]{\prod x_i} = \sqrt[n]{x_1 x_2 \ldots x_n}$ (6.3)	$\bar{x}_0 = \sqrt[n]{\prod_{i=1}^m x_i^{f_i}}, \; n = \sum_{i=1}^m f_i$ (6.4)
1	Арифметична	$\bar{x} = \frac{\sum_1^n x_i}{n}$ (6.5)	$\bar{x} = \frac{\sum_{i=1}^m x_i f_i}{\sum_{i=1}^m f_i}, \sum_{i=1}^m f_i = n$ (6.6)
2	Квадратична	$\bar{x} = \sqrt{\frac{\sum_1^n x_i^2}{n}}$ (6.7)	$\bar{x}_2 = \sqrt{\frac{\sum_{i=1}^m x_i^2 f_i}{n}}$ (6.8)

Гармонійна середня застосовується при розрахунку середніх обернених показників $\frac{1}{x_1},\ldots,\frac{1}{x_n}$. Необхідно, щоб при усереднюванні незмінної залишалася сума величин, обернених індивідуальним значенням ознаки. Зважену гармонійну застосовують тоді, коли показник, що виступає вагою f, відсутній і його слід додатково визначити на основі відомих варіант x і добутку варіант на частоту f (z=xf).

Геометрична середня застосовується, якщо визначальна особливість сукупності формується як добуток індивідуальних значень ознаки, тобто при заміні індивідуальних величин ознаки на середню необхідно зберегти незмінним добуток величин. Застосовується тільки для обчислення середніх темпів динаміки. Якщо часові інтервали неоднакові, використовують зважену середню.

Арифметична середня використовується у випадках, коли обсяг варіюючої ознаки для всієї сукупності формується як сума індивідуальних значень її окремих одиниць. Для ряду розподілу використовують зважену середню.

Квадратична середня застосовується, якщо при заміні індивідуальних величин ознаки на середню необхідно зберегти незмінною суму квадратів початкових величин. Застосовується для розрахунків показників варіації.

Якщо розрахувати всі види середніх для одних і тих же початкових даних, то діє правило мажорантності середніх: із збільшенням показника ступеня збільшується і відповідна середня величина:

$$\bar{x}_{\text{гарм}} \le \bar{x}_{\text{геом}} \le \bar{x}_{\text{арифм}} \le \bar{x}_{\text{квадр}} \le \bar{x}_{\text{куб}}.$$

Використовування кожного виду середніх залежить від характеру індивідуальних значень ознаки (прямі, обернені, квадратичні, відносні) і від характеру алгебраїчного зв'язку між індивідуальними значеннями ознаки і загальним обсягом (сума, добуток, ступінь, квадратичний корінь). Цей зв'язок є визначальною властивістю сукупності й відображається в логічній формулі усереднюваної ознаки (його економічному змісті). На підставі логічної формули вибирають вид середньої. Формально цей принцип можна записати у вигляді табл. 6.2.

Таблиця 6.2

Показники	Прямі	Зворотні
Первинні	Проста арифметична	Проста гармонійна
Вторинні	Зважена арифметична	Зважена гармонійна

У статистичній практиці найчастіше використовують середні арифметичні й середні гармонійні зважені.

У структурованій сукупності при розрахунку середньої зваженої варіантами можуть бути як окремі значення ознаки, так і групові середні $\bar{x}_i$.

Припустимо, що всі значення сукупності поділені на *к* груп. Розглядаючи кожну групу як самостійну сукупність, можна знайти її ***загальну середню*** як середню арифметичну ***групових середніх***:

$$\bar{x} = \frac{\sum_{i=1}^{k} \bar{x}_i f_i}{\sum_{i=1}^{k} f_i}, \sum_{i=1}^{k} f_i = n, \qquad (6.9)$$

де – $\bar{x}_i$ групова середня i-ї групи, обсяг якої дорівнює f_i:

$$\bar{x}_i = \frac{\sum_{j=1}^{l} x_j m_j}{n_i}, \text{где} \sum_{j=1}^{l} m_j = f_i.$$

Особливості обчислення середньої для інтервального ряду розподілу:

1. Обчислити середину кожного інтервалу (варіанту) як напівсуму нижньої і верхньої меж інтервалу.
2. Якщо ряд розподілу має відкритий інтервал (перший або останній), то його розмір умовно приймається таким, що дорівнює розміру сусіднього закритого інтервалу.
3. Середня обчислюється як середня зважена.

Обчислення нормованого середнього балуа для значень описових порядкових ознак:

1. Ранжуємо значення ознаки в порядку зростання якісної сторони ознаки. Наприклад, найнижчий ранг дають відповідям «немає», найвищий – відповідям «так». Варіанти ознаки необхідно оцифрувати порядковими рангами x_i = 1,2 ..., n або центрованими x_i = -2, -1, 0, 1, 2.
2. Нормований середній бал обчислюють за формулою

$$\bar{K} = \frac{\bar{x} - \frac{R}{2}}{x'}, \qquad (6.10)$$

де $\bar{x}$ – середньозважений ранг $\bar{x} = \frac{\sum xf}{\sum f}$; x – ранг ознаки; f – частка якісної сторони ознаки, причому $\sum f=1$ або $\sum f=100\%$; R – розмах шкали рангів, $R=x_{max}-x_{min}$, де x_{max}, x_{min} – максимальне і мінімальне значення рангів; x' – середина шкали рангів $x' = \frac{x_{max} + x_{min}}{2}$.

6.3. Структурні середні

Структурні середні застосовують для характеристики структури сукупності. До них відносять показники моди, медіани, квартилі й децилі.

Мода (Мо) – величина, яка найчастіше трапляється в даній сукупності, або варіанта, що найчастіше повторюється в ряді.

Зустрічаються ряди, які мають дві моди – бімодальний ряд або декілька мод – полімодальний ряд.

Медіана (Ме) – варіанта, що ділить ранжований ряд на дві рівні за чисельністю частини, внаслідок чого у однієї половини одиниць сукупності значення ознаки не перевищує медіанного рівня, а у іншої – не менше нього.

Особливості обчислень моди і медіани:

1. Для *дискретного* варіаційного ряду:

мода – це варіанта з максимальною частотою;

медіана при парному числі одиниць сукупності – арифметична середня величина з двох центральних варіантів, при непарному числі – це центральна варіанта, розташована в центрі ряду.

2. Для *інтервального ряду розподілу :*

Мода розраховується тільки для рівних інтервалів, оскільки від цього залежить показник повторюваності значень ознаки Х.

Для обчислення моди необхідно визначити модальний інтервал, тобто інтервал зі значенням ознаки, що найчастіше повторюється:

$$Mo = x_{mo} + h\frac{f_{mo} - f_{mo-1}}{(f_{mo} - f_{mo-1}) + (f_{mo} - f_{mo+1})}, \qquad (6.10)$$

де $X_{мо}$ і h – відповідно нижня межа і ширина модального інтервалу; f_{mo} – частоти модального інтервалу; f_{mo-1}, f_{mo+1} – частоти попереднього і наступного інтервалів відносно модального.

Для обчислення медіани визначають медіанний интервал – це интервал, кумулятивна частота (сума накопичених частот, передуючих медіанному інтервалу) якого дорівнює чи перевищує половину суми частот $0{,}5\sum_{1}^{m} f_j$. За допомогою інтерполяції в цьому медіанному інтервалі знаходять значення медіани:

$$Me = x_{Me} + h\frac{0{,}5\sum_{1}^{m} f_j \quad S_{Me\;\;1}}{f_{Me}}, \qquad (6.12)$$

де X_{Me} і h – відповідно нижня межа і ширина медіанного інтервалу; f_{me} – частота медіанного інтервалу; S_{Me-1} – кумулятивна частота передмедіанного інтервалу.

Квартилі – це варіанти, які ділять обсяг сукупності на чотири рівні частини, ***децилі*** – на десять рівних частин, **процентилі** – на 100. Ці характеристики визначаються на основі кумулятивних частот аналогічно медіані, яка є другим квартилем або п'ятим децилем.

6.4. Розв'язання типових задач

Приклад 6.1

Розрахувати середній вік студентів у групі з 10 чоловік, дані про яких наведено в табл. 6.3.

Таблиця 6.3

№ п/п	1	2	3	4	5	6	7	8	9	10
Вік	22	18	19	20	19	20	19	21	19	20

Розв'язання.

Середній вік можливо розрахувати за допомогою середніх арифметичної простої і зваженої.

Спочатку середній вік розрахуємо за формулою простої середньої арифметичної (6.5):

$$\bar{x} = \frac{22+18+19+20+19+20+19+19+21+20}{10} = \frac{197}{10} = 19{,}7 \text{ (років)}.$$

Для використання середньої арифметичної зваженої за формулою (6.6) згрупуємо початкові дані. Отримаємо дискретний ряд розподілу, наведений в табл. 6.4:

Таблиця 6.4

Вік X_i, років	18	19	20	21	22	Усього
Число студентів f_i	1	4	3	1	1	10

У результаті групування одержуємо новий показник – частоту f_i, вона показує число студентів у віці X_i років. Отже, середній вік студентів групи розраховуватиметься за формулою (6.6):

$$\bar{x} = \frac{18*1+19*4+20*3+21*1+22*1}{10} = \frac{197}{10} = 19{,}7 \text{ (років)}.$$

Приклад 6.2

За даними табл. 6.5 визначити середню частку забракованої продукції в загальному обсязі перевіреної.

Таблиця 6.5

Продовольчі товари	Забраковано товарів $z_i=x_i f_i$, т	Частка забракованої продукції в загальному обсязі перевіреної x_i,%
Товар 1	473,1	6,1
Товар 2	107,3	11,1
Товар 3	153,4	13,5
Разом	813,0	-

Розв'язання.

При визначенні середньої частки слід усвідомити економічний зміст усереднюваного показника, тобто

$$\text{частка забракованої продукції} = \frac{\text{обсяг забракованої продукції}}{\text{загальний обсяг перевіреної продукції}}.$$

Знаменник цього співвідношення відіграє роль ваги при визначенні середньої. Числове значення знаменника відсутнє і потребує розрахунку. Виходячи з наявної інформації:

$$\text{обсяг перевіреної продукції} = \frac{\text{обсяг забракованої продукції}}{\text{частка забракованої продукції}},$$

тобто $f_i = \frac{z_i}{x_i}$. Тому для обчислення середньої частки забракованої продукції треба використати формулу середньої гармонійної зваженої (6.2):

$$\bar{x} = \frac{473{,}1 + 107{,}3 + 153{,}4}{\frac{473{,}1}{0{,}061} + \frac{107{,}3}{0{,}111} + \frac{153{,}4}{0{,}135}} = \frac{733{,}8}{9858{,}7} = 0{,}074, \text{ або } 7{,}4\%.$$

Приклад 6.3

Розрахувати середню ціну товарів, реалізованих у трьох містах області, за даними табл. 6.6.

Таблиця 6.6

Місто	Ціна x_i, грн	Сума реалізації z_i, тис. грн	Частоти, $f_i = \frac{z_i}{x_i}$
А	30	600	20
Б	20	1000	50
В	35	350	10
Разом	-	1950	80

Розв'язання.

Розрахунок середньої ціни:

$$\text{середня ціна} = \frac{\text{сума реалізації}}{\text{кількість реалізованих одиниць}}.$$

Суму реалізації відомо. Оскільки вагою f_i є кількість реалізованих одиниць, яка невідома, то застосовують середню гармонійну зважену. Щоб знайти кількість реалізованих одиниць f_i, необхідно окремо розділити суму реалізації кожного виду товару на ціну x_i: $f_i = \frac{z_i}{x_i}$. Тоді середня ціна товарів за формулою (6.2):

$$\bar{x} = \frac{600 + 1000 + 350}{\frac{600}{30} + \frac{1000}{20} + \frac{350}{35}} = 24{,}3 \text{ грн.}$$

Приклад 6.4

Унаслідок інфляції споживацькі ціни за три роки зросли в 2,7 раза, у тому числі за перший рік у 1,8 раза, за другий – у 1,2, за третій – у 1,25 раза. Визначити середньорічний темп зростання цін.

Розв'язання.

Згідно з головною вимогою до формули розрахунку середнього значення, отриману середню величину слід обчислювати так, щоб при заміні кожного індивідуального значення усереднюваного показника його середньою величиною залишався без зміни деякий визначальний підсумковий показник. Середня арифметична проста, яка обчислена за формулою (6.5) (1,8+1,2+1,25)/3=1,416, не забезпечує заданих вимог: за 3 роки за цією середньою ціни виросли б у 1,416*1,416*1,416=2,84, а не в 2,7 раза. Задану умову $\prod_{i=1}^{m} x_i = 2{,}7$ може забезпечити лише геометрична середня проста за формулою (6.3):

$$\bar{x} = \sqrt[3]{1.8 * 1.2 * 1.25} = 1{,}394\,.$$

Приклад 6.5

Визначити середній бал задоволення населення якістю медичного обслуговування за наведеними в табл. 6.7 даними.

Таблиця 6.7

Задоволені ви якістю медичного обслуговування?	Кількість відповідей, %, до підсумкового значення
Так, повністю	15
Частково	49
Ні	36
Разом	100

Розв'язання.

Нормований середній бал знаходять для значень описових порядкових ознак. Ранжуємо значення ознаки в порядку зростання якісної сторони ознаки. Найнижчий ранг «один» дамо відповідям «немає», найвищий «три» – відповідям «так».

Нормований середній бал $\bar{K}$ обчислюють за формулою (6.9).

Середньозважений ранг:

$$\bar{x} = \frac{\sum xw}{\sum w} = \frac{15 * 3 + 49 * 2 + 36 * 1}{100} = 1{,}79\,.$$

Розмах шкали рангів: R=xmax-xmin=3-1=2.

Середина шкали рангів:

$$x' = \frac{x_{max} + x_{min}}{2} = \frac{3+1}{2} = 2\,.$$

Отже, $$\bar{K} = \frac{1{,}79 - 1}{2} = 0{,}395, \text{ або } 39{,}5\%.$$

Таким чином, рівень задоволення якістю медичного обслуговування населення становить в середньому 39,5%.

Приклад 6.6

За даними табл. 6.8 визначити середні значення собівартості виробу, моду і медіану.

Таблиця 6.8

Групи підприємств	Собівартість одного виробу, тис. грн	Число підприємств, %	Обсяг продукції, %	Витрати на виробництво, %
1	110-115	8	9	8,2
2	115-120	16	18	17,2
3	120-125	24	24	23,9
4	125 і вище	52	49	50,7
Разом	-	10	100	100

Розв'язання.

Маємо інтервальний ряд. Обчислимо середину кожного інтервалу як напівсуму його нижньої і верхньої меж. Оскільки останній інтервал відкритий, то його розмір умовно беремо таким, що дорівнює попередньому інтервалу. Таким чином, маємо 112,5; 117,5;122,5; 127,5.

Обсяг продукції розглянемо як вагу. Таким чином, середнє значення собівартості одного виробу обчислюється як середня зважена арифметична за формулою (6.6):

$$\bar{x} = \frac{112{,}5 * 9 + 117{,}5 * 18 + 122{,}5 * 24 + 127{,}5 * 49}{100} = 123{,}15 \text{ тис. грн.}$$

Медіану обчислимо за формулою (6.11). У прикладі можуть бути отримані три медіанних значення: згідно з ознаки – за кількістю підприємств, обсягом продукції і загальною сумою витрат на виробництво:

1. Медіанний інтервал – 125 і вище, тоді

$$Me1 = 125 + 5\frac{50 \quad (8 + 16 + 24)}{52} = 125{,}19 \text{ тис. грн.}$$

Таким чином, у половини підприємств рівень собівартості одиниці продукції перевищує 125,19 тис. грн.

2. Медіанний інтервал – 120-125:

$$Me2 = 120 + 5\frac{50 - (9 + 18)}{24} = 24{,}79 \text{ тис. грн.}$$

Отже, половина всього обсягу продукції має рівень витрат на виріб більше 124,79 тис. грн.

3. Медіанний інтервал – 125 і вище:

$$Me3 = 125 + 5\frac{50 \quad 49{,}3}{50{,}7} = 125{,}07 \text{ тис. грн.}$$

Таким чином, утворюється 50 % загальної суми витрат, якщо рівень собівартості одного виробу вище 125,07 тис. крб.

Для нашого прикладу можна розрахувати три модальних значення за формулою (6.10) виходячи з ознак числа підприємств, обсягу продукції і суми витрат. У всіх трьох випадках модальний інтервал один і той же, оскільки для одного і того ж інтервалу виявляються найбільшими і число підприємств, і обсяг продукції, і загальна сума витрат на виробництво:

$$Mo1 = 125 + 5\frac{52-24}{(52-24)+(52-0)} = 125{,}75 \text{ тис. грн;}$$

$$Mo2 = 125 + 5\frac{49-24}{(52-24)+(52-0)} = 126{,}69 \text{ тис. грн;}$$

$$Mo3 = 125 + 5\frac{50{,}7-23{,}9}{(50{,}7-23{,}9)+(50{,}7-0)} = 126{,}73 \text{ тис. грн.}$$

Таким чином, найчастіше фіксують підприємства з рівнем собівартості 126,75 тис. грн, частіше за все випускається продукція з рівнем витрат 126,69 тис. грн і найчастіше витрати на виробництво пояснюються рівнем собівартості в 123,73 тис. грн.

Контрольні запитання

1. Що таке середня величина?
2. Які є види середніх величин?
3. Коли використовують середню арифметичну?
4. Які розрізняють види середньої арифметичної?
5. Які умови використання середньої гармонійної?
6. Які розрізняють види середньої гармонійної?
7. Які види узагальнювальних величин називають структурними середніми?
8. Що таке мода і медіана?
9. Як визначають моду в дискретному та інтервальному рядах?
10. Назвіть особливості визначення медіани в дискретному й інтервальному рядах.

Тема 7. ПОКАЗНИКИ ВАРІАЦІЇ

7.1. Поняття варіації і її основні показники

Середня величина – це узагальнювальна характеристика ознаки статистичної сукупності. Проте вона не пояснює, як групуються навколо неї окремі значення, лежать вони поблизу або значно відхиляються від середньої. Коливність окремих значень ознаки характеризують показники варіації.

Варіація – це кількісні зміни ознаки в межах однорідної сукупності, обумовлені впливом різних факторів.

Ступінь близькості даних окремих одиниць до середньої вимірюється системою показників варіації, до якої відносяться абсолютні, середні та відносні показники. До абсолютних відносяться: варіаційний розмах, середнє лінійне і середнє квадратичне відхилення, дисперсії. Відносні характеристики надані рядом коефіцієнтів варіації, нерівномірності, локалізації, концентрації.

Абсолютні та середні показники варіації

Розмах варіації (варіаційний розмах) R – різниця між максимальним (X_{max}) і мінімальним (X_{min}) значеннями варіант:

$$R = X_{max} - X_{min}. \tag{7.1}$$

Розмах варіації фіксує лише крайні відхилення ознаки. Повторюваність проміжних значень тут не враховується.

Для узагальнюючої характеристики розподілу відхилень обчислюють ***середнє лінійне відхилення*** варіаційного ряду – середню арифметичну з абсолютних величин відхилень варіантів від їх середньої:

$$\overline{d} = \frac{\sum_{i=1}^{n} |x_i - \overline{x}|}{n} \text{ (дані незгруповані),} \tag{7.2}$$

$$\text{або } \overline{d} = \frac{\sum_{i=1}^{m} |x_i - \overline{x}| f_i}{\sum_{i=1}^{m} f_i} \text{ (дані згруповані).} \tag{7.3}$$

За допомогою середнього лінійного відхилення аналізують, наприклад, склад працюючих, ритмічність виробництва, рівномірність поставок матеріалів; розробляють системи матеріального стимулювання.

На практиці ступінь варіації об'єктивніше відображає ***дисперсія*** σ^2 – середня арифметична з суми квадратів відхилень окремих варіантів від їх середньої:

$$\sigma^2 = \frac{\sum_{i=1}^{n} (x_i - \overline{x})^2}{n} \text{(дані незгруповані);} \tag{7.4}$$

$$\text{або } \sigma^2 = \frac{\sum_{i=1}^{m} (x_i - \overline{x})^2 f_i}{\sum_{i=1}^{m} f_i} \text{ (дані згруповані).} \tag{7.5}$$

Обчислення дисперсії можливе за формулою

$$\sigma^2 = \overline{x^2} - \overline{x}^2, \tag{7.6}$$

де $\overline{x^2} = \frac{\sum_{i=1}^{m} x^2{}_i f_i}{\sum_{i=1}^{m} f_i}$.

Середнє квадратичне відхилення σ – корінь квадратний з дисперсії:

$$\sigma = \sqrt{\sigma^2}\,. \tag{7.7}$$

Середнє квадратичне відхилення є мірилом надійності середньої: чим менше σ, тим краще середня арифметична відображає собою всю сукупність.

Дисперсія частки як альтернативної ознаки обчислюється як добуток часток:

$$D = d_1 d_0 = d_1(1 - d_1),$$

де d_0 – частка елементів сукупності, яким відповідає ознака; d_1 – частка решти елементів, d_0=1- d_1.

Для забезпечення порівняння абсолютних показників варіації у варіаційних рядах різних явищ обчислюють відносні показники – ***коефіцієнти варіації***. Вони дозволяють порівнювати характер розсіювання в різних розподілах (різні одиниці спостереження однієї й тієї ж ознаки в двох сукупностях, при різних значеннях середніх, при порівнянні різнойменних сукупностей).

Квадратичний коефіцієнт варіації є найпоширенішим показником, що використовується для оцінки однорідності сукупності, тобто надійності та типовості середньої величини:

$$V = \frac{\sigma}{\bar{x}} \cdot 100\%,\ (\bar{x} \neq 0). \tag{7.8}$$

Сукупності, які мають коефіцієнт варіації більше 30...35 %, прийнято вважати неоднорідними.

Коефіцієнт осциляції відображає відносну коливальність крайніх значень ознаки навколо середньої:

$$K_o = \frac{R}{\bar{x}} 100\%,\ (\bar{x} \neq 0).$$

Відносне лінійне відхилення (лінійний коефіцієнт варіації) характеризує частку усередненого значення ознаки абсолютних відхилень від середньої величини

$$K_d = \frac{\bar{d}}{\bar{x}} 100\%,\ (\bar{x} \neq 0).$$

Види дисперсій і правила їх додавання

Варіація ознаки формується під впливом різних причин, серед яких можна виділити випадкові й систематичні. Отже, варіація може бути випадковою і систематичною. Визначити кожну з них можна за допомогою дисперсійного аналізу.

Дисперсія є аддитивною величиною, тобто в структурованій сукупності, яка була розподілена на групи за факторною ознакою *x*, дисперсія результативної ознаки (***загальна***) може бути розкладена на ***середню з групових десперсій*** і ***міжгрупову*** дисперсію:

$$\sigma^2 = \overline{\sigma^2{}_i} + \delta^2. \tag{7.9}$$

Загальна дисперсія характеризує загальну варіацію ознаки під впливом усіх умов і причин, що зумовили цю варіацію:

$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2 f_i}{\sum f_i}, \tag{7.10}$$

де $\bar{x}$ – загальна середня для всієї сукупності.

Середня внутрішньогрупових дисперсій $\overline{\sigma^2{}_i}$ характеризує випадкову варіацію в кожній окремій групі. Ця варіація виникає під впливом інших, не врахованих в групуванні факторів, і не залежить від ознаки-фактора, покладеної в основу групування. Обчислюється як середня арифметична зважена з групових дисперсій $\sigma^2{}_i$:

$$\overline{\sigma^2{}_i} = \frac{\sum \sigma^2{}_i f_i}{\sum f_i}, \tag{7.11}$$

де f_i – обсяг i-ї групи, $\sum\limits_{i=1}^{m} f_i = n$;

$\sigma^2{}_i$ – групова дисперсія, що відображає варіацію ознаки лише за рахунок умов, що діють усередині i-ї групи:

$$\sigma_i^2 = \frac{\sum\limits_{j=1}^{f_i} (x_{ji} - \bar{x}_i)^2}{f_i}. \tag{7.12}$$

Міжгрупова дисперсія характеризує варіацію результативної ознаки під впливом груповальної ознаки-фактора X і є дисперсією групових середніх щодо загальної середньої:

$$\delta^2 = \frac{\sum\limits_{i=1}^{m} (\bar{x}_i - \bar{x}_o)^2 f_i}{\sum\limits_{i=1}^{m} f_i}, \tag{7.13}$$

де f_i – обсяг i-ї групи, $\sum_{i=1}^{m} f_i = n$; $\bar{x}_i$ – середня кожної окремої групи; $\bar{x}_o$ – загальна середня всієї сукупності.

7.2. Розв'язання типових задач

Приклад 7.1

Є дані про розподіл робітників за тарифними розрядами (табл. 7.1).

Таблиця 7.1

Тарифний розряд	2	3	4	5	6
Число робітників	1	2	6	8	3

Визначіть:

а) дисперсію;

б) середнє квадратичне відхилення;

в) квадратичний коефіцієнт варіації.

Розв'язання.

Для розрахунку показників варіації необхідно попередньо визначити середню величину $\bar{x}$ за формулою середньої арифметичної зваженої (6.6). Дисперсію σ^2 визначимо за формулою (7.5). Обчисливши зазначені показники варіації, необхідні розрахунки зведемо до табл. 7.2.

Таблиця 7.2

Тарифний розряд x	Число робітників f, чол.	xf	$x-\bar{x}$	$(x-\bar{x})f$	$(x-\bar{x})^2 f$
2	1	2	-2,5	-2,5	6,25
3	2	6	-1,5	-3,0	4,50
4	6	24	-0,5	-3,0	1,50
5	8	40	0.5	4,0	2.00
6	3	18	1,5	4,5	6,75
Разом	20	90	-	-	21,00

Визначимо показники:

$$\bar{x} = \frac{90}{20} = 4{,}5 \text{ розряду;}$$

$$\sigma^2 = \frac{21}{20} = 1{,}05;$$

$$\sigma = \sqrt{\sigma^2} = 1{,}025 \text{ розряду.}$$

Квадратичний коефіцієнт варіації обчислимо за формулою (7.8):

$$V = \frac{1{,}025}{4{,}5} 100 = 22{,}7\,\%.$$

Він свідчить про однорідність сукупності робітників.

Контрольні запитання

1. Що розуміють під варіацією ознаки?
2. Як вимірюють варіацію ознаки?
3. Які використовують показники для вимірювання варіації?
4. Яке місце в цій системі належить показникам дисперсії, середнього квадратичного відхилення та коефіцієнт варіації?
5. Які спрощені способи визначення дисперсії?
6. Які розрізняють види дисперсії?
7. Назвіть сутність кожного виду дисперсії та послідовність їх визначення.
8. Чим відрізняються внутрішньогрупова та міжгрупова дисперсії від загальної дисперсії?
9. У чому полягає суть правила додавання дисперсій?

Тема 8. ВИБІРКОВЕ СПОСТЕРЕЖЕННЯ

8.1. Вибіркове спостереження та його основні задачі

Вибіркове спостереження **–** це вид несуцільного спостереження, за характеристикою відібраної частини одиниць якого оцінюють усю сукупність.

Сукупність, з якої відбираються елементи для обстеження, називають ***генеральною,*** а сукупність, яку безпосередньо обстежують, – ***вибірковою.*** Статистичні характеристики вибіркової сукупності розглядають як ***оцінки*** відповідних характеристик генеральної сукупності. Оскільки вибіркова сукупність неточно відтворює структуру генеральної, то вибіркові оцінки також не збігаються з характеристиками генеральної сукупності. Розбіжності між ними називають ***помилками репрезентативності (граничними похибками).*** Залежно від причин виникнення ці похибки поділять на систематичні та випадкові. ***Систематичні похибки*** виникають унаслідок порушення принципів випадковості відбору. ***Випадкові похибки*** –- це слідство випадковості відбору елементів сукупності для обстеження.

При організації вибіркового обстеження важливо запобігти виникненню систематичних похибок. Що стосується випадкових похибок, то уникнути їх неможливо, проте на підставі теорії вибіркового методу можна визначити їх розмір і в міру можливості регулювати.

У генеральної сукупності частка одиниць, якій властива ознака, що вивчається, називається ***генеральною часткою*** р, а середня величина ознаки – ***генеральною середньою***. У вибіркової сукупності частку одиниць, якій властива ознака, що вивчається, називають ***вибірковою ча-***

стокою (частістью) w, а середню величину ознаки у виборці – ***вибірковою середньою*** $\overline{x}$.

Розрізняють такі види вибіркового спостереження:

1) ***власне випадкова вибірка*** передбачає випадковий відбір одиниць з генеральної сукупності. Це класичний спосіб формування вибіркової сукупності й саме на ньому ґрунтується теорія вибіркового методу;

2) ***механічна вибірка*** – це послідовний вибір одиниць через рівні проміжки за їхнім розташуванням у генеральній сукупності або в будь-якій іншій послідовності. Відбір елементів здійснюється через однакові інтервали, крок інтервалу залежить від частості вибірки. Так, при n/N =0,05 (де n, N – обсяг вибіркової та генеральної сукупності) крок інтервалу дорівнює 1/0,05=20;

3) ***типова (районована) вибірка*** передбачає попередню структуризацію генеральної сукупності на однорідні групи за певною ознакою і незалежний відбір елементів у кожній складовій частині випадковим або механічним способом. Обсяг розшарованої вибірки n – це сума часткових вибірок n_i, тобто $n = \sum_1^m n_i$, де m – число складових частин (груп, типів, районів тощо);

4) ***серійна (гніздова) вибірка*** складається з серій елементів сукупності, зв'язаних територіально (райони, селища), організаційно (фірми, акціонерні суспільства) і тому подібне. Серії відбираються за схемою механічної або простої випадкової вибірки, обстеженню підлягають усі елементи серії.

8.2. Знаходження середньої й граничної похибок і необхідної чисельності вибірки

Можливі розходження між середніми величинами або частками ознаки вибіркової і генеральної сукупності вимірюються ***середніми похибками репрезентативності (стандартом) μ*** (табл. 8.1):

Таблиця 8.1

Спосіб відбору	Повторна вибірка	Безповторна вибірка
Визначення середньої	$\mu = \sqrt{\frac{\sigma^2}{n}}$ (8.1)	$\mu = \sqrt{\frac{\sigma^2}{n}\left(1 - \frac{n}{N}\right)}$ (8.2)
Визначення частки	$\mu = \sqrt{\frac{w(1-w)}{n}}$ (8.3)	$\mu = \sqrt{\frac{w(1-w)}{n}\left(1 - \frac{n}{N}\right)}$ (8.4)

де σ^2 – вибіркова дисперсія; n – чисельність вибірки; N – чисельність генеральної сукупності; w – частка одиниць, які мають певну ознаку.

При практичному використовуванні наведених формул слід враховувати, що:

1) дисперсія частості є добутком частостей $\sigma_w{}^2 = w(1-w)$;

2) якщо вибірка малочисельна (менш 30 одиниць), то в формулу помилки вибірки вноситься поправка $\frac{n}{n-1}$;

3) коректуючий множник для безповторної вибірки $\sqrt{\frac{N-n}{N-1}}$ при малих величинах $\frac{n}{N} \le 0{,}05$ наближається до 1, а тому при 1–5%-ній вибірці розрахунок ведеться за формулою для повторної вибірки.

Гранична похибка вибірки Δ – це максимально можлива похибка для прийнятої імовірності р, якій відповідає коефіцієнт довіри – t-разове значення μ. Коефіцієнт довіри вказує, скільки середніх помилок міститься в граничній помилці. Так, t=1 – для імовірності 0,683; t=2 – для імовірності 0,954; t=3 – для 0,997, t=4 – для 0,999.

Формула граничної похибки має вигляд

$$\Delta = t\mu. \tag{8.5}$$

Для розрахунку граничної похибки вибірки застосовують формули, наведені в табл. 8.2.

Таблиця 8.2

	Повторна вибірка	Безповторна вибірка
Визначення середньої	$\Delta_{\overline{x}} = t\sqrt{\frac{\sigma^2}{n}}$ (8.6)	$\Delta_{\overline{x}} = t\sqrt{\frac{\sigma^2}{n}\left(1-\frac{n}{N}\right)}$ (8.7)
Визначення частки	$\Delta_w = t\sqrt{\frac{w(1-w)}{n}}$ (8.8)	$\Delta_w = t\sqrt{\frac{w(1-w)}{n}\left(1-\frac{n}{N}\right)}$ (8.9)

Похибку механічної вибірки знаходять за формулою (8.7).

При обчисленні похибки районованої вибірки застосовують середню із групових дисперсій:

$$\overline{\sigma}^2 = \frac{\sum_1^m \sigma_i^2 n_i}{\sum_1^m n_i}, \text{ тоді } \Delta = t\sqrt{\frac{\overline{\sigma}^2}{n}(1-\frac{n}{N})}.$$

Як правило, похибка розшарованої вибірки менша, ніж механічної або випадкової вибірки. Частіше за все використовують відбір, пропорційний чисельності складових сукупності, тобто частість вибірки для всіх складових однакова.

При обчисленні похибки серійної вибірки враховують міжсерійну варіацію:

$$\Delta = t\sqrt{\frac{\delta^2}{n}(1-\frac{s}{S})},$$

де δ^2 – міжсерійна дисперсія:

$$\delta^2 = \frac{\sum_1^m (\bar{x}_k - \bar{x})^2 n_k}{\sum_1^m n_k},$$

де n_k і $\bar{x}_k$ – відповідно обсяг і середня к-ї серії s_k.

За допомогою формул граничної похибки вибірки визначають:

1) довірчі межі генеральної середньої і частки з певною ймовірністю:

– для середньої генеральної сукупності

$$\bar{x} - \Delta \le \bar{x} \le \bar{x} + \Delta; \quad (8.10)$$

– для частки одиниць w, яким притаманна ця ознака в генеральній сукупності:

$$w - \Delta \le w \le w + \Delta; \quad (8.11)$$

2) ймовірність того, що відхилення між вибірковими та генеральними характеристиками не перевищує визначену величину;

3) необхідну чисельність вибірки n, яка із заданою ймовірністю забезпечує очікувану точність вибіркових показників і при якій вибіркові оцінки мали б основні властивості генеральної сукупності (табл. 8.3):

Таблиця 8.3

	Повторна вибірка	Безповторна вибірка
Визначення середньої	$n = \frac{t^2\sigma^2}{\Delta_{\bar{x}}^2}$ (8.12)	$n = \frac{t^2\sigma^2 N}{\Delta_{\bar{x}}^2 N + t^2\sigma^2}$ (8.13)
Визначення частки	$n = \frac{t^2 w(1-w)}{\Delta_w^2}$ (8.14)	$n = \frac{t^2 w(1-w)N}{\Delta_w^2 N + t^2 w(1-w)}$ (8.15)

Для визначення обсягу вибірки n використовують оцінки дисперсій σ^2 аналогічних пробних обстежень. Якщо такі обстеження відсутні, можна скористатися співвідношенням $\sigma = \frac{1}{6}(x_{max} - x_{min})$, а для частості узяти найбільше значення дисперсії $\sigma^2 = 0{,}25$.

Якщо в основу розрахунку n покласти відносну похибку вибірки, формули відповідно модифікуються:

- для середньої $$n = \frac{t^2 V_x^2}{V_\Delta^2}; \quad (8.16)$$

- для частки $$n = \frac{t^2 q}{V_\Delta^2 p}.\qquad(8.17)$$

При порівнянні точності вибіркових оцінок використовують ***відносну похибку вибірки*** V_μ, яка показує, на скільки відсотків вибіркова оцінка відхиляється від параметра генеральної сукупності:

$$V_\mu = \frac{\mu}{\bar{x}} \cdot 100.$$

Відносну похибку вибірки можна розрахувати на підставі коефіцієнта варіації ознаки V_x :

– для повторної вибірки $$V_\mu = 100 \cdot \frac{V_x}{\sqrt{n-1}},$$

– для безповторної вибірки $$V_\mu = 100 \cdot \frac{V_x}{\sqrt{n-1}} \sqrt{1 - \frac{n}{N}}.$$

Аналогічно розраховують відносну похибку вибірки для частості:

$$V_\mu = \frac{\mu_p}{p} = \frac{\sqrt{pq/n}}{p} = \sqrt{\frac{q}{np}}.$$

8.3. Розв'язання типових задач

Приклад 8.1

Для дослідження оснащення заводів основними виробничими фондами було проведено 10%-не вибіркове обстеження, у результаті якого отримано такі дані про розподіл заводів за вартістю основних виробничих фондів (табл. 8.4).

Таблиця 8.4

Середньорічна вартість основних виробничих фондів, млн грн	До 2	2,0-4,0	4,0-6,0	Понад 6,0	Разом
Число заводів	5	12	23	10	50

Обчисліть: 1) з імовірністю 0,997 граничну похибку вибіркової середньої і межі, у яких буде знаходитися середньорічна вартість основних виробничих фондів усіх заводів генеральної сукупності;

2) з імовірністю 0,954 граничну похибку вибірки при визначенні частки і межі, у яких буде знаходитися питома вага заводів з вартістю основних виробничих фондів понад 4 млн грн;

3) відповісти, який має бути обсяг вибіркової сукупності за умови:

а) гранична похибка вибірки при визначенні середньорічної вартості основних виробничих фондів з імовірністю 0,997 була б не більшою 0,5 млн грн;

б) гранична похибка частки з імовірністю 0,954 була не більш 15%.

Розв'язання.

Дано: n=50, N=500 (оскільки в умові задачі оговорено, що проведено 10%-не вибіркове обстеження, отже, генеральна сукупність – це 100% чи 500 заводів).

1. Для визначення меж генеральної середньої необхідно обчислити середню вибіркову $\bar{x}$ за формулою арифметичної зваженої (6.6) і дисперсію σ^2 за формулою (7.5), техніка разрахунку яких – у табл. 8.5.

Таблиця 8.5

Середньорічна вартість основних виробничих фондів, млн грн	Число заводів f	Середина інтервалу x	xf	$x-\bar{x}$	$(x-\bar{x})$	$(x-\bar{x})^2 f$
До 2	5	1	5	-3,52	12,39	61,95
2,0 - 4,0	12	3	36	-1,52	2,31	27,72
4,0 - 6,0	23	5	115	0,48	0,23	5,29
Понад 6,0	10	7	70	2,48	6,15	61,5
Разом	50	-	226	-	-	156,46

Тоді $\bar{x} = \frac{226}{50} = 4{,}52$ млн грн, $\sigma^2 = \frac{156.46}{50} = 3{,}13$.

Середня похибка вибірки при визначенні середньорічної вартості основних фондів за формулами (8.1)-(8.2) становить:

при повторному відборі: $\mu = \sqrt{\frac{3{,}13}{50}} \approx 0{,}25$ млн грн;

при безповторному відборі: $\mu = \sqrt{\frac{3{,}13}{50}\left(1 - \frac{50}{500}\right)} \approx 0{,}24$ млн грн.

Отже, при визначенні середньорічної вартості основних виробничих фондів можна було б припуститися середньої помилки репрезентативності в 0,25 млн грн при повторному і 0,24 млн грн при безповторному відборі в той чи іншій бік від середньорічної вартості основних виробничих фондів, що припадають на один завод у вибірковій сукупності. Обчислені дані свідчать, що при безповторній вибірці середня помилка репрезентативності (0,24) менша, ніж за тих самих умов при повторному відборі (0,25).

Обчислимо граничну помилку вибіркової середньої $\Delta_{\bar{x}}$ за формулою (8.5). У нашому прикладі: p=0,997, отже, t=3. Таким чином, $\Delta_{\bar{x}} = 3 * \mu = 3 * 0{,}25 = 0{,}75$ млн грн при повторному відборі;

$\Delta_{\bar{x}} = 3 * \mu = 3 * 0{,}24 = 0{,}72$ млн грн при безповторному відборі.

Порядок встановлення меж, у яких знаходиться середня величина досліджуваного показника в генеральній сукупності, тобто довірчий інтервал, можна визначити за формулою (8.10). Таким чином, при повторному відборі маємо

$4{,}52 - 0{,}75 \le \bar{x} \le 4{,}52 + 0{,}75$ або $3{,}77 \le \bar{x} \le 5{,}27$;

при безповторному відборі:

$4{,}52 - 0{,}72 \le \bar{x} \le 4{,}52 + 0{,}72$ або $3{,}80 \le \bar{x} \le 5{,}24$.

Ці межі можна гарантувати з імовірністю 0,997.

2. Частка заводів у вибірковій сукупності з вартістю основних виробничих фондів понад 4 млн грн становить

$$w = \frac{23+10}{50} = 0{,}66, \text{ або } 66\%.$$

Визначаємо граничну помилку для цієї частки. За умовою задачі відомо, що для ймовірності p=0,954 маємо t=2.

При повторному відборі граничну помилку для частки обчислимо за формулою (8.8):

$$\Delta_w = 2\sqrt{\frac{0{,}66(1-0{,}66)}{50}} \approx 0{,}134, \text{ або } 13{,}4\%;$$

при безповторному відборі – за формулою (8.9):

$$\Delta_w = 2\sqrt{\frac{0{,}66(1-0{,}66)}{50}\left(1-\frac{50}{500}\right)} \approx 0{,}127, \text{ або } 12{,}7\%.$$

При встановленні частки межі обчислюють аналогічно знаходженню меж для середньої величини за формулою (8.11).

Отже, з імовірністю 0,954 частка заводів з вартістю основних виробничих фондів понад 4 млн грн у генеральній сукупності буде знаходитися в межах:

при повторному відборі: $66 - 13{,}4 \le w \le 66 + 13{,}4$ або $52{,}6 \le w \le 79{,}4$;

при безповторному відборі: $66 - 12{,}7 \le w \le 66 + 12{,}7$ або $53{,}3 \le w \le 78{,}7$.

Розрахунки свідчать про те, що при безповторному відборі похибка вибірки менша, ніж за тих самих умов при повторному відборі.

3. а) Відомо, що N=500; $\Delta_{\bar{x}}$=0,5 млн грн; σ^2=3,13; p=0,997; t=3.

Знайдемо обсяг вибірки для розрахунку середнього розміру ознаки:

при повторному відборі – за формулою (8.12):

$$n = \frac{3^2 \cdot 3{,}13}{0{,}5^2} \approx 113 \text{ заводів};$$

при безповторному відборі – за формулою (8.13):

$$n = \frac{3^2 \cdot 3{,}13^2 \cdot 500}{0{,}5^2 \cdot 500 + 3^2 \cdot 3{,}13^2} \approx 92 \text{ заводи}.$$

Висновки: чисельність вибірки збільшиться, якщо за інших рівних умов зменшити граничну помилку.

б) Відомо, що N=500; Δ_w=0,15 млн грн; σ^2=3,13; p=0,954; t=2.

Обчислимо обсяг вибірки для розрахунку похибки частки:

при повторному відборі – за формулою (8.14):

$$n = \frac{2^2 \cdot 0{,}66 \cdot 0{,}34}{0{,}15^2} \approx 40 \text{ заводів;}$$

при безповторному відборі – за формулою (8.15):

$$n = \frac{2^2 \cdot 0{,}66 \cdot 0.34 \cdot 500}{0{,}15^2 \cdot 500 + 2^2 0.66 \cdot 0.34} \approx 37 \text{ заводів.}$$

Висновки: чисельність вибірки збільшиться, якщо за інших рівних умов збільшити граничну похибку вибірки.

Контрольні запитання

1. Яке спостереження називають вибірковим? Його суть і завдання.
2. У чому переваги вибіркового спостереження порівняно з суцільним?
3. Яких умов необхідно дотримуватися у разі відбору одиниць при вибірковому спостереженні?
4. Чому при вибірковому спостереженні завжди виникають похибки? Як їх класифікують і що вони характеризують?
5. Які ви знаєте види і способи відбору до вибіркової сукупності?
6. Як здійснюють простий випадковий, механічний, районований відбір?
7. На що вказує коефіцієнт довіри і як його позначають?
8. Чи впливає обсяг вибірки на її точність і якою мірою?
9. Як визначають похибку вибірки для середньої і частки?
10. Як визначають необхідну чисельність вибірки для середньої і частки в разі повторного і безповторного відбору?

Тема 9. СТАТИСТИЧНІ МЕТОДИ ВИВЧЕННЯ ВЗАЄМОЗВ'ЯЗКІВ

9.1. Основні поняття і категорії

Усі соціально-економічні явища взаємозв'язані. Зв'язок між ними має причиново-наслідковий характер. Ознаки, які характеризують причини і умови зв'язку, називаються ***факторними x,*** а ті, які характеризують наслідки зв'язку, – ***результативними y.*** Між ознаками *x* і *y* виникають різні за природою і характером зв'язки, зокрема: функціональні та стохастичні. При ***функціональному зв'язку*** певному значенню факторної ознаки x відповідає чітко визначене значення результативної ознаки *y*. Цей зв'язок виявляється однозначно у кожному конкретному випадку. При ***стохастичному зв'язку*** кожному значенню ознаки *x* відповідає певна множина значень у, які створюють так званий умовний розподіл. Як закон цей зв'язок виявляється тільки в масі випадків і характеризується зміною умовних розподілів у. Якщо замінити умовний розподіл середньою величиною *у*, то утворюється різновид стохастичного зв'язку – ***кореляційний.*** У разі кореляційного зв'язку кожному значенню ознаки *x* відповідає середнє значення результативної ознаки у.

Умовні розподіли можна замінити середніми значеннями результативної ознаки, які обчислюються як середня арифметична зважена.

Поступова зміна середніх $\bar{y}_j$ від однієї групи до іншої свідчить про наявність кореляційного зв'язку між ознаками.

Характеристикою кореляційного зв'язку є ***лінія регресії*** *у* на *x* – це функція, яка зв'язує середні значення ознаки *у* зі значеннями ознаки x. Лінія регресії розглядається в двох моделях: аналітичного групування і регресійного аналізу. В моделі ***аналітичного групування*** – це емпірична лінія регресії, яка складається з групових середніх значень результативної ознаки $\bar{y}_j$, для кожного значення (інтервалу) x_j.

Ефекти дії *x* на *y* визначаються як відношення приростів середніх групових значень

$$\Delta_y / \Delta_x, \text{ де } \Delta_y = \bar{y}_j - \bar{y}_{j-1}, \quad \Delta_x = x'_j - x'_{j-1}. \tag{9.1}$$

Оцінка щільності зв'язку ґрунтується на правилі складання дисперсій. В моделі аналітичного групування мірою щільності зв'язку є відношення міжгрупової дисперсії δ^2 до загальної σ^2, яке називають ***кореляційним відношенням:***

$$\eta^2 = \frac{\delta^2}{\sigma^2}, \tag{9.2}$$

де σ^2 – загальна дисперсія, яка вимірює варіацію результативної ознаки у, обумовлену дією всіх можливих чинників:

$$\sigma^2 = \frac{\sum (y_i - \bar{y})^2 f_i}{\sum f_i}, \tag{9.3}$$

або за формулою

$$\sigma^2 = \overline{y^2} - \bar{y}^2 = \frac{\sum (y_i)^2 f_i}{\sum f_i} - \left(\frac{\sum y_i f_i}{\sum f_i} \right)^2,$$

де δ^2 – міжгрупова дисперсія, яка вимірює варіацію результативної ознаки *у* за рахунок дії тільки груповальної ознаки *x*:

$$\delta^2 = \frac{\sum\left(\bar{y}_j - \bar{y}\right)^2 f_j}{\sum f_j}. \tag{9.4}$$

Кореляційне відношення коливається від 0 до 1 або від 0 до 100%. За відсутності зв'язку маємо η^2=0, а за умови функціонального зв'язку – η^2=1. Чим більше це відношення наближається до одиниці, тим більший щільний зв'язок.

Щільний зв'язок може виникнути випадково, тому необхідно перевірити наявність його щільності, тобто довести невипадковість зв'язку. ***Перевірка щільності зв'язку*** – це порівняння фактичного значення η^2, обчисленого за формулою (9.2), з його критичним значенням $\eta^2_{1-\alpha}(k_1,k_2)$ для певного рівня щільності α і числа ступенів свободи k_1=m-1 і k_2=n-m, де m – число груп; n – обсяг сукупності.

Критичне значення є тим максимально можливим значенням кореляційного відношення, яке може виникнути випадково за відсутності кореляційного зв'язку.

Якщо $\eta^2 > \eta^2_{1-\alpha}(k_1,k_2)$, то зв'язок між результативною і факторною ознаками вважається істотним. Якщо $\eta^2 < \eta^2_{1-\alpha}(k_1,k_2)$, то наявність кореляційного зв'язку між результативною і факторною ознаками не доказана, і зв'язок вважається неістотним.

Критичне значення вибирають таким чином, щоб імовірність отримання значення η^2, більшого ніж критичне (за умови відсутності зв'язку між ознаками), була достатньо малою. Таку імовірність називають рівнем істотності α. Найчастіше зостосовують такі рівні істотності, як α=0,05 і α=0,01. Макет таблиці критичних значень η^2 для рівня істотності α=0,05 подано в табл. 9.1.

Таблиця 9.1

Критичні значення кореляційного відношення η^2
і коефіцієнта детермінації R^2 для рівня істотності α=0,05

k2	k1							
	1	2	3	4	5	6	8	10
5	0,569	0,699	0,764	0,806	0,835	0,854	0,885	0,904
6	0,500	0,632	0,704	0,751	0,785	0,811	0,847	0,871
7	0,444	0,775	0,651	0,702	0,739	0,768	0,810	0,839
9	0,362	0,362	0,362	0,362	0,362	0,362	0,362	0,749
10	0,332	0,451	0,527	0,582	0,624	0,659	0,711	0,332
20	0,179	0,259	0,318	0,384	0,404	0,432	0,495	0,540
30	0,122	0,182	0,227	0,264	0,297	0,326	0,373	0,419
40	0,093	0,139	0,176	0,207	0,234	0,259	0,304	0,342
…	...	...	...	...	...	...	...	...

Для перевірки істотності зв'язку використовують також функціонально звязану з η^2 характеристику F-критерію (критерій Фішера), який обчислюють за формулами

$$F = \frac{\eta^2}{1-\eta^2} \cdot \frac{k_2}{k_1} \quad \text{або} \quad F = \frac{\delta^2}{\overline{\sigma}^2} \cdot \frac{k_2}{k_1}.$$

Існують таблиці критичних значень F-критерію. Перевірку істотності зв’язку за його допомогою здійснюють аналогічно описаній вище для кореляційного відношення η^2.

9.2. Кореляційний і регресійний методи аналізу зв’язку

У ***моделі регресійного аналізу*** характеристикою кореляційного зв'язку є теоретична лінія регресії, що описується функцією Y=f(x)***,*** яка називається ***рівнянням регресії.***

На підставі рівняння регресії визначаються теоретичні значення Y, тобто значення результативної ознаки за умови дії тільки чинника *x* при незмінному рівні інших чинників.

Залежно від характеру зв'язку використовують:

лінійні рівняння $Y=a_0+a_1x$, коли зі зміною *x* ознака *y* змінюється більш-менш рівномірно;

нелінійні рівняння, коли зміна взаємозв'язаних ознак відбувається нерівномірно (з прискоренням, уповільненням або зі змінним напрямом зв'язку), зокрема: степенева $Y=ax^b$, гіперболічне $Y=a+b/x$, параболічне $Y=a_0+ a_1x+ a_2x^2$ і тому подібне.

Найчастіше застосовують лінійні рівняння, тобто рівняння, що були приведені до лінійного вигляду. В лінійному рівнянні параметр a_1 – ***коефіцієнт регресії*** – позначає, на скільки одиниць в середньому зміниться *y* із зміною *x* на одиницю. Він має одиницю вимірювання результативної ознаки. У разі прямого зв'язку a_1 – величина додатна, а при зворотному – від’ємна. Параметр a_0 – вільний член рівняння регресії, тобто це значення y при x=0. Якщо *x* не придбає нульового значення, то цей параметр буде тільки розрахунковим. Параметри визначаються ***методом найменших квадратів (МНК),*** згідно з яким сума квадратів відхилень емпіричних значень *y* від теоретичних *У* мінімальна: $\sum(y-Y)^2 \to \min$. Відповідно до умови мінімізації параметри лінійного рівняння регресії обчислюються за системою нормальних рівнянь:

$$\begin{cases} na_0 + a_1\sum x = \sum y, \\ a_0\sum x + a_1\sum x^2 = \sum xy. \end{cases}$$

Звідси

$$a_1 = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - \sum x \sum y}; \quad (9.5)$$

$$a_0 = \bar{y} - a_1\bar{x}. \quad (9.6)$$

Коефіцієнт регресії у невеликих за обсягом сукупностях (n<30) схильний до випадкових коливань. Тому здійснюється перевірка його істотності за допомогою t-критерію Стьюдента. При цьому фактичні значення t-критерію розраховують за формулою

$$t_{a_1} = \frac{a_1}{\mu_{a_1}}, \quad (9.7)$$

де a_1 – коефіцієнт регресії; μ_{a_1} – власне стандартна похибка, яка розраховується за формулою

$$\mu_{a_1} = \frac{\sigma_e}{\sigma_x}\sqrt{(n-2)}, \quad (9.8)$$

де n – обсяг сукупності; σ_x – середнє квадратичне відхилення факторної ознаки x_i від загальної середньої $\bar{x}$:

$$\sigma_x = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n}}; \quad (9.9)$$

σ_e – середнє квадратичне відхилення результативної ознаки y_i від теоретичних значеньY:

$$\sigma_e = \sqrt{\frac{\sum (y_i - Y)^2}{n}}. \quad (9.10)$$

Розраховане за формулою (9.7) фактичне значення t-критерію t_{a_1} порівнюють з критичним значенням t_k, що отримують за таблицею Стьюдента з урахуванням рівня істотності α і числа ступенів свободи k=n-2. Якщо $t_k < t_{a_1}$, то коефіцієнт регресії вважається істотним.

Характеристикою відносної зміни *у* за рахунок *х* є ***коефіцієнт еластичності***

$$K_{эл} = a_1 \frac{\bar{x}}{\bar{y}}, \quad (9.11)$$

який показує, на скільки відсотків у середньому змінюється результативна ознака зі зміною чинника на 1%.

Для статистичної оцінки щільності зв'язку використовують такі показники варіації:

1) загальна дисперсія результативної ознаки, яка відображає дію усих факторів:

$$\sigma_y^2 = \frac{1}{n}\sum_1^n \left(y - \overline{Y}\right)^2; \tag{9.12}$$

2) факторна дисперсія результативної ознаки, яка відображає варіацію *y,* обумовлену дією тільки чинника *x***:**

$$\sigma_Y^2 = \frac{1}{n}\sum_1^n (Y - \overline{y})^2 ; \tag{9.13}$$

3) залишкова дисперсія, яка відображає дію на результативну ознаку всіх інших чинників, окрім *x*:

$$\sigma^2{}_e = \frac{\sum (y_i - Y)^2}{n}.$$

Частка факторної дисперсії в загальній характеризує щільність зв'язку і називається ***коефіцієнтом детермінації***:

$$R^2 = \frac{\sigma_Y^2}{\sigma_y^2}. \tag{9.14}$$

Він має такий же сенс, інтерпретацію і цифрові межі, як і η^2.

Щільність зв'язку оцінюється також ***індексом кореляції*** $R = \sqrt{R^2}$***,*** проте інтерпретується тільки R^2. Для лінійного зв'язку використовують лінійний коефіцієнт кореляції (Пірсона) r:

$$r = \frac{\sum_1^n xy - n\overline{x}\cdot\overline{y}}{n\sqrt{\sigma_x^2 \cdot \sigma_y^2}}, \text{ або } r = \frac{\overline{xy} - \overline{x}\cdot\overline{y}}{\sigma_x \cdot \sigma_y}, \tag{9.15}$$

який набуває значення в межах ±1, тому характеризує не тільки щільність, але і напрям зв'язку. Додатне значення свідчить про прямий зв'язок, а від'ємне – про зворотний.

Абсолютне значення r дорівнює індексу кореляції:

$$|r| = R = \sqrt{R^2}.$$

Перевірка істотності зв'язку здійснюється таким же чином, як і в моделі аналітичного угрупування, шляхом порівняння фактичного і критичного значень. Відмінності мають тільки визначення k_1=m-1 і k_2=n-m, в яких m – число параметрів рівняння регресії. Для лінійної моделі k_2=2-1=1.

Істотність зв'язку в обох моделях можна здійснювати також за критерієм Фішера, який функціонально зв'язаний з R^2 і η^2. Тоді фактичне значення визначають за формулою

$$F = \frac{R^2}{1 - R^2}\cdot\frac{k_2}{k_1} \quad \text{або} \quad F = \frac{\eta^2}{1 - \eta^2}\cdot\frac{k_2}{k_1},$$

тому процедури перевірки і висновки ідентичні.

9.3. Розв'язання типових задач

Приклад 9.1

Відомі дані про розподіл проданих на біржі нерухомості однокімнатних квартир за їх середньою вартістю *y* і розміром загальної площі *x* (табл. 9.2, стовпчики 1-3), а також про загальну дисперсію $\sigma^2=7$. За допомогою аналітичного групування встановити кількісні співвідношення між середньою вартістю *y* і розміром загальної площі проданих квартир, оцінити щільність зв'язку за допомогою кореляційного відношення і перевірити істотність зв'язку між ознаками з імовірністю 0,95.

Таблиця 9.2

Загальна площа квартири x_j, м^2	Кількість квартир f_j	Середня вартість квартири $\bar{y}_j$, тис. грн	$\bar{y}_j f_j$	$\bar{y}_j - \bar{y}$	$(\bar{y}_j - \bar{y})^2 f_j$
1	2	3	4	5	6
До 25	40	10,8	432	- 2,2	193,6
25–30	30	13,2	396	0,2	1,2
30–35	24	15,2	364	2,2	116,2
35 і більше	6	18,0	108	5,0	150,0
В цілому	100	13,0	1300	X	461,0

Розв'язання.

У табл. 9.1 наведено приклад аналітичного групування проданих квартир, що описує залежність їх вартості від загальної площі. Маємо стохастичний і зокрема кореляційний зв'язок.

Аналітичне групування дає змогу встановити за формулами (9.1) кількісні співвідношення між ознаками, що вивчаються. За даними табл. 9.1 прирости Δ_x у всіх групах однакові – 5 м^2, а середня вартість проданих квартир збільшується по групах таким чином:

Δ_{y_2} =13,2-10,8=2,4 тис. грн;

Δ_{y_3} =15,2-13,2=2,0 тис. грн;

Δ_{y_4} = 18,0-15,2=2,8 тис. грн.

Отже, із збільшенням розміру загальної площі квартир на 1 м^2 їх вартість у середньому росте на:

Δ_{y2}/Δ_{x_2} =2,4 : 5 =0,48 тис. грн і на 0,4 і 0,56 тис. грн відповідно.

Оцінимо щільність зв'язку за формулою (9.2). Деякі проміжкові розрахунки наведено в табл. 9.2. Насамперед розрахуємо середню загальну вартість квартир за формулою (6.9):

$$\bar{y} = \frac{1300}{100} = 13{,}0 \text{ тис. грн.}$$

Міжгрупова дисперсія за формулою (9.5):

$$\delta^2 = \frac{461}{100} = 4{,}61.$$

Загальна дисперсія σ^2=7 відома, тому кореляційне відношення становить:

$$\eta^2 = \frac{4{,}61}{7{,}00} = 0{,}659.$$

Отже, варіація вартості проданих квартир на 65,9% пояснюється варіацією їхньої загальної площі та на 34,1% – варіацією інших чинників. Тобто зв'язок між ознаками достатньо щільний.

Проте щільний зв'язок може виникнути випадково, тому необхідно перевірити його щільність, тобто довести невипадковість зв'язку.

У нашому прикладі числа ступенів свободи k_1=4–1=3, k_2=100–4=96. Критичні значення кореляційного відношення для α=0,05 наведено в табл. 9.1. Але через відсутність у таблиці критичних значень k_2=96 використовуємо найближче (k_2=100), тоді $\eta^2_{0,95}(3, 100)$=0,075. Оскільки η^2=0,659>0,075, то зв'язок визнається істотним з вірогідністю 0,95.

Приклад 9.2

Дані, що відображають зв'язок між добовою вартістю туристичних путівок в одному з туристичних агентств і тривалістю відпочинку (днів), наведено в табл. 9.3 (1-3 стовпчики).

Таблиця 9.3

Номер путівки	Тривалість відпочинку, днів	Добова вартість путівки, грн	xy	x^2	y^2	Y=113,25–4,34x	$(y-Y)^2$
1	5	*78*	390	25	6084	91,6	185,0
2	14	55	770	196	3025	52,5	6,2
3	7	95	665	49	9025	82,9	146,4
4	18	30	540	324	900	35,1	126,0
5	14	53	742	196	2809	52,5	0,2
6	20	26	520	400	676	26,4	0,2
7	7	85	595	49	7225	82,9	4,4
8	15	50	750	225	2500	48,1	3,6
Всього	100	472	4972	1464	32244	472,0	372,0

Розрахувати параметрі лінійного рівняння регресії, дати їм економічну інтерпретацію. Визначити коефіцієнт еластичності.

Перевірити істотність коефіцієнта регресії за допомогою t-критерия Ст'юдента.

Розв'язання.

Розрахуємо параметри лінійного рівняння регресії за формулами (9.5)-(9.6). Величини, на підставі яких обчислюються параметри, дорівнюють: n=8; $\sum x=100$; $\sum y=472$; $\sum xy=4972$; $\sum x^2=1464$; $\bar{x}=100:8=12{,}5$; $\bar{y}=472:8=59$. Отже, параметри становлять

$$a_1 = \frac{8\cdot 4972 - 100\cdot 472}{8\cdot 1464 - 100\cdot 100} = -\frac{7424}{1712} = -4{,}34 \text{ грн;}$$

$$a_0 = 59 - (-4{,}34) * 12{,}5 = 113{,}25.$$

Тоді рівняння регресії має вигляд Y=113,25–4,34x, тобто зі збільшенням тривалості відпочинку на один день добова вартість туристичної путівки дешевшає в середньому на 4,34 грн.

Розрахуємо фактичне значення t-критерію t_{a_1} за формулою (9.7). Для цього, використавши розрахунки, наведені в табл. 9.3, обчислимо за формулою (9.9) – $\sigma_x = \sqrt{26{,}75}$=26,75, за формулою (9.10) – $\sigma_e = \sqrt{46{,}5}$. Тоді стандартна похибка, знайдена за формулою (9.8), становить

$$\mu_{a_1} = \sqrt{\frac{46{,}5}{26{,}75(8-2)}} = 0{,}54 \text{ грн, а } t_{a_1} = \frac{|-4{,}34|}{0{,}54} = 8,$$

що значно перевищує критичне значення $t_{0.95}(6)=2{,}54$, отримане за таблицею Стьюдента з урахуванням рівня істотності α=0,95 і числа ступенів свободи k=n-2=8-2.

Таким чином, з вірогідністю 0,95 дія тривалості відпочинку на добову вартість путівок визнається істотною. З вірогідністю 0,95 довірчі межі коефіцієнта регресії становлять: - 4,34±2,54 * 0,54 або - 4,34±1,37 грн.

Коефіцієнт еластичності за формулою (9.11):

$$K = -4{,}34\frac{12{,}5}{59} = -0{,}9.$$

Отже, зі збільшенням тривалості відпочинку на 1% добова вартість путівок зменшується в середньому на 0,9%.

Визначимо теоретичні значення Y – це очікувана вартість путівок за рахунок дії тільки тривалості відпочинку (табл. 9.3, стовпчик 6). Так, для x=5 днів добова вартість путівки становить Y=113,2–54,34 • 5=91,6 грн, це дещо відхиляється від емпіричного значення.

Теоретичні значення Y використємо для обчислення коефіцієнта детермінації за формулою (9.14). Загальна дисперсія результативної ознаки за формулою (9.12)

$$\sigma_Y^2 = \frac{4396}{8} = 549{,}5;$$

факторна дисперсія результативної ознаки за формулою (9.13):

$$\sigma_{\hat{Y}}^2 = \frac{4031}{8} = 503{,}86.$$

Тоді R^2=503б86:549,5=0,917, тобто 91,7% варіації добової вартості путівок лінійно пов'язано з варіацією тривалості відпочинку, а 8,3% варіації доводиться на решту чинників. Тому зв'язок дуже щільний.

Визначимо лінійний коефіцієнт кореляції (Пірсона) r за формулою (9.15):

$$r = \frac{4972 - 8 \cdot 12{,}5 \cdot 59}{8\sqrt{26{,}75 \cdot 549{,}5}} = -0{,}957.$$

Отже, зв'язок між добовою вартістю путівок і терміном відпочинку є щільним і зворотним. Абсолютне значення r дорівнює індексу кореляції:

$$|r| = R = \sqrt{R^2} = |-0{,}957| = \sqrt{0{,}915} = 0{,}957.$$

Перевіримо щільність зв'язку: k_1=2–1=1, а k_2=8–2=6; критичне значення $R^2_{0,95}(1,6)$ = 0,5 значно менше фактичного R^2=0,915.

Зв'язок між добовою вартістю путівок і тривалістю відпочинку визнається істотним з імовірністю 0,95.

Приклад 9.3

За даними табл. 9.3 оцінити щільність зв'язку між рівнем ефективності економіки і надійністю ділового партнерства для семи країн Східної Європи.

Таблиця 9.3

Країна	Інтегральні показники		Ранги показників		Відхилення рангів d_j	d_j^2
	ефективності економіки (max=10)	надійності ділового партнерства (max=100)	R_x	R_y		
A	5,9	54,9	6	7	- 1	1
B	7,1	54,8	7	6	1	1
C	4,2	45,3	4	5	- 1	1
D	3,4	36,9	3	4	- 1	1
K	4,9	35,8	5	3	2	4
M	2,7	26,4	1	2	- 1	1
P	2,9	24,8	2	1	1	1
Усього	X	X	X	X	0	10

Розв'язання.

Оскільки інформацію подано у формі інтегральних показників (бальної оцінки), необхідно провести ранжування країн. Найменшому значенню інтегрального показника надається ранг 1, найбільшому – ранг n=7. Сума квадратів відхилень рангів: $\sum_{1}^{n} d_j^2 = 10,$ а коефіцієнт рангової кореляції:

$$\rho = 1 - \frac{6 \cdot 10}{7\left(7^2 - 1\right)} = 1 - \frac{60}{336} = 0{,}82 .$$

Значення коефіцієнта рангової кореляції свідчить про наявність прямого і достатньо помітного зв'язку між названими параметрами ризику іноземного інвестування економіки. Згідно з таблицею критичне значення коефіцієнта рангової кореляції для α =0,05 і n=7 становить $\rho_{0,95}(7)=0,71$, що значно менше фактичного. Отже, істотність зв'язку було доведено з вірогідністю 0,95.

Контрольні запитання

1. У чому полягають основні завдання статистичного вимірювання взаємозв'язків між явищами?
2. Який зв'язок називають функціональним і як він проявляється?
3. Який зв'язок називають стохастичним і в чому полягає його суть?
4. Які ви знаєте форми кореляційної залежності? Як вони співвідносяться зі стохастичним зв'язком?
5. У чому полягає суть лінії регресії?
6. За допомогою якого методу оцінюють параметри регресійної моделі?
7. Як розраховують кореляційне відношення і що воно відображає?
8. Як перевіряють істотність зв'язку в аналітичному групуванні?
9. Як використовують для перевірки істотності зв'язку критерій Фішера?
10. У чому виражається відмінність між кореляційною і регресійною моделями?
11. Які завдання вирішують за допомогою регресійної моделі?
12. Які функції найчастіше застосовують для побудови регресійних моделей?
13. Які показники використовують для вимірювання щільності зв'язку в регресійній моделі та як їх розраховують?
14. Як обчислюють лінійний коефіцієнт кореляції? У чому виявляється зв'язок з коефіцієнтом детермінації?

Тема 10. РЯДИ ДИНАМІКИ

10.1. Ряди динаміки. Класифікація динамічних рядів

Динаміка – це зміна статистичних показників у часі.

Основна мета статистичного вивчення динаміки комерційної діяльності полягає у виявленні та вимірюванні закономірностей розвитку соціально-економічних явищ у часі.

Динамічний ряд – це статистичні дані, що відображають зміну того чи іншого соціально-економічного явища і процесу в часі.

Усякий ряд динаміки включає два обов'язкових елементи:

1) ***показники часу*** t – моменти часу або окремі періоди (роки, квартали, місяці, доба);

2) ***рівні*** розвитку явища y – конкретні значення статистичних показників, що відносяться до певних дат або окремих періодів.

Залежно від характеру рівнів ряду розрізняють моментні й інтервальні ряди.

Інтервальний ряд динаміки характеризує розміри суспільних явищ за певні періоди часу (день, місяць, квартал, рік тощо): наприклад, виробництво електроенергії за рік.

Кожний рівень інтервального ряду складається з даних за більш короткі інтервали (субперіоди) часу.

Моментний ряд динаміки характеризує обсяг явища на певні моменти часу t: наприклад, послідовності показників чисельності населення на початок року, величину запасу якого-небудь матеріалу на початок періоду і т.ін. Підсумовування рівнів моментного ряду не має сенсу, оскільки може виникнути повторний облік.

За способом вираження рівнів ряди поділяють на ***ряди абсолютних, відносних і середніх величин.***

За повнотою часу динамічні ряди поділяються на повні та неповні.

Повні ряди динаміки мають місце, коли дати реєстрації або закінчення періодів слідують одна за одною з рівними інтервалами. Це рівновіддалені ряди динаміки.

У неповних рядах *у* послідовності показників спостерігають нерівні часові інтервали.

За кількістю змінних показників ряди бувають одновимірні та багатовимірні.

Одновимірні ряди динаміки характеризують зміну в часі одного показника, а багатовимірні – двох і більше показників.

У свою чергу, багатовимірні ряди поділяють на:

- ***паралельні ряди*** динаміки – ряди, що відображають зміну в часі або одного показника для різних об'єктів (чисельність населення в різних країнах), або різних показників одного об'єкта (валовий збір пшениці, буряка та картоплі в районі);

- ***ряди взаємозв'язаних показників*** характеризують залежність одного явища від іншого (залежність зарплати від тарифного розряду).

10.2. Характеристики динамічних рядів

Для кількісної оцінки динаміки соціально-економічних явищ застосовуються статистичні показники: абсолютний приріст, темпи зростання і приросту, абсолютне значення одного відсотка приросту і т. ін.

Розрахунок показників динаміки грунтується на порівнянні рівнів ряду. База порівняння може бути постійною або змінною. Залежно від цього існують базисні та ланцюгові показники.

Рівень, який зіставляють з іншими рівнями, називається ***поточним***, а рівень, з яким зіставляють інші рівні, – ***базисним***.

Якщо кожний наступний рівень зіставляють з попереднім, то одержують ***ланцюгові показники динаміки***, а якщо кожний наступний рівень зіставляють з рівнем, узятим за базу зіставлення, то отримані показники називають ***базисними***.

Розрахунок показників динаміки наведений у табл. 10.1.

Таблиця 10.1

Показник	Базисний	Ланцюговий
Абсолютний приріст Δ_t виражає абсолютну швидкість зміни рівнів ряду динаміки і показує, на скільки одиниць підвищився чи зменшився рівень щодо базисного за певний період часу	$\Delta_{tбаз}=yt-y_1$ (10.1)	$\Delta_t^{лан}=y_t-y_{t-1}$ (10.2)
Коефіцієнт зростання Кр виражає відносну швидкість динаміки і показує, в скільки разів порівнюваний рівень більше (Кр>1) або менше (Кр<1) від базисного. Відсоткове подання коефіцієнта зростання – ***темп зростання*** (Тр)	$Кзр_{баз}=y_t / y_1$ (10.3) $Тзр_{баз}=(y_t / y_1)*100$ (10.4)	$Кзр^{лан}=y_t / y_{t-1}$ (10.5) $Тзр^{лан}=(yt / yt-1)*100$ (10.6)

Закінчення табл. 10.1

Показник	Базисний	Ланцюговий
Коефіцієнт приросту ($К_{пр}$) виражає відносну швидкість зростання і показує, в скільки разів порівнюваний рівень більше або менше від рівня, взятого за базу порівняння. Відсоткове подання коефіцієнта приросту називається ***темпом приросту*** ($Т_{пр}$)	$К_{пр} = \frac{\Delta_t^б}{у_1} = К_р^б - 1;$ (10.7) $Т_{пр} = 100К_{пр} =$ $= Т_р^б - 100$ (10.9)	$К_{пр} = \frac{\Delta_t^{лан}}{у_{t-1}} = К_р^л - 1;$ (10.8) $Т_{пр}^{лан} = 100К_{пр} =$ $= Т_р^{лан} - 100$ (10.10)
Абсолютне значення 1% приросту (At) показує вагомість одного відсотка приросту і знаходиться як співвідношення абсолютного приросту і темпу приросту за один і той же період	$А_t = \frac{\Delta_t^б}{Т^б_{пр}} = \frac{у_1}{100}$ (10.11)	$А_t = \frac{\Delta_t^{лан}}{Т_{пр}} = \frac{у_{t-1}}{100}$ (10.12)

Співвідношення між ланцюговими і базисними показниками:

1) ланцюгові та базисні абсолютні прирости адитивно зв'язані: сума ланцюгових приростів дорівнює базисному приросту останнього періоду ряду динаміки: $\Delta_{tб} = \sum_1^n \Delta_{tлла} = (у_2 - у_1) + (у_3 - у_2) + ... + (у_n - у_{n-1}) = у_n - у_1;$

2) ланцюгові та базисні коефіцієнти (темпи) зростання мультиплікативно зв'язані: добуток послідовних ланцюгових коефіцієнтів (темпів) зростання дорівнює базисному коефіцієнту (темпу) зростання:

$$К_р^{баз} = \prod_{i=1}^n К_{рі}^{лан}.$$

Взаємозв'язок абсолютної і відносної швидкостей динаміки

Якщо порівнюваний рівень *у* виразити через рівень попереднього року плюс приріст $yt = у_{t-1} + \Delta_{tлан}$ або через рівень базисного року плюс базисна абсолютна зміна $yt = у_1 + \Delta_{tбаз}$, отримаємо

$$К_р^л = \frac{у_{t-1} + \Delta_t^л}{у_{t-1}} = 1 + \frac{\Delta_t^л}{у_{t-1}} \text{ або } Т_р^{лан} = 100\% + \frac{\Delta_t^л}{у_{t-1}}100;$$

$$К_t^б = \frac{у_1 + \Delta_t^б}{у_1} = 1 + \frac{\Delta_t^б}{у_1} \text{ або } Т_{рбаз} = 100\% + \frac{\Delta_t^б}{у_1}100.$$

Якщо швидкість розвитку в межах періоду, що вивчається, неоднакова, то зіставленням однойменних характеристик швидкості визначають прискорення чи уповільнення зростання.

Абсолютне прискорення зростання $\delta_t = \Delta_t - \Delta_{t-1}$ характеризується додатною величиною $\delta_t > 0$, ***уповільнення*** – від'ємною величиною

$\delta_t < 0$. Якщо інтервали часу однакові, зіставляють базисні характеристики швидкості, якщо неоднакові – використовують середні абсолютні прирости:

$$\delta_t = \overline{\Delta}_t - \overline{\Delta}_{t-1}. \qquad (10.13)$$

Темп прискорення (уповільнення) динаміки визначають, зіставляючи абсолютні прирости: $K\Delta = \Delta_t / \Delta_{t-1}$, або якщо інтервали часу неоднакові: $K\Delta = \overline{\Delta}_t - \overline{\Delta}_{t-1}$.

Коефіцієнт прискорення (уповільнення) відносної швидкості розвитку γ_t одержуємо відношенням коефіцієнтів або темпів зростання (приросту):

$$\gamma_t = K_з^1 / K_з^2 \text{ або } \gamma_t = K_{пр}^1 / K_{пр}^2. \qquad (10.14)$$

Для порівняння інтенсивності динаміки в багатомірних динамічних рядах однакового змісту по різних об'єктах (регіони, країни і т.п.) або в рядах різного змісту по одному об'єкту використовують ***коефіцієнт випередження***, який одержують відношенням базисних чи середніх темпів зростання $T'р:T''р$.

Темп нарощування в інтенсифікації економіки вимірює нарощування в часі економічного потенціалу. Обчислюються темпи нарощування $T_н$ відношенням ланцюгових абсолютних приростів на рівень, прийнятий за постійну базу порівняння, або безпосередньо визначанням за базисними темпами зростання:

$$T_н = \frac{\Delta y_t^{лан}}{y_1} 100\% = \frac{y_t - y_{t-1}}{y_1} 100 = T_{зpt}^{баз} - T_{зpt-1}^{баз}.$$

10.3. Аналіз структурних зрушень

Структура будь-якої статистичної сукупності динамічна. Змінюються склад і технічний рівень виробничих фондів, вікова структура робітників, склад і якість залучених до виробництва природних ресурсів, асортимент і якість продукції тощо. Зміна частин окремих складових сукупності – це наслідок структурних зрушень.

Структурні зрушення – зміни у складі сукупності за певний період часу.

Структурні зрушення оцінюють за допомогою абсолютних і відносних характеристик динаміки:

1) абсолютного приросту j-ї частки в процентних пунктах (п.п.)

$$\Delta d_j = d_{j1} - d_{j0};$$

2) коефіцієнта (темпу) зростання j-ї частки $Kpd_j = d_{j1} / d_{j0}$.

Характеристики структурних зрушень взаємозв'язані:

$$\Delta d_j = d_{j0}(Kd_j - 1).$$

Як узагальнюючі характеристики інтенсивності структурних зрушень застосовують лінійний $\bar{l}_d$ і квадратичний σ_d коефіцієнти. Їх обчислюють на основі абсолютних приростів часток Δd, тобто

$$\bar{l}_d = \frac{\sum_{j=1}^{n} \left| d_{j1} - d_{j0} \right|}{n}; \; \sigma_d = \sqrt{\frac{\sum_{j=1}^{n} (d_{j1} - d_{j0})^2}{n}}.$$

Знаючи темпи зростання часток, обчислюють квадратичний коефіцієнт, який порівняно з іншими сильніше реагує на зміни в структурі:

$$\sigma_{Kd} = \sqrt{\frac{\sum (d_{j1} - d_{j0})^2}{d_{j0}}} = \sqrt{\sum_{j=1}^{n} (K_{d_j} - 1)^2 d_{j0}}$$

(якщо d_{j1}, d_{j0} надані не в %, то результат необхідно помножити на 100).

10.4. Середні показники динаміки

Рівні динамічних рядів з часом змінюються, отже, варіюють і обчислені на їх основі абсолютні прирости і темпи зростання, виникає необхідність узагальнення властивостей ряду динаміки – отримання узагальнюючих показників динаміки соціально-економічних явищ у вигляді середніх показників: середнього рівня ряду, середнього абсолютного приросту, середнього темпу зростання і приросту.

Середній рівень ряду – це показник, що узагальнює підсумки розвитку явища за одиничний інтервал або момент часу.

Середній рівень інтервального ряду з рівними періодами часу – середня арифметична проста:

$$\overline{Y} = \frac{1}{n} \sum_{1}^{n} y_t \, , \tag{10.15}$$

де n – число рівнів ряду.

Середній рівень інтервального або моментного динамічного ряду з нерівними періодами часу розраховується як середня арифметична зважена:

$$\overline{Y} = \frac{\sum_{1}^{n} y_i t_i}{\sum_{1}^{n} t_i}, \tag{10.16}$$

де n – число рівнів ряду, y_i – рівні ряду, t_i – проміжки часу.

Для наближеної оцінки середнього рівня ряду при умові про рівномірну зміну показника між датами середня обчислюється як напівсума значень на початок і кінець періоду:

$$\bar{y} = \frac{y_1 + y_n}{2},$$

де y_n – кінцевий рівень ряду динаміки.

Якщо в моментному ряді кількість рівнів n>2 і між суміжними датами рівні або приблизно рівні інтервали, розрахунок виконують за формулою середньої хронологічної:

$$\bar{y} = \frac{\frac{y_1 + y_n}{2} + \sum_{2}^{n-1} y_t}{n-1}. \tag{10.17}$$

Середній абсолютний приріст показує, на скільки одиниць у середньому змінився рівень порівняно з попереднім; він обчислюється як середня арифметична проста з ланцюгових абсолютних приростів за певні періоди:

$$\bar{\Delta} = \frac{\sum_{t=1}^{n} \Delta_t^{л}}{n} = \frac{y_n - y_1}{n-1} = \frac{\Delta_{базt}}{n-1}, \tag{10.18}$$

де n – кількість приростів.

Середній коефіцієнт (темп) зростання – узагальнююча характеристика індивідуальних коефіцієнтів (темпів) зростання ряду динаміки, яка обчислюється:

1) за формулою середньої геометричної простої з ланцюгових коефіцієнтів (темпів) зростання:

$$\bar{K}_{зр} = \sqrt[n]{K_{зр1}^{л} K_{зр2}^{л} \ldots K_{зрn}^{л}} = \sqrt[n]{\prod_{t=1}^{n} K_{зрt}^{л}} = \sqrt[n]{K_{зрt}^{лан}} \tag{10.19}$$

$$\text{або } \bar{T}_{з} = \bar{K}_{з} \cdot 100, \tag{10.20}$$

де n – число коефіцієнтів (темпів) зростання;

2) за формулою середньої геометричної зваженої

$$\bar{T}_{з} = \sqrt[\sum t]{T_{з1}^{t_1} T_{з2}^{t_2} \ldots T_{зn}^{t_n}},$$

де t – інтервал часу, в перебігу якого зберігався певний темп зростання; $\sum t$ – сума відрізків часу періоду;

3) з урахуванням мультиплікативності зв'язку ланцюгових і базисних коефіцієнтів (темпів) зростання

$$\bar{K}_{з} = \sqrt[n]{K_n} \text{ або } \bar{T}_{з} = \sqrt[n]{T_n};$$

4) на основі кінцевого y_n і початкового y_1 рівнів ряду

$$\bar{K}_{з} = \sqrt[n-1]{\frac{y_n}{y_1}} \text{ або } \bar{T}_{з} = \sqrt[n-1]{\frac{y_n}{y_1}}100,$$

де n – кількість рівнів динамічного ряду.

Середній темп приросту визначають як різницю між середнім темпом зростання і 100 %: $\bar{T}_{пр} = \bar{T}_{з} - 100\%$.

10.5. Характеристика основної тенденції розвитку

Загальна тенденція (тренд) – певна закономірність зміни рівнів, при цьому одним рядам властива тенденція до зростання, іншим – до зниження рівнів.

У рядах динаміки з чітко виявленою тенденцією її описують аналітично у вигляді відповідної функції часу, яка називається рівнянням тренду

$$Y_t = f(t),$$

де f(t) – основна тенденція, зумовлена впливом постійно діючих чинників.

Тенденція f(t) виявляється при заміні фактичних (емпіричних) рівнів y_t динамічного ряду іншими Y_t – розрахунковими або теоретичними рівнями ряду, обчисленими за певною методикою.

Визначення теоретичних рівнів Y_t проводиться на основі так званої адекватної математичної функції, яка оптимально відображає основну тенденцію ряду динаміки.

У практиці статистичного вивчення тренду розрізняють п'ять еталонних типів розвитку соціально-економічних явищ і процесів у часі, тенденція розвитку яких виражається основними типами рівнянь тренду:

1. *Рівномірний розвиток*. Для цього типу динаміки властиві майже постійні абсолютні прирости $\Delta y^{л} \cong const$, рівні ряду змінюються в арифметичній прогресії або наближені до неї. Наприклад, тенденції динаміки врожайності для масштабу області, республіки, країни в цілому.

Основна тенденція розвитку відображається рівнянням лінійної функції – лінійної форми тренду:

$$Y_t=a_0+a_1t,$$

де Y_t – теоретичні рівні, звільнені від коливань, вирівняні по прямій;

t – порядковий номер періоду або моменту часу; a_0 і a_1 – параметри прямої (початковий рівень ряду при t=0 і приріст щороку).

Якщо $a_1>0$, то рівні ряду динаміки рівномірно зростають, а при $a_1<0$ відбувається їх рівномірне зниження.

2. *Рівноприскорений (рівноуповільнений) розвиток*. Рівні таких рядів динаміки змінюються з постійними (стабільними) темпами приросту $\Delta T^{л}пр \cong const$, а абсолютні ланцюгові прирости абсолютних ланцюгових приростів не виявляють тенденції до розвитку.

Основна тенденція розвитку відображається функцією параболи другого порядку – параболічної форми тренду:

$$Y_t=a_0 + a_1t + a_2t^2,$$

де t , a_0 і a_1 – мають той самий зміст;

a_2 – константа тренду, що характеризує постійну зміну інтенсивності розвитку в одиницю часу. При $a_2>0$ відбувається прискорення розвитку.

При $a_2<0$ йде процес уповільнення зростання, причому з часом веде не тільки до припинення зростання рівня, але і до його зниження все з більшою швидкістю (випуск застарілої продукції).

3. *Розвиток зі змінним прискоренням (уповільненням).* Для цього типу динаміки основна тенденція розвитку виражається функцією параболи третього порядку:

$$Y_t=a_0+a_1t+a_2t^2+a_3t^3.$$

У рівнянні параметр a_3 відображає зміну прискорення. При $a_3>0$ прискорення зростає, а при $a_3<0$ воно сповільнюється.

4. *Розвиток за експонентою*. Цей тип динаміки характеризується відносно стабільними ланцюговими темпами зростання $\Delta Тр^{л} \cong const$; ряд розвивається за геометричною прогресією. Основна тенденція відображається показовою функцією – експоненціальною формою тренду:

$$Y_t=a_0a_1^t.$$

Якщо $a_1>1$, то тренд виражає тенденцію прискорюваного зростання рівнів. Наприклад, розмноження організмів за відсутності обмежень з боку середовища. Якщо $a_1<1$, то тренд виражає тенденцію все більш уповільненого зниження рівнів.

Параметри трендових кривих визначають методом найменших квадратів (МНК), згідно з яким сума квадратів відхилень фактичних рівнів ряду y_t від теоретичних рівнів ряду Y_t має бути мінімальною:

$$\sum_i (y_i - Y_{ti})^2 = \min.$$

Для лінійної функції умова мінімізації:

$$\sum_{t=1}^{n}(y_t - Y_t)^2 = \sum_{t=1}^{n}(y_t - (a_0 + a_1t))^2 = \min.$$

Параметри трендових кривих обчислюють, розв’язуючи системи нормальних рівнянь. Для лінійної функції вона записується так:

$$\begin{cases} a_0n + a_1\sum t = \sum y, \\ a_0\sum t + a_1\sum t^2 = \sum yt. \end{cases}$$

Якщо відлік часу t перенести в середину динамічного ряду, то значення t, розміщені вище за середину, будуть від’ємними, а нижче – додатними. Тоді $\sum t = 0$. При непарному числі членів ряду, наприклад, n=5 t набуває значення -2, -1, 0, 1, 2; при парному n=6: -5, -3, -1, 1, 3, 5. В обох випадках система рівнянь спрощується:

Звідси

$$a_0 = \frac{\sum y}{n} = \overline{Y};\ a_1 = \frac{\sum yt}{\sum t^2}. \tag{10.21}$$

Значення $\sum t^2$ можна знайти за формулами:

- для непарного числа членів ряду $\sum_{i=1}^{n} t^2{}_i = \frac{n(n^2-1)}{12}$;

- для парного числа членів ряду $\sum_{i=1}^{n} t^2{}_i = \frac{n(n^2-1)}{3}$.

10.6. Вимірювання сезонних коливань

Якщо в аналізованій тимчасовій послідовності спостерігається стійке відхилення від тенденції, то можна припустити наявність у ряді динаміки деяких коливальних процесів.

Сезонними коливаннями називають більш-менш стійкі внутрішньорічні коливання рівнів розвитку соціально-економічних явищ у рядах динаміки, обумовлені специфікою виробництва або споживанням певного виду продукції. Вони виявляються або аналізуються на основі рядів щомісячних або щоквартальних даних і описуються сезонною хвилею.

Для вимірювання сезонних коливань обчислюють індекси сезонності.

Індекс сезонності I_s – в загальному вигляді це процентне відношення однойменних (місячних, квартальних) фактичних рівнів ряду динаміки y_i до їх середньорічних або теоретичних рівнів $Y_t = f(t)$ (якщо ряд виявляє тенденцію):

$$I_{Si} = \frac{y_i}{Y_{ti}} 100\% .$$

За допомогою індексів сезонності за декілька років можна виявити постійну ***сезонну хвилю*** – це сукупність індексів сезонності.

Оскільки сезонні коливання з року в рік не залишаються незмінними, то для їх усунення проводиться усереднювання індивідуальних індексів однойменних внутрішньорічних періодів аналізованого ряду динаміки. Тому для кожного періоду річного циклу визначаються узагальнені показники у вигляді середніх індексів сезонності:

$$\bar{I}_{Si} = \frac{\sum_{1}^{n} I_{Si}}{n}, \qquad (10.22)$$

де n – число років.

Обчислені за формулою (10.1) середні індекси сезонності вільні від впливу основної тенденції і випадкових відхилень.

За допомогою методів вимірювання сезонних коливань, залежно від характеру тренду формула (10.1) набуває таких форм:

1) для рядів внутрішньорічної динаміки зі значно вираженою основною тенденцією розвитку (ланцюгові та базисні темпи зростання високі) індекс сезонності розрахуємо ***методом змінної середньої***:

$$\bar{I}_{Si} = \frac{\sum \frac{y_i}{Y_{ti}}}{n} 100\%,$$

де y_t – фактичні (емпіричні) рівні ряду, Y_t – теоретичні рівні ряду, n – кількість років;

2) для рядів внутрішньорічної динаміки, в яких річні рівні стабільні, тенденція відсутня або незначна, індекс сезонності розрахуємо ***методом постійної (простої) середньої***:

$$\bar{I}_{Si} = \frac{\bar{y}_i}{\bar{y}} 100\%,$$

де $\bar{y}_i = \frac{\sum_{j=1}^{n} y_{ij}}{n}$ – середні місячні або квартальні рівні ряду, обчислені по рівнях y_{ij} за i-й місяць (квартал) для j-го року, n – кількість років;

$\bar{y} = \frac{\sum_{i=1}^{m} y_i}{m}$ – загальна середня (місячна або квартальна) по всіх рівнях,

m – загальна кількість місяців або кварталів.

Якщо є дані за один рік, то індекс сезонності розрахуємо за формулою

$$\bar{I}_{Si} = \frac{y_i}{\bar{y}} 100\%, \tag{10.23}$$

де y_i – місячні (квартальні) рівні.

Для порівняння інтенсивності сезонних коливань різних явищ або одного явища, але за різні періоди (роки), використовують узагальнювальніі характеристики варіації індексів сезонності:

1) амплітуда сезонних коливань (вимірюються в п.п.) $R_t = I_S^{max} - I_S^{min}$;

2) середнє лінійне відхилення $\bar{I}_S = \frac{1}{12}\sum_{1}^{12}|I_S - 100|$;

3) середнє квадратичне відхилення

$$\sigma_t = \sqrt{\frac{1}{12}\sum_{1}^{12}(I_S - 100)^2}\,. \tag{10.24}$$

10.7. Екстраполяція в рядах динаміки і прогнозування

Визначувані в аналізі рядів динаміки показники зміни рівнів, тренду, сезонної хвилі широко використовують при ***прогнозуванні*** – отриманні статистичної оцінки можливої міри розвитку соціально-економічних явищ на майбутнє.

Складання надійних прогнозів динаміки попиту і пропозиції товарів є необхідною умовою регулювання ринкових відносин. Згладжування динамічних рядів використовують також для визначення невідомих членів ряду за допомогою інтерполяції та екстраполяції.

Інтерполяція – визначення невідомого показника в середині ряду динаміки.

Екстраполяція – розповсюдження виявлених в аналізі рядів динаміки закономірностей розвитку явища, що вивчається, на майбутнє, тобто продовження виявленої тенденції за межі ряду динаміки, а також визначення невідомих рівнів у кінці або напочатку динамичного ряду.

Основою прогнозування є припущення, що закономірність, діюча всередині аналізованого ряду динаміки, зберігається і надалі, тобто залишається незмінним причиновий комплекс, що формує тенденцію. Точність прогнозу залежить від того, наскільки обгрунтованими виявляться припущення про збереження на майбутнє дій тих чинників, які сформували в базисному ряді динаміки його основні компоненти.

Встановлення термінів прогнозування m залежить від задачі дослідження. Але слід мати на увазі, що чим коротше терміни попередження прогнозу, тим надійніше результати екстраполяції.

Вживання методів екстраполяції залежить від характеру змін в базисному ряді динаміки і зумовлюється постановкою задачі дослідження:

1. При екстраполяції рівнів розвитку явища, що вивчається, на базі ряду динаміки з постійними абсолютними приростами (Δуцеп≅const) застосовується формула

$$y_{n+m} = y_n + \Delta y \cdot m, \qquad (10.25)$$

де y_{n+m} – рівень, що екстраполюється; y_n – кінцевий рівень базисного ряду динаміки; *m* — термін прогнозу (період попередження).

2. На базі ряду динаміки із стабільними темпами зростання (ΔТрцеп≅const) невідомі рівні обчислюють за формулою

$$y_{n+m} = y_n(T_p)^m.$$

На практиці результат екстраполяції прогнозованих рівнів соціально-економічних явищ оцінюють не точковими (дискретними), а інтервальними оцінками.

10.8. Розв'язання типових задач

Приклад 10.1

Випуск промислових роботів в об'єднанні характеризується даними, наведеними в табл. 10.2. Визначити:

1) базисні та ланцюгові характеристики динаміки: абсолютні прирости, коефіцієнти і темпи зростання, темпи приросту, абсолютне значення 1% приросту;

2) середньорічні коефіцієнти і темпи зростання, абсолютні прирости, абсолютне прискорення (уповільнення) зростання за 2000-2005 і 2006-2007 роки;

3) абсолютне прискорення (уповільнення) та коефіцієнт прискорення (уповільнення) зростання.

Таблиця 10.2

Рік	2000	2005	2007
Кількість, шт.	60	114	126

Розв'язання.

Абсолютний приріст обчислимо за формулами (10.1)-(10.2), коефіцієнти і темпи зростання – за формулами (10.3)-(10.6), абсолютне значення 1% приросту – за формулами (10.11)-(10.12). Результати розрахунків наведено в табл. 10.3.

Таблиця 10.3

Показники динаміки	2000-2005	2006-2007	За весь період
Абсолютний приріст	$\Delta_t^{лан}$ =114-60=54	$\Delta_t^{лан}$ =126-114=12	Δ_tбаз=126-60=66
Коефіцієнт зростання	$Кр^{лан}$ =114/60=1,9	$Кр^{лан}$=126/114=1,1	$Кр_{баз}$=126/60=2,1
Темп зростання	$Тр^{лан}$=$Кр^{лан}$*100=190	$Тр^{лан}$=Крл*100=110	$Тр_б$=Крб*100=210
Темп приросту	$Т_{пр}^{лан}$=190-100=90%	$Т_{пр}^{лан}$=110-100=10%	$Т_{пр}^{б}$= 210-100=110%
Абсолютне значення 1% приросту	$А_t$=60/100=0,6	$А_t$=114/100=1,14	$А_t$=60/100=0,6
Середньорічний коефіцієнт зростання	$\overline{К}_р = \sqrt[5]{\frac{114}{60}} = 1{,}137$	$\overline{К}_р = \sqrt{\frac{126}{114}} = 1{,}049$	
Середньорічний темп зростання	$\overline{Т}_р$=1,137*100=113,7	$\overline{Т}_р = \overline{К}_р 100\% = 104{,}9$	
Середній темп приросту	$\overline{Т}_{пр} = 113{,}7 - 100\% = 13{,}7\%$	$\overline{Т}_{пр} = \overline{Т}_р - 100\%$=4,9%	

Середньорічні коефіцієнти і темпи зростання обчислимо за формулами (10.19)-(10.20), абсолютне прискорення (уповільнення) зростання за 2000-2005 і 2006-2007 роки – за формулою (10.13), коефіцієнт прискорення (уповільнення) зростання – за формулою (10.14).

Абсолютне прискорення (уповільнення) δ_t = 6 - 10,8 = - 4,8. Значення – 4,8 свідчить про уповільнення зростання.

Коефіцієнт прискорення зростання $\gamma_t = 1{,}137 / 1{,}049 = 1{,}083$.

Приклад 10.2

Є дані про заборгованість по кредитах районному відділенню банка (табл. 10.4):

Таблиця 10.4

Дата	Залишки заборгованості, тис. грн	Загальна заборгованість, тис. грн
01.01.2008	12	500
01.02.2008	10	650
01.03.2008	10	600
01.04.2008	8	500

1. Визначити середній залишок простроченої заборгованості за перший квартал 2008 року.
2. Обчислити середній залишок заборгованості по всіх кредитах за перший квартал 2008 року.

Розв'язання.

1. Середній залишок простроченої заборгованості, використавши середню хронологічну для моментного ряду, визначимо за формулою (10.17):

$$\bar{y} = \frac{12/2 + 10 + 10 + 8/2}{4 - 1} = 10 \text{ тис. грн.}$$

2. Аналогічно обчислимо середній залишок заборгованості за всіма кредитами:

$$\bar{y} = \frac{500/2 + 650 + 600 + 500/2}{4 - 1} = 583{,}33 \text{ тис. грн.}$$

Приклад 10.3

За наведеними в табл. 10.5 даними виробництва товарів народного споживання в регіоні описати тенденцію ряду лінійним трендом, дати економічну інтерпретацію параметрів трендового рівняння. Зробити прогноз на 2009 рік, припускаючи, що умови формування тенденції виробництва товарів не зміняться.

Таблиця 10.5

Рік	2001	2002	2003	2004	2005	2006	2007
Виробництво товарів, тис. т	68	74	77	80	86	94	97

Розв'язання.

Параметри трендового рівняння визначимо за формулою (10.21). Розрахунок параметрів наведено в табл. 10.6.

Таблиця 10.6

Рік	Y	t	t^2	yt	Y_t=83+5t
2001	68	-3	9	-204	
2002	74	-2	4	-148	
2003	77	-1	1	-77	
2004	82	0	0	0	
2005	89	1	1	89	
2006	94	2	4	188	
2007	97	3	9	291	98
Усього	581	0	28	139	581

Тоді

$$a_0 = \frac{581}{7} = 83; \;\; a_1 = \frac{139}{28} \approx 5.$$

Отже, річний обсяг виробництва товарів народного споживання становив у середньому 83 тис. т, а середньорічний приріст – 5 тис. т.

В останньому стовпці для кожного року наведені теоретичні рівні Y_t, тобто очікувані рівні обсягу виробництва товарів народного споживання в t-му році, обумовлені впливом основних причин розвитку виробництва.

Зробимо прогноз на 2009 рік, припускаючи, що умови формування тенденції виробництва товарів не зміняться. Базою прогнозу являється теоретичний рівень 2007 року, період випередження m=2. Тоді очікуваний у 2009 році рівень виробництва товарів Y_{t+m}=98 +5*2=108.

Приклад 10.4

За наведеними в табл. 10.7 даними про споживання електроенергії комунальним господарством району визначити сезонну хвилю на базі постійної середньої й оцінити сезонність за допомогою амплітуди коливань і середнього квадратичного відхилення.

Таблиця 10.7

Місяць року	1	2	3	4	5	6	7	8	9	10	11	12	Разом
Виробництво товарів, тис. т	172	161	158	151	147	130	124	146	149	155	168	187	1848

Розв'язання.

Середнє місячне споживання електроенергії

$$\overline{Y} = 1848/12 = 154 \text{ (млн кВт*г)}.$$

Сезонна хвиля – це сукупність помісячних індексів сезонності, обчислених за формулою (10.23). Індекси сезонності наведено в табл. 10.8.

Таблиця 10.8

	Спожито електроенергії y_t, млн кВт*г	Індекс сезонності I_s, %	I_s -100	$(I_s -100)^2$
1	172	111,7		
2	161	104,5		
3	158	102,6		
4	151	98,0		
5	147	95,5		
6	130	84,4		
7	124	80,5		
8	146	94,9		
9	149	96,8		
10	155	100,6		
11	168	109,1		
12	187	121,4		
Разом	1848			

Амплітуда сезонних коливань становить R_t=121,4-80,5 = 40,9 п.п.

Згідно з розрахунками середнє квадратичне відхилення обчислимо за формулою

$$\sigma_t = \sqrt{\frac{}{12}} = \%.$$

Характер сезонної хвилі схематично ілюструє рис. 10.1.

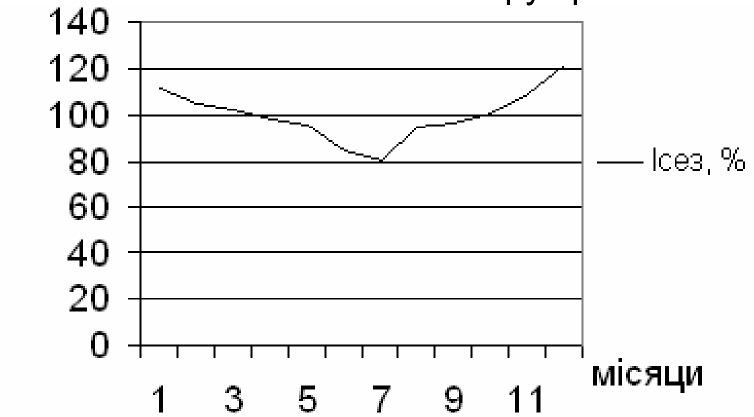

Рис. 10.1. Сезонна хвиля

Контрольні запитання

1. Що називають рядом динаміки?
2. З яких елементів складаються ряди динаміки та що вони виражають?
3. Яких умов потрібно дотримуватися при побудові рядів динаміки?
4. За яких причин виникає непорівнюваність рівнів ряду динаміки?
5. Які ви знаєте види рядів динаміки?

6. Які ряди динаміки називають моментними та чому їхні рівні не можна підсумовувати?

7. Чому ряди динаміки, виражені абсолютними величинами, є первинними, а виражені відносними та середніми величинами – вторинними?

8. Як розраховують середній рівень для інтервального і моментного рядів динаміки?

9. Як визначають абсолютний приріст, темп зростання і приросту?

10. Як визначають середній темп зростання на основі даних першого й останнього рівнів ряду динаміки?

11. Як розраховують середній темп зростання за ланцюговими коефіцієнтами зростання?

12. Які способи обробки і аналізу рядів динаміки ви знаєте?

13. Що таке інтерполяція і екстраполяція рядів динаміки, їхні значення та застосування?

Тема 11. СТАТИСТИЧНІ ІНДЕКСИ

11.1. Суть і функції індексів

Статистичний індекс – це відносна величина, що характеризує зміну соціально-економічного явища в часі, просторі або ступінь відхилення значення показника від певного стандарту (нормативу, середньої). Індекс показує, в скільки разів рівень явища, що вивчається, в певних умовах відрізняється від рівня того ж явища в інших умовах.

Форми виразу індексів – коефіцієнти, відсотки, промілле.

За характером досліджуваних об'єктів розрізняють:

- ***індекси об'ємних показників***, які характеризують зміни обсягу того або іншого явища, який вимірюється в певних одиницях (індекси фізичного обсягу продукції, роздрібного товарообігу, споживання продуктів тощо);
- ***індекси якісних показників***, які характеризують зміни якісної ознаки (індекси цін, собівартості продукції, продуктивності праці тощо).

За ступенем охоплення одиниць сукупності індекси поділяють на:

- ***індивідуальні*** (і), що характеризують зміни окремих елементів складної сукупності або однорідних груп;
- ***зведені або загальні*** (І), що характеризують зміну сукупності, до якої входять різнорідні елементи. Такими елементами є окремі продукти у складі продукції або обсяг товарів, реалізованих магазинами.

Якщо індекси охоплюють не всі одиниці сукупності, то їх називають груповими, або субіндексами.

Залежно від бази порівняння розрізняють ***ланцюгові, базисні індекси***.

Залежно від методології обчислення загальні індекси поділяють на ***агрегатні, середні з індивідуальних індексів.***

За характером і умовами порівнянь (час, простір, стандарт) індекси поділяють на:

- ***динамічні***, які характеризують інтенсивність динаміки;
- ***територіальні***, які визначають ступінь відхилення значень показника в просторі – між об'єктами, країнами, регіонами, які ідентифікуються відповідними буквами. Вибір бази порівняння довільний;
- ***індекси виконання зобов'язань***, що характеризують відхилення від стандарту (еталонного, максимального або мінімального значення) або середнього рівня за сукупністю в цілому.

Вибір базисного періоду завжди був аргументований тією задачею, для якої будується індекс. При цьому керуються двома правилами:

1) база порівняння являє собою стабільний рівень;
2) база порівняння як екстремальний рівень – вище досягнення або низький рівень (у разі падіння економічних показників).

Індекс поєднує в собі такі аспекти:

- якісний, в якому назва відображає соціально-економічний зміст показника;
- кількісний, в якому числове значення відображає інтенсивності змін або ступінь відхилення.

Кожний індекс містить два види даних: оцінювані дані, звані поточними (звітними) (позначають значком «1»), і дані, що використовуються як база порівняння, – базисні (позначають значком «0»).

Основний елемент індексного відношення – ***індексована величина*** – показник, динаміку або співвідношення якого характеризує індекс.

У міжнародній статистиці використовують таку систему позначень індексованих величин: p – ціна; q – фізичний обсяг (кількість товарів, продукції тощо в натуральних одиницях); Q (або pq) – товарообіг; z – собівартість одиниці продукції; t – трудомісткість праці; w – середній випуск продукції з розрахунку на одного працівника; y – врожайність культури з 1 га; n – розмір посівної площі.

11.2. Індивідуальні індекси

Індивідуальним індексом (i) називається відносна величина, яка характеризує зміни окремих елементів складної сукупності або однорідних груп.

За індивідуальні індекси приймають відносну величину динаміки або порівняння.

Методологія побудови індивідуальних індексів проста.

Рівень товарообігу Q_1 в умовах звітного року подамо у вигляді суми виручки від продажу товару: $Q_1 = p_1 \cdot q_1$. Аналогічно рівень товарообігу базисного року набуває вигляду $Q_0 = p_0 \cdot q_0$. Тоді отримаємо ***індивідуальний індекс товарообігу***

$$i_Q = \frac{Q_1}{Q_0}. \tag{11.1}$$

Аналогічні індивідуальні індекси можна розрахувати для будь-якого показника:

- ***індивідуальний індекс ціни,*** який показує, в скільки разів змінилася загальна сума виручки під впливом зміни ціни товару:

$$i_p = \frac{p_1}{p_0}; \tag{11.2}$$

- ***індивідуальний індекс кількості проданих товарів***, який показує, в скільки разів збільшилася (або зменшилася) загальна сума виручки під впливом зміни обсягу продажу в натуральних одиницях:

$$i_q = \frac{q_1}{q_0}. \tag{11.3}$$

Очевидно, що $i_Q = \frac{Q_1}{Q_0} = \frac{q_1}{q_0} \cdot \frac{p_1}{p_0} = i_q i_p$ або $Q_1 = Q_0 i_q i_p$.

Друга формула являє собою індексну мультиплікативну модель підсумкового показника (в цьому випадку – обсягу товарообігу), за допомогою якої знаходять приріст підсумку під впливом кожного чинника окремо.

11.3. Загальні індекси

Якщо відомо, що явище, яке вивчається, неоднорідне і порівняння рівнів можна провести тільки після приведення їх до загальної міри, економічний аналіз виконують за допомогою загальних індексів.

Загальний індекс (I) характеризує зміну сукупності, до якої входять різнорідні елементи.

Важливою особливістю загальних індексів є те, що вони виконують дві взаємозв'язані функції:

1) синтетичну – це узагальнююча характеристика зміни явища;
2) аналітичну **–** вивчення дії окремих чинників на зміну явища.

Основною формою загальних індексів є агрегатні індекси.

У складних статистичних сукупностях зіставності різнорідних одиниць досягають введенням в індексні відношення спеціальних додаткових показників – сумірників (ваг), які економічно щільно пов’язані з індексованою величиною. Таким сумірником може бути ціна, собівартість чи трудомісткість одиниці продукції. Перемноживши обсяг продукції кожного виду на відповідний сумірник, отримують показники, які можна підсумувати.

Сума добутку кількості продукції q_j на його сумірник, наприклад ціну p_j, створює з'єднання – так звані агрегати $\sum qp$ (aggrego (лат.) – приєдную).

Агрегат – добуток з'єднаних величин, одна з яких називається індексованою, інша – сумірником.

Агрегатні індекси розраховуються як відношення агрегатів, побудованих для різних умов.

Індексована величина, зміну якої характеризує даний індекс, в чисельнику і знаменнику агрегату знаходиться в різних періодах – звітне значення зіставляється з базисним.

Сумірник можна фіксувати на рівні як базисного, так і поточного періодів. Залежно від цього розглядаються формули індексів цін і фізичного обсягу в різних системах вагів, поданих у табл. 11.1.

Таблиця 11.1

Базисно-вагова система (індекси Ласпейреса)	Поточно-вагова система (індекси Пааше)
$I_p = \frac{\sum p_1 q_0}{\sum p_0 q_0}$ (11.4)	$I_p = \frac{\sum p_1 q_1}{\sum p_0 q_1}$ (11.5)
$I_q = \frac{\sum q_1 p_0}{\sum q_0 p_0}$ (11.6)	$I_q = \frac{\sum q_1 p_1}{\sum q_0 p_1}$ (11.7)

Вибір форми індексів залежить від мети дослідження і наявної інформації. Обидві системи індексів рівноправні.

При синтезі загального індексу цін замість фактичної кількості товарів (у звітний або базисний періоди) як сумірник можуть застосовуватися середні величини реалізації товарів за два або більше число періодів. При такому способі розрахунку (***індекс Лоу***) формула загального індексу синтезується у вигляді:

$$I_p = \frac{\sum p_1 \overline{q}}{\sum p_0 \overline{q}},$$

де $\overline{q}$ – середня кількість товарів, реалізованих за аналізований період.

Індекс товарообігу у фактичних цінах в агрегатній формі:

$$I_Q = \frac{\sum q_1 p_1}{\sum q_0 p_0}, \qquad (11.8)$$

де n – кількість товарів; p_1q_1, p_0q_0 – товарообіг окремих видів товарів у базисному та звітному періодах.

11.3. Середньозважені індекси

Крім запису загальних індексів в агрегатній формі на практиці часто використовують формули розрахунку загальних індексів як середніх з індивідуальних індексів – ***середньозважені індекси*** (табл. 11.2).

У базисно-ваговій системі загальні індекси обчислюють як середню арифметичну з вагами q_0p_0. У поточно-ваговій системі загальні індекси обчислюють як середню гармонійну з вагами q_1p_1.

Таблиця 11.2

Базисно-вагова система (середня арифметична форма)	Поточно-вагова система (середня гармонійна форма)
$I_q = \frac{\sum i_q q_0 p_0}{\sum q_0 p_0}$ (11.9)	$I_p = \frac{\sum p_1 q_1}{\sum \frac{1}{i_p} p_1 q_1}$ (11.10)
$I_p = \frac{\sum i_p q_0 p_0}{\sum q_0 p_0}$ (11.11)	$I_q = \frac{\sum p_1 q_1}{\sum \frac{1}{i_q} p_1 q_1}$ (11.12)
$I_Q = \frac{\sum i_Q p_0 q_0}{\sum q_0 p_0}$ (11.13)	$I_Q = \frac{\sum p_1 q_1}{\sum \frac{1}{i_Q} p_1 q_1}$ (11.14)

При побудові середньозважених індексів вартісні ваги можна замінити відносними величинами структури d, сума яких $\sum d = 1$. У цьому випадку середньозважені індекси набувають вигляду

$$I_q = \sum i_q d_0 , \quad I_p = \frac{1}{\sum \frac{1}{i_p} d_1} . \qquad (11.15)$$

11.4. Взаємозв'язок індексів та індексні системи

Практично кожний індекс є складовою певної індексної системи, а його зв'язки з іншими індексами цієї системи відображають зв'язки між відповідними показниками. Так, товарообіг в поточному періоді порівняно з базовим збільшується (зменшується) під впливом двох чинників – фізичного обсягу проданого товару q і цін p. Відповідно індекс товарообігу I_Q у формі мультиплікативної індексної моделі можна виразити як добуток індексів:

$$I_Q = I_q I_p .$$

Аналогічно грошові витрати на виробництво можна передати як функцію фізичного обсягу виробництва q і собівартості z, тоді індекс грошових витрат на виробництво:

$$I_{qz}=I_q U_z.$$

Обсяг виробництва – функція трудових витрат qt і продуктивності роботи w, тобто $I_{qt}=I_t I_w$.

Таким чином, у будь-якій системі індекс добутку зв'язаних величин дорівнює добутку індексів цих величин.

Взаємозв'язані індекси прямих і зворотних показників, наприклад, споживацьких цін і купівельної здатності грошової одиниці, продуктивності роботи і трудомісткості продукції I_p:

$$I_{1/t} = \frac{1}{I_t};\ \ I_{1/p} = \frac{1}{I_p}.$$

Треба зауважити, що в системі індексів обидва індекси-співмножники мають бути різної ваги: вага одного з них фіксується на рівні базисного періоду, іншого – на рівні поточного. Тоді:

$$\frac{\sum q_1 p_0}{\sum q_0 p_0}\cdot\frac{\sum p_1 q_1}{\sum p_0 q_1}=\frac{\sum p_1 q_1}{\sum p_0 q_0},\ \text{або}\ \frac{\sum q_1 p_1}{\sum q_0 p_1}\cdot\frac{\sum p_1 q_0}{\sum p_0 q_0}=\frac{\sum p_1 q_1}{\sum p_0 q_0}. \qquad (11.16)$$

Таким чином, загальні індекси цін I_p і товарної маси I_q, маючи самостійне значення, одночасно виконують аналітичну функцію, тобто оцінюють вплив відповідного фактора на динаміку обороту.

У межах індексної системи можна виявити:

- ступінь впливу факторів на результат, який характеризують ***темпи приросту факторів:***

$$T_{пр} = I - 100;$$

- ***абсолютний приріст*** як різницю між чисельником і знаменником відповідних індексів. Так, абсолютний вплив чинників на приріст результату – товарообігу виразимо через абсолютний приріст:

$$\Delta Q = \sum q_1 p_1 - \sum q_0 p_0 = \Delta q + \Delta p.$$

Абсолютний вплив кожного чинника окремо в системі :

$$\Delta q = \sum q_1 p_0 - \sum q_0 p_0 = \sum (q_1 - q_0) p_0,$$

$$\Delta p = \sum p_1 q_1 - \sum p_0 q_1 = \sum (p_1 - p_0) q_1.$$

Наведемо формули розрахунку деяких самих споживаних агрегатних індексів і їх взаємозв'язок. Треба зауважити, що в системі індексів обидва агрегатні індекси-співмножники мають бути різної ваги: вага одного з них фіксується на рівні базисного періоду (індекс Ласпейреса), іншого – на рівні поточного (індекс Пааше). Добуток індексів у системі будується аналогічно формулам (11.4)-(11.7).

Індекс зміни загальної суми витрат на виробництво продукції залежно від обсягу виробництва (q) і собівартості (z):

$$I_c = \frac{\sum z_1 q_1}{\sum z_0 q_0} = \frac{\sum q_1 z_0}{\sum q_0 z_0} \cdot \frac{\sum z_1 q_1}{\sum z_0 q_1} = \frac{\sum z_1 q_1}{\sum z_1 q_0} \cdot \frac{\sum z_1 q_0}{\sum z_0 q_0} = I_q I_z .$$

Індекс зміни обсягу продукції Q у зв'язку зі зміною чисельності працюючих (t) і рівня їх продуктивності праці (w):

$$I_Q = \frac{\sum w_1 t_1}{\sum w_0 t_0} = \frac{\sum w_0 t_1}{\sum w_0 t_0} \frac{\sum w_1 t_1}{\sum w_0 t_1} = \frac{\sum w_1 t_1}{\sum w_1 t_0} \frac{\sum w_1 t_0}{\sum w_0 t_0} = I_t I_w . \quad (11.17)$$

Індекс зміни загального фонду оплати праці у зв'язку із зміною загальної чисельності працюючих (t) і заробітної платні (f):

$$I_f = \frac{\sum f_1 t_1}{\sum f_0 t_0} = \frac{\sum f_0 t_1}{\sum f_0 t_0} \cdot \frac{\sum f_1 t_1}{\sum f_0 t_1} = \frac{\sum f_1 t_1}{\sum f_1 t_0} \cdot \frac{\sum f_1 t_0}{\sum f_0 t_0} = I_t I_f .$$

Аналогічно обчислюють загальні агрегатні індекси і для багатьох інших економічних показників.

11.5. Індекси середніх величин

Разом із агрегатними індексами в статистиці використовують індекси середніх величин (наприклад, індекси середньої заробітної платні, середньої врожайності тощо).

Як відомо, рівень середньої величини залежить від значень ознаки x_j і структури сукупності:

$$\bar{x} = \frac{\sum_{i=1}^{m} x_i f_i}{\sum_{i=1}^{m} f_i} = \sum_{i=1}^{m} x_i d_i,$$

де f_j - частота; $n = \sum_{i=1}^{m} f_i$ – обсяг сукупності; m – кількість складових сукупності; d_j – частка j-ї складової сукупності.

Очевидно, що на середню величину впливають як значення ознаки, яку усереднюють, так і чисельність окремих варіантів сукупності – структурні зрушення.

Вплив кожного з чинників на динаміку середньої оцінюється з допомогою системи індексів середніх величин – змінного та фіксованого і системи індексів структурних зрушень.

Індексом змінного складу називається індекс середньої величини, який відображає зміну індексованої величини *x* і зміну структури явища, яку вивчають:

$$I_{\bar{x}} = \frac{\bar{x}_1}{\bar{x}_0} = \frac{\sum_{i=1}^{m} x_{i1} f_{i1}}{\sum_{i=1}^{m} f_{i1}} : \frac{\sum_{i=1}^{m} x_{i0} f_{i0}}{\sum_{i=1}^{m} f_{i0}}. \qquad (11.18)$$

Якщо в початкових формулах загальних індексів показники ваг f_i заміни-ти відносними, тобто питомою вагою $d = \frac{f_i}{\sum f}$, то отримаємо $I_{\bar{x}} = \frac{\sum_{i=1}^{m} x_{i1} d_{i1}}{\sum_{i=1}^{m} x_{i0} d_{i0}}$.

Наведена форма запису індексів має деяку перевагу над розгорненою, тому її ширше застосовують на практиці. Вона компактна, чіткіше виділяє чинники, які впливають на зміну середньої величини показника.

Якщо за величину *x* прийняти ціну *p*, то отримаємо індекс цін:

$$I_{\bar{p}} = \frac{\bar{p}_1}{\bar{p}_0} = \frac{\sum_{i=1}^{m} p_{i1} q_{i1}}{\sum_{i=1}^{m} q_{i1}} : \frac{\sum_{i=1}^{m} p_{i0} q_{i0}}{\sum_{i=1}^{m} q_{i0}}.$$

Індекс змінного складу   розкладається на два субіндекси:

- індекс фіксованого складу I_x свідчить, як в середньому змінилося значення показника при незмінній, фіксованій структурі суккупності:

$$I_x = \frac{\sum_{i=1}^{m} x_{i1} f_{i1}}{\sum_{i=1}^{m} f_{i1}} : \frac{\sum_{i=1}^{m} x_{i0} f_{i1}}{\sum_{i=1}^{m} f_{i1}} = \frac{\sum_{i=1}^{m} x_{i1} f_{i1}}{\sum_{i=1}^{m} x_{i0} f_{i1}} \text{ або } I_x = \sum_{i=1}^{m} x_{i1} d_{i1} : \sum_{i=1}^{m} x_{i0} d_{i1}. \qquad (11.19)$$

$$I_p = \frac{\sum_{i=1}^{m} p_{i1} q_{i1}}{\sum_{i=1}^{m} q_{i1}} : \frac{\sum_{i=1}^{m} p_{i0} q_{i1}}{\sum_{i=1}^{m} q_{i1}} = \frac{\sum_{i=1}^{m} p_{i1} q_{i1}}{\sum_{i=1}^{m} p_{i0} q_{i1}}.$$

- індекс структурних зрушень I_d показує, як змінилася середня внаслідок структурних зрушень. Значення показника *x* фіксується на рівні базисного періоду:

$$I_d = \frac{\sum_{i=1}^{m} x_{i0} f_{i1}}{\sum_{i=1}^{m} f_{i1}} : \frac{\sum_{i=1}^{m} x_{i0} f_{i0}}{\sum_{i=1}^{m} f_{i0}} \text{ або } I_d = \sum_{i=1}^{m} x_{i0} d_{i1} : \sum_{i=1}^{m} x_{i0} d_{i0};$$

$$I_d = \frac{\sum_{i=1}^{m} p_{i0} q_{i1}}{\sum_{i=1}^{m} q_{i1}} : \frac{\sum_{i=1}^{m} p_{i0} q_{i0}}{\sum_{i=1}^{m} q_{i0}} \text{ або } I_d = \sum_{i=1}^{m} x_{i0} d_{i1} : \sum_{i=1}^{m} x_{i0} d_{i0}. \qquad (11.20)$$

Формули індексів фіксованого складу і структурних зрушень різноважові: в I_x вагу фіксують на рівні поточного періоду, в I_d – значення показника x – на рівні базисного періоду. Саме такий варіант зважування забезпечує зв'язок цих індексів у системі

$$I_{\bar{x}} = I_x I_d. \qquad (11.21)$$

У межах індексної системи можна виявити абсолютний приріст ***середньої за рахунок кожного чинника.***

11.6. Територіальні індекси

Територіальні індекси використовують як інструмент порівняння соціально-економічних показників у просторі: по окремих країнах, територіях, регіонах, об'єктах.

Територіальні індекси – відносні величини, які дають порівняльну характеристику в межах територій і об'єктів.

Їх особливість – рівноправність порівнюваних об'єктів А і В. Кожний з них може бути базою порівняння, отже, рівноправні індекси як з базою А, так і з базою В:

$$I_{x,\frac{A}{B}} = \frac{\sum x_A f_A}{\sum x_B f_A}; \quad I_{x,\frac{B}{A}} = \frac{\sum x_B f_B}{\sum x_A f_B}, \qquad (11.22)$$

де x – індексована величина, f – вага індексованої величини.

При фіксованих значеннях вагів індекси I_A й I_B обернено пропорційні.

Побудову найпростіших територіальних індексів розглянемо на прикладі показника товарообігу Q для двох районів: А і В.

Територіальний індекс товарообігу – це відношення суми виручки від продажу в одному з районів до аналогічного показника в іншому районі:

$$I_{Q,\frac{A}{B}} = \frac{\sum p_A q_A}{\sum p_B q_B}. \qquad (11.23)$$

Для характеристики співвідношення рівнів цін на товари, реалізовані у районі *А* порівняно з районом В, розраховують загальний індекс цін, в якому за вагу індексованих величин p_A і p_B беруть кількість товарів, реалізованих у районі *А*:

$$I_{p_{A/B}} = \frac{\sum q_A p_A}{\sum q_A p_B}, \qquad (11.24)$$

або в місті В:

$$I_{p_{B/A}} = \frac{\sum q_B p_B}{\sum q_B p_A}. \qquad (11.\ 22)$$

Аналогічно будуються індекси обсягів реалізованих товарів.

Територіальний індекс змінного складу по об'єктах А і В обчислюють за формулою

$$I_{\overline{x}} = \frac{\sum x_A f_A}{\sum f_A} : \frac{\sum x_B f_B}{\sum f_B} = \sum x_A d_A : \sum x_B d_B .$$

Він показує, в скільки разів середній рівень ознаки об'єкта *А* більший або менший об'єкта *В.*

Територіальний індекс фіксованого складу:

$$I_x = \frac{\sum x_A f^{st}}{\sum f^{st}} : \frac{\sum x_B f^{st}}{\sum f^{st}} = \sum x_A d^{st} : \sum x_B d^{st} ,$$

де f_{st} – частота; d_{st} – частка стандартної структури сукупності.

Замість стандартної структури сукупності можна використовувати середню структуру. Індекс I_x показує співвідношення середніх значень ознаки при фіксованій структурі сукупності.

11.7. Розв'язання типових задач

Приклад 11.1

Розглянемо групу з двох підприємств, що виробляють різну продукцію. Кожне підприємство має дані за два суміжних роки (базисний і звітний) про чисельність працюючих і середню продуктивність праці на одну людину (табл. 11.3).

Таблиця 11.3

Номер підприємства	Базисний рік		Звітний рік	
	Продуктивність праці w_0, тис. грн на чол.	Середня чисельність працюючих t_0, чол.	Продуктивність праці w_1, тис. грн на чол.	Середня чисельність працюючих t_1, чол.
1	14,3	1500	14,5	1510
2	59,6	423	60,0	420
Разом	73,9	1923	24,4	1930

Визначити:

1) індивідуальні індекси зміни обсягу виробленої продукції, чисельності працюючих; продуктивності праці;

2) загальні індекси обсягу виробленої продукції, зміни чисельності працюючих, продуктивності праці; показати взаємозв'язок загальних індексів;

3) загальні індекси обсягу виробленої продукції, зміни чисельності працюючих, продуктивності праці як середньозважені індекси;

4) абсолютний приріст обсягу виробленої продукції в цілому і за рахунок кожного чинника;

Зробити висновки з кожного пункту.

Розв'язання.

1. Індивідуальні індекси визначимо за формулами (11.2)-(11.3), тільки замість *p* і *q* використаємо *t* і *w*. Розрахунок наданий в табл. 11.4.

Таблиця 11.4

Індивідуальні індекси	По 1-му підприємству	По 2-му підприємству
Індекси зміни чисельності працюючих	$i_t^1 = \frac{t_1}{t_0} = \frac{1510}{1500} = 1{,}0067$, або 100,67%	$i_t^2 = \frac{420}{423} = 0{,}9929$, або 99,29%
Індекси рівня виробітку	$i_w^1 = \frac{w_1}{w_0} = \frac{14{,}5}{14{,}3} = 1{,}0139$, або 101,39%	$i_w^2 = \frac{60}{59{,}9} = 1{,}0067$, або 100,67%
Індекси обсягу виробленої продукції	$i_Q^1 = \frac{w_1 t_1}{w_0 t_0} = \frac{21895}{21450} = 1{,}02$, або 102%	$i_Q^2 = \frac{25200}{25210{,}8} = 0{,}9995$, або 99,95%

Висновок: значення індивідуальних індексів показують, що чисельність працюючих на першому підприємстві збільшилась на 0,67%, на другому – зменшилась на 0,71%; рівень виробітку на першому підприємстві збільшився на 1,39%, на другому – на 0,67%; обсяг виробленої продукції на першому підприємстві збільшився на 2,0%, на другому – зменшився на 0,05%.

2. Загальний індекс обсягу виробленої продукції обчислимо за формулою агрегатного індексу:

$$I_Q = \frac{14{,}5 \cdot 1510 + 60{,}0 \cdot 420}{14{,}3 \cdot 1500 + 59{,}6 \cdot 423} = \frac{47095{,}0}{46660{,}8} = 1{,}009305 \text{, або } 100{,}93\%.$$

Загальний індекс зміни чисельності працюючих обчислимо за формулою агрегатного індексу Ласпейреса:

$$I_t = \frac{\sum t_1 w_0}{\sum t_0 w_0} = \frac{14{,}3 \cdot 1510 + 59{,}6 \cdot 420}{14{,}3 \cdot 1500 + 59{,}6 \cdot 423} = \frac{46625{,}0}{46660{,}8} = 0{,}999233 \text{, або } 99{,}92\%.$$

Висновок: у зв'язку зі зміною чисельності працюючих обсяг виробленої продукції зменшився на 0,08%.

Загальний індекс продуктивності праці обчислимо за формулою агрегатного індексу Пааше:

$$I_w = \frac{\sum w_1 t_1}{\sum w_0 t_1} = \frac{14{,}5 \cdot 1510 + 60{,}0 \cdot 420}{14{,}3 \cdot 1510 + 59{,}6 \cdot 420} = \frac{47095{,}0}{46625{,}0} = 1{,}01008 \text{, або } 101{,}008\%.$$

Висновок: у зв'язку зі зміною рівнів продуктивності праці на підприємствах обсяг виробленої продукції збільшився на 0,008%.

Зауважимо, що майже такий результат можна отримати, якщо загальний індекс I_t обчислити за формулою агрегатного індексу Пааше, а индекс I_w – за формулою індексу Ласпейреса.

3. Обчислимо загальні індекси як середньозважені. Формули побуду-

ємо аналогично формулам (11.9)-(11.10), (11.13), тільки замість p і q використаємо t і w:

$$I_Q = \frac{\sum i_Q w_0 t_0}{\sum w_0 t_0} = \frac{1{,}02 * 21450 + 0{,}9995 * 25210{,}8}{21450 + 25210{,}8} = 1{,}009305;$$

$$I_t = \frac{\sum i_t t_0 w_0}{\sum t_0 w_0} = \frac{1{,}0067 * 21450 + 0{,}9929 * 25210{,}8}{21450 + 25210{,}8} = 0{,}999233;$$

$$I_w = \frac{\sum t_1 w_1}{\sum \frac{1}{i_w} t_1 w_1} = \frac{21825 + 25200}{\frac{21895}{1{,}0139} + \frac{25200}{1{,}0067}} = 1{,}01008.$$

4. Обчислимо абсолютний приріст як різницю між чисельником і знаменником відповідних агрегатних індексів.

Абсолютний приріст обсягу виробленої продукції:

$$\Delta Q = \sum w_1 t_1 - \sum w_0 t_0 = 47095 - 46660{,}8 = +433{,}2;$$

за рахунок факторів:

$$\Delta t = \sum t_1 w_0 - \sum t_0 w_0 = 46625 - 46660{,}8 = -35{,}8;$$

$$\Delta w = \sum w_1 t_1 - \sum w_0 t_1 = 47095 - 46625 = +470.$$

Висновок: обсяг виробленої продукції збільшився в цілому на 433,2 тис. грн, а завдяки збільшенню продуктивності на – 470 тис. грн. Оскільки чисельність працюючих зменшилась, обсяг виробленої продукції зменшився на 35 тис. грн.

Приклад 11.2

У регіоні виробництво продовольчих і непродовольчих товарів співвідноситься як 6:4. За деякий період воно зменшилося: продовольчих товарів – на 3%, промислових – на 7%, а ціни виросли на 4 і 6%. Через нерівномірність динаміки виробництва по групах товарів змінилася їхня структура: на 2 п.п. збільшилась частка продовольчих товарів і на стільки ж зменшилася частка непродовольчих.

Розрахувати середньозважені індекси цін та обсягу товарів.

Розв'язання.

Середньозважені індекси розрахуємо за формулою (11.5). Розрахунки індивідуальних індексів обсягу і цін наведено в табл. 11.5.

Таблиця 11.5

Товарні групи	Структура виробництва		Індивідуальні індекси		Розрахункові величини	
	d_0	d_1	i_q	i_p	$i_q\ d_0$	d_1/i_p
Продовольчі	0,60	0,62	0,97	1,04	0,582	0,596
Непродовольчі	0,40	0,38	0,93	1,06	0,372	0,358
Усього	1,00	1,00	-	-	0,954	0,954

Тоді

$I_q = 0{,}97 * 0{,}60 + 0{,}93 * 0{,}40 = 0{,}954$, тобто в середньому обсяги виробництва зменшилися на 4,6%;

$I_p = \dfrac{1}{\dfrac{0{,}62}{1{,}04} + \dfrac{0{,}38}{1{,}06}} = 1{,}048$, тобто ціни в середньому виросли на 4,8%.

Приклад 11.3

За даними табл. 11.6 визначити індекси середнього розміру тарифу страховки при страхуванні легкових автомобілів з терміном експлуатації до трьох років

Таблиця 11.6

Автомобіль	Тариф страхування, %		Сума страхування, тис. грн		Сума відшкодування страхування, тис. грн		
	базисний період, x_0	поточний період, x_1	базисний період, f_0	поточний період, f_0	x_0f_0	x_1f_1	x_0f_1
Вітчизняний	2,5	3,0	520	750	13,0	22,5	18,75
Зарубіжний	5,0	6,0	380	850	19,0	51,0	42,50
Всього	--	--	900	1600	32,0	73,5	61,25

Розв'язання.

Індекс змінного складу розрахуємо за формулою (11.18):

$$I_{\bar{x}} = \frac{73{,}5}{1600} : \frac{32}{900} = 0{,}046 : 0{,}036 = 1{,}278.$$

Він показує, що середній тариф страхування поточному періоді порівняно з базисним підвищився на 27,8%. Індекс фіксованого складу визначимо за формулою (11.19):

$$I_x = \frac{73{,}5}{1600} : \frac{61{,}25}{1600} = 0{,}046 : 0{,}038 = 1{,}211.$$

Внаслідок підвищення тарифу страхування по кожній групі автомобілів його середній тариф збільшився на 21,1%. Індекс структурних зрушень становить: I_d=0,038:0,036=1,056. Отже, середній тариф страхування збільшився на 5,6% за рахунок зміни у складі об'єктів страхування, а саме – збільшення частки суми страхування зарубіжних автомобілів з вищою ставкою страхування.

Індекс структурних зрушень знайдемо за допомогою формули (11.20):

$$I_d = 1{,}278/1{,}211 = 1{,}056.$$

Приклад 11.4

Відомо, що на ринках двох міст у звітному періоді реалізація товарів характеризувалася певними даними табл. 11.8.

Таблиця 11.8

Товар	Місто *А*		Місто В		Індивідуальний індекс цін	
	Модальна ціна за 1 ц, p_A	Кількість Т, q_A	Модальна ціна за 1 ц, p_B	Кількість, т, q_B	$i_{p_{A/B}} = \frac{p_A}{p_B}$	$i_{p_{B/A}} = \frac{p_B}{p_A}$
А	50	4	40	5	1,25	0,8
Б	70	7	60	8	1,17	0,86
В	120	3	130	2	0,92	1,08

Розв'язання.

Для характеристики співвідношення рівнів цін на товари, реалізованих у місті *А* порівняно з містом В, розраховують загальний індекс цін за формулою (11.24). Знаменник формули відображає умовний товарообіг, який міг бути у разі реалізації асортименту товарів за цінами, що склалися в місті В:

$$I_{p_{A/B}} = \frac{\sum q_A p_A}{\sum q_A p_B} = \frac{4*50+7*70+3*120}{4*40+7*60+3*130} = \frac{1050}{970} = 1{,}082.$$

Результат свідчить про те, що якби товари цього асортименту продавалися за цінами міста *А,* то їх рівень був би вище за рівень цін міста в середньому на 8,2 %.

Різниця між чисельником і знаменником розрахованого індексу відображає результат від різниці цін в порівнюваних містах:

$$\sum q_A p_A - \sum q_A p_B = 1050 - 970 = 80 \text{ грн.}$$

Отже, при продажу товарів зазначеного асортименту за цінами ринку міста В грошова виручка була б на 80 грн вищою за фактичний обсяг їх товарообігу в місті *А*.

Контрольні запитання

1. Суть і функції індексів.
2. Індивідуальні індекси і їх використання в економічному аналізі.
3. Загальні індекси і їх використання в економічному аналізі.
4. Агрегатні індекси. Формули розрахунку самих споживаних агрегатних індексів і їх взаємозв'язок.
5. Загальні індекси в базисно-ваговій системі (Ласпереса).
6. Загальні індекси в поточно-ваговій системі (Пааше).
7. Загальні індекси як середні з індивідуальних індексів.
8. Взаємозв'язки індексів.
9. Індекси середніх величин і структурних змін.
10. Територіальні індекси.

Бібліографічний список

1. Кулинич О. І. *Теорія статистики /* О.І. Кулинич. – К.: Вища шк., 1992. –135 с.

2. Общая теория статистики; под ред. А.А. Спириной, О.Э. Башиной. – М.: Финансы и статистика, 1996. – 296 с.

3. Статистика: підруч. для вузів; за ред. А. В. Головача. – К.: Вища шк., 1993. – 623 с.

4. Статистика: Збірник задач; за ред. А. В. Головача. – К.: Вища шк., 1994. – 448 с.

5. Теорія статистики: навч. посіб. / П. Г. Вашків, П. І. Пастер, В.П. Сторожук, Є.І. Ткач. – К.: Либідь, 2001. – 320 с.

Зміст

Петрик Валерія Леонідівна

СТАТИСТИКА

Редактор Є.О. Александрова

Зв. план, 2009
Підписано до друку 12.06.2009
Формат $60 \times 84^1/_{16}$. Папір офс. № 2. Офс. друк
Ум. друк. арк. 5,5. Обл.-вид. арк. 6,25. Замовлення 216. Наклад 60 прим.
Ціна вільна.

Національний аерокосмічний університет ім. М.Є. Жуковського
«Харківський авіаційний інститут»
61070, Харків-70, вул. Чкалова, 17
http://www.khai.edu
Видавничий центр «ХАІ»
61070, Харків-70, вул. Чкалова, 17
izdat@khai.edu