

**В. Л. Петрик, М. А. Голованова**

**СТАТИСТИКА**

2023

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
Національний аерокосмічний університет ім. М. Є. Жуковського  
«Харківський авіаційний інститут»

**В. Л. Петрик, М. А. Голованова**

## **СТАТИСТИКА**

Навчальний посібник

Харків «ХАІ» 2023

УДК 311(075.8)  
П30

Рецензенти: д-р екон. наук, проф. А. В. Федоренко,  
канд. екон. наук, доц. О. В. Дудник

**Петрик, В. Л.**

П30            Статистика [Електронний ресурс] : навч. посіб. / В. Л. Петрик,  
М. А. Голованова. – Харків : Нац. аерокосм. ун-т ім. М. Є. Жуковського  
«Харків. авіац. ін-т», 2023. – 137 с.

Розглянуто комплекс статистичних методів збирання, оброблення й аналізу статистичної інформації, теоретичні й методичні основи будовання системи статистичних показників, які використовують для вивчення закономірностей суспільних явищ. Особливу увагу приділено статистичній методології, можливостям її використання в умовах суттєвих змін в економіці.

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

Іл. 10. Табл. 51. Бібліогр.: 6 назв

**УДК 311(075.8)**

© Петрик В. Л., Голованова М. А., 2023  
© Національний аерокосмічний  
університет ім. М. Є. Жуковського  
«Харківський авіаційний інститут», 2023

# 1. МЕТОДОЛОГІЧНІ ЗАСАДИ СТАТИСТИКИ. СТАТИСТИЧНЕ СПОСТЕРЕЖЕННЯ

## 1.1. Предмет і метод статистичної науки

**Статистика** (від лат. status – стан) – наука, у якій вивчають розміри й кількісні співвідношення масових суспільно-економічних явищ і процесів у нерозривному зв'язку з їх змістом.

Статистика має багатовікову історію. Її виникнення і розвиток були зумовлені такими суспільними потребами: підрахунок населення, худоби, облік земельних угідь, майна і т. ін. Найперші відомості про такі роботи в Китаї припадають на XV ст. до н. е.

З розвитком суспільного виробництва, внутрішньої й зовнішньої торгівлі збільшувалася потреба в статистичній інформації. Це розширювало сферу діяльності статистики, вдосконалювало її прийоми й методи.

Уважається, що основи статистичної науки було закладено англійським економістом У. Петті (1623–1687). Послідовники У. Петті започаткували науковий напрям, який дістав назву «політична арифметика».

Засновником іншого напрямку розвитку статистичної науки був визначний німецький учений Г. Конрінг (1606–1681), який розробив систему опису держави. Його послідовник професор філософії й права Г. Ахенваль (1719–1772) уперше в Марбурзькому університеті (1746 р.) почав викладати нову дисципліну, названу ним статистикою. Цей напрям розвитку статистики дістав назву *описового*.

У XIX ст. бельгійський статистик А. Кетле заклав основу вченню про *середні величини*.

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

**Предмет статистики** – розміри і кількісні співвідношення масових суспільних явищ, закономірності їх формування, розвитку та взаємозв'язку.

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

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

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

**Статистичне спостереження** – це сплановане, систематичне й науково організоване збирання масових даних про різноманітні соціально-економічні явища й процеси.

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

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

Під час **аналізу статистичної інформації** передбачається проведення аналізу даних на основі обчислення узагальнювальних 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 тощо.

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

Основні категорії теорії статистики:

1. **Статистична сукупність** – велика кількість одиниць явища, об'єднаних відповідно до задачі дослідження єдиною якісною основою, загальним зв'язком, але таких, що різняться окремими ознаками (наприклад, сукупність промислових підприємств, фірм, сукупність сімей тощо).

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

2. **Одиниця статистичної сукупності** – первинний елемент статистичної сукупності, який є носієм ознак, що підлягають реєстрації. Наприклад, при проведенні перепису 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 фондів, кількість персоналу та ін.; демографічні й соціально-економічні ознаки людини: вік, рівень освіти, професія, стать і т. ін.

4. **Статистичний показник** – узагальнювальна характеристика соціально-економічного явища або процесу, у якій поєднуються якісна й кількісна визначеність останнього (наприклад, кількість населення, продукція промислового підприємства, рівень продуктивності праці, рівень рентабельності й т. ін.).

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

## 1.2. Статистичне спостереження

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

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

**Мета статистичного спостереження** – отримання статистичних даних, які є основою для узагальненої характеристики стану й розвитку явища або процесу.

**Об'єкт спостереження** – сукупність соціально-економічних явищ і процесів, які підлягають дослідженню, або точні межі, у яких реєструватимуться статистичні дані.

Статистичне спостереження може бути:

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

Про статистичне спостереження можна казати лише тоді, коли:

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

Тому статистичне спостереження має 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 таких пунктів:

- 1) формулювання мети обстеження й конкретних завдань;
- 2) визначення об'єкта обстеження;
- 3) наведення складових частин об'єкта (одиниць сукупності);
- 4) наведення джерел інформації (одиниць спостереження);
- 5) розроблення програми спостереження, вибір виду й способу спостереження.

**Одиниця статистичного спостереження** – це та первинна ланка, з якої має бути отримано необхідні статистичні дані.

Наприклад, при проведенні перепису населення об'єктом спостереження є сукупність людей, що проживають у певній країні на цей час. Одиниця сукупності – окрема людина, і в цьому випадку така одиниця збігається з одиницею спостереження. При визначенні обсягу роздрібного товарообігу

одинацями спостереження є торговельні підприємства, одинацями сукупності – акти купівлі-продажу товарів населенню.

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

Для забезпечення однакового тлумачення програми в обліковому формулярі, у якому її розміщено, даються пояснення, викладені в документі, який має назву **інструкції**.

До програми спостереження також включають: розроблення статистичного інструментарію, визначення виду й способу обстеження.

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

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

Під час складання формулярів ураховують не тільки зміст та інформативність ознак, але й можливість їх статистичного оброблення, яке забезпечується завдяки вживанню системи шкал.

**Шкала** – це засіб упорядкування й кількісного вираження ознак. Використовують такі види шкал:

1) номінальна – шкала найменувань, у якій встановлено відношення подібності елементів, порядок розташування ознак при цьому значення не має. Наприклад, класифікатор сфер економічної діяльності, перелік форм власності, видів підприємницької діяльності й т. ін.;

2) порядкова (рангова) – шкала, у якій встановлено послідовність інтенсивності виявлення ознаки. Її застосовують під час визначення ступеня економічного ризику підприємництва, рівня кваліфікації робітників або ставлення 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. За метричною шкалою визначають обсяги виробництва й реалізації продукції, розміри капіталу, кількість зайнятих у виробництві осіб, кількість і вартість приватизованих об'єктів тощо.

**Організаційні питання** статистичного спостереження містять 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** – це офіційний документ, що містить статистичні дані про роботу підприємств і фірм у вигляді обов'язкових звітів про їх роботу за встановленими формами й у відповідних термінах (у вигляді формулярів регламентованого зразка).

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

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

**Перепис** – це спеціально організоване статистичне спостереження, основне завдання якого полягає в повному обліку кількості й характеристиці складу явища шляхом запису до статистичного формуляра даних про досліджувані одиниці статистичної сукупності.

**Одноразовий облік** – це перепис, у якому статистичні формуляри заповнюють на основі матеріалів первинного обліку. Наприклад, перепис залишків сировини, устаткування, облік посівних площ.

**Способи спеціального статистичного обстеження:**

1. Безпосереднє спостереження, коли відомості одержують шляхом особистого обліку одиниць сукупності: перерахунку, вимірювання і т. д.

2. Документальний спосіб збирання статистичної інформації, який базується на систематичних записах у первинних документах, що підтверджують той чи інший факт.

3. Опитування, під час якого передбачається, що відомості фіксуються зі слів опитуваного. Розрізняють три способи такого обстеження:

– експедиційний – призначена особа опитує іншу і з його слів заповнює бланк дослідження;

– анкетний;

– самореєстрація – відповідні документи заповнюють опитувані;

– кореспондентський спосіб – до органів, які здійснюють спостереження, надсилають відомості їх кореспонденти.

Залежно від завдань 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** поділяють за такими ознаками: повнота охоплення одиниць сукупності, час реєстрації фактів, джерела відомостей, способи реєстрації та збору даних.

З огляду на повноту охоплення фактів (одиниць) статистичне спостереження може бути суцільним і несуцільним.

**Суцільне спостереження** є повним обліком усіх одиниць сукупності, що вивчається (наприклад, перепис населення).

**Несуцільне спостереження** організують як облік частини одиниць сукупності, на основі якої можна отримати узагальнювальну характеристику всієї сукупності.

Види несуцільного спостереження:

1) **спосіб основного масиву** – вибирають найкрупніші одиниці спостереження, у яких зосереджено значну частку всіх фактів, що підлягають вивченню;

2) **вибіркові спостереження** – випадково вибирають зі всієї сукупності обмежену кількість об'єктів;

3) **монографічні описи**, які 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 виторг магазинів від реалізації товарів, реєстрацію актів цивільного стану або роботу транспорту).

**Періодичним** вважається спостереження, що повторюється 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 узагалі лише один раз (наприклад, вивчення думки читачів або глядачів з приводу того чи іншого питання, думки покупців про якість товарів і т. ін.).

Під час **безпосереднього обліку** представники органів статистики одержують необхідну інформацію про досліджувану сукупність шляхом особистого підрахунку. Наприклад, під час інвентаризації основних засобів до облікового формуляра відомості заносять на основі особистого огляду реєстратором кожного об'єкта.

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

## Приклади розв'язання типових задач

### Задача 1.1

Дослідити успішність студентів другого курсу економічного факультету й фактори, що впливають на результати зимової екзаменаційної сесії. Установити об'єкт спостереження, одиницю сукупності, скласти програму спостереження.

#### *Розв'язання*

Об'єктом спостереження є студенти другого курсу економічного факультету.

Одиницею сукупності є окремих студент.

Одиницею звітності є деканат економічного факультету.

Програма статистичного спостереження містить такі пункти:

- 1) оцінки, які отримали студенти з кожного предмета;
- 2) інформація для виявлення факторів, що впливають на успішність:
  - а) середній бал атестата;
  - б) відвідування занять;
  - в) своєчасність виконання письмових робіт;
  - г) сімейний стан;
  - д) бюджет часу.

### Задача 1.2

Визначити перелік питань програми проведення статистичного спостереження:

- 1) вибіркоче обстеження плинності працівників у будівельних організаціях;
- 2) обстеження зайнятих працівників на фірмі за рівнем освіти й спеціальності;
- 3) вибіркоче обстеження заробітної плати працівників підприємств промисловості.

#### *Розв'язання*

Програма спостереження – перелік запитань, відповіді на які отримують під час спостереження. Для вирішення організаційних питань складається організаційний план статистичного спостереження, де визначають мету, вид, форму, спосіб спостереження, місце й терміни його проведення. Склад і зміст питань програми спостереження залежать від завдань дослідження й особливостей досліджуваного соціального явища.

Згідно з умовами цього завдання можна запропонувати такі питання:

1. Статистичне спостереження вибіркового обстеження плинності працівників у будівельних організаціях.

Мета: провести вибіркоче обстеження плинності працівників у будівельних організаціях.

Об'єкт спостереження: сукупність працівників у будівельних організаціях.

Одиниця спостереження: кожен окремий працівник.

Форма: спеціально організоване спостереження.

Вид: періодичне несучільне спостереження.

Спосіб: вибіркоче спостереження.

2. Статистичне спостереження обстеження зайнятих працівників на фірмі за рівнем освіти й спеціальності.

Мета: провести обстеження зайнятих працівників на фірмі за рівнем освіти й спеціальності.

Об'єкт спостереження: сукупність працівників фірми.

Одиниця спостереження: кожен окремий працівник.

Форма: звітність.

Вид: поточне суцільне спостереження.

3. Статистичне спостереження вибіркового обстеження заробітної плати працівників підприємств промисловості.

Мета: провести вибіркоче обстеження заробітної плати працівників підприємств промисловості.

Об'єкт спостереження: сукупність працівників підприємств промисловості.

Одиниця спостереження: кожен окремий працівник.

Форма: спеціально організоване спостереження.

Вид: періодичне несучільне спостереження.

Спосіб: вибіркоче спостереження.

### **Завдання для самостійної роботи**

1. Сформулювати об'єкт, одиницю, мету спостереження й розробити програму обстеження вузів міста.

2. Визначити об'єкт, одиницю спостереження й одиницю сукупності обстежень:

а) оцінювання якості підготовки студентів з фахових дисциплін у державних економічних навчальних закладах;

б) перепис виробничих площ на державних підприємствах промисловості;

в) облік наявності касових апаратів у торгових пунктах міста.

3. Для проведення обстеження інвестиційної привабливості об'єктів готельного комплексу регіону визначити:

а) мету спостереження;

б) об'єкт спостереження;

в) одиницю сукупності;

г) одиницю спостереження.

**4.** Скласти перелік ознак для характеристики:

- промислового підприємства;
- сільськогосподарського підприємства;
- торгового підприємства;
- комерційного банку;
- бібліотеки;
- транспортного підприємства.

**5.** Визначити, до якого виду належать наведені нижче спостереження за формою спостереження, ступенем охоплення одиниць сукупності, часом обліку фактів, способом отримання інформації:

- а) перепис населення області;
- б) обстеження сімейного бюджету;
- в) щомісячний звіт про собівартість продукції;
- г) опитування молодих сімей регіону з питань планування сім'ї;
- д) анкетне опитування пасажирів авіакомпанії, яка обслуговує міжнародні рейси;
- е) обстеження цін на ринках міста;
- ж) облік клієнтів комерційного банку.

**6.** Визначити місце, час та органи проведення статистичних спостережень:

- а) опитування учасників виробничо-торгового ярмарку, яка проводилася з 1.04 по 9.04;
- б) облік доходів громадян і джерел їх надходжень, який здійснюється податковими інспекціями за підсумками року.

### **Контрольні запитання**

1. Хто і коли започаткував описовий напрям у статистиці? У чому полягає суть цього напрямку?

2. З якого часу статистика почала розвиватися як наука? Який внесок зробив У. Петті у розвиток статистики?

3. Наведіть означення предмета статистики. Чим відрізняється предмет статистики від предметів інших суспільних наук?

4. Назвіть коло суспільних явищ, що вивчаються в статистиці.

5. Чи обов'язково входять до статистичної сукупності якісно однорідні елементи 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?
11. Що таке статистичне спостереження? У чому полягає його суть?
12. Назвіть організаційні форми статистичного спостереження та їхні особливості.
13. У чому полягає суть статистичної звітності? Які існують види звітності?
14. Що таке спеціально організоване статистичне спостереження?
15. Назвіть види спеціально організованих статистичних спостережень.
16. Які види статистичного спостереження за ступенем охоплення одиниць сукупності ви знаєте?
17. Яке значення для організації статистичною дослідження має програма спостереження?
18. Що саме містять програмно-методологічні та організаційні питання статистичного спостереження?
19. Поясніть, що таке одиниця статистичного спостереження та одиниця статистичного обліку.
20. Що таке об'єкт статистичного спостереження?

## 2. ЗВЕДЕННЯ ТА ГРУПУВАННЯ СТАТИСТИЧНИХ ДАНИХ

### 2.1. Поняття статистичного зведення та групування

**Статистичне зведення** – це упорядкування, систематизація й наукове оброблення даних з метою отримання узагальненої характеристики досліджуваного об'єкта або процесу.

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

За складністю побудови зведення поділяють на прості й групові.

При **простому підсумковому зведенні** не передбачається попереднього розподілу на групи отриманих відомостей, а визначається загальний підсумок усіх одиниць сукупності або загальний обсяг досліджуваного показника. Наприклад, щоб визначити загальну кількість студентів в Україні, достатньо скласти дані про кількість студентів у всіх вищих навчальних закладах.

При **груповому зведенні** передбачається попередній поділ одиниць на групи (наприклад, рентабельні й збиткові підприємства), що дає змогу підрахувати кількість одиниць сукупності й обсяг досліджуваної ознаки в кожній групі.

За способом здійснення статистичне зведення поділяють на централізоване й децентралізоване.

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

При **децентралізованому зведенні** матеріал узагальнюється від низу до верху по ієрархічних сходинках керування з відповідним його обробленням на кожній з них.

**Групування** – це розбиття сукупності на однорідні групи на основі поділу всієї сукупності досліджуваного явища на окремі групи за найістотнішими ознаками.

Основні задачі, які вирішуються з допомогою статистичних групувань: виокремлення соціально-економічних типів явищ, вивчення структури досліджуваних явищ і структурних змін, дослідження взаємозв'язку й залежності між ознаками суспільних явищ. Згідно з цими задачами розрізняють типологічні, структурні й аналітичні групування.

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

**Структурне групування** характеризує структуру якісно однорідної сукупності за якою-небудь однією варіювальною ознакою, яка дає змогу описати складові частини сукупності та проаналізувати структурні зрушення (наприклад, вивчення підприємств за галузями виробництва, розмірами основних виробничих фондів, 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. 2.1).

Таблиця 2.1

Межі групування за істотною ознакою	Кількість одиниць сукупності	Система показників		
Разом				

**Аналітичне групування** дає змогу оцінювати взаємозв'язки між двома (і більше) взаємодійними ознаками, з яких одна розглядається як наслідок, а інша – як фактор. У кожній групі факторної ознаки визначається середня величина результативної ознаки (табл. 2.2).

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

Таблиця 2.2

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

**Комбінаційне групування** містить групи, виокремлені за однією ознакою, які поділяються на підгрупи за іншою ознакою.

Особливим видом групування, яке часто використовується в статистиці, є класифікація – стійке розмежування об'єктів, що ґрунтується на найістотніших ознаках (наприклад, класифікація галузей народного господарства, класифікація основних фондів і т. д.).

## 2.2. Правила проведення групування

При проведенні групування доводиться вирішувати такі завдання:

- 1) виокремлення групувальної ознаки;
- 2) визначення кількості груп і величини інтервалів;
- 3) опис того, як декілька групувальних ознак комбінуються між собою;
- 4) установлення показників, за якими характеризують групи.

За **групувальною ознакою** об'єднують окремі одиниці сукупності в однорідні групи.

Ознаки розрізняють за способами їх вимірювання й іншими особливостями, що впливають на прийоми статистичного вивчення. У табл. 2.3 наведено класифікацію ознак.

Таблиця 2.3

За формою виразу	За способом обчислення	Відносно об'єкта	За характером варіації	За ознакою часу	За функцією 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, їх можна виміряти, порахувати, зважити. Ці ознаки 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, урожайність).

**Прямі ознаки** – це властивості самого об'єкта (обсяг продукції заводу, кількість його робітників тощо).

**Непрямі ознаки** – це властивості, що є характерними не для самого об'єкта, а для інших сукупностей, які належать до об'єкта (оплата праці робітників – непрямая 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 (кількість працівників).

**Моментні ознаки** характеризують об'єкт у якийсь момент часу, установлений планом статистичного дослідження (кількість населення, 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 такі ознаки, розмір і динаміка яких формуються під впливом інших ознак.

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

**Інтервал** окреслює кількісні межі груп і є зазвичай проміжком між максимальними й мінімальними значеннями ознаки в групі. Інтервали поділяють на такі:

– *відкриті*, коли є тільки або верхня, або нижня межа;

– *закриті*, коли є і нижня, і верхня межі.

Закриті інтервали поділяють на такі:

– *однакові*, коли різниця між максимальним і мінімальним значеннями в кожному з інтервалів є однаковою;

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

Кількість груп **при однакових інтервалах** визначають за формулою Стерджеса

$$n = 1 + 3,322 \lg n, \quad (2.1)$$

де  $n$  – кількість елементів сукупності.



Користуватися цією формулою можна лише в тих випадках, коли досліджувана сукупність є досить великою і змінення груп ознаки має відносно повільний характер.

Величину однакових інтервалів можна визначити за формулою

$$l = \frac{X_{\max} - X_{\min}}{n}, \quad (2.2)$$

де  $X_{\max}$ ,  $X_{\min}$  – відповідно максимальне й мінімальне значення ознаки;  
 $n$  – кількість елементів сукупності.

Нижня межа першого інтервалу  $a_0$  може збігатися з мінімальним значенням ознаки, її також можна визначити за формулою

$$a_0 = X_{\min} - \frac{l}{2}. \quad (2.3)$$

Верхню межу першого інтервалу й нижню межу другого розраховують за формулою

$$a_i = a_0 + l. \quad (2.4)$$

Межі подальших інтервалів знаходять за формулою

$$a_{j+1} = a_j + l. \quad (2.5)$$

Основою будь-якого групування є ряд розподілу.

**Ряд розподілу** – це групування, у якому для характеристики груп, упорядкованих за значенням групувальної ознаки, застосовується один показник – чисельність групи.

Ряд розподілу складається з двох елементів – варіантів і частот.

**Варіанти** – це окремі значення групувальної ознаки.

**Частоти** – це числа, що відображають кількість повторень окремих значень варіантів або кожної групи варіаційного ряду.

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

Нагромаджену частоту, або частість, називають **кумулятивною**.

**Обсяг сукупності** – сума всіх частот, що визначає кількість елементів усієї сукупності.

Ряди розподілу, побудовані за атрибутивною ознакою, називають **атрибутивними** (наприклад, розподіл населення за статтю, зайнятістю, національністю, професією тощо).

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

Варіаційні ряди залежно від способу побудови поділяють на дискретні й інтервальні.

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

Якщо ознака може набувати хоча й дискретних значень, але їх кількість є великою, або дробових значень в області свого існування, то будують **інтервальний варіаційний ряд** – таблицю з двох рядків або граф: інтервалів ознаки, варіація якої вивчається, і частот.

Засобом наочного вираження результатів статистичного дослідження є статистичні таблиці, які призначено для подання результатів зведення й групування матеріалів спостереження.

**Підмет таблиці** – статистична сукупність, об'єкти або їх частини, які характеризуються кількома числовими показниками.

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

Складену таблицю, але не заповнену цифрами, називають **макетом статистичної таблиці** (табл. 2.4).

Таблиця 2.4

Загальний заголовок

Підмет	Присудок						
	Заголовки присудка (верхні заголовки)						
А	1	2	3	4	5	6	7
Підмет (бокові заголовки)							
Підсумок							

У комерційній роботі складаються різноманітні таблиці, які залежно від побудови підмета поділяють на прості, групові й 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 кількома показниками.

Найпростіший вид групових таблиць – ряди розподілу, побудовані за атрибутивними або кількісними ознаками. Присудок таких таблиць має лише один показник – кількість одиниць сукупності, які належать 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 – у присудку. Такий вигляд мають комбінаційні ряди розподілу.

## Приклад розв'язання типових задач

Є дані про робітників підприємства (табл. 2.5).

1. Побудувати ряд розподілу робітників за стажем, утворивши п'ять груп з однаковими інтервалами. Зажною групою визначити кількість робітників та їх питому вагу до загальної кількості робітників.

2. Для вивчення залежності між стажем роботи і місячним виробітком працівників виконати:

а) аналітичне групування робітників за стажем, утворивши п'ять груп з однаковими інтервалами. Кожну групу охарактеризувати за кількістю робітників; середнім стажем роботи; місячним виробітком продукції – усього й у середньому на одного робітника;

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

Таблиця 2.5

№ п/п	Стаж роботи, кількість років	Місячний виробіток робітником нормативно-чистої продукції (НЧП), грн
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

### Розв'язання

1. Обчислимо величину інтервалу ознаки групування (стажу роботи) за формулою (2.2):

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

Отже, перша група робітників має стаж 1–4 роки; друга – 4–7 і т. д. Для кожної групи потрібно підрахувати кількість робітників і оформити результати у вигляді табл. 2.6. Для проведення подальшого аналізу кількість робітників можна подати у відсотках (табл. 2.6, графа 4).

Таблиця 2.6

Номер групи	Групи робітників за стажем, кількість років	Кількість робітників, чол.	Кількість робітників, %
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

З результатів групування випливає, що більше половини робітників, тобто 60,1 %, мають стаж роботи від 1 до 7 років.

2а. Застосовуючи метод групувань для взаємозв'язку, визначимо факторну ознаку, яка впливає на взаємозалежні ознаки. Такою ознакою в цьому прикладі є стаж роботи, який має бути покладено в основу групування. За умовою потрібно виокремити п'ять груп робітників за стажем з однаковими інтервалами. За основу аналітичного групування візьмемо ці ж групи.

Для будування й оформлення результатів групування спочатку оформимо робочу табл. 2.7, де для кожної групи робітників за стажем підрахуємо кількість робітників, їх загальний стаж і місячний виробіток.

Таблиця 2.7

№ п/п	Групи робітників за стажем, кількість років	Номер робітника	Стаж, кількість 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
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
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 до відповідних граф зведеної аналітичної табл. 2.8.

Таблиця 2.8

Номер групи	Групи робітників за стажем, кількість років	Кількість робітників, чол.	Середній стаж роботи, кількість років	Місячний виробіток НЧП, грн	
				усього	на одного робітника
1	2	3	4	5	6
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 4 і 6 табл. 2.8, бачимо, що зі збільшенням стажу робітників збільшується місячний виробіток продукції, отже, між досліджуваними ознаками існує пряма залежність.

2б. Щоб зробити комбінаційне групування за двома ознаками (табл. 2.9) – стажем роботи і середнім місячним виробітком продукції – потрібно в кожній групі робітників за стажем виокремити групи за другою ознакою (середнім місячним виробітком продукції на одного робітника) та охарактеризувати групи за необхідними показниками.

Таблиця 2.9

№ п/п	Групи робітників		Кількість робітників, чол.	Середній стаж роботи, кількість років	Середньомісячний виробіток продукції, грн	
	За стажем, кількість 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	

## Закінчення табл. 2.9

№ п/п	Групи робітників		Кількість робітників, чол.	Середній стаж роботи, кількість років	Середньомісячний виробіток продукції, грн	
	За стажем, кількість років	За середньомісячним виробітком продукції, грн			усього	на одного робітника
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	

З даних таблиці випливає, що виробіток продукції робітників має пряму залежність від стажу.

Ще один вид комбінаційного групування наведено в табл. 2.10.

Таблиця 2.10

Групи робітників за стажем, кількість років	Групи робітників за середньомісячним виробітком продукції, грн					Разом
	200–224	224–248	248–272	272–296	296–320	
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

## Завдання для самостійної роботи

1. Відомо такі дані бюджетних обстежень 20 домогосподарств району:

№ п/п	Кількість членів, осіб	Грошовий дохід, грн	№ п/п	Кількість членів, осіб	Грошовий дохід, грн
1	2	1850	11	3	5160
2	3	2680	12	3	3740
3	4	5390	13	4	4500
4	2	1930	14	3	6030
5	3	4730	15	3	2290
6	3	3240	16	2	3680
7	4	7100	17	4	3130
8	3	1720	18	3	3460
9	4	2480	19	3	4470
10	2	3500	20	4	3920

Згрупувати домогосподарства за такими ознаками:

1) кількість членів кожного домогосподарства; за кожною групою визначити кількість домогосподарств і загальну кількість членів;

2) грошовий дохід кожного домогосподарства, утворивши чотири інтервали (менше 2000, 2000–4000, 4000–6000, 6000 і більше); за кожною групою визначити кількість домогосподарств і загальний грошовий дохід;

3) кількість членів кожного домогосподарства і його грошовий дохід, зробивши висновок про взаємозв'язок цих ознак.

Результати групування викласти у формі статистичних таблиць.

На основі даних про загальну кількість членів і загальний грошовий дохід домогосподарств кожної групи визначити середній дохід на одне домогосподарство й на одного члена домогосподарства.

2. Є такі дані про суму активів та кредитні вкладення 20 комерційних банків:

Номер банку	Кредитні вкладення, млрд грн	Сума активів, млрд грн	Номер банку	Кредитні вкладення, млрд грн	Сума активів, млрд грн
1	311	518	11	3900	6728
2	658	1194	12	305	497
3	2496	3176	13	799	1732
4	1319	1997	14	914	2002
5	783	2941	15	1039	2295
6	1962	3066	16	2822	5636
7	1142	1865	17	1589	2998
8	382	602	18	1012	1116
9	853	1304	19	1350	2482
10	2439	4991	20	3500	6453

З метою вивчення залежності суми активів і кредитних вкладів комерційних банків здійснити аналітичне групування банків за кредитними вкладками (факторна ознака), утворивши п'ять груп з однаковими інтервалами. За кожною групою й сукупністю банків підрахувати:

- 1) кількість банків;
- 2) кредитні вклади – усього й у середньому на один банк;
- 3) суму активів – усього й у середньому на один банк.

Результати подати у вигляді групувальної таблиці. Зробити висновки.

**3.** За наведеними даними про порушення технологічної дисципліни й утрати від браку продукції на 22 виробничих ділянках скласти:

а) аналітичне групування, яке б відобразило залежність утрат від порушення технологічної дисципліни, визначити ефекти впливу порушення технологічної дисципліни на втрати від браку продукції;

б) комбінаційний поділ виробничих ділянок за цими ознаками, утворивши три групи з однаковими інтервалами.

Зробити висновки про наявність і напрямок зв'язку між ознаками.

Номер ділянки	Порушення технологічної дисципліни, %	Утрати від браку продукції, тис. грн	Номер ділянки	Порушення технологічної дисципліни, %	Утрати від браку продукції, тис. грн
1	1,2	1,0	12	1,7	1,5
2	2,0	1,6	13	2,1	1,7
3	1,4	1,2	14	1,3	1,4
4	1,9	1,5	15	2,0	1,8
5	1,6	1,4	16	2,3	1,6
6	2,4	1,9	17	2,5	2,0
7	1,8	1,4	18	2,7	2,1
8	2,6	2,1	19	2,6	2,0
9	2,0	1,7	20	1,7	1,4
10	1,5	1,2	21	1,5	1,3
11	1,2	0,9	22	2,1	1,6

Результати групування викласти у вигляді статистичних таблиць, проаналізувати їх.

### Контрольні запитання

1. Що таке другий етап статистичного дослідження?
2. Які види зведення ви знаєте? Стисло їх охарактеризуйте.
3. Що саме називають статистичним групуванням і групувальними ознаками?
4. Які завдання вирішуються в статистиці з допомогою методу групування?



5. Назвіть основні види статистичних групувань.
6. У чому полягають принципи вибору груповальної ознаки й утворення груп?
7. Які угруповання називають типологічними, структурними, аналітичними?
8. У чому полягає взаємозв'язок усіх зазначених видів групування?
9. Як визначають кількість груп і межі інтервалів між ними?
10. Що таке ряди розподілу і за якими ознаками їх утворюють?
11. За якими принципами поділяють 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 полягає перевага групових і комбінаційних таблиць порівняно з простими?

### 3. СТАТИСТИЧНІ ГРАФІКИ

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

**Основні елементи графіка:**

- **поле графіка** – простір, на якому розміщують графічне зображення;
- **графічний образ** – сукупність геометричних або графічних знаків, з допомогою яких відображаються статистичні дані;
- **масштабні орієнтири** – масштаб, масштабні шкали, масштабні знаки для визначення розмірів;
- **просторові орієнтири** – 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 виокремлюють 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. ін.
3. За способом будування (за виглядом поля) графіки поділяють на діаграми й статистичні карти (картодіаграми й картограми).

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

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

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

6. За функціонально-цільовим призначенням розрізняють графіки:

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

– структури й структурних зрушень (структурні діаграми – стовпцеві, стрічкові, секторні);

– зображення динаміки статистичних показників (динамічні графіки, стовпцеві, стрічкові, квадратні, кругові, фігурні, діаграми темпів, лінійні діаграми (їх різновид – радіальні));

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

– поширення в просторі (картограма, картодіаграма, центрограма);

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

– взаємозв'язку й взаємозалежності (графіки кореляційної залежності).

**Діаграма** – це креслення, на якому статистична інформація зображується з допомогою геометричних фігур, ліній або символічних знаків.

**Стовпцева діаграма** – це графік, на якому статистичні дані зображено у вигляді стовпців-прямокутників однакової ширини й будь-якої висоти, розташованих вертикально на осі абсцис. Кожний стовпець характеризує окремий об'єкт.

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

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

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

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

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

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

**Динамічні графіки** призначено для зображення економічних явищ, що відбуваються в часі. У динамічних діаграмах об'єктом відображення є процеси.

**Лінійні графіки** характеризують змінення явищ у часі, залежність між двома показниками. У статистиці комерційної діяльності на ринку товарів і послуг ці графіки є найбільш поширеними.

Різновидом лінійних діаграм є **радіальні**, 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 штрихуванням, забарвленням, крапками.

**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 застосовують лінійні та площинні діаграми, побудовані в прямокутній системі координат.

При дискретній варіації ознаки графіком варіаційного ряду є **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** – це графічне зображення інтервального варіаційного ряду, у якому на осі абсцис відкладаються варіанти, а прямокутники є пропорційними за висотою частотам значень ознаки для кожного інтервалу.

**Кумуляти (кумулятивні діаграми)** використовують для графічного порівняння двох або більше варіаційних розподілів з однаковими або неоднаковими інтервалами. На осі абсцис відкладають відрізки інтервалів групувань, на осі ординат – нагромаджені частоти або частоті.

## Приклади розв'язання типових задач

### Задача 3.1

Показники виробництва телевізорів деякими заводами країни за певний період часу наведено в табл. 3.1.

Таблиця 3.1

Завод	Загальний випуск телевізорів, тис. шт.	Середня ціна одного телевізора, грн	Загальна вартість виготовлених телевізорів, тис. грн
I	30	650	19500
II	25	450	11250
III	15	670	10050

Побудувати графіки порівняння статистичних величин для діаграм:

- 1) стовпцевої;
- 2) стрічкової;
- 3) прямокутної.

#### Розв'язання

1. Під час будування стовпцевої діаграми на осі абсцис на однаковій відстані відкладають три відрізки однакової довжини – основи стовпців. Назви заводів відображають на графіку ранжовано. Діаграма дає наочне порівняння стовпців за висотою відповідно до кількості телевізорів. Внизу під стовпцями вказують назви об'єктів порівняння – заводів за їх номером (рис. 3.1).

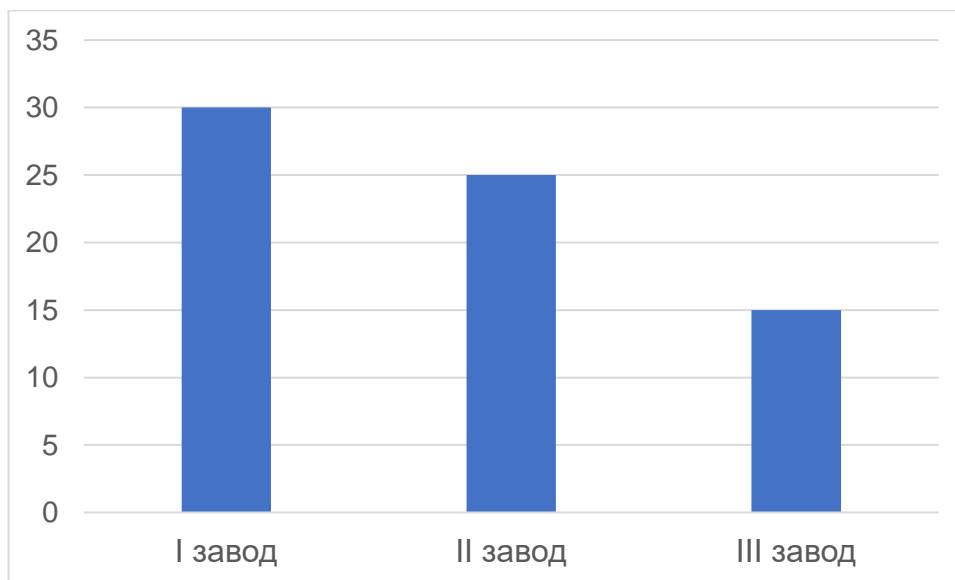


Рис. 3.1. Стовпцева діаграма виробництва телевізорів заводами країни за деякий час, тис. шт.

2. Для стрічкової діаграми характерними є горизонтальна орієнтація стовпців і вертикальне розташування базової лінії (рис. 3.2).

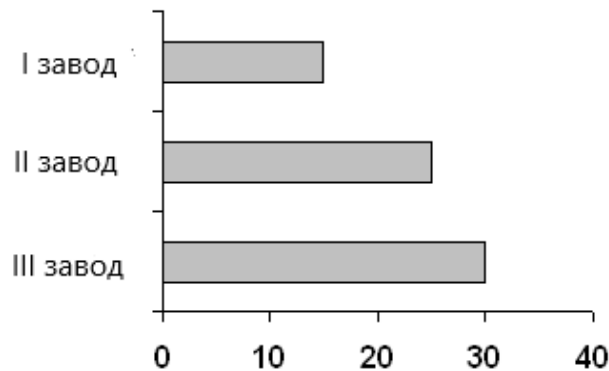


Рис. 3.2. Стрічкова діаграма виробництва телевізорів заводами країни за деякий час, тис. шт.

3. Для будування прямокутної діаграми (рис. 3.3) 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 = 1 см).

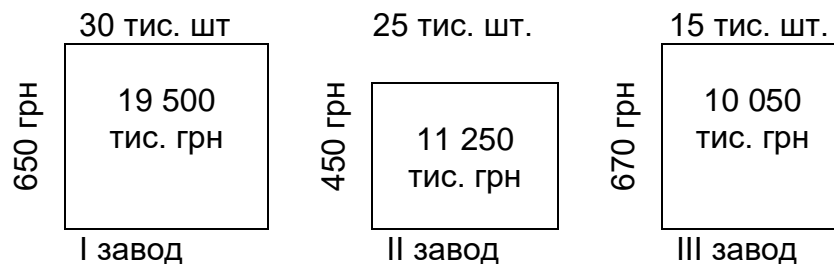


Рис. 3.3. Прямокутна діаграма: порівняльні обсяги роботи деяких заводів за кількістю виготовлених телевізорів, середньою ціною й вартістю

### Задача 3.2

Для ряду розподілу робітників за стажем (табл. 3.2) побудувати полігон і гістограму розподілу та кумулятивну гістограму.

Таблиця 3.2

Номер групи	Групи робітників за стажем, кількість років	Кількість робітників, чол.	Кількість робітників, %
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. Щоб побудувати полігон для інтервального ряду по осі  $x$ , необхідно відкласти середини інтервалів – груп робітників за стажем, а на осі ординат – кількість робітників у відсотках (рис. 3.4).

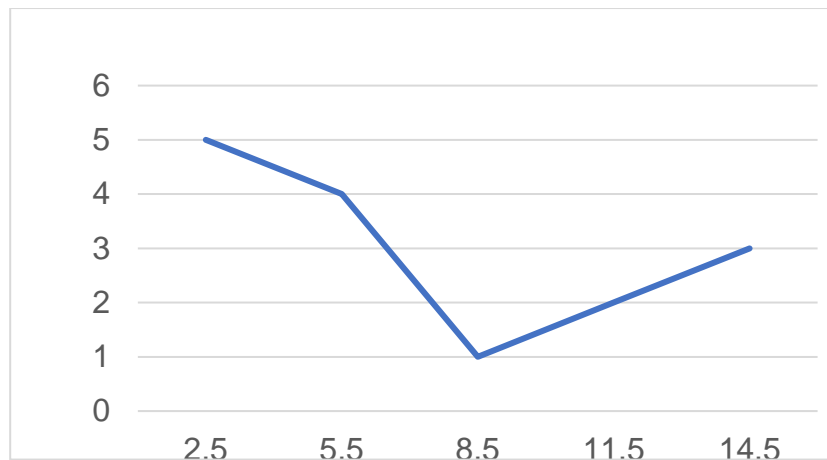


Рис. 3.4. Полігон розподілу робітників за стажем

2. Для будування гістограми розподілу робітників за стажем (рис. 3.5) на осі абсцис відкладають межі інтервалів, на осі ординат – кількість робітників у відсотках.

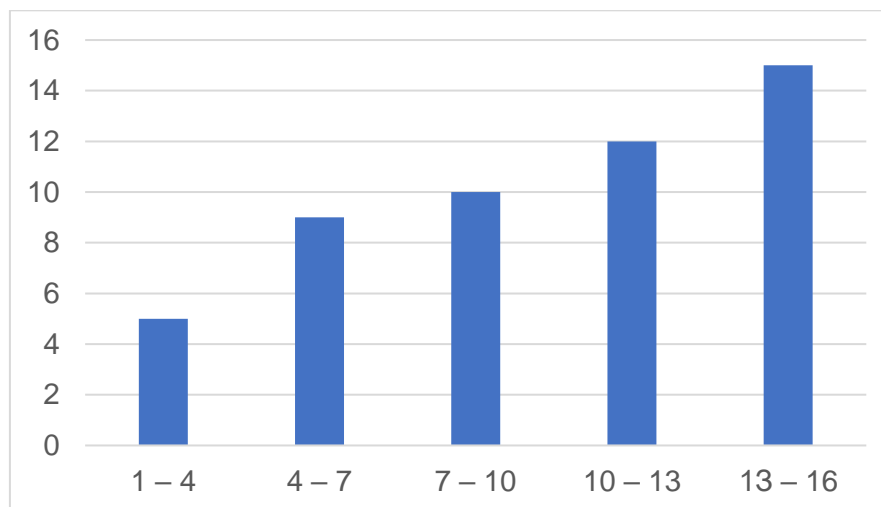


Рис. 3.5. Гістограма розподілу робітників за стажем

Утворені прямокутники є пропорційними за висотою частоті значень ознаки для кожного інтервалу.

3. При будуванні 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. 3.6).

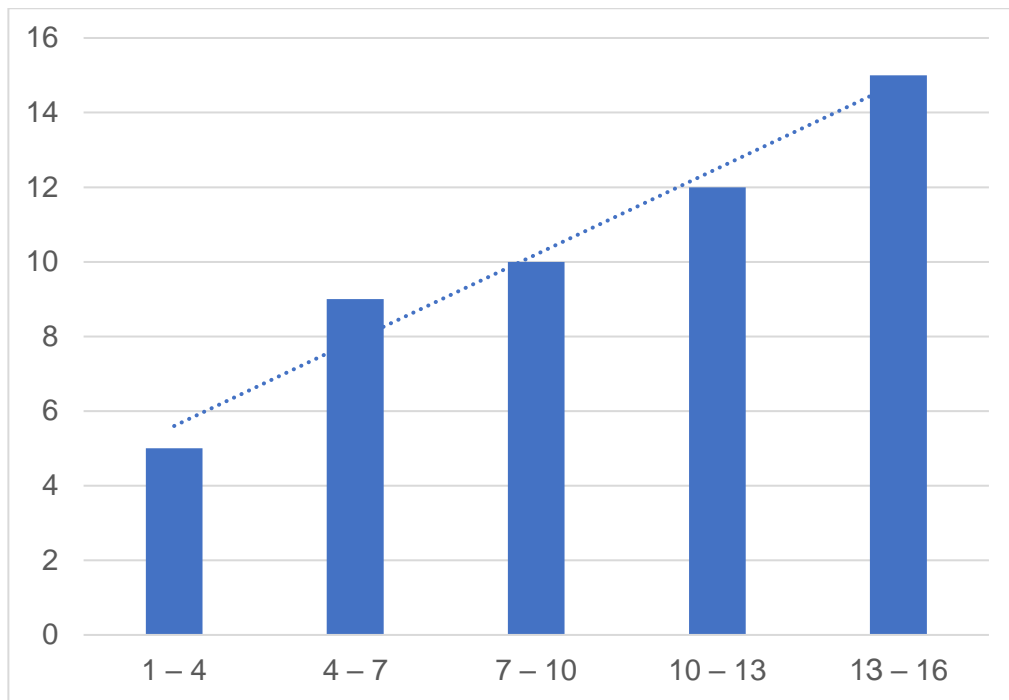


Рис. 3.6. Кумулятивна гістограма розподілу робітників за стажем

### Задача 3.3

За даними обсягів виробництва деякої продукції чотирма підприємствами регіону за кілька років (табл. 3.3) побудувати зображення структури явищ і структурних зрушень у вигляді стовпцевої, стрічкової й секторної діаграм.

Таблиця 3.3

Рік	Обсяг продукції, тис. т				
	Підприємство 1	Підприємство 2	Підприємство 3	Підприємство 4	Загальний обсяг
1997	1,9	1,7	0,7	0,4	4,7
2007	3,3	3,3	1,0	1,1	8,7
2017	6,7	5,1	0,8	2,4	15,0

### Розв'язання

1. Для характеристики та ілюстрації обсягу й структури виробництва продукції в регіоні побудуємо стовпцеву діаграму (рис. 3.7). На осі абсцис на однаковій відстані побудуємо стовпці, розбивши їх на частини відповідно до обсягу виробництва.

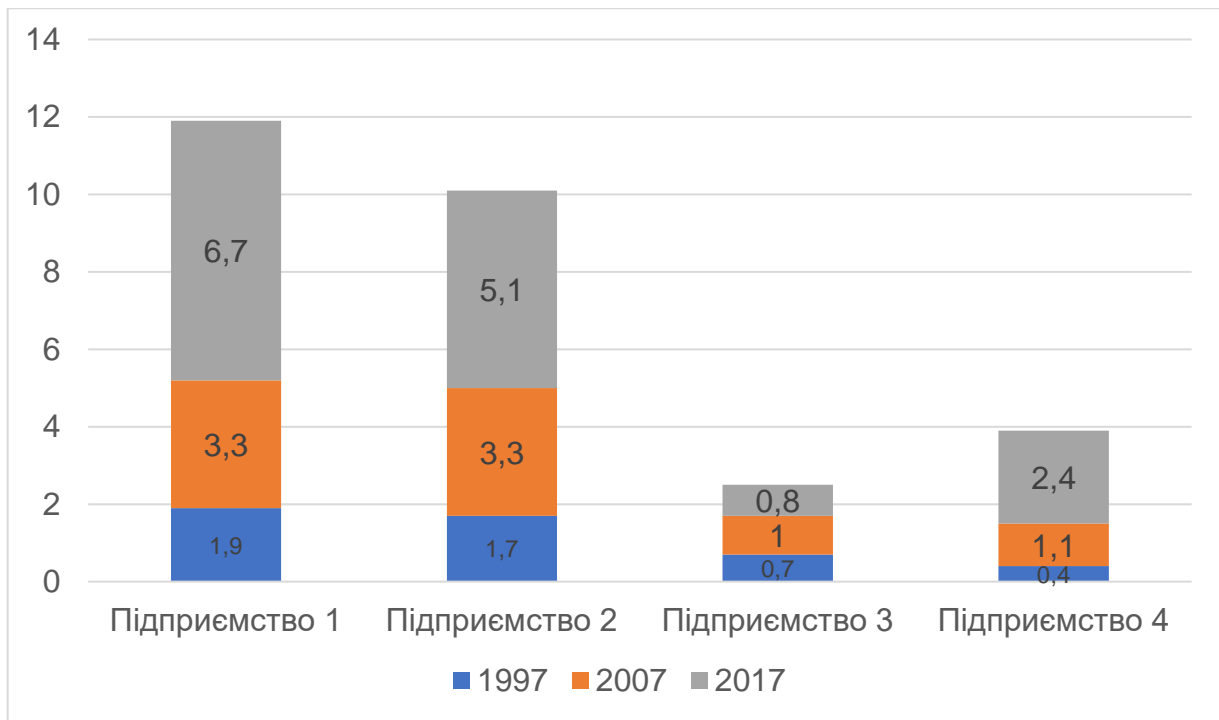


Рис. 3.7. Стовпцева діаграма обсягу й структури виробництва продукції в регіоні

2. Аналогічно будуємо стрічкову діаграму (рис. 3.8). Статистичні показники відображаємо перпендикулярно до осі ординат.

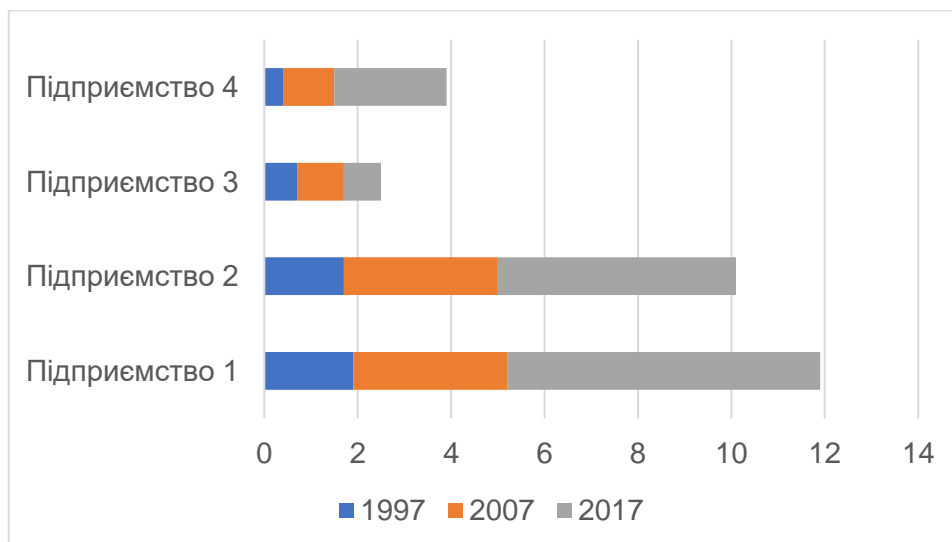


Рис. 3.8. Стрічкова діаграма обсягу й структури виробництва продукції в регіоні

3. Для будування секторної діаграми, у якій ураховано лише питому вагу частин виробництва, його загальний розмір беруть за 100 %, далі розраховують частки окремих частин у відсотках. Круг поділяють на сектори пропорційно до частин зображуваного цілого (рис. 3.9).



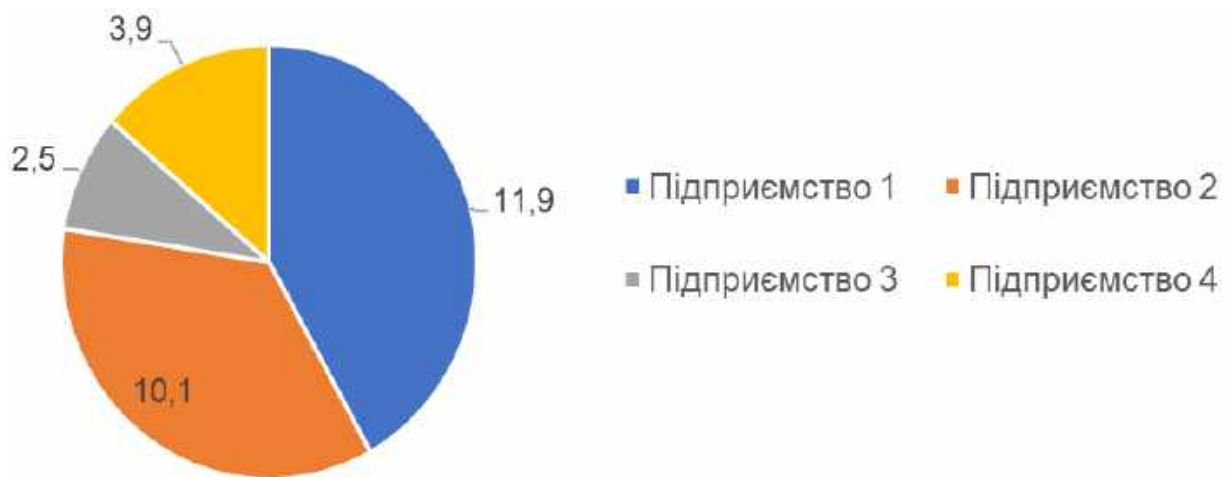


Рис. 3.9. Секторна діаграма

### Завдання для самостійної роботи

1. Побудувати макет статистичної таблиці для розміщення результатів комбінованого групування студентів своєї групи за рівнем успішності (середнього балу) і кількістю пропущених занять без поважних причин. Сформууйте по три групи за кожною ознакою.

За результатами групування (див. табл. 2.5) побудувати графіки (полігон і кумуляту).

2. За наведеними даними скласти статистичні таблиці, озаглавити кожную з них, визначити підмет і присудок, а також вид наведеного в таблиці угруповання за аналітичною функцією:

а) за рік у регіоні введено в дію основних фондів на суму 26,5 млн грн, у тому числі на об'єктах виробничого призначення – 18,2 млн грн; невиробничого – 8,3 млн грн. Обсяг капіталовкладень за той же період загалом по регіону становив 28,6 млн грн, з них на об'єктах виробничого призначення – 20,4 млн грн, невиробничого – 8,2 млн грн;

б) сума кредитів, наданих комерційними банками міста, на початок року становила 600 млн грн, на кінець року – відповідно 1080, 800 і 280 млн грн. За рік загальний обсяг банківських послуг з надання кредитів збільшився в 1,8 раза, зокрема на ринку короткострокових кредитів – у два рази, на ринку довгострокових кредитів – в 1,4 раза;

в) домогосподарства, бюджети яких досліджують, поділили на три групи: з високим, середнім і низьким доходом. Частота цих груп у загальному обсязі покупок відповідно становить 18, 52 і 30 % а середній бал якості придбаних товарів – 2,8; 2,2 і 1,4.

3. Динаміка капітальних вкладень характеризується такими даними в порівнянних цінах, млрд грн:

Показник	Рік				
	1-й	2-й	3-й	4-й	5-й
Капітальні вкладення усього	136,95	112,05	84,66	74,7	62,3
у тому числі: виробничого призначення	97,35	79,65	60,18	53,10	41,40
невиробничого призначення	39,6	32,4	24,48	21,6	20,9

Побудувати зображення структури явищ і структурних зрушень у вигляді стовпцевої, стрічкової й секторної діаграм.

4. Є такі дані про розподіл 143 господарств щодо врожайності зернових культур:

Урожайність, ц/га	10–15	15–20	20–25	25–30	30–35	35–40	40–45	45–50	50–55
Кількість господарств	6	9	20	41	26	21	14	5	1

За результатами групування побудувати графіки (полігон і кумуляту).

### Контрольні запитання

1. Що таке статистичні графіки? Якими є вимоги до їх будування?
2. Для чого призначено графіки?
3. Якими є основні елементи графіків?
4. Що відображує шкала графіка? Назвіть основні її види.
5. Назвіть основні види графіків.
6. Як будують стовпцеві й стрічкові діаграми?
7. Як будують квадратні й кругові діаграми?
8. Які явища характеризують секторні діаграми? Як їх будують?
9. Що таке прямокутні діаграми?
10. У чому полягає перевага лінійних діаграм, для чого їх використовують і якими є правила їх будування?

## 4. УЗАГАЛЬНЮВАЛЬНІ СТАТИСТИЧНІ ПОКАЗНИКИ

### 4.1. Суть і класифікація статистичних показників

**Статистичний показник** – узагальнювальна характеристика соціально-економічного явища або процесу, у якій об'єднуються якісна й кількісна визначеність явища або процесу.

**Якісний** зміст показника залежить від суті явища (процесу) і відображається в назві (наприклад, народжуваність, прибутковість тощо).

**Кількісний** зміст показника відображають число і його вимірник.

Статистичні показники можна поділити на два основних види:

– обліково-оцінні показники (розміри, обсяги, рівні явища);

– аналітичні показники (відносні й середні величини, показники варіації тощо).

Показники розрізняють за способом обчислення, ознакою часу (табл. 4.1) й аналітичними функціями.

Таблиця 4.1

За якісною ознакою	За кількісною ознакою	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** визначають шляхом зведення даних статистичного спостереження й подають у вигляді абсолютних величин (наприклад, кількість і сума внесків в Ощадбанку).

**Вторинні показники** обчислюють на базі первинних 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, кількість мешканців і т. д.).

**Інтервальні показники** – підсумковий нагромаджений результат за період у 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 тривалості періоду, за який їх обчислюють. Особливістю первинних інтервальних показників є адитивність – можливість підсумовування. Вторинні показники в основному є неадитивними.

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

## 4.2. Абсолютні величини

Результати статистичного спостереження реєструються насамперед у формі первинних абсолютних величин.

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

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

– **індивідуальні абсолютні величини**, які характеризують кількісні ознаки окремих одиниць (рівень виробітку окремого робітника за конкретний період);

– **групові й загальні абсолютні величини**, які виражають значення тієї чи іншої ознаки всіх одиниць певної сукупності, разом узятих, або 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 змогу знайти показник абсолютного розміру площі району).

У статистиці всі абсолютні величини є іменованими, вимірюються в конкретних одиницях і можуть бути як додатними, так і від'ємними (збитки, спад, утрати і т. ін.).

### 4.3. Види і взаємозв'язки відносних величин

**Відносна величина** – це узагальнювальний показник, який є вираженням кількісних співвідношень між соціально-економічними явищами або процесами. Оскільки багато абсолютних величин явищ є взаємозв'язаними, то й відносні величини одного типу в деяких випадках можна визначати через відносні величини іншого типу.

Загальні принципи розрахунку відносної величини:

1) між порівнюваними у відносній величині показниками мають бути реальні зв'язки (наприклад, при обчисленні відносного показника, що характеризує писемність населення, зі всього населення необхідно виключити дітей до 6 років та інвалідів;

2) порівнювані початкові показники можуть різнитися 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 відображає, у скільки разів порівнювана величина перевищує базисну або яку частину перша становить від другої, іноді – скільки одиниць однієї величини припадає на 100, 1000 і далі одиниць іншої – величини, яка є базою порівняння.

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

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

### **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, що спостерігається в певному періоді, до запланованої (середньої, максимальної або мінімальної) величини ознаки.

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

На практиці розрізняють два методи знаходження відносних показників виконання плану:

1) порівняння фактичних і планових рівнів;

2) установлення абсолютної величини збільшення або зменшення показника, і, відповідно, перевірка ступеня виконання плану за цією величиною (у плановому завданні).

**Відносна величина динаміки**  $i$  характеризує динаміку процесу, напрям та інтенсивність змінення явища в часі; її визначають шляхом співвідношення значення ознаки в певний (поточний) період або момент часу до значення ознаки в попередній (базисний) період або момент часу.

Якщо значення показника зменшується, то відносна величина буде меншою за одиницю.

**Відносні величини структури**  $i$  характеризують склад, структуру, частку (відношення частини до цілого), питому вагу складових елементів сукупності за тією чи іншою ознакою. Зазвичай їх розраховують у формі відсотків:

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

Відносні величини структури називають частками, які є адитивними, причому сума відносних величин структури дорівнює одиниці або 100 %.

Сукупність відносних величин структури відображає будову явища, що вивчається. Скільки є складових величин, стільки ж є і відносних величин структури.

Різницю між відповідними частинами (частками) називають відсотковим пунктом (в. п.).

З допомогою відносних величин структури можна оцінити **структурні зрушення** – зміни в складі сукупності за певний період часу. Оцінювання 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:

$$i = \frac{\text{величина } j\text{-ї частки сукупності}}{\text{величина } k\text{-ї частки сукупності}} \cdot 100 \% = \frac{d_j}{d_k} \cdot 100 \%$$

Відносні величини координації є відображенням того, 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 і далі одиниць іншої частини, узяті за базу порівняння.

**Відносні величини порівняння** – це співвідношення однойменних абсолютних величин, що належать до одного й 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 і завжди є відношенням двох різнойменних величин. У чисельнику – величина явища (показник), ступінь поширення якого вивчають, а в знаменнику – величина того середовища, у якому розвивається (поширюється) це явище. У кожному конкретному випадку таке співвідношення характеризує інтенсивність поширення, розвиток явища в певному середовищі. Відносні величини інтенсивності називають одиницями вимірювання чисельника і знаменника співвідношення.

## Приклади розв'язання типових задач

**Задача 4.1.** Розрахунок відносних величин виконання плану, планового завдання та динаміки

2020 року кількість персоналу підприємства становила 120 осіб. 2021 року планувалося зменшення виробництва й доведення кількості працівників до 100 осіб. Однак їх кількість за 2021 рік збільшилася до 130 осіб. Провести аналіз руху кількості співробітників підприємства.

### *Розв'язання*

Маємо:  $x_0 = 120$ ;  $x_1 = 130$ ;  $x_n = 100$ .

Відносна величина планового завдання (відповідає на запитання «Що планується порівняно з тим, що було?»):

$$i = \frac{x_n}{x_0} = \frac{100}{120} = 0,83, \text{ або } 83 \%$$

Відносна величина виконання плану (відповідає на запитання «Чи було виконано те, що планувалося?»):

$$i = \frac{x_1}{x_n} = \frac{130}{100} = 1,3, \text{ або } 130 \%$$

**Висновок:** підприємство планувало зменшення кількості персоналу на 17 % ( $83 - 100 = -17$ ), але фактична кількість працівників перевищила запланований рівень на 30 % ( $130 - 100 = 30$ ).

Відносні величини динаміки:

$$i_d = i_{\text{пл.з}} \cdot i_{\text{вик.пл}} = 0,83 \cdot 1,3 = 1,08, \text{ або } 108 \%$$

$$i_d = 130/120 = 1,08, \text{ або } 108 \%$$

**Висновок:** кількість персоналу підприємства 2021 року збільшилася на 8 %.

### **Задача 4.2**

Розрахунок відносних величин виконання плану, планового завдання та динаміки

Страхова компанія 2020 року уклала договорів на суму 500 тис. грн. 2021 року вона планувала укласти договори на суму 510 тис. грн, однак під впливом різних факторів фактично уклала на суму 400 тис. грн.

Провести аналіз діяльності підприємства.

### *Розв'язання*

Маємо:  $x_1 = 400$  тис. грн,  $x_0 = 500$  тис. грн,  $x_n = 510$  тис. грн.

Відносна величина планового завдання (відповідає на запитання «Що планується порівняно з тим, що було?»):

$$i = \frac{x_n}{x_0} = \frac{510}{500} = 1,02, \text{ або } 120 \%$$

Відносна величина виконання плану (відповідає на запитання «Чи було виконано те, що планувалося?»):

$$i = \frac{x_1}{x_n} = \frac{400}{510} = 0,784, \text{ або } 78,4\%$$

**Висновок:** страхова компанія планувала збільшити суму договорів на 2 % (102 – 100 = 2 %), але фактично сума страхових договорів зменшилася на 21,6 % (78,4 – 100 = –21,6 %).

Відносні величини динаміки:

$$i_d = i_{пл.з} \cdot i_{вик.пл} = 1,02 \cdot 0,784 = 0,8, \text{ або } 80 \%$$

**Висновок:** сума страхових договорів 2021 року зменшилася на 20 % (80 – 100 = –20).

**Задача 4.3.** Розрахунок відносних величин виконання плану, планового завдання та динаміки

У плані на 2021 р. передбачено підвищення продуктивності праці робітників заводу на 5 %. Фактично у звітному періоді збільшення відбулося на 8 % порівняно з 2020 р. Визначити відносну величину виконання плану, яка характеризує підвищення продуктивності праці робітників заводу.

**Розв'язання**

Насамперед задане в плані й фактичне підвищення продуктивності праці у звітному періоді необхідно виразити у відсотках, включаючи базу порівняння. Припустимо, що базисний рівень продуктивності праці 2020 року – 100 %. Плановий відносний рівень продуктивності 2021 року буде становити 100 + 5 = 105 % (1,05), а її фактичне підвищення – 100 + 8 = 108 % (1,08). Відносна величина виконання плану підвищення продуктивності праці обчислюється так:

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

**Висновок:** план підвищення продуктивності праці робітників заводу перевиконано на 2,9 %.

**Задача 4.4.** Розрахунок відносних величин динаміки, структури та координації

Дані про віковий склад населення регіону, тис. чол., наведено в табл. 4.2.

Обчислити відносні величини, які б характеризували:

- 1) динаміку кількості населення;
- 2) його структуру за віком у кожному році та структурні зрушення;
- 3) співвідношення працездатного населення з кількістю допрацездатного віку.

Таблиця 4.2

Вікова група	2018 рік	2020 рік
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.

Динаміка змінення кількості всього населення регіону (коефіцієнт зростання):

$$962 : 740 = 1,3.$$

Темп зростання:

$$1,3 \cdot 100 \% = 130 \%$$

Темп приросту:

$$130 - 100 = 30 \%$$

**Висновок:** загальна кількість населення регіону 2020 року збільшилася порівняно з 2018 р. у 1,3 раза, або на 30 %.

За окремими віковими групами відносні величини динаміки наведено в табл. 4.3.

Таблиця 4.3

Вікова група	Відносні величини					
	динаміки (темпи зростання)	структури, %		структурних зрушень, в.п.	координації	
		2018	2020		2018	2020
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 й цілого. Частка населення допрацездатного віку 2018 р. становить

$$192,4 : 740 = 0,26, \text{ або } 26 \%$$

Частка населення допрацездатного віку 2020 року:

$$211,6 : 962 = 0,22, \text{ або } 22 \%$$

**Висновок:** частка населення допрацездатного віку 2018 року становить 26 %, а 2020 року – 22 %.

Структурні зрушення:

$$22 \% - 26 \% = -4 \text{ в. п. (відсоткові пункти)}$$

**Висновок:** частка населення допрацездатного віку зменшилася на чотири відсоткові пункти.

Результати розрахунків за іншими віковими групами наведено в табл. 4.4.

3. Співвідношення кількості населення окремих вікових груп називають відносною величиною координації.

Розрахунки для 2018 року:

$$469,9 / 192,4 = 2,44;$$

$$192,4 / 469,9 = 0,409, \text{ або } 192,4 / 469,9 \cdot 100 = 40,9 \text{ чол.}$$

*Висновок:* 2018 року людей працездатного віку було в 2,44 раза більше, ніж допрацездатного віку, або на 100 чол. працездатного віку припадало 40,9 чол. допрацездатного віку.

Розрахунки для 2020 року:

$$586,8 / 211,6 = 2,77;$$

$$211,6 / 586,8 = 0,36, \text{ або } 211,6 / 586,8 \cdot 100 = 36 \text{ чол.}$$

*Висновок:* 2020 року людей працездатного віку було в 2,77 раза більше, ніж непрацездатного віку, або на 100 чол. працездатного віку припадало 36 чол. допрацездатного віку.

**Задача 4.5.** Розрахунок відносних величин інтенсивності й порівняння

За даними про територію, кількість населення і валовий внутрішній продукт двох країн (табл. 4.4) обчислити відносні величини:

- 1) інтенсивності;
- 2) порівняння.

Таблиця 4.4

Країна	Територія, тис. км <sup>2</sup>	Кількість 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:

$$i = \frac{\text{Кількість населення, тис. чол.}}{\text{Територія, тис. км}^2} = \text{Густота населення, чол./км}^2,$$

$$i = \frac{\text{ВВП, млн ум. од.}}{\text{Кількість населення, тис. чол.}} = \text{ВВП на душу населення, ум. од.}$$

Результати розрахунків наведено в табл. 4.5.

Таблиця 4.5

Країна	Густота населення, кількість чол./км <sup>2</sup>	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. Вибір бази порівняння є довільним, але зручніше для висновків ділити більше число на менше:

$$18 / 14,6 = 1,2, \text{ або } 14,6 / 18 = 0,81;$$

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

*Висновки:*

1. У країні А порівняно з країною В густота населення є більшою в 1,2 раза, або в країні В порівняно з країною А густота населення є меншою в 0,81 раза.

2. Виробництво ВВП на душу населення в країні А порівняно з країною В є більшим у 2,2 раза.

**Задача 4.6.** Розрахунок відносної величини інтенсивності

Кількість підприємств роздрібної торгівлі регіону на кінець року становила 6324. Кількість населення цього регіону на ту саму дату – 234,2 тис. чол. Обчислити, скільки підприємств припадає на 10 000 чоловік.

*Розв'язання*

Маємо відносну величину інтенсивності

$$i = \frac{6324 \cdot 10000}{234200} = 270.$$

*Висновок:* 270 підприємств припадає на 10 000 чол. регіону.

**Завдання для самостійної роботи**

1. Для питань анкети визначити тип ознаки й можливі варіанти її кількісного виміру:

1) вік;

2) освіта;

3) стаж роботи;

4) задоволення від вибору професії (задоволений, не зовсім задоволений, незадоволений);

5) задоволення відносинами, що складаються на роботі (задоволений, незадоволений).

**2.** Здійснити класифікацію статистичних показників на первинні й вторинні, інтервальні й моментні, адитивні й неадитивні. Для показників, що мають форму відносних величин, визначити їх вид і форму відображення:

- 1) обсяг платних послуг населенню Києва за рік, млн грн;
- 2) кількість людей з вищою освітою на 1 000 чол. населення на початок року;
- 3) уведення в дію корисної площі житлових будинків за рік, млн м<sup>2</sup>;
- 4) частка робітників, зайнятих ручною працею, %;
- 5) виробництво електроенергії на душу населення протягом року, кВт·ч;
- 6) довжина електрифікованих ліній залізниць наприкінці року, тис. км;
- 7) частка молоді, що працює (віком до 30 років), у загальній кількості працівників науки на початок року;
- 8) кількість зареєстрованих протягом року шлюбів на 1 000 чол. населення.

**3.** Споживання палива тепловими електростанціями регіону становило:

Вид палива	Попередній рік	Поточний рік	Коефіцієнт переведення в умовне паливо
Вугілля, млн т	8,2	22,6	0,90
Газ природний, млн м <sup>3</sup>	11,4	21,6	1,20

Визначити обсяги спожитого умовного палива за кожний рік, відносні величини динаміки й структури. Зробити висновки про структурні зрушення.

**4.** Плановий і фактичний обсяги реалізації продукції трьох ліспромгоспів характеризуються такими даними:

Номер ліспромгоспу	Реалізація продукції, млн грн		
	Фактично попереднього року	Планове завдання на звітний рік	Фактично звітного року
1	22,1	22,4	23,5
2	32,4	32,5	35,7

Визначити для кожного підприємства окремо й загалом для сукупності господарств відносні величини планового завдання, виконання плану, динаміки. Показати взаємозв'язок між обчисленими величинами. Розв'язок оформити у вигляді таблиці. Зробити висновки.

**5.** За даними про територію, кількість населення й валовий внутрішній продукт двох країн обчислити такі відносні величини:

- 1) інтенсивність;
- 2) порівняння.

Країна	Територія, тис. км <sup>2</sup>	Кількість населення, тис. осіб	ВВП, млрд грн
А	912	16390	32483
В	1285	18710	16866

**6.** За наведеними даними обчислити відносні величини інтенсивності й порівняння:

Показник	Завод № 1	Завод № 2
Випуск продукції, тис. грн	432	510
Кількість працівників, чол.	380	414

**7.** За наведеними даними для кожної країни обчислити відносні величини, які характеризували б рівень економічного розвитку країни й ступінь її фінансової залежності. Зробити висновки.

Країна	Кількість населення, тис. чол.	Валовий внутрішній продукт, млрд дол. США	Зовнішній борг, млрд дол. США
А	23200	129,9	18,2
Б	15400	73,1	16,1
В	7800	30,6	8,6

**8.** З допомогою всіх можливих видів відносних величин проаналізувати наведені показники кількості професорсько-викладацького персоналу й студентів.

Показники	На початок року	
	2022	2023
Професорсько-викладацький персонал	114	120
Студенти	2100	2020
у тому числі:		
– очна форма навчання	1040	1120
– заочна форма навчання	1060	900

**9** Собівартість одиниці продукції на підприємстві у звітному періоді становила 550 грн, що на 150 грн більше, ніж попереднього року, і на 50 грн більше, ніж на аналогічному підприємстві. Чому дорівнює відносна величина порівняння?

**10.** Звітнього року прибуток становив 4 млн грн, попереднього – 3,7 млн грн. У плані було передбачено обсяг прибутку збільшити на 3,72 % порівняно з попереднім роком. Обчислити рівень виконання плану.

### Контрольні запитання

1. Що таке абсолютні статистичні величини? Яким є їх значення в статистиці?

2. Які види абсолютних величин можна виокремити 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 й координації?
10. Для характеристики яких явищ використовують відносні величини інтенсивності та порівняння в просторі?

## 5. СЕРЕДНІ ВЕЛИЧИНИ

### 5.1. Суть і види середньої величини

**Середня величина** – це узагальнювальна кількісна характеристика варьовальної ознаки в статистичній сукупності при конкретних умовах місця й часу, яка характеризує її рівень з розрахунку на одиницю сукупності.

Середня величина відображає те загальне, що є властивим для всіх одиниць досліджуваної сукупності. Однак у ній не враховано індивідуальні відмінності окремих одиниць сукупності, зумовлені 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 основі 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, що має назву визначального, пов'язаний так чи інакше з усереднюваним показником.

**Види середніх величин:**

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

2) **структурні**, до яких належать мода, медіана, квартиль, дециль.

## 5.2. Види степеневих середніх величин і способи їх обчислення

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

$$\bar{x}_k = \sqrt[k]{\frac{\sum_{i=1}^n x_i^k}{n}} = \left( \frac{\sum_{i=1}^n x_i^k}{n} \right)^{\frac{1}{k}}, \quad (5.1)$$

де  $x_i$  – рівень (варіанта) усереднюваної ознаки;

$k$  – показник степеня середньої величини;

$n$  – кількість варіант.

**Зважена середня величина** обчислюється за згрупованими даними і має загальний вигляд

$$\bar{x}_k = \sqrt[k]{\frac{\sum_{i=1}^m x_i^k f_i}{\sum_{i=1}^m f_i}} = \left( \frac{\sum_{i=1}^m x_i^k f_i}{\sum_{i=1}^m f_i} \right)^{\frac{1}{k}}, \quad (5.2)$$

де  $x_i$  – рівень (варіанта) усереднюваної ознаки або середнє значення інтервалу, у якому вимірюється варіанта;

$k$  – показник степеня середньої величини;

$n$  – кількість варіант;

$f_i$  – відповідні частоти (ваги) – кількість одиниць сукупності в різних групах, інтервалах, що є відображенням того, скільки разів трапляється  $i$ -те значення усереднюваної ознаки, причому  $\sum_{i=1}^m f_i = n$ .

**Гармонійна середня величина** ( $k = -1$ ) застосовується при розрахунку середніх обернених показників  $\frac{1}{x_1}, \dots, \frac{1}{x_n}$ . Необхідно, щоб при усередненні незмінною залишалася сума величин, обернених до індивідуальних значень ознаки.

### **Гармонійна середня проста величина**

$$\bar{x} = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}}. \quad (5.3)$$

### **Гармонійна середня зважена величина**

$$\bar{x}_{-1} = \frac{\sum_{i=1}^m z_i}{\sum_{i=1}^m \frac{z_i}{x_i}}, z_i = x_i f_i \quad (5.4)$$

Зважену гармонійну застосовують тоді, коли показника, що є вагою  $f$ , немає і його слід додатково визначити на основі відомих варіант  $x$  і добутку варіант на частоту  $f$  ( $z = xf$ ).

**Геометрична середня величина** ( $k = 0$ ) 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.

### **Геометрична середня проста величина**

$$\bar{x} = \sqrt[n]{\prod x_i} = \sqrt[n]{x_1 x_2 \dots x_n} \quad (5.5)$$

### **Геометрична середня зважена величина**

$$\bar{x}_0 = \sqrt[n]{\prod_{i=1}^m x_i^{f_i}}, n = \sum_{i=1}^m f_i \quad (5.6)$$

Якщо часові інтервали є неоднаковими, то використовують зважену середню величину.

**Арифметична середня величина** ( $k = 1$ ) використовується у випадках, коли обсяг варіювальної ознаки для всієї сукупності формується як сума індивідуальних значень її окремих одиниць. Для ряду розподілу використовують зважену середню величину.

### **Арифметична середня проста величина**

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \quad (5.7)$$

### **Арифметична середня зважена величина**

$$\bar{x} = \frac{\sum_{i=1}^m x_i f_i}{\sum_{i=1}^m f_i}, \sum_{i=1}^m f_i = n \quad (5.8)$$



**Квадратична середня величина** ( $k = 2$ ) застосовується, якщо при заміні індивідуальних величин ознаки на середню необхідно зберегти незмінною суму квадратів початкових величин.

**Квадратична середня проста величина**

$$\bar{x} = \sqrt{\frac{\sum_{i=1}^n x_i^2}{n}}. \quad (5.9)$$

**Квадратична середня зважена величина**

$$\bar{x}_2 = \sqrt{\frac{\sum_{i=1}^m x_i^2 f_i}{n}}. \quad (5.10)$$

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

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

Використання кожного виду середніх величин залежить від характеру індивідуальних значень ознаки (прямі, обернені, квадратичні, відносні) і від характеру алгебричного зв'язку між індивідуальними значеннями ознаки й загальним обсягом (сума, добуток, степінь, квадратичний корінь). Цей зв'язок є визначальною властивістю сукупності й відображається в логічній формулі усереднюваної ознаки (його економічному змісті). На основі логічної формули вибирають вид середньої величини. Формально цей принцип можна записати у вигляді табл. 5.1.

Таблиця 5.1

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

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

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

Припустимо, що всі значення сукупності поділено на  $k$  груп. Розглядаючи кожен групу як самостійну сукупність, можна знайти її **загальну середню величину** як середнє арифметичне **групових середніх величин**:

$$\bar{x} = \frac{\sum_{i=1}^k x_i f_i}{\sum_{i=1}^k f_i}, \quad (5.11)$$

де  $f_i$  – частота  $i$ -ї групи,  $\sum_{i=1}^k f_i = n$ ;

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

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

де  $\sum_{j=1}^l m_j = f_i$ .

*Особливості обчислення середньої величини для інтервального ряду розподілу:*

1. Середина кожного інтервалу (варіанта) обчислюється як півсума нижньої й верхньої меж інтервалу.

2. Якщо ряд розподілу має відкритий інтервал (перший або останній), то його розмір умовно береться таким, що дорівнює розміру сусіднього закритого інтервалу.

3. Середня величина обчислюється як середня зважена.

*Обчислення нормованого середнього бала для значень описових порядкових ознак:*

1. Значення ознаки ранжують у порядку збільшення значення якісної ознаки. Наприклад, найнижчий ранг присвоюють відповідям «ні», найвищий – відповідям «так». Варіанти ознаки оцифровують порядковими рангами  $x_i = 1, 2, \dots, n$  або центрованими  $x_i = -2, -1, 0, 1, 2$ .

2. Нормований середній бал обчислюють за формулою

$$\bar{K} = \frac{\bar{x} - \frac{R}{2}}{x'}, \quad (5.12)$$

де  $\bar{x}$  – середньозважений ранг,  $\bar{x} = \frac{\sum x f}{\sum f}$ ;

$x$  – ранг ознаки;

$f$  – частка значення якісної ознаки, причому  $\sum f = 1$ , або 100 %;

$R$  – розмах шкали рангів,  $R = x_{\max} - x_{\min}$ , де  $x_{\max}$ ,  $x_{\min}$  – максимальне й мінімальне значення рангів;

$x'$  – середина шкали рангів  $x' = \frac{x_{\max} + x_{\min}}{2}$ .

### 5.3. Структурні середні величини

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

**Мода** ( $M_o$ ) – величина, яка найчастіше трапляється в сукупності, або варіанта, що найчастіше повторюється в ряді.

Трапляються ряди, які мають дві моди (бімодальний ряд) або декілька мод (полімодальний ряд).

**Медіана** ( $M_e$ ) – варіанта, що поділяє ранжований ряд на дві однакові за кількістю одиниць сукупності частини, унаслідок чого значення ознаки однієї її половини не перевищує медіанного рівня, а іншої – не менше за нього.

#### **Особливості обчислення моди й медіани:**

1. Для *дискретного* варіаційного ряду:

– мода – це варіанта з максимальною частотою;

– медіана при парній кількості одиниць сукупності – це арифметична середня величина з двох центральних варіант, при непарній – це центральна варіанта, розташована в центрі ряду.

2. Для *інтервального* ряду розподілу:

– моду розраховують тільки для однакових інтервалів, оскільки від цього залежить показник повторюваності значень ознаки  $X$ ;

– для обчислення моди визначають модальний інтервал, тобто інтервал зі значенням ознаки, що найчастіше повторюється:

$$M_o = x_{M_o} + h \frac{f_{M_o} - f_{M_o-1}}{(f_{M_o} - f_{M_o+1}) + (f_{M_o} - f_{M_o-1})}, \quad (5.13)$$

де  $x_{M_o}$  і  $h$  – відповідно нижня межа й ширина модального інтервалу;

$f_{M_o}$  – частоти модального інтервалу;

$f_{M_o-1}$ ,  $f_{M_o+1}$  – частоти попереднього й наступного інтервалів відносно модального.

Для обчислення медіани визначають медіанний інтервал – такий, кумулятивна частота (сума нагромаджених частот, що передують медіанному інтервалу) якого

дорівнює або перевищує половину суми частот  $0,5 \sum_{j=1}^m f_j$ . З допомогою інтерполяції

в цьому медіанному інтервалі знаходять значення медіани:

$$M_e = x_{M_e} + h \frac{0,5 \sum_{j=1}^m f_j - S_{M_e-1}}{f_{M_e}}, \quad (5.14)$$

де  $x_{M_e}$  і  $h$  – відповідно нижня межа й ширина медіанного інтервалу;

$f_{M_e}$  – частота медіанного інтервалу;

$S_{M_e-1}$  – кумулятивна частота передмедіанного інтервалу.

**Квартилі** – це варіанти, які поділяють обсяг сукупності на чотири однакові частини, **децилі** – на десять рівних частин, **процентилі** – на 100. Ці характеристики визначають на основі кумулятивних частот, як і медіану, яка є другим квартилем або п'ятим децилем.

### Приклади розв'язання типових задач

**Задача 5.1.** Середні арифметичні проста та зважена величини для дискретних рядів

Розрахувати середній вік студентів у групі з 10 чоловік, дані про яких наведено в табл. 5.2.

Таблиця 5.2

№ 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

Скористаємося логічною формулою визначення середнього віку на основі змісту показника:

*Середній вік = Сумарний вік усіх студентів / Кількість студентів.*

Середній вік за формулою (5.7) простої середньої арифметичної величини:

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

Для використання середньої арифметичної зваженої величини за формулою (5.8) згрупуємо початкові дані. Отримаємо дискретний ряд розподілу (табл. 5.3).

Таблиця 5.3

Вік $X_i$	18	19	20	21	22
Кількість студентів $f_i$	1	4	3	1	1

Після групування одержимо новий показник – частоту  $f_i$ , що відображає кількість студентів віком  $X_i$  років. Отже, середній вік студентів групи розраховуватиметься за формулою

$$\bar{x} = \frac{18 \cdot 1 + 19 \cdot 4 + 20 \cdot 3 + 21 \cdot 1 + 22 \cdot 1}{10} = \frac{197}{10} = 19,7 \text{ року.}$$

**Висновок:** середній вік студентів групи становить 19,7 року.

**Задача 5.2.** Середня арифметична зважена величина для дискретних рядів  
За даними розподілу працівників цеху залежно від їхньої зарплати  
(табл. 5.4), визначити середню зарплату працівника цеху.

Таблиця 5.4

Місячна зарплата $x_i$ , грош. од.	Кількість працівників $f_i$ , чол.
300	2
330	4
360	8
390	20
420	16

*Розв'язання*

Логічна формула визначення середньої зарплати на основі економічного змісту показника:

$$\text{Середня зарплата робітника} = \frac{\text{Фонд заробітної плати}}{\text{Кількість працівників}}$$

Середня зарплата одного працівника цеху за формулою середньої арифметичної зваженої величини, грош. од.,

$$\bar{x} = \frac{\sum_{i=1}^5 x_i f_i}{\sum_{i=1}^5 f_i} = \frac{300 \cdot 2 + 330 \cdot 4 + 360 \cdot 8 + 390 \cdot 20 + 420 \cdot 16}{2 + 4 + 8 + 20 + 16} = \frac{19320}{50} = 386,40.$$

**Висновок:** середня зарплата одного працівника цеху становить 386,4 грош. од.

**Задача 5.3.** Середня гармонійна зважена величина

За даними табл. 5.5 визначити середню частку забракованої продукції в загальному обсязі перевіреної.

Таблиця 5.5

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. Числового значення знаменника немає, його необхідно розрахувати. 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

необхідно використати формулу середньої гармонійної зваженої (5.4):

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

**Висновок:** середня частка забракованої продукції в загальному обсязі перевіреної становить 7,4 %.

### Задача 5.4. Середня гармонійна зважена величина

Розрахувати середню ціну товарів, реалізованих у трьох містах області, за даними табл. 5.6.

Таблиця 5.6

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:

$$\text{Середня ціна} = \frac{\text{Сума реалізації}}{\text{Кількість реалізованих одиниць}}$$

Сума реалізації є відомою. Оскільки вагою  $f_i$  є невідома кількість реалізованих одиниць, застосуємо середню гармонійну зважену.

Щоб знайти кількість реалізованих одиниць, необхідно окремо поділити суму реалізації кожного виду товару на ціну:

$$\begin{aligned} & \text{Кількість реалізованих одиниць кожного товару } f_i = \\ & = \text{Сума реалізації кожного товару } z_i / \text{Ціна кожного товару } x_i. \end{aligned}$$

Тоді середня ціна товарів

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

*Висновок:* середня ціна товарів становить 24,3 грн.

### **Задача 5.5.** Геометрична середня проста величина

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

#### *Розв'язання*

Згідно з основною 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.

Використаємо для розрахунку середню арифметичну просту (формула (5.7)):

$$(1,8 + 1,2 + 1,25) / 3 = 1,416.$$

Отримане значення не забезпечує заданих вимог: за три роки за цією середньою величиною ціни збільшилися би так:

$$1,416 \cdot 1,416 \cdot 1,416 = 2,84,$$

а не в 2,7 раза.

Задану умову  $\prod_{i=1}^m x_i = 2,7$  може забезпечити лише геометрична середня проста величина за формулою (5.5):

$$\bar{x} = \sqrt[3]{1,8 \cdot 1,2 \cdot 1,25} = 1,394.$$

*Висновок:* середньорічний темп збільшення цін становить 139,4 %, тобто ціни в середньому збільшувалися на 39,4 %.

**Задача 5.6.** Середня квадратична величина

Є три квадратні ділянки земельної площі зі сторонами  $x_1 = 100$  м,  $x_2 = 200$  м,  $x_3 = 300$  м. Розрахувати середню довжину сторін ділянок.

*Розв'язання*

Перевіримо виконання основної вимоги до формули розрахунку середньої величини.

Спочатку використаємо середню арифметичну величину:

$$\bar{x}_1 = \frac{\sum x}{n} = \frac{100 + 200 + 300}{3} = 200 \text{ м,}$$

потім – середню квадратичну величину:

$$\bar{x}_2 = \sqrt{\frac{100^2 + 200^2 + 300^2}{3}} = 216 \text{ м.}$$

При заміні вихідних значень довжини сторін ділянок на середню величину необхідно виходити зі збереження загальної площі всіх ділянок:

$$S = 100^2 + 200^2 + 300^2 = 140\,000 \text{ м}^2.$$

Арифметична середня величина 200 м не задовольняє цю умову, тому що загальна площа  $S$  трьох ділянок зі стороною 200 м становила би  $3 \cdot 200^2 = 120\,000 \text{ м}^2$ .

Загальна площа трьох ділянок із середньою квадратичною величиною 216 м

$$S = 216^2 + 216^2 + 216^2 = 140\,000 \text{ м}^2,$$

що відповідає реальній загальній площі всіх ділянок.

*Висновок:* середня довжина сторін ділянок становить 216 м.

**Задача 5.7.** Обчислення нормованого середнього бала для значень описових порядкових ознак

Визначити середній бал задоволення населення якістю медичного обслуговування за наведеними в табл. 5.7 даними.

Таблиця 5.7

Чи задоволені ви якістю медичного обслуговування?	Кількість відповідей, %
Так, повністю	15
Частково	49
Ні	36
Разом	100

*Розв'язання*

Нормований середній бал обчислюємо для значень описових порядкових ознак. Ранжуємо значення ознаки в порядку збільшення



значення якісної ознаки. Найнижчий ранг «один» дамо відповідям «ні», найвищий «три» – відповідям «так».

Нормований середній бал  $\bar{K}$  обчислимо за формулою (5.12).

Середньозважений ранг:

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

Розмах шкали рангів

$$R = x_{\max} - x_{\min} = 3 - 1 = 2.$$

Середина шкали рангів

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

Отже,

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

*Висновок:* рівень задоволення якістю медичного обслуговування населення становить у середньому 39,5 %.

### Задача 5.8. Структурні середні

Розподіл населення країни за віком на початок 2022 р. наведено в табл. 5.8.

Таблица 5.8

Вік	0–14	15–24	25–44	45–64	65 і більше
Кількість населення, тис. осіб	6501,1	7103,1	13206,6	11874,8	7506,7

Визначити моду й медіану в цьому ряді розподілу.

*Розв'язання*

Маємо інтервальний ряд.

$$M_o = x_{M_o} + h \frac{f_{M_o} - f_{M_o-1}}{(f_{M_o} - f_{M_o+1}) + (f_{M_o} - f_{M_o+1})}.$$

Модальний інтервал – 25–44 років, оскільки 13 206,6 тис. осіб – найбільша частота.

Ширина модального інтервалу

$$h = 44 - 25 = 19.$$

Нижня межа модального інтервалу

$$X_{Mo} = 25.$$

Частота модального інтервалу

$$f_{Mo} = 13\,206,6.$$

Частота попереднього інтервалу відносно модального

$$f_{Mo-1} = 7\,103,1.$$

Частота наступного інтервалу відносно модального

$$f_{Mo+1} = 11\,874,8.$$

Мода

$$Mo = 25 + 19 \frac{13206,6 - 7103,1}{(13206,6 - 7103,1) + (13206,6 - 11874,8)} = 41 \text{ рік.}$$

Медіану обчислимо за формулою (5.14).

Медіанний інтервал становить 25–44 років, оскільки половина обсягу сукупності

$$23096,15 = 0,5 \cdot (6501,1 + 7103,1 + 13206,6 + 11874,8 + 7506,7)$$

не перевищує нагромаджену частоту цього інтервалу

$$S = 6501,1 + 7103,1 + 13206,6 = 26810,8.$$

Ширина медіанного інтервалу

$$h = 44 - 25 = 19.$$

Нижня межа медіанного інтервалу

$$X_{Me} = 25.$$

Кумулятивна частота передмедіанного інтервалу  $S_{Me-1} = 7103,1$ .

Медіана

$$Me = 25 + 19 \frac{23 - 96,15 - (6501,1 + 7103,1)}{13206,6} = 39 \text{ років.}$$

*Висновки:*

1. Середній вік більшості населення становить 41 рік.
2. Вік однієї половини населення є меншим від 39 років, другої – більшим.

### Задача 5.9. Структурні середні величини

За даними табл. 5.9 визначити середні значення собівартості виробу, моду й медіану.

Таблиця 5.9

Групи підприємств	Собівартість одного виробу, тис. грн	Кількість 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 більше	52	49	50,7
Разом	–	100	100	100

#### Розв'язання

Маємо інтервальний ряд. Обчислимо середину кожного інтервалу як напівсуму його нижньої й верхньої меж. Оскільки останній інтервал є відкритим, то його розмір умовно беремо таким, що дорівнює попередньому інтервалу. Отже, 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 (5.8):

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

Медіану обчислимо за формулою (5.14). У прикладі може бути отримано три медіанних значення згідно з ознакою: за кількістю підприємств, обсягом продукції і загальною сумою витрат на виробництво.

1. Медіанний інтервал становить 125 і більше, тоді

$$Me_1 = 125 + 5 \frac{50 - (8 + 16 + 24)}{52} = 125,19 \text{ тис. грн.}$$

Отже, рівень собівартості одиниці продукції половини підприємств перевищує 125,19 тис. грн.

2. Медіанний інтервал становить 120–125:

$$Me_2 = 120 + 5 \frac{50 - (9 + 18)}{24} = 124,79 \text{ тис. грн.}$$

Таким чином, половина всього обсягу продукції має рівень витрат на виріб понад 124,79 тис. грн.

3. Медіанний інтервал становить 125 і більше:

$$Me_3 = 125 + 5 \frac{50 - 49,3}{50,7} = 125,07 \text{ тис. грн.}$$

Отже, утворюється 50 % від загальної суми витрат, якщо рівень собівартості одного виробу перевищує 125,07 тис. грн.

Для цього прикладу можна розрахувати три модальних значення за формулою (5.13) виходячи з ознак кількості підприємств, обсягу продукції й суми витрат. У всіх трьох випадках модальний інтервал є одним і тим самим, оскільки для одного й того ж інтервалу виявляються найбільшими і кількість підприємств, і обсяг продукції, і загальна сума витрат на виробництво:

$$Mo_1 = 125 + 5 \frac{52 - 24}{(52 - 24) + (52 - 0)} = 125,75 \text{ тис. грн.};$$

$$Mo_2 = 125 + 5 \frac{49 - 24}{(52 - 24) + (52 - 0)} = 126,69 \text{ тис. грн.};$$

$$Mo_{13} = 125 + 5 \frac{50,7 - 23,9}{(50,7 - 23,9) + (50,7 - 0)} = 126,73 \text{ тис. грн.}$$

Таким чином, більшість підприємств мають рівень собівартості 126,75 тис. грн, найчастіше за все випускається продукція з рівнем витрат 126,69 тис. грн і витрати на виробництво більшості підприємств пояснюються рівнем собівартості в 123,73 тис. грн.

### Завдання для самостійної роботи

1. У двох продавців куплено товару на однакову суму – 1 грн, але за різною ціною: по 3 грн за 1 кг у першого продавця та по 2 грн – у другого. Якою є середня ціна товару?

2. Бригада токарів з п'яти осіб виготовляє за зміну один вид продукції. На виготовлення одиниці виробленої продукції кожен із них витрачає певний час:

Номер робітника	1	2	3	4	5
Витрати часу на виготовлення однієї деталі, хв	12	15	14	16	14

Розрахувати середні витрати часу, необхідні для виготовлення однієї деталі.

3. За наведеними даними визначити середній відсоток виконання договірних зобов'язань постачальниками комплектувальних виробів. Обґрунтуйте вибір середньої величини договірних зобов'язань.

Виріб	Виконано зобов'язань на суму, млн грн	Виконання договірних зобов'язань, %
А	31,2	97,5
Б	17,1	95,0

4. На ринках міста зафіксовано певні ціни на сільськогосподарські продукти, грн:

Товар	Ціна за товар				
	1	2	3	4	5
Яловичина (за 1 кг)	50,0	45,0	40,0	50,0	55,0
Молоко (за 1 л)	8,0	10,0	9,0	7,0	6,0
Яйця (за 10 шт.)	15,0	16,0	14,0	13,0	17,0

Визначити максимальну, мінімальну й середню ціни кожного продукту.

5. Є дані про заробітну плату й кількість працівників філій АТ.

Визначити показники середньої заробітної плати по кожній філії та загалом по підприємству. Показати динаміку середньої заробітної плати.

Філіал	Місячний фонд заробітної плати, тис. грн		Середня облікова кількість працівників, чол	
	Жовтень	Листопад	Жовтень	Листопад
1	400	260	150	140
2	720	740	220	335
Разом	14700	1610	475	475

6. Є дані щодо двох заводів, які виробляють однорідну продукцію.

Обчислити середні витрати часу на виготовлення одиниці продукції по двох заводах з 2022 р. по 2023 р. Обґрунтувати вибір середньої величини.

Номер заводу	2022 рік		2023 рік	
	Витрати часу на одиницю продукції, год	Виготовлено продукції, шт.	Витрати часу на одиницю продукції, год	Витрати часу на всю продукцію, год
1	2,0	150	1,9	380
2	3,0	250	3,0	840

7. У таблиці наведено кількість студентів, прийнятих до навчальних закладів країни, у тому числі на денне відділення у 2021/2022 та 2022/2023 навчальних роках.

Навчальні заклади	2020/2021 навч. рік		2021/2022 навч. рік	
	Усього прийнято студентів, тис. чол.	Частка прийнятих на денне відділення, %	Прийнято студентів на денне відділення, тис. чол.	Частка прийнятих на денне відділення, %
Виші	181,7	58,8	113,4	65,3
Середні технічні	264,6	62,8	157,4	66,3

Визначити середню частку студентів, яких прийнято на денне відділення за роками.

**8.** Розподіл професорсько-викладацького персоналу навчального закладу за стажем роботи:

Стаж роботи, кількість років	Кількість робітників, чол.
До 10	64
10–15	52
15–20	40
20–25	22
Понад 25	24
Усього	202

Визначити середній стаж роботи професорсько-викладацького персоналу.

**9.** Визначити за даними, наведеними нижче, середній бал якості відеокліпів, що демонструються на телебаченні.

Рівень якості відеокліпів	Кількість відповідей, %
Дуже високий	6
Високий	10
Середній	43
Нижчий від середнього	25
Низький	16
Разом	100

**10.** Визначити середній бал задоволення населення якістю медичного обслуговування за наведеними даними.

Чи задоволені ви якістю медичного обслуговування?	Кількість відповідей, %
Так, повністю	15
Частково	49
Не задоволений	36
Разом	100

**11.** На основі даних про розподіл магазинів за розміром товарообігу визначити моду й медіану.

Групи магазинів за розміром товарообігу, тис. грн	До 5	5–10	10–15	15–20	20–25	25 та більше
Кількість магазинів	6	12	20	11	7	4

**12.** Обчислити середній стаж роботи працівників цеху, модальний і медіанний рівні стажу працівників, перший і третій квартилі, перший і дев'ятий децилі, використовуючи наведені дані.

Стаж роботи, кількість років	До 5	5–10	10–15	15–20	20 і більше
Частка робітників, %	10	44	30	10	6

## Контрольні запитання

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

## 6. ПОКАЗНИКИ ВАРІАЦІЇ

### 6.1. Поняття варіації та її основні показники

Середня величина – це узагальнювальна характеристика ознаки статистичної сукупності яка, однак, не пояснює, як групуються навколо неї окремі значення, знаходяться вони поблизу або значно відхиляються від середньої величини. Змінність окремих значень ознаки характеризують показники варіації.

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

Ступінь близькості окремих одиниць до середньої вимірюється системою показників варіації, до якої належать абсолютні, середні й відносні показники. Абсолютними показниками є варіаційний розмах, середнє лінійне й середнє квадратичне відхилення, дисперсії. Відносні характеристики: коефіцієнти варіації, нерівномірності, локалізації, концентрації.

#### **Абсолютні й середні показники варіації**

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

$$R = X_{max} - X_{min}. \quad (6.1)$$

Тут фіксуються лише крайні відхилення ознаки, повторюваність проміжних значень не враховується.

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

$$\bar{d} = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{n} \quad (\text{дані є незгрупованим}); \quad (6.2)$$

$$\bar{d} = \frac{\sum_{i=1}^m |x_i - \bar{x}| f_i}{\sum_{i=1}^m f_i} \quad (\text{дані є згрупованими}). \quad (6.3)$$

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

На практиці ступінь варіації об'єктивніше відображає **дисперсія**  $\sigma^2$  – середня арифметична величина із суми квадратів відхилень окремих варіантів від їх середньої величини:

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n} \quad (\text{дані є незгрупованими}); \quad (6.4)$$

$$\sigma^2 = \frac{\sum_{i=1}^m (x_i - \bar{x})^2 f_i}{\sum_{i=1}^m f_i} \quad (\text{дані є згрупованими}). \quad (6.5)$$

Обчислити дисперсію можна за формулою

$$\sigma^2 = \overline{x^2} - \bar{x}^2, \quad (6.6)$$

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

**Середнє квадратичне відхилення**  $\sigma$  – корінь квадратний з дисперсії:

$$\sigma = \sqrt{\sigma^2}. \quad (6.7)$$

Середнє квадратичне відхилення є мірилом надійності середньої величини: чим меншим є  $\sigma$ , тим краще середня арифметична величина буде відображати 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 інших елементів,  $d_0 = 1 - d_1$ .



Для забезпечення порівняння абсолютних показників варіації у варіаційних рядах різних явищ обчислюють відносні показники – **коефіцієнти варіації**, які дають змогу 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 оцінювання однорідності сукупності, тобто надійності й типовості середньої величини:

$$V = \frac{\sigma}{\bar{x}} \cdot 100 \% (\bar{x} \neq 0). \quad (6.8)$$

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

**Коефіцієнт осциляції** відображає відносну змінюваність крайніх значень ознаки навколо середньої величини:

$$K_o = \frac{R}{\bar{x}} \cdot 100 \% (\bar{x} \neq 0). \quad (6.9)$$

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

$$K_d = \frac{\bar{d}}{\bar{x}} \cdot 100 \% (\bar{x} \neq 0). \quad (6.10)$$

## 6.2. Види дисперсій і правила їх додавання

Варіація ознаки формується під впливом різних причин, серед яких можна виокремити випадкові й систематичні. Отже, варіація може бути випадковою й систематичною. Визначити кожен з них можна шляхом дисперсійного аналізу.

Дисперсія є адитивною величиною, тобто в структурованій сукупності, яку було поділено на групи за факторною ознакою  $x$ , дисперсію результативної ознаки (**загальну**) можна розкласти на **середню величину з групових дисперсій** і **міжгрупову** дисперсії:

$$\sigma^2 = \overline{\sigma_i^2} + \delta^2. \quad (6.11)$$

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

$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2 f_i}{\sum f_i}, \quad (6.12)$$

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

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

$$\overline{\sigma^2_i} = \frac{\sum \sigma^2_i f_i}{\sum f_i}, \quad (6.13)$$

де  $f_i$  – обсяг  $i$ -ї групи,  $\sum_{i=1}^m f_i = n$ ;  $\sigma_i^2$  – групова дисперсія, що відображає варіацію ознаки лише з урахуванням умов, що діють усередині  $i$ -ї групи,

$$\sigma_i^2 = \frac{\sum_{j=1}^{f_i} (x_{ji} - \bar{x}_i)^2}{f_i}. \quad (6.14)$$

**Міжгрупова дисперсія** характеризує варіацію результативної ознаки під впливом груповальної ознаки-фактора  $X$  і є дисперсією групових середніх величин відносно загальної середньої величини:

$$\delta^2 = \frac{\sum_{i=1}^m (\bar{x}_i - \bar{x}_0)^2 f_i}{\sum_{i=1}^m f_i}, \quad (6.15)$$

де  $f_i$  – обсяг  $i$ -ї групи,  $\sum_{i=1}^m f_i = n$ ;  $\bar{x}_i$  – середня кожної окремої групи;  $\bar{x}_0$  – загальна середня величина всієї сукупності.

### Приклад розв'язання типових задач

Є дані про розподіл робітників за тарифними розрядами (табл. 6.1).

Таблиця 6.1

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

Визначити:

- дисперсію;
- середнє квадратичне відхилення;
- квадратичний коефіцієнт варіації.

### Розв'язання

Для розрахунку показників варіації попередньо визначимо середню величину  $\bar{x}$  за формулою середнього арифметичного зваженого (5.8). Дисперсію  $\sigma^2$  обчислимо за формулою (6.5). Обчисливши зазначені показники варіації, необхідні розрахунки зведемо в табл. 6.2.

Таблиця 6.2

Тарифний розряд x	Кількість робітників f, чол.	xf	x - $\bar{X}$	(x - $\bar{X}$ )f	(x - $\bar{X}$ ) <sup>2</sup> 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 (6.8):

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

**Висновок:** середній розряд становить 4,5. Сукупність робітників є однорідною, тому що квадратичний коефіцієнт варіації є меншим за 33 %.

### Завдання для самостійної роботи

1. Є така сукупність:

Номер одиниці сукупності	1	2	3	4	5	6	7	8	9	10
Значення ознаки	19,2	17,8	20,1	16,9	20,4	18,7	18,3	19,6	19,8	17,5

Визначити показники варіації: розмах варіації; загальний обсяг варіації; середній розмір відхилень у розрахунку на одиницю сукупності; середнє лінійне відхилення; дисперсію; середнє квадратичне відхилення; коефіцієнт варіації.

2. Є такі дані про товарообіг по 20 магазинах продовольчих товарів: 2200, 3000, 1000, 2000, 4000, 3300, 2500, 2320, 1500, 2100, 2700, 30, 00, 200, 30 3500, 2380, 2600 тис. грн. Побудуйте інтервальний ряд розподілу торгових підприємств за товарообігом, утворивши п'ять груп з однаковими інтервалами. Розрахуйте характеристики ряду розподілу. Зробіть висновки.

3. Наведено такі дані про виробництво продукції на двох підприємствах галузі за п'ятиденним робочим тижнем звітного місяця, тис. грн:

Підприємство	Кількість днів				
	10	12	11	14	12
1	10	12	11	14	12
2	12	13	14	12	14

З допомогою показників варіації оцінити ритмічність виробництва на кожному підприємстві (розрахувати розмах варіації, середнє лінійне відхилення, дисперсію, квадратичний коефіцієнт варіації).

4. Є дані про розподіл вкладів населення до банків двох районів:

Розмір вкладу, грн	Кількість вкладників, %	
	1-й район	2-й район
До 800	12	10
800–1200	28	10
1200–1600	35	30
1600–2000	13	40
Понад 2000	12	10
Усього	100	100

Визначити дисперсію, використовуючи правило додавання дисперсій.

5. Для вивчення рівня заробітної плати робітників на підприємстві обстежено 500 чоловіків і 300 жінок. З результатів дослідження випливає, що у чоловіків середня заробітна плата становила 1200 у. о. при середньоквадратичному відхиленні 200 у. о., у жінок відповідно 800 у. о. та 150 у. о.

Визначити: 1) середню заробітну плату працівників; 2) дисперсії заробітної плати й коефіцієнт варіації.

6. За результатами зимової сесії середній бал успішності становив: група 621е – 4,2 бала, група 621ме – 4,0 бала. Дисперсія успішності відповідно становила 0,81 та 0,15. Визначити для кожної групи квадратичний коефіцієнт варіації та зробити висновок, у якій групі успішність є стабільнішою.

7. Місячний прибуток банку у I півріччі характеризується такими показниками: 3,7; 4,3; 6,7; 5,6; 5,1; 7.1 млн грн. У II півріччі місячний прибуток становив: 4,6; 5,7; 6,4; 5,9; 5,2; 6,2 млн грн. Визначити, у якому півріччі прибуток банку був стабільнішим.

## Контрольні запитання

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

## 7. ВИБІРКОВЕ СПОСТЕРЕЖЕННЯ

### 7.1. Вибіркове спостереження та його основні завдання

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

Сукупність, з якої вибираються 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**. Залежно від причин виникнення ці помилки поділяють на систематичні й випадкові. **Систематичні помилки** виникають унаслідок порушення принципів випадковості відбору. **Випадкові помилки** – це наслідок випадковості відбору елементів сукупності для обстеження.

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

Частку одиниць генеральної сукупності, якій властива ознака, що вивчається, називають **генеральною часткою**  $p$ , а середня величина ознаки – **генеральною середньою**. У вибіркової сукупності частку одиниць, якій властива ознака, що вивчається, називають **вибірковою часткою** (частістю)  $w$ , а середню величину ознаки у вибірці – **вибірковою середньою величиною**  $\bar{x}$ .

Розрізняють такі види вибіркового спостереження:

- 1) **власне випадкова вибірка** – передбачається випадковий відбір одиниць з генеральної сукупності. Це – класичний спосіб формування

вибіркової сукупності, і саме на ньому ґрунтується теорія вибіркового методу;

2) **механічна вибірка** – це послідовний вибір одиниць через однакові проміжки за їх розташуванням у генеральній сукупності або в будь-якій іншій послідовності. Вибір елементів здійснюється через однакові інтервали, крок інтервалу залежить від частоти вибірки. Так, при  $\frac{n}{N} = 0,05$  (де  $n, N$  – обсяг вибіркової та генеральної сукупностей) крок інтервалу становить  $1 / 0,05 = 20$ ;

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

4) **серійна (гніздова) вибірка** складається із серій елементів сукупності, пов'язаних територіально (райони, селища), організаційно (фірми, акціонерні суспільства) і т. ін. Серії добираються за схемою механічної або простої випадкової вибірки, обстеженню підлягають усі елементи серії.

## 7.2. Знаходження середньої й граничної помилок і необхідного обсягу вибірки

Можливі розходження між середніми величинами або частками ознаки вибіркової й генеральної сукупностей вимірюються **середніми (стандартними) помилками  $\mu$**  (табл. 7.1).

Таблиця 7.1

Помилка вибірки	Метод відбору	
	Повторна вибірка	Безповторна вибірка
Для середньої величини	$\mu = \sqrt{\frac{\sigma^2}{n}}$	$\mu = \sqrt{\frac{\sigma^2}{n} \left(1 - \frac{n}{N}\right)}$
Для частки	$\mu = \sqrt{\frac{w(1-w)}{n}}$	$\mu = \sqrt{\frac{w(1-w)}{n} \left(1 - \frac{n}{N}\right)}$

Тут  $\sigma^2$  – вибіркова дисперсія;  $n$  – обсяг вибірки;  $N$  – чисельність генеральної сукупності;  $w$  – частка одиниць сукупності, які мають певну ознаку.

При практичному використанні наведених формул слід урахувати таке:

1) дисперсія частоти є добутком частотей  $\sigma_w^2 = w(1-w)$ ;

2) якщо вибірка є нечисленною (меншою від 30 одиниць), то у формулу

помилки вибірки вноситься поправка  $\frac{n}{n-1}$ ;

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

**Гранична помилка вибірки**  $\Delta$  – це максимально можлива помилка для прийнятої ймовірності  $p$ , якій відповідає коефіцієнт довіри –  $t$ -разове значення  $\mu$ . Коефіцієнт довіри відображає кількість середніх помилок у граничній помилці. Так,  $t = 1$  – для ймовірності 0,683;  $t = 2$  – для 0,954;  $t = 3$  – для 0,997,  $t = 4$  – для 0,998.

Формула граничної помилки має вигляд

$$\Delta = t\mu. \quad (7.1)$$

Для розрахунку граничної помилки вибірки застосовують формули, наведені в табл. 7.2.

Таблиця 7.2

Помилка вибірки	Метод відбору	
	Повторна вибірка	Безповторна вибірка
Для середньої величини	$\Delta_{\bar{x}} = t\sqrt{\frac{\sigma^2}{n}}$	$\Delta_{\bar{x}} = t\sqrt{\frac{\sigma^2}{n} \left(1 - \frac{n}{N}\right)}$
Для частки	$\Delta_w = t\sqrt{\frac{w(1-w)}{n}}$	$\Delta_w = t\sqrt{\frac{w(1-w)}{n} \left(1 - \frac{n}{N}\right)}$

Помилку механічної вибірки знаходять за формулами з табл. 7.2.

При типовому відборі для визначення помилок вибірки замість  $\sigma^2$  використовують середню з групових дисперсій:

$$\bar{\sigma}^2 = \frac{\sum_1^m \sigma_i^2 n_i}{\sum_1^m n_i},$$

тоді

$$\Delta = t\sqrt{\frac{\bar{\sigma}^2}{n} \left(1 - \frac{n}{N}\right)}.$$

Зазвичай помилка районованої вибірки є меншою, ніж механічної або випадкової вибірки. Найчастіше використовують добір, пропорційний кількості складових сукупності, тобто частість вибірки для всіх складових є однаковою.

З допомогою формул граничної помилки вибірки визначають:

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

– для середньої генеральної сукупності

$$\bar{x} - \Delta \leq \bar{x} \leq \bar{x} + \Delta; \quad (7.2)$$

– для частки одиниць, для яких властивою є ця ознака в генеральній сукупності,

$$w - \Delta \leq w \leq w + \Delta; \quad (7.3)$$

2) імовірність того, що відхилення між вибірковими й генеральними характеристиками не перевищує певну 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. 7.3).

Таблиця 7.3

Чисельність вибірки	Метод відбору	
	Повторна вибірка	Безповторна вибірка
Для середньої величини	$n = \frac{t^2 \sigma^2}{\Delta_x^2}$	$n = \frac{t^2 \sigma^2 N}{\Delta_x^2 N + t^2 \sigma^2}$
Для частки	$n = \frac{t^2 w(1-w)}{\Delta_w^2}$	$n = \frac{t^2 w(1-w)N}{\Delta_w^2 N + t^2 w(1-w)}$

Для визначення обсягу вибірки  $n$  використовують оцінки дисперсій  $\sigma^2$  пробних обстежень. Якщо таких обстежень немає, то можна скористатися співвідношенням  $\sigma = \frac{1}{6}(x_{\max} - x_{\min})$ , а для частоти взяти найбільше значення дисперсії  $\sigma^2 = 0,25$ .

Якщо в основу розрахунку  $n$  покласти відносну помилку вибірки, то формули відповідно модифікуються:

– для середньої  $n = \frac{t^2 V_x^2}{V_\Delta^2}; \quad (7.4)$

– для частки  $n = \frac{t^2 q}{V_\Delta^2 p}. \quad (7.5)$

При порівнянні точності вибіркових оцінок використовують **відносну помилку вибірки**  $V_\mu$ , яка є відображенням того, на скільки відсотків вибіркова оцінка відхиляється від параметра генеральної сукупності:

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

Відносну помилку вибірки можна розрахувати на основі коефіцієнта варіації ознаки  $V_x$ :



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

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

Аналогічно розраховують відносну помилку вибірки для частоти:

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

### Приклад розв'язання типових задач

Для дослідження оснащення заводів основними виробничими фондами було проведено 10-відсоткове вибіркове обстеження, унаслідок якого отримано дані про розподіл 50 заводів за вартістю основних виробничих фондів (табл. 7.4).

Таблиця 7.4

Середньорічна вартість основних виробничих фондів, млн грн	До 2	2,0–4,0	4,0–6,0	Понад 6,0
Кількість заводів	5	12	23	10

Обчислити:

1) з імовірністю 0,997 граничну помилку вибіркової середньої й межі, у яких буде знаходитися значення середньорічної вартості основних виробничих фондів усіх заводів генеральної сукупності;

2) з імовірністю 0,954 граничну помилку вибірки при визначенні частки й межі, у яких буде знаходитися значення питомої ваги заводів з вартістю основних виробничих фондів понад 4 млн грн;

3) обсяг вибіркової сукупності за умов:

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

б) гранична помилка частки з імовірністю 0,954 є не більшою за 15 %.

#### Розв'язання

Задано:  $n = 50$ ,  $N = 500$  (за умовою задачі проведено 10-відсоткове вибіркове обстеження, отже, генеральна сукупність – це 100 %, або 500 заводів).

1. Для визначення меж генеральної середньої обчислимо середню вибіркочну  $\bar{x}$  за формулою арифметичної зваженої (5.8) і дисперсію  $\sigma^2$  за формулою (6.5). Техніку розрахунку наведено в табл. 7.5.

Таблиця 7.5

Середньорічна вартість основних виробничих фондів, млн грн	Кількість 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 помилка вибірки при визначенні середньорічної вартості основних фондів становить:

– при повторному доборі

$$\mu = \sqrt{\frac{\sigma^2}{n}} \quad \mu = \sqrt{\frac{3,13}{50}} \approx 0,25 \text{ млн грн};$$

– при безповторному доборі

$$\mu = \sqrt{\frac{\sigma^2}{n} \left(1 - \frac{n}{N}\right)} \quad \mu = \sqrt{\frac{3,13}{50} \left(1 - \frac{50}{500}\right)} \approx 0,24 \text{ млн грн}.$$

Отже, при визначенні середньорічної вартості основних виробничих фондів можна було би припуститися середньої помилки репрезентативності 0,25 млн грн при повторному і 0,24 млн грн при безповторному доборі в той чи іншій бік від середньорічної вартості основних виробничих фондів, що припадають на один завод у вибірковій сукупності. З обчислених даних випливає, що при безповторній вибірці середня помилка репрезентативності (0,24) є меншою, ніж за тих самих умов при повторному доборі (0,25).

Обчислимо граничну помилку вибіркової середньої  $\Delta_{\bar{x}}$  за формулою (7.1). У цьому прикладі:  $p = 0,997$ , отже,  $t = 3$ . Таким чином:

–  $\Delta_{\bar{x}} = 3 \cdot m = 3 \cdot 0,25 = 0,75$  млн грн при повторному доборі;

–  $\Delta_{\bar{x}} = 3 \cdot m = 3 \cdot 0,24 = 0,72$  млн грн при безповторному доборі.

Порядок установлення меж, у яких знаходиться середня величина досліджуваного показника в генеральній сукупності, тобто довірчий інтервал, можна визначити за формулою

$$\Delta_{\bar{x}} = t \sqrt{\frac{\sigma^2}{n} \left(1 - \frac{n}{N}\right)}.$$

Таким чином, при повторному доборі маємо

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

при безповторному –

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

Ці межі можна гарантувати з імовірністю 0,996.

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

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

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

При повторному доборі гранична помилка для частки

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

при безповторному доборі

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

При визначенні частки межі обчислюють аналогічно до меж для середньої величини за формулою (7.3).

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

– при повторному доборі

$$66 - 13,4 \leq w \leq 66 + 13,4, \text{ або } 52,6 \leq w \leq 79,4;$$

– при безповторному доборі

$$66 - 12,7 \leq w \leq 66 + 12,7, \text{ або } 53,3 \leq w \leq 78,6.$$

З розрахунків випливає, що при безповторному доборі помилка вибірки є меншою, ніж за тих самих умов при повторному доборі.

3. Обчислимо обсяг вибіркової сукупності:

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

Знайдемо обсяг вибірки для розрахунку середнього розміру ознаки:

– при повторному доборі

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

– при безповторному доборі

$$n = \frac{t^2 \sigma^2 N}{\Delta_{\bar{x}}^2 N + t^2 \sigma^2} = \frac{3^2 \cdot 3,13^2 \cdot 500}{0,5^2 \cdot 500 + 3^2 \cdot 3,13^2} \approx 92 \text{ заводи.}$$

Висновок: чисельність вибірки збільшиться, якщо за інших однакових умов зменшити граничну помилку;

б) відомо, що  $N = 500$ ;  $\Delta_w = 0,15$  млн грн;  $\sigma^2 = 3,13$ ;  $p = 0,954$ ;  $t = 2$ .

Обчислимо обсяг вибірки для розрахунку помилки частки:

– при повторному доборі

$$n = \frac{t^2 w(1-w)}{\Delta_w^2} = \frac{2^2 \cdot 0,66 \cdot 0,34}{0,15^2} \approx 40 \text{ заводів};$$

– при безповторному доборі

$$n = \frac{t^2 w(1-w)N}{\Delta_w^2 N + t^2 w(1-w)} = \frac{2^2 \cdot 0,66 \cdot 0,34 \cdot 500}{0,15^2 \cdot 500 + 2^2 \cdot 0,66 \cdot 0,34} \approx 37 \text{ заводів.}$$

Висновок: чисельність вибірки збільшиться, якщо за інших однакових умов збільшити граничну помилку вибірки.

### Завдання для самостійної роботи

1. Для встановлення середнього часу горіння електроламп було проведено 5-відсоткове обстеження партії електроламп, унаслідок якого було отримано такі дані:

Час горіння електролампи, тис. год	Кількість електроламп
3–5	10
5–7	36
7–9	72
9–11	64
11–13	12
13–15	6

Визначити:

1) з імовірністю 0,99 граничну помилку вибіркової середньої та межі, у яких знаходиться значення середнього часу горіння електролампи й усієї партії електроламп (генеральної сукупності);

2) з імовірністю 0,954 граничну помилку вибірки відносно частки й меж, у яких знаходиться частина електроламп з часом горіння трохи більше 11 тис. год.

2. Для встановлення середнього обсягу реалізації в організаціях роздрібно торгівлі міста було проведено 2-відсоткову типову вибірку з пропорційним добором усередині типових груп. У наслідок добору отримано такі дані:

Тип торговельної організації	Кількість організацій	Середній обсяг реалізації, млн грн	Середнє квадратичне відхилення, млн грн
Продовольчі	20	60	5
Непродовольчі	10	100	9

3 імовірністю 0,954 визначити межі, у яких знаходиться значення середнього обсягу реалізації в організаціях роздрібно торгівлі міста.

3. Обчислити, скільки магазинів необхідно добрати для вибіркового спостереження, щоб визначити середню кількість покупців. Помилка вибірки з імовірністю 0,997 має бути  $\pm 15$  чол. при середньому квадратичному відхиленні  $\pm 70$  чол.

4. Розміри костюмів визначають, виходячи з антропометричних показників. Скільки необхідно добрати дорослих для обчислення їх середнього росту, щоб помилка вибірки з імовірністю 0,954 не перевищила 1 см при дисперсії 36?

5. З метою вивчення витрат часу на виготовлення однієї деталі робітниками заводу проведено 10-відсоткову випадкову неповторну вибірку, унаслідок якої отримано такий розподіл:

Витрати часу на одну деталь, хв	Кількість деталей, од.
До 20	10
Від 20 до 24	20
Від 24 до 28	50
Від 28 до 32	15
Понад 32	5
Разом	100

Обчислити: середні витрати часу виготовлення однієї деталі; дисперсію та середнє квадратичне відхилення, коефіцієнт варіації; з імовірністю 0,954 граничну помилку вибіркової середньої величини та можливі межі, у яких передбачаються середні витрати часу на виготовлення іншої деталі на заводі; з імовірністю 0,95 граничну помилку вибіркової частки та межі частки від загальної кількості деталей з витратами часу на їх виготовлення від 20 до 28 хв. Зробити висновки.

## Контрольні запитання

1. Яке спостереження називають вибіркоvim? Якими є його суть і завдання?
2. У чому полягають переваги вибіркового спостереження порівняно із суцільним?
3. Яких умов необхідно дотримуватися під час відбору одиниць при вибіркоvimу спостереженні?
4. Чому при вибіркоvimу спостереженні завжди виникають помилки? Як їх класифікують і що вони характеризують?
5. Які ви знаєте види і способи добору до вибіркової сукупності?
6. Як здійснюють простий випадковий, механічний, районований відбір?
7. Про що свідчить коефіцієнт довіри і як його позначають?
8. Чи впливає обсяг вибірки на її точність і якою мірою?
9. Як визначають помилку вибірки для середньої й частки?
10. Як визначають необхідну чисельність вибірки для середньої й частки в разі повторного й безповторного добору?

## 8. СТАТИСТИЧНІ МЕТОДИ ВИВЧЕННЯ ВЗАЄМОЗВ'ЯЗКІВ

### 8.1. Основні поняття й категорії

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

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

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

Характеристикою кореляційного зв'язку є **лінія регресії**  $y$  на  $x$  – функція, яка пов'язує середні значення ознаки  $y$  зі значеннями ознаки  $x$ . Лінія регресії розглядається в двох моделях: аналітичного групування й регресійного аналізу. У 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:

$$\frac{\Delta y}{\Delta x}, \quad (8.1)$$

де  $\Delta y = \bar{y}_j - \bar{y}_{j-1}$ ,  $\Delta x = x'_j - x'_{j-1}$ .

**Оцінювання щільності зв'язку** ґрунтується на правилі складання дисперсій. В моделі аналітичного групування мірою щільності зв'язку є відношення міжгрупової дисперсії  $\delta^2$  до загальної  $\sigma^2$ , яке називають **кореляційним відношенням**:

$$\eta^2 = \frac{\delta^2}{\sigma^2}. \quad (8.2)$$

Тут  $\sigma^2$  – загальна дисперсія, за якою вимірюють варіацію результативної ознаки  $y$ , зумовлену дією всіх можливих чинників,

$$\sigma^2 = \frac{\sum (y_i - \bar{y})^2 f_i}{\sum f_i}, \quad (8.3)$$

або

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

де  $\delta^2$  – міжгрупова дисперсія, за якою вимірюють варіацію результативної ознаки  $y$  під дією тільки групувальної ознаки  $x$ ,

$$\delta^2 = \frac{\sum (\bar{y}_j - \bar{y})^2 f_j}{\sum f_j}. \quad (8.4)$$

Кореляційне відношення змінюється від 0 до 1 або від 0 до 100 %. При відсутності зв'язку маємо  $\eta^2 = 0$ , а при функціональному зв'язку –  $\eta^2 = 1$ . Чим більше це відношення наближається до одиниці, тим більшою буде щільність зв'язку.

Щільний зв'язок може виникнути випадково, тому необхідно перевірити наявність його щільності, тобто довести невідповідність зв'язку. **Перевірка**

**щільності зв'язку** – це порівняння фактичного значення  $\eta^2$ , обчисленого за формулою (8.2), з його критичним значенням  $\eta_{1-\alpha}^2(k_1, k_2)$  для певного рівня щільності  $\alpha$  і кількості ступенів свободи  $k_1 = m - 1$  і  $k_2 = n - m$ , де  $m$  – кількість груп;  $n$  – обсяг сукупності.

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

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

Критичне значення вибирають таким чином, щоб імовірність отримання значення  $\eta^2$ , більшого за критичне (при відсутності зв'язку між ознаками), була досить малою. Таку ймовірність називають рівнем істотності  $\alpha$ . Найчастіше зостосовують такі рівні істотності, як  $\alpha = 0,05$  і  $\alpha = 0,01$ . Критичні значення кореляційного відношення  $\eta^2$  і коефіцієнта детермінації  $R^2$  для рівня істотності  $\alpha = 0,05$  наведено в табл. 8.1.

Таблиця 8.1

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 зв'язану з  $\eta^2$  характеристику F-критерію (критерію Фішера), який обчислюють за формулою

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

Існують таблиці критичних значень F-критерію. Перевірку істотності зв'язку з його допомогою здійснюють аналогічно описаній вище для кореляційного відношення  $\eta^2$ .



## 8.2. 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 має назву **рівняння регресії**.

На основі рівняння регресії визначають 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 змінням  $x$  ознака  $y$  змінюється більш-менш рівномірно;

– **нелінійні рівняння**, коли взаємозв'язані ознаки змінюються нерівномірно (з прискоренням, уповільненням або зі змінним напрямком зв'язку), зокрема степенева  $Y = ax^b$ , гіперболічна  $Y = a + b/x$ , параболічна  $Y = a_0 + a_1x + a_2x^2$  тощо.

Найчастіше застосовують лінійні рівняння, тобто рівняння, які було зведено до лінійного вигляду. У лінійному рівнянні параметр  $a_1$  – **коефіцієнт регресії** – є відображенням того, на скільки одиниць у середньому зміниться  $y$  зі змінням  $x$  на одиницю. Цей коефіцієнт має одиницю вимірювання результативної ознаки. При прямому зв'язку  $a_1$  є величиною додатною, а при зворотному – від'ємною. Параметр  $a_0$  – вільний член рівняння регресії, тобто це значення, якого набуває  $y$  при  $x = 0$ . Якщо  $x$  не набуває нульового значення, то цей параметр буде тільки розрахунковим. Параметри визначають **методом найменших квадратів (МНК)**, згідно з яким сума квадратів відхилень емпіричних значень  $y$  від теоретичних  $Y$  є мінімальною:  $\sum (y - Y)^2 \rightarrow \min$ . Відповідно до умови мінімізації параметри лінійного рівняння регресії обчислюють 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

$$a_1 = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - \sum x \sum y}; \quad (8.5)$$

$$a_0 = \bar{y} - a_1 \bar{x}. \quad (8.6)$$

Коефіцієнт регресії в невеликих за обсягом сукупностях ( $n < 30$ ) може випадково змінюватися. Тому перевіряють його істотність з допомогою  $t$ -критерію Стьюдента. При цьому фактичні значення  $t$ -критерію розраховують за формулою

$$t_{a_1} = \frac{a_1}{\mu_{a_1}}. \quad (8.7)$$

Тут  $a_1$  – коефіцієнт регресії;  $\mu_{a_1}$  – власне стандартна помилка, яку розраховують за формулою

$$\mu_{a_1} = \frac{\sigma_e}{\sigma_x} \sqrt{(n-2)}, \quad (8.8)$$

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

$$\sigma_x = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n}}; \quad (8.9)$$

$\sigma_e$  – середнє квадратичне відхилення результативної ознаки  $y_i$  від теоретичних значень  $Y$ ,

$$\sigma_e = \sqrt{\frac{\sum (y_i - Y)^2}{n}}. \quad (8.10)$$

Розраховане за формулою (8.7) фактичне значення  $t_{a_1}$   $t$ -критерію порівнюють з критичним значенням  $t_k$ , яке отримують за таблицею Стьюдента з урахуванням рівня істотності  $\alpha$  і кількості ступенів свободи  $k = n - 2$ . Якщо  $t_k < t_{a_1}$ , то коефіцієнт регресії вважається істотним.

Характеристикою відносного змінення  $y$  завдяки  $x$  є **коефіцієнт еластичності**

$$K_{ел} = a_1 \frac{\bar{x}}{y}, \quad (8.11)$$

який є відображенням того, на скільки відсотків у середньому змінюється результативна ознака зі змінням фактора на 1 %.

Для статистичного оцінювання щільності зв'язку використовують такі показники варіації:

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

$$\sigma_y^2 = \frac{1}{n} \sum_1^n (y - \bar{y})^2; \quad (8.12)$$

2) факторна дисперсія результативної ознаки, яка є відображенням варіації  $y$ , зумовленої дією тільки фактора  $x$ :

$$\sigma_Y^2 = \frac{1}{n} \sum_1^n (Y - \bar{y})^2; \quad (8.13)$$

3) залишкова дисперсія, яка є відображенням дії на результативну ознаку всіх інших факторів, окрім  $x$ :

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

Частка факторної дисперсії в загальній є характеристикою щільності зв'язку й має назву **коефіцієнта детермінації**:

$$R^2 = \frac{\sigma_Y^2}{\sigma_y^2}. \quad (8.14)$$

Він має такі жє суть, інтерпретацію й цифрові межі, як і  $\eta^2$ .

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

$$r = \frac{\sum xy - n\bar{x}\bar{y}}{n\sqrt{\sigma_x^2\sigma_y^2}},$$

або

$$r = \frac{\overline{xy} - \bar{x}\bar{y}}{\sigma_x\sigma_y}, \quad (8.15)$$

який набуває значення в межах  $\pm 1$ , тому характеризує не тільки щільність, але й 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 групування, шляхом порівняння фактичного й критичного значень. Відмінності мають тільки визначення  $k_1 = m - 1$  і  $k_2 = n - m$ , у яких  $m$  – кількість 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 пов'язаний з  $R^2$  і  $\eta^2$ . Тоді фактичне значення обчислюють 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 й висновки є ідентичними.

## Приклади розв'язання типових задач

### Задача 8.1

Відомо дані про розподіл проданих на біржі нерухомості однокімнатних квартир за їх середньою вартістю  $y$  і розміром загальної площі  $x$  (табл. 8.2, стовпці 1–3), а також про загальну дисперсію  $\sigma^2 = 6$ . З допомогою аналітичного групування встановити кількісні співвідношення між середньою вартістю  $y$  і розміром загальної площі проданих квартир, оцінити щільність зв'язку з допомогою кореляційного відношення й перевірити істотність зв'язку між ознаками з імовірністю 0,95.

Таблиця 8.2

Загальна площа квартири $X_j$ , м <sup>2</sup>	Кількість квартир $f_j$	Середня вартість квартири $\bar{y}_j$ , тис. грн	$\bar{y}_j f_j$	$\bar{y}_j - \bar{y}$	$(\bar{y}_j - \bar{y})^2 f_j$
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
У цілому	100	13,0	1300	X	461,0

### Розв'язання

У табл. 8.2 наведено приклад аналітичного групування проданих квартир, де описано залежність їх вартості від загальної площі. Маємо стохастичний і зокрема кореляційний зв'язок.

Аналітичне групування дає змогу встановити за формулами (8.1) кількісні співвідношення між ознаками, що вивчаються. За даними табл. 8.2 приріст  $\Delta_x$  у всіх групах є однаковим – 5 м<sup>2</sup>, а середня вартість проданих квартир збільшується по групах таким чином:

$$\Delta_{y_2} = 13,2 - 10,8 = 2,4 \text{ тис. грн};$$

$$\Delta_{y_3} = 15,2 - 13,2 = 2,0 \text{ тис. грн};$$

$$\Delta_{y_4} = 18,0 - 15,2 = 2,8 \text{ тис. грн}.$$

Отже, із збільшенням розміру загальної площі квартир на 1 м<sup>2</sup> їх вартість у середньому збільшується на таку величину:

$$\frac{\Delta_{y_2}}{\Delta_{x_2}} = 2,4 : 5 = 0,48 \text{ тис. грн і на } 0,4 \text{ і } 0,56 \text{ тис. грн відповідно.}$$

Оцінимо щільність зв'язку за формулою (8.2). Деякі проміжні розрахунки наведено в табл. 8.2.

Насамперед розрахуємо середню загальну вартість квартир за формулою (5.9):

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

міжгрупову дисперсію – за формулою (8.5):

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

Загальна дисперсія  $\sigma^2 = 7$  є відомою, тому кореляційне відношення

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

Отже, варіація вартості проданих квартир на 65,9 % пояснюється варіацією їх загальної площі та на 34,1 % – варіацією інших факторів, тобто зв'язок між ознаками є досить щільним.

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

У цьому прикладі кількість ступенів свободи є такою:

$$k_1 = 4 - 1 = 3, k_2 = 100 - 4 = 95.$$

Критичні значення кореляційного відношення для  $\alpha = 0,05$  наведено в табл. 8.1, але через відсутність у таблиці критичних значень  $k_2 = 96$  використовуємо найближче ( $k_2 = 100$ ), тоді

$$\eta_{0,95}^2(3, 100) = 0,075.$$

Оскільки  $\eta^2 = 0,659$ , що є більшим за 0,075, то зв'язок визнається істотним з імовірністю 0,95.

## **Задача 8.2**

Дані, що відображають зв'язок між добовою вартістю туристичних путівок в одному з туристичних агентств і тривалістю відпочинку (кількість днів), наведено в табл. 8.3 (1–3 стовпці).

Розрахувати параметри лінійного рівняння регресії, економічно інтерпретувати. Визначити коефіцієнт еластичності.

Перевірити істотність коефіцієнта регресії з допомогою  $t$ -критерію Ст'юдента.

Таблиця 8.3

Номер путівки	Тривалість відпочинку, кількість днів	Добова вартість путівки, грн	$xu$	$x^2$	$y^2$	$Y = 113,25 - 4,34x$	$(y - Y)^2$
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
Усього	100	472	4972	1464	32244	472,0	372,0

### Розв'язання

Розрахуємо параметри лінійного рівняння регресії за формулами (8.5), (8.6). Величини, на основі яких обчислюються параметри:

$n = 8$ ;  $\sum x = 100$ ;  $\sum y = 472$ ;  $\sum xy = 4972$ ;  $\sum x^2 = 1464$ ;  $\bar{x} = 100 : 8 = 12,5$ ;  $\bar{y} = 472 : 8 = 58$ . Отже, параметри будуть такими:

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

$$a_0 = 58 - (-4,34) \cdot 12,5 = 113,25.$$

Тоді рівняння регресії набуде вигляду  $Y = 113,25 - 4,34x$ , 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 (8.7). Для цього, використавши розрахунки, наведені в табл. 8.3, обчислимо за формулою (8.9) середнє квадратичне відхилення факторної ознаки від загальної середньої  $\sigma_x = \sqrt{26,75} = 5,17$ , за формулою (8.10) – середнє квадратичне відхилення результативної ознаки від теоретичних значень  $Y$   $\sigma_e = \sqrt{46,5}$ . Тоді знайдена за формулою (8.8) стандартна помилка

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

$$t_{a_1} = \frac{|-4,34|}{0,54} = 8,$$

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

Таким чином, з імовірністю 0,95 дія тривалості відпочинку на добову вартість путівок визнається істотною.

З імовірністю 0,95 довірчі межі коефіцієнта регресії будуть такими:

$$-4,34 \pm 2,54 \cdot 0,54, \text{ тобто } -4,34 \pm 1,37 \text{ грн.}$$

Коефіцієнт еластичності розрахуємо за формулою (8.11):

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

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

Визначимо теоретичні значення  $Y$  – очікувану вартість путівок з урахуванням тільки тривалості відпочинку (табл. 8.3, стовпець 6). Так, для  $x = 5$  днів добова вартість путівки

$$Y = 113,2 - 54,34 \cdot 5 = 91,6 \text{ грн.}$$

Це значення відхиляється від емпіричного.

Теоретичні значення  $Y$  використаємо для обчислення коефіцієнта детермінації за формулою (8.14), загальну дисперсію результативної ознаки – за формулою (8.12):

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

факторну дисперсію результативної ознаки – за формулою (8.13):

$$\sigma_Y^2 = \frac{4031}{8} = 503,86.$$

Тоді  $R^2 = 503,86 : 549,5 = 0,917$ , тобто 91,7 % варіації добової вартості путівок лінійно пов'язані з варіацією тривалості відпочинку, а 8,3 % варіації припадає на інші фактори, тому зв'язок є дуже щільним.

Визначимо лінійний коефіцієнт кореляції (Пірсона)  $r$  за формулою (8.15):

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

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  $R_{0,95}^2(1,6) = 0,5$  є набагато меншим від фактичного:  $R^2 = 0,915$ .

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

### Задача 8.3

За даними табл. 8.4 оцінити щільність зв'язку між рівнем ефективності економіки й надійністю ділового партнерства для семи країн Східної Європи.

Таблиця 8.4

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  $n = 1$ , найбільшому –  $n = 6$ . Сума квадратів відхилень рангів  $\sum_1^n d_j^2 = 10$ , а коефіцієнт рангової кореляції

$$\rho = 1 - \frac{\sum_1^n d_j^2}{n(n^2 - 1)} = 1 - \frac{10}{7(7^2 - 1)} = 1 - \frac{10}{336} = 0,82.$$

Значення коефіцієнта рангової кореляції свідчить про наявність прямого й досить помітного зв'язку між наведеними параметрами ризику іноземного інвестування економіки. Згідно з табл. 8.4 критичне значення коефіцієнта рангової кореляції  $\rho_{0,95}(7)$  для  $\alpha = 0,05$  і  $n = 7$  становить 0,71, що значно є меншим від фактичного. Отже, істотність зв'язку було доведено з імовірністю 0,95.



## Завдання для самостійної роботи

1. В одному з відділень банку досліджувався зв'язок між наявністю вкладу й сімейним становищем потенційних вкладників на 01.01.2021. Результати обстеження характеризуються такими даними:

Сімейний стан	Кількість вкладників, чол.		
	усього, чол.	що мають заощадження	що не мають заощадження
Самотні	400	250	150
Сімейні	1250	800	450
Усього	1650	1050	600

Визначити коефіцієнт асоціації.

2. Є вибіркві дані про випуск продукції та суму прибутку, млн грн:

Номер підприємства	Випуск продукції	Прибуток	Номер підприємства	Випуск продукції	Прибуток
1	65	15,7	16	52	14,6
2	78	18	17	62	14,8
3	41	12,1	18	69	16,1
4	54	13,8	19	85	16,7
5	66	15,5	20	70	15,8
6	80	16,9	21	71	16,4
7	45	12,8	22	64	15
8	57	14,2	23	72	16,5
9	67	15,9	24	88	18,5
10	81	16,6	25	73	16,4

Методом аналітичного групування встановити наявність і характер кореляційного зв'язку між вартістю виробленої продукції та сумою прибутку на одне підприємство (результати оформити робочою й аналітичною таблицями).

Виміряти щільність кореляційного зв'язку між вартістю виробленої продукції та сумою прибутку емпіричним кореляційним відношенням. Зробити висновки.

3. Є такі умовні дані:

Групи підприємств за формою власності	Кількість підприємств	Середній випуск продукції, млн грн	Середнє квадратичне відхилення
Державні	10	500	100
Комерційні	20	800	120

Виміряти щільність кореляційного зв'язку між випуском продукції та формою власності підприємства емпіричним кореляційним відношенням.

4. За наведеними даними про вартість основних фондів і випуск продукції підприємствами галузі наведіть кількісну характеристику зв'язку між цими підприємствами, використовуючи метод аналітичного групування, дисперсійний аналіз, кореляційний аналіз.

Номер підприємства	Вартість основних фондів, млн грн	Випуск продукції, млн грн
1	5,5	12,0
2	6,0	12,4
3	7,0	12,9
4	7,5	13,1
5	8,5	13,5
6	8,7	13,6
7	9,4	13,8
8	10,2	14,0
9	10,3	14,1
10	12,1	14,4

5. Унаслідок обстеження виробності 62 верстатів з різними термінами експлуатації виявлено:

Термін експлуатації верстатів, кількість років	Кількість верстатів	Виробництво деталей за зміну в перерахунку на один верстат, шт.
До 7	10	110
7–14	15	96
14–20	25	70
20 і більше	12	66
Разом	62	82

Загальна дисперсія виробності верстатів становить 470. Визначити міжгрупову дисперсію та кореляційне відношення. Перевірити суттєвість зв'язку з імовірністю 0,95. Зробити висновки.

6. За результатами перевірки якості 20 партій твердих сирів виявлено залежність якості від терміну зберігання:

Термін зберігання, міс.	Кількість партій	Зменшення якості, бали	Групова дисперсія зменшення якості
До 2	7	1,3	0,08
2–4	8	2,8	0,13
4 й більше	5	4,1	0,20
Разом	20	2,6	–

Визначити міжгрупову, середню з групових і загальну дисперсію зменшення якості сиру, відобразити їх зв'язок. Розрахувати кореляційне відношення та пояснити його зміст. Перевірити суттєвість зв'язку з імовірністю 0,95.

7. За даними таблиці оцінити щільність зв'язку між виконанням норм виробітку молодими працівниками та рівнем їх освіти.

Група робітників	Виконують норму виробітку	Не виконують норму виробітку	Разом
Закінчили середню школу	78	22	100
Не закінчили середню школу	32	68	100
Разом	110	90	200

8. За даними про розподіл шкіл міста за типами та оцінкою складності навчального предмета «Основи інформатики», тис. чол., розрахувати коефіцієнти взаємного зв'язку Пірсона й Чупрова. Зробити висновки.

Тип школи	Відмінне освоєння курсу	Середнє освоєння курсу	Проблеми з освоєнням курсу	Разом
А	74	12	5	91
Б	71	11	12	94
В	62	21	23	106
Разом	207	44	40	291

### Контрольні запитання

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

## 9. РЯДИ ДИНАМІКИ

### 9.1. Ряди динаміки. Класифікація динамічних рядів

**Динаміка** – це змінення статистичних показників у часі.

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

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

Будь-який ряд динаміки містить два обов'язкових елементи:

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

2) **рівні** розвитку явища  $y$  – конкретні значення статистичних показників, що належать 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 менші інтервали (субперіоди) часу.

**Моментний ряд** динаміки характеризує обсяг явища на певні моменти часу  $t$ , наприклад, послідовності показників кількості населення на початок року, величину запасу деякого матеріалу на початок періоду й т. ін. Підсумовування рівнів моментного ряду не має сенсу, оскільки може виникнути повторне врахування.

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

За повнотою часу динамічні ряди поділяють на повні й неповні.

**Повні ряди** динаміки мають місце, коли дати реєстрації або закінчення періодів ідуть одна за одною з однаковими інтервалами. Це – рівновіддалені ряди динаміки.

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

За кількістю змінних показників ряди поділяють на одновимірні й багатовимірні.

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

Своєю чергою багатовимірні ряди поділяють на такі:

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

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

## 9.2. Характеристики динамічних рядів

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

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

Рівень, який порівнюють з іншими рівнями, називають **поточним**, а з яким порівнюють інші рівні – **базисним**.

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

**Абсолютний приріст**  $\Delta_t$  є вираженням абсолютної швидкості змінення рівнів ряду динаміки і відображенням того, на скільки одиниць підвищився чи зменшився рівень відносно базисного за певний період часу:

– базисний

$$\Delta_t^b = y_t - y_0; \quad (9.1)$$

– ланцюговий

$$\Delta_t^l = y_t - y_{t-1}, \quad (9.2)$$

де  $y_i$ ,  $y_{i-1}$ ,  $y_0$  – рівні ряду в аналізовному, попередньому або базисному періоді.

**Коефіцієнт зростання**  $K_{зр}$  є вираженням відносної швидкості динаміки і відображенням того, у скільки разів порівнюваний рівень є більшим ( $K_{зр} > 1$ ) або меншим ( $K_{зр} < 1$ ) від базисного. Відсоткове подання коефіцієнта зростання – **темп зростання**  $T_{зр}$ :

– базисний

$$K_{зр_t}^b = \frac{y_t}{y_0}; \quad (9.3)$$

$$T_{зр_t}^b = K_{зр_t}^b \cdot 100; \quad (9.4)$$

– ланцюговий

$$K_{зр_t}^l = \frac{y_t}{y_{t-1}}; \quad (9.5)$$

$$T_{зр_t}^l = K_{зр_t}^l \cdot 100. \quad (9.6)$$

**Коефіцієнт приросту**  $K_{пр}$  є вираженням відносної швидкості зростання і відображенням того, у скільки разів порівнюваний рівень є більшим або меншим від рівня, узятото за базу порівняння. Відсоткове подання коефіцієнта приросту називають **темпом приросту**  $T_{пр}$ :

– базисний

$$K_{пр} = \frac{\Delta_t^б}{y_1} = K_p^б - 1; \quad (9.7)$$

$$T_{пр} = 100K_{пр} = T_{зр}^б - 100; \quad (9.8)$$

– ланцюговий

$$K_{пр} = \frac{\Delta_t^л}{y_{t-1}} = K_p^л - 1; \quad (9.9)$$

$$T_{пр}^л = 100K_{пр} = T_p^л - 100. \quad (9.10)$$

**Абсолютне значення 1 % приросту**  $A_t$  відображає вагомість одного відсотка приросту й обчислюється як співвідношення абсолютного приросту й темпу приросту за один і той же період:

– базисний

$$A_t = \frac{\Delta_t^б}{T_{пр}^б} = \frac{y_1}{100}; \quad (9.11)$$

– ланцюговий

$$A_t = \frac{\Delta_t^к}{T_{пр}^к} = \frac{y_{t-1}}{100}. \quad (9.12)$$

**Співвідношення між ланцюговими й базисними показниками:**

1) ланцюгові й базисні абсолютні прирости є адитивно пов'язаними: сума ланцюгових приростів дорівнює базисному приросту за весь період ряду динаміки:

$$\Delta_t^б = \sum_1^n \Delta_t^л = \sum_1^n (y_t - y_{t-1}) = y_n - y_0;$$

2) ланцюгові й базисні коефіцієнти (темпи) зростання є мультиплікативно пов'язаними: добуток послідовних ланцюгових коефіцієнтів (темпів) зростання дорівнює кінцевому базисному коефіцієнту (темпу) зростання:

$$K_{зр}^б = \prod_1^n K_{зпт}^л = \frac{y_1}{y_0} \cdot \frac{y_2}{y_1} \cdot \dots \cdot \frac{y_n}{y_{n-1}} = \frac{y_n}{y_0}.$$

### **Взаємозв'язок абсолютної й відносної швидкостей динаміки**

Якщо порівнюваний рівень  $y_t$  виразити через рівень попереднього року плюс приріст ( $y_{t-1} + \Delta_t^n$ ) або через рівень базисного року плюс базисна абсолютна змінна ( $y_1 + \Delta_t^b$ ), отримаємо

$$K_{зр}^n = \frac{y_{t-1} + \Delta_t^n}{y_{t-1}} = 1 + \frac{\Delta_t^n}{y_{t-1}}, \text{ або } T_{зр}^n = 100\% + \frac{\Delta_t^n}{y_{t-1}} \cdot 100;$$

$$K_{зр}^b = \frac{y_1 + \Delta_t^b}{y_1} = 1 + \frac{\Delta_t^b}{y_1}, \text{ або } T_{зр}^b = 100\% + \frac{\Delta_t^b}{y_1} \cdot 100.$$

Якщо швидкість розвитку в межах періоду, що вивчається, є неоднаковою, то порівнянням однойменних характеристик швидкості визначають прискорення або 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 є однаковими, то порівнюють базисні характеристики швидкості, якщо неоднаковими – використовують середні абсолютні прирости:

$$\delta_t = \bar{\Delta}_t - \bar{\Delta}_{t-1}. \quad (9.13)$$

**Темп прискорення (уповільнення) динаміки** визначають, порівнюючи абсолютні прирости:  $K\Delta = \Delta_t / \Delta_{t-1}$ , а коли інтервали часу є неоднаковими, то застосовують формулу  $K\Delta = \bar{\Delta}_t - \bar{\Delta}_{t-1}$ .

**Коефіцієнт прискорення (уповільнення) відносної швидкості розвитку**  $\gamma_t$  дорівнює відношенню коефіцієнтів або темпів зростання (приросту):

$$\gamma_t = \frac{K_{зр}^1}{K_{зр}^2} \text{ або } \gamma_t = \frac{K_{пр}^1}{K_{пр}^2}. \quad (9.14)$$

Для порівняння інтенсивності динаміки в багатовимірних динамічних рядах однакового змісту по різних об'єктах (регіони, країни і т. ін.) або в рядах різного змісту по одному об'єкту використовують **коефіцієнт випередження**, який дорівнює відношенню базисних або середніх темпів зростання:

$$\frac{T'_{зр}}{T''_{зр}}.$$

### 9.3. Аналіз структурних зрушень

Структура будь-якої статистичної сукупності є динамічною. Змінюються склад і технічний рівень виробничих фондів, вікова структура робітників, склад і якість залучених до виробництва природних ресурсів, асортимент і якість продукції тощо. Змінення частин окремих складових сукупності – це наслідок **структурних зрушень** – змін у складі сукупності за певний період часу.

Структурні зрушення оцінюють з допомогою абсолютних і відносних характеристик динаміки:

1) абсолютного приросту  $j$ -ї частки у відсоткових пунктах (в.п.)

$$\Delta d_j = d_{j_1} - d_{j_0};$$

2) коефіцієнта (темпу) зростання  $j$ -ї частки

$$K_{зр_{d_j}} = d_{j_1} / d_{j_0}.$$

Характеристики структурних зрушень є взаємозв'язаними:

$$\Delta d_j = d_{j_0} (K_{зр_{d_j}} - 1).$$

Як узагальнювальні характеристики інтенсивності структурних зрушень застосовують лінійний  $\bar{l}_d$  і квадратичний  $\sigma_d$  коефіцієнти. Їх обчислюють на основі абсолютних приростів часток  $\Delta d$ :

$$\bar{l}_d = \frac{\sum_{j=1}^n |d_{j_1} - d_{j_0}|}{n}; \sigma_d = \sqrt{\frac{\sum_{j=1}^n (d_{j_1} - d_{j_0})^2}{n}}.$$

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

$$\sigma_{K_d} = \sqrt{\frac{\sum_{j=1}^n (d_{j_1} - d_{j_0})^2}{d_{j_0}}} = \sqrt{\sum_{j=1}^n d_{j_0} (K_{зр_{d_j}} - 1)^2}.$$

(якщо  $d_{j_1}$ ,  $d_{j_0}$  наведено не у відсотках, то результат необхідно помножити на 100).



#### 9.4. Середні показники динаміки

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

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

Середній рівень інтервального ряду з однаковими періодами часу – середня арифметична проста

$$\bar{Y} = \frac{1}{n} \sum_{t=1}^n y_t, \quad (9.15)$$

де  $n$  – кількість рівнів ряду.

Середній рівень інтервального або моментного динамічного ряду з неоднаковими періодами часу розраховують як середню арифметичну зважену величину:

$$\bar{Y} = \frac{\sum_{i=1}^n y_i t_i}{\sum_{i=1}^n t_i}, \quad (9.16)$$

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

Для наближеного оцінювання середнього рівня ряду за умови рівномірного змінення показника між датами середню величину обчислюють як напівсуму 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 однакові або приблизно однакові інтервали, то розрахунок виконують за формулою середньої хронологічної величини:

$$\bar{y} = \frac{\frac{y_1 + y_n}{2} + \sum_{t=2}^{n-1} y_t}{n-1}, \quad (9.17)$$

де  $n$  – кількість рівнів ряду динаміки.

**Середній абсолютний приріст** є відображенням того, на скільки одиниць у середньому змінився рівень порівняно з попереднім; його обчислюють як середню арифметичну просту величину з ланцюгових абсолютних приростів за певні періоди:

$$\bar{\Delta} = \frac{\sum_{t=1}^n \Delta_t^{\text{п}}}{n} = \frac{y_n - y_1}{n} = \frac{\Delta_t^{\text{б}}}{n}, \quad (9.18)$$

де  $n$  – кількість ланцюгових абсолютних приростів.

**Середній коефіцієнт (темп) зростання** – узагальнювальна характеристика індивідуальних коефіцієнтів (темнів) зростання ряду динаміки, яка обчислюється:

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

$$\bar{K}_{\text{зр}} = \sqrt[n]{K_{\text{зр}1}^{\text{п}} K_{\text{зр}2}^{\text{п}} \dots K_{\text{зр}n}^{\text{п}}} = \sqrt[n]{\prod_{t=1}^n K_{\text{зр}t}^{\text{п}}} = \sqrt[n]{K_{\text{зр}n}^{\text{б}}} = \sqrt[n]{\frac{y_n}{y_0}}, \quad (9.19)$$

$$\bar{T}_3 = \bar{K}_3 \cdot 100, \quad (9.20)$$

де  $n$  – кількість ланцюгових коефіцієнтів (темнів) зростання;

2) за формулою середньої геометричної зваженої з ланцюгових коефіцієнтів (темнів) зростання для рядів з **різними** інтервалами або проміжками часу між моментами

$$\bar{K}_{\text{зр}} = \sqrt[\sum t]{K_{\text{зр}1}^{t_1} K_{\text{зр}2}^{t_2} \dots K_{\text{зр}n}^{t_n}};$$

де  $t$  – інтервал часу, під час якого зберігався певний темп зростання;

$\sum t$  – сума проміжків часу періоду.

**Середній темп приросту** визначають як різницю між середнім темпом зростання і 100 %:

$$\bar{T}_{\text{пр}} = \bar{T}_{\text{зр}} - 100 \text{ \%}.$$

## 9.5. Характеристика основної тенденції розвитку

**Загальна тенденція (тренд)** – певна закономірність змінення рівнів, при цьому для одних рядів властивою є тенденція до збільшення, іншим – до зменшення рівнів.

У рядах динаміки з чітко виявленою тенденцією її описують аналітично у вигляді відповідної функції часу, яку називають рівнянням тренду:

$$Y_t = f(t),$$

де  $f(t)$  – основна тенденція, зумовлена впливом постійних чинників.

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

Теоретичні рівні  $Y_t$  визначають на основі так званої адекватної математичної функції, яка є оптимальним відображенням основної тенденції 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 властивими є майже постійні абсолютні прирости  $\Delta y^n \cong \text{const}$ , 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:

$$Y_t = a_0 + a_1 t,$$

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 збільшуються, а при  $a_1 < 0$  – так само рівномірно зменшуються.

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

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

$$Y_t = a_0 + a_1 t + a_2 t^2,$$

де  $t$ ,  $a_0$  і  $a_1$  – мають ті ж самі значення;  $a_2$  – константа тренду, що характеризує постійне змінення інтенсивності розвитку за одиницю часу. При  $a_2 > 0$  відбувається прискорення розвитку. При  $a_2 < 0$  процес зростання сповільнюється, причому з часом призводить не тільки до припинення зростання рівня, але й 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:

$$Y_t = a_0 + a_1 t + a_2 t^2 + a_3 t^3.$$

У рівнянні параметр  $a_3$  відображає змінення прискорення. При  $a_3 > 0$  розвиток прискорюється, а при  $a_3 < 0$  – уповільнюється.

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

$$Y_t = a_0 a_1^t.$$

Якщо  $a_1 > 1$ , то тренд виражає тенденцію прискорюваного зростання рівнів, наприклад: розмноження організмів, коли немає 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 є такою:

$$\sum_{t=1}^n (y_t - Y_t)^2 = \sum_{t=1}^n (y_t - (a_0 + a_1 t))^2 = \min.$$

Параметри трендових кривих обчислюють, розв'язуючи системи нормальних рівнянь. Для лінійної функції її записують так:

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

Якщо відлік часу  $t$  перенести в середину динамічного ряду, то його значення, розміщені вище за середину, будуть від'ємними, а нижче – додатними, тоді  $\sum t = 0$ . При непарній кількості членів ряду (наприклад  $n = 5$ ),  $t = -2, -1, 0, 1, 2$ ; при парному (наприклад  $n = 6$ )  $t = -5, -3, -1, 1, 3, 5$ . В обох випадках система рівнянь спрощується, після чого параметри рівняння набувають вигляду

$$a_0 = \frac{\sum y}{n} = \bar{y}; \quad a_1 = \frac{\sum yt}{\sum t^2}. \quad (9.21)$$

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

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

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

## 9.6. Аналіз 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, зумовлені специфікою виробництва або споживанням певного виду продукції. Їх виявляють або аналізують на основі рядів щомісячних або щоквартальних даних і описують сезонною хвилею.

Для вимірювання сезонних коливань обчислюють індекси сезонності.

**Індекс сезонності**  $I_s$  – у загальному вигляді є відсотковим відношенням однойменних (місячних, квартальних) фактичних рівнів ряду динаміки  $y_i$  до їх середньорічних або теоретичних рівнів  $Y_t = f(t)$  (якщо ряд виявляє тенденцію):

$$I_{s_i} = \frac{y_i}{Y_{t_i}} 100 \%$$

З 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 усереднення індивідуальних індексів однойменних внутрішньорічних періодів аналізованого ряду динаміки. Тому для кожного періоду річного циклу визначають узагальнені показники у вигляді середніх індексів сезонності:

$$\bar{I}_{s_i} = \frac{\sum_{i=1}^n I_{s_i}}{n}, \quad (9.22)$$

де  $n$  – кількість років.

Обчислені за формулою (9.22) середні індекси сезонності не залежать від впливу основної тенденції й випадкових відхилень.

Індекс сезонності можна розрахувати різними методами:

1) для рядів внутрішньорічної динаміки зі значно вираженою основною тенденцією розвитку (ланцюгові та базисні темпи зростання є високими) – **методом змінної середньої**:

$$\dot{I}_{S_i} = \frac{\sum Y_i}{Y_{t_i}} \cdot 100 \%,$$

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

2) для рядів внутрішньорічної динаміки, у яких річні рівні є стабільними, тенденції немає або вона є незначною – **методом постійної (простой) середньої**:

$$\dot{I}_{S_i} = \frac{\bar{y}_i}{\bar{y}} \cdot 100 \%,$$

де  $\bar{y}_i = \frac{\sum_{j=1}^n y_{ij}}{n}$  – середні місячні або квартальні рівні ряду, обчислені за рівнями  $y_{ij}$  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 розраховують за формулою

$$\dot{I}_{S_i} = \frac{y_i}{\bar{y}} \cdot 100 \%, \quad (9.23)$$

де  $y_i$  – місячні (квартальні) рівні.

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

1) амплітуда сезонних коливань (вимірюються у в.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

$$\sigma_t = \sqrt{\frac{1}{12} \sum_1^{12} (I_S - 100)^2}. \quad (9.24)$$

## 9.7. Екстраполяція в рядах динаміки та прогнозування

Визначувані в аналізі рядів динаміки показники змінення рівнів, тренду, сезонної хвилі широко використовують при **прогнозуванні** – статистичному оцінюванні можливої міри розвитку соціально-економічних явищ на майбутнє.

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

**Інтерполяція** – визначення невідомого показника в середині ряду динаміки.

**Екстраполяція** – поширення 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 наприкінці або на початку динамічного ряду.

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

Установлення термінів прогнозування  $m$  залежить від завдання дослідження. Однак слід мати на увазі, що чим коротшими є терміни попередження прогнозу, тим надійнішими будуть результати екстраполяції.

Уживання методів екстраполяції залежить від характеру змін в базисному ряді динаміки й зумовлюється постановкою завдання дослідження:

1. При екстраполяції рівнів розвитку явища, що вивчається, на базі ряду динаміки з постійними абсолютними приростами ( $\Delta y^n \cong \text{const}$ ) застосовують формулу

$$y_{n+m} = y_n + \Delta y m, \quad (9.25)$$

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

2. На базі ряду динаміки зі стабільними темпами зростання ( $\Delta T p^n \cong \text{const}$ ) невідомі рівні обчислюють за формулою

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

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

## Приклади розв'язання типових задач

### Задача 9.1

Випуск промислових роботів в об'єднанні характеризується даними, наведеними в табл. 9.1. Визначити:

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

2) середньорічні коефіцієнти й темпи зростання, абсолютні прирости, абсолютне прискорення (уповільнення) зростання за 2015–2020 і 2020-2022 роки;

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

Таблиця 9.1

Рік	2015	2020	2022
Кількість, шт.	60	114	126

### Розв'язання

Абсолютний приріст обчислимо за формулами (9.1), (9.2), коефіцієнти й темпи зростання – за формулами (9.3)–(9.6), абсолютне значення 1 % приросту – за формулами (9.11), (9.12). Результати розрахунків зведемо в табл. 9.2.

Таблиця 9.2

Показники динаміки	2015–2020 рр.	2020–2022 рр.	За весь період
Абсолютний приріст	$\Delta_t^n = 114 - 60 = 54$	$\Delta_t^n = 126 - 114 = 12$	$\Delta_n^6 = 126 - 60 = 66$
Коефіцієнт зростання	$Kp_t^n = 114 / 60 = 1,9$	$Kp_t^n = 126 / 114 = 1,1$	$Kp_n^6 = 126 / 60 = 2,1$
Темп зростання	$Tr_t^n = Kp_t^n \cdot 100 = 190$	$Tr_t^n = Kp_t^n \cdot 100 = 110$	$Tr_n^6 = Kp_n^6 \cdot 100 = 210$
Темп приросту	$T_{np}^n = 190 - 100 = 90 \%$	$T_{np}^n = 110 - 100 = 10 \%$	$T_{np}^6 = 210 - 100 = 110 \%$
Абсолютне значення 1 % приросту	$A_t = 60 / 100 = 0,6$	$A_t = 114 / 100 = 1,14$	$A_t = 60 / 100 = 0,6$
Середньорічний коефіцієнт зростання	$\bar{K}_p = \sqrt[5]{\frac{114}{60}} = 1,137$	$\bar{K}_p = \sqrt{\frac{126}{114}} = 1,049$	–
Середньорічний темп зростання	$\bar{T}_p = 1,137 \cdot 100 = 113,7$	$\bar{T}_p = \bar{K}_p \cdot 100 \% = 1,049 \cdot 100 = 104,9$	–
Середній темп приросту	$\bar{T}_{np} = 113,7 - 100 \% = 13,7 \%$	$\bar{T}_{np} = \bar{T}_p - 100 \% = 4,9 \%$	–

Середньорічні коефіцієнти й темпи зростання обчислимо за формулами (9.19), (9.20), абсолютне прискорення (уповільнення)



зростання за 2015–2020 і 2020–2022 роки – за формулою (9.13), коефіцієнт прискорення (уповільнення) зростання – за формулою (9.14).

Абсолютне прискорення (уповільнення)  $\delta_t = 12 - 54 = -42$ . Значення  $-42$  свідчить про сповільнення зростання.

Коефіцієнт прискорення зростання  $\gamma_t = 1,1 / 1,9 = 0,58$ .

### Задача 9.2

Є дані про заборгованість по кредитах районному відділенню банку (табл. 9.3).

Таблиця 9.3

Дата	Залишки заборгованості, тис. грн	Загальна заборгованість, тис. грн
01.01.2022	12	500
01.02.2022	10	650
01.03.2022	10	600
01.04.2022	8	500

1. Визначити середній залишок простроченої заборгованості за перший квартал 2022 року.

2. Обчислити середній залишок заборгованості по всіх кредитах за перший квартал 2022 року.

#### Розв'язання

1. Середній залишок простроченої заборгованості, використавши середню хронологічну для моментного ряду, визначимо за формулою (9.17):

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

2. Аналогічно обчислимо середній залишок заборгованості за всіма кредитами:

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

### Задача 9.3

За даними виробництва товарів народного споживання в регіоні (табл. 9.4) описати тенденцію ряду лінійним трендом, навести економічну інтерпретацію параметрів трендового рівняння. Зробити прогноз на 2022 рік, припускаючи, що умови формування тенденції виробництва товарів не зміняться.

Таблиця 9.4

Рік	2014	2015	2016	2017	2018	2019	2020
Виробництво товарів, тис. т	68	74	77	80	86	94	97

### Розв'язання

Параметри трендового рівняння визначимо за формулою (9.21). Розрахунок параметрів наведено в табл. 9.5.

Таблиця 9.5

Рік	Y	t	t <sup>2</sup>	y <sub>t</sub>	Y <sub>t</sub> = 83 + 5t
2014	68	-3	9	-204	
2015	74	-2	4	-148	
2016	77	-1	1	-77	
2017	82	0	0	0	
2018	89	1	1	89	
2019	94	2	4	188	
2020	97	3	9	291	98
Усього	581	0	28	139	581

Тоді за формулами (9.21)

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

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 наведено теоретичні рівні Y<sub>t</sub>, тобто очікувані рівні обсягу виробництва товарів народного споживання в t-му році, зумовлені впливом основних причин розвитку виробництва.

Зробимо прогноз на 2022 рік, припускаючи, що умови формування тенденції виробництва товарів не зміняться. Базою прогнозу є теоретичний рівень 2020 року, період випередження m = 2. Тоді очікуваний 2022 року рівень виробництва товарів Y<sub>t+m</sub> = 98 + 5 · 2 = 107.

### Приклад 9.4

За наведеними в табл. 9.6 даними про споживання електроенергії комунальним господарством району визначити сезонну хвилю на базі постійної середньої величини й оцінити сезонність з допомогою амплітуди коливань і середнього квадратичного відхилення.

Таблиця 9.6

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 за i-й місяць за формулою середньої арифметичної простої величини можна визначити так:

$$\bar{y} = \frac{\sum_{i=1}^{12} y_i}{12} = \frac{1848}{12} = 154 \text{ млн кВт·год.}$$

Сезонна хвиля – це сукупність помісячних індексів сезонності. Маємо дані за один рік, тому індекси сезонності за весь період обчислимо за формулою

$$I_{s_i} = \frac{y_i}{\bar{y}} 100 \text{ \%}.$$

Розрахунок індексів сезонності зведемо в табл. 9.7.

Таблиця 9.7

Місяць року	Спожито електроенергії $y_t$ , млн кВт·год	Індекс сезонності $I_s$ , %	$I_s - 100$	$(I_s - 100)^2$
1	172	$172 / 154 \cdot 100 = 111,7$	11,7	136,6
2	161	$161 / 154 \cdot 100 = 104,5$	4,5	20,7
3	158	102,6	2,6	6,7
4	151	98,1	-1,9	3,8
5	147	95,5	-4,5	20,7
6	130	84,4	-15,6	242,9
7	124	80,5	-19,5	379,5
8	146	94,8	-5,2	27,0
9	149	96,8	-3,2	10,5
10	155	100,6	0,6	0,4
11	168	109,1	9,1	82,6
12	187	121,4	21,4	459,2
Разом	1848	–	–	1390,6

Амплітуда сезонних коливань

$$R_t = 121,4 - 80,5 = 40,9 \text{ в.п.}$$

Згідно з розрахунками середнє квадратичне відхилення обчислимо за формулою

$$\sigma_t = \sqrt{\frac{1}{12} \sum_1^{12} (I_{\text{сез}} - 100)^2} = \sqrt{\frac{1390,6}{12}} = 10,76 \text{ \%}.$$

Характер сезонної хвилі схематично зображено на рис. 9.1.

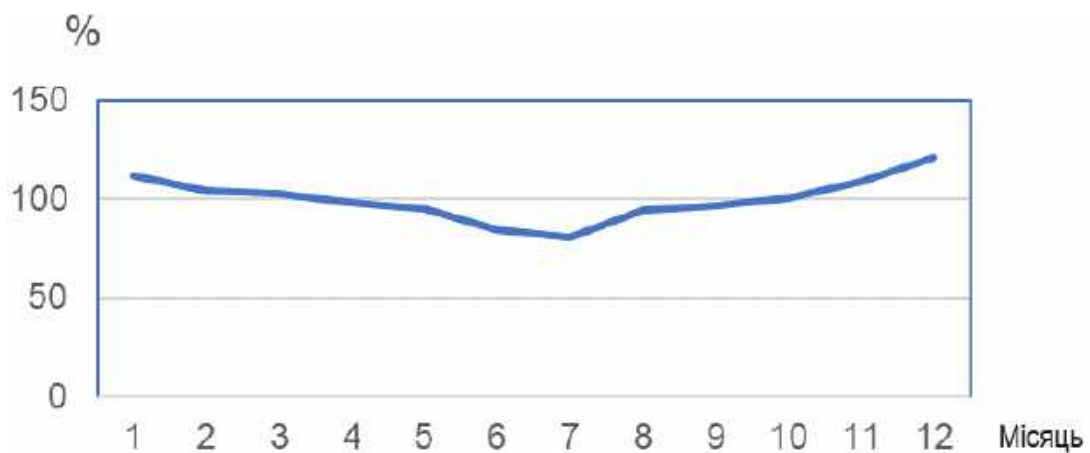


Рис. 9.1. Сезонна хвиля споживання електроенергії комунальним господарством району

### Завдання для самостійної роботи

1. Обсяг реалізації платних послуг на одного мешканця регіону характеризується такими даними:

Рік	2015	2016	2017	2018	2019	2020
Разом, грн	208,1	223,5	237,5	274,6	285,5	323,9

Для аналізу динаміки платних послуг обчислити:

- 1) абсолютні прирости, темпи зростання й темпи приросту за роками й за весь період, абсолютний вміст одного відсотка приросту (отримані показники подайте у вигляді таблиці);
- 2) середні показники динаміки.

2. Кількість працівників заводу з початку місяця першого півріччя становило:

Показник	Дата						
	01.01	01.02	01.03	01.04	01.05	01.06	01.07
Кількість працівників	860	862	866	874	868	870	874

Обчислити:

- 1) середньомісячну кількість працівників у I й II кварталах;
- 2) абсолютний приріст кількості у II кварталі відносно I.

3. Динаміка капітальних вкладень характеризується такими даними в порівнянних цінах, млрд грн:

Показник	Рік				
	1-й	2-й	3-й	4-й	5-й
Капітальні вкладення всього	136,95	112,05	84,66	74,7	62,3
У тому числі:					
виробничого призначення	97,35	79,65	60,18	53,10	41,40
невиробничого призначення	39,6	32,4	24,48	21,6	20,9

Для вивчення інтенсивності змінення обсягу капітальних вкладень обчислити:

1) абсолютні прирости, темпи зростання й приросту загального обсягу капітальних вкладень (результати подати у вигляді таблиці);

2) для загального обсягу капітальних вкладень, у тому числі виробничого й невиробничого призначення, середній рівень ряду динаміки, середньорічний темп зростання й приросту.

**4.** Прямі іноземні інвестиції у розвиток металургії становили, млн дол.:

Рік	Чорна металургія	Кольорова металургія
2019	7,5	10,0
2020	12,3	14,2
2021	24,3	26,8

Для кожної галузі визначити базові й ланцюгові абсолютні прирости, оцінити прискорення процесу інвестування металургії.

**5.** Динаміка заборгованості промислових підприємств регіону за короткостроковими кредитами характеризується такими даними:

Галузь промисловості	Заборгованість на 01.01, млн грн	Абсолютне збільшення заборгованості за місяць, млн грн		
		Січень	Лютий	Березень
А	320	48	60	84
В	510	184	118	106
С	170	18	9	7

Для кожної галузі визначити:

а) абсолютний приріст заборгованості за перший квартал та розмір заборгованості на 1 квітня;

б) темп приросту заборгованості за перший квартал.

**6.** Станом на 01.01 поточного року у списках установи було 286 осіб. У січні відбулися такі зміни в обліковому складі працівників:

- вибуло з 05.05 – 14 осіб;
- зараховано з 12.01 – 6 осіб;
- зараховано з 20.01 – 2 особи;
- вибуло з 25.01 – 3 особи.

Визначити середньооблікову кількість працівників установи за січень поточного року.

**7.** Протягом трьох років рівень інфляції становив 160, 110 та 70 %. Визначити абсолютне прискорення й коефіцієнт прискорення.

**8.** Випуск продукції підприємством за кварталами року характеризується такими даними:

- 1-й квартал – збільшився на 5 %;

- 2-й квартал – збільшився на 3,5 % (порівняно з I кварталом);
- 3-й квартал – зменшився на 1 % (порівняно з II кварталом);
- 4-й квартал – збільшився на 2,5 % (порівняно з III кварталом).

Визначити, на скільки відсотків змінювався випуск продукції поквартально.

**9.** Планом технологічного розвитку на 2018–2022 рр. передбачено збільшення виробництва в порівняних цінах в 1,3 раза. Визначити, на скільки відсотків має збільшуватися випуск продукції щорічно.

**10.** Показники зовнішньоекономічної діяльності країни, млн дол., характеризуються такими даними:

Показник	Рік						
	2015	2016	2017	2018	2019	2020	2021
Експорт	18970,3	16457,2	15195,5	18059,3	19809,1	22012,4	27328,2
Імпорт	18546,3	16106,5	12959,3	15106,9	16922,6	18168,6	24409,2

Порівняти показники темпів збільшення експорту й імпорту і визначити коефіцієнт випередження.

### Контрольні запитання

1. Що називають рядом динаміки?
2. З яких елементів складаються ряди динаміки? Що вони виражають?
3. Яких умов необхідно дотримуватися під час будування рядів динаміки?
4. За наявності яких причин виникає непорівнюваність рівнів ряду динаміки?
5. Які ви знаєте види рядів динаміки?
6. Які ряди динаміки називають моментними? Чому їхні рівні не можна підсумовувати?
7. Чому ряди динаміки, виражені абсолютними величинами, є первинними, а відносними та середніми величинами – вторинними?
8. Як розраховують середній рівень для інтервального й моментного рядів динаміки?
9. Як визначають абсолютний приріст, темпи зростання й 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 оброблення й аналізу рядів динаміки ви знаєте?
13. Що таке інтерполяція та екстраполяція рядів динаміки? Якими є їх значення й застосування?

## 10. СТАТИСТИЧНІ ІНДЕКСИ

### 10.1. Суть і функції індексів

**Статистичний індекс** – це відносна величина, що характеризує змінення соціально-економічного явища в часі, просторі або ступінь відхилення значення показника від певного стандарту (нормативу, середньої). Індекс є показником того, у скільки разів рівень явища, що вивчається, у певних умовах відрізняється від рівня того ж явища в інших умовах.

Форми вираження індексів – коефіцієнти, відсотки, проміле.

За характером досліджуваних об'єктів розрізняють:

– **індекси об'ємних показників**, які характеризують змінення обсягу того чи іншого явища, що вимірюється в певних одиницях (індекси фізичного обсягу продукції, роздрібного товарообігу, споживання продуктів тощо);

– **індекси якісних показників**, які характеризують змінення якісної ознаки (індекси цін, собівартості продукції, продуктивності праці тощо).

За ступенем охоплення одиниць сукупності індекси поділяють на такі:

– **індивідуальні  $i$** , що характеризують змінення окремих елементів складної сукупності або однорідних груп;

– **зведені або загальні  $I$** , що характеризують змінення сукупності, до якої належать різні елементи. Такими елементами є окремі продукти в 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** й **базисні індекси**.

Залежно від методології обчислення загальні індекси поділяють на **агрегатні** й **середні з індивідуальних індексів**.

За характером та умовами порівнянь (час, простір, стандарт) індекси поділяють на такі:

– **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 є довільним;

– **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 аргументується тією задачею, для якої будується індекс. При цьому дотримуються двох правил:

1) база порівняння являє собою стабільний рівень;

2) база порівняння як екстремальний рівень – найвищий або найнижчий рівень (у разі зменшення економічних показників).

Індекс поєднує в собі такі аспекти:

– якісний, у якому назва відображає соціально-економічний зміст показника;

– кількісний, у якому числове значення відображає інтенсивність змінення або ступінь відхилення.

Кожний індекс містить два види даних: оцінювані дані, які називають поточними (звітними) (позначають значком «1»), і дані, що використовують як базу 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;  $y$  – урожайність культури з 1 га;  $n$  – розмір посівної площі.

## 10.2. Індивідуальні індекси

**Індивідуальним індексом**  $i$  називають відносну величину, яка характеризує змінення окремих елементів складної сукупності або однорідних груп.

За індивідуальні індекси беруть відносну величину динаміки або порівняння.

Методологія будування індивідуальних індексів є простою.

Рівень товарообігу  $Q_1$  в умовах звітного року розраховуємо у вигляді суми виторгу від продажу товару:  $Q_1 = p_1q_1$ . Аналогічно рівень товарообігу базисного року набуває вигляду  $Q_0 = p_0q_0$ . Тоді **індивідуальний індекс товарообігу**

$$i_Q = \frac{Q_1}{Q_0}. \quad (10.1)$$

Аналогічні індивідуальні індекси можна розрахувати для будь-якого показника:

– **індивідуальний індекс ціни**, який є відображенням того, у скільки разів змінилася загальна сума виторгу під впливом змінення ціни товару:

$$i_p = \frac{p_1}{p_0}; \quad (10.2)$$

– **індивідуальний індекс кількості проданих товарів**, який є відображенням того, у скільки разів збільшилася (або зменшилася) загальна сума виторгу під впливом змінення обсягу продажу в натуральних одиницях:



$$i_q = \frac{q_1}{q_0}. \quad (10.3)$$

Очевидно, що

$$i_Q = \frac{Q_1}{Q_0} = \frac{q_1 p_1}{q_0 p_0} = i_q i_p, \text{ або } Q_1 = Q_0 i_q i_p.$$

Друга формула являє собою індексну мультиплікативну модель підсумкового показника – обсягу товарообігу, з допомогою якої знаходять приріст підсумку під впливом кожного чинника окремо.

### 10.3. Загальні індекси

Якщо відомо, що явище, яке вивчається, є неоднорідним і порівняння рівнів можна провести тільки після зведення їх до загальної міри, економічний аналіз виконують з допомогою загальних індексів.

**Загальний індекс**  $I$  характеризує змінення сукупності, до якої належать різні елементи.

Важливою особливістю загальних індексів є те, що вони виконують дві взаємозв'язані функції:

- 1) синтетичну – узагальнювальну характеристику змінення явища;
- 2) аналітичну – вивчення дії окремих чинників на змінення явища.

Основною формою загальних індексів є агрегатні індекси.

У складних статистичних сукупностях порівняння різнірідних одиниць досягають уведенням в індексні відношення спеціальних додаткових показників – співмножників (ваг), які економічно щільно пов'язані з індексованою величиною. Таким співмножником може бути ціна, собівартість або трудомісткість одиниці продукції. Перемноживши обсяг продукції кожного виду на відповідний співмножник, отримують показники, які можна підсумувати.

Сума добутків кількості продукції  $q_j$  на його співмножник, наприклад ціну  $p_j$ , створює з'єднання – так звані агрегати  $\sum q_j p_j$  (від лат. *aggregare* – приєднувати).

**Агрегат** – добуток величин, одна з яких має назву індексованої, інша – співмножника (ваги).

**Агрегатні індекси** розраховують як відношення агрегатів, побудованих для різних умов.

Індексована величина, змінення якої характеризує певний індекс, у чисельнику й знаменнику агрегата знаходиться в різних періодах – звітне значення зіставляється з базисним.

Співмножник можна фіксувати на рівні як базисного, так і поточного періодів. Залежно від цього розглядаються формули індексів цін і фізичного обсягу в різних системах вагів, наведених у табл. 10.1.

Таблиця 10.1

Базисно-вагова система (індекси Ласпейреса)	Поточно-вагова система (індекси Пааше)
$I_p = \frac{\sum p_1 q_0}{\sum p_0 q_0}$	$I_p = \frac{\sum p_1 q_1}{\sum p_0 q_1}$
$I_q = \frac{\sum q_1 p_0}{\sum q_0 p_0}$	$I_q = \frac{\sum q_1 p_1}{\sum q_0 p_1}$

Вибір форми індексів залежить від мети дослідження й наявної інформації. Обидві системи індексів є рівноправними.

Під час синтезу загального індексу цін замість фактичної кількості товарів (у звітний або базисний періоди) як співмножник можуть застосовуватися середні величини реалізації товарів за два або більше періодів. При такому способі розрахунку (*індекс Лоу*) формула загального індексу синтезується у вигляді

$$I_p = \frac{\sum p_1 \bar{q}}{\sum p_0 \bar{q}},$$

де  $\bar{q}$  – середня кількість товарів, реалізованих за аналізований період.

**Індекс товарообігу** у фактичних цінах в агрегатній формі

$$I_Q = \frac{\sum q_1 p_1}{\sum q_0 p_0}, \quad (10.4)$$

де  $n$  – кількість товарів;  $p_1 q_1$ ,  $p_0 q_0$  – товарообіг окремих видів товарів у базисному та звітному періодах.

#### 10.4. 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. 10.2).

У базисно-ваговій системі загальні індекси обчислюють як середню арифметичну з вагами  $q_0 p_0$ . У поточно-ваговій системі загальні індекси обчислюють як середню гармонійну з вагами  $q_1 p_1$ .

Таблиця 10.2

Базисно-вагова система (середня арифметична форма)	Поточно-вагова система (середня гармонійна форма)
$I_q = \frac{\sum i_q q_0 p_0}{\sum q_0 p_0}$	$I_p = \frac{\sum p_1 q_1}{\sum \frac{1}{i_p} p_1 q_1}$
$I_p = \frac{\sum i_p q_0 p_0}{\sum q_0 p_0}$	$I_q = \frac{\sum p_1 q_1}{\sum \frac{1}{i_q} p_1 q_1}$
$I_Q = \frac{\sum i_Q p_0 q_0}{\sum q_0 p_0}$	$I_Q = \frac{\sum p_1 q_1}{\sum \frac{1}{i_Q} p_1 q_1}$

При будівництві середньозважених індексів вартісні ваги можна замінити відносними величинами структури  $d$ , сума яких  $\sum d = 1$ . У цьому випадку середньозважені індекси набувають вигляду

$$I_q = \sum i_q d_0, \quad I_p = \frac{1}{\sum \frac{1}{i_p} d_1}.$$

### 10.5. Взаємозв'язок індексів та індексні системи

Майже кожний індекс є складовою певної індексної системи, а його зв'язки з іншими індексами цієї системи відображають зв'язки між відповідними показниками. Так, товарообіг у поточному періоді порівняно з базовим збільшується (зменшується) під впливом двох чинників – фізичного обсягу проданого товару  $q$  і ціни  $p$ . 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.

Індекси прямих і зворотних показників, наприклад споживацьких цін і купівельної спроможності грошової одиниці, продуктивності роботи й трудомісткості продукції є взаємозв'язані  $I_p$ :

$$I_{1/t} = \frac{1}{I_t}; \quad I_{1/p} = \frac{1}{I_p}.$$

Слід 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

$$\frac{\sum q_1 p_0}{\sum q_0 p_0} \cdot \frac{\sum p_1 q_1}{\sum p_0 q_1} = \frac{\sum p_1 q_1}{\sum p_0 q_0}, \quad \text{або} \quad \frac{\sum q_1 p_1}{\sum q_0 p_1} \cdot \frac{\sum p_1 q_0}{\sum p_0 q_0} = \frac{\sum p_1 q_1}{\sum p_0 q_0}.$$

Таким чином, загальні індекси цін  $I_p$  і товарної маси  $I_q$ , маючи самостійне значення, одночасно виконують аналітичну функцію, тобто з їх допомогою оцінюють вплив відповідного фактора на динаміку обороту.

У межах індексної системи можна виявити абсолютний вплив факторів на приріст результату у вигляді **абсолютного приросту**, який є різницею між чисельником і знаменником відповідних індексів. Абсолютний вплив факторів на приріст товарообігу:

$$\Delta Q = \sum q_1 p_1 - \sum q_0 p_0 = \Delta q + \Delta p.$$

Абсолютний вплив кожного фактора окремо в системі буде таким:

$$\Delta q = \sum q_1 p_0 - \sum q_0 p_0 = \sum (q_1 - q_0) p_0,$$

$$\Delta p = \sum p_1 q_1 - \sum p_0 q_1 = \sum (p_1 - p_0) q_1.$$

Наведемо формули розрахунку деяких самих найспоживаніших агрегатних індексів і їх взаємозв'язок. Слід 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).

Індекс змінення загальної суми витрат на виробництво продукції залежно від обсягу виробництва  $q$  і собівартості  $z$

$$I_c = \frac{\sum z_1 q_1}{\sum z_0 q_0} = \frac{\sum z_0 q_1}{\sum z_0 q_0} \cdot \frac{\sum z_1 q_1}{\sum z_0 q_1} = \frac{\sum z_1 q_1}{\sum z_1 q_0} \cdot \frac{\sum z_1 q_0}{\sum z_0 q_0} = I_z I_q.$$

Індекс змінення обсягу продукції  $Q$  у зв'язку зі зміненням кількості працівників  $t$  і рівня їх продуктивності праці  $w$

$$I_Q = \frac{\sum w_1 t_1}{\sum w_0 t_0} = \frac{\sum w_0 t_1}{\sum w_0 t_0} \cdot \frac{\sum w_1 t_1}{\sum w_0 t_1} = \frac{\sum w_1 t_1}{\sum w_1 t_0} \cdot \frac{\sum w_1 t_0}{\sum w_0 t_0} = I_w I_t.$$

Індекс змінення загального фонду оплати  $\Phi$  праці у зв'язку зі зміненням загальної кількості працівників  $t$  і заробітної плати  $f$

$$I_\Phi = \frac{\sum f_1 t_1}{\sum f_0 t_0} = \frac{\sum f_0 t_1}{\sum f_0 t_0} \cdot \frac{\sum f_1 t_1}{\sum f_0 t_1} = \frac{\sum f_1 t_1}{\sum f_1 t_0} \cdot \frac{\sum f_1 t_0}{\sum f_0 t_0} = I_f I_t.$$

Аналогічно обчислюють загальні агрегатні індекси і для багатьох інших економічних показників.

## 10.6. 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 кількість окремих варіантів сукупності – структурні зрушення.

Вплив кожного з чинників на динаміку середньої оцінюють з допомогою системи індексів середніх величин – змінного й фіксованого та системи індексів структурних зрушень.

**Індексом змінного складу** називають індекс середньої величини, який відображає змінення індексованої величини  $x$  і структури явища, яку вивчають:

$$I_x = \frac{\bar{x}_1}{\bar{x}_0} = \frac{\sum_{i=1}^m x_{i1} f_{i1}}{\sum_{i=1}^m f_{i1}} : \frac{\sum_{i=1}^m x_{i0} f_{i0}}{\sum_{i=1}^m f_{i0}}. \quad (10.5)$$

Якщо в початкових формулах загальних індексів показники ваг  $f_i$  замінити

відносними, наприклад питомою вагою  $d = \frac{f_i}{\sum f}$ , то отримуємо  $I_x = \frac{\sum_{i=1}^m x_{i1} d_{i1}}{\sum_{i=1}^m x_{i0} d_{i0}}$ .

Наведена форма запису індексів має деяку перевагу над розгорненою, тому її частіше застосовують на практиці. Вона є компактною, отже, чіткіше виокремлює чинники, які впливають на змінення середньої величини показника.

Якщо за величину  $x$  узяти ціну  $p$ , то отримуємо індекс цін

$$I_p = \frac{\bar{p}_1}{\bar{p}_0} = \frac{\sum_{i=1}^m p_{i1} q_{i1}}{\sum_{i=1}^m q_{i1}} : \frac{\sum_{i=1}^m p_{i0} q_{i0}}{\sum_{i=1}^m q_{i0}}.$$

Індекс змінного складу розкладається на два субіндекси:

– **індекс фіксованого складу**  $I_x$  відображає середнє змінення значення показника при незмінній, фіксованій структурі сукупності:

$$I_x = \frac{\sum_{i=1}^m x_{i1} f_{i1}}{\sum_{i=1}^m f_{i1}} : \frac{\sum_{i=1}^m x_{i0} f_{i1}}{\sum_{i=1}^m f_{i1}} = \frac{\sum_{i=1}^m x_{i1} f_{i1}}{\sum_{i=1}^m x_{i0} f_{i1}} \text{ або } I_x = \sum_{i=1}^m x_{i1} d_{i1} : \sum_{i=1}^m x_{i0} d_{i1}; \quad (10.6)$$

$$I_p = \frac{\sum_{i=1}^m p_{i1} q_{i1}}{\sum_{i=1}^m q_{i1}} : \frac{\sum_{i=1}^m p_{i0} q_{i1}}{\sum_{i=1}^m q_{i1}} = \frac{\sum_{i=1}^m p_{i1} q_{i1}}{\sum_{i=1}^m p_{i0} q_{i1}};$$

– **індекс структурних зрушень**  $I_d$  відображає змінення середньої внаслідок структурних зрушень. Значення показника  $x$  фіксується на рівні базисного періоду:

$$I_d = \frac{\sum_{i=1}^m x_{i0} f_{i1}}{\sum_{i=1}^m f_{i1}} : \frac{\sum_{i=1}^m x_{i0} f_{i0}}{\sum_{i=1}^m f_{i0}} \text{ або } I_d = \sum_{i=1}^m x_{i0} d_{i1} : \sum_{i=1}^m x_{i0} d_{i0};$$

$$I_d = \frac{\sum_{i=1}^m p_{i0} q_{i1}}{\sum_{i=1}^m q_{i1}} : \frac{\sum_{i=1}^m p_{i0} q_{i0}}{\sum_{i=1}^m q_{i0}} \text{ або } I_d = \sum_{i=1}^m x_{i0} d_{i1} : \sum_{i=1}^m x_{i0} d_{i0}. \quad (10.7)$$

Формули індексів фіксованого складу й структурних зрушень є різноваговими: в  $I_x$  вагу фіксують на рівні поточного періоду, в  $I_d$  – значення показника  $x$  фіксують на рівні базисного періоду. Саме такий варіант зважування забезпечує зв'язок цих індексів у системі:

$$I_{\bar{x}} = I_x I_d. \quad (10.8)$$

У межах індексної системи можна виявити абсолютний приріст **середньої з урахуванням кожного фактора**.

### 10.7. 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 та об'єктів. Їх особливість – рівноправність порівнюваних об'єктів А і В. Кожен з них може бути базою порівняння. Отже, рівноправні індекси як з базою А, так і з базою В будуть такими:

$$I_{x, \frac{A}{B}} = \frac{\sum x_A f_A}{\sum x_B f_A}; \quad I_{x, \frac{B}{A}} = \frac{\sum x_B f_B}{\sum x_A f_B}, \quad (10.9)$$

де  $x$  – індексована величина,  $f$  – вага індексованої величини.

При фіксованих значеннях вагів індекси  $I_A$  й  $I_B$  є обернено пропорційними.

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:

$$I_{Q, \frac{A}{B}} = \frac{\sum p_A q_A}{\sum p_B q_B}. \quad (10.10)$$

Для характеристики співвідношення рівнів цін на товари, реалізовані в районі А порівняно з районом В, розраховують загальний індекс цін, у якому за вагу індексованих величин  $p_A$  і  $p_B$  беруть кількість товарів, реалізованих у районі А:

$$I_{p_{A/B}} = \frac{\sum q_A p_A}{\sum q_A p_B} \quad (10.11)$$

або в районі В:

$$I_{p_{B/A}} = \frac{\sum q_{Bp_B}}{\sum q_{Bp_A}}. \quad (10.12)$$

Аналогічно будуються індекси обсягів реалізованих товарів.

**Територіальний індекс змінного складу по об'єктах А і В** обчислюють за формулою

$$I_x = \frac{\sum x_A f_A}{\sum f_A} : \frac{\sum x_B f_B}{\sum f_B} = \sum x_A d_A : \sum x_B d_B.$$

Цей індекс є відображенням того, у скільки разів середній рівень ознаки об'єкта А є більшим або меншим від середнього рівня ознаки 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  $I_x$  відображає співвідношення середніх значень ознаки при фіксованій структурі сукупності.

### Приклади розв'язання типових задач

#### Задача 10.1

Розглянемо групу з двох підприємств, що виробляють різну продукцію. Кожне підприємство має дані за два суміжних роки (базисний і звітний) про кількість працівників і середню продуктивність праці на одну людину (табл. 10.3).

Таблиця 10.3

Номер підприємства	Базисний рік		Звітний рік	
	Продуктивність праці $w_0$ , тис. грн на 1 людину	Середня кількість працівників $t_0$ , чол.	Продуктивність праці $w_1$ , тис. грн на 1 людину	Середня кількість працівників $t_1$ , 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 змінення обсягу виробленої продукції, кількості працівників; продуктивності праці;
- 2) загальні індекси обсягу виробленої продукції, змінення кількості працівників, продуктивності праці; навести взаємозв'язок загальних індексів;



3) загальні індекси обсягу виробленої продукції, змінення кількості працівників, продуктивності праці як середньозважені індекси;

4) абсолютний приріст обсягу виробленої продукції в цілому й з урахуванням кожного фактора;

Зробити висновки з кожного пункту.

### Розв'язання

1. Індивідуальні індекси визначимо за формулами (10.2), (10.3), тільки замість  $p$  і  $q$  використаємо  $t$  і  $w$ . Розрахунок зведемо в табл. 10.4.

Таблиця 10.4

Індивідуальні індекси	По 1-му підприємству	По 2-му підприємству
Змінення кількості працівників	$i_t^1 = \frac{t_1}{t_0} = \frac{1510}{1500} = 1,2267,$ або 100,67 %	$i_t^2 = \frac{420}{423} = 0,9929,$ або 99,29 %
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 %
Обсягу виробленої продукції	$i_Q^1 = \frac{w_1 t_1}{w_0 t_0} = \frac{21895}{21450} = 1,02,$ або 102,0 %	$i_Q^2 = \frac{25200}{25210,8} = 0,9995,$ або 99,95 %

**Висновок:** з розрахунків значень індивідуальних індексів випливає, що кількість працівників на першому підприємстві збільшилася на 0,67 %, на другому – зменшилася 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 (10.4):

$$I_Q = \frac{14,5 \cdot 1510 + 60,0 \cdot 420}{14,3 \cdot 1500 + 59,6 \cdot 423} = \frac{47095,0}{46660,8} = 1,009305, \text{ або } 100,93 \%$$

Загальний індекс змінення кількості працівників 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 зміненням кількості працівників 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 зміненням 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 індекс  $I_w$  – за формулою індексу Ласпейреса.

3. Обчислимо загальні індекси як середньозважені. Формули побудуємо аналогічно формулам з таблиці 10.2, тільки замість  $p$  і  $q$  використаємо  $t$  і  $w$ :

$$I_Q = \frac{\sum i_Q w_0 t_0}{\sum w_0 t_0} = \frac{1,02 \cdot 21450 + 0,9995 \cdot 25210,8}{21450 + 2510,8} = 1,009305;;$$

$$I_t = \frac{\sum i_t t_0 w_0}{\sum t_0 w_0} = \frac{1,0067 \cdot 21450 + 0,9929 \cdot 25210,8}{21450 + 2510,8} = 0,999233;;$$

$$I_w = \frac{\sum t_1 w_1}{\sum \frac{1}{i_w} t_1 w_1} = \frac{21825 + 25200}{\frac{21895}{1,0139} + \frac{25200}{1,0067}} = 1,01008.$$

4. Обчислимо абсолютний приріст як різницю між чисельником і знаменником відповідних агрегатних індексів.

Абсолютний приріст обсягу виробленої продукції

$$\Delta Q = \sum w_1 t_1 - \sum w_0 t_0 = 47095 - 46660,8 = +433,2;$$

з урахуванням 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 кількість працівників зменшилась, обсяг виробленої продукції зменшився на 35 тис. грн.

### Задача 10.2

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 в. п. збільшилася частка продовольчих товарів і на стільки ж зменшилася частка непродовольчих.

Розрахувати середньозважені індекси цін та обсягу товарів.

### Розв'язання

Розрахунки індивідуальних індексів обсягу і цін зведемо в табл. 10.5.

Таблиця 10.5

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

$$I_q = \sum i_q d_0 = 0,97 \cdot 0,60 + 0,93 \cdot 0,40 = 0,954,$$

тобто в середньому обсяги виробництва зменшилися на 4,6 %;

$$I_p = \frac{1}{\sum \frac{1}{i_p} d_1} = \frac{1}{\frac{0,62}{1,04} + \frac{0,38}{1,06}},$$

тобто ціни в середньому збільшилися на 4,8 %.

### Задача 10.3

За даними табл. 10.6 визначити індекси середнього розміру тарифу страхування легкових автомобілів з терміном експлуатації до трьох років

Таблиця 10.6

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 $f_1$	$x_0 f_0$	$x_1 f_1$	$x_0 f_1$
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
Усього	–	–	900	1600	32,0	73,5	61,25

### Розв'язання

Індекс змінного складу розрахуємо за формулою (10.5):

$$I_x = \frac{73,5}{1600} : \frac{32}{900} = 0,046 : 0,036 = 1,278.$$

Цей індекс є відображенням того, що середній тариф страхування в поточному періоді порівняно з базисним підвищився на 27,8 %. Індекс фіксованого складу визначимо за формулою (10.6):

$$I_x = \frac{73,5}{1600} : \frac{61,25}{1600} = 0,046 : 0,038 = 1,211.$$

Унаслідок підвищення тарифу страхування по кожній групі автомобілів його середній тариф збільшився на 21,1 %. Індекс структурних зрушень  $I_d = 0,038 : 0,036 = 1,055$ . Отже, середній тариф страхування збільшився на 5,6 % через змінення у складі об'єктів страхування, а саме – збільшення частки суми страхування зарубіжних автомобілів з вищою ставкою страхування.

Індекс структурних зрушень знайдемо з допомогою формули (10.7):

$$I_d = 1,278 / 1,211 = 1,055.$$

#### Задача 10.4

Відомо, що на ринках двох міст у звітному періоді реалізація товарів характеризувалася певними даними (табл. 10.7).

Таблиця 10.7

Товар	Місто А		Місто В		Індивідуальний індекс цін	
	Модальна ціна $p_A$ , грн, за 1 ц	Кількість товару $q_A$ , т	Модальна ціна $p_B$ , грн, за 1 ц,	Кількість товару $q_B$ , т	$i_{p_{A/B}} = \frac{p_A}{p_B}$	$i_{p_{B/A}} = \frac{p_B}{p_A}$
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, розрахуємо загальний індекс цін за формулою (10.11). Знаменник формули відображає умовний товарообіг, який міг бути в разі реалізації асортименту товарів за цінами, що склалися в місті В:

$$I_{p_{A/B}} = \frac{\sum q_A p_A}{\sum q_A p_B} = \frac{4 \cdot 50 + 7 \cdot 70 + 3 \cdot 120}{4 \cdot 40 + 7 \cdot 60 + 3 \cdot 130} = \frac{1050}{970} = 1,082.$$

Результат свідчить про те, що якби товари цього асортименту продавалися за цінами міста А, то їх рівень був би вищим 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.

## Завдання для самостійної роботи

1. Є такі дані про зміни в реалізації верхнього одягу в комерційному магазині за два місяці:

Пальто демісезонні	Ціна за одиницю товару, тис. грн		Кількість реалізованого товару, шт.	
	Базовий період $p_0$	Поточний період $p_1$	Базовий період $q_0$	Поточний період $q_1$
Чоловічі	22	30	50	60
Жіночі	26	38	45	30

1. Визначити індивідуальні індекси цін, фізичного обсягу реалізованого товару й товарообігу.

2. Розрахувати загальні індекси товарообігу, цін, фізичного обсягу реалізованого товару. Показати взаємозв'язок загальних індексів.

3. Обчислити абсолютний приріст товарообігу загалом унаслідок змінення кожного фактора. Зробити висновки.

2. Розглядається група з двох підприємств, які виробляють різну продукцію. По кожному підприємству є дані за два суміжні роки (базисний і звітний) про кількість працівників і середній рівень виробітку на одну особу:

Номер підприємства	Базовий рік		Поточний рік	
	Середній виробіток, млн грн на 1 людину	Середня кількість працівників, чол.	Середній виробіток, млн грн на 1 людину	Середня кількість працівників, чол.
1	14,3	1500	14,5	1510
2	59,6	423	60,0	420
Разом	24,264	1923	24,4	1930

1. Визначити індивідуальні індекси змінення кількості працівників, рівня виробітку й обсягу виробленої продукції.

2. Знайти загальні індекси обсягу виробленої продукції, змінення кількості працівників, рівня виробітку. Показати взаємозв'язок індексів.

3. Обчислити загальні індекси середніх величин.

4. Розрахувати абсолютний приріст товарообігу загалом і внаслідок змінення кожного чинника.

5. Зробити висновки.

3. Є такі дані щодо двох підприємств галузі:

Підприємство	Реалізовано продукції, тис. грн		Середньооблікова кількість робочих, чол.	
	I квартал	II квартал	I квартал	II квартал
I	540	544	100	80
II	450	672	100	120

1. Визначити рівні й динаміку продуктивності праці робітників кожного підприємства.

2. Для кожного підприємства обчислити:

- індекс продуктивності праці змінного складу;
- індекс продуктивності праці фіксованого складу;
- індекс впливу структурних змін кількості робітників на динаміку середньої продуктивності праці;
- абсолютне й відносне змінення обсягу реалізації продукції у II кварталі (на одному з підприємств) унаслідок змінення кількості робітників, рівня продуктивності праці, двох факторів одночасно.

3. Показати зв'язок між обчисленими показниками.

4. Є такі умовні дані про ціни й реалізацію товару А двома торговими підприємствами:

Підприємство	Базовий період		Поточний період	
	Ціна за одиницю товару, грн	Кількість проданого товару, од.	Ціна за одиницю товару, грн	Кількість проданого товару, од.
I	4,5	750	5	200
II	3,5	450	4	300

Визначити:

- індекси цін змінного й постійного складу, структурних зрушень;
- абсолютне змінення середньої ціни товару під впливом окремих чинників.

Зробити висновки.

5. Наведено такі дані про товарообіг за трьома товарними групами:

Група товарів	Роздрібний товарообіг у фактичних цінах, тис. грн		Змінення цін у звітному періоді порівняно з базовим, %
	Базовий період	Поточний період	
1	2500	3000	+6
2	2000	2200	+12
3	3000	3180	-4

Визначити за трьома групами товарів разом:

- індекси товарообігу у фактичних і порівнянних цінах, індекс цін;
- абсолютне змінення товарообігу, у тому числі внаслідок змінення фізичного обсягу товарообігу й цін.

Зробити висновки.

6. У регіоні виробництво продовольчих і непродовольчих товарів співвідноситься як 6 : 4. За деякий період виробництво зменшилося: продовольчих товарів – на 3 %, промислових – на 7 %, а ціни збільшилися відповідно на 4 і 6 %. Унаслідок нерівномірності динаміки виробництва за групами

товарів змінилася їхня структура: на 2 в. п. збільшилася частка продовольчих товарів і на стільки ж зменшилася частка промислових. Розрахувати загальні індекси фізичного обсягу й цін.

7. За наведеними нижче даними розрахувати середньозважені індекси цін і кількості проданих акцій на фондовому ринку. Зробити висновки.

Ринок	Обсяг торгів, млн грн		Темп приросту, %	
	Базовий період	Поточний період	Ціна на акції	Кількість акцій
Первинний	80	190	+90	+35
Вторинний	40	120	+150	+20
Разом	120	310		

9. Динаміка ефективності виробничо-фінансової діяльності малих підприємств з різною формою власності характеризується такими даними:

Форма власності	Витрати на виробництво в порівнянних цінах, млн грн.		Окупність витрат на виробництво, %	
	Базовий період	Поточний період	Базовий період	Поточний період
Державна	45	70	30	26
Колективна	450	580	28	32
Приватна	325	800	32	38

Визначити індекси середньої окупності виробничих витрат змінного складу, фіксованого складу, структурних зрушень. Навести їх економічну інтерпретацію, показати взаємозв'язок.

10. У звітному періоді продано шкіряного взуття на 50 тис. грн, гумового – на 20 тис. грн, комбінованого – на 10 тис. грн. Визначити загальний індекс цін, якщо відомо, що ціна на шкіряне взуття збільшилася на 17 %, на гумове – зменшилася на 7 %, на комбіноване – залишилася без змін.

11. У звітному періоді порівняно з базисним ціни на продукцію зменшилися на 5 %, а вартість виготовленої продукції збільшилася на 5 %. Визначити фізичний обсяг продукції у звітному періоді.

12. Випуск продукції в лютому проти січня зменшився на 5 %, а в березні проти лютого збільшився на 10 %. Визначити, як змінився випуск продукції в березні порівняно із січнем.

13. У звітному періоді порівняно з базисним змінення собівартості продукції загалом у галузі характеризується такими показниками: загальний індекс собівартості змінного складу – 0,985, фіксованого – 0,937. Визначити, як змінилася середня собівартість продукції внаслідок структурних зрушень, які сталися у виробництві продукції на підприємствах галузі.

## Контрольні запитання

1. Що характеризує індекс? Які функції виконують індекси в статистичному аналізі?
2. Що таке індивідуальні індекси? Як їх використовують в економічному аналізі?
3. Що таке загальні індекси? Як їх використовують в економічному аналізі?
4. Поясніть суть і методику побудови індексів агрегатної форми. Які системи зважування індексів ви знаєте?
5. Як визначити абсолютний вплив факторів-співмножників індексної системи на динаміку результативного показника?
6. Поясніть суть середньозважених індексів.



## ТЕСТИ

### Модуль № 1

1. До основних категорій теорії статистики належать:

- а) статистична сукупність, статистична закономірність, статистичний показник;
- б) одиниці виміру, технологія збирання даних, обчислювальні операції;
- в) теорія статистики, економічна статистика, соціальна статистика.

2. Статистична сукупність – це:

- а) безліч одиниць явища, що вивчається, об'єднаних єдиною якісною основою, але відрізняються один від одного окремими ознаками;
- б) безліч одиниць явища, що вивчається, об'єднаних єдиною кількісною основою, але відрізняються один від одного окремими ознаками;
- в) сукупність соціально-економічних явищ і процесів, що підлягають дослідженню, або точні межі, у яких реєструватимуться статистичні дані.

3. Статистичне дослідження містить:

- а) статистичне спостереження, групування й зведення, оброблення й аналіз даних;
- б) статистичне спостереження;
- в) групування й зведення статистичних даних;
- г) статистичне спостереження, групування й зведення, будування таблиць і графіків.

4. Статистичне спостереження – це:

- а) облік та нагромадження даних про одиниці сукупності масових явищ;
- б) науково організоване збирання даних про масові явища й процеси за певною програмою;
- в) контроль виконання будь-якої роботи.

5. Об'єктом статистичного спостереження є:

- а) сукупність соціально-економічних явищ і процесів, що підлягають дослідженню, або точні межі, у яких реєструватимуться статистичні дані;
- б) статистичні дані, які є основою для узагальненої характеристики стану й розвитку явища або процесу;
- в) безліч одиниць явища, що вивчається, об'єднаних єдиною якісною основою, але відрізняються один від одного окремими ознаками.

5. У чому полягає мета статистичного спостереження:

- а) отримання статистичних даних, які є основою для узагальненої характеристики стану й розвитку явища або процесу;
- б) вивчення кількісного боку масових суспільних явищ у нерозривному зв'язку з їх якісним боком;
- в) проведення аналізу статистичних даних з урахуванням обчислення узагальнювальних показників?

6. Одиницею статистичного спостереження називають:
- а) складовий елемент об'єкта спостереження, який є носієм ознак, що підлягають реєстрації під час спостереження;
  - б) соціально-економічне явище або процес, що вивчається під час статистичного спостереження;
  - в) звітна одиниця, що є джерелом інформації.
7. Суть статистичного зведення полягає:
- а) у розподілі суспільних явищ на однорідні групи за певними ознаками;
  - б) в обліку фактів про явища життя;
  - в) у науковому обробленні статистичного матеріалу, виявленні типових характеристик і закономірностей досліджуваних явищ;
  - г) у перевірці й контролі статистичного матеріалу.
8. За способом отримання даних спостереження може бути:
- а) поточним, періодичним, одноразовим;
  - б) у вигляді звітності, спеціально організованим спостереженням;
  - в) безпосереднім, документальним, як опитування.
9. З метою вивчення досвіду зроблено опис господарської діяльності сільськогосподарського підприємства. За охопленням одиниць сукупності це:
- а) суцільне спостереження;
  - б) вибіркоче спостереження;
  - в) обстеження основного масиву.
10. Під час статистичного аналізу:
- а) виявляють закономірності й взаємозв'язки суспільних явищ шляхом використання узагальнювальних абсолютних, відносних, середніх величин, індексів і статистичних коефіцієнтів;
  - б) одержують первинну інформаційну базу для статистичного дослідження;
  - в) одержують підсумкові дані матеріалів статистичного спостереження.
12. Безперервною з кількісних ознак є:
- а) середньодобовий пробіг автомобіля;
  - б) загальний результат балів;
  - в) середній бал;
  - г) випуск одиниць продукції.
13. З допомогою якого виду графіків рядів розподілу зображують дискретні варіаційні ряди:
- а) полігон;
  - б) гістограма;
  - в) кумуляти?
14. Для обчислення відносних величин використовують арифметичну дію:
- а) розподіл;
  - б) різницю;
  - в) множення.

15. Відносні величини структури характеризують:

- а) відношення частини сукупності до всієї сукупності;
- б) відношення однієї частини сукупності до іншої частини цієї сукупності;
- в) відношення явищ у часі.

16. Відносні величини координації характеризують:

- а) співвідношення між окремими складовими частинами сукупності;
- б) ступінь поширення явища в певному середовищі;
- в) співвідношення розмірів явищ, що належать до різних об'єктів;
- г) питома вага окремих елементів сукупності в загальному обсязі.

17. Якщо всі індивідуальні значення ознаки збільшити в 3 рази, а частоти зменшити в 3 рази, то середня величина:

- а) не зміниться;
- б) зменшиться в 3 рази;
- в) збільшиться в 3 рази;
- г) змінення середньої величини передбачити неможливо.

18. Якщо частоти всіх значень ознаки зменшити на 10 одиниць, то середня величина:

- а) зменшиться на 10;
- б) зменшиться в 10 разів;
- в) не зміниться;
- г) змінення середнього розміру визначити неможливо.

19. Середнє лінійне відхилення є:

- а) відхиленням індивідуальних значень ознаки середньої величини;
- б) абсолютним відхиленням індивідуальних значень ознаки середньої величини;
- в) квадратом відхилень індивідуальних значень ознаки середньої величини.

20. Дисперсія – це:

- а) квадрат відхилень індивідуальних значень ознаки середньої величини;
- б) середнє відхилення індивідуальних значень ознаки середньої величини;
- в) середній квадрат відхилень індивідуальних значень ознаки середньої величини.

21. Абсолютний вимір варіації ознаки у статистичній сукупності характеризує:

- а) середнє квадратичне відхилення;
- б) розмах варіації;
- в) середній квадрат відхилення.

22. Якщо всі частоти дискретного ряду розподілу зменшити втричі, а індивідуальні значення ознаки залишити без змін, то мода:

- а) не зміниться;
- б) зменшиться на три;
- в) змінення визначити неможливо.

## Модуль № 2

1. Рядом розподілу називають:
  - а) ранжований у порядку збільшення або зменшення значень ознаки ряд варіантів з відповідними частотами;
  - б) розподіл цифрових даних у збільшному або спадному порядку;
  - в) сукупність даних, одержаних унаслідок спостереження.
2. При якому розподілі середня арифметична, мода й медіана будуть однаковими:
  - а) при нормальному симетричному;
  - б) при помірно асиметричному;
  - в) при асиметричному?
3. У чому полягає головне завдання дисперсійного аналізу:
  - а) статистичне виявлення впливу факторів, що зумовлюють мінливість ознаки;
  - б) статистичне вивчення варіації середньої величини ознаки;
  - в) визначення вибіркової дисперсії?
4. Внутрішньогруповою дисперсією називають:
  - а) середню арифметичну зважену з групових дисперсій;
  - б) дисперсію групових середніх відносно загальної середньої величини;
  - в) дисперсію значень ознаки, що включається до групи, відносно групової середньої величини;
  - г) дисперсію значень ознаки відносно загальної середньої величини?
5. Загальна дисперсія ознаки дорівнює:
  - а) сумі внутрішньогрупової й міжгрупової дисперсій;
  - б) сумі групової й міжгрупової дисперсій;
  - в) сумі групової й внутрішньогрупової дисперсій;
  - г) різниці міжгрупової й внутрішньогрупової дисперсій.
6. Що таке регресія:
  - а) лінія залежності середньої величини результативної ознаки від факторної;
  - б) щільність зв'язку;
  - в) вид пропорційної залежності двох змінних?
7. Головною умовою застосування методу найменших квадратів для визначення параметрів рівняння регресії є:
  - а) мінімізація суми квадратів відхилень емпіричних значень результативної ознаки від теоретичних;
  - б) мінімізація суми відхилень емпіричних значень результативної ознаки від теоретичних;
  - в) максимізація суми квадратів відхилень емпіричних значень результативної ознаки від теоретичних.

8. Стохастичний зв'язок – це:
- а) зв'язок між випадковими величинами, коли змінення однієї з них приводить до змінення розподілу інших;
  - б) вид кореляційного зв'язку;
  - в) форма кореляційного зв'язку.
9. При функціональному зв'язку:
- а) кожному можливному значенню факторної ознаки відповідає певне значення результативної ознаки;
  - б) кожному можливному значенню факторної ознаки може відповідати кілька значень результативної ознаки;
  - в) кожному можливному значенню факторної ознаки відповідає деяке середнє значення результативної ознаки.
10. Кореляційне відношення відображає:
- а) співвідношення загальної й внутрішньогрупової дисперсій;
  - б) яка частина загальної дисперсії результативної ознаки пояснюється варіацією ознаки, за якою групують;
  - в) яка частина загальної варіації результативної ознаки пояснюється неврахованими факторами.
11. Коефіцієнт детермінації відображає:
- а) частку загальної варіації результативної ознаки, що пояснюється факторною ознакою;
  - б) загальну варіацію результативної ознаки;
  - в) частку загальної варіації результативної ознаки, яка пояснюється ознаками, що не увійшли до рівняння регресії.
12. Укажіть, які з наведених динамічних рядів є моментними:
- а) валютні резерви банківської системи на початок року;
  - б) заборгованість комерційних банків акціонерам і бюджету станом на 1 квітня;
  - в) експорт товарів за квартал року.
13. Який аналітичний показник ряду динаміки характеризує абсолютну величину розміру змінення явища:
- а) абсолютний приріст;
  - б) коефіцієнт зростання;
  - в) темп приросту?
14. Яку назву має середня величина для розрахунку середньорічної кількості працівників, якщо відомою є їх кількість на початок кожного місяця року:
- а) хронологічна;
  - б) арифметична;
  - в) гармонійна?

15. Який тип аналітичної функції використовують для вирівнювання ряду динаміки у випадках, коли абсолютні прирости є майже постійними:
- а) рівняння прямої;
  - б) рівняння параболи;
  - в) рівняння показової функції?
16. Екстраполяцією тренду називають:
- а) один із методів статистичного прогнозування;
  - б) продовження виявленої тенденції межі ряду динаміки;
  - в) вирівнювання сезонних коливань.
17. За рівнем охоплення одиниць сукупності індекси поділяють:
- а) на індивідуальні й загальні;
  - б) на ланцюгові й базисні;
  - в) на динамічні й територіальні.
18. Залежно від методології обчислення загальні індекси поділяють:
- а) на агрегатні й середні з індивідуальних;
  - б) на ланцюгові й базисні;
  - в) на динамічні й територіальні.
19. Залежно від бази порівняння розрізняють:
- а) ланцюгові й базисні індекси;
  - б) агрегатні й середні з індивідуальних індексів;
  - в) динамічні й територіальні індекси.
20. Залежно від характеру досліджуваних об'єктів розрізняють:
- а) індекси об'ємних і якісних показників;
  - б) ланцюгові й базисні індекси;
  - в) агрегатні й середні з індивідуальних індексів.
21. Функції загальних індексів:
- а) синтетична й аналітична;
  - б) факторна й результативна;
  - в) узагальнювальна й усереднення даних.

## БІБЛІОГРАФІЧНИЙ СПИСОК

Городянська, Л. В. Статистика для економістів : навч. посіб. / Л. В. Городянська, А. І. Сизов. – Київ : Київ. нац. ун-т ім. Т. Шевченка, 2019. – 350 с.

1. Мармоза, А. Т. Теорія статистики : підручник / А. Т. Мармоза. – 2-ге вид., перероб. і доп. – Київ : Центр учбової літератури, 2013. – 592 с.

Петрик, В. Л. Статистика : навч. посіб. / В. Л. Петрик. – Харків : ХАІ, 2009. – 100 с.

2. Рощик, І. А. Соціально-економічна статистика : навч. посіб. / І. А. Рощик. – Рівне : НУВГП, 2010. – 338 с.

Статистика : підручник / С. І. Пирожков, В. В. Рязанцева, Р. М. Моторин та ін. – Київ : Київ. нац. торг.-екон. ун-т, 2020. – 328 с.

Чекотовський, Е. В. Статистичні методи. Історія і теорія : навч. посіб. / Е. В. Чекотовський. – Київ : Знання, 2016. – 191 с.

## ЗМІСТ

1. Методологічні засади статистики. Статистичне спостереження.....	3
1.1. Предмет і метод статистичної науки.....	3
1.2. Статистичне спостереження.....	5
Приклади розв'язання типових задач.....	9
Завдання для самостійної роботи.....	10
Контрольні запитання.....	11
2. Зведення та групування статистичних даних.....	12
2.1. Поняття статистичного зведення та групування.....	12
2.2. Правила проведення групування.....	14
Приклад розв'язання типових задач.....	18
Завдання для самостійної роботи.....	22
Контрольні запитання.....	23
3. Статистичні графіки.....	24
Приклади розв'язання типових задач.....	27
Завдання для самостійної роботи.....	32
Контрольні запитання.....	33
4. Узагальнювальні статистичні показники.....	33
4.1. Суть і класифікація статистичних показників.....	33
4.2. Абсолютні величини.....	34
4.3. Види і взаємозв'язки відносних величин.....	35
Приклади розв'язання типових задач.....	37
Завдання для самостійної роботи.....	42
Контрольні запитання.....	44
5. Середні величини.....	45
5.1. Суть і види середньої величини.....	45
5.2. Види степеневих середніх і способи їх обчислення.....	46
5.3. Структурні середні величини.....	50
Приклади розв'язання типових задач.....	51
Завдання для самостійної роботи.....	59
Контрольні запитання.....	62
6. Показники варіації.....	62
6.1. Поняття варіації та її основні показники.....	62
6.2. Види дисперсій і правила їх додавання.....	64
Приклад розв'язання типових задач.....	65
Завдання для самостійної роботи.....	66
Контрольні запитання.....	68
7. Вибіркове спостереження.....	68
7.1 Вибіркове спостереження та його основні завдання.....	68



7.2. Знаходження середньої й граничної помилок і необхідного обсягу вибірки.....	69
Приклад розв'язання типових задач.....	72
Завдання для самостійної роботи.....	75
Контрольні запитання.....	77
8. Статистичні методи вивчення взаємозв'язків.....	77
8.1. Основні поняття й категорії.....	77
8.2. Кореляційний і регресійний методи аналізу зв'язку.....	80
Приклади розв'язання типових задач.....	83
Завдання для самостійної роботи.....	88
Контрольні запитання.....	90
9. Ряди динаміки.....	91
9.1. Ряди динаміки. Класифікація динамічних рядів.....	91
9.2. Характеристики динамічних рядів.....	92
9.3. Аналіз структурних зрушень.....	95
9.4. Середні показники динаміки.....	96
9.5. Характеристика основної тенденції розвитку.....	97
9.6. Аналіз сезонних коливань.....	100
9.7. Екстраполяція в рядах динаміки та прогнозування.....	102
Приклади розв'язання типових задач.....	103
Завдання для самостійної роботи.....	107
Контрольні запитання.....	109
10. Статистичні індекси.....	110
10.1. Суть і функції індексів.....	110
10.2. Індивідуальні індекси.....	111
10.3. Загальні індекси.....	112
10.4. Середньозважені індекси.....	113
10.5. Взаємозв'язок індексів та індексні системи.....	114
10.6. Індекси середніх величин.....	116
10.7. Територіальні індекси.....	118
Приклади розв'язання типових задач.....	119
Завдання для самостійної роботи.....	124
Контрольні запитання.....	127
Тести.....	128
Бібліографічний список.....	134

Навчальне видання

**Петрик Валерія Леонідівна  
Голованова Майя Анатоліївна**

## **СТАТИСТИКА**

Редактор О. Ф. Серьожкіна

Зв. план, 2023

Підписано до видання 22.01.2024

Ум. друк. арк. 7,6. Обл.-вид. арк. 8,56. Електронний ресурс

---

Видавець і виготовлювач

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](mailto:izdat@khai.edu)

Свідоцтво про внесення суб'єкта видавничої справи  
до Державного реєстру видавців, виготовлювачів і розповсюджувачів  
видавничої продукції сер. ДК № 391 від 30.03.2001