

ПЕДЧЕНКО Г.П.

СТАТИСТИКА

НАВЧАЛЬНИЙ ПОСІБНИК

*Рекомендовано Міністерством
освіти та науки України*

2018

УДК 311.1 (075.8)
ББК 60.6я73
П 24

Педченко Г. П. Статистика: Навчальний посібник / Г. П. Педченко. — Мелітополь: Колор Принт, 2018. — 266 с. — ISBN 978-966-2489-67-5.

Рецензенти:

В.В. Липчук - доктор економічних наук, завідувач кафедри статистики та аналізу Львівського національного аграрного університету, професор;

І.В. Охріменко - доктор економічних наук, декан факультету аграрного менеджменту Національного університету біоресурсів і природокористування України, професор;

О.В. Кочетков - кандидат економічних наук, завідувач кафедри статистики та економічного аналізу Луганського національного аграрного університету, доцент.

У навчальному посібнику розглянуті питання організації статистичного спостереження, в тому числі вибіркового, групування та подання статистичних даних, викладені методики розрахунку узагальнюючих показників, аналізу варіації, диференціації, концентрації та подібності в рядах розподілу, аналізу інтенсивності та тенденцій розвитку та коливань в рядах динаміки, статистичного вимірювання взаємозв'язків та індексного аналізу. Теоретичні відомості підкріплені наведеними прикладами та розв'язаннями типових задач, а також надані завдання для самостійного розв'язання та тести для самоконтролю.

Навчальний посібник призначений для студентів економічних напрямів аграрних вищих навчальних закладів III-IV рівнів акредитації.

*Рекомендовано Міністерством освіти і науки України
(лист №1/11 - 11469 від 22 липня 2014 р.)*

ISBN 978-966-2489-67-5

© Г.П.Педченко, 2018
© Колор Принт, 2018

ЗМІСТ

ВСТУП.....	6
1. МЕТОДОЛОГІЧНІ ОСНОВИ СТАТИСТИКИ.	
СТАТИСТИЧНЕ СПОСТЕРЕЖЕННЯ.....	7
1.1 Предмет і метод статистики.....	7
1.2 Статистичне спостереження.....	8
1.3 Помилки статистичного спостереження.....	12
Завдання для самостійного розв'язання.....	14
Тести для самоконтролю.....	17
2. ПОДАННЯ СТАТИСТИЧНИХ ДАНИХ:	
ТАБЛИЦІ, ГРАФІКИ, КАРТИ.....	23
2.1 Статистичні таблиці.....	23
2.2 Статистичні графіки та їх класифікація.....	32
2.3 Діаграми.....	35
2.4 Статистичні карти.....	52
Завдання для самостійного розв'язання.....	57
Тести для самоконтролю.....	60
3. ЗВЕДЕННЯ ТА ГРУПУВАННЯ	
СТАТИСТИЧНИХ ДАНИХ.....	64
3.1 Статистичне зведення та групування. Види групувань.....	64
3.2 Попередні відомості про ряди розподілу.....	66
3.3 Техніка проведення групування.....	73
3.4 Порівнянність статистичних групувань.....	79
Завдання для самостійного розв'язання.....	84
Тести для самоконтролю.....	88
4. УЗАГАЛЬНЮЮЧІ СТАТИСТИЧНІ ПОКАЗНИКИ.....	92
4.1 Сутність та види статистичних показників.....	92
4.2 Абсолютні величини.....	93
4.3 Відносні величини: сутність, види, одиниці виміру.....	94
4.4 Середні величини.....	100
Завдання для самостійного розв'язання.....	108
Тести для самоконтролю.....	112

5. АНАЛІЗ РЯДІВ РОЗПОДІЛУ	119
5.1 Статистичні характеристики розподілів.....	119
5.2 Графічне зображення варіаційних рядів розподілу.....	121
5.3 Показники варіації рядів розподілу.....	127
Завдання для самостійного розв'язання.....	133
Тести для самоконтролю.....	135
6. АНАЛІЗ ДИФЕРЕНЦІАЦІЇ, КОНЦЕНТРАЦІЇ ТА ПОДІБНОСТІ РОЗПОДІЛІВ.....	141
6.1 Структурні характеристики розподілів.....	141
6.2 Показники диференціації розподілів.....	148
6.3 Показники концентрації розподілів.....	149
6.4 Подібність розподілів.....	154
Завдання для самостійного розв'язання.....	155
Тести для самоконтролю.....	157
7. ВИБІРКОВИЙ МЕТОД.....	162
7.1 Поняття вибіркового методу (спостереження).....	162
7.2 Основні способи формування вибіркової сукупності і статистична оцінка вибірових характеристик.....	163
7.3 Визначення необхідного обсягу вибірки та поширення її характеристик на генеральну сукупність.....	172
7.4 Мала вибірка.....	175
Завдання для самостійного розв'язання.....	177
Тести для самоконтролю.....	180
8. АНАЛІЗ РЯДІВ ДИНАМІКИ.....	187
8.1 Сутність та види рядів динаміки.....	187
8.2 Визначення середнього рівня динамічного ряду.....	188
8.3 Показники інтенсивності рядів динаміки.....	189
Завдання для самостійного розв'язання.....	193
Тести для самоконтролю.....	194

9. АНАЛІЗ ТЕНДЕНЦІЙ РОЗВИТКУ ТА КОЛИВАНЬ.....	199
9.1 Методи визначення основної тенденції розвитку у рядах динаміки.....	199
9.2 Прогнозування.....	204
9.3 Статистичне вивчення сезонних коливань.....	207
Завдання для самостійного розв'язання.....	211
Тести для самоконтролю.....	213
10. СТАТИСТИЧНІ МЕТОДИ ВИМІРЮВАННЯ ВЗАЄМОЗВ'ЯЗКІВ.....	217
10.1 Види та форми взаємозв'язків.....	217
10.2 Сутність та етапи кореляційно-регресійного аналізу.....	219
10.3 Парна кореляція.....	221
10.3 Множинна кореляція.....	226
Завдання для самостійного розв'язання.....	231
Тести для самоконтролю.....	232
11. ІНДЕКСИ.....	239
11.1 Індеси та їх класифікація.....	239
11.2 Індивідуальні індеси.....	240
11.3 Загальні індеси.....	242
11.4 Аналіз співвідношень середніх рівнів.....	246
Завдання для самостійного розв'язання.....	248
Тести для самоконтролю.....	299
ДОДАТКИ.....	257
ПРЕДМЕТНИЙ ПОКАЖЧИК.....	259
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	264

ВСТУП

В сучасному суспільстві ефективність управління економікою значною мірою зумовлюється якістю інформаційної бази, основу якої становить статистика. Статистична грамотність є важливою рисою підприємця, економіста, менеджера, маркетолога, фінансиста, бухгалтера та інших професіоналів своєї справи. Тому у системі економічної освіти значна роль відводиться статистиці як фундаментальній дисципліні з циклу професійної та практичної підготовки фахівців економічних напрямів, робота яких постійно пов'язана з використанням і аналізом статистичних даних.

Статистика вивчає кількісну сторону масових суспільних явищ у нерозривному зв'язку з їх якісним змістом у конкретних умовах місця і часу. Вона здійснює збір, обробку та аналіз даних, що характеризують всі сторони суспільного життя, вивчає закономірності їх зміни і прогнозування подальшого розвитку. Оволодіння статистичної методологією - одна з умов проведення статистичних досліджень, спрямованих на пізнання закономірностей розвитку соціально-економічних явищ і процесів та прийняття оптимальних рішень фахівцем будь-якого рівня у сфері виробництва, маркетингу, фінансів, тощо.

Мета навчального посібника - продемонструвати використання основних методів статистичного дослідження, знаходження статистичних показників та інтерпретацію їх значень при вирішенні економічних завдань, сформувати у студентів теоретичні знання і практичні навички для аналізу досліджуваного об'єкта, розробки обґрунтованих рішень та рекомендацій, а також полегшити самостійну роботу над курсом, для чого в ньому викладені основні теоретичні відомості, наведені приклади та розв'язання типових задач, а також надані завдання для самостійного розв'язання та тести для самоконтролю.

Вивчаючи курс «Статистика» з допомогою даного посібника, студенти повинні вміти проводити збір та статистичну обробку даних, обчислювати узагальнюючі характеристики сукупності, оцінювати варіацію ознак, диференціацію та концентрацію та подібність розподілів, визначати інтенсивність динаміки явищ та основні тенденції їх розвитку, визначати фактори, які впливають на розвиток того чи іншого явища, вимірювати силу такого впливу, а також оцінювати надійність одержаних результатів.

1. МЕТОДОЛОГІЧНІ ОСНОВИ СТАТИСТИКИ. СТАТИСТИЧНЕ СПОСТЕРЕЖЕННЯ

1.1 Предмет і метод статистики

Статистика - це суспільна наука і, одночасно, галузь практичної діяльності, що вивчає кількісну сторону якісно визначених масових соціально-економічних явищ і процесів, їх структуру і розподіл, розміщення в просторі, рух в часі, виявляючи діючі кількісні залежності, тенденції і закономірності в конкретних умовах місця і часу за допомогою сукупності методів і прийомів по збору, обробці, представленню і аналізу числових даних.

Статистична сукупність - це множина об'єктів або явищ, досліджуваних статистикою, які мають одну або кілька загальних ознак і розрізняються між собою по інших ознаках.

Об'єкт дослідження - сукупність явищ, що підлягає обстеженню і про які потрібно зібрати дані. Об'єктом спостереження може бути сукупність фізичних осіб (населення окремого регіону, країни; особи, зайняті на підприємствах галузі), фізичні одиниці (верстати, машини, житлові будинки), юридичні особи (підприємства, фермерські господарства, комерційні банки, навчальні заклади).

Об'єкт дослідження складається з **одиниць сукупності** - первинних елементів, які є носіями ознак, що підлягають реєстрації. Саме одиниці сукупності становлять основу обліку і підлягають безпосередньому обстеженню, тому що сума таких одиниць обов'язково дає статистичну сукупність явищ.

Одиниця спостереження - джерело здобуття відомостей, первинна одиниця (особа, підприємство, заклад), від якої одержують інформацію, елемент спостереження, що відбирається для дослідження на різних етапах цього процесу.

Приклад. При проведенні перепису торгівельного устаткування в магазинах міста одиницею спостереження є кожний магазин у місті як джерело одержання інформації, а одиницею сукупності – кожна одиниця торгівельного встаткування (прилавки, холодильні агрегати й т.д.), що переписується. Іноді поняття одиниці сукупності та одиниці спостереження можуть збігатися, як, зокрема, під час перепису населення ними є кожна окрема людина.

Загальний розмір явища (земельна площа, обсяг продукції) для всієї сукупності називають **обсягом явища**, а кількість одиниць, які мають дану ознаку (кількість підприємств, працівників, худоби) - **обсягом сукупності**.

Приклад. Вивчається поголів'я корів в аграрних підприємствах області.

Обсяг явища - загальне поголів'я корів у аграрних підприємствах області;

Обсяг сукупності - кількість аграрних підприємств області.

Досліджується продуктивність корів в аграрних підприємствах області.

Обсяг явища - валове виробництво молока в аграрних підприємствах області;

Обсяг сукупності - поголів'я корів в цих підприємствах.

Ознака - це характерна риса, властивість досліджуваного явища, що відрізняє його від інших явищ.

Приклад. *Об'єктом вивчення є підприємство, його ознаки - спеціалізація, форма власності, обсяг продукції, чисельність працюючих, спеціалізація, форма власності, рентабельність виробництва і т.д.*

Об'єктом дослідження є окрема людина, ознаки - стать, вік, національність, зріст, вага, професія й т.д.

Якісна ознака (атрибутивна) - ознака, окремі значення якої виражаються у вигляді понять, найменувань.

Приклад. *Професія - лікар; форма власності - приватна; колір очей - блакитний, національність - українець і т.д.*

Кількісна ознака - ознака, певні значення якої мають кількісне вираження.

Приклад. *Зріст - 185 см; вага - 90 кг; ціна одиниці певного товару - 5 грн.; стаж роботи - 10 років і т.д.*

Предметом статистики є **статистичний показник** - узагальнена істотна ознака якого-небудь масового явища в її кількісній і якісній визначеності для конкретних умов місця і часу.

Головні риси методу статистики:

- масове спостереження;
- зведення та групування результатів спостереження;
- визначення та аналіз узагальнюючих показників.

1.2 Статистичне спостереження

Статистичне спостереження — це початкова стадія економіко-статистичного спостереження. Вона являє собою науково організовану роботу зі збирання масових первинних даних про явища й процеси суспільного життя.

Статистичне спостереження має відповідати наступним вимогам:

1. Спостережувані явища повинні мати наукову й практичну цінність, виражати певні соціально-економічні типи явищ.

2. Безпосередній збір масових даних повинен забезпечити повноту фактів, що ставляться до розглянутого питання, тому що явища перебувають у постійній зміні, розвитку. У тому випадку, якщо відсутні повні дані, аналіз і висновки можуть бути помилковими.

3. Для забезпечення вірогідності статистичних даних необхідна ретельна всебічна перевірка (контроль) якості фактів, що збирають.

4. Для створення найкращих умов одержання об'єктивних матеріалів, необхідна наукова організація статистичного спостереження.

Організаційні форми статистичного спостереження:

А) звітність;

Б) спеціально організовані спостереження;

В) реєстри.

А) **Звітністю** називають таку організовану форму статистичного спостереження, при якій відомості надходять у вигляді обов'язкових звітів у визначений термін і по затверджених формах.

При цьому джерелом відомостей, як правило, є первинні облікові записи в документах бухгалтерського й оперативного обліку.

Первинний облік являє собою реєстрацію різних фактів (подій, процесів і т.д.), по мірі їхнього здійснення.

Приклад. Свідоцтво про народження дитини, наряди на відпуск товарів, рахунки-фактури, накладні й т.д.

Класифікація звітності.

За змістом:

- **типова (загальна)** — це звітність, що містить одні й ті самі дані для певної галузі народного господарства й для підприємств усього народного господарства;

- **спеціалізована** - це звітність, що містить специфічні показники окремих галузей народного господарства.

За періодичністю подання:

- **річна** - відомості представляються за рік;

- **поточна** - звітність за всі інші періоди в межах менш року (відповідно квартальна, місячна, тижнева й т.п.).

За способом подання:

- **поштова** - подаються посланцем або надсилаються поштою;

- **термінова** - подаються органам державної статистики у термін, зазначений на бланку форми оперативними засобами зв'язку - телефон, факс, модем, електронна пошта, з подальшим обов'язковим письмовим підтвердженням поштою або кур'єром.

За порядком проходження:

- **централізована** - проходить через систему державної статистики, де обробляється і передається відповідним органам управління;

- **децентралізована** - опрацьовується у відповідних міністерствах чи відомствах, а зведення подають статистичним органам.

Б) **Спеціально організоване статистичне спостереження** - збір відомостей про сфери життя, що не уловлюються звітністю, за допомогою переписів, одноразових обліків й обстежень.

Переписи проводяться періодично чи одноразово і дають повну характеристику масового явища станом на якусь дату або певний момент часу.

Приклад. Перепис населення, сільськогосподарських тварин, плодово-ягідних насаджень і виноградників і т.д.

Одноразові обліки проводяться на містах згідно з інструкцією статистичного органу.

Приклад. Облік залишків сировини й матеріалів, посівних площ, невстановленого устаткування і т.д.

Обстеження - спеціально організований збір певних даних за програмою розробленою дослідниками.

Приклад. Обстеження бюджетів родин, бюджету часу працівника, рівня цін на продукцію і т.д.

В) Реєстри - списки або переліки одиниць об'єкта спостереження із зазначенням ознак, які постійно оновлюються та поповнюються.

Приклад. Реєстр населення - поіменний список мешканців регіону з зазначенням їх паспортних даних; реєстр суб'єктів діяльності; реєстр домашніх господарств; реєстр земельного фонду.

Види статистичного спостереження.

За часом реєстрацій фактів:

- **поточне** - безперервна реєстрація фактів по мірі виникнення. При поточному спостереженні не можна допускати значного розриву між моментом виникнення факту й моментом його реєстрації;

Приклад. Реєстрація актів цивільного стану, облік виробленої продукції, відпуск матеріалів зі складу, виручка магазину, підсумки біржових торгів, продажу на аукціонах.

- **періодичне** - повторюється через певні проміжки часу;

Приклад. Перепис населення - кожні 10 років, перепис обладнання - 2 роки, звітність підприємств - щороку, щокварталу, щомісяця і т.д.

- **одноразове** - проводиться в міру потреби, час від часу, без дотримання строгої періодичності або взагалі проводиться один раз.

Приклад. Маркетингове дослідження щодо адаптації товару до місцевого ринку, перепис плодоягідних насаджень, перепис худоби населення, обстеження думки населення щодо впровадження страхової медицини і т.д.

За способом одержання статистичних даних:

- **безпосередній облік** - передбачає огляд, перелік, вимірювання, зважування;

Приклад. Облік товарних залишків на складі, інвентаризація майна, реєстрація товарних потоків, що перетинають митні кордони, облік грошової маси в банках, оцінка рівня екологічного забруднення в регіоні і т.д.

- **документальний облік** - ґрунтується на даних документів первинного обліку;

Приклад. Складання статистичної звітності на підприємствах, визначаються всі економічні показники макро- та мікрорівнів: обсяги матеріальних, трудових і фінансових ресурсів, розмір доходів, капітальних вкладень, обсяги експорту та імпорту товарів і т.д.

- **опитування респондентів** - спостереження, при якому відповіді на питання формуляра записують зі слів респондента. Буває:

а) **експедиційне** - спеціально підготовлені реєстратори заповнюють формуляри спостереження і одночасно перевіряють правдивість відповідей на питання;

б) **самореєстрація** - респонденти самі заповнюють статистичні формуляри. Працівники статистичних органів лише інструктують їх;

в) **кореспондентське** - здійснюють спеціально підготовленими добровільні кореспонденти, які заповнюють формуляри згідно з інструкцією і передають відомості статистичним органам;

г) **анкетне** - респонденти заповнюють анкети, які їм вручають особисто або надсилають поштою;

д) **інтерв'ю** - респонденти дають довільні відповіді на питання.

За повнотою охоплення одиниць сукупності:

- **суцільне** - обстеженню піддаються все без винятку одиниці досліджуваної сукупності;

Приклад. Всеукраїнський перепис населення, одержання звітності від підприємств й установ і т.д.

- **несуцільне** обстеженню піддаються не всі одиниці досліджуваної сукупності, а тільки заздалегідь встановлена їхня частина. Воно поділяється на спостереження:

а) **основного масиву** - охоплює переважну частину елементів сукупності, використовується, коли з організаційних позицій значно легше провести дослідження на декількох великих об'єктах, чим на сотні дрібних;

Приклад. Дослідження кон'юнктури товарообороту і динаміки цін на продуктових ринках великих міст з метою вивчення їх впливу на купівельну спроможність міського населення; вивчення здоров'я дітей, батьки яких працюють на текстильних підприємствах, коли для аналізу варто відібрати тільки великі комбінати, виключивши дрібні об'єкти і т.д.

б) **вибіркове** - характеристика всієї сукупності фактів дається по деякій їхній частині, відібраної у випадковому порядку;

Приклад. Контроль якості продукції, виявлення продуктивності худоби, попит населення й ступінь його задоволення, обстеження бюджетів родин, екзит-поли (післявиборчі опитування) і т.д.

в) **монографічне** - детальне обстеження лише окремих елементів сукупності.

Приклад. Опис нових технологій, виробництва окремих видів продукції, вивчення і популяризація передового досвіду, а також процесу розвитку окремого трудового колективу, недоліків у роботі і т.д.

Час спостереження.

Встановлення часу спостереження є дуже важливим, і в статистику розрізняють *час*:

- **об'єктивний** - час, до якого відносяться дані спостереження, характеризує той момент, станом на який фіксується розмір явища, або період часу, за який були одержані необхідні дані;

Приклад. Дані про обсяг виробництва та реалізації продукції, розміри споживання, тощо, можна одержати тільки за певний період. Відомості про чисельність населення, вартість основних засобів, поголів'я тварин, кількість лікарняних ліжок і т.д. можна зібрати тільки за станом на певну дату.

- **критичний момент** - певний момент, до якого призначене спостереження. Цей момент обирають, якщо об'єктом спостереження є певний стан явища.

- **суб'єктивний** – час здійснення спостереження, тобто період, протягом якого проводиться реєстрація одиниць сукупності.

Приклад. Критичним моментом перепису населення є час 00.00 годин з 26 на 27 січня.

Очевидно, неможливо переписати все населення країни за один момент, тому крім об'єктивного часу спостереження (в даному випадку - критичного моменту) встановлюється суб'єктивний час, протягом якого реєструються дані, скажімо, 10 днів (з 27 січня по 5 лютого включно).

Тоді, наприклад, 2 лютого, всі дані про членів родини реєструються станом на початок перепису - 27 січня. Коли в родині навіть на день реєстрації, але після початку перепису народилася дитина, переписом вона не враховується.

1.3 Помилки статистичного спостереження

Завданням всякого статистичного спостереження є одержання таких даних, які точніше б відображали дійсність. Відхилення, або різниця між обчисленими показниками й дійсними величинами досліджуваних явищ знайшли відображення в показниках, що називають **помилками, або похибками**.

Залежно від характеру й ступеня впливу на кінцеві результати спостереження, а також виходячи із джерел і причин виникнення неточностей, що допускають у процесі статистичного спостереження, зазвичай виділяють:

1) **Помилки реєстрації** виникають внаслідок неправильного встановлення фактів у процесі спостереження або неправильного їхнього запису й можуть бути як при суцільному, так і несучільному спостереженні. Вони підрозділяються на:

а) випадкові помилки — помилки реєстрації, які можуть бути допущені як опитуваними в їхніх відповідях, так і реєстраторами при заповненні бланків;

Приклад. Описка, обмовка і т.д. Респондент може обмовитись, а реєстратор недочути чи випадково переставити місцями цифри, замість віку 34 роки записати 43 і навпаки.

б) систематичні помилки, які пов'язані з однобічним викривленням ознаки, що фіксується. Залежно від причини можуть бути:

- **навмисні** - утворюються у результаті того, що опитуваний, знаючи дійсне положення справи, свідомо повідомляє невірні дані;

Приклад. Помилки при фіксації сімейного положення по причині різного підходу до відповіді: при переписах у разі цивільного шлюбу чисельність одружених чоловіків завжди менша за чисельність заміжніх жінок; різні види викривлення даних в статистичних звітах, які зроблені навмисно, зокрема, так звані «приписки», і т.д.

- **ненавмисні** - викликаються різними випадковими причинами (через необґрунтованість програми спостереження, нечітке формулювання питань, некомпетентність реєстраторів, неосвіченість респондентів, несправність вимірювальних приладів і т.д.).

Приклад. Помилки пов'язані зі звичками: більшість населення на запитання про вік завжди округлюють його до чисел, які кратні 5 або 10, внаслідок цього виходить, що 35, 40, 45, 50-річних громадян значно більше, ніж 34, 41, 46, 51-річних.

2) Помилки репрезентативності притаманні тільки несучільному спостереженню і виникають у результаті того, що склад відібраної для обстеження частини одиниць сукупності недостатньо повно відображає склад всієї досліджуваної сукупності, хоча реєстрація відомостей по кожній відібраній для обстеження одиниці була проведена точно. Помилки репрезентативності можуть бути:

а) випадкові помилки виникають через те, що при несучільному спостереженні сукупність відібраних одиниць спостереження не відтворює всю сукупність у цілому і навіть при правильно організованому відборі випадково відбираються одиниці з характеристиками у середньому відмінними від всієї сукупності. Випадкові помилки неминуче виникають навіть при самому строгому дотриманні принципів і правил відбору одиниць у вибірккову сукупність.

б) систематичні помилки виникають внаслідок порушення принципів випадкового відбору одиниць досліджуваної сукупності, коли до вибірки потрапили найкращі або найгірші одиниці.

Приклад. Класичним прикладом краху дослідження з причини систематичних помилок є передвиборче опитування, проведене журналом «Літерарі Дайджест» (Literary Digest) в 1936 році. За його результатами на виборах президента США повинен був перемогти Альфред Лендон. Показово те, що у дослідженні, проведеному «Літерарі Дайджест», взяли участь більше 2 млн. респондентів.

На виборах переміг Теодор Рузвельт, перемогу якого передбачили соціологи Дж. Геллап й Ел. Роупер на основі опитування лише 4000 чоловік.

Помилка «Літерарі Дайджест» полягала в тому, що основою для відбору респондентів виступили телефонні книги та списки володарів автомобілів. Телефони та авто в 1936 році мали переважно заможні верстви населення США, більшість яких збиралася голосувати за Альфреда Лендона. Отже отримана вибірка відображала не всіх виборців США, а лише їх специфічну групу.

Вибірка ж Геллапа і Роупера носила випадковий характер і відображала все населення США, що дозволило йому зробити правильний прогноз.

Для виявлення й усунення допущених при реєстрації помилок може застосовуватися **контроль зібраного матеріалу**:

- **арифметичний** - полягає в перевірці точності арифметичних розрахунків, що застосовувалися при складанні звітності або заповненні формулярів обстеження;

- **логічний** - полягає в перевірці відповідей на питання програми спостереження шляхом їхнього логічного осмислення або шляхом порівняння отриманих даних з іншими джерелами по цьому ж питанню.

Завдання для самостійного розв'язання.

Завдання 1.1. Укажіть об'єкт, одиницю сукупності та одиницю спостереження:

- А) проводиться перепис художньої літератури в бібліотеках області;
- Б) аналіз зареєстрованих злочинів, скоєних на території міста;
- В) дослідження кількості серцево-судинних захворювань в області;
- Г) перепис виробничого устаткування на державних підприємствах промисловості;
- Д) проводиться перепис населення України;
- Е) облік наявності касових апаратів у комерційних торговельних пунктах міста.

Завдання 1.2. Визначте обсяг явища та обсяг сукупності:

- А) дослідження продуктивність праці робітників підприємств тваринницької галузі області;
- Б) обстеження проданих квартир на біржах нерухомості міста;
- В) визначення кількості студентів у аграрних вищих навчальних закладах;
- Г) визначається середній розмір грошових вкладів в банках України;
- Д) досліджується рівень заробітної плати робітників торговельних підприємств району;
- Е) дослідження залучення сезонних робітників в сільськогосподарських підприємствах району.

Завдання 1.3. Визначте організаційну форму статистичних спостережень та їх види за ступенем охоплення одиниць сукупності та часом реєстрації фактів:

- А) анкетне опитування пасажирів авіакомпанії, яка обслуговує щотижневі рейси;
- Б) перелік всіх землевласників країни;
- В) обстеження фінансової діяльності лідируючої інвестиційної компанії;
- Г) дослідження рівня цін на сільськогосподарські продукти, що продаються на ринках великих міст;
- Д) реєстрація шлюбів у РАГСах;
- Е) опитування окремих учасників презентації фірми.

Завдання 1.4. Визначте організаційну форму та спосіб спостереження:

- А) інвентаризація нероздержавленого майна на підприємствах громадського харчування;
- Б) перелік усіх зареєстрованих економічних структур із зазначенням їх реквізитів;
- В) обстеження відгуків клієнтів на станціях автосервісу;
- Г) бухгалтерські баланси підприємств народного господарства;
- Д) щомісячний облік обсягів виробництва та реалізації продукції підприємств промисловості;
- Е) маркетингові дослідження насиченості ринку аудіо- та відеотехнікою.

Завдання 1.5. Визначте об'єктивний, суб'єктивний час і критичний момент спостереження.

- А) дані обліку дебіторської заборгованості банків на початок року мають бути подані протягом 15 днів з початку наступного року;
- Б) термін подання річних звітів за 2011 рік - 1-25 січня 2012 р.;
- В) Перший Всесоюзний перепис населення 1926 р. проводився за станом на 12 год. ночі з 16 на 17 грудня й тривала в містах 7 днів, а в сільській місцевості - 14 днів;
- Г) складання платіжного балансу країни за 2010 рік здійснювалося з 10 по 17 січня 2011 року
- Д) строк подання місячних звітів про виробничо-фінансову діяльність державних підприємств - не пізніше 10 числа наступного місяця.

Завдання 1.6. Охарактеризуйте допущені помилки:

- А) у відомостях заводу про випуск продукції за лютий була включена й продукція, що випущена 1 й 2 березня;
- Б) по причині несправності терезів невірно вказується вага тварин;
- В) у переписному листі записано: «років - __, місяців - 10» замість «років - 10, місяців - __ »;

Г) при вибіркового обстеження середнього розміру банківських вкладів громадян серед одиниць сукупності випадково опинився вклад, розмір якого становив 5 млн. доларів;

Д) в регіоні чисельність заміжніх жінок складає 15 млн. осіб, одружених чоловіків - 14,2 млн. осіб;

Е) при визначенні рівня залишкових знань студентів факультету для тестування відібрані групи з найвищими показниками успішності.

Завдання 1.7. Здійснить арифметичний контроль та внесіть виправлення:

А) Маємо наступні дані господарства про посів зернових культур:

Номер бригади	Посівна площа, га			
	пшениця озима	пшениця яра	жито озиме	Всього
1	130	170	150	450
2	90	150	200	440
3	200	320	180	800
4	170	250	300	720
Разом	590	990	830	2410

Б) По заводу в результаті статистичного спостереження одержані наступні дані:

Номер Цеху	Продукція, тис. шт.		Відхилення факту від плану, тис. шт.	Середньорічна чисельність працівників, осіб		Відхилення факту від плану, осіб
	план	факт		план	факт	
1	7750	8050	+300	1070	1090	+20
2	9860	10200	+340	883	901	+28
3	3300	3400	+100	540	526	-14
Разом	20910	21650	+740	2493	2517	+34

Завдання 1.8. Проведіть логічний контроль правильності заповнення переписних листів. У відповідях на які питання допущені помилки?

Переписний лист №1

1. Прізвище, ім'я та по-батькові - Сидоров Петро Іванович
2. Стать - чоловіча
3. Вік (кількість років, що виповнилася) - 19
4. Чи перебуває в шлюбі в даний час - ні
5. Освіта - середня
6. Рід занять - робочий цегляного заводу
7. Скільки років проживає в даному населеному пункті - 42

Переписний лист №2

1. Прізвище, ім'я та по-батькові - Петрова Олена Сергіївна
2. Стать - жіноча
3. Вік (кількість років, що виповнилася) - 51
4. Освіта - середня спеціальна
5. Джерело коштів для існування - стипендія
6. Заняття за містом роботи (посада або робота, що виконується) - бухгалтер

Переписний лист №3

1. Прізвище, ім'я та по-батькові - Васін Іван Іванович
2. Стать - чоловіча
3. Вік (кількість років, що виповнилася) - 3
4. Чи перебуває в шлюбі в даний час - так.
5. Освіта - вища
6. Заняття за містом роботи - зоотехнік

Переписний лист № 4

1. Прізвище, ім'я та по-батькові - Іванова Марія Павлівна
2. Стать - жіноча.
3. Відношення до голови родини - внучка.
4. Вік (кількість років, що виповнилася) - 4.
5. Чи перебуває в шлюбі в даний час - так.
6. Освіта - не має початкової.

Тести для самоконтролю.

1. Окремі об'єкти або явища, що утворюють статистичну сукупність, називаються:

- А) одиниці сукупності
- Б) обсяг сукупності
- В) обсяг явища
- Г) одиниці спостереження

2. Звітність, що опрацьовується у відповідних міністерствах чи відомствах, а зведення подають статистичним органам, називають:

- А) типова
- Б) поточна
- В) децентралізована
- Г) централізована

3. За повнотою охоплення одиниць сукупності спостереження, що охоплює переважну частину елементів сукупності, називають:

- А) монографічне
- Б) основного масиву
- В) вибіркове

Г) суцільне

4. Опитування респондентів, при якому спеціально підготовлені реєстратори заповнюють формуляри спостереження і одночасно перевіряють правдивість відповідей на питання, називають:

- А) самореєстрація
- Б) кореспондентське
- В) анкетне
- Г) експедиційне

5. За часом реєстрації фактів спостереження, що передбачає безперервну реєстрація фактів по мірі виникнення, називають:

- А) поточне
- Б) багаторазове
- В) періодичне
- Г) одноразове

6. Звітність, що містить одні й ті самі дані для певної галузі народного господарства й для підприємств усього народного господарства, називають:

- А) спеціалізована
- Б) поточна
- В) типова
- Г) централізована

7. Ознака, окремі значення якої виражаються у вигляді понять, найменувань, називається:

- А) атрибутивна ознака
- Б) кількісна ознака
- В) типологічна ознака
- Г) статистичний показник

8. За способом одержання статистичних даних спостереження, що передбачає огляд, перелік, вимірювання, зважування, називають:

- А) документальний облік
- Б) перепис
- В) безпосередній облік
- Г) спеціально організоване спостереження

9. Період, протягом якого проводиться реєстрація одиниць сукупності, називається:

- А) контрольний момент
- Б) суб'єктивний час
- В) критичний момент
- Г) об'єктивний час

10. Помилки, які виникають внаслідок порушення принципів випадкового відбору одиниць досліджуваної сукупності, це:

- А) систематичні помилки реєстрації
- Б) систематичні помилки репрезентативності

В) випадкові помилки репрезентативності

Г) випадкові помилки реєстрації

11. Списки або переліки одиниць певного об'єкта спостереження із зазначенням ознак, які постійно оновлюються та поповнюються, називають:

А) переписи

Б) реєстри

В) звітність

Г) обстеження

12. Кількість одиниць, які мають дану ознаку, називають:

А) обсяг явища

Б) обсяг сукупності

В) одиниці спостереження

Г) одиниці сукупності

13. За часом реєстрації фактів спостереження, що повторюється через певні проміжки часу, називають:

А) одноразове

Б) багаторазове

В) періодичне

Г) поточне

14. Ознака, певні значення якої мають кількісне вираження, називається:

А) статистичний показник

Б) атрибутивна ознака

В) кількісна ознака

Г) типологічна ознака

15. За повнотою охоплення одиниць сукупності спостереження, при якому характеристика всієї сукупності фактів дається по деякій їхній частині, відібраної у випадковому порядку, називають:

А) монографічне

Б) суцільне

В) основного масиву

Г) вибіркоче

16. Помилки, що утворюються у результаті того, що опитуваний, знаючи дійсне положення справи, свідомо повідомляє неправильні дані, називають:

А) випадкові помилки реєстрації

Б) випадкові помилки репрезентативності

В) систематичні помилки реєстрації

Г) систематичні помилки репрезентативності

17. Час, до якого відносяться дані спостереження, характеризує той момент або період часу, станом на який були зібрані й зареєстровані дані, називається:

А) критичний момент

- Б) об'єктивний час
- В) контрольний момент
- Г) суб'єктивний час

18. Опитування респондентів, при якому респонденти самі заповнюють статистичні формуляри, а працівники статорганів лише інструктують їх, називають:

- А) анкетне
- Б) кореспондентське
- В) самореєстрація
- Г) експедиційне

19. За способом одержання статистичних даних спостереження, що ґрунтується на даних документів первинного обліку називають:

- А) звітність
- Б) документальний облік
- В) безпосередній облік
- Г) реєстр

20. Помилки, які виникають через те, що при правильно організованому відборі одібраними випадково виявилися одиниці з характеристиками у середньому відмінними від всієї сукупності, називають:

- А) систематичні помилки репрезентативності
- Б) систематичні помилки реєстрації
- В) випадкові помилки репрезентативності
- Г) випадкові помилки реєстрації

21. Що з нижченаведеного **не є** організаційною формою статистичного спостереження:

- А) спеціально організоване спостереження
- Б) перепис
- В) реєстр
- Г) звітність

22. За повнотою охоплення одиниць сукупності спостереження, при якому обстеженню піддаються всі без винятку одиниці досліджуваної сукупності, називають:

- А) перепис
- Б) суцільне
- В) поточне
- Г) реєстр

23. Певний момент, до якого призначене статистичне спостереження, називається:

- А) суб'єктивний час
- Б) об'єктивний час
- В) контрольний момент
- Г) критичний момент

24. Перевірка точності арифметичних розрахунків, що застосовувалися при складанні звітності або заповненні формулярів обстеження:

- А) об'єктивний контроль
- Б) суб'єктивний контроль
- В) арифметичний контроль
- Г) логічний контроль

25. Опитування респондентів, при якому респонденти дають довільні відповіді на поставлені питання, називають:

- А) анкетне
- Б) інтерв'ю
- В) кореспондентське
- Г) самореєстрація

26. Звітність, що проходить через систему державної статистики, де обробляється і передається відповідним органам управління називають:

- А) типова
- Б) поточна
- В) централізована
- Г) спеціалізована

27. За часом реєстрації фактів спостереження, що проводиться в міру потреби, час від часу, без дотримання строгої періодичності, називають:

- А) поточне
- Б) періодичне
- В) одноразове
- Г) багаторазове

28. Помилки, які можуть бути допущені як опитуваними в їхніх відповідях, так і реєстраторами при заповненні бланків, називають:

- А) випадкові помилки репрезентативності
- Б) випадкові помилки реєстрації
- В) систематичні помилки репрезентативності
- Г) систематичні помилки реєстрації

29. За повнотою охоплення одиниць сукупності спостереження, при якому проводиться детальне обстеження лише окремої одиниці сукупності, називають:

- А) суцільне
- Б) вибіркоче
- В) основного масиву
- Г) монографічне

30. Організаційну форму статистичного спостереження, при якій відомості надходять у визначений термін і по затверджених формах, називають:

- А) організоване спостереження
- Б) реєстр

В) звітність

Г) перепис

31. Загальний розмір явищ для всієї сукупності називають:

А) обсяг явища

Б) спостереження

В) статистична сукупність

Г) обсяг сукупності

32. Контроль зібраного матеріалу, що полягає в перевірці відповідей на питання програми спостереження шляхом їхнього логічного осмислення або шляхом порівняння отриманих даних з іншими джерелами по цьому ж питанню

А) об'єктивний

Б) суб'єктивний

В) арифметичний

Г) логічний

33. Який вид статистичного спостереження за способом одержання статистичних даних застосовується економістом крупного сільськогосподарського підприємства при підготовці річного звіту:

А) безпосередній облік

Б) спеціально організоване спостереження

В) документальний облік

Г) перепис

34. Який вид статистичного спостереження за повнотою охоплення одиниць сукупності застосовується при Всеукраїнському перепису населення:

А) монографічне

Б) основного масиву

В) суцільне

Г) вибіркоче

2. ПОДАННЯ СТАТИСТИЧНИХ ДАНИХ: ТАБЛИЦІ, ГРАФІКИ, КАРТИ

2.1 Статистичні таблиці

Таблиця є найбільш раціональною, наочною і компактною формою подання статистичного матеріалу. Проте не всяка таблиця є статистичною. Таблиця множення, опитувальний лист соціологічного обстеження і т.д. можуть носити табличну форму, але не бути статистичними таблицями.

В економічній й управлінській роботі статистичні таблиці застосовуються дуже часто. Тому необхідно навчитися правильно їх складати й аналізувати.

Значення статистичних таблиць полягає в тому, що вони дозволяють охопити матеріали статистичного зведення в цілому. Статистична таблиця, власне кажучи, є системою думок про досліджуваний об'єкт, що викладають цифрами на основі визначеного порядку в розташуванні систематизованої інформації.

Статистична таблиця – форма наочного і систематизованого зображення результатів зведення і обробки статистичних матеріалів, яка являє собою ряд горизонтальних і вертикальних ліній, що пересікаються та утворюють по горизонталі рядки, а по вертикалі - графи (стовпці, колонки), які в сукупності становлять остов таблиці.

Таблиця, що складається з рядків і граф, які ще не заповнені цифрами, називається **макетом таблиці** (рис. 2.1).

Кожна статистична таблиця має *підмет і присудок*.

Підмет таблиці - це та статистична сукупність, ті об'єкти або їх частини, які характеризують статистичну сукупність.

Присудок таблиці - це система показників, якими характеризується підмет таблиці.

Примітка. Зазвичай підмет розташовується ліворуч, у вигляді найменування горизонтальних рядків, а присудок - праворуч, у вигляді найменування вертикальних граф.

У таблиці можуть бути підведені підсумки по графах і рядках.

Обов'язкова частина таблиці - заголовок, що показує, про що йде мова в таблиці, до якого місця й часу вона відноситься.

Залежно від побудови підмета таблиці поділяють на:

1) **Прості** - таблиці, у підметі яких немає групувань, а дається лише об'єкт чи перелік будь-яких об'єктів, територіальних одиниць, періодів часу;



Рис. 2.1 Макет таблиці

За характером матеріалу прості таблиці поділяють на:

- **монографічні** – характеризують не всю сукупність одиниць досліджуваного об’єкту, а тільки одну яку-небудь групу з нього, яка виділяється за певною, заздалегідь сформульованою ознакою (табл.2.2);

Таблиця 2.1

Виробництво основних сільськогосподарських культур у 2012 році

Показники	Зернові та зернобобові культури	Цукрові буряки	Насіння соняшнику	Картопля	Овочі відкритого ґрунту
Валовий збір, тис. т	29287,6	17064,8	4173,7	19102,3	6835,1

- **перелікові** – у підметі подається перелік показників або одиниць досліджуваної сукупності (табл. 2.2);

Таблиця 2.2

Переробка винограду на виноматеріали за сортами в Україні в 2012 році

Найменування сортів винограду	Фактично перероблено винограду на виноматеріали		Середня масова концентрація цукрів, г/дм ³
	тон	у % до загального підсумку	
Аліготе	47493,2	14,4	186,0
Бастардо магарацький	8100,4	2,5	197,0
Каберне Совіньйон	37092,2	11,3	196,5
Кокур білий	2768,9	0,8	205,0
Мерло	8895,3	2,7	195,8
Мускат (білий, рожевий і т.п.)	26076,0	7,9	190,2
Одеський чорний	5119,6	1,6	200,5
Піно (білий, чорний, сірий)	9074,6	2,8	190,1
Рислінг	13281,2	4,0	182,6
Ркацителі	30830,1	9,4	191,6
Сапераві північний	7310,0	2,2	198,8
Совіньйон	18186,7	5,5	185,9
Сухоліманський	3259,9	1,0	179,0
Трамінер рожевий	2874,4	0,9	189,7
Фетяска	4778,2	1,4	177,4
Шардоне	30683,0	9,3	185,6
Інші	73806,0	22,3	183,9

- **хронологічні** - у підметі наводяться періоди часу або дати, а в присудку - ряд показників (табл. 2.3);

Таблиця 2.3.

Вищі навчальні заклади (дані на початок навчального року)

Навчальний рік	Кількість закладів		Кількість студентів у закладах, тис. осіб	
	I-II рівнів акредитації	III-IV рівнів акредитації	I-II рівнів акредитації	III-IV рівнів акредитації
2004/05	619	347	548,5	2026,7
2005/06	606	345	505,3	2203,8
2006/07	570	350	468,0	2318,6
2007/08	553	351	441,3	2372,5
2008/09	528	353	399,3	2364,5
2009/10	511	350	354,2	2245,2
2010/11	505	349	361,5	2129,8
2011/12	501	345	356,8	1954,8
2012/13	489	334	345,2	1824,9

- територіальні - у підметі наводиться перелік територій (районів, областей и т.п.) (табл. 2.4);

Таблиця 2.4

Зовнішня торгівля України послугами з країнами ЄС за 2011 рік,
млн. дол. США

Держава	Експорт	Імпорт	Сальдо
Австрія	200,0	234,1	-34,1
Бельгія	327,4	55,8	271,6
Болгарія	37,9	20,4	17,5
Велика Британія	707,9	647,8	60,1
Греція	80,8	31,1	49,7
Данія	86,7	46,2	40,5
Естонія	165,8	25,0	140,8
Ірландія	60,4	20,9	39,5
Іспанія	85,8	32,4	53,4
Італія	121,8	59,9	61,9
Кіпр	439,3	830,7	-391,4
Латвія	42,9	27,2	15,7
Литва	28,1	29,6	-1,5
Люксембург	7,8	11,0	-3,2
Мальта	85,2	2,4	82,8
Нідерланди	108,7	128,0	-19,3
Німеччина	416,4	436,2	-19,8
Польща	134,6	142,9	-8,3
Португалія	9,9	44,3	-34,4
Румунія	28,2	13,0	15,2
Словаччина	49,4	33,6	15,8
Словенія	22,3	4,0	18,3
Угорщина	84,2	101,0	-16,8
Фінляндія	46,7	46,1	0,6
Франція	113,0	171,9	-58,9
Чехія	50,6	59,4	-8,8
Швеція	67,9	140,0	-72,1
Всього	3609,8	3394,9	214,9

- територіально-хронологічні – підмет містить перелік територій, а присудок - періоди або моменти часу (табл. 2.5).

Таблиця 2.5.

Доходи населення за регіонами України, млн. грн.

Області	Роки				
	2007	2008	2009	2010	2011
АР Крим	22878	30386	32046	39393	45562
Вінницька	19264	25170	26813	33602	38990
Волинська	10943	14662	15177	19137	22584
Дніпропетровська	50955	68420	72138	88922	101868
Донецька	67057	92213	96596	118223	135599
Житомирська	15132	20216	20655	26124	30069
Закарпатська	12103	16033	16492	20841	24446
Запорізька	26424	35628	37019	45779	52272
Івано-Франківська	14779	20143	21023	26504	31224
Київська	23145	31666	34358	42732	48990
Кіровоградська	11602	15452	16149	20213	23443
Луганська	28967	39861	41916	51338	58619
Львівська	31361	42220	43813	54838	63602
Миколаївська	14772	19950	20723	26034	29800
Одеська	29186	39367	42422	52924	61435
Полтавська	20283	27024	28239	34462	39299
Рівненська	12555	16865	17458	22362	26144
Сумська	14965	19819	20318	24918	28347
Тернопільська	11367	14999	15608	19587	22712
Харківська	37835	51971	54519	67102	77316
Херсонська	12070	16206	16731	20978	24096
Хмельницька	15551	20596	21526	26987	31336
Черкаська	15177	20424	21351	26194	29646
Чернівецька	9178	12214	12619	16114	18682
Чернігівська	13560	17908	18917	23179	26277
м. Київ	77435	109781	122516	143903	164057
м. Севастополь	4745	6447	7144	8785	10338

- **переліково-хронологічні** - у підметі таблиці наводиться ряд показників, а присудок – періоди або моменти часу (табл. 2.6);

Таблиця 2.6

Споживання продуктів харчування в домогосподарствах у середньому за місяць у розрахунку на одну особу

Продукти	Роки				
	2007	2008	2009	2010	2011
М'ясо і м'ясопродукти, кг	5,1	5,1	4,8	5,1	5,1
Молоко і молочні продукти, кг	22,1	22,6	19,8	19,1	18,9
Яйця, шт.	20	20	20	20	20
Риба і рибопродукти, кг	1,9	2,1	1,8	1,8	1,7
Цукор, кг	3,2	3,4	3,2	3,0	3,1
Олія та інші рослинні жири, кг	1,7	1,8	1,9	1,8	1,8
Картопля, кг	8,3	8,4	8,0	7,6	7,7
Овочі та баштанні, кг	8,7	9,3	10,1	9,5	10,1
Фрукти, ягоди, горіхи, виноград, кг	3,6	3,7	3,6	3,7	3,7
Хліб і хлібні продукти, кг	9,6	9,6	9,3	9,3	9,2

2) **Групові** – таблиці, в яких підмет розподілено на групи за однією кількісною (табл. 2.7) або якісною ознакою (табл. 2.8).

Таблиця 2.7

Групування заводів галузі промисловості за потужністю

Групи заводів за вартістю основних засобів, млн. грн.	Кількість заводів	Основні засоби, млн. грн.	Чисельність робітників, осіб	Валова продукція, млн. грн.
1,0 - 2,2	3	5,0	820	5,6
2,2 - 3,4	9	27,4	3150	26,5
3,4 - 4,6	5	19,4	1945	23,0
4,6 - 5,8	3	15,2	1295	15,9
5,8 - 7,0	4	27,1	1420	43,8
Разом	24	94,1	8630	114,8

Таблиця 2.8

Забезпеченість населених пунктів водою

Показники	2011
Кількість населених пунктів, які мають водопровід	7543
в тому числі міст	456
селищ міського типу	775
сільських населених пунктів	6312

Такі таблиці дають більше інформативний матеріал для аналізу досліджуваних явищ, завдяки утвореним у їхньому підметі групам по істотній ознаці або виявленню зв'язку між рядом показників.

3) **Комбінаційні** - таблиці, підмет яких розподілено на групи за двома і більше ознаками, і кожна з груп, що побудована за однією ознакою, розбивається у свою чергу на підгрупи по якому-небудь іншому признаку.

Якщо процес групування відбувається з розрахунком системи показників, то групи, виділені за однією ознакою розміщують в лівій частині в підмет таблиці, а підгрупи за іншою ознакою розташовують у додатковій графі підмета (табл. 2.9).

Таблиця 2.9

Комбінаційне групування підприємств галузі промисловості за потужністю та рівнем використаної праці

Групи заводів за вартістю основних засобів, млн. грн.	Групи заводів за чисельністю робітників, осіб	Кількість заводів	Валова продукція, млн. грн.	
			всього	в середньому на 1 завод
1,0 - 2,2	До 300	2	4	2
	300-400	1	1,6	1,6
	більше 400	-	-	-
2,2 - 3,4	До 300	3	6,4	2,13
	300-400	4	8,6	2,15
	більше 400	1	5,5	2,25
3,4 - 4,6	До 300	-	-	-
	300-400	2	5,3	2,65
	більше 400	3	17,7	5,9
4,6 - 5,8	До 300	-	-	-
	300-400	1	3,5	3,5
	більше 400	2	12,4	6,2
5,8 - 7,0	До 300	2	24,8	12,4
	300-400	1	9,6	9,6
	більше 400	1	9,4	9,4
Разом		24	114,8	4,78

Якщо завдання групування обмежуються лише розподілом сукупності по певних групах, то групи, виділені за однією ознакою розміщують в лівій частині таблиці як підмет а підгрупи за іншою ознакою розташовують у верхній частині таблиці як присудок (табл. 2.10).

Таблиця 2.10

Розподіл автомобілів за пробігом і тривалістю експлуатації, %

Загальний пробіг автомобілів, тис. км	Тривалість експлуатації				Разом
	До 3-х років	Від 3-х до 8-ми включно	Від 8-ми до 10-ти включно	Понад 10 років	
До 100	28,4	1,0	-	-	29,4
101...250	4,4	15,1	1,9	-	21,4
251...400	-	14,8	3,4	0,8	19,0
Понад 400	-	9,3	7,0	13,9	30,2
Разом	32,8	40,2	12,3	14,7	100

За структурною побудовою присудка розрізняють таблиці:

1) **З простою розробкою присудка** – показник, що визначає присудок, не підрозділяють на підгрупи, і підсумкові значення одержують шляхом простого сумування значень за кожною ознакою окремо, незалежно одна від іншої (табл. 2.11);

Таблиця 2.11

Загальні обсяги надходження цукрових буряків на переробні підприємства в 2012 році, тис. т

Показники	Усього	У тому числі від:		
		сільськогосподарських підприємств	господарств населення	інших господарських структур
Надійшло цукрових буряків – усього	17380,8	12221,9	327,9	398,1
у тому числі: закуплено переробними підприємствами	6441,2	6202,6	126,2	112,4
надійшло на давальницьких засадах	6506,7	6019,3	201,7	285,7
надійшло власно вирощених буряків	4432,9	х	х	х
Середня ціна закупівлі, грн. за т	431,2	431,4	443,9	406,0

2) **Зі складною розробкою присудка** – ознака, що формує присудок, поділяється на підгрупи. Виходить більш повна та детальна характеристика об'єкту, але може привести до безмірного збільшення розмірності статистичних таблиць, що, в свою чергу, знижує наочність, читання та аналіз (табл. 2.12).

Таблиця 2.12

Розподіл населення за сімейним станом за підсумками
Всеукраїнського перепису населення 2001 року

	Частка осіб даного сімейного стану у відповідному віці, %							
	ЧОЛОВІКИ				ЖІНКИ			
	перебувають у зареєстрованому або незареєстрованому шлюбі	ніколи не перебували у шлюбі	удівці	розлучені (розійшлися)	перебувають у зареєстрованому або незареєстрованому шлюбі	ніколи не перебували у шлюбі	удови	розлучені (розійшлися)
Постійне населення у віці 15 років і старші	64,6	24,3	3,9	6,8	54,0	16,1	19,2	10,4
15-19	1,0	98,6	0,0	0,1	7,3	92,1	0,0	0,4
20-24	24,8	73,1	0,0	1,6	48,1	45,8	0,3	5,3
25-29	60,8	32,0	0,1	6,6	70,1	16,3	0,9	12,4
30-39	77,4	10,9	0,4	10,9	75,2	6,1	2,5	15,8
40-49	82,5	4,9	1,2	11,1	73,6	3,6	6,4	16,1
50-59	84,5	2,7	3,4	9,0	66,0	3,1	16,8	13,8
60-69	83,6	1,6	9,3	5,3	51,4	3,6	36	8,9
70 років і старші	71,8	0,8	25,1	2,1	23,6	5,6	66,4	4,2

При складанні таблиці рекомендовано дотримуватися *ряду правил*:

- чітко формулювати найменування, що повинно точно відображати мету складання таблиці;
- ясно й коротко формулювати назву рядків і граф таблиці;
- дотримуватися послідовності розташування показників присудка;

- вказувати одиниці виміру; якщо вони однакові, то одиниці виміру виносяться в заголовок;
- нумерувати графи;
- мати підсумкові показники;
- якщо в таблиці проводиться зіставлення з яким-небудь роком, то в заголовку, у дужках, відбивається рік зіставлення;
- територіальні, адміністративні утворення перелічуються за алфавітом;
- дані за багато періодів розташовуються в хронологічному порядку;
- якщо в таблиці надані абсолютні й відносні показники за кілька років, то спочатку приводяться абсолютні, потім відносні показники за один рік, потім так само за наступний рік;
- якщо показник не може мати значення, ставиться знак «х»;
- якщо значення ознаки в якій-небудь клітинці невідомо, ставляться знаки «...» або «н.в.» (немає відомостей);
- якщо значення показнику дорівнює нулю, ставиться знак «-».

Дотримання наведених правил побудови та оформлення статистичних таблиць робить їх основним засобом представлення, обробки та узагальнення статистичної інформації про стан і розвиток аналізованих соціально-економічних явищ.

2.2 Статистичні графіки та їх класифікація

Важливе значення при вивченні економічної діяльності має графічне зображення статистичної інформації. Правильно побудований графік робить статистичну інформацію більш виразною, вона краще запам'ятовується й зручно сприймається. У економічній діяльності графічний метод знаходить широке застосування для ілюстрації сформованого становища на ринку товарів і послуг, кон'юнктури попиту та пропозиції, реклами товарів.

Застосування графіків у статистиці нараховує більш ніж двохсотлітню історію. Основою положником графічного методу в економічній статистиці вважають англійського економіста У. Плейфейра (1731- 1798). У своїх роботах він уперше застосував способи графічного зображення статистичних даних (лінійні, стовпчикові, секторні й інші діаграми).

Статистичні графіки – це наочне зображення статистичних даних та їх співвідношень за допомогою точок, географічних карт або геометричних фігур і ліній, яким надано умовного значення.

У статистичному графіку виділяють наступні *основні елементи*:

- поле графіка;
- графічний образ;
- просторові й масштабні орієнтири;
- експлікація графіка.

Поле графіка є місце, на якому він виконується. Це аркуші паперу, географічні карти, план місцевості й т.п. Поле графіка характеризується його форматом (розмірами й пропорціями сторін). Розмір поля графіка залежить від його призначення.

Графічний образ — це символічні знаки, за допомогою яких зображуються статистичні дані (лінії, точки, прямокутники, квадрати, кола й т.д.). Як графічний образ виступають й об'ємні фігури. Іноді в графіках використовуються негеометричні фігури у вигляді силуетів або малюнків предметів.

Просторові орієнтири визначають розміщення графічних образів на полі графіка. Вони задаються координатною сіткою або контурними лініями й ділять поле графіка на частини, що відповідають значенням досліджуваних показників. Найчастіше використовують прямокутну систему координат, де положення кожної точки визначається довжиною двох перпендикулярів, опущених з цієї точки на вісі абсцис та ординат.

Масштабні орієнтири (масштаб, масштабна шкала, масштабний знак) застосовують для вивчення розмірів геометричних та інших знаків та надають графічним образам кількісну значимість, що передається за допомогою системи масштабних шкал.

Масштаб графіка - умовна міра переведення чисельної величини в графічну (наприклад, 1 см = 100 тис. грн.). При цьому, чим довший відрізок лінії, прийнятої за числову одиницю, тим крупніший масштаб.

Масштабною шкалою є лінія, окремі точки якої читаються як певні числа.

Вона складається з 3 елементів:

- лінія, що є носієм або опорою шкали
- позначки шкали - риски або точки, розміщені на носії шкали
- числові позначення, які відповідають поділу шкали

Шкала, по якій відраховують рівні досліджуваних показників, як правило, починається з 0. Останнє число, що наноситься на шкалу, трохи перевищує максимальний рівень, відлік якого проводиться по цій шкалі. При побудові графіка допускається розрив масштабної шкали. Цей прийом використовується для зображення статистичних даних, що мають значення лише в певних значеннях.

Експлікація графіка — це пояснення його змісту. Вона включає в себе:

- **заголовок графіка** - у короткій і чіткій формі пояснює основний зміст зображуваних даних.

- **написи вдовж масштабних шкал** - цифрові позначення шкали доповнюються назвою показників з вказівкою одиниць виміру.

- **пояснювальні написи окремих елементів графічного образу** — умовні позначення, тобто конфігурація, штрихування або колір та текст, що робить можливим читання графіка.

Класифікація статистичних графіків.

При всьому своєму різноманітті графіки в статистиці класифікуються за рядом ознак.

За загальним призначенням статистичні графіки бувають:

- аналітичні;
- ілюстративні;
- інформаційні.

За видом поля графіку розрізняють:

- діаграми
- статистичні карти.

За формою графічного образу виділяють:

- **лінійні** - графічними образами є лінії;
- **плоскі** - графічними образами є геометричні фігури: прямокутники, окружності, трикутники, багатокутники;
- **об'ємні** - графічними образами є стереометричні фігури: паралелепіпеди, конуси, піраміди, циліндри, багатогранники;
- **точкові** - як графічні образи застосовується сукупність точок;
- **зображувальні (фігурні)** - як графічні образи застосовуються будь-які художні зображення;

За типом шкали статистичні графіки бувають:

- лінійні рівномірні (арифметичні);
- лінійні нерівномірні (функціональні);
- криволінійні.

Узагальнення багатогранної практики використання графічного методу в зображенні показників комерційної діяльності дозволяє сформулювати *ряд вимог* до методики побудови статистичних графіків:

1. При графічному зображенні кількісних показників економічної діяльності (обсяг, склад і динаміка товарообороту, стан товарної пропозиції, товарних запасів, витрат обігу, прибутку й т.д.) переважніше використовують лінійні, стовпчикові або кругові діаграми, що мають найбільшу в порівнянні з об'ємними або площинними фігурами наочність і дохідливість.

2. У загальному розташуванні на полі графіка графічних образів останні з метою правильного читання й розуміння досліджуваного показника розміщуються зліва направо. При цьому масштабні орієнтири графіка по горизонтальній шкалі (вісь абсцис), як правило, розміщуються від його нижньої частини. Для вертикальної шкали (вісь ординат) масштабні орієнтири звичайно розміщуються в лівій частині графіка.

3. До графіків по можливості варто включати вихідні дані до їхньої побудови. Якщо це недоцільно, то вихідні дані повинні в табличній формі супроводжувати графік. Це зумовлює довіру до графічного зображення

показників економічної діяльності, підвищує пізнавальне значення статистичних графіків.

4. Всі буквені й цифрові значення повинні розташовуватися на графіку так, щоб їх легко можна було відлічити від початку масштабної шкали. Ряди цифрових даних, що відображають зміни показників економічної діяльності в часі, розміщуються в строгій хронологічній послідовності й обов'язково по осі абсцис.

5. Загальною вимогою графічного методу зображення статистичних показників є те, що факторні ознаки розміщуються на горизонтальній шкалі графіка і їхні зміни читаються зліва направо, а результативні ознаки - по вертикальній шкалі й читаються знизу вгору. Це підвищує аналітичне значення статистичних графіків. При цьому важливо, щоб заголовок графіка був би коротким, але досить чітко пояснював основний його зміст.

2.3 Діаграми

Діаграма (від грец. *diagramma* - зображення, малюнок, креслення) - графічне зображення, що наочно показує співвідношення між порівнюваними величинами.

Вибір та способи побудови діаграм

1. Діаграми порівняння - включають групу діаграм, що застосовуються для співставлення величин за їх кількісними даними за допомогою прямокутної системи координат. Найбільш поширеними є:

а) **стовпчикова** – відображує статистичні дані у вигляді прямокутників однакової ширини, розміщених вертикально на горизонтальній прямій.

Спосіб побудови. Висота кожного стовпчика пропорційна величині даних, що відображуються, та характеризує порівнювані величини. У середині стовпчиків або поряд можуть бути записані цифри. Ширина стовпчиків має бути однаковою, а відстань між ними має дорівнювати половині (іноді 1/4) ширини самих стовпчиків. Зазвичай стовпчики роблять або чорними чи кольоровими, або заштрихованими. Вертикальна шкала діаграми завжди починається з нуля та охоплює весь діапазон даних.

Приклад. Середню зарплату штатних працівників за видами економічної діяльності у січні 2013 року можна порівняти між собою та з прожитковим мінімумом на одну працездатну особу і з середнім рівнем по економіці (рис.2.2).

Стовпчикова діаграма може бути *об'ємною* (рис. 2.3а), у т.ч. стовпчики можуть зображуватися у вигляді циліндрів, пірамід чи конусів і утворювати відповідно *циліндричну* (рис. 2.3б), *пірамідальну* (рис. 2.3в) або *конусну* (рис. 2.3г) діаграму.

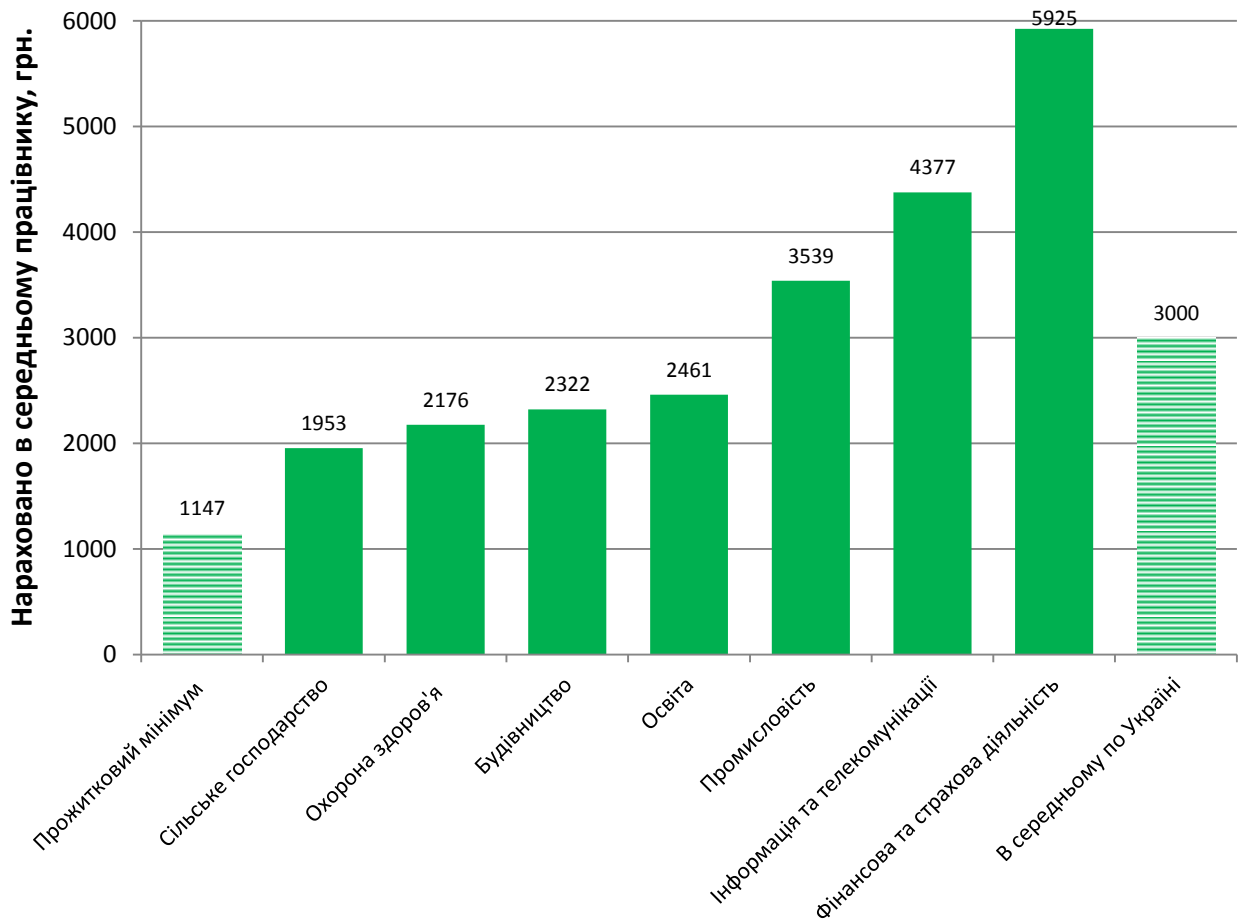


Рис. 2.2. Заробітна плата працівників за деякими видами економічної діяльності за січень 2013 року

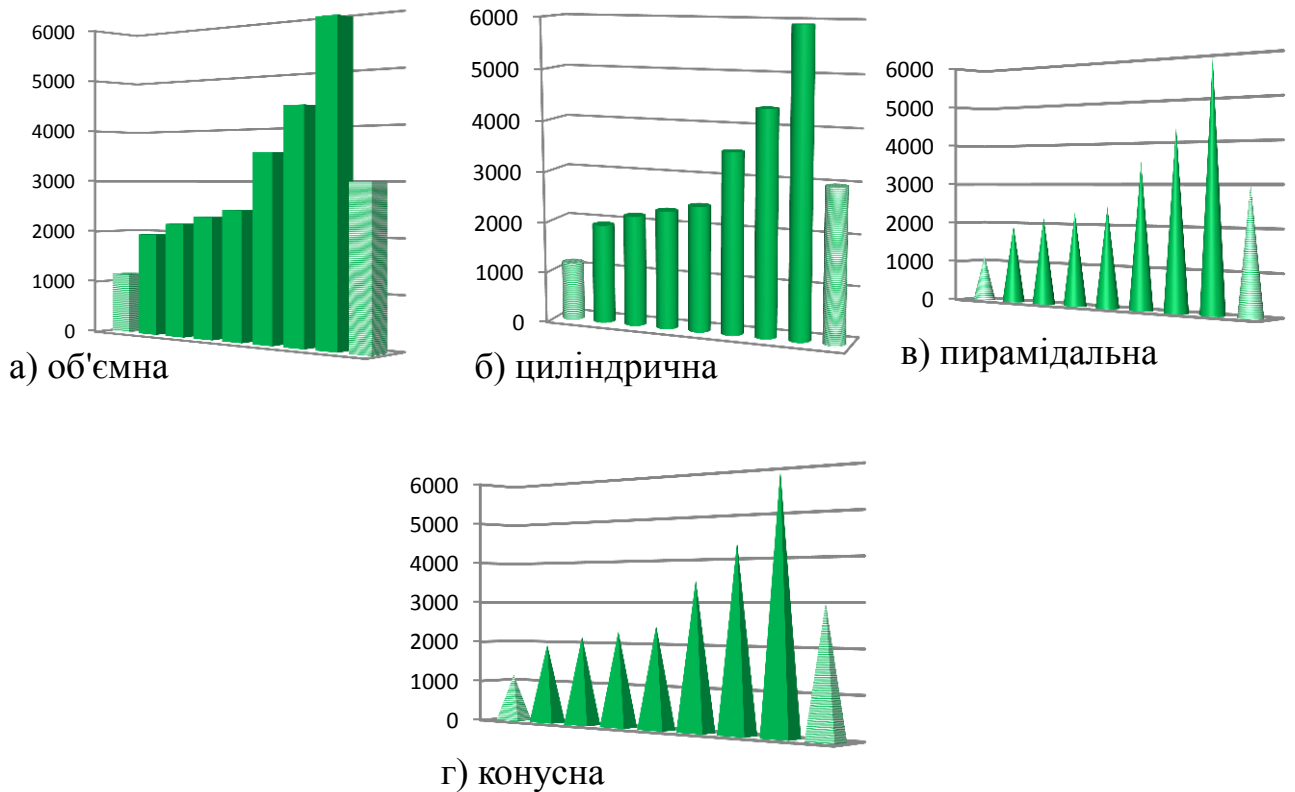


Рис. 2.3. Різновиди стовпчикової діаграми.

Якщо окрім порівняння загальної величини показника необхідно порівняти його складові частини, використовують *накопичувальну* стовпчикову діаграму.

Приклад. При аналізі зовнішньоторговельного балансу України порівняння загальних обсягів експорту та імпорту можна здійснити з одночасним порівнянням обсягів експортованих та імпортованих товарів та послуг за допомогою *накопичувальної* діаграми (рис. 2.4)

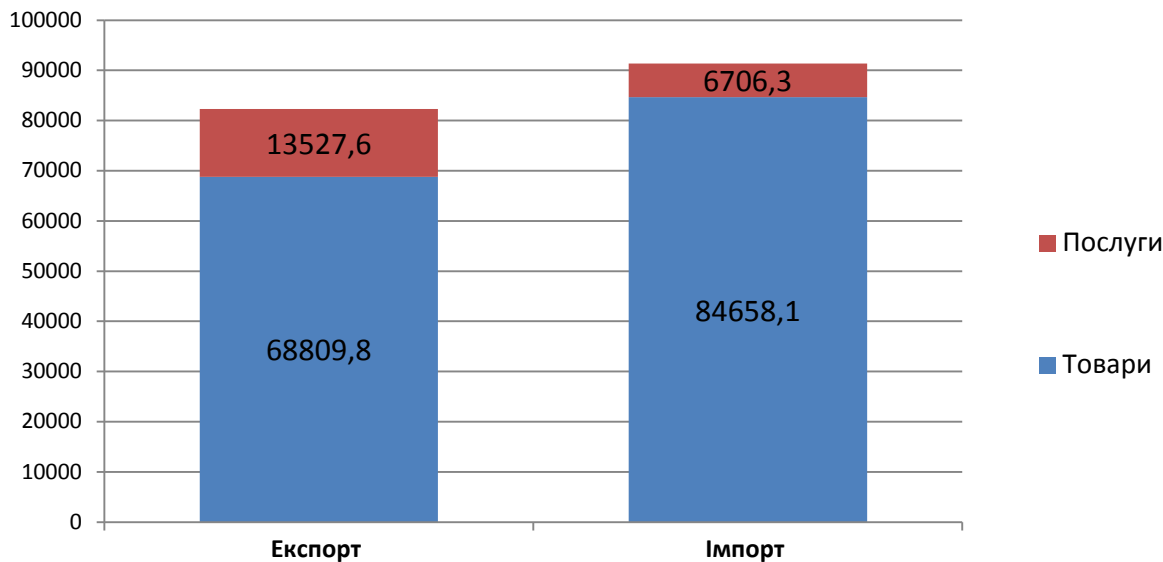


Рис. 2.4. Зовнішньоторговельний баланс України в 2012 р., млн. дол. США

б) **стрічкова, або смугова** - на відміну від стовпчикової, смуги розміщуються горизонтально на вертикальній прямій.

Спосіб побудови. Довжина відображує величину даних. Горизонтальна шкала починається з нуля, її розрив не допускається.

Приклад. Прямі інвестиції нерезидентів в Україні (акціонерний капітал та боргові інструменти) на 1 січня 2012 р склали 58624,9 млн. дол. США. Порівняти величину прямих інвестицій з країн Євросоюзу, СНД та інших країн світу можна за допомогою стрічкової діаграми (рис. 2.5).

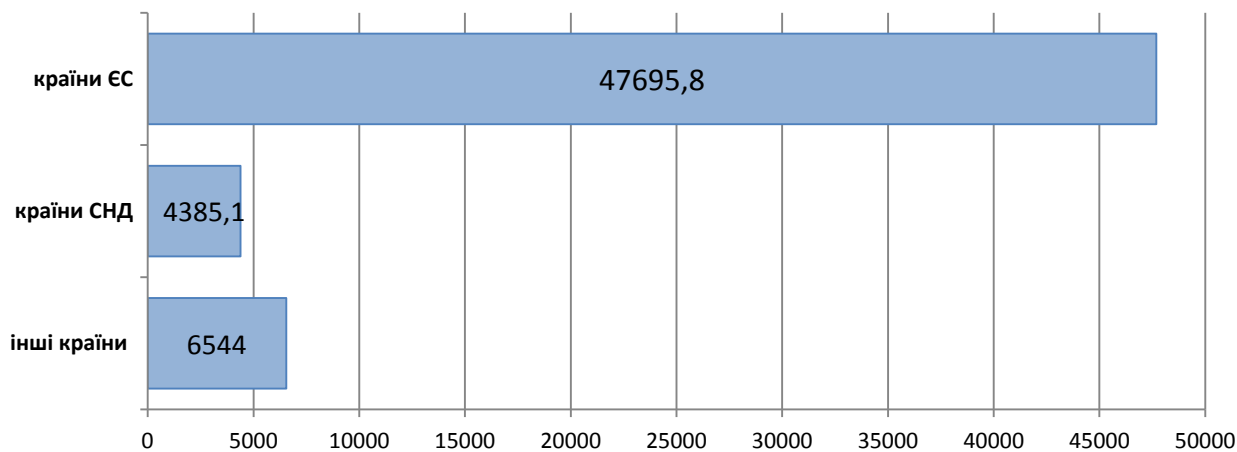


Рис. 2.5. Прямі іноземні інвестиції в економіку України на початок 2012 року, млн. дол. США

в) **квадратна** – величина зображеного явища виражається розміром площі квадрату.

Спосіб побудови. Із порівнюваних статистичних величин добути квадратні корені, а потім побудувати квадрати зі сторонами, пропорційними одержаним результатам.

Приклад. Вантажоборот залізничного транспорту в Україні в 2012 році склав 237,3 млрд. т-км, трубопровідного – 112,5 млрд. т-км, автомобільного – 39,2 млрд. т-км, водного 5,3 млрд. т-км.

Квадратні корені з величин вантажооборотів складуть:

- для залізничного транспорту – 15,4;
- для трубопровідного транспорту – 10,6;
- для автомобільного транспорту – 6,3;
- для водного транспорту – 2,3.

Відповідно до масштабу *сторони квадратів* дорівнюватимуть:

- для залізничного транспорту – 3,1 см;
- для трубопровідного транспорту – 2,1 см;
- для автомобільного транспорту – 1,3 см;
- для водного транспорту – 0,46 см (рис.2.6).

Масштаб: 1 см = 5 млрд. т-км

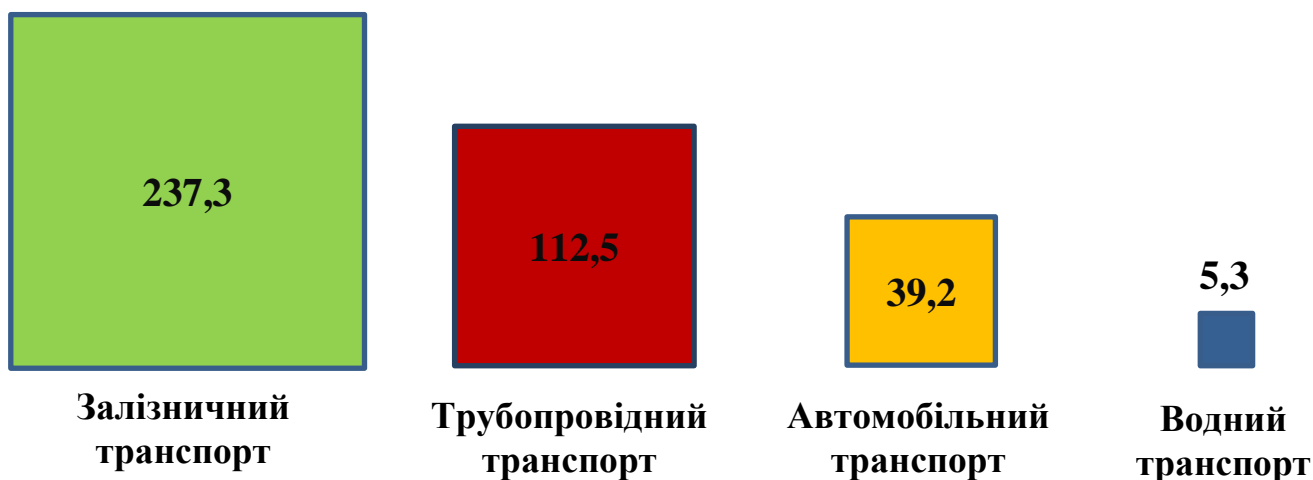


Рис. 2.6. Вантажоборот основних видів транспорту України у 2012 році, млрд. т-км

г) **кругова** – за змістом аналогічна квадратній.

Спосіб побудови. На графіку зображують кола, площа яких пропорційна квадратним кореням із зображених величин. Для цього добувається квадратний корінь із розміру явища та будується радіус, пропорційний обчисленій величині.

Приклад. Skorистуємося попередніми даними про вантажооборот різних видів транспорту. Квадратні корені з величин вантажооборотів вже визначені у попередньому прикладі.

Відповідно до масштабу *радіуси кіл* дорівнюватимуть:

- для залізничного транспорту – 1,54 см;
- для трубопровідного транспорту – 1,06 см;
- для автомобільного транспорту – 0,63 см;
- для водного транспорту – 0,23 см (рис.2.7).

Масштаб: 1 см = 10 млрд. т-км

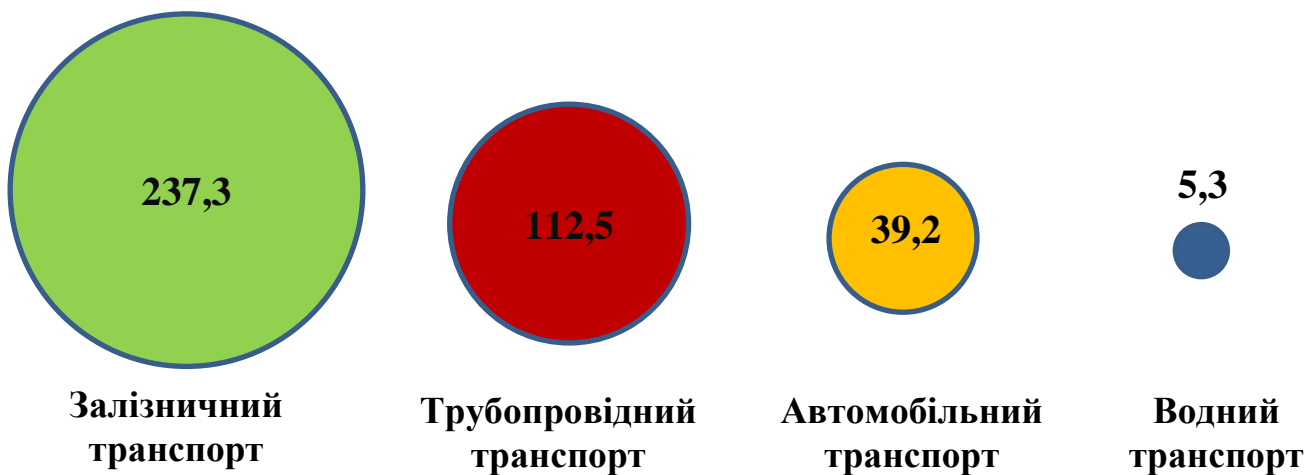


Рис. 2.7. Вантажооборот основних видів транспорту України у 2012 році, млрд. т-км

д) **фігурна (зображувальна)** - відображує статистичні дані у вигляді певних довільно обраних символів.

Спосіб побудови 1. Статистичні величини зображуються фігурами різних розмірів. Для цього необхідно точно визначити, що відповідає зображеним числам: лінійний розмір фігури (її висота, довжина) чи її площа.

Приклад. Порівняння поголів'я великої рогатої худоби та свиней можна збільшувати та зменшувати розміри фігурок тварин (рис. 2.8).

Примітка. Фігурні діаграми, на яких розміри фігурок збільшуються чи зменшуються довільно у всіх напрямках, не можуть вважатися статистично коректними. Якщо зміна розмірів фігурок відбувається в довжину або в висоту, їх основа повинна бути однаковою.

Приклад. Популярне видання «USA Today» в 2004 році представило своїм читачам діаграму, яка порівнювала кількість книг, що написані про американські бейсбольні команди (рис. 2.9).

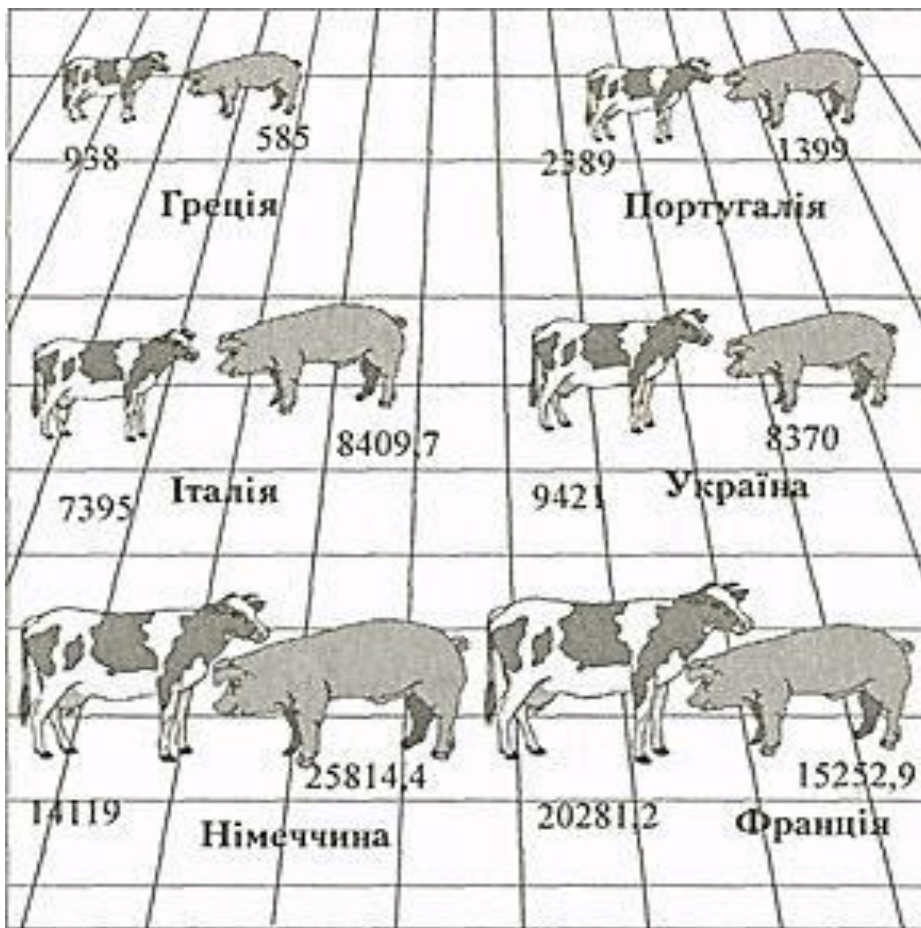


Рис. 2.8. Поголів'я ВРХ та свиней у 2001 р., тис. голів

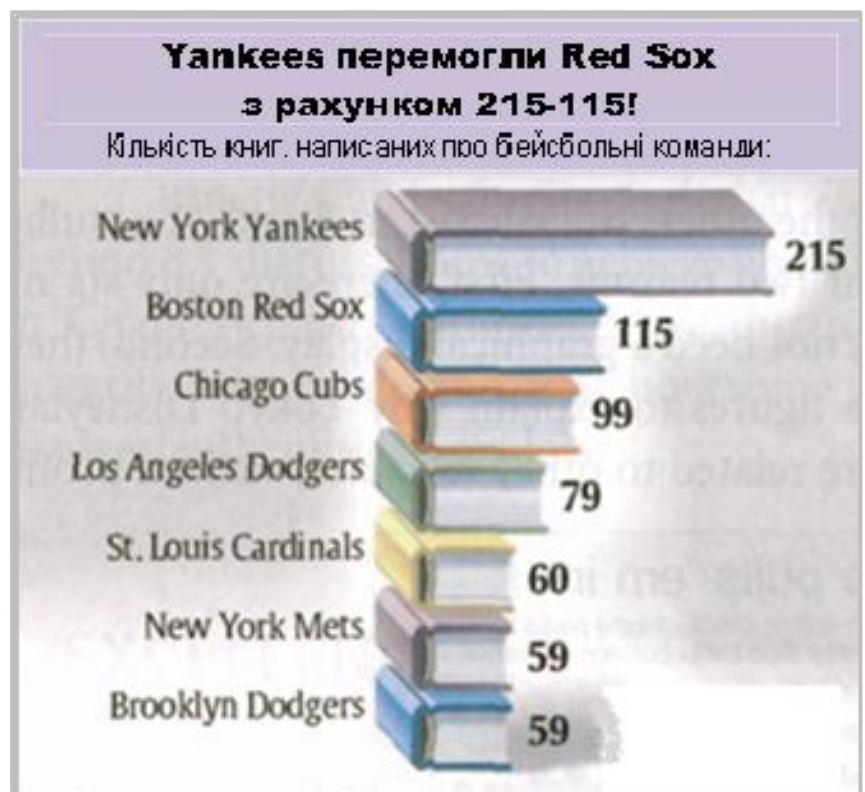


Рис. 2.9. Книги про бейсбольні команди США

Спосіб побудови 2. Статистичні величини зображуються різною чисельністю фігур, кожна з яких дорівнює певному числу одиниць і порівняння здійснюється за чисельністю фігурок. При цьому допускається подрібнення символу до половини та навіть до чверті. (населення - фігурки людей, виробництво зерна – снопи, колоси; гроші – монети, банкноти; вантажообіг - вагони, мішки; і т.п.

Приклад. Населення трьох країн логічно буде зобразити за допомогою чоловічків, кількість яких згідно до умовних позначень, і характеризуватиме чисельність населення (рис. 2.10).

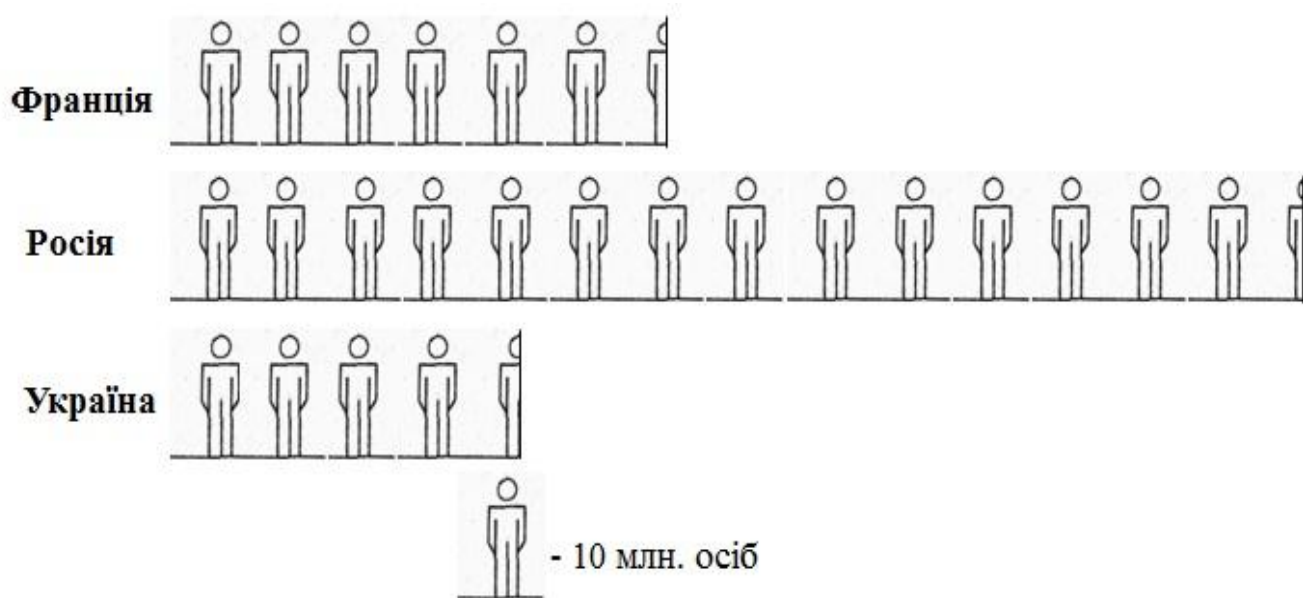


Рис. 2.10. Населення держав світу у 2012 році

Примітка. Недоліком даного способу є дещо приблизна точність визначення розміру показника при подрібненні художнього образу.

2. Діаграми структури – групи діаграм, що зображують склад цілого, поділеного на частини.

а) кругова секторна

Спосіб побудови. Площа кола приймається за 100%. Частки цілого (явища) виражають у градусах кола: оскільки 1% дорівнює 3,6 градуси ($360^\circ : 100\% = 3,6^\circ$), то відповідні показники, виражені у відсотках до підсумку, треба помножити на 3,6 для визначення центральних кутів. Зазвичай сектори розміщують за розміром, а інші - послідовно, за годинниковою стрілкою. Для того, щоб відрізнити сектори використовують штрихування або різні кольори.

Приклад. Питому вагу окремих видів продукції у загальній вартості реалізованої продукції господарства відображено за допомогою кругової секторної діаграми (рис. 2.11).

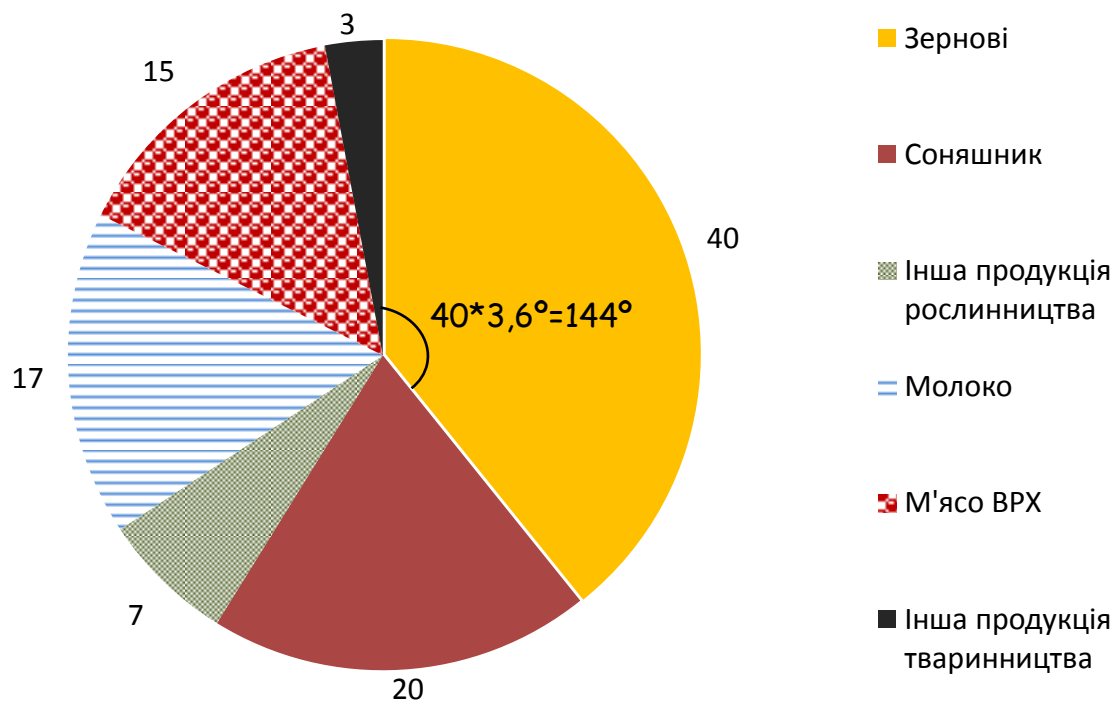


Рис. 2.11. Структура товарної структури господарства, у % до загального підсумку

Кругова секторна діаграма також може бути *об'ємна* (рис. 2.12а), *розрізана* (рис. 2.12б), або мати форму тора – *кільцева* (рис. 2.12в).

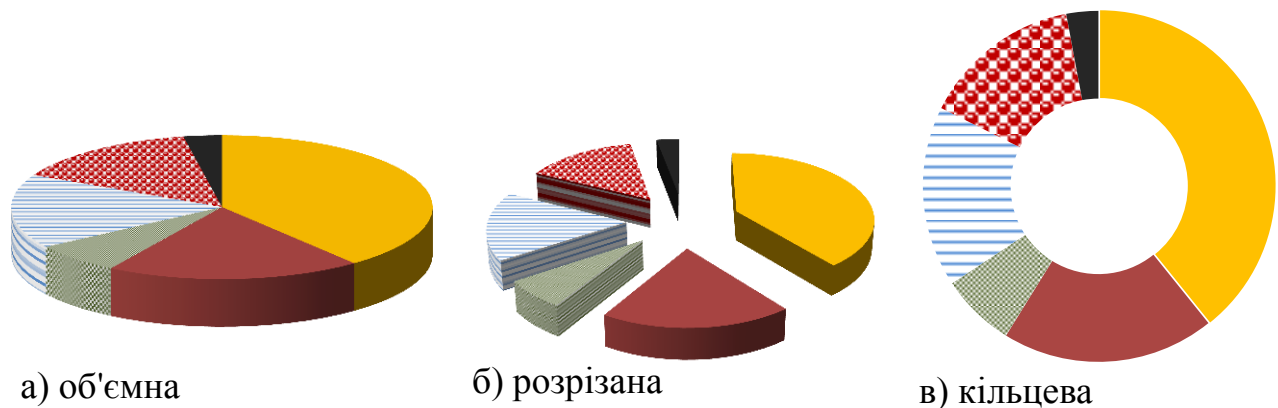


Рис. 2.12. Різновиди кругової секторної діаграми

б) **стовпчикова** – призначена для відображення великої кількості складових частин, а також дозволяє їх порівняння в часі або за різними об'єктами на одному полі.

Спосіб побудови. Висота стовпчика приймається за 100%, а його частини виражають питому вагу. Їх також штрихують або фарбують.

Приклад. Структуру зовнішньої торгівлі послугами, зокрема частки основних видів імпортованих послуг та їхні структурні зміни за останні два роки, зображено за допомогою стовпчикової діаграми структури (рис. 2.13).

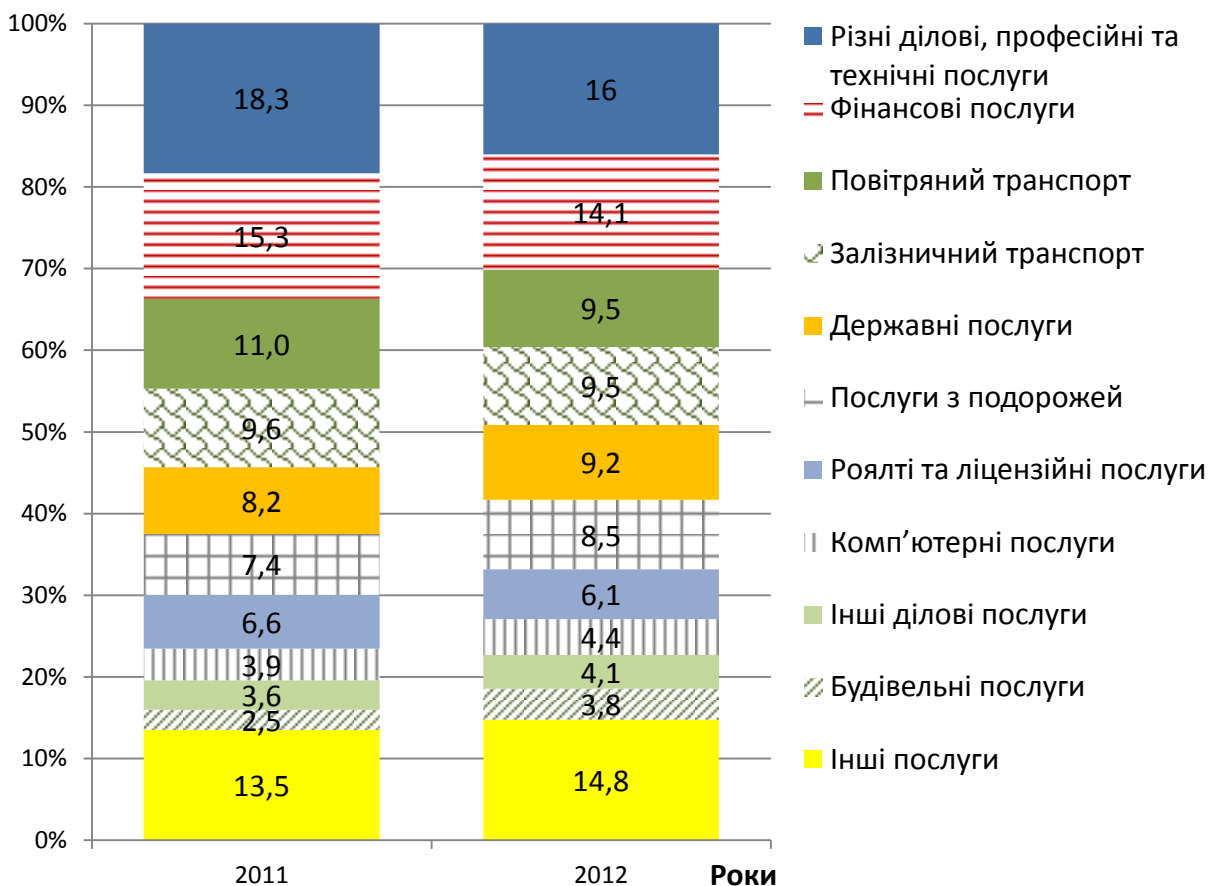


Рис. 2.13. Структура імпорту послуг в Україну

в) стрічкова (смугова)

Спосіб побудови. За 100% приймається довжина горизонтальної смуги.

Приклад. Структурні зміни, що відбулися в 2011-2012 рр. в галузі рибальства та добування інших водних живих ресурсів відображені за допомогою стрічкової діаграми (рис. 2.14).

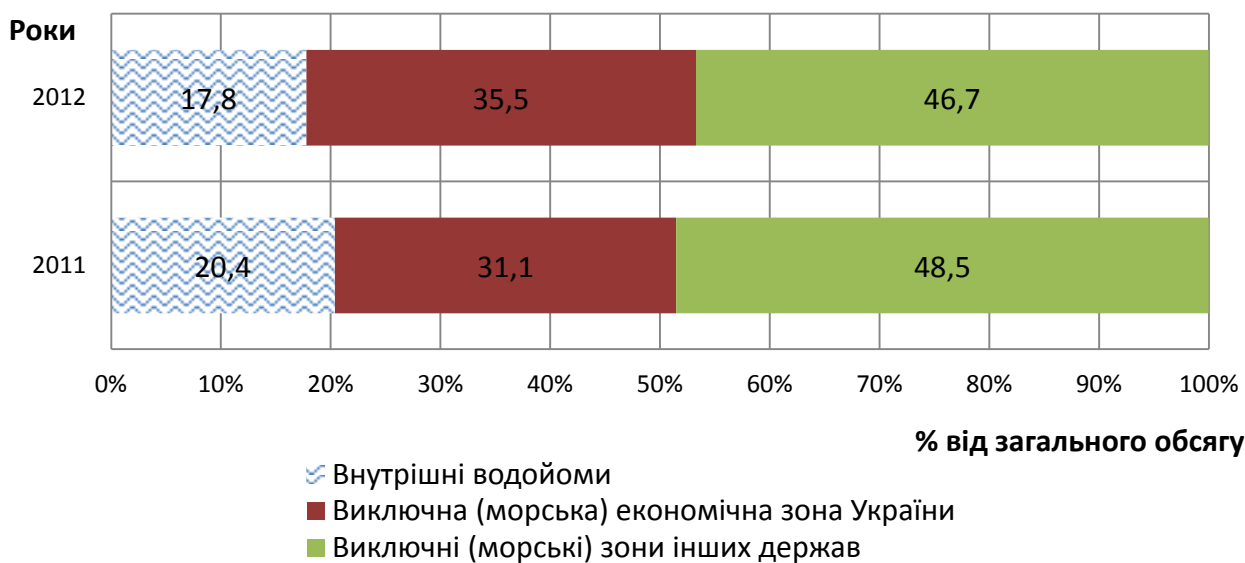


Рис. 2.14. Структура вилову риби та добування інших водних живих ресурсів за видами водоймищ

3. Діаграми динаміки - графічне зображення змін в часі (динаміки) досліджуваних явищ.

а) **лінійна** – динаміку явища зображують за допомогою ліній, побудованих в системі прямокутних координат.

Спосіб побудови. На горизонтальній вісі ОХ відкладають однакові відрізки, що становлять періоди або моменти часу (роки, дні), а на вертикальну вісь ОУ в певному масштабі наносять величину, що характеризує рівні ряду. Далі всі точки з'єднують, одержуючи лінію, що відображає розвиток явища за даний період. На графіку можна відобразити закономірності розвитку декількох явищ з однаковими одиницями виміру. Лінії можуть відрізнятися за малюнком (суцільна, пунктирна, жирна, тощо), маркером (точки помічають квадратами, кружечками, трикутниками, тощо), кольором.

Приклад. Дані 8 переписів щодо кількості населення України представлені за допомогою лінійної діаграми (рис. 2.15).

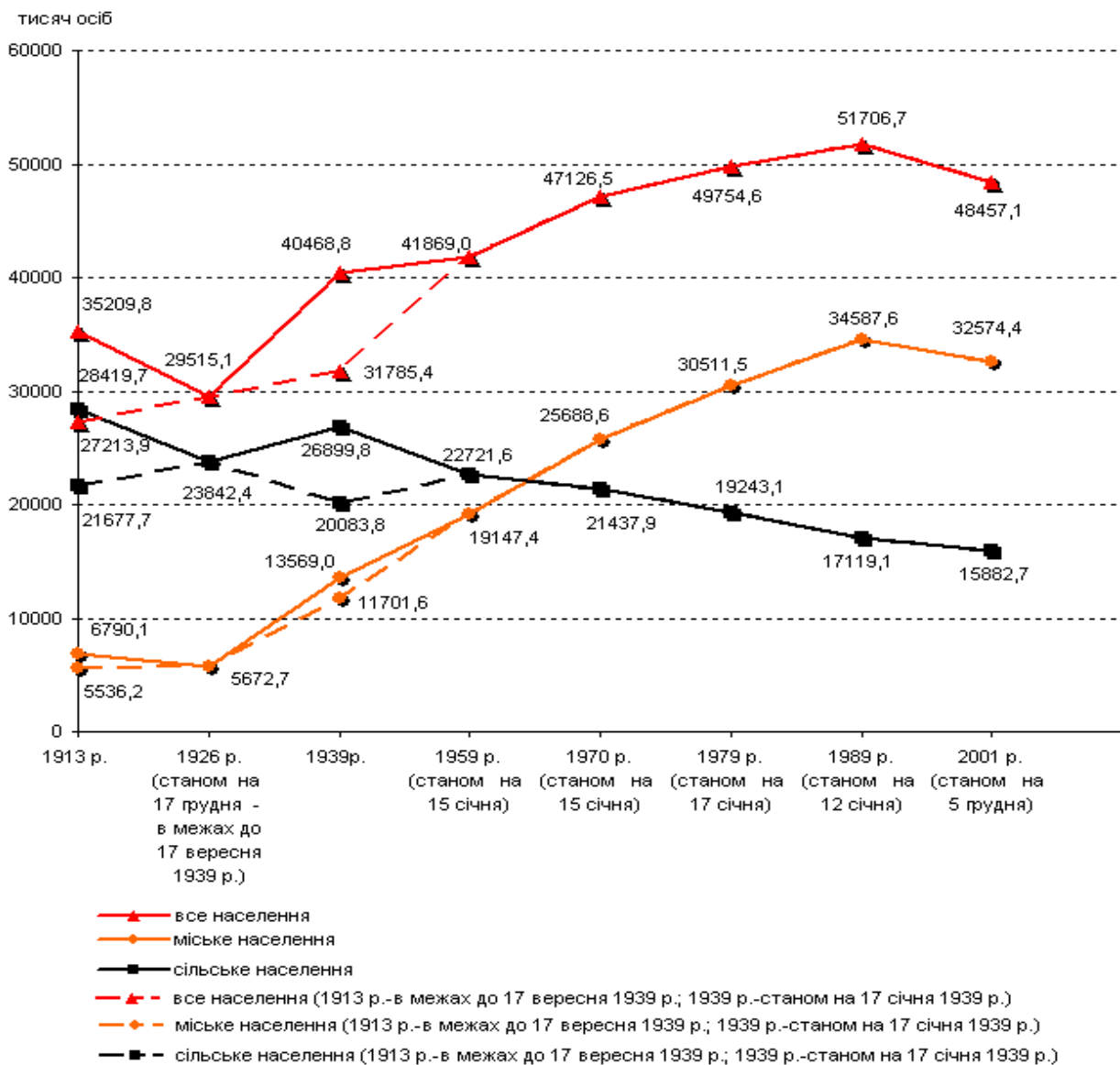


Рис. 2.15. Динаміка кількості наявного населення України за даними переписів населення різних років

б) **стовпчикова чи фігурна** – для графічного зображення динамічних рядів з невеликою кількістю даних.

Спосіб побудови. Аналогічний діаграмі порівняння, але на вісі ОХ відкладений масштаб часу.

Приклад. Маємо наступні дані про виробництво овочів в Україні:

Роки	2000	2001	2002	2003	2004	2005
Валовий збір овочів, тис. т	5821	5907	5827	6538	6964	7295
Роки	2006	2007	2008	2009	2010	2011
Валовий збір овочів, тис. т	8058	6835	7965	8341	8122	9833

Діаграма динаміки матиме вигляд (рис. 2.16):

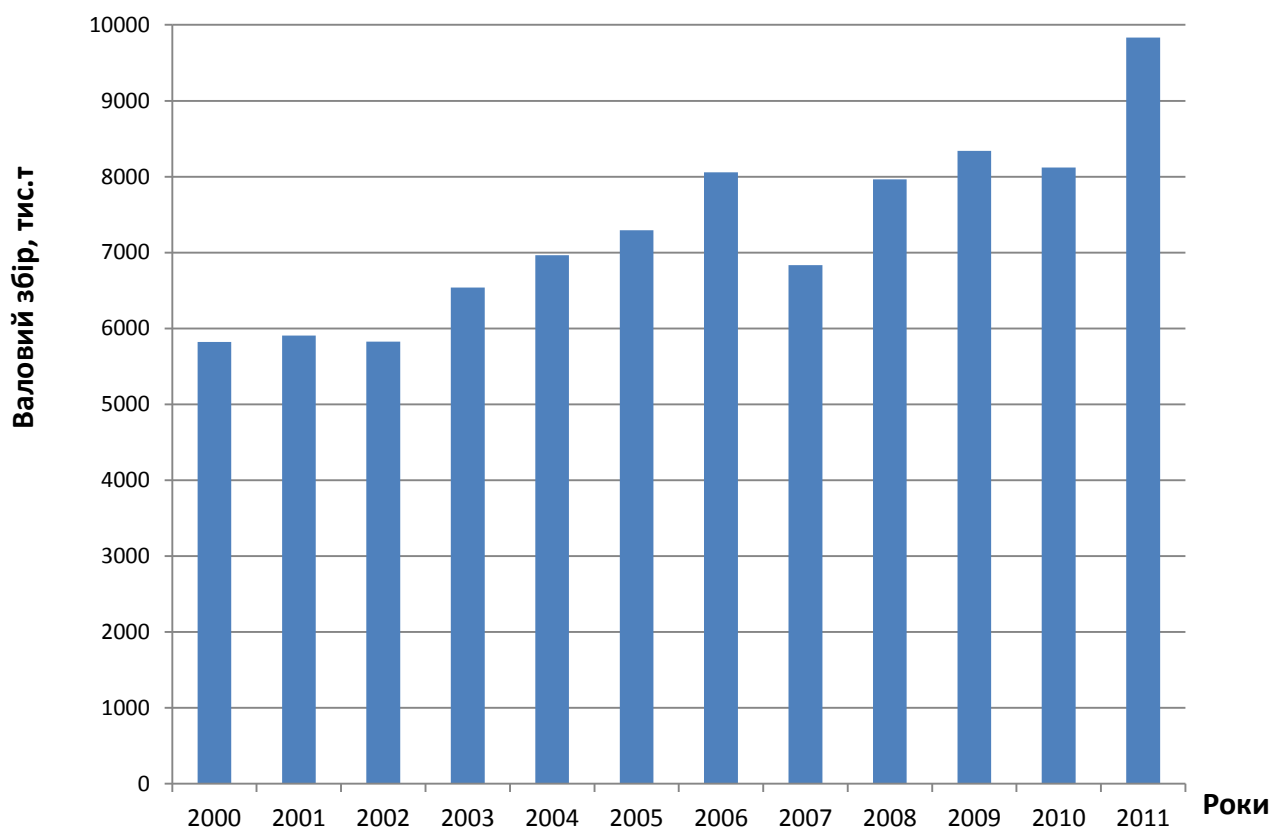


Рис. 2.16. Динаміка виробництва овочів в Україні

Приклад. Департамент з питань навколишнього середовища Канадського уряду «Environment Canada» сповістив про підвищення рівнів випадіння снігу взимку 1992-1993 рр. за допомогою фігурної діаграми (рис. 2.17). А відоме фінансове видання «The Value Line Investment Survey» в 1993 році опублікувало дані про ефективність інвестування в акції компанії «Coca-Cola». Динаміку також було зображено за допомогою фігурної діаграми (рис. 2.18).



Рис. 2.17. Динаміка рівня випадіння снігу в Торонто.



Рис. 2.18. Ефективність акцій компанії «Coca-Cola».

Примітка. Треба зауважити, що при всій своїй кумедності діаграма 2.17 є некоректною з точки зору зображення динаміки статистичних показників – фігурки сніговиків довільно зростають і в ширину, і в висоту, і навіть мають різну кількість «сегментів» - снігових куль. Тоді як діаграма 2.18 ефективно відображає динаміку прибутковості акцій, використовуючи пропорційне зростання висоти пляшечок «Coca-Cola».

в) **радіальна, або пелюсткова** – основою слугує полярна система координат, що використовується для зображення явищ, які періодично змінюються в часі (переважно сезонних коливань)

Спосіб побудови.

- За вісь ОХ приймають коло, за ОУ - радіуси, пункт відліку - центр кола.
- Кількість періодів визначає кількість радіусів.
- Довжина радіусу кола визначається середньою величиною явища.
- З центру на радіусах періодів, де величина ознаки менша за середню, відкладають відповідні відрізки, а радіуси періодів, де величина ознаки більша за середню, подовжують до потрібної довжини.
- Послідовно з'єднуються точки кінців відрізків.
- На одній системі координат можна зобразити кілька явищ.

Приклад. Маємо наступні дані про виробництво м'яса в регіоні:

Місяць	Вироблено м'яса, млн. т	Місяць	Вироблено м'яса, млн. т	Місяць	Вироблено м'яса, млн. т
Січень	35	Травень	18	Вересень	76
Лютий	32	Червень	40	Жовтень	79
Березень	43	Липень	35	Листопад	70
Квітень	18	Серпень	56	Грудень	50

Радіальна діаграма матиме вигляд (рис. 2.19):

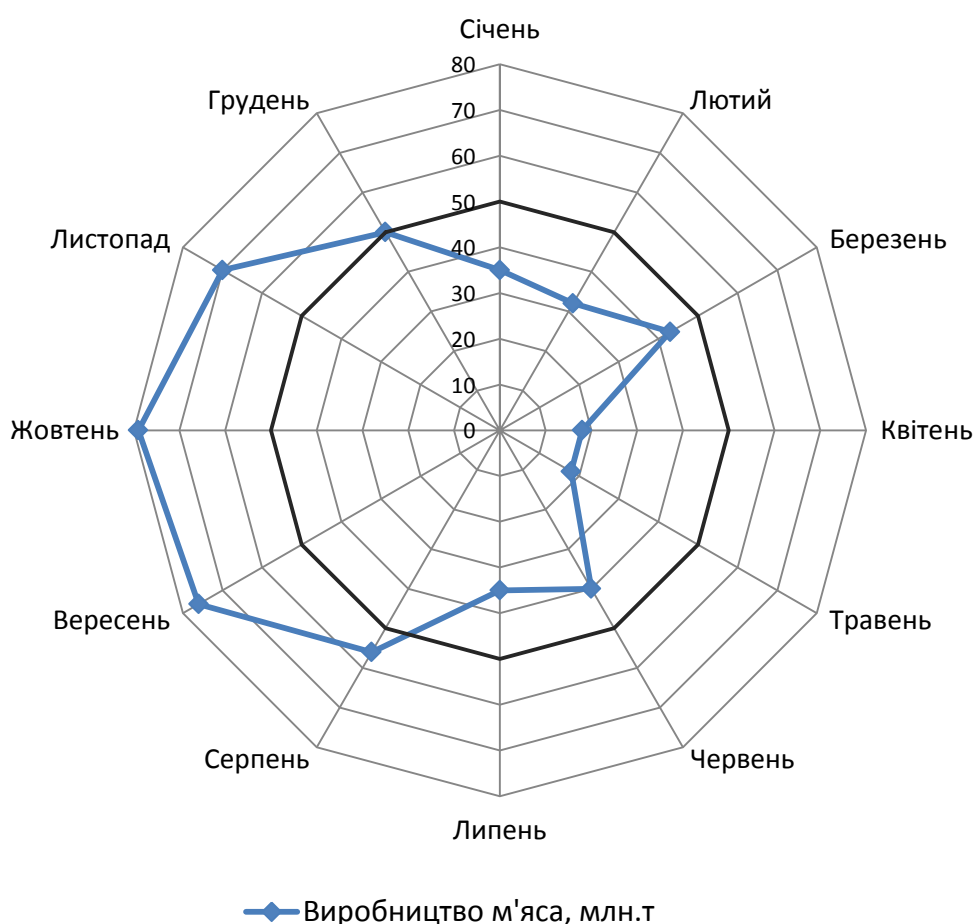


Рис. 2.19. Сезонні коливання виробництва м'яса в регіоні

4. Діаграми взаємозв'язку – показують залежність однієї ознаки від іншої.

а) **точкова** - показує фактичну залежність однієї ознаки від іншої.

Спосіб побудови. На горизонтальній вісі ОХ відкладають фактичні значення факторної ознаки, а на вертикальну вісь ОУ в певному масштабі наносять фактичні величини результативної ознаки.

б) **лінійна** - показує функціональну залежність однієї ознаки від іншого (лінію тренду), може бути представлена у вигляді прямої або кривої (параболи, гіперболи, тощо).

Спосіб побудови. На горизонтальній вісі ОХ відкладають значення фактичні значення факторної ознаки, а на вертикальну вісь ОУ в певному масштабі наносять вирівняні (теоретичні) за відповідним рівнянням величини результативної ознаки.

Примітка. Частіше за все точкова та лінійна діаграма надаються на одному полі графіку.

Приклад. Маємо наступні дані про собівартість та ціну зерна у господарстві за 7 років:

Ознаки	Роки						
	2006	2007	2008	2009	2010	2011	2012
Виробнича собівартість зерна, грн./ц	32,3	28,79	59,54	93,94	88,75	101,4	193,5
Середня ціна реалізації, грн./ц	51,5	80,9	83	79,4	98,2	135,3	154,7

Діаграма взаємозв'язку матиме вигляд (рис. 2.20):

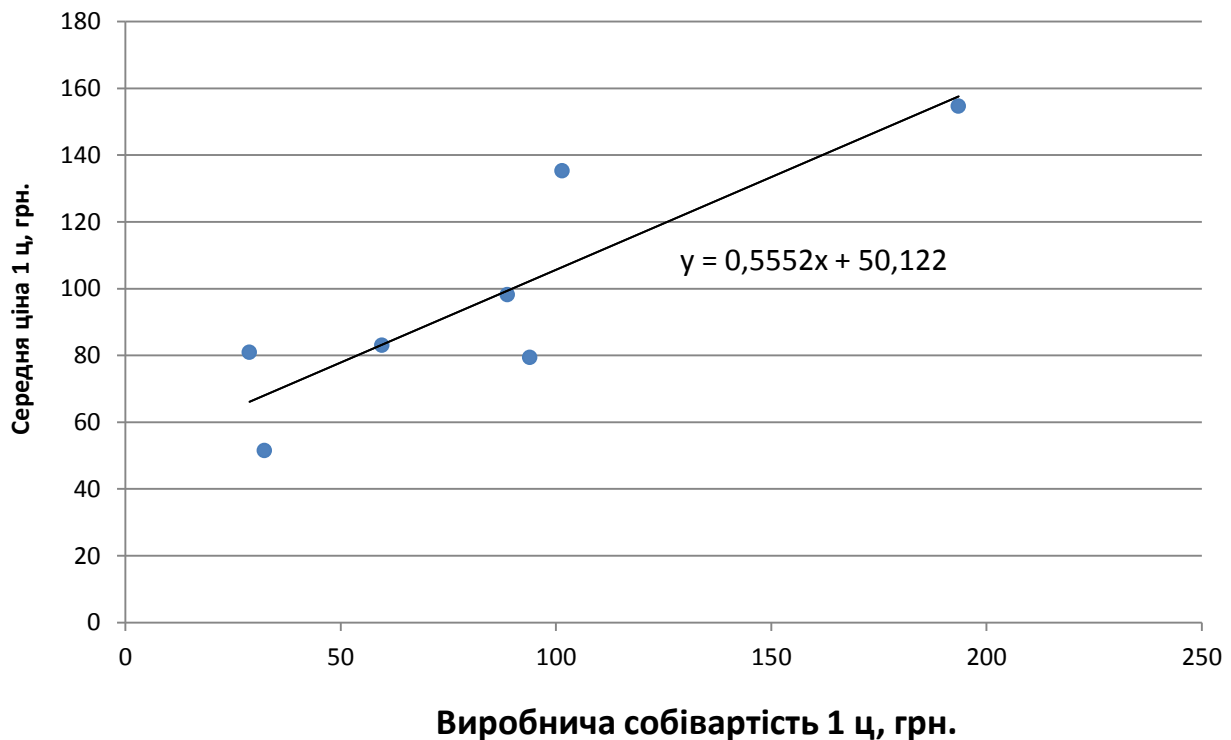


Рис. 2.20. Залежність ціни зерна від його собівартості

5. Діаграми розподілу сукупностей - графічне зображення згрупованих за певною ознакою даних на прямокутній системі координат.

Використовуються:

- а) полігон;
- б) гістограма;
- в) кумулятивні криві.

Більш детально вищеназвані діаграми розглянуті в главі 5.2 «Графічне зображення варіаційних рядів».

6. Знаки Варзара (по імені російського статистика В.Е. Варзара (1851-1940) - прямокутні фігури, що використовуються для графічного зображення трьох показників, один із яких є добутком двох інших.

а) **простий** – для зображення величини складних явищ.

Спосіб побудови. Одна з ознак явища зображується основою прямокутника (**a**), інша - його висотою (**h**), третя - розміром одержаної площі (**S=a*h**). Розташовуючи рядом кілька прямокутників, що стосуються різних показників, можна порівнювати не тільки розміри показника-добутку, але й значення показників-співмножників.

Приклад. В господарстві 300 га засіяно соняшником, а 600 – зерновими. Урожайність соняшника становить 20 ц/га, зернових – 30 ц/га.

$$\text{Валовий збір} = \text{Урожайність} * \text{Посівна площа}$$

Приймаємо:

a – урожайність, ц/га;

h – площа посіву, га;

S – валовий збір, ц.

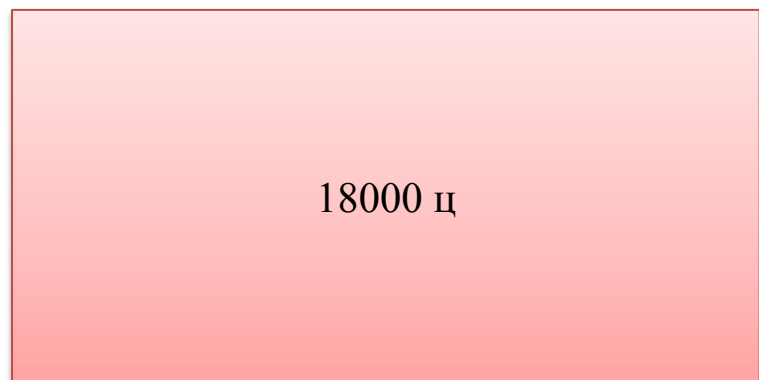
Масштаб:

1 см = 10 га

1 см = 10 ц/га



Соняшник



Зернові

Рис 2.19. Виробництво основних культур в господарстві в 2012 р.

На цьому графіку можна порівняти між собою:

- урожайність (по довжині бокової сторони);
- посівні площі (по довжині основи);
- валовий збір (по площі прямокутника).

б) **складний** – дозволяє за допомогою різної штрихування показати структуру цілого.

Приклад. Середня продуктивність праці на заводі становить 3000 грн./роб., при цьому в цеху 1 працює 30 робітників, в цеху 2 – 40 робітників, в цеху 3 – 25 робітників.

*Вартість продукції = кількість робітників * продуктивність праці*

Приймаємо:

a – кількість робітників цеху;

h – середня продуктивність праці, грн./роб.;

S - вартість продукції, грн.;

$$S = S_1 + S_2 + S_3.$$

Масштаб:

1 см = 10 роб.

1 см = 1000 грн./роб.

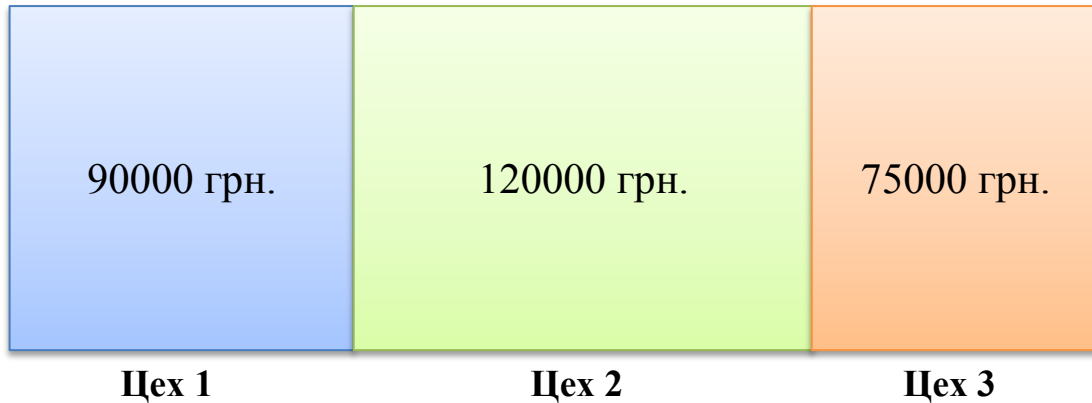


Рис. 2.20. Цехова структура вартості продукції заводу

Діаграми можуть:

- вирішувати одразу декілька завдань зображення;

- поєднувати декілька графічних образів;

- поєднувати декілька масштабних шкал та зображувати різновимірні показники.

Приклад. На рис. 2.21 одночасно вирішуються завдання порівняння рентабельності галузей рослинництва і тваринництва, а також відображується їхня зміна у часі – динаміка.

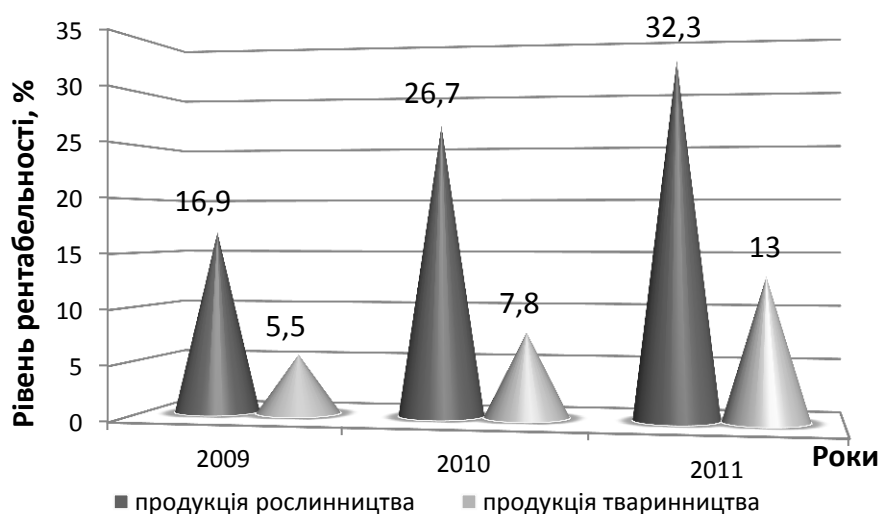


Рис. 2.21. Динаміка рентабельності сільськогосподарської продукції

Приклад. Середньомісячне споживання продуктів харчування одною особою в домогосподарствах України за I півріччя 2011 та 2012 року для наочності зобразили за допомогою різних графічних образів - стовпчиків та ліній (рис. 2.22).

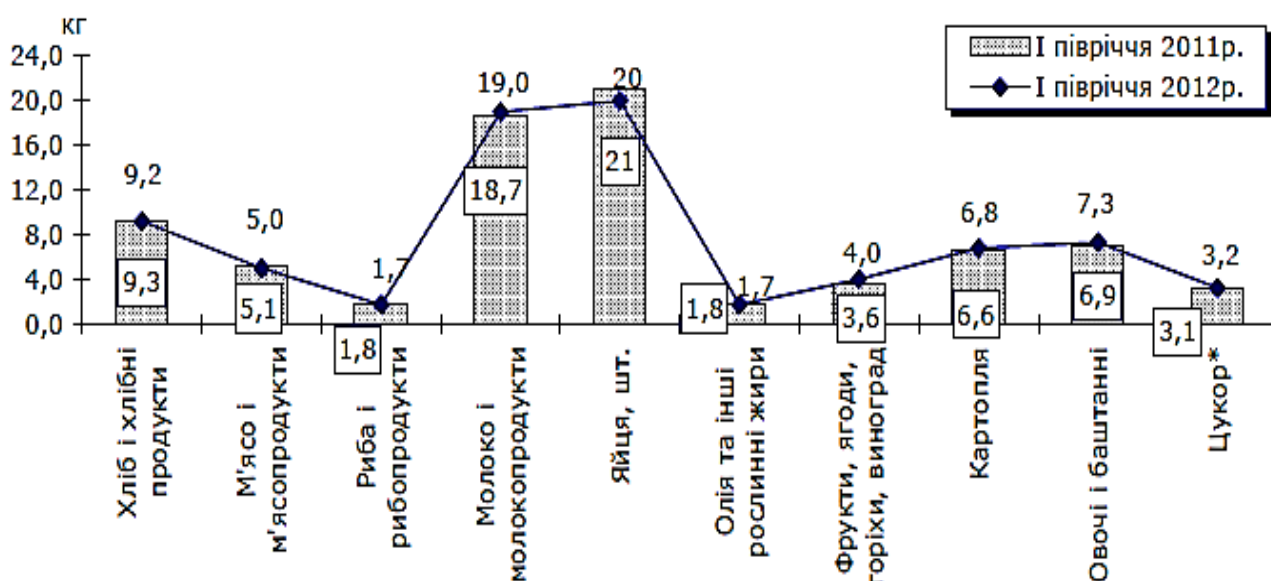


Рис. 2.22. Споживання продуктів харчування в домогосподарствах (в середньому за місяць у розрахунку на одну особу)

Приклад. При аналізі динаміки рівня бідності населення України поряд з часткою найбідніших верств населення, вираженою у відсотках, на діаграмі також зображено зміну у часі розміру прожиткового мінімуму, вираженого у гривнях, для чого було введено додаткову вісь ординат (рис. 2.23).

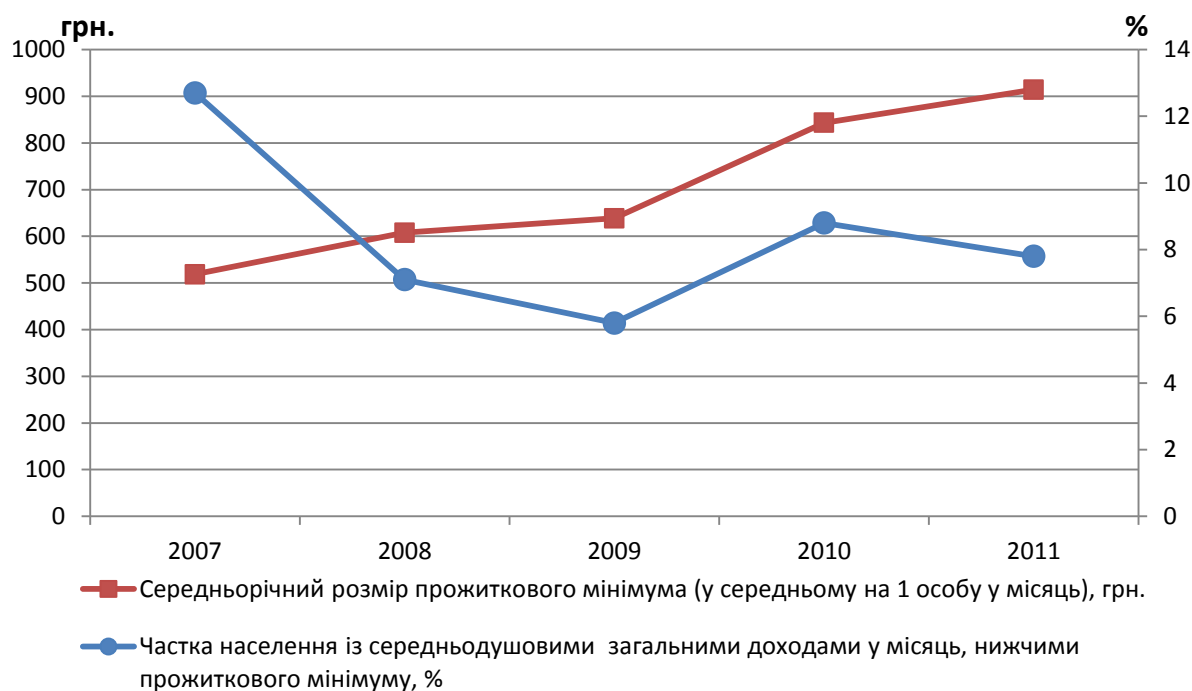


Рис. 2.23. Динаміка рівня бідності населення України

В даний час, у зв'язку з широким впровадженням в економіко-статистичну роботу та дослідження електронно-обчислювальної техніки і пакетів прикладних програм комп'ютерної графіки, відбуваються суттєві якісні зміни в застосуванні статичних діаграм.

Завдяки спеціальним графічним засобам, які мають в розпорядженні персональні комп'ютери, з'явилася можливість побудови різних типів діаграм з використанням широкого набору образотворчих засобів безпосередньо на екрані, що дозволяє оперативно змінювати при необхідності вихідні дані.

Безсумнівно, реалізація подібних перетворень розширює не тільки наочні, але і пізнавальні, аналітичні можливості діаграм, підвищуючи тим самим їхню роль і значення в обробці та аналізі статистичної інформації.

2.4 Статистичні карти

Статистична карта - вид графічного зображення статистичних даних на схематичній географічній карті, що характеризує рівень та поширення того чи іншого явища на певній території.

Розрізняють:

1) **Картограма** – за допомогою штрихування, ліній, точок, кольору різного ступеня насиченості показує порівняльну інтенсивність певного показника в межах кожної одиниці нанесеного на карту територіального поділу (щільність населення). Зокрема дає можливість одночасно дослідити події в часі та просторі.

а) **фонова** – для характеристики показника використовують штрихування та фарбування в межах територіальної одиниці (рис. 2.24);

б) **точкова** – рівень явища зображується за допомогою точок, кожна з яких означає одиницю сукупності або деяку їх кількість, щоб показати на географічній карті щільність або частоту появи певної ознаки (рис. 2.25);

в) **ізолінійна** (ізолінія від грец. *isos* – рівний, однаковий, подібний) – замкненими лініями позначаються контури приблизно однакового статистичного показника(рис. 2.26);

2) **Картодіаграма** - на фоні карти, присутні елементи діаграмних фігур. Перевага картодіаграми перед діаграмою полягає в тому, що вона не тільки дає уявлення про величину досліджуваного показника на різних територіях, але й зображує просторове розміщення досліджуваного показника (рис. 2.27).

Статистичні карти застосовують при картографуванні природних та соціально-економічних явищ, вони можуть бути використані для одержання кількісної характеристики явищ і процесів та для аналізу взаємозв'язків між ними.

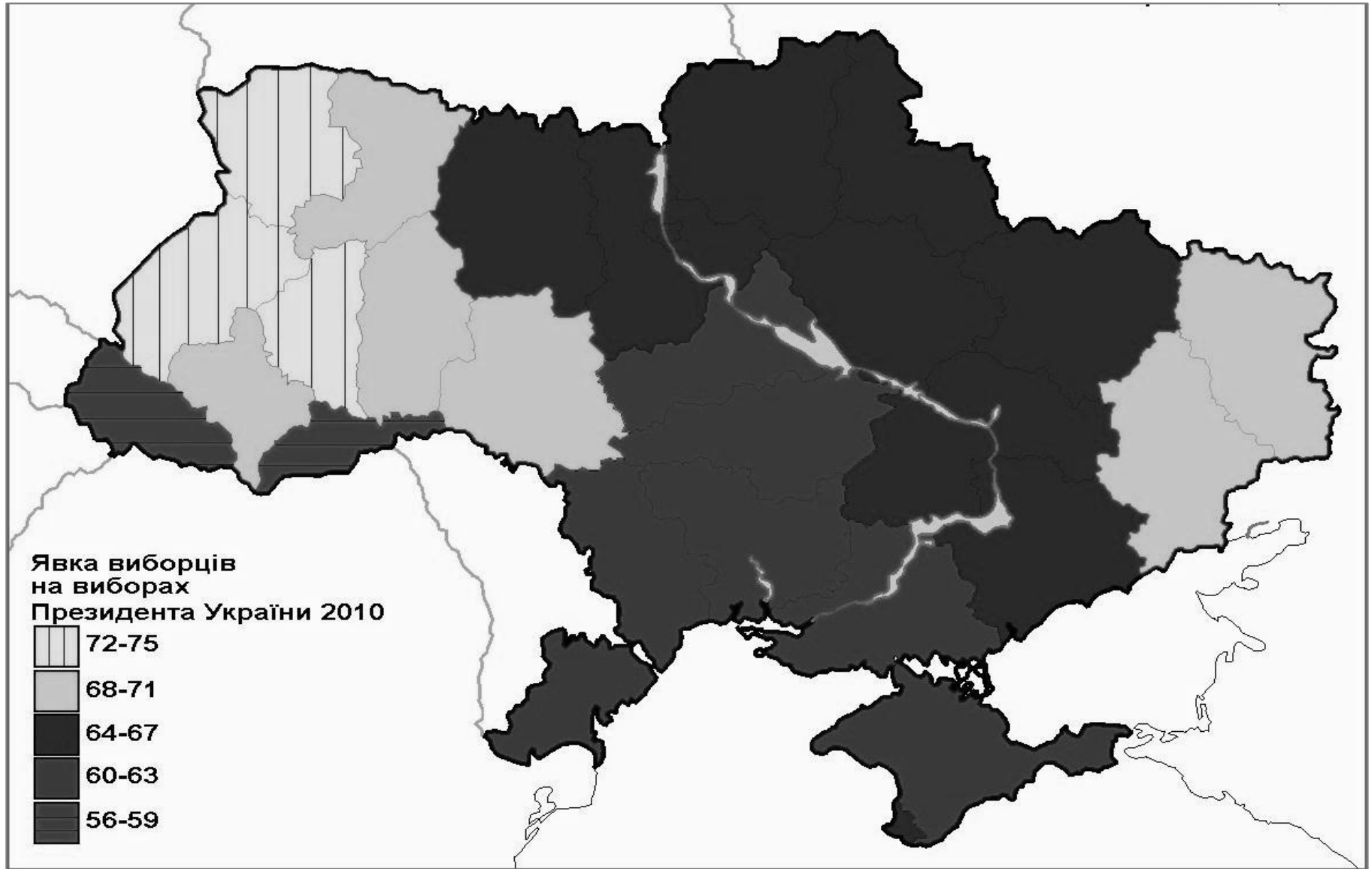


Рис. 2.24. Явка виборців на виборах Президента України 2010 року

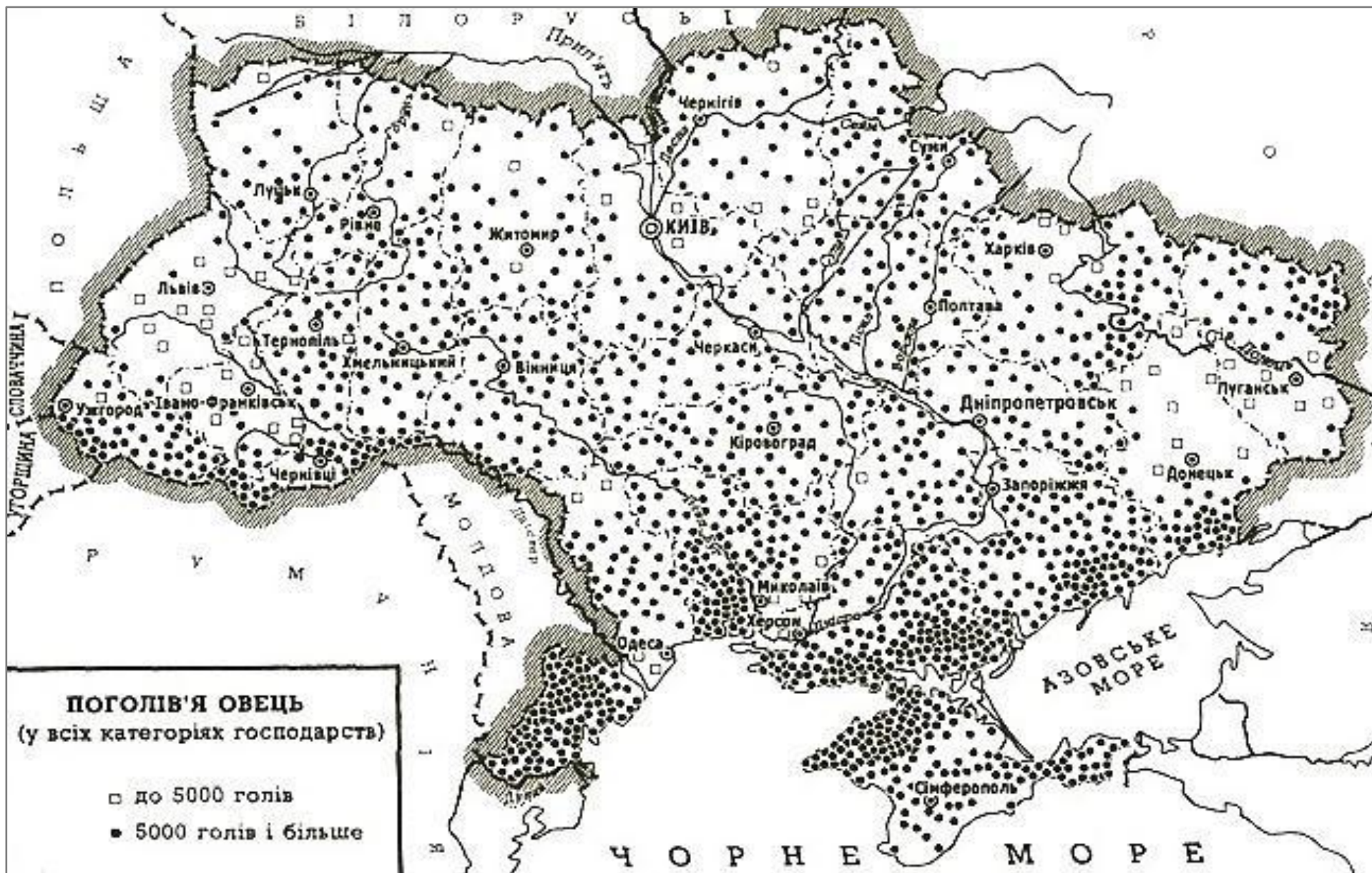


Рис. 2.25. Розподіл овець за територією України

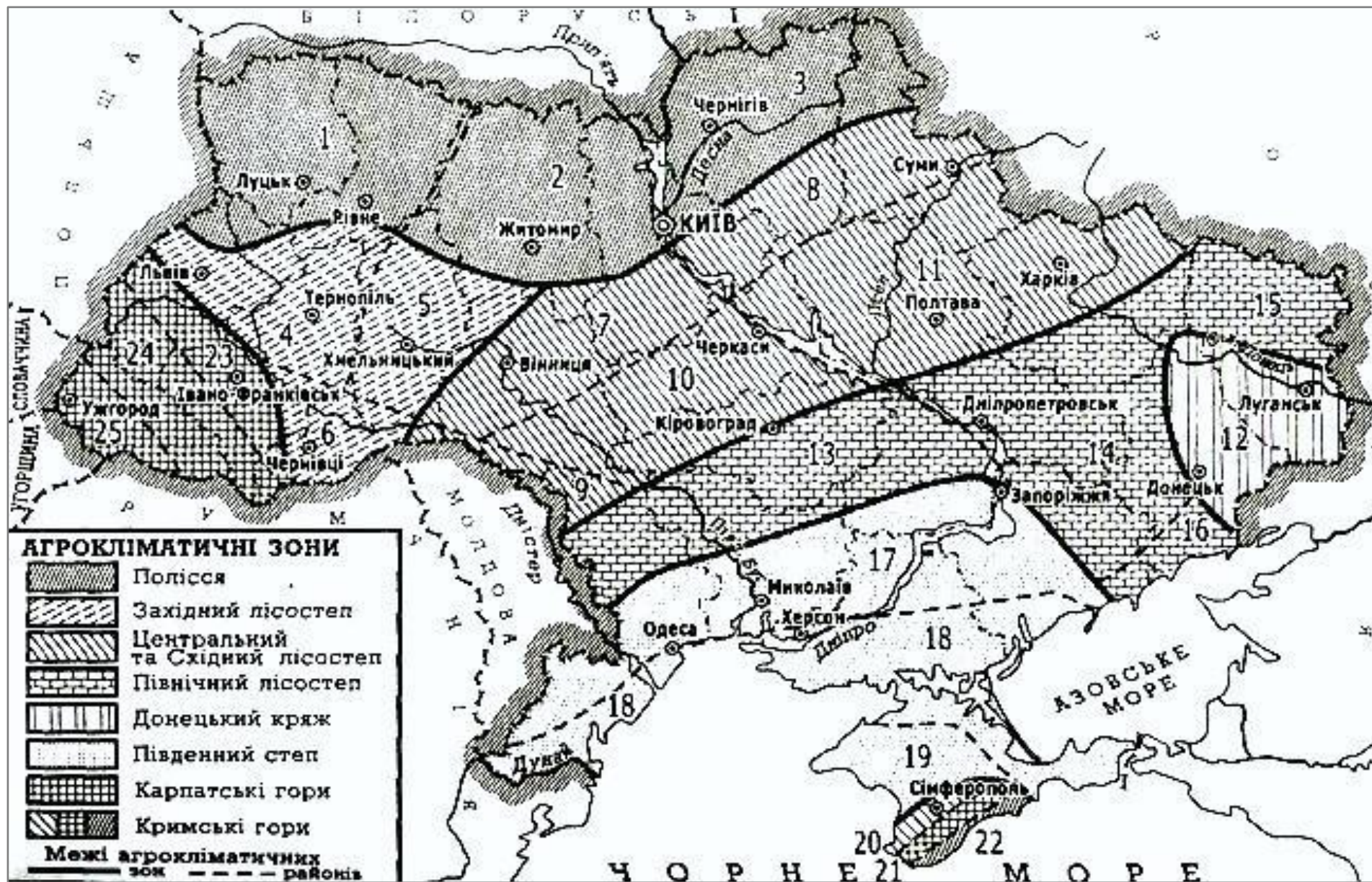


Рис. 2.26. Агрокліматичне районування України (за В.В.Поповим)

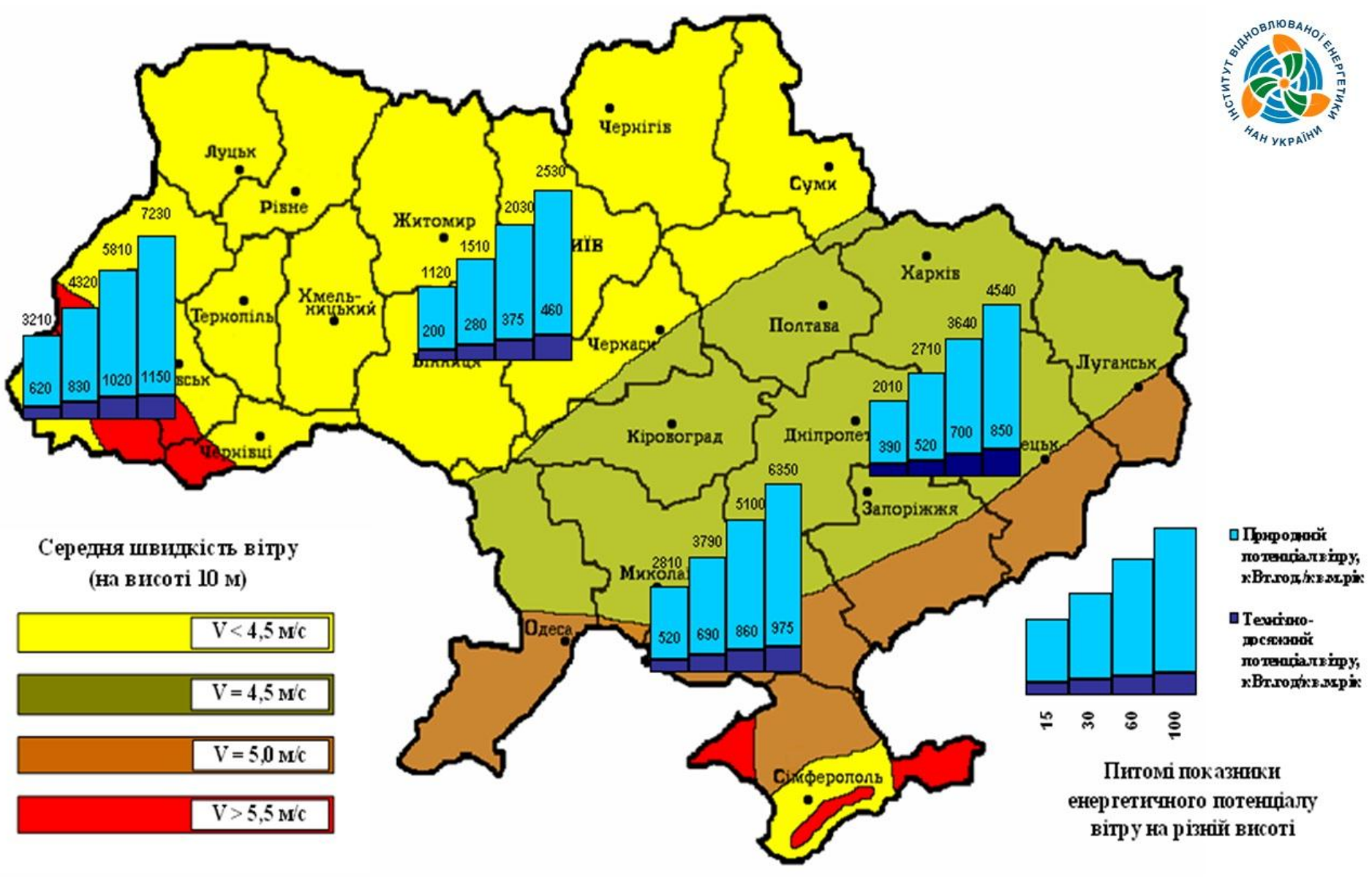


Рис. 2.28. Вітроенергетичний потенціал України

Завдання для самостійного розв'язання.

Завдання 2.1. Спроектуйте макети таблиць, які б містили наступні дані:

А) обсяги (млн. грн.) та структуру виробництва (%) товарів народного споживання (продовольчі, непродовольчі та алкогольні вироби) за регіонами протягом останнього року;

Б) обсяги (млн. дол. США) та темпи зростання (%) прямих іноземних інвестицій у галузі промисловості (паливна, енергетика, металургія, нафтопереробна, харчова, легка) за два останніх роки;

В) населення за статтю та рівнем освіти в динаміці за 2000-2012 роки;

Вкажіть підмет та присудок та визначте вид таблиці.

Завдання 2.2. За наведеними даними складіть статистичні таблиці, озаглавьте кожну з них, визначте підмет та присудок, а також вид таблиці за побудовою підмету, присудку та характером матеріалу:

А) За останній рік у регіоні N введено в дію основних засобів на суму 26,5 млн. грн., в т.ч. на об'єктах виробничого призначення — 18,2; невиробничого призначення - 8,3 млн. грн. Обсяг капітальних вкладень за той же період в цілому по регіону становив 28,6 млн. грн., з них на об'єктах виробничого призначення 20,4, невиробничого призначення - 8,2 млн. грн.;

Б) За даними перепису 2001 року в Запорізькій області знаходяться 37 міських поселень, в яких проживає 1452828 осіб, в тому числі 14 міст та 23 селища міського типу з населенням відповідно 1331965 та 120863 осіб;

В) В майбутньому році в Україні очікувана тривалість життя складає 71 рік, у тому числі для чоловіків – 66 років, для жінок – 75 років. В Росії очікувана тривалість життя складає для всього населення 69 років, у Франції – 77 років, у ФРН – 75 років, в Ірані - 55. Очікувана тривалість життя чоловіків складає у цих країнах відповідно 63.5, 73, 72 та 56 років, а жінок – 74.3, 81,79 та 55 років відповідно.

Завдання 2.3. За даними таблиць побудувати статистичні графіки:

А) Щоденне виробництво нафти у деяких державах світу, млн. барелів

Держава	США	Канада	Мексика	Венесуела	Велика Британія	Норвегія	Росія	Китай	Іран	Саудівська Аравія
Обсяг виробництва	8,1	2,3	2,9	3,2	2,9	3,3	5,8	3,2	3,7	7,5

Окремо за допомогою квадратної або кругової діаграми порівняти щоденне виробництво нафти в США, Росії та Китаї.

Б) Структура посівних площ господарства

Культури	Частка посівних площ, %		
	2010 рік	2011 рік	2012 рік
Зернові та зернобобові культури	40	30	25
Цукрові буряки	5	10	15
Соняшник	15	15	5
Картопля	10	5	20
Овочі відкритого ґрунту	10	15	25
Кормові культури	20	25	10
Разом	100	100	100

В) Сезонність реалізації молока на ринку міста у 2012 році

Місяць	1	2	3	4	5	6
Продано молока, т	35	40	43	48	55	58
Місяць	7	8	9	10	11	12
Продано молока, т	65	63	60	50	45	38

Г) Урожайність основних сільськогосподарських культур, ц/га

Роки	зернові та зернобобові культури (у вазі після доробки)	цукрові буряки (фабричні)	насіння соняшнику (у вазі після доробки)
2002	27,3	189	12,0
2003	18,2	201	11,2
2004	28,3	238	8,9
2005	26,0	248	12,8
2006	24,1	285	13,6
2007	21,8	294	12,2
2008	34,6	356	15,3
2009	29,8	315	15,2
2010	26,9	279	15,0
2011	37,0	363	18,4

Д) Залежність між урожайністю картоплі і кількістю внесених мінеральних добрив

№ спостереження	Внесено мінеральних добрив, ц діючої речовини на 1 га (x)	Урожайність, ц/га (y)
1	2,3	270
2	2,5	290
3	3	320
4	3,1	325
5	2,7	305
6	2,9	310
7	1,5	190
8	1,9	210
9	1,7	210
10	2,4	275

На графіку відобразити фактичні дані та пряму, одержану за рівнянням залежності: $y=57,5+88,8x$.

Завдання 2.4. Побудувати знаки Варзара, які б характеризували:

А) річну виручку від реалізації продукції на одному з ринків міста, якщо відома кількість продавців – 450 та середня виручка одного продавця в рік – 25 тис. грн., а також виручку на сусіднім ринку, де кількість продавців становить 500 осіб та середня виручка одного продавця в рік – 20 тис. грн.

Б) загальну річну виручку від реалізації продукції на ринках міста, якщо відома середня виручка одного продавця в рік – 25 тис. грн., а кількість продавців на ринку А складає 450 осіб, на ринку Б – 600, на ринку В – 300.

Завдання 2.5. Надати зазначену інформацію у вигляді статистичної карти.

А) В Україні можна виділити 8 великих економічних районів:

- 1) Донецький (Донецька і Луганська області);
- 2) Придніпровський (Дніпропетровська, Запорізька і Кіровоградська області);
- 3) Східний (Харківська, Полтавська і Сумська області);
- 4) Центральний (Київська і Черкаська обл., м. Київ);
- 5) Поліський (Волинська, Житомирська, Рівненська, Чернігівська області);
- 6) Подільський (Вінницька, Тернопільська, Хмельницька області);
- 7) Причорноморський (АР Крим, Миколаївська, Одеська, Херсонська області, м. Севастополь);
- 8) Карпатський (Закарпатська, Львівська, Івано-Франківська, Чернівецька області).

Б) Маємо наступні дані про щільність населення по регіонах України:

Області	Щільність населення, осіб/кв. км	Області	Щільність населення, осіб/кв. км
АР Крим	78	Миколаївська	51
Вінницька	67	Одеська	74
Волинська	53	Полтавська	57
Дніпропетровська	112	Рівненська	58
Донецька	183	Сумська	55
Житомирська	46	Тернопільська	83
Закарпатська	98	Харківська	93
Запорізька	71	Херсонська	41
Івано-Франківська	101	Хмельницька	69
Київська	65	Черкаська	67
Кіровоградська	46	Чернівецька	114
Луганська	95	Чернігівська	39
Львівська	120	Україна	80

Тести для самоконтролю.

1. Частина таблиці, що містить статистичну сукупність, ті об'єкти або їх частини, які характеризують статистичну сукупність, називається:

- А) заголовок
- Б) макет
- В) підмет
- Г) присудок

2. За характером матеріалу прості таблиці, що характеризують не всю сукупність одиниць досліджуваного об'єкту, а тільки одну групу з нього, яка виділяється за певною ознакою, називають:

- А) монографічні
- Б) групові
- В) перелікові
- Г) комбінаційні

3. Знайдіть зайвий елемент серед масштабних та просторових орієнтирів:

- А) система координат
- Б) поле графіку
- В) масштабна шкала
- Г) масштаб

4. За характером матеріалу прості таблиці, підмет яких містить показники по районах, областях, країнах, тощо, а присудок - періоди або моменти часу:

- А) переліково-хронологічні
- Б) територіально-хронологічні
- В) територіальні
- Г) хронологічні

5. Статистична карта, на якій для характеристики показника використовують штрихування та фарбування в межах територіальної одиниці, називається:

- А) картодіаграма
- Б) фонові картограма
- В) точкова картограма
- Г) ізолінійна картограма

6. Для зображення сезонності явища чи процесу використовується:

- А) стрічкова діаграма
- Б) секторна діаграма
- В) радіальна діаграма
- Г) знак Варзара

7. Прямокутна фігура, що використовується для графічного зображення трьох показників, один із яких є добутком двох інших, називається:

- А) знак Варзара
- Б) секторна діаграма
- В) радіальна діаграма
- Г) квадратна діаграма

8. Знайдіть зайвий елемент серед складових експлікації графіка:

- А) пояснювальні написи окремих елементів графічного образу
- Б) написи вдовж масштабних шкал
- В) заголовок графіка
- Г) графічний образ

9. Таблиці, підмет яких розподілено на групи за двома і більше ознаками, і кожна з груп, побудованих за однією ознакою розбивається у свою чергу на підгрупи по якому-небудь іншому признаку називають:

- А) прості
- Б) комбінаційні
- В) групові
- Г) складні

10. Таблиця, що складається з рядків і граф, які ще не заповнені цифрами, називається

- А) підмет
- Б) макет таблиці

- В) присудок
- Г) статистична таблиця

11. За характером матеріалу прості таблиці, у підметі яких подається перелік показників або одиниць досліджуваної сукупності, називають:

- А) монографічні
- Б) хронологічні
- В) перелікові
- Г) впорядковані

12. Статистична карта, на якій на фоні карти, присутні елементи діаграмних фігур, називається:

- А) піктограма
- Б) ізолінійна картограма
- В) фонові картограма
- Г) картодіаграма

13. За характером матеріалу прості таблиці, у підметі яких наводиться ряд показників, а присудок – періоди або моменти часу

- А) переліково-хронологічні
- Б) хронологічні
- В) комбінаційні
- Г) перелікові

14. Частина таблиці, що містить систему показників, якими характеризується статистична сукупність:

- А) підмет
- Б) присудок
- В) макет
- Г) заголовок

15. За характером матеріалу прості таблиці, у підметі яких наводиться перелік районів, областей и т.п., називають:

- А) монографічні
- Б) територіальні
- В) географічні
- Г) перелікові

16. Статистична карта, на якій замкненими лініями позначаються контури приблизно однакового статистичного показника, називається:

- А) контурна картограма
- Б) картодіаграма
- В) фонові картограма
- Г) ізолінійна картограма

17. Статистична карта, на якій рівень явища зображується за допомогою точок, кожна з яких означає одиницю сукупності або деяку їх кількість, щоб показати на географічній карті щільність або частоту появи певної ознаки, називається:

- А) картодіаграма
- Б) точкова картограма
- В) фоновна картограма
- Г) піктограма

18. Таблиці, в яких підмет розподілено на групи за однією кількісною або якісною ознакою, називають:

- А) складні
- Б) прості
- В) комбінаційні
- Г) групові

19. Діаграма, що зображується на полярній системі координат:

- А) секторна
- Б) знак Варзара
- В) точкова
- Г) радіальна

20. Залежно від побудови підмета, таблиці у підметі яких немає групувань, а дається об'єкт чи перелік будь-яких об'єктів, територіальних одиниць, періодів часу, називають:

- А) прості
- Б) групові
- В) складні
- Г) комбінаційні

21. За характером матеріалу прості таблиці, у підметі яких наводяться періоди часу або дати, а в присудку - ряд показників, називають:

- А) періодичні
- Б) хронологічні
- В) перелікові
- Г) монографічні

3. ЗВЕДЕННЯ ТА ГРУПУВАННЯ СТАТИСТИЧНИХ ДАНИХ

3.1 Статистичне зведення та групування. Види групувань

У результаті першої стадії статистичного дослідження - статистичного спостереження - одержують відомості про кожну одиницю сукупності. Завдання другої стадії статистичного дослідження полягає в тому, щоб упорядкувати й узагальнити первинний матеріал, звести його в групи й на цій основі дати узагальнюючу характеристику сукупності. Цей етап у статистиці називається зведенням.

Статистичне зведення – це комплекс послідовних операцій по узагальненню конкретних одиничних фактів, що утворюють сукупність, для виявлення типових рис і закономірностей, властивих досліджуваному явищу в цілому, тобто систематизація, обробка і підрахунок групових і загальних підсумків даних статистичного спостереження.

Розрізняють *зведення*:

- **просте** – виконують без розподілу одиниць сукупності на групи, підрахунок тільки загальних підсумків;
- **статистичне групування** (групове або складне зведення) - розподіл сукупності на групи за істотною для одиниць сукупності ознакою.

За організацією робіт розрізняють зведення:

- **централізоване** – матеріали статистичного спостереження надсилають у центральні органи державної статистики, де їх обробляють;
- **децентралізоване** – дані спостереження підраховують на місцях, а в центральний орган надсилають уже зведені підсумки.

За способом виконання бувають зведення:

- **ручне** – всі операції виконують ручним способом за допомогою карток або списків;
- **машинне (комп'ютерне)** – застосовують для обробки масового матеріалу і здійснюють за допомогою електронно-обчислювальних машин.

Щоб вирішити ряд конкретних завдань, виявити особливості в розвитку явищ, виявити тенденції, установити залежності, необхідно зробити групування статистичних даних. З цією метою обирається групувальна ознака й розробляється система показників, якими будуть характеризуватися виділені групи. Визначення й обґрунтування показників цілком залежать від мети дослідження й поставленого завдання.

Ознаки в статистиці – це властивості, характерні риси або особливості явищ, які можна виразити рядом статистичних величин. Ознаки, покладені в основу групування, називаються **групувальними**.

Ознаки бувають *первинні й вторинні*. Первинні ознаки характеризують абсолютні розміри досліджуваних явищ (наприклад, чисельність робітників), вторинні є похідними від первинних і показують структуру явищ, що групують (фондоозброєність, собівартість і т.д.).

Види групувань.

Залежно від мети й завдань дослідження:

- **типологічне** - групування, що характеризує якісні особливості й розходження між типами явищ. Воно передбачає поділ якісно неоднорідної сукупності на окремі якісно однорідні групи і виявлення на цій основі економічних типів явищ. Таким чином, основне завдання такого групування - ідентифікація типів соціально-економічних явищ, тому важливе значення при її побудові повинно надаватися вибору групувальної ознаки (групування населення за класами, сферою діяльності; підприємств за формою власності або господарювання, за виробничим напрямом і т.п.);

- **структурне** - групування, що передбачає виявлення закономірностей розподілу одиниць однорідної сукупності за варіюючим значенням досліджуваної ознаки. Воно дозволяє вивчити структуру сукупності і зрушень, що в ній відбуваються. Структурні групування відрізняються від типологічних не стільки за зовнішнім виглядом, скільки за цілями, тобто відрізняються за рівнем якісних відмінностей між групами (групування підприємств за потужністю, по числу робітників і т.д.; склад населення може бути згрупований за статтю, за віком, за рівнем освіти, за родом занять, доходом і т.д.);

- **аналітичне** - групування, що застосовується для дослідження взаємозв'язку між окремими ознаками одиниць статистичної сукупності (між розміром торгівельної площі та товарооборотом, фондоозброєністю і продуктивністю праці, тощо).

Використовуючи аналітичні групування, визначають факторні й результативні ознаки досліджуваних явищ. **Факторні** - це ознаки, що впливають на інші, пов'язані з ними ознаки. **Результативні** - це ознаки, які змінюються під впливом факторних. Відповідно: якщо групувальною є факторна ознака - групування називають **факторним**, якщо групувальною є результативна ознака - **результативним**.

За кількістю ознак, покладених в основу:

- **просте** - групування за однією ознакою.

- **комбінаційне** - одночасне групування за двома і більше ознаками.

Тобто групи, утворені за однією ознакою, підрозділяються на підгрупи за іншою, а отримані в результаті цього підгрупи підрозділяються (кожна окремо) ще на підгрупи й т.д. При цьому групувальні ознаки прийнято розташовувати, починаючи з атрибутивної, у певній послідовності, виходячи з

логіки взаємозв'язку показників. Застосування комбінованих групувань зумовлено різноманіттям економічних явищ, а також необхідністю їхнього всебічного вивчення. Але збільшення числа групувальних ознак обмежується зменшенням наочності, що знижує ефективність використання статистичної інформації.

Прикладом комбінованого групування може служити поділ груп підприємств, утворених за вартістю основних засобів, на підгрупи за чисельністю працівників або по інших ознаках (за площею, обсягом виробленої продукції і т.д.).

3.2 Попередні відомості про ряди розподілу

Результати зведення можуть бути представлені у вигляді статистичних рядів розподілу.

Статистичний ряд розподілу являє собою найпростіше групування, у якому кожна виділена група характеризується одним показником – чисельністю одиниць спостереження, що потрапили в кожену групу.

Залежно від ознаки ряди розподілу поділяють на:

- **варіаційні** – ряди, утворені за кількісними ознаками;
- **атрибутивні** - ряди, утворені за якісними ознаками.

Розподіл елементів сукупності за двома і більше ознаками називається **комбінаційним**.

Числові значення кількісної ознаки у варіаційному ряді називаються **варіантами** (x) і розташовуються в певному порядку; чисельність одиниць у кожній групі називається **частотою** (f). Частота, виражена у відсотках до підсумку, називається **часткою** (d). Сума всіх частот називається **обсягом** (чисельністю) ряду розподілу.

Варіаційні ряди поділяють на:

- **дискретні (перервні)** – ряди, у яких числовий розподіл ознаки виражений певними, зазвичай цілими, числами (розподіл робочих за розрядами, родин за кількістю дітей);
- **інтервальні (неперервні)** - ряди, у яких числові значення ознаки можуть набувати будь-яких значень в певних межах, тобто в певному інтервалі (зріст людини, урожайність культур, заробітна плата й т.д.)

Інтервал - це різниця між максимальним та мінімальним значенням ознаки.

За величиною інтервали групування можуть бути:

- **рівні** - різниця між верхньою і нижньою межею однакова. Використовуються, коли зміна ознаки усередині сукупності відбувається рівномірно, або якщо далі планується наступна математична обробка згрупованих даних;

- **нерівні (прогресивні)** – різниця між верхньою і нижньою межею неоднакова. Використовується, якщо для більших значень ознаки зміна його величини менш суттєва, чим для малих. Наприклад, підприємства з чисельністю робітників 100 і 300 значно більше відрізняються між собою, ніж підприємства з 1100 та 1300 робітників.

За визначеністю меж бувають інтервали:

- **закриті** – у яких відомі нижня та верхня межа інтервалу;
- **відкриті** – перший та останній інтервал; для першого вказується тільки верхня межа, а для останнього – тільки нижня.

Число груп в групуванні вибирається за наступних передумов: мінливість ознаки, число спостережень, однорідність груп.

Якщо групувальна ознака має плавний характер і застосовуються рівні інтервали, кількість груп орієнтовно можна визначити за формулою *Стерджеса*:

$$n = 1 + 3,322 \lg N,$$

де n – кількість груп;

N - чисельність сукупності.

Користуючись цією формулою, можна скласти номограму, яка показує залежність кількості груп від чисельності ряду розподілу (табл. 3.1).

Таблиця 3.1

Номограма Стерджеса

N	15-24	25-44	45-89	90-179	180-359	360-719	720-1449
n	5	6	7	8	9	10	11

Однак ця формула здобута емпіричним шляхом і не враховує ні складу, ні характеру коливань досліджуваної сукупності. Тому механічне її застосування може призвести до незадовільних наслідків.

При групуванні із застосуванням *рівних закритих інтервалів* величина інтервалу визначається за формулою:

$$h = \frac{x_{\max} - x_{\min}}{n},$$

де h – величина інтервалу;

x_{\max} - максимальне значення ознаки;

x_{\min} - мінімальне значення ознаки;

n - кількість груп.

Якщо максимальне і мінімальне значення сильно відрізняються від суміжних з ними значень варіантів в упорядкованому ряду значень групувальної ознаки, то для визначення величини інтервалу слід використовувати не максимальне і мінімальне значення, а значення, дещо вищі за мінімум і дещо менші за максимум.

Задача 3.1. На основі зведених даних про 21 сільськогосподарського підприємства району (табл. 3.2.) побудувати варіаційний ряд їх розподілу за розміром.

Розв'язання.

Розмір сільськогосподарських підприємств визначається багатьма показниками, зокрема: обсягом виробництва валової продукції, земельною площею, основними засобами, кількістю середньорічних працівників та ін. (табл. 3.2)

Таблиця 3.2

Розміри сільськогосподарських підприємств району

№	Назва підприємства	Площа с.-г. угідь, га	Середньорічна кількість працівників, осіб	Вартість основних засобів, тис. грн.	Валова продукція (в постійних цінах 2010 року), тис. грн.
1	ПСП «Дружба»	2788	56	3000	6855
2	ДПДГ «Сади»	1319	279	24000	7500
3	ПСП «Україна»	4286	81	6640	9954
4	ТОВ «Агромир»	2871	91	4300	6397
5	ПРАТ «Зоря»	6248	106	10900	10490
6	ПСП «Лідер»	3560	193	5500	11742
7	ТОВ «Нива»	3475	43	2100	2749
8	ТОВ «Лан»	8170	64	3600	8960
9	ДПДГ «Фруктове»	1645	337	25500	12720
10	ПСП «Мир»	2790	52	2850	6139
11	ПСП «Колос»	4980	130	6653	9882
12	ТОВ «Сонячне»	3050	115	3851	6643
13	ПРАТ «Агросвіт»	8045	111	13800	12282
14	ПСП «Вікторія»	2966	192	5200	7773
15	ТОВ «Перемога»	3475	51	2300	3552
16	ФГ «Ясне»	2956	59	2105	9279
17	ФГ «Господар»	2231	30	1357	4287
18	ФГ «Світоч»	2195	28	1241	5009
19	ПАТ «Еліта»	1130	42	1053	1251
20	ТОВ «Південне»	2745	57	3474	8955
21	ФГ «Урожай»	2810	37	887	2843

Основним показником розміру сільськогосподарських підприємств є **вартість валової продукції**, оскільки вона визначає кінцевий результат виробництва. При порівнянні розмірів різних господарств за ряд років для оцінки валової продукції використовують порівнянні ціни на продукцію.

Однак, хоча показник обсягу валової продукції є основним, однак при визначенні розміру сільськогосподарських підприємств більш зручно користуватися показником **розміру земельної площі**, зокрема сільськогосподарських угідь. Це пояснюється тим, що земельна площа є найбільш стабільним по роках показником, тоді як показник валової продукції значно коливається, оскільки залежить від багатьох факторів (погодні умови, різниця цін на продукцію, зміна структури виробництва тощо).

Тому ми проведемо групування підприємств саме за площею сільськогосподарських угідь.

Згідно номограми Стерджеса (табл. 3.1) для обсягу сукупності $N=21$ кількість груп $n=5$. Отже, утворимо 5 груп з рівними інтервалами.

Величина інтервалу визначається по формулі

$$h = \frac{x_{\max} - x_{\min}}{5},$$

де x_{\max} й x_{\min} - максимальна й мінімальна площа сільськогосподарських угідь, а саме 8170 га (ТОВ «Лан») та 1130 га (ПАТ «Еліта») відповідно.

Тоді величина інтервалу дорівнює:

$$h = \frac{8170 - 1130}{5} = 1408$$

Тепер утворимо групи підприємств, які відрізняються друг від друга по площі сільськогосподарських угідь на цю величину:

(1130+1408 = 2538)	1130 - 2538
(2538+1408 = 3946)	2538 - 3946
(3946+1408 = 5354)	3946 - 5354
(5354+1408 = 6762)	5354 - 6762
(6762+1408 = 8170)	6762 - 8170

Техніка підрахунку 1. Можна вручну зробити вибірку потрібних значень і занести їх попередньо в робочу таблицю (табл. 3.3), де кожна риса відповідає одиниці сукупності, тобто одному заводу, а рахунок ведеться п'ятірками - кожні чотири риски перекреслюються п'ятою.

Таблиця 3.3

Робоча таблиця для побудови варіаційного ряду розподілу

Групи підприємств за площею сільськогосподарських угідь, га	Кількість підприємств
1130 - 2538	###
2538 - 3946	### ### /
3946 - 5354	//
5354 - 6762	/
6762 – 8170	//
Разом	21

На підставі робочої таблиці складемо відповідний варіаційний ряд розподілу (табл. 3.4).

Таблиця 3.4

Розподіл підприємств за площею сільськогосподарських угідь

Групи підприємств за площею сільськогосподарських угідь, га	Кількість підприємств	Питома вага підприємств групи, в % до підсумку
1130 - 2538	5	23,8
2538 - 3946	11	52,4
3946 - 5354	2	9,5
5354 - 6762	1	4,8
6762 – 8170	2	9,5
Разом	21	100

Таким чином, ряд розподілу підприємств за площею сільськогосподарських угідь показує, що найбільш характерними для сільського господарства району є підприємства з площею сільськогосподарських угідь від 2538 до 3946 га, які становлять 52,4 %, тобто більше половини, всіх господарств, і що в 76,2% підприємств площа сільськогосподарських угідь не перевищує 3946 га.

Техніка підрахунку 2. Складається **ранжований ряд** – ряд, у якому значення ознаки розташовується в зростаючому або убутному порядку (табл. 3.5). В даному випадку доцільно використовувати комп'ютер. За даними ранжованого ряду добре видно зміну площі сільськогосподарських угідь і легко встановити межі груп.

Таблиця 3.5

Ранжований ряд підприємств за площею сільськогосподарських угідь

№ групи	№ підприємства	Площа сільськогосподарських угідь, га	№ групи	№ підприємства	Площа сільськогосподарських угідь, га
I	19	1130	II	14	2966
	2	1319		12	3050
	9	1645		7	3475
	18	2195		15	3475
	17	2231		6	3560
II	20	2745	III	3	4286
	1	2788		11	4980
	10	2790	IV	5	6248
	21	2810	V	13	8045
	4	2871		8	8170
	16	2956		Разом	21

Задача 3.2. за даними табл. 3.2 побудувати атрибутивний ряд розподілу сільськогосподарських підприємств району за організаційно-правовими формами.

Розв'язання.

Можна скористатися *Технікою підрахунку 1* (табл. 3.6).

Таблиця 3.6

Робоча таблиця для побудови атрибутивного ряду розподілу

Групи підприємств за площею сільськогосподарських угідь, га	Кількість підприємств
Державні підприємства дослідні господарства	//
Публічні акціонерні товариства	/
Приватні акціонерні товариства	//
Приватне сільськогосподарські товариства	### /
Товариства з обмеженою відповідальністю	### /
Фермерські господарства	////

На підставі робочої таблиці складемо відповідний атрибутивний ряд розподілу (табл. 3.7).

Таблиця 3.7

Розподіл підприємств за організаційно-правовими формами

Організаційно-правові форми	Кількість підприємств	Питома вага підприємств, %
Державні підприємства дослідні господарства	2	9,5
Публічні акціонерні товариства	1	4,8
Приватні акціонерні товариства	2	9,5
Приватні сільськогосподарські товариства	6	28,6
Товариства з обмеженою відповідальністю	6	28,6
Фермерські господарства	4	19,0
Разом	21	100

Таким чином, найбільш розповсюдженими організаційно-правовими формами серед сільськогосподарських підприємств району є приватні сільськогосподарські товариства та товариства з обмеженою відповідальністю, які складають 57,2% (по 28,6% кожний) у загальній кількості господарств.

Задача 3.3. Побудувати комбінаційний розподіл підприємств району за площею сільськогосподарських угідь та організаційно-правовими формами.

Розв'язання.

Можна знову скористатися *Технікою підрахунку 1* (табл. 3.8).

На підставі робочої таблиці складемо відповідний комбінаційний ряд розподілу (табл. 3.9).

Таблиця 3.8

Робоча таблиця для побудови комбінаційного ряду розподілу

Організаційно-правові форми	Групи підприємств за площею сільськогосподарських угідь, га				
	1130-2538	2538-3946	3946-5354	5354-6762	6762-8170
Державні підприємства дослідні господарства	//				
Публічні акціонерні товариства	/				
Приватні акціонерні товариства				/	/
Приватні сільськогосподарські товариства		////	//		
Товариства з обмеженою відповідальністю		###			/
Фермерські господарства	//	//			

Таблиця 3.9

Розподіл підприємств за площею сільськогосподарських угідь
та організаційно-правовими формами

Організаційно-правові форми	Групи підприємств за площею сільськогосподарських угідь, га					Разом
	1130-2538	2538-3946	3946-5354	5354-6762	6762-8170	
Державні підприємства дослідні господарства	2	-	-	-	-	2
Публічні акціонерні товариства	1	-	-	-	-	1
Приватні акціонерні товариства	-	-	-	1	1	2
Приватні сільськогосподарські товариства	-	4	2	-	-	6
Товариства з обмеженою відповідальністю	-	5	-	-	1	6
Фермерські господарства	2	2	-	-	-	4
Разом	5	11	2	1	2	21

3.3 Техніка проведення групування

Метод групування - головний метод статистичного вивчення суспільних явищ - слугує передумовою для застосування різних статистичних прийомів і методів аналізу, в тому числі для використання узагальнюючих показників.

Розглянуті раніше статистичні ряди розподілу дозволяють систематизувати та узагальнювати статистичний матеріал. Однак вони не дають всебічної характеристики одержаних груп. Щоб вирішити ряд конкретних завдань, виявити особливості та тенденції в розвитку явищ, встановити залежності, необхідно провести групування статистичних даних. На відміну від ряду розподілу, де є лише чисельний розподіл значень ознаки, групування дозволяє зробити більш конкретні і змістовні висновки.

З цією метою обирається групувальна ознака та розробляється система показників зведення, якими будуть користуватися виділені групи. Визначення та обґрунтування показників цілком залежить від мети дослідження та поставленої задачі.

Етапи проведення групування:

1. Виявляються групувальні ознаки.
2. Установити по кожній з них інтервали групування.
3. Установити для кожного групування показники, якими повинні характеризуватися групи.
4. Визначити, які з ознак повинні комбінуватися одна з одною.

Задача 3.4. За даними табл. 3.2 згрупувати підприємства за розміром та охарактеризувати виділені групи за іншими показниками зведення.

Розв'язання.

1. Перш за всі необхідно обрати групувальну ознаку. Ми вже з'ясували, що при визначенні розміру сільськогосподарських підприємств доцільно користуватися показником площі сільськогосподарських угідь.

Тобто, щоб зробити розподіл підприємств за розміром, треба розбити сукупність за площею сільськогосподарських угідь. Це зроблено в **задачі 3.1.**

2. Також у **задачі 3.1.** виділені 5 груп підприємств та визначена їх кількість у цих групах (частота).

Щоб зрозуміти різницю між рядами розподілу та групуванням, візьмемо в основу групування ті ж самі групи підприємств.

3. Завдання групування зводиться до визначення показників для характеристики виділених груп. Показниками, що характеризують групи будуть: середньорічна кількість працівників, вартість основних засобів та валова продукція підприємства.

Зроблене нами групування буде *структурним простим*. -

Складемо макет таблиці із системою показників, куди занесемо результати групування підприємств за площею сільськогосподарських угідь.

Таблиця 3.11 (макет)

Групування підприємств за площею сільськогосподарських угідь

Групи підприємств в за площею с.-г. угідь, га	Підприємства		Площа с.-г. угідь		Середньорічна кількість працівників		Вартість основних засобів		Валова продукція (в постійних цінах 2010 року)	
	кількість	%	га	%	осіб	%	тис. грн.	%	тис. грн.	%
1130 - 2538										
2538 - 3946										
3946 - 5354										
5354 - 6762										
6762 – 8170										
Всього										

Для заповнення макету попередньо складемо робочу таблицю 3.10.

Групові показники робочої таблиці занесемо до відповідних рядків і граф макета таблиці та отримаємо кінцеву зведену групову таблицю з результатами групування (табл. 3.11).

Таблиця 3.10

Робоча таблиця для побудови структурного групування

Групи підприємств за площею с.-г. угідь	Підприємства	Площа с.-г. угідь, га	Середньорічна кількість працівників, осіб	Вартість основних засобів, тис. грн.	Валова продукція тис. грн.
1130 - 2538	19	1130	42	1053	1251
	2	1319	279	24000	7500
	9	1645	337	25500	12720
	18	2195	28	1241	5009
	17	2231	30	1357	4287
Разом	5	8520	716	53151	30767
2538 - 3946	20	2745	57	3474	8955
	1	2788	56	3000	6855
	10	2790	52	2850	6139
	21	2810	37	887	2843
	4	2871	91	4300	6397
	16	2956	59	2105	9279
	14	2966	192	5200	7773
	12	3050	115	3851	6643
	7	3475	43	2100	2749
	15	3475	51	2300	3552
	6	3560	193	5500	11742
Разом	11	33486	946	35568	72925
3946 - 5354	3	9266	211	13293	19836
	11	9266	211	13293	19836
Разом	2	2810	37	887	2843
5354-6762	5	6248	106	10900	10490
Разом	1	6248	106	10900	10490
6762 – 8170	13	8045	111	13800	12282
	8	8170	64	3600	8960
Разом	2	16215	175	17400	21242
Всього	21	73734	2154	130312	155259

Таблиця 3.11

Групування підприємств за площею сільськогосподарських угідь

Групи підприємств за площею сільськогоспо- дарських угідь, га	Підприємства		Площа с.-г. угідь		Середньорічна кількість працівників		Вартість основних засобів		Валова продукція (в постійних цінах 2010 року)	
	кількість	%	га	%	осіб	%	тис. грн.	%	тис. грн.	%
1130 - 2538	5	23,8	8520	11,6	716	33,2	53151	40,8	30767	19,8
2538 - 3946	11	52,4	33486	45,4	946	43,9	35568	27,3	72925	47,0
3946 - 5354	2	9,5	9266	12,6	211	9,8	13293	10,2	19836	12,8
5354 - 6762 ^з	1	4,8	6248	8,5	106	4,9	10900	8,4	10490	6,8
6762 – 8170	2	9,5	16215	22,0	175	8,1	17400	13,4	21242	13,7
Всього	21	100	73734	100	2154	100	130312	100	155259	100

76

Дане групування показує, що більше половини аграрних підприємств району (52,4%) мають площу сільськогосподарських угідь від 2538 до 3946 га, в них зайняті 43,9% працівників галузі. Маючи лише 27,3% від загальної вартості основних засобів, вони виробляють 47% валової продукції сільського господарства району .

Задача 3.5. За даними табл. 3.2 методом аналітичного групування дослідити характер залежності між основними показниками розміру підприємства - площею сільськогосподарських угідь та валовою продукцією.

Розв'язання.

Факторна ознака – площа сільськогосподарських угідь, **результативна** – валова продукція (в постійних цінах 2010 року) в середньому на 1 підприємство, тис. грн.

Групування проведемо за факторною ознакою, тобто здійснюємо *факторне аналітичне групування*, (табл. 3.12).

Таблиця 3.12

Залежність між площею сільськогосподарських угідь та вартістю валової продукції

Групи підприємств за площею сільськогосподарських угідь, га	Кількість підприємств	Валова продукція (в постійних цінах 2010 року), тис. грн.	Валова продукція в середньому на 1 підприємство, тис. грн.
1130 - 2538	5	53151	6153,4
2538 - 3946	11	35568	6629,6
3946 - 5354	2	13293	9917,9
5354 - 6762	1	10900	10489,8
6762 – 8170	2	17400	10620,8
Разом	21	130312	7393,3

Таким чином, наявна пряма залежність між площею сільськогосподарських угідь та середньою вартістю валової продукції, тобто із збільшенням основного ресурсу аграрного підприємства - площі сільськогосподарських угідь - зростає і кінцевий результат виробництва – обсяг валової продукції.

Задача 3.6. За даними табл. 3.2 виявити розподіл підприємств за площею сільськогосподарських угідь та рівнем використаної праці, а також вплив цих ознак на обсяг виробництва.

Розв'язання.

Ми вже маємо розподіл підприємств за площею сільськогосподарських угідь. Рівень використаної праці визначається середньорічною кількістю працівників. Таким чином, необхідно сформувати групи підприємств за кількістю працівників

При визначенні розміру сільськогосподарських підприємств одним з критеріїв є саме середньорічна кількість працівників, яка не повинна перевищувати 50 осіб - для малих і 250 осіб - для середніх підприємств.

Таким чином має сенс створити 3 ключові групи з відкритими інтервалами: до 50 робітників, 50-250, більше 250 осіб.

Комбінаційне групування представлене в табл. 3.13.

Таблиця 3.13

Комбінаційне групування підприємств за площею сільськогосподарських угідь та рівнем використаної праці

Групи підприємств за площею с.-г. угідь, га	Групи підприємств за чисельністю робітників, осіб.	Кількість підприємств	Валова продукція, тис. грн.	
			всього	в середньому на одне підприємство
1130 - 2538	до 50	3	10547	3516
	50-250	-	-	-
	більше 250	2	20220	10110
2538 - 3946	до 50	2	5592	2796
	50-250	9	67333	7481
	більше 250	-	-	-
3946 - 5354	до 50	-	-	-
	50-250	2	19836	9918
	більше 250	-	-	-
5354 - 6762	до 50	-	-	-
	50-250	1	10490	10490
	більше 250	-	-	-
6762 – 8170	до 50	-	-	-
	50-250	2	21242	10621
	більше 250	-	-	-
Разом		24	130312	7393,3

Найбільшу сукупну вартість валової продукції виробляють 9 підприємств з площею сільськогосподарських угідь від 2538 до 3946 га та кількістю робітників від 50 до 250 осіб, кожне з яких виробляє продукції в середньому майже на 7,5 млн. грн.

Однак найбільш продуктивними є 2 підприємства з максимальною площею сільськогосподарських угідь (6762 – 8170 га) та середньою кількістю робітників (від 50 до 250 осіб), кожне з яких виробляє продукції в середньому на 10,6 млн. грн.

3.4 Порівнянність статистичних групувань

Групування, побудовані за один і тієї самий період години, але в різних регіонах або для одного регіону, але за два різних періоди можуть оказатися неспівставними по причині різної кількості виділених груп або неоднаковості границь інтервалів.

Для того, щоб привести такі групування до співставного виду (що дозволяє провести їх порівняльний аналіз), використовується метод вторинного групування. Сутність методу полягає в перегрупуванні одиниць об'єкту без звернення до первинних даних.

Вторинне групування - операція по створенню нових груп на основі раніше побудованого групування.

Способи вторинного групування:

1) **об'єднання первинних інтервалів** – укрупнення інтервалів шляхом сумування відповідних показників первинних груп, використовується у випадку переходу від дрібних до більш крупних інтервалів, а також, коли границі нових та старих інтервалів співпадають;

Задача 4.7. Маємо наступні дані групування магазинів за товарооборотом (табл. 3.14) Необхідно ущільнити ряди розподілу, створивши 5 груп.

Розв'язання.

Ущільнимо ряди розподілу, утворивши п'ять груп. Нові групи утворені шляхом підсумовування первісних груп (табл. 3.15).

Таблиця 3.14

Групування магазинів за товарооборотом

Групи магазинів за товарооборотом за квартал, тис. грн.	Кількість магазинів	Товарооборот за квартал, тис. грн.	Середній товарооборот, тис. грн.
До 10	20	120	6
10 — 15	10	110	11
15 — 20	15	250	16,6
20 — 30	5	125	25
30 — 50	10	400	40
50 — 60	7	400	57,1
60 — 70	5	320	64
70 — 100	8	650	81,25
100 — 200	25	3000	120
Понад 200	15	4500	300
Разом	120	9875	82,3

Таблиця 3.15

Вторинне групування магазинів за товарооборотом

Групи магазинів за товарооборотом, тис. грн.	Кількість магазинів	Товарооборот за квартал, тис. грн.	Середній товарооборот тис. грн.
До 20	$20+10+15=45$	$120+110+250=480$	10,7
20 — 50	$5+10=15$	$125+400=525$	35
50 — 100	$7+5+8=20$	$400+320+650=1370$	68,5
100 — 200	25	3000	120
Понад 200	15	4500	300
Разом	120	9875	82,3

2) **часткове перегрупування** – створення нових груп на основі закріплення за кожною групою певної частини одиниць сукупності, використовується у випадку, коли необхідно в ході перегрупування даних визначити, яка частина одиниць перейде із старих груп до нових.

Задача 3.8. Маємо дані про розподіл сільських поселень за числом особистих селянських господарств (ОСГ) (табл. 3.16). Необхідно привести ряди розподілу до співставного виду.

Таблиця 3.16

Розподіл сільських поселень двох районів за числом ОСГ

№	Групи поселень за числом ОСГ району А	Частка господарств групи у % до підсумку	№	Групи поселень за числом ОСГ району Б	Частка поселень групи в % до підсумку
1	До 100	5,0	1	до 50	1,0
2	100 — 200	18,0	2	50 - 70	2,0
3	200 — 300	20,0	3	70 - 100	3,0
4	300 — 500	30,0	4	100 - 150	10,0
5	Понад 500	27,0	5	150 - 250	18
			6	250 - 400	21
			7	400 - 500	22
			8	понад 500	23
	Разом	100		Разом	100

Розв'язання.

Наведені в табл. 3.16 дані не дозволяють провести порівняння розподілу сільських поселень в 2-х районах по числу особистих селянських господарств, тому що в цих районах є різне число груп поселень. Необхідно ряди розподілу привести до порівнянного виду, обравши за основу розподіл одного з районів.

Варіант 1.

За основу порівняння візьмемо розподіл поселень району А.

Отже, по району Б треба зробити вторинне групування, щоб утворити таке ж число груп і з тими ж інтервалами, як й у першому районі.

1) У першу створену групу з числом ОСГ до 100 ввійдуть три групи сільських поселень району Б, сума часток яких складе:

$$1+2+3=6$$

2) Тепер треба створити другу групу сільських поселень з числом ОСГ 100-200. До неї входить:

- четверта група 100-150 з часткою 10%;

- певна частина п'ятої групи.

Для визначення числа сільських поселень, які треба взяти з п'ятої групи в знову утворену, умовно приймемо, що це число поселень повинне бути пропорційно питомій вазі відібраних ОСГ у групі.

Визначаємо питома вагу 50 ОСГ у п'ятій групі:

$$\frac{100-50}{250-150} = \frac{1}{2}.$$

Тобто, у нову групу 2 треба взяти з п'ятої групи:

$$\frac{1}{2} * 18 = 9.$$

Таким чином, питома вага нової групи 2 складе:

$$10+9=19.$$

3) У нову групу 3 з числом ОСГ 200-300 входить:

- залишок частки групи 5:

$$18 - 9 = 9;$$

- частина групи 6 району Б – 50 ОСГ.

Визначаємо їхню частку:

$$\frac{300-250}{400-250} = \frac{1}{3}.$$

Тобто, у нову групу 3 треба взяти з шостої групи:

$$\frac{1}{3} * 21 = 7.$$

Таким чином, питома вага нової групи 3 складе:

$$9+7=16.$$

4) У нову групу 4 з числом ОСГ 300-500 входить:

- залишок частки групи 6:

$$21 - 7 = 14;$$

- питома вага групи 7 – 22%.

Таким чином, питома вага нової групи 3 складе:

$$14+22=36.$$

5) Нова група 5 включає 23% восьмої групи поселень району Б.

Провівши вторинне групування одержимо цілком співставні розподіли сільських поселень за числом ОСГ по обох районах (табл. 3.17).

Таблиця 3.17.

Розподіл сільських поселень за числом ОСГ
(база - розподіл поселень району А)

№	Групи поселень за числом ОСГ	Питома вага поселень групи в % до підсумку		Розрахунки
		Район А	Район Б	
1	до 100	5,0	6,0	$1+2+3=6$
2	100 - 200	18,0	19,0	$10+1/5*18=19$
3	200 - 300	20,0	16,0	$9+1/3*7=16$
4	300 - 500	30,0	36,0	$(21-7)+22=36$
5	понад 500	27,0	23,0	23
	Разом	100,0	100,0	x

Варіант Б.

За основу групування візьмемо розподіл поселень району Б.

Інтервал групи до 100 ОСГ буде розбитий на 3 нові групи:

1) у створену групу 1 з числом ОСГ до 50 ввійде половина питомої ваги першої групи сільських поселень району А:

$$\frac{50}{100} * 5 = \frac{1}{2} * 5 = 2,5.$$

2) питома вага нової групи 2 з інтервалом 50-70 складе:

$$\frac{70 - 50}{100} * 5 = \frac{1}{5} * 5 = 1.$$

3) питома вага нової групи 3 з інтервалом 70-100 складе:

$$\frac{100 - 70}{100} * 5 = \frac{3}{10} * 5 = 1,5.$$

4) частка нової групи 4 з інтервалом 100-150 складе половину питомої ваги другої групи сільських поселень району А:

$$\frac{150 - 100}{300 - 200} * 18 = \frac{1}{2} * 18 = 9.$$

5) у частку нової групи 5 з інтервалом 150-250 входить:

- залишок частки групи 2:

$$18 - 9 = 9;$$

- частина групи 3 району А - 50 ОСГ.

Визначаємо їхню частку:

$$\frac{250 - 200}{300 - 200} * 20 = 0,5 * 20 = 10$$

Таким чином, питома вага нової групи 5 складе:

$$9 + 10 = 19$$

б) у частку нової групи 5 з інтервалом 250-400 входить:

- залишок частки групи 3:

$$20 - 10 = 10;$$

- частина групи 4 району А - 100 ОСГ. Визначаємо їхню частку:

$$\frac{500 - 400}{500 - 300} = \frac{1}{2} * 30 = 15$$

Таким чином, питома вага нової групи 6 складе:

$$10 + 15 = 25$$

7) питому вагу нової групи 7 складе залишок групи 4 району А:

$$30 - 15 = 15$$

8) питома ваги нової групи 8 співпадає з часткою п'ятої групи району А, так як співпадають їхні границі, і становить 27% (табл. 3.18).

Таблиця 3.18

Розподіл сільських поселень за числом ОСГ
(база - розподіл поселень району Б)

№	Групи поселень за числом ОСГ	Питома вага поселень групи в % до підсумку		Розрахунки
		Район Б	Район А	
1	до 50	1,0	2,5	$1/2 * 5 = 2,5$
2	50 - 70	1,0	1,0	$1/5 * 5 = 1$
3	70 - 100	2,0	1,5	$3/10 * 5 = 1,5$
4	100 - 150	10,0	9,0	$0,5 * 18 = 9$
5	150 - 250	18,0	19,0	$9 + 1/2 * 20 = 19$
6	250 - 400	21,0	25,0	$(20 - 10) + 1/2 * 30 = 25$
7	400 - 500	23,0	15,0	$30 - 15 = 15$
8	понад 500	24,0	27,0	27
Разом		100,0	100,0	x

Завдання для самостійного розв'язання.

Завдання 3.1. При вивченні думки населення щодо ринку цінних паперів серед запитань анкети були самооцінка схильності до ризику та пріоритетність вкладання вільних коштів. Відповіді респондентів на ці запитання наведені у таблиці:

№	Вік, років	Схильність до ризику*	Напрямок вкладання коштів**	№	Вік, років	Схильність до ризику	Напрямок вкладання коштів
1	68	1	2	16	45	3	2
2	29	2	4	17	23	2	2
3	43	2	1	18	61	1	3
4	32	2	2	19	33	2	2
5	38	2	3	20	42	1	3
6	62	1	3	21	27	3	1
7	35	1	3	22	37	2	3
8	22	3	1	23	65	1	4
9	56	3	2	24	40	2	4
10	50	2	1	25	49	2	2
11	25	3	2	26	57	2	3
12	67	1	2	27	19	3	2
13	21	3	3	28	70	2	3
14	39	1	4	29	36	2	2
15	28	2	1	30	52	3	2

* коди самооцінки схильності до ризику:

- ризику уникаю взагалі – 1
- готовий ризикувати з певними гарантіями – 2
- люблю ризикувати – 3

**коди напрямків вкладання коштів:

- цінні папери – 1
- нерухомість – 2
- валюта – 3
- товар – 4

Необхідно: 1) згрупувати респондентів:

- а) за віком, виділивши 3 групи: до 30 років, 30-60 років, 60 років та старше;
- б) за схильністю до ризику, виділивши 3 групи ризикових, обережних та неризикових потенційних інвесторів;
- в) за пріоритетними напрямками вкладання вільних коштів;

- 2) скласти комбінаційні розподіли респондентів:
- а) за віком та за схильністю до ризику;
 - б) за віком та пріоритетними напрямками вкладання вільних коштів;
 - в) за схильністю до ризику та пріоритетними напрямками вкладання вільних коштів.

Зробити висновки.

Завдання 3.2. Здійснити типологічне групування підприємств регіону за ознакою «галузь промисловості» (енергетика, легка промисловість, харчова промисловість, машинобудування, хімічна) за табличними даними:

№	Підприємство	Обсяг реалізованої продукції, млн. грн.	
		за планом	фактично
1	Кондитерська фірма	20,2	21,3
2	Фармакологічний комбінат	28,9	30,2
3	Швейна фабрика	8,6	8,6
4	ТЕЦ	137,5	133,2
5	Автобусний завод	238,0	176,0
6	Консервний завод	55,0	60,0
7	Бавовно-прядильний комбінат	35,2	13,1
8	Взуттєва фабрика	165,1	130,4
9	Гідроелектростанція	102,1	105,0
10	Хлібозавод	63,1	67,2
11	Олієжировий комбінат	45,6	48,0
12	Комбінат побутової хімії	15,4	20,5
13	Інструментальний завод	45,8	12,3
14	Лакофарбовий завод	20,8	20,5
15	Завод безалкогольних напоїв	6,8	7,6
16	Тракторний завод	300,6	300,0
17	М'ясокомбінат	88,6	90,4
18	Трикотажна фабрика	10,3	8,9
19	Завод мінеральних добрив	34,5	42,6
20	Молокозавод	38,5	40,5

- Визначити: 1) обсяг реалізованої продукції для кожної галузі;
 2) відсоток виконання плану в цілому для галузі;
 3) питому вагу реалізованої продукції (за планом та фактом) для кожної галузі.

Результати зобразити у табличній формі та зробити висновки.

Завдання 3.3. Маємо дані про товарооборот і торговельну площу 18 магазинів за звітний період:

№	Товарооборот, тис. грн.	Торговельна площа, кв. м.	№	Товарооборот, тис. грн.	Торговельна площа, кв. м.
1	192	163	10	165,4	113
2	37,3	32	11	140,2	30
3	164,4	101	12	114,6	118
4	32,7	36	13	57,8	20
5	161,5	180	14	73,4	115
6	67,9	117	15	87,5	103
7	64,9	96	16	50,8	48
8	102,3	96	17	111	110
9	89,2	105	18	93,5	86

1) Зробити групування магазинів за торговельною площею, розбивши сукупності на чотири групи з рівними інтервалами;

2) По кожній групі й підсумку в цілому підрахуйте: кількість магазинів; торговельну площу: всього й в середньому на один магазин; товарооборот: всього, в середньому на один магазин і в середньому на 1 кв. м торговельної площі.

Завдання 3.4. Маємо дані даними про якість ґрунтів та урожайність озимої пшениці в сільськогосподарських підприємствах регіону:

№	Середня якість ґрунту (балів)	Урожайність пшениці, ц/га	№	Середня якість ґрунту (балів)	Урожайність пшениці, ц/га
1	50	41	11	54	39
2	53	30	12	43	40
3	42	33	13	47	25
4	48	34	14	49	37
5	45	41	15	52	32
6	53	43	16	55	41
7	40	39	17	47	24
8	46	35	18	51	28
9	48	28	19	43	29
10	49	33	20	48	35

Виконати аналітичне групування для вивчення залежності урожайності озимої пшениці від якості ґрунтів, утворивши чотири групи з рівними інтервалами за результативною ознакою. По кожній групі визначити кількість сільськогосподарських підприємств, середню якість ґрунтів та середню урожайність озимої пшениці. Результати розрахунків подати у вигляді таблиці, зробити висновки.

Завдання 3.5. Маємо дані про кількість пропусків годин занять за семестр 24 студентів однієї з академічних груп:

№	Пропуски, год.	Стать	№	Пропуски, год.	Стать	№	Пропуски, год.	Стать
1	8	Ж	9	36	Ч	17	10	Ж
2	6	Ж	10	18	Ж	18	6	Ч
3	-	Ж	11	10	Ч	19	8	Ч
4	4	Ч	12	14	Ж	20	38	Ч
5	12	Ч	13	-	Ж	21	54	Ч
6	20	Ч	14	-	Ж	22	24	Ж
7	16	Ж	15	2	Ж	23	32	Ж
8	42	Ж	16	30	Ч	24	26	Ж

За цими даними провести комбінаційне групування, виділивши 3 рівні групи по пропусках занять і дві групи за статтю. Для кожної виділеної групи й по академічній групі в цілому визначити число студентів, загальну кількість пропусків і в середньому на 1 студента.

Завдання 3.6. Використовуючи дані таблиці:

№	Країна	Територія, тис. кв. км	Чисельність населення, млн. осіб	Валовий внутрішній продукт, млрд. дол.
1	Куба	111	11,5	108
2	Ісландія	103	0,3	12
3	Азербайджан	87	8	72
4	Албанія	29	4	20
5	Польща	313	39	625
6	Румунія	238	22	247
7	Данія	43	5,5	205
8	Франція	547	64	2097
9	Нідерланди	41,5	17	670
10	Бельгія	31	10	495
11	Фінляндія	337	5	195
12	Австрія	84	8	325
13	Великобританія	245	61	2231
14	Таджикистан	143	7	12
15	Туркменістан	488	5	47
16	Греція	132	11	343
17	Португалія	92	11	237
18	Японія	378	127	4348
19	Ізраїль	21	7	201
20	Узбекистан	447	28	62

Здійснити: 1) групування країн за територією, сформувавши групи з рівними інтервалами.

2) комбіноване групування країн за площею територій і валовим внутрішнім продуктом на душу населення, ураховуючи, що високим є середньодушовий ВВП у розмірі вище 10,5 тис. дол., середнім - від 9,5 до 10,5 тис. дол., низьким - до 9,5 тис. дол.

Для кожної групи встановити: число країн, площу території, чисельність населення (усього й у середньому на одну країну), а також щільність населення. Результати групування оформити у вигляді таблиці. Зробити висновки.

Завдання 3.7. Маємо наступні дані про розподіл робітників за рівнем праці на двох підприємствах:

Підприємство А		Підприємство Б	
Продуктивність праці, шт.	Питома вага робітників, %	Продуктивність праці, шт.	Питома вага робітників, %
100-120	2	-	-
120-140	3	130-160	15
140-160	12	160-190	25
160-180	19	190-220	26
180-200	25	220-250	15
200-220	30	250-280	12
220-240	4	280-300	7
240-260	3	-	-
260-300	2	-	-
Разом	100	Разом	100

Здійснити вторинне перегрупування, утворивши по обох підприємствах групи з інтервалами: 100-140, 140-180, 180-220, 220-260, 260-300.

Тести для самоконтролю.

1. Систематизація обробка і підрахунок групових і загальних підсумків даних статистичного спостереження називається:

- А) статистичне групування
- Б) статистичний аналіз
- В) статистичне зведення
- Г) статистичне спостереження

2. Групування, що застосовується для дослідження взаємозв'язку між окремими ознаками одиниць статистичної сукупності, називається:

- А) структурне
- Б) типологічне

- В) аналітичне
- Г) комбіноване

3. Ряди розподілу, утворені за кількісними ознаками, називаються:

- А) кількісні
- Б) атрибутивні
- В) варіаційні
- Г) прості

4. Чисельність одиниць у кожній групі називається:

- А) частота
- Б) частка
- В) інтервал
- Г) обсяг групи

5. За кількістю ознак, покладених в основу, групування за однією ознакою називається:

- А) просте
- Б) групове
- В) комбіноване
- Г) множинне

6. Якщо групувальна ознака має плавний характер і застосовуються рівні інтервали, то кількість груп орієнтовно можна визначити за формулою:

- А) Ст'юдента
- Б) Варзара
- В) Стерджеса
- Г) Пірсона

7. Групування, що характеризує якісні особливості й розходження між типами явищ, називається:

- А) аналітичне
- Б) комбіноване
- В) типологічне
- Г) структурне

8. За величиною інтервали групування, в яких різниця між верхньою і нижньою межею однакова, називаються:

- А) цілісні
- Б) рівні
- В) обмежені
- Г) подібні

9. Найпростіше групування, у якому кожна виділена група характеризується чисельністю одиниць сукупності, що потрапили в кожну групу, називається:

- А) просте групування
- Б) статистичне зведення
- В) груповий розподіл

Г) ряд розподілу

10. Числові значення кількісної ознаки у варіаційному ряді називаються:

А) показниками

Б) варіантами

В) частками

Г) частотами

11. Варіаційні ряди, у яких числові значення ознаки можуть набувати будь-яких значень в певних межах, границях, називаються:

А) інтервальні

Б) дискретні

В) прості

Г) рівні

12. Операція по створенню нових груп на основі раніше побудованого групування називається:

А) побудова ряду розподілу

Б) статистичний аналіз

В) вторинне групування

Г) статистичне зведення

13. Сума всіх частот - це:

А) частка

Б) обсяг ряду розподілу

В) інтервал

Г) обсяг явища

13. Якщо зведення виконують без розподілу одиниць сукупності на групи, здійснюється підрахунок тільки загальних підсумків, воно називається:

А) просте

Б) групове

В) загальне

Г) складне

14. За величиною інтервали групування, в яких різниця між верхньою і нижньою межею неоднакова, називаються:

А) нерівні

Б) неадекватні

В) прості

Г) необмежені

15. Вторинне групування, що здійснюється шляхом сумування відповідних показників первинних груп і використовується у випадку переходу від дрібних до більш крупних інтервалів

А) укрупнення інтервалів

Б) об'єднання первинних інтервалів

В) часткове перегрупування

Г) укрупнене перегрупування

16. Варіаційні ряди, у яких числовий розподіл ознаки виражений певними, зазвичай цілими, числами, називаються:

- А) прості
- Б) дискретні
- В) рівні
- Г) інтервальні

17. Групування, що передбачає виявлення закономірностей розподілу одиниць однорідної сукупності за варіюючим значенням ознаки, називається:

- А) аналітичне
- Б) типологічне
- В) комбіноване
- Г) структурне

18. За визначеністю меж інтервали групування, у яких відомі нижня та верхня межа інтервалу, називаються:

- А) прості
- Б) обмежені
- В) граничні
- Г) закриті

19. За кількістю ознак, покладених в основу, одночасне групування за двома і більше ознаками, називається:

- А) комбінаційне
- Б) просте
- В) групове
- Г) множинне

20. Якщо здійснюється розподіл сукупності на групи за істотною для одиниць сукупності ознакою, процес називається:

- А) загальне зведення
- Б) просте зведення
- В) статистичний аналіз
- Г) статистичне групування

21. За визначеністю меж інтервали групування, у яких відомі або нижня, або верхня межа інтервалу, називаються:

- А) відкриті
- Б) безграничні
- В) необмежені
- Г) розірвані

23. Метод вторинного групування, при якому певна частина одиниць переходить із старих груп до нових:

- А) укрупнення інтервалів
- Б) об'єднання первинних інтервалів
- В) укрупнене перегрупування
- Г) часткове перегрупування

4. УЗАГАЛЬНЮЮЧІ СТАТИСТИЧНІ ПОКАЗНИКИ

4.1 Сутність та види статистичних показників

Статистичний показник – це кількісно та якісно виражена характеристика певної властивості сукупності в цілому чи її частин в умовах конкретного місця і часу.

У відповідності з визначенням статистичний показник має певну структуру, в ньому розрізняють *якісну і кількісну сторони*. *Якісна сторона* статистичного показника визначається ознакою, яка підлягає вивченню і відбивається в назві показника, *кількісна сторона* - в чисельному значенні показника.

Величина показника зазвичай варіюється в просторі і коливається в часі, тому обов'язковим атрибутом статистичного показника є також зазначення території або об'єкту та моменту або періоду часу.

Статистичне дослідження незалежно від його масштабів і цілей завжди передбачає розрахунок і аналіз різних за видом і формою вираження статистичних показників (рис. 4.1).

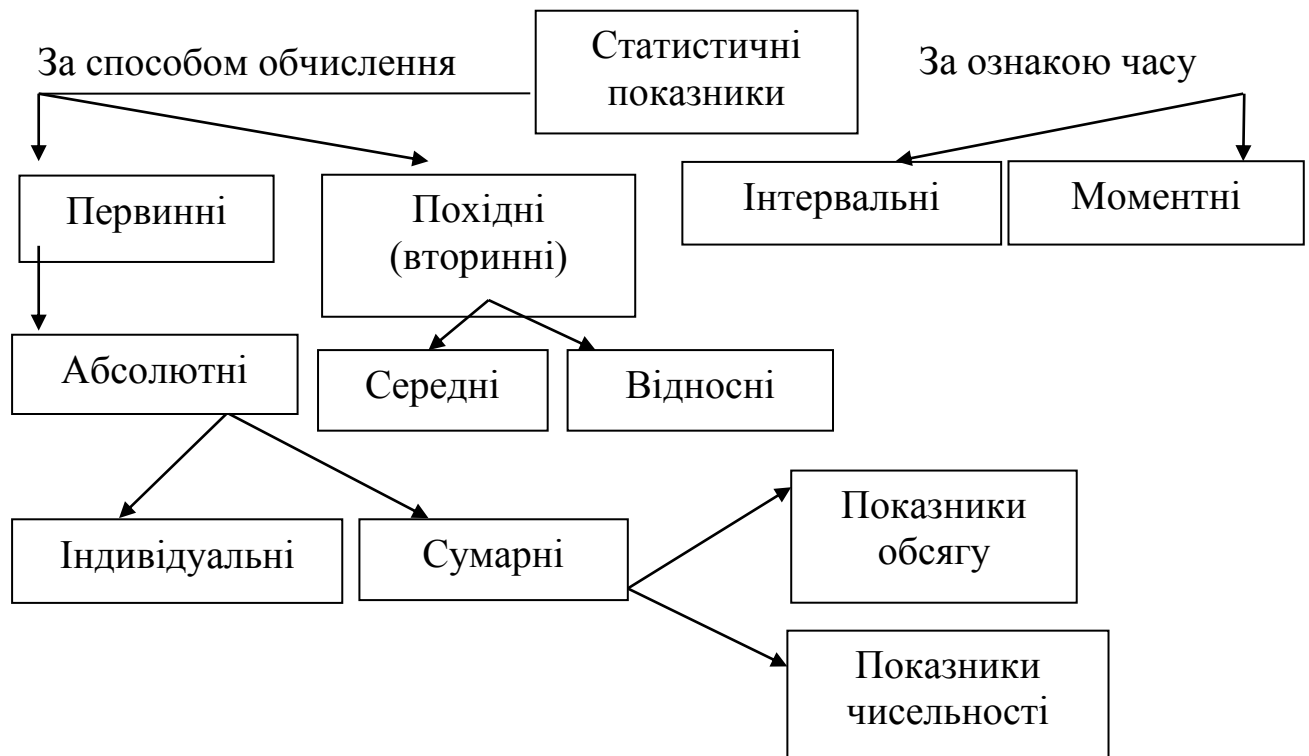


Рис. 4.1 Класифікація статистичних показників

За способом обчислення статистичні показники бувають:

- **первинні** – визначаються шляхом зведення та групування даних і подаються у формі абсолютних величин
- **похідні** – обчислюються на базі первинних і мають форму середніх чи відносних величини.

За ознакою часу розрізняють статистичні показники:

- **інтервальні** - характеризують явище за певний час (тиждень, місяць, квартал, рік, тощо)

- **моментні** - характеризують явище на певний момент часу (дату, кінець місяця, початок року і т.п.)

Вимоги до найменування статистичного показника.

У ньому повинні бути вказані:

1. *Статистична структура показника* – сума, середнє, відхилення і т.д.;

2. *Зміст* – обсяг виробленої продукції, чисельність робітників, вартість товарної продукції і т.д.;

3. *Сукупність об'єктів, класифікаційна рубрика, група* – промислові підприємства Запорізької області, фермерські господарства Мелітопольського району і т.д.;

4. *Час* – на кінець 2012 року, за 2001-2010 роки, 12.03.2013 і т.д.;

5. *Одиниця виміру* – центнерів, тис. осіб, млн. грн., відсотків і т.д.;

6. *Спеціальні уточнення* – наприклад, у постійних цінах 2010 року.

Статистичні показники повинні правильно кількісно вимірюватися з урахуванням рівня, масштабів і якісних ознак стану або розвитку відповідного економічного чи соціального явища (галузевий та регіональний рівні, окреме підприємство, працівник і т.п.).

При цьому побудова показників повинна носити наскрізний характер, що дозволяє не тільки підсумовувати відповідні показники, але й забезпечувати їх якісну однорідність в групах і сукупностях, перехід від одного показника до іншого для повної характеристики обсягу і структури більш складної категорії або явища.

Нарешті, побудова статистичного показника, його структура та сутність повинні передбачати можливість всебічно аналізувати досліджуване явище чи процес, характеризувати особливості його розвитку, визначати основні фактори впливу.

Таким чином, обчислення статистичних величин і аналіз даних про досліджуваних явищах - це третій і завершальний етап статистичного дослідження.

4.2 Абсолютні величини

Первинна статистична інформація виражається насамперед у вигляді абсолютних показників, які є кількісною базою всіх форм обліку.

Абсолютною величиною називається статистичний показник, що виражає абсолютні розміри досліджуваного явища.

Абсолютні величини виражають або чисельність одиниць досліджуваної сукупності, її окремих складових частин, або абсолютні

розміри в натуральних одиницях, що впливають з фізичних властивостей (вага, довжина і т. п.), або в одиницях виміру, які впливають з економічних властивостей (вартості, витрат праці).

За способом вираження розмірів досліджуваних явищ абсолютні величини підрозділяються на:

- **індивідуальні** - відображають абсолютні розміри кожної окремої одиниці сукупності;

- **сумарні (групові)** - утворюються підсумовуванням індивідуальних абсолютних величин.

Абсолютні величини - завжди іменовані числа, що мають певну розмірність, одиниці виміру.

Залежно від різних причин і цілей аналізу застосовують такі **одиниці виміру абсолютних показників**:

- **натуральні** - відповідають природним або споживчим властивостям предмета вивчення й виражаються у фізичних мірах ваги, довжини й т.д. У ряді випадків використовуються складені (комбіновані) одиниці виміру, що представляють собою добуток двох величин, виражених в різних розмірностях. Наприклад, вантажообіг визначається відстанню (в км), яку пройшов транспорт з вантажем (т) і виражається в тонно-кілометрах (т-км);

- **умовно-натуральні** - одержують, приводячи різні натуральні одиниці до однієї, прийнятої за базу. В них вимірюється однорідна, але неоднакова продукція (умовні голови, умовні трактори, умовні еталонні гектари, умовні консервні банки, кормові одиниці, калорії, тощо);

- **трудові** - одиниці, що використовують для характеристики показників, які дозволяють оцінити витрати праці, відображають наявність, розподіл і використання трудових ресурсів;

- **вартісні** – грошове вираження обсягу показника.

Абсолютні показники не дають відповіді на питання, яку частку має та чи інша частина в загальній сукупності, не можуть охарактеризувати рівні планового завдання, ступінь виконання плану, інтенсивність того чи іншого явища, так як вони не завжди придатні для порівняння і тому часто використовуються лише для розрахунку відносних величин.

4.3 Відносні величини: сутність, види, одиниці виміру

Відносна величина – кількісне співвідношення між показниками, що утворюється у результаті ділення однієї величини на іншу, прийняту за основу порівняння, і показує, у скільки разів одна величина більша за іншу або яку частку другої становить перша.

Порівнювальний показник - показник, що відображає явище, що вивчається, і порівнюється з іншим.

База порівняння - показник, з яким здійснюється порівняння.

Примітка. Порівнювальний показник завжди знаходиться в чисельнику, а база порівняння - в знаменнику відношення.

Порівнювальний показник та база порівняння можуть бути *однойменними* або *різнойменними*.

Одиниці виміру відносних показників.

- при порівнянні *однойменних показників*:

а) **коефіцієнт** - якщо значення бази приймається за одиницю, тобто відбувається простий розподіл двох величин;

б) **відсоток (процент)** - за базу приймають 100, позначається «%»;

в) **проміле** - за базу приймають 1000, позначається «‰». Використовується при розрахунку коефіцієнтів народжуваності або смертності, у фармації, при визначенні вмісту алкоголю у крові й т.д.;

г) **продециміле** - за базу приймають 10 000, позначається «⁰/₀₀₀». У продециміле зазвичай виражають забезпеченість населення лікарняними ліжками, місцями у вищих навчальних закладах і т.д.; індекс наповнення травного тракту при вивченні харчування риб, тощо;

д) **просантиміле** — за базу приймають 100 000, позначається «⁰/₀₀₀₀». В просантиміле, наприклад, виражають захворюваність відносно рідкісними хворобами;

- при порівнянні *різнойменних показників*:

- **іменовані величини** – сполучення вимірників порівнювального показника та бази порівняння (щільність населення - осіб/км², вартість продукції в розрахунку на 100 га земельної площі - грн./100 га, продуктивність праці по окремих видах продукції - ц/люд.-год., і т.д.).

Види відносних величин.

1. Відносна величина планового завдання (намірів, договірних зобов'язань) - співвідношення планової (очікуваної, договірної) величини показника до його фактично досягнутого рівня в попередньому періоді або в базисному періоді:

$$ВВПЗ = \frac{\text{Планове завдання}}{\text{Фактичний рівень в періоді, прийнятому за базу}}$$

Відносна величина планового завдання показує, в скільки разів рівень плану більший або менший за фактичний рівень.

Задача 4.1. Розмір мінімальної заробітної плати в Україні з 1 січня 2013 року становив 1147 гривень, а 1 грудня цього ж року було заплановано ще одне її підвищення - до 1218 гривень. Визначити відсоток підвищення заробітної плати.

Розв'язання.

Поділимо планову мінімальну зарплату на фактичну й одержимо:

$$ВВПЗ = \frac{1218}{1147} * 100 = 106,2\%$$

Планується підвищити зарплату на 6,2 %.

2. Відносна величина виконання плану (реалізації намірів, виконання договірних зобов'язань) - співвідношення величини показника, досягнутої за якийсь час або який-небудь момент, і величини його, установленної за планом (наміром, договором) за цей же час. Дає кількісну характеристику виконання планових завдань.

а) процентне відношення фактично досягнутої у звітному періоді абсолютної (або середньої) величини рівня до абсолютної (або середньої) величини рівня планового завдання:

$$ВВВП = \frac{\text{Фактичне виконання}}{\text{Планове завдання}} * 100\%$$

Задача 4.2. За планом передбачалося зібрати по 30 ц озимої пшениці з гектара, а зібрано 33 ц/га. Визначити відсоток виконання плану по врожайності озимої пшениці.

Розв'язання.

Поділимо фактичну врожайність на планову й одержимо:

$$\frac{33}{30} * 100 = 110\%$$

План по врожайності виконаний на 110%, або перевиконаний на 10%.

б) порівняння фактично досягнутого та запланованого (очікуваного) зниження рівня:

$$ВВВП = \frac{\text{Фактична величина зниження}}{\text{Планова величина зниження}}$$

Задача 4.3. За планом в 2012 р. намічено зниження собівартості виробу на 5 грн. при рівні собівартості 175 грн. Фактично в 2012 р. собівартість виробу склала 169,8 грн. Визначити відносну величину виконання плану по зниженню собівартості виробу „А” в 2012 р.

Розв'язання.

Фактична величина зниження собівартості виробу:

$$175 - 169,8 = 5,2 \text{ (грн.)}$$

$$ВВВП = \frac{5,2}{5} * 100 = 104,0\%$$

Планове завдання по зниженню собівартості перевиконано на 4%.

в) порівняння відносних величин, що характеризують фактичний ріст явища та частку виконання планового завдання (намірів, договірних зобов'язань):

$$ВВВП = \frac{\text{Коефіцієнт (відсоток) фактичного зростання}}{\text{Коефіцієнт (відсоток) планового завдання}}$$

Задача 4.4. По договорах приріст обсягу реалізації продукції підприємством на 2012 р. повинний скласти 7,5%. Фактично зростання випуску продукції в цьому році становило 105,5%. Визначити відносну величину виконання договірних зобов'язань щодо реалізації продукції підприємством.

Розв'язання.

Плановий коефіцієнт зростання випуску продукції:

$$1 + 0,075 = 1,075 \quad \text{або} \quad 100\% + 7,5\% = 107,5\%$$

$$ВВВП = \frac{1,055}{1,075} * 100\% = \frac{105,5}{107,5} * 100\% = 98,1\%$$

Договірні зобов'язання щодо обсягу реалізації продукції підприємством виконані на 98,1%, тобто обсяг реалізації продукції на 1,9% менший за вказаний в договорі.

3. Відносна величина динаміки - співвідношення величини показника за даний час і величини його за який-небудь попередній час, прийнятий за базу порівняння. Характеризує ступінь, швидкість зміни показника в часі.

а) порівняння ознаки кожного періоду з ознакою попереднього для нього періоду (ланцюговий метод):

$$ВВД = \frac{\text{Рівень ознаки в поточному періоді}}{\text{Рівень ознаки в попередньому періоді}}$$

б) порівняння ознак періодів з ознакою одного й того самого період, прийнятого за базу (базисний):

$$ВВД = \frac{\text{Рівень ознаки в поточному періоді}}{\text{Рівень ознаки в базисному періоді}}$$

Задача 4.5. Маємо наступні дані про зміну середньої ціни реалізації зерна в господарстві:

Показники	Роки		
	2010	2011	2012
Середня ціна реалізації зерна, грн. /ц	110,8	137,5	135,0

Визначити відносні показники динаміки ціни ланцюговим та базисним методами.

Розв'язання.

Ланцюговий метод: $VVD_{2011} = \frac{137,5}{110,8} * 100\% = 124,1\%$

$$VVD_{2012} = \frac{135,0}{137,5} * 100\% = 98,2\%$$

Базисний метод: $VVD_{2011} = 124,1\%$

$$VVD_{2012} = \frac{135,0}{110,8} * 100\% = 121,8\%$$

В період з 2010 по 2012 р. середня ціна 1 ц зерна в господарстві зросла на 21,8%, при цьому в 2011 р. вона підвищилася на 24,1%, але в 2012 р відбулося зниження ціни 1 ц зерна на 1,8% порівняно з показником 2011 р.

4. Відносна величина структури – відношення значення кожної частини сукупності до загального підсумку, який прийнято за базу:

$$BBC = \frac{\text{Частина сукупності}}{\text{Вся сукупність}}$$

Характеризує склад досліджуваної сукупності й показує яку питому вагу (частку) у загальному підсумку становить кожна її частина.

Задача 4.6. За даними тестування якості сиру 30 осіб з 50 учасників тестування дали позитивний відгук про якість продукту, 5 осіб - негативний відгук, 15 осіб - не визначилися. Визначте структуру результатів тестування якості сиру.

Розв'язання.

Позитивний відгук: $BBC = \frac{20}{50} * 100\% = 60\%$

Негативний відгук: $BBC = \frac{5}{50} * 100\% = 10\%$

Не визначилися: $BBC = \frac{15}{50} * 100\% = 30\%$

За результатами тестування якість сиру задовольнила 60% учасників, не задовольнила - 10%, а 30% - не змогли відповісти однозначно.

5. Відносна величина порівняння - співвідношення однойменних показників, що відносяться до різних об'єктів або територій, але за той самий період часу. величина ознаки одного об'єкту

$$VVPor = \frac{\text{Величина ознаки одного об'єкту}}{\text{Величина ознаки іншого об'єкту}}$$

Задача 4.7. Середня зарплата в 2012 році склала в Україні -378 дол. США на місяць, у Росії – 762, в Білорусі – 323 дол. США. Порівняти середній рівень зарплат в 2012 р. по країнах, беручи за базу дані по Україні.

Розв'язання.

Порівняємо середню зарплату в Росії та Україні:

$$ВВПор = \frac{762}{378} * 100\% = 201,6\%$$

Порівняємо середню зарплату в Білорусі та Україні:

$$ВВПор = \frac{323}{378} * 100\% = 85,45\%$$

Порівняно з середньомісячною зарплатою українців зарплата росіян є вищою на 101,6 %, а зарплата білорусів - нижчою на 14,55%.

6. Відносна величина координації характеризує співвідношення окремих частин цілого, одна з яких приймається за базу порівняння, і іноді вважається доповненням до відносної величини структури.

$$ВБК = \frac{\text{Одна зі складових сукупності}}{\text{Інша частина цієї сукупності}}$$

Показує скільки одиниць певної частини цілого припадає на одиницю іншої його частини

Задача 4.8. За даними перепису 2001 р. в Україні в сільській місцевості проживало 15,9 млн. осіб, а в міській – 32,6 млн. осіб. Визначити кількість міських жителів, що приходить на 100 сільських жителів.

Розв'язання.

$$ВБК = \frac{32,6}{15,9} * 100 = 205 \frac{\text{міських жителів}}{100 \text{ сільських жителів}}$$

На кожних 100 сільських жителів приходить 205 міських жителів.

6. Відносна величина інтенсивності - відношення двох різно-йменних, але пов'язаних між собою абсолютних величин. Переважно вона характеризує ступінь насиченості, поширеності даного явища в певному середовищі:

$$ВВІ = \frac{\text{Величина одного показника}}{\text{Величина іншого показника}}$$

Метод розрахунку відносних величин інтенсивності застосовується при визначенні середніх рівнів (середнього рівня продуктивності праці, середніх затрат праці, середньої собівартості продукції, середньої ціни товару, середньої продуктивності тварин і т.д.). Тому поширена думка, що відносні величини інтенсивності - це один із способів вираження середніх величин.

Задача 4.9. Чисельність населення України на 1 січня 2013 р. становила 45553 тис. осіб, а площа України становить 603,7 тис км². Визначити щільність населення України.

Розв'язання.

$$BVI = \frac{45553}{603,7} = 75,5 \frac{\text{осіб}}{\text{км}^2}$$

На кожні 10 км² площі України приходить майже 755 осіб.

Різновидом відносної величини інтенсивності є **відносна величина рівня економічного розвитку** – співвідношення величини найважливіших економічних показників (країни, регіону, галузі) і чисельності населення (так називані показники на душу населення).

$$BBPER = \frac{\text{Значення економічного показника}}{\text{Чисельність населення}}$$

Задача 4.10. Рівень ВВП в Україні в 2012 р. сягнув 344,7 млрд. дол. США, населення складало 45,553 млн. осіб. Визначити середньодушовий ВВП.

Розв'язання.

$$BBPER = \frac{347000}{45,553} = 7567 \frac{\text{дол.США}}{\text{особу}}$$

Рівень ВВП в розрахунку на душу населення в Україні становив близько 7,6 тис. дол. США.

Відносні величини широко використовуються при порівняльній оцінці різних показників роботи окремих підприємств, міст, регіонів, країн. При цьому, наприклад, результати роботи конкретного підприємства і т.п. приймаються за базу порівняння і послідовно співвідносяться з результатами аналогічних підприємств інших галузей, регіонів, країн і т.д.

4.4 Середні величини

Середньою величиною називається показник, що дає узагальнену характеристику мінливої ознаки одиниць однорідної сукупності.

Середня відображає те спільне, що приховується в кожній одиниці сукупності, уловлює загальні риси, загальну тенденцію, закономірність, властивому даному розподілу, вона є рівнодіючою, тому що в ній знаходить своє відображення, підсумовується й синтезується вплив всієї сукупності факторів, під впливом яких формується ряд розподілу. Середня дає також характеристику центра розподілу.

Необхідно врахувати, що середня величина повинна обчислюватися з урахуванням економічного змісту ознаки, що визначається.

Форми середніх величин:

- проста – якщо середня обчислюється за первинними незгрупованими даними;
- зважена - якщо середня обчислюється за згрупованими даними.

Види середніх величин

Об'ємні (ступеневі) середні.

1. Середня арифметична:

а) проста – використовується для осереднення прямих значень ознак шляхом їх підсумовування:

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

де $x_1, x_2, x_3, \dots, x_n$ – кількісні значення ознаки;
 n – кількість цих значень.

Задача 4.11. Маємо наступні дані про склад родин робітників малого підприємства:

№ родини	1	2	3	4	5	Разом
Кількість дітей у родині	2	3	1	2	2	10

Розв'язання.

$$\bar{x} = \frac{\sum x_i}{n} = \frac{10}{5} = 2(\text{дитини})$$

Кожний з робітників даного малого підприємства має в середньому по 2 дитини.

б) зважена – використовується, коли значення ознаки зустрічається кілька разів:

$$\bar{x} = \frac{\sum xf}{\sum f}$$

де f – кількість одиниць сукупності, що мають однакове значення ознаки (частота).

Задача 4.12. Маємо наступні дані про результати іспиту зі статистики:

Оцінка (x)	Кількість студентів (f)
2	12
3	20
4	64
5	24
Разом	120

Розв'язання.

$$\bar{x} = \frac{\sum xf}{\sum f} = \frac{2*12 + 3*20 + 4*64 + 5*24}{12 + 20 + 64 + 24} = 3,8$$

Середній бал успішності зі статистики за національної п'ятибальною шкалою складає 3,8.

Задача 4.13. Бригада токарів була зайнята обточкою однакових деталей протягом 8-годинної робочої зміни:

Номер токаря	1	2	3	4	5
Витрати часу на 1 деталь, хв.	12	15	11	16	14

Визначити середній час, необхідний на виготовлення 1 деталі.

Розв'язання.

На перший погляд, здається, що завдання легко вирішується за формулою середньої арифметичної простої:

$$\bar{x} = \frac{\sum x}{n} = \frac{12 + 15 + 11 + 16 + 14}{5} = \frac{68}{5} = 13,6(\text{хв.})$$

Здобута середня була б правильною, якби кожний токар зробив тільки по 1 деталі.

Але протягом зміни окремим токарем була виготовлена різна кількість деталей.

$$\text{Середні витрати часу на 1 деталь} = \frac{\text{Загальні витрати часу бригади}}{\text{Загальна кількість деталей}}$$

Відомо, що тривалість зміни - 8 часів (1 час = 60 хв.)

$$\begin{aligned} \bar{x} &= \frac{\frac{8*60}{12} + \frac{8*60}{15} + \frac{8*60}{11} + \frac{8*60}{16} + \frac{8*60}{14}}{5} = \frac{(8*60)*(1+1+1+1+1)}{(8*60)*\left(\frac{1}{12} + \frac{1}{15} + \frac{1}{11} + \frac{1}{16} + \frac{1}{14}\right)} = \\ &= \frac{5}{\frac{1}{12} + \frac{1}{15} + \frac{1}{11} + \frac{1}{16} + \frac{1}{14}} = 13,3(\text{хв.}) \end{aligned}$$

Таким чином, кожний токар на обточку 1 деталі витратив в середньому по 13,3 хвилини.

Для розв'язання задачі 4.13 ми скористалися середньою гармонійною простою.

2. Середня гармонійна:

а) **проста** – обернена середній арифметичній простій та розраховується для обернених значень ознаки.

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

де x – окремі варіанти (значення ознаки);

n - число спостережень (загальне число ознак або варіантів).

Середня гармонійна проста використовується, коли обсяги явища W , тобто добутку $x \cdot f$ однакові ($x_1 \cdot f_1 = x_2 \cdot f_2 = \dots = x_n \cdot f_n$).

Якщо ці добутки за кожною ознакою нерівні, використовують *середню гармонійну зважену*.

б) **зважена** – використовується, коли відомі значення індивідуальної ознаки x й обсяг явища $W = x \cdot f$, але значення частот невідомі:

$$\bar{x} = \frac{\sum W}{\sum \frac{W}{x}}$$

де W – обсяг явища, тобто добуток значення ознаки на частоту.

Задача 4.14. Витрати виробництва та собівартість одиниці продукції 3 заводів характеризуються наступними даними:

Номер заводу	Витрати виробництва, тис. грн.	Собівартість одиниці продукції, грн.
1	200	20
2	460	23
3	110	22

Визначити середню собівартість виробу.

Розв'язання.

Середня собівартість одиниці продукції = $\frac{\text{Витрати виробництва}}{\text{Кількість продукції}}$

Кількість продукції $f = \frac{W}{x}$, звідси:

$$\bar{x} = \frac{\sum W}{\sum \frac{W}{x}} = \frac{200000 + 460000 + 110000}{\frac{200000}{20} + \frac{460000}{23} + \frac{110000}{22}} = \frac{770000}{35000} = 22,0(\text{грн.})$$

Середня собівартість одиниці продукції по 3 заводах становить 22 грн.

3. Середня геометрична - відображає середній показник зростання показника в часі:

а) проста:

$$\bar{x} = \sqrt[n]{x_1 * x_2 * x_3 * \dots * x_n} = \sqrt[n]{\prod x_i}$$

де $\prod x_i$ – добуток відносних величин динаміки;

n – кількість осереднюваних ознак.

б) зважена:

$$\bar{x} = \sqrt[f]{x_1^{f_1} * x_2^{f_2} * x_3^{f_3} * \dots * x_n^{f_n}} = \sqrt[f]{\prod x_i^{f_i}}$$

Задача 4.15. Виробництво продукції сільського господарства в Україні за останні чотири роки зросла в 1,11 рази, у тому числі темп його зростання в 2009 р. становив 0,982, в 2010 р. – 0,985, в 2011 р. - 1,199, в 2012 - 0,955. Визначити середньорічний темп зростання виробництва продукції сільського господарства

Розв'язання.

$$\bar{x} = \sqrt[4]{0,982 * 0,985 * 1,199 * 0,955} = \sqrt[4]{1,11} = 1,026$$

Виробництво продукції сільського господарства в Україні протягом останніх чотирьох років зросло щорічно в середньому на 2,6%.

4. Середня квадратична – використовується для оцінки варіації (мінливості) ознак, а також для узагальнення ознак, виражених лінійними розмірами яких-небудь площ (середніх діаметрів кіл, стовбурів дерев, труб, сторін квадратів тощо)

$$\text{а) проста } \bar{x} = \sqrt{\frac{\sum x^2}{n}}, \quad \text{б) зважена } \bar{x} = \sqrt{\frac{\sum x^2 f}{\sum f}}$$

Задача 4.16. Діаметри 6 кошиків соняшнику становили 13, 13, 13, 16, 16 та 18 см. Визначити середній діаметр кошика.

Розв'язання.

Можна скористатися як формулою *середньої квадратичної простої*:

$$\bar{x} = \sqrt{\frac{\sum x^2}{n}} = \sqrt{\frac{13^2 + 13^2 + 13^2 + 16^2 + 16^2 + 18^2}{6}} = 15,0(\text{см}),$$

так і формулою *середньої квадратичної зваженої*:

$$\bar{x} = \sqrt{\frac{\sum x^2 f}{\sum f}} = \sqrt{\frac{13^2 * 3 + 16^2 * 2 + 18^2}{6}} = 15,0(\text{см}),$$

Середній діаметр кошика соняшнику становить 15 см.

5. Середня кубічна - застосовується для узагальнення ознак, виражених лінійними розмірами об'ємних фігур (середніх діаметрів плодів, насіння, яєць, тощо)

$$\text{а) проста } \bar{x} = \sqrt{\frac{\sum x^3}{n}}, \quad \text{б) зважена } \bar{x} = \sqrt{\frac{\sum x^3 f}{\sum f}}.$$

Задача 4.17. Діаметри плодів черешні сорту Мелітопольська чорна становлять 18, 19, 19, 20, 20, 20, 21 мм.

Визначити середній діаметр плодів.

Розв'язання.

Можна скористатися як формулою *середньої кубічної простої*:

$$\bar{x} = \sqrt[3]{\frac{18^3 + 19^3 + 19^3 + 20^3 + 20^3 + 20^3 + 21^3}{7}} = 19,6(\text{мм})$$

так і формулою *середньої кубічної зваженої*:

$$\bar{x} = \sqrt[3]{\frac{18^3 + 19^3 * 2 + 20^3 * 3 + 21^3}{7}} = 19,6(\text{мм})$$

Середній діаметр плодів черешні даного сорту 19,6 мм

6. Середня хронологічна - використовується для визначення середнього рівня моментного ряду динаміки з рівними проміжками часу між датами:

$$\bar{x} = \frac{\frac{1}{2} * x_1 + x_2 + \dots + x_{n-1} + \frac{1}{2} x_n}{n - 1},$$

де n – кількість дат.

Задача 4.18. У комерційному банку сума кредиторської заборгованості на кінець кожного кварталу становила: 1.01 - 20 млн. грн., 1.04. - 26 млн. грн., 1.07 - 32 млн. грн., 1.10 - 29 млн. грн., 1.01 наступного року - 22 млн. грн. Визначити середньоквартальну суму кредиторської заборгованості.

Розв'язання.

$$\bar{X} = \frac{\frac{1}{2} * 20 + 26 + 32 + 29 + \frac{1}{2} * 22}{5 - 1} = 27(\text{млн.грн.})$$

В середньому сума кредиторської заборгованості за квартал становить 27 млн. грн.

Структурні середні.

Мода – найбільш поширене значення ознаки.

Медіана – значення ознаки, що ділить ранжовану сукупність навпіл.

Детально мода та медіана розглянуті в главі 6.1 «Структурні характеристики розподілів» даного навчального посібника.

Досить часто в статистиці доводиться обчислювати, середню із групових середніх або часток, середню з відносних величин і середню в інтервальному ряді.

При обчисленні середньої із групових (або часток) середніх за індивідуальне значення ознаки береться відповідна групова середня.

Задача 4.19. За наступним даними визначите середньодобову продуктивність праці робітників підприємства

Цех	Добова продуктивність праці, грн./роб.
Цех №1	300
Іванов	300
Петров	200
Сидоров	400
Цех №2	270
Соколов	300
Воробйов	200
Орлов	250
Лебедєв	200
Сініцин	400

Розв'язання.

Знаходимо середню продуктивність праці в кожному з цехів:

$$\bar{x}_1 = \frac{300 + 200 + 400}{3} = 300(\text{грн./роб.})$$
$$\bar{x}_2 = \frac{300 + 200 + 250 + 200 + 400}{5} = 270(\text{грн./роб.})$$

Використовуючи дані про середньоцехову продуктивність праці та кількість робітників в кожному з цехів визначимо середньодобову продуктивність праці на підприємстві:

$$\bar{x} = \frac{\sum xf}{\sum f} = \frac{300 \cdot 3 + 270 \cdot 5}{3 + 5} = 281,25(\text{грн./роб.})$$

Середньодобова продуктивність праці на підприємстві становить 281,25 грн./роб.

При обчисленні середньої з *відносних величин* як вагу (або частоту) беруть основу відносної величини, а як значення ознаки - відповідну цій основі відносну величину.

Задача 4.20. Маємо наступні дані про урожайність та посівні площі пшениці в господарстві:

Номер бригади	Урожайність, ц/га	Посівна площа, га
1	22,0	200
2	24,0	280
3	26,0	350
4	23,0	170

Визначити середню урожайність пшениці по господарству.

Розв'язання.

$$\text{Урожайність} = \frac{\text{Валовий збір}}{\text{Площа посіву}}$$

Основою показника урожайності є площа посіву, яку і приймаємо за частоту f . Сама урожайність є осереднюваною ознакою x .

Таким чином, середня урожайність дорівнює:

$$x = \frac{22 * 200 + 24 * 280 + 26 * 350 + 23 * 170}{200 + 280 + 350 + 170} = 24,13(\text{ц} / \text{га})$$

Середня урожайність пшениці в господарстві становить 24,13 ц/га.

При обчисленні середньої для *інтервального ряду* необхідно знайти середину кожного інтервалу. Якщо інтервали відкриті, то, за правилами, приймаємо величину першого інтервалу рівною другому, а останнього передостанньому.

Задача 4.21. За наступним даними необхідно визначити середній стаж робітників заводу:

Загальний стаж роботи, років	до 5	5-10	10-20	20-30	40 і більше	Разом
Кількість робітників	75	68	25	22	10	200

Розв'язання.

Ознакою в даному завданні є загальний стаж робітника, а частотами відповідно кількість робітників, що мають той або інший стаж.

Ряд розподілу - інтервальный, з відкритими першим й останнім інтервалами.

За дискретне значення ознаки в кожній групі беремо середини інтервалу. Для цього ділимо величину інтервалу навпіл і одержану половину додаємо до нижньої межі відповідного інтервалу.

Для першої групи робітників величина інтервалу дорівнюватиме величині інтервалу другої групи – 5, а для останньої - величині інтервалу передостанньої групи – 10.

$$\bar{x} = \frac{\sum xf}{\sum f} = \frac{2,5 * 75 + 7,5 * 68 + 15 * 225 + 25 * 22 + 35 * 10}{200} = 9,9(\text{років})$$

Середній стаж робітника заводу становить 9,9 років.

При визначенні середніх величин необхідно дотримуватись двох *головних вимог*: сукупність повинна бути якісно однорідною, а кількість одиниць у сукупності - достатньо великою.

Завдання для самостійного розв'язання.

Завдання 4.1. Класифікуйте наведені статистичні показники як абсолютні чи відносні, інтервальні чи моментні. Для відносних величин зазначте вид та одиниці виміру:

- А) виробництво електроенергії (кВт-год.) на душу населення за рік;
- Б) довжина електрифікованих ліній залізниць на кінець року, км;
- В) кількість лікарів на 10 000 населення на початок року;
- Г) введення в дію загальної площі житлових будинків за рік, тис. кв. м;
- Д) співвідношення роздрібного товарообороту продовольчих та промислових товарів за місяць;
- Е) індекс роздрібних цін на споживчі товари за перший квартал, %;
- Є) національний дохід в розрахунку на 1 жителя країни;
- Ж) частка інвестицій на охорону навколишнього середовища у загальному обсязі капіталовкладень за останнє десятиріччя;
- З) зростання витрат на заходи з охорони праці у плановому році порівняно з поточним, %;
- И) кількість зареєстрованих за рік шлюбів на 1000 осіб населення;
- І) вартість основних засобів на 100 га сільськогосподарських угідь на початок року, тис. грн.

Завдання 4.2. За планом на поточний рік підприємство намітило випуск продукції на 350 млн. грн. при чисельності працівників 1300 осіб. Фактично обсяг продукції склав 320 млн. грн. при кількості працівників 1000 осіб. Визначити показники виконання плану підприємством за обсягом продукції, чисельністю працівників і продуктивністю праці. Проаналізувати отримані результати та зробити висновки.

Завдання 4.3. У Китаї в 2012 році рівень валового внутрішнього продукту сягнув 12380 млрд. дол., а чисельність населення склала 1,34 млрд. осіб. ВВП Японії в цьому ж році становив 4617 млрд. дол., а чисельність населення 127,4 млн. осіб. Визначте середньодушові ВВП обох держав та порівняйте одержані показники. Визначте вид обчислених відносних величин.

Завдання 4.4. Господарство має 350 корів. За рік вироблено 15000 ц молока. Витрати кормів наведені у таблиці:

Види кормів	Витрачено кормів, ц	Коефіцієнт переводу в кормові одиниці
Грубі	9600	0,43
Соковиті	43000	0,12
Концентровані	1770	1,00
Зелені	42800	0,18

- Визначити: 1) загальний обсяг і структуру витрачених кормів;
 2) витрати кормів (у кормових одиницях) на 1 голову і на 1 ц молока;
 3) продуктивність корів.

Класифікувати обчислені показники. Зробити висновки.

Завдання 4.5. Маємо наступні дані про поголів'я сільськогосподарських тварин та розмір земельних угідь по господарству:

Види тварин та угідь	Наявність у господарстві	Коефіцієнт переводу в умовні голови
Велика рогата худоба - основне стадо, гол.	400	1,0
Молодняк ВРХ, гол.	250	0,5
Свині, гол.	300	0,3
Віці та кози, гол.	100	0,3
Птиця, гол.	2000	0,01
Площа сільськогосподарських угідь, га	5000	х
Площа ріллі, га	4000	х
Площа посіву зернових, га	3000	х

Визначити: 1) умовне поголів'я тварин у господарстві;

2) поголів'я тварин в розрахунку на 100 га земельної площі, в тому числі по видах тварин.

Класифікувати обчислені показники.

Примітка. Умовне поголів'я, поголів'я ВРХ, овець та кіз розраховується на одиницю сільськогосподарських угідь; поголів'я свиней – на одиницю ріллі, поголів'я птиці - на одиницю посівної площі зернових культур.

Завдання 4.6. Маємо наступні дані по переписах населення України:

Рік перепису	Чисельність населення, млн. осіб		
	жіночого	чоловічого	Разом
1959	23,299	18,570	41,869
1970	25,823	21,304	47,127
1979	27,070	22,685	49,755
1989	27,802	23,905	51,707
2001	26,016	22,441	48,457

Визначити: 1) питому вагу чоловіків і жінок по кожному перепису;

2) кількість жінок, що припадає на 100 чоловіків, по кожному перепису;

3) зміну чисельності населення України від перепису до перепису, а також у порівнянні з показниками 1959 року.

Визначити вид обчислених відносних величин. Зробити висновки

Завдання 4.7. П'ять підприємств виробляють мікрохвильові печі. Середня ціна реалізації на першому підприємстві склала 60 доларів, на другому - 70 доларів, на третьому – 50 доларів, на четвертому - 65 доларів, на п'ятому – 80 доларів.

Визначити середню ціну мікрохвильової печі, якщо відомо, що на кожному з підприємств виручка від їх реалізації становила 900 тис. доларів.

Завдання 4.8. Маємо наступні результати вимірювання діаметрів страусових яєць по фермі:

<i>№ яйця</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>6</i>	<i>7</i>	<i>8</i>	<i>9</i>	<i>10</i>
Діаметр, см	14	15	15	16	15	16	15	15	15	14
<i>№ яйця</i>	<i>11</i>	<i>12</i>	<i>13</i>	<i>14</i>	<i>15</i>	<i>16</i>	<i>17</i>	<i>18</i>	<i>19</i>	<i>20</i>
Діаметр, см	13	16	16	13	15	16	14	15	15	16

Визначити середній по фермі діаметр страусових яєць.

Завдання 4.9. На одному з ринків міста зафіксовані такі ціни на сільськогосподарські продукти, грн.:

Продукт	Прода- вець 1	Прода- вець 2	Прода- вець 3	Прода- вець 4	Прода- вець 5
Яловичина (1 кг)	53,5	52	51,5	52,5	54,0
Масло вершкове (1 кг)	49,5	51	50	20,5	20
Молоко (1 л)	8	8,5	7,5	7	7,5
Яйця (10 шт.)	9,5	10,5	10	9	10

Визначити максимальну, мінімальну та середню ціну по кожному виду продуктів на ринку.

Завдання 4.10. Маємо наступні дані про розподіл магазинів міста за розміром річного товарообороту:

Товарооборот, тис. грн.	Число магазинів
До 100	25
100 - 250	40
250 - 400	55
400 - 600	30
600 - 1000	15
Більше 1000	5

Визначити середній товарооборот в розрахунку на один магазин.

Завдання 4.11. Маємо дані про вартість основних засобів на підприємстві:

Період	на 1.01.2009	на 1.01.2010	на 1.01.2011	на 1.01.2012	на 1.01.2013
Вартість основних засобів, тис. грн.	10200	12300	11555	11900	11000

Визначити середньорічну вартість основних засобів.

Завдання 4.12. Маємо наступні дані про витрати сировини на виробництво продукції цехами заводу:

Номер цеху	Витрати матеріалу, кг	
	на 1 кг продукції	на всю продукцію
1	1,2	360
2	1,4	280
3	1,8	180
4	1,5	250
5	1,1	330
6	1,3	260
7	1,4	420
8	2,0	300

Визначити витрати сировини на 1 кг продукції в середньому по заводу.

Завдання 4.13. Маємо наступні дані про валовий збір і урожайність зернових культур у господарствах району:

Господарства	Урожайність, ц/га	Валовий збір, ц
ТОВ "Нива"	24,0	4800
ПСП "Надія"	20,0	6000
ТОВ "Промінь"	20,0	5000
СБК "Україна"	24,5	4500
ПСП "Ранок"	23,0	4600
Разом	x	

Визначити середню урожайність зернових культур та середній валовий збір.

Тести для самоконтролю.

1. Кількісно та якісно виражена характеристика певної властивості сукупності в цілому чи її частин в умовах конкретного місця і часу:

- А) статистичний показник
- Б) абсолютна величина
- В) ознака
- Г) відносна величина

2. Одиниці виміру абсолютних показників, що відображають результати праці, називаються:

- А) натуральні
- Б) умовно-натуральні
- В) трудові
- Г) вартісні

3. Показник з яким здійснюється порівняння, називається:

- А) відносний показник
- Б) база порівняння
- В) порівнюваний показник
- Г) середня величина

4. Співвідношення планової величини показника до його фактично досягнутого рівня в попередньому періоді або в періоді називають:

- А) відносна величина виконання плану
- Б) величина планового завдання
- В) відносна величина порівняння
- Г) відносна величина динаміки

5. Частка сільськогосподарської продукції у загальному обсязі виробленої в регіоні продукції є відносною величиною:

- А) інтенсивності
- Б) динаміки
- В) порівняння
- Г) структури

6. Співвідношення обсягів виробництва сільськогосподарської продукції підприємствами Мелітопольського та Якимівського районів у 2012 році є відносною величиною:

- А) порівняння
- Б) динаміки
- В) структури
- Г) інтенсивності

7. Одиниці виміру абсолютних показників, що відповідають природним або споживчим властивостям предмета вивчення, називаються:

- А) натуральні
- Б) умовно-натуральні

В) споживчі

Г) вартісні

8. Кількісне співвідношення між показниками, що утворюється у результаті ділення однієї величини на іншу, називається:

А) відносна величина

Б) середня величина

В) абсолютна величина

Г) порівняна величина

9. При порівнянні різнойменних показників одиницею виміру відносних показників є:

А) коефіцієнт

Б) відсоток

В) іменована величина

Г) проміле

10. Співвідношення величини показника, досягнутої за якийсь час, і величини його, установлені за планом за цей же час, називають:

А) відносна величина виконання плану

Б) відносна величина планового завдання

В) відносна величина порівняння

Г) відносна величина динаміки

11. Співвідношення величини показника за даний час і величини його за який-небудь попередній час, прийнятий за базу порівняння, називають:

А) відносна величина динаміки

Б) відносна величина порівняння

В) відносна величина планового завдання

Г) відносна величина виконання плану

12. Одиниці виміру абсолютних показників, що одержують, приводячи різні натуральні одиниці до однієї, прийнятої за базу, називаються:

А) натуральні

Б) умовно-натуральні

В) трудові

Г) вартісні

13. При порівнянні однойменних показників одиницею виміру відносних показників не може бути:

А) іменована величина

Б) проміле

В) відсоток

Г) коефіцієнт

14. Показник, що порівнюється з іншим, взятим за основу, називається:

А) відносний показник

- Б) база порівняння
- В) порівнюваний показник
- Г) середня величина

15. Відношення значення кожної частини сукупності до загального підсумку, який прийнято за базу, називають:

- А) відносна величина структури
- Б) відносна величина динаміки
- В) відносна величина порівняння
- Г) відносна величина координації

16. Ріст інвестицій у нафтодобувну промисловість за рік є відносною величиною:

- А) структури
- Б) координації
- В) динаміки
- Г) порівняння

17. Забезпеченість населення житлом, що становить 19 кв. м на особу, є відносною величиною:

- А) інтенсивності
- Б) порівняння
- В) структури
- Г) динаміки

18. Частка соняшнику у загальній виручці господарства є відносною величиною:

- А) порівняння
- Б) структури
- В) координації
- Г) рівня економічного розвитку

19. Кількість дівчат у класі, що приходиться на одного хлопця є відносною величиною:

- А) порівняння
- Б) інтенсивності
- В) координації
- Г) структури

20. Одиниці виміру абсолютних показників, що передбачають грошове вираження обсягу показника, називаються:

- А) натуральні
- Б) умовно-натуральні
- В) трудові
- Г) вартісні

21. Співвідношення окремих частин цілого, одна з яких приймається за базу порівняння, називають:

- А) відносна величина координації

- Б) відносна величина структури
- В) відносна величина інтенсивності
- Г) відносна величина порівняння

22. Питома вага ріллі у загальній земельній площі господарства є відносною величиною:

- А) інтенсивності
- Б) порівняння
- В) структури
- Г) координації

23. Щільність населення України є відносною величиною:

- А) координації
- Б) порівняння
- В) рівня економічного розвитку
- Г) інтенсивності

24. Середньодушовий національний дохід є відносною величиною:

- А) порівняння
- Б) координації
- В) структури
- Г) рівня економічного розвитку

25. Співвідношення обсягів реалізації продукції супермаркетами «Амстор» і «Сільпо» у 2012 році є відносною величиною:

- А) порівняння
- Б) структури
- В) координації
- Г) інтенсивності

26. Відношення запланованих на майбутній рік виробничих витрат до фактичних витрат це:

- А) відносна величина виконання плану
- Б) величина планового завдання
- В) відносна величина порівняння
- Г) відносна величина динаміки

27. Середня величина, що використовується для визначення середнього рівня моментного ряду динаміки з рівними проміжками часу між датами, називається:

- А) арифметична
- Б) гармонійна
- В) хронологічна
- Г) геометрична

28. Відношення фактичних виробничих витрат 2012 року до витрат, запланованих на цей рік, це:

- А) відносна величина виконання плану
- Б) величина планового завдання

В) відносна величина порівняння

Г) відносна величина динаміки

29. Середня величина, що використовується, коли відомі значення індивідуальної ознаки й обсяг явища, але значення частот невідомі, називається:

А) гармонійна зважена

Б) арифметична зважена

В) геометрична зважена

Г) квадратична зважена

30. Прибуток, що приходиться на 100 га сільськогосподарських угідь, є відносною величиною:

А) інтенсивності

Б) порівняння

В) структури

Г) динаміки

31. Середня величина, що використовується для узагальнення ознак, виражених лінійними розмірами яких-небудь площ, називається:

А) геометрична

Б) квадратична

В) арифметична

Г) гармонійна

32. Відсоток зменшення обсягів виробництва продукції підприємства у порівнянні з попереднім роком це:

А) відносна величина виконання плану

Б) величина планового завдання

В) відносна величина порівняння

Г) відносна величина динаміки

33. Відсоток позитивних результатів дослідів у загальній їх кількості є відносною величиною:

А) координації

В) інтенсивності

Б) структури

Г) порівняння

34. Середня величина, що застосовується для узагальнення ознак, виражених лінійними розмірами об'ємних фігур, називається:

А) кубічна

Б) квадратична

В) геометрична

Г) арифметична

35. Співвідношення негативних та позитивних результатів дослідів є відносною величиною:

А) порівняння

- Б) структури
- В) координації
- Г) інтенсивності

36. Співвідношення населення Донецької та Волинської областей за останній рік це відносна величина:

- А) структури
- Б) динаміки
- В) інтенсивності
- Г) порівняння

37. Середня величина, що відображає середній показник росту показника в часі, називається:

- А) гармонійна
- Б) геометрична
- В) хронологічна
- Г) арифметична

38. Зростання фактичних обсягів виробництва продукції підприємства у порівнянні з планом є відносною величиною:

- А) виконання плану
- Б) планового завдання
- В) порівняння
- Г) динаміки

39. Обсяг виробництва видобувної промисловості в розрахунку на жителя Донецької області є відносною величиною:

- А) координації
- Б) динаміки
- В) рівня економічного розвитку
- Г) порівняння

40. Середня величина, обернена середній арифметичній простій та розраховується для обернених значень ознаки, називається:

- А) гармонійна проста
- Б) кубічна проста
- В) квадратична проста
- Г) арифметична зважена

41. Кількість зернових комбайнів на 100 га площі зернових це:

- А) відносна величина координації
- Б) відносна величина динаміки
- В) відносна величина інтенсивності
- Г) відносна величина порівняння

42. Середня величина, що використовується, коли значення ознаки зустрічається кілька разів, називається:

- А) арифметична зважена
- Б) гармонійна зважена

В) квадратична проста

Г) арифметична проста

43. Зменшення рівня інфляції в плановому році у порівнянні з поточним роком є відносною величиною:

А) виконання плану

Б) планового завдання

В) порівняння

Г) динаміки

44. Середня величина, що використовується для осереднення прямих значень ознак шляхом їх підсумовування, називається:

А) арифметична проста

Б) арифметична зважена

В) гармонійна зважена

Г) гармонійна проста

45. Виробництво зерна в Україні в розрахунку на 1 жителя є відносною величиною:

А) порівняння

Б) структури

В) координації

Г) рівня економічного розвитку

5. АНАЛІЗ РЯДІВ РОЗПОДІЛУ

5.1 Статистичні характеристики розподілів

Сутність та види рядів розподілу вже розглядалися нами в главі 3.2. «Попередні відомості про ряди розподілу».

Ряд розподілу характеризує склад, структуру сукупності за певною ознакою. Його будують, щоб мати можливість судити про однорідність сукупності, межі її зміни, виявити характер розподілу одиниць сукупності за варіюючою ознакою, визначити закономірності в цьому розподілі. Елементами ряду розподілу є варіанти та частоти, у співвідношенні яких і проявляється закономірність розподілу.

Для аналізу ряду розподілу використовують ряд *статистичних характеристик*:

- частотні характеристики рядів розподілу;
- показники варіації;
- структурні характеристики розподілу;
- характеристики диференціації, концентрації та подібності розподілів.

Всі вищевказані характеристики будуть розглянуті у даному та наступному розділах навчального посібника, а набір частотних характеристик розглянемо просто зараз.

Частотні характеристики рядів розподілу:

1) **частота** f_i - абсолютна чисельність i -ї групи, тобто $\sum f_i = n$, де n - обсяг сукупності;

2) **частка** d_i – відносна частота, або питома вага абсолютних частот в обсязі сукупності, виражена у долях одиниці або відсотках, тобто $\sum d_i = 1$ або $\sum d_i = 100\%$;

3) **кумулятивна (накопичена) частота** S_{f_i} – утворюється послідовним підсумовуванням частот ($S_1=f_1$; $S_2=f_1+f_2$; $S_n=f_1+f_2+\dots+f_n$);

4) **кумулятивна (накопичена) частка** S_{d_i} – утворюється аналогічно послідовним підсумовуванням часток;

5) **щільність розподілу** – число одиниць сукупності, що припадає в середньому на одиницю величини інтервалу - використовується переважно для характеристики ряду розподілу з *нерівними інтервалами*, але при цьому є універсальною частотною характеристикою:

а) **абсолютна:** $g_{f_i} = \frac{f_i}{h_i}$,

б) **відносна:** $g_{d_i} = \frac{d_i}{h_i}$

де h_i – величина i -го інтервалу.

Задача 5.1. Маємо наступні дані про кількість філій кожного з двадцяти банків у місті М: 2, 4, 3, 5, 4, 4, 6, 5, 4, 3, 4, 3, 4, 5, 3, 4, 6, 3, 5, 4.

За наявними даними побудувати ряд розподілу та розрахувати частотні характеристики.

Розв'язання.

Дискретний ряд розподілу, побудований за наявними даними, та необхідні частотні характеристики представлені в табл. 5.1.

Таблиця 5.1

Розрахунок частотних характеристик розподілу банків за кількістю філій в місті М

Кількість філій банку у місті	Число банків (частота) f	Питома вага банків, % (частка) d	Кумулятивна (накопичена) частота S_f	Кумулятивна (накопичена) частка S_d
2	1	$1*100/20=5$	1	5
3	5	$5*100/20=25$	$1+5 = 6$	$5+25 = 30$
4	8	$8*100/20=40$	$6+8 = 14$	$30+40= 70$
5	4	$4*100/20=20$	$14+4 = 18$	$70+20 = 90$
6	2	$2*100/20=10$	$18+2 = 20$	$90+10 = 100$
Разом	20	100	x	x

Задача 5.2. На основі згрупованих даних про розподіл сільських магазинів району за кількістю секцій розрахувати відповідні частотні характеристики.

Розв'язання.

Інтервальний ряд розподілу з *нерівними інтервалами*, побудований за наявними даними, та необхідні частотні характеристики представлені в табл. 5.2.

Таблиця 5.2

Розрахунок частотних характеристик сільських магазинів району за числом секцій

Групи магазинів за числом секцій	Величина інтервалу h	Кількість магазинів (частота) f	Частка магазинів, % d	Абсолютна щільність g_f	Відносна щільність g_f
2-4	2	20	28,6	$20/2=10$	$28,6/2=14,3$
4-5	1	35	50,0	$35/1=35$	$50/1=50,0$
5-8	3	15	21,4	$15/3=5$	$21,4/3=7,1$

5.2 Графічне зображення варіаційних рядів розподілу

Зручніше за все ряди розподілу аналізувати за допомогою їх графічного зображення, що дозволяє судити про форму розподілу. Способи побудови графіків різні для інтервальних і дискретних рядів.

Наочне уявлення про характер зміни частот варіаційного ряду дають полігон та гістограма.

1) **Полігон** – використовується для зображення дискретних варіаційних рядів.

Спосіб побудови. На прямокутній системі координат по вісі абсцис (OX) в однаковому масштабі відкладають ранжовані значення варіюючої ознаки, а по вісі ординат наноситься шкала для зображення величини частот або часток. Одержані на перетині точки поєднуються прямими лініями, в результаті чого одержують ломану лінію (рис.5.1а), що називається *полігоном частот*.

Іноді для замикання полігону пропонується крайні точки (ліворуч та праворуч на ломаній лінії) з'єднати з точками на вісі абсцис. В цьому випадку виходить багатокутник (рис.5.1б).

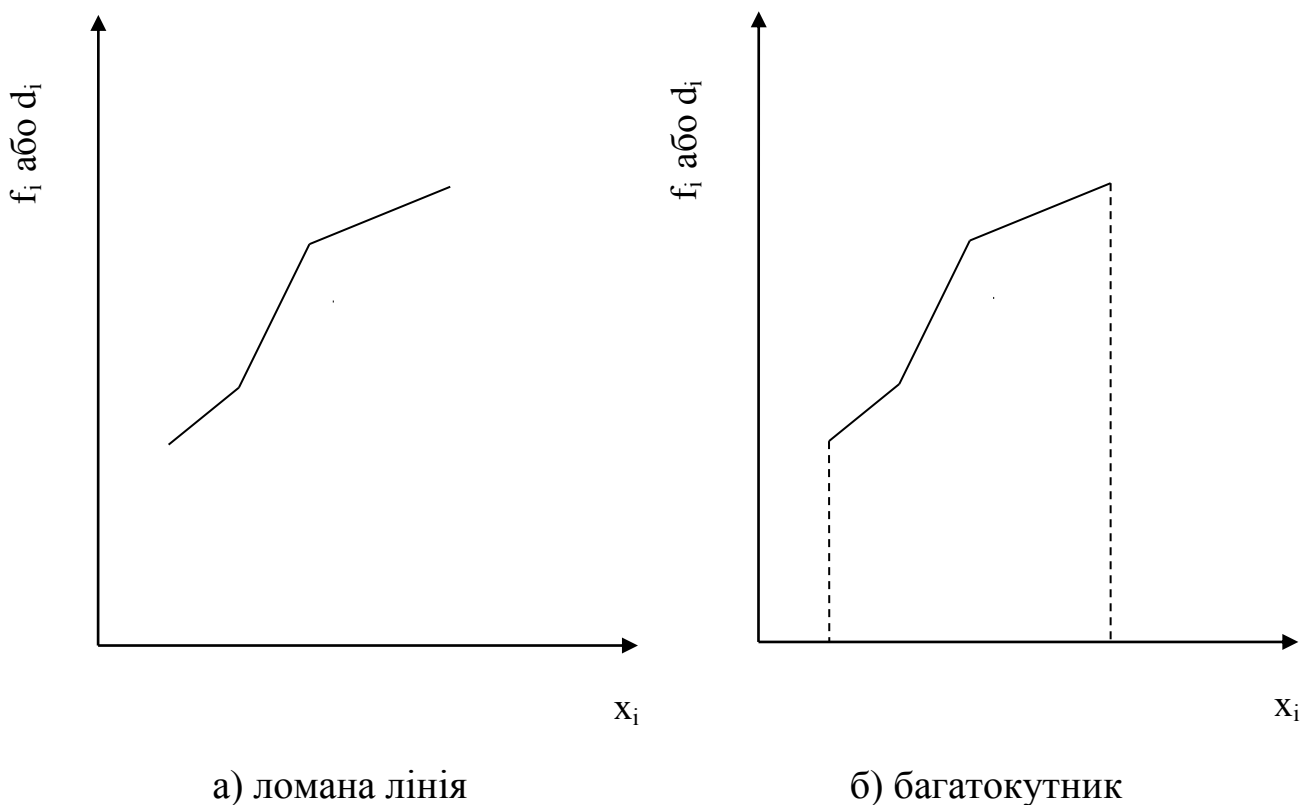


Рис.5.1 Загальний вигляд полігону

Приклад. На рис. 5.2 та 5.3 зображено розподіл житлового фонду одного з районів міста за числом кімнат у квартирах, представлений в табл. 5.1.

Таблиця 5.3

Розподіл житлового фонду міського району за типом квартир

Групи квартир за числом кімнат	Кількість квартир	Питома вага відповідних квартир у їх загальній кількості, %
1	150	19,4
2	300	38,7
3	250	32,3
4	50	6,5
5	25	3,2
Разом	775	100

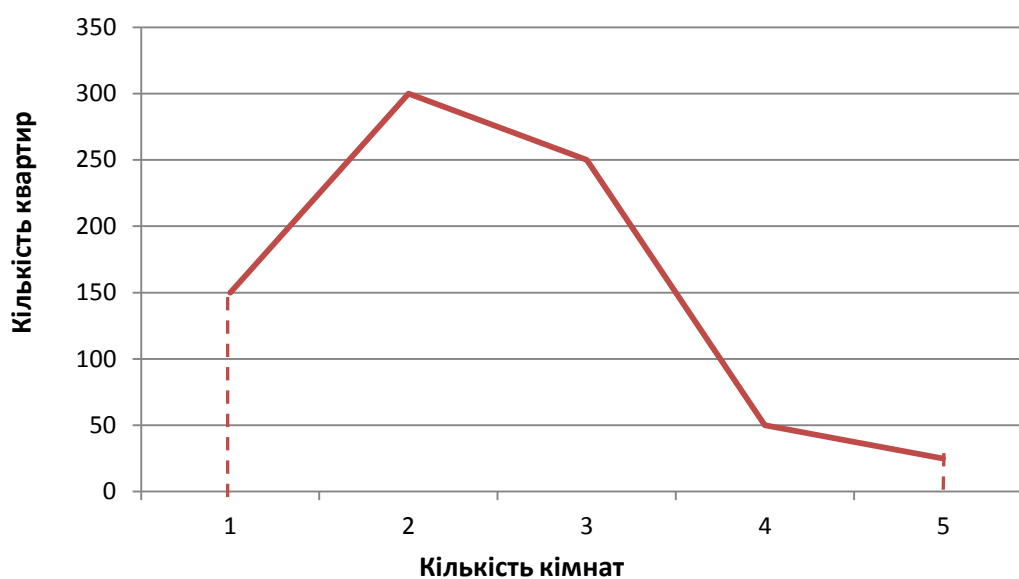


Рис. 5.2 Полігон розподілу житлового фонду міського району за типом квартир (частотна характеристика - частота)

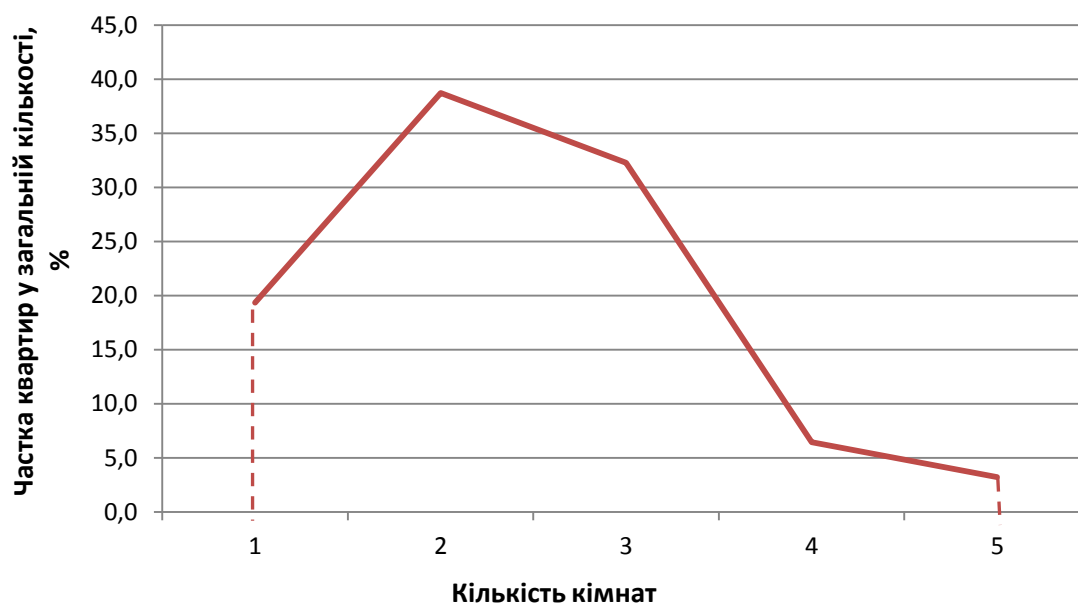
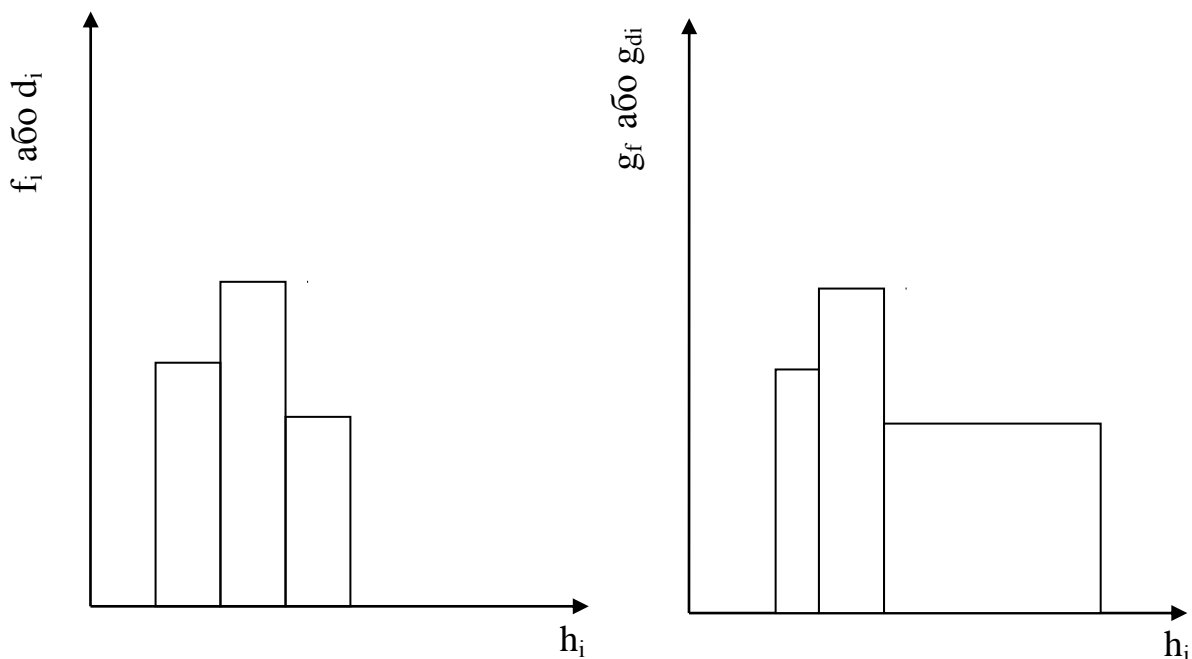


Рис. 5.3 Полігон розподілу житлового фонду міського району за типом квартир (частотна характеристика - частка)

Як бачимо, незалежно від того, яку частотну характеристику - частоту (рис. 5.2) або частку (рис. 5.3) – ми використовуємо для зображення полігону розподілу сукупності, форма розподілу залишається незмінною.

2) **Гістограма** – застосовується для зображення інтервального варіаційного ряду.

Спосіб побудови. На вісі абсцис (ОХ) в однаковому масштабі відкладають величини інтервалів, на вісі ординат (ОУ): при рівних інтервалах - частоти або частки (рис. 5.4а), а при нерівних інтервалах - абсолютну або відносну щільність розподілу (рис. 5.4б), які зображуються прямокутниками, побудованими на відповідних інтервалах. В результаті одержуємо графік, на якому ряд розподілу зображений у вигляді суміжних один з одним стовпчиків.



а) для рівних інтервалів

б) для нерівних інтервалів

Рис.5.4 Загальний вигляд гістограми

Приклад. На рис. 5.5 зображено розподіл родин міського району за розміром житлової площі, що приходить на 1 особу, представлений в табл. 5.1.

Примітка. Гістограма може бути перетворена у полігон, якщо з'єднати середини вершин прямокутників. Однак такий полігон неможна вважати цілком коректним, так як перша половина одиниць, які входять до першого інтервалу, та друга половина одиниць, які входять до останнього інтервалу, залишаться за межами цього полігону, тобто не увійде до графіку розподілу.

Розподіл родин за розміром житлової площі,
що приходить на 1 особу

Групи за розміром житлової площі, що приходить на 1 особу, м ²	Кількість родин з даним розміром житлової площі
До 6	100
6-8	200
8-10	400
10-12	300
Від 12	150
Разом	1150

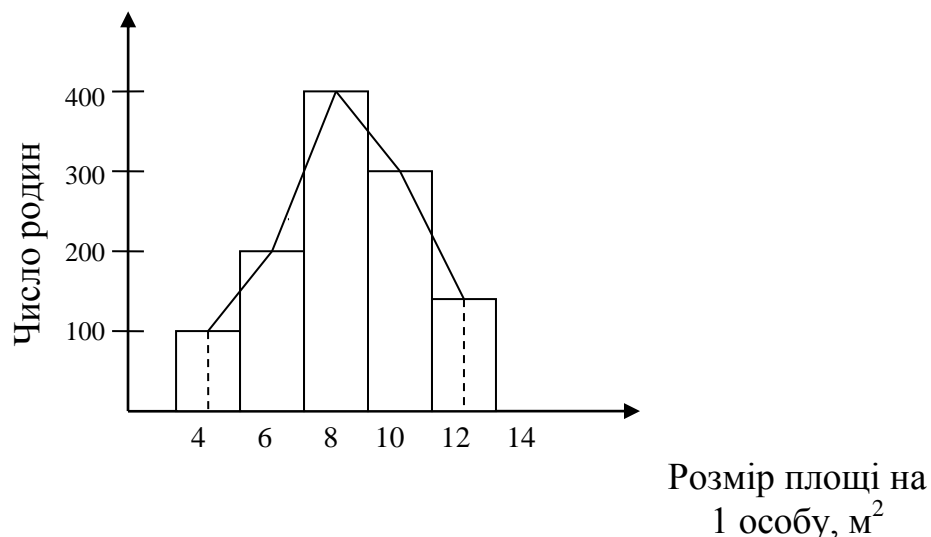


Рис. 5.5 Гістограма розподілу родин за розміром житлової площі,
що приходить на 1 особу

3) **кумулятивні криві** – криві, побудовані з використанням накопичених частот:

а) **кумулята**

Спосіб побудови. На вісі абсцис (ОХ) відкладають варіанти дискретного або інтервали інтервального ряду розподілу, а на вісі ординат – кумулятивні (накопичені) частоти (частки). Знаходяться точки перетину перпендикулярів відповідних варіантів та накопичених частот (для дискретного ряду) або верхніх меж інтервалів та накопичених частот (для інтервального ряду). Точки послідовно з'єднуються, утворюючи ломану лінію (рис. 5.6а).

б) **огіва.**

Спосіб побудови. Змінюємо містами вісі кумуляти (рис. 5.6б).

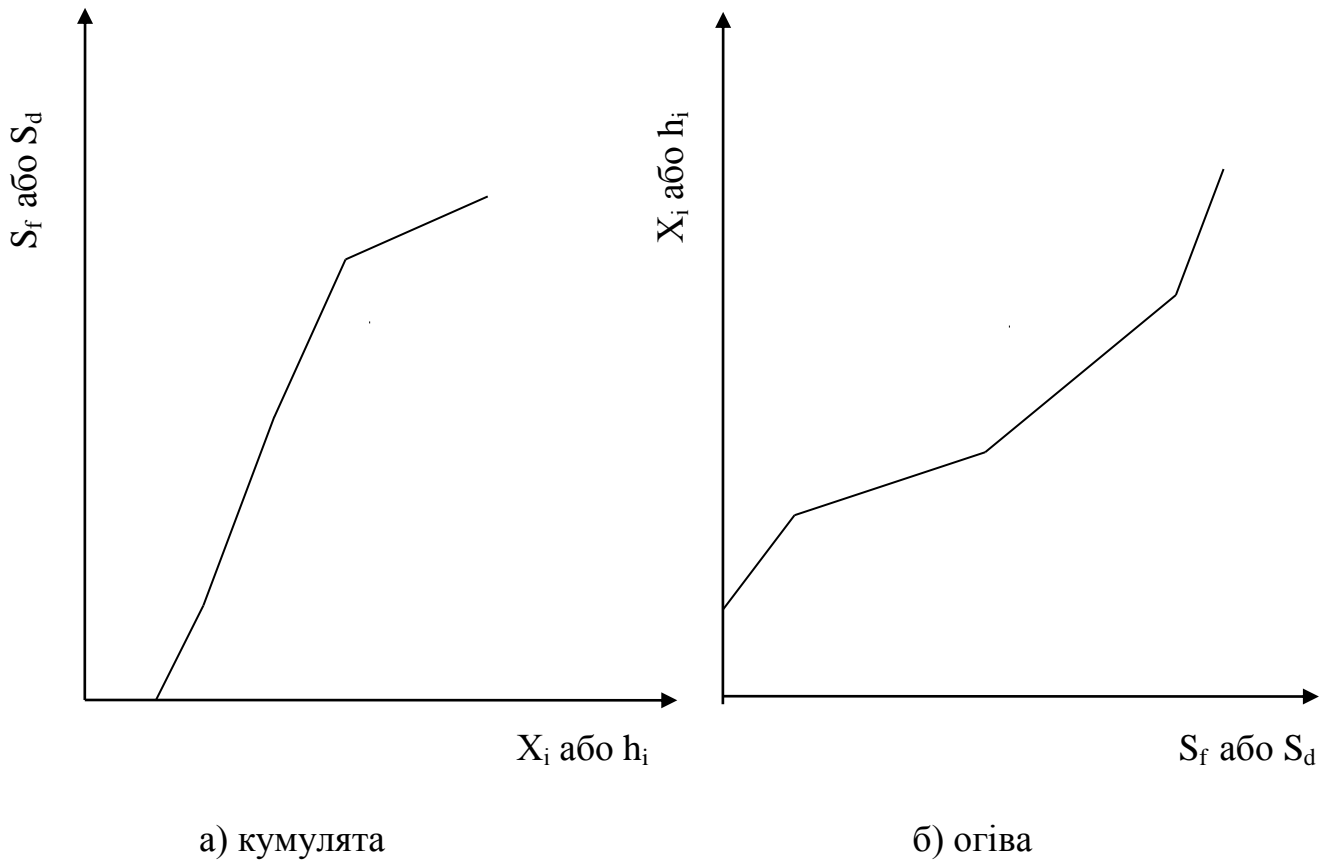


Рис. 5.6 Загальний вид кумулятивних кривих

Приклад. Скористуємося даними табл. 5.4 та розрахуємо накопичені частоти для побудови кумуляти (табл. 5.5), яку зображено на рис. 5.7.

Таблиця 5.5

Вихідні та розрахункові дані для побудови кумуляти

Групи за розміром житлової площі, що приходить на 1 особу, м ²	Кількість родин з даним розміром житлової площі	Накопичена кількість родин
До 6	100	100
6-8	200	300
8-10	400	700
10-12	300	1000
Від 12	150	1150
Разом	1150	-

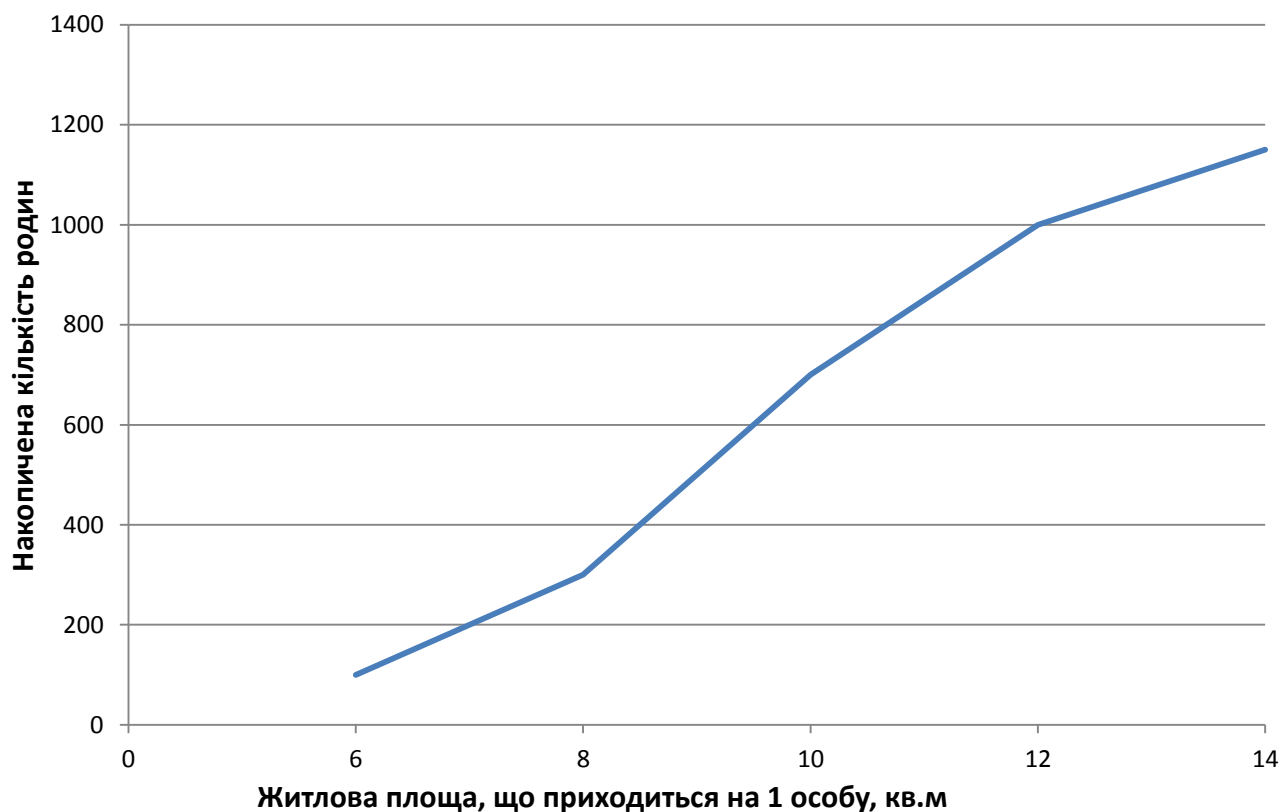


Рис. 5.7. Кумулята розподілу родин за розміром житлової площі, що приходить на 1 особу

Приклад. Скористуємося даними, представленими у табл. 5.3 та розрахуємо накопичені частоти для побудови огіви (табл. 5.6), яку зображено на рис. 5.8.

Таблиця 5.6

Вихідні та розрахункові дані для побудови огіви

Групи квартир за числом кімнат	Кількість квартир	Накопичена кількість квартир
1	150	150
2	300	450
3	250	700
4	50	750
5	25	775
Разом	775	-

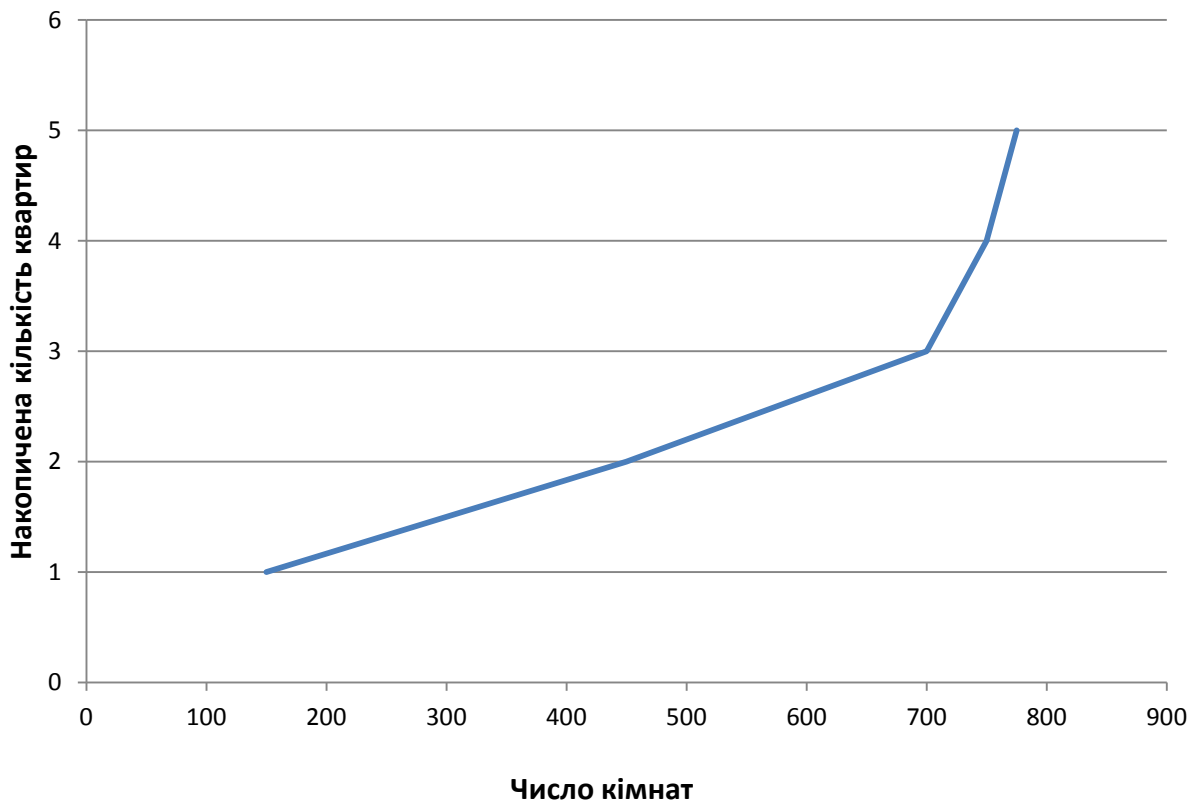


Рис. 5.8 Огіва розподілу житлового фонду міського району за типом квартир

5.3 Показники варіації рядів розподілу

Термін «варіація» відбувся від латинського *variatio* – «зміна, коливання, розходження». Однак не всякі розходження прийнято називати варіацією.

Під **варіацією** в статистиці розуміють такі кількісні зміни величини досліджуваної ознаки в межах однорідної сукупності, які обумовлені перехресним впливом дії різних факторів.

Різниця індивідуальних значень ознаки усередині досліджуваної сукупності, тобто варіація ознаки, виникає в результаті того, що індивідуальні значення ознаки складаються під сукупним впливом різноманітних факторів, які по-різному сполучаються в кожному окремому випадку.

Середня величина - це абстрактна, узагальнююча характеристика ознаки досліджуваної сукупності, але вона не показує складу сукупності, що є досить суттєвим для її пізнання. Середня не дає уявлення про те, як окремі значення досліджуваної ознаки групуються навколо середньої, чи зосереджені вони поблизу або значно відхиляються від неї.

У деяких випадках окремі значення ознаки близько примикають до середньої арифметичної й мало від неї відрізняються. У таких випадках середня добре представляє всю сукупність. В інших, навпаки, окремі

значення сукупності далеко відстають від середньої, і середня погано представляє всю сукупність.

Розрізняють варіацію ознаки: випадкову й систематичну. Аналіз систематичної варіації дозволяє оцінити ступінь залежності змін у досліджуваній ознаці від визначальних її факторів. Наприклад, вивчаючи силу й характер варіації у сукупності, можна оцінити, наскільки однорідною є дана сукупність у кількісному, а іноді і якісному відношенні, а отже, наскільки характерною є обчислена середня величина.

Ступінь близькості даних окремих одиниць x_i до середнього вимірюється поруч абсолютних, середніх і відносних показників. Для характеристики сукупностей й обчислених величин важливо знати, яка варіація досліджуваної ознаки приховується за середнім значенням.

Абсолютні показники варіації.

1) **Розмах варіації** - це різниця між найбільшим (x_{max}) і найменшим (x_{min}) значеннями варіантів:

$$R = x_{max} - x_{min}$$

Цей показник уловлює тільки крайні відхилення й не відображає відхилень всіх варіант у ряді, тому він є головним показником у тих випадках, коли варіанти зустрічаються один раз. В інших випадках розмах варіації застосовується для того, щоб одержати загальне уявлення про варіацію ознаки у всієї сукупності.

2) **Середнє лінійне відхилення** визначається як середня арифметична з відхилень індивідуальних значень від середньої, без урахування знака цих відхилень. Залежно від вихідних даних воно може обчислюватися за формулою середньої арифметичної простої або зваженої:

а) для незгрупованих даних:

$$\bar{d} = \frac{\sum |x - \bar{x}|}{n} = \frac{|x_1 - \bar{x}| + |x_2 - \bar{x}| + \dots + |x_n - \bar{x}|}{n}$$

б) для згрупованих даних:

$$\bar{d} = \frac{\sum |x_i - \bar{x}| f_i}{\sum f_i} = \frac{|x_1 - \bar{x}| f_1 + |x_2 - \bar{x}| f_2 + \dots + |x_n - \bar{x}| f_n}{f_1 + f_2 + \dots + f_n}$$

Середнє лінійне відхилення характеризує середній розмір коливань значень ознаки навколо середнього рівня.

Недоліком даного показника є те, що при його обчисленні не враховують знаки відхилень, тобто їх спрямованість. Тому цей показник варіації використовується рідко, зокрема, для аналізу складу працюючих, ритмічності виробництва, обертання коштів у зовнішній торгівлі, для характеристики виконання договірних зобов'язань по рівномірності

постачання, при оцінці якості продукції з урахуванням технологічних особливостей виробництва тощо).

3) **Дисперсія** - це середня арифметична квадратів відхилень кожного значення ознаки від загальної середньої. Дисперсія звичайно називається середнім квадратом відхилень. Залежно від вихідних даних дисперсія може обчислюватися за формулою середньої арифметичної простої або зваженої:

а) для незгрупованих даних:

$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2}{n} \text{ — дисперсія незважена (проста);}$$

б) для згрупованих даних:

$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2 f_i}{\sum f_i} \text{ — дисперсія зважена.}$$

Дисперсію важко інтерпретувати змістовно, але вона є основою для розрахунку наступного показника – стандартного відхилення, який вважається основним абсолютним показником варіації.

Також, треба зауважити, що дисперсію використовують не лише для оцінки варіації, а й для вимірювання зв'язків між досліджувальними факторами; розклад дисперсії на складові дозволяє оцінити вплив різних факторів, які обумовлюють варіацію ознаки.

4) **Середнє квадратичне (стандартне) відхилення** являє собою корінь квадратний з дисперсії. Залежно від вихідних даних стандартне відхилення може обчислюватися за формулою середньої квадратичної простої або зваженої:

а) для незгрупованих даних:

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n}} \text{ - стандартне відхилення незважене;}$$

б) для згрупованих даних:

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum (x_i - \bar{x})^2 f_i}{\sum f_i}} \text{ - стандартне відхилення зважене.}$$

Інтерпретація середнього квадратичного відхилення є такою самою, як і для лінійного відхилення: воно показує, на скільки в середньому відхиляються індивідуальні значення ознаки від їх середнього значення.

Перевага цього показника порівняно із середнім лінійним відхиленням полягає у відсутності умовного припущення з сумування відхилень без врахування їх знаків, бо відхилення використовуються у квадратній степені.

Крім зазначеного, перевагою даного показника у порівнянні з дисперсією є те, що стандартне відхилення виражається в тих же одиницях вимірювання, що і значення досліджуваної (у метрах, тоннах, відсотках, гектарах, грошових одиницях і т.д.).

Середнє квадратичне відхилення є мірилом надійності середньої. Чим менше стандартне відхилення, тим краще середня арифметична відображає собою всю представлену сукупність, тим більш однорідною є сукупність і більш типовою є середня величина

Крім вищевказаних для характеристики міри коливання досліджуваної ознаки обчислюються показники коливання у відносних величинах. Вони дозволяють порівнювати характер розсіювання в різних розподілах (різні одиниці спостереження того самого ознаки у двох сукупностях, при різних значеннях середніх, при порівнянні різнойменних сукупностей).

Розрахунок показників міри відносного розсіювання здійснюють як відношення абсолютного показника розсіювання до середньої величини.

Відносні показники варіації:

1. **Коефіцієнт осциляції** є відсотковим співвідношенням розмаху варіації та середньої величини:

$$V_R = \frac{R}{\bar{x}} * 100\%$$

Коефіцієнт осциляції відображає відносні коливання крайніх значень ознаки навколо середньої.

2. **Лінійний коефіцієнт варіації** розраховується як відношення середнього лінійного відхилення до середньої, вираженим у відсотках:

$$V_{\bar{d}} = \frac{\bar{d}}{\bar{x}} * 100\%$$

Лінійний коефіцієнт варіації характеризує частку усередненого значення абсолютних відхилень від середньої величини.

3. **Коефіцієнт варіації** визначається співвідношенням стандартного відхилення та середньої величини:

$$V_{\sigma} = \frac{\sigma}{\bar{x}} * 100\%$$

Коефіцієнт варіації характеризує відносну міру впливу випадкових факторів порівняно з основними визначальними умовами сукупності, які формують середню.

З огляду на те, що середнє квадратичне відхилення дає узагальнюючу характеристику коливання всіх варіантів сукупності, коефіцієнт варіації є найпоширенішим показником вимірювання коливання, що використовується для оцінки типовості середніх величин.

При $V_{\sigma} \leq 33\%$ - сукупність є однорідною, а середня величина є типовою та надійною характеристикою сукупності.

При цьому, якщо: $V_\sigma \leq 10\%$ - коливання ознаки незначні;

$10\% < V_\sigma \leq 30$ - коливання ознаки середні;

$V_\sigma > 30\%$ - коливання ознаки великі.

При $V_\sigma > 33\%$ - сукупність не є однорідною, тобто середні значення не можуть слугувати надійними характеристиками сукупності.

Задача 5.3. Визначити показники варіації на основі даних про продуктивність праці робітників у двох бригадах (табл. 5.7)

Таблиця 5.7

Вихідні та розрахункові дані для визначення показників варіації продуктивності праці робітників

№ робітника	Обсяг продукції, виготовленої 1 робітником, шт.		Відхилення					
			Бригада 1		Бригада 2		Бригада 1	Бригада 2
	Бригада 1 x_1	Бригада 2 x_2	$x_1 - \bar{x}$	$ x_1 - \bar{x} $	$x_2 - \bar{x}$	$ x_2 - \bar{x} $	$(x_1 - \bar{x})^2$	$(x_2 - \bar{x})^2$
1	2	8	- 8	8	-2	2	64	4
2	3	9	- 7	7	-1	1	49	1
3	12	10	2	2	0	0	4	0
4	15	11	5	5	1	1	25	1
5	18	12	8	8	2	2	64	4
Σ	50	50	0	30	0	6	216	10

Розв'язання.

1) Розмах варіації:

$$R = x_{max} - x_{min} :$$

$$R_1 = 18-2=16 (шт.)$$

$$R_2 = 12-8=4 (шт.)$$

2) Середня арифметична проста для кожної бригади:

$$\bar{x}_1 = \bar{x}_2 = \frac{\sum x}{n} = \frac{50}{5} = 10(шт.)$$

3) Середнє лінійне відхилення для незгрупованих даних:

$$\bar{d} = \frac{\sum |x - \bar{x}|}{n} = \frac{|x_1 - \bar{x}| + |x_2 - \bar{x}| + \dots + |x_n - \bar{x}|}{n}$$

$$\bar{d}_1 = \frac{30}{5} = 6(шт.)$$

$$\bar{d}_2 = \frac{6}{5} = 1,2(шт.)$$

4) Дисперсія незважена:

$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2}{n}$$

$$\sigma_1^2 = \frac{216}{5} = 43,2$$

$$\sigma_2^2 = \frac{10}{5} = 2$$

5) Середнє квадратичне відхилення:

$$\sigma = \sqrt{\sigma^2}$$

$$\sigma_1 = \sqrt{\sigma_1^2} = \sqrt{43,2} = 6,6 \text{ (шт.)}$$

$$\sigma_2 = \sqrt{\sigma_2^2} = \sqrt{2} = 1,41 \text{ (шт.)}$$

6) Коефіцієнт осциляції:

$$V_R = \frac{R}{\bar{x}} * 100\%$$

$$V_{R1} = \frac{16}{10} * 100\% = 160\%$$

$$V_{R2} = \frac{4}{10} * 100\% = 40\%$$

7) Лінійний коефіцієнт варіації:

$$V_{\bar{d}} = \frac{\bar{d}}{\bar{x}} * 100\%$$

$$V_{d1} = \frac{6}{10} * 100\% = 60\%$$

$$V_{d2} = \frac{1,2}{10} * 100\% = 12\%$$

8) Коефіцієнт варіації:

$$V_{\sigma} = \frac{\sigma}{\bar{x}} * 100\%$$

$$V_{\sigma1} = \frac{6,6}{10} * 100\% = 66\%$$

$$V_{\sigma2} = \frac{1,41}{10} * 100\% = 14,1\%$$

сукупність не є однорідною,
середня величина - ненадійна

сукупність є однорідною,
коливання ознаки середні,
середня - надійна

Задача 5.4. Визначити показники варіації на основі даних про розподіл посівної площі господарства за урожайністю ярої пшениці:

Урожайність пшениці, ц/га	Посівна площа, га f_i	x_i	$x_i f_i$	$ x_i - \bar{x} $	$(x_i - \bar{x})f$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^2 f$
14 - 16	100	15	1500	3,4	340	11,56	1156
16 - 18	300	17	5100	1,4	420	1,96	588
18 - 20	400	19	7600	0,6	240	0,36	144
20 - 22	200	21	4200	2,6	520	6,76	1352
Разом	1000	72	18400	8	1520	20,64	3240

Розв'язання.

1) У подібних прикладах насамперед визначається дискретне значення ознаки x_i в кожному інтервалі, за яке береться середина інтервалу.

2) Розмах варіації:

$$R = x_{max} - x_{min} = 22 - 14 = 8 \text{ у/га}$$

3) Середня арифметична зважена:

$$\bar{x} = \frac{\sum xf}{\sum f} = \frac{18400}{1000} = 18,4 (\text{у/га})$$

4) Середнє лінійне відхилення для згрупованих даних:

$$\bar{d} = \frac{\sum |x_i - \bar{x}| / f_i}{\sum f_i} = \frac{|x_1 - \bar{x}| / f_1 + |x_2 - \bar{x}| / f_2 + \dots + |x_n - \bar{x}| / f_n}{f_1 + f_2 + \dots + f_n} = \frac{1520}{1000} = 1,52 (\text{у/га})$$

5) Дисперсія зважена:

$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2 f_i}{\sum f_i} = \frac{3240}{1000} = 3,24$$

6) Середнє квадратичне відхилення:

$$\sigma = \sqrt{\sigma^2} = \sqrt{3,24} = 1,8 (\text{у/га})$$

7) Коефіцієнт осциляції:

$$V_R = \frac{R}{\bar{x}} * 100\% = \frac{8}{18,4} * 100 = 43,5\%$$

8) Лінійний коефіцієнт варіації:

$$V_{\bar{d}} = \frac{\bar{d}}{\bar{x}} * 100\% = \frac{1,52}{18,4} * 100 = 8,3\%$$

8) Коефіцієнт варіації:

$$V_{\sigma} = \frac{\sigma}{\bar{x}} * 100\% = \frac{1,8}{18,4} * 100 = 9,8\%$$

Тобто, сукупність є однорідною, коливання ознаки незначні, середня величина є надійною характеристикою сукупності.

Завдання для самостійного розв'язання.

Завдання 5.1. За наведеними даними про кількість дітей в домогосподарствах: 3; 3; 4; 6; 3; 3; 5; 3; 5; 2; 4; 3; 4; 5; 2; 4; 3; 1; 2; 1.

1) Побудувати дискретний ряд розподілу;

2) Побудувати полігон та кумуляту розподілу.

3) Визначити середню кількість дітей в домогосподарствах та оцінити надійність середньої величини.

Завдання 5.2 Розподіл студентів одного з факультетів за віком характеризується наступними даними:

Вік студентів, років	17	18	19	20	21	22	23	24	Усього
Число студентів	20	80	90	110	130	170	90	60	750

Обчисліть: а) розмах варіації; б) середнє квадратичне відхилення; в) коефіцієнт варіації віку студентів. Побудуйте полігон розподілу студентів за віком. Зробіть відповідний висновок.

Завдання 5.3. Маємо наступні рейтингові оцінки комерційних банків міста:

№	Оцінка, балів	№	Оцінка, балів	№	Оцінка, балів
1	2,8	7	3,3	13	2,9
2	3,0	8	2,9	14	5,0
3	3,7	9	4,8	15	4,2
4	2,3	10	2,0	16	2,0
5	1,2	11	3,2	17	2,7
6	4,0	12	2,6	18	1,5

- Виділити групи банків, для яких характерні:
 - стабільний фінансовий стан (1-2,4 бали);
 - фінансові та операційні недоліки у межах допустимого (2,5-3,4 бали);
 - серйозні фінансові проблеми (3,5 – 4,4);
 - критичний фінансовий стан (4,5 – 5 балів).
- Побудувати ряд розподілу с визначенням частот та часток.
- Визначити абсолютну та відносну щільність розподілу.
- Зобразити ряд розподілу графічно у вигляді гістограми.

Завдання 5.4. За даними таблиці варіацію оборот роздрібної торгівлі по областях визначивши:

- розмах варіації
- середнє лінійне відхилення
- дисперсію
- середнє квадратичне відхилення
- коефіцієнт варіації.

Оборот роздрібної торгівлі по областях за січень-жовтень 2012 р., млн. грн.	Кількість областей
До 10000	12
10000 – 15000	5
15000 – 20000	4
20000 – 25000	1
Більше 25000	5

Зробити висновки.

Завдання 5.4. Для характеристики виробничого стажу робітників однієї з галузей промисловості проведено обстеження різних категорій робітників. Результати обстеження систематизовані у вигляді таблиці:

Групи робітників за стажем роботи, років	Питома вага робітників за стажем, в % до підсумку		
	Робочі	Майстри	Технологи
До 2	7	1	-
2 - 4	15	10	3
4 - 6	20	22	20
6 - 8	30	20	10
8 - 10	10	23	32
10 - 12	8	7	20
12 - 14	2	6	10
Більше 14	8	11	5

1) Визначити показники варіації (для розрахунку оберати будь-які дві категорії робітників). Зробити висновки.

2) Зобразити обидва ряди розподілу графічно у вигляді огіві.

Завдання 5.5. За даними про рівень річного доходу в різних групах населення міста необхідно:

1) розрахувати характеристики відповідного ряду розподілу: середній рівень доходу, розмах варіації, середнє квадратичне відхилення, коефіцієнт варіації.

2) визначити за допомогою гістограми та кумуляти.

Доходи, тис. грн.	Чисельності груп, тис. осіб
До 10	42
10 – 15	90
15 – 25	246
25 – 40	132
40 – 60	60
60 і вище	30

Зробіть висновки.

Тести для самоконтролю.

1. Що з нижчепереліченого є частотною характеристикою рядів розподілу?

А) кумулята

Б) щільність розподілу

В) стандартне відхилення

Г) розмах варіації

2. Графік, на вісі абсцис якого відкладають варіанти, а на вісі ординат – кумулятивні частоти, називається:

А) гістограма

Б) полігон

В) кумулята

Г) огіва

3. Різниця між найбільшим і найменшим значеннями варіантів – це:

А) коефіцієнт варіації

Б) варіація ознаки

В) розмах варіації

Г) середнє лінійне відхилення

4. Дисперсія дорівнює 4, яким є значення стандартного відхилення?

А) 2

Б) 4

В) 8

Г) 16

5. Різниця, коливання індивідуальних значень ознаки усередині досліджуваної сукупності, називається:

А) варіація ознаки

Б) розмах варіації

В) дисперсія

Г) ряд розподілу

6. Коефіцієнт варіації дорівнює 20%. Що можна сказати про однорідність сукупності?

А) коливання ознаки незначні

Б) коливання ознаки середні

В) коливання ознаки великі

Г) сукупність не є однорідною

7. Середня арифметична квадратів відхилень кожного значення ознаки від загальної середньої називається:

А) дисперсія

Б) середнє лінійне відхилення

В) розмах варіації

Г) стандартне відхилення

8. Графік, що використовується для зображення дискретних варіаційних рядів, називається:

А) гістограма

Б) полігон

В) кумулята

Г) огіва

9. Стандартне відхилення дорівнює 4, яким є значення дисперсії?

- А) 2
- Б) 4
- В) 8
- Г) 16

10. Коефіцієнт варіації дорівнює 10%. Що можна сказати про однорідність сукупності?

- А) коливання ознаки незначні
- Б) коливання ознаки середні
- В) коливання ознаки великі
- Г) сукупність не є однорідною

11. Що з нижчепереліченого є частотною характеристикою рядів розподілу?

- А) кумулята
- Б) стандартне відхилення
- В) частка
- Г) дисперсія

12. Графік, побудований з використанням накопичених частот, називається:

- А) гістограма
- Б) кумулята
- В) полігон
- Г) крива Варзара

13. Графік для зображення інтервального варіаційного ряду, на якому ряд розподілу зображений у вигляді суміжних один з одним стовпчиків, називається:

- А) гістограма
- Б) полігон
- В) кумулята
- Г) огіва

14. Середня арифметична з відхилень індивідуальних значень від середньої, без урахування знака цих відхилень, називається:

- А) розмах варіації
- Б) дисперсія
- В) коефіцієнт варіації
- Г) середнє лінійне відхилення

15. Дисперсія дорівнює 9, яким є значення стандартного відхилення?

- А) 3
- Б) 9
- В) 18
- Г) 81

16. При якому значенні коефіцієнта варіації сукупність є якісно неоднорідною?

- А) до 10%
- Б) до 30 %
- В) більше 30%
- Г) більше 33%

17. Що з нижчепереліченого **не є** частотною характеристикою рядів розподілу?

- А) розмах варіації
- Б) частка
- В) щільність розподілу
- Г) кумулятивна частка

18. Коефіцієнт варіації дорівнює 33%. Що можна сказати про однорідність сукупності?

- А) коливання ознаки незначні
- Б) коливання ознаки середні
- В) коливання ознаки великі
- Г) сукупність не є однорідною

19. Графік, при побудові якого змінюють місцями вісі кумуляти, називається:

- А) гістограма
- Б) полігон
- В) крива Лоренца
- Г) огіва

20. Яка частотна характеристика використовується для характеристики ряду розподілу з нерівними інтервалами?

- А) кумулятивна частота
- Б) щільність розподілу
- В) частка
- Г) розмах варіації

21. При якому значенні коефіцієнта варіації середня **не є** надійною характеристикою сукупності?

- А) до 10%
- Б) до 30 %
- В) більше 30%
- Г) більше 33%

22. Коефіцієнт варіації дорівнює 30%. Що можна сказати про однорідність сукупності?

- А) коливання ознаки незначні
- Б) коливання ознаки середні
- В) коливання ознаки великі
- Г) сукупність не є однорідною

23. Що з нижчепереліченого є частотною характеристикою рядів розподілу?

- А) кумулята
- Б) кумулятивна частка
- В) стандартне відхилення
- Г) розмах варіації

24. Показник, що визначається, як корінь квадратний з дисперсії, називається:

- А) розмах варіації
- Б) коефіцієнт варіації
- В) середнє квадратичне відхилення
- Г) середнє лінійне відхилення

25. Графік, на вісі абсцис якого відкладають кумулятивні частоти, а на вісі ординат – варіанти, називається:

- А) гістограма
- Б) полігон
- В) кумулята
- Г) огіва

26. Показник, що відображає відносні коливання крайніх значень ознаки навколо середньої, називається:

- А) коефіцієнт осциляції
- Б) розмах варіації
- В) коефіцієнт варіації
- Г) середнє лінійне відхилення

27. Стандартне відхилення дорівнює 9, яким є значення дисперсії?

- А) 3
- Б) 9
- В) 18
- Г) 81

28. Що з нижчепереліченого не є частотною характеристикою рядів розподілу?

- А) кумулята
- Б) частка
- В) щільність розподілу
- Г) кумулятивна частка

29. Коефіцієнт варіації дорівнює 9%. Що можна сказати про однорідність сукупності?

- А) коливання ознаки незначні
- Б) коливання ознаки середні
- В) коливання ознаки великі
- Г) сукупність не є однорідною

30. Графік, побудований з використанням накопичених частот, називається:

- А) крива Варзара
- Б) полігон
- В) огіва
- Г) гістограма

31. Показник, що характеризує частку усередненого значення абсолютних відхилень від середньої величини, називається:

- А) лінійний коефіцієнт варіації
- Б) коефіцієнт осциляції
- В) середнє лінійне відхилення
- Г) стандартне відхилення

32. Коефіцієнт варіації дорівнює 35%. Що можна сказати про однорідність сукупності?

- А) коливання ознаки незначні
- Б) коливання ознаки середні
- В) коливання ознаки великі
- Г) сукупність не є однорідною

33. Показник, що використовується для оцінки надійності середніх величин, називається:

- А) коефіцієнт варіації
- Б) коефіцієнт осциляції
- В) розмах варіації
- Г) стандартне відхилення

34. Який з перелічених показників варіації виражається в тих же одиницях виміру, що й ознака?

- А) коефіцієнт варіації
- Б) коефіцієнт осциляції
- В) дисперсія
- Г) стандартне відхилення

6. АНАЛІЗ ДИФЕРЕНЦІАЦІЇ, КОНЦЕНТРАЦІЇ ТА ПОДІБНОСТІ РОЗПОДІЛІВ

6.1 Структурні характеристики розподілів

Характеристикою варіаційних рядів також є особливий вид середніх величин - структурні середні, які застосовуються для вивчення внутрішньої будови рядів розподілу значень ознаки, а також для оцінки ступеневої середньої величини, якщо за наявними статистичними даними її розрахунок не може бути виконаний (наприклад, якщо відсутні дані і про обсяг виробництва, і про суму витрат по групах підприємств). Найвідомішими зі структурних середніх є мода та медіана.

Мода – це найпоширеніше значення ознаки, тобто варіант, який в ряду розподілу має найбільшу частоту (частку) або щільність.

Медіана – це варіант, який припадає на середину упорядкованого (ранжованого) ряду розподілу і ділить його на дві рівні частини.

Якщо мода виражає типовий, найбільш розповсюджений варіант значення признаку, то медіана практично виконує функції середньої для неоднорідної сукупності.

Приклад. Необхідно дати характеристику середнього доходу групи людей із 100 осіб, з яких 99 мають доходи в інтервалі від 100 до 200 дол. в місяць, а місячні доходи останнього складають 50000 дол.(табл.6.1).

Таблиця 6.1

Ранжування громадян за щомісячним доходом

№ особи	1	2	3	4	...	50	51	...	99	100
Дохід, дол.	100	104	104	107	...	162	164	...	200	50000

Якщо ми скористаємося середньою арифметичною, то отримаємо середній дохід близько 600-700 дол., який не тільки у кілька разів менший від доходу 100-ї людини, але й має мало східного з доходами решти групи. Медіана ж, яка в даному випадку дорівнює 163 дол., дозволить дати об'єктивну характеристику рівня доходів 99% даної групи людей.

В дискретному ряді

- мода визначається візуально за максимальною частотою або часткою.
- медіана визначається в 2 етапи:

1) знаходиться номер медіанної одиниці ряду за формулою:

$$N_{Me} = \frac{\sum f_i + 1}{2}$$

2) знаходиться варіант, для якого кумулятивна частота дорівнює або перевищує половину обсягу сукупності.

Приклад. Маємо наступні дані про розподіл робітників за тарифним розрядом на одному з підприємств міста (табл.6.2).

Таблиця 6.2

Розподіл робітників підприємства за тарифним розрядом

Тарифний розряд	2	3	4 (Me)	5 (Mo)	6	Разом
Чисельність робітників	22	38	56	60	14	190

Визначення моди не становить складності – найбільшу частоту (60 осіб) має 5 тарифний розряд, тобто він і є модою.

Знаходимо медіану:

1) Визначаємо номер медіанної одиниці ряду

$$N_{Me} = \frac{190 + 1}{2} = 95,5$$

Тобто точна середина сукупності знаходиться між 95-м та 96 робітниками.

2) Визначаємо, до якої групи відносяться робітники з цими порядковими номерами:

- очевидно, що робочих з цими номерами немає в першій групі, де всього 12 осіб;

- кумулятивна частота для другої групи складає $22+38=60 < 95,5$ – не медіана

- кумулятивна частота для третьої групи складає $22+38+56=116 > 95,5$ – медіаною є 4 тарифний розряд.

В інтервальному ряді:

Мода:

- для рядів з рівними інтервалами визначається за найбільшою частотою за формулою:

$$Mo = x_0 + h * \frac{f_{Mo} - f_{Mo-1}}{(f_{Mo} - f_{Mo-1}) + (f_{Mo} - f_{Mo+1})},$$

де x_0 – нижня межа модального інтервалу (модальним є інтервал, що має найбільшу частоту або частку);

h - величина модального інтервалу;

f_{Mo} – частота (частка) модального інтервалу;

f_{Mo-1} - частота (частка) інтервалу, що передує модальному;

f_{Mo+1} - частота (частка) інтервалу, наступного за модальним.

- для рядів з нерівними інтервалами визначається за найбільшою щільністю розподілу за формулою:

$$Mo = x_0 + h * \frac{g_{Mo} - g_{Mo-1}}{(g_{Mo} - g_{Mo-1}) + (g_{Mo} - g_{Mo+1})},$$

де x_0 – нижня межа модального інтервалу (модальним є інтервал, що має найбільшу щільність);

h - величина модального інтервалу;

g_{Mo} – щільність модального інтервалу;

g_{Mo-1} - щільність інтервалу, що передує модальному;

g_{Mo+1} - щільність інтервалу, наступного за модальним.

Медіана визначається за формулою:

$$Me = x_0 + h * \frac{\frac{1}{2} \sum f_i - S_{Me-1}}{f_{Me}}$$

де x_0 – нижня межа медіанного інтервалу (медіанним є інтервал, кумулятивна частота (частка) якого перевищує половину загальної суми частот (часток));

h - величина медіанного інтервалу;

S_{Me-1} – кумулятивна частота (частка) інтервалу, що передує медіанному;

f_{Me} – частота (частка) медіанного інтервалу.

Задача 6.1. Маємо 50 банків, ранжованих за величиною їх активів (табл. 6.3). Визначити моду та медіану.

Таблиця 6.3

Ранжування банків за величиною активів

№ банку	Активи, млн. гр. од.	№ банку	Активи, млн. гр. од.	№ банку	Активи, млн. гр. од.	№ банку	Активи, млн. гр. од.	№ банку	Активи, млн. гр. од.
1	330,0	11	14,1	21	8,0	31	3,5	41	2,3
2	29,3	12	13,3	22	7,8	32	3,3	42	2,3
3	27,1	13	12,0	23	6,6	33	3,1	43	2,3
4	25,5	14	10,2	24	6,2	34	2,9	44	2,2
5	23,1	15	9,9	25	5,1	35	2,8	45	2,2
6	20,8	16	9,7	26	4,9	36	2,7	46	2,1
7	19,0	17	9,5	27	4,7	37	2,5	47	1,9
8	17,1	18	9,2	28	4,5	38	2,4	48	1,8
9	15,9	19	8,9	29	4,0	39	2,4	49	1,7
10	14,8	20	8,5	30	3,8	40	2,4	50	1,5

Розв'язання.

Побудуємо інтервальний варіаційний ряд з нерівними інтервалами (табл. 6.4).

1) Мода:

$$Mo = 2 + 1 * \frac{14 - 8}{(14 - 8) + (14 - 3,5)} = 2,36(\text{млн.гр.од.})$$

Таким чином, найбільш типовим є уставний фонд банків у розмірі 2,36 млн. гр. од.

2) Медіана:

$$Me = 3 + 2 * \frac{25 - 18}{7} = 5(\text{млн.гр.од.})$$

Тобто, в половині досліджених банків величина активів не перевищує 5 млн. гр. од.

Таблиця 6.4

Розподіл банків за величиною активів

Величина активів, млн. гр. од. x_i	Кількість банків (частота) f_i	Частка банків, % d_i	Накопичена (кумулятивна) частота S_{fi}	Щільність розподілу $g_i = \frac{f_i}{h_i}$
1,5-2	4	8	4	8
2-3	14	28	18	14
3-5	7	14	25	3,5
5-10	11	22	36	2,2
10-20	8	16	44	0,8
20-30	5	10	49	0,5
30-330	1	2	50	0,0035
Разом	50	100	x	x

Аналогічно з знаходженням медіани у варіаційних рядах розподілу можна визначити значення ознаки у одиниць, що ділять ряд на чотири, п'ять, десять, двадцять та сто частин.

Квантилі, або градієнти – такі значення ознаки, що ділять всі одиниці розподілу на рівні чисельності (рис.6.1).

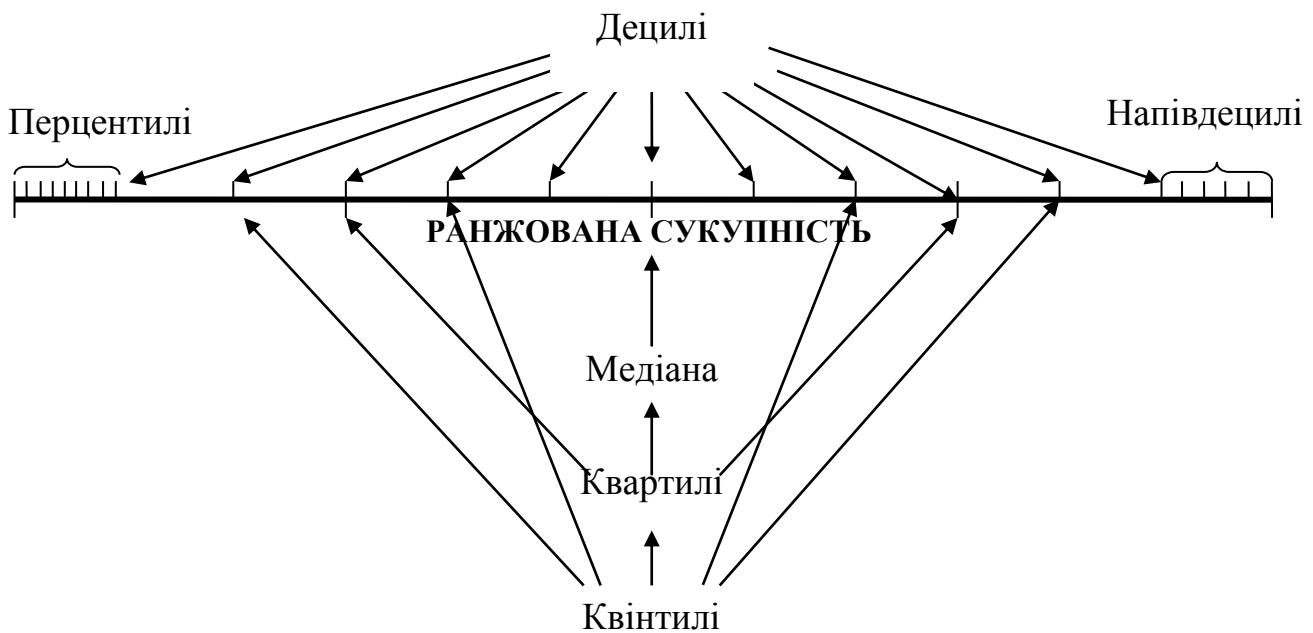


Рис. 6.1. Види квантилів

Види квантилів:

- **Квартилі (Q)** – це значення варіантів, які ділять упорядкований ряд за обсягом на чотири рівні частини (в рядах розподілу визначаються 3 квартилі):

- перший (нижній) квантиль

$$Q_1 = x_{Q_1} + h * \frac{\frac{1}{4} \sum f - S_{Q_1-1}}{f_{Q_1}}$$

- третій (верхній) квантиль

$$Q_3 = x_{Q_3} + h * \frac{\frac{3}{4} \sum f - S_{Q_3-1}}{f_{Q_3}},$$

де x_{Q_1} , x_{Q_3} – нижня межа інтервалу, що містить відповідно перший та третій квантиль;

h - величина інтервалу, що містить даний квантиль;

m – кількість груп;

S_{Q_1-1} , S_{Q_3-1} - кумулятивна частота (частка) інтервалу, що передуює інтервалу, який містить відповідно перший та третій квантиль;

f_{Q_1} , f_{Q_3} – частота (частка) інтервалу, що містить відповідно перший та третій квантиль;

- **Квінтилі** – ділять сукупність на п'ять рівних частин;

- **Децилі (DL)** - це значення варіантів, які ділять упорядкований ряд за обсягом на десять рівних частин (в рядах розподілу визначаються 9 децилів):

- перший дециль:

$$DL_1 = x_{D_1} + h * \frac{\frac{1}{10} \sum f - S_{D_1-1}}{f_{D_1}},$$

- дев'ятий дециль:

$$DL_9 = x_{D_9} + h * \frac{\frac{9}{10} \sum f - S_{D_9-1}}{f_{D_9}},$$

де x_{Q1} , x_{Q9} – нижня межа інтервалу, що містить відповідно перший та дев'ятий дециль;

h - величина інтервалу, що містить даний дециль;

S_{D1-1} , S_{D9-1} - кумулятивна частота (частка) інтервалу, що передуює інтервалу, який містить відповідно перший та дев'ятий дециль;

f_{D1} , f_{D9} – частота (частка) інтервалу, що містить відповідно перший та дев'ятий дециль;

- **Напівдецилі** - значення ознаки, що ділять ряд на двадцять частин;

- **Перцентилі** – значення ознаки, що ділять ряд на сто частин.

Примітка. Медіана є водночас другим квантилем, третім квінтилем, п'ятим децилем, десятим напівдецилем, п'ятидесятим перцентилем.

Задача 6.2. За даними задачі 6.1. визначити нижні та верхні квантилі та децилі.

Розв'язання.

Для розрахунку відповідних квантилів необхідно визначити певні частотні характеристики в ряду розподілу (табл. 6.5).

1) Квантилі:

а) нижній (перший):

$$Q_1 = x_{Q_1} + h * \frac{\frac{1}{4} \sum f - S_{Q_1-1}}{f_{Q_1}} = 2 + 1 * \frac{0,25 * 50 - 4}{14} = 2,6(\text{млн.грод})$$

або

$Q_1 = 2 + 1 * \frac{0,25 * 100 - 8}{28} = 2,6(\text{млн.грод.})$ – максимальна величина активів 25% найменших банків

Таблиця 6.5

Вихідні та розрахункові дані для визначення квантилів
в ряду розподілу банків за величиною активів

Величина активів, млн.гр.од, x_i	Кількість банків		Накопичена (кумулятивна)	
	Частота f_i	Частка, % d_i	частота S_{fi}	частка S_{di}
1,5-2	4	8	4	8
2-3	14	28	18	36
3-5	7	14	25	50
5-10	11	22	36	72
10-20	8	16	44	88
20-30	5	10	49	98
30-330	1	2	50	100
Разом	50	100	X	X

б) верхній (третій):

$$Q_3 = x_{Q_3} + h * \frac{\frac{3}{4} \sum f - S_{Q_3-1}}{f_{Q_3}} = 10 + 10 * \frac{0,75 * 50 - 36}{8} = 11,875(\text{млн.гр.од.})$$

або

$$Q_3 = 10 + 10 * \frac{0,75 * 100 - 72}{16} = 11,875(\text{млн.гр.од.}) - \text{мінімальна величина}$$

активів 25% найбільших банків

2) Децилі:

а) нижній (перший):

$$DL_1 = x_{D_1} + h * \frac{\frac{1}{10} \sum f - S_{D_1-1}}{f_{D_1}} = 2 + 1 * \frac{0,1 * 50 - 4}{14} = 2,07(\text{млн.гр.од.})$$

або

$$DL_1 = 2 + 1 * \frac{0,1 * 100 - 8}{28} = 2,07(\text{млн.гр.од.}) - \text{максимальна величина активів}$$

10% найменших банків

б) верхній (дев'ятий):

$$DL_9 = x_{D_9} + h * \frac{\frac{9}{10} \sum f - S_{D_9-1}}{f_{D_9}} = 20 + 10 * \frac{0,9 * 50 - 44}{5} = 22(\text{млн.гр.од.}) \text{ або}$$

$$DL_9 = 20 + 10 * \frac{0,9 * 100 - 88}{10} = 22(\text{млн.гр.од.}) - \text{мінімальна величина}$$

активів 10% найбільших банків

6.2 Показники диференціації розподілів

Диференціація являє собою розшарування сукупності за рівнем досліджуваної ознаки. Вона оцінюється рядом показників:

1) **коефіцієнти диференціації** - у розподілі ознаки використовуються співвідношення між квантилями (точками ряду розподілу, що ділять його в заданих співвідношеннях). Як квантилі можуть використатися квартилі (чверті), квінтилі (п'яті), децилі (десяті), напівдецилі (двадцяті) і перцентилі (соті).

Найпоширенішим є **коефіцієнт децильної диференціації**:

$$K_{д} = \frac{DL_9}{DL_1},$$

де DL_9 – дев'ятий дециль; DL_1 – перший дециль.

Аналогічно розраховують інші коефіцієнти диференціації.

Задача 6.3. За даними задачі 6.1. визначити коефіцієнт децильної диференціації.

Розв'язання.

Коефіцієнт децильної диференціації:

$$K_{д} = \frac{DL_9}{DL_1} = \frac{22}{2,07} = 11(\text{разів})$$

Одержаний коефіцієнт показує високий рівень диференціації, так як мінімальна величина активів 10% найбільших банків у 11 разів вища, ніж максимальна величина активів 10% найменших банків

2) **фондові коефіцієнти** – визначають співвідношення середніх рівнів ознак в групах з максимальними та мінімальними їх значеннями.

Так, **децильний фондовий коефіцієнт** визначає співвідношення між середніми значеннями ознаки 10% одиниць сукупності з найвищими варіантами та 10% одиниць сукупності з найменшими варіантами.

$$K_{\phi} = \frac{\bar{X}_{10\% \max}}{\bar{X}_{10\% \min}}$$

Квінтильний - визначає співвідношення між середніми значеннями ознаки 20% одиниць сукупності з найвищими варіантами та 20% одиниць сукупності з найменшими варіантами і так далі.

Задача 6.4. За даними задачі 6.1. визначити децильний фондний коефіцієнт.

Розв'язання.

1) 10 % найбільших та найменших банків складають по:

$$50 * 10\% = 5 \text{ (банків)}$$

2) децильний фондний коефіцієнт:

$$K_{\phi} = \frac{\bar{X}_{10\% \max}}{\bar{X}_{10\% \min}} = \frac{(330 + 29,3 + 27,1 + 25,5 + 23,1) / 5}{(1,5 + 1,7 + 1,8 + 1,9 + 2,1) / 5} = 48,3 \text{ (рази)}$$

Одержаний коефіцієнт показує, що рівень диференціації досить високий, середня величина активів 10% найбільших банків у 48,3 рази перевищує середню величину активів 10% найменших банків.

6.3 Показники концентрації розподілів

До показників диференціації близькі за значенням показники концентрації розподілів.

Концентрація – нерівномірність розподілів досліджуваної ознаки, що не пов'язана ні з обсягом сукупності, ні з чисельністю окремих груп.

Оцінка нерівномірності розподілу значень ознаки між окремими складовими сукупностей ґрунтується на порівнянні часток двох розподілів – за кількістю елементів d_i та за обсягом значень явища D_i . Якщо розподіл значень ознаки рівномірний, то $d_i = D_i$.

Відхилення часток свідчить про певну нерівномірність розподілу, яка вимірюється наступними показниками:

1) **коефіцієнт локалізації** L_i – означає зосередження обсягу явища у окремих групах одиниць сукупності; розраховується для кожної i -ї складової сукупності:

$$L_i = \frac{D_i}{d_i}$$

де D_i – частка i -ї групи у загальному обсязі явища;

d_i - частка i -ї групи у загальному обсязі сукупності.

При $L_i = 1$ – рівномірний розподіл ознаки;

$L_i > 1$ – наявність концентрації значень ознаки в i -ій складовій;

$L_i < 1$ - відсутність концентрації.

2) **коефіцієнт Лоренца** K_L є узагальнюючою характеристикою відхилення фактичного розподілу від рівномірного. Даний коефіцієнт ще називають *коефіцієнтом концентрації* і визначають за формулою:

$$K_L = \frac{1}{2} \sum_1^m |D_i - d_i|$$

Значення його коливаються у межах від 0 до 1, або від 0 до 100%.

У рівномірному розподілі $K=0$. Чим помітніша концентрація, тим більше значення K_L відхиляється від 0.

При $K_L \leq 0,25$ ($K_L \leq 25\%$) - низька концентрація ознаки у обмеженої кількості одиниць або груп сукупності;

при $0,25 < K_L \leq 0,35$ ($25\% < K_L \leq 35\%$) - помірна концентрація ознаки;

при $K_L > 0,35$ (від 35%) - висока концентрація ознаки.

Задача 6.5. За даними задачі 6.1. визначити коефіцієнти локалізації та коефіцієнт концентрації банківських активів.

Розв'язання.

Для розрахунку відповідних коефіцієнтів необхідно визначити частку обсягу явища та обсягу сукупності у кожній складовій ряду розподілу (табл. 6.6).

Таблиця 6.6

Вихідні та розрахункові дані для визначення коефіцієнтів локалізації та концентрації активів в групах банків

Величина активів, млн.гр.од, x_i	Частота f_i	Активи груп банків $x_i f_i$	% до підсумку		$L_j = \frac{D_i}{d_i}$	$ D_i - d_i $
			Кількість банків d_i	Величина активів D_i		
1,5-2	4	7	8	1,2	0,15	6,80
2-3	14	35	28	6,1	0,22	21,90
3-5	7	28	14	4,8	0,34	9,20
5-10	11	82,5	22	14,3	0,65	7,70
10-20	8	120	16	20,8	1,30	4,80
20-30	5	125	10	21,6	2,16	11,60
30-330	1	180	2	31,2	15,60	29,20
Разом	50	577,5	100	100	x	91,20

Коефіцієнт Лоренца:

$$K_L = \frac{1}{2} \sum_1^m |D_i - d_i| = \frac{1}{2} * 91,2 = 45,9\%, \text{ або } 0,495$$

Значення K_L свідчить про відносно високий рівень концентрації банківських активів. При цьому, розрахунки L_i показують, що найвища концентрація активів спостерігається у найбільшому банку №1, а також в 5 банках з величиною активів 20-30 млн. гр. од. та 8 банках з величиною активів 10-20 млн. гр. од.

3) **коефіцієнт Джині** K_G – відображає ступінь відхилення фактичного розподілу значень ознаки за чисельно рівними групами одиниць сукупності від лінії їх рівномірного розподілу і розраховується за формулами:

$$K_G = 1 - 2 \frac{\sum d_i S_{Di}}{\sum d_i D_i} - \text{якщо частки виражені коефіцієнтами,}$$

$$K_G = 1 - 2 \frac{\sum d_i S_{Di}}{10000} + \frac{\sum d_i D_i}{10000} - \text{якщо частки виражені у відсотках,}$$

де S_{Di} – накопичена частка i -ї групи у загальному обсязі явища.

Значення його коливаються у межах від 0 до 1.

При $K_G < 0,29$ – низька концентрація ознаки у обмеженій кількості одиниць або груп сукупності;

$0,3 < K_G < 0,44$ – помірна концентрація ознаки;

$K_G > 0,45$ – висока концентрація ознаки.

Коефіцієнт Джині базується на **кривій концентрації Лоренца**, яка графічно відображає нерівномірність розподілу певної ознаки.

Спосіб побудови. Крива будується у прямокутній системі координат. На вісі абсцис (OX) відкладають накопичені (кумулятивні) частоти обсягу сукупності (у %), а на вісі ординат – накопичені (кумулятивні) частоти обсягу явища (у %). Також будується пряма, що проходить з нижнього лівого кута до верхнього правого – лінія рівномірного розподілу (рис. 6.2).

Чим сильніша концентрація досліджуваної ознаки, тим помітніше крива Лоренца відхиляється від лінії рівномірного розподілу, та навпаки, чим слабша концентрація ознаки, тим ближче буде крива до прямої.

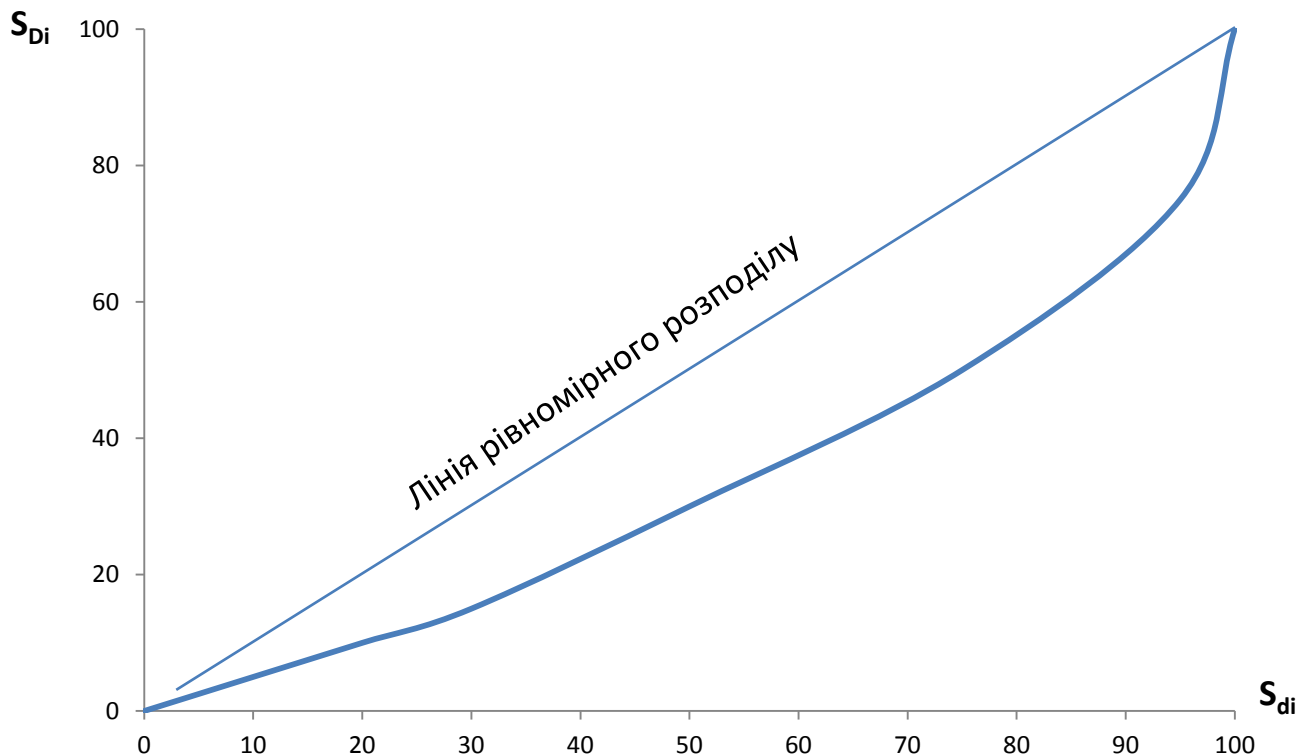


Рис. 6.2 Загальний вигляд кривої концентрації Лоренца

Задача 6.6. За даними задачі 6.1. визначити коефіцієнт Джині та побудувати криву концентрації Лоренца.

Розв'язання.

Для визначення коефіцієнту та побудови кривої необхідно здійснити певні розрахунки (табл. 6.7).

Таблиця 6.7

Вихідні та розрахункові дані для визначення коефіцієнту Джині та побудови кривої Лоренца

Величина активів, млн.гр.од. x_i	% до підсумку				$d_i S_{D_i}$	$d_i D_i$
	Кількість банків (частка) d_i	Величина активів D_i	Накопичена частка банків d_i^H	Накопичена частка активів S_{D_i}		
1,5-2	8	1,2	8	1,2	9,6	9,6
2-3	28	6,1	36	7,3	204,4	170,8
3-5	14	4,8	50	12,1	169,4	67,2
5-10	22	14,3	72	26,4	580,8	314,6
10-20	16	20,8	88	47,2	755,2	332,8
20-30	10	21,6	98	68,8	688	216
30-330	2	31,2	100	100	200	62,4
Разом	100	100	-	-	2607,4	1173,4

1) Коефіцієнт Джині:

$$K_G = 1 - 2 \frac{\sum d_i S_{D_i}}{10000} + \frac{\sum d_i D_i}{10000} = 1 - 2 * 0,26074 + 0,11734 = 0,596,$$

Величина K_G свідчить про високий ступень нерівномірності розподілу банківських активів, тобто їх високу концентрацію в окремих групах банків.

2) За розрахунковими даними графі 4 та 5 таблиці 6.7 побудуємо криву Лоренца (рис. 6.3)

4) **Коефіцієнт Херфіндаля** K_H – визначає вплив лише домінуючих одиниць або груп сукупності у загальному обсязі. Чим більшим є значення K_H , тим вища концентрація ознаки у домінуючих одиниць або груп.

$$K_H = \sum \left(\frac{x_i}{\sum x_i} \right)^2 \quad \text{або} \quad K_H = \sum \left(\frac{x_i f_i}{\sum x_i f_i} \right)^2$$

Якщо ж були проведені попередні розрахунки часток обсягу явища в окремих групах, то можна скористатися формулою:

$$K_H = \sum D_i^2$$

Значення його коливаються у межах від 0 до 1, або від 0 до 10000 (якщо частка у %).

При $K_H \leq 0,1$ ($K_H \leq 1000$) – низька концентрація ознаки;

при $0,1 < K_H \leq 0,2$ ($1000 < K_H \leq 2000$) – помірна концентрація ознаки;

при $K_H > 0,2$ ($K_H > 2000$) – висока концентрація ознаки, тобто наявне домінування окремих одиниць або груп сукупності.

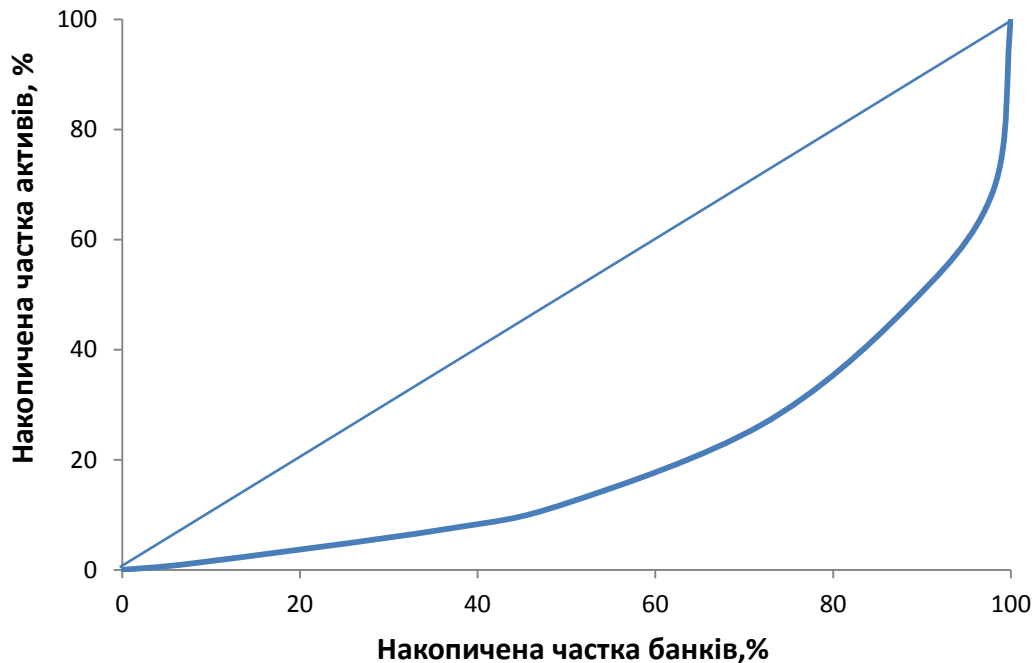


Рис. 6.3. Крива концентрації банківських активів

Задача 6.7. За даними задачі 6.1. оцінити концентрацію активів банків за коефіцієнтом Херфіндаля.

Розв'язання.

Для визначення коефіцієнту необхідно розрахувати суму квадратів часток обсягу явища (табл. 6.8).

Таблиця 6.8

Вихідні та розрахункові дані для визначення коефіцієнту Херфіндаля

Величина активів, млн. гр. од. x_i	Кількість банків (частота) f_i	Активи груп банків, млн. гр. од. $x_i f_i$	$\left(\frac{x_i f_i}{\sum x_i f_i} \right)^2$
1,5-2	4	7	0,0001
2-3	14	35	0,004
3-5	7	28	0,002
5-10	11	82,5	0,02
10-20	8	120	0,043
20-30	5	125	0,047
30-330	1	180	0,097
Разом	50	577,5	0,213

Коефіцієнт Херфіндаля:

$$K_H = \sum \left(\frac{x_i f_i}{\sum x_i f_i} \right)^2 = 0,213$$

Величина K_H свідчить про високу концентрацію банківських активів, що підтверджує домінуючу роль банку №1 та декількох крупних банків.

6.4 Подібність розподілів

Подібність розподілів - схожість структур двох об'єктів або одного об'єкта за двома ознаками

Для виявлення схожості структур (за аналогією з коефіцієнтом концентрації) розраховується **коефіцієнт подібності P** :

$$P = 1 - \frac{1}{2} \frac{\sum |d_i - d_k|}{100}$$

де d_i - частка i -ї групи у загальному обсязі сукупності для одного об'єкту чи ознаки, %;

d_k - частка k -ї групи у загальному обсязі сукупності для іншого об'єкту чи ознаки, %.

Значення P коливається від 0 до 1. Чим більше відхилення структур, тем менше значення коефіцієнта P .

Якщо структури однакові, то $P=1$.

При $P \leq 0,3$ - низький рівень подібності структур;

при $0,3 < P \leq 0,7$ - середній рівень подібності структур;

при $P > 0,7$ - високий рівень подібності структур.

Задача 6.8. Маємо дані про розподіл спеціалістів за галузями економіки по 2 регіонах. Оцінити ступінь подібності структур зайнятості спеціалістів.

Розв'язання.

Для визначення коефіцієнту необхідно розрахувати суму абсолютних відхилень часток обсягів сукупності (табл. 6.9).

Таблиця 6.9

Вихідні та розрахункові дані для визначення коефіцієнту подібності

Галузь економіки	Структура зайнятості, %		$ d_i - d_k $
	Регіон А d_i	Регіон Б d_k	
Матеріальне виробництво	25	33	8
Освіта	31	21	10
Охорона здоров'я	18	19	1
Інші сфери	26	27	1
Разом	100	100	20

Коефіцієнт подібності:

$$P = 1 - \frac{1}{2} \frac{\sum |d_i - d_k|}{100} = 1 - \frac{1}{2} * \frac{20}{100} = 0,9$$

Величина P свідчить про високий рівень подібності структур, схожість структур зайнятості в регіонах становить 90%.

Завдання для самостійного розв'язання.

Завдання 6.1. Маємо інтервальний ряд розподілу господарств району за урожайністю зернових культур:

Урожайність, ц/га	Кількість господарств
16-20	3
20-24	7
24-28	12
28-32	15
32-36	5
36-40	4

Визначити моду та медіану.

Завдання 6.2. Маємо дані про структуру реалізованої сільськогосподарської продукції по двох господарствах:

Господарства	Товарна продукція, % до підсумку						
	Зерно	Соняшник	Овочі	Плоди	Молоко	М'ясо ВРХ	Разом
А	25	45	2	10	10	8	100
Б	35	35	5	5	10	10	100

Оцінити ступінь подібності структур товарної продукції.

Завдання 6.3. Є дані про розподіл внесків фізичних осіб в комерційному банку:

Величина внеску, грн.	Число внесків
До 2000	2200
2000-4000	1320
4000-6000	840
6000-8000	360
8000-10000	280
Разом	5000

Визначте нижній і верхній кватиль і розрахуйте коефіцієнт кватильної диференціації. Зробіть відповідний висновок.

Завдання 6.4. Маємо наступні дані обстеження аграрних вищих навчальних закладів (за формою 2-ЗНК станом на січень 2013 р.):

№	Назва ВНЗ	Кількість студентів, тис. осіб
1	Білоцерківський державний аграрний університет	7,49
2	Вінницький національний аграрний університет	7,81
3	Житомирський національний агроекологічний університет	8,13
4	Дніпропетровський державний аграрний університет	6,19
5	Керченський державний морський технологічний університет"	1,88
6	Луганський національний аграрний університет	9,1
7	Львівський національний аграрний університет	3,97
8	Львівський національний університет ветеринарної медицини та біотехнологій	7,86
9	Миколаївський національний аграрний університет	5,98
10	Одеський державний аграрний університет	3,79
11	Подільський державний аграрно-технічний університет	5,04
12	Полтавська державна аграрна академія	6,87
13	Сумський національний аграрний університет	8,68
14	Таврійський державний агротехнологічний університет	5,54
15	Уманський національний університет садівництва	5,83
16	Харківська державна зооветеринарна академія	2,58
17	Харківський національний аграрний університет	5,07
18	Харківський національний технічний університет сільського господарства	5,66
19	Херсонський державний аграрний університет	3,93
20	Національний університет біоресурсів і природокористування України	10,3
21	Південна філія НУБіП " Кримський агротехнологічний університет "	2,61
22	Бережанський агротехнічний інститут НУБіП	1,95
23	Ніжинський агротехнічний інститут НУБіП	1,26

Згрупувати вищі навчальні заклади за кількістю студентів, утворивши групи с інтервалами: до 2 тис. осіб, 2-4 тис., 4-6 тис., 6-8 тис., більше 8 тис. осіб.

- Визначити: 1) моду та медіану;
 2) верхні та нижні квартилі та децилі;
 3) показники диференціації: коефіцієнти децильної та квартильної диференціації, децильний та квартильний фондові коефіцієнти;
 4) показники концентрації: коефіцієнти локалізації, Лоренца, Джині, Херфіндаля.

Побудувати криву концентрації Лоренца. Зробити висновки.

Завдання 6.5. За даними про рівень річного доходу в різних групах населення міста необхідно визначити модальне й медіанне значення доходу.

Доходи, тис. грн.	Чисельності груп, тис. осіб
До 10	42
10 – 15	90
15 – 25	246
25 – 40	132
40 – 60	60
60 і вище	30

Завдання 6.6. Собівартість одиниці однойменної продукції по підприємствах галузі характеризують наступні показники:

Групи підприємств за собівартістю одиниці продукції, тис. грн.	Число підприємств
16-20	3
20-24	5
24-28	6
28-32	8
32-36	10
36-40	9

Визначити нижній і верхній дециль і розрахувати коефіцієнт децильної диференціації. Зробити висновки.

Тести для самоконтролю.

1. Варіант, який в ряду розподілу має найбільшу частоту або щільність, називається:

- А) мода
- Б) медіана
- В) квартиль
- Г) дециль

2. Значення варіантів, які ділять упорядкований ряд за обсягом на п'ять рівних частин, називаються:

- А) кuartилі
- Б) децилі
- В) квінтилі
- Г) перцентилі

3. Що з нижченаведеного є показником диференціації?

- А) коефіцієнт Лоренца
- Б) коефіцієнт Джині
- В) коефіцієнт локалізації
- Г) фондовий коефіцієнт

4. Значення коефіцієнта Лоренца $K_L = 0,3$ свідчить про:

- А) рівномірний розподіл ознаки
- Б) помірну концентрацію ознаки
- В) низьку концентрацію ознаки
- Г) високу концентрацію ознаки

5. Значення коефіцієнта локалізації $L_i = 1$ свідчить про:

- А) рівномірний розподіл ознаки
- Б) нерівномірний розподіл ознаки
- В) наявність концентрації значень ознаки в i -ій складовій сукупності
- Г) відсутність концентрації значень ознаки в i -ій складовій сукупності

6. Значення коефіцієнта Херфіндаля $K_H = 0,5$ свідчить про:

- А) рівномірний розподіл ознаки
- Б) помірну концентрацію ознаки
- В) низьку концентрацію ознаки
- Г) високу концентрацію ознаки

7. Варіант, який припадає на середину упорядкованого ряду розподілу, називається:

- А) мода
- Б) медіана
- В) кuartиль
- Г) дециль

8. Значення варіантів, які ділять упорядкований ряд за обсягом на чотири рівних частини, називаються:

- А) кuartилі
- Б) децилі
- В) квінтилі
- Г) перцентилі

9. Для визначення якого показника використовується співвідношення між квантилями?

- А) фондовий коефіцієнт
- Б) коефіцієнт локалізації
- В) коефіцієнт подібності
- Г) коефіцієнт диференціації

10. Значення коефіцієнта локалізації $L_i=2$ свідчить про:
- А) наявність концентрації значень ознаки в i -ій складовій сукупності
 - Б) відсутність концентрації значень ознаки в i -ій складовій сукупності
 - В) помірну концентрацію ознаки в i -ій складовій сукупності
 - Г) рівномірний розподіл ознаки
11. Значення коефіцієнта Джині $K_G=0,2$ свідчить про:
- А) рівномірний розподіл ознаки
 - Б) помірну концентрацію ознаки
 - В) низьку концентрацію ознаки
 - Г) високу концентрацію ознаки
12. Нерівномірність розподілів досліджуваної ознаки, що не пов'язана ні з обсягом сукупності, ні з чисельністю окремих груп, - це:
- А) диференціація
 - Б) концентрація
 - В) варіація
 - Г) подібність
13. Значення коефіцієнта Херфіндаля $K_H=900$ свідчить про:
- А) рівномірний розподіл ознаки
 - Б) помірну концентрацію ознаки
 - В) низьку концентрацію ознаки
 - Г) високу концентрацію ознаки
14. Значення варіантів, які ділять упорядкований ряд за обсягом на десять рівних частин, називаються:
- А) кuartилі
 - Б) децилі
 - В) квінтилі
 - Г) перцентилі
15. Значення коефіцієнта подібності $P=0,9$ свідчить про:
- А) однаковість структур
 - Б) високий рівень подібності структур
 - В) середній рівень подібності структур
 - Г) низький рівень подібності структур
16. Значення коефіцієнта Лоренца $K_L=0,5$ свідчить про:
- А) рівномірний розподіл ознаки
 - Б) помірну концентрацію ознаки
 - В) низьку концентрацію ознаки
 - Г) високу концентрацію ознаки
17. Значення коефіцієнта локалізації $L_i=0,9$ свідчить про:
- А) рівномірний розподіл ознаки
 - Б) нерівномірний розподіл ознаки
 - В) наявність концентрації значень ознаки в i -ій складовій сукупності
 - Г) відсутність концентрації значень ознаки в i -ій складовій сукупності

18. Значення коефіцієнта Джині $K_G=0,35$ свідчить про:
- А) рівномірний розподіл ознаки
 - Б) помірну концентрацію ознаки
 - В) низьку концентрацію ознаки
 - Г) високу концентрацію ознаки
19. Мода для рядів з нерівними інтервалами визначається:
- А) за найбільшою щільністю розподілу
 - Б) за найбільшою величиною інтервалу
 - В) за найбільшою частотою
 - Г) за серединним інтервалом
20. Значення варіантів, які ділять упорядкований ряд за обсягом на сто рівних частин, називаються:
- А) квартилі
 - Б) децилі
 - В) квінтилі
 - Г) перцентилі
21. Схожість структур двох об'єктів або одного об'єкта за двома ознаками називається:
- А) варіація
 - Б) подібність
 - В) диференціація
 - Г) концентрація
22. Значення коефіцієнта Лоренца $K_L = 0,2$ свідчить про:
- А) рівномірний розподіл ознаки
 - Б) помірну концентрацію ознаки
 - В) низьку концентрацію ознаки
 - Г) високу концентрацію ознаки
23. Значення коефіцієнта подібності $P=0,1$ свідчить про:
- А) однаковість структур
 - Б) високий рівень подібності структур
 - В) середній рівень подібності структур
 - Г) низький рівень подібності структур
24. Значення коефіцієнта Херфіндаля $K_H = 0,18$ свідчить про:
- А) рівномірний розподіл ознаки
 - Б) помірну концентрацію ознаки
 - В) низьку концентрацію ознаки
 - Г) високу концентрацію ознаки
25. Мода для рядів з рівними інтервалами визначається:
- А) за найбільшою щільністю розподілу
 - Б) за найбільшою частотою
 - В) за найбільшою величиною інтервалу
 - Г) за серединним інтервалом

26. Значення коефіцієнта Лоренца $K_L = 0$ свідчить про:

- А) рівномірний розподіл ознаки
- Б) помірну концентрацію ознаки
- В) низьку концентрацію ознаки
- Г) високу концентрацію ознаки

27. Значення коефіцієнта Джині $K_G = 0,5$ свідчить про:

- А) рівномірний розподіл ознаки
- Б) помірну концентрацію ознаки
- В) низьку концентрацію ознаки
- Г) високу концентрацію ознаки

28. Який показник характеризує зосередження обсягу явища у окремих групах одиниць сукупності і розраховується для кожної складової сукупності?

- А) фондовий коефіцієнт
- Б) коефіцієнт диференціації
- В) коефіцієнт локалізації
- Г) коефіцієнт подібності

29. Для визначення якого показника використовується співвідношення середніх рівнів ознак в групах з максимальними та мінімальними їх значеннями?

- А) фондовий коефіцієнт
- Б) коефіцієнт диференціації
- В) коефіцієнт локалізації
- Г) коефіцієнт подібності

30. Значення коефіцієнта подібності $P = 1$ свідчить про:

- А) однаковість структур
- Б) високий рівень подібності структур
- В) середній рівень подібності структур
- Г) низький рівень подібності структур

31. Розшарування сукупності за рівнем досліджуваної ознаки - це:

- А) диференціація
- Б) концентрація
- В) варіація
- Г) подібність

7. ВИБІРКОВИЙ МЕТОД

7.1 Поняття вибіркового методу (спостереження)

Статистичне дослідження може здійснюватися за даними несучільного спостереження, основна мета якого складається в одержанні характеристик досліджуваної сукупності по обстеженій її частині.

Вибіркове спостереження – таке несучільне спостереження, при якому статистичному дослідженню підлягають одиниці сукупності, відібрані випадковим способом.

Вибірковий метод - сукупність математичних засобів та обґрунтувань, які використовують при застосуванні вибіркового спостереження. При вибіркового методі обстеженню підлягає порівняно невелика частина всієї досліджуваної сукупності (звичайно до 5-10%, рідше до 15-25%).

Вибірковий метод має ряд *переваг*:

- менша вартість: витрати на отримання даних лише відносно невеликої частини всієї сукупності менше, ніж при суцільній перепису;

- коротші терміни: з тих же причин дані вибіркового обстеження можна зібрати і узагальнити швидше, ніж при суцільному перепису, що особливо важливо, коли відомості потрібні терміново;

- ширша область застосування: при деяких видах обстежень для збору даних необхідно залучити висококваліфікований персонал або скористатися спеціальним обладнанням; як правило, і те й інше обмежена. У цих випадках суцільне обстеження неможливо, а вибіркові обстеження дають велику можливість отримувати відомості найрізноманітнішого характеру;

- більша достовірність - це твердження може здатися парадоксальним, проте, якщо загальний обсяг роботи менше, то можна залучити більш кваліфікований персонал, краще його підготувати, більш ретельно контролювати проведення обстеження та обробку його результатів. Тому вибіркове обстеження може дати більш достовірні відомості, ніж відповідне суцільне обстеження.

Таким чином, значення вибіркового методу полягає в тому, що при мінімальній чисельності обстежуваних одиниць проведення дослідження здійснюється в більш короткий термін і з мінімальними витратами праці й коштів. Це підвищує оперативність статистичної інформації, зменшує помилки реєстрації.

У здійсненні ряду досліджень вибірковий метод є єдино можливим, наприклад, при контролі якості продукції (товару), якщо перевірка супроводжується знищенням або розкладанням на складові частини обстежуваних зразків (визначення цукристості фруктів, клейковини печеного хліба, встановлення носкості взуття, міцності тканин на розрив і т.д.).

Загальна сукупність одиниць, з яких здійснюється відбір називається **генеральною сукупністю**.

Сукупність відібраних для дослідження одиниць називається **вибірковою сукупністю** або просто **вибіркою**.

Основні характеристики генеральної і вибіркової сукупності представлені в табл. 7.1.

Таблиця 7.1

Основні характеристики та символіка вибіркового методу

Характеристика	Генеральна сукупність	Вибіркова сукупність
Обсяг сукупності (чисельність одиниць)	N	n
Чисельність одиниць, які мають досліджувану ознаку	M	m
Частка одиниць, які мають досліджувану ознаку	$P = \frac{M}{N}$	$w = \frac{m}{n}$
Середній розмір ознаки	$\bar{x} = \frac{\sum x_i}{N}$	$\tilde{x} = \frac{\sum x_i}{n}$
Дисперсія кількісної ознаки	$\sigma_{\bar{x}}^2 = \frac{\sum (x_i - \bar{x})^2}{N}$	$\sigma_{\tilde{x}}^2 = \frac{\sum (x_i - \tilde{x})^2}{n}$
Дисперсія частки	$\sigma_p^2 = P(1 - P)$	$\sigma_w^2 = W(1 - W)$

Проведення вибіркового спостереження складається з ряду послідовних етапів, представлених на рис. 7.1.

7.2 Основні способи формування вибіркової сукупності і статистична оцінка вибірових характеристик

Для того щоб на основі вибірових характеристик можна було робити висновок про властивості генеральної сукупності, вибірка має бути репрезентативною (представницькою), тобто, вона повинна правильно представляти пропорції генеральної сукупності. Репрезентативність вибірки може бути забезпечена тільки при об'єктивності відбору даних і кожний елемент вибірки буде відібраним випадково, якщо всі елементи мають однакову ймовірність потрапити до вибірки.

Вибіркова сукупність формується за принципом масових імовірнісних процесів без яких би то не було виключень від прийнятої схеми відбору; необхідно забезпечити відносну однорідність вибіркової сукупності або її розділення на однорідні групи одиниць.

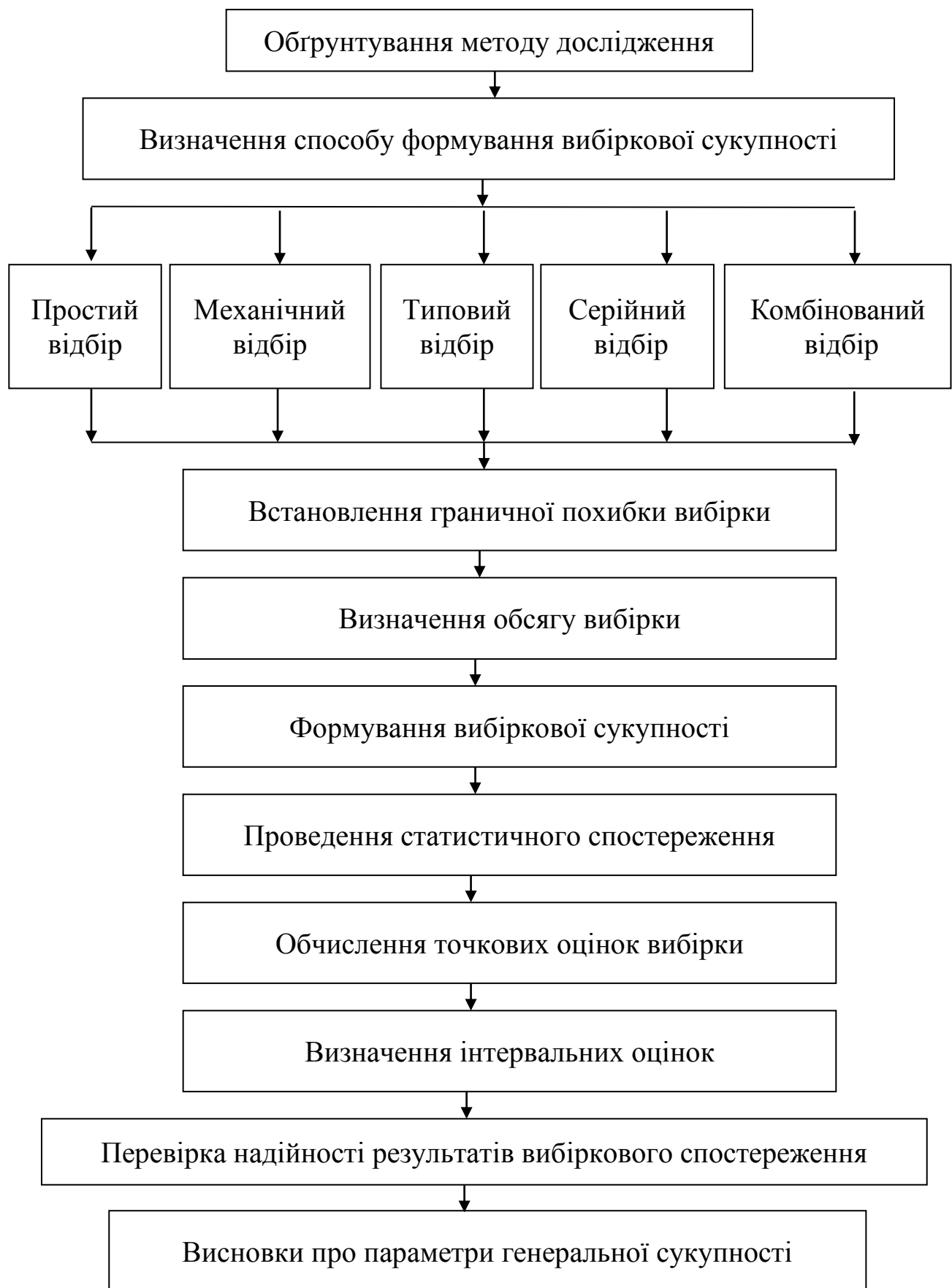


Рис. 7.1. Послідовність етапів вибіркового спостереження

При формуванні вибіркової сукупності має бути дано чітке визначення одиниці відбору. Бажаним є приблизно однаковий розмір одиниць відбору, причому результати будуть тим точніше, чим менше одиниця відбору.

Різноманіття видів, методів та способів вибіркового спостереження обумовлено особливостями розв'язуваних проблем, своєрідністю генеральних сукупностей і оптимальним за конкретних умов вибором організаційно-методологічних принципів обстеження.

Види відбору передбачають участь у вибіркового спостереженні одиниць генеральної сукупності або їх груп.

Методи відбору визначають можливість продовження участі відібраної одиниці у процедурі відбору.

Способи відбору визначають конкретну процедуру вибірки одиниць з генеральної сукупності.

Види відбору:

- А) **індивідуальний** – відбір окремих одиниць генеральної сукупності;
- Б) **груповий** - відбір груп (серій) одиниць;
- В) **комбінований** – комбінація групового та індивідуального відбору.

Методи відбору

А) **повторний** – відбір, при якому відібрана раніше одиниця повертається до генеральної сукупності і може повторно брати участь у виборці (вибіркові дослідження населення, як покупців, пацієнтів, виборців, абітурієнтів і т.п.)

Б) **безповторний** - відбір, при якому відібрана раніше одиниця не повертається до сукупності, з якої здійснюється подальший відбір.

Безповторний відбір дає більш точні результати в порівнянні з повторним, так як при одному і тому ж обсязі вибірки спостереження охоплює більше одиниць генеральної сукупності. Тому він знаходить більш широке застосування в статистичній практиці. І тільки в тих випадках, коли безповторний відбір провести не можна, використовується повторна вибірка (при обстеженні споживчого попиту, пасажирообороту і т.п.).

Основні способи відбору:

1) **власне випадковий (простий) відбір** – кожна одиницю з генеральної сукупності відбирають у вибірку випадково, ненавмисно без якої-небудь схеми або системи, можливо, за допомогою жеребкування або таблиці випадкових чисел.

2) **механічний відбір** - кожна одиницю відбирають через рівні проміжки з упорядкованої генеральної сукупності (сукупність населених пунктів – за географічним положенням, працівників, виборців – за алфавітом, однойменні товари – по мірі їх виробництва або надходження).

Крок інтервалу h прямо пропорціональний частці вибірки $\frac{n}{N}$.

Приклад. Якщо частка вибірки дорівнює 0,02, тобто нам потрібно створити 2%-ну вибірку, то крок становитиме: $h = \frac{1}{0,02} = 50$, тобто відбиратиметься кожна 50 одиниця упорядкованої генеральної сукупності. Таким чином, 5%-на вибірка передбачає відбір кожної 20-ї одиниці, 10%-на – кожної 10-ї, 20%-на – кожної 5-ї і т.п.

3) **типовий (районований, розшарований) відбір** – генеральну сукупність поділяють на однорідні групи за певною ознакою (для населення – райони, соціальні, вікові, освітні групи, для підприємств – галузі та підгалузі, форми власності і т.п.), потім з кожної групи випадковим або механічним способом здійснюється відбір певної кількості одиниць, пропорційної обсягу типових груп:

$$n_i = n \frac{N_i}{N},$$

де n_i – обсяг вибірки з i -ї типової групи;

N_i – обсяг i -ї типової групи.

4) **серійний (гніздовий) відбір** – власне випадковим або механічним способом відбираються не окремі одиниці генеральної сукупності, а цілі групи (серії) таких одиниць, у відібраних серіях досліджуються всі одиниці без винятку. В якості серій виступають: упаковки продукції, партії товару, студентські групи, робочі бригади, тощо).

5) **комбінований відбір** – комбінація вищезазначених способів. Наприклад, можна комбінувати типову та серійну вибірки, коли серії відбираються з декількох типових груп. Можлива комбінація серійного та власне випадкового відбору, при якому окремі одиниці відбираються в серії у простому випадковому порядку.

Помилки вибірки

Оскільки вибіркова сукупність не точно відтворює структуру генеральної, то вибіркові характеристики також не збігаються з характеристиками генеральної сукупності.

Розбіжності між характеристиками генеральної та вибіркової сукупності називають **помилками репрезентативності**.

Визначення величини випадкових помилок репрезентативності є одним з головних завдань теорії випадкового методу.

Для узагальненої характеристики помилки вибірки розраховують середню (стандартну) помилку вибірки.

Середня (стандартна) помилка вибірки μ – середнє квадратичне відхилення вибірових середніх або часток від генеральних середніх або часток. Її використовують для визначення можливих відхилень показників вибіркової сукупності від відповідних показників генеральної сукупності.

Формула визначення середньої залежить від методу та способу формування вибіркової сукупності (табл.7.2).

Таблиця 7.2

Середня помилка вибірки для деяких способів формування вибіркової сукупності

Метод Спосіб	Повторний відбір		Безповторний відбір	
	для середньої	для частки	для середньої	для частки
Власне випадковий та механічний відбір	$\mu = \sqrt{\frac{\sigma_{\tilde{x}}^2}{n}}$	$\mu = \sqrt{\frac{w(1-w)}{n}}$	$\mu = \sqrt{\frac{\sigma_{\tilde{x}}^2}{n} \left(1 - \frac{n}{N}\right)}$	$\mu = \sqrt{\frac{w(1-w)}{n} \left(1 - \frac{n}{N}\right)}$
Типовий відбір	$\mu = \sqrt{\frac{\bar{\sigma}_{\tilde{x}}^2}{n}}$	$\mu = \sqrt{\frac{w_i(1-w_i)}{n}}$	$\mu = \sqrt{\frac{\bar{\sigma}_{\tilde{x}}^2}{n} \left(1 - \frac{n}{N}\right)}$	$\mu = \sqrt{\frac{w_i(1-w_i)}{n} \left(1 - \frac{n}{N}\right)}$
Серійний відбір	$\mu = \sqrt{\frac{\delta_x^2}{r}}$	$\mu = \sqrt{\frac{\delta_w^2}{r}}$	$\mu = \sqrt{\frac{\delta_x^2}{r} \left(1 - \frac{r}{R}\right)}$	$\mu = \sqrt{\frac{\delta_w^2}{r} \left(1 - \frac{r}{R}\right)}$

де $\sigma_{\tilde{x}}^2$ - дисперсія вибіркової сукупності

де $\bar{\sigma}_{\tilde{x}}^2$ середня з внутрішньогрупових дисперсій типових груп: $\bar{\sigma}_{\tilde{x}}^2 = \frac{\sum \sigma_i^2 n_i}{\sum n_i}$

де δ_x^2 та δ_w^2 - міжсерійна дисперсія середніх або часток; r - кількість відібраних серій; R - загальна кількість серій

$$\delta_x^2 = \frac{\sum (\tilde{x}_i - \tilde{x})^2}{r}$$

$$\delta_w^2 = \frac{\sum (w_i - w)^2}{r}$$

де \tilde{x}_i та w_i - відповідно середня та частка i -ї серії; \tilde{x} та w - загальна середня та частка по всій вибірковій сукупності

Однак стверджувати, що генеральні показники не виходять за межі середньої помилки вибірки можна не з абсолютною вірогідністю, а лише з певним ступенем імовірності. Тому, поряд із стандартною помилкою, визначають граничну помилку вибірки.

Гранична помилка вибірки - максимально можлива помилка для прийнятної імовірності p , що дає змогу визначити, в яких межах знаходиться величина генеральної середньої або частки:

$$\Delta = t\mu,$$

де μ – середня (стандартна) помилка вибірки;

t - коефіцієнт довіри (надійності) для заданого значення імовірності, який указує на імовірність того, на яку величину генеральна середня або частка буде відрізняться від вибіркової середньої або частки. Коефіцієнт довіри обчислені й знаходяться в таблиці інтегральних функцій нормального розподілу p (Додаток А).

Приклад. Для імовірності $p = 0,683$ коефіцієнт довіри $t=1$, тобто з імовірністю 0,683 (в 68,3% випадків) різниця вибірковою та генеральною середньою не перевищить однієї величини стандартної помилки вибірки ($\pm\mu$).

Для імовірності $p=0,954$ коефіцієнт довіри $t=2$, тобто в 95,4% випадків різниця вибірковою та генеральною середньою не перевищить подвійної величини стандартної помилки вибірки ($\pm 2\mu$).

Оцінки результатів вибіркового спостереження

При застосуванні вибіркового методу завжди відбувається погашення особливостей окремих одиниць генеральної сукупності. Саме тому передбачається невідповідність параметрів генеральної сукупності параметрам вибіркової, тобто наявність великих або менших помилок спостереження. Щоб виключити таку невідповідність параметри генеральної сукупності зазвичай представляють не за допомогою окремого значення, а у вигляді меж інтервалу, в межах якого можуть відбуватися коливання параметрів.

У процесі вибіркового спостереження розглядають два типи **вибіркових оцінок**:

- **точкові** – характеризує значення параметра, обчислене на основі даних вибірки (вибіркова середня \tilde{x} або вибіркова частка w)

- **інтервальні** - інтервал значень параметра для даної імовірності (довірчий інтервал).

Межі довірчого інтервалу визначаються на основі точкової оцінки та граничної помилки вибірки:

- для середньої: $\tilde{x} - \Delta_{\tilde{x}} \leq \bar{x} \leq \tilde{x} + \Delta_{\tilde{x}}$

- для частки: $w - \Delta_w \leq P \leq w + \Delta_w$

Задача 7.1. Під час простої безповторної вибірки з 1000 працівників агрохолдингу було відібрано 100 осіб для вивчення середнього стажу їх роботи (років):

Стаж роботи	До 5	5-10	10-15	15-20	20-25	25 і більше
Кількість працівників	5	10	20	30	25	10

З імовірністю 0,954 визначити: 1) межі середнього стажу працівника; 2) частку працівників із стажем понад 20 років та її границі в генеральній сукупності

Розв'язання

1) а) Вибіркова середня для розрахунку дисперсії:

$$\tilde{x} = \frac{\sum xf}{\sum f} = \frac{2,5*5 + 7,5*10 + 12,5*20 + 17,5*30 + 22,5*25 + 27,5*10}{100} = 17 \text{ років}$$

б) Вибіркова дисперсія:

$$\sigma_{\tilde{x}}^2 = \frac{\sum (x - \tilde{x})^2 f}{\sum f} = \frac{(2,5 - 17)^2 * 5 + (7,5 - 17)^2 * 10 + (12,5 - 17)^2 * 20 + (17,5 - 17)^2 * 30 + (22,5 - 17)^2 * 25 + (27,5 - 17)^2 * 10}{100} = 42,25$$

в) Середня помилка вибірки:

$$\mu = \sqrt{\frac{\sigma_{\tilde{x}}^2}{n} \left(1 - \frac{n}{N}\right)} = \sqrt{\frac{42,25}{100} \left(1 - \frac{100}{1000}\right)} = 0,62 \text{ (року)}$$

г) Гранична помилка вибірки:

З таблиці нормального розподілу (Додаток А) визначаємо для імовірності $p=0,954$, що $t=2$.

$$\Delta_x = t\mu = 2*0,62 = 1,24 \text{ роки}$$

д) Довірчий інтервал: $\tilde{x} - \Delta_{\tilde{x}} \leq \bar{x} \leq \tilde{x} + \Delta_{\tilde{x}}$

$$17 - 1,24 \leq \bar{x} \leq 17 + 1,24 \quad \text{або} \quad 15,76 \leq \bar{x} \leq 18,24$$

Отже, середній стаж працівників агрохолдингу у 95,4 % буде не меншим від 15,76 і не більшим, ніж 18,24 років.

2) а) Частка працівників із стажем понад 20 років дорівнює:

$$w = \frac{25 + 10}{100} = 0,35$$

б) Середня помилка вибірки:

$$\mu = \sqrt{\frac{w(1-w)}{n} \left(1 - \frac{n}{N}\right)} = \sqrt{\frac{0,35(1-0,35)}{100} \left(1 - \frac{100}{1000}\right)} = 0,045$$

в) Гранична помилка вибірки:

$$\Delta_w = t\mu = 2 * 0,045 = 0,09$$

д) Довірчий інтервал: $W - \Delta_w \leq P \leq W + \Delta_w$

$$0,35 - 0,09 \leq P \leq 0,35 + 0,09 \quad \text{або} \quad 0,26 \leq P \leq 0,44$$

З ймовірністю 95,4% можна стверджувати, що частка працівників зі стажем понад 20 років знаходиться в межах від 26 до 44%.

Задача 7.2. З метою оцінки втрат по причині тимчасової непрацездатності проведено 10%-ний безповторний типовий відбір робітників агрофірми. Визначити межі середнього терміну тимчасової непрацездатності кожного з робітників підприємства з ймовірністю 0,866.

Галузь	Всього робітників, осіб	Обстежено, осіб	Кількість днів тимчасової непрацездатності	
			середня	дисперсія
Рослинництво	140	14	12	25
Тваринництво	100	10	18	49
Переробна	80	8	15	16

Розв'язання

1) Вибіркова середня

$$\tilde{x} = \frac{\sum x_i n_i}{\sum n_i} = \frac{18 * 10 + 12 * 14 + 15 * 8}{10 + 14 + 8} = 14,6 (\text{днів})$$

2) Середня з внутрішньогрупових дисперсій

$$\bar{\sigma}_{\tilde{x}}^2 = \frac{\sum \sigma_i^2 n_i}{\sum n_i} = \frac{49 * 10 + 25 * 14 + 16 * 8}{10 + 14 + 8} = 30,25$$

в) Середня помилка вибірки:

$$\mu = \sqrt{\frac{\bar{\sigma}_{\tilde{x}}^2}{n} \left(1 - \frac{n}{N}\right)} = \sqrt{\frac{30,25}{32} \left(1 - \frac{32}{320}\right)} = 0,92$$

3) Гранична помилка вибірки:

З таблиць нормального розподілу визначаємо для ймовірності $p=0,866$, що $t=1,5$

$$\Delta_{\tilde{x}} = t\mu = 1,5 * 0,92 = 1,38$$

4) Довірчий інтервал: $\tilde{x} - \Delta_{\tilde{x}} \leq \bar{x} \leq \tilde{x} + \Delta_{\tilde{x}}$

$$14,6 - 1,38 \leq \bar{x} \leq 14,6 + 1,38 \quad \text{або} \quad 13,22 \leq \bar{x} \leq 15,98 \text{ днів}$$

З ймовірністю 86,6% можна стверджувати, що середній термін тимчасової непрацездатності одного робітника в цілому по агрофірмі знаходиться у межах від 13,22 до 15,98 днів.

Задача 7.3. Для перевірки фахових знань працівників консалтингової служби з 14 відділів випадковим чином відібрано чотири. Результати тестування:

Підрозділ	Кількість працівників, осіб	Кількість працівників з позитивним результатом тестування, осіб
1	13	11
2	12	10
3	15	14
4	10	8
Разом	50	43

З імовірністю 0,997 встановити межі частки працівників служби, що відповідають фаховим вимогам.

Розв'язання

Відразу визначимо спосіб відбору як серійний безповторний.

Далі відмітимо, що загальна кількість серій (загальна кількість відділів служби) $R=14$, а кількість відібраних серій (відібраних відділів служби) $r=4$.

1) Частка персоналу, що успішно витримала тестування у кожній з обстежених груп:

$$w_1 = \frac{11}{13} = 0,846, \quad w_2 = \frac{10}{12} = 0,833, \quad w_3 = \frac{14}{15} = 0,933, \quad w_4 = \frac{8}{10} = 0,8$$

2) Вибіркова частка працівників з позитивним результатом тестування:

$$w = \frac{43}{50} = 0,86$$

3) Міжсерійна дисперсія:

$$\delta_w^2 = \frac{\sum (w_i - w)^2}{r} = \frac{(0,846 - 0,86)^2 + (0,833 - 0,86)^2 + (0,933 - 0,86)^2 + (0,8 - 0,86)^2}{4} = 0,0025$$

4) Середня помилка вибірки:

$$\mu = \sqrt{\frac{\delta_w^2}{r} \left(1 - \frac{r}{R}\right)} = \sqrt{\frac{0,0025}{4} \left(1 - \frac{4}{14}\right)} = 0,021$$

5) Гранична помилка вибірки:

З таблиць нормального розподілу визначаємо для імовірності $p=0,997$, що $t=3$.

$$\Delta_w = t\mu = 3 * 0,021 = 0,063$$

6) Довірчий інтервал $W - \Delta_w \leq P \leq W + \Delta_w$

$$0,86 - 0,063 \leq P \leq 0,86 + 0,063 \quad \text{або} \quad 0,797 \leq P \leq 0,923$$

З імовірністю 0,997 можна стверджувати, що фаховим вимогам відповідає не менше ніж 79,7% і не більше від 92,3% працівників консалтингової служби.

7.3 Визначення необхідного обсягу вибірки та поширення її характеристик на генеральну сукупність

При організації вибіркового спостереження, особливо, коли воно проводиться вперше, важливо правильно встановити чисельність одиниць вибіркової сукупності. Якщо спостереженню підлягає недостатня кількість одиниць, то знайдені результати будуть неточні. Якщо відбирається занадто велика кількість одиниць, це призводить до зайвих затрат праці і коштів, втрат продукції.

Чисельність вибіркової сукупності залежить від способу відбору одиниць, рівня варіації досліджуваної ознаки, розміру граничної помилки вибірки, а також рівня імовірності, з якою потрібно гарантувати результати вибіркового спостереження (табл. 7.3).

Таблиця 7.3

Необхідний обсяг вибірки для деяких способів формування вибіркової сукупності

Метод Спосіб	Повторний відбір		Безповторний відбір	
	для середньої	для частки	для середньої	для частки
Власне випадковий та механічний відбір	$n = \frac{t^2 \sigma_{\bar{x}}^2}{\Delta_x^2}$	$n = \frac{t^2 w(1-w)}{\Delta_w^2}$	$n = \frac{t^2 \sigma_{\bar{x}}^2 N}{\Delta_x^2 N + t^2 \sigma_{\bar{x}}^2}$	$n = \frac{t^2 w(1-w)N}{\Delta_w^2 N + t^2 w(1-w)}$
Типовий відбір	$n = \frac{t^2 \overline{\sigma_{\bar{x}}^2}}{\Delta_x^2}$	$n = \frac{t^2 \overline{w(1-w)}}{\Delta_w^2}$	$n = \frac{t^2 \overline{\sigma_{\bar{x}}^2} N}{\Delta_x^2 N + t^2 \overline{\sigma_{\bar{x}}^2}}$	$n = \frac{t^2 \overline{w(1-w)}N}{\Delta_w^2 N + t^2 \overline{w(1-w)}}$
Серійний відбір	$n = \frac{t^2 \delta_{\bar{x}}^2}{\Delta_x^2}$	$n = \frac{t^2 w_r(1-w_r)}{\Delta_w^2}$	$n = \frac{t^2 \delta_{\bar{x}}^2 R}{\Delta_x^2 R + t^2 \delta_{\bar{x}}^2}$	$n = \frac{t^2 w_r(1-w_r)R}{\Delta_w^2 R + t^2 w_r(1-w_r)}$

При визначенні необхідної чисельності вибірки дисперсія є невідомою. Для її наближеного знаходження можна скористатися одним з наведених нижче способів:

- 1) скористатися співвідношенням:

$$\sigma = \frac{x_{\max} - x_{\min}}{6}$$

- 2) для частки взяти максимальне значення дисперсії $\sigma^2=0,25$
- 3) провести пробне обстеження і знайти:

$$\sigma^2 = \frac{\sum (x - \bar{x}_{\text{проб}})^2}{n_{\text{проб}} - 1}$$

4) скористатися значенням дисперсії минулих спостережень

5) використати нормативне значення

Задача 7.4. Для визначення середньої довжини деталі потрібно провести вибіркоче обстеження способом випадкового повторного відбору. Визначити кількість деталей, яку потрібно відібрати, щоб гранична помилка вибірки не перевищувала 2 мм з імовірністю 0,987 при середньому квадратичному відхиленні 5 мм (помилка та середнє квадратичне відхилення надані виходячи з технічних нормативів)?

Розв'язання.

Для $p=0,987$, $t=2,5$

$$n = \frac{t^2 \sigma_x^2}{\Delta_x^2} = \frac{2,5^2 * 5^2}{2^2} = 69,4 \approx 70 (\text{деталей})$$

Задача 7.5. На підприємстві числиться 1000 працівників. З метою виявлення мотивації праці на підприємстві необхідно провести соціологічне дослідження. Яким повинен бути обсяг простої неповторної вибірки, щоб з ймовірністю 0,990 гранична помилка частки не перевищувала 10%?

Розв'язання.

Оскільки дисперсія частки невідома, то приймаємо:

$$\sigma^2 = w(1-w) = 0,25.$$

Загальна кількість працівників на підприємстві $N = 1000$;

гранична помилка $\Delta_w = 10\% = 0,1$;

для $P=0,990$, $t=2,58$.

$$n = \frac{t^2 w(1-w)N}{\Delta_w^2 N + t^2 w(1-w)} = \frac{2,58^2 * 0,25 * 1000}{0,1^2 * 1000 + 2,58^2 * 0,25} = 143 (\text{працівники})$$

Задача 7.6. У фермерських господарствах області 20 000 корів. З них в районі А – 10000, в районі Б – 6000, в районі В – 4000. З метою визначення середньої продуктивності (удійності) планується провести типову неповторну вибірку корів. Яку кількість всього та зокрема по районах треба відібрати, щоб з ймовірністю 0,954 гранична помилка вибірки не перевищила 5 л, якщо з попередніх обстежень відомо, що дисперсія дорівнює 1600?

Розв'язання.

1) Для $P=0,954$, $t=2$

$$n = \frac{t^2 \sigma_x^2 N}{\Delta_x^2 N + t^2 \sigma_x^2} = \frac{2^2 * 1600 * 20000}{5^2 * 20000 + 2^2 * 1600} = 249,6 \approx 250 (\text{корів})$$

2) Необхідно відібрати 250 корів з них: $n_i = n \frac{N_i}{N}$

$$\text{в районі А: } n_A = 250 \frac{5000}{10000} = 125(\text{корів})$$

$$\text{в районі Б: } n_B = 250 \frac{3000}{10000} = 75(\text{корів})$$

$$\text{в районі В: } n_B = 250 \frac{2000}{10000} = 50(\text{корів})$$

Способи поширення даних вибіркового спостереження на генеральну сукупність

1) **спосіб прямого перерахунку** – на основі вибірки розраховують показники обсягу генеральної сукупності, використовуючи вибірку середню або частку. Перерахунок здійснюють:

а) множенням вибіркової середньої на кількість одиниць генеральної сукупності:

Задача 7.7 Вибіркове спостереження молочної продуктивності в особистих селянських господарствах(ОСГ) області показало, що середній річний удій на 1 корову становить 2500 кг, з урахуванням похибки вибірки він коливається у межах від 2480 до 2520. Всього в подібних господарствах області 20000 корів. Визначити валовий надій молока в особистих селянських господарствах області.

Розв'язання.

1) Валовий надій молока в ОСГ області становить:

$$2500 * 20000 = 50000000 \text{ кг або } 50 \text{ тис. т}$$

2) Довірчий інтервал:

$$2480 * 20000 \leq \text{ВН} \leq 2520 * 20000 \quad \text{або} \quad 49,6 \leq \text{ВН} \leq 50,4 \text{ (тис. т)}$$

Загальний валовий надій молока в особистих селянських господарствах області коливається від 49,6 до 50,4 тис. т.

б) співвідношенням чисельностей вибіркової і генеральної сукупностей:

Задача 7.8. При 2%-ному обстеженні якості продукції партії з 5000 одиниць у вибірку потрапило 100 одиниць, з яких 20 одиниць виявилися неякісними. Гранична помилка складає 2 одиниці. Визначити кількість бракованої продукції.

Розв'язання.

1) Співвідношення обсягів вибіркової і генеральної сукупностей:

$$2\% = \frac{100}{5000} = \frac{1}{50}$$

2) Число бракованих одиниць: $20 \cdot 50 \pm 2 \cdot 50$,
тобто $1000 - 100 \leq \text{КБО} \leq 1000 + 100$ або $900 \leq \text{КБО} \leq 1100$

Кількість бракованої продукції в партії коливається в межах від 900 до 1100 одиниць.

2) спосіб поправочних коефіцієнтів – зіставляючи дані вибіркового спостереження із суцільним, обчислюють коефіцієнт $\frac{y_1}{y_0}$, який використовують для внесення поправок у матеріали суцільного спостереження. При цьому застосовується наступна формула:

$$Y_1 = Y_0 \frac{y_1}{y_0}$$

де Y_1 – чисельність сукупності з поправкою;

Y_0 - чисельність сукупності без поправки;

y_0 - чисельність сукупності у контрольних точках за первинними даними;

y_1 - чисельність сукупності у контрольних точках за даними контрольних заходів.

Задача 7.9. За переписом худоби, що є приватною власністю населення району, є 15000 свиней, у тому числі в населених пунктах, де проведено вибірку – 1200 свиней. Внаслідок контрольних обходів у цих самих населених пунктах було обчислено 1215 свиней. Визначити загальну кількість свиней, що є приватною власністю населення району.

Розв'язання.

$$Y_1 = 15000 * \frac{1215}{1200} = 15187 (\text{свиней})$$

Уточнена чисельність свиней, що є приватною власністю населення району, становила 15187 свиней.

7.4 Мала вибірка

Для того щоб вибіркова сукупність правильно відображувала генеральну сукупність, вона повинна охоплювати достатню кількість одиниць спостереження. Чим більше одиниць спостереження включає вибірка, тим правильніший буде висновок про розмір ознаки у генеральній сукупності. Однак на практиці не завжди доцільно одержувати великі вибіркові сукупності. Так, при перевірці якості продукції обмежуються порівняно невеликими за обсягом вибірковими сукупностями.

Вибірки, чисельність яких не перевищує 20-30 одиниць спостереження, називають **малими вибірками**. Невеликий обсяг малої вибірки деякою мірою знижує її точність порівняно із звичайною вибіркою, чисельність якої

перевищує 30 одиниць спостереження. Математична статистика розробила способи, які дають змогу вірогідно оцінювати результати малої вибірки і поширювати їх на генеральну сукупність. При цьому розрахунок середньої і граничної помилок має деякі особливості.

Основи теорії малої вибірки розробив англійський математик-статистик Вільям Госсет (псевдонім Ст'юдент). В. Госсет обґрунтував закон розподілу відхилень вибірових середніх від генеральної середньої для малих вибірок. Згідно з цим законом імовірна оцінка того, що гранична помилка не перевищить t -кратну середню помилку в малих вибірках, залежить не тільки від значення t , а й від обсягу вибірки n . Теоретичне нормоване відхилення для малих вибірок дістало назву **критерію t - Ст'юдента**.

У невеликих за чисельністю сукупностях ($n < 20$) дисперсію в малих вибірках обчислюють за формулою:

$$\sigma_{MB}^2 = \frac{\sum (x - \bar{x})^2}{n}$$

Середня помилка малої вибірки $\mu_{M.B}$ обчислюється за формулою:

$$\mu_{MB} \approx \sqrt{\frac{\sigma_{MB}^2}{n-1}},$$

де $n-1$ — кількість ступенів свободи (вільності) варіації.

Під *ступенем свободи* розуміють кількість варіантів, які можуть приймати довільні значення, не змінюючи їх загальної характеристики.

Гранична помилка малої вибірки $\Delta_{M.B}$ визначається за формулою::

$$\Delta_{M.B} = t\mu_{MB},$$

де t — коефіцієнт довіри, що визначається за спеціальною таблицею «Критичні точки розподілу Ст'юдента, t -розподіл» (Додаток Б) при рівнях імовірності помилки 0,10, 0,05 і 0,01.

Примітка. На практиці при визначенні значень коефіцієнт довіри по критерію t - Ст'юдента найчастіше імовірність помилки приймається рівною 0,05. Рівень імовірності 0,05 показує, що тільки у 5 випадках із 100 значення коефіцієнта довіри t може перевищити зазначену в таблиці відповідну величину.

Задача 7.10. Вибіркове обстеження підприємств-виробників зерна показало, що виробнича собівартість 1 ц зерна в них становить (грн.): 105,5; 100,4; 95,5; 90,7; 110,2; 107,3; 97,5; 100,5; 102,8; 97,9. Визначити середню собівартість одиниці продукції.

Розв'язання

1) Вибіркова середня:

$$\tilde{x} = \frac{105,5 + 100,4 + 95,5 + 90,7 + 110,2 + 107,3 + 97,5 + 100,5 + 102,8 + 97,9}{10} = 100,8(\text{грн})$$

2) Вибіркова дисперсія:

$$\sigma_{MB}^2 = \frac{\sum (x - \bar{x})^2}{n-1} = \frac{(105,5 - 110,8)^2 + (100,4 - 110,8)^2 + (95,5 - 110,8)^2 + (90,7 - 110,8)^2 + (110,2 - 110,8)^2 + (107,3 - 110,8)^2 + (97,5 - 110,8)^2 + (100,5 - 110,8)^2 + (102,8 - 110,8)^2 + (97,9 - 110,8)^2}{10} = 30,63$$

в) Середня помилка малої вибірки:

$$\mu_{MB} \approx \sqrt{\frac{\sigma_{MB}^2}{n-1}} = \sqrt{\frac{30,63}{10-1}} = 1,85(\text{грн.})$$

г) Гранична помилка:

Для ступеню вільності $V = n-1 = 10-1 = 9$ при імовірності помилки $p = 0,05$ критерій Ст'юдента $t = 2,26$.

$$\Delta_{M.B} = t\mu_{M.B} = 2,26 * 1,85 = 4,2$$

Д) Довірчий інтервал: $\bar{x} - \Delta_{\bar{x}} \leq \bar{x} \leq \bar{x} + \Delta_{\bar{x}}$

$$100,8 - 4,2 \leq \bar{x} \leq 100,8 + 4,2 \quad \text{або} \quad 96,6 \leq \bar{x} \leq 105,0$$

Отже, з імовірністю 95% на виробництво 1 ц зерна в середньому витрачається не менше ніж 96,6 грн. і не більше ніж 105 грн.

Завдання для самостійного розв'язання.

Завдання 7.1. В результаті вибіркового обстеження житлових умов жителів міста (населення 100 000 осіб), який здійснено на основі власне випадкової повторної вибірки, був одержаний наступний ряд розподілу:

Загальна (корисна) площа житла, що припадає на 1 особу, м ²	До 5	5-10	10-15	15-20	20-25	25-30	30 і більше
Чисельність жителів	18	150	205	270	210	130	95

Визначити: 1) з імовірністю 0,991 межі середнього розміру загальної площі, що припадає на 1 жителя міста;

2) довірчий інтервал частки осіб, забезпеченість житлом яких складає менше 10 м²;

3) способом прямого перерахунку визначити загальний житловий фонд міста.

Завдання 7.2. З метою дослідження безробіття в регіоні проведена 5% механічна неповторна вибірка, в результаті чого отримано розподіл:

Групи безробітних за тривалістю відсутності роботи, місяців	До 4	4-8	8-12	12-16	16-20	20-24	24 і більше
Кількість безробітних	50	75	130	160	110	60	55

Визначити: 1) з імовірністю 0,867 границі, в яких знаходиться середня тривалість безробіття в регіоні;

2) необхідну чисельність вибірки для визначення середньої тривалості безробіття, щоб з імовірністю 0,972 гранична похибка вибірки не перевищувала 1 місяця.

Завдання 7.3. Для перевірки залишкових знань зі статистики в університеті із 12 груп випадковим чином були відібрані 4 групи, всі студенти яких пройшли тестування. Результати перевірки наведені у таблиці:

Група	Кількість студентів в групі	Кількість студентів, що отримали позитивні результати	Середній результат тестування, балів за 100-бальною шкалою
21ЕП	30	24	73
21ОА	25	18	67
22МК	23	15	63
22ФК	29	22	70

Визначити: 1) з імовірністю 0,954 встановити довірчий інтервал середнього балу тестування студентів;

2) з імовірністю 0,997 виявити частку студентів університету, що мають достатні знання зі статистики.

Завдання 7.4. У господарстві є три отари овець чисельністю відповідно 300, 500 і 200 голів. Для визначення середнього настригу вовни з однієї вівці проведена 10%-на типова вибірка. Попередні результати обробки вибірових даних надані в таблиці:

Характеристики вибірки	1-а отара	2-а отара	3-я отара
Середній настриг вовни з вівці, кг	4,3	4,2	4,5
Середнє квадратичне відхилення, кг	1,23	1,14	1,27

Визначити: 1) з імовірністю 0,683 встановити довірчий інтервал середнього настригу від 1 вівці;

2) якою повинна бути чисельність овець у виборці в цілому і окремо по отарах, якщо гранична помилка не перевищуватиме 100 г?

Завдання 7.5. Виробнича лінія призначена для розливу молока в тару місткістю 2 л. Контрольна перевірка 12 вибірових пляшок принесла такі результати:

Номер пляшки	1	2	3	4	5	11
Об'єм молока, л	1,98	2,003	2,0	2,01	1,99	2,002
Номер пляшки	6	7	8	9	10	12
Об'єм молока, л	2,004	2,001	2,0	1,998	1,999	2,005

Визначити з ймовірністю 0,05 середній об'єм молока в пляшці.

Завдання 7.6. При проведенні обліку комерційних лотків у місті було зареєстровано наступну їх кількість по районах: А – 2000, Б – 1500; В – 750. З ціллю перевірки даних суцільного обліку проведені контрольні обходи частини обстежених районів. В результаті одержані наступні дані:

Район	Зареєстровано при суцільному обліку	Встановлено при контрольному обході
А	400	415
Б	300	310
В	150	160

Визначити кількість комерційних лотків у місті з поправкою на недооблік.

Завдання 7.7. У місті 400000 жителів. За матеріалами обліку міського населення було обстежено 40000 жителів методом випадкового неповторного відбору. У результаті обстеження встановлене, що в місті 25% жителів старше 60 років.

Визначити: 1) з ймовірністю 0,806 межі, у яких перебуває частка жителів у місті у віці старше 60 років;

2) чисельність вибірки за умови, що гранична помилка вибіркової середньої не повинна перевищувати 10 осіб з ймовірністю 0,995 і при середньому квадратичному відхиленні 20 осіб.

Завдання 7.8. При контрольній перевірці якості ковбаси, що поставляється в торговельну мережу, отримані наступні дані:

№ проби	1	2	3	4	5	6	7	8	9	10
Вміст солі, %	4,3	4,2	3,8	4,3	3,7	3,9	4,5	4,4	4,0	3,9

Визначити, в яких межах знаходиться середній вміст солі в ковбасі в генеральній сукупності.

Завдання 7.9. Маємо дані 2%-ної вибірки внесків населення в комерційному банку:

Внесок, тис. грн.	Число вкладників, осіб
до 1	10
1 – 3	25
3 – 9	70
9 – 15	40
15 і більше	8

За цими даними розрахувати: 1) з імовірністю 0,984 межі коливання середнього рівня внеску у всій сукупності;

2) чисельність вибірки, яку необхідно взяти, щоб з імовірністю 0,999 гранична помилка не перевищували 500 грн.

Тести для самоконтролю.

1. Вибіркове спостереження – це:

- А) загальна сукупність одиниць, з яких здійснюється відбір
- Б) сукупність відібраних для дослідження одиниць
- В) сукупність математичних засобів та обґрунтувань, які використовують при застосуванні несущільного обстеження
- Г) несущільне спостереження, при якому статистичному дослідженню підлягають одиниці сукупності, відібрані випадковим способом

2. Вибірковий метод – це:

- А) відбір, коли кожен однуцю відбирають через рівні проміжки з упорядкованої сукупності
- Б) сукупність відібраних для дослідження одиниць
- В) сукупність математичних засобів та обґрунтувань, які використовують при застосуванні вибіркового спостереження
- Г) несущільне спостереження, при якому статистичному дослідженню підлягають одиниці сукупності, відібрані випадковим способом

3. Метод відбору, при якому відібрана раніше одиниця не повертається до сукупності, з якої здійснюється подальший відбір, називається:

- А) загальний
- Б) повторний
- В) стандартний
- Г) без повторний

4. Спосіб відбору, коли кожен однуцю відбирають через рівні проміжки з упорядкованої генеральної сукупності, називають:

- А) власне випадковий
- Б) типовий
- В) механічний
- Г) серійний

5. Вид відбору окремих одиниць генеральної сукупності називається:

- А) індивідуальний
- Б) загальний
- В) груповий
- Г) комбінований

6. Для вивчення вмотивованості праці робітників підприємства відібрано 5 підрозділів, всі робітники яких заповнили відповідні анкети. Визначте спосіб відбору:

- А) власне випадковий
- Б) типовий
- В) механічний
- Г) серійний

7. При визначенні необхідної чисельності вибірки, коли дисперсія невідома, можна для частки взяти максимальне значення дисперсії:

- А) $\sigma^2=0,15$
- Б) $\sigma^2=0,25$
- В) $\sigma^2=0,35$
- Г) $\sigma^2=0,55$

8. Методи відбору:

А) визначають можливість продовження участі відібраної одиниці у процедурі відбору

Б) передбачають відбір окремих одиниць чи груп генеральної сукупності

В) визначають конкретну процедуру вибірки одиниць з генеральної сукупності

Г) визначають необхідний обсяг вибірки

9. Загальна сукупність одиниць, з яких здійснюється відбір називається:

- А) основна сукупність
- Б) генеральна сукупність
- В) вибіркова сукупність
- Г) випадкова сукупність

10. Спосіб відбору, коли власне випадковим способом або механічним відбираються не окремі одиниці генеральної сукупності, а цілі групи, в яких досліджуються всі одиниці без винятку:

- А) власне випадковий
- Б) механічний
- В) типовий
- Г) серійний

11. Середнє квадратичне відхилення вибіркових середніх від генеральних середніх, називається:

- А) помилка репрезентативності

- Б) середня помилка
- В) помилка реєстрації
- Г) гранична помилка

12. Всім бажаючим покупцям супермаркету пропонується протестувати новий сорт сиру. Визначте спосіб відбору:

- А) власне випадковий
- Б) типовий
- В) механічний
- Г) серійний

13. При вибірковому спостереженні коефіцієнт довіри $t=3$ для заданого значення імовірності 0,997 означає, що:

- А) в 99,7% випадків різниця між вибірковою та генеральною середньою не перевищить величини стандартної помилки вибірки
- Б) в 99,7% випадків різниця між вибірковою та генеральною середньою не перевищить величини граничної помилки вибірки
- В) в 99,7% випадків різниця між вибірковою та генеральною середньою не перевищить потрібної величини стандартної помилки вибірки
- Г) в 99,7% випадків різниця між вибірковою та генеральною середньою не перевищить потрібної величини граничної помилки вибірки

14. Сукупність відібраних для дослідження одиниць називається:

- А) вибіркова сукупність
- В) випадкова сукупність
- Б) генеральна сукупність
- Г) основна сукупність

15. Генеральною сукупністю називається:

- А) сукупність відібраних для дослідження одиниць
- Б) загальна сукупність одиниць, з яких здійснюється відбір
- В) сукупність математичних засобів та обґрунтувань, які використовують при застосуванні вибіркового спостереження
- Г) несучільне спостереження, при якому статистичному дослідженню підлягають одиниці сукупності, відібрані випадковим способом

16. Вид відбору серій одиниць генеральної сукупності називається:

- А) загальний
- Б) індивідуальний
- В) груповий
- Г) комбінований

17. Для вивчення задоволення населення послугами житлово-комунального господарства з було випадково відібрано 10 домів, що відносяться до даного ЖКГ, всі жильці яких заповнили відповідні анкети. Визначте спосіб відбору:

- А) власне випадковий
- В) механічний

- Б) типовий
- Г) серійний

18. За формулою $\Delta = t_{\alpha}$ визначається:

- А) стандартна помилка вибірки
- Б) середня помилка вибірки
- В) гранична помилка вибірки
- Г) коефіцієнт довіри

19. Для вивчення вмотивованості праці робітників підприємства з кожного підрозділу за допомогою жеребкування відібрано по 10% робітників, які заповнили відповідні анкети. Визначте спосіб відбору:

- А) власне випадковий
- Б) типовий
- В) механічний
- Г) серійний

20. Теоретичне нормоване відхилення для малих вибірок дістало назву критерію:

- А) Джині
- Б) Херфіндаля
- Б) Лоренца
- Г) Ст'юдента

21. Малими вибірками називають вибірки, чисельність одиниць спостереження в яких не перевищує:

- А) 10
- Б) 30
- В) 50
- Г) 100

22. Спосіб відбору, коли генеральну сукупність поділяють на однорідні групи за певною ознакою, потім з кожної групи випадковим або механічним способом здійснюється відбір певної кількості одиниць, пропорційної обсягу цих груп, називають:

- А) власне випадковий
- В) механічний
- Б) типовий
- Г) серійний

23. Максимально можлива помилка для прийнятної імовірності p , що дає змогу визначити, в яких межах знаходиться величина генеральної середньої, називається:

- А) стандартна помилка вибірки
- Б) середня помилка вибірки
- В) гранична помилка вибірки
- Г) помилка репрезентативності

24. Вибірковою сукупністю називається:

- А) загальна сукупність одиниць, з яких здійснюється відбір
- Б) сукупність відібраних для дослідження одиниць
- В) сукупність математичних засобів та обґрунтувань, які використовують при застосуванні несущільного обстеження
- Г) несущільне спостереження, при якому статистичному дослідженню підлягають одиниці сукупності, відібрані випадковим способом

25. При вибіркового спостереженні коефіцієнт довіри $t=2$ для заданого значення імовірності 0,954 означає, що:

- А) в 68,3% випадків різниця між вибірковою та генеральною середньою не перевищить величини стандартної помилки вибірки
- Б) в 68,3% випадків різниця між вибірковою та генеральною середньою не перевищить величини граничної помилки вибірки
- В) в 68,3% випадків різниця між вибірковою та генеральною середньою не перевищить величини подвійної стандартної помилки вибірки
- Г) в 68,3% випадків різниця між вибірковою та генеральною середньою не перевищить подвійної величини граничної помилки вибірки

26. Телефонна компанія цікавиться думками про якість своїх послуг, для чого обдзвонює кожного 50-го абонента зі списку, складеного в алфавітному порядку. Визначте спосіб відбору:

- А) власне випадковий
- Б) типовий
- В) механічний
- Г) серійний

27. Метод відбору, при якому відібрана раніше одиниця повертається до генеральної сукупності і може повторно брати участь у виборці, називається:

- А) стандартний
- Б) загальний
- В) повторний
- Г) відновлений

28. Спосіб відбору, коли кожен одну одиницю з генеральної сукупності відбирають у за допомогою жеребкування або таблиці випадкових чисел, називають:

- А) власне випадковий
- Б) типовий
- В) механічний
- Г) серійний

29. Для перевірки безпечності молока в пакетах відбирається кожна 50-та коробочка, з якої перевіряється якість молока в кожному з 20 пакетів. Визначте спосіб відбору:

- А) власне випадковий

- Б) типовий
- В) механічний
- Г) серійний

30. Розбіжності між характеристиками генеральної та вибіркової сукупності називають називаються:

- А) помилки вибіркової
- Б) помилки реєстрації
- В) помилки репрезентативності
- Г) помилки генерації

31. Способи відбору:

А) визначають конкретну процедуру вибірки одиниць з генеральної сукупності

Б) визначають необхідний обсяг вибірки

В) визначають можливість продовження участі відібраної одиниці у процедурі відбору

Г) передбачають відбір окремих одиниць чи груп генеральної сукупності

32. Для перевірки безпечності молока в пакетах відбирається по 2 зі 100 пакетів з кожної коробки. Визначте спосіб відбору:

- А) власне випадковий
- Б) типовий
- В) механічний
- Г) серійний

33. Види відбору:

А) передбачають відбір окремих одиниць чи груп генеральної сукупності

Б) визначають конкретну процедуру вибірки одиниць з генеральної сукупності

В) визначають необхідний обсяг вибірки

Г) визначають можливість продовження участі відібраної одиниці у процедурі відбору

34. Для виявлення залишкового рівня знань до тестування залучається кожний другий студент зі списку академічної групи. Визначте спосіб відбору:

- А) власне випадковий
- Б) типовий
- В) механічний
- Г) серійний

35. Формула визначення середньої помилки вибірки залежить від:

- А) методу та способу формування вибіркової сукупності
- Б) виду та способу формування вибіркової сукупності
- В) виду та методу формування вибіркової сукупності

Г) способу формування вибіркової сукупності

36. При визначенні необхідної чисельності вибірки, коли дисперсія невідома, для її наближеного знаходження можна скористатися співвідношенням:

А)
$$\sigma = \frac{x_{\max} - x_{\min}}{6}$$

Б)
$$\sigma = \frac{x_{\max} - x_{\min}}{10}$$

В)
$$\sigma = \frac{6}{x_{\max} - x_{\min}}$$

Г)
$$\sigma = \frac{10}{x_{\max} - x_{\min}}$$

37. Для вивчення задоволення населення послугами житлово-комунального господарства з кожного дому, що відноситься до даного ЖКГ було випадково відібрано по 10% квартир, жильці яких заповнили відповідні анкети. Визначте спосіб відбору:

А) власне випадковий

Б) типовий

В) механічний

Г) серійний

38. Для формування складу учасників забігу, відбирається кожний третій з шеренги учнів, вишикуваної за зростом. Визначте спосіб відбору:

А) власне випадковий

В) механічний

Б) типовий

Г) серійний

8. АНАЛІЗ РЯДІВ ДИНАМІКИ

8.1 Сутність та види рядів динаміки

Суспільні явища безперервно змінюються. Протягом певного часу змінюється їхній розмір, обсяг та структура. Процес розвитку суспільних явищ у часі називається **динамікою**.

Ряд динаміки – розміщені у хронологічній послідовності значення певного статистичного показника.

Ряд динаміки складається з двох *елементів*:

- періодів або моментів часу t
- рівнів ряду y – числових значень показників

За ознакою часу ряди динаміки поділяють на:

- 1) **інтервальні (періодичні)** – показники характеризують величину явища за відповідні періоди часу (місяць, квартал, рік, тощо);
- 2) **моментні** - показники характеризують стан явища на відповідний момент часу (дату).

Показники інтервального ряду можна додавати, в результаті одержимо новий ряд динаміки, кожний показник якого характеризує величину явища за збільшені періоди часу.

За формою подання рівнів розрізняють ряди абсолютних, відносних і середніх величин.

За відстанню між датами або інтервалами часу виділяють динамічні ряди:

- **повні** - ряди, в яких дати реєстрації або закінчення періодів слідує один за одним з рівними інтервалами;
- **неповні** - ряди, в яких не дотримано принципу рівних інтервалів.

Статистичні дані мають бути співставними за територією, повнотою охоплених об'єктів, одиницями виміру, часом реєстрації, цінами, методологією розрахунку.

Порівнянність по території означає, що дані по країнах і регіонах, межі яких змінилися, повинні бути перераховані в старих межах. Порівнянність за повнотою охоплених об'єктів означає умову співставлення сукупностей з рівним числом елементів. Територіальна та об'ємна порівнянність забезпечується змиканням рядів динаміки, при цьому або абсолютні рівні замінюються відносними, або робиться перерахунок в умовні абсолютні рівні. Не виникає особливих складнощів при забезпеченні порівнянності даних за одиницями виміру; вартісна порівнянність досягається системою порівнянних цін.

Числові рівні рядів динаміки повинні бути впорядкованими в часі. Не допускається аналіз рядів з пропусками окремих рівнів, якщо ж такі пропуски неминучі, то їх заповнюють умовними розрахунковими значеннями.

Аналіз рядів динаміки дозволяє отримати характеристики:

- інтенсивності зміни окремих рівнів явища;
- середнього рівня і середньої інтенсивності розвитку явища;
- тенденції закономірності розвитку явища;
- сезонність зміни явища.

8.2 Визначення середнього рівня динамічного ряду

Середній рівень інтервального ряду динаміки з рівними інтервалами визначають за формулою середньої арифметичної простої:

$$\bar{y} = \frac{\sum y}{n},$$

де y – рівень ряду динаміки,
 n – кількість рівнів.

Задача 8.1. Маємо наступні дані про виробництво продукції за I півріччя:

Місяці	Січень	Лютий	Березень	Квітень	Травень	Червень
Виробництво продукції, тис. шт.	250	274	260	300	298	280

Визначити середньомісячний обсяг виробництва.

Розв'язання.

$$\bar{y} = \frac{\sum y}{n} = \frac{250 + 274 + 260 + 300 + 298 + 280}{6} = 277 \text{ (тис. шт.)}$$

Середній рівень моментного ряду динаміки з рівними інтервалами визначають за формулою середньої хронологічної:

$$\bar{y} = \frac{\frac{1}{2}y_1 + y_2 + y_3 + \dots + y_{n-1} + \frac{1}{2}y_n}{n-1}$$

Задача 8.2. Маємо наступні дані про наявність поголів'я ВРХ у господарстві на початок року:

Дата	1.01.09	1.01.10	1.01.11	1.01.12	1.01.13
Поголів'я ВРХ, гол.	200	180	210	190	160

Визначити середнє поголів'я за 5 років.

Розв'язання.

$$\begin{aligned} \bar{y} &= \frac{\frac{1}{2}y_1 + y_2 + y_3 + \dots + y_{n-1} + \frac{1}{2}y_n}{n-1} = \\ &= \frac{\frac{1}{2}200 + 180 + 210 + 190 + \frac{1}{2}160}{5-1} = 190(\text{зол.}) \end{aligned}$$

Середній рівень моментного або інтервального ряду динаміки з нерівними інтервалами визначають за формулою середньої арифметичної зваженої:

$$\bar{y} = \frac{\sum yt}{\sum t},$$

де t – інтервал часу між окремими датами (періодами)

Задача 8.3. Маємо наступні дані про вартість основних засобів підприємства на певні дати:

Дата	1.01.12	1.03.12	1.07.12	1.08.12	1.01.12
Вартість ОВФ, тис. грн.	254	260	268	280	270

Визначити середньорічну вартість основних засобів.

Розв'язання.

$$\bar{y} = \frac{\sum yt}{\sum t} = \frac{254 * 2 + 260 * 4 + 268 * 1 + 280 * 5}{2 + 4 + 1 + 5} = 268 \text{ (тис. грн.)}$$

8.3 Показники інтенсивності рядів динаміки

При вивченні особливостей розвитку соціально-економічних явищ визначають певні характеристики інтенсивності динаміки, розрахунок яких ґрунтується на порівнянні рівнів динамічного ряду.

Показники динаміки необхідні для економічного аналізу процесів і явищ, вони дозволяють визначити, як розвиваються явища, зростають чи зменшуються їх розміри, швидко чи повільно відбуваються зміни.

Під час аналізу рядів динаміки обчислюють і використовують такі аналітичні показники динаміки: абсолютний приріст, темп зростання, темп приросту і абсолютне значення одного процента приросту.

Обчислення цих показників ґрунтується на абсолютному або відносному зіставленні рівнів ряду динаміки.

За базу зіставлення приймають або попередній, або якийсь постійний (зазвичай початковий) рівень ряду динаміки.

Рівень, який є загальною основою для порівняння, називається **базисним** (y_0).

Рівень, який порівнюють з базисним або попереднім, називається **поточний** (y_t).

Якщо кожний рівень порівнюють з одним і тим самим рівнем, що є постійною базою, показники називаються **базисними**.

Якщо кожний рівень порівнюють з попереднім, показники називаються **ланцюговими**.

Способи розрахунку показників інтенсивності рядів динаміки розглянемо і засвоїмо, розв'язуючі відповідну задачу.

Задача 8.4. На основі даних про виробництво продукції на підприємстві у 2008-2012 рр. розрахувати базисні, ланцюгові та узагальнюючі показники інтенсивності в рядах динаміки.

Розв'язання.

Базисні та ланцюгові показники інтенсивності в рядах динаміки виробництва продукції наведені в таблиці 8.1.

Формули та приклади їх розрахунку наведено під вищевказаною таблицею.

Таблиця 8.1

Вихідні та розрахункові дані для визначення показники інтенсивності в рядах динаміки обсягів виробництва продукції

Роки (t)	Виробництво продукції, тис. т (y)	Абсолютний приріст, тис. т (Δ)		Темп зростання, % (ТЗ)		Темп приросту, % (ТП)		Абсолютне значення 1% приросту, тис. т ($\Delta\%$)
		Базисний	Ланцюговий	Базисний	Ланцюговий	Базисний	Ланцюговий	
2008	290	-	-	-	-	-	-	-
2009	310	20	20	106,9	106,9	6,9	6,9	2,9
2010	330	40	20	113,8	106,5	13,8	6,5	3,1
2011	315	25	-15	108,6	95,4	8,6	-4,6	3,3
2012	320	30	5	110,3	101,6	10,3	1,6	3,15

Показники інтенсивності рядів динаміки:

1) **Абсолютний приріст** Δ – різниця рівнів ряду динаміки:

Базисний

$$\Delta_{\text{б}} = y_t - y_0$$

$$\Delta_{\text{б}2009} = 310 - 290 = 20 \text{ (тис. т)}$$

Ланцюговий

$$\Delta_{\text{л}} = y_t - y_{t-1}$$

$$\Delta_{\text{л}2010} = 330 - 310 = 20 \text{ (тис. т)}$$

Сума ланцюгових абсолютних приростів дорівнює кінцевому базисному:

$$\Sigma (y_t - y_{t-1}) = y_n - y_0$$
$$20 + 20 - 15 + 5 = 30$$

2) **Темп зростання** TЗ – відношення рівнів ряду, яке виражається коефіцієнтом або в процентах:

Базисний

$$\text{TЗ}_{\text{б}} = \frac{y_t}{y_0} \text{ або } \text{TЗ}_{\text{б}} = \frac{y_t}{y_0} * 100$$

$$\text{TЗ}_{\text{б}2009} = \frac{310}{290} * 100 = 106,9\%$$

Ланцюговий

$$\text{TЗ}_{\text{л}} = \frac{y_t}{y_{t-1}} \text{ або } \text{TЗ}_{\text{л}} = \frac{y_t}{y_{t-1}} * 100$$

$$\text{TЗ}_{\text{л}2010} = \frac{330}{310} * 100 = 106,5\%$$

Добуток ланцюгових темпів зростання, виражених у коефіцієнтах, дорівнює кінцевому базисному:

$$\frac{y_1}{y_0} * \frac{y_2}{y_1} * \dots * \frac{y_n}{y_{n-1}} = \frac{y_n}{y_0}$$
$$1,069 * 1,065 * 0,954 * 1,016 = 1,103$$

3) **Темп приросту** – показує, на скільки відсотків поточний рівень ряду більший або менший за базу порівняння:

Базисний

$$\text{ТП}_{\text{б}} = \frac{\Delta_t}{y_0} * 100$$

$$\text{ТП}_{\text{б}2009} = \frac{20}{290} * 100 = 6,9\%$$

$$\text{ТП}_{\text{б}} = \frac{y_t - y_0}{y_0} * 100$$

$$\text{ТП}_{\text{б}2009} = \frac{310 - 290}{290} * 100 = 6,9\%$$

$$\text{ТП}_{\text{б}} = \text{TЗ}_{\text{б}} - 100$$

$$\text{ТП}_{\text{б}2009} = 106,9 - 100 = 6,9\%$$

Ланцюговий

$$\text{ТП}_{\text{л}} = \frac{\Delta_t}{y_{t-1}} * 100$$

$$\text{ТП}_{\text{л}2010} = \frac{20}{310} * 100 = 6,5\%$$

$$\text{ТП}_{\text{л}} = \frac{y_t - y_{t-1}}{y_{t-1}} * 100$$

$$\text{ТП}_{\text{л}2010} = \frac{330 - 310}{310} * 100 = 6,5\%$$

$$\text{ТП}_{\text{л}} = \text{TЗ}_{\text{л}} - 100$$

$$\text{ТП}_{\text{л}2010} = 106,5 - 100 = 6,5\%$$

4) **Абсолютне значення 1% приросту** – показує, скільки одиниць абсолютного показника містить 1% темпу приросту.

Примітка. Розраховується тільки ланцюговий показник.

$$A\% = \frac{\Delta}{T\Pi_n}$$

$$A\%_{2010} = \frac{20}{6,5} = 3,1 \text{ (тис. т)}$$

$$A\% = \frac{y_t - y_{t-1}}{y_{t-1}} * 100$$

$$A\%_{2010} = \frac{330 - 310}{310} = 3,1 \text{ (тис. т)}$$

$$A\% = \frac{y_{t-1}}{100}$$

$$A\%_{2010} = \frac{310}{100} = 3,1 \text{ (тис. т)}$$

Узагальнюючі показники інтенсивності динаміки:

1) Середній абсолютний приріст:

- по рівнях ряду:

$$\bar{\Delta} = \frac{y_n - y_0}{n - 1} = \frac{320 - 290}{5 - 1} = 7,5 \text{ (тис. т)}$$

- по ланцюгових показниках:

$$\bar{\Delta} = \frac{\sum \Delta_t}{n - 1} = \frac{20 + 20 - 15 + 5}{5 - 1} = 7,5 \text{ (тис. т)}$$

2) Середній темп зростання:

- по рівнях ряду:

$$\overline{TЗ} = \sqrt[n-1]{\frac{y_n}{y_0}} * 100 = \sqrt[5-1]{\frac{320}{290}} * 100 = 102,5\%$$

- по базисних показниках:

$$\overline{TЗ} = \sqrt[n-1]{TЗ_{\bar{0}_n}} * 100 = \sqrt[5-1]{1,103} * 100 = 102,5\%$$

- по ланцюгових показниках:

$$\begin{aligned} \overline{TЗ} &= \sqrt[n-1]{TЗ_{л1} * TЗ_{л2} * \dots * TЗ_{лn}} * 100 = \\ &= \sqrt[5-1]{1,069 * 1,065 * 0,954 * 1,016} * 100 = 102,5\% \end{aligned}$$

3) Середній темп зростання:

$$\overline{T\Pi} = \overline{TЗ} - 100 = 102,5 - 100 = 2,5\%$$

4) Середнє абсолютне значення 1% приросту:

$$\bar{A}\% = \frac{\bar{\Delta}}{\overline{T\Pi}} = \frac{7,5}{2,5} = 3 \text{ (тис. т)}$$

Таким чином, у 2008-2012 роках середньорічний приріст виробництва продукції становив 7,5 тис. т, тобто щороку виробництво продукції збільшувалося в середньому на 2,5%, а кожний процент цього приросту складав 3 тис. т.

Завдання для самостійного розв'язання.

Завдання 8.1. Чисельність працівників фірми на 01.01.2012 р. становила 50 осіб. Протягом 3-х місяців відбувалися такі зміни:

12.01.2012 р. – прийнято на роботу 4 особи;

16.01. 2012 р. – прийнято на роботу 1 особу;

21.01. 2012 р. – звільнено з роботи 3 особи;

28.01. 2012 р. – прийнято на роботу 2 особи;

08.02. 2012 р. – звільнено з роботи 5 осіб;

15.02. 2012 р. – звільнено з роботи 1 особу;

05.03. 2012 р. - прийнято на роботу 6 осіб;

25.03. 2012 р. - звільнено з роботи 2 особи.

Визначити середню чисельність працівників фірми.

Завдання 8.2 Маємо наступні дані про капітал та прибуток комерційного банку (млн. грн.):

Показники	Минулий рік, квартали				Поточний рік, I квартал
	I	II	III	IV	
Капітал на початок періоду	790	880	945	1050	1200
Прибуток за квартал	200	235	310	345	370

Визначити середньоквартальну вартість капіталу та середній прибуток за квартал.

Завдання 8.3 Маємо наступні дані даними про собівартість 1 ц молока у господарстві за 2008-2012 роки:

Роки	2008	2009	2010	2011	2012
Собівартість 1 ц, грн.	174,5	192,0	236,5	285,8	312,0

Визначити: 1) показники інтенсивності ряду динаміки за базисним і ланцюговим методами;

2) середні показники інтенсивності динаміки.

Зробити висновки.

Завдання 8.4. Є наступні дані про випуск фахівців вищими навчальними закладами міста:

Роки	2008	2009	2010	2011	2012
Число фахівців, осіб	1580	1650	1845	2118	2468

Для аналізу динаміки випуску фахівців вишами визначити:

- 1) ланцюгові абсолютні прирости, темпи зростання й приросту, абсолютне значення одного відсотка приросту;
 - 2) середнє абсолютне значення одного відсотка приросту.
- Зробити висновки.

Завдання 8.5. Є дані про доходи торговельного підприємства за перше півріччя:

Місяці	Січень	Лютий	Березень	Квітень	Травень	Червень
Доходи, тис. грн.	600	690	710	760	740	745

- Визначити: 1) середній рівень доходу підприємства;
- 2) абсолютні прирости, темпи зростання й темпи приросту доходів базисним методом по місяцях;
 - 3) середньомісячний абсолютний приріст і темп приросту доходів за півріччя.
- Зробити висновки.

Завдання 8.7. Дані про чисельність безробітних в країні на кінець кожного року становили:

Рік	1	2	3	4	5	6
Кількість безробітних, тис. осіб	5704,4	6711,9	6732,3	8058,8	8876,1	9000,5

- Визначити: 1) середній рівень безробіття;
- 2) абсолютні прирости, темпи росту і приросту (ланцюгові та базисні) та абсолютне значення 1 % приросту;
 - 3) середній абсолютний приріст, середній темп росту і приросту рівнів динамічного ряду, середнє абсолютне значення 1 % приросту.
- Зробити висновки.

Тести для самоконтролю.

1. Динамікою називається
 - А) процес розвитку суспільних явищ у часі
 - Б) розшарування сукупності за рівнем досліджуваної ознаки
 - В) кількісні зміни величини досліджуваної ознаки в межах однорідної сукупності
 - Г) різниця індивідуальних значень ознаки усередині досліджуваної сукупності

2. Ряд динаміки, в якому показники характеризують величину явища за відповідні періоди часу, називається:

- А) нормальний
- Б) моментний
- В) інтервальний
- Г) стандартний

3. Ряд динаміки –

А) розміщені у хронологічній послідовності значення певного статистичного показника.

Б) найпростіше групування, у якому кожна група характеризується чисельністю одиниць, що потрапили в кожную групу

В) послідовність елементів, що характеризує склад, структуру сукупності за певною ознакою

Г) ряд показників, що утворюється послідовним підсумовуванням частот

4. Середній рівень інтервального ряду динаміки з рівними інтервалами визначають за формулою:

- А) середньої арифметичної простої
- Б) середньої арифметичної зваженої
- В) середньої хронологічної простої
- Г) середньої гармонійної зваженої

5. За відстанню між інтервалами часу ряди динаміки, в яких закінчення періодів слідує один за одним з рівними інтервалами, називають:

- А) повні
- Б) неповні
- В) стандартні
- Г) нестандартні

6. Різниця рівнів ряду динаміки називається:

- А) абсолютний приріст
- Б) темп зростання
- В) темп приросту
- Г) абсолютне значення 1% приросту

7. Елементами ряду динаміки можуть бути:

- А) періоди часу
- Б) моменти часу
- В) рівні ряду
- Г) всі варіанти вірні

8. За відстанню між датами ряди динаміки, в яких дати реєстрації слідує одна за одною з рівними інтервалами; називають:

- А) повні
- Б) неповні
- В) стандартні
- Г) нестандартні

9. Темп зростання становить 95%. Значення темпу приросту дорівнює:

- А) 0,5
- Б) 0,95
- В) 5
- Г) 95

10. Середній рівень моментного ряду динаміки з рівними інтервалами визначають за формулою: середньої арифметичної простої

- А) середньої арифметичної зваженої
- Б) середньої хронологічної простої
- В) середньої гармонійної зваженої
- Г) середньої арифметичної простої

11. Рівень ряду динаміки, який є загальною основою для порівняння, називається:

- А) базисний
- Б) загальний
- В) порівняний
- Г) постійний

12. Якщо кожний рівень ряду динаміки порівнюють з попереднім, показники називають:

- А) базисними.
- Б) ланцюговими
- В) постійними
- Г) порівняними

13. За відстанню між датами або інтервалами часу ряди динаміки, в яких не дотримано принципу рівних інтервалів.

- А) повні
- Б) неповні
- В) стандартні
- Г) нестандартні

14. Середній рівень моментного ряду динаміки з нерівними інтервалами визначають за формулою: середньої арифметичної простої

- А) середньої арифметичної зваженої
- Б) середньої хронологічної простої
- В) середньої гармонійної зваженої
- Г) середньої арифметичної простої

15. Відношення рівнів ряду динаміки називається:

- А) абсолютний приріст
- Б) темп зростання
- В) темп приросту
- Г) абсолютне значення 1% приросту

16. Ряд динаміки, в якому показники характеризують величину явища на відповідну дату, називається:

- А) нормальний
- Б) моментний
- В) інтервальний
- Г) стандартний

17. Темп зростання виражається:

- А) тільки коефіцієнтом
- Б) тільки в процентах
- В) коефіцієнтом або в процентах
- Г) натуральними одиницями

18. Темп приросту становить 90%. Значення темпу зростання дорівнює:

- А) 10
- Б) 90
- В) 190
- Г) 0,9

19. Якщо кожний рівень ряду динаміки порівнюють з одним і тим самим рівнем, що є постійною базою, показники називаються:

- А) базисними.
- Б) ланцюговими
- В) постійними
- Г) порівняними

20. Показник, що показує, на скільки відсотків поточний рівень ряду більший або менший за базу порівняння, називають:

- А) абсолютний приріст
- Б) темп зростання
- В) темп приросту
- Г) абсолютне значення 1% приросту

21. Рівень ряду в 2010 році становив 5000 од., в 2011 році – 6000 од., в 2012 році – 7000 од. Визначте абсолютне значення 1% приросту для 2011 року:

- А) 10
- Б) 50
- В) 60
- Г) 70

22. Рівень ряду динаміки, що порівнюють з попереднім, називають:

- А) поточний
- Б) наступний
- В) порівняний
- Г) основний

23. При визначенні абсолютного значення 1% приросту розраховують:

- А) тільки базисний показник
- Б) тільки ланцюговий показник
- В) базисний та ланцюговий показник
- Г) базисний, ланцюговий та сумарний показник

24. Показник, що показує, скільки одиниць абсолютного показника містить 1% темпу приросту.

- А) абсолютний приріст
- Б) темп зростання
- В) темп приросту
- Г) абсолютне значення 1% приросту

25. Середній рівень інтервального ряду динаміки з нерівними інтервалами визначають за формулою середньої арифметичної простої

- А) середньої арифметичної зваженої
- Б) середньої хронологічної простої
- В) середньої гармонійної зваженої
- Г) середньої арифметичної простої

26. Добуток ланцюгових темпів зростання, виражених у коефіцієнтах, дорівнює:

- А) першому базисному
- Б) першому ланцюговому
- В) кінцевому базисному
- Г) кінцевому ланцюговому

27. Темп зростання становить 105%. Значення темпу приросту дорівнює:

- А) 0,5
- Б) 1,05
- В) 5
- Г) 105

28. Рівень ряду в 2010 році становив 100 грн., в 2011 році – 200 грн., в 2012 році – 150 грн. Визначте абсолютне значення 1% приросту для 2012 року:

- А) 1
- Б) 2
- В) 5
- Г) 50

9. АНАЛІЗ ТЕНДЕНЦІЙ РОЗВИТКУ ТА КОЛИВАНЬ

9.1 Методи визначення основної тенденції розвитку у рядах динаміки

В ході обробки динамічного ряду найважливішим завданням є виявлення основної тенденції розвитку явища і згладжування випадкових коливань.

Тенденція (тренд) – це зміна, що визначає загальний напрямок розвитку явища.

Закономірності зміни явища в часі не проявляються в кожному конкретному рівні ряду. Це пов'язано з дією на явища довготривалих факторів. Тому в статистиці для виявлення закономірності або тенденції розвитку явища використовують певні методи обробки рядів динаміки.

Як показує практика, в одних рядах основна тенденція розвитку проявляється досить чітко на основі інтенсивності розвитку, в інших рядах вона може бути виявлена з використанням спеціальних методів аналізу динаміки. Вибір конкретних методів статистики для цієї мети залежить від характеру вихідної інформації і зумовлюється завданнями аналізу.

Розглянемо *основні методи* перетворення рядів динаміки, що дозволяють виявити і обґрунтувати основну тенденцію в розвитку явища:

1) **Метод укрупнення інтервалів** – полягає в тому, що дані динамічного ряду об'єднують у групи за періодами. Вибір величини нових періодів залежить від конкретних особливостей ряду, зокрема від величини їх в первинному ряді. Нові періоди характеризуються сумарними або середніми величинами.

Задача 9.1. Методом укрупнення періодів необхідно виявити тенденції зміни урожайності озимої пшениці в господарстві на основі даних за 12 років.

Розв'язання.

Для визначення тенденції зміни урожайності пшениці об'єднаємо роки у групи за триріччями та розрахуємо сумарну (яке має суто розрахункове значення) та середню урожайність за укрупненими періодами (табл. 9.1).

Аналіз тенденції показників в укрупнених періодах свідчить про зростання урожайності озимої пшениці в господарстві.

2) **Метод згладжування за допомогою ковзної середньої** – передбачає укрупнення інтервалів за допомогою послідовного зсування та один період (рік, місяць, тиждень) при збереженні сталого інтервалу періоду (триріччя, п'ятиріччя, десятиріччя і т.д.). Теоретичний (вирівняний) ряд значень дає можливість погасити індивідуальні коливання і виявити загальну тенденцію розвитку явищ у вигляді плавної лінії

Таблиця 9.1

Вихідні та розрахункові дані для визначення тенденції розвитку урожайності озимої пшениці методом укрупнення періодів

Періоди	Урожайність, ц/га	Сумарна урожайність за триріччя, ц/га	Середня урожайність за триріччя, ц/га
2001	22,3	64,6	21,5
2002	21,5		
2003	20,8		
2004	22,9	72	24
2005	21,9		
2006	27,2		
2007	26,5	84,7	28,2
2008	29,4		
2009	28,8		
2010	29,6	91,7	30,6
2011	32,7		
2012	29,4		

Задача 9.2. За даними задачі 9.1 виявити тенденції зміни урожайності озимої пшениці в господарстві за допомогою методу ковзної середньої

Розв'язання.

Сумарну та середню урожайність за укрупненими періодами розраховуватимемо для кожного триріччя, кожного разу зсуваючись на один рік (табл. 9.2).

Таблиця 9.2

Вихідні та розрахункові дані для визначення тенденції розвитку урожайності озимої пшениці методом згладжування за допомогою ковзної середньої

Періоди	Урожайність, ц/га	Сумарна урожайність за триріччя, ц/га	Середня урожайність за триріччя, ц/га
2001	22,3	-	-
2002	21,5	64,6	21,5
2003	20,8	65,2	21,7
2004	22,9	65,6	21,9
2005	21,9	72	24
2006	27,2	75,6	25,2
2007	26,5	83,1	27,7
2008	29,4	84,7	28,2
2009	28,8	87,8	29,3
2010	29,6	91,1	30,4
2011	32,7	91,7	30,6
2012	29,4	-	-

Згідно до величин згладжених рівнів динаміки, спостерігається тенденція до поступового збільшення урожайності озимої пшениці в господарстві.

Графічно тенденцію, описану методом ковзної середньої, зображено на рис. 9.1

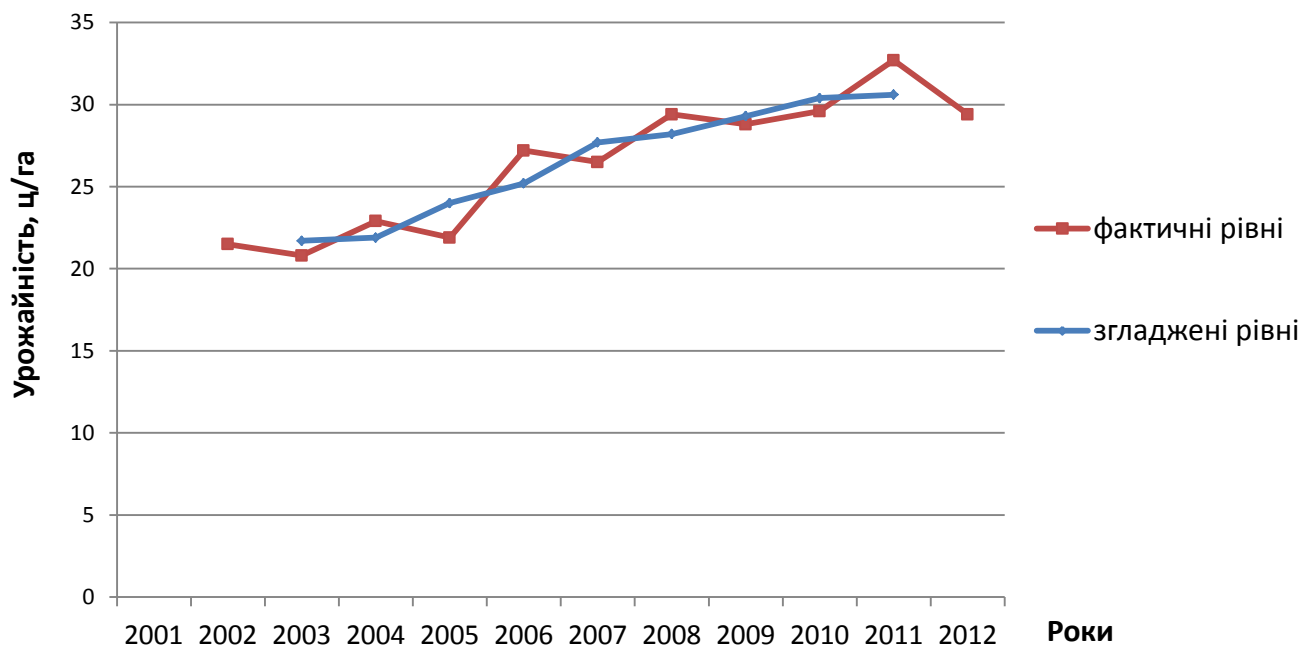


Рис. 9.1. Згладжування урожайності озимої пшениці за допомогою ковзної середньої

3) **Аналітичне вирівнювання способом найменших квадратів** – метод, суть якого полягає в знаходженні такої математичної лінії, ординати точок якої були б найближчі до фактичних значень ряду динаміки, тобто сума квадратів відхилень вирівняних рівнів від фактичних повинна бути мінімальною:

$$(y - y_t)^2 \rightarrow \min$$

Аналітичне вирівнювання способом найменших квадратів можна зробити за рівнянням прямої або параболи, які виражають функціональну залежність рівнів ряду від часу.

1) Рівняння прямої.

Рівняння прямої має вигляд:

$$y_t = a_0 + a_1 t ,$$

де y_t – вирівняний (теоретичний рівень ряду динаміки;

a_0 – вирівняний рівень ряду динаміки, при умові, що $t=0$;

a_1 – середній щорічний абсолютний приріст (або зниження) рівнів вирівняного ряду динаміки;

t – порядковий номер року (умовне позначення часу).

Параметри a_0 та a_1 знаходять розв'язанням системи рівнянь:

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

Щоб спростити розрахунки при аналітичному вирівнюванні рядів динаміки, відлік часу (при непарній кількості періодів) необхідно починати з середини ряду, умовно позначивши центральний період за 0. Тоді вгору від нуля будуть числа -1, -2, -3 і т.д., а вниз - 1, 2, 3 і т.д.

При парній кількості періодів необхідно, щоб умовні значення приймали вигляд -5, -3, -1, 1, 3, 5. Тоді $\sum t = 0$.

Якщо $\sum t = 0$, система матиме вигляд:

$$\begin{cases} a_0 n = \sum y \\ a_1 \sum t^2 = \sum yt \end{cases}$$

тобто параметри a_0 та a_1 обчислюються за формулами:

$$a_0 = \frac{\sum y}{n},$$

$$a_1 = \frac{\sum yt}{\sum t^2}$$

2) Рівняння параболи.

Рівняння параболи II порядку, яка має вигляд:

$$y = a_0 + a_1 t + a_2 t^2$$

де a_2 – середнє прискорення або сповільнення зростання (зниження) рівня досліджуваного явища

Невідомі параметри a_0 , a_1 і a_2 знайдемо, розв'язавши системи нормальних рівнянь:

$$\begin{cases} a_0 n + a_1 \sum t + a_2 \sum t^2 = \sum y \\ a_0 \sum t + a_1 \sum t^2 + a_2 \sum t^3 = \sum yt \\ a_0 \sum t^2 + a_1 \sum t^3 + a_2 \sum t^4 = \sum yt^2 \end{cases}$$

Після спрощення системи параметри a_0 , a_1 , та a_2 обчислюються за формулами:

$$a_0 = \frac{\sum t^4 \sum y - \sum t^2 \sum yt^2}{n \sum t^4 - \sum t^2 \sum t^2},$$

$$a_1 = \frac{\sum yt}{\sum t^2},$$

$$a_2 = \frac{n \sum yt^2 - \sum t^2 \sum y}{n \sum t^4 - \sum t^2 \sum t^2}$$

Задача 9.3. Методом аналітичного вирівнювання способом найменших квадратів необхідно виявити тенденції зміни урожайності озимої пшениці в господарстві на основі даних за 5 років.

Розв'язання.

Дані розрахунків для визначення тенденції зміни урожайності пшениці наведені в таблиці 9.3.

Таблиця 9.3

Вихідні та розрахункові дані для визначення тенденції розвитку урожайності озимої пшениці методом аналітичного вирівнювання способом найменших квадратів

Роки	Урожайність, ц/га	Умовні позначення						Вирівняні значення за рівнянням	
		t	t^2	t^3	t^4	yt	yt^2	прямої	параболи
2008	29,4	-2	4	-8	16	-58,8	117,6	29,2	28,76
2009	28,8	-1	1	-1	1	-28,8	28,8	29,59	29,81
2010	29,6	0	0	0	0	0	0	29,98	30,42
2011	32,7	1	1	1	1	32,7	32,7	30,37	30,59
2012	29,4	2	4	8	16	58,8	117,6	30,76	30,32
<i>Разом</i>	149,9	0	10	0	34	3,9	296,7	149,9	149,9

Проведемо аналітичне вирівнювання способом найменших квадратів за рівнянням прямої та визначимо параметри a_0 та a_1 :

$$a_0 = \frac{\sum y}{n} = \frac{149,9}{5} = 29,98$$

$$a_1 = \frac{\sum yt}{\sum t^2} = \frac{3,9}{10} = 0,39$$

Таким чином, рівняння прямої матиме вигляд:

$$y_t = 29,98 + 0,39t$$

Економічний зміст цього рівняння наступний: в 2010 році вирівняна урожайність озимої пшениці становила 29,98 ц/га, а протягом досліджуваного періоду (2008-2012 роки) вона збільшувалася щорічно в середньому на 0,39 ц/га.

Далі проведемо аналітичне вирівнювання способом найменших квадратів за рівнянням параболи та визначимо параметри a_0 , a_1 , та a_2 :

$$a_0 = \frac{\sum t_4 \sum y - \sum t^2 \sum yt^2}{n \sum t^4 - \sum t^2 \sum t^2} = \frac{34 * 149,9 - 10 * 296,7}{5 * 34 - 10 * 10} = 30,42$$

$$a_1 = \frac{\sum yt}{\sum t^2} = \frac{3,9}{10} = 0,39$$

$$a_2 = \frac{n \sum yt^2 - \sum t^2 \sum y}{n \sum t^4 - \sum t^2 \sum t^2} = \frac{5 * 296,7 - 10 * 149,9}{5 * 34 - 10 * 10} = -0,22$$

Таким чином, рівняння параболи матиме вигляд:

$$y_t = 30,42 + 0,39t - 0,22t^2$$

Економічний зміст рівняння полягає в тому, що в 2010 році теоретична урожайність озимої пшениці становила 30,42 ц/га, протягом досліджуваного періоду вона збільшувалася щороку в середньому на 0,39 ц/га із сповільненням цього росту на 0,22 ц/га.

Графічно тенденцію зміни урожайності озимої пшениці, одержана за допомогою аналітичного вирівнювання способом найменших квадратів зображено на рис. 9.2.

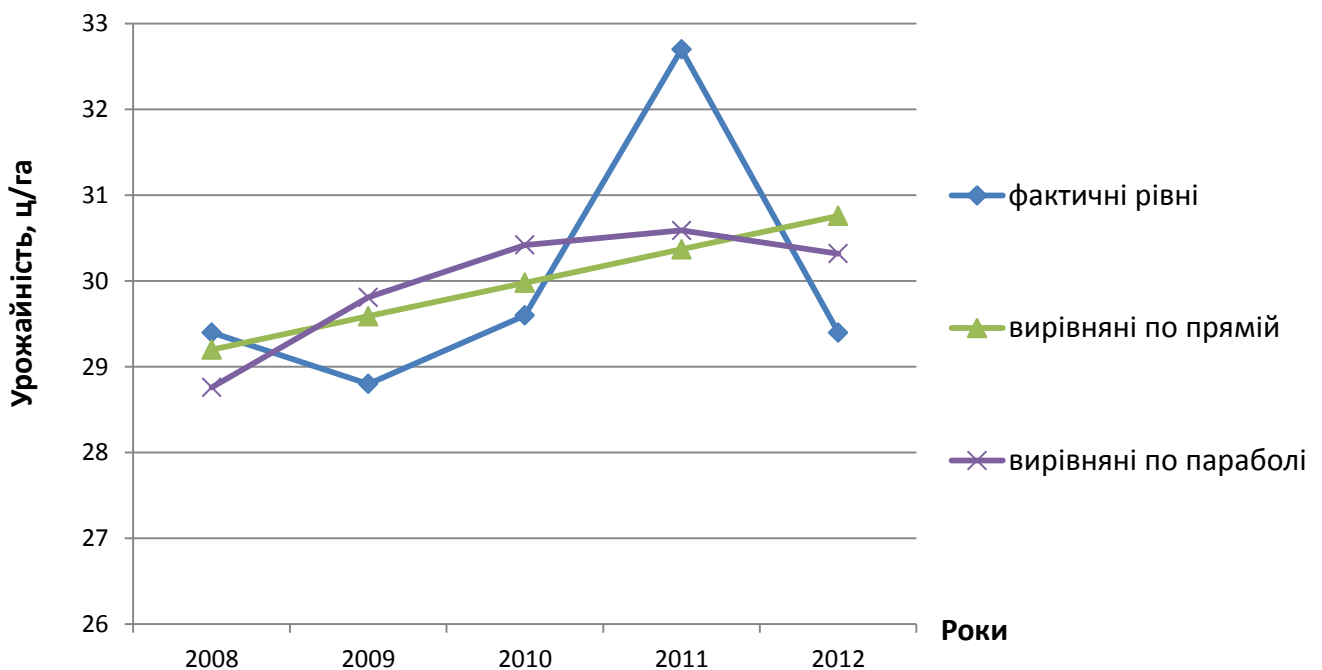


Рис. 9.2. Аналітичне вирівнювання урожайності озимої пшениці способом найменших квадратів.

9.2 Прогнозування

Аналітичне вирівнювання дозволяє не тільки визначити загальну тенденцію зміни явища в досліджуваній період часу, але і провести розрахунки відсутніх рівнів рядів динаміки.

Визначення за наявними даними за певний період часу відсутніх значень ознаки всередині періоду називається **інтерполяцією**.

Знаходження значень ознаки за межами аналізованого періоду називається **екстраполяцією**. Екстраполяція може здійснюватися як в минуле, так і в майбутнє. Тобто ми можемо використати аналітичне вирівнювання способом найменших квадратів для подальшого прогнозування рівня показника - екстраполяції тренду.

Для прогнозування використовується модель тренду, яка найбільш точно виявляє динаміку явищ, тобто відповідає умові:

$$\sum(y-y_t)^2 \rightarrow \min.$$

Задача 9.4. За даними задачі 9.3 зробити прогноз урожайності озимої пшениці на 2013 рік.

Розв'язання.

Для того, щоб виявити, яке рівняння (пряма або парабола) найбільш точно описує тенденцію, необхідно розрахувати квадрати відхилень фактичних рівнів від теоретичних (табл. 9.4).

Таблиця 9.4

Вихідні та розрахункові дані для прогнозування
урожайності озимої пшениці

Роки	Фактична урожайність, ц/га	Вирівняні значення за рівнянням		Відхилення за рівнянням			
				прямої		параболи	
		прямої	параболи	$(y-y_t)$	$(y-y_t)^2$	$(y-y_t)$	$(y-y_t)^2$
2008	29,4	29,2	28,76	0,2	0,04	0,6	0,36
2009	28,8	29,59	29,81	-0,8	0,64	-1,0	1
2010	29,6	29,98	30,42	-0,4	0,16	-0,8	0,64
2011	32,7	30,37	30,59	2,3	5,29	2,1	4,41
2012	29,4	30,76	30,32	-1,4	1,96	-0,9	0,81
<i>Разом</i>	149,9	149,9	149,9	0,0	8,09	0	7,22

Так як сума квадратів відхилень фактичних рівнів від теоретичних порівняно параболи є меншою (табл. 9.4), то рівняння параболи найбільш точніше описує тенденцію розвитку досліджуваного явища. Таким чином, прогнозування виконаємо за рівнянням параболи.

Підставивши відповідний порядковий номер року t в обрану модель тренду одержимо дискретне значення прогнозованого показника.

Урожайність озимої пшениці прогнозується на наступний за останнім рівень ряду динаміки, тобто порядковий номер року (умовне позначення часу) $t=3$.

Таким чином, дискретна оцінка майбутнього рівня урожайності озимої пшениці у 2013 році за рівнянням параболи становитиме:

$$y_{2013} = 30,42 + 0,39*3 - 0,22*9 = 29,6(\text{ц/га})$$

Тобто, в 2013 році урожайність озимої пшениці прогнозується на рівні 29,6 ц/га

Однак статистика рідко зупиняється на визначені дискретних оцінок. В більшості випадків вони виступають в якості вихідних даних для інтервальної оцінки.

Таким чином, наступним кроком прогнозування стане знаходження меж **інтервалу екстраполяції**, в якому буде знаходитись прогнозоване значення досліджуваного явища, по формулі:

$$\hat{y}_t \pm t_{теор} \delta_{\Sigma t},$$

зокрема $\hat{y}_t - t_{теор} \delta_{\Sigma t}$ - нижня межа,

$\hat{y}_t + t_{теор} \delta_{\Sigma t}$ - верхня межа,

де y_t – дискретна оцінка майбутнього рівня ряду динаміки, визначена по моделі тренду;

$t_{теор}$ - коефіцієнт довіри по критерію Ст'юдента (вибирається з відповідної таблиці «Критичні точки розподілу Ст'юдента, t -розподіл» (Додаток Б)); .

$\delta_{\Sigma t}$ - середньоквадратичне відхилення скориговане за числом ступенів вільності.

Середньоквадратичне відхилення скориговане за числом ступенів вільності, в свою чергу, визначається за формулою:

$$\delta_{\Sigma t} = \sqrt{\frac{\sum (y - y_t)^2}{n - m}},$$

де n – кількість рівнів ряду динаміки ;

m – кількість параметрів відповідної моделі тренду;

$n-m$ - ступінь вільності.

$$\delta_{\Sigma t} = \sqrt{\frac{7,22}{5-3}} = 1,9$$

Знаходимо межі інтервалу екстраполяції:

$$\hat{y}_{2013} \pm 4,3 * 1,9$$

$$29,6 - 4,3 * 1,9 = 21,3 \text{ (ц/га)}$$

$$29,6 + 4,3 * 1,9 = 37,9 \text{ (ц/га)}$$

[21,3; 37,9] - інтервал екстраполяції.

Таким чином, урожайність озимої пшениці у 2013 році прогнозується в межах від 21,3 до 37,9 ц/га.

9.3 Статистичне вивчення сезонних коливань

Багато процесів господарської діяльності, торгівлі, сільського господарства й інших сфер людської діяльності піддаються сезонним змінам, наприклад, продаж морозива, споживання електроенергії, виробництво молока, цукру, продаж сільгосппродукції й ін.

Для аналізу рядів динаміки, що піддаються сезонним змінам, використовуються спеціальні методи, що дозволяють установити й описати особливості зміни рівнів ряду.

Перш, ніж використати методи вивчення сезонності, необхідно підготувати дані, приведені до порівняного вигляду, за кілька років спостереження по місяцях або кварталам. Зміни сезонних коливань досліджуються за допомогою коефіцієнтів сезонності. Залежно від існуючих у ряді динаміки тенденцій використовуються різні правила розрахунку таких коефіцієнтів.

1. Ряд динаміки не має загальної тенденції розвитку або вона є невеликою.

Коефіцієнт сезонності знаходять за формулою:

$$K_{s_i} = \frac{\bar{y}_i}{\bar{y}},$$

де \bar{y}_i - середній рівень ряду, отриманий у результаті осереднення рівнів ряду за однойменні періоди часу (наприклад, середній рівень січня за всі роки спостереження);

\bar{y} - загальний середній рівень ряду за увесь час спостереження.

Висновок про наявність або відсутність у ряді динаміки яскраво вираженої тенденції може робитися, наприклад, за допомогою методу укрупнення інтервалів.

Задача 9.5. На підставі даних про укладання шлюбу в місті М за 3 роки спостереження (табл.9.5) дослідити зміну сезонних коливань, розрахувавши коефіцієнти сезонності та побудувавши «сезонну хвилю».

Розв'язання.

При переході від місячних до річних рівнів можна встановити, що тенденція росту дуже незначна.

Загальний середній рівень ряду:

$$\bar{y} = \frac{1940 + 2008 + 2088}{3 * 365} = 5,51 - \text{середнє число шлюбів,}$$

що укладають за один день.

Середній рівень січня:

$$\bar{y}_1 = \frac{y_{1_{2010}} + y_{1_{2011}} + y_{1_{2012}}}{3 * 31} = \frac{173 + 183 + 178}{93} = 5,74 \text{ — середнє число шлюбів за}$$

один день січня.

Таблиця 9.5

Щомісячна кількість укладених шлюбів в місті М за 2010-2012 рр.

Місяці	2010 рік	2011 рік	2012 рік
Січень	173	183	178
Лютий	184	185	179
Березень	167	162	161
Квітень	142	160	184
Травень	137	143	151
Червень	145	150	156
Липень	153	167	177
Серпень	171	173	181
Вересень	143	150	157
Жовтень	162	165	174
Листопад	178	181	193
Грудень	185	189	197
Разом за рік	1940	2008	2088

Аналогічно розраховуються середні рівні лютого, березня й т.д. Результати розрахунків зведені в табл. 9.6.

Таблиця 9.6

Коефіцієнти сезонності укладання шлюбів в місті М

Місяць	\bar{y}_i	$K_{s_i} * 100\%$	Місяць	\bar{y}_i	$K_{s_i} * 100\%$
січень	5,74	104,2	липень	5,34	96,9
лютий	6,45	117,1	серпень	5,64	102,4
березень	5,27	95,6	вересень	5,0	90,7
квітень	5,4	88,0	жовтень	5,39	97,8
травень	4,63	84,0	листопад	6,13	111,3
червень	5,01	91,0	грудень	6,14	111,4

Отримані коефіцієнти сезонності дають оцінку того, як в окремі місяці року кількість укладених шлюбів відхиляється від середнього значення.

Побудований по отриманих коефіцієнтах сезонності лінійний графік – «сезонна хвиля» - наочно покаже сезонність розглянутого процесу (рис. 9.3).

Правила побудови. На вісі ОХ відкладають періоди часу на протязі року, а на вісі ОУ— значення коефіцієнтів сезонності. Відхилення коефіцієнтів сезонності від середнього рівня унаочнюється прямою на рівні 100%.

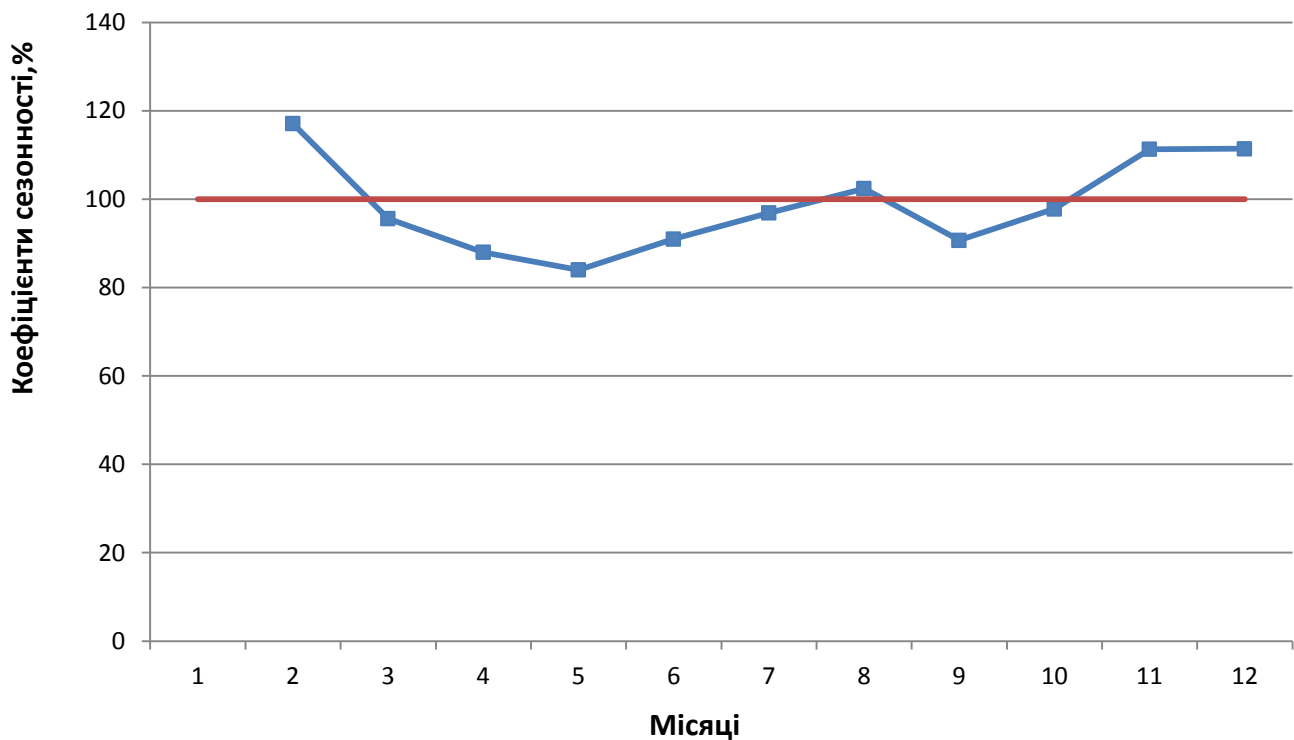


Рис.9.3. Сезонна хвиля укладання шлюбів в місті М

2. Ряд динаміки має загальну тенденцію, і вона визначена або методом ковзного середнього, або методом аналітичного вирівнювання.

Коефіцієнт сезонності обчислюють за формулою:

$$K_{s_i} = \left[\sum \frac{y_i}{y'_i} \right] : n,$$

де y_i — вихідні рівні ряду;

y'_i — рівні ряду, отримані в результаті визначення ковзних середніх для тих же періодів часу, що й вихідні рівні:

i - номер місяця (кварталу), для якого визначається індекс сезонності;

n - число років спостереження за процесом.

У випадку, якщо тенденція розвитку визначалася методом аналітичного вирівнювання, розрахункова формула одержання коефіцієнтів сезонності зовсім аналогічна попередньої, але замість y'_i - рівнів, отриманих методом ковзних середніх, використовуються y'_i , отримані методом аналітичного вирівнювання.

Задача 9.6. На основі вихідних даних про реалізації цукру в продовольчих магазинах міста в 2010 - 2012 р. (т), визначені ковзні середні по трьох рівнях ряду (табл. 9.7). Необхідно дослідити зміну сезонних коливань, розрахувавши коефіцієнти сезонності та побудувавши «сезонну хвилю».

Таблиця 9.7

Обсяги реалізації цукру в продовольчих магазинах міста, т

Місяць	2010 рік		2011 рік		2012 рік	
	Вихідні рівні y_i	Згладжені рівні y'_i	Вихідні рівні y_i	Згладжені рівні y'_i	Вихідні рівні y_i	Згладжені рівні y'_i
Січень	78,9	-	108,6	106,2	129,1	131,3
Лютий	78,1	81,0	107,9	107,8	128,6	129,5
Березень	86,0	87,2	106,8	115,4	130,7	137,4
Квітень	97,5	88,9	132,1	117,3	152,8	141,1
Травень	83,3	88,9	113,0	119,0	139,8	146,7
Червень	86,0	86,6	111,8	116,4	147,4	150,3
Липень	90,6	87,6	124,4	116,8	163,8	152,5
Серпень	86,1	86,0	114,1	115,6	146,3	149,3
Вересень	81,3	90,8	108,4	115,6	137,8	145,4
Жовтень	105,1	94,5	124,0	117,0	152,2	144,4
Листопад	97,2	101,5	118,0	126,2	143,2	150,6
Грудень	102,1	102,6	136,3	128,0	156,5	-

Розв'язання.

На основі вихідних і згладжених рівнів ряду будуються коефіцієнти сезонності.

Так для січня:

$$K_{s_i} = \left[\frac{y_{1,2011}}{y'_{1,2011}} + \frac{y_{1,2012}}{y'_{1,2012}} \right] : 2 = \left[\frac{108,6}{106,2} + \frac{129,1}{131,3} \right] : 2 = 1,0$$

Для лютого:

$$K_{s_i} = \left[\frac{y_{2,2010}}{y'_{2,2010}} + \frac{y_{2,2011}}{y'_{2,2011}} + \frac{y_{2,2012}}{y'_{2,2012}} \right] : 3 = \left[\frac{78,1}{81,0} + \frac{107,9}{107,8} + \frac{128,6}{129,5} \right] : 3 = 0,98 \text{ і т.д.}$$

Коефіцієнти сезонності по місяцях зведені в табл. 9.8.

Таблиця 9.8

Коефіцієнти сезонності продажу цукру

№ місяць	1	2	3	4	5	6	7	8	9	10	11	12
$K_{s_i} * 100\%$	100	98	96	110	95	98	106	96	93	107	95	103

Побудувавши лінійний графік, можна побачити закономірності зміни обсягу продажів цукру по місяцях року (рис. 9.4).

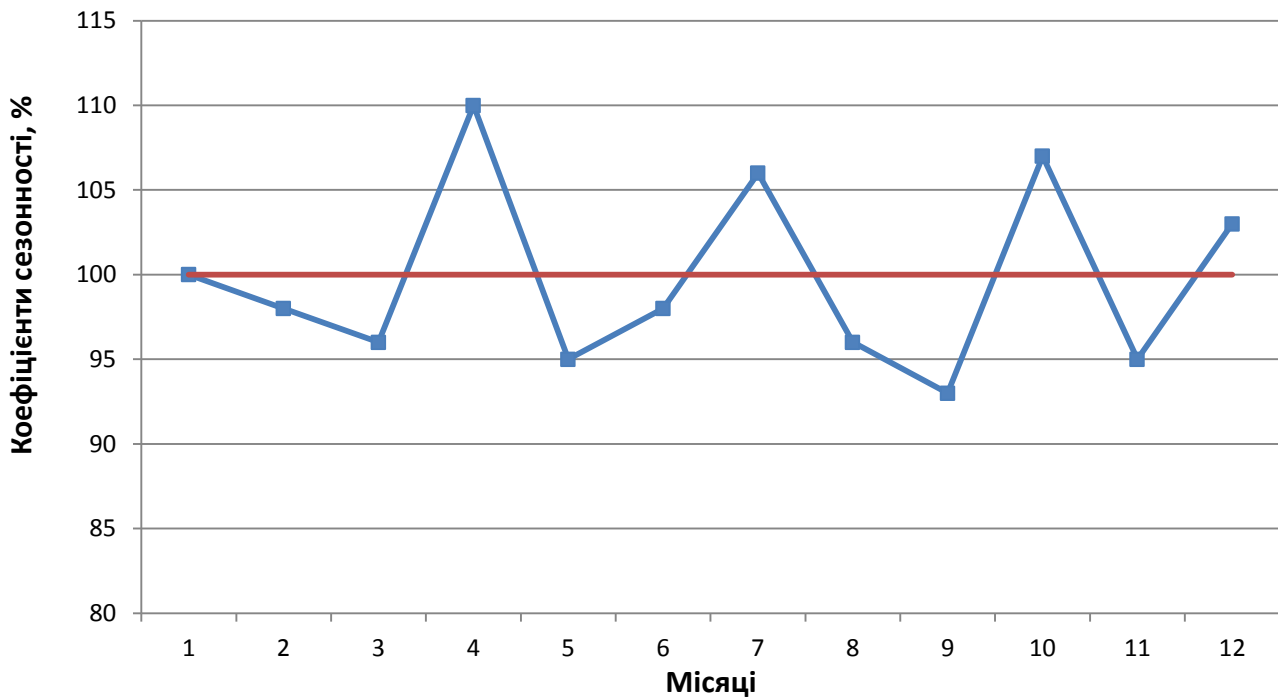


Рис.9.4. Сезонна хвиля продажу цукру

Завдання для самостійного розв'язання.

Завдання 9.1. За даними завдання 8.3:

1) провести згладжування ряду динаміки за допомогою ковзної середньої;

2) описати тенденцію собівартості 1 ц молока лінійним трендом та параболою II порядку;

3) зробити прогноз собівартості 1 ц молока на 2013 рік.

Побудувати графіки фактичної собівартості 1 ц молока та виявлених тенденцій. Зробити висновки.

Завдання 9.2. Є дані про реалізацію телевізорів за 7 років:

Роки	2006	2007	2008	2009	2010	2011	2012
Реалізація телевізорів, тис. шт.	366	310	296	380	337	296	280

Здійснити аналітичне вирівнювання по рівняннях прямої і параболі. Зробити висновки. Побудувати графік фактичних і розрахункових рівнів реалізації телевізорів.

Завдання 9.3. Є дані про реалізацію автомобілів за 6 років:

Роки	2007	2008	2009	2010	2011	2012
Реалізація автомобілів, тис. шт.	200	210	206	180	237	215

Здійснити аналітичне вирівнювання по рівнянню параболі і прогноз на 2015 рік. Зробити висновки.

Завдання 9.4. Маємо наступні дані про пасажирські перевезення залізницею в країні динаміці за 3 роки:

Місяць	Перевезення, тис. пасажирів		
	2010	2011	2012
Січень	526	540	536
Лютий	564	605	557
Березень	646	800	88
Квітень	662	650	663
Травень	927	953	937
Червень	1136	955	1102
Липень	1025	1002	1050
Серпень	1007	976	1012
Вересень	905	860	932
Жовтень	651	655	666
Листопад	554	560	601
Грудень	650	656	539

Визначити наявність сезонних коливань та побудувати сезонну хвилю.

Завдання 9.5 За даними таблиці:

1) описати тенденцію зміни ціни 1 ц зерна за допомогою рівняння прямої лінії;

2) зробити прогноз ціни 1 ц зерна на 2014 рік.

Роки	2008	2009	2010	2011	2012
Середня ціна 1 ц, грн.	80,9	83,0	79,4	98,2	135,3

Побудувати графіки фактичної ціни 1 ц зерна та виявлених тенденцій. Зробити висновки.

Завдання 9.6. Продаж молока та молочних продуктів (тис. т) у регіоні характеризується наступними даними за 4 роки:

Квартали	2009	2010	2011	2012
I	72,0	76,3	77,9	60,7
II	60,1	66,7	65,7	71,0
III	61,7	63,5	66,9	71,7
IV	77,9	76,0	60,0	62,7

Для аналізу внутрішньорічної динаміки реалізації молока та молочних продуктів по кварталах 2009-2012 років визначити коефіцієнти сезонності із застосуванням чотирьохчленної ковзної середньої та побудувати графік сезонної хвилі.

Тести для самоконтролю.

1. Метод визначення основної тенденції розвитку у рядах динаміки, який полягає в тому, що дані динамічного ряду об'єднують у групи за періодами, називається:

- А) укрупнення інтервалів
- Б) згладжування за допомогою ковзної середньої
- В) аналітичне вирівнювання
- Г) екстраполяція тренда

2. При аналізі рядів динаміки параметр a_0 у рівнянні прямої характеризує:

- А) вирівняний рівень ряду динаміки при $t=0$
- Б) порядковий номер року
- В) середній щорічний абсолютний приріст рівнів вирівняного ряду динаміки
- Г) середнє прискорення зміни рівня досліджуваного явища

3. При аналізі динаміки виробництва зерна параметр $a_2=2000$ у рівнянні параболи означає:

- А) середній щорічний абсолютний приріст зерна становив 2000 ц
- Б) середнє прискорення зміни валового збору зерна становило 2000 ц
- В) середнє сповільнення зміни валового збору зерна становило 2000 ц
- Г) параметр відсутній

4. Прогнозування здійснюється за рівнянням прямої на основі 20 рівнів ряду динаміки. Визначте ступінь вільності:

- А) 17
- Б) 18
- В) 22
- Г) 23

5. Модель тренду, що використовується для прогнозування, визначається:

- А) за найменшою сумою абсолютних відхилень теоретичних рівнів ряду динаміки від фактичних
- Б) за найбільшою сумою абсолютних відхилень теоретичних рівнів ряду динаміки від фактичних
- В) за найменшою сумою квадратів відхилень теоретичних рівнів ряду динаміки від фактичних
- Г) за найбільшою сумою квадратів відхилень теоретичних рівнів ряду динаміки від фактичних

6. Метод визначення основної тенденції розвитку у рядах динаміки, який полягає в знаходженні такої математичної лінії, ординати точок якої були б найближчі до фактичних значень ряду динаміки, називається:

- А) укрупнення інтервалів

Б) згладжування за допомогою ковзної середньої

В) аналітичне вирівнювання

Г) інтерполяція

7. При аналізі динаміки виробництва зерна параметр $a_1 = -2000$ у рівнянні параболи означає:

А) теоретичний валовий збір зерна у році з номером 0 становив 2000 ц

Б) середній щорічний абсолютний приріст зерна становив 2000 ц

В) середнє сповільнення зміни валового збору зерна становило 2000 ц

Г) середнє щорічне абсолютне зменшення виробництва зерна становило 2000 ц

8. Прогнозування здійснюється за рівнянням параболи на основі 13 рівнів ряду динаміки. Визначте ступінь вільності:

А) 16

Б) 15

В) 11

Г) 10

9. Інтервалом екстраполяції називають:

А) інтервал, в якому знаходяться всі рівні досліджуваного ряду динаміки

Б) інтервал, в якому знаходяться теоретичні рівні досліджуваного ряду динаміки

В) інтервал, в якому буде знаходитись прогнозоване значення досліджуваного явища

Г) інтервал, в якому знаходяться параметри моделі тренду

10. При аналізі рядів динаміки параметр a_1 у рівнянні прямої характеризує:

А) порядковий номер року

Б) середній щорічний абсолютний приріст рівнів вирівняного ряду динаміки

В) середнє прискорення зміни рівня досліджуваного явища

Г) вирівняний рівень ряду динаміки при $t=0$

11. Знаходження значень ознаки за межами аналізованого періоду називається

А) сезонність

Б) тенденція

В) інтерполяція

Г) екстраполяція

12. При аналізі динаміки виробництва зерна параметр $a_2 = -1000$ у рівнянні прямої означає:

А) середнє прискорення зміни валового збору зерна становило 1000 ц

Б) середнє сповільнення зміни валового збору зерна становило 1000 ц

В) середнє щорічне абсолютне зменшення виробництва зерна становило 1000 ц

Г) параметр відсутній

13. Прогнозування здійснюється за рівнянням прямої на основі 10 рівнів ряду динаміки. Визначте ступінь вільності:

А) 7

Б) 8

В) 12

Г) 13

14. При аналізі рядів динаміки параметр a_2 у рівнянні параболи характеризує:

А) вирівняний рівень ряду динаміки при $t=0$

Б) середній щорічний абсолютний приріст рівнів вирівняного ряду динаміки

В) середнє прискорення зміни рівня досліджуваного явища

Г) порядковий номер року

15. При аналізі динаміки виробництва зерна параметр $a_1=5000$ у рівнянні прямої означає:

А) теоретичний валовий збір зерна у році з номером 0 становив 5000 ц

Б) середній щорічний абсолютний приріст зерна становив 5000 ц

В) середнє прискорення зміни валового збору зерна становило 5000 ц

Г) середнє щорічне абсолютне зменшення виробництва зерна становило 5000 ц

16. Зміна, що визначає загальний напрямок розвитку явища називається:

А) тренд,

Б) інтерполяція

В) екстраполяція

Г) сезонність

17. Прогнозування здійснюється за рівнянням прямої на основі 15 рівнів ряду динаміки. Визначте ступінь вільності:

А) 12

Б) 13

В) 17

Г) 18

18. При аналізі динаміки виробництва зерна параметр $a_0=8000$ у рівнянні прямої означає:

А) теоретичний валовий збір зерна у році з номером 0 становив 8000 ц

Б) фактичний валовий збір зерна у початковому році становив 8000 ц

В) середній щорічний абсолютний приріст виробництва зерна становив 8000 ц

Г) середнє прискорення зміни валового збору зерна становило 8000 ц

19. Метод визначення основної тенденції розвитку у рядах динаміки, який передбачає послідовне зсування та один період при збереженні сталого інтервалу періоду, називається:

- А) аналітичне вирівнювання
- Б) згладжування за допомогою ковзної середньої
- В) екстраполяція тренда
- Г) укрупнення інтервалів

20. При аналізі рядів динаміки параметр a_2 у рівнянні прямої характеризує:

- А) середній щорічний абсолютний приріст рівнів вирівняного ряду динаміки
- Б) середнє прискорення зміни рівня досліджуваного явища
- В) вирівняний рівень ряду динаміки при $t=0$
- Г) параметр відсутній

21. Прогнозування здійснюється за рівнянням параболи на основі 7 рівнів ряду динаміки. Визначте ступінь вільності:

- А) 4
- Б) 5
- В) 9
- Г) 10

22. Побудований по отриманих індексах сезонності лінійний графік має назву:

- А) сезонний стрибок
- Б) місячний стрибок
- В) сезонна хвиля
- Г) місячна хвиля

23. При аналізі динаміки виробництва зерна параметр $a_2 = -500$ у рівнянні параболи означає:

- А) середнє щорічне абсолютне зменшення виробництва зерна становило 500 ц
- Б) середнє прискорення зміни валового збору зерна становило 500 ц
- В) середнє сповільнення зміни валового збору зерна становило 500 ц
- Г) параметр відсутній

24. Визначення за наявними даними за певний період часу відсутніх значень ознаки всередині періоду називається:

- А) інтерполяція
- Б) екстраполяція
- В) сезонність
- Г) тенденція

10. СТАТИСТИЧНІ МЕТОДИ ВИМІРЮВАННЯ ВЗАЄМОЗВ'ЯЗКІВ

10.1 Види та форми взаємозв'язків

Досліджуючи різні процеси і явища, що відбуваються в природі та суспільстві, необхідно враховувати їх причинно-наслідкові взаємозв'язки, які визначаються певними кількісними характеристиками. Оцінка найбільш істотних з них, а також впливу одних факторів на інші є однією з основних завдань статистики.

Форми прояву взаємозв'язків досить різноманітні, що зумовлює їх розгляд за наступними класифікаційними ознаками.

За ступенем залежності одного явища від іншого розрізняють зв'язок:

- **функціональний (повний)** - зв'язок, при якому кожному значенню факторної ознаки, що характеризує певне явище, відповідає одне значень результативної ознаки (функції);

- **кореляційний (неповний, або статистичний)** - зв'язок, який носить імовірнісний характер, й кожному певному значенню факторної ознаки відповідає кілька значень результативної ознаки.

Функціональні зв'язки зазвичай виражаються формулами і досліджуються в математиці і фізиці. Наприклад, площа круга - результативний ознака - прямо пропорційна його радіусу - факторній ознаці. Однак, функціональні зв'язки мають місце і в економіці. Так, заробітна плата робітника погодинній оплаті дорівнює добутку годинної тарифної ставки на число відпрацьованих годин. Функціональний зв'язок є точним і повним, тому зазвичай відомі всі фактори, що впливають на результативну ознаку. При функціональних зв'язках величина результативної ознаки повністю показується факторними ознаками.

Однак, в масових явищах суспільного життя по причині крайньої різноманітності чинників та їх взаємозв'язку і суперечливої дії цих факторів, що не піддаються суворому обліку і контролю, виникає широке варіювання результативної ознаки. Це свідчить про те, що зв'язок між ознаками неповна, а виявляється лише в загальному і середньому. Наприклад, кореляційний зв'язок між впливом добрива і врожайністю культур, між продуктивністю і енергооснащеністю підприємства.

Важлива особливість кореляційних зв'язків полягає в тому, що вони виявляються не в окремих випадках, а в масових суспільних явищах. Прояв кореляційних залежностей підлягає дії закону великих чисел: лише в досить великому числі фактів індивідуальні особливості та другорядні факти згладяться і залежність проявиться досить чітко.

За напрямом кореляційний зв'язок може бути:

- **прямий** - факторна ознака змінюється в тому самому напрямі, що й результативна;

- **зворотний** - із збільшенням факторної ознаки результативна ознака зменшується або, навпаки, із зменшенням факторної ознаки результативна ознака збільшується.

Прямий кореляційний зв'язок спостерігається між внесенням добрив та урожайністю, фондоозброєністю і продуктивністю праці, собівартістю і ціною товару і т.д.

Прикладом же зворотного зв'язку може бути залежність між урожайністю і собівартістю продукції, собівартістю продукції і рентабельністю виробництва, продуктивністю праці і собівартістю продукції, тощо.

Крім того, розрізняють наступні **форми кореляційного зв'язку**:

- **прямолінійний зв'язок** - із зростанням величини факторної ознаки відбувається безупинне зростання результативної ознаки і навпаки. Математично така залежність представляється рівнянням прямої. Графік представлений у вигляді прямої лінії. Цю залежність називають лінійною.

- **криволінійний зв'язок** - із зростанням величини факторної ознаки зміна результативної ознаки відбувається нерівномірно, її напрямок може навіть змінюватися. Математично така залежність представляється рівняннями кривих ліній: Графічно цей процес може бути представлений гіперболою, параболою чи ламаною лінією.

Прямолінійний зв'язок спостерігається між вартістю основного капіталу та випуском продукції, фондоозброєністю праці та середньодобовим виробітком одного працівника, споживанням окремих товарів та рівнем доходів населення, а також між урожайністю сільськогосподарських культур та дозами внесенням добрив або продуктивністю тварин та рівнем їх годівлі у межах діючих науково обґрунтованих норм.

Криволінійна залежність має форму параболи другого порядку при аналізі впливу кількості опадів за вегетаційний період на урожайність сільськогосподарських культур, віку робітників на продуктивність їхньої праці і т.п. За допомогою криволінійної ступеневої функції часто визначають, наприклад, залежність між фондом заробітної плати та випуском продукції, витратами праці та випуском продукції і т.д. Прикладом використання гіперболічної функції може бути зв'язок між випуском продукції та собівартістю, товарними запасами та товарооборотом, вплив продуктивності тварин на рівень годівлі, тощо.

Для кореляційних зв'язків є відмінності в тому випадку, якщо: досліджується зв'язок між однією ознакою - фактором і результативною ознакою чи досліджується зв'язок між декількома ознаками - факторами і

результативною ознакою. У першому випадку має місце парний зв'язок і парна кореляція, у другому випадку багатofакторний зв'язок і множинна кореляція.

Кореляція (лат. *correlatio* - співвідношення) - поняття, що вказує на статистичний зв'язок, який існує між досліджуваними явищами.

Таким чином, *за кількістю досліджуваних ознак* розрізняють кореляцію:

- **парну (просту)** - аналізують зв'язок між факторною і результативною ознаками;

- **множинну** - аналізують залежність результативної ознаки від двох і більше факторних ознак.

Вочевидь, вся безліч факторів, які впливають на величину результативного показника, реально не може бути введена до розгляду, та практично в цьому і немає необхідності, тому що їх роль і значення в формуванні величини результативного показника можуть мати істотні відмінності. Тому при обмеженні числа факторів, які включаються до вивчення, поряд з якісним аналізом доцільно використовувати і певні кількісні оцінки, що дозволяють конкретно охарактеризувати вплив факторів на результативний показник.

Фактори, які включаються в дослідження, повинні бути незалежними один від одного, так як наявність тісного зв'язку між ними свідчить про те, що вони характеризують одні й ті ж сторони досліджуваного явища і значною мірою дублюють одне одного.

10.2 Сутність та етапи кореляційно-регресійного аналізу

У найбільш загальному вигляді завдання статистики в області вивчення взаємозв'язків полягає в кількісній оцінці їх наявності та напрямку, а також характеристиці сили та форми впливу одних факторів на інші. Для її вирішення застосовуються дві групи методів, одна з яких включає в себе методи кореляційного аналізу, а інша - регресійний аналіз.

У той же час ряд дослідників об'єднує ці методи в **кореляційно-регресійний аналіз** - метод кількісної оцінки взаємозалежностей між статистичними ознаками, які характеризують окремі суспільно-економічні явища і процеси, що має під собою деякі підстави: наявність цілого ряду спільних обчислювальних процедур, взаємодоповнення при інтерпретації результатів, тощо.

Завдання власне *кореляційного аналізу* зводяться до виміру щільності зв'язку між варіюючими ознаками, визначенню невідомих причинних зв'язків і оцінці факторів, які найбільше впливають на результативну ознаку.

Завдання *регресійного аналізу* лежать у сфері встановлення форми залежності, визначення функції регресії, використання рівняння для оцінки невідомих значенні залежної змінної.

Рівняння, за допомогою яких визначають статистичний зв'язок між корелюючими величинами, називають **рівнянням регресії** (кореляційними рівняннями), а лінії, побудовані на їх основі, називають **лініями регресії**.

Розв'язання названих завдань спирається на відповідні прийоми, алгоритми, показники, застосування яких дає підставу говорити про статистичне вивчення взаємозв'язків.

Етапи кореляційно-регресійного аналізу:

1. Попередній (апріорний) аналіз. Він дає непогані результати якщо проводиться досить кваліфікованим дослідником.
2. Збір інформації та її первинна обробка.
3. Побудова моделі (рівняння регресії).
4. Оцінка тісноти зв'язків ознак, оцінка рівняння регресії та аналіз моделі.
5. Прогнозування розвитку аналізованої системи за рівнянням регресії.

На першому етапі формулюється завдання дослідження, визначається методика вимірювання показників або збору інформації, визначається число факторів, виключаються дублюючі фактори, тощо.

На другому етапі аналізується обсяг одиниць: сукупність повинна бути досить великою за кількістю одиниць і спостережень. Дані повинні бути кількісно і якісно однорідними.

На третьому етапі визначається напрям, форма зв'язку і тип аналітичної функції (парабола, гіпербола, пряма) і знаходяться її параметри. Напрямок і форму зв'язку встановлюють за допомогою статистичних групувань, а також графіків, побудованих у системі прямокутних координат на основі емпіричних даних.

Графічне зображення статистичних величин дає наочне уявлення про наявність зв'язку між досліджуваними ознаками. При побудові графіка на горизонтальній вісі відкладають значення факторної ознаки (x). Відклавши на перетині відповідних значень x і y точки, дістають кореляційне поле. За характером розміщення точок на кореляційному полі роблять висновок про напрям і форму зв'язку. Якщо точки безладно розкидані по всьому полю, то це свідчить, що зв'язку між досліджуваними ознаками немає. Якщо точки концентруються навколо осі від нижнього лівого кута до верхнього правого, то між результативною і факторною ознаками існує прямий зв'язок. Якщо ж точки концентруються навколо осі від верхнього лівого кута до нижнього правого, то між ознаками існує зворотний зв'язок. Характер розподілу точок по кореляційному полю вказує також на наявність прямолінійної чи криволінійної залежності між факторною і результативною ознаками.

На четвертому етапі оцінюється достовірність усіх характеристик кореляційного зв'язку і рівняння регресії використовуючи критерій достовірності Фішера або Ст'юдента, проводиться економіко-технологічний аналіз параметрів.

На п'ятому етапі здійснюється прогноз можливих значень результату по кращим значенням факторних ознак, включених в модель. Тут вибираються найкращі і найгірші значення факторів і результату.

10.3 Парна кореляція

Парна кореляція або парна регресія можуть розглядатися як окремий випадок відображення зв'язку деякої залежної змінної, з одного боку, і однією з безлічі незалежних змінних - з іншого.

Ознаки, які впливають на інші і зумовлюють їх зміну, називають **факторними (x)**, ознаки, які змінюються під впливом факторних називають **результативними (y)**.

1) Парна прямолінійна кореляція.

При аналізі *прямолінійної залежності* використовується рівняння прямої:

$$y_x = a_0 + a_1 x,$$

де y_x – теоретичні (обчислені за рівнянням регресії) значення результативної ознаки;

a_0 - вільний член рівняння, має тільки розрахункове та не має економічного значення;

a_1 - **коефіцієнт регресії** вказує, як змінюється результативна ознака y внаслідок зміни факторної ознаки x на одиницю;

x — значення факторної ознаки.

Параметри a_0 та a_1 знаходять розв'язанням системи рівнянь:

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

де n - кількість спостережень.

Таким чином, параметри a_0 і a_1 обчислюють за формулами:

$$a_0 = \frac{\sum x^2 \sum y - \sum x \sum xy}{n \sum x^2 - (\sum x)^2};$$
$$a_1 = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - (\sum x)^2}.$$

При *прямому* зв'язку між корелюючими ми ознаками коефіцієнт регресії має *додатне* значення, при *зворотному* - *від'ємне*.

При прямолінійній формі щільність зв'язку визначається за формулою *лінійного коефіцієнта парної кореляції*.

$$r = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sigma_y \cdot \sigma_x},$$

де σ_y, σ_x - відповідно середні квадратичні відхилення факторної і результативної ознак.

Однак лінійний коефіцієнт парної кореляції також можна визначити за наступною формулою:

$$r = \frac{n \sum xy - \sum x \sum y}{\sqrt{[n \sum x^2 - (\sum x)^2] \cdot [n \sum y^2 - (\sum y)^2]}}$$

У разі парної залежності коефіцієнт кореляції при прямому зв'язку коливається від 0 до 1, при зворотному від 0 до -1.

Чим ближчим є абсолютне значення коефіцієнту кореляції до 1, тим щільніший зв'язок між y і x .

При $r < \pm 0,1$ зв'язку немає;

при $r = \pm 0,1-0,3$ - зв'язок слабкий;

при $r = \pm 0,3-0,7$ - зв'язок середній;

при $r > 0,7$ - зв'язок щільний;

при $r = 1$ - зв'язок між ознаками функціональний.

Примітка. Коефіцієнт кореляції має такий самий знак, як і коефіцієнт регресії у рівняння зв'язку.

Коли отримані оцінки кореляції та регресії, необхідно перевірити їх на відповідність істинним параметрам взаємозв'язку.

Для оцінки значимості лінійного коефіцієнта кореляції використовується t -критерій (критерій Ст'юдента), значення якого розраховується за формулою:

$$t_{\text{факт.}} = \frac{|r| \cdot \sqrt{n-2}}{\sqrt{1-r^2}},$$

де $t_{\text{факт.}}$ - фактичне значення критерію Ст'юдента; .

r - лінійний коефіцієнт кореляції;

n - кількість спостережень.

Теоретичне значення t -критерію визначається за відповідною таблицею «Критичні точки розподілу Ст'юдента, t -розподіл» (Додаток Б)

Якщо $t_{\text{факт.}} > t_{\text{теор.}}$, то це свідчить про суттєвість (значимість) коефіцієнта кореляції. Якщо $t_{\text{факт.}} < t_{\text{теор.}}$, то величина коефіцієнта кореляції вважається несуттєвою.

Для визначення, яка частина загальної варіації результативної ознаки, залежить від факторної ознаки розраховується *коефіцієнт детермінації*:

$$D = r^2 \cdot 100\%$$

Задача 10.1. Здійснити кореляційно-регресійний аналіз впливу внесених доз мінеральних добрив на урожайність озимої пшениці на підставі даних за 7 років.

Розв'язання.

Спочатку встановлюється адекватна математична формула, яка характеризує цей зв'язок та здійснюються відповідні розрахунки за способом найменших квадратів (табл. 10.1)

Таблиця 10.1

Вихідні та розрахункові дані для визначення залежності між внесенням мінеральних добрив та урожайністю озимої пшениці

Роки	Мінеральні добрива, ц д.р./га	Урожайність, ц/га	Розрахункові величини			Теоретичне значення результативної ознаки
	x	y	xy	x^2	y^2	$y_x = a_0 + a_1x$
2006	1,5	27,6	41,4	2,25	761,76	27,7
2007	1,4	25,0	35	1,96	625	26,2
2008	1,2	24,4	29,28	1,44	595,36	23,3
2009	1,4	26,7	37,38	1,96	712,89	26,2
2010	1,3	23,1	30,03	1,69	533,61	24,7
2011	1,4	27,2	38,08	1,96	739,84	26,2
2012	1,6	29,5	47,2	2,56	870,25	29,2
Разом	9,8	183,5	258,37	13,82	4838,71	183,5

Параметри a_0 і a_1 рівняння парної прямої регресії обчислюються за формулами:

$$a_0 = \frac{\sum x^2 \sum y - \sum x \sum xy}{n \sum x^2 - (\sum x)^2} = \frac{13,82 * 183,5 - 9,8 * 258,37}{7 * 13,82 - 9,8^2} = 5,63$$

$$a_1 = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - (\sum x)^2} = \frac{7 * 258,37 - 9,8 * 183,5}{7 * 13,82 - 9,8^2} = 14,7$$

Отже, рівняння кореляційного зв'язку між урожайністю озимої пшениці і внесенням мінеральних добрив матиме такий вигляд:

$$y_x = 5,63 + 14,7x;$$

Економічний зміст рівняння: коефіцієнт регресії показує, що із збільшенням дози внесення мінеральних добрив на 1 ц д.р./га урожайність озимої пшениці зростає в середньому на 14,7 ц/га.

Графічне зображення фактичного та функціонального зв'язку (лінія регресії) представлено на рис. 10.1.

Для визначення щільності прямолінійного парного зв'язку між корелюючими величинами обчислюється лінійний коефіцієнт парної кореляції:

$$r = \frac{n \sum xy - \sum x \sum y}{\sqrt{[n \sum x^2 - (\sum x)^2] \cdot [n \sum y^2 - (\sum y)^2]}} =$$

$$= \frac{7 * 258,37 - 9,8 * 183,5}{\sqrt{(7 * 13,82 - 9,8^2)(7 * 4838,71 - 183,5^2)}} = 0,87$$

Коефіцієнт лінійної кореляції показує, що між дозами внесених мінеральних добрив і урожайністю озимої пшениці зв'язок прямий і щільний.

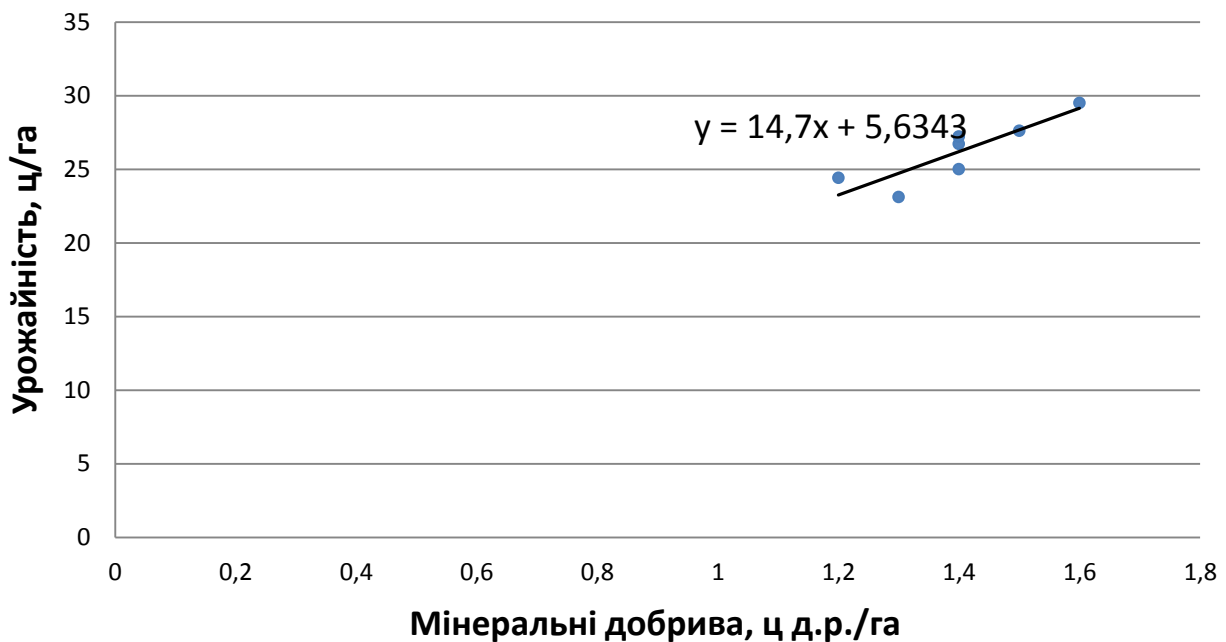


Рис.10.1. Залежність між дозами внесення мінеральних добрив та урожайністю озимої пшениці

Для оцінки значимості лінійного коефіцієнта кореляції використаємо *t*-критерій Ст'юдента:

$$t_{\text{факт}} = \frac{|r| * \sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0,87 * \sqrt{7-2}}{\sqrt{1-0,87^2}} = 5,46$$

Згідно таблиці «Критичні точки розподілу Ст'юдента, *t*-розподіл» теоретичне значення критерію $t_{\text{теор}} = 2,57$.

Так як $t_{\text{факт}} (5,46) > t_{\text{теор}} (2,57)$, то це свідчить про суттєвість коефіцієнта кореляції.

Долю варіації результативної ознаки признаку під впливом факторної визначаємо за допомогою **коефіцієнту детермінації**:

$$D = r^2 * 100\% = 0,87^2 * 100\% = 76,1\%$$

Урожайність озимої пшениці на 76,1% залежить від внесення мінеральних добрив, а на 23,9% - від впливу неврахованих факторів.

2) Парна криволінійна кореляція.

Коли із зміною факторної ознаки змінюється не тільки результативна ознака, а й її приріст, використовують параболічну і гіперболічну нелінійні форми зв'язку.

Якщо криволінійна залежність має форму параболи другого порядку, зв'язок виражають рівнянням:

$$y_x = a_0 + a_1x + a_2x^2.$$

На відміну від прямолінійної залежності коефіцієнти регресії криволінійного зв'язку не можна інтерпретувати однозначно, оскільки зміна результативної ознаки при різних значеннях факторної неоднакова.

Приклад. Рівняння параболи другого порядку, що характеризує залежність урожайності насіння багаторічних трав (в ц/га) від глибини зрошення (в см):

$$y_x = 3,6 + 0,42x - 0,007x^2$$

показує, що при збільшенні глибини зрошення на 1 см урожайність насіння багаторічних трав зростатиме в середньому на 0,42 ц/га, але це підвищення відбуватиметься до певного рівня зрошення, після чого знижуватиметься в середньому на 0,007 ц/га.

Коли залежність має гіперболічний вигляд (зворотна залежність), її обчислюють за допомогою рівняння гіперболи:

$$y_x = a_0 + \frac{a_1}{x}$$

Приклад. Підставивши у рівняння гіперболи, яке характеризує залежність витрат кормів на 1 ц молока (ц кормових одиниць) від виходу продукції на 1 голову (ц):

$$y_x = 0,3 + \frac{78,5}{x}$$

значення факторної ознаки x , матимемо теоретичні рівні витрат кормів на виробництво 1 ц молока.

Щільність зв'язку при криволінійній залежності визначають за допомогою **індексу кореляції i** (кореляційного відношення) за формулою:

$$i = \sqrt{\frac{\sigma_{yx}^2}{\sigma_y^2}},$$

де σ_{yx}^2 - міжгрупова дисперсія, обчислена за рівнянням регресії;

σ_y^2 - загальна дисперсія.

Індекс кореляції змінюється в межах від 0 до 1, тобто завжди є додатною величиною.

Досліджуючи кореляційні зв'язки, слід уникати необґрунтованого ускладнення застосовуваних криволінійних рівнянь. Ускладнення рівнянь зв'язку, якщо воно не ґрунтується на достатньому логічному і економічному змісті досліджуваних явищ, утруднює інтерпретацію показників регресії і позбавляє аналіз цілеспрямованості.

10.3 Множинна кореляція

Багатофакторні регресійні моделі дають змогу оцінювати вплив на досліджувану результативну ознаку кожного із факторів рівняння при фіксованому значенні (на середньому рівні) інших факторів. При цьому важливою умовою множинної кореляції є відсутність між факторами функціонального зв'язку.

Найбільш складним питанням при множинній кореляції є вибір форми зв'язку і відповідного математичного рівняння множинної регресії. Вибір типу функції повинен ґрунтуватися на теоретичному аналізі досліджуваного явища або на досвіді попередніх аналогічних досліджень.

Враховуючи, що будь-яку функцію багатьох змінних логарифмуванням можна звести до лінійного виду, рівняння множинної регресії частіше будують у лінійній формі.

У загальному вигляді формула лінійного рівняння множинної регресії така:

$$y_x = a_0 + a_1x_1 + a_2x_2 + \dots + a_nx_n$$

де x_1, x_2, \dots, x_n — факторні ознаки.

a_1, a_2, \dots, a_n - окремі коефіцієнти регресії цього рівняння, що характеризують ступінь впливу відповідного фактору на результативний показник при фіксованому значенні інших факторів, вони показують, наскільки зміниться результативний показник при зміні відповідного фактору на одиницю;

a_0 - вільний член рівняння, який не має економічного змісту і не інтерпретується.

Рівняння множинної регресії обчислюють способом найменших квадратів розв'язанням системи рівнянь:

$$\begin{cases} \Sigma y = na_0 + a_1 \Sigma x_1 + a_2 \Sigma x_2 + \dots + a_n \Sigma x_n \\ \Sigma yx_1 = a_0 \Sigma x_1 + a_1 \Sigma x_1^2 + a_2 \Sigma x_1x_2 + \dots + a_n \Sigma x_1x_n \\ \Sigma yx_2 = a_0 \Sigma x_2 + a_1 \Sigma x_1x_2 + a_2 \Sigma x_2^2 + \dots + a_n \Sigma x_2x_n \end{cases}$$

$$\Sigma yx_n = a_0 \Sigma x_n + a_1 \Sigma x_1x_n + a_2 \Sigma x_2x_n + \dots + a_n \Sigma x_n^2$$

Показниками щільності зв'язку при множинній кореляції є парні, часткові і множинні (сукупні) коефіцієнти кореляції та множинний коефіцієнт детермінації.

Парні коефіцієнти кореляції використовують для вимірювання щільності зв'язку між двома досліджуваними ознаками і розраховують аналогічно однофакторному зв'язку.

Часткові коефіцієнти кореляції характеризують щільності зв'язку між результативною і однією факторною ознакою при умові, що інші факторні ознаки перебувають на постійному рівні.

При значній кількості факторів частковий коефіцієнт кореляції можна визначити за формулою:

$$R_{yx(z...v)} = \sqrt{\frac{R_{yxz...v}^2 - R_{yz...v}^2}{1 - R_{yz...v}^2}},$$

де $R_{yxz...v}$ – коефіцієнт множинної кореляції;

$R_{yz...v}$ – коефіцієнт множинної кореляції результативного фактора (y) зі всіма за виключенням досліджуваного.

Основним показником щільності зв'язку при множинній кореляції є **коефіцієнт множинної (сукупної) кореляції**, який може мати значення від 0 до 1 і знаходиться за формулою:

$$R = \sqrt{\frac{\sigma_{обчисл}^2}{\sigma_{заг}^2}},$$

де $\sigma_{обчисл}^2$ - дисперсія результативного показника, обчислена за рівнянням множинної регресії;

$\sigma_{заг}^2$ - загальна дисперсія результативного показника.

Для оцінки значимості коефіцієнта множинної кореляції також використовується *t*-критерій (критерій Ст'юдента), значення якого розраховується за формулою: .

$$t_{факт} = \frac{R * \sqrt{n - m - 1}}{1 - R^2},$$

де n – кількість спостережень;

m- кількість факторів

Множинний коефіцієнт детермінації показує, яка частка варіації досліджуваного результативного признаку зумовлена впливом факторів, включених у рівняння.

$$R^2 = \frac{\sigma_{обчисл}^2}{\sigma_{заг}^2} * 100$$

Найбільш простим видом рівняння множинної регресії є *лінійне рівняння з двома незалежними змінними*:

$$y_{x_1x_2} = a_0 + a_1x_1 + a_2x_2,$$

де a_1, a_2 - параметри рівняння, які характеризують вплив кожного з факторів при елімінаванні іншого.

Параметри цього рівняння визначаються розв'язанням системи рівнянь:

$$\begin{cases} \Sigma y = na_0 + a_1 \Sigma x_1 + a_2 \Sigma x_2 \\ \Sigma ux_1 = a_0 \Sigma x_1 + a_1 \Sigma x_1^2 + a_2 \Sigma x_1 x_2 \\ \Sigma ux_2 = a_0 \Sigma x_2 + a_1 \Sigma x_1 x_2 + a_2 \Sigma x_2^2 \end{cases}$$

Задача 10.2. На підставі даних про склад, доходи та витрати на харчування 7 родин розрахувати показники множинної регресії та кореляції та зробити належні висновки.

Розв'язання.

Обчислюємо необхідні елементи системи рівнянь (табл.10.2).

Таблиця 10.2

Вихідні та розрахункові дані для визначення залежності між щомісячними доходами, величиною родини та її витратами на харчування

№	Доходи родини в місяць, дол. (x_1)	Кількість членів родини, осіб (x_2)	Витрати на харчування родини в місяць, дол. (y)	ux_1	ux_2	x_1x_2	x_1^2	x_2^2	y^2
1	90	1	25	2250	25	90	8100	1	625
2	110	1	28	3080	28	110	12100	1	784
3	120	2	31	3720	62	240	14400	4	961
4	130	2	32	4160	64	260	16900	4	1024
5	180	3	36	6480	108	540	32400	9	1296
6	200	3	42	8400	126	600	40000	9	1764
7	280	4	55	15400	220	1120	78400	16	3025
Σ	1110	16	249	43490	633	2960	202300	44	9479

Параметри рівняння множинної кореляції a_0, a_1 та a_2 знаходяться шляхом вирішення системи з 3 нормальних рівнянь.

Система рівнянь для нашого прикладу:

$$\begin{cases} 249=7a_0+1110 a_1+16a_2 \\ 43\ 490=1110a_0+202\ 300 a_1+2\ 960a_2 \\ 633=16a_0+2\ 960 a_1+44a_2 \end{cases}$$

Для розв'язання системи розділимо всі члени рівнянь на коефіцієнти при a_0 :

$$\begin{array}{l|l} 249=7a_0+1110 a_1+16a_2 & : 7 \\ 43\ 490=1110a_0+202\ 300 a_1+2\ 960a_2 & :1110 \\ 633=16a_0+2\ 960 a_1+44a_2 & : 16 \end{array}$$

Віднімемо перше рівняння від другого та друге від третього:

$$\begin{array}{r} _35,57=a_0+158,571 a_1+2,29a_2 \\ _39,18=a_0+182,25a_1+2,67a_2 \\ 39,56=a_0+185 a_1+2,75a_2 \end{array}$$

Розділимо всі члени рівнянь на коефіцієнти при a_1 та віднімемо друге рівняння від першого:

$$\begin{array}{l|l} 3,61= 23,68 a_1+0,38a_2 & : 23,68 \\ 0,38=2,75a_1+0,08a_2 & :2,75 \\ _0,1524 = a_1+0,016a_2 \\ \underline{0,1382 =a_1+0,0291a_2} \\ 0,0142 = -0,0131 a_2 \\ a_2= -1,084 \end{array}$$

Підставивши значення параметру a_2 в рівняння, отримаємо параметри a_0 та a_1 :

$$\begin{array}{l} a_1= 0,1697 \\ a_0= 11,149 \end{array}$$

Рівняння регресії має вигляд:

$$y_{x_1x_2} = 11,149 + 0,1697x_1 - 1,084x_2$$

Економічний зміст рівняння наступний: із зростанням місячного доходу родини на 1 дол. витрати на харчування збільшуються на 0.17 дол., а із збільшення розміру родини на 1 особу витрати на харчування зменшуються на 1,08 дол.

Далі розраховуємо показники щільності багатомірного зв'язку.

Парні коефіцієнти кореляції:

$$\begin{aligned} r_{yx_1} &= \frac{n \sum yx_1 - \sum y \sum x_1}{\sqrt{(n \sum x_1^2 - (\sum x_1)^2)(n \sum y^2 - (\sum y)^2)}} = \\ &= \frac{7 * 43490 - 249 * 1110}{\sqrt{(7 * 202300 - 1110^2) * (7 * 9479 - 249^2)}} = 0,991 \end{aligned}$$

тобто між доходами родини та витратами на її харчування існує дуже щільний прямий зв'язок.

$$r_{yx_2} = \frac{n \sum yx_2 - \sum y \sum x_2}{\sqrt{(n \sum x_2^2 - (\sum x_2)^2)(n \sum y^2 - (\sum y)^2)}} =$$

$$= \frac{7 * 633 - 249 * 16}{\sqrt{(7 * 44 - 16^2)(7 * 9479 - 249^2)}} = 0,94$$

тобто між кількістю членів родини та витратами на її харчування існує щільний прямий зв'язок.

$$r_{x_1x_2} = \frac{n \sum x_1x_2 - \sum x_1 \sum x_2}{\sqrt{(n \sum x_1^2 - (\sum x_1)^2)(n \sum x_2^2 - (\sum x_2)^2)}} =$$

$$= \frac{7 * 2960 - 1110 * 16}{\sqrt{(7 * 202300 - 1110^2)(7 * 44 - 16^2)}} = 0,957$$

тобто між доходами та кількістю членів родини існує щільний прямий зв'язок.

Часткові коефіцієнти кореляції:

- між ознаками у та x_1 без урахування впливу ознаки x_2 :

$$r_{yx_1(x_2)} = \frac{r_{yx_1} - r_{yx_2} * r_{x_1x_2}}{\sqrt{(1 - r_{yx_2}^2)(1 - r_{x_1x_2}^2)}} = \frac{0,991 - 0,94 * 0,957}{\sqrt{(1 - 0,94^2)(1 - 0,957^2)}} = 0,923$$

тобто без урахування впливу розміру родини між її доходами та витратами на харчування існує щільний прямий зв'язок.

- між ознаками у та x_2 без урахування впливу ознаки x_1 :

$$r_{yx_2(x_1)} = \frac{r_{yx_2} - r_{yx_1} * r_{x_1x_2}}{\sqrt{(1 - r_{yx_1}^2)(1 - r_{x_1x_2}^2)}} = \frac{0,94 - 0,991 * 0,957}{\sqrt{(1 - 0,991^2)(1 - 0,957^2)}} = -0,22$$

тобто без урахування впливу доходів родини тобто між кількістю членів родини та витратами на її харчування існує слабкий зворотний зв'язок.

Вочевидь, величина парного коефіцієнта кореляції r_{yx_2} , яка свідчила про щільний прямий зв'язок між другим фактором та результативною ознакою, була зумовлена щільним взаємозв'язком між обома факторами.

Коефіцієнт множинної (сукупної) кореляції:

$$R_{yx_1x_2} = \sqrt{\frac{r_{yx_1}^2 + r_{yx_2}^2 - 2r_{yx_1} * r_{yx_2} * r_{x_1x_2}}{1 - r_{x_1x_2}^2}} =$$

$$= \sqrt{\frac{0,991^2 + 0,94^2 - 2 * 0,991 * 0,94 * 0,957}{1 - 0,957^2}} = 0,991$$

тобто між витратами на харчування, доходом родини та кількістю її членів існує дуже щільний зв'язок.

За допомогою t-критерію, який обчислюється за формулою:

$$t_{\text{факт}} = \frac{R * \sqrt{n - m - 1}}{1 - R^2} = \frac{0,991 * \sqrt{7 - 2 - 1}}{1 - 0,991^2} = 110,7$$

робимо висновок про значимість множинного коефіцієнта кореляції ($t_{\text{факт}}(110,7) > t_{\text{теор}}(2,78)$).

Множинний коефіцієнт детермінації:

$$R^2 = 0,991^2 * 100 = 98,3\%$$

Витрати родини на харчування на 98,3% залежать від її доходу та кількості членів, а на 1,7% - від інших факторів.

Завдання для самостійного розв'язання.

Завдання 10.1. По десяти промисловим підприємствам однієї галузі є наступні дані за рік:

Показники	Номер підприємства									
	1	2	3	4	5	6	7	8	9	10
Основні засоби, млн. грн.	10	13	15	19	22	26	27	30	34	35
Прибуток, млн. грн.	4	5	6	8	9	12	13	15	16	17

За цими даними знайдіть рівняння регресії, що виражає залежність прибутку підприємств від вартості їх основних засобів. Визначте напрямок і тісноту зв'язку. Зробіть висновки. Зобразіть зв'язок графічно.

Завдання 10.2. Визначити рівняння регресії й коефіцієнт кореляції між зростом і вагою студентів:

№ пп	1	2	3	4	5	6	7	8	9	10
Зріст, см	160	161	171	178	176	181	185	182	184	183
Вага, кг	60	55	75	70	72	75	76	80	78	80

Оцінити значимість коефіцієнтів. Зробити висновки. Зобразити зв'язок графічно.

Завдання 10.3. Є дані про рівень виробітку й стаж робітників однієї з молодіжних бригад машинобудівного підприємства:

Номер робітника	А	Б	В	Г	Д	Е	Ж	З
Стаж, повних років	1	3	4	2	5	7	8	6
Виробіток, шт./ год.	8	9	12	10	11	15	16	18

За цими даними визначте напрямок і силу зв'язку, оцініть значимість коефіцієнтів. Побудуйте рівняння регресії. Зобразіть зв'язок графічно.

Завдання 10.4. За даними таблиці:

- 1) описати залежність між кількістю бухгалтерів на підприємстві та його товарооборотом за допомогою рівняння прямої ;
- 2) за допомогою лінійного коефіцієнта кореляції оцінити щільність зв'язку між ознаками;
- 3) перевірити суттєвість коефіцієнта кореляції за формулою ;
- 4) зобразити залежність графічно та зробити відповідні висновки.

№ підприємства	Чисельність бухгалтерів, осіб	Товарооборот, тис. грн.
1	6	8,2
2	5	8,2
3	7	9,4
4	17	9,5
5	8	12,4
6	6	16,2
7	9	16,8
8	8	15,9
9	7	16,4
10	17	17,2

Завдання 10.6. Маємо наступні дані про продуктивність корів і собівартість 1 ц молока в господарстві:

Роки	2006	2007	2008	2009	2010	2011	2012
Продуктивність 1 корови, ц	24,3	25,5	23,3	28,9	24,5	26,2	31,5
Собівартість 1 ц молока, грн.	154,5	162,0	174,5	192,0	236,5	285,8	302,0

- Визначити: 1) рівняння регресії, що виражає зв'язок між продуктивністю корів і собівартістю 1 ц молока;
- 2) щільність зв'язку між продуктивністю корів і собівартістю 1 ц молока;
 - 3) значимість коефіцієнта кореляції за критерієм Ст'юдента;
 - 4) коефіцієнт детермінації.
- Зробити висновки. Побудувати графік залежності.

Завдання 10.6. Маємо наступні дані по 8 заводам галузі:

№ заводу	1	2	3	4	5	6	7	8
Випуск продукції А, тис.шт.	12	20	28	25	18	15	9	23
Випуск продукції Б, тис. м ²	12	15	19	30	25	18	20	21
Витрати електроенергії, тис. кВт-год.	130	150	120	160	150	120	135	145

Визначити: 1) рівняння регресії, що виражає зв'язок між випуском двох видів продукції та витратами електроенергії;

2) парні та часткові коефіцієнти кореляції;

3) множинні коефіцієнти кореляції та детермінації;

4) значимість множинного коефіцієнта кореляції.

Зробити висновки.

Тести для самоконтролю.

1. Для якого з видів зв'язку є характерною повна відповідність між факторною і результативною ознаками?

А) для регресійного

Б) для кореляційного

В) для стохастичного

Г) для функціонального

2. Який показник, вказує на скільки одиниць змінюється середнє значення результативної ознаки зі збільшенням факторної ознаки на одиницю?

А) коефіцієнт регресії

Б) коефіцієнт кореляції

В) коефіцієнт детермінації

Г) коефіцієнт зростання

3. Параметр a_0 у рівнянні множинної регресії характеризує:

А) щільність кореляційного зв'язку

Б) вплив одного з факторів при елімінаванні іншого

В) вирівняний рівень ряду динаміки

Г) не інтерпретується

4. Охарактеризуйте щільність зв'язку при $r = -0,26$

А) слабкий

Б) середній

В) щільний

Г) функціональний

5. Коефіцієнт регресії дорівнює $(-0,9)$. Це означає, що:

А) при зменшенні факторної ознаки на одиницю результативна зростає на 0,9

Б) при збільшенні факторної ознаки на одиницю результативна зростає на 0,9

В) між факторною та результативною ознакою існує щільний прямий зв'язок

Г) між факторною та результативною ознакою існує щільний зворотний зв'язок

6. Яке з наведених значень лінійного коефіцієнта кореляції відповідає прямому зв'язку між ознаками x та y ?

А) $-0,85$

Б) $-0,74$

В) $0,69$

Г) $-0,98$

7. Зв'язок між корелюючими величинами, при якому факторна ознака змінюється в тому самому напрямі, що й результативна, називається:

- А) щільний
- Б) пропорційний
- В) прямий
- Г) зворотній

8. Коефіцієнт, що показує частку варіації результативного признаку під впливом факторного признаку,

- А) коефіцієнт кореляції
- Б) коефіцієнт регресії
- В) коефіцієнт детермінації
- Г) коефіцієнт варіації

9. Параметр a_1 у рівнянні множинної регресії характеризує:

- А) вплив одного з факторів при елімінуванні іншого
- Б) абсолютний приріст рівнів ряду динаміки
- В) щільність зв'язку
- Г) не інтерпретується

10. Охарактеризуйте щільність зв'язку при $r = - 0,55$

- А) слабкий
- Б) середній
- В) щільний
- Г) функціональний

11. Яке з наведених значень лінійного коефіцієнта кореляції відповідає найбільш щільному зв'язку між ознаками x та y ?

- А) 0,7
- Б) - 0,5
- В) - 0,9
- Г) 0,8

12. Зв'язок між корелюючими величинами, при якому із збільшенням факторної ознаки результативна ознака зменшується, називається:

- А) щільний
- Б) пропорційний
- В) прямий
- Г) зворотній

13. Щільність зв'язку при криволінійній залежності визначають за допомогою:

- А) індексу кореляції
- Б) коефіцієнту кореляції
- В) коефіцієнту детермінації
- Г) коефіцієнту регресії

14. Параметр a_2 у рівнянні множинної регресії характеризує:

- А) вплив одного з факторів при елімінуванні іншого
- Б) середнє прискорення зміни рівня досліджуваного явища
- В) щільність зв'язку
- Г) не інтерпретується

15. Охарактеризуйте щільність зв'язку при $r = - 0,09$:

- А) слабкий
- Б) середній
- В) щільний
- Г) відсутній

16. Кореляція між факторною і результативною ознаками називається:

- А) одинична
- Б) парна
- В) пропорційна
- Г) множинна

17. Кореляційний зв'язок, що характеризується рівномірним збільшенням або зменшенням результативної ознаки під впливом відповідної зміни факторної ознаки, називається:

- А) рівномірний
- Б) пропорційний
- В) прямолінійний
- Г) криволінійний

18. Яке з наведених значень лінійного коефіцієнта кореляції відповідає найбільш щільному зв'язку між ознаками x та y ?

- А) - 0,5
- Б) 0,8
- В) 0,7
- Г) - 0,9

19. Рівняння, за допомогою яких визначають статистичний зв'язок між корелюючими величинами, називають

- А) рівняння тренду
- Б) рівняння кореляції
- В) рівняння регресії
- Г) рівняння детермінації

20. Параметр a_0 у рівнянні регресії характеризує:

А) середній щорічний абсолютний приріст рівнів вирівняного ряду динаміки

- Б) зміну результативної ознаки при кожній зміні факторної на одиницю
- В) вирівняний рівень ряду динаміки
- Г) не інтерпретується

21. Охарактеризуйте щільність зв'язку при $r = - 0,75$

- А) слабкий
- Б) середній
- В) щільний
- Г) відсутній

22. Яке з наведених значень лінійного коефіцієнта кореляції відповідає оберненому зв'язку між ознаками x та y ?

- А) 0,9
- Б) - 0,8

В) 0,7

Г) 0,5

23. Залежність результативної ознаки від двох і більше факторних ознак називається:

А) функціональний зв'язок

Б) парна кореляція

В) множинна кореляція

Г) балансова ув'язка

24. Коефіцієнт кореляції дорівнює (- 0,8). Це означає, що:

А) між факторною та результативною ознакою існує щільний зворотний зв'язок

Б) між факторною та результативною ознакою існує щільний прямий зв'язок

В) при збільшенні факторної ознаки на одиницю результативна зростає на 0,8

Г) при зменшенні факторної ознаки на одиницю результативна зростає на 0,8

25. Коефіцієнт регресії дорівнює 0,9. Це означає, що:

А) при збільшенні факторної ознаки на одиницю результативна зростає на 0,9

Б) між факторною та результативною ознакою існує щільний зв'язок

В) результативна ознака на 90% залежить від факторної, а на 10% - від інших факторів, не включених у модель

Г) результативна ознака на 0,9% залежить від факторної, а на 99,1% - від інших факторів, не включених у модель

26. Кореляційний зв'язок, при якому рівним змінам середніх значень факторної ознаки відповідають нерівні зміни середніх значень результативної ознаки, називається:

А) пропорційний

Б) прямолінійний

В) криволінійний

Г) нерівномірний

27. Параметр a_1 у рівнянні регресії характеризує:

А) зміну результативної ознаки при кожній зміні факторної на одиницю

Б) середній щорічний абсолютний приріст рівнів вирівняного ряду динаміки

В) вирівняний рівень ряду динаміки

Г) не інтерпретується

28. Коефіцієнт a_0 дорівнює 0,99. Це означає, що:

А) при збільшенні факторної ознаки на одиницю результативна зростає на 0,99

Б) між факторною та результативною ознакою існує щільний зв'язок

В) результативна ознака на 99% залежить від факторної, а на 1% - від інших факторів, не включених у модель

Г) не інтерпретується

29. Яке з наведених значень лінійного коефіцієнта кореляції відповідає найбільш щільному зв'язку між ознаками x та y ?

А) - 0,98

- Б) 0,69
- В) - 0,85
- Г) 0,74

31. При прямому зв'язку між корелюючими ознаками коефіцієнт регресії має:

- А) додатне значення
- Б) від'ємне значення
- В) дорівнює 0
- Г) напрям зв'язку не впливає на знак коефіцієнта

32. Охарактеризуйте щільність зв'язку при $r=1$:

- А) слабкий
- Б) середній
- В) щільний
- Г) функціональний

33. Щільність зв'язку при криволінійній залежності визначають за допомогою:

- А) індексу кореляції
- Б) коефіцієнту детермінації
- В) коефіцієнту кореляції
- Г) коефіцієнту регресії

34. При зворотному зв'язку між корелюючими ознаками коефіцієнт регресії має:

- А) додатне значення
- Б) від'ємне значення
- В) дорівнює 0
- Г) напрям зв'язку не впливає на знак коефіцієнта

35. Коефіцієнт кореляції дорівнює 0,5. Це означає, що:

- А) між факторною та результативною ознакою існує середній зв'язок
- Б) результативна ознака на 0,5% залежить від факторної, а на 99,5% - від інших факторів, не включених у модель
- В) при збільшенні факторної ознаки на одиницю результативна зростає на 0,5
- Г) не має економічного змісту

36. Параметр a_2 у рівнянні множинної регресії характеризує:

- А) вплив одного з факторів при елімінаванні іншого
- Б) середнє прискорення зміни рівня досліджуваного явища
- В) щільність зв'язку
- Г) не інтерпретується

37. Яке з наведених значень лінійного коефіцієнта кореляції відповідає найбільш щільному прямому зв'язку між ознаками x та y ?

- А) 0,74
- Б) - 0,98
- В) - 0,85
- Г) 0,69

38. Охарактеризуйте щільність зв'язку при $r=0,6$:

- А) слабкий

- Б) середній
- В) щільний
- Г) функціональний

39. Який показник, вказує на скільки одиниць змінюється середнє значення результативної ознаки зі збільшенням факторної ознаки на одиницю?

- А) коефіцієнт регресії
- В) коефіцієнт детермінації
- Б) коефіцієнт кореляції
- Г) коефіцієнт зростання

40. Параметр a_0 у рівнянні регресії характеризує:

- А) середній щорічний абсолютний приріст рівнів вирівняного ряду динаміки
- Б) зміну результативної ознаки при кожній зміні факторної на одиницю
- В) вирівняний рівень ряду динаміки
- Г) не інтерпретується

41. Яке з наведених значень лінійного коефіцієнта кореляції відповідає найбільш щільному оберненому зв'язку між ознаками x та y ?

- А) - 0,8
- Б) 0,9
- В) - 0,5
- Г) 0,7

42. Охарактеризуйте щільність зв'язку при $r=0,25$:

- А) слабкий
- Б) середній
- В) щільний
- Г) відсутній

43. Коефіцієнт детермінації дорівнює 90. Це означає, що:

- А) результативна ознака на 90% залежить від факторної, а на 10% - від інших факторів, не включених у модель
- Б) між факторною та результативною ознакою існує щільний зв'язок
- В) при збільшенні факторної ознаки на одиницю результативна зростає на 90
- Г) не має економічного змісту

44. Яке з наведених значень лінійного коефіцієнта кореляції відповідає прямому зв'язку між ознаками x та y ?

- А) 0,69
- Б) - 0,74
- В) - 0,98
- Г) - 0,85

11. ІНДЕКСИ

11.1 Індеси та їх класифікація

Індеси - це відносні показники, що характеризують результат зміни у соціально-економічному явищі, а також порівняння двох сукупностей які складаються з різнорідних елементів у часі, у просторі або порівняно з планом.

Індеси обчислюють у вигляді:

- коефіцієнта, який показує, у скільки разів величина порівняння (чисельник) більша чи менша, ніж база порівняння (знаменник);
- процента, що показує, скільки відсотків від рівня бази порівняння складає величина порівняння

Індеси використовують:

- для характеристики рівня виконання плану;
- для характеристики зміни явищ у часі;
- для порівняння явищ у просторі;
- для аналізу впливу окремих факторів на зміну розміру.

Для характеристики рівня виконання плану або зміни явищ у часі величину порівняння називають **величиною звітного періоду** (позначають **1**), а базу порівняння - **величиною базисного (планового) періоду** (позначають **0**).

Залежно від змісту і характеру явищ, що вивчаються, розрізняють індеси:

- **кількісних (об'ємних) показників**, які можна підсумувати (індекси фізичного обсягу, площ, чисельності робітників, поголів'я тварин, витрат кормів і т.д.);

- **якісних показників**, які характеризуються середніми величинами і не можуть бути підсумовані (індекси урожайності, ціни, собівартості, продуктивності праці, продуктивності тварин і т.д.).

За ступенем охоплення елементів досліджуваного явища розрізняють:

- **індивідуальні індеси** - виражають співвідношення величин окремого явища складної сукупності;

- **групові індеси** - виражають співвідношення частини елементів складної сукупності;

- **загальні індеси** - характеризують зміни складного явища, що включає окремі елементи, які не можна підсумувати.

Індеси обчислені порівнянням даних кожного періоду з даними попереднього періоду, називають **ланцюговими**:

$$i = \frac{y_1}{y_0}; i = \frac{y_2}{y_1}; \dots i = \frac{y_n}{y_{n-1}};$$

Індекси, обчислені порівнянням даних кожного періоду з даними будь-якого періоду, взятого за постійну базу порівняння, називають **базисними**:

$$i = \frac{y_1}{y_0}; i = \frac{y_2}{y_0}; \dots i = \frac{y_n}{y_0};$$

Між ланцюговими і базисними індексами існує певний зв'язок, що дозволяє здійснити перехід від одного виду індексів до іншого:

- добуток ланцюгових індексів дорівнює базисному індексу останнього періоду:

$$\frac{y_1}{y_0} * \frac{y_2}{y_1} * \frac{y_3}{y_2} * \frac{y_4}{y_3} = \frac{y_4}{y_0}$$

- ланцюгові індекси визначаються з базисних діленням відповідного базисного індексу на попередній базисний індекс:

$$\frac{y_4}{y_0} \div \frac{y_3}{y_0} = \frac{y_4}{y_3}$$

Індексний метод має свою *символіку*:

q - кількість виготовленої (проданої) продукції в натуральних вимірниках;

c - собівартість одиниці продукції;

p - ціна одиниці продукції;

t - витрати часу на одиницю продукції;

v - виробіток продукції в натуральному виразі на одного робітника або в одиницю часу;

w - виробіток продукції у вартісному виразі на одного робітника або в одиницю часу;

П - посівна площа;

У - урожайність культур.

11.2 Індивідуальні індекси

Якщо нас не цікавить структура досліджуваного явища і кількісна оцінка рівня в даних умовах порівнюється з такою ж конкретною величиною рівня цього явища в інших умовах, то відносна величина, що отримується при порівнянні рівнів, називається індивідуальним індексом.

Загальний вид індивідуального індексу:

кількісного показника:
$$i_f = \frac{f_1}{f_0}$$

якісного показника:
$$i_x = \frac{x_1}{x_0}$$

Приклади індивідуальних індексів:

індекс цін: $i_p = \frac{p_1}{p_0}$

індекс собівартості: $i_c = \frac{c_1}{c_0}$

індекс фізичного обсягу: $i_q = \frac{q_1}{q_0}$

індекс трудомісткості: $i_t = \frac{t_1}{t_0}$

індекс продуктивності праці: $i_v = \frac{v_1}{v_0}$

індекс продуктивності праці (по витратам праці): $i_v = \frac{t_0}{t_1}$

індекс урожайності: $i_y = \frac{y_1}{y_0}$

індекс площі: $i_{\pi} = \frac{\pi_1}{\pi_0}$ і т.д.

Задача.11.1. На основі даних про реалізацію товарів підприємством розрахувати індивідуальні індекси фізичного обсягу та цін по окремих видах товарів.

Розв'язання.

Розрахунок індивідуальних індексів показав, що в звітному періоді кількість проданого товару А збільшилася на 8,9%, тоді як обсяги продажів товарів Б та В знизилися відповідно на 14,3% та 13,6%. Крім того в звітному періоді ціни на товари А, Б та В зросли на 11,1%, 11,8% та 16,7% відповідно (табл. 11.1).

Таблиця 11.1

Розрахунок індивідуальних індексів

Товар	Кількість проданого товару, кг		Ціна за 1 кг, грн.		Індивідуальні індекси	
	Базисний період	Звітний період	Базисний період	Звітний період	Фізичного обсягу	Цін
	q_0	q_1	p_0	p_1	$i_q = \frac{q_1}{q_0}$	$i_p = \frac{p_1}{p_0}$
А	9000	9800	1,62	1,8	1,089	1,111
Б	700	600	2,2	2,46	0,857	1,118
В	1100	950	1,8	2,1	0,864	1,167

11.3 Загальні індекси

Якщо відомо, що досліджуване явище неоднорідне і порівняння рівнів можна провести тільки після приведення їх до загальної основи, економічний аналіз виконують за допомогою загальних індексів.

В залежності від наявних даних загальні індекси можуть обчислюватись у формі *агрегатного* або *середнього* індексу.

Агрегатний індекс - основна форма загальних індексів, що являє собою відношення сум добутків індексованих величин (величин, зміну яких визначають індексом) та їх співвимірників (ознак, які застосовують, як постійну величину).

Примітка. В добутках чисельника та знаменника агрегатного індексу змінюється індексована величина, а співвимірники - незмінні.

Примітка. Індексовані величини в формулі зазвичай пишуть на першому місці, а співвимірники - на другому.

Загальний вид агрегатних індексів:

Індекс явища:

$$I_{xf} = \frac{\sum x_1 f_1}{\sum x_0 f_0} - \text{показує зміну явища у часі}$$

Індекс кількісного показника:

$$I_f = \frac{\sum f_1 x_0}{\sum f_0 x_0} - \text{показує зміну явища за рахунок зміни кількісної ознаки}$$

Індекс якісного показника:

$$I_x = \frac{\sum x_1 f_1}{\sum x_0 f_1} - \text{показує зміну явища за рахунок зміни якісної ознаки}$$

Приклади агрегатних індексів:

Індекс товарообороту: $I_{pq} = \frac{\sum p_1 q_1}{\sum p_0 q_0}$

Індекс цін: $I_p = \frac{\sum p_1 q_1}{\sum p_0 q_1}$

Індекс фізичного обсягу (проданого товару): $I_q = \frac{\sum q_1 p_0}{\sum q_0 p_0}$

Індекс валового збору: $I_{\Pi Y} = \frac{\sum \Pi_1 Y_1}{\sum \Pi_0 Y_0}$

$$\text{Індекс урожайності: } I_y = \frac{\sum y_1 \Pi_1}{\sum y_0 \Pi_1}$$

$$\text{Індекс площі: } I_{\Pi} = \frac{\sum \Pi_1 y_0}{\sum \Pi_0 y_0}$$

$$\text{Індекс виробничих витрат: } I_{cq} = \frac{\sum c_1 q_1}{\sum c_0 q_0}$$

$$\text{Індекс собівартості: } I_c = \frac{\sum c_1 q_1}{\sum c_0 q_1}$$

$$\text{Індекс фізичного обсягу (виробленої продукції): } I_q = \frac{\sum q_1 c_0}{\sum q_0 c_0} \text{ і т.д.}$$

Задача 11.2. Визначити агрегатні індекси товарообороту, цін та фізичного обсягу (проданого товару) на основі наступних даних:

Товар	Кількість проданого товару, кг		Ціна за 1 кг, грн.		Товарооборот, грн.		
	Базисний період	Звітний період	Базисний період	Звітний період	Базисний	Звітний	Умовний
	q_0	q_1	p_0	p_1	$p_0 q_0$	$p_1 q_1$	$q_1 p_0$
А	9000	9800	1,62	1,8	14580	17640	15876
Б	700	600	2,2	2,46	1540	1476	1320
В	1100	950	1,8	2,1	1980	1995	1710
Разом	10800	11350	х	х	18100	21111	18906

Розв'язання.

Індекс товарообороту:

$$I_{pq} = \frac{\sum p_1 q_1}{\sum p_0 q_0} = \frac{\sum 21111}{\sum 18100} = 1,166, \text{ або } 116,6\%$$

Абсолютна зміна товарообороту:

$$\Delta_{pq} = \sum p_1 q_1 - \sum p_0 q_0 = 21111 - 18100 = 3011 (\text{грн.})$$

Виручка від реалізації продукції у звітному році порівняно з базисним збільшилася на 16,6% або на 3011 грн.

Індекс цін:

$$I_p = \frac{\sum p_1 q_1}{\sum p_0 q_1} = \frac{21111}{18906} = 1,117, \text{ або } 111,7\%$$

Абсолютна зміна товарообороту за рахунок ціни:

$$\Delta_p = \sum p_1 q_1 - \sum p_0 q_1 = 21111 - 18906 = 2205 (\text{грн.})$$

За рахунок зростання цін виручка від реалізації продукції у звітному році порівняно з базисним збільшилася на 11,7% або на 2205 грн.

Індекс фізичного обсягу (проданого товару):

$$I_q = \frac{\sum q_1 p_0}{\sum q_0 p_0} = \frac{18906}{18100} = 1,045, \text{ або } 104,5\%$$

Абсолютна зміна товарообороту за рахунок фізичного обсягу товару:

$$\Delta_q = \sum q_1 p_0 - \sum q_0 p_0 = 18906 - 18100 = 806 (\text{грн.})$$

За рахунок зростання кількості товару виручка від реалізації продукції у звітному році порівняно з базисним збільшилася на 4,5% або на 806 грн.

Таким чином, індексний аналіз показав, що товарооборот у звітному році порівняно з базисним збільшився на 16,6% або на 3011 грн., в тому числі за рахунок зростання цін він збільшився на 11,7% або на 2205 грн., а за рахунок зростання кількості товару - на 4,5% або на 806 грн.

Взаємозв'язок агрегатних індексів.

Економічні індекси тісно пов'язані між собою і утворюють індексні системи.

$$\frac{\sum x_1 f_1}{\sum x_0 f_1} * \frac{\sum f_1 x_0}{\sum f_0 x_0} = \frac{\sum x_1 f_1}{\sum x_0 f_0} \quad \text{або} \quad I_x * I_f = I_{xf}$$

$$(\sum x_1 f_1 - \sum x_0 f_1) + (\sum f_1 x_0 - \sum f_0 x_0) = \sum x_1 f_1 - \sum x_0 f_0$$

або

$$\Delta_x + \Delta_f = \Delta_{xf}$$

Задача 11.3. Провести балансову ув'язку індексів, одержаних в задачі 11.2.

Розв'язання.

Індекс цін пов'язаний з індексом фізичного обсягу (проданого товару) утворюючи індексну систему:

$$I_p * I_q = I_{pq}$$

$$1,117 * 1,045 = 1,116$$

$$\Delta_p + \Delta_q = \Delta_{pq}$$

$$2205 + 806 = 3011$$

Балансова ув'язка підтверджує правильність розрахунків індексів в задачі 11.2.

Середні індекси

У деяких випадках загальні індекси обчислюють як *середні* з індивідуальних індексів окремих елементів: агрегатний індекс перетворюють у середні з індивідуальних індексів, підставляючи у чисельник або знаменник агрегатного індексу замість індексованого показника його вираз, що виводиться з формули відповідного індивідуального індексу.

Примітка. Якщо таку заміну роблять у чисельнику - агрегатний індекс перетворюється у середній арифметичний, якщо заміна відбувається у знаменнику - у середній гармонійний.

Задача 11.4. Обчислити середній арифметичний індекс фізичного обсягу (проданого товару) товарообороту на основі наступних даних:

Товарні групи	Товарооборот в базисному періоді, тис. грн.	Індивідуальні індекси фізичного обсягу (проданого товару) товарообороту
		$P_0 q_0$
Продовольчі товари А	9600	1,045
Непродовольчі товари Б	12750	0,931

Розв'язання.

З формули індивідуального індексу фізичного обсягу $i_q = \frac{q_1}{q_0}$ випливає:

$$q_1 = i_q * q_0$$

Підставимо у чисельник агрегатного індексу замість q_1 величину $i_q * q_0$, тоді дістанемо середній арифметичний індекс фізичного обсягу (проданого товару) товарообороту:

$$I_q = \frac{\sum i_q q_0 P_0}{\sum q_0 P_0} = \frac{1,045 * 9600 + 0,931 * 12750}{9600 + 12750} = 0,98$$

Обчислений індекс показує, що за рахунок зменшення кількості товару товарооборот у звітному періоді знизився на 2%.

Задача 11.5. Обчислити середній гармонійний індекс цін:

Товарні групи	Товарооборот в звітному періоді, тис. грн.	Індивідуальні індекси цін
	p_0q_0	$i_q = \frac{p_1}{p_0}$
Молоко	480	1,24
Сметана	360	1,22
Сир	150	1,36

Розв'язання.

З формули індивідуального індексу цін $i_p = \frac{p_1}{p_0}$ випливає: $p_0 = \frac{p_1}{i_p}$

Підставимо у знаменник агрегатного індексу замість p_0 величину $\frac{p_1}{i_p}$,

тоді дістанемо середній гармонійний індекс цін:

$$I_p = \frac{\sum p_1q_1}{\sum \frac{p_1q_1}{i_p}} = \frac{480 + 360 + 150}{\frac{480}{1,24} + \frac{360}{1,22} + \frac{150}{1,36}} = 1,249$$

Обчислений індекс показує, що за рахунок зростання цін виручка від реалізації продукції у звітному році збільшилася на 24,9%.

11.4 Аналіз співвідношень середніх рівнів

Економічні явища часто характеризуються за допомогою середніх величин. Зокрема, всі якісні показники, як правило, виражаються у вигляді середніх: середня ціна одиниці продукції, середня собівартість одиниці виробу, середня заробітна плата одного робітника, виробіток продукції в середньому на одного працівника, середня трудомісткість одного виробу і т.п. Для вивчення динаміки таких показників у статистичній практиці застосовуються індекси середніх рівнів (середніх величин).

Загальний вид індексів середніх рівнів:

Індекс змінного складу має вигляд:

$$I_{зм.скл.} = \bar{x}_1 \div \bar{x}_0 = \frac{\sum x_1f_1}{\sum f_1} \div \frac{\sum x_0f_0}{\sum f_0}$$

Цей індекс характеризує вплив на динаміку середніх показників зміни усередненої ознаки та структури явища.

Індекс фіксованого складу має вигляд:

$$I_{ф.скл.} = \frac{\sum x_1 f_1}{\sum f_1} \div \frac{\sum x_0 f_1}{\sum f_1} = \frac{\sum x_1 f_1}{\sum x_0 f_1}$$

Цей індекс характеризує вплив на динаміку середніх показників зміни усередненої ознаки. У даному індексі вплив структурного фактора виключено.

Індекс структури розраховують за формулою:

$$I_{стр.} = I_{зм.скл.} \div I_{ф.скл.}$$

Даний індекс характеризує вплив на динаміку середніх показників зміни структури явища.

Задача 11.6. На основі даних про виробництво зерна в господарстві обчислити індекси середньої урожайності зернових:

Культура	Базисний період			Звітний період		
	Площа, га	Урожайність, ц/га	Валовий збір, ц	Площа, га	Урожайність, ц/га	Валовий збір, ц
Озима пшениця	500	50	25000	600	52	31200
Ярий ячмінь	500	40	20000	400	42	16800
Разом	1000	х	45000	1000	х	48000

Індекс змінного складу:

$$I_{зм.скл.} = \bar{Y}_1 \div \bar{Y}_0 = \frac{\sum Y_1 \Pi_1}{\sum \Pi_1} \div \frac{\sum Y_0 \Pi_0}{\sum \Pi_0}$$

$$I_{зм.скл.} = \frac{48000}{1000} \div \frac{45000}{1000} = 48 \div 45 = 1,067, \text{ або } 106,7\%$$

У звітному періоді середня урожайність зернових порівняно з базисним періодом зросла на 6,7%.

Індекс фіксованого складу:

$$I_{ф.скл.} = \frac{\sum Y_1 \Pi_1}{\sum \Pi_1} \div \frac{\sum Y_0 \Pi_1}{\sum \Pi_1} = \frac{\sum Y_1 \Pi_1}{\sum Y_0 \Pi_1}$$

$$I_{ф.скл.} = \frac{52 * 600 + 42 * 400}{50 * 600 + 40 * 400} = 1,043, \text{ або } 104,3\%$$

За рахунок підвищення урожайності окремих культур середня урожайність зернових у звітному періоді зросла на 4,3%.

Індекс структури:

$$I_{стр.} = I_{зм.скл.} \div I_{ф.скл.} = 1,067 \div 1,043 = 1,023, \text{ або } 102,3\%$$

За рахунок змін у структурі посівів середня урожайність зернових у звітному періоді зросла на 2,3%.

Завдання для самостійного розв'язання.

Завдання 11.1. Маємо наступні дані про виробництво зернових у господарстві:

Групи культур	Площа, га		Урожайність, ц/га	
	Базисний рік	Звітний рік	Базисний рік	Звітний рік
Озимі зернові	900	1100	32,5	26,5
Ярові зернові	480	450	20,0	22,0

Визначити:

- 1) індивідуальні індекси валового збору, урожайності і посівної площі;
- 2) загальні (агрегатні) індекси валового збору, урожайності, площі;
- 3) абсолютний приріст валового збору, в тому числі за рахунок кожного з його факторів.

Зробити висновки.

Завдання 11.2. Маємо наступні дані про реалізацію плодів:

Види продукції	Обсяг продажу, кг		Ціна за 1 кг, грн.	
	Базисний рік	Звітний рік	Базисний рік	Звітний рік
Яблука	5000	4500	12,5	14,5
Персики	6200	6000	6,5	9,0
Сливи	5850	6500	4,5	5,0

Визначити:

- 1) агрегатні індекси цін, фізичного обсягу і товарообороту;
- 2) абсолютну зміну товарообороту за рахунок кожного з факторів;
- 3) індекси ціни змінного і постійного (фіксованого) складу, структури реалізації.

Зробити висновки.

Завдання 11.3. Є наступні дані по підприємству:

Професія	Базисний період		Звітний період	
	Чисельність робітників, осіб	Середня заробітна плата за місяць, грн.	Чисельність робітників, осіб	Середня заробітна плата за місяць, грн.
Маляр	35	3200	40	3320
Штукатур	55	3350	55	3440
Монтажник	40	4500	45	4960

Визначити:

- 1) індивідуальні індекси заробітної плати та чисельності окремих категорій робітників;

2) індекси середньої заробітної плати змінного й фіксованого складу та індекс структури.

Зробити висновки.

Завдання 11.4. Маємо дані про обсяги виробництва продукції і її собівартість на підприємстві за два періоди:

Товарні групи	Собівартість одиниці продукції, грн.		Виробництво продукції, тис. т	
	Базисний період	Звітний період	Базисний період	Звітний період
А	180	185	200	180
Б	130	135	250	260
В	160	150	170	170

Визначити:

1) загальні індекси виробничих витрат, собівартості, фізичного обсягу виробленої продукції;

2) абсолютний розмір зміни витрат виробництва: усього, у тому числі за рахунок зміни собівартості продукції й фізичного обсягу продукції.

Зробити висновки.

Завдання 11.5. За даними про урожайність озимої пшениці у господарстві, визначте базисні і ланцюгові індекси. Зробіть перевірку зв'язку між ланцюговими і базисними індексами:

Роки	2008	2009	2010	2011	2012
Урожайність, ц/га	24,0	22,5	24,8	27,5	26,0

Тести для самоконтролю.

1. Величина базисного періоду в індексному аналізі позначається:

- А) 0
- Б) 1
- В) б
- Г) і

2. Індекси, що характеризують зміни складного явища, що включає окремі елементи, які не можна підсумувати, називаються:

- А) загальні
- Б) індивідуальні
- В) сукупні
- Г) групові

3. Індекс, що характеризує вплив на динаміку середніх показників зміни усередненої ознаки та структури явища, називається:

- А) індекс змінного складу

Б) індекс фіксованого складу

В) індекс структури

Г) індекс структурних зрушень

4. Абсолютна зміна товарообороту за рахунок зміни обсягу товару знаходиться:

А) $\Delta = \sum q_1 p_0 - \sum q_0 p_0$

Б) $\Delta = \sum p_0 q_0 - \sum p_1 q_1$

В) $\Delta = \sum p_1 q_1 - \sum p_0 q_0$

Г) $\Delta = \sum q_1 p_1 - \sum q_1 p_0$

5. Індивідуальний індекс урожайності має вигляд:

А) $i = \frac{y_1}{y_0}$

Б) $i = \frac{c_1}{c_0}$

В) $i = \frac{q_1}{q_0}$

Г) $i = \frac{p_1}{p_0}$

6. Агрегатний індекс площі має вигляд:

А) $I = \frac{\sum \Pi_1}{\sum \Pi_0}$

Б) $I = \frac{\sum y_1 \Pi_1}{\sum y_0 \Pi_1}$

В) $I = \frac{\sum \Pi_1 y_0}{\sum \Pi_0 y_0}$

Г) $I = \frac{\sum \Pi_1 y_1}{\sum \Pi_0 y_0}$

7. Величина звітного періоду в індексному аналізі позначається:

А) 0

Б) 1

В) \bar{b}

Г) i

8. Абсолютна зміна товарообороту за рахунок зміни ціни:

А) $\Delta = \sum p_1 q_1 - \sum p_0 q_0$

Б) $\Delta = \sum q_1 p_0 - \sum q_0 p_0$

$$\text{В) } \Delta = \sum p_1 q_1 - \sum p_0 q_1$$

$$\text{Г) } \Delta = \sum p_0 q_0 - \sum p_1 q_1$$

9. Агрегатний індекс урожайності має вигляд:

$$\text{А) } I = \frac{\sum \Pi_1 Y_1}{\sum \Pi_0 Y_0}$$

$$\text{Б) } I = \frac{\sum \Pi_1 Y_0}{\sum \Pi_0 Y_0}$$

$$\text{В) } I = \frac{\sum Y_1 \Pi_1}{\sum Y_0 \Pi_1}$$

$$\text{Г) } I = \frac{Y_1}{Y_0}$$

10. Індивідуальний індекс собівартості має вигляд:

$$\text{А) } i = \frac{y_1}{y_0}$$

$$\text{Б) } i = \frac{c_1}{c_0}$$

$$\text{В) } i = \frac{q_1}{q_0}$$

$$\text{Г) } i = \frac{p_1}{p_0}$$

11. Індокси цін, собівартості, продуктивності праці, урожайності тощо, відносяться до:

- А) індоксів якісних показників
- Б) індоксів об'ємних показників
- В) індоксів обсягу явища
- Г) територіальних індоксів

12. Згідно символіки індоксного методу фізичний обсяг продукції позначається:

- А) q
- Б) p
- В) y
- Г) п

13. Індокси, що виражають співвідношення величин окремого явища складної сукупності, називаються:

- А) загальні
- Б) сукупні
- В) індивідуальні

Г) групові

14. Індивідуальний індекс цін має вигляд:

А) $i = \frac{Y_1}{Y_0}$

Б) $i = \frac{c_1}{c_0}$

В) $i = \frac{q_1}{q_0}$

Г) $i = \frac{p_1}{p_0}$

15. Агрегатний індекс виробничих витрат має вигляд:

А) $I = \frac{\sum c_1 q_1}{\sum c_0 q_1}$

Б) $I = \frac{\sum c_1 q_1}{\sum c_0 q_0}$

В) $I = \frac{\sum c_0 q_0}{\sum c_1 q_1}$

Г) $I = \frac{\sum q_1 c_0}{\sum q_0 c_0}$

16. Індекс, що характеризує вплив на динаміку середніх показників зміни складу та структури явища, називається:

А) складний індекс

Б) індекс змінного складу

В) індекс структури

Г) індекс фіксованого складу

17. Абсолютна зміна валового збору за рахунок зміни урожайності:

А) $\Delta = \sum Y_1 P_1 - \sum Y_0 P_1$

Б) $\Delta = \sum Y_0 P_0 - \sum Y_1 P_1$

В) $\Delta = \sum P_1 Y_0 - \sum P_0 Y_0$

Г) $\Delta = \sum Y_1 P_1 - \sum Y_0 P_0$

18. Згідно символіки індексного методу ціна одиниці продукції позначається:

А) q

Б) p

В) t

Г) с

20. Індеси фізичного обсягу, розміру посівних площ, поголів'я тварин, витрат кормів тощо, відносяться до:

А) індесів об'ємних показників

Б) індесів якісних показників

В) територіальних індесів

Г) індесів середніх рівнів

21. Індивідуальний індес фізичного обсягу продукції має вигляд:

А) $i = \frac{v_1}{v_0}$

Б) $i = \frac{t_1}{t_0}$

В) $i = \frac{q_1}{q_0}$

Г) $i = \frac{p_1}{p_0}$

22. Добуток ланцюгових індесів дорівнює:

А) базисному індесу останнього періоду

Б) сумі базисних індесів

В) добутку базисних індесів

Г) кореню з добутку базисних індесів

23. Агрегатний індес товарообороту має вигляд:

А) $I = \frac{q_1}{q_0}$

Б) $I = \frac{\sum p_1 q_1}{\sum p_0 q_1}$

В) $I = \frac{\sum p_1 q_1}{\sum p_0 q_0}$

Г) $I = \frac{\sum q_1 p_0}{\sum q_0 p_0}$

24. Індес, що характеризує вплив на динаміку середніх показників зміни тільки усередненої ознаки, називається:

А) індес фіксованого складу

Б) індес структури

В) індес змінного складу

Г) індес структурних зрушень

25. Абсолютна зміна валового збору за рахунок зміни площі:

А) $\Delta = \sum Y_1 \Pi_1 - \sum Y_0 \Pi_1$

Б) $\Delta = \sum \Pi_1 Y_1 - \sum \Pi_0 Y_0$

В) $\Delta = \sum \Pi_1 Y_0 - \sum \Pi_0 Y_0$

Г) $\Delta = \sum \Pi_0 Y_0 - \sum \Pi_1 Y_1$

26. Згідно символіки індексного методу собівартість одиниці продукції позначається:

А) w

Б) с

В) q

Г) p

27. Індеси, обчисленні порівнянням даних кожного періоду з даними попереднього періоду, називають:

А) динамічними

Б) ланцюговими

В) базисними

Г) загальними

28. Агрегатний індекс цін має вигляд:

А) $I = \frac{\sum q_1 p_0}{\sum q_0 p_0}$

Б) $I = \frac{\sum p_1 q_1}{\sum p_0 q_1}$

В) $I = \frac{p_1}{p_0}$

Г) $I = \frac{\sum p_1 q_1}{\sum p_0 q_0}$

29. Абсолютна зміна валового збору:

А) $\Delta = \sum \Pi_1 Y_0 - \sum \Pi_0 Y_0$

Б) $\Delta = \sum Y_1 \Pi_1 - \sum Y_0 \Pi_0$

В) $\Delta = \sum Y_0 \Pi_0 - \sum Y_1 \Pi_1$

Г) $\Delta = \sum Y_1 \Pi_1 - \sum Y_0 \Pi_1$

30. Індеси, обчисленні порівнянням даних кожного періоду з даними будь-якого періоду, взятого за постійну базу порівняння, називають:

А) динамічними

Б) базисними

- Б) основними
- Г) ланцюговими

31. Агрегатний індекс кількості проданого товару має вигляд:

$$A) I = \frac{\sum p_1 q_1}{\sum p_0 q_1}$$

$$B) I = \frac{\sum q_1 p_0}{\sum q_0 p_0}$$

$$B) I = \frac{\sum p_0 q_0}{\sum p_1 q_1}$$

$$Г) I = \frac{\sum p_1 q_1}{\sum p_0 q_0}$$

32. Згідно символіки індексного методу витрати часу на виробництво одиниці продукції позначається:

- А) w
- Б) q
- В) t
- Г) v

33. Абсолютна зміна витрат виробництва за рахунок зміни собівартості одиниці продукції:

$$A) \Delta = \sum c_0 q_0 - \sum c_1 q_1$$

$$B) \Delta = \sum c_1 q_1 - \sum c_0 q_1$$

$$B) \Delta = \sum c_1 q_1 - \sum c_0 q_0$$

$$Г) \Delta = \sum q_1 c_0 - \sum q_0 c_0$$

34. Індеси, що виражають співвідношення частини елементів складної сукупності, називаються:

- А) групові
- Б) загальні
- В) агрегатні
- Г) індивідуальні

35. Абсолютна зміна товарообороту визначається:

$$A) \Delta = \sum q_1 p_0 - \sum q_0 p_0$$

$$B) \Delta = \sum p_1 q_1 - \sum p_0 q_0$$

$$B) \Delta = \sum p_0 q_0 - \sum p_1 q_1$$

$$Г) \Delta = \sum p_1 q_1 - \sum p_0 q_1$$

36. Індивідуальний індекс продуктивності праці (по витратам праці) має вигляд:

$$\text{А) } i = \frac{v_1}{v_0}$$

$$\text{Б) } i = \frac{t_1}{t_0}$$

$$\text{В) } i = \frac{w_1}{w_0}$$

$$\text{Г) } i = \frac{t_0}{t_1}$$

37. Агрегатний індекс собівартості має вигляд:

$$\text{А) } I = \frac{\sum q_1 c_0}{\sum q_0 c_0}$$

$$\text{Б) } I = \frac{\sum c_1 q_1}{\sum c_0 q_0}$$

$$\text{В) } I = \frac{\sum c_1 q_1}{\sum c_0 q_1}$$

$$\text{Г) } I = \frac{\sum c_0 q_0}{\sum c_1 q_1}$$

ДОДАТКИ.

Додаток А.

Інтегральна функція нормального розподілу F(t)

$$F(t) = \frac{2}{\sqrt{2\pi}} \int_0^t e^{-\frac{t^2}{2}} dt$$

t	F(t)	t	F(t)	t	F(t)	t	F(t)
0,0	0,0000	1,0	0,6827	2,0	0,9545	3,0	0,9973
0,1	0,0797	1,1	0,7287	2,1	0,9643	3,1	0,9981
0,2	0,1585	1,2	0,7699	2,2	0,9722	3,2	0,9986
0,3	0,2358	1,3	0,8064	2,3	0,9786	3,3	0,9990
0,4	0,3108	1,4	0,8385	2,4	0,9836	3,4	0,9993
0,5	0,3829	1,5	0,8664	2,5	0,9876	3,5	0,9995
0,6	0,4515	1,6	0,8904	2,6	0,9907	3,6	0,9997
0,7	0,5161	1,7	0,9109	2,7	0,9931	3,7	0,9998
0,8	0,5763	1,8	0,9281	2,8	0,9949	3,8	0,9999
0,9	0,6319	1,9	0,9426	2,9	0,9963	3,9	0,9999

Значення критерію t-Ст'юдента

V*	P			V*	P		
	0,1	0,05	0,01		0,1	0,05	0,01
1	6,3130	12,7060	63,6560	23	1,7139	2,0687	2,8073
2	2,9200	4,3020	9,9240	24	1,7109	2,0639	2,7969
3	2,35340	3,1820	5,8400	25	1,7081	2,0595	2,7874
4	2,13180	2,7760	4,6040	26	1,7050	2,0590	2,7780
5	2,01500	2,5700	4,0321	27	1,7033	2,0518	2,7707
6	1,9430	2,4460	3,7070	28	1,7011	2,0484	2,7633
7	1,8946	2,3646	3,4995	29	1,6991	2,0452	2,7564
8	1,8596	2,3060	3,3554	30	1,6973	2,0423	2,7500
9	1,8331	2,2622	3,2498	40	1,6839	2,0211	2,7045
10	1,8125	2,2281	3,1693	50	1,6759	2,0086	2,6778
11	1,7950	2,2010	3,1050	60	1,6706	2,0003	2,6603
12	1,7823	2,1788	3,0845	70	1,6689	1,9944	2,6479
13	1,7709	2,1604	3,1123	80	1,6640	1,9900	2,6380
14	1,7613	2,1448	2,9760	90	1,6620	1,9867	2,6316
15	1,7530	2,1314	2,9467	100	1,6602	1,9840	2,6259
16	1,7450	2,1190	2,9200	120	1,6577	1,9719	2,6174
17	1,7396	2,1098	2,8982	150	1,6551	1,9759	2,6090
18	1,7341	2,1009	2,8784	200	1,6525	1,9719	2,6006
19	1,7291	2,0930	2,8609	250	1,6510	1,9695	2,5966
20	1,7247	2,0860	2,8453	300	1,6499	1,9679	2,5923
21	1,7200	2,0790	2,8310	400	1,6487	1,9659	2,5882
22	1,7117	2,0739	2,8188	500	1,6470	1,9640	2,7850

* V – ступінь вільності варіації

ПРЕДМЕТНИЙ ПОКАЖЧИК

А

Абсолютна величина, 93
- індивідуальна, 94
- сумарна (групова), 94
Абсолютне значення 1% приросту, 191
- середнє, 192
Абсолютний приріст, 191
- середній, 192
Аналітичне вирівнювання способом найменших квадратів, 201

Б

База порівняння, 95
Безпосередній облік, 10

В

Варіація, 127
Вибірковий метод, 162
Вибіркові оцінки
- інтервальні, 168
- точкові, 168
Відбір
- безповторний, 165
- власне випадковий (простий), 165
- груповий, 165
- індивідуальний, 165
- комбінований, 165, 166
- механічний, 165
- повторний, 165
- серійний (гніздовий), 166
- типовий (районований, розшарований), 166
Відносна величина, 94
- виконання плану (реалізації намірів, виконання договірних зобов'язань), 96
- динаміки, 97
- інтенсивності, 99

- координації, 99
- планового завдання (намірів, договірних зобов'язань), 95
- порівняння, 98
- рівня економічного розвитку, 100
- структури, 98
Вторинне групування, 79
- об'єднання первинних інтервалів, 79
- часткове перегрупування, 80

Г

Гістограма, 123
Графічний образ, 33

Д

Динаміка, 187
Дисперсія, 129
Диференціація розподілів, 148
Діаграма
- взаємозв'язку лінійна, 48
- взаємозв'язку точкова, 47
- динаміки лінійна, 44
- динаміки радіальна (пелюсткова), 46
- динаміки стовпчикова, 45
- динаміки фігурна, 45
- порівняння квадратна, 38
- порівняння кругова, 38
- порівняння стовпчикова, 35
- порівняння стрічкова (смугова), 37
- порівняння фігурна (зображувальна), 39
- структури кругова секторна, 41
- структури стовпчикова, 42
- структури стрічкова (смугова), 43
- взаємозв'язку, 47
- динаміки, 44
- порівняння, 35

- структури, 41
Документальний облік, 10

Е

Експлікація статистичного графіка, 33
Екстраполяція, 205

З

Звітність, 9
- децентралізована, 9
- поточна, 9
- поштова, 9
- річна, 9
- спеціалізована, 9
- термінова, 9
- типова (загальна), 9
- централізована, 9
Зв'язок
- зворотний, 218
- кореляційний (неповний, або статистичний), 217
- криволінійний, 218
- прямий, 218
- прямолінійний, 218
- функціональний (повний), 217
Згладжування за допомогою ковзної середньої, 199

І

Індекс кореляції, 225
Індекси статистичні, 239
- агрегатні, 242
- базисні, 240
- групові, 239
- загальні, 239, 242
- змінного складу, 246
- індивідуальні, 239, 240
- кількісних (об'ємних) показників, 239
- ланцюгові, 239

- середні, 245
- структури, 247
- фіксованого складу, 247
- якісних показників, 239
інтервал екстраполяції, 206
Інтервали групування, 66
- відкриті, 67
- закриті, 67
- нерівні (прогресивні), 67
- рівні, 66
Інтерполяція, 205

К

Квантилі (градієнти), 144
- децилі, 145
- квантилі, 145
- квінтилі, 146
- напівдецилі, 146
- перцентилі, 146
Коефіцієнт
- варіації, 130
- варіації лінійний, 130
- детермінації, 222
- детермінації множинний, 227, 231
- Джині, 151
- кореляції, 222
- кореляції парний, 227, 229
- кореляції частковий, 227, 230
- локалізації, 149
- Лоренца, 149
- множинної (сукупної) кореляції, 227, 230
- осциляції, 130
- подібності, 154
- регресії, 221
- сезонності, 207, 209
- Херфіндаля, 152
- диференціації, 148
- фондовий, 148
Контроль зібраного матеріалу, 14
- арифметичний, 14
- логічний, 14

Концентрація розподілів, 149
Кореляційно-регресійний
аналіз, 219
Кореляція, 219
- множинна, 219, 226
- парна (проста), 219, 221
Крива концентрації Лоренца, 151
Критерій t- Ст'юдента, 176, 206,
222, 227, 258
Кумулята, 124

М

Макет таблиці, 23, 24
Мала вибірка, 175
Масштабні орієнтири
статистичного графіка, 33
Медіана, 106, 141, 143
Мода, 106, 141, 142

О

Об'єкт дослідження, 7
Обстеження, 10
Обсяг
- вибірки, 172
- сукупності, 7
- явища, 7
Огіва, 124
Одиниці виміру абсолютних
показників, 94
- вартісні, 94
- натуральні, 94
- умовно-натуральні, 94
Одиниці виміру відносних
показників, 95
- відсоток (процент), 95
- іменовані величини, 95
- коефіцієнт, 95
- проценти, 95
- проміле, 95
- просантиміле, 95
Одиниця
- спостереження, 7

- сукупності, 7
Одноразові обліки, 10
Ознака, 8
- групувальна, 64
- кількісна, 8
- результативна, 65
- факторна, 65
- якісна (атрибутивна), 8
Опитування респондентів, 11
- інтерв'ю, 11
- самореєстрація, 11
- анкетне, 11
- експедиційне, 11
- кореспондентське, 11
Організаційні форми
статистичного спостереження, 9

П

Перепис, 10
Підмет таблиці, 23
Подібність розподілів, 154
Показники динаміки
- базисні, 190
- ланцюгові, 190
Поле графіка, 33
Полігон, 121
Помилка вибірки
- гранична, 168
- середня (стандартна), 166, 167
Помилки статистичного
спостереження, 12
- реєстрації, 12
- реєстрації випадкові, 12
- реєстрації систематичні, 13
- реєстрації систематичні
навмисні, 13
- реєстрації систематичні
ненавмисні, 13
- репрезентативності, 13, 166
- репрезентативності випадкові, 13
- репрезентативності
систематичні, 13

Порівнювальний показник, 94
Присудок таблиці, 23
Прогнозування, 204
Просторові орієнтири
статистичного графіка, 33

Р

Реєстр, 10
Рівень ряду динаміки, 187
- базисний, 190
- поточний, 190
Розмах варіації, 128
Ряд динаміки, 187
- інтервальний (періодичний), 187
- моментний, 187
- неповний, 187
- повний, 187
Ряд розподілу, 66
- атрибутивний, 66, 71
- варіаційний, 66, 68
- варіаційний дискретний
(перервний), 66
- варіаційний інтервальні
(неперервні), 66
- комбінаційний, 66, 72

С

Сезонна хвиля, 208, 211
Середнє квадратичне (стандартне)
відхилення, 129
Середнє лінійне відхилення, 128
Середня величина, 100
- арифметична зважена, 101, 189
- арифметична проста, 101, 188
- гармонійна зважена, 103
- гармонійна проста, 103
- геометрична зважена, 104
- геометрична проста, 104
- зважена, 101
- квадратична зважена, 104
- квадратична проста, 104
- кубічна зважена, 105

- кубічна проста, 105
- проста, 101
- хронологічна, 105, 188
Спеціально організоване
статистичне спостереження, 9
Статистика, 7
Статистична карта, 52
- картограма, 52
- картограма ізолінійна, 52
- картограма точкова, 52
- картограма фоновна, 52
- картодіаграма, 52
Статистична сукупність, 7
Статистичне групування, 64
- аналітичне, 65
- аналітичне результативне, 65
- аналітичне факторне, 65, 77
- комбінаційне, 65
- просте, 65
- структурне, 65, 74
- типологічне, 65
Статистичне зведення, 64
- децентралізоване, 64
- машинне (комп'ютерне), 64
- просте, 64
- ручне, 64
- централізоване, 64
Статистичне спостереження, 8
- вибіркове, 11, 162
- монографічне, 11
- несущільне, 11
- одноразове, 10
- основного масиву, 11
- періодичне, 10
- поточне, 10
- суцільне, 11
Статистичні графіки, 32
- зображувальні (фігурні), 34
- лінійні, 34
- об'ємні, 34
- плоскі, 34
- точкові, 34

Статистичні показники, 92

- інтервальні, 93

- моментні, 93

- первинні, 92

- похідні, 92

Статистична таблиця, 23

- групова, 28

- з простою розробкою
присудка, 30

- зі складною розробкою
присудка, 31

- комбінаційна, 29

- проста, 23

- проста монографічна, 24

- проста перелікова, 24

- проста переліково-
хронологічна, 28

- хронологічні, 25

Стерджеса

- номограма, 67

- формула, 67

Сукупність

- вибіркова (виборка), 163

- генеральна, 163

Т

Темп зростання, 191

- середній, 192

Темп приросту, 191

- середній, 192

Тенденція (тренд), 199

У

Укрупнення інтервалів, 199

Ч

Час спостереження, 12

- критичний момент, 12

- об'єктивний, 12

- суб'єктивний, 12

Частка, 66

- кумулятивна (накопичена), 119

Частота, 66, 119

- кумулятивна (накопичена), 119

Щ

Щільність розподілу, 119

- абсолютна, 119

- відносна, 119

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Бек В.Л. Теорія статистики: Навч. посібник/ В.Л. Бек. – К.: Центр учбової літератури, 2003. – 482 с.
2. Бугуцький О.А. Сільськогосподарська статистика з основами економічної статистики / О.А. Бугуцький, А.Т. Опря, М.В. Степаненко та ін. – К.: Вища школа, 1984. – 240 с.
3. Горькавий В.К. Статистика: Навчальний посібник/ В.К. Горькавий. - К.: Вища школа, 2000 - 415с.
4. Доценко О.С. Практикум з загальної теорії статистики: Навч. посібник/ О.С.Доценко. – Севастополь: Вид-во СевНТУ, 2010. – 204 с.
5. Елисеєва И.И., Юзбашев М.М. Общая теория статистики: Учебник/ Под ред. чл.-корр. РАН И.И. Елисеевой. - 4-е изд., перераб. и доп., — М.: Финансы и статистика, 2001. - 480 с.
6. Ефимова М. Р., Ганченко О.И., Петрова Е.В. Практикум по общей теории статистики: Учеб. пособие. - 2-е изд., перераб. и доп. - М.: Финансы и статистика, 2005. - 336 с.
7. Єріна А.М., Пальян З.О. Теорія статистики: Практикум/ А.М. Єріна, З.О. Пальян – К.: Товариство „Знання”, КОО, 1997. – 325 с.
8. Иода Е.В., Герасимов Б.И. Статистика: Учеб. пособие / Под общей ред. Е.В. Иода. - Тамбов: Изд-во Тамб. гос. техн. ун-та, 2004.- 104 с.
9. Ковалевський Г.В. Практикум та тренінг зі статистики: Навч. посібник/ Г.В. Ковалевський, Т.М. Колесник, Г.Б. Тихонова; за ред. Г.В. Ковалевського; Харк. нац. акад. міськ. госп-ва. – Х.: ХНАМГ, 2012. – 156 с.
10. Кулинич О.І. Теорія статистики: Підручник/ О.І. Кулініч, Р.О. Кулинич – 3-тє вид., переробл. і доповн. – К.: Знання, 2006. - 294 с.
11. Моторин Р.М. Статистика для економістів: навч. посібник/ Р.М. Моторин, Е.В. Чеботовський. - К.: Знання, 2011. - 429 с.
12. Общая теория статистики/ Под редакцией А.Я. Боярского, Г.Л. Громыко – М.: Изд-во Моск. Ун-та, 1985. – 376 с.
13. Опря А.Т. Статистика. (модульний варіант з програмованою формою контролю знань). Навч. посібник/ А.Т. Опря. – К.: Центр учбової літератури, 2012. – 448 с.
14. Практикум по теории статистики: Учеб. пособие/ Р.А. Шмойлова, В.Г. Минашкин, Н.А. Садовникова; Под ред. Р.А. Шмойловой. - 2-е изд., перераб. и доп. - М.: Финансы и статистика, 2006.- 416 с.
15. Рудакова Р.П., Букин Л.Л., Гаврилов В.В. Практикум по статистике/ Р.П. Рудакова, Л.Л. Букин, В.В. Гаврилов. – Спб.: Питер, 2007. – 288 с.
16. Статистика: Навчально-методичний посібник для самостійного вивчення дисципліни /Кушнір Н. Б., Кузнецова Т. В., Красовська Ю. В. та інші/— К.:Центр учбової літератури, 2009. — 208 с.

17. Статистика: Підруч. для студ. вищ. навч. закл./ В.Б. Захожай, І.І. Попов. – К.: МАУП, 2006. – 536 с.
18. Статистика: Підручник/ С.С. Герасименко, А.В. Головач, А.М.Еріна та ін.. За наук. ред. С.С. Герасименка. - 2-ге вид., перероб. і доп. - К.: КНЕУ, 2000.- 467 с.
19. Статистика: теоретичні засади та прикладні аспекти. Навч. посібник/ Р.В. Фещур, А.Ф. Барвінський, В.П. Кічор та ін. За наук. ред. Р.В. Фещура. – Львів: «Інтелект-Захід», 2003. – 576 с.
20. Теория статистики: Учебник/ Под ред. проф. Г.Л. Громько. - М.: ИНФРА-М, 2002. – 414 с.
21. Теория статистики: Учебник/ Под ред. Р.А. Шмойловой. - 4-е изд. - М.: Финансы и статистика, 2004. – 656 с.
22. Теорія статистики: Навч. посібник/ Вашків П.Г., Пастер П.І., Сторожук В.П., Ткач Є.І. – К.: Либідь, 2001. – 320 с.
23. Тринько Р.І. Основи теоретичної і прикладної статистики: Навч. посіб. / Р.І. Тринько, М.Є. Стадник. — К. : Знання, 2011. — 397 с.
24. Кустовська О.В, Матійчук Л.П., Солтис В.В., Ціщик Р.В., Чорний В.С. Практикум із дисципліни «Статистика» з використанням Excel/ О.В. Кустовська, Л.П. Матійчук, В.В. Солтис, Р.В. Ціщик, В.С. Чорний. – Тернопіль: ТНЕУ, 2009. – 216 с.
25. Чекотовський Е.В. Статистика сільського господарства: Підручник/ Е.В. Чекотовський. – К.: КНЕУ, 2008. - 504 с.
26. Чернова Т.В. Экономическая статистика: Учеб. пособие/ Т.В. Чернова. - Таганрог: Изд-во ТРТУ, 1999. - 140 с.
27. Мармоза А.Т. Практикум із статистики / А.Т. Мармоза. – К.: Кондор, 2005. – 512 с.
28. Ткач. Є.І. Загальна теорія статистики: Підручник [для студ. вищ. навч. закл.]/ Є.І.Ткач, В.П.Сторожук. – [3-тє вид.]. – К.: Центр учбової літератури, 2009. – 442 с.
29. Keller G. Statistics for Managment and Economics. – Seventh edition. – Thomson South-Western, 2005. – 848 p.
30. Anderson D.R., Sweeney D.J., Williams T.A., Freeman J., Shoemith E. Statistics for Business and Economics. – Thomson Learning, 2007. – 904 p.
31. Офіційний сайт Державної служби статистики України [Електронний ресурс]. Режим доступу: <http://www.ukrstat.gov.ua>
32. Офіційний сайт ДУ «Науково-методичний центр інформаційно-аналітичного забезпечення діяльності ВНЗ «Агроосвіта» [Електронний ресурс]. Режим доступу: <http://www.smcae.kiev.ua>
33. CIA - The World Factbook. [Електронний ресурс]. Режим доступу: <https://www.cia.gov/library/publications/the-world-factbook>

Навчальне видання

Педченко Ганна Павлівна

СТАТИСТИКА

Навчальний посібник

Видавець та виготовлювач ТОВ «Колор Принт»
72312, Запорізька обл., м. Мелітополь, вул. Університетська, 44/7.
Тел. (0619) 46-50-20

Свідоцтво Державного комітету телебачення та радіомовлення України
про внесення суб'єкта видавничої справи до державного реєстру видавців
та виготівників видавничої продукції
Серія ДК № 3782 от 12.05.2010р.

Підписано до друку 05.09.2018р.
Папір офсетний. Формат 60x90/16. Гарнітура Arial.
Друк цифровий. Умов.-друк. аркушів 9,38.
Тираж 300 прим. Замовлення № 45.
ТОВ «Колор Принт»
72312, Запорізька обл., м. Мелітополь, вул. Університетська, 44/7.
Тел. (0619) 46-50-20