

Міністерство освіти і науки України
Івано-Франківський
університет права ім. Короля Данила Галицького

М.В. Щурик
А.В. Ключенко

Статистика

**Для студентів економічних спеціальностей
вузів усіх рівнів акредитації**

Івано-Франківськ - 2016

УДК 31 (076)
ББК 60.6я73
Щ 98

Щурик М.В., Ключенко А.В.

Щ98 Статистика: Навч. посіб. для студ. вищ. навч. закл. усіх рівнів акредит. – 3-тє вид., оновлене і доповнене/ М.В. Щурик, А.В. Ключенко. – Івано-Франківськ: НАІР, 2016. – 274 с.

Навчальний посібник підготовлено відповідно до вимог освітньо-кваліфікаційної програми дисципліни "Статистика", яку включено до навчального плану підготовки бакалаврів з економіки і підприємництва. У ньому висвітлені теоретичні, методологічні і методичні засади теорії статистики, які використовуються в процесі дослідження явищ суспільного життя з урахуванням вимог міжнародних стандартів статистики та обліку, задекларованого вступу України до ЄС.

Рецензенти:

Я.С. Витвицький – доктор економічних наук, професор
І.А. Іртищева – доктор економічних наук, професор
Г.В. Черевко – доктор економічних наук, професор

Схвалено Вченою радою Івано-Франківського університету права ім. Короля Данила Галицького(протокол №7 від 25.04.2016р.)

ISBN 966-7740-70-6

© М.В.Щурик, Ключенко А.В., 2016

Зміст

| | |
|---|----|
| Передмова | 5 |
| Розділ 1. Методологічні засади Державної статистики України .. | 8 |
| 1.1. Історія розвитку статистики | 8 |
| 1.2. Предмет статистичної науки | 11 |
| 1.3. Категорійний і понятійний апарат статистики | 16 |
| 1.4. Методи статистичних досліджень | 28 |
| 1.5. Організація та основні функції органів Державної служби статистики України..... | 20 |
| 1.6. Тестові завдання..... | |
| 1.7. Питання для самоконтролю..... | |
| Розділ 2. Статистичне спостереження | 27 |
| 2.1. Організаційні форми статистичного спостереження | 27 |
| 2.2. План статистичного спостереження | 32 |
| 2.3. Види та способи статистичного спостереження | 34 |
| 2.4. Помилки статистичного спостереження | 36 |
| 2.5. Тестові завдання..... | |
| 2.6. Питання для самоконтролю..... | |
| Розділ 3. Зведення і групування статистичних даних | 40 |
| 3.1. Статистичне зведення | 40 |
| 3.2. Статистичні групування | 42 |
| 3.3. Методологія статистичного групування | 48 |
| 3.4. Тестові завдання..... | |
| 3.5. Питання для самоконтролю..... | |
| Розділ 4. Узагальнюючі статистичні показники | 70 |
| 4.1. Загальна характеристика статистичних показників | 70 |
| 4.2. Система статистичних показників | 72 |
| 4.3. Абсолютні статистичні показники | 73 |
| 4.4. Відносні показники | 74 |
| 4.5. Загальна характеристика середніх величин | 80 |
| 4.6. Узагальнюючі степеневі середні показники | 82 |
| 4.7. Тестові завдання..... | |
| 4.8. Питання для самоконтролю..... | |

| | |
|--|------------|
| Розділ 5. Аналіз рядів розподілу | 113 |
| 5.1. Найважливіші показники рядів розподілу | 113 |
| 5.2. Показники варіації | 120 |
| 5.3. Тестові завдання..... | |
| 5.4. Питання для самоконтролю..... | |
| | |
| Розділ 6. Аналіз концентрації диференціації та подібності розподілів..... | 133 |
| 6.1. Характеристика форм розподілу | 133 |
| 6.2. Криві розподілу | 140 |
| 6.3. Дисперсійний аналіз | 143 |
| 6.4. Тестові завдання..... | |
| 6.5. Питання для самоконтролю..... | |
| | |
| Розділ 7. Статистичні методи вимірювання взаємозв'язків ... | 152 |
| 7.1. Поняття та види взаємозв'язків | 152 |
| 7.2. Парна кореляція | 154 |
| 7.3. Непараметричні методи оцінювання взаємозв'язку між ознаками..... | 163 |
| 7.4. Поняття множинної кореляції | 171 |
| 7.5. Тестові завдання..... | |
| 7.6. Питання для самоконтролю..... | |
| | |
| Розділ 8. Аналіз інтенсивності динаміки | 180 |
| 8.1. Загальна характеристика рядів динаміки | 180 |
| 8.2. Найважливіші характеристики динамічних рядів | 186 |
| 8.3. Тестові завдання..... | |
| 8.4. Питання для самоконтролю..... | |
| | |
| Розділ 9. Аналіз тенденцій розвитку та коливань | 197 |
| 9.1. Дослідження тенденцій динамічних рядів | 197 |
| 9.2. Аналітичне вирівнювання | 200 |
| 9.3. Статистичне вивчення стійкості динамічних рядів та сезонності | 203 |
| 9.4. Тестові завдання..... | |
| 9.5. Питання для самоконтролю..... | |
| | |
| Розділ 10. Індексний метод | 214 |

| | |
|--|------------|
| 10.1. Суть, значення та види індексів | 214 |
| 10.2. Агрегатна форма індексів | 217 |
| 10.3. Середньозважені індекси. Факторний аналіз | 221 |
| 10.4. Індеси змінного, постійного складу та структурних зрушень | 226 |
| 10.5. Тестові завдання..... | |
| 10.6. Питання для самоконтролю..... | |
| Розділ 11. Вибірковий метод | 236 |
| 11.1. Суть і значення вибіркового спостереження | 236 |
| 11.2. Похибки вибіркового спостереження. Способи відбору | 237 |
| 11.3. Основні характеристики вибіркового спостереження | 242 |
| 11.4. Поширення результатів вибірки на генеральну сукупність .. | 249 |
| 11.5. Тестові завдання..... | |
| 11.6. Питання для самоконтролю..... | |
| Розділ 12. Подання статистичних даних: таблиці, графіки, карти..... | 258 |
| 12.1. Статистичні таблиці | 258 |
| 12.2. Статистичні графіки | 259 |
| 12.3. Тестові завдання..... | |
| 12.4. Питання для самоконтролю..... | |
| Додатки..... | 270 |
| Список використаних джерел | 272 |

ПЕРЕДМОВА

Нинішній розвиток економічних, соціальних, екологічних, національних та інших важливих напрямків відтворювального процесу значно відрізняється від того, що був сформований командно-адміністративною системою, а також на початковому етапі проголошення незалежності нашої держави.

Динамічність, яка властива сучасному суспільству потребує наукової її оцінки, що досягається завдяки застосуванню економічного інструментарію адекватного вимогам та особливостям формування моделі ринкової економіки. Суть нинішнього етапу соціально-економічного відтворювального процесу полягає в тому, що він значно відрізняється від соціалістичного і від капіталістичного, оскільки поєднує елементи періоду командної та ринкової економіки. Сформовані економічні відносини вже не є соціалістичними, але й не капіталістичними, що потребує врахування їх в процесі дослідження економічних, екологічних, соціальних, національних та інших найважливіших складових процесу відтворення. Складність дослідження полягає також у тому, що на відміну від стабільно сформованих економічних відносин, властивих соціалістичному та капіталістичному суспільству, нинішній етап потребує не лише констатації й розрахунку критеріїв та показників явищ і процесів, що мають місце в процесі суспільного відтворення, а й пошук основних напрямків, шляхів його поліпшення. Інакше кажучи, статистичне дослідження динамічних процесів, властивих нині Україні потребує вироблення конкретних пропозицій спрямованих на розв'язання існуючих проблем.

Мета пропонованого навчального посібника полягає в наданні допомоги студентам у процесі освоєння основних теоретичних, методологічних та прикладних засад застосування статистичного аналізу та оцінки найважливіших трансформаційних процесів, які мають місце в суспільстві, економіці, природі тощо.

Відповідно до зазначеної мети побудовано зміст і структуру тем, за допомогою яких студенти одержать можливість оволодіти прийомами та методами статистичного аналізу, статистичним інструментарієм дослідження явищ і процесів суспільного життя, а також виявити основні закономірності розвитку суспільства та складових, що його опосередковують.

Навчальне видання «Статистика» побудоване відповідно до ОПП, затвердженого Наказом Міністра освіти та науки України від 2 лютого 2010р. і призначене для студентів економічних спеціальностей вузів України всіх рівнів акредитації, крім випускників за фаховим спрямуванням «Статистика».

Структура навчального видання включає дві частини, кожна з яких органічно доповнює одна одну. Перша частина «Загальна теорія статистики» включає перелік тем, за допомогою яких забезпечується можливість оволодіння традиційними статистичними прийомами та методами аналізу, зокрема: «Статистичне спостереження», «Зведення і групування статистичних даних», «Середні величини», «Ряди динаміки», «Індекси» тощо. Крім цього значне місце в структурі видання займають теми з математичної статистики та теорії ймовірностей: «Статистичні методи вимірювання взаємозв'язку», «Статистичні методи вимірювання варіації та форм розподілу», «Вибіркове спостереження» та ін.

З метою закріплення та контролю знань у кінці кожної теми приведені тестові завдання, типові задачі, а також питання для здійснення самоконтролю знань студентом.

Тематика навчального видання побудована з урахуванням змін, які властиві ринковій моделі господарювання в Україні, вимогам країн ЄС. Разом з тим, автори усвідомлюють, що явища та процеси, що часто виникають можуть бути надзвичайно динамічними та мінливими. Безперечно, що час, затрачений на написання, друкування та публікацію навчального видання не може бути співставленим зі змінами законодавчої та нормативної бази, і тому не всі положення, що викладені у виданні будуть відповідати

динаміці часу. В цілому навчальне видання підготовлено відповідно до нормативів кваліфікаційної характеристики освітньо-кваліфікаційного рівня "бакалавр" та враховує перехід України на Європейські стандарти навчання.

Автори будуть вдячні усім тим хто надішле свої пропозиції та відгуки щодо вдосконалення змісту і структури посібника, що в кінцевому рахунку буде враховано в процесі подальшої роботи над навчальним виданням.

РОЗДІЛ 1. МЕТОДОЛОГІЧНІ ЗАСАДИ ДЕРЖАВНОЇ СТАТИСТИКИ УКРАЇНИ

План (логіка) викладу та засвоєння матеріалу:

- 1.1. Історія розвитку статистики
 - 1.2. Історія розвитку статистики
 - 1.3. Предмет статистичної науки Категорійний і понятійний апарат статистики
 - 1.4. Методи статистичних досліджень
 - 1.5. Організація та основні функції органів Державної служби статистики в Україні
- Тестові завдання
- Питання для самоконтролю

1.1. Історія розвитку статистики

Статистика як наука виникла з потреб суспільного життя. Її розвиток як науки пов'язується із суспільним поділом праці, створенням держави, розвитком приватної форми власності.

Пізнання статистики, як науки потребує ідентифікації найважливіших дефініцій і понять, які використовуються з метою пізнання предмета статистики.

Слово «статистика» латинського походження (від «status» – стан, становище).

В науковій термінології воно вперше було використано німецьким ученим Готфрідом Ахенвалем у середині XVII ст. Його застосовували із метою характеристики політичного стану держави: «stato» – держава; «statista» – статистик-знавець держави; «statistika» – певна сума знань про державу.

Статистика як наука зазнавала в своєму розвитку злетів і падінь. Первинною формою зародження статистики вважають проведення господарського обліку, що виник у глибоку давнину. Прикладом тут може слугувати перепис населення в Китаї, що проводився більш як за 4 тис. років до нашої ери. Періодичні війни,

збір податків, управління державою, торгівля зумовлювали потребу володіти інформацією, котра стосувалася чисельності населення, розміру земель, лісів, поголів'я худоби тощо. Роботи зі збору такої інформації в давнину і середньовіччя проводились примітивно, від випадку до випадку, тобто в міру потреби.

Точнішою стає статистика з початком створення держави і особливо приватної форми власності на землю, ліс, водні ресурси, корисні копалини, фабрики і заводи тощо.

Зародження статистики значна частина вчених пов'язує з країнами Стародавнього Сходу. Саме тут, як засвідчують ретроспективні дослідження вперше склалися розвинені системи адміністративного та державного обліку.

Втім, перші статистичні дослідження зустрічаються значно раніше, про що засвідчує інформація про них у числених священних книгах. У Четвертій книзі Мойсея «Числа» в Святому Письмі записано: «І сказав Господь Мойсеєві у пустелі Синайській, у скинії зібрання, у перший (день) другого місяця, у другий рік після виходу їх із землі Єгипетської, говорячи: підрахуйте всю громаду синів Ізраїлевих за родами їх, за сімействами їх, за числом імен, усіх чоловічої статі поголовно: від двадцяти років і вище, усіх придатних для війни у Ізраїля за ополченнями їх порахуйте...»¹ Тут же міститься інформація щодо підрахунку осіб чоловічої статі від 20 років і вище, усіх придатних до війни по кожному коліні, племені, сімейству.

З середньовіччя до наших днів збереглося унікальне видання «Книга страшного суду», в якому достатньо повно відображено дані перепису населення і майна в Англії.

Широкий розвиток одержала статистика як наука в таких країнах як Китай, Єгипет, Вавилон, Рим, Стародавня Греція. Разом з тим, вона не була позбавлена недоліків, оскільки у той час статистичні роботи носили описовий характер, тому точність і

¹ Біблія. Книга священного писання Старого та Нового Завіту. Видання Київської Патріархії Української Православної церкви Київського Патріархату. – с. 136

достовірність їх була невисокою. Відомо, що перший статистичний орган був утворений у Стародавньому Римі.

Формування і становлення статистики як науки найбільшого поширення набуло в середині XVII століття. В цей час створюються два статистичних напрями (школи) досліджень – описова німецька та англійська школа політичних арифметиків. І перша, і друга внесли значний вклад у розвиток статистики.

Засновниками німецької описової школи вважають Г. Конрінга (1606-1681), Г. Ахенваля (1719-1772), А. Брюшинга (1724-1793) та інших. Основні завдання статистики вони вбачили у систематизованому описуванні фактів, які засвідчували могутність та міцність держави, її процвітання. Дослідження представниками даної школи організовувалися в абсолютній більшості за допомогою описової, словесної форми, без цифр і в «статистиці», тобто без достатнього урахування історичних особливостей розвитку держави за той чи інший період, а лише на момент спостереження. Відомо, що описова школа розвитку статистики у XVII ст. одержала найбільшого поширення в університетах міст Галле, Гельмштадта, Гетінгенда. Зокрема, Г. Конрінг робив виклади у німецьких університетах, читаючи курс "Universitatstätistik" - «Університетська статистика». З розвитком статистики як науки даний напрям з часом втратив своє значення й актуальність.

Натомість засновники школи політичних арифметиків при статистичному дослідженні суспільних явищ і процесів віддавали перевагу числовим показникам, цифрам, графікам, таблицям. Найяскравішими представниками зазначеної школи були англійські вчені Уільям Петті (1623-1687рр.) та Джон Граунт (1620-1674рр.). Основне призначення статистики ними вбачалося у вивченні суспільних явищ, а не лише у возвеличенні держави використовуючи при цьому найновіші надбання інших точних наук. Насамперед це стосувалося використання вимог закону великих чисел, згідно з яким статистичні закономірності можуть проявлятися лише при достатньо великому числі одиниць статистичної сукупності. Таким чином, ще задовго до

опублікування математичної теореми Я. Бернуллі про закон великих чисел, статистика значно раніше використовувала основні положення цього закону в процесі статистичного вивчення суспільних явищ.

Школа політичних арифметиків на той час поділялась на два напрямки: демографічний (Д. Граунт і Е. Гален) і статистико-економічний (В. Петті). Спільним об'єктом дослідження для державознавчої школи і школи політичних арифметиків було суспільство. В той же час методи дослідження були різні – описовий і вимірювальний.

Домінуючим напрямком розвитку статистики в XVIII – XX ст. став напрям започаткований школою політичних арифметиків, який поступово витіснив описовий, державотворчий. Статистика все більше використовує в процесі дослідження свого предмета кількісні показники, що зумовлює значне зростання й розширення методів та прийомів статистичного дослідження.

Основою, базою становлення та розвитку статистики є теорія статистики. Відомо, що її фундатором був бельгійський вчений А. Кетле, якого вважають основоположником вчення про середні величини. Предметом статистики, на його переконання слугує суспільство, людина, а методологічною основою – закон великих чисел, що потребує застосування середніх величин як узагальнюючих показників сукупності.

Серед вітчизняних учених значний вклад у розвиток статистики, внесли М. В. Птуха (1894-1961), І. Я. Франко (1856-1916), Ю. А. Корчак-Чепурківський (1896-1967), К. Г. Воблий (1876-1947) та інші.

1.2. Предмет статистичної науки

Пізнання змістовного навантаження тієї чи іншої науки, включаючи статистику забезпечується, передусім, через з'ясування предмета дослідження, а також методів, прийомів і показників, які використовуються при цьому. Слід зазначити, що змістовне

навантаження «предмета» досить складне, дискусійне і пройшло в своєму розвитку тривалий час. Навіть нині продовжуються точитися дискусії серед науковців щодо визначення предмета статистики. Складність полягає, перш за все у тому, що статистика часто досліджує кількісні явища та процеси суспільного життя, що є предметом дослідження інших наук.

Така посилена увага до визначення предмета статистики пов'язана, насамперед із важливістю науки як загальнонаукового та світоглядного характеру.

При тлумаченні поняття предмета статистики слід усвідомити, що статистика є суспільною наукою, оскільки досліджує соціальні явища й процеси. Принагідно зазначимо, що суспільство є об'єктом дослідження не лише статистики, а й багатьох інших суспільних наук. Разом з тим, статистика на відміну від них має свої особливості котрі дають можливості виокремити її предмет.

Предметом статистики є кількісна сторона масових суспільних явищ і процесів у нерозривному зв'язку з їх якісним змістом, а також кількісне визначення закономірностей суспільного розвитку в конкретних умовах місця та часу.

Статистика є самостійною суспільною наукою. Таке тлумачення предмета статистики дозволяє стверджувати, що статистичні управлінські органи завжди виконують волю та розв'язують завдання тих державних управлінських структур, котрі причетні до її формування та потребують адекватної статистичної інформації.

Приведене визначення предмета статистики спрямоване, передусім на одержання кількісної характеристики досліджуваного суспільного процесу й підрахунок показників за допомогою яких здійснюється оцінювання досліджуваного явища або процесу.

Кількісна сторона розвитку суспільних явищ вимірюється за допомогою показників їх обсягу та розміру. Наприклад, у першому кварталі поточного року фірма по виготовленню мобільних телефонів експортувала за межі країни 10000 мобільних телефонів загальною вартістю 500000 ум. одиниць (собівартість одного

телефону – 400 ум. од.). Звісно, власників заводу цікавить не лише обсяг реалізації (кількість реалізованих телефонів), а й економічна ефективність (продуктивність) результатів. Йдеться про статистичне оцінювання кількісної сторони цього процесу, що забезпечується розрахунком співвідношення розмірів. У нашому прикладі слід порівняти ціну реалізації (500 ум.од.) та собівартість (400 ум. од.). Ціна реалізації перевищує собівартість в 1.25 рази, що засвідчує про успішну діяльність фірми в першому кварталі поточного року.

Досліджуючи предмет статистики, можна виокремити такі чотири основні принципові положення:

– по-перше, статистика досліджує суспільні процеси. Явища, що знаходяться поза сферою суспільної діяльності людства, не підлягають статистичному обліку й вивченню;

– по-друге, статистика вивчає кількісну сторону масових суспільних явищ, а не поодинокі явища;

– по-третє, явища і процеси вивчаються в конкретних умовах місця і часу;

– по-четверте, статистика є самостійною суспільною наукою.

Принагідно зазначимо, що явища й процеси суспільного життя слугують також предметом дослідження багатьох інших суспільних наук – економіки, політичної економії, історії, соціології, політології тощо. Однак, статистика вивчає лише ту сторону суспільних явищ, які властиві лише цій дисципліні. Це має важливе принципове значення при визначенні об'єкта і предмета науки. Об'єкт у декількох науках може бути один, а предмети дослідження різні.

Статистика, як зазначається вище вивчає кількісну сторону суспільних явищ. Читач може заперечити. Кількісна сторона є також предметом дослідження інших наук, зокрема бухгалтерського обліку. Втім, у приведеному визначенні наголошується, що предметом вивчення статистики є кількісна сторона масових суспільних явищ. Натомість предметом дослідження бухгалтерського обліку слугують господарські засоби, їх стан, рух у процесі господарської діяльності, розподіл і обіг, а також джерела

утворення засобів. У більшості випадків, предметом дисципліни бухгалтерського обліку є поодинокі явища і процеси, за допомогою яких оцінюється діяльність суб'єктів підприємницької діяльності. Принагідно зазначимо, що інші суспільні науки часто використовують результати предмета дослідження статистики .

Кількісна сторона явищ вивчається також математичними дисциплінами, але предмет і метод математики значно відрізняється від статистики. Візьмемо простий приклад і проаналізуємо динаміку виконання договорів постачання продукції двома фірмами за звітний період. На фірмі № 1 у поточному році відсоток виконання договорів постачання зріс з 80 до 95 %. На фірмі № 2 він не змінився і склав 100 %. З позиції математики виконання договорів постачання на фірмі № 1 зросло на 15 %, на фірмі № 2 змін не відбувалося. Для статистики важливими при цьому є не тільки кількісне співвідношення виконання договорів постачання, а й їх якісна характеристика, котра могла вплинути на результати діяльності фірм. Перша фірма залишилась у групі підприємств, які не виконали договорів постачання у поточному році з усіма впливаючими звідси економічними, соціальними та іншими складовими діяльності. Друга, продовжувала працювати стабільно. Очевидно, споживачами будуть враховані результати діяльності обидвох фірм при укладанні нових договорів постачання. Таким чином, поряд із установленням відсотка виконання договорів постачання до уваги слід привернути якісну сторону, котра засвідчує про стабільну діяльність фірми №1 та невиконання договорів постачання фірмою №2.

Статистика як суспільна наука досліджує явища й процеси суспільного життя в динаміці, взаємозв'язку, взаємозалежності, територіальному співвідношенні тощо. Процеси суспільного життя розглядаються статистичною наукою в конкретних умовах місця та часу.

Нині статистика, досліджуючи динаміку суспільних явищ, характеризує їх у різноманітних зв'язках з екологією, навколишнім середовищем, з процесами глобалізації, асоційованого членства

України в ЄС тощо. Таке вивчення проводиться крізь призму взаємозв'язку, взаємозалежності, в постійній зміні в часі та просторі, що пояснюється зростаючою складністю дослідження предмета статистики.

Динамічність та особливість предмета дослідження статистики в умовах інтегрування нашої країни до вимог і стандартів багатьох європейських держав, запровадження якісно нових фінансово-кредитних відносин та посилення їх значущості в суспільному відтворенні зумовлює необхідність поділу статистичної науки на окремі самостійні частини:

1. Загальна теорія статистики - дає тлумачення категорій статистичної науки, правил і методів, які є спільними і можуть бути використані для дослідження кількісної та якісної сторін масових суспільних явищ.
2. Соціально-економічна (макроекономічна) статистика - вивчає явища й процеси в сфері економіки, розробляє систему економічних показників та методів вивчення економіки на мікро-, мезо- та макрорівнях.
3. Національне рахівництво – метод, який запроваджений у міжнародній практиці з метою забезпечення комплексної характеристики виробництва, освіти, розподілу та використання доходів, накопичення національного багатства тощо.
4. Статистика сфер діяльності (промисловість, будівництво, сільське господарство, торгівлі, банківська статистика тощо). Ними розробляється предмет дослідження конкретної сфери з використанням властивих їм методів та прийомів дослідження. В загальному переліку сфер діяльності окремо виділяють фінансову статистику, за допомогою якої забезпечується статистичний аналіз та оцінювання рівня функціонування фінансово-кредитних установ та нефінансових інституцій.

Статистика як суспільна наука в своєму розвитку ґрунтується на теоретичних, методологічних та прикладних дослідженнях інших

дисциплін. При цьому обов'язково забезпечуються умови зворотного зв'язку, котрі полягають в тому, що пізнання предмета іншими науками забезпечується завдяки використанню статистичної інформації.

Йдеться про використання статистикою в процесі статистичних досліджень надбань інших наук, зокрема політичної економії, соціології, політології, конкретних економічних дисциплін тощо. Базуючись на суті, якісній природі явищ, через узагальнення масових даних статистика вивчає характер і дію законів об'єктивного розвитку суспільства в реальному житті.

Завершуючи характеристику предмета вивчення статистики, зазначимо, що статистичні дослідження охоплюють не тільки явища й процеси, котрі належать безпосередньо до її предмета, але й такі, завдяки яким можна одержати цілісну характеристику дослідженого явища або процесу суспільного життя.

Статистика в процесі дослідження свого предмета вивчає також природні багатства, навколишнє середовище, технологію і технічний процес, незважаючи на те, що вони самі по собі не належать до економічних категорій. Втім, беззаперечним є те, що інноватика генерує ті чинники, котрі безпосередньо впливають на характер кількісних і якісних змін процесу суспільного відтворення в зв'язку з формуванням в нашій країні проєвропейського шляху розвитку.

1.3. Категорійний і понятійний апарат статистики

Пізнання предмета статистики потребує застосування відповідного понятійного апарату, котрий використовується для тлумачення понять, дефініцій тощо. Найбільш широко під час статистичного дослідження використовують такі поняття та категорії: статистична сукупність, одиниця сукупності, варіація, статистична закономірність тощо.

Статистична сукупність – це множина елементів соціально-економічних об'єктів або явищ суспільного життя, що поєднані між

собою єдиною якісною основою існування та розвитку, але мають поряд із цим і відмінні ознаки. Прикладом сукупності, котра має спільні ознаки може бути загальна кількість студентів, сімей, підприємств промисловості тощо. Сукупність буває однорідною й неоднорідною. Однорідність засвідчується поєднанням між собою властивостями декількох ознак, які мають єдину спільну якісну основу. Неоднорідна статистична сукупність характеризується значними відмінностями складу досліджуваних одиниць.

Важливим при вивченні статистичної сукупності є встановлення одиниці сукупності.

Одиниця сукупності – первинна частина (елемент) сукупності, що є носієм характеристик ознаки.

Наступною важливою складовою, за допомогою якої характеризується сукупність є ідентифікація поняття «ознака». **Ознака** слугує якісною характеристикою одиниці сукупності. Кожна одиниця сукупності може мати багато ознак. Ознака, котра приймає в межах сукупності різні значення називається **варіюючою**, а коливання значень – **варіацією**.

Неоднакові властивості одиниць сукупності потребують їх поділу на **кількісні** та **якісні**. Перші, мають кількісний вираз. Наприклад, розмір заробітної плати, стаж роботи, кількість виготовлених деталей і т. д. Вони можуть бути **дискретними** (перервними) і **безперервними**. Ознаки, котрі не мають кількісного тлумачення, називаються якісними, або атрибутивними («атрибут» – невід’ємна властивість предмета). Наприклад, склад населення за статтю, освітнім рівнем та ін. Якщо ознака задається лише двома протилежними за значеннями результатами (так, ні, задоволені, незадоволені) називається альтернативною.

Статистична сукупність не може бути механічним об’єднанням елементів, а є упорядкованою системою, котра має як спільні та відмінні риси, що проявляються в конкретних умовах місця та часу. Сукупність вважається статистичною лише за умови коли її властиві такі риси: об’єктивність існування; якісна однорідність явищ, що утворюються нею; варіація в часі та просторі тощо.

Статистичне дослідження базується на статистичному спостереженні за якомога великою кількістю явищ і процесів суспільного життя. Результати його проведення дають змогу звільнитися від впливу випадковостей і виокремити найбільш важливі та суттєві риси досліджуваних явищ, встановити істотні взаємозв'язки, а також тенденції розвитку, закономірності трансформаційних процесів у часі та просторі.

Важливою статистичною характеристикою предмета статистичної науки є **статистична закономірність**, за допомогою якої оцінюється повторювальність і порядок у досліджуваних явищах і процесах. Прикладом статистичних закономірностей може слугувати закон руху населення.

Статистичні закономірності завжди є незалежними. В своїй основі вони відображають характер дії об'єктивних законів розвитку явищ і процесів суспільного життя. Вони постійно оцінюються в динаміці, взаємозв'язку і взаємозалежності, структурних зрушеннях. Статистична закономірність втілюється крізь призму закону великих чисел, який проявляється через статистичну сукупність.

1.4. Методи статистичних досліджень

Пізнання свого предмета статистика здійснює за допомогою статистичних методів, способів та прийомів.

Саме поняття «метод» означає спосіб (прийом), який використовується для вивчення конкретною наукою. Статистика для пізнання предмета використовує комплекс спеціальних, властивих лише їй методів і способів дослідження.

Методологічною основою статистики як суспільної науки, є теорія пізнання в основі якої покладено загальнонаукові і загальнофілософські принципи, основа суть яких полягає у наступному:

1. Всі явища і процеси обов'язково досліджуються у взаємній залежності і взаємозумовленості.

2. Статистичні ознаки оцінюються крізь призму динаміки, оновлення та розвитку.
3. Кількісні зміни зумовлюють якісні.
4. Статистична методологія пізнає та вивчає явища і процеси використовуючи принципи філософських законів єдності і боротьби протилежностей, заперечення, суперечностей між новим і старим тощо.

Методи статистики ґрунтуються на поєднанні аналізу і синтезу, індукції та дедукції.

Виходячи з методологічних основ дослідження, статистика розробила свої, властиві лише її предмету методи дослідження.

Взагалі, будь-яке пізнання предмета статистикою здійснюється у такій послідовності: статистичне спостереження, зведення й групування статистичних даних і розрахунок узагальнюючих зведених показників. На кожному конкретному етапі використовують насамперед ті методи та прийоми, котрі забезпечують повну, глибоку і всебічну характеристику явищ і процесів, що підлягають дослідженню.

Статистичне спостереження – перша стадія статистичного дослідження. При його проведенні статистикою обліковується кожна одиниця сукупності та вивчаються індивідуальні значення властивих їй ознак.

У процесі проведення статистичного спостереження використовують результати масового дослідження одиниць досліджуваної ознаки. Такі засади пояснюються тим, що результативність, об'єктивність має безперечну доказову силу лише тоді, коли вона спирається не на окремі факти, а на їх масовість. В свою чергу, статистичні закономірності, як правило відображають об'єктивність лише завдяки дослідженню достатньо великої кількості одиниць спостереження. Використання масовості в процесі проведення статистичного спостереження дає можливість абстрагуватись від випадкового і виявити типове, закономірне. Означене засвідчує про потребу застосування закону великих чисел під час проведення першої стадії статистичного дослідження.

Об'єктивність цього закону досягається лише при достатній кількості спостережень, а вплив випадкових величин, які визначають величину ознаки окремих одиниць сукупності, в цілому взаємно погашається. У зведених характеристиках втілюється дія основних причин, тобто проявляється закономірність.

Другою стадією статистичного дослідження є **зведення і групування** матеріалів статистичного спостереження. На цьому етапі одиниці масового спостереження групують за певними ознаками, наприклад, населення можна класифікувати за статтю, віком, зайнятістю по сферах діяльності, джерелах доходів. Впорядковану таким чином статистичну сукупність називають статистичним рядом розподілу, або рядом динаміки. Перший є результатом групування одиниць сукупності в статистиці, другий описує динаміку розвитку досліджуваної ознаки.

Розрахунок і аналіз узагальнюючих зведених показників забезпечується на третій стадії статистичного дослідження. Для його проведення використовують узагальнюючі показники – підсумкових, відносних і середніх величин, статистичних коефіцієнтів. Крім цього в процесі аналізу застосовують прийоми порівняння показників, індексний, балансовий, кореляційний, дисперсійний та інші методи. Для унаочнення й подання результатів аналізу в найбільш компактній та раціональній формах використовують статистичні таблиці та графіки.

1.5. Організація та основні функції органів Державної служби статистики України

Діюча структура організації сучасних статистичних органів в Україні значною мірою успадкована від старої системи з усіма її позитивними і негативними рисами.

В колишньому СРСР статистичні органи були підпорядковані існуючому тоді адміністративно-командному прокомуністичному апарату та їх ідеології. Про це переконливо доводить застосування донині значної кількості показників, які використовуються в процесі

проведення статистичних досліджень. Методи і методологія досліджень залишаються також підпорядкованими єдиній меті – довести велич, могутність, непогрішність і справедливість діючого державного чиновницького апарату на макро-, меза- і макрорівнях.

Слід зазначити, такі засади організації Державної служби статистики не забезпечують порівняльність соціально-економічного розвитку України з країнами ЄС, республіками колишнього Радянського Союзу. Більше того, значна частина з них продовжує носити закритий і недоступний для усього суспільства характер.

Проголошення нашою країною курсу на інтеграцію соціально-економічного розвитку до країн ЄС потребує радикальних змін у організації Державної служби статистики в Україні. Виникла потреба створення таких організаційно-управлінських державних статистичних структур, які б гармоніювали з методологічними і методичними засадами функціонування статистичних органів у країнах Європейського співтовариства, а також Організації Об'єднаних Націй. Передусім це стосується досягнення порівняльності показників соціально-економічного розвитку, завдань і мети статистичних органів, як найбільш об'єктивного інформаційно-комунікаційного ресурсу держави. Змістом діяльності органів Державної статистики в Україні повинно стати збирання, обробка та подання об'єктивної й неупередженої інформації з метою одержання даних щодо соціально-економічного розвитку країни у період інтеграції її до країн Європейського співтовариства, формування й імплементації методичних, методологічних основ на яких організовані статистичні органи а ООН в умовах запровадження ринкових економічних відносин.

Слід констатувати, що сьогодні в Україні, як і в період командно-адміністративної системи функціонує виключно державна служба статистики. Створення приватних статистичних структур чинним вітчизняним законодавством не передбачено. На наше переконання нині на часі постає завдання створення адекватної ринковим відносинам статистичної інституції, котра би забезпечувала не лише збір та обробку статистичної інформації на

старих теоретичних і методологічних засадах, а й розробляла та пропонувала найважливіші напрямки та шляхи розв'язання економічних, екологічних, соціальних, національних та інших важливих завдань, які постають нині в державі в зв'язку з інтеграційними, глобалізаційними процесами.

Загальне керівництво ведення статистики в нашій країні здійснює Державна служба статистики України, діяльність якої підпорядкована Кабінету Міністрів України. Основним документом, що регламентує повноваження та функції органів державної статистики, є Закон України «Про державну статистику» (1992 р.). Статті цього закону регулюють правові відносини в галузі статистики і ведення первинного обліку, визначають організацію та найважливіші завдання державної статистики, порядок подання і використання статистичних даних, а також права, обов'язки та відповідальність органів Державної служби статистики України.

Центральному статистичному органу держави підпорядковуються головні управління державної статистики в областях, а також районні й міські відділи статистики.

Відповідно до Закону України «Про державну статистику» найважливішим завданням державної статистики є збір, обробка та аналіз статистичної інформації, котра стосується явищ і процесів, які мають місце в економічному й соціальному житті держави.

Аналітична статистична інформація подається у встановлені терміни до Верховної Ради України, Адміністрації Президента, Кабінету Міністрів, місцевим органам управління, адекватним міністерствам, агентствам, іншим керівним органам держави.

Державна служба статистики України та її представницькі органи на місцях розробляють інструкції та вказівки з питань організації звітності й обліку, котрі є обов'язковими для виконання усіма міністерствами і відомствами. Відповідно до наказів та інструктивних документів підприємства, організації незалежно від форм власності зобов'язані подавати органам статистики статистичну інформацію у відповідні строки, керуючись при цьому встановленими вимогами.

Обов'язком статистичних органів є перевірка стану, об'єктивності та достовірності звітних даних. Вони мають право відмінити статистичну звітність, яка є надуманою тими чи іншими управлінськими структурами і не затвердженою у встановленому порядку Державною службою статистики України.

Найважливішими функціями статистики та її органів на місцях слугує організація та здійснення статистичного спостереження за ходом проведення соціально-економічних реформ у цілому та в розділі окремих регіонів; забезпечення збирання статистичної інформації; ведення сімейних бюджетів; проведення переписів, одноразових обліків; опитувань; вибіркового демографічного та інших обстежень. Прикладом організації роботи державної статистичної служби України є проведення перепису населення, який здійснювався в ніч з 4 на 5 січня 2002 року. Разом з тим відомо, що періодичність проведення переписів населення встановлена в термін 1 раз на 10 років. В той же час в нашій країні встановлена періодичність порушена.

Функціональні обов'язки статистичних органів є надзвичайно динамічними в умовах побудови незалежної України, вступу нашої держави до ЄС, запроваджені засад сталого розвитку. Зазначені трансформації потребують адекватного статистичного супроводу таких змін і відповідної інформації. Статистичні органи повинні більш конструктивно й доступно організовувати спостереження за процесами формування ринкових засад в сферах діяльності, ринковою інфраструктурою, розвитком нових форм власності, роздержавленням і приватизацією, дієвістю фінансових інституцій тощо.

Предметом аналізу статистичних органів в умовах сьогодення слугує також виконання міждержавних угод та договорів, зміни кількісних та якісних характеристик населення країни, дослідження впливу позитивних і негативних процесів ринку на розв'язання соціальних завдань.

Не менш важливими є функції статистики, пов'язані з дослідженням якості продукції, визначенням її

конкурентоспроможності, показниками праці, впровадження досягнень науково-технічного прогресу в економічну і соціальну сферу, визначення рівня життя та добробуту населення країни.

Наведені нами основні функції статистики та її органів в умовах інтеграції нашої країни до ЄС, посилення глобалізації можна оцінити як перехідні та динамічні. Тому їх постійно потрібно доповнювати з урахуванням взаємозв'язку, змінами, котрі мають місце в економічних, соціальних та інших сферах діяльності як всередині держави, так і поза її межами.

Тестові завдання

1. Слово «статистика» походить від:

- а) латинського;
- б) грецького;
- в) римського;
- г) китайського.

2. В наукову термінологію слово «статистика» було введено вперше:

- а) Г. Конрінгом;
- б) Г. Ахенвалем;
- в) А. Брюшингом;
- г) Д. Граунтом.

3. У XVII столітті було утворено дві статистичні школи:

- а) грецька та китайська;
- б) єгипетська та вавилонська;
- в) німецька та англійська;
- г) російська та американська.

4. Домінуючим напрямком розвитку статистики в XVII – XX столітті вважають напрям започаткований:

- а) школою політичних арифметиків;
- б) описовою школою;
- в) демографічною школою;
- г) вавилонською школою.

5. Предметом статистики є:

- а) матеріали статистичних переписів ;
- б) кількісна сторона масових суспільних явищ у нерозривному зв'язку з їх якісним змістом, а також кількісне вираження закономірностей суспільного розвитку в конкретних умовах місця і часу;
- в) статистична звітність;
- г) розрахунок статистичних показників.

6. Статистична сукупність – це:

- а) множина елементів, які досліджуються статистикою;
- б) явища і процеси суспільного життя;
- в) множина елементів соціально-економічних об'єктів суспільного життя, що поєднані між собою якісною основою існування та розвитку але мають поряд з цим і відмінні ознаки;
- г) кількісна і масова сторона суспільних явищ і процесів.

7. Ознака, що приймає в межах сукупності різні значення, називається:

- а) кількісною;
- б) якісною;
- в) альтернативною;
- г) варіюючою.

8. Ознаки, що не мають безпосередньої кількісної характеристики, називають:

- а) атрибутивними;
- б) альтернативними;
- в) варіативними;
- г) дискретними.

9. Якщо ознака подається лише за двома протилежними за значення-ми результатами, то її називають:

- а) варіюючою;
- б) альтернативною;
- в) атрибутивною;
- г) кількісною.

10. Першою стадією статистичного дослідження слугує:

- а) розрахунок і аналіз узагальнюючих зведених показників;
- б) зведення і групування;
- в) статистичне спостереження;
- г) збір статистичних даних.

11. Другою стадією статистичного дослідження слугує:

- а) збір статистичних даних;
- б) статистичне спостереження;
- в) розрахунок і аналіз узагальнюючих зведених показників;
- г) зведення і групування.

Питання для самоконтролю

1. Засновником статистичної науки вважають?
2. В яких країнах одержала найбільший розвиток статистика в давнину?
3. Охарактеризуйте «школу політичних арифметиків».
4. Сформулюйте предмет статистики.
5. Які завдання поставлені перед статистикою в Україні?
6. Які основні складові частини статистики?
7. Що таке «статистична закономірність»?
8. Як Ви розумієте поняття «статистична сукупність»?
9. Які є основні методи статистичного дослідження?
10. Які основні функції органів статистики?
11. Дайте характеристику статистичної ознаки.

РОЗДІЛ 2. СТАТИСТИЧНЕ СПОСТЕРЕЖЕННЯ

План (логіка) викладу та засвоєння матеріалу:

- 2.1. Організаційні форми статистичного спостереження
 - 2.2. План статистичного спостереження
 - 2.3. Види та способи статистичного спостереження
 - 2.4. Помилки статистичного спостереження
- Тестові завдання
Питання для самоконтролю

2.1. Організаційні форми статистичного спостереження

Характеристика явищ і процесів, які досліджуються статистичною наукою забезпечується завдяки використанню відповідної інформації, котру одержують за результатами статистичного спостереження.

Першою стадією статистичного дослідження слугує статистичне спостереження, повнота й достовірність якого значною мірою впливає на кінцеві результати.

Статистичне спостереження є не що інше, як науково організований, систематизований, плановий збір статистичної інформації за допомогою якої оцінюється економічна, екологічна, соціальна та інші найважливіші складові діяльності суспільства та його інституціональних одиниць на макро-, меза- та мікрорівнях. Організація статистичного спостереження здійснюється завдяки використанню адекватних методів і статистичної методології. Лише завдяки правильному, науково-обґрунтованому їх застосуванню й науковій обробці можна забезпечити досягнення основної цілі – неупередженості в процесі збору необхідної інформації. Основне завдання статистичного спостереження полягає в отриманні новітньої, повної і, головне об'єктивної інформації.

Статистична інформація – це сукупність результатів статистичного дослідження, котрі використовуються для потреб

оцінювання рівня, стану, закономірностей розвитку економічних, соціальних, екологічних, національних та інших важливих складових суспільного відтворювального процесу.

Об'єктивність інформаційного забезпечення слугує основою статистичного оцінювання розвитку явищ і процесів суспільного життя.

Особливо важливе місце належить інформації в умовах сьогодення. Хто володіє інформацією, той володіє світом. Без інформації суспільний розвиток часто унеможлиблюється й гальмується. Інформаційний продукт відіграє надзвичайно важливе значення для розв'язання завдань на мікро-, мезо- й макрорівнях. При цьому користувачі, як правило вимагають різносторонньої інформації.

Належне інформаційне забезпечення слугує першопричиною досягнення динамічного й ефективного економічного та соціального піднесення багатьох держав. Водночас ті держави, котрі не володіють у достатній мірі обсягом інформаційних даних, або вони мають суб'єктивний характер, об'єктивно зазнавали і зазнають невдач у своєму розвитку. Сьогодні як ніколи актуальним слід вважати вислів С. Лемма: «Інформація – це влада».

Прийняття того чи іншого рішення, тієї чи іншої дії має забезпечуватись завдяки використанню сучасної, об'єктивної інформації. Особливо зростає роль інформаційного забезпечення та інформатизації в період інтеграції економіки нашої держави до країн ЄС. Зазначимо, що проголошення Україною проєвропейських засад розвитку докорінно змінило існуючу раніше дію інформаційного простору. Йдеться про його розширення та диверсифікацію. Поряд із державними суб'єктами підприємницької діяльності нині в нашій країні успішно функціонують підприємства приватної та комунальної форм власності.

Розширення інформатизації в Україні засвідчує про створення нової, досить потужної ринкової структури, котра представлена тіньовим капіталом – фінансовими олігархами. Означене призвело до виникнення певних ускладнень щодо отримання інформації через

численні штучні перепони: комерційна таємниця, швидкий ріст і банкрутство малих і середніх підприємств, недосконалість податкового законодавства, порушення окремих законодавчих норм тощо.

Складність полягає також у тому, що виникла потреба в розширенні інформаційних даних, які по суті є новими в порівнянні з традиційно існуючими. Зокрема це стосується отримання інформації щодо встановлення партнерів та конкурентів по бізнесу, розробки та запровадження європейських стандартів формування споживчого кошика, визначення механізму індексації доходів населення, забезпечення ринкової рівноваги, імплементація показників, які характеризують банкрутство підприємств, рівень безробіття тощо.

Зростання вимог та якісна зміна характеру одержання необхідної інформації зумовлює необхідність об'єктивного наукового обґрунтування мети та способу отримання інформації, механізму її обробки та використання. Таким чином, формування сучасної інформаційної бази потребує застосування якісно нових теоретичних і методологічних засад до збору даних уже на першому етапі статистичного дослідження, котрим слугує статистичне спостереження.

Статистичне спостереження умовно можна поділити на такі етапи:

- 1) підготовка до проведення спостереження;
- 2) реєстрація статистичних даних та формування інформаційної бази;
- 3) контроль результатів спостереження.

Перший етап – найвідповідальніший, оскільки підготовка до проведення статистичного спостереження передбачає розв'язання широкого кола питань як методологічного так і організаційного характеру. Саме тут забезпечується конкретна відповідь: що і як вивчатиметься? який об'єкт буде носієм інформації та її характер? хто, де, коли і за які кошти проводитиме спостереження; матеріально-технічна та фінансова база спостереження? період

спостереження? порядок розв'язання методологічних, організаційних та інших питань.

На другому етапі відбувається накопичення бази даних відповідною статистичною інформацією. Первинна інформація систематизується за певними ознаками й піддається обробітку.

На третьому етапі результати спостереження уточнюються і підлягають відповідному арифметичному й логічному контролю.

Статистичне спостереження як метод дослідження відрізняється від інших методів збирання даних, передусім характером і масовістю інформації та способами їх отримання. Воно здійснюється за допомогою двох організаційних форм: **статистичної звітності та спеціально організованого статистичного спостереження.**

Звітність є найбільш поширеною формою статистичного спостереження. Вона (форма, зміст, порядок і строки подання) затверджується Державною службою статистики України. Джерелом звітності слугують документи оперативно-технічного і бухгалтерського обліку. Кожний звітний документ має свої реквізити: заголовок, номер і дату затвердження, назву підприємства та міністерства, підписи осіб, які юридично відповідають за подану інформацію тощо.

За рівнем підпорядкування звітність поділяється на **загальнодержавну та внутрівідомчу**; за способами подання – **поштовою та електронну.**

Крім цього, розрізняють типову й спеціалізовану звітність. Перша, має єдину форму та зміст для всіх юридичних і фізичних осіб. Спеціалізовану звітність подають лише окремі суб'єкти діяльності, що мають певні особливості.

За часом подання звітність поділяють на **щоденну, тижневу, декадну, місячну, квартальну, піврічну та річну.** Використовують також і інший поділ звітності в часі, котра обумовлюється специфікою діяльності юридичної або фізичної особи.

До спеціально організованого статистичного спостереження відносять переписи, одноразові обліки, опитування, вибірккові,

монографічні та інші обстеження. Їх проводять у випадках недоцільності використання звітності або в зв'язку з потребою її уточнення. Потреба використання тієї чи іншої форми регулюється необхідністю більш конкретного вивчення окремих явищ та процесів.

Переписи проводяться періодично або одноразово з метою одержання повної характеристики досліджуваного масового явища на заданий момент часу. Перепис населення проводиться через кожні десять років станом на визначену дату. Перепис є особливим видом одноразового спостереження, використання якого пов'язане насамперед із встановленням чисельності населення в країні, регіоні, населеному пункті тощо. Досвід проведення перепису населення в Україні в січні 2002 року показує, що основними методичними вимогами щодо його проведення були:

- 1) співпадання критичної точки і термінів перепису в розрізі всіх об'єктів і територій спостереження;
- 2) визначення такого періоду в році, який би сповна дозволив виконати завдання перепису;
- 3) перепис потрібно проводити в максимально короткий строк;
- 4) при організації перепису доцільно запозичувати практику його проведення тих країн, які мають досвід та розраховують показники, за допомогою яких забезпечується об'єктивне оцінювання демографічних процесів.

У практиці проведення переписів населення інколи з метою апробації методики і методів дослідження організації перепису проводять «пробні» переписи, тобто спостереження вибіркової частини об'єкта.

Одноразові обліки часто проводяться по конкретних суб'єктах підприємницької діяльності з використанням інструктивного матеріалу статистичного органу. Наприклад, переписи промислового устаткування, залишків сировини і матеріалів, плодово-ягідних насаджень, облік худоби і птиці, посівних площ та інше.

Спеціальні статистичні спостереження, як правило, здійснюють за допомогою вибірових спостережень. Прикладом їх проведення можуть слугувати обстеження бюджетів сімей, бюджетів часу населення, рівня цін, громадської думки. Щодо інших форм спеціально організованого статистичного спостереження то вони будуть більш детально розглянуті в наступних темах навчального посібника.

2.2. План статистичного спостереження

Організація процесу статистичного спостереження повинна передувати складання плану, використання котрого повинно забезпечити вирішення завдань двох типів: програмно-методологічного та організаційного.

Завдання програмно-методологічного характеру включають:

- 1) формулювання мети спостереження;
- 2) встановлення об'єкта, одиниці та елементів спостереження;
- 3) складання програми статистичного спостереження.

Мета спостереження визначається конкретними потребами в статистичній інформації. Тому її слід формулювати чітко, доступно, лаконічно та по суті.

Об'єктом спостереження слугує вся сукупність явищ і процесів, які підлягають дослідженню. При цьому важливим заходом слід вважати встановлення істотних ознак об'єкта та найбільш суттєвих рис. Зокрема, при дослідженні складу та структури корпоративних підприємств визначають перелік ознак, які слугують критерієм приналежності їх до суб'єктів підприємницької діяльності зазначеного типу.

Одиниця спостереження є органічною складовою об'єкта спостереження, оскільки слугує джерелом інформації та може охоплювати одну або декілька ознак дослідження. Одиницею

статистичного спостереження, залежно від поставлених завдань можуть бути банк, організація, сім'я, кооператив, орендне підприємство тощо.

Носіями ознак, які підлягають реєстрації в статистичних формулярах, є **елементи сукупності**. Зокрема, при вивченні бюджету сімей одиницею спостереження може бути сім'я, а елементом сукупності – член сім'ї.

Програма статистичного спостереження містить перелік ознак, які підлягають реєстрації суб'єктами дослідження. Вона складається з урахуванням мети та завдань дослідження й повинна включати лише ті питання (ознаки), за допомогою яких оцінюється сутність досліджуваного явища. Програма спостереження може мати форму анкети, переписного аркуша, відомості або простого бланка.

Відповіді на запитання програми спостереження записують у статистичних формулярах. Правильність одержаних відповідей на питання, поставлених в формулярі забезпечується складанням інструкції, котра містить сукупність роз'яснень і вказівок щодо програми спостереження.

Формуляри бувають **індивідуальні** та **спискові**. Індивідуальний – містить інформацію, що стосується окремої одиниці спостереження. В списковому формулярі реєструють не одну, а декілька одиниць спостереження.

Організаційна частина плану використовується для розв'язання завдань, які стосуються визначення місця, часу і способів спостереження, а також питань, які потрібно вирішити в зв'язку з забезпеченням підготовки й інструктування кадрів, матеріально-технічною базою, комплексом підготовчих робіт тощо.

Місцем спостереження є спеціально визначений пункт, в якому безпосередньо реєструються ознаки окремих одиниць сукупності в статистичних формулярах. Його обирають залежно від особливостей і способу спостереження. Зокрема, при дослідженні на підприємстві процесів виконання норм виробітку окремими робітниками, інформацію реєструють безпосередньо біля місця

прикладання праці робітника, якість знань студентів оцінюють за результатами сесії - на основі екзаменаційних відомостей.

Час спостереження – період часу, протягом якого проводиться спостереження. Час (година, день) на який проводиться реєстрація ознак елементів сукупності, називається критичним моментом спостереження. Так, критичним моментом перепису населення 2001 року була 24 година в ніч з 4 на 5 січня 2002 року.

Статистичні спостереження, як правило здійснюються державними органами статистики. Інколи ініціаторами статистичного спостереження можуть бути громадські організації, відомства, інституціональні одиниці тощо.

Одним із організаційних завдань статистичного спостереження є підготовка кадрів і їх інструктаж. Перед проведенням статистичного спостереження виконують ряд підготовчих заходів, які включають навчання кадрів, їх добір і розстановку, інструктаж.

Матеріально-технічне забезпечення передбачає вирішення питань транспорту, оргтехніки, фінансування, інструментарію тощо.

Важливою складовою ведення статистичного спостереження є організація підготовки документації, що включає складання списків, друкування пояснювальних матеріалів, стендів та інше.

Складовою організаційних завдань плану статистичного спостереження є поліпшення його та висвітлення в засобах масової інформації, комунікативних джерелах.

2.3. Види та способи статистичного спостереження

Статистичне спостереження поділяється на два види: **суцільне і несучільне**. При **суцільному** – реєстрації та обстеженню підлягають усі одиниці сукупності. Наприклад, перепис населення, подання обов'язкової статистичної звітності інституціональними одиницями, декларування доходів тощо.

При **несучільному спостереженні** вивчається лише частина сукупності, що відібрана за певними правилами. Воно

поділяється на спостереження основного масиву, вибіркоче, анкетне, монографічне.

Спостереження основного масиву охоплює дослідження переважної частини сукупності, яка складає понад 50% загальної її кількості. Наприклад, рівень цін ринку на продовольчу групу товарів, обсяг доходів окремих категорій населення.

Вибіркове спостереження охоплює дослідження певної частини сукупності, відібраної за встановленими правилами та способами.

Анкетне спостереження базується на зборі інформації за допомогою анкети. На питання анкет відповідають окремі респонденти, думка яких може враховуватись при прийнятті певного рішення.

Монографічне спостереження вирізняється значно більшим рівнем глибини дослідження значень ознаки окремої одиниці (одиниць).

За часом проведення статистичні спостереження поділяють на **поточні, періодичні та одноразові**. При поточному явища реєструються в міру їх здійснення. Наприклад, реєстрація шлюбів, розлучень, народжень.

Періодичні проводяться регулярно, через певні проміжки часу. Наприклад, інвентаризація матеріальних цінностей, переписи.

Одноразові статистичні спостереження проводяться з метою розв'язання певних соціально-економічних завдань. Наприклад, вивчення громадської думки щодо діяльності правоохоронних органів, формування цін на окремі види продовольчих товарів тощо.

Статистичне спостереження може здійснюватися трьома способами: **безпосередній та документальний облік і опитування респондентів**.

Безпосередній облік базується на тому, що статистичні формуляри заповнюються на основі особистого підрахунку,

зважування, вимірювання. **При документальному обліку** інформацію одержують на основі даних документів оперативно-технічного, бухгалтерського, управлінського обліку. Опитування респондентів забезпечується зі слів осіб, залучених до спостереження. Воно поділяється на експедиційне, самореєстрації, кореспондентське та анкетне. Експедиційне виконують спеціально підготовлені реєстратори, котрими заповнюються формуляри спостереження, а також перевіряється правильність заповнення формулярів.

Самореєстрація здійснюється шляхом самостійного заповнення респондентами статистичних формулярів.

Кореспондентське опитування виконують самі дописувачі.

При анкетному спостереженні статистичні дані одержують за допомогою спеціальних анкет.

2.4. Помилки статистичного спостереження

В процесі проведення статистичного спостереження можуть виникати помилки, які можуть спричинити розбіжності між фактично отриманими результатами і даними спостереження. Їх називають **помилками спостереження**. Розрізняють **помилки реєстрації і репрезентативні**. Перші, виникають унаслідок неправильного встановлення фактів або недостовірних записів у статистичних формулярах. Вони поділяються на **випадкові і систематичні**.

Випадкові виникають у результаті дії випадкових причин і тому вони, як правило, не є суттєвими.

Систематичні помилки виникають у результаті нечіткого формулювання мети та програми. Вони поділяються на **навмисні й ненавмисні**. Навмисні помилки можуть суттєво негативно вплинути на результати спостереження, оскільки засвідчують про свідоме перекручення статистичної інформації.

Ненавмисні є результатом допущення помилок програми спостереження, недосвідченості та низької професійності

реєстраторів або респондентів. Прикладом їх виникнення може слугувати недостатнє матеріальне, часове і просторове обстеження об'єкта та одиниць спостереження, відсутність логіки та послідовності при складанні формулярів. Ці помилки виникають часто при заокругленні віку до чисел, кратних 5 чи 10 (збільшення) у дитячому віці і відповідно (зменшення) у похилому віці. Ненавмисні помилки виникають також у формі описок, обмовок.

Помилки репрезентативності властиві лише несущільному спостереженню і часто зумовлені недостатньою кількістю відбору одиниць сукупності у вибірку. Крім того, вони можуть виникнути внаслідок порушення правил відбору.

Виправлення помилок спостереження здійснюється за допомогою логічного та арифметичного контролю.

Логічний контроль проводять шляхом зіставлення відповідей на взаємопов'язані у формулярі питання. Неузгодженість засвідчує про наявність помилки. Наприклад, помилковим є запис у формулярі про те, що дев'ятирічна дитина заміжня або, що основним джерелом її доходів є стипендія.

Арифметичний контроль полягає у перерахунку статистичних даних за допомогою виконання арифметичних дій на основі інформації, що міститься в статистичному формулярі. Наприклад, дохід домогосподарства можна обчислити з певною точністю, підсумовуючи при цьому основні дохідні джерела.

Тестові завдання

- 1. Основне завдання статистичного спостереження полягає в:**
 - а) отриманні новітньої, повної та об'єктивної інформації;
 - б) зборі та систематизації статистичної інформації;
 - в) оцінці мінливості значень варіюючої ознаки;
 - г) підрахунку та аналізі результатів статистичного спостереження.
- 2. Статистичне спостереження здійснюється за допомогою:**
 - а) чотирьох організаційних форм;

- б) трьох організаційних форм;
- в) двох організаційних форм;
- г) жодна з відповідей не є вірною.

3. За рівнем підпорядкування звітність поділяється:

- а) місцеву та централізовану;
- б) поштову та телеграфну;
- в) щоденну та місячну;
- г) загальнодержавну та внутрівідомчу.

4. До спеціально організованого статистичного спостереження відносять:

- а) звітність суб'єктів підприємницької діяльності;
- б) переписи, одноразові обліки, опитування, вибіркові та монографічні обстеження;
- в) внутрівідомчу звітність ;
- г) спеціалізовану звітність.

5. «Пробні» переписи – це спостереження:

- а) повторно всієї статистичної сукупності;
- б) вибіркової частини об'єкта;
- в) що передують процесу суцільного перепису;
- г) основного масиву.

6. Програма статистичного спостереження містить:

- а) перелік ознак, які підлягають реєстрації;
- б) вимоги щодо проведення статистичного дослідження;
- в) перелік завдань дослідження;
- г) вимоги до завдань і мети дослідження.

7. Несуцільне спостереження поділяється на:

- а) вибіркоче;
- б) спостереження основного масиву, вибіркоче, анкетне, спеціальне;
- в) спостереження основного масиву, вибіркоче, анкетне, монографічне;
- г) одноразове та періодичне.

8. Статистичне спостереження можна проводити за допомогою такої кількості способів:

- а) двох;
- б) трьох;
- в) одного;
- г) чотирьох.

9. Помилки реєстрації виникають унаслідок:

- а) відсутності плану та програми спостереження;
- б) недостовірних записів у статистичних формулярах;
- в) неправильного встановлення фактів або недостовірних записів у статистичних формулярах;
- г) допущення помилок програми спостереження.

10. Помилки репрезентативності є результатом:

- а) свідомого перекручування фактів;
- б) недостатньої кількості відбору одиниць сукупності;
- в) низької професійності реєстраторів;
- г) неузгодженості в часі та просторі.

Питання для самоконтролю

1. Дайте визначення статистичного спостереження.
2. Яка роль інформації в умовах сучасності?
3. Які існують основні форми статистичного спостереження?
4. Дайте характеристику спеціально організованого статистичного спостереження.
5. З яких основних розділів складається план статистичного спостереження?
6. Які основні завдання програмно-методологічної частини плану статистичного спостереження?
7. Які основні завдання організаційної частини плану статистичного спостереження?
8. Як слід розуміти суцільне і несучільне спостереження?
9. З чого складається програма статистичного спостереження?
10. Дайте тлумачення критичного моменту статистичного спостереження.
11. Охарактеризуйте помилки спостереження.
12. Яка різниця між помилками реєстрації і помилками

репрезентативності?

13. Як слід розуміти логічний і арифметичний контроль помилок?

РОЗДІЛ 3. ЗВЕДЕННЯ І ГРУПУВАННЯ СТАТИСТИЧНИХ ДАНИХ

3.1. Статистичне зведення

Одним із найбільш значущих етапів статистичного дослідження слугує зведення та групування результатів спостереження. Інформацію, котру одержують при цьому упорядковують, підсумовують і узагальнюють. Упорядкування, систематизація та наукова обробка даних статистичного дослідження називається **статистичним зведенням**. Воно буває **просте і складне**. При **простому** – підсумовують лише одиниці спостереження. **Складне зведення** здійснюється крізь поділ сукупності на групи або підгрупи. При цьому проводять також підрахунок групових і загальних підсумків, а результати дослідження подають у вигляді статистичних таблиць.

Зведення результатів статистичного спостереження залежить від поставленої мети і завдань дослідження. Обов'язковою складовою зведення слугує встановлення групувальних ознак, визначення кількості груп, а також особливостей розміщення та подання даних.

Взагалі сам процес зведення умовно можна поділити на дві частини:

- 1) розробка програми статистичного зведення;
- 2) пошук і використання узагальнюючих показників результатів спостереження з метою встановлення характерних рис та істотних властивостей досліджуваного явища, а також виявлення

закономірностей;

3) визначення наочних способів подання результатів зведення;

4) встановлення порядку розрахунку одержаних результатів статистичного спостереження;

5) безпосереднє виконання статистичного зведення згідно приведеного переліку в пунктах 1, 2, 3, 4.

Програма зведення має чітко регулювати встановлені територіальні межі, адміністративне та відомче підпорядкування, сфера приналежності, форми власності та організаційні принципи виробничої й технологічної діяльності.

Процесу проведення статистичного зведення має передувати обов'язково логічний та арифметичний контроль. Контролю підлягає як сам процес так і результати зведення.

Інколи для узагальнення результатів статистичного зведення використовують спеціальні показники за допомогою яких характеризується специфіка та особливості досліджуваних явищ і процесів.

Результати зведення подають у формі статистичних таблиць, макети котрих розробляються в процесі складання програми. Обов'язковими складовими статистичних формулярів (таблиць) має бути посилання на об'єкт та час спостереження, одиниці вимірювання, територіальну поширеність зведення, тощо.

Статистичне зведення за організацією та формою обробки поділяється на централізоване та децентралізоване. Перше, здійснюється шляхом подання матеріалів статистичного спостереження у вищестоящі органи статистики (Державна служба статистики України). Тут їх обробляють за певною завчасно встановленою програмою. Прикладом такого зведення можуть бути матеріали переписів, соціологічних опитувань, які надсилають до центрального органу, де вони обробляються.

Особливо зросла роль централізованого зведення зросла в умовах інтеграції Державної служби статистики України до вимог і стандартів статистичних служб країн ЄС. Йдеться про потребу прискорення та завершення в Україні запровадження Єдиної

статистичної інформаційної системи (ЄСІС). Означене потребує проведення адекватних дій та стандартів, які стосуються організації роботи зі збору, розробки, узагальнення та аналізу статистичної інформації, обґрунтування достовірних висновків та пропозицій. Запровадження системи передбачає широке використання засобів найсучаснішої електронної та комп'ютерної техніки, новітніх технологій, прогресивних економіко-математичних методів моделювання, комунікативних засобів та поліграфічної техніки тощо.

Більш широке використання в умовах інтеграції вітчизняної економіки до стандартів, вимог і принципів до країн ЄС одержало в статистичній практиці децентралізоване зведення, котре організоване в ієрархічній підпорядкованості.

Зокрема, при обробці статистичної звітності, зведення спочатку здійснюється первинними ланками управління (мікрорівні). Одержані результати зведення наростаючим підсумком зводяться по всій сукупності на мікро-, меза- та макрорівнях. При цьому загальні результати зведення подаються в державну службу статистики України, Тут вони узагальнюються по країні в цілому.

3.2. Статистичні групування

Результати статистичного зведення не завжди в достатній мірі віддзеркалюють суттєві властивості досліджуваних явищ, характеризують закономірності їх розвитку. Розв'язання цього завдання в статистиці забезпечується завдяки використанню методу групувань.

Групування – розподіл статистичної сукупності на групи, підгрупи за істотними ознаками з метою встановлення зв'язку та закономірностей розвитку досліджуваних явищ і процесів предмета статистики.

Метод групувань є одним із найбільш поширених інструментаріїв наукових досліджень й обробки інформації за результатами статистичного спостереження. Основні функції методу групування полягають у:

- 1) виявленні однорідних соціально-економічних явищ, шляхом поділу їх на однорідні групи і підгрупи;
- 2) статистичній оцінці структури явищ та структурних зрушень;
- 3) встановленні взаємозв'язку та взаємозалежності між досліджуваними ознаками.

Практична реалізація приведених функцій у статистиці забезпечується завдяки використанню трьох видів групувань: **типологічного, структурного та аналітичного.**

Поділ статистичних групувань на окремі види є певною мірою умовним, оскільки застосування одного з них може забезпечити розв'язання декількох завдань: визначити тип ознаки, встановити структуру та взаємозв'язок між досліджуваними явищами та процесами.

Організація та сам процес проведення статистичного групування повинні забезпечувати одержання об'єктивних показників, за допомогою яких здійснюється всебічний аналіз і формулюються висновки результатів дослідження. Об'єктивність результатів групування забезпечується через дотримання певних правил, принципів статистичного групування, а саме

- мінімальна кількість одиниць спостереження повинна становити не менше 20;
- статистична сукупність, яка використовується для проведення статистичного групування за своїм складом має бути однотипною (за винятком типологічного);
- в основу групування потрібно покласти ознаки, котрі найбільш об'єктивно та повно характеризують внутрішні особливості досліджуваних явищ.

Типологічне групування – розподіл якісно неоднорідної досліджуваної статистичної сукупності на соціально-економічні типи, однорідні групи та підгрупи. Найбільш часто даний вид групування використовують при дослідженні соціально-економічних явищ і процесів. Типологічне групування можна використовувати при статистичному дослідженні соціуму завдяки поділу населення на окремі соціальні групи, оцінці типових явищ, характеристиці форм

власності, а також визначенні ознак, властивості котрих є спільними для всіх груп.

Зокрема, статистичні групування за формами власності суб'єктів підприємницької діяльності, зайнятих у виробничих, обслуговуючих та сферах надання послуг класифікують та відносять до типологічних. Взагалі типологічні групування одержали достатньо широке використання в умовах командно-адміністративної прокомуністичної моделі господарювання. Вони часто штучно виокремлювались із метою штучного утвердження величності прокомуністичної ідеології. З цією метою часто проводилось групування населення на класи, соціально-економічні групи. Підприємства групувались за формами власності. При цьому результати групування практично завжди мали домінуючу роль, пріоритетність при характеристиці державної форми власності, вказуючи на переваги соціалістичного ладу над капіталістичним. Часто такий вид групування носив надуманий, псевдооб'єктивний характер.

Запровадження ринкової моделі соціально-економічного розвитку дещо обмежило використання під час статистичних досліджень типологічного виду групування. Нині його використовують поєднуючи типологічні і структурні види групувань. Зокрема до типологічного виду, зважаючи на змістове його тлумачення можна віднести кількість об'єктів ЄДРПОУ за організаційно-правовими формами господарювання.

Таблиця 3.1

Кількість об'єктів ЄДРПОУ за організаційно-правовими формами господарювання

(на початок року)

| Організаційно-правова форма господарювання | 2007 | 2010 | 2013 | 2014* | 2015* |
|---|-------------|-------------|-------------|--------------|--------------|
| Усього | 1070705 | 1258513 | 1341781 | 1372177 | 1331230 |
| у тому числі | | | | | |
| фермерське господарство | 48918 | 49764 | 49095 | 49169 | 47356 |

| | | | | | |
|---|--------|--------|--------|--------|--------|
| приватне підприємство | 248211 | 283697 | 280073 | 278227 | 259772 |
| державне підприємство | 7562 | 6811 | 6075 | 5933 | 5559 |
| казенне підприємство | 50 | 44 | 38 | 33 | 32 |
| комунальне підприємство | 16688 | 15994 | 15207 | 14974 | 13778 |
| дочірнє підприємство | 23127 | 20408 | 18594 | 18173 | 16981 |
| іноземне підприємство | 2315 | 2272 | 2176 | 2096 | 1988 |
| підприємство об'єднання громадян (релігійної організації, профспілки) | 4369 | 4568 | 4498 | 4470 | 4056 |
| підприємство споживчої кооперації | 860 | 1291 | 1360 | 1325 | 1299 |
| акціонерне товариство | 33976 | 30169 | 25531 | 24813 | 23110 |
| товариство з обмеженою відповідальністю | 325925 | 418145 | 488781 | 515371 | 519607 |
| товариство з додатковою відповідальністю | 700 | 782 | 1539 | 1583 | 1589 |
| повне товариство | 1997 | 2084 | 2074 | 2062 | 1941 |
| командитне товариство | 690 | 631 | 638 | 629 | 595 |
| кооператив | 30790 | 35063 | 33664 | 33806 | 29681 |
| з нього | | | | | |
| виробничий | 2556 | 2461 | 2459 | 2647 | 2646 |
| обслуговуючий | 16645 | 22620 | 20619 | 20620 | 17899 |
| споживчий | 215 | 591 | 771 | 809 | 726 |
| сільськогосподарський виробничий | 1324 | 1401 | 1224 | 1192 | 1064 |
| Сільськогосподарський обслуговуючий | 372 | 759 | 947 | 1035 | 928 |

| | | | | | |
|---|-------|--------|--------|--------|--------|
| організація (установа, заклад) | 85821 | 108552 | 115042 | 114018 | 109545 |
| з неї | | | | | |
| державна | 19313 | 15786 | 16479 | 14941 | 13744 |
| комунальна | 53009 | 65308 | 66378 | 67023 | 64975 |
| приватна | 641 | 907 | 1000 | 1027 | 1005 |
| організація об'єднання громадян | 570 | 1393 | 1579 | 1608 | 1618 |
| асоціація | 2751 | 3125 | 3194 | 3234 | 3108 |
| корпорація | 824 | 866 | 852 | 836 | 777 |
| консорціум | 80 | 92 | 93 | 94 | 90 |
| концерн | 406 | 390 | 359 | 350 | 327 |
| інші об'єднання юридичних осіб | 1971 | 1634 | 1412 | 1365 | 1275 |
| філія (інший відокремлений підрозділ) | 51429 | 52973 | 49014 | 49717 | 47507 |
| представництво | 3810 | 4330 | 4744 | 4843 | 4647 |
| товарна біржа | 221 | 559 | 598 | 601 | 593 |
| кредитна спілка | 737 | 1127 | 1154 | 1171 | 1105 |
| споживче товариство | 5656 | 5252 | 5044 | 5162 | 5026 |
| спілка споживчих товариств | 524 | 404 | 363 | 378 | 363 |
| недержавний пенсійний фонд | 11 | 83 | 75 | 74 | 72 |
| політична партія | 13976 | 17375 | 19222 | 19166 | 18237 |
| громадська організація | 46682 | 63899 | 74500 | 77286 | 75828 |
| релігійна організація | 18617 | 22343 | 24720 | 25475 | 24957 |
| профспілка, об'єднання профспілок | 15639 | 24649 | 28852 | 29724 | 28890 |
| творча спілка (інша професійна організація) | 180 | 240 | 278 | 298 | 277 |
| благодійна організація | 9590 | 12267 | 14055 | 14999 | 15934 |
| об'єднання | 4159 | 10329 | 15018 | 16213 | 15992 |

| | | | | | |
|---|-----|------|------|------|------|
| співвласників багатоквартирного будинку | | | | | |
| орган самоорганізації населення | 425 | 1152 | 1426 | 1503 | 1372 |

* Дані наведені без урахування тимчасово окупованої території Автономної Республіки Крим

Суть структурного групування полягає в розподілі якісно однорідної сукупності на групи за певними ознаками, що слугує передумовою оцінювання структури досліджуваної сукупності. На основі даних табл. 1 проведемо структуру прикладу і застосування структурних групувань (рис. 1).



Рис. 3.1. Структура об'єктів ЄДРПОУ за видами економічної діяльності

Наведені приклади типологічного й структурного групування засвідчують про взаємозв'язок між ними, оскільки вони повністю збігаються за формою і відрізняються лише за метою дослідження.

Структурні групування проведені за територіальною ознакою або в територіальному розрізі в статистиці мають назву **територіальних**.

Типологічні й структурні групування в абсолютній своїй більшості мають описовий характер. Їх використання дає можливість

виявити характерні риси та особливості досліджуваних ознак, а також проаналізувати структуру сукупності.

Більш дієвим методичним інструментом організації наукових досліджень і узагальнення матеріалів статистичного спостереження в статистичній науці одержали аналітичні групування.

Їх використовують із метою встановлення взаємозв'язку між факторними та результативними ознаками. За допомогою цього виду групування оцінюють наявність та напрям зв'язку. Аналітичний зв'язок буває прямим і зворотним. Наприклад, із збільшенням величини депозитних вкладів збільшується сума нарахованих відсотків – прямий зв'язок; із зростанням обсягу товарообігу витрати обігу зменшуються - зворотний зв'язок. Прикладом аналітичного групування є також взаємозв'язок між кількістю народжених дітей та віком матері, що сформувався в Україні за останні 15 років (табл.3.2).

Таблиця 3.2

Коефіцієнти народжуваності за віковими групами

(кількість народжених за рік на 1000 жінок відповідної вікової групи)

| | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 |
|---------------------------------------|-------|-------|-------|-------|-------|-------|-------|-------|
| Жінки у віці 15-49 років у тому числі | 38,8 | 42,4 | 43,1 | 42,5 | 43,6 | 45,9 | 45 | 44,5 |
| молодшому за 20 | 30,3 | 32 | 31,2 | 28,8 | 28,1 | 28,7 | 27,2 | 27 |
| 20-24 | 92,2 | 97,5 | 94,8 | 90,1 | 89,9 | 93,6 | 91 | 89,9 |
| 25-29 | 81,3 | 87,8 | 89 | 87,9 | 89,2 | 93,6 | 91,5 | 91,3 |
| 30-34 | 45,4 | 51,1 | 54,1 | 55,1 | 58 | 61,4 | 61,2 | 60,6 |
| 35-39 | 16,8 | 19,7 | 21,5 | 22,3 | 24,6 | 26,4 | 27,2 | 27,6 |
| 40-44 | 2,9 | 3,3 | 3,8 | 4,2 | 4,6 | 5 | 5,2 | 5,5 |
| 45-49 | 0,1 | 0,2 | 0,2 | 0,2 | 0,2 | 0,3 | 0,3 | 0,4 |
| Сумарний коефіцієнт народжуваності | 1,345 | 1,458 | 1,473 | 1,443 | 1,459 | 1,531 | 1,506 | 1,498 |

Коефіцієнти народжуваності, наведені в таблиці 3.2 засвідчують, що в Україні найбільш дітородними є жінки у віці 20-24 та 25-29 роки.

Про це переконливо засвідчує динаміка коефіцієнтів народжуваності. Після 29 років, тобто із зростанням віку кількість дітей народжуваних стабільно знижується.

3.3. Методологія статистичного групування

Статистичне групування здійснюється з використанням наступної методичної послідовності: вибір групувальної ознаки, визначення кількості груп та встановлення розміру інтервалу. Взаємозв'язані ознаки, котрі використовуються в процесі проведення статистичного групування, поділяються на факторні та результативні. Групувальну, як правило, використовують факторну ознаку. Ознака, варіація значень якої зумовлює коливання інших ознак, називається **факторною**. Ознака, значення котрої змінюється під дією впливу на неї інших чинників які покладено в основу групування називається **результативною**. Наприклад, за даними табл.2.3. до факторної ознаки буде віднесена вікова група осіб жіночої статі, а результативною – кількість народжених дітей на 1000 жінок.

Поділ ознак на факторні (причина) й результативні (наслідок) є дещо умовним та значною мірою залежить від завдання статистичного групування. Одна і та ж ознака в одних випадках може слугувати в якості результативної, а в інших – причиною.

Наприклад, розмір заробітної плати робітників – відрядників (результативна ознака) залежить від обсягу виконаних ними робіт, виготовленої продукції, надання послуг тощо (факторна ознака). Порівняння заробітної плати з розміром цін на продовольчі товари в умовах інфляції засвідчує, що їх купівельна спроможність регулюється інфляційними процесами, тобто результативність регулюється величиною інфляції.

Групувальна ознака в кожному конкретному випадку статистичного групування визначається за допомогою логічного економічного аналізу. Групування в статистиці проводиться за кількісними, а також атрибутивними ознаками. **Кількісні** – передбачають використання цифрових характеристик ознаки (зарплата, ціна, виробіток, відсоток, дивіденди). **Атрибутивні** – мають словесне вираження (національність, освіта, професія).

Різновидом атрибутивної ознаки може бути **альтернативна**, котра може набувати лише два взаємовиключні значення «так» або «ні»; «задоволений» - «незадоволений»; «чоловік» - «жінка»; «одружений» - «неодружений» тощо.

Кількісні, в свою чергу, поділяються на **дискретні** (переривчасті) і безпереривчасті.

Дискретні подаються за допомогою цілих чисел (кількість членів сім'ї, дерев, ринків, акцій тощо).

Безпереривчасті можуть набувати цифрового значення з використанням двох і більше співвимірників. Наприклад, вік людини визначатись числом років, місяців та днів; швидкість вимірюється одночасно в кілометрах та метрах руху за годину; вартість товару - в гривнях і копійках тощо.

Черговою складовою проведення статистичного групування постає потреба встановлення кількості груп і меж кожної з них.

Визначення кількості групи здійснюється за допомогою логічного аналізу, змістове навантаження котрою полягає в тому, що їх число повинно забезпечити відтворення відповідної закономірності. Наприклад, із зростанням енергоозброєності має зростати продуктивність праці; працівники вищої кваліфікації виробляють більше продукції тощо.

Кількість груп залежить також від характеру ознак, що покладені в основу групування. Зокрема, при групуванні за атрибутивною ознакою кількість груп відповідає кількості її різновидів. Наприклад, групування населення за освітою, статтю, місцем проживання тощо.

Встановлення числа груп із використанням кількісної ознаки значною мірою залежить від розмаху варіації групувальної ознаки та обсягу сукупності. Зокрема, для дискретних ознак, ступінь варіації яких відносно невеликий, число груп дорівнює кількості варіант ознаки. Наприклад, кількість працездатних або пенсіонерів у сім'ї, дітей у сім'ї, позитивних оцінок успішності тощо.

Практика статистичних досліджень засвідчує також, що кількість груп статистичного дослідження залежить від обсягу досліджуваної сукупності.

Взагалі рекомендується збільшувати їх число з ростом загальної кількості одиниць статистичного дослідження, при умові, що коливання значень ознак є близьким до нормального. Однак такі твердження не є догматичними, оскільки критерій в кінцевому рахунку регулюється тими завданнями, розв'язання яких забезпечується завдяки використанню методу групувань.

У великих статистичних сукупностях із близькими до нормального розподілу, кількість груп орієнтовно можна встановити за допомогою формули відомого американського вченого Г.А.Стерджеса:

$$n = 1 + 3.322 \lg N, \quad (3.1)$$

де N - число одиниць сукупності.

Після розрахунку числа груп у кожному конкретному статистичному дослідженні, приступають до визначення величини інтервалів.

Інтервал групування розраховується як різниця між максимальним і мінімальним значеннями групувальної ознаки. Одержаний результат ділять на число груп (n).

Інтервали бувають рівні та нерівні. Рівні інтервали зазвичай використовують у структурних та аналітичних групуваннях. Розмір інтервалу при цьому визначається за допомогою наступної формули

$$i = \frac{x_{\max} - x_{\min}}{n}, \quad (3.2)$$

де i - величина інтервалу;

x_{\max} - максимальне значення групувальної ознаки;

x_{\min} - мінімальне значення групувальної ознаки;

n - кількість груп.

При рівних інтервалах проміжок між двома значеннями у всіх групах однаковий, а при нерівних змінюється при переході від однієї групи до другої.

Обов'язковою умовою побудови рівних інтервалів є однорідність досліджуваної статистичної сукупності, а також достатньо велике число одиниць спостереження, що підтверджуватиме про наявність певної закономірності.

Інтервали при групуванні за кількісними ознаками поділяються також на закриті та відкриті, зростаючі або спадаючі. Іноколи використовують кратні інтервали.

Відсутність верхньої або нижньої межі в групі засвідчує про її відкритість. Наприклад, рівень рентабельності підприємств нафтопереробної сфери, котрі віднесено до першої групи, становить до 10 %, а тих, що віднесені до останньої групи – понад 50%.

Закриті інтервали мають верхню і нижню межу. Наприклад, розмір депозитних вкладів клієнтів банку двох групах знаходиться в межах: перший, - від 20000 до 30000 гривень; друга, від 30000 до 40000 гривень і т.д.

Використання нерівних інтервалів засвідчує про адекватну тенденцію до їх зростання або спадання.

Зростаючі інтервали характеризуються збільшенням їх розміру при переході від однієї групи до наступної. Спадаючі - навпаки.

Величина, на яку збільшується або зменшується інтервал групування від однієї групи до наступної, називається **кроком інтервалу**. Крок інтервалу може бути сталим або змінюватись.

Часто в практиці статистичного дослідження використовують нерівні інтервали групування за однією і тією самою ознакою для різних сукупностей. Такі інтервали називають **спеціалізованими**. Прикладом їх застосування може бути порівняння кількості працюючих у гірничодобувній і харчовій промисловості. Зрозуміло, що число працюючих у гірничодобувній промисловості значно

більше, аніж на харчових підприємствах. Саме тому з метою їх порівняння використовують спеціалізовані інтервали.

Досить часто при обробці статистичної інформації проводять групування за допомогою методу рівних частот, який вимагає кратної кількості одиниць спостереження.

Розглянемо на конкретному прикладі порядок утворення та використання рівних інтервалів груп, для чого використаємо дані статистичного щорічника України про середньомісячну номінальну заробітну плату найманих працівників за регіонами в 2005 році (табл.3.3.). Проведено групування регіонів України за розміром середньомісячної номінальної заробітної плати, утворивши п'ять груп (n) з рівними інтервалами. Найперше визначимо розмір інтервала. Для цього від найбільшого розміру заробітної плати віднімаємо найменший і одержаний результат поділимо на п'ять.

Таблиця 3.3.

**Середньомісячна номінальна заробітна плата
найманих працівників за регіонами**

| | 1995 | 2000 | 2005 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 |
|--|------|------|------|------|------|------|------|------|------|
| | крб. | грн. | | | | | | | |
| Україна | 73 | 230 | 806 | 2239 | 2633 | 3026 | 3265 | ... | ... |
| Україна (без урахування АР Крим і м. Севастополя) | ... | ... | ... | 2250 | 2648 | 3041 | 3282 | 3480 | 4195 |
| Автономна Республіка Крим | 70 | 225 | 730 | 1991 | 2295 | 2654 | 2850 | ... | ... |
| Області | | | | | | | | | |
| Вінниця | 58 | 159 | 597 | 1782 | 2074 | 2432 | 2651 | 2810 | 3396 |
| Волинська | 53 | 150 | 591 | 1692 | 1994 | 2339 | 2580 | 2721 | 3291 |
| Дніпропетровська | 91 | 273 | 913 | 2369 | 2790 | 3138 | 3336 | 3641 | 4366 |
| Донецька | 97 | 292 | 962 | 2549 | 3063 | 3496 | 3755 | 3858 | 4980 |
| Житомирська | 61 | 164 | 602 | 1785 | 2071 | 2369 | 2561 | 2763 | 3271 |
| Закарпатська | 50 | 172 | 665 | 1846 | 2069 | 2351 | 2553 | 2744 | 3381 |
| Запорізька | 84 | 289 | 860 | 2187 | 2607 | 2927 | 3142 | 3432 | 4200 |
| Івано-Франківська | 65 | 188 | 718 | 1927 | 2213 | 2539 | 2679 | 2875 | 3402 |
| Київська | 78 | 241 | 811 | 2295 | 2761 | 3157 | 3351 | 3489 | 4153 |
| Кіровоградська | 58 | 170 | 624 | 1815 | 2114 | 2428 | 2608 | 2789 | 3282 |
| Луганська | 82 | 232 | 805 | 2271 | 2742 | 3090 | 3337 | 3377 | 3427 |

| | | | | | | | | | |
|---------------|-----|-----|------|------|------|------|------|------|------|
| Львівська | 62 | 196 | 713 | 1941 | 2244 | 2578 | 2789 | 2961 | 3646 |
| Миколаївська | 68 | 227 | 744 | 2122 | 2448 | 2822 | 3094 | 3344 | 3984 |
| Одеська | 66 | 236 | 768 | 2046 | 2387 | 2700 | 2947 | 3129 | 3897 |
| Полтавська | 76 | 220 | 758 | 2102 | 2481 | 2850 | 2988 | 3179 | 3783 |
| Рівненська | 61 | 173 | 685 | 1960 | 2211 | 2575 | 2844 | 3033 | 3573 |
| Сумська | 66 | 194 | 663 | 1866 | 2177 | 2503 | 2702 | 2877 | 3449 |
| Тернопільська | 53 | 135 | 553 | 1659 | 1871 | 2185 | 2359 | 2527 | 2994 |
| Харківська | 72 | 230 | 759 | 2060 | 2407 | 2753 | 2975 | 3143 | 3697 |
| Херсонська | 59 | 173 | 625 | 1733 | 1970 | 2269 | 2464 | 2617 | 3123 |
| Хмельницька | 55 | 156 | 584 | 1786 | 2075 | 2425 | 2641 | 2878 | 3371 |
| Черкаська | 63 | 175 | 642 | 1835 | 2155 | 2508 | 2682 | 2829 | 3360 |
| Чернівецька | 55 | 157 | 621 | 1772 | 1985 | 2329 | 2484 | 2578 | 3050 |
| Чернігівська | 57 | 177 | 602 | 1711 | 1974 | 2308 | 2504 | 2690 | 3295 |
| Міста | | | | | | | | | |
| Київ | 100 | 405 | 1314 | 3431 | 4012 | 4607 | 5007 | 5376 | 6732 |
| Севастополь | 83 | 251 | 803 | 2167 | 2476 | 2891 | 3114 | ... | |

$$i = \frac{x_{\max} - x_{\min}}{n} = \frac{6732 - 2994}{5} \approx 748 \text{ грн.}$$

Таким чином, величина інтервалу становить 748 гривні.

Проведемо групування та визначимо межі груп:

| Групи регіонів за розміром середньомісячної номінальної заробітної плати найманих працівників, грн. | Кількість регіонів |
|---|--------------------|
| I – від 2994 до 3742 | 17 |
| II – від 3742 до 4490 | 7 |
| III – від 4490 до 5237 | 1 |
| IV – від 5237 до 5984 | 0 |
| V – від 5984 до 6732 | 1 |
| Разом | 26 |

Розрахунки при групуванні за кількісною ознакою з використанням рівних інтервалів та меж груп проведені у такій послідовності: до нижньої межі першої групи (2994 грн.) додаємо величину інтервалу ($i = 748$), в результаті одержимо величину верхньої межі першої групи, яка є одночасно нижньою межею другої групи; до нижньої межі другої групи знову додаємо розмір інтервалу ($i = 748$), утворивши при цьому верхню межу другої групи, яка одночасно виступає межею наступної групи і т.д. При

розв'язанні даної задачі нами використано рівні інтервали для кожної групи.

Одним із важливих методичних завдань при проведенні групування є встановлення границь інтервалів. Границі інтервалів мають бути визначені з встановленою точністю і визначеністю, щоб було зрозуміло до якої групи слід віднести ту чи іншу одиницю сукупності. Зокрема, регіон, де середньомісячна номінальна заробітна плата найманих працівників в 2015 році складала 635грн., слід віднести не до першої, а до другої групи, регіони, в яких місячна заробітна плата починається з 4490 грн., - до третьої в т.д. В нашій задачі побудова кожної групи регулюється за допомогою слів «від – до», що дозволяє однозначно віднести кожний регіон до відповідної групи. Такі групи називають у статистиці закритими. Наведений приклад статистичного групування засвідчує, що майже в половині регіонів України заробітна плата найманих працівників коливається в межах 2994 до 3742 гривень. Найвищу заробітну плату в 2015 році мали наймані працівники в місті Києві – 6732.

Крім закритих меж груп у статистиці використовують також відкриті. Інтервали при цьому позначаються за допомогою додаткових вказівок до них "до", "більше", "менше" і т.д. В першій групі обмежена верхня, а в останній - нижня межа. Нижче наводимо приклади застосування таких інтервалів.

| Інтервали рівні | Інтервали нерівні | |
|---|---|---|
| | зростаючі | спадаючі |
| Урожайність озимої пшениці по районах регіону, ц/га | Групи комерційних банків за прибутковіс-тю активів, % | Групи домогосподарств за відсот-ком втрат електроенергії при виробництві продукції «А», кВт/год |
| випадок I | випадок II | випадок III |
| від 20 до 25 | менше 10 | 20 і більше |
| 25-30 | 11-20 | 10-20 |
| 30-35 | 21-40 | 5-10 |
| 35-40 | 41-80 | 3-5 |
| 40 і більше | 81 і більше | менше 3 |

Наступним важливим методичним питанням у процесі проведення статистичного групування є встановлення числа груп (n).

При проведенні типологічного групування число груп, а також межі інтервалів встановлюють виходячи із соціально-економічного змісту досліджуваних ознак. Межа інтервалу має розглядатися як поріг переходу з кількості в нову якість. Число груп залежить від кількості соціально-економічних типів явищ, що вивчаються. Наприклад, при групуванні страхових компаній за формами власності використовують такі групи:

- акціонерні компанії;
- командитні товариства;
- повні товариства;
- товариства з додатковою відповідальністю тощо.

Інколи в процесі проведення аналітичного групування використовують принцип рівних частот. Проілюструємо його застосування такому прикладі: витрати електричної енергії за травень поточного року по 20 домогосподарствах одного з населених пунктів району становив: 140,3; 134,2; 139,8; 141,3; 128,3; 118,6; 124,5; 122,3; 124,8; 120,5; 128,4; 131,6; 130,3; 137,8; 139,6; 119,5; 125,8; 126,7; 119,3; 128,4 кВт/год.

Побудувати ранговий ряд споживання електроенергії, а також провести групування домогосподарств, утворивши 4 групи з рівними інтервалами, використовуючи при цьому метод рівних частот.

Спочатку побудуємо ранговий ряд споживання електроенергії: 118,6; 119,3; 119,5; 120,5; 122,3; 124,5; 124,8; 125,8; 126,7; 128,3; 128,4; 128,4; 130,3; 131,6; 134,2; 137,8; 139,6; 139,8; 140,3; 141,3 кВт/год.

Використавши метод рівних частот проведемо групування помістивши його результати в таблицю 3.4.

Таблиця 3.4

Розподіл домогосподарств за обсягом спожитої електроенергії в травні-місяці, кВт/год

| № групи | Групи домогосподарств за | Кількість |
|---------|--------------------------|-----------|
|---------|--------------------------|-----------|

| | обсягом спожитої електроенергії, кВт/год | домогосподарств |
|-------|--|-----------------|
| I | 118,6-122,3 | 5 |
| II | 124,5-128,3 | 5 |
| III | 128,4-134,2 | 5 |
| IV | 137,8-141,3 | 5 |
| Разом | | 20 |

Саме таким чином забезпечується практичне застосування методу рівних частот у процесі проведення статистичного групування.

Групування може проводитись за однією або кількома ознаками. Групування за однією ознакою – **просте**, за кількома – **складне**.

Складне групування може бути комбінаційним і багатомірним. Зокрема, групування населення за віком і статтю відноситься до комбінаційного, а поділ сімей за рівнем споживання продовольчих і непродовольчих товарів одночасно - багатомірним.

При комбінаційному групуванні первинні групи будуються за однією ознакою, а підгрупи – за іншою.

Наприклад, комерційні банки можна згрупувати за розміром статутного капіталу. Одночасно можна побудувати підгрупи, в яких можуть бути використані показники прибутковості активів, капіталу, обсягу зовнішньоекономічних операцій тощо. Число груп при такому групуванні дорівнює добутку їх числа, утворених за кожною ознакою окремо. Якщо групування проведено за чотирма ознаками і за кожною з них побудовано по чотири групи, то загальне число груп становитиме 64 (4x4x4). Звісно, що аналіз на основі такої кількості груп є надзвичайно складним і трудомістким процесом.

Багатомірне групування на відміну від комбінаційного забезпечується завдяки утворенню груп за певною множиною ознак одночасно. При цьому найчастіше використовують групування за інтегральними (синтезуючими) показниками.

Інколи в практиці статистичного дослідження виникає потреба порівняти два групування з неоднаковим числом груп. Для

забезпечення порівняльності цих групувань використовують вторинне групування. Його суть полягає в проведенні перегрупування нових даних із метою утворення нових груп використовуючи при цьому результати первинного групування.

Вторинне групування можна проводити двома способами перегрупування:

- 1) за величиною інтервалів первинного групування;
- 2) за питомою вагою окремих груп у загальному їх підсумку.

Розглянемо методику проведення вторинного групування на такому прикладі:

Таблиця 3.5

Розподіл робітників за рівнем середньомісячної заробітної плати

| Завод №1 | | Завод №2 | |
|-------------------------------|--|-------------------------------|--|
| заробітна плата, умовних г.о. | чисельність робітників у % до підсумку | заробітна плата, умовних г.о. | чисельність робітників у % до підсумку |
| 1400-1500 | 6 | 1400-1500 | 6 |
| 1500-1700 | 20 | 1500-1600 | 8 |
| 1700-1900 | 18 | 1600-1700 | 15 |
| 1900-2200 | 25 | 1700-2000 | 15 |
| 2200-2700 | 15 | 2000-2200 | 14 |
| 2700-3000 | 10 | 2200-2500 | 10 |
| 3000-3400 | 6 | 2500-2700 | 12 |
| | | 2700-2800 | 10 |
| | | 2800-3000 | 6 |
| | | 3000 і більше | 4 |
| Разом: | 100 | Разом: | 100 |

Використавши метод вторинного групування для порівняння розподілу робітників заводів за розміром заробітної плати,

збільшимо інтервали в ряді другого заводу, взявши за основу величини інтервалів розподілу робітників за заробітною платою першого заводу.

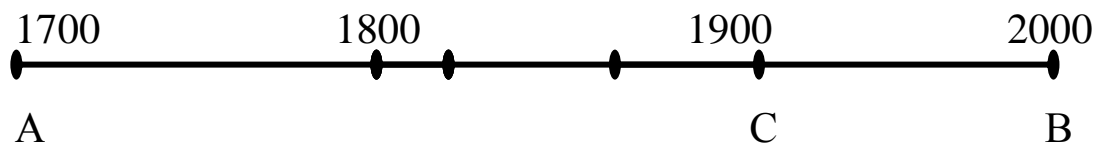
В результаті проведеного перегрупування одержимо:

Таблиця 3.6

Розподіл робітників за рівнем середньомісячної заробітної плати (вторинне групування)

| Групи робітників за розміром заробітної плати, умовних г.о | Чисельність робітників, % до підсумку | | Пояснення техніки розрахунку |
|--|---------------------------------------|----------|--------------------------------------|
| | завод №1 | завод №2 | |
| 1400-1500 | 6 | 6 | - |
| 1500-1700 | 20 | 23 | 8+15 |
| 1700-1900 | 18 | 10 | $15 - 15 \cdot \frac{1}{3}$ |
| 1900-2200 | 25 | 19 | $15 \cdot \frac{1}{3} + 14$ |
| 2200-2700 | 15 | 22 | 10+12 |
| 2700-3000 | 10 | 17 | $10 + 6 + (4 - 4 \cdot \frac{3}{4})$ |
| 3000-3400 | 6 | 3 | $4 \cdot \frac{1}{3}$ |
| Разом | 100 | 100 | - |

Пояснимо порядок розрахунку. Перші групи робітників збігаються. До другої новоствореної групи другого заводу із заробітною платою робітників 1500-1700 г.о. ввійдуть друга і третя групи робітників, сума чисельності яких дорівнює 23(8+15). Потім потрібно створити третю групу робітників із заробітною платою 1700-1900 г. о. До цієї групи ввійде $\frac{2}{3}$ четвертої групи другого заводу, тобто



Відрізок $AC = \frac{2}{3} \times AB$, що рівнозначно $AC = AB - CB = 20 - 20 \cdot \frac{1}{3} = 13,4$. Аналогічно проводимо інші розрахунки.

Тестові завдання

1. Статистичним зведенням називається:

- а) пошук та використання узагальнюючих показників результатів спостереження;
- б) визначення наочних способів подання результатів спостереження;
- в) встановлення групувальних ознак та кількості груп;
- г) упорядкування, систематизація та наукова обробка статистичних даних.

2. Статистичне зведення за організацією та формою обробки поділяється на:

- а) первинне, вторинне;
- б) суцільне, несуцільне;
- в) централізоване, децентралізоване;
- г) анкетне та спеціально організоване.

3. Групування – розподіл статистичної сукупності на :

- а) групи за певними ознаками з метою встановлення їх зв'язку;
- б) групи, підгрупи за істотними ознаками з метою встановлення зв'язку та закономірностей розвитку досліджуваних явищ і процесів;
- в) однорідні соціально-економічні явища;
- г) групи за певною типовою ознакою.

4. Типологічне групування – розподіл:

- а) якісно неоднорідної досліджуваної статистичної сукупності на соціально-економічні типи, однорідні групи та підгрупи;

- б) якісно однорідної досліджуваної статистичної сукупності на соціально-економічні типи, однорідні групи та підгрупи;
- в) якісно однорідної сукупності на групи за певними ознаками, завдяки яким оцінюється структура досліджуваної сукупності;
- г) статистичної сукупності з метою встановлення зв'язку між факторними та результативними ознаками.

5. Суть структурного групування полягає в розподілі:

- а) якісно неоднорідної досліджуваної статистичної сукупності на соціально-економічні типи;
- б) якісно однорідної статистичної сукупності на соціально-економічні типи;
- в) статистичної сукупності з метою встановлення зв'язку між факторними та результативними ознаками;
- г) якісно однорідної сукупності на групи за певними ознаками завдяки яким оцінюється структура досліджуваної ознаки.

6. Аналітичні групування використовують для:

- а) встановлення зв'язку між факторними та результативними ознаками;
- б) встановлення структури досліджуваної сукупності;
- в) встановлення типових груп;
- г) оцінки результатів статистичного зведення.

7. За допомогою аналітичного групування оцінюють:

- а) тісноту зв'язку між ознаками;
- б) наявність та напрям зв'язку між ознаками;
- в) ефективність структурного групування;
- г) ефективність типологічного групування.

8. Статистичне групування передбачає наступну послідовність його проведення:

- а) визначення кількості груп, встановлення розміру інтервалу та вибір груповальної ознаки;
- б) встановлення розміру інтервалу, визначення кількості груп і вибір груповальної ознаки;

- в) вибір результативної ознаки, визначення кількості груп і встановлення розміру інтервалу;
- г) встановлення загального числа досліджуваної сукупності, побудова розміру інтервалів та безпосереднє здійснення групування.

9. Альтернативна групувальна ознака може поділятися на таке число груп:

- а) дві;
- б) три;
- в) чотири;
- г) п'ять.

10. При групуванні за атрибутивною ознакою число груп складатиме:

- а) дві;
- б) три;
- в) відповідає кількості її різновидів;
- г) залежить від завдання статистичного групування.

11. Розмір інтервалу групування розраховується як:

- а) сума максимального та мінімального значення групувальної ознаки та поділити на число груп;
- б) різниця між максимальним і мінімальним значенням групувальної ознаки;
- в) відношення максимального значення досліджуваної ознаки до мінімального;
- г) різниця між максимальним і мінімальним значенням групувальної ознаки, поділеної на кількість груп.

12. Кроком інтервалу називається:

- а) величина, на яку збільшується інтервал групування;
- б) величина, на яку зменшується інтервал групування;
- в) величина, на яку збільшується або зменшується інтервал групування;
- г) розмір верхньої межі групи.

13. Багатомірне групування передбачає:

- а) утворення груп за двома ознаками одночасно;

- б) утворення груп за певною множиною ознак одночасно;
- в) утворення однієї групи;
- г) утворення груп за допомогою альтернативних ознак.

Типові задачі

Задача 1. За наведеними даними (табл.) виконайте:

а) типологічне групування підприємств за собівартістю одиниці продукції, утворивши три групи – з низькою собівартістю (до 70 грн.), з середньою (70-90 грн.) та високою (90 грн. і вище). По кожній групі визначіть кількість підприємств та вартість виробленої продукції;

б) структурне групування підприємств за середньою списковою чисельністю працюючих, утворивши чотири групи з рівними інтервалами;

в) аналітичне групування з метою вивчення залежності між вартістю основних виробничих засобів і виробництвом продукції, утворивши три групи з рівними інтервалами за факторною ознакою. У кожній групі обчисліть кількість підприємств, вартість основних виробничих засобів і виробництво продукції загалом та в середньому на одне підприємство. Результати розрахунків подайте у вигляді таблиці.

| № з/п | Середньорічна вартість основних виробничих засобів (млн. грн.) | Вироблено продукції за звітний місяць (млн. грн.) | Собівартість одиниці продукції (грн.) | Середня спискова чисельність працюючих (осіб) |
|-------|--|---|---------------------------------------|---|
| 1. | 26,1 | 33,3 | 71 | 360 |
| 2. | 21,6 | 29,8 | 75 | 275 |
| 3. | 22,0 | 21,6 | 72 | 313 |
| 4. | 35,0 | 44,4 | 67 | 475 |
| 5. | 13,4 | 16,5 | 86 | 200 |
| 6. | 22,9 | 27,9 | 66 | 286 |
| 7. | 16,7 | 19,6 | 76 | 291 |
| 8. | 26,8 | 22,0 | 65 | 418 |
| 9. | 32,1 | 36,7 | 64 | 375 |

| | | | | |
|-----|------|------|----|-----|
| 10. | 24,8 | 23,6 | 71 | 340 |
| 11. | 12,9 | 14,5 | 92 | 109 |
| 12. | 11,4 | 16,6 | 95 | 257 |
| 13. | 18,0 | 21,5 | 93 | 215 |
| 14. | 33,2 | 43,0 | 60 | 421 |
| 15. | 23,1 | 32,6 | 73 | 240 |
| 16. | 11,5 | 18,0 | 86 | 120 |
| 17. | 9,2 | 13,7 | 89 | 115 |
| 18. | 15,7 | 28,5 | 74 | 252 |
| 19. | 13,6 | 12,5 | 94 | 280 |
| 20. | 31,4 | 42,9 | 61 | 410 |

Задача 2. Для вивчення залежності між виробництвом продукції та її собівартістю за даними попередньої таблиці виконайте аналітичне групування підприємств за факторною ознакою, утворивши чотири групи з рівними інтервалами. У кожній групі визначіть: а) кількість підприємств; б) середнє виробництво продукції та середню собівартість. Результати розрахунків подайте у вигляді таблиці. Зробіть висновки.

Задача 3. З метою вивчення залежності між чисельністю працюючих і виробництвом продукції виконайте аналітичне групування за факторною ознакою, утворивши три групи з рівними інтервалами. У кожній групі обчисліть кількість підприємств, чисельність працюючих та виробництво продукції загалом та в середньому на одне підприємство. Результати розрахунків подайте у вигляді таблиці. Зробіть висновки.

Задача 4.

Маємо дані про кількість суб'єктів ЄДРПОУ по районах Тернопільської області:

(на початок 2015 року)

| Райони | Кількість суб'єктів | На 1000 жителів |
|--------------|---------------------|-----------------|
| Бережанський | 645 | 15 |
| Борщівський | 935 | 13 |
| Бучацький | 679 | 11 |
| Гусятинський | 806 | 12 |
| Заліщицький | 664 | 13 |
| Збаразький | 807 | 14 |
| Зборівський | 714 | 16 |

| | | |
|-----------------|------|----|
| Козівський | 542 | 13 |
| Кременецький | 788 | 11 |
| Лановецький | 464 | 15 |
| Монастириський | 405 | 12 |
| Підволочиський | 698 | 16 |
| Підгаєцький | 336 | 16 |
| Теребовлянський | 827 | 12 |
| Тернопільський | 1046 | 17 |
| Чортківський | 950 | 12 |
| Шумський | 492 | 14 |

За наведеними даними виконайте структурне групування районів: а) за кількістю суб'єктів ЄДРПОУ, утворивши чотири групи (до 500, 500-700, 700-900, 900 і більше); б) за кількістю суб'єктів ЄДРПОУ на 1000 жителів.

Задача 5. За даними про якість ґрунтів та урожайність озимої пшениці по тридцяти сільськогосподарських підприємствах регіону (табл.) виконайте:

а) типологічне групування підприємств за якістю ґрунтів, утворивши три групи – з низькою, середньою та високою якістю. По кожній групі обчисліть кількість підприємств та середню урожайність озимої пшениці. Результати розрахунків подайте у вигляді таблиці, зробіть висновки;

б) структурне групування сільськогосподарських підприємств за урожайністю озимої пшениці, утворивши п'ять груп з рівними інтервалами;

в) аналітичне групування для вивчення залежності урожайності озимої пшениці від якості ґрунтів, утворивши чотири групи з рівними інтервалами за другою ознакою. По кожній групі визначіть кількість сільськогосподарських підприємств, середню якість ґрунтів (у балах) та середню урожайність озимої пшениці (ц/га). Результати розрахунків подайте у вигляді таблиці, зробіть висновки.

| № з/п | Середня якість ґрунту | Урожайність озимої пшениці, ц/га | № з/п | Середня якість ґрунту | Урожайність озимої пшениці, ц/га |
|-------|-----------------------|----------------------------------|-------|-----------------------|----------------------------------|
|-------|-----------------------|----------------------------------|-------|-----------------------|----------------------------------|

| | (балів) | | | (балів) | |
|-----|---------|----|----|---------|----|
| 1. | 41 | 29 | 16 | 41 | 27 |
| 2. | 46 | 36 | 17 | 46 | 32 |
| 3. | 50 | 39 | 18 | 54 | 41 |
| 4. | 53 | 40 | 19 | 43 | 30 |
| 5. | 42 | 25 | 20 | 47 | 33 |
| 6. | 48 | 37 | 21 | 49 | 34 |
| 7. | 45 | 32 | 22 | 52 | 41 |
| 8. | 53 | 41 | 23 | 55 | 43 |
| 9. | 40 | 24 | 24 | 47 | 39 |
| 10. | 46 | 28 | 25 | 51 | 35 |
| 11. | 48 | 29 | 26 | 43 | 28 |
| 12. | 49 | 35 | 27 | 48 | 33 |
| 13. | 44 | 30 | 28 | 45 | 37 |
| 14. | 48 | 37 | 29 | 47 | 34 |
| 15. | 49 | 34 | 30 | 42 | 31 |

Задача 6. За результатами вибіркового обстеження умов життя домогосподарств у регіоні виконайте:

а) групування домогосподарств за кількістю дітей віком до 15 років. У кожній групі обчисліть кількість домогосподарств, середній грошовий місячний дохід та середньодушові грошові витрати в місяць у розрахунку на одне домогосподарство. Результати групування подайте у вигляді статистичної таблиці, зробіть висновки;

б) за даними табл. 1.3 виконайте структурне групування домогосподарств за розміром грошового місячного доходу, утворивши п'ять груп з рівними інтервалами;

в) за даними табл. 1.3 виконайте типологічне групування домогосподарств за середньодушовими грошовими витратами у місяць, утворивши три групи – з низьким, середнім та високим рівнем витрат. По кожній групі обчисліть кількість домогосподарств, середній грошовий місячний дохід та середні середньодушові грошові витрати у місяць.

| № домогосподарства | Кількість дітей до 15 років | Грошовий місячний дохід, грн | Середньодушові грошові витрати у місяць, грн | № домогосподарства | Кількість дітей до 15 років | Грошовий місячний дохід, гр.од. | Середньодушові грошові витрати у місяць, грн |
|--------------------|-----------------------------|------------------------------|--|--------------------|-----------------------------|---------------------------------|--|
| 1 | 3 | 1675 | 258 | 16 | 2 | 2950 | 639 |
| 2 | 2 | 2446 | 540 | 17 | 0 | 1512 | 507 |
| 3 | 2 | 2172 | 456 | 18 | 2 | 2467 | 466 |
| 4 | 1 | 2517 | 561 | 19 | 1 | 2326 | 572 |
| 5 | 3 | 1390 | 249 | 20 | 1 | 2004 | 538 |
| 6 | 2 | 1464 | 335 | 21 | 3 | 2403 | 454 |
| 7 | 0 | 1526 | 496 | 22 | 2 | 2037 | 433 |
| 8 | 3 | 1485 | 262 | 23 | 1 | 1704 | 505 |
| 9 | 2 | 1950 | 377 | 24 | 1 | 1488 | 436 |
| 10 | 1 | 1496 | 374 | 25 | 2 | 1454 | 350 |
| 11 | 2 | 2475 | 441 | 26 | 2 | 1956 | 389 |
| 12 | 1 | 1076 | 292 | 27 | 1 | 1616 | 472 |
| 13 | 3 | 1735 | 307 | 28 | 3 | 2446 | 437 |
| 14 | 4 | 1625 | 230 | 29 | 1 | 2715 | 646 |
| 15 | 0 | 1654 | 641 | 30 | 2 | 3340 | 750 |

Задача 7. Маємо дані про середньомісячну заробітну плату працівників у 2016 р. по регіонах України:

| | Середньомісячна заробітна плата, грн. | | Середньомісячна заробітна плата, грн |
|-------------------|---------------------------------------|---------------|--------------------------------------|
| Україна | 4362 | Миколаївська | 4120 |
| Вінницька | 3412 | Одеська | 4265 |
| Волинська | 3357 | Полтавська | 3850 |
| Дніпропетровська | 4445 | Рівненська | 3583 |
| Донецька | 5142 | Сумська | 3455 |
| Житомирська | 3297 | Тернопільська | 3008 |
| Закарпатська | 3419 | Харківська | 3797 |
| Запорізька | 4341 | Херсонська | 3249 |
| Івано-Франківська | 3567 | Хмельницька | 3394 |
| Київська | 4453 | Черкаська | 3484 |
| Кіровоградська | 3286 | Чернівецька | 3139 |
| Луганська | 3642 | Чернігівська | 3365 |
| Львівська | 3806 | м. Київ | 7126 |

За наведеними даними виконайте: а) структурне групування регіонів України за рівнем середньомісячної заробітної плати, утворивши чотири групи з рівними інтервалами; б) типологічне групування з утворенням трьох груп – з низькою, середньої та високою середньомісячною заробітною платою працівників; в) альтернативне групування з виділенням двох груп регіонів – із середньомісячною заробітною платою нижче та вище середньої по Україні. Зробіть висновки

Задача 8. Маємо результати групування сільськогосподарських підприємств регіону за розміром сільськогосподарських угідь (табл.). Виконайте вторинне групування методом збільшення інтервалів, утворивши чотири групи сільськогосподарських підприємств за площею сільгоспугідь: дуже малі (до 100 га); малі (100-500); середні (500-1000); великі (1000 і більше). По кожній групі обчисліть кількість сільськогосподарських підприємств, площу сільськогосподарських угідь по групі та в середньому на одне підприємство.

| | Кількість сільськогосподарських підприємств | Площа сільськогосподарських угідь, га |
|------------------------------------|--|--|
| Сільськогосподарські угіддя, га | 1220 | 340869 |
| до 10 | 312 | 2184 |
| 10-50 | 240 | 8160 |
| 50-100 | 156 | 12792 |
| 100-200 | 108 | 19008 |
| 200-500 | 91 | 28574 |
| 500-700 | 155 | 99355 |
| 700-1000 | 96 | 85440 |
| 1000-1200 | 32 | 36704 |
| 1200-1500 | 17 | 22984 |
| 1500-2000 | 8 | 13968 |
| понад 2000 | 5 | 11700 |

Задача 9. За наведеними вище вихідними даними (попередня табл.) виконайте вторинне групування сільськогосподарських підприємств таким чином, щоб у кожній групі було по 305 одиниць або по 25%. По кожній групі визначіть інтервал групування, площу сільськогосподарських угідь по групі загалом та в середньому на одне підприємство.

Задача 10. Маємо результати групування магазинів за середнім місячним роздрібним товарооборотом (табл.).

| Групи магазинів за середнім місячним роздрібним товарооборотом (тис. грн.) | Кількість магазинів | Місячний роздрібний товарооборот (тис. грн.) |
|--|---------------------|--|
| До 10 | 5 | 40 |
| 10 – 15 | 12 | 162 |
| 15 - 20 | 18 | 315 |
| 20 – 25 | 15 | 345 |
| 25 - 30 | 20 | 480 |
| 30 – 35 | 19 | 684 |
| 35 – 40 | 32 | 1254 |
| 40 -50 | 42 | 1932 |
| 50 – 70 | 38 | 2223 |
| 70 - 90 | 10 | 840 |
| 90 – 100 | 22 | 2208 |
| більше 100 | 7 | 840 |

За даними табл. виконайте вторинні групування магазинів, утворивши:

а) чотири групи за середнім місячним роздрібним товарооборотом – до 25 тис. грн., 25-50, 50-75, 75 і більше;

б) три інтервали таким чином, щоб у кожній групі було по 80 магазинів.

Задача 11. За даними, наведеними у табл. 1.6, виконайте вторинне групування робітників, сформувавши три групи таким чином, щоб у кожену групу увійшла однакова кількість робітників.

| Місячна заробітна плата робітників, | До 1000 | 1000 - 1200 | 1200 - 1500 | 1500 - 2000 | більше 2000 |
|-------------------------------------|---------|-------------|-------------|-------------|-------------|
| | | | | | |

| | | | | | |
|--|-----------|-----------|------------|-----------|-----------|
| грн. | | | | | |
| Кількість робітників (осіб) | 38 | 73 | 157 | 22 | 10 |

Питання для самоконтролю

1. Дайте визначення статистичного зведення.
2. Яка різниця між простим і складним зведенням?
3. Як слід розуміти централізоване і децентралізоване зведення?
4. Дайте визначення статистичного групування.
5. Дайте визначення типологічного групування.
6. Яка різниця між структурним і аналітичним групуванням?
7. Які є види групування?
8. Порядок визначення групувальної ознаки.
9. Порядок визначення інтервалу групування.
10. Порядок визначення числа груп.
11. Яка різниця між закритими та відкритими інтервалами?
12. Яка різниця між простим і складним групуванням?
13. Поясніть суть вторинного групування.

РОЗДІЛ 4. УЗАГАЛЬНЮЮЧІ СТАТИСТИЧНІ ПОКАЗНИКИ

4.1. Загальна характеристика статистичних показників

Оцінювання явищ і процесів суспільного життя, які є предметом статистичного дослідження забезпечується завдяки використанню адекватних статистичних показників, перелік і зміст яких повинен відображати характер й особливості розвитку економічних і соціальних процесів. Саме тому одним із головних завдань статистичної науки має стати розробка, запровадження й ідентифікація показників, які слугуватимуть для потреб об'єктивного оцінювання суті та основного змісту досліджуваних явищ і процесів.

Статистичний показник – це науково обґрунтований критерій, використання якого дає можливість одержати об'єктивну оцінку кількісної сторони предмета дослідження в поєднанні з якісною, співвідношення ознак суспільних явищ. Кількісна характеристика кожного показника обов'язково має поєднувати якісну сторону явища з обов'язковим урахуванням місця й часу його дії.

Статистичний показник, побудований із урахуванням якісного критерію складається з моделі кількісної сторони явища, що вивчається, або його числового значення. Модель дає можливість оцінювати структуру показника, встановити «що? де? коли? яким чином?» підлягає кількісній оцінці. Обов'язковим при цьому має бути встановлення одиниці виміру, методики розрахунку, способів одержання інформації тощо.

Статистичні показники поділяються на **об'ємні** (обсяг депозитних вкладень, розмір фонду заробітної плати) та **розрахункові** (середні та відносні величини). Крім цього, в аналітичних цілях їх можуть поділяти на **планові** та **звітні**. Перші, є конкретними завданнями з розвитку економіки в цілому по країні, в розрізі окремих економічних сфер, регіонів, підприємств і ін. Звітні - відображають фактично досягнуті результати.

Поряд із кількісною стороною кожен показник, як зазначено вище має якісну визначеність і відповідну методологію його розрахунку. Наприклад, вимірювання продуктивності праці потребує конкретизації питання щодо виробітку на одну людину-годину, людину-день, в середньому за рік тощо.

Числове значення статистичного показника, зазначається у відповідних одиницях виміру, називається його величиною.

Статистичні показники відрізняються від статистичних даних, оскільки останні відображають їх конкретні числові значення. Вони завжди подаються кількісною та якісною стороною, а також залежать від умов місця й часу.

Статистикою використовується також поділ показників на **первинні** і **вторинні**. **Первинні** (кількісні, об'ємні) охоплюють загальне число одиниць сукупності, або підсумок значень ознаки. **Вторинні** (якісні, розрахункові) передбачають використання середніх та відносних величин. Розмір складного соціально-економічного явища оцінюється за допомогою синтетичних (узагальнюючих) показників.

Залежно від одиниць виміру показники поділяються на натуральні, умовно натуральні, вартісні та трудові. За обсягом і змістом об'єкта статистичного вивчення розрізняють індивідуальні й узагальнюючі статистичні показники.

По відношенню до часу їх поділяють на **моментні** та **інтервальні**. Моментний показник показує розмір явища на певний момент часу. Прикладом може бути кількість дітей народжених у місті Києві в ніч з 31 на 1 число місяця станом на 24 годину, залишки товарних запасів на перше число відповідного відрізка часу тощо.

Інтервальні, характеризують обсяг ознаки за період часу, котрий прийнятий у розрахунок. Наприклад, розмір одержаних дивідендів по вкладках за термінами визначених у договорі, величина прибутку за видами діяльності за рік та інше. Особливістю первинних інтервальних показників є те, що вони в абсолютній своїй більшості є адитивними, тобто їх можна підсумовувати.

Зміст, структура та значення статистичних показників із часом перманентно змінюється та удосконалюється. Зокрема, інтеграція національної економіки до країн ЄС, посилення глобалізаційних процесів потребує використання принципово нових кількісних і якісних показників, які би реально відображали процеси соціально-економічного розвитку країни. Слід зазначити, що новітні показники, які генеруються вітчизняною статистикою не можуть бути надуманими або догматичними, їх розробка і впровадження має об'єктивно відображати процеси імплементації новацій у економічне і соціальне життя країни. Важливим є також дотримання встановленої єдиної статистичної методології розрахунку, імплементації та використання статистичних показників.

4.2. Система статистичних показників

Явища й процеси об'єктивного світу вирізняються надзвичайною різноманітністю та складністю. З розвитком суспільства виникає потреба одержання всебічної оцінки процесів та глибинного їх аналізу. Найбільш часто статистичною наукою досліджуються зміни, котрі стосуються динаміки сфер матеріального і нематеріального виробництва, обслуговування демографії населення, праці, соціального рівня життя суспільства, екології та інше. Кожна з сфер, яка вивчається статистикою має властиві для неї ознаки і показники.

Найбільш повну, синтезуючу характеристику в процесі статистичного дослідження одержують за допомогою системи статистичних показників. Коло властивостей, які вивчає статистична наука, а звідси й система показників, котрі використовуються при цьому, залежить від мети. Вона в свою чергу оцінюється певною їх множиною, використання яких дає можливість об'єктивно відтворити результати дослідження.

Система статистичних показників постійно оновлюється й тому має свої особливості. Йдеться про те, що з розвитком суспільства показники потребують модифікації та вдосконалення.

Зокрема, при статистичній оцінці обсягу виробництва затратні показники змінюються на такі, що характеризують організацію виробничого процесу з позицій "мінімум затрат - максимум прибутку". Появляються також нові показники, пов'язані з оцінюванням добробуту, обслуговуванням населення, екологією, інфляцією, фінансовими операціями тощо.

Система статистичних показників є не що інше як ієрархічна структура, на вершині якої розміщений узагальнюючий інтегральний показник (наприклад, обсяг валового внутрішнього продукту), на нижчому рівні розміщуються показники, що використовуються для оцінювання окремих локальних процесів соціально-економічного розвитку (добробут населення, екологія тощо).

Інтенсифікація інтеграційних процесів в Україні до моделі соціально-економічного розвитку країн ЄС потребує комбінування показників, розробки інтегральної та комплексної оцінки явищ і процесів, які властиві процесам суспільного розвитку.

4.3. Абсолютні статистичні показники

Узагальнення підсумків статистичного дослідження забезпечується завдяки розрахунку абсолютних показників, які дають змогу одержати узагальнені рівні та розміри суспільних явищ і процесів.

Величини, котрі характеризують розмір того чи іншого явища або процесу суспільного життя за допомогою іменованих (метри, тонни, штуки) вимірників, називаються абсолютними.

Абсолютні величини (показники) розраховують використовуючи додавання (віднімання) первинних статистичних даних або їх отримують за результатами вимірювання, перерахунку або обчислень досліджуваної сукупності.

Розрізняють **індивідуальні, групові і загальні** абсолютні величини. **Індивідуальні** – це показники, котрі використовують для оцінювання кількісний розміру ознак окремих одиниць сукупності.

Групові одержують у процесі поділу статистичної сукупності на групи, підгрупи використовуючи при цьому іменовані одиниці виміру.

Статистичні показники, котрі одержують шляхом додавання або кількості одиниць сукупності, або значень варіюючої ознаки називають загальними.

Абсолютні статистичні величини можуть розраховуватися в натуральних (умовно-натуральних), трудових і вартісних одиницях виміру.

Натуральні одиниці вимірювання використовуються для оцінювання фізичних властивостей окремих явищ (міри ваги, довжини, часу). В деяких сферах діяльності застосовуються умовно-натуральні одиниці виміру (паливно-енергетична, переробна сфери діяльності).

Трудові вимірники слугують для потреб вимірювання затрат праці (людино-година, людино-день) тощо.

Вартісні показники розраховуються для оцінки обсягів діяльності (виконаних робіт, наданих послуг), які безпосередньо не можна додавати. Наприклад, загальний обсяг виготовленого взуття фабрикою для різних статево вікових груп населення за рік.

4.4. Відносні показники

Відносними величинами називаються узагальнюючі показники, застосування яких дає можливість оцінювати кількісні співвідношення соціально-економічних явищ і процесів. Вони обчислюються як відношення двох чисел. Величину, котру порівнюють, записують у чисельнику. Вона може бути поточною або звітною величиною. В величину, з якою порівнюють, записують у знаменнику як основу порівняння.

Використання для потреб порівняння числа слугують для одержання відповіді на запитання: на скільки число чисельника, більше за знаменник. Це коефіцієнт. Наприклад, за даними Державного комітету статистики на 1 січня 2016 року кількість

наявного населення в Україні становила 42,7млн, а за результатами перепису в ніч з 10 на 11 січня 2010 року – 45,9млн. чоловік. Таким чином населення за досліджуваний період зменшилась і становило до рівня базового року – 3,2. Одержаний результат характеризує і кількісну і якісну сторону досліджуваної ознаки.

Відносні величини вимірюються також у математичних відсотках (%) - база порівняння прийнята за 100 %; промілях (‰) – база порівняння є число 1000; продецепромілях (‱) - база порівняння є число 10000; просантипромілях (‱‱) - база порівняння є число 100000.

Прикладами тут можуть бути відповідно: відсоток збільшення обсягу виробництва продукції; чисельність народжених дітей в розрахунку на 1000 чоловік населення; кількість вищих навчальних закладів у регіоні на 10000 учнів; зафіксовано хворих на СНІД на 100000 чол.

Об'єктом дослідження статистичної науки слугує надзвичайно широка різноманітність явищ і процесів суспільного життя, що потребує використання різних за змістом та аналітичними властивостями видів відносних величин.

У статистиці виділяють такі види відносних величин:

- 1) динаміки;
- 2) структури;
- 3) просторових порівнянь;
- 4) координації;
- 5) інтенсивності.

Відносна величина динаміки використовується з метою оцінювання зміни однойменних показників у часі. Її одержують у результаті співвідношення показників даного періоду з попереднім, початковим, або ще іншим, який прийнято за базу порівняння. В першому випадку одержують відносні величини із змінною базою порівняння (їх називають ланцюговими), у другому і третьому - відносні величини динаміки з постійною базою порівняння (їх називають базовими).

Для характеристики відносних величин динаміки використаємо офіційні дані Статистичного щорічника України за 2014 рік про рівень безробіття в Україні за причинами незайнятості (табл.4.1).

Таблиця 4.1

Динаміка безробітних в Україні

| Безробітні за методологією МОП за причинами незайнятості (у віці 15-70 років) | | | | | |
|---|---------|---------|---------|---------|---------|
| | 2010 | 2011 | 2012 | 2013 | 2014 |
| Всього безробітних, тис. | 1 713,9 | 1 661,9 | 1 589,8 | 1 510,4 | 1 847,6 |
| За причинами незайнятості, відсотків | | | | | |
| вивільнені з економічних причин | 39,3 | 34,6 | 32,3 | 30,3 | 31,6 |
| звільнені за власним бажанням | 29,6 | 32,6 | 36,5 | 38,4 | 36,4 |
| демобілізовані з військової строкової служби | 1,3 | 1,2 | 1,4 | 1 | 0,3 |
| не працевлаштовані після закінчення загальноосвітніх та вищих навчальних закладів I-IV рівнів акредитації | 15,7 | 18,3 | 17 | 17 | 17,3 |
| звільнені за станом здоров'я, через оформлення пенсії за віком, інвалідністю | 0,9 | 0,9 | 1,5 | 1,2 | 1,4 |
| звільнення у зв'язку з закінченням строку контракту | 10 | 9,2 | 8,4 | 8,1 | 7,8 |
| інші причини безробіття | 3,2 | 3,2 | 2,9 | 4 | 5,2 |

Спочатку розрахуємо базові показники динаміки безробітних, прийнявши за основу порівняння 2010рік.

$$\frac{1\,661,9}{1\,713,9} = 0,970 ; \quad \frac{1\,589,8}{1\,713,9} = 0,928 ; \quad \frac{1\,510,4}{1\,713,9} = 0,881 ; \quad \frac{1\,847,6}{1\,713,9} = 1,078 .$$

Наведені розрахунки засвідчують про сформовану позитивну динаміку зменшення безробітних в Україні до 2013 року і негативну – в 2014 році.

Розрахуємо відносні ланцюгові величини динаміки, для чого поділимо кожне наступне абсолютне значення кількості безробітних на попереднє.

$$\frac{1\,661,9}{1\,713,9} = 0,970 ; \quad \frac{1\,589,8}{1\,661,9} = 0,957 ; \quad \frac{1\,510,4}{1\,589,80} = 0,950 ; \quad \frac{1\,847,6}{1\,510,4} = 1,223 .$$

Таким чином, в Україні спостерігається щорічне зниження кількості безробітних по відношенню як до базового, так і до попередніх років.

Об'єктивність розрахунку відносних величин динаміки значною мірою залежить від дотримання вимог методології. У нашому випадку це стосується дотримання переліку причин звільнення, що регулюються Міжнародним статистичним органом праці і методологією МОП.

Відносні величини структури розраховуються як відношення величини окремих складових частин (часток) сукупності до загального підсумку (одиниці). Кожна частина називається часткою, або питомою вагою й вимірюється за допомогою дробових чисел або відсотків.

Зокрема, згідно даних табл. 4.1 в структурі безробітних у 2010 році найбільшу питому вагу по причині незайнятості зайняла група вивільнених з економічних причин. В той же час, у 2014 році питома вага звільнених за цією причиною зменшилась, оскільки пройшов перерозподіл на користь звільнених за власним бажанням.

Відносні величини просторових порівнянь визначаються як співвідношення однойменних показників, які відносяться до різних об'єктів або територій і мають однакову часову визначеність.

Прикладом просторового порівняння є інформація, наведена в таблиці 4.2.

Таблиця 4.2

Індекси споживчих цін за регіонами у 2015 році

(грудень до грудня попереднього року; відсотків)

| | Всі товари та послуги |
|----------------------|-----------------------|
| Україна ¹ | 143,3 |
| області | |
| Вінницька | 138,8 |
| Волинська | 143,3 |
| Дніпропетровська | 142,6 |

| | |
|-------------------|-------|
| Донецька | 146,9 |
| Житомирська | 143,7 |
| Закарпатська | 144 |
| Запорізька | 142,1 |
| Івано-Франківська | 143,4 |
| Київська | 144,3 |
| Кіровоградська | 141,3 |
| Луганська | 138,8 |
| Львівська | 145,2 |
| Миколаївська | 143,5 |
| Одеська | 144,6 |
| Полтавська | 145 |
| Рівненська | 144,9 |
| Сумська | 145,5 |
| Тернопільська | 145,1 |
| Харківська | 144,2 |
| Херсонська | 145,8 |
| Хмельницька | 142,5 |
| Черкаська | 143,8 |
| Чернівецька | 142 |
| Чернігівська | 145,6 |
| Київ | 138,6 |

¹ без урахування тимчасово окупованої території Автономної Республіки Крим, м. Севастополя та частини зони проведення антитерористичної операції;

Наведені дані у 2005 році засвідчують, що в усіх областях держави зросли споживчі ціни на товари та послуги.

Відносні величини координації використовують для аналізу співвідношення окремих частин сукупності з однією з них, яка прийнята за базу порівняння. Така відносна величина передбачає оцінювання співвідношень, пропорцій між складовими цілого. При цьому розраховують скільки одиниць одної частини сукупності припадає відповідно на 1, 100 і 1000 одиниць, які прийняті за базу порівняння.

Прикладом відносної величини координації може бути частка осіб чоловічої статі, що припадає на 100 або 1000 осіб жіночої статі; частка власних інвестицій по відношенню до всього обсягу залучених інвестицій тощо. Зазначимо, що охарактеризовані нами

відносні величини забезпечували оцінку того чи іншого показника за допомогою коефіцієнтів, відсотків, проміле та ін.

Відносна величина інтенсивності - відношення двох різнойменних абсолютних показників, які між собою пов'язані логічно, тобто слугують для оцінювання економічних процесів. Їх рівень характеризується за допомогою іменованих одиниць виміру.

Чисельник такого показника показує обсяги певного явища, в знаменнику - обсяг (величина) середовища, якому це явище за логічним змістом властиве. Відносні величини інтенсивності дають змогу охарактеризувати технічний рівень виробництва (енергоозброєність, фондоозброєність, фондозабезпеченість, обсяг вкладених інвестицій). Особливо широке застосування вони одержали в демографічній статистиці при підрахунку показників народжуваності, смертності, шлюбності, розлучень, пенсійного навантаження тощо. Зокрема, дані, поміщені в табл.4.3 засвідчують про негативну їх динаміку, що склалася в Україні за останні 15 років.

Таблиця 4.3

Коефіцієнти народжуваності, смертності та природного приросту населення

(на 1000 наявного населення)

| | 1990 | 1995 | 2000 | 2005 | 2010 | 2011 | 2012 | 2013 | 2014 |
|-------------------------------------|------|------|------|------|------|------|------|------|------|
| Кількість живонароджених | 12,7 | 9,5 | 7,8 | 9,0 | 10,8 | 11,0 | 11,4 | 11,1 | 10,3 |
| Кількість померлих | 12,1 | 15,3 | 15,3 | 16,5 | 15,2 | 14,5 | 14,5 | 14,5 | 13,9 |
| Природний приріст, скорочення (-) | 0,5 | -5,8 | -7,5 | -7,5 | -4,4 | -3,5 | -3,1 | -3,5 | -3,7 |
| Міграційний приріст, скорочення (-) | 1,5 | -2,5 | -2,7 | 0,1 | 0,4 | 0,4 | 1,4 | 0,7 | 0,5 |
| Кількість зареєстрованих шлюбів | 9,3 | 8,3 | 5,6 | 7,0 | 6,7 | 7,8 | 6,1 | 6,7 | 6,5 |
| Кількість зареєстрованих розлучень | 3,7 | 3,8 | 4,0 | 3,9 | 2,7 | 4,0 | 3,7 | 3,6 | 2,9 |

З наведених вище прикладів бачимо, що при проведенні статистичних досліджень у кожному конкретному випадку необхідно керуватись доцільністю та об'єктивністю застосування тих чи інших абсолютних і відносних величин. Найбільш ефективним є їх комплексне поєднання.

4.5 Загальна характеристика середніх величин

Часто в практиці статистичних досліджень виникає потреба одержати узагальнюючу характеристику явищ і процесів суспільного життя. Найбільш поширеними узагальнюючими статистичними показниками при цьому слугують середні величини.

Середня величина - узагальнюючий показник, за допомогою якого оцінюється типовий рівень варіюючої ознаки в розрахунку на одиницю однорідної сукупності в конкретних умовах місця і часу. Середня поєднує в собі як спільні для всієї сукупності типові риси, так і властиві лише для індивідуальних одиниць досліджуваного масиву. Середня величина характеризує загальне, типове, спільне, котре об'єднує всю масу елементів, тобто статистичну сукупність.

Властивості середніх найбільш повно проявляються в поєднанні з законом великих чисел, який діє в масових статистичних сукупностях. Саме завдяки використанню основних правил цього закону за допомогою середньої нівелюються (погашаються) нетипові індивідуальні значення ознаки одиниць досліджуваної сукупності, що досліджуються.

Критичні зауваження, щодо достовірності середніх величин, є результатом ігнорування окремих правил дії закону великих чисел. Не можна, наприклад, розраховувати середню величину з двох індивідуальних значень досліджуваної ознаки, розміри яких значно відрізняються між собою. Непорівнянним є розмір заробітної плати в сумі 600 і 6000 гривень, віком людини 20 і 60 років. Не можна включати в розрахунок середньої величини окремі числові значення ознаки, котрі в декілька раз перевищують або є надзвичайно малими величинами в порівнянні з абсолютною більшістю індивідуальних

значень статистичної сукупності. Нетипові значення одиниць усереднюваної сукупності, необхідно виключати при розрахунках середньої. Ігнорування цим правилом призведе до того, що розрахована середня не буде типовою для досліджуваної сукупності. Саме тому розрахунок середньої потребує обов'язковості дотримання таких принципів:

по-перше, середня величина може бути розрахована для достатньо великого числа одиниць сукупності;

по-друге, при визначенні середньої має бути врахований якісний зміст усереднюваної ознаки, а також взаємозв'язок між досліджуваними ознаками;

по-третє, розрахунок середньої проводиться з урахуванням якісної однорідності усереднюваної сукупності;

по-четверте, загальні середні повинні доповнюватися груповими;

по-п'яте, вибір одиниць сукупності при розрахунку середньої, повинен базуватися на попередньому логічному аналізі.

Дотримання принципів розрахунку узагальнюючих середніх статистичних показників є однією з найважливіших передумов забезпечення об'єктивності результатів дослідження. Зокрема, використання при дослідженні достатнього числа одиниць сукупності відповідатиме основним вимогам застосування закону великих чисел.

Урахування якісного змісту, а також взаємозв'язку ознак, які досліджуються забезпечує адекватність використання методу аналітичних групувань та об'єктивності розрахунку загальної та групових середніх.

Реалізація принципів розрахунку середньої за якісно однорідними явищами, в однорідній сукупності означає, що усереднюваний показник буде об'єктивно характеризувати сукупність лише завдяки однорідності ознак. Саме тому алогічним є розрахунок середнього рівня заробітної плати працівників нафтогазової та сільськогосподарської сфери діяльності (ознака однорідна – рівень заробітної плати).

Принципово важливим при розрахунку середньої величини є використання формули розрахунку, оскільки вона може значно вплинути на об'єктивність одержаного результату. Застосування формули розрахунку має базуватися на логічному аналізі заданої інформації. Чисельник формули, як правило, відображає обсяг значень досліджуваної ознаки, а знаменник – величину статистичної сукупності. Характер використаної формули залежить також від місця і часу покладеної в розрахунок усереднюваної ознаки. В статистиці середня позначається символом \bar{x} . Вона вимірюється в тих же одиницях виміру, що й ознака.

4.6. Узагальнюючі степеневі середні показники

Узагальнюючі середні величини в статистиці поділяють на два класи:

- 1) клас степеневих;
- 2) клас структурних.

До середніх величин класу степеневих належать: середня арифметична (проста і зважена), середня гармонічна (проста і зважена), середня геометрична, квадратична, кубічна та хронологічна. Підкреслимо, що різні види середніх величин, обчислені на основі однієї і тієї ж вихідної інформації не співпадають за підсумками розрахунку.

Найбільше поширення в практиці статистичних досліджень одержала середня арифметична.

Середня арифметична проста застосовується в тому випадку, коли значення варіюючої ознаки зустрічається один раз або однакове число разів. Значення ознаки називають варіантою, а число, котре показує скільки разів зустрічається варіанта – частотою, або «вагою».

Середня арифметична проста розраховується за формулою:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{x_1 + x_2 + \dots + x_n}{n}, \quad (4.1)$$

де \bar{x} – середня; x_i – індивідуальні значення ознаки; n – кількість одиниць сукупності;

\sum - знак суми (грецька буква «сігма»).

Наприклад, у червні поточного року обсяг споживання природного газу п'ятьма споживачами становив м³: 200, 220, 240, 180, 210. Середній обсяг спожитого газу становитиме:

$$\bar{x} = \frac{200 + 220 + 240 + 180 + 210}{5} = 210 \text{ м}^3.$$

Середня арифметична зважена розраховується за формулою:

$$\bar{x} = \frac{x_1 f_1 + x_2 f_2 + \dots + x_n f_n}{f_1 + f_2 + \dots + f_n} = \frac{\sum x_i f_i}{\sum f_i}, \quad (4.2)$$

де f - значення частот, ваг.

Її застосовують у випадках, коли варіанта (x_i) зустрічається неоднакове число разів. Вагами при цьому можуть бути частоти або частки (відносні величини структури).

Наприклад, за робочий день виробництво мобільних телефонів, на якому працюють 50 робітників характеризується такими даними.

| Виготовлено за день телефонів SAMSUNG, x | Число робітників, f |
|--|-----------------------|
| 14 | 17 |
| 15 | 19 |
| 16 | 20 |
| 18 | 12 |
| 19 | 10 |
| 20 | 8 |

Розрахуємо середній виробіток робітника за зміну:

$$\bar{x} = \frac{\sum xf}{\sum f} = \frac{14 \cdot 17 + 15 \cdot 19 + 16 \cdot 20 + 18 \cdot 12 + 19 \cdot 10 + 20 \cdot 8}{17 + 19 + 20 + 12 + 10 + 8} = \frac{1409}{86} \approx 16,3 \text{ штук.}$$

Таким чином, середній обсяг виробництва телефонів відділу за день становив 16,3 телефонів SAMSUNG. Одержаний результат не збігається з жодним значенням варіант, але він є типовим для

досліджуваної сукупності, оскільки розрахунок відповідає встановленим правилам.

На перший погляд, між середніми арифметичними простою і зваженою істотних відмінностей немає. Разом з тим, середня арифметична зважена є більш навантажена, оскільки враховує поширеність варіант і певною мірою склад сукупності.

Величина середньої величини значно залежить від співвідношення ваг (f). Наприклад, у випадку рівності фондів заробітної плати у двох відділах за абсолютною величиною, середня заробітна плата буде вища в тому відділі, де чисельність робітників є меншою. Зазначену особливість слід обов'язково враховувати при порівняльному аналізі явищ, склад яких істотно відрізняється між собою.

При розрахунку середньої зваженої в структурованій сукупності варіантами можуть бути як окремі значення ознаки, так і групові середні \bar{x} . Груповій середній відповідає відповідна вага групових частот. Загальну середню із групових середніх обчислюють за формулою:

$$\bar{x}_{\text{зв}} = \frac{\sum \bar{x}_i f_i}{\sum f_i}, \quad (4.3)$$

Середня арифметична може бути розрахована для дискретного і інтервального ряду. Розрахунок середньої для інтервального ряду передбачає, насамперед, знаходження для кожної групи середини інтервалу як півсуми його меж. Наведемо приклад розрахунку такої середньої.

Обсяг виготовлених телефонів SAMSUNG за робочий день робітниками відділу становив:

| Групи робітників за кількістю виготовлених телефонів, шт. | Кількість робітників |
|---|----------------------|
| 1-4 | 7 |

| | |
|-------------|----|
| 4-7 | 5 |
| 7-10 | 5 |
| 10-13 | 3 |
| 13-16 | 2 |
| 16 і більше | 1 |
| Усього: | 23 |

Визначити середню кількість телефонів, виготовлених робітниками за день. Для розв'язання задачі використовуємо розрахунки, наведені в таблиці (4.4.).

Таблиця 4.4.

Розрахунок середнього виробітку деталей робітниками за зміну

| Групи робітників за кількістю виготовлених телефонів, SAMSUNG шт. (x) | Кількість робітників, чол. (f) | Розрахункові дані | |
|---|------------------------------------|--|--------------|
| | | середня кількість виготовлених телефонів по групі, (x) | \bar{x}, f |
| 1-4 | 7 | $\frac{1+4}{2} = 2,5$ | 17,5 |
| 4-7 | 5 | $\frac{4+7}{2} = 5,5$ | 27,5 |
| 7-10 | 5 | $\frac{7+10}{2} = 8,5$ | 42,5 |
| 10-13 | 3 | $\frac{10+13}{2} = 11,5$ | 34,5 |
| 13-16 | 2 | $\frac{13+16}{2} = 14,5$ | 29 |
| 16 і більше | 1 | $\frac{16+3}{2} = 9,5$ | 9,5 |
| Усього: | 23 | - | 160,5 |

Отже, $\bar{x} = \frac{160,5}{23} = 7$, тобто середня кількість телефонів SAMSUNG, виготовлених одним робітником за зміну становить 7 штук.

З наведеного прикладу видно, що верхня межа шостої групи відкрита. Її закривають шляхом поширення розміру інтервалу, що є в п'ятій (попередній) групі, тобто $16-13=3$ деталі. Таким чином, верхня межа шостої групи становитиме $16+3=19$.

Розрахунок середньої можна значно спростити, використавши її математичні властивості, найважливішими з яких є:

1. Алгебраїчна сума відхилень варіант від середньої завжди дорівнює нулю:

$$\sum (x_i - \bar{x}) f_i = 0, \quad (4.4).$$

2. Добуток середньої на суму частот завжди дорівнює сумі добутоків варіант на частоти:

$$\bar{x} \sum f = \sum x f \quad (4.5).$$

3. Якщо до кожної варіанти (x_i) додати (відняти) будь-яке довільне число (A), то одержана середня збільшиться (зменшиться) на те ж число (A):

Зокрема, для першого випадку:

$$\frac{\sum (x_i + A) f_i}{\sum f_i} = \bar{x} + A, \text{ звідки } \bar{x} = \frac{\sum (x_i + A) f_i}{\sum f_i} - A; \quad (4.6).$$

для другого випадку:

$$\frac{\sum (x_i - A) f_i}{\sum f_i} = \bar{x} - A, \text{ звідки } \bar{x} = \frac{\sum (x_i - A) f_i}{\sum f_i} + A \quad (4.7).$$

4. Якщо кожному варіанту (x_i) поділити (помножити) на будь-яке довільне число (A), то середня арифметична зменшиться (збільшиться) в стільки ж разів.

Таке твердження можна записати так:

для першого випадку:

$$\frac{\sum \frac{x_i}{A} f_i}{\sum f_i} = \frac{\bar{x}}{A}, \text{ звідки } \bar{x} = \frac{\sum \frac{x_i}{A} f_i}{\sum f_i} \cdot A; \quad (4.8).$$

для другого випадку:

$$\frac{\sum (x_i \cdot A) f_i}{\sum f_i} = \bar{x} \cdot A, \text{ звідки } \bar{x} = \frac{\sum (x_i \cdot A) f_i}{\sum f_i} \cdot \frac{1}{A}. \quad (4.9).$$

5. Сума квадратів відхилень варіант ознаки від середньої арифметичної завжди менша, будь-якої іншої величини, тобто

$$\sum (x_i - \bar{x})^2 f = \min \quad (4.10).$$

6. Якщо всі частоти поділити або помножити на будь-яке число A , то середня арифметична від цього не зміниться. В даному

випадку, при розрахунку середньої беруть не абсолютні значення частот, а питому вагу, котру мають окремі варіанти в досліджуваній сукупності. Практичне використання властивостей покажемо на такому прикладі.

В процесі обстеження ринкових умов в регіонах було зафіксовано такий рівень цін на цемент (табл. 4.5).

Таблиця 4.5

Економічні показники від реалізації цементу по регіонах

| Регіон | Ціна однієї тонни цементу, у.о. (x_i) | Обсяг продажу цементу, тонн (f_i) | Обсяг виручки, грн. ($x_i f_i$) |
|--------|---|---------------------------------------|-----------------------------------|
| 1 | 1200 | 750 | 900000 |
| 2 | 1350 | 800 | 1080000 |
| 3 | 1400 | 840 | 1176000 |
| 4 | 1100 | 950 | 1045000 |
| 5 | 1250 | 750 | 937500 |
| 6 | 1300 | 640 | 832000 |
| Разом | 1266,7 | 4730 | 35948000 |

Узагальнюючий показник оптової середньої ціни однієї тонни цукру становить:

$$\bar{x} = \frac{\sum x_i f_i}{\sum f_i} = \frac{35948000}{4730} = 1266,7 \text{ у.о.}$$

$$\bar{x} \sum f = \sum x f = 1266,7 \cdot 4730 = 35948000.$$

Проведені розрахунки підтверджують правильність другої властивості середньої арифметичної.

Для характеристики першої властивості побудуємо розрахункову таблицю (табл. 4.6.):

Таблиця 4.6.

Розрахунок відхилень варіант від середньої арифметичної

| Регіон | Ціна однієї | Обсяг | Розрахункові величини |
|--------|-------------|-------|-----------------------|
|--------|-------------|-------|-----------------------|

| | тонни цементу, у.о. (x_i) | продажу цементу, тонн (f_i) | $x_i - \bar{x}$ | $(x_i - \bar{x})f_i$ |
|-------|-------------------------------------|---------------------------------------|-----------------|----------------------|
| 1 | 1200 | 750 | -66,7 | -50025 |
| 2 | 1350 | 800 | 83,3 | 66640 |
| 3 | 1400 | 840 | 133,3 | 111972 |
| 4 | 1100 | 950 | -166,7 | -158365 |
| 5 | 1250 | 750 | -16,7 | -12525 |
| 6 | 1300 | 640 | 33,3 | 21312 |
| Разом | 1266,7 | 4730 | - | -20991 |

За даними таблиці 4.6 знаходимо, що $\sum (x_i - \bar{x})f_i = 1266,7$.

Практичну реалізацію інших (3, 4, 5, 6) математичних властивостей середньої покажемо знову ж таки за допомогою розрахунків, наведених в допоміжних таблицях (4.7, 4.8, 4.9, 4.10, 4.11).

Таблиця 4.7

Розрахунок середньої при $x_i + A$, якщо ($A=1100$)

| Регіон | $x_i + A$ | Обсяг продажу, тонн (f_i) | $(x_i + A)f_i$ |
|--------|-----------|----------------------------------|----------------|
| 1 | 2300 | 750 | 1725000 |
| 2 | 2450 | 800 | 1960000 |
| 3 | 2500 | 840 | 2100000 |
| 4 | 2200 | 950 | 2090000 |
| 5 | 2350 | 750 | 1762500 |
| 6 | 2400 | 640 | 1536000 |
| Разом | X | 4730 | 11173500 |

$$\frac{\sum (x_i + A)f_i}{\sum f_i} = \bar{x} + A, \text{ звідки}$$

$$\bar{x} = \frac{\sum (x_i + A)f_i}{\sum f_i} - A = \frac{11173500}{4730} - 1100 = 1262,26 \text{ т.};$$

Розрахунок середньої при $x_i - A$, якщо ($A=1200$)

| Регіон | $x_i - A$ | Обсяг продажу, тонн (f_i) | $(x_i - A)f_i$ |
|--------|-----------|----------------------------------|----------------|
| 1 | 0 | 750 | 0 |
| 2 | 150 | 800 | 120000 |
| 3 | 200 | 840 | 168000 |
| 4 | -100 | 950 | -95000 |
| 5 | 50 | 750 | 37500 |
| 6 | 100 | 640 | 64000 |
| Разом | - | 4730 | 294500 |

$$\bar{x} = \frac{\sum (x_i + A)f_i}{\sum f_i} + A = 62.26 + 2200 = 1262,26 \text{ óì .îä .}$$

Розрахунок середньої при діленні варіант на 10 ($A=10$)

| Регіон | $\frac{x_i}{A}$ | Обсяг продажу, тонн (f_i) | $\frac{x_i}{A} f_i$ |
|--------|-----------------|----------------------------------|---------------------|
| 1 | 120 | 750 | 90000 |
| 2 | 135 | 800 | 108000 |
| 3 | 140 | 840 | 117600 |
| 4 | 110 | 950 | 104500 |
| 5 | 125 | 750 | 93750 |
| 6 | 130 | 640 | 83200 |
| Разом | - | 4730 | 597050 |

$$\frac{\sum \frac{x_i}{A} f_i}{\sum f_i} = \frac{597050}{4730} = 126,22 \text{ ;}$$

$$\bar{x} = \frac{\sum \frac{x_i}{A} f_i}{\sum f_i} \cdot A = 126,22 \cdot 10 = 1262,26 \text{ óì .îä .}$$

Таблиця 4.10

Розрахунок середньої при множенні варіант на 10 (A=10)

| Регіон | $x_i \cdot A$ | Обсяг продажу, тонн (f_i) | $(x_i \cdot A) f_i$ |
|--------|---------------|----------------------------------|---------------------|
| 1 | 12000 | 750 | 9000000 |
| 2 | 13500 | 800 | 10800000 |
| 3 | 14000 | 840 | 11760000 |
| 4 | 11000 | 950 | 10450000 |
| 5 | 12500 | 750 | 9375000 |
| 6 | 13000 | 640 | 8320000 |
| Разом | - | 4730 | 59705000 |

$$\bar{x} = \frac{\sum (x_i \cdot A) f_i}{\sum f_i} = \bar{x} \cdot A = \frac{59705000}{4730} \cdot 10 = 1262,26 \text{ óì .îä.}$$

Таблиця 4.11

Розрахунок середньої шляхом визначення питомої ваги кожної частоти в загальній їх сукупності

| Регіон | Ціна однієї тонни цементу, ум.од. | Обсяг продажу, тонн (f_i) | $\frac{f_i}{4730} \cdot 100$ | $x_i \left(\frac{f_i}{4730} \cdot 100 \right)$ |
|--------|--------------------------------------|----------------------------------|------------------------------|---|
| 1 | 1200 | 750 | 10,86 | 23913,04 |
| 2 | 1350 | 800 | 14,49 | 34057,97 |
| 3 | 1400 | 840 | 17,39 | 41739,13 |
| 4 | 1100 | 950 | 21,73 | 45652,17 |
| 5 | 1250 | 750 | 18,11 | 40760,87 |
| 6 | 1300 | 640 | 17,39 | 40000 |
| Разом | 2261,23 | 4730 | 100 | 226123,2 |

$$\bar{x} = \frac{\sum x_i \left(\frac{f_i}{4730} \cdot 100 \right)}{100} = \frac{226123,2}{100} = 2261,23.$$

Таким чином, усі властивості середньої підтверджуються практичними розрахунками.

Використання математичних властивостей середньої дає можливість спростити її розрахунок, застосувавши "спосіб моментів" або "спосіб відрахунку від умовного нуля". Суть його полягає в тому, що від всіх варіантів інтервального ряду віднімають постійне число A (для розрахунку приймають середину центрального інтервалу або інтервалу, що має найбільшу частоту). Потім одержані різниці ділять на інше постійне число - ширину інтервалу (k). Таким чином, розраховують відповідний момент. Зокрема, момент першого порядку:

$$m_1 = \frac{\sum \left(\frac{x_i - A}{k} \right) f_i}{\sum f_i} = \frac{\sum x' f_i}{\sum f_i}; \quad (4.11).$$

момент другого порядку:

$$m_2 = \frac{\sum \left(\frac{x_i - A}{k} \right)^2 f_i}{\sum f_i} = \frac{\sum (x')^2 f_i}{\sum f_i} \text{ і т.д.} \quad (4.12).$$

Звідки середня дорівнює:

$$\bar{x} = A + km_1. \quad (4.13).$$

Наведемо приклад. Вік робітників ПАТ "Укртрансгаз" у 2015 році розподілявся таким чином (табл.4.12).

Таблиця 4.12

**Розподіл робітників за віком на ПАТ "Укртрансгаз" у
2015 році**

| Групи робітників за віком, років, (x_i) | Середнє значення інтервалів, \bar{x}_i | Кількість робітників, f_i | $x_i - A$ ($A=43$) | $x' = \frac{x_i - A}{k}$ ($k=10$) | f, % до підсумку | $x' f$ |
|---|--|-----------------------------|-------------------------|--|------------------|--------|
| 18-28 | 23 | 125 | -20 | -2 | 12,5 | -25 |

| | | | | | | |
|----------------|----|------|-----|----|------|-------|
| 28-38 | 33 | 397 | -10 | -1 | 39,7 | -39,7 |
| 38-48 | 43 | 248 | 0 | 0 | 24,8 | 0 |
| 48-58 | 53 | 203 | 10 | 1 | 20,3 | 20,3 |
| 58 і більше | 63 | 27 | 20 | 2 | 2,7 | 5,4 |
| Разом | | 1000 | | | 100 | -39 |

Визначити середній стаж роботи.

$$\text{Спочатку знаходимо } m_1 = \frac{\sum x' f_i}{\sum f_i} = -0.39.$$

Звідки середня $x = A + km_1 = 43 + (-0.39 \cdot 10) = 43 + (-3.9) = 39.1$ року.

Таким чином, на ПАТ "Укртрансгаз" у 2015 році вік переважної більшості робітників становив 39 років.

Крім середньої арифметичної в практиці статистичних досліджень використовують узагальнені показники, які розраховують за допомогою середньої гармонійної.

Середня гармонійна - це величина, обернена середній арифметичній з обернених значень ознаки. Її використовують у випадку відсутності частоти. Натомість відомий лише добуток варіант на частоту, тобто w де $w = xf$. У практиці статистичних досліджень використовують середню гармонійну просту та зважену. Середня гармонійна проста має такий вигляд:

$$\bar{x}_{\text{гарм}} = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_n}} = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}}. \quad (4.14).$$

Для прикладу, на придбання мобільних телефонів SAMSUNG на трьох ринках міста покупці витратили відповідно: $\frac{1}{3}$; $\frac{1}{4}$ і $\frac{1}{5}$ грошової суми, встановленої за еталон. Це означає, що за один еталон грошової одиниці покупці на ринках можуть придбати відповідно телефонів SAMSUNG:

$$\text{перший} - 1 : \frac{1}{3} = 3 \text{ шт.};$$

другий – $1 : \frac{1}{4} = 4$ шт.;

третій – $1 : \frac{1}{5} = 5$ шт.;

Середня вартість придбання одиниці товару А покупцем становитиме:

$$\bar{x} = \frac{n}{\sum \frac{1}{x_n}} = \frac{3}{3+4+5} = \frac{1}{4} \text{ г.о.}$$

Середня гармонійна зважена розраховується за такою формулою:

$$\bar{x} = \frac{\sum w}{\sum \frac{w}{x}}. \quad (4.15).$$

Наприклад, прибуток п'яти підприємств готельно-ресторанної справи за період дії ліцензії характеризується такими даними (табл. 4.13):

Таблиця 4.13

Прибутковість підприємств готельно-ресторанної справи

| № підприємства | Місячний прибуток, тис. грн., (x) | Загальний прибуток за період дії ліцензії, тис. грн. (w) |
|----------------|-----------------------------------|--|
| 1 | 12 | 450 |
| 2 | 15 | 375 |
| 3 | 19 | 720 |
| 4 | 23 | 547 |
| 5 | 26 | 952 |

Визначити середній місячний прибуток підприємств готельно-ресторанної справи.

В умові задачі відсутні дані щодо терміну дії ліцензії кожного підприємства (f), натомість відомо що: $w = xf$. Для розрахунку середнього прибутку використаємо формулу середньої гармонійної зваженої.

$$\bar{x}_{\text{аарі}} = \frac{\sum w}{\sum \frac{w}{x}} = \frac{450 + 375 + 720 + 547 + 952}{\frac{450}{12} + \frac{3375}{15} + \frac{720}{19} + \frac{547}{23} + \frac{952}{26}} = \frac{3044}{160,79} = 18,9 \text{ тис. грн/міс.}$$

Середній місячний прибуток не лише від мінливості значень ознаки(x), а й зміни структури статистичної сукупності (f). Зокрема, у випадку збільшення терміну дії ліцензії на підприємстві за №1 у два рази – загальний прибуток становитиме 900 тис. грн, а середній місячний прибуток зменшиться:

$$\bar{x}_{\text{аарі}} = \frac{\sum w}{\sum \frac{w}{x}} = \frac{900 + 375 + 720 + 547 + 952}{\frac{900}{12} + \frac{3375}{15} + \frac{720}{19} + \frac{547}{23} + \frac{952}{26}} = \frac{3494}{198,29} = 17,6 \text{ тис. грн/міс.}$$

Таким чином, за рахунок структурних зрушень терміну ліцензії (f) середній місячний прибуток зменшився на 1,3 тис. грн.

Середня геометрична використовується, за умови, що задається послідовність ланцюгових відносних величин за допомогою середньорічних коефіцієнтів. Для її розрахунку використовують такі формули:

$$\bar{x}_{\text{геом}} = \sqrt[n]{k_1 \cdot k_2 \cdot \dots \cdot k_n}; \quad (4.16).$$

$$\bar{x}_{\text{геом}} = \sqrt[n]{\prod k_i}; \quad (4.17).$$

$$\bar{x}_{\text{геом}} = \sqrt[f]{\prod k^f} \text{ - середня геометрична зважена,} \quad (4.18).$$

де k - коефіцієнт росту; Д - символ добутку.

Приклад. Частка дітей, народжених жінками у віці від 30 до 40 років за останні 7 років зростає у 1,35 рази, у тому числі за 2009– у 1,05, 2010 – у 1,72, 2011 – у 1,56, 2012 – у 1,37, 2013 – у 1,41, 2014 – у 1,23 та 2015 рік – у 1,19 рази. Середній темп зростання кількості народжених дітей жінками у віці від 30 до 40 років становить:

$$\bar{x}_{\text{аарі}} = \sqrt[7]{1,05 \cdot 1,72 \cdot 1,56 \cdot 1,37 \cdot 1,41 \cdot 1,23 \cdot 1,19} = \sqrt[7]{12,78} = 1,439;$$

Якщо часові інтервали між коефіцієнтами різні, для розрахунку використовують середню геометричну зважену.

Середня квадратична буває простою та зваженою. Використовують її за умови, що визначальна властивість вимірюється величиною другого порядку.

Наприклад, потрібно визначити середню сторону 5 квадратів. Прийmemo в розрахунок сторони п'ятьох квадратів, розміри яких становлять: $x_1=4$, $x_2=6$, $x_3=10$, $x_4=15$, $x_5=16$. Загальна площа квадратів формується з величин другого порядку та в нашому прикладі дорівнює:

$$x_1^2 + x_2^2 + \dots + x_n^2 = 4^2 + 6^2 + 10^2 + 15^2 + 16^2 = 670 \text{ м}^2.$$

Середня площа середнього квадрата становить:

$$\bar{x} = \frac{x_1^2 + x_2^2 + \dots + x_n^2}{n} = \frac{4^2 + 6^2 + 10^2 + 15^2 + 16^2}{5} = 134 \text{ м}^2.$$

Враховуючи зв'язок між стороною квадрата та його площею, одержимо середній розмір сторони квадрата.

$$\bar{x}_{\text{кв}} = \sqrt{\frac{x_1^2 + x_2^2 + \dots + x_n^2}{n}} = \sqrt{\frac{\sum x_i^2}{n}} = \sqrt{\frac{4^2 + 6^2 + 10^2 + 15^2 + 16^2}{5}} = 11,58 \text{ м}^2.$$

Для розв'язання даної задачі нами використана формула середньої квадратичної простої. В даному випадку $\bar{x}_{\text{кв}}$ замінює конкретні значення сторін квадратів, беручи до уваги загальну площу п'ятьох квадратів.

При розрахунку середньої квадратичної обов'язковою умовою є визначення форми зв'язку між визначальним показником (в нашому випадку загальна площа п'ятьох квадратів) і ознакою, котру усереднюють (значення сторін квадратів). Тільки тоді можна одержати дійсно достовірні дані. Перевіримо, чи дійсно воно так. Для розрахунку середнього розміру сторони квадрата використаємо арифметичну просту.

$$\bar{x} = \frac{\sum x}{n} = \frac{5 + 8 + 10 + 15 + 16}{5} = \frac{54}{5} = 11 \text{ метрів.}$$

Для визначення розміру площі п'ятьох квадратів піднесемо одержаний результат до квадрата і помножимо на п'ять:

$$11^2 \cdot 5 = 605 \text{ м}^2, \text{ а не } 670 \text{ м}^2$$

Якщо варіанти зустрічаються декілька разів, для розрахунку використовують середню квадратичну зважену:

$$\bar{x}_{кв} = \sqrt{\frac{x_1^2 f_1 + x_2^2 f_2 + \dots + x_n^2 f_n}{f_1 + f_2 + \dots + f_n}} = \sqrt{\frac{\sum x_i^2 f_i}{\sum f_i}}. \quad (4.19).$$

Середня кубічна розраховується аналогічно середній квадратичній. При цьому визначальною властивістю є загальний об'єм n кубів.

Середня кубічна може бути розрахована як проста або зважена. Для її розрахунку використовують формули:

$$\bar{x}_{куб} = \sqrt{\frac{x_1^3 + x_2^3 + \dots + x_n^3}{n}} = \sqrt[3]{\frac{\sum x^3}{n}} - \text{проста}; \quad (4.20).$$

$$\bar{x}_{куб} = \sqrt{\frac{x_1^3 f_1 + x_2^3 f_2 + \dots + x_n^3 f_n}{f_1 + f_2 + \dots + f_n}} = \sqrt[3]{\frac{\sum x_i^3 f_i}{\sum f_i}} - \text{зважена}. \quad (4.21).$$

Середні квадратична і кубічна в практиці статистичних досліджень мають обмежене застосування. Використовують їх переважно при розрахунку середніх сторін квадратів, середніх діаметрів труб, запасів деревини тощо. Більш детально застосування середньої квадратичної нами буде розглянуто при вивченні показників варіації.

Середня хронологічна використовується як узагальнююча характеристика для ознаки, котра змінюється в часі. Для моментного ряду динаміки вона розраховується за формулою:

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

Наприклад, згідно даних державної служби статистики в Україні середньомісячна заробітна плата на підприємствах промислової сфери у першому кварталі 2015 року становили (грн., у розрахунку на одного штатного працівника):

| на 1.01 | на 1.02 | на 1.03 | на 1.04 |
|---------|---------|---------|---------|
| 4034 | 4041 | 4236 | 4321 |

Визначити середню заробітну плату за перший квартал.

$$\bar{x} = \frac{\frac{1}{2}x_1 + x_2 + \dots + \frac{1}{2}x_n}{n-1} = \frac{\frac{4034}{2} + 4041 + 4236 + \frac{4321}{2}}{4-1} = \frac{12454,5}{3} = 4151,5 \text{ тис. грн.}$$

Середня прогресивна. В процесі планування або нормування часто вдаються до розрахунку середньої прогресивної. Вона розраховується як середня величина з значень ознаки, котрі за абсолютною величиною є кращими у порівнянні з загальною середньою. Пояснимо техніку розрахунку середньої прогресивної.

На початковому етапі будують ранговий ряд з усіх варіант усереднюваної ознаки. Потім знаходимо середнє значення варіант, яке ділить цей ряд на дві частини. Значення однієї з цих частин за абсолютною величиною варіант більші за середню, а другої - менші. В зв'язку з цим розглядають два випадки середньої прогресивної.

Перший випадок – значення варіант більші за розраховану загальну середню (наприклад, виробництво продукції за зміну, обсяг реалізації продукції, середньодобовий приріст тварин на відгодівлі тощо). Другий випадок – значення варіант, котрі розміщені в ранговому ряді, менші за середню й з точки зору підприємницької діяльності кращі (наприклад, собівартість одиниці продукції, матеріаломісткість, енергомісткість і ін.).

Після розрахунку загальної середньої з усіх варіант ознаки відбирають окремі значення, котрі є кращими (більші або менші) користуючись логікою доцільності їх застосування. На основі відібраних варіант розраховують нову середню, котру називають середньою прогресивною.

Приклад 1. Середній місячний приріст заробітної плати працівників одного з готельних господарств у 2015 році становив: 730; 625; 820; 540; 1256; 900; 947; 528; 1340; 1520; 850; 657; 1840; 2150; 1205.

Визначити середній приріст заробітної плати працівників дослідженого готельного господарства. Спочатку будемо ранговий ряд:

730; 625; 820; 540; 1256; 900; 947; 528; 1340; 1520; 850; 657; 1840; 2150; 1205. Визначимо загальну середню:

$$\bar{x} = \frac{730+625+820+540+1256+900+947+528+1340+1520+850+657+1840+2150+1205}{15} = \frac{15908}{15} = 1061 \text{ грн.}$$

Визначимо середній прогресивний приріст з тих варіант значень рангового ряду, які є більшими за 1000 грн., що відповідає рівню інфляції в даному регіоні:

$$\bar{x} = \frac{1256+1340+1520+1840+2150+1205}{6} = \frac{9311}{6} = 1552 \text{ грн.}$$

Отже, при плануванні приросту заробітної плати працівників дослідженого готельного господарства в наступному році слід орієнтуватись саме на цифру 1552 грн.

Приклад 2. На підприємстві витрати електроенергії станками однакової потужності за становлять (кв/год.):

15; 18; 14; 19; 20; 21; 16; 17; 18; 14; 13; 16; 19; 18; 20.

Визначити середню прогресивну величину витрат електроенергії на один станок.

Діємо аналогічно, як і в попередньому прикладі. Будуємо ранговий ряд витрат електроенергії за одну годину роботи 15 станками:

15; 18; 14; 15; 16; 16; 17; 18; 18; 18; 19; 19; 20; 20; 21.

Визначаємо загальну середню витрат електроенергії за одну годину роботи станка.

$$\bar{x} = \frac{13+14+14+15+16+16+17+18+18+18+19+19+20+20+21}{15} = \frac{258}{15} = 17.2 \text{ кВ}$$

Визначаємо середню прогресивну з варіант ознаки, які менші від загальної середньої:

$$\bar{x}_{\text{прогр}} = \frac{13+14+14+15+16+16+17}{7} = \frac{105}{7} = 15,0 \text{ кВ}$$

Таким чином, витрачання 15 кВ електроенергії за годину роботи станка є тією індикативною величиною, якою слід керуватися при плануванні операцій з витрат електроенергії на даному підприємстві.

Всі вищенаведені формули розрахунку степеневих середніх побудовані з урахуванням показника степеня (n). Залежно від того, в

якій степені вони знаходяться, розрізняють такі види степеневих середніх:

середня арифметична, $n=1$;
 середня гармонійна, $n=-1$;
 середня геометрична, $n=0$;
 середня квадратична, $n=2$;
 середня кубічна, $n=3$.

Якщо для розрахунку середньої використати одну і ту саму інформацію але різні види середніх, то одержимо неоднаковий результат. Це пояснюється правилом мажорантності середніх: збільшення показника степені веде до збільшення розміру середньої, тобто

$$\bar{x}_{\text{куб}} \geq \bar{x}_{\text{кв}} \geq \bar{x}_{\text{ариф}} \geq \bar{x}_{\text{геом}} \geq \bar{x}_{\text{гарм}} .$$

Тестові завдання

1. Узагальнюючі середні величини в статистиці поділяють на:

- а) два класи;
- б) три класи;
- в) клас степеневих і структурних;
- г) усі відповіді неправильні.

2. Середня арифметична проста застосовується у випадках коли:

- а) варіююча ознака зустрічається в ряді багато разів;
- б) варіююча ознака зустрічається в ряді один раз або однакове число разів;
- в) частота зустрічається багато разів;
- г) відсутні частоти.

3. Число, що показує скільки разів зустрічається варіанта називається:

- а) варіантою;
- б) степенем;
- в) частотою;
- г) ознакою.

4. Середню арифметичну зважену застосовують у випадках за умови:

- а) варіанта зустрічається один раз;
- б) частота зустрічається один раз;
- в) усереднена ознака задається за допомогою коефіцієнтів;
- г) варіанта зустрічається неоднакове число разів.

5. Розрахунок середньої для інтервального ряду передбачає знаходження:

- а) для кожної групи середини інтервалу як півсуми його меж;
- б) верхньої межі інтервалу;
- в) нижньої межі інтервалу;
- г) півсуми частот.

6. Середня гармонійна використовується у випадку:

- а) відсутності варіант;
- б) відсутності частоти;
- в) коли усереднюваний показник розраховують на основі даних моментного ряду;
- г) коли задається послідовність ланцюгових відносних величин за допомогою середньорічних коефіцієнтів.

7. Середня геометрична використовується у випадку:

- а) відсутності варіант;
- б) відсутності частот;
- в) коли задається послідовність ланцюгових відносних величин за допомогою середньорічних коефіцієнтів;
- г) інформація задається на певний момент часу.

8. Середня хронологічна використовується як узагальнююча характеристика ознаки, котра змінюється:

- а) в просторі;
- б) в середині сукупності;
- в) поза межами середньої;
- г) в часі.

9. Статистичний показник – це науково-обґрунтований критерій за допомогою якого:

- а) характеризується кількісна сторона предмета дослідження в поєднанні з якісною, співвідношення ознак суспільних явищ;
- б) визначається результатом кінцевої діяльності суб'єкта господарювання;
- в) оцінюється розвиток досліджуваної ознаки;
- г) оцінюється кількісна сторона явища, що досліджується статистикою.

10. Первинні статистичні показники характеризують:

- а) рівень розвитку явища;
- б) рівень участі суб'єкта підприємницької діяльності на первинному ринку;
- в) загальне число одиниць сукупності або підсумок значень ознаки;
- г) результат господарювання за рік.

11. Вторинні статистичні показники передбачають використання:

- а) кількісних результатів діяльності досліджуваних одиниць сукупності;
- б) середніх та відносних величин;
- в) трудових вимірників;
- г) вартісних вимірників.

12. Синтетичні показники характеризують:

- а) первинну діяльність суб'єктів дослідження;
- б) вторинну діяльність суб'єктів дослідження;
- в) діяльність суб'єктів діяльності на певний момент часу;
- г) узагальнюючу діяльність суб'єктів дослідження.

13. Залежно від одиниць виміру показники бувають:

- а) натуральні, умовно-натуральні, вартісні, трудові;
- б) кількісні;
- в) якісні;
- г) кількісні та якісні.

14. За обсягом і змістом об'єкта статистичного дослідження розрізняють показники:

- а) моментні;

- б) інтервальні;
- в) синтетичні;
- г) індивідуальні та узагальнюючі.

15. Величини, котрі характеризують розмір того чи іншого явища або процесу суспільного життя за допомогою іменованих вимірників називають:

- а) відносними;
- б) абсолютними;
- в) динамічними;
- г) вартісними.

16. Статистичні показники, котрі одержують шляхом додавання або кількості одиниць сукупності, або значень варіюючої ознаки називають:

- а) вартісними;
- б) трудовими;
- в) загальними;
- г) розрахунковими.

17. Відносними величинами називаються узагальнюючі показники, за допомогою яких характеризують:

- а) обсяг виробництва;
- б) підсумки діяльності;
- в) результати трудової діяльності;
- г) кількісні співвідношення соціально-економічних явищ.

18. Відносна величина динаміки визначається як:

- а) співвідношення показників даного періоду з попереднім або початковим;
- б) відношення розмірів складових частин (часток) сукупності до загального підсумку;
- в) співвідношення однойменних показників, які відносяться до різних об'єктів або територій і мають однакову часову визначеність;
- г) співвідношення окремих частин сукупності з однією, що прийнята за базу порівняння.

19. Відносна величина структури визначається як співвідношення:

- а) показників даного періоду з попереднім або початковим;
- б) розмірів складових частин (часток) сукупності до загального підсумку;
- в) різнойменних показників, які відносяться до різних об'єктів або територій і мають однакову часову визначеність;
- г) окремих частин сукупності з однією, що прийнята за базу порівняння.

20. Відносна величина просторових порівнянь визначається як співвідношення:

- а) окремих частин сукупності з однією, що прийнята за базу порівняння;
- б) двох різнойменних абсолютних показників, які між собою пов'язані логічно;
- в) однойменних показників, що відносяться до різних об'єктів або територій і мають однакову часову визначеність;
- г) окремих частин сукупності з однією, що прийнята за базу порівняння.

21. Відносні величин координації визначаються як відношення:

- а) двох різнойменних абсолютних показників, які між собою пов'язані логічно;
- б) однойменних показників, які відносяться до різних суб'єктів або територій;
- в) розмірів складових частин (часток) сукупності до загального підсумку;
- г) окремих частин сукупності з однією, що прийнята за базу порівняння .

22. Відносна величина інтенсивності розраховується як відношення:

- а) двох різнойменних абсолютних показників, які між собою пов'язані логічно;
- б) розмірів складових частин (часток) сукупності до загального підсумку;

- в) окремих частин сукупності з однією, що прийнята за базу порівняння;
- г) показників даного періоду з попереднім або початковим.

Типові задачі

Задача 1. Споживання окремих видів енергетичних матеріалів і продуктів перероблення нафти у регіоні характеризуються такими даними:

| | Спожито у 2015 р. | Коефіцієнт переведення в умовне паливо |
|-----------------------------------|----------------------|--|
| Всього (т умовного палива) | 1358235 | x |
| вугілля кам'яне, т | 46432 | 0,742 |
| торф неагломерований паливний, т | 15063 | 0,341 |
| газ природний, тис. куб. м | 982388 | 1,150 |
| дрова для опалення, щільн. куб. м | 47331 | 0,266 |
| бензин моторний, т | 22764 | 1,490 |
| масла мастильні, т | 2565 | 1,000 |

За наведеними даними здійсніть переведення обсягів споживання окремих видів енергетичних матеріалів і продуктів перероблення нафти в умовне паливо та визначіть їх частку або питому вагу в загальному обсязі споживання.

Задача 2. За наведеними даними визначіть обсяг експорту та імпорту чавуну в регіоні в умовно-натуральному вираженні, а також питому вагу кожного виду чавуну в розрізі імпорту та експорту. Зробіть висновки.

(тис. т)

| Вид чавуну | Експорт | Імпорт | Коефіцієнт переведення |
|-----------------|---------|--------|---------------------------|
| Передільний | 1200 | 410 | 1,00 |
| Ливарний | 700 | 200 | 1,15 |
| Хромонікелевий | 100 | 850 | 1,50 |
| Феромарганцевий | 60 | 570 | 2,50 |

| | | | |
|---------------|----|-----|------|
| Ферофосфорний | 85 | 430 | 4,00 |
|---------------|----|-----|------|

Задача 3. Маємо дані про кредитування банками області.

(залишки заборгованості по позиках на кінець року, тис. грн.)

| | 2011 | 2013 | 2015 |
|-----------------------------|--------|--------|---------|
| Комерційні кредити у валюті | 164187 | 454819 | 1875499 |
| національній | 107943 | 314957 | 971318 |
| іноземній | 56244 | 139862 | 904181 |

За наведеними даними обчисліть у кожному році відносні величини структури, координації та динаміки. Зробіть висновки.

Задача 4. Маємо дані про валову додану вартість за видами економічної діяльності у Тернопільській області (табл.).

(у фактичних цінах, млн. грн.)

| | 2011 | 2015 |
|---|------|------|
| Всього | 2347 | 4738 |
| Сільське господарство, мисливство, лісове господарство | 810 | 1158 |
| Добувна промисловість | 4 | 17 |
| Обробна промисловість | 302 | 651 |
| Виробництво і розподілення електроенергії, газу та води | 96 | 122 |
| Будівництво | 86 | 199 |
| Оптова і роздрібна торгівля; торгівля транспортними засобами; послуги з ремонту | 280 | 698 |
| Готелі та ресторани | 11 | 24 |
| Транспорт і зв'язок | 257 | 515 |
| Фінансова діяльність | 41 | 121 |
| Операції з нерухомістю, здавання під найом та послуги юридичним особам | 62 | 166 |
| Державне управління | 94 | 306 |
| Освіта | 187 | 447 |
| Охорона здоров'я та соціальна допомога | 94 | 230 |
| Колективні, громадські та особисті послуги | 23 | 84 |

За даними таблиці, використовуючи відповідні види відносних величин, охарактеризуйте структуру валової доданої вартості, її динаміку в розрізі видів економічної діяльності. Сформулюйте висновки.

Задача 5. За наведеними даними по регіону визначіть у кожному році: а) відносні величини структури населення за статтю;

б) відносні величини координації (на 1000 осіб відповідної статі); в) відносні величини інтенсивності народжуваності та смертності (на 1000 осіб); г) відносні величини динаміки народжуваності та смертності. Зробіть висновки.

| | 2000 | 2005 | 2010 | 2016 |
|-------------------------------------|-------|-------|-------|-------|
| Середньорічна чисельність, тис осіб | | | | |
| – чоловіків | 538,0 | 542,4 | 533,9 | 514,3 |
| – жінок | 628,6 | 629,7 | 617,5 | 594,6 |
| Кількість народжених, тис осіб | 16,7 | 14,1 | 10,6 | 11,0 |
| Кількість померлих, тис. осіб | 15,0 | 16,1 | 15,7 | 16,8 |

Задача 6. Маємо дані про валовий збір та посівні площі окремих сільськогосподарських культур у Тернопільській області в 2015 р.:

| | Валовий збір, тис. т | Посівні площі, тис. га |
|------------------------------------|-------------------------|---------------------------|
| Озимі зернові | 367,4 | 167,2 |
| у т.ч. пшениця | 351,9 | 157,2 |
| Ярі зернові | 537,9 | 266,7 |
| у т.ч. пшениця | 106,1 | 50,4 |
| ячмінь | 242,5 | 125,3 |
| Технічні культури | х | 112,5 |
| у т.ч. цукрові буряки | 2237,1 | 76,2 |
| Картопля і овоче-баштанні культури | х | 67,3 |
| у т.ч. картопля | 806,6 | 57,9 |
| Кормові культури | | 84,7 |
| у т.ч. кормові коренеплоди | 475,7 | 14,1 |
| кукурудза на силос і зелений корм | 241,8 | 11,6 |

За наведеними даними охарактеризуйте:

- структуру посівних площ;
- урожайність окремих сільськогосподарських культур.

Зробіть висновки.

Задача 7. Маємо дані про кількість студентів у вищих навчальних закладів III-IV рівнів акредитації в регіоні на початок навчального року:

| | 2005/06 | 2010/11 | 2015/16 |
|---------------------------|---------|---------|---------|
| Кількість студентів, осіб | 17496 | 43797 | 54844 |

| | | | |
|---------------------------------|-------|-------|-------|
| у тому числі за формою навчання | | | |
| денною | 10723 | 23613 | 26985 |
| вечірньою | 946 | 664 | 832 |
| заочною | 5827 | 19520 | 27027 |

Обчисліть відносні величини динаміки, структури і координації. Результати розрахунків подайте у вигляді таблиці. Зробіть висновки.

Задача 8. За наведеними даними визначіть середній грошовий місячний дохід домогосподарств (табл.):

| № з/п | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|-------------------------------|------|------|------|------|------|------|------|------|------|------|
| Грошовий місячний дохід, грн. | 1120 | 2150 | 3010 | 2220 | 1790 | 1980 | 1100 | 2560 | 2320 | 2760 |

Задача 9.

Маємо дані про розподіл домогосподарств регіону за рівнем середньодушових грошових витрат, відсотків (табл.):

(відсотків)

| | Всі Домогосподарства | у тому числі, які проживають у | |
|--|----------------------|--------------------------------|----------------------|
| | | міських поселеннях | сільській місцевості |
| Всі домогосподарства | 100,0 | 100,0 | 100,0 |
| із середньодушовими витратами у місяць, грн. | | | |
| до 180,0 | 17,4 | 15,0 | 19,2 |
| 180,0-240,0 | 10,6 | 9,7 | 11,4 |
| 240,0-300,0 | 13,4 | 13,9 | 12,9 |
| 300,0-360,0 | 15,7 | 11,1 | 19,4 |
| 360,0-420,0 | 10,3 | 11,5 | 9,4 |
| 420,0-480,0 | 7,5 | 5,0 | 9,4 |
| 480,0-540,0 | 5,8 | 7,1 | 4,9 |
| 540,0-600,0 | 4,3 | 6,5 | 2,5 |
| 600,0-660,0 | 2,0 | 3,6 | 0,7 |
| 660,0-720,0 | 2,1 | 2,2 | 2,1 |
| понад 720,0 | 10,9 | 14,4 | 8,1 |

За наведеними у табл. даними обчисліть середні місячні грошові витрати на одне домогосподарство по регіону загалом, а також у міських поселеннях і сільській місцевості: а) традиційним способом; б) способом «моментів». Порівняйте одержані середні значення та зробіть висновки.

Задача 10. Маємо дані про розподіл груп за кількістю студентів денної форми навчання по двох факультетах ВНЗ (табл.). Визначіть середню та модальну кількість студентів у групі по факультетах, порівняйте одержані результати та зробіть висновки.

| Кількість груп | Кількість студентів у групі | | | | | | | | | |
|---------------------------|-----------------------------|----|----|----|----|----|----|----|----|----|
| | 20 | 21 | 25 | 26 | 28 | 29 | 30 | 31 | 33 | 35 |
| Факультет обліку і аудиту | 2 | 4 | 3 | 6 | 7 | 3 | 2 | 4 | 3 | 1 |
| Факультет фінансів | 3 | 1 | 5 | 7 | 3 | 6 | 3 | 2 | 4 | 2 |

Задача 11. Маємо дані про реалізацію товарів протягом трьох місяців:

| Товар | Травень | | Червень | | Липень | |
|-------|----------------|----------------|------------|----------------|----------------|------------|
| | Вартість, грн. | Кількість, од. | Ціна, грн. | Кількість, од. | Вартість, грн. | Ціна, грн. |
| А | 37500 | 150 | 260 | 130 | 48000 | 240 |
| Б | 59400 | 220 | 280 | 240 | 67500 | 270 |
| В | 17600 | 80 | 210 | 100 | 27600 | 230 |

Для виявлення тенденції зміни цін за наведеними даними визначіть середню ціну товару протягом травня, червня та липня. Порівняйте одержані результати, зробіть висновки.

Питання для самоконтролю

1. Дайте визначення статистичного показника.
2. Яка різниця між первинними і вторинними статистичними показниками?
3. Які є види вторинних статистичних показників?

4. Яка різниця між моментними і інтервальними статистичними показниками?
5. Дайте характеристику системи статистичних показників.
6. Як слід розуміти "інтегральні" статистичні показники?
7. Дайте характеристику абсолютних статистичних показників.
8. Дайте визначення відносних величин.
9. Які є види відносних величин?
10. Для чого використовують узагальнюючі статистичні показники?
11. Привести перелік узагальнюючих статистичних показників.
12. Які найважливіші принципи застосування середніх величин?
13. Які середні відносять до класу степеневих?
14. Яка різниця між загальною середньою та середньою з групових?
15. Основні математичні властивості середньої.
16. В чому суть «способу моментів»?
17. Які умови застосування середньої гармонійної?
18. Приведіть характеристику умов застосування середньої геометричної.
19. Основні умови застосування середніх квадратичної та кубічної.

Змістовний модуль 1

1. Методологічні засади статистики

Розвиток статистики як науки.

Найважливіші напрямки зародження статистики.

Предмет та завдання статистики.

Статистична сукупність та її найважливіші характеристики.

Статистична закономірність.

Методи статистичних досліджень.

Історія розвитку статистики в Україні.

Статистичні органи, їх структура та основні завдання.

2. Статистичне спостереження.

Статистична інформація.

Найважливіші етапи статистичного спостереження.

Статистична звітність та спеціально організоване статистичне спостереження.

Загальна характеристика плану статистичного спостереження.

Завдання програмно-методологічного характеру статистичного спостереження.

Організаційна частина плану статистичного спостереження.

Статистична характеристика суцільного спостереження.

Несуцільне спостереження.

Помилки статистичного спостереження.

3. Зведення і групування статистичних даних.

3.1 Найважливіші складові статистичних зведень.

3.2. Централізоване та децентралізоване зведення.

3.3. Найважливіші види групувань.

3.4. Методична послідовність організації статистичного групування.

3.5. Групувальні ознаки.

3.6. Методичні основи визначення кількості груп та інтервалів групування.

3.7. Просте та складне групування.

3.8. Багатомірне групування.

3.9. Вторинне групування.

4. Узагальнюючі статистичні показники.

4.1. Загальна характеристика статистичних показників.

4.2. Склад, види та структура статистичних показників.

4.3. Система статистичних показників.

4.4. Абсолютні статистичні показники.

4.5. Відносні статистичні показники.

4.6. Роль та умови застосування середніх величин.

4.7. Середня арифметична.

4.8. Спрощений спосіб розрахунку середньої арифметичної.

4.9. Середня гармонійна, геометрична та хронологічна.

Завдання для самостійного вивчення

1. Історія розвитку статистики.
2. Напрямки (школи) розвитку статистики.
3. Статистика як суспільна наука.
4. Теоретичні та методологічні засади статистики.
5. Особливості типологічного групування в умовах інтеграції України до економіки країн ЄС.
6. Нові завдання та функції органів статистики в нинішніх умовах.
7. Роль і значення статистичної інформації.
8. Розширення можливостей та об'єктивна потреба поліпшення інформаційного забезпечення економічного відтворювального процесу.
9. Особливості використання статистичного групування: умови та принципи.
10. Практика та особливості використання нерівних інтервалів у процесі проведення статистичного групування.
11. Комбінаційне та багатомірне групування.
12. Об'єктивна потреба розширення та оновлення статистичних показників в умовах сьогодення та на перспективу.
13. Проблеми гармонізації та співставності показників у вітчизняній статистиці до вимог стандартів і нормативів ЄС.
14. Математичні властивості середньої.
15. Середня квадратична, кубічна, прогресивна.

РОЗДІЛ 5. АНАЛІЗ РЯДІВ РОЗПОДІЛУ

5.1. Найважливіші показники рядів розподілу

Склад і структура досліджуваної ознаки оцінюються статистикою за допомогою **ряду розподілу**, складовими якого є варіанти та частоти. Залежно від виду досліджуваної ознаки ряди поділяються на кількісні та атрибутивні.

Інколи виникає потреба статистичної оцінки явищ і процесів, мінливість яких залежить від співвідношення варіантів і частот. Виникає потреба одержати характеристику закономірності розподілу, котру статистика може описати за допомогою певних показників, найважливішими з яких є:

- 1) частотні показники;
- 2) показники центру розподілу;
- 3) показники варіації;
- 4) показники нерівномірності розподілу.

Зупинимось коротко на характеристиці приведених показників.

Частотні показники ряду розподілу досліджуються за допомогою абсолютних та відносних величин. Зокрема, абсолютна чисельність певної групи є сума частот $\sum f$. Відносним частотним показником ряду розподілу слугує частка d . Зазначимо, що загальна сума частот $\sum f$ дорівнює їх кількості – n , а загальний підсумок часток $d=1$ або 100%.

Інколи в процесі статистичного дослідження рядів розподілу використовують кумулятивні частоти S_f (частки S_d), які в загальному підсумку не повинні перевищувати значення варіант x . Кумулятивні частотні показники одержують шляхом послідовного додавання частот (часток). Наприклад, є такі дані про розподіл робітників цеху за віком (табл.5.1).

Часто в практиці статистичних досліджень варіаційні ознаки рядів розподілу мають нерівні інтервали. В такому випадку проводять розрахунок щільності частоти (часток) на одиницю інтервалу. Для цього величину частоти (частки) ділять на ширину відповідного інтервалу.

Розподіл робітників за віком

| Вік, роки | Кількість робітників, f | Накопичені частоти, S_f |
|-----------|-------------------------|---------------------------|
| 18-28 | 30 | 30 |
| 28-38 | 25 | 55 |
| 38-48 | 50 | 105 |
| 48-58 | 15 | 120 |
| 58-68 | 30 | 150 |
| Разом | 150 | - |

Показниками центру розподілу слугують також **середня, мода та медіана**. Середні величини (степеневі) охарактеризовані нами досить детально в розділі IV. Тому зупинимось на характеристиці різновиду середніх, які в статистиці називають структурними середніми.

Структурні середні застосовуються для вивчення внутрішньої побудови рядів розподілу значень ознаки, а також для оцінювання середньої величини, розрахунок якої за вихідною інформацією не є неможливим.

Структурними середніми слугують **мода і медіана**. **Мода** - це варіанта, котра найбільш часто повторюється в ряді розподілу. В дискретному варіаційному ряді модою є варіанта, яка має найбільшу частоту (табл. 5.2).

Таблиця 5.2

Розподіл підприємств за рівнем використання природного газу

| | | | | | | |
|---|----|----|----|----|----|-------------|
| Середній обсяг використаного газу, $\text{м}^3/\text{добу}$, (x) | 15 | 18 | 20 | 35 | 40 | 50 і більше |
| Кількість підприємств, (f) | 12 | 37 | 15 | 22 | 17 | 35 |

Виходячи з інформації, приведеної в табл. 5.2, модою є середній обсяг використаного газу, що становить $18 \text{ м}^3/\text{добу}$, оскільки їй відповідає найбільша частота – 37 підприємств.

Кількісні варіаційні ознаки бувають дискретні та інтервальні. В дискретному ряді варіанти приводяться за допомогою цілих чисел (табл.5.3).

Таблиця 5.3

Розподіл постійного населення України за віковими групами

| | 1990 | 1995 | 2000 | 2005 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015¹ |
|-------------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------------------|
| всього населення | 51556 | 51300 | 49115 | 47101 | 45783 | 45598 | 45453 | 45373 | 45246 | 42760 |
| 0-14 років | 11084 | 10529 | 8781 | 6990 | 6484 | 6496 | 6532 | 6621 | 6711 | 6449 |
| 0-15 років | 11814 | 11248 | 9572 | 7665 | 6983 | 6976 | 6993 | 7048 | 7120 | 6816 |
| 0-17 років | 13305 | 12668 | 11116 | 9129 | 8081 | 8003 | 7972 | 7990 | 8010 | 7615 |
| 16-59 років | 30291 | 30596 | 29353 | 29656 | 29329 | 29090 | 28842 | 28623 | 28373 | 26613 |
| 15-64 роки | 34298 | 33811 | 33515 | 32604 | 32130 | 32137 | 31993 | 31847 | 31606 | 29635 |
| 18 років і старше | 38252 | 38632 | 37999 | 37971 | 37702 | 37595 | 37482 | 37382 | 37236 | 35145 |
| 60 років і старше | 9451 | 9456 | 10190 | 9779 | 9471 | 9532 | 9618 | 9702 | 9753 | 9330 |
| 65 років і старше | 6175 | 6961 | 6819 | 7507 | 7169 | 6965 | 6929 | 6905 | 6929 | 6676 |

¹ Без урахування тимчасово окупованої території Автономної Республіки Крим і м. Севастополя.

В інтервальному варіаційному ряді значення ознаки записуються за допомогою інтервалів (табл.5.4).

Таблиця 5.4

Безробітні за методологією МОП за тривалістю незайнятості

(у віці 15-70 років)

| | Безробітне населення у віці 15-70 років, усього, тис. осіб | З них особи, які раніше мали роботу | | у тому числі за тривалістю незайнятості | | | | | | Середня тривалість незайнятості, місяців |
|------|--|-------------------------------------|-------------------------|---|--------------------|--------------------|--------------------|---------------------|---------------------|--|
| | | | | до 1 місяця | від 1 до 3 місяців | від 3 до 6 місяців | від 6 до 9 місяців | від 9 до 12 місяців | 12 місяців і більше | |
| | | всього, тис. осіб | у % до всіх безробітних | у відсотках до підсумку | | | | | | |
| 2010 | 1 713,90 | 1 359,2 | 79,3 | 7,2 | 18,2 | 21 | 11,4 | 7,5 | 34,7 | 13 |
| 2011 | 1 661,90 | 1 272,1 | 76,5 | 7,7 | 22 | 23,1 | 10,8 | 7,2 | 29,2 | 11 |
| 2012 | 1 589,80 | 1 233,3 | 77,6 | 9,8 | 21,2 | 22,3 | 10,8 | 7,2 | 28,7 | 11 |
| 2013 | 1 510,40 | 1 148,2 | 76 | 8,2 | 22,5 | 23,9 | 10,5 | 6,7 | 28,2 | 11 |
| 2014 | 1 847,60 | 1 411,1 | 76,4 | 7,7 | 25,4 | 26,4 | 13,1 | 7,3 | 20,1 | 9 |

При розрахунку величини модального значення в інтервальному ряді до уваги приймають розміри інтервалу. Вони повинні бути однаковими, оскільки від цього залежить показник повторюваності значень ознаки x . В інтервальному ряді з рівними інтервалами модальна величина визначається за допомогою наступної формули:

$$M_0 = x_0 + d \frac{f_2 - f_1}{(f_2 - f_1) + (f_2 - f_3)}, \quad (5.1).$$

де x_0 - нижня межа модального інтервалу;

d - розмір модального інтервалу;

f_2 - частота модального інтервалу;

f_1 і f_3 - частоти відповідно передмодального і післямодального інтервалів.

Використавши дані табл. 5.1 розрахуємо моду для інтервального варіаційного ряду заробітної плати робітників цеху. Вона становить:

$$M_0 = 38 + 10 \frac{50 - 25}{(50 - 25) + (50 - 15)} = 43 \text{ днів}.$$

Таким чином, в приведеному інтервальному варіаційному ряді найчастіше зустрічається вік 43 роки.

Медіана - це значення ознаки, котре займає середнє положення в зростаючому або спадному ряді розподілу. Якщо непарне число варіант записати у порядку зростання або спадання, то центральне його значення буде медіаною.

У випадку, коли число варіант парне, медіана розраховується як середня арифметична проста з двох центральних значень ознаки.

Приклад 1. Річний валовий прибуток одинадцяти підприємств становив, тис. грн. 75; 90; 70; 95; 63; 87; 90; 95; 58, 76, 49. Визначити медіану середньомісячного доходу в розрахунку на сім'ю. В даному ранговому ряді медіаною є значення 87 тис. грн., оскільки воно займає середнє положення.

Приклад 2. Ранговий ряд торгівельних марок за кількістю найменувань продукції складає: 20, 28, 36, 44, 52, 60 од.

$$M_e = \frac{36+44}{2} = 40 \text{ од.}$$

В інтервальному варіаційному ряді з рівними інтервалами медіану розраховують за допомогою формули:

$$M_0 = x_0 + d \frac{\frac{\sum f}{2} - S_{me-1}}{f_{me}}, \quad (5.2).$$

де x_0 - нижня межа медіанного інтервалу;

d - розмір медіанного інтервалу;

$\sum f$ - сума частот ряду;

S_{me-1} - сума нагромаджених частот перед медіанним інтервалом;

f_{me} - частота медіанного інтервалу.

Використавши дані, наведені в таблиці 5.1, розрахуємо медіану. Спочатку визначимо $\frac{\sum f}{2}$, яке в наведеному прикладі становить:

$$\frac{\sum f}{2} = \frac{150}{2} = 75, \quad M_e = 38 + 10 \frac{75-55}{50} = 42 \text{ од.}$$

Розрахунок медіани показує, що на підприємстві половина робітників молодші 42 річного віку, а половина - більше.

Показниками центру розподілу рядів можуть бути також **квартилі (Q)** та **децилі (D)**. Розрахунок першого здійснюється шляхом розподілу значень варіант упорядкованого ряду на чотири рівних частини. Визначення децилів передбачає поділ ряду розподілу на десять рівних частин. Квартилі та децилі розраховуються з використанням накопичених (кумулятивних) частот (часток). Для прикладу, відомі дані про розподіл підприємців за розміром місячних податкових виплат (табл. 5.5).

Таблиця 5.5

Розподіл підприємців за розміром місячних податкових виплат

| Місячні податкові виплати, грн, (x) | Кількість підприємців, чол., (f) | Кумулятивна частота, чол. S_f |
|---|--------------------------------------|---------------------------------|
| 250-270 | 20 | 20 |
| 270-290 | 30 | 50 |
| 290-310 | 58 | 108 |
| 310-330 | 90 | 198 |
| 330-350 | 52 | 250 |
| Разом | 250 | - |

Використавши дані табл. 5.5 розрахуємо квартилі та децилі. Розрахунок проведемо за допомогою наступних формул:

$$Q = x_0 + d \frac{1,2,3,4, \sum f_i - S_{Q-1}}{f_Q}; \quad (5.3).$$

$$D = x_D + d \frac{1,2...10 \sum f_i - S_{D-1}}{f_D}. \quad (5.4).$$

Побудуємо допоміжну таблицю 5.6.

Таблиця 5.6

Вихідні дані для розрахунку квартилів та децилів

| Місячні | Кількість | Кумулятивна | Середина | xf |
|---------|-----------|-------------|----------|------|
|---------|-----------|-------------|----------|------|

| податкові виплати, грн, (x) | підприємців, чол., (f) | частота, чол. S _f | інтервалу, (x) | |
|-----------------------------|------------------------|------------------------------|----------------|-------|
| 250-270 | 20 | 20 | 260 | 5200 |
| 270-290 | 30 | 50 | 280 | 8400 |
| 290-310 | 58 | 108 | 300 | 17400 |
| 310-330 | 90 | 198 | 320 | 28800 |
| 330-350 | 52 | 250 | 340 | 17680 |
| Разом | 250 | - | - | 77480 |

Розрахуємо квартиль для:

перший квартиль

$$Q_1 = x_{0_1} + d \frac{\frac{1}{4} \sum f_i - S_{Q_1-1}}{f_{Q_1}} = 260 + 20 \frac{\frac{1}{4} \cdot 250 - 20}{30} = 288,5 \text{ грн.};$$

третій квартиль

$$Q_3 = x_{0_3} + d \frac{\frac{3}{4} \sum f_i - S_{Q_3-1}}{f_{Q_3}} = 310 + 20 \frac{\frac{3}{4} \cdot 250 - 108}{90} = 327,7 \text{ грн.}.$$

Розрахунок децилів здійснюється за формулою:

$$D = x_D + d \frac{1,2...10 \sum f_j - S_{D-1}}{f_{D_1}}$$

Звідси, перший децил (D_1)

$$D_1 = x_{D_1} + d \frac{\frac{1}{10} \sum f_j - S_{D_1-1}}{f_{D_1}} = 250 + 20 \frac{\frac{1}{10} \cdot 250}{10} = 275 \text{ грн.},$$

дев'ятий децил

$$D_9 = x_{D_9} + d \frac{\frac{9}{10} \sum f_j - S_{D_9-1}}{f_{D_9}} = 310 + 20 \frac{\frac{9}{10} \cdot 250 - 108}{90} = 310 \text{ грн.}$$

Таким чином, у наведеному ряді розподілу підприємців за розміром місячних податкових виплат перший квартиль становить 288,5 грн., а третій – 327,7 грн., тобто у 15% розмір місячних

податкових виплат не перевищує 288,5 грн., а у 85% складає 327,7 грн.

Розрахунок децилів у наведеному прикладі дає можливість стверджувати про те, що серед 10% підприємців розмір місячних податкових виплат не перевищує 275 грн., а 10% підприємців виплачують близько 310 грн. податків в місяць.

5.2. Показники варіації

Значення будь-якої ознаки досліджуваного варіаційного ряду, як правило, характеризується мінливістю, варіацією. Взагалі варіація, мінливість значень ознаки є властивістю статистичної сукупності. Вона функціонує під дією багатьох взаємопов'язаних причин, серед яких виділяють основні та другорядні. Основні причини формують середню величину (центр розподілу), другорядні - варіацію, сукупна їх дія - форму розподілу. Наприклад, на продуктивність праці впливають: електро- та енергоозбросність, кваліфікація, стаж роботи, розмір оплати праці та інші фактори. Всі вони зумовлюють коливання значень ознаки навколо їх середньої величини. В одних сукупностях індивідуальні значення ознаки суттєво відхиляються від центру розподілу, в інших практично перебувають у безпосередній близькості від нього. Такі коливання зумовлюють потребу оцінити центр розподілу, а також величину й ступінь варіації. Визначення розміру варіації коливання значень ознаки навколо середньої здійснюють за допомогою таких показників: розмах варіації, середнє лінійне відхилення (просте і зважене): середній квадрат відхилення (дисперсія) простий і зважений, середнє квадратичне відхилення (просте і зважене), коефіцієнти варіації.

Найпростішим є розрахунок показника розмаху варіації R як різниці між максимальним x_{\max} і мінімальним x_{\min} значеннями ознаки:

$$R = x_{\max} - x_{\min} \quad (5.5).$$

Інколи максимальне і мінімальне значення досліджуваної ознаки ряду розподілу не відображають типовості закономірності. З цією метою проводять розрахунок квартального та децильного розмахів. Квартильний розмах охоплює 50% загального обсягу

досліджуваної сукупності, тобто $R_Q = Q_3 - Q_1$. Децильний становить 60%, тобто $D_8 - D_2 = 60\%$, або 80% ($D_9 - D_1$).

Середнє лінійне відхилення є середнім арифметичним значенням абсолютних відхилень ознаки від її середнього рівня:

$$\bar{d} = \frac{\sum |x - \bar{x}|}{n} \text{ - просте (для незгрупованих даних)}. \quad (5.6).$$

При повторюваності окремих значень x використовують середнє зважене лінійне відхилення:

$$\bar{d} = \frac{\sum |x - \bar{x}| f}{\sum f} \text{ - зважене (дані згруповані)}. \quad (5.7).$$

Дисперсія ознаки (σ^2) розраховується завдяки використанню квадратичної степеневі середньої:

$$\sigma^2 = \frac{\sum (x - \bar{x})^2}{n} \text{ - проста}; \quad (5.8).$$

$$\sigma^2 = \frac{\sum (x - \bar{x})^2 f}{\sum f} \text{ - зважена}. \quad (5.9).$$

Середнє квадратичне відхилення розраховують шляхом добування квадрата-тичного кореня з дисперсії, тобто $\sigma = \sqrt{\sigma^2}$. Крім цього, його можна розраховувати:

$$\sigma = \sqrt{\frac{\sum (x - \bar{x})^2}{n}} \text{ - просте}; \quad \sigma = \sqrt{\frac{\sum (x - \bar{x})^2 f}{\sum f}} \text{ зважене}. \quad (5.10).$$

При розрахунках середнього лінійного відхилення \bar{d} використовують модулі відхилень. Це пов'язано з тим, що алгебраїчна сума відхилень

$$\sum (x - \bar{x})^2 f = 0. \quad (5.11).$$

Середнє лінійне та квадратичне відхилення, а також розмах варіації іменовані величини. Перший і другий показники ідентичні за змістом, проте завдяки їх математичним відхиленням $\sigma > \bar{d}$.

Статистичні показники варіації між собою пов'язані. При достатньо великому обсязі сукупності і розподілі близькому до нормального коливання значень ознаки незначні, тобто

$$\sigma = 1,25 \cdot \bar{d}. \quad (5.12).$$

Зв'язок між розмахом варіації і середнім квадратичним відхиленням дорівнює:

$$R = 6\sigma = \bar{x} \pm 3\sigma. \quad (5.13).$$

При вивченні зв'язку між R і σ враховують обсяг сукупності, вносячи при цьому певну поправку на коефіцієнт k .

Р. Пірсон обчислив значення коефіцієнтів k , використання яких дає можливість визначити середнє квадратичне відхилення (σ) за розмахом варіації (R): $\sigma = kR$. Нижче наводимо значення коефіцієнтів k для різних обсягів сукупності n (табл. 5.7.):

Таблиця 5.7

Значення коефіцієнта k для різного обсягу сукупності

| | | | | | | | | | | |
|-----|------|------|------|------|------|------|------|-----|------|-------|
| n | 6 | 8 | 10 | 20 | 30 | 40 | 50 | 100 | 200 | |
| k | 0,39 | 0,35 | 0,32 | 0,27 | 0,24 | 0,23 | 0,22 | 0,2 | 0,18 | |

Коефіцієнт варіації - це відсоткове (%) відношення середнього квадрата-тичного, лінійного відхилення, або величини розмаху варіації, до середньої арифметичної. Його позначають буквою V і розраховують за допомогою таких формул:

$$\text{лінійний} - V_{\bar{d}} = \frac{\bar{d}}{\bar{x}} \cdot 100; \quad (5.14).$$

$$\text{квадратичний} - V_{\sigma} = \frac{\sigma}{\bar{x}} \cdot 100; \quad (5.15).$$

$$\text{осциляції} - V_R = \frac{R}{\bar{x}}. \quad (5.16).$$

Прийнято вважати, що сукупність є однорідною, а середня типовою, у випадку, коли квадратичний коефіцієнт варіації не перевищує 33 %.

Наведемо приклад розрахунку показників варіації. Підприємства по виготовленню автомобілів у минулому місяці розподілялись за обсягом виробництва автомобілів (одиниць) наступним чином:

| | | | | | |
|--|----|----|----|----|----|
| Обсяг виробництва автомобілів, од. (x_i) | 12 | 15 | 10 | 18 | 14 |
| Кількість підприємств, (f_i) | 3 | 5 | 4 | 2 | 6 |

Визначити середній обсяг виробництва автомобілів, а також розрахувати основні показники варіації. середній обсяг виробництва визначаємо за допомогою формули середньої арифметичної зваженої, тобто:

$$\bar{x} = \frac{\sum xf}{\sum f} = \frac{12 \cdot 3 + 15 \cdot 5 + 10 \cdot 4 + 18 \cdot 2 + 14 \cdot 6}{3 + 5 + 4 + 2 + 6} = \frac{271}{20} = 13,5.$$

Для розрахунку показників варіації використаємо табличний метод та алгоритм.

Таблиця 5.8

Розрахунок показників варіації

| Вихідні дані | | Розрахункові дані | | |
|--|----------------------------|-------------------|------------------|---------------------|
| Обсяг виробництва автомобілів, од. (x_i) | Кількість підприємств, f | $ x - \bar{x} $ | $ x - \bar{x} f$ | $(x - \bar{x})^2 f$ |
| 12 | 3 | 1,5 | 4,5 | 6,75 |
| 15 | 5 | 1,5 | 7,5 | 11,25 |
| 10 | 4 | 3,5 | 14 | 49 |
| 18 | 2 | 4,5 | 9 | 40,5 |
| 14 | 6 | 0,5 | 3 | 1,5 |
| Всього | 20 | X | 38 | 109 |

Середнє лінійне відхилення:

$$\bar{d} = \frac{\sum |x - \bar{x}|f}{\sum f} = \frac{38}{20} = 1,9;$$

дисперсія:

$$\sigma^2 = \frac{\sum (x - \bar{x})^2 f}{\sum f} = \frac{109}{20} = 5,45;$$

середнє квадратичне відхилення:

$$\sigma = \sqrt{\frac{\sum (x - \bar{x})^2 f}{\sum f}} = \sqrt{5,45} = 2,33;$$

коефіцієнт варіації (квадратичний):

$$V_{\sigma} = \frac{\sigma}{x} \cdot 100\% = \frac{2,33}{13,5} \cdot 100\% = 17,2\%;$$

коефіцієнт варіації (лінійний):

$$V_{\sigma} = \frac{\bar{d}}{\bar{\sigma}} \cdot 100\% = \frac{1,9}{13,5} \cdot 100\% = 14,02\%.$$

Вимірність дисперсії відрізняється від вимірності варіюючої ознаки. Вона визначається за допомогою коефіцієнтів або відсотків.

В практиці статистичних досліджень розрахунок дисперсії здійснюється також завдяки використанню наступних формул:

$$\sigma^2 = \overline{x^2} - (\bar{x})^2, \quad (5.17)$$

де $\overline{x^2}$ - середній квадрат варіюючої ознаки; $(\bar{x})^2$ - квадрат середньої величини. В свою чергу:

$$\overline{x^2} = \frac{\sum x^2}{n} \quad \text{і} \quad \bar{x}^2 = \left(\frac{\sum x}{n} \right)^2. \quad (5.18)$$

Звідки

$$\sigma^2 = \frac{\sum x^2}{n} - \left(\frac{\sum x}{n} \right)^2.$$

З введенням частот формула дисперсії має такий вигляд:

$$\sigma^2 = \frac{\sum x^2 f}{\sum f} - \left(\frac{\sum x f}{\sum f} \right)^2. \quad (5.19)$$

Дисперсія альтернативної ознаки є добутком часток:

$\sigma^2 = d_1 d_0 = d_1(1 - d_1)$, де $d_1 d_0$ - показники, котрі характеризують структуру сукупності.

Максимальне значення дисперсії становить 0,25 у випадку коли $d_1 = d_0 = 0,5$, тобто одиниці сукупності розділені порівну між двома її значеннями.

Інколи емпірична ознака набуває більше двох значень. У такому випадку оцінка варіації розраховується за формулою:

$\sigma^2 = d_1 d_2 \dots d_j \dots d_m$, де - $d_1 d_2 \dots d_j$ частка j -ої групи; m - кількість груп.

Продуктивність праці станків у двох цехах заводу за минулий місяць характеризувалась наступним чином (табл. 5.9):

Таблиця 5.9

Продуктивність праці станків у двох цехах заводу

| Рівень виконання плану | Число станків, % до загального числа | |
|---------------------------|--------------------------------------|---------|
| | цех № 1 | цех № 2 |
| перевиконано | 20 | 25 |
| виконано | 60 | 55 |
| частково не виконано | 15 | 12 |
| Всього | 100 | 100 |

Варіація успішності складає:

для цеху № 1 $\sigma_1^2 = 0,20 \cdot 0,60 \cdot 0,15 \cdot 0,05 = 0,0009$;

для цеху № 2 $\sigma_2^2 = 0,25 \cdot 0,55 \cdot 0,12 \cdot 0,08 = 0,00132$.

Продуктивність праці станків першого цеху є більш стабільною, оскільки у другому варіація навколо середньої оцінки є вищою.

Використовуючи математичні властивості дисперсії, котрі притаманні середній арифметичній, можна застосувати спрощений спосіб розрахунку цього показника. Мова йде про "спосіб моментів", або "спосіб розрахунку від умовного нуля". Розрахунок дисперсії та середнього квадратичного відхилення "способом моментів" проводиться за допомогою таких формул:

$$\sigma^2 = k^2(m_2 - m_1^2); \quad (5.20).$$

$$\sigma = k\sqrt{m_2 - m_1^2}, \quad (5.21).$$

де m_1 - момент першого порядку, m_2 - момент другого порядку, k - величина інтервалу.

Наведемо приклад розрахунку дисперсії та середнього квадратичного відхилення, використавши дані табл. 5.10

**Розподіл сільськогосподарських господарств регіону за
площею ріллі, км. кв.**

| Площа ріллі, км. кв., x | Кількість господарств в, од., f | Розрахункові величини | | | | | | |
|------------------------------|---|----------------------------|------|--------------------------|--|-------|----------|------------|
| | | середина інтервалу, x | xf | $x - A$ ($A = 225$) | $x' = \frac{x - A}{k}$, ($k = 50$) | $x'f$ | $(x')^2$ | $(x')^2 f$ |
| 50-100 | 5 | 75 | 375 | -60 | -3 | -15 | 9 | 45 |
| 100-150 | 7 | 125 | 875 | -40 | -2 | -14 | 4 | 28 |
| 150-200 | 10 | 175 | 1750 | -20 | -1 | -10 | 1 | 10 |
| 200-250 | 6 | 225 | 1350 | 0 | 0 | 0 | 0 | 0 |
| 250-300 | 3 | 275 | 825 | 20 | 1 | 3 | 1 | 3 |
| 300-350 | 2 | 325 | 650 | 40 | 2 | 4 | 4 | 8 |
| 350-400 | 1 | 375 | 375 | 60 | 3 | 3 | 9 | 9 |
| Всього | 100 | - | 6200 | - | - | -29 | - | 103 |

Знаходимо середню звичайним способом:

$$\bar{x} = \frac{\sum xf}{\sum f} = \frac{6200}{100} = 62,0 \text{ км. кв.}$$

Розрахуємо середню «способом моментів»:

$$m_1 = \frac{\sum x'f}{\sum f} = -\frac{29}{100} = -0,29.$$

$$\bar{x} = m_1 k + A = -(0,29) \cdot 50 + 225 = 210,5 \text{ км. кв.}$$

Таким чином, розрахунок середньої двома способами збігається. Обчислення дисперсії та середнього квадратичного відхилення потребує розрахунку моменту другого порядку.

Як відомо, $m_1 = \frac{\sum x'f}{\sum f}$.

Оскільки при розрахунку σ^2 та σm подається в квадраті, то відповідно m_1 можна записати як квадрат моменту першого порядку.

$$(\bar{x}')^2 = \left(\frac{\sum x'f}{\sum f} \right)^2. \quad (5.22).$$

Момент другого порядку можна записати:

$$m_2 = \frac{\sum x'^2 f}{\sum f} = \overline{(x')^2}. \quad (5.23)$$

За даними таблиці 5.10 одержимо:

$$\overline{(x')^2} = \frac{103}{100} = 1,03 ;$$

$$(\overline{x'})^2 = -\left(\frac{-29}{100}\right)^2 = 0,084.$$

Звідки

$$\sigma^2 = 50^2 \cdot (1,03 - 0,084) = 2364,75 .$$

$$\sigma = \sqrt{2364,75} = 48,62 \text{ км.кв.}$$

Тестові завдання

1. Кумулятивні частотні показники одержують шляхом:

- а) послідовного додавання варіант;
- б) послідовного віднімання варіант;
- в) послідовного віднімання частот;
- г) послідовного додавання частот.

2. Для розрахунку щільності частоти на одиницю інтервалу величину частоти:

- а) множать на ширину відповідного інтервалу;
- б) додають до розміру ширини відповідного інтервалу;
- в) ділять на ширину відповідного інтервалу;
- г) множать на розмір інтервалу.

3. Структурні середні застосовують для:

- а) оцінки структури явища;
- б) вивчення внутрішньої побудови рядів розподілу значень ознаки, а також для оцінки середньої величини, якщо за вихідною інформацією її розрахунок є неможливим;
- в) розрахунку середньої;
- г) розрахунку середнього значення атрибутивної ознаки.

4. В дискретному варіаційному ряді модою є:

- а) варіанта, що має найбільшу частоту;
- б) варіанта, що має найменшу частоту;
- в) результативна ознака;
- г) факторна ознака.

5. Медіана – це значення ознаки, яке:

- а) займає початкове положення в зростаючому або спадному ряді розподілу;
- б) займає середнє положення в зростаючому або спадному ряді розподілу;
- в) займає кінцеве положення в зростаючому або спадному ряді розподілу;
- г) займає середнє положення в ряді після розрахунку моди.

6. Розрахунок квартилів здійснюється шляхом розподілу значень варіант упорядкованого ряду на:

- а) частини, виходячи з завдань дослідження;
- б) дві частини;
- в) чотири частини;
- г) десять частин.

7. Визначення децилів передбачає поділ ряду розподілу на:

- а) дві частини;
- б) три частини;
- в) чотири частини;
- г) десять частин.

8. Варіація функціонує під дією багатьох взаємозв'язаних причин, серед яких виділяють:

- а) поні, неповні;
- б) цілі, часткові;
- в) основні, другорядні;
- г) звітні, планові.

9. Середнє лінійне відхилення визначається як:

- а) різниця між мінімальним і максимальним значеннями ознаки-причини;
- б) середнім арифметичним значенням абсолютних відхилень ознаки від її середнього рівня;
- в) середнім арифметичним значенням відхилень ознаки від її середнього рівня;
- г) корінь квадратний з дисперсії.

10. Дисперсія ознаки визначається:

- а) як корінь квадратний з степеневій середній;
- б) шляхом розрахунку середнього арифметичного значення абсолютних відхилень ознаки від її середнього рівня;
- в) на основі квадратичної степеневій середній;
- г) як різниця між мінімальним і максимальним значенням результативної ознаки.

11. Середнє квадратичне відхилення розраховують через:

- а) добування кореня з середній;
- б) добування квадратного кореня з середній;
- в) ділення середнього лінійного відхилення на середню варіанту;
- г) добування квадратного кореня з дисперсії.

12. Коефіцієнт варіації – це відсоткове відношення:

- а) середнього квадратичного, лінійного відхилення або величини розмаху варіації до середній арифметичній;
- б) середній арифметичній до середнього квадратичного відхилення;
- в) середній арифметичній до середнього лінійного відхилення;
- г) середнього квадратичного відхилення до середнього лінійного.

13. Дисперсія вимірюється за допомогою:

- а) іменованих величин;
- б) трудових вимірників;
- в) вартісних вимірників;
- г) коефіцієнтів або відсотків.

14. Дисперсія альтернативної ознаки є:

- а) добутком цілих чисел;
- б) добутком часток;
- в) різниця між цілим числом та його частинами;
- г) сумою двох частин цілого.

Типові задачі

Задача 1. Заданий інтервальний ряд розподілу працівників ВАТ за рівнем місячної заробітної плати у 2016 році

| Групи працівників за рівнем середньомісячної заробітної плати, грн | Кількість працівників (f) |
|--|---------------------------|
|--|---------------------------|

| | |
|---------------|----|
| 800-1000 | 10 |
| 1000-1200 | 27 |
| 1200-1400 | 36 |
| 1400-1600 | 41 |
| 1600-1800 | 20 |
| 1800-2000 | 13 |
| 2000 і більше | 7 |

Розрахувати характеристики центру розподілу – середню, моду і медіану.

Задача 2. В таблиці приводиться ряд розподілу робітників підприємства за відсотком виконання ними норм виробітку.

| | | | | | | |
|--|--------|---------|---------|---------|---------|-------|
| Групи робітників за відсотком виконання норм виробітку | до 100 | 100-110 | 110-120 | 120-130 | 130-140 | Разом |
| Кількість робітників, % до підсумку | 9 | 45 | 25 | 15 | 6 | 100 |

Визначити середній відсоток виконання норм виробітку, моду та медіану.

Задача 3. Використовуючи дані ряду розподілу за розміром місячного товарообігу розрахувати дисперсію, використовуючи при цьому формулу

$$\sigma^2 = \overline{x^2} - (\bar{x})^2.$$

| | | | | | | |
|-----------------------|------|------|-------|-------|-------|-------------|
| Товарообіг, млн. грн. | до 5 | 5-10 | 10-15 | 15-20 | 20-25 | 25 і більше |
| Число магазинів | 10 | 13 | 10 | 7 | 5 | 5 |

Задача 4. Визначити середнє квадратичне відхилення, якщо відомо, що середня величина ознаки – 260, а коефіцієнт варіації становить 30%.

Задача 5. Маємо дані про розподіл сімей за місячним доходом.

| | |
|---|--------------------|
| Місячний дохід на одного члена сім'ї, грн. | Число сімей |
| До 2000 | 2242 |
| 2000 – 4000 | 7160 |

| | |
|---------------|------|
| 4000 – 6000 | 5382 |
| 6000 – 10000 | 2178 |
| 10000 – 15000 | 838 |

Визначіть: 1. Щільність розподілу сімей за місячним доходом на одного члена сім'ї. 2. Відносні та нагромаджені частоти. 3. Скільки сімей мають дохід менше: а) 5000 грн. на одного члена сім'ї; б) 8500 грн. на одного члена сім'ї. 4. Менше якого рівня мають дохід 30 % сімей, 75% сімей. Перетворіть інтервальний ряд розподілу в дискретний. Зробіть висновки.

Задача 6. Розподіл безробітних у регіоні за тривалістю безробіття та статтю характеризується такими даними (%):

| | Жінки | Чоловіки |
|----------------------------------|--------------|-----------------|
| Всього безробітних. | 100 | 100 |
| з них за тривалістю незайнятості | | |
| до 1 місяця | 6,1 | 4,5 |
| 1-3 | 18,5 | 15,3 |
| 4-6 | 30,9 | 36,2 |
| 7-9 | 23,6 | 28,1 |
| 10-12 | 18,3 | 10,4 |
| більше 1 року | 2,6 | 5,5 |

Визначіть: а) перший, другий і третій квартилі; б) квартильний коефіцієнт диференціації. Зробіть висновки.

Питання для самоконтролю

1. Поняття варіації.
2. Охарактеризуйте показник розмаху варіації.
3. Середнє лінійне відхилення.
4. Середнє квадратичне відхилення та дисперсія.
5. Коефіцієнти варіації.

6. «Спосіб моментів» в процесі дослідження варіації ознаки ряду розподілу.
7. Яке значення моди в процесі проведення статистичних досліджень рядів розподілу?
8. Медіана як узагальнююча характеристика ряду розподілу.
9. Показник ряду розподілу
10. Дайте тлумачення поняття квантилів та децилів.

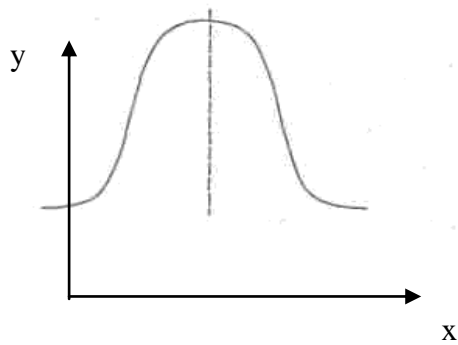
РОЗДІЛ 6. АНАЛІЗ КОНЦЕНТРАЦІЇ, ДИФЕРЕНЦІАЦІЇ ТА ПОДІБНОСТІ РОЗПОДІЛІВ

6.1. Характеристика форм розподілу

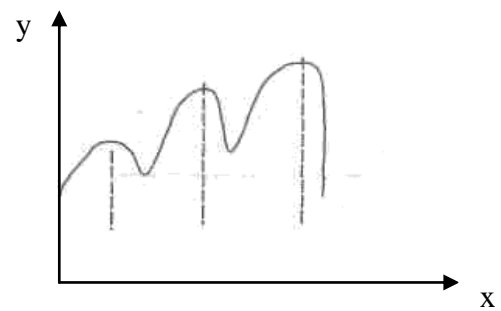
Варіація значень досліджуваної ознаки ряду розподілу, а також співвідношення варіант і частот (часток) в статистиці досліджуються за допомогою оцінювання форм розподілу.

Залежно від форми, розподіли одиниць статистичної сукупності поділяють на одно-, дво- і багатoverшинні. При наявності двох і більше вершин сукупність вважається неоднорідною. Одновершинні розподіли є якісно однорідними. Вони поділяються на **симетричні** й **асиметричні (скошені), гостро- і плосковершинні**. В **симетричному розподілі** значення ознаки рівновіддалені від центру та мають однакові частоти, в асиметричному - вершина зміщена вправо або вліво. Напрямок асиметрії протилежний напрямку зміщення вершини. Зміщення вершини вліво засвідчує про правосторонню асиметрію, і навпаки. Величина асиметрії залежить від результату обмеженої варіації в одному напрямку, що веде до зміщення центру розподілу. Найбільш наочною мірою асиметрії слугує величина відхилення між середньою арифметичною і модою або медіаною.

У симетричному розподілі показники центру розподілу рівні між собою, тобто $\bar{x} = M_0 = M_e$. В асиметричному між ними існують певні розбіжності. Правостороння асиметрія оцінюється нерівністю $-\bar{x} > M_0 > M_e$, лівостороння $-\bar{x} < M_0 < M_e$. Приведемо приклади розподілу за допомогою графіків (Рис.1,2,3,4,5,6).



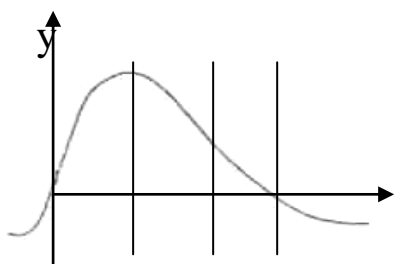
$$\bar{X} = M_0 = M_e$$



$$M_{01} \quad M_{02} \quad M_{03}$$

Рис. 6.1. Симетрична крива

Рис. 6.2. Багатoverшинний розподіл



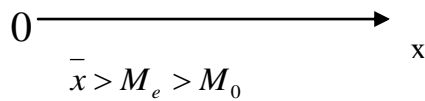


Рис. 6.3 Правостороння асиметрія

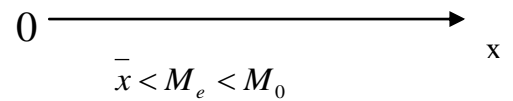


Рис. 6.4 Лівостороння асиметрія

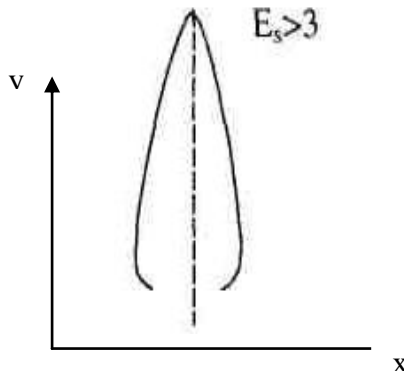


Рис. 6.5. Гостровершинний розподіл

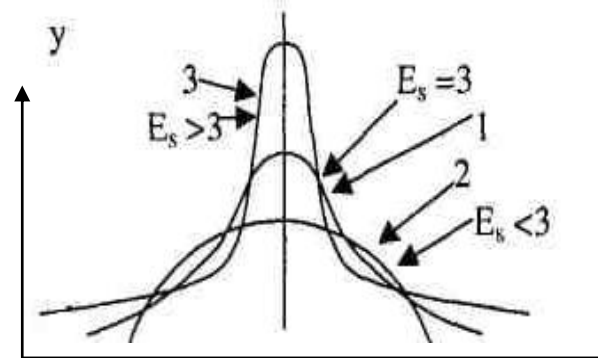


Рис. 6.6. Екссеси розподілу:

- 1 - нормальний;
- 2 - нижче нормального;
- 3 - вище нормального

У статистиці напрям і міру скошеності вивчають за допомогою стандартизованих відхилень A_s , тобто:

$$A_s = \frac{\bar{x} - M_0}{\sigma} \text{ або } A_s = \frac{\bar{x} - M_e}{\sigma}. \quad (6.1).$$

При симетричному розподілі $A_s = 0$ (рис. 6.1), при правосторонній асиметрії $A_s > 0$ (рис.6.3), при лівосторонній $A_s < 0$ (рис. 6.4). Симетричний і асиметричні розподіли показано на малюнках 6.1; 6.3; 6.4.

Гостровершинність розподілу відбиває скупченість значень ознаки навколо центру розподілу і називається **екссесом**.

Узагальнюючими характеристиками розподілу слугують моменти розподілу. В математичній статистиці під моментом k -го порядку (M_k) розуміють середню арифметичну k -ої степені відхилення окремих варіант від якої-небудь постійної величини A . За умови, що A - число, відмінне від нуля момент k -степені можна записати так:

$$M_k = \frac{\sum (x_i - A)^k f}{\sum f}. \quad (6.2).$$

Якщо $A=0$, моменти мають назву первинних або початкових і їх можна розрахувати за допомогою формули:

$$m_k = \frac{\sum (x_i)^k f}{\sum f}. \quad (6.3).$$

Використовуючи базову формулу (6.2) розраховують початкові моменти:

а) першого порядку

$$m_1 = \frac{\sum xf}{\sum f} = \bar{x} \text{ (середня арифметична);} \quad (6.4).$$

б) другого порядку

$$m_2 = \frac{\sum x^2 f}{\sum f} = \bar{x}^2; \quad (6.5).$$

в) третього порядку

$$m_3 = \frac{\sum x^3 f}{\sum f} = \bar{x}^3 \text{ і т.д.} \quad (6.6).$$

Практично при аналізі варіаційних рядів обмежуються розрахунком моментів перших чотирьох порядків. Момент називається центральним якщо $A = \bar{x}$. В загальному вигляді його записують так:

$$M_k = \frac{\sum (x_i - \bar{x})^k f}{\sum f}. \quad (6.7).$$

Використання формули 6.7. дає можливість визначити моменти перших чотирьох порядків:

$$\text{а). } M_1 = \frac{\sum (x_i - \bar{x}) f}{\sum f} = 0; \quad (6.8). \quad \text{б). } M_2 = \frac{\sum (x_i - \bar{x})^2 f}{\sum f} = \sigma^2; \quad (6.9).$$

$$\text{в). } M_3 = \frac{\sum (x_i - \bar{x})^3 f}{\sum f}; \quad (6.10). \quad \text{г). } M_4 = \frac{\sum (x_i - \bar{x})^4 f}{\sum f}. \quad (6.11).$$

| на одне домогосподарство за місяць, кВт. Г, x | x | дарств, f | | Г | кВт.Г | кВт.Г | кВт.Г | кВт.Г | кВт.Г | кВт.Г |
|---|----|-----------|------|------|-------|--------|--------|---------|--------|---------|
| до 5 | 4 | 6 | 24 | -8,6 | 73,9 | 443,8 | -636,1 | -3816,6 | 5470,1 | 32820,5 |
| 5-7 | 6 | 10 | 60 | -6,6 | 43,5 | 435,6 | -287,5 | -2875,0 | 1897,5 | 18975,0 |
| 7-9 | 8 | 12 | 96 | -4,6 | 21,1 | 253,9 | -97,3 | -1167,6 | 447,7 | 5372,9 |
| 9-11 | 10 | 8 | 80 | -2,6 | 6,7 | 54,1 | -17,6 | -140,8 | 45,7 | 365,6 |
| 11-13 | 12 | 14 | 168 | -0,6 | 0,3 | 5,0 | -0,2 | -2,8 | 0,1 | 1,8 |
| 13-15 | 14 | 18 | 252 | 1,4 | 1,9 | 35,3 | 2,7 | 48,6 | 3,8 | 69,1 |
| 15-17 | 16 | 12 | 192 | 3,4 | 11,5 | 138,7 | 39,3 | 471,6 | 133,6 | 1603,6 |
| 17-19 | 18 | 9 | 162 | 5,4 | 29,1 | 262,4 | 157,5 | 1417,5 | 850,3 | 7652,7 |
| 19-21 | 20 | 8 | 160 | 7,4 | 54,7 | 438,1 | 405,2 | 3241,6 | 2998,7 | 23989,3 |
| 21 і більше | 22 | 3 | 66 | 9,4 | 88,3 | 265,1 | 830,6 | 2491,8 | 7807,5 | 23422,5 |
| Всього | - | 100 | 1260 | - | - | 2332,0 | - | -331,7 | - | 11427,3 |

Насамперед розраховуємо середню величину витрат електричної енергії.

$$\bar{x} = \frac{1260}{100} = 12,6 \text{ кВт.}$$

Таким чином, середній розмір втрат електричної енергії в розрахунку на одне домогосподарство за місяць становив 12,6 кВт. Використовуючи дані табл.6.1 розрахуємо дисперсію, середнє квадратичне відхилення, асиметрію, ексцес та інші показники за допомогою яких оцінюються форми розподілу.

$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2 f}{\sum f} = \frac{2332,0}{100} = 23,32 ;$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{23,32} = 4,83 \text{ кВт.Г};$$

$$\sigma^3 = 2332,0 \cdot 4,83 = 11263,6 ;$$

$$\sigma^4 = 2332,0^2 = 543822,4 ;$$

$$M_3 = \frac{\sum (x_i - \bar{x})^3 f}{\sum f} = -\frac{331,7}{100} = -3,317 ;$$

$$M_4 = \frac{\sum (x_i - \bar{x})^4 f}{\sum f} = \frac{1427,3}{100} = 14,27 ;$$

$$A_s = \frac{M_3}{\sigma^3} = \frac{-3,317}{11263,6} = -0,0003 ;$$

$$E_s = \frac{M_4}{\sigma^4} - 3 = \frac{114,27}{5438224} - 3 = 0,00002 - 3 = -2,999.$$

В нашому прикладі ексцес значно нижчий від нормального, а сам розподіл – плосковершинний.

Інколи при статистичному дослідженні рядів розподілу виникає потреба оцінювання концентрації значень ознаки в окремих її частинах (наприклад, розмір заробітної плати між окремими працівниками сфери національного господарства України, обсяг інвестицій отриманих підприємствами, розмір площі земельного фонду землекористувачів). Міра концентрації оцінюється за допомогою **коефіцієнта концентрації**. Механізм його розрахунку забезпечується завдяки порівнянню часток розподілу елементів сукупності d_j та обсягу ознаки xd_j . Рівномірний розподіл забезпечується рівністю: $x_j = xd_j$. Значні відхилення між xd_j та d_j засвідчують про існуючу концентрацію елементів в ряді розподілу.

Міра концентрації розраховується як півсума модулів відхилень, тобто

$$k = \frac{1}{2} \sum |xd_j - d_j|, \quad (6.14).$$

де k - коефіцієнт концентрації. При рівномірній концентрації він дорівнює нулю. Якщо $k = 1$, то має місце повна концентрація.

Нерівномірність розподілу характеризується також за допомогою коефіцієнта локалізації L_j , який розраховується для кожної j -ї складової сукупності. Рівномірність розподілу окремих значень в ряді розподілу засвідчується рівністю: $L_j = 1$.

$L_j > 1$ - означає, що концентрація значень ознаки в j -ій складовій, і навпаки.

$$L_j = \frac{xd_j}{d_j} \cdot 100, \quad (6.15).$$

де xd_j - частка обсягу значень ознаки;

d_j - частка елементів сукупності.

Приведемо приклад розрахунку коефіцієнтів концентрації (k) та локалізації (L_j), використавши дані, наведені в табл. 6.2.

Таблиця 6.2

Дані для розрахунку коефіцієнтів концентрації та локалізації

| Обсяг реалізованих товарів за день продовольчими базами регіону | % до підсумку | | $L_j = \frac{xd_j}{d_j}$ | $ xd_j - d_j $ |
|---|-----------------------------------|--|--------------------------|----------------|
| | кількість продовольчих баз, d_j | обсяг реалізованих продовольчих xd_j | | |
| До 200 | 6 | 3 | 0,50 | 3 |
| 200-500 | 30 | 9 | 0,30 | 21 |
| 500-800 | 42 | 14 | 0,33 | 28 |
| 800-1100 | 15 | 18 | 1,20 | 3 |
| 1100-1400 | 5 | 27 | 5,40 | 22 |
| 1400-1700 | 2 | 29 | 14,50 | 27 |
| Разом | 100 | 100 | x | 104 |

Визначимо коефіцієнт концентрації за допомогою формули 6.14.

$$k = \frac{1}{2} \cdot 104 = 52\% \text{ або } 0,52.$$

Одержаний результат засвідчує про відносно високий рівень концентрації обсягу реалізованих продовольчих товарів на базах, які входять до п'ятої та шостої груп.

Об'єктивність такого висновку підтверджується також також L_j , розмір якого в останній групі становить 14,5.

6.2. Криві розподілу

Використання розрахунків коефіцієнтів, наведених у 6.1 дає можливість оцінити характер варіації та форму розподілу. Зазначене дослідження можна значно розширити, описавши закономірність співвідношення варіантів і частот певною теоретичною функцією. Серед значної кількості кривих розподілу найбільш поширеною є

| | | | | | | | | | |
|-------|----|----|------|------------|----|----------|-------|-------|----|
| 40-44 | 5 | 42 | 210 | 8820 | 44 | -1,38 | 0,084 | 0,084 | 7 |
| 44-48 | 26 | 46 | 1196 | 55016 | 48 | -0,33 | 0,372 | 0,288 | 23 |
| 48-52 | 32 | 50 | 1600 | 80000 | 52 | 0,72 | 0,764 | 0,393 | 31 |
| 52-56 | 13 | 54 | 702 | 37908 | 56 | 1,78 | 0,962 | 0,198 | 16 |
| 56-60 | 4 | 58 | 232 | 13456 | 60 | ∞ | 1,000 | 0,038 | 3 |
| Разом | 80 | X | 3940 | 19520 0 | X | X | X | X | 80 |

Насамперед розраховуємо середню врожайність пшениці.

$$\bar{x} = 3940 : 80 = 49,25 \text{ (ц/га)}.$$

Визначимо $\sigma = \sqrt{(195200 : 80) - 49,25^2} = 3,8 \text{ (ц/га)}$.

Функція $F_{(x)}$ базується на стандартизованих відхиленнях t :

$$t_1 = \frac{44 - 49,25}{3,8} = -1,38 ;$$

$$t_4 = \frac{56 - 49,25}{3,8} = -1,78 .$$

За даними $F_{(x)}$ для додатних $t, F(1,78) = 0,96$. При від'ємних значеннях $t, F(-1,38) = 1 - 0,916 = 0,084$. Теоретичні частоти (f') становлять для першої групи

$$f'_1 = 80 \cdot 0,084 = 7, \text{ для другої } f'_2 = (80 \cdot 0,372) - 0,084 = 23.$$

Для оцінки наближеності емпіричних (f) і теоретичних (f') частот можна використати один із критеріїв узгодженості Пірсона (χ^2 - «хі-квадрат»), і критерій Колмогорова (λ - «лямбда»).

Критерій Персона χ^2 - сума відхилень квадратів між емпіричними і теоретичними частотами:

$$\chi^2 = \sum \frac{(f - f')^2}{f'} . \quad (6.19).$$

Використавши дані таблиці 6.4 розрахуємо критерій Пірсона (χ^2)

Таблиця 6.4

Розрахунок критерію Пірсона

| Номер групи | Частоти | | Відхилення $f - f'$ | $(f - f')^2$ | $\frac{(f - f')^2}{f'}$ |
|-------------|---------|------|------------------------|--------------|-------------------------|
| | f | f' | | | |
| 1 | 5 | 7 | -2 | 4 | 0.57 |

| | | | | | |
|-------|----|----|----|---|------|
| 2 | 26 | 23 | 3 | 9 | 0.39 |
| 3 | 32 | 31 | 1 | 1 | 0.03 |
| 4 | 13 | 16 | -3 | 9 | 0.56 |
| 5 | 4 | 3 | 1 | 1 | 0.33 |
| Разом | 80 | 80 | 0 | X | 1.88 |

В нашому прикладі $\chi^2 = 1,88$ Порівняємо фактичне значення χ^2 з критичним для імовірності $1-\lambda$ і числа ступенів вільності $K = m - r - 1$, де m - число груп, r - число параметрів функції. Критичне значення $\chi^2_{1-\lambda(k)}$ - максимально можливе значення χ^2 за умови випадкового походження відхилення $f - f'$. Якщо фактичне значення перевищує критичне $\chi^2 > \chi^2_{1-\lambda(k)}$ то відхилення між емпіричними і теоретичними частотами є істотними, і навпаки $\chi^2 < \chi^2_{1-\lambda(k)}$.

Критичні значення χ^2 для ймовірності 0,90 становлять:

| Число ступенів вільності | $\chi^2_{0,90(k)}$ |
|--------------------------|--------------------|
| 2 | 4,6 |
| 3 | 6,3 |
| 4 | 7,8 |
| 5 | 9,2 |

У нашому прикладі $k = (5-2-1)$. Йому відповідає критичне значення $\chi^2_{0,90(k)} = 4,6$. Фактичне значення $\chi^2 = 1,88$ менше критичного. Отже, з ймовірністю 0,90 можна стверджувати, що розподіл фермерських господарств за урожайністю озимої пшениці відповідає теоретичному нормальному розподілу.

Перевіримо нашу гіпотезу за допомогою критерію Колмогорова (табл.6.5):

Таблиця 6.5

Розрахунок критерію Колмогорова (λ)

| Номер групи | Нагромаджені частоти | | Відхилення $ S_f - S'_f $ |
|-------------|----------------------|----------------------|------------------------------|
| | емпіричні S_f | теоретичні S'_f | |
| 1 | 5 | 7 | -2 |
| 2 | 31 | 30 | +1 |
| 3 | 63 | 61 | +2 |

| | | | |
|---|----|----|----|
| 4 | 76 | 77 | -1 |
| 5 | 80 | 80 | 0 |

Його розрахунок ґрунтується на максимальних відхиленнях D між нагромадженими частотами або частками емпіричного та теоретичного розподілів:

$$D_{\max} = |S_f - S'_f| \text{ для частот; } \quad (6.20).$$

$$\lambda = \frac{D}{\sqrt{N}}, \text{ де } N - \text{сума частот; } \quad (6.21).$$

Для часток

$$D_{\max} = |S_d - S'_d|. \quad (6.22).$$

В нашому прикладі $\lambda = 0,233$, тобто $\frac{2}{\sqrt{80}}$.

При випадковому походженні відхилення $f - f'$ ймовірність того, що λ не перевищить 0,223, близька до 1.

6.3. Дисперсійний аналіз

Статистичне дослідження варіації варіант ряду розподілу пов'язано насамперед з практичними потребами. В кожному конкретному випадку їх роль і значення визначаються вихідною інформацією, а також соціально-економічним значенням досліджуваного показника. Щодо першого, то тут дати однозначну відповідь неможливо. Зокрема, як бути у випадку, коли у досліджуваній сукупності значення окремих варіант істотно відрізняються від всієї сукупності і не є типовими. Візьмемо для прикладу ряд розподілу, котрий характеризує річний рівень доходів у розрахунку на одне домогосподарство. В такому ряді можуть перебувати домогосподарства з розміром річного доходу, що не перевищує 10000 гривень, і такі, розмір доходів яких перевищує мільйон і більше грошових одиниць.

В економічній літературі з цього приводу немає єдиної думки. Одні вчені пропонують виключати нетипові рівні досліджуваної ознаки, інші такі дії вважають незаконними по відношенню до фактичних даних.

На нашу думку, в кожному конкретному випадку необхідно керуватися соціально-економічним змістом досліджуваного явища. Дійсно, у випадку, коли розмір середньої заробітної плати працівників установи за місяць склав відповідно: 1200, 1700, 2200, 2700, 10000 умовних одиниць, то для статистичної оцінки обсягу місячного фонду заробітної плати в цілому не можна виключати працівника, котрий отримав заробітну плату 10000 ум.од. В той же час при визначенні типової для даної групи працівників середньої заробітної плати за місяць, заробітну плату працівника, котрий отримав 10000 ум.од. потрібно виключити з досліджуваного ряду розподілу.

Варіацію ознаки в кожному конкретному випадку можна оцінити позитивно або негативно, є бажаною або небажаною.

Для практичної і пізнавальної діяльності важливим є пізнання ролі та значення тих чи інших факторів у формуванні варіації ознаки. З цією метою використовують дисперсійний аналіз. В основі такого аналізу лежить закон розкладання загальної дисперсії σ^2 на систематичну δ^2 (дельта) і випадкову $\overline{\sigma_{a.a}^2}$, тобто

$$\sigma^2 = \delta^2 + \overline{\sigma_{a.a}^2} = 1, \quad (6.23).$$

де σ^2 - загальна дисперсія;

δ^2 - міжгрупова дисперсія (систематична, факторна);

$\overline{\sigma_{a.a}^2}$ - випадкова дисперсія.

Загальна варіація варіює під впливом систематичних, та випадкових факторів.

Систематична варіація (факторна) є результатом дії постійних, систематичних факторів.

Випадкова варіація - частина загальної варіації, викликана дією випадкових факторів.

Дисперсійний метод аналізу тісно пов'язаний із аналітичним групуванням і будується з урахуванням теоретичних і методологічних засад проведення аналітичного групування. Насамперед використовуючи механізм логічного економічного аналізу визначають факторні та результативні ознаки. Порівнюючи групові середні результативної ознаки з факторною, встановлюють

наявність та величину зв'язку, взаємозалежності між причинами та наслідками.

Поєднання дисперсійного та групувального методів вимірювання взаємозв'язку потребує розрахунку поряд із загальною дисперсією (6.23), міжгрупової (факторної) (6.24) та внутрішньогрупової (випадкової) (6.25). Міжгрупову варіацію називають також систематичною. Вона характеризує вплив на результативну ознаку систематичних факторів, а її розмір визначають за допомогою формули міжгрупової дисперсії:

$$\delta^2 = \frac{\sum (\bar{y}_i - \bar{y})^2 f}{\sum f}, \quad (6.24).$$

де δ^2 - міжгрупова дисперсія; \bar{y}_i - середнє значення груп; \bar{y} - загальна середня; f - частота.

Зазначимо, що міжгрупова дисперсія показує відмінності, варіацію

групових середніх \bar{y}_i навколо загальної середньої \bar{y} .

Випадкова варіація, що викликана дією всіх інших факторів, за винятком фактора, що покладений в основу групування, оцінюється за допомогою внутрішньогрупової дисперсії. Вона є середньою величиною з групових дисперсій, і розраховується за формулою:

$$\sigma_{e.z.}^2 = \frac{\sum \sigma_i^2 f_i}{\sum f_i}. \quad (6.25).$$

Внутрішньогрупова дисперсія розраховується окремо для кожної групи (6.25):

$$\sigma_i^2 = \frac{\sum (y - \bar{y}_i)^2}{n}, \quad (6.26).$$

де y - значення ознаки окремих елементів сукупності.

Для встановлення тісноти зв'язку між ознаками міжгрупову дисперсію порівнюють із загальною. Відношення міжгрупової дисперсії до загальної називається кореляційним. Його позначають грецькою буквою η^2 (ета) і розраховують за формулою:

$$\eta^2 = \frac{\delta^2}{\sigma^2}. \quad (6.27).$$

Кореляційне відношення може змінюватись від 0 до ± 1 . У випадку якщо $\eta^2 = 0$ - міжгрупова дисперсія дорівнює нулю, а зв'язок між факторною і результативною ознаками відсутній.

Якщо $\eta^2 = 1$, то міжгрупова дисперсія рівна загальній, а внутрішньогрупова дорівнює нулю. Зв'язок функціональний. Чим ближче значення η^2 до 1, тим тіснішим є зв'язок.

Розглянемо приклад розрахунку дисперсій, використавши дані табл. 6.6.

Таблиця 6.6

Матеріали для дисперсійного аналізу

| Сорт пшениці | Кількість сортодільниць, (f) | Урожайність у кожній сортодільниці, ц (y) | $\sum y$ | \bar{y}_i |
|-----------------|----------------------------------|---|----------|-------------|
| Миронівська-808 | 6 | 25; 27; 30; 35; 30; 30 | 177 | 29,5 |
| Поліська | 4 | 28; 24; 30; 30 | 112 | 28,0 |
| Київська | 10 | 40; 36; 35; 42; 45; 42; 42; 44; 40; 39 | 405 | 40,5 |
| В цілому | 20 | | 694 | 34,7 |

Варіація урожайності окремо для кожного сорту пшениці визначається шляхом розрахунку трьох внутрішньогрупових дисперсій. Для сорту Миронівська-808 вона становитиме:

$$\sigma_1^2 = \frac{(25 - 29,5)^2 + (27 - 29,5)^2 + (30 - 29,5)^2 + (35 - 29,5)^2 + (30 - 29,5)^2 + (30 - 29,5)^2}{6} =$$

$$= \frac{20,25 + 6,25 + 0,25 + 30,25 + 0,25 + 0,25}{6} = \frac{57,5}{6} = 9,58.$$

Внутрішньогрупова дисперсія для сорту пшениці "Поліська" становитиме:

$$\sigma_2^2 = \frac{(28 - 28,0)^2 + (24 - 28,0)^2 + (30 - 28,0)^2 + (30 - 28,0)^2}{4} = \frac{16 + 4 + 4}{4} = 6.$$

Відповідно для сорту "Київська"

$$\begin{aligned}\sigma_3^2 &= \frac{(40-40,5)^2 + (36-40,5)^2 + (35-40,5)^2 + (42-40,5)^2 + (45-40,5)^2 + (42-40,5)^2 + (42-40,5)^2}{10} + \\ &+ \frac{(44-40,5)^2 + (40-40,5)^2 + (39-40,5)^2}{10} = \frac{0,25 + 20,25 + 30,25 + 2,25 + 20,25 + 2,25 + 2,25}{10} + \\ &+ \frac{12,25 + 0,25 + 2,25}{10} = \frac{92,5}{10} = 9,25.\end{aligned}$$

Середня з групових дисперсій $\bar{\sigma}^2$ дорівнює:

$$\bar{\sigma}^2 = \frac{9,58 \cdot 6 + 6 \cdot 4 + 9,25 \cdot 10}{20} = \frac{173,98}{20} = 8,7.$$

Розрахуємо міжгрупову дисперсію:

$$\delta^2 = \frac{(29,5 - 34,7)^2 \cdot 6 + (28,0 - 34,7)^2 \cdot 4 + (40,5 - 34,7)^2 \cdot 10}{20} = \frac{162,24 + 179,56 + 336,4}{20} = \frac{678,2}{20} = 33,91,$$

$$\text{де } \bar{y} = \frac{\sum \bar{y}_i f_i}{\sum f_i} = \frac{694}{20} = 34,7.$$

Використовуючи взаємозв'язок загальної, міжгрупової та внутрігрупової дисперсій (правило додавання дисперсій), визначимо загальну дисперсію $\sigma^2 = 33,91 + 1,24 = 35,15$, $\sigma^2 = 33,91 + 8,7 = 42,61$.

Перевіримо наші розрахунки, обчисливши загальну дисперсію за спрощеною формулою:

$$\begin{aligned}\sigma^2 &= \bar{y}^2 - (\bar{y})^2 = \frac{25^2 + 27^2 + 30^2 + 35^2 + 30^2 + 30^2 + 28^2 + 24^2 + 30^2 + 30^2 + 40^2 + 36^2 + 35^2}{20} + \\ &+ \frac{42^2 + 45^2 + 42^2 + 42^2 + 44^2 + 40^2 + 39^2}{20} - 34,7^2 = \frac{625 + 729 + 900 + 1225 + 900 + 900 + 784}{20} + \\ &+ \frac{576 + 900 + 900 + 1600 + 1296 + 1225 + 1764 + 2025 + 1764 + 1764 + 1936 + 1600 + 1521}{20} - 1204,09 = \\ &= \frac{24934}{20} - 1204,09 = 1246,7 - 1204,09 = 42,61.\end{aligned}$$

Правильність розрахунків підтверджується.

Тестові завдання

1. Варіація значень досліджуваної ознаки ряду розподілу, а також співвідношення варіант і частот (часток) характеризується за допомогою:

- діаграм;
- таблиць;
- моди;
- форм розподілу.

2. У симетричному розподілі значення ознаки:

- а) асиметричні;
- б) атрибутивні;
- в) рівновіддалені від центру і мають однакову частоту;
- г) вершини зміщені вправо або вліво.

3. Найбільш наочною мірою асиметрії є величина відхилення між:

- а) варіантами та частотами;
- б) середньою арифметичною, модою або медіаною;
- в) центрами розподілу;
- г) середньою та стандартизованими відхиленнями.

4. У статистиці напрям і міру скошеності визначають за допомогою:

- а) стандартизованих відхилень;
- б) ексцесу;
- в) асиметрії;
- г) моди або медіани.

5. Гостровершинність розподілу відображає скупченість значень ознаки навколо центру розподілу і називається:

- а) симетрією;
- б) асиметрією;
- в) стандартизованим відхиленням;
- г) ексцесом.

6. Критерій χ^2 Пірсона це сума відхилень:

- а) між емпіричними та теоретичним варіантами;
- б) квадратів між емпіричними та теоретичним варіантами;
- в) квадратів між емпіричними та теоретичним частотами;
- г) між середньою та варіантами.

7. Загальна дисперсія зумовлюється дією:

- а) міжгрупової дисперсії;
- б) систематичних і випадкових факторів;
- в) тільки систематичних факторів;
- г) тільки випадкових факторів.

8. Міжгрупова варіація є систематичною варіацією, котра характеризує вплив на:

- а) факторну ознаку систематичних чинників;
- б) систематичну ознаку випадкових величин;
- в) результативну ознаку систематичних факторів;
- г) всі відповіді неправильні.

9. Випадкову варіацію, котра викликана дією всіх інших факторів, за винятком того, що покладений в основу групування, характеризує:

- а) загальна дисперсія;
- б) міжгрупова дисперсія;
- в) внутрішньогрупова;
- г) дисперсія результативної ознаки.

10. Кореляційне відношення розраховують шляхом ділення:

- а) міжгрупової дисперсії на загальну;
- б) загальної дисперсії на міжгрупову;
- в) міжгрупової на внутрішньогрупову;
- г) внутрішньогрупової на загальну.

11. Кореляційне відношення може набувати значень:

- а) 0;
- б) 1;
- в) від 0 до 1;
- г) від 0 до ± 1 .

12. Якщо між групова дисперсія тотожна загальній, то внутрішньогрупова дорівнює:

- а) 1;
- б) 0;
- в) менше 1;
- г) більше 1.

13. Правило додавання дисперсії полягає в тому, що:

- а) загальна дисперсія дорівнює міжгруповій;
- б) загальна дисперсія дорівнює нулю;
- в) загальна дисперсія дорівнює одиниці;
- г) випадкова дисперсія є меншою за міжгрупову.

Типові задачі

Задача 1. За результатами обстеження підприємств міста одержано дані про терміни експлуатації обладнання:

| Термін експлуатації, років | До 5 | 5-10 | 10-15 | 15-20 | 20-25 | 25-30 | понад 30 |
|----------------------------|------|------|-------|-------|-------|-------|----------|
| Кількість одиниць | 120 | 280 | 385 | 560 | 210 | 143 | 57 |

Визначить коефіцієнти асиметрії та ексцесу.

Задача 2. Розподіл безробітних у регіоні за тривалістю безробіття та статтю характеризується такими даними (%):

| | Жінки | Чоловіки |
|----------------------------------|--------------|-----------------|
| Всього безробітних. | 100 | 100 |
| з них за тривалістю незайнятості | | |
| до 1 місяця | 6,1 | 4,5 |
| 1-3 | 18,5 | 15,3 |
| 4-6 | 30,9 | 36,2 |
| 7-9 | 23,6 | 28,1 |
| 10-12 | 18,3 | 10,4 |
| більше 1 року | 2,6 | 5,5 |

Визначіть коефіцієнти асиметрії та ексцесу.

Задача 3. Маємо дані про розподіл магазинів за обсягом річного товарообороту.

| Обсяг товарообороту, млн. грн. | Кількість магазинів |
|---------------------------------------|----------------------------|
| До 10 | 8 |
| 10 – 15 | 25 |
| 15 – 20 | 32 |
| 20 – 25 | 81 |
| 25 – 30 | 40 |
| 30 – 35 | 20 |
| 35 – 40 | 18 |
| 40 – 45 | 14 |
| 45 – 50 | 9 |
| 50 і більше | 6 |

Визначіть коефіцієнт асиметрії. Побудуйте гістограму розподілу магазинів за обсягом товарообороту.

Задача 4. Маємо дані про розподіл магазинів за торговою площею на 1 січня звітного року. Використовуючи центральні моменти першого, другого, третього і четвертого порядку, визначіть коефіцієнти асиметрії та ексцесу. Побудуйте графік розподілу. Перевірте гіпотезу про наявність нормального розподілу магазинів за допомогою критеріїв К. Пірсона (χ^2) та А. Колмогорова (λ).

| Торговельна площа магазину, м² | Кількість магазинів |
|--|----------------------------|
| До 200 | 10 |
| 200-300 | 14 |
| 300-400 | 21 |
| 400-500 | 39 |
| 500-600 | 18 |

| | |
|-----------|----|
| 600-700 | 11 |
| 700-800 | 7 |
| понад 800 | 5 |

Питання для самоконтролю

1. Показники, що використовуються для характеристики варіації ознаки.
2. Коефіцієнти варіації.
3. Використання "способу моментів" при вивченні варіації.
4. Види дисперсій.
5. Правило додавання дисперсій.
6. Різниця між систематичною і випадковою варіацією.
7. Кореляційне відношення.
8. Відмінність між одно-, дво- і багатoverшинними розподілами.
9. Симетричний розподіл.
10. Види розподілу.
11. Моменти розподілу.
12. Значення коефіцієнтів асиметрії і ексцесу.
13. Показники, що характеризують концентрацію і локалізацію.
14. Значення аналізу кривих розподілу.
15. Різниця між теоретичними і емпіричними частотами.
16. Для чого використовують критерій χ^2 - Пірсона?
17. Для чого використовують критерій λ - Колмогорова?

РОЗДІЛ 7. СТАТИСТИЧНІ МЕТОДИ ВИМІРЮВАННЯ ВЗАЄМОЗВ'ЯЗКІВ

7.1 Поняття та види взаємозв'язків

Всі явища об'єктивного світу, що нас оточують, перебувають не ізольовано один від одного, а в тісному взаємозв'язку й взаємозалежності. При цьому найбільш важлива роль у статистиці

відводиться причинній залежності. Зв'язок може виникати лише тоді, коли існує причина та наслідок, тобто присутні фактори, котрі взаємозв'язані між собою. Ознаку, що характеризує наслідок називають **результативною**, а та ознака, котра на неї впливає називається **факторною**. Наприклад, у промисловості будь-яка результативна ознака (продуктивність праці, собівартість, рентабельність, прибуток) зазнає впливу багатьох причин-факторів: забезпеченість виробничими ресурсами, рівень використання виробничих потужностей, технічна озброєність, кваліфікація працюючих, імплементація інноваційних процесів тощо. Вивчення взаємозв'язку між ознаками дає змогу пізнати суть явищ і процесів, які досліджуються та встановити закономірність їх розвитку.

Завдання статистики полягають у вивченні зв'язків між явищами з метою встановлення їх напрямку та характеру, а також оцінити зв'язок кількісно.

Статистика досліджує **функціональні та стохастичні види зв'язку**. При першому кожному значенню факторної ознаки x відповідає одне значення результативної ознаки y , інакше кажучи, функціональний зв'язок передбачає повну відповідність між причиною та наслідком, факторною й результативною ознаками. Такі зв'язки вивчаються точними науками: фізикою, хімією, математикою. Наприклад, площа кола залежить від величини радіуса $S_{kr} = \pi R^2$, розмір трикутника від розміру його сторін.

Стохастичний - це зв'язок, при якому кожному значенню x відповідає не одне, а декілька можливих значень y , які варіюють і утворюють ряд розподілу за даною ознакою. Розподіл одиниць сукупності за одною ознакою при фіксованому значенні іншої називається умовним. Підвидом стохастичного зв'язку є кореляційний зв'язок.

Прикладом стохастичного зв'язку може слугувати розподіл безробітного населення (за методологією МОП) за статтю та віковими групами (табл. 7.1).

Таблиця 7.1

Розподіл безробітного населення (за методологією МОП) за статтю та віковими групами
(у середньому за період, у % до кількості економічно активного населення відповідної вікової групи)

| Вік | жінки | | | | | | чоловіки | | | | | |
|-------|-------|------|------|------|------|------|----------|------|------|------|------|------|
| | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 |
| 15–24 | 15,1 | 6,8 | 18,7 | 16,1 | 16,3 | 22,4 | 15,1 | 17,8 | 18,6 | 18,1 | 18,2 | 23,7 |
| 25–29 | 9,8 | 16,7 | 8,2 | 8 | 7,4 | 9,7 | 9,8 | 11,2 | 9,9 | 10,6 | 9,6 | 12,2 |
| 30–39 | 7,1 | 8,1 | 6,1 | 6 | 6 | 6,9 | 7,1 | 8,7 | 8,4 | 7,4 | 6,7 | 9,3 |
| 40–49 | 6,8 | 6,8 | 5,7 | 5,8 | 5,3 | 5,6 | 6,8 | 7,6 | 7 | 6,7 | 7,2 | 9,1 |
| 50–59 | 4,3 | 6 | 4,2 | 4,1 | 4,4 | 4,4 | 4,3 | 6,8 | 6 | 6,5 | 5,7 | 7,6 |
| 60–70 | 0,1 | 3,9 | 0,1 | 0,1 | - | 0,1 | 0,1 | 0,1 | 0 | 0 | | 0 |

* - дані наведені без урахування тимчасово окупованої території Автономної Республіки Крим, м.Севастополя та частини зони проведення антитерористичної операції.

Термін "**кореляція**" (correlation) англійського походження і означає дослівно - співвідношення, відповідність. Даний вид зв'язку проявляється у тому, що зі зміною факторної ознаки x змінюються групові середньої - результативної ознаки y , тобто замість умовних розподілів порівнюються їх середні значення. В нашому прикладі (табл.7.1) кожній групі за факторною ознакою віку (x) відповідає свій розподіл (y) кількість жінок і чоловіків.

Кореляційний зв'язок є неповним, неточним і простежується лише при великій кількості випробувань і лише через середні величини. Зокрема, на загальний добробут соціуму впливає платоспроможність гривні, але не тільки. Розмір заробітної плати, пенсії, стипендії, допомога малозабезпеченим, неблагополучним сім'ям, ціни на продукти й послуги також впливають на соціальне забезпечення. Як бачимо з прикладу, на результативну ознаку (загальний добробут соціуму) має вплив не лише один, а багато чинників, ступінь дії яких є різним. Одні з них суттєво впливають на результат - інші навпаки.

За напрямком зв'язки бувають прямими та зворотними. В першому випадку зростання фактора-причини зумовлює зростання результативної ознаки. При зворотному зв'язку зростання причини-

ознаки спричиняє до її зменшення. Також зв'язки поділяють на лінійні й нелінійні. Перші, проявляються між ознаками в середньому через лінійні співвідношення. Нелінійний взаємозв'язок описується нелінійною функцією, а змінні при цьому пов'язані між собою в середньому нелінійно.

Кореляційний зв'язок може бути описаний за допомогою лінії регресії. Лінія регресії (y від x) - це функція, котра зв'язує середні значення ознаки "у" зі значеннями ознаки "х".

Лінія регресії може бути подана таблично, аналітично, графічно. Останні дві з них слугують основними моделями кореляційного зв'язку - аналітичного групування та регресивну.

Суть побудови таких моделей полягає в теоретичному обґрунтуванні та оцінці лінії регресії, вимірюванні тісноти зв'язку між ознаками, перевірці істотності зв'язку щоби довести не випадковий характер закономірностей, які досліджуються.

7.2 Парна кореляція

Інколи в практиці статистичного дослідження доводиться визначати взаємозв'язок між двома ознаками, оцінювання якого потребує побудови моделі парної кореляції. При виборі її функціонального виду використовують графіки, аналітичні групування, балансовий метод та паралельні ряди.

Часто кореляційний зв'язок охарактеризовують завдяки побудови кореляційного поля. При цьому значення факторної ознаки відкладаються на осі абсцис, а результативної на осі ординат. Місце їх перетину позначають крапками. Відповідно до характеру розміщення крапок на кореляційному полі ідентифікуються висновки щодо форми лінії регресії й тісноти зв'язку (рис. 7.1, 7.2, 7.3).

Використання методу аналітичних групувань із метою оцінювання зв'язку між факторною й результативною ознаками нами охарактеризовано раніше. Тому зупинимось на використанні

балансового методу та паралельних рядів в процесі встановлення рівня залежності ознак, які досліджуються.

Балансовий метод допомагає встановити зв'язок між ознаками шляхом побудови балансу показників, які логічно між собою взаємопов'язані.

Використання методу паралельних рядів полягає в побудові двох рядів, які логічно взаємодіють один із одним.

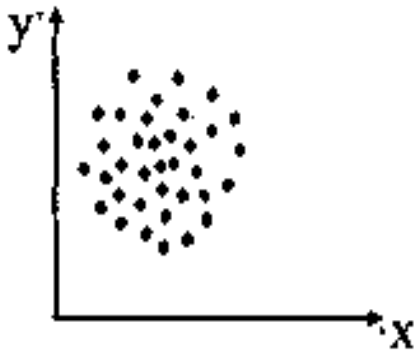


Рис.7.1 Зв'язок відсутній

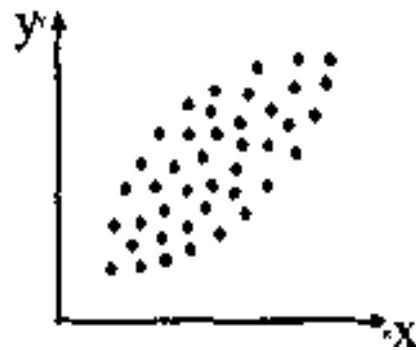


Рис. 7.2. Зв'язок прямий

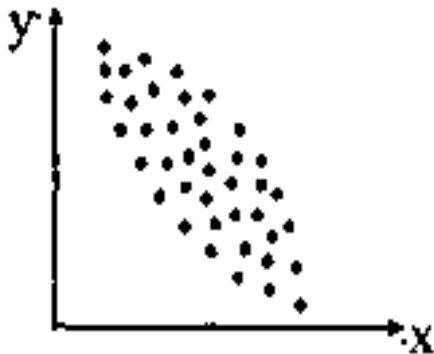


Рис. 7.3. Зв'язок зворотний

Зважаючи на характер варіації значень ознак, які покладені в основу дослідження, встановлюють наявність зв'язку, його тісноту та напрямок.

У практиці статистичних досліджень значно більше використання одержав кореляційний метод зв'язку, оскільки завдяки йому забезпечується кількісна оцінка взаємозв'язку ознак.

Основними завданнями при використанні кореляційного методу зв'язку є, передусім пошук математичної формули, за допомогою якої описується залежність y від x та забезпечується

вимірювання тісноти й напрямку зв'язку. Розв'язання першого завдання здійснюється завдяки встановленню форми зв'язку з наступним знаходженням параметрів рівняння регресії.

Найбільш широке використання при застосуванні кореляційного зв'язку в статистиці одержали такі моделі:

1) прямолінійна:

$$y = a_0 + a_1x, \quad (7.1)$$

де a_0 - початок прямої на графіку; a_1 - коефіцієнт регресії, за допомогою якого оцінюється зміна в середньому результативної ознаки y при зміні x на одиницю;

2) криволінійна у вигляді:

а) параболи другого порядку:

$$y = a_0 + a_1x + a_2x^2 \quad (7.2)$$

(або вищих порядків);

б) гіперболи:

$$y = a_1 + \frac{a_1}{x}; \quad (7.3)$$

в) степеневої функції

$$y = a_0x^{a_1}. \quad (7.4).$$

Прямолінійна форма зв'язку використовується за умови, якщо коливання змінної x та результативної ознаки " y " є незначними. При прямому зв'язку a_1 - має додатне значення, при зворотному - від'ємне.

Значене коливання варіант взаємозв'язаних досліджуваних ознак визначається завдяки побудови нелінійних (криволінійних) рівнянь регресії.

Вибору тієї чи іншої моделі криволінійної регресії повинен передувати попередній теоретичний аналіз суті зв'язку. Зокрема, вивчення взаємозв'язку між собівартістю та обсягом продукції оцінюється шляхом побудови рівняння гіперболи. Взаємозалежність між кількістю внесених добрив і урожайністю пшениці описується параболою другого порядку.

Пошук параметрів рівнянь зв'язку найчастіше здійснюють за допомогою побудови системи нормальних рівнянь, які відповідають вимогам способу найменших квадратів.

При прямолінійній залежності система нормальних рівнянь має вигляд:

$$\begin{cases} \sum y = na_0 + a_1 \sum x \\ \sum xy = a_0 \sum x + a_1 \sum x^2 \end{cases} \quad (7.5).$$

Завдання зводиться до визначення параметрів a_0 та a_1 , які знаходять шляхом розв'язання системи двох нормальних рівнянь, завдяки чому одержують необхідні підсумки $\sum y; \sum x; \sum xy; \sum x^2$.

Приведемо приклад розв'язку прямолінійного рівняння. Витрати умовного палива (y) та обсяг виробництва продукції (x) на десяти однорідних підприємствах характеризується такими даними (див. табл. 7.2).

Завдання зводиться до оцінювання кореляційного зв'язку між досліджуваними ознаками. Використовуючи розрахункові дані табл. 7.2, побудуємо систему рівнянь:

$$\begin{cases} 10a_0 + 125a_1 = 80 \\ 125a_0 + 1961a_1 = 1218 \end{cases}$$

Розв'язавши рівняння одержимо:

$$a_0 = 1,16, \quad a_1 = 0,547x.$$

Звідки

$$\bar{y}_x = 1,16 + 0,547x.$$

Таблиця 7.2

Дані для кореляційного аналізу

| № під-приємств | Обсяг виробництва продукції, тис. одиниць, x | Витрати умовного палива, тонн, y | x^2 | xy | $\bar{y}_x = 1,16 + 0,547x$ | y^2 |
|----------------|--|------------------------------------|------------------|------------------|-----------------------------|------------------|
| 1 | 5 | 4 | 25 | 20 | 3,9 | 16 |
| 2 | 6 | 4 | 36 | 24 | 4,4 | 16 |
| 3 | 8 | 6 | 64 | 48 | 5,5 | 36 |
| 4 | 8 | 5 | 64 | 40 | 5,5 | 25 |
| 5 | 10 | 7 | 100 | 70 | 6,6 | 49 |
| 6 | 10 | 8 | 100 | 80 | 6,6 | 64 |
| 7 | 14 | 8 | 196 | 112 | 8,8 | 64 |
| 8 | 20 | 10 | 400 | 200 | 12,1 | 100 |
| 9 | 20 | 12 | 400 | 240 | 12,1 | 144 |
| 10 | 24 | 16 | 576 | 384 | 14,3 | 256 |
| | $\sum x = 125$ | $\sum y = 80$ | $\sum x^2 = 196$ | $\sum xy = 1218$ | $\sum \bar{y}_x = 79,8$ | $\sum y^2 = 770$ |

Дане рівняння читають так: з ростом обсягу виробництва на одну тисячу одиниць витрати пального зростають на 0,547 тонни.

Оцінка тісноти зв'язку між приведеними ознаками в статистиці забезпечується за допомогою розрахунку коефіцієнта парної кореляції r . При цьому для вимірювання тісноти (щільності) зв'язку при лінійній залежності найбільш широко використовуються формули:

$$r = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sigma_x \sigma_y}; \quad (7.6).$$

$$r = \frac{\sum (x - \bar{x}) \cdot (y - \bar{y})}{n \sigma_x \sigma_y}; \quad (7.7).$$

$$r = \frac{\sum (x - \bar{x}) \cdot (y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}}, \quad (7.8).$$

де r - лінійний коефіцієнт кореляції.

Значення r може змінюватись у межах від 0 до ± 1 . Числове значення r та знак \pm використовують для оцінки тісноти та напрямку зв'язку.

Використаємо дані, наведені в табл. 7.2, розрахуємо коефіцієнт кореляції за допомогою формули:

$$r = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sigma_x \sigma_y}.$$

Знайдемо значення чисельника:

$$\overline{xy} = 121,8; \bar{x} = 12,5; \bar{y} = 8.$$

Розрахуємо середні квадратичні відхилення σ_x та σ_y :

$$\sigma_x = \sqrt{\overline{x^2} - (\bar{x})^2} = \sqrt{196,1 - 12,5^2} = \sqrt{39,85} = 6,31;$$

$$\sigma_y = \sqrt{\overline{y^2} - (\bar{y})^2} = \sqrt{77 - 8^2} = \sqrt{13} = 3,6.$$

$$\text{Знайдемо коефіцієнт кореляції } r = \frac{121,8 - 12,5 \cdot 8}{6,31 \cdot 3,6} = 0,959.$$

Одержаний результат свідчить про досить тісний зв'язок між показниками обсягу виробництва та витратами умовного палива.

Щільність нелінійного зв'язку оцінюється співвідношенням варіацій емпіричних та теоретичних значень результативної ознаки (y). Відхилення індивідуальних значень результативної ознаки (y) від (\bar{y}) можна представити як відхилення від лінії регресії ($y - \bar{Y}$) і відхилення лінії регресії від середньої ($\bar{Y} - \bar{y}$). В першому випадку одержаний результат характеризує вплив фактора x , а відхилення ($y - \bar{Y}$) - дію всіх інших факторів.

Тіснота зв'язку в статистиці розраховується також на основі оцінки взаємозв'язку дисперсій, використовуючи при цьому правило додавання дисперсій. На відміну від аналітичного групування, при якому оцінками ліній регресії є групові середні результативної ознаки (міжгрупова дисперсія), в кореляційно-дисперсійному аналізі – теоретичні значення останньої. У відповідності з правилом додавання дисперсії можна справедливо наступне:

$$\delta_y^2 = \sigma_y^2 - \sigma_{\text{залиш.}}^2;$$

$$\delta_y^2 = \frac{1}{n} \sum (Y - \bar{y})^2 = \frac{1}{n} (a_0 \sum x + a_1 \sum xy) - \bar{y}^2 - \text{факторна дисперсія} \quad (7.9);$$

$$\sigma_y^2 = \frac{1}{n} \sum (y - \bar{y})^2 - \text{загальна дисперсія} \quad (7.10);$$

$$\sigma_{\text{залиш.}}^2 = \frac{1}{n} \sum (y - Y)^2 - \text{залишкова дисперсія} \quad (7.11).$$

Не виникає сумнівів, що розмір δ_y^2 буде зростати з зростанням впливу x на y . Відношення міжгрупової факторної дисперсії δ_y^2 до загальної σ_y^2 називається **коефіцієнтом детермінації**, за допомогою якого встановлюють щільність зв'язку

$$R^2 = \frac{\delta_Y^2}{\sigma_y^2}. \quad (7.12).$$

Корінь квадратний з коефіцієнта детермінації називається **індексом кореляції R** .

Зазначимо, що при використанні аналітичного групування, мірою щільності зв'язку було кореляційне відношення η^2 , яке за змістом є ідентичним коефіцієнту детермінації R^2 .

$$\eta^2 = \frac{\delta_y^2}{\sigma_y^2}, \quad (7.13),$$

$$\text{звідки } \delta^2 = \frac{\sum (\bar{y}_i - \bar{y})^2 f}{\sum f}; \quad (7.14).$$

$$\sigma^2 = \frac{\sum (y - \bar{y})^2 f}{\sum f}, \text{ де } \bar{y}_i - \text{групові середні.} \quad (7.15).$$

Використовуючи зазначені теоретичні положення, розрахуємо вказані величини, на основі інформації табл. 7.2,7.3:

Таблиця 7.3

Дані для розрахунку щільності зв'язку

| № п-ва | Обсяг виробництва продукції, тис. шт., x | Витрати умовного палива тонн, y | Y_x | $y - \bar{Y}$ | $(y - \bar{Y})^2$ | $Y_x - \bar{Y}$ | $((Y_x - \bar{Y})^2)$ | $y - \bar{Y}_x$ | $(y - \bar{Y}_x)^2$ |
|--------|--|-----------------------------------|------------|---------------|------------------------|-----------------|-----------------------------------|-----------------|----------------------------------|
| 1. | 2. | 3. | 4. | 5. | 6. | 7. | 8. | 9. | 10. |
| 1 | 5 | 4 | 3,9 | -4 | 16 | -4,1 | 16,81 | 0,1 | 0,01 |
| 2 | 6 | 4 | 4,4 | -4 | 16 | -3,6 | 12,96 | -0,4 | 0,16 |
| 3 | 8 | 6 | 5,5 | -2 | 4 | -2,5 | 6,25 | 0,5 | 0,25 |
| 4 | 8 | 5 | 5,5 | -3 | 9 | -2,5 | 6,25 | -0,5 | 0,25 |
| 5 | 10 | 7 | 6,6 | -1 | 1 | -1,4 | 1,96 | 0,4 | 0,16 |
| 6 | 10 | 8 | 6,6 | 0 | 0 | -1,4 | 1,96 | 1,4 | 1,96 |
| 7 | 14 | 8 | 8,8 | 0 | 0 | 0,8 | 0,64 | -0,8 | 0,64 |
| 8 | 20 | 10 | 12,1 | 2 | 4 | 4,1 | 16,81 | -2,1 | 4,41 |
| 9 | 20 | 12 | 12,1 | 4 | 16 | 4,1 | 16,81 | -0,1 | 0,01 |
| 10 | 24 | 16 | 14,3 | 8 | 64 | 6,3 | 39,69 | 1,7 | 2,89 |
| | $\sum x = 125$ | $\sum y = 80$ | $\sum y_x$ | - | $\sum (y - Y)^2 = 130$ | - | $\sum (Y_x - \bar{Y})^2 = 120,14$ | - | $\sum (y - \bar{Y}_x)^2 = 10,74$ |

В результаті проведення розрахунків одержимо:

$$\sigma^2 = \frac{\sum (Y_x - \bar{Y})^2}{n} = \frac{120,14}{10} = 12,014;$$

$$\sigma^2 = \frac{\sum (y - \bar{Y})^2}{n} = \frac{130}{10} = 13;$$

$$R = \sqrt{\frac{\delta^2}{\sigma^2}} = \sqrt{\frac{12,014}{13}} = 0,961;$$

$$\eta^2 = \sqrt{1 - \frac{\delta^2_{\text{залиш.}}}{\sigma^2}} = \sqrt{1 - \frac{1,074}{13}} = 0,958.$$

Піднесемо $\eta^2=0,958$ до квадрату в результаті одержимо 0,918. Таким чином, 91,8% варіації витрат палива залежить від обсягу виробництва продукції, а 8,2% припадає на інші фактори.

Часто в процесі оцінювання тісноти зв'язку використовують невеликі за обсягом статистичні сукупності, що може спричинити до випадкових коливань коефіцієнта регресії. Перевірка достовірності (істотності) одержаного результату забезпечується за допомогою t -критерію (Стюдента), як відношення коефіцієнта регресії a_1 , до власної стандартної похибки μ_{a_1} , тобто

$$t = \frac{a_1}{\mu_{a_1}}. \quad (7.16).$$

В свою чергу похибка коефіцієнта регресії залежить від розміру варіації факторної ознаки σ_x^2 , залишкової дисперсії $\sigma^2_{\text{залиш.}}$ і числа ступенів свободи $k = n - 2$:

$$\mu = \sqrt{\frac{\sigma^2_{\text{залиш.}}}{\sigma_x^2 (n - 2)}}. \quad (7.17).$$

У випадку, коли фактичне значення перевищує t критичне, то зв'язок між факторною і результативною ознаками є істотним, і навпаки. Підкреслимо, що мова йде про використання t -критерію (Стюдента) за умови оцінювання лінійного зв'язку.

Істотність кореляційного зв'язку досліджується також завдяки порівнянню фактичних значень R^2 і η^2 з їх критичним значенням.

Знову ж таки, при умові, що фактичні значення коефіцієнта детермінації, індексу кореляції чи кореляційне відношення за

абсолютною величиною є більшими за критичні, зв'язок оцінюється як істотний, не випадковий.

Критичне значення вибирають таким чином, щоби ймовірність отримання значень R^2 , η^2 і μ більшого від критичного (за умови відсутності зв'язку між ними) була достатньо мала. Таку ймовірність називають рівнем істотності α .

В статистичній практиці ймовірність найбільш часто оцінюється за допомогою таких рівнів істотності: $\alpha = 0,05$ і $\alpha = 0,01$ (див. табл.7.4)

Таблиця 7.4

**Критичні значення коефіцієнта детермінації R^2 і
кореляційного відношення η^2 для рівня істотності $\alpha = 0,05$**

| К ₂ | К ₁ | | | | |
|----------------|----------------|-------|-------|-------|-------|
| | 1 | 2 | 3 | 4 | 5 |
| 5 | 0,569 | 0,699 | 0,764 | 0,806 | 0,835 |
| 6 | 0,500 | 0,632 | 0,704 | 0,751 | 0,785 |
| 7 | 0,444 | 0,575 | 0,651 | 0,702 | 0,739 |
| 8 | 0,399 | 0,527 | 0,604 | 0,657 | 0,697 |
| 9 | 0,362 | 0,488 | 0,563 | 0,618 | 0,659 |
| 10 | 0,332 | 0,451 | 0,527 | 0,582 | 0,624 |
| 12 | 0,283 | 0,394 | 0,466 | 0,521 | 0,564 |
| 14 | 0,247 | 0,348 | 0,417 | 0,471 | 0,514 |
| 16 | 0,219 | 0,312 | 0,378 | 0,429 | 0,477 |
| 18 | 0,197 | 0,283 | 0,345 | 0,394 | 0,435 |
| 20 | 0,179 | 0,259 | 0,318 | 0,364 | 0,404 |
| 24 | 0,151 | 0,221 | 0,273 | 0,316 | 0,353 |
| 28 | 0,130 | 0,193 | 0,240 | 0,279 | 0,314 |
| 32 | 0,115 | 0,171 | 0,214 | 0,250 | 0,282 |
| 36 | 0,102 | 0,153 | 0,192 | 0,226 | 0,256 |
| 40 | 0,093 | 0,139 | 0,176 | 0,207 | 0,234 |
| 50 | 0,075 | 0,113 | 0,143 | 0,170 | 0,194 |
| 60 | 0,063 | 0,095 | 0,121 | 0,144 | 0,165 |
| 80 | 0,047 | 0,072 | 0,093 | 0,110 | 0,127 |
| 100 | 0,038 | 0,058 | 0,075 | 0,090 | 0,103 |
| і т.д. | | | | | |

Розподіл R^2 та η^2 , наведений в табл. 7.3, залежить від числа ступенів свободи для міжгрупової k_1 і залишкової k_2 дисперсій. k_1 , і k_2 обчислюють: $k_1 = m - 1$, $k_2 = n - m$, де m - число груп або параметрів функції m ; n - кількість одиниць сукупності.

В нашому прикладі значення коефіцієнта детермінації для $k_1 = 2 - 1 = 1$ і $k_2 = 10 - 2 = 8$ становить $R_{0,95}^2(1,8) = 0,399$.

Обчислений за даними таблиці 7.3 результат щільності зв'язку $R^2 = 0,924$ перевищує критичне значення, що з імовірністю 0,95 підтверджує істотність зв'язку між ознаками, котрі досліджуються. Інакше кажучи, в 95 випадках із 100 фактичне значення перевищить критичне.

Перевіряють істотність зв'язку також за допомогою F -критерію (критерію Фішера). Критерій Фішера функціонально пов'язаний з кореляційним відношенням η^2 .

$$F = \frac{\eta^2}{1 - \eta^2} \cdot \frac{k_2}{k_1}, \quad (7.18)$$

або

$$F = \frac{\delta^2}{\sigma^2} \cdot \frac{k_2}{k_1}. \quad (7.19).$$

У статистиці побудовані таблиці критичного значення F - критерію. Перевірка істотності проводиться аналогічно.

7.3 Непараметричні методи оцінювання взаємозв'язку між ознаками

Застосування методів кореляційного та дисперсійного аналізу базується на використанні таких основних параметрів розподілу, як середні величини й дисперсії, тому їх називають параметричними. Проте в практиці статистичних досліджень є випадки, коли параметричне оцінювання зв'язку унеможливлено або його недоцільно використовувати. В зв'язку з цим статистичною наукою розроблені методи, завдяки яким оцінка взаємозв'язку між явищами забезпечується без використання кількісних значень ознаки, а також параметрів розподілу. Такі методи називаються непараметричними. Використання непараметричних методів оцінки тісноти

стохастичного зв'язку ґрунтується на порівнюванні частот або часток умовних розподілів у таблицях співзалежності, наприклад (табл.7.5).

Таблиця 7.5

**Розподіл випускників школи за наміром вступити до вузів,
залежно від оцінок незалежного тестування**

| Групи випускників, за оцінками незалежного тестування, x | Кількість випускників, що мають намір вступити до (f): | | | |
|--|--|-----------------|---------|----------------|
| | державного вузу | приватного вузу | коледжу | Разом f_{oi} |
| відмінні | 30 | 10 | 2 | 42 |
| переважно відмінні | 20 | 15 | 4 | 39 |
| переважно добрі і відмінні | 5 | 20 | 8 | 33 |
| добрі | - | 5 | 10 | 15 |
| добрі і задовільні | - | - | 8 | 8 |
| Разом f_{oj} | 55 | 50 | 32 | 137 |

Розподіл частот у таблиці 7.5 засвідчує про наявність стохастичного взаємозв'язку між оцінками незалежного тестування та бажанням абітурієнтів навчатися у навчальному закладі відповідного рівня акредитації та форми власності.

Вимірювання тісноти стохастичного зв'язку забезпечується завдяки розрахунку **коефіцієнта взаємного узгодження**. Його обчислення базується на розбіжностях частот (часток) умовного та безумовного розподілів і здійснюється за формулою:

$$\chi^2 = f_{i0} \sum \frac{(d_{ij} - d_j)^2}{d_j}, \quad (7.20).$$

де χ^2 - квадратична спряженість Пірсона;

f_{i0} - підсумкова частота по i -му рядку за ознакою x ;

d_{ij} - частка j -го стовпця по i -тому рядку;

d_j - частка j -го стовпця безумовного розподілу.

Величину χ^2 можна обчислити, порівнюючи фактичні частоти таблиць співзалежності f_{ij} з теоретичними частотами f'_{ij} , тобто

$$\chi^2 = \sum_i \sum_j \frac{(f_{ij} - f'_{ij})^2}{f'_{ij}} \quad (7.22);$$

$$f'_{ij} = \frac{f_{i0} \cdot f_{0j}}{n} \quad (7.23),$$

де f_{0j} - підсумкові частоти за ознакою y ;

f_{i0} - підсумкові частоти за ознакою x ;

n - обсяг сукупності.

Якщо ознаки незалежні, то частки умовних і безумовних розподілів збігаються: $d_{ij} = d_j$. Однаковими будуть також фактичні і теоретичні частоти $f_{ij} = f'_{ij}$, тоді $\chi^2 = 0$. Критичні значення χ^2 для $\alpha = 0,05$ і числа ступенів свободи $k = (m_x - 1)(m_y - 1)$ наведено в табл. 7.6.

Таблиця 7.6

Критичні значення $\chi^2_{0,95}(k)$

| k | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|----------|------|------|------|------|-------|-------|-------|-------|
| χ^2 | 3.84 | 5.99 | 7.81 | 9.49 | 11.07 | 12.59 | 14.07 | 15.51 |

Наведемо приклад розрахунку χ^2 . Для цього скористаємось даними табл. 7.7, в якій містяться дані щодо віку матері та кількості народжених дітей. Між досліджуваними ознаками існує тісний взаємозв'язок, який полягає в тому, що саме вік жінки є тим чинником, який суттєво впливає на народжуваність. Особливо чітко це проявляється в категорії жінок у віці 19-25 років.

Таблиця 7.7

Розподіл жінок за віком і бажанням народжувати дітей

| Вік, x , років | Проживають у | | | Разом, f_{i0} |
|---------------------|--------------|----------------------------|------|-----------------|
| | місті | селищі міського типу | селі | |
| 15-18 | 12 | 6 | 2 | 20 |
| 19-25 | 10 | 25 | 15 | 50 |
| 25-33 | 3 | 9 | 18 | 30 |

| | | | | |
|-----------------|----|----|----|-----|
| Разом, f_{0j} | 25 | 40 | 35 | 100 |
|-----------------|----|----|----|-----|

Проведемо відповідні розрахунки для кожної вікової групи. В результаті одержимо:

$$\text{перша група } f'_{11} = \frac{f_{i0} \cdot f_{0j}}{n} = \frac{20 \cdot 25}{100} = 5;$$

$$\text{у другій віковій групі } f'_{22} = \frac{50 \cdot 40}{100} = 20;$$

$$\text{у третій віковій групі } f'_{33} = \frac{30 \cdot 35}{100} = 10,5.$$

Розрахуємо абсолютну величину відхилень f_{ij} від f'_{ij} за допомогою

квадратичної спряженості χ^2 Пірсона:

$$\begin{aligned} \chi^2 &= \sum_i \sum_j \frac{(f_{ij} - f'_{ij})^2}{f'_{ij}} = n \left[\sum_i \sum_j \frac{f_{ij}^2}{f_{i0} f_{0j}} - 1 \right] = \\ &= \left[\frac{12^2}{20 \cdot 25} + \frac{6^2}{20 \cdot 40} + \frac{2^2}{20 \cdot 35} + \frac{10^2}{50 \cdot 25} + \frac{25^2}{50 \cdot 40} + \frac{15^2}{50 \cdot 35} + \frac{3^2}{30 \cdot 25} + \frac{9^2}{30 \cdot 40} + \frac{18^2}{30 \cdot 35} - 1 \right] = \\ &= 100 \left[\frac{144}{500} + \frac{36}{800} + \frac{4}{700} + \frac{100}{1250} + \frac{625}{2000} + \frac{225}{1750} + \frac{9}{750} + \frac{81}{1200} + \frac{324}{1050} - 1 \right] = \\ &= 100 [0,288 + 0,045 + 0,005 + 0,08 + 0,312 + 0,129 + 0,012 + 0,068 + 0,309 - 1] = 24,8. \end{aligned}$$

За одержаними розрахунками, фактичне значення $\chi^2 = 24,8$.

Порівняємо одержаний результат із критичним значенням для $\alpha = 0,05$. З урахуванням числа ступенів свободи k . При $k = (3-1)(3-1) = 4$ критичне значення $\chi^2_{0,095(4)} = 9,49$. Отже, з імовірністю 0,95 нами доведена істотність зв'язку між віком матері й кількістю народжених дітей.

Відносною мірою щільності стохастичного зв'язку може слугувати також розрахунок **коефіцієнтів співзалежності Чупрова та Крамера**.

Коефіцієнт Чупрова

$$C = \sqrt{\frac{\chi^2}{n \sqrt{(m_1 - 1)(m_2 - 1)}}}, \quad (7.23).$$

де n - число елементів сукупності;

m_1 і m_2 - число груп за ознакою "x" і ознакою "y".

$$C = \sqrt{\frac{24,8}{100\sqrt{(3-1)(3-1)}}} = \sqrt{\frac{24,8}{100\sqrt{4}}} = \sqrt{\frac{24,8}{0,124}} = 0,352.$$

У випадку коли ознаки є незалежними, то і χ^2 і C дорівнюють нулю. При функціональному зв'язку коефіцієнт співзалежності Чупрова дорівнює одиниці лише за умови рівності числа груп, тобто $m_1 = m_2$. Нерівність $m_1 \neq m_2$ призведе, що $C < 1$ навіть при функціональному зв'язку. Коефіцієнт співзалежності при цьому розраховують за допомогою формули Крамера:

$$C = \sqrt{\frac{\chi^2}{n(m_{\min} - 1)}}, \text{ де } m_{\min} - \text{мінімальне число груп (} m_1 \text{ або } m_2 \text{)}.$$

З цього виходить, що коли $m_1 = m_2$ - значення коефіцієнтів, обчислені за формулами Чупрова і Крамера збігаються.

В нашому прикладі (табл. 7.7) $m_1 = m_2 = 3$, а тому результат буде однаковим:

$$C = \sqrt{\frac{24,8}{100(3-1)}} = \sqrt{0,124} = 0,352. \text{ Одержаний результат більший від}$$

нуля, що засвідчує про наявність стохастичного зв'язку. Розрахунок коефіцієнта співзалежності Крамера при умові, що $m_1 \neq m_2$ проведемо, використавши дані таблиці 7.5. Порівняємо емпіричні частоти з теоретичними.

Теоретичні частоти для першого рядка будуть становити:

$$f'_{11} = \frac{42 \cdot 55}{137} = 17; f'_{12} = \frac{42 \cdot 50}{137} = 15; f'_{13} = \frac{42 \cdot 32}{137} = 10.$$

$$\text{Для другого рядка: } f'_{21} = \frac{39 \cdot 55}{137} = 16; f'_{22} = \frac{39 \cdot 50}{137} = 14; f'_{23} = \frac{39 \cdot 32}{137} = 9.$$

$$\text{Для третього рядка: } f'_{31} = \frac{33 \cdot 55}{137} = 13; f'_{32} = \frac{33 \cdot 50}{137} = 12; f'_{33} = \frac{33 \cdot 32}{137} = 8.$$

$$\text{Для четвертого рядка: } f'_{41} = \frac{15 \cdot 55}{137} = 6; f'_{42} = \frac{15 \cdot 50}{137} = 6; f'_{43} = \frac{15 \cdot 32}{137} = 3.$$

$$\text{Для п'ятого рядка: } f'_{51} = \frac{8 \cdot 55}{137} = 3; f'_{52} = \frac{8 \cdot 50}{137} = 3; f'_{53} = \frac{8 \cdot 32}{137} = 2.$$

Знайдемо абсолютне значення χ^2 , використавши розрахункові дані. Побудуємо табл. 7.8

Таблиця 7.8

Розрахунок коефіцієнта співзалежності Крамера

| Групи випускників за оцінками незалежного тестування, x | Кількість випускників, що мають намір вступити до навчального закладу, f | | | | | | Разом |
|---|--|-----------|-----------------|-----------|----------|-----------|-------|
| | державного вузу | | приватного вузу | | коледжу | | |
| | f_{ij} | f'_{ij} | f_{ij} | f'_{ij} | f_{ij} | f'_{ij} | |
| відмінні | 30 | 17 | 10 | 15 | 2 | 10 | 42 |
| переважно відмінні | 20 | 16 | 15 | 14 | 4 | 9 | 39 |
| переважно добрі | 5 | 13 | 20 | 12 | 8 | 8 | 33 |
| добрі | - | 6 | 5 | 6 | 10 | 3 | 15 |
| добрі і заловільні | - | 3 | - | 3 | 8 | 2 | 8 |
| Разом | 55 | 55 | 50 | 50 | 32 | 32 | 137 |

$$\chi^2 = \sum_i \sum_j \frac{(f_{ij} - f'_{ij})^2}{f'_{ij}} = \frac{(30-17)^2}{17} + \frac{(10-15)^2}{15} + \frac{(2-10)^2}{10} + \frac{(20-16)^2}{16} + \frac{(15-14)^2}{14} +$$

$$+ \frac{(4-9)^2}{9} + \frac{(5-13)^2}{13} + \frac{(20-12)^2}{12} + \frac{(8-8)^2}{8} + \frac{(0-6)^2}{6} + \frac{(5-6)^2}{6} + \frac{(10-3)^2}{3} + \frac{(0-3)^2}{3} +$$

$$+ \frac{(0-3)^2}{3} + \frac{(8-2)^2}{2} = 77.$$

Звідси коефіцієнт співзалежності Крамера буде дорівнювати:

$$C = \sqrt{\frac{\chi^2}{n(m_{\min} - 1)}} = \sqrt{\frac{77}{137(3-1)}} = \sqrt{0,281} = 0,530.$$

Перевіримо істотність зв'язку. Число ступенів свободи становить:

$$K = (m_1 - 1) \cdot (m_2 - 1) = (5 - 1) \cdot (3 - 1) = 8.$$

Критичне значення χ^2 для рівня істотності $\alpha = 0,05$ і $K = 8$ становить $\chi^2_{0,095(8)} = 15,51$. Фактичне значення дорівнює 77, що значно більше за критичне. Істотність зв'язку між оцінками незалежного тестування і наміром абітурієнтів піти вступити вчитися до вузів різних форм власності та акредитації доведено.

Тіснота зв'язку альтернативних ознак визначається за допомогою **коефіцієнта асоціації (конвергенції)**. При цьому використовують таку формулу:

$$A = \frac{ad - bc}{\sqrt{(a+b) \cdot (c+d) \cdot (a+c) \cdot (b+d)}}. \quad (7.24).$$

Приклад. Встановити взаємозв'язок між відвідуванням студентами занять і успішністю (табл.7.9)

Таблиця 7.9

Дані для розрахунку коефіцієнта асоціації

| Показники | Кількість студентів, які заняття: | | Всього |
|----------------------------|-----------------------------------|----------------|--------|
| | відвідували | не відвідували | |
| Склали екзамени успішно | 86 | 14 | 100 |
| Не склали екзамени успішно | 22 | 28 | 50 |
| Всього | 108 | 42 | 150 |

$$A = \frac{86 \cdot 28 - 22 \cdot 14}{\sqrt{108 \cdot 42 \cdot 100 \cdot 50}} = \frac{2100}{\sqrt{22680000}} = 0,441.$$

Розрахований коефіцієнт асоціації засвідчує про слабкий зв'язок між ознаками, що досліджувалися.

За змістом коефіцієнт асоціації є ідентичним коефіцієнту взаємної спряженості (залежності), а з χ^2 пов'язаний функціонально: $\chi^2 = nC^2$. Це впливає з наступного:

$$\chi^2 = n \frac{(f_{11}f_{22} - f_{12}f_{21})^2}{f_{01}f_{02}f_{10}f_{20}}, \quad (7.25).$$

$$C = \sqrt{\frac{\chi^2}{n}} = \frac{f_{11}f_{22} - f_{12}f_{21}}{\sqrt{f_{01}f_{02}f_{10}f_{20}}}, \quad \text{звідки} \quad (7.26).$$

$$\chi^2 = nC^2 = 150 \cdot 0,441^2 = 150 \cdot 0,194 = 29. \quad \text{Критичне } \chi^2_{0,95(1)} = 3,89.$$

Таким чином істотність зв'язку доведено з імовірністю 0,95. Наведені вище коефіцієнти використовуються для оцінювання узгодженості варіації атрибутивних ознак.

Часто в практиці статистичного дослідження взаємозв'язок між ознаками можна встановити за допомогою **рангового**

коефіцієнта кореляції Спірмена ρ . Його розрахунок проводиться для ознак, значення яких слід розмістити по відповідному рангу.

Розрахунок коефіцієнта здійснюється за формулою:

$$\rho = 1 - \frac{6 \sum d^2}{n(n^2 - 1)}, \quad (7.27).$$

де ρ - ранговий коефіцієнт кореляції;

$d = R_x - R_y$ - різниця рангів між ознаками;

n - число значень ознаки.

Коефіцієнт кореляції рангів може набувати значень від 0 до ± 1 . Якщо дві ознаки рангів повністю збігаються, то $\sum d^2 = 0$. В цьому випадку говорять про повний функціональний зв'язок.

Приклад розрахунку рангового коефіцієнта кореляції Спірмена наведемо за допомогою інформації табл. 7.10.

Таблиця 7.10

Розрахунок коефіцієнта рангової кореляції Спірмена

| Рівень забезпеченості підприємств сировиною, млн. грн., x | Обсяг виготовленої продукції, млн. грн., y | R_x | R_y | $d = R_x - R_y$ | d^2 |
|---|--|----------|----------|-----------------|-------------------|
| 1.5 | 3.9 | 1 | 3 | -2 | 4 |
| 1.5 | 4.4 | 2 | 5 | -3 | 9 |
| 2.0 | 3.8 | 3 | 2 | 1 | 1 |
| 2.2 | 3.5 | 4 | 1 | 3 | 9 |
| 2.3 | 4.8 | 5 | 6 | -1 | 1 |
| 2.6 | 4.3 | 6 | 4 | 2 | 4 |
| 3.0 | 7.0 | 7 | 9 | -2 | 4 |
| 3.1 | 6.5 | 8 | 8 | 0 | 0 |
| 3.5 | 6.1 | 9 | 7 | 2 | 4 |
| 3.8 | 8.2 | 10 | 10 | 0 | 0 |
| Σ | Σ | Σ | Σ | Σ | $(\Sigma d^2=36)$ |

$$R = 1 - \frac{6 \cdot 36}{10 \cdot 99} = 0,782.$$

Розрахований за даними таблиці 7.10 ранговий коефіцієнт кореляції Спірмена свідчить про тісний зв'язок між рівнем забезпеченості підприємств сировиною та обсягом виготовленої продукції.

Порівняємо фактичне значення розрахункового коефіцієнта з критичним (табл. 7.11) для рівня істотності $\alpha = 0,05$ і $n = 10$.
 $P_{0,95(10)} = 0,563$.

Таблиця 7.11

Критичні значення коефіцієнта рангової кореляції Спірмена при $\alpha = 0,05$

| Обсяг вибірки n | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|----------------------|-------|-------|-------|-------|-------|-------|-------|-------|
| $P_{0,95(n)}$ | 0,900 | 0,828 | 0,714 | 0,642 | 0,600 | 0,563 | 0,527 | 0,497 |

Таким чином, з імовірністю 0,95 істотність зв'язку доведено.

7.4 Поняття множинної кореляції

Парна кореляція або парна регресія можуть розглядатись як окремий випадок відображення зв'язку деякої залежної змінної, з однієї сторони, та однієї з багатьох незалежних змінних - з другої. Коли ж потрібно охарактеризувати зв'язок всієї множини незалежних змінних із результативною ознакою, то ведуть мову про множинну кореляцію або множинну регресію.

Для встановлення зв'язку при багатофакторній кореляції використовують багатофакторні рівняння, котрі приводять до лінійного виду за результатами певних перетворень, тобто

$$Y = a_0 + a_1x_1 + a_2x_2 + \dots + a_mx_m. \quad (7.28).$$

Параметри $a_0, a_1, a_2, \dots, a_m$ визначають шляхом розв'язання системи нормальних рівнянь.

Тестові завдання

1. У статистиці розрізняють такі види зв'язку:

- а) функціональні та стохастичні;
- б) функціональні та кореляційні;
- в) факторні та результативні;
- г) повні та неповні.

2. При функціональному зв'язку кожному значенню факторної ознаки відповідає:

- а) декілька значень результативної ознаки;
- б) одне значення результативної ознаки;
- в) значення результативної ознаки, яке дорівнює нулю;
- г) усі відповіді вірні.

3. Стохастичний – це зв'язок, при якому кожному значенню “х” відповідає:

- а) одне значення “у”;
- б) два значення “у”;
- в) декілька можливих значень “у”, які варіюють і утворюють розподіл за даною ознакою;
- г) одне фіксоване значення результативної ознаки “у”.

4. Кореляційний зв'язок є:

- а) повним функціональним;
- б) середнім;
- в) умовним;
- г) неповним, що виявляється при великій кількості випробувань і лише через середні величини.

5. За напрямками кореляційні зв'язки бувають:

- а) повними;
- б) неповними ;
- в) прямими та оберненими;
- г) функціональними.

6. Прямолінійна форма зв'язку використовується у випадках, коли:

- а) зі зміною ознаки “х” адекватно змінюється ознака “у”;

- б) зі зміною ознаки “х” значення ознаки “у” змінюється не адекватно;
- в) між “х” та “у” відсутній зв’язок;
- г) значення “х” зростає, а “у” зменшується.

7. Оцінка тісноти зв’язку при парній кореляції забезпечується за допомогою розрахунку:

- а) параметрів a_0 та a_1 в прямолінійному рівнянні;
- б) коефіцієнта кореляції;
- в) коефіцієнта множинної кореляції;
- г) параметрів a_0 та a_1 в криволінійному рівнянні.

8. Оцінка щільності нелінійного зв’язку ґрунтується на співвідношенні варіацій:

- а) емпіричних і теоретичних значень факторної ознаки;
- б) теоретичних і емпіричних значень результативної ознаки;
- в) емпіричних і вторинних значень результативної ознаки;
- г) теоретичних і емпіричних значень досліджуваної сукупності.

9. Відношення міжгрупової факторної дисперсії до загальної називається:

- а) кореляційним рівнянням;
- б) коефіцієнтом кореляції;
- в) кореляційним відношенням;
- г) коефіцієнтом детермінації.

10. Корінь квадратний із коефіцієнта детермінації називається:

- а) індексом кореляції;
- б) коефіцієнтом кореляції;
- в) кореляційним відношенням;
- г) кореляційним рівнянням.

11. Похибка коефіцієнта регресії залежить від розміру варіації:

- а) результативної ознаки, залишкової дисперсії та числа ступенів свободи;
- б) факторної ознаки залишкової дисперсії та числа ступенів свободи;
- в) факторної ознаки, міжгрупової дисперсії та числа ступенів свободи;

г) результативної ознаки, залишкової дисперсії та числа ступенів свободи.

12. Коефіцієнт взаємного узгодження базується на:

- а) розбіжностях варіант умовного та безумовного розподілів;
- б) розбіжностях умовного та безумовного розподілів досліджуваної ознаки;
- в) розбіжностях частот (часток) умовного та безумовного розподілів;
- г) розподілі частот стохастичного зв'язку.

13. Частки умовних і безумовних розподілів збігаються за умови, що ознаки:

- а) тісно пов'язані між собою;
- б) частково пов'язані між собою;
- в) перебувають у функціональному зв'язку;
- г) незалежні.

14. Коефіцієнт кореляції рангів може набувати значень:

- а) від 0 до ± 1 ;
- б) від 0 до $+ 1$;
- в) від 0 до -1 ;
- г) 1 і більше.

15. Ознака зв'язку всієї множини незалежних змінних з результативною ознакою забезпечується шляхом використання:

- а) парної кореляції;
- б) множинної кореляції;
- в) часткової кореляції;
- г) непараметричних методів зв'язку.

Задача 1. Відомі дані про вартість основних виробничих засобів та випуск продукції по 10 підприємствах:

| № з/п | Вартість основних виробничих засобів, млн. грн. (X) | Випуск продукції, млн. грн., (Y) | № з/п | Вартість основних виробничих засобів, млн. грн. (X) | Випуск продукції, млн. грн., (Y) |
|--------|---|----------------------------------|-------|---|----------------------------------|
| 1 | 5,3 | 5,8 | 11 | 3,5 | 4,6 |
| 2 | 6,4 | 7,6 | 12 | 3,2 | 1,6 |
| 3 | 7,9 | 8,7 | 13 | 3,3 | 3,2 |
| 4 | 8,3 | 9,1 | 14 | 3,3 | 2,7 |
| 5 | 9,2 | 11,9 | 15 | 3,7 | 2,7 |
| 6 | 10,1 | 12,3 | 16 | 4,2 | 3,0 |
| 7 | 12,5 | 13,8 | 17 | 1,2 | 1,8 |
| 8 | 13,0 | 14,0 | 18 | 7,2 | 13,1 |
| 9 | 14,6 | 15,2 | 19 | 6,8 | 12,1 |
| 10 | 15,7 | 17,6 | 20 | 2,2 | 2,7 |
| Разом: | 103,0 | 116,0 | x | x | x |

1. Побудувати рівняння кореляційної залежності між вартістю основних виробничих засобів підприємства і обсягом виготовленої продукції.
2. Розрахувати коефіцієнт кореляції.
3. Оцінити надійність коефіцієнта кореляції.
4. На основі одержаних даних зробити висновки.

Задача 2. Задані дані про стаж роботи та розмір місячної заробітної плати працівників промислового підприємства:

| Номер працівника | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|--|------|------|------|------|------|------|------|------|------|------|
| Стаж роботи, років | 16,0 | 3,5 | 6,0 | 5,0 | 4,0 | 9,5 | 10,0 | 8,0 | 7,0 | 6,5 |
| Розмір місячної заробітної плати, грн. | 2500 | 1360 | 1440 | 1480 | 1400 | 1800 | 1910 | 1760 | 1670 | 1630 |

1. Побудувати рівняння кореляційного зв'язку між стажем роботи та розміром місячної заробітної плати (зв'язок вигляді параболі другого порядку).

2. Розрахувати кореляційне відношення та критерій його надійності.
3. Зробити висновки.

Задача 3. Приводяться дані щодо обсягу реалізованої продукції (x) та рівня накладних витрат у зв'язку з її реалізацією (y):

| № підприємства | Обсяг реалізованої продукції, x, млн. грн. | Обсяг накладних витрат, y, тис. грн. |
|----------------|--|--------------------------------------|
| 1 | 12,0 | 46,2 |
| 2 | 18,8 | 93,9 |
| 3 | 11,0 | 90,6 |
| 4 | 29,0 | 110,8 |
| 5 | 17,5 | 87,2 |
| 6 | 23,4 | 76,5 |
| 7 | 35,6 | 136,8 |
| 8 | 30,2 | 81,0 |
| 9 | 26,1 | 99,8 |
| 10 | 20,7 | 80,4 |

Розрахувати ранговий коефіцієнт кореляції Спірмена.

Задача 4. Відомі такі дані про виконання норм виробітку робітниками-відрядниками та їхню кваліфікацію:

| Група робітників-відрядників | Виконують і перевиконують норми виробітку | Не виконують норм виробітку | Всього |
|------------------------------|---|-----------------------------|--------|
| Закінчили ПТУ | 86 | 24 | 100 |
| Не мають спеціальної освіти | 48 | 52 | 100 |
| Всього | 134 | 66 | 200 |

Для оцінювання тісноти між кваліфікацією та виконанням норм виробітку визначіть коефіцієнти асоціації.

Задача 5. Відомі такі дані про розподіл страхувальників життя за страховими сумами і середньомісячною заробітною платою:

| Страхова сума, | Середньомісячна заробітна плата, грн. |
|----------------|---------------------------------------|
|----------------|---------------------------------------|

| грн. | до 1500 | 1500 - 2000 | 2000 і більше |
|---------------|---------|-------------|---------------|
| До 3000 | 171 | 63 | 21 |
| 3000 – 5000 | 109 | 332 | 75 |
| 5000 і більше | 65 | 147 | 415 |

Для оцінювання тісноти зв'язку між середньомісячною заробітною платою і страховою сумою обчисліть коефіцієнт співзалежності Чупрова. Зробіть висновки.

Задача 6. Є такі дані розподілу робітників за рівнем освіти і задоволеністю роботою.

| Групи робітників за рівнем освіти | Чисельність робітників, що дали відповідь на запитання щодо задоволеності роботою, % до підсумку | | | | | |
|-----------------------------------|--|-------------------------------------|----------------|-------------------------------------|------------|-------|
| | не задоволені | швидше незадоволені, ніж задоволені | значення немає | швидше задоволені, ніж незадоволені | задоволені | разом |
| початкова | 0,050 | 0,100 | 0,150 | 0,367 | 0,333 | 1,0 |
| н/середня | 0,133 | 0,092 | 0,173 | 0,337 | 0,265 | 1,0 |
| середня | 0,163 | 0,402 | 0,120 | 0,185 | 0,130 | 1,0 |
| | 0,124 | 0,208 | 0,148 | 0,288 | 0,232 | 1,0 |

Розрахувати коефіцієнти співзалежності Чупрова та Крамера.

Питання для самоконтролю

1. Поняття факторної ознаки.
2. Тлумачення результативної ознаки.

3. Види зв'язку, котрі розраховуються статистикою.
4. Різниця між функціональним та стохастичним видами зв'язку.
5. Змістове навантаження терміну «кореляція».
6. Характерні риси кореляційного зв'язку.
7. Основні риси парної кореляції.
8. Основні завдання оцінювання кореляційної залежності.
9. Форми зв'язку.
10. Показники, котрі використовуються для визначення тісного зв'язку.
11. Щільність зв'язку.
12. Показники, котрі використовуються для вимірювання щільності нелінійного зв'язку.
13. Мета використання коефіцієнта детермінації.
14. Різниця між коефіцієнтом детермінації і індексом кореляції.
15. Перевірка істотності зв'язку.
16. Різниця між парною й множинною кореляцією.
17. Випадки використання коефіцієнта взаємного узгодження.
18. Випадки використання непараметричних методів оцінки зв'язку.
19. Основне призначення коефіцієнтів співзалежності Чупрова та Крамера.
20. Випадки використання коефіцієнта кореляції.

РОЗДІЛ 8. АНАЛІЗ ІНТЕНСИВНОСТІ ДИНАМІКИ

8.1. Загальна характеристика рядів динаміки

Об'єктивність досліджуваних явищ і процесів забезпечується статистикою завдяки оцінюванню їх крізь призму часу. Інакше кажучи, пізнання предмета здійснюється в динаміці шляхом розбудови й аналізу динамічних рядів.

Динамічний, хронологічний, часовий ряд - це послідовність упорядкованих числових показників, які характеризують рівень розвитку явищ і процесів у часі. Кожний динамічний ряд складається з двох обов'язкових елементів: періодів часу (t) та рівнів ряду (y).

За часом динамічні ряди поділяються на інтервальні (періодичні) й моментні. Інтервальний (періодичний) ряд показує розвиток явища за певний проміжок часу. Рівні моментного ряду характеризують значення показника на певний момент часу. За допомогою інтервальних рядів досліджуються обсяги виробництва продукції за місяцями року, валовий збір сільськогосподарських культур за роками, динаміка валового внутрішнього продукту тощо. Прикладом моментного динамічного ряду можуть бути дані щодо залишків депозитів населення в комерційних банках на конкретну дату кожного місяця року, наявність товарно-матеріальних цінностей за результатами інвентаризації, залишки грошей в касі на кінець дня, місяця тощо.

Рівні моментних рядів підсумовувати не можна, оскільки такий одержаний результат позбавлений змісту. Наприклад, якщо на початок року в комерційному банку загальне число депозитних вкладів складало три тисячі, а на кінець - чотири, то загальна сума цих показників (сім тисяч) позбавлена логічного змісту, оскільки депозитних рахунків стало не сім тисяч, а лише чотири, тобто протягом року кількість вкладів зросла на одну тисячу.

На відміну від моментних, рівні інтервальних рядів можна підсумувати, оскільки одержаний підсумок характеризує зміст реального показника.

Динамічні ряди поділяють також на первинні, похідні, ряди абсолютних, середніх і відносних величин. Об'єктивна статистична оцінка інформації, котра задається динамічними рядами забезпечується завдяки дотриманню певних передумов. Передусім це стосується порівнянності за територією, колом об'єктів, одиницями виміру, часом реєстрації, цінами, методологією розрахунку тощо.

Порівнянність рівнів динамічних рядів є однією з найважливіших передумов досягнення об'єктивності висновків, отриманих у результаті статистичного дослідження. Розглянемо більш предметно окремі випадки порівнянності рівнів динамічного ряду.

Порівнянність за територією вказує на потребу врахування змін, які стосуються кількісних та якісних ознак досліджуваних об'єктів. Зокрема, при оцінці динаміки двох рядів, які засвідчують розвиток одного і того ж явища в старих і нових адміністративних границях за одним і тим же колом об'єктів, то їх порівняння досягається завдяки використанню прийому змикання динамічних рядів. Приведемо приклад розрахунку. Кількість іноземних громадян, які відвідали Україну до тимчасової окупації території АР Крим і м. Севастополя та після того характеризується такими даними (див. табл. 8.1).

Потрібно побудувати динамічний ряд числа туристів без урахування територіальних змін.

Таблиця 8.1

Кількість іноземних громадян, які відвідали Україну
(млн. осіб)

| | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 |
|---|------|------|------|------|------|------|
| З урахуванням території АР Крим та м. Севастополя | 21,2 | 21,4 | 23,0 | 24,8 | | |
| Без урахування території тимчасово окупованої АР Крим та м. Севастополя | | | | 13,5 | 12,7 | 12,4 |

В даному прикладі використання емпіричних даних рівнів динамічних рядів у приведеному вигляді буде помилковим і методично невірним.

З метою порівнянності числа туристів за територією використовуємо спосіб змикання рядів. Для цього рівні ряду за 2010-2015рр. множать на коефіцієнт змикання (перерахунку) динамічних рядів:

$$K_z = \frac{P_n}{P_c}, \quad (8.1).$$

де K_z - коефіцієнт змикання динамічного ряду;

P_c - рівень показника в старих межах;

P_n - рівень показника в нових межах.

В нашому прикладі розмір коефіцієнта змикання становитиме:

$$\hat{E}_z = \frac{24,8}{13,5} = 1,83.$$

Побудуємо новий ряд з врахуванням розміру коефіцієнта змикання в результаті чого одержимо рівні:

в 2014р. $12,7 \cdot 1,83 = 23,3$ млн. осіб;

в 2015р. $12,4 \cdot 1,83 = 22,7$ млн. осіб;

В результаті проведених розрахунків динамічний ряд числа іноземних туристів без урахування змін території набуде такого вигляду (табл.8.2).

Таблиця 8.2

Динаміка кількості іноземних громадян, які відвідали Україну
(млн.осіб)

| | 2010р. | 2011р | 2012р. | 2013р. | 2014р. | 2015р. |
|---|--------|-------|--------|--------|--------|--------|
| Кількість іноземних громадян, які відвідали Україну | 21,2 | 21,4 | 23,0 | 24,8 | 23,3 | 22,7 |

Крім вимог, які стосуються порівнянності рівнів динамічного ряду за територією, порівнянність має забезпечуватися також за іншими критеріями. Зокрема, це стосується кола об'єктів. Йдеться про те, що в розрахунок повинні прийматися сукупності з однаковим числом одиниць.

Порівнянність за часом потребує використання однакових часових відрізків між рівнями ряду. Умови порівнянності забезпечуються також при використанні інтервальних або моментних рядів. Вартісна порівнянність динамічних рядів досягається завдяки використанню порівняльних цін.

Часом виникає потреба одержати узагальнюючу характеристику ознаки динамічного ряду. Такою величиною є середня.

Розрахунок тієї чи іншої середньої динамічного ряду залежить від його виду: інтервальний, дискретний або моментний. Залежно від величини інтервалів розміру проміжків часу в інтервальному варіаційному ряді узагальнена величина обчислюється з використанням формул середньої арифметичної простої або зваженої.

$$\bar{y} = \frac{y_1 + y_2 + \dots + y_n}{n} = \frac{\sum y}{n} - \text{проста.}$$

(8.2).

$$\bar{y} = \frac{y_1 t_1 + y_2 t_2 + \dots + y_n t_n}{t_1 + t_2 + \dots + t_n} = \frac{\sum y t}{\sum t} \text{ зважена,}$$

(8.3).

де y_n - рівні ряду, t - періоди часу.

Розрахунок середньої для моментного ряду відрізняється від інтервального, оскільки рівні досліджуваної ознаки задаються на певний момент часу.

$$\bar{y} = \frac{\frac{1}{2}y_1 + y_2 + \dots + \frac{1}{2}y_n}{n-1}, \quad (8.4),$$

де n число рівнів ряду (середня хронологічна).

Приведемо приклад розрахунку середніх, які найбільш часто використовуються під час статистичного дослідження динамічних рядів.

Обсяг виробництва продукції А на підприємстві за місяцями поточного року характеризується наступним чином (табл. 8.3).

Таблиця 8.3

Обсяги оптового товарообороту підприємств продовольчих товарів, млн. грн.

| Місяць | 2009р. | 2010р. | 2011р. | 2012р. | 2013р. | 2014р |
|----------------------------|--------|--------|--------|--------|--------|--------|
| Обсяги товарообороту, млн. | 168299 | 193849 | 194516 | 215169 | 199674 | 194669 |

Визначити обсяг виробництва продукції А в середньому за перше півріччя поточного року:

$$\bar{y} = \frac{\sum y}{n} = \frac{168299 + 193849 + 194516 + 215169 + 199674 + 194669}{6} = \frac{1166175}{6} = 194362 \text{ млн.}$$

грн.

Узагальнюючий рівень динамічного ряду розраховано з використанням формули середньої арифметичної простої.

Інший приклад. Обсяг реалізації цукру підприємством за місяцями поточного року задається наступним чином (табл. 8.4).

Таблиця 8.4

Обсяг реалізації цукру за місяцями року, тис. грн.

| | |
|----------------------|-----------------------|
| Січень - 200 | Квітень-липень – 210 |
| Лютий-березень – 215 | Серпень-грудень – 220 |

Визначити середній рівень обсягу реалізації цукру за рік

$$\bar{y} = \frac{\sum yt}{t} = \frac{200 \cdot 1 + 215 \cdot 2 + 210 \cdot 4 + 220 \cdot 5}{12} = \frac{2570}{12} = 214,1 \text{ тис.грн.}$$

Середньомісячне виробництво продукції за рік нами визначене за формулою середньої арифметичної зваженої, оскільки рівні ряду (ваги) задаються за неоднакові проміжки часу.

Приведемо приклад розрахунку середньої для моментного ряду. Залишок товарних запасів торгівельної бази на початок кожного місяця першого кварталу становили (див. табл. 8.5)

Таблиця 8.5

Залишки товарів на початок кожного місяця кварталу

| Місяць | На 1.01. | на 1.02. | на 1.03. | на 1.04. |
|---------------------------------|----------|----------|----------|----------|
| Обсяг запасу товарів, млн. грн. | 1400 | 1550 | 1270 | 1600 |

Визначити середні залишки товарів торгівельної бази в I-му кварталі.

Для розрахунку середніх запасів за досліджуваний період використовуємо формулу середньої хронологічної:

$$\bar{y} = \frac{\frac{1}{2}y_1 + y_2 + \dots + \frac{1}{2}y_n}{n-1} = \frac{\frac{1400}{2} + 1550 + 1270 + \frac{1600}{2}}{4-1} = 1400 \text{ млн. грн.}$$

Ряд динаміки, що характеризує зміну лише однієї ознаки, називається **одномірним**; двох і більше ознак - **багатомірним**. Багатомірні ряди поділяються на **паралельні** та **взаємозв'язаних показників**.

Перші (паралельні) характеризують динаміку одного і того ж показника для різних об'єктів (чисельність робітників, зайнятих в різних сферах діяльності), або різних показників щодо одного і того ж об'єкта (фінансове становище суб'єкта підприємницької діяльності).

Статистичне дослідження динаміки передбачає також оцінювання взаємозв'язку між ознаками. В багатомірному динамічному ряді він буває функціональним (адитивним, мультиплікативним) або кореляційним. Прикладом адитивного зв'язку в динамічних рядах може слугувати динаміка цілого і його окремих складових частин (загальний обсяг виробництва комп'ютера

та виробництво його комплектуючих). Мультиплікативність динамічного ряду характеризується взаємозв'язком між розміром фонду заробітної плати, чисельністю працюючих та середньомісячною зарплатою. Прикладом кореляційного зв'язку є взаємозалежність технічного рівня виробничого процесу й продуктивності праці за ряд років.

Використання в процесі статистичного дослідження динамічних фондів зумовлює розв'язання ряду завдань. При цьому до основних із них належить встановлення закономірностей соціально-економічного розвитку, явищ і процесів, які досліджуються статистикою; виявлення і описання тенденцій; оцінювання структурних зрушень; характеристика сталості та коливань рівнів рядів; статистичне дослідження факторів соціально-економічного економічного розвитку тощо.

8.2. Найважливіші характеристики динамічних рядів

Статистичне дослідження динамічних рядів забезпечується завдяки застосування системи показників, завдяки яким забезпечується всебічна характеристика напряму й інтенсивності змін явищ у часі. Найважливішими показниками оцінювання ряду динаміки слугують розрахунки:

- 1) абсолютного приросту;
- 2) коефіцієнта росту (темпер зростання (зниження)
- 3) темпу приросту;
- 4) абсолютного значення одного відсотка приросту.

Абсолютний приріст - це різниця між даним рівнем і тим, із яким порівнюють. Порівняння лише з одним рівнем, який прийнято за базу, дає можливість одержати базові показники динаміки. Порівняння кожного наступного рівня з попереднім забезпечує розрахунок ланцюгових показників динаміки, тобто:

$$\Delta_y = y_t - y_0 - \text{базові абсолютні прирости}; \quad (8.5).$$

$$\Delta_y = y_t - y_{t-1} - \text{ланцюгові абсолютні прирости}, \quad (8.6).$$

де y_t - даний рівень ряду, y_{t-1} - попередній рівень ряду, y_0 - базовий рівень ряду.

Коефіцієнт росту також розраховується базовим і ланцюговим способами. Перший – розраховують як відношення кожного наступного рівня до того, що прийнятий за базу. Ланцюговий одержують шляхом ділення кожного наступного рівня до попереднього. Коефіцієнт росту позначають буквою K і вимірюють у коефіцієнтах або відсотках. Ідентифікація коефіцієнта залежить від одержаного результату. За умови, що він більший за одиницю або ста відсотків, то це засвідчує про зростання рівня, і навпаки. Розрахунок коефіцієнтів зростання здійснюється за допомогою таких формул:

$$K_p = \frac{y_t}{y_0} - \text{базові коефіцієнти росту}; \quad (8.7).$$

$$K_p = \frac{y_t}{y_{t-1}} - \text{ланцюгові коефіцієнти росту}. \quad (8.8).$$

Ланцюгові та базові коефіцієнти росту знаходяться між собою в мультиплікативному зв'язку, тобто базовий коефіцієнт зростання можна одержати як добуток ланцюгових.

$$K_1 \cdot K_2 \cdot K_3 \cdot \dots \cdot K_n = \overset{t}{D}, \quad (8.9).$$

де $K_1 \cdot K_2 \cdot K_3 \cdot \dots \cdot K_n$ - ланцюгові коефіцієнти зростання, $\overset{t}{D}$ - умовне позначення добутку чисел.

Темп приросту розраховують шляхом ділення абсолютного приросту до попереднього або початкового рівня. Одержаний результат множать на 100%, тобто

$$T = \frac{\Delta_t}{y_0} \cdot 100 - \text{базові темпи приросту}; \quad (8.10).$$

$$T = \frac{\Delta_t}{y_{t-1}} \cdot 100 - \text{ланцюгові темпи приросту}. \quad (8.11).$$

Більш практичне використання має розрахунок темпу приросту як різницю між коефіцієнтом росту і «1». Якщо він обрахований у відсотках, тоді віднімають сто і одержаний результат множать на 100%.

Абсолютне значення одного відсотка приросту - це відношення абсолютно-го приросту за аналізований період до темпу приросту за цей самий період, тобто

$$A_t = \frac{\Delta_t}{T_t}. \quad (8.12).$$

Величина одержаного результату засвідчує вагомість одного відсотка приросту. Його можна визначити також шляхом ділення абсолютного рівня попереднього періоду на 100. Наведемо приклад розрахунку вищенаведених показників динамічного ряду (табл. 8.6).

Таблиця 8.6

До розрахунку показників динамічного ряду обсягу інвестицій

| Порядковий номер року, t | Обсяг інвестицій, y_t , млн. грн. | Абсолютний приріст | | Коефіцієнти зростання | | Темпи приросту, % | | Абсолютне значення одного відсотка приросту, млн. грн |
|----------------------------|-------------------------------------|--------------------|---------|-----------------------|---------|-------------------|---------|---|
| | | ланцюговий | базовий | ланцюговий | базовий | ланцюговий | базовий | |
| 1 (базовий) | 80 | - | - | - | 1,0 | - | - | - |
| 2 | 150 | 70 | 70 | 1,875 | 1,875 | 87,5 | 87,5 | 0,8 |
| 3 | 200 | 50 | 120 | 1,333 | 2,500 | 33,3 | 1,5 | 1,5 |
| 4 | 250 | 50 | 170 | 1,250 | 3,125 | 25,0 | 2,125 | 2,0 |

Розраховані нами показники динаміки в таблиці 8.6 засвідчують про зростання динаміки обсягу інвестицій.

У практиці статистичного дослідження інколи виникає потреба порівняти інтенсивність розвитку ознак у різних рядах. З цією метою розраховують коефіцієнт випередження, як відношення коефіцієнта зростання даного ряду до порівнюваного. Наприклад, щільність населення в розрахунку на 1 км² між двома переписами населення зросла в одному регіоні на 35, в іншому - на 30 %.

Коефіцієнт випередження темпу зростання щільності населення в першому регіоні порівняно з другим становить $1,35 : 1,30 = 1,038$.

Взаємозв'язок двох ознак динамічних рядів можна охарактеризувати використовуючи значення темпів приросту. Мова йде про застосування **коефіцієнта еластичності**, котрий показує, на скільки відсотків змінюється результативна ознака у при зміні факторної x на один відсоток. Наприклад, продуктивність праці зросла на 4%, а розмір заробітної плати - на 2 %. Поділивши відсоток приросту заробітної плати на відповідний приріст продуктивності праці, отримаємо: $2:4 = 0,5$. Це означає, що при зростанні продуктивності праці на один відсоток, заробітна плата збільшується на 0,5%.

Для узагальнення оцінок швидкості та інтенсивності розвитку ознак у часі розраховують середні значення показників ряду.

Завдяки такій властивості як адитивність **середній абсолютний приріст** обчислюється шляхом ділення суми абсолютних приростів до їх числа, тобто

$$\bar{\Delta} = \frac{\sum_{t=1}^n \Delta t}{n}. \quad (8.13).$$

Середній коефіцієнт росту обчислюється за допомогою формули середньої геометричної:

$$\bar{K} = \sqrt[n]{K_1 \cdot K_2 \cdot \dots \cdot K_n}, \quad (8.14)$$

де K_1, K_2, \dots, K_n - ланцюгові темпи зростання; n - кількість ланцюгових темпів зростання.

Враховуючи мультиплікативність зв'язку ланцюгових і базових темпів зростання, середній коефіцієнт росту можна обчислити за формулою:

$$\bar{K} = \sqrt[n]{\frac{y_n}{y_0}}, \text{ де } \frac{y_n}{y_0} - \text{кінцевий базовий темп зростання.} \quad (8.15).$$

Середній темп приросту розраховують за формулою:

$$\bar{T} = (\bar{K} - 1) \cdot 100. \quad (8.16).$$

Наприклад, за останні 5 місяців динаміка темпів зростання цін на автомобільний бензин марки А-95 до попереднього місяця була такою: січень-1,10; лютий-1,15; березень - 1,18; квітень - 1,21; травень -1,30.

Середньомісячний темп зростання цін на бензин становить:

$$\bar{K} = \sqrt[5]{1,10 \cdot 1,15 \cdot 1,18 \cdot 1,21 \cdot 1,30} = \sqrt[5]{2,348} = 1,186 \text{ або } 118,6\%.$$

Середньомісячний темп приросту становить:

$$\bar{T} = (118,6 - 100) = 18,6\% .$$

Важливим при аналізі динамічних рядів є встановлення, як за той чи інший період змінилася структура ознаки. Структурні зміни можуть характеризуватись за допомогою як абсолютних, так і відносних показників динаміки, зокрема абсолютного приросту j -ї частки:

$$\Delta d_j = d_{j1} - d_{j0}; \quad (8.17).$$

темпу зростання j -ї частки:

$$kd_j = \frac{d_{j1}}{d_{j0}}. \quad (8.18).$$

Узагальнююча характеристика інтенсивності структурних зрушень оцінюється статистикою завдяки розрахунку середнього лінійного та квадратичного відхилення, тобто

$$\bar{d}_x = \frac{\sum |d_{j1} - d_{j0}|}{n} - \text{середнє лінійне відхилення}; \quad (8.19)$$

$$\sigma_d = \sqrt{\frac{\sum (d_{j1} - d_{j0})^2}{n}} - \text{середнє квадратичне відхилення}. \quad (8.20).$$

Тестові завдання

1. Кожний динамічний ряд складається з такої кількості обов'язкових елементів:

- а) двох;
- б) одного;
- в) трьох;
- г) залежно від виду динамічного ряду.

2. За часом динамічні ряди поділяють на:

- а) відкриті та закриті;
- б) інтервальні та моментні;
- в) додатні та від'ємні;
- г) повні та неповні.

3. Рівні моментного ряду характеризують значення показника:

- а) за рік;
- б) за певний період часу;
- в) на певний момент часу;
- г) в середньому за аналізований період.

4. Узагальнюючою характеристикою ознаки динамічного ряду є:

- а) рівень ряду;
- б) час;
- в) варіанта;
- г) середня величина.

5. Залежно від величин інтервалів в інтервальному варіаційному ряді узагальнююча величина обчислюється за допомогою:

- а) середньої арифметичної простої або зваженої;
- б) середньої хронологічної;
- в) середньої гармонійної;
- г) моди.

6. Середня величина ознаки для моментного динамічного ряду розраховується за допомогою:

- а) моди;
- б) середньої хронологічної;
- в) середньої прогресивної;

г) середньої геометричної.

7. Багатомірні ряди поділяються на:

- а) одномірні;
- б) двомірні;
- в) паралельні та ряди взаємозв'язаних показників;
- г) відкриті та закриті.

8. Зв'язок між ознаками багатомірного динамічного ряду буває:

- а) функціональним;
- б) кореляційним;
- в) функціональним або кореляційним;
- г) динамічним.

9. Абсолютний приріст рівня динамічного ряду це:

- а) відношення кожного наступного рівня до попереднього;
- б) відношення попереднього рівня до кожного наступного;
- в) різниця між попереднім і даним рівнем;
- г) різниця між даним рівнем і тим, із яким порівнюють.

10. Ланцюгові та базові коефіцієнти росту знаходяться між собою у зв'язку:

- а) мультиплікативному;
- б) адитивному;
- в) функціональному;
- г) кореляційному.

11. Порівняння інтенсивності розвитку ознак у двох рядах розраховують за допомогою коефіцієнта:

- а) кореляції;
- б) детермінації;
- в) абсолютного приросту;
- г) випередження.

12. Коефіцієнт еластичності характеризує:

- а) на скільки відсотків змінюється результативна ознака при зміні факторної на 1%;
- б) на скільки відсотків змінюється факторна ознака при зміні результативної на 1%;
- в) середньорічне зростання рівнів динамічного ряду;

г) вплив факторної ознаки на результативну.

Типові задачі

Задача 1. Є такі дані про динаміку виробництва продукції підприємством протягом 2007 – 2015 рр.:

| Рік | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 |
|---------------------------------|------|------|------|------|------|------|------|------|------|
| Виробництво продукції, тис. од. | 6,2 | 6,9 | 7,2 | 7,6 | 8,4 | | | | |
| До об'єднання | | | | | 9,2 | 10,3 | 11,2 | 12,4 | 13,1 |
| Після об'єднання | | | | | | | | | |

З метою забезпечення зіставності даних виконайте зімкнення ряду динаміки.

Задача 2. Маємо такі дані про використання валового внутрішнього продукту на споживання і нагромадження за 2008–2016 рр.:

(у фактичних цінах, млрд. грн.)

| Рік | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 |
|---|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| ВВП, використаний на споживання і нагромадження | 285,0 | 363,0 | 451,1 | 477,9 | 512,9 | 536,4 | 559,0 | 568,7 | 576,0 |

Розрахувати:

- 1) абсолютний приріст, коефіцієнти росту та приросту базовим і ланцюговим способом;
- 2) абсолютне значення одного відсотка приросту;
- 3) середньорічні абсолютні прирости.

Задача 3. Відомі такі дані про кількість працівників фірми:

| Дата | 1.01 | 1.02 | 1.03 | 1.04 | 1.05 | 1.06 | 1.07 | 1.08 | 1.09 | 1.10 | 1.11 | 1.12 | 1.01 |
|-----------------------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| Кількість працівників | 62 | 64 | 64 | 66 | 67 | 68 | 69 | 73 | 75 | 79 | 80 | 82 | 88 |

Визначіть вид ряду динаміки. Обчисліть середню кількість працівників за кожний квартал, перше і друге півріччя та за рік в цілому. Зробіть висновки.

Задача 4. Приводяться дані про розмір капітальних вкладень і введення в експлуатацію основних засобів за роками (млн. грн.):

| Показники | 2011 | 2012 | 2013 | 2014 | 2015 |
|---|------|------|------|------|------|
| Капітальні вкладення | 26,4 | 29,5 | 31,2 | 25,1 | 40,4 |
| Введено в експлуатацію основних засобів | 25,1 | 28,2 | 29,0 | 20,4 | 37,7 |

Привести ряд динаміки до однієї основи. За допомогою коефіцієнта випередження порівняти динаміку зазначених показників. Зробити висновки.

Питання для самоконтролю

1. Дайте визначення динамічного ряду.
2. Складові елементи динамічного ряду.
3. Види динамічних рядів.
4. Дайте тлумачення „змикання динамічних рядів”.
5. Найважливіші умови розрахунку середніх величин ряду динаміки.
6. Багатомірні динамічні ряди.
7. Найважливіші характеристики динамічних рядів.

Змістовий модуль 2

5. Аналіз рядів розподілу

Склад і структура ряду розподілу.

Структурні середні.

Квартилі та децилі як показники центру розподілу рядів.

Загальна характеристика показників варіації.

Розмах варіації, середнє лінійне відхилення, дисперсія, середнє квадратичне відхилення, коефіцієнт варіації.

«Спосіб моментів» при розрахунку дисперсії та середнього квадратичного відхилення.

6. Аналіз концентрації, диференції та подібності розподілів.

- 6.1 Загальна характеристика найважливіших форм розподілу.
- 6.2 Моменти розподілу.
- 6.3 Коефіцієнти асиметрії та ексцесу.
- 6.4 Оцінка концентрації розподілу.
- 6.5 Теоретичні частоти розподілу.
- 6.6 Критерії Пірсона та Колмогорова.
- 6.7 Загальна, систематична та випадкова дисперсія.

7. Статистичні методи вимірювання взаємозв'язків.

- 7.1 Причинні та результативні ознаки зв'язку.
- 7.2 Функціональні та стахостичні зв'язки
- 7.3 Парна кореляція.
- 7.4 Прямолінійна форма зв'язку.
- 7.5 Криволінійна форма зв'язку.
- 7.6 Оцінка лінійного та нелінійного зв'язку.
- 7.7 Коефіцієнт регресії у невеликих за обсягом сукупностях.
- 7.8 Перевірка істотності зв'язку.
- 7.9 Коефіцієнт взаємного узгодження.
- 7.10 Коефіцієнт спряженості Пірсона.
- 7.11 Коефіцієнт співзалежності Крамера.
- 7.12 Коефіцієнти асоціації та Спірмена.

8. Аналіз інтенсивності динаміки.

- 8.1 Порівнянність рівнів динамічних рядів.
- 8.2 Середні характеристики рядів динаміки.
- 8.3 Одномірні та багатомірні динамічні ряди.
- 8.4 Найважливіші характеристики динамічних рядів.
- 8.5 Середні значення показників ряду.

Завдання для самостійного вивчення

- 1. Частотні показники ряду розподілу.
- 2. Особливості розрахунку моди та медіани для дискретного та варіаційного рядів.
- 3. Взаємозв'язок показників варіації.
- 4. Дисперсія альтернативної ознаки в рядах розподілу.
- 5. Симетричний та асиметричний розподіли.
- 6. Стандартизовані відхилення.

7. Критерії оцінки нерівномірності розподілу.
8. Нормальна крива та нормальний розподіл.
9. Міжгрупова та середнє з групових дисперсії.
10. Кореляційне відношення.
11. Основні недоліки кореляційного методу аналізу.
12. Метод паралельних рядів і кореляційне поле як інструментарій пізнання зв'язку. Лінійний коефіцієнт кореляції.
13. Використання правила додавання дисперсій про оцінці кореляційного зв'язку.
14. Коефіцієнт детермінації та індекс кореляції.
15. Поняття множинності кореляції.
16. Динамічні ряди та завдання статистичного аналізу.
17. Коефіцієнт зникання динамічних рядів. Коефіцієнт еластичності.

РОЗДІЛ 9. АНАЛІЗ ТЕНДЕНЦІЙ РОЗВИТКУ ТА КОЛИВАНЬ

9.1 Дослідження тенденцій динамічних рядів

Інколи виникає потреба встановити напрям і тенденцію розвитку динамічного ряду. Використання первинної інформації не завжди робить це можливим. У зв'язку з цим ряди динаміки піддаються певній обробці. Зазначимо, що розвиток тенденції ряду залежить від рівня впливу на неї постійних $f(t)$ і залишкових (t) факторів.

Тенденція розвитку $f(t)$ забезпечується завдяки побудові поряд з емпіричним динамічним рядом теоретичного. Інакше кажучи, ряд динаміки піддається спеціальній статистичній обробці, з використанням адекватних способів способи згладжування та аналітичного вирівнювання.

Згладжування здійснюють використовуючи такі прийоми: збільшення періодів з подальшим розрахунком його середньої величини; переведення абсолютних показників динамічних рядів у відносні, в результаті чого досягається порівнянність динамічних рядів; розрахунок середньої плинної; вирівнювання з використанням середнього абсолютного приросту. Коротко охарактеризуємо кожний із наведених прийомів.

Вирівнювання за допомогою збільшення періодів проводять шляхом додавання рівнів аналізованого ряду за однакові проміжки часу (два, три, п'ять років тощо). Одержану в результаті додавання суму ділять на кількість рівнів ряду. Розраховані таким чином середні рівні, як правило, чіткіше відображають загальну тенденцію розвитку ряду.

Визначення тенденції динамічного ряду за допомогою середньої плинної полягає в тому, що кожний наступний інтервал утворюється на основі попереднього шляхом заміни одного рівня. Зокрема, початком першого такого інтервалу буде початковий рівень динамічного ряду, другого - другий, третього - третій і т.д. Теоретичний ряд плинних середніх завжди коротший за первинний на $(n-1)$ рівнів, що потребує логічного осмислення вибору ширини інтервалу n . В практичній діяльності розрахунок вказаним вище

способом проводять із непарним числом рівнів, оскільки в цьому випадку середня плинна буде віднесена до якогось конкретного часового періоду, тобто до медіанного рівня збільшеного інтервалу.

Схематично розрахунок середніх плинних можна подати таким чином:

$$\bar{a}_1 = \frac{a_1 + a_2 + a_3}{3} = \bar{a}_2 = \frac{a_2 + a_3 + a_4}{3} = \bar{a}_3 = \frac{a_3 + a_4 + a_5}{3} \text{ і т.д.} \quad (9.1).$$

Формула 9.1 засвідчує, що збільшений інтервал охоплює три рівні ряду.

Дію механізму згладжування методом середньої плинної розглянемо на прикладі динамічного ряду урожайності зернових культур (табл. 9.1). Ширина інтервалу згладжування $n = 3$. Емпіричний ряд складається із дев'яти рівнів, ряд плинних середніх на два рівні коротший.

Таблиця 9.1

Розрахунок плинних середніх урожайності зернових культур

| Порядковий номер року | Емпіричний ряд (y_t), ц/га | Плинна середня, \bar{y}_j | Техніка розрахунку, \bar{y}_j |
|-----------------------|--------------------------------|-----------------------------|---------------------------------|
| 1 | 28 | --- | |
| 2 | 30 | 31,0 | $(28+30+35):3=31,0$ |
| 3 | 35 | 34,3 | $(30+35+38):3=34,3$ |
| 4 | 38 | 38,0 | $(35+38+41):3=38,0$ |
| 5 | 41 | 39,0 | $(38+41+38):3=39,0$ |
| 6 | 38 | 39,0 | $(41+38+38):3=39,0$ |
| 7 | 38 | 39,6 | $(38+38+43):3=39,6$ |
| 8 | 43 | 42,0 | $(38+43+45):3=42,0$ |
| 9 | 45 | X | X |

Вирівняний динамічний ряд урожайності чітко засвідчує стабільний ріст урожайності зернових з 1 га посівної площі.

Вирівнювання за середнім абсолютним приростом полягає у тому, що кожний наступний рівень більший від попереднього на рівновелику величину, що дорівнює середньорічному абсолютному приросту, тобто

$$\bar{\Delta} = \frac{\sum \Delta_t}{n}, \text{ або } \bar{\Delta} = \frac{y_n - y_0}{n}, \quad (9.2).$$

де y_n - кінцевий рівень динамічного ряду; y_0 - початковий рівень динамічного ряду; n - число рівнів ряду.

В нашому прикладі середньорічний абсолютний приріст урожайності зернових в розрахунку на 1 га площі становить:

$$\bar{\Delta} = \frac{420 - 28}{9} = 1,55 \text{ ц/га.}$$

Практичне використання прийому згладжування за допомогою переведення абсолютних рівнів динамічних рядів у відносні дає можливість забезпечити порівнянність динамічних рядів непорівняних величин (табл.9.2)

Таблиця 9.2

Динаміка інвестицій в Україні за джерелами фінансування
(відсотків)

| | 2010 | 2011 | 2012 | 2013 | 2014 |
|---|------|------|------|------|------|
| Усього | 100 | 100 | 100 | 100 | 100 |
| у т.ч. за рахунок | | | | | |
| коштів державного бюджету | 5,7 | 7,2 | 6,0 | 2,5 | 1,2 |
| коштів місцевих бюджетів | 3,2 | 3,2 | 3,1 | 2,7 | 2,7 |
| власних коштів підприємств та організацій | 61,7 | 61,2 | 62,6 | 66,3 | 70,5 |
| кредитів банків та інших позик | 12,7 | 15,2 | 14,5 | 13,9 | 9,9 |
| коштів іноземних інвесторів | 2,1 | 2,1 | 1,8 | 1,7 | 2,6 |
| коштів населення на будівництво житла | 10,5 | 7,3 | 8,3 | 9,6 | 10,1 |
| інших джерел фінансування | 4,3 | 3,9 | 3,7 | 3,2 | 3,0 |

Приведені відносні показники динаміки інвестицій в Україні за джерелами фінансування засвідчують про зростання інвестицій за рахунок власних коштів підприємств та організацій. В той же час значно зменшилися інвестиції за рахунок коштів державного бюджету.

Подання показників динаміки через відносні величини називають в статистиці – **приведення рядів динаміки до однакової основи.**

9.2. Аналітичне вирівнювання

Часто з метою визначення тенденції (тренда) розвитку динамічного ряду використовують метод аналітичного вирівнювання. Його суть полягає в тому, що поряд з емпіричними рівнями ряду розраховують теоретичні рівні (ймовірні), які розглядаються як функції (тренди) часу, тобто:

$$Y_t = f(t). \quad (9.3).$$

Вирівнювання рядів динаміки можна проводити за допомогою таких функцій: лінійної, квадратичної, показникової тощо. Вибір типу функції, за допомогою якої буде описана тенденція ряду, залежить від варіації рівнів ряду.

Для лінійної функції графіком, який описує тенденцію ряду слугує пряма лінія. Лінійна функція характеризується стабільною швидкістю. Вона може бути описана рівнянням виду $\bar{y}_t = a_0 + a_1 t$, де \bar{y}_t - рівні вирівняного ряду динаміки, a_0 і a_1 - параметри рівняння прямої, t - час.

Для вирівнювання ряду за рівнянням прямої розв'язують систему двох нормальних рівнянь методом найменших квадратів:

$$\begin{cases} na_0 + a_1 \sum t = \sum y; \\ a_0 \sum t + a_1 \sum t^2 = \sum yt. \end{cases} \quad (9.4).$$

Система рівнянь значно спрощується, якщо відлік часу значень t перенести всередину динамічного ряду. Тоді $\sum t = 0$. В результаті перетворень одержимо:

$$\begin{cases} na_0 = \sum y; \\ a_1 \sum t^2 = \sum yt. \end{cases} \quad (9.5).$$

Використовуючи формулу 9.5 визначимо a_0 і a_1

$$a_0 = \frac{\sum y}{n}, \quad a_1 = \frac{\sum yt}{\sum t^2}. \quad (9.6).$$

Розрахунок $\sum t^2$ залежить від парності або непарності числа рівнів вирівнюваного динамічного ряду. Зокрема, при парному числі рівнів ряду використовують наступну формулу:

$$\sum t^2 = \frac{n(n^2 - 1)}{3}; \quad (9.7).$$

при непарному:

$$\sum t^2 = \frac{n(n^2 - 1)}{12}. \quad (9.8).$$

Наведемо приклад обчислення параметрів лінійної функції (табл. 9.3) Виходячи з інформації, наведеної в табл. 9.3, визначимо значення $\sum t^2$ для непарного числа рівнів ряду.

$$\sum t^2 = \frac{7(7^2 - 1)}{12} = 28.$$

Використавши дані таблиці та розраховане $\sum t^2$, обчислимо значення a_0 і a_1

Таблиця 9.3

Динаміка споживання дизельного палива на підприємстві

| Роки | Спожито дизельного палива, тис. м ³ , y_t | Умовні позначення часу, t | $y_t t$ | $Y_t = 428.57 + 47.88 t$ |
|-------|--|-----------------------------|---------|--------------------------|
| 2008 | 300 | -3 | -900 | 285,00 |
| 2009 | 330 | -2 | -660 | 332,86 |
| 2010 | 380 | -1 | -380 | 380,71 |
| 2011 | 410 | 0 | 0 | 428,57 |
| 2012 | 470 | 1 | 470 | 476,43 |
| 2013 | 520 | 2 | 1040 | 524,29 |
| 2014 | 590 | 3 | 1770 | 572,14 |
| Разом | 3000 | 0 | 1340 | 3000 |

$$a_0 = 3000 : 7 = 428,57; \quad a_1 = 1340 : 28 = 47,88.$$

Розрахований лінійний тренд має такий вигляд: $\bar{y}_t = 428,57 + 47,857t$. Економічна інтерпретація приведенного тренду означає, що середній рівень споживання дизельного палива за останні сім років на підприємстві становив 428,57 тис. м³, а середньорічний приріст споживання - 47,857 тис. м³.

Підсумкові дані емпіричного ряду (y_t) і теоретичного (Y_t) (див.табл.9.3) за величиною ϵ ідентичними, що дає можливість стверджувати про правильність проведених нами розрахунків.

Значне коливання абсолютних приростів ряду динаміки, а також не чітко описана їх спрямованість до зростання або зменшення потребує вирівнювання його за допомогою рівнянь параболи, гіперболи тощо. Правильність вибору рівняння перевіряється сумою квадратів відхилень вирівняних рівнів від фактичних, яка має бути мінімальною, тобто

$$\sum (y - \bar{y}_y)^2 = \min. \quad (9.9).$$

| Вирівнювання за рівнянням параболи другого порядку | Вирівнювання за рівнянням прямої лінії |
|---|---|
| $\bar{y}_1 = a_0 + a_1 t + a_2 t^2$ | $\bar{y}_1 = a_0 + a_1 t$ |
| Система нормальних рівнянь для визначення параметрів a_0, a_1, a_2 | Система нормальних рівнянь для визначення параметрів a_0 і a_1 |
| $\begin{cases} \sum y = na_0 + a_1 \sum t + a_2 \sum t^2 \\ \sum yt = a_0 \sum t + a_1 \sum t^2 + a_2 \sum t^3 \\ \sum yt^2 = a_0 \sum t^2 + a_1 \sum t^3 + a_2 \sum t^4 \end{cases}$ | $\begin{cases} \sum y = na_0 + a_1 \sum t \\ \sum yt = a_0 \sum t + a_1 \sum t^2 \end{cases}$ |

Тенденція динамічного ряду оцінюється в статистиці також за допомогою вирівнювання ряду динаміки за середнім абсолютним приростом та середнім коефіцієнтом зростання.

1) за середнім абсолютним приростом

$$\bar{y}_t = y_0 + \bar{A}t; \quad (9.10).$$

2) за середнім коефіцієнтом зростання

$$\bar{y}_t = y_0 \bar{K}^t, \quad (9.11).$$

де

$$\bar{A} = \frac{y_n - y_0}{n-1}; \quad \bar{K} = \sqrt[n-1]{\frac{y_n}{y_0}}. \quad (9.12).$$

Інколи в динамічному ряді відсутні значення рівнів у середині або в кінці ряду. Знаходження невідомого рівня в середині ряду називається **інтерполяцією**, в кінці - **екстраполяцією**. Розрахунок

їх значення ґрунтується на даних попереднього (наступного) рівнів і може бути обчислений виходячи з тенденції розвитку, що склалася раніше.

9.3. Статистичне вивчення стійкості рівнів динамічних рядів та сезонності

Рівні динамічних рядів під впливом різноманітних причин варіюють і часто відхиляються від загальної тенденції розвитку. При цьому в одних випадках відхилення мають систематичний характер, в інших - випадковий. Систематичність проявляється у повторюваності коливань. Випадковість такої повторюваності не має. Коливання рівнів динамічних рядів може бути результатом дії як систематичних, так і випадкових причин.

Варіація рівнів динамічних рядів часто зумовлюється також сезонними коливаннями ознак досліджуваних динамічних рядів. Сезонні коливання, як правило, впливають на рівномірність використання ресурсів, часу, праці, грошових коштів тощо. Інколи вони спричиняють до втрат, нераціонального використання ресурсного потенціалу. Саме тому вони потребують статистичного дослідження та регулювання.

Мірою сезонних коливань є "хвиля сезонності", яка оцінюється за допомогою індексів сезонності. Інтенсивність сезонних коливань вимірюється також за допомогою розрахунку таких основних показників: розмах варіацій (R_t); середнє лінійне відхилення (\bar{d}_t); середнє квадратичне відхилення (σ_t); коефіцієнт варіації (v_t). При цьому:

$$R_t = l_{\max} - l_{\min} - \text{розмах варіації, амплітуда коливання}; \quad (9.13).$$

$$\bar{d}_t = \frac{1}{12} \sum |I_{\text{сез}} - 100| - \text{середнє лінійне відхилення}; \quad (9.14).$$

$$\sigma_t = \sqrt{\frac{1}{12} \sum (I_{\text{сез}} - 100)^2} - \text{середнє квадратичне відхилення}. \quad (9.15).$$

Мірою стійкості рівнів динамічного ряду може бути величина, котра дорівнює одиниці (100%) мінус значення коефіцієнта

варіації v_t . Чим ближчий цей коефіцієнт до одиниці, тим більша стійкість динамічного ряду.

Статистичним показником вимірювання сезонних коливань є також індекс сезонності (I) - відношення фактичного рівня (y_t) за певний часовий період року до середньомісячного рівня за рік \bar{y} .

$$I_{сез} = \frac{y_t}{\bar{y}} \cdot 100. \quad (9.16).$$

Розрахунок показників сезонності проведемо покажемо за допомогою інформації, поміщеної в табл.9.4. В таблиці приведені дані про затрати часу на сільськогосподарських роботах за місяцями минулого року.

Таблиця 9.4

Дані для розрахунку показників сезонних коливань

| Місяць року | $I_{сез}, \%$ | $I_{сез} - 100$ | $(I_{сез} - 100)^2$ |
|-------------|---------------|-----------------|---------------------|
| січень | 96,9 | -3,1 | 9,61 |
| лютий | 95,6 | -4,4 | 19,36 |
| березень | 84,3 | -15,7 | 246,49 |
| квітень | 95,0 | -5,0 | 25,00 |
| травень | 80,7 | -19,3 | 372,49 |
| червень | 100,6 | +0,6 | 0,36 |
| липень | 108,6 | +8,6 | 73,96 |
| серпень | 122,4 | +22,4 | 501,76 |
| вересень | 111,3 | + 11,3 | 127,69 |
| жовтень | 104,2 | +4,2 | 17,64 |
| листопад | 102,4 | +2,4 | 5,76 |
| грудень | 98,0 | -2,0 | 4,00 |
| Разом: | 100,0 | 0 | 1404,12 |

Приведемо необхідні розрахунки.

Амплітуда сезонних коливань R_t , дорівнює 41,7 (122,4 - 80,7). Розрахунок засвідчує, що піковий період затрат праці припадав на серпень. Найменше часу було затрачено в травні – 80,7 тис.людино-годин. Середнє квадратичне відхилення σ становить - 10,82%

$\left(\sqrt{\frac{1}{12} \cdot 1404,12}\right)$ і дає можливість оцінити величину коливань затрат праці навколо її середньої величини. Найбільший розмір $I_{сез}$ склав у серпні (1.224). Означене пояснюється тим, що в цей період закінчуються жнива та розпочинаються осінньо-польові роботи.

Інколи досліджувані статистикою динамічні ряди мають тенденцію до зростання або зменшення, що може спричинити до значного відхилення окремих значень від постійного середнього рівня ознаки. В таких випадках застосовують метод помісячних відношень. Суть цього методу полягає в тому, що спочатку обчислюють по кожному періоду ланцюгові темпи зростання, а потім на основі одержаних результатів визначають їх середню арифметичну величину.

Показником, за допомогою якого досліджують коливання в рядах динаміки за відносно невеликі проміжки часу слугує **коефіцієнт нерівномірності**. Його розраховують як відношення відповідно максимального і мінімального рівнів динамічного ряду до середнього. Значна відмінність між розрахованими коефіцієнтами засвідчує про розмір нерівномірності.

Статистична оцінка двох і більше динамічних рядів інколи засвідчує про наявність взаємозв'язку між їх рівнями. Дослідження зв'язку, як зазначалося раніше здійснюється за допомогою кореляційного методу аналізу. Його використання передбачає розв'язання кореляційно-регресійних рівнянь, а також встановлення тісноти зв'язку. Спочатку на основі статистичного аналізу оцінюють якісний зміст досліджуваних ознак динамічних рядів. Відомо, що два випадкові динамічні ряди, які мають однакову трендову спрямованість (обидва зростають або зменшуються), завжди будуть перебувати один з одним у прямій кореляційній залежності. Зокрема, у випадку, коли рівні одного ряду мають тенденцію до постійного зростання, а іншого - до зменшення, то вони будуть перебувати в достатньо тісному зворотному зв'язку. Завдання полягає в тому, щоби на основі логічного аналізу виявити справжню природу кореляційного зв'язку. Мова йде про те, що інколи при

використанні кореляційних методів аналізу динамічних рядів може виникнути **автокореляція** (неправильна кореляція), котра може бути через нехтування логічності взаємозв'язку окремих ознак. Наприклад, зв'язок між надоем молока і кількістю продемонстрованих кінофільмів; урожайністю цукрових буряків і чисельністю політичних партій засвідчує про алогічність ознак.

Статистика для усунення автокореляції використовує декілька способів. Один із них передбачає розрахунок одночасно коефіцієнта кореляції різниць, який розраховується як різниця між даним рівнем і попереднім за обома ознаками.

$$r_{\Delta_x \Delta_y} = \frac{\sum \Delta_x \cdot \Delta_y}{\sqrt{\sum \Delta_x^2 \cdot \Delta_y^2}}. \quad (9.17).$$

Другий спосіб усунення автокореляції передбачає розрахунок відхилень не від попереднього рівня, а від тренда - основної лінії розвитку явища. Тренд може бути поданий рівнянням прямої або кривої лінії. Рівняння розв'язують для ряду динаміки як факторної, так і результативної ознаки, тобто для \bar{x}_i і \bar{y}_i . Розрахунок коефіцієнта кореляції за допомогою цього способу здійснюється завдяки використанню наступної формули

$$r = \frac{\sum (x - \bar{x}_i)(y - \bar{y}_i)}{\sqrt{\sum (x - \bar{x}_i)^2 \sum (y - \bar{y}_i)^2}}. \quad (9.18).$$

Статистичне дослідження динамічних рядів інколи пов'язане з явищем лагу. **Лаг** - це період часу відставання в розвитку одного із двох взаємозв'язаних рядів. Зокрема, ефективність капітальних вкладень проявляється, як правило, через декілька років; урожайність кормових культур даного року впливає на кількість поголів'я в наступному році, тобто лаг дорівнює одному року.

Розглянемо приклад наявності кореляційного зв'язку між рядами динаміки обсягу споживання дизельного палива та природного газу на одному із фермерських господарств (табл.9.5).

Таблиця 9.5

Обсяг споживання дизельного палива та природного газу фермерським господарством за 2008-2014р.

| Рік | Обсяг споживання дизельного палива, тис. м ³ , x | Обсяг споживання газу, тис. м ³ , y | Розрахункові значення | | | | |
|-------|---|--|-----------------------|---------------|-------------------------------------|-------------------|-------------------|
| | | | $x - \bar{x}$ | $y - \bar{y}$ | $(x - \bar{x}) \cdot (y - \bar{y})$ | $(x - \bar{x})^2$ | $(y - \bar{y})^2$ |
| 2008 | 300 | 20 | -139 | -20 | 2817 | 19346 | 410 |
| 2009 | 330 | 35 | -109 | -5 | 573 | 11901 | 28 |
| 2010 | 380 | 60 | -59 | 20 | -1167 | 3492 | 390 |
| 2011 | 410 | 70 | -29 | 30 | -865 | 846 | 885 |
| 2012 | 470 | 0 | 31 | -40 | -1244 | 955 | 1620 |
| 2013 | 520 | 20 | 81 | -20 | -1638 | 6546 | 410 |
| 2014 | 590 | 70 | 151 | 30 | 4490 | 22774 | 885 |
| Разом | 3000 | 275 | - | - | 2964 | 65860 | 4628 |

Обчислений коефіцієнт кореляції $r_{xy} = 0,17$ вказує на не тісний прямий зв'язок між динамікою обсягу спожитого дизельного палива та газу. Разом з тим, логіка розвитку цих двох рядів засвідчує про помилковість та алогічність такого твердження, тобто має місце автокореляція. Використаємо перший спосіб усунення автокореляції. Для цього побудуємо таблицю 9.6.

Таблиця 9.6

Розрахунок коефіцієнта кореляції різниць

| Роки | Обсяг споживання дизельного палива, тис. м ³ , | Обсяг споживання газу, тис м ³ , у | Різниця між наступним і попереднім рівнями | | Δx^2 | Δy^2 | $\Delta_x \Delta_y$ |
|-------|---|---|--|------------|--------------|--------------|---------------------|
| | | | Δx | Δy | | | |
| 2008 | 300 | 20 | - | - | - | - | - |
| 2009 | 330 | 35 | +30 | +15 | 900 | 225 | 450 |
| 2010 | 380 | 60 | +50 | +25 | 2500 | 625 | 1250 |
| 2011 | 410 | 70 | +30 | +10 | 900 | 100 | 300 |
| 2012 | 470 | 0 | +60 | +30 | 3600 | 900 | 1800 |
| 2013 | 520 | 20 | +50 | +20 | 2500 | 400 | 1000 |
| 2014 | 590 | 70 | +70 | +50 | 4900 | 2500 | 3500 |
| Разом | 3000 | 275 | - | - | 15300 | 4750 | 8300 |

$$r_{\Delta_x \Delta_y} = \frac{\sum \Delta_x \cdot \Delta_y}{\sqrt{\sum \Delta_x^2 \cdot \Delta_y^2}} = \frac{8300}{\sqrt{15300 \cdot 4750}} = \frac{8300}{\sqrt{72675000}} = \frac{8300}{8525} = 0,973.$$

При порівнянні коефіцієнтів кореляції двох рядів динаміки видно, що в даному прикладі має місце автокореляція, оскільки одержаний результат не є однаковим.

Другий спосіб усунення автокореляції, як нами зазначалось вище, полягає в обчисленні відхилень не від попереднього рівня, а від тренда - основної лінії розвитку явища. В залежності від напрямку руху динамічного ряду тренд може бути побудований за допомогою рівняння прямої або кривої лінії. Використання цього способу забезпечується з урахуванням величини лага. Початковий етап полягає в побудові рівняння тренда для кожного ряду. Потім, виходячи із приведеної формули, знаходять відповідно \bar{x}_i і \bar{y}_i . Одержаний результат r інтерпретують аналогічно, як при використанні першого способу.

Тестові завдання

1. Практичне використання прийому згладжування за допомогою переведення абсолютних рівнів динамічних рядів у відносні дає можливість:

- а) забезпечити порівнянність динамічних рядів непорівняних величин;
- б) встановити напрям розвитку досліджуваної ознаки;
- в) визначити коефіцієнти росту рівнів ряду;
- г) розрахувати середній приріст рівнів ряду.

2. Суть аналітичного вирівнювання полягає в тому, що:

- а) тенденцію динамічного ряду визначають шляхом укрупнення періодів;
- б) поряд з емпіричними рівнями ряду розраховують теоретичні, які розглядаються як функції часу;
- в) поряд з теоретичними рівнями ряду розраховують емпіричні, котрі розглядаються як тренди часу;
- г) знаходять значення невідомих рівнів в середині ряду.

3. Вибір функції за допомогою якої проводять вирівнювання динамічного ряду залежить від:

- а) числа рівнів динамічного ряду;
- б) виду динамічного ряду;
- в) варіації рівнів ряду;
- г) середнього лінійного відхилення.

4. Правильність вибору аналітичного рівняння за допомогою якого здійснюється вирівнювання динамічного ряду перевіряється:

- а) укрупненням періодів;
- б) за допомогою середньої ковзної;
- в) знаходження невідомого рівня в кінці ряду;
- г) сумою квадратів відхилень вирівняних рівнів від фактичних.

5. Знаходження невідомого рівня в середині ряду називається:

- а) інтерполяцією;
- б) екстраполяцією;
- в) трендом;
- г) ковзною.

6. Знаходження невідомого рівня в кінці ряду називається:

- а) інтерполяцією;
- б) екстраполяцією;
- в) трендом;
- г) ковзною.

7. Коливання рівнів динамічних рядів може бути результатом дій:

- а) систематичних причин;
- б) випадкових причин;
- в) систематичним і випадкових причин;
- г) зовнішніх і внутрішніх причин.

8. Мірою сезонних коливань є:

- а) середнє лінійне відхилення;
- б) розмах варіації;
- в) коефіцієнт варіації;
- г) хвиля сезонності.

9. Показником, за допомогою якого розраховується коливання в рядах динаміки за відносно невеликі проміжки часу є:

- а) коефіцієнт нерівномірності;
- б) індекс сезонності;
- в) розмах варіації;
- г) коефіцієнт варіації.

10. Коефіцієнт нерівномірності, за допомогою якого оцінюються коливання в рядах динаміки за відносно невеликі проміжки часу розраховують шляхом ділення:

- а) максимального рівня ряду на мінімальний;
- б) мінімального рівня ряду на максимальний;
- в) мінімального і максимального рівнів динамічного ряду на середній ;
- г) середнього рівня ряду на величину розмаху варіації.

11. Автокореляція може виникнути внаслідок:

- а) значного коливання рівнів динамічного ряду;
- б) нехтування логічністю взаємозв'язку окремих ознак;
- в) дослідження незначного обсягу сукупності;

г) помилок, допущених при розрахунку коефіцієнта кореляції.

12. Лаг – це:

- а) початковий рівень динамічного ряду;
- б) середній рівень динамічного ряду;
- в) період часу відставання ознак у розвитку одного з двох взаємозв'язаних рядів;
- г) підсумок відхилень теоретичних і емпіричних рівнів ряду.

Типові задачі

Задача 1. Приводиться динаміка кількості дорожньо-транспортних випадків у регіоні за рік, чол.

| Місяць | Осіб | Місяць | Осіб |
|----------|------|----------|------|
| січень | 190 | липень | 100 |
| лютий | 150 | серпень | 130 |
| березень | 110 | вересень | 110 |
| квітень | 80 | жовтень | 140 |
| травень | 90 | листопад | 160 |
| червень | 120 | грудень | 180 |

Визначити середню хвилю дорожньо-транспортних випадків і оцінити сезонність за допомогою амплітуди коливань і середнього квадратичного відхилення.

Задача 2. Маємо наступні дані про обсяг наданих послуг:

(тис. грн.)

| Квартал | Роки | | | |
|---------|------|------|------|------|
| | 2012 | 2013 | 2014 | 2015 |
| 1 | 648 | 709 | 515 | 766 |
| 2 | 312 | 206 | 186 | 284 |
| 3 | 123 | 94 | 109 | 157 |
| 4 | 1013 | 841 | 973 | 1251 |

За наведеними даними необхідно:

- 1) побудувати лінійний графік та сформулювати гіпотезу про наявність сезонних коливань;
- 2) розрахувати індекси сезонності;
- 3) побудувати графік сезонної хвилі.

Задача 3. Приводяться такі дані про динаміку виробництва товарної продукції підприємством:

| | | | | | | | | | | | | |
|------------------------------|-----|-----|------|------|------|------|------|------|------|------|------|------|
| Місяць | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| Обсяг виробництва, млн. грн. | 6,3 | 9,3 | 10,2 | 11,7 | 12,6 | 11,7 | 14,0 | 12,6 | 13,0 | 14,3 | 13,5 | 14,5 |

Провести згладжування динамічного ряду, використовуючи трьохмісячну ковзну середню.

Задача 4. Динаміка чисельності населення (дані умовні) двох країн характеризується наступним чином (див. табл.), тис. чол.

| Роки | А | Б |
|------|------|------|
| 2007 | 2414 | 3620 |
| 2008 | 2558 | 3631 |
| 2009 | 2586 | 3682 |
| 2010 | 2596 | 3690 |
| 2011 | 2595 | 3695 |
| 2012 | 2610 | 3701 |
| 2013 | 2611 | 3715 |
| 2014 | 2600 | 3711 |
| 2015 | 2595 | 3707 |

Для кожного ряду динаміки:

- 1) описати тенденцію зростання за допомогою кривої;
- 2) обчислити параметри трендових рівнянь, дати їм економічну інтерпретацію;
- 3) приймаючи, що виявлена тенденція збережеться, визначити очікувану чисельність населення кожної країни в 2016-2017 році.

Питання для самоконтролю

1. Порядок розрахунку коефіцієнт випередження.

2. Найважливіші показники інтенсивності структурних зрушень рівнів динамічного ряду.
3. Способи згладжування динамічних рядів.
4. Суть методу плинних середніх.
5. Суть методу аналітичного групування.
6. Дефініція «екстраполяція» та «інтерполяція».
7. Показники сезонності.
8. Порядок розрахунку індексу сезонності.
9. Поняття та суть автокореляції.

10.1. Суть, значення та види індексів

Слово «**index**» латинського походження і означає «узагальнюючий показник». У статистиці **індексом** називається відносна величина, яка характеризує зміну в часі, співвідношення в просторі рівня явища суспільного життя, або ступінь виконання наміченого завдання. Його одержують шляхом порівняння числових значень однойменних показників у часі або просторі. Значення індексів полягає насамперед у тому, що за їх допомогою вивчають динаміку обсягу виробництва, рівень цін, продуктивність праці, собівартість та інше. Крім цього, встановлюють взаємозв'язок між показниками, визначають вплив факторів на зміну складного явища. Статистичне вивчення індексів забезпечується завдяки їх поділу на окремі види. Зокрема, за способом побудови загальні індекси поділяють на агрегатні та середні. За характером порівнянь (у часі, просторі) їх поділяють на динамічні, територіальні, міжгрупові.

За ступенем агрегованості інформації виокремлюють індивідуальні та зведені індекси.

Індивідуальні індекси характеризують співвідношення окремих елементів. Вони поділяються на ланцюгові та базові. Це темп росту, рівень виконання наміченого завдання або відносні показники територіального порівняння.

Індивідуальний індекс позначається буквою «*i*». Його розрахунок п забезпечується кризь обов'язкове зіставлення рівнів двох періодів. Показники базового періоду позначають підрядковим знаком «0», а поточного або звітного - «1». Оцінка розвитку явища, котре охоплює більше як два періоди, забезпечується завдяки позначенню їх підрядковими знаками «2», «3» і т. д.

Статистичні ознаки прийнято позначати такими буквами: кількість продукції *q*, ціна *p*, собівартість одиниці товару *z*, продуктивність праці *w*, трудомісткість *t*.

Символи *q* та *p* запозичені від англійських слів quantity (кількість) та price (ціна). Найбільш широке використання в статистиці одержав розрахунок таких індивідуальних індексів:

$$i_q = \frac{q_1}{q_0} - \text{фізичного обсягу продукції}; \quad (10.1).$$

$$i_p = \frac{p_1}{p_0} - \text{ціни}; \quad (10.2).$$

$$i_z = \frac{z_1}{z_0} - \text{собівартості}; \quad (10.3).$$

$$i_w = \frac{w_1}{w_0} - \text{продуктивності праці натуральний}; \quad (10.4).$$

$$i_t = \frac{t_0}{t_1} - \text{продуктивності праці трудовий}; \quad (10.5).$$

$$i_{\frac{1}{t}} = \frac{t_1}{t_0} - \text{трудомісткості} \quad (10.6).$$

Індекси вимірюють у відсотках або коефіцієнтах. Якщо результат більше одиниці, це свідчить про ріст явища, і навпаки. Індекси можуть розраховувати базовим і ланцюговим способами. Індивідуальні ланцюгові та базові індекси пов'язані між собою за умови, що вони охоплюють один і той же період. Тоді добуток ланцюгових індексів дорівнює базовому:

$$i = \frac{q_1}{q_0} \cdot \frac{q_2}{q_1} \cdot \frac{q_3}{q_2} \cdot \frac{q_4}{q_3} = \frac{q_4}{q_0}. \quad (10.7).$$

Частка від ділення рівня наступного базового індексу на попередній дорівнює відповідному ланцюговому індексу: $\frac{q_4}{q_0} : \frac{q_3}{q_0} = \frac{q_4}{q_3}$.

Статистичне дослідження показників інколи подається як добуток індексів - співмножників. Такі індекси називаються "взаємозалежними". Наприклад, індекс затрат дорівнює добутку індексів собівартості та фізичного обсягу продукції ($i_{zq} = i_z \cdot i_q$), індекс товарообігу одержують шляхом множення індексу фізичного обсягу продукції на індекс ціни ($i_{pq} = i_p \cdot i_q$).

Наведені вище взаємозалежні індекси використовують при необхідності визначення третього, коли величина двох із них відома. Наприклад, валовий прибуток підприємства за звітній період зріс на 20,5%, а обсяг виробництва на - 15%. Індекс валового прибутку в цьому випадку становитиме $1,205 : 1,15 = 1,04$.

Часто в процесі статистичного дослідження виникає потреби оцінювання явища, котре безпосередньо підсумувати не можна, оскільки така дія позбавлена логічності. Зокрема, на швейній фабриці шиють не тільки жіночий одяг, але і чоловічий, дитячий, для літнього сезону, зимового і т. д. Розрахунок динаміки або порівняння досліджуваного показника потребує в цьому випадку використання загального або групового індексу. Узагальнююча характеристика обсягу виробництва за всім асортиментом одягу забезпечується шляхом використання загального індексу. Оцінка динаміки за окремими групами забезпечується завдяки побудові зведених групових індексів.

Загальний індекс позначається буквою I . Він складається з двох елементів: індексованої величини і «ваги», або співвимірника.

Одним із принципово важливих питань при побудові індексів є визначення співвимірника. Зокрема, для визначення вартості різних видів продукції як співвимірник використовується відповідна ціна. Оцінка динаміки виробництва також забезпечується шляхом застосування в якості ваги відповідної ціни. Співвимірником загальної собівартості або виробничих затрат слугує кількість виробленої продукції (виконаних робіт, наданих послуг). Взагалі при побудові загальних індексів необхідно виходити з наступного: якщо визначається індекс якісного показника, то співвимірник (кількісний) беруть за звітний період, а якщо кількісний, то співвимірник береться на рівні базового періоду.

10.2 Агрегатна форма індексів

Найбільш широке використання в практиці статистичних досліджень одержала агрегатна форма індексів. Агрегатний індекс -

співвідношення двох агрегатів, які наповнені конкретним змістом і часом. Індексована величина агрегату розміщена після значка Σ . В чисельнику та знаменнику вона подається в різних періодах. Після індексованої величини розміщується вага (співвимірник), котра подається на рівні звітного або базового періодів. Зокрема, при розрахунку загальних індексів ціни, фізичного обсягу продукції індексованою величиною слугують відповідно ціна p_1 та обсяг продукції q_1 , вагами відповідно для ціни - продукція звітного періоду, для індексу фізичного обсягу виробництва - ціна базового періоду.

Агрегатна форма індексу використовується для одержання узагальнюючої характеристики неоднорідної сукупності. Прикладом такої може бути загальна сума реалізованих продовольчих та промислових товарів або їх частини. При цьому суму виручки товарообігу при цьому записують у вигляді агрегату (сума добутків ваг (ціни) та кількості проданих товарів). Агрегатну форму використовують при обчисленні загальних індексів товарообігу, фізичного обсягу продукції, собівартості, затрат, продуктивності праці, трудомісткості. Загальний індекс товарообігу в фактичних цінах розраховують:

$$I_{pq} = \frac{\sum p_1 q_1}{\sum p_0 q_0}, \quad (10.8).$$

де $\sum p_1 q_1$ - сума товарообігу в звітному періоді, $\sum p_0 q_0$ - сума товарообігу в базовому періоді.

Приведений індекс засвідчує про те, як змінився товарообіг у звітному періоді в порівнянні з базовим. Його величина залежить від двох складових: кількості реалізованих товарів і ціни за одиницю товару. Оцінка впливу кожної з них на загальний товарообіг визначається шляхом розрахунку загальних індексів фізичного обсягу товарообігу та ціни. Індекс фізичного обсягу визначають за допомогою формули:

$$I_q = \frac{\sum q_1 p_0}{\sum q_0 p_0}. \quad (10.9).$$

Чисельником цього індексу слугує сума вартості продукції (товару) звітного періоду за цінами базового періоду, а знаменником - сума вартості продукції (товару) базового періоду за цінами базового періоду. Використовуючи результати розрахунку індексу фізичного обсягу продукції можна визначити величину абсолютного приросту продукції (товарообігу) в грошовому виразі в звітному періоді в порівнянні з базовим за рахунок зміни кількості виробленої продукції (проданого товару). Абсолютний приріст дорівнює різниці між чисельником і знаменником вищенаведеного індексу тобто,

$$\pm \Delta q = \sum q_1 p_0 - \sum q_0 p_0. \quad (10.10)$$

Формула загального агрегатного індексу цін має такий вигляд:

$$I_p = \frac{\sum p_1 q_1}{\sum p_0 q_1}. \quad (10.11)$$

Чисельник і знаменник цього індексу складається з двох співмножників. Один із них є змінною індексованою величиною (p), а другий приймається умовно як постійна величина (q_1). Різниця між знаменником і чисельником індексу цін $\sum p_0 q_1 - \sum p_1 q_1$ є фактичною абсолютною економією (перевитратою) від зміни цін. Приведемо приклад (табл. 10.1).

Таблиця 10.1

Обсяг та ціни реалізованої сільськогосподарської продукції

| Продукти | Продано, кг (q) | | Ціна одиниці, грн. (p) | |
|----------|---------------------------|---------------------------|----------------------------|---------------------------|
| | в базовому періоді, q_0 | в звітному періоді, q_1 | в базовому періоді, p_0 | в звітному періоді, p_1 |
| Цукор | 30000 | 40000 | 15 | 12 |
| Пшениця | 20000 | 30000 | 4 | 3 |
| М'ясо | 8500 | 9000 | 50 | 40 |

Визначити:

1. Загальну зміну фізичного обсягу продажу.
2. Загальну зміну цін на продукти.
3. Абсолютну економію населення від зниження цін.

Розв'язок:

1. Загальна зміна фізичного обсягу продаж становить:

$$I_q = \frac{\sum q_1 p_0}{\sum q_0 p_0} = \frac{40000 \cdot 15 + 30000 \cdot 4 + 9000 \cdot 50}{30000 \cdot 15 + 20000 \cdot 4 + 8500 \cdot 50} = \frac{1170000}{955000} = 1,225.$$

2. Загальний індекс цін на продукти:

$$I_p = \frac{\sum p_1 q_1}{\sum p_0 q_1} = \frac{12 \cdot 40000 + 3 \cdot 30000 + 40 \cdot 9000}{15 \cdot 40000 + 4 \cdot 30000 + 50 \cdot 9000} = \frac{930000}{1170000} = 0,795.$$

3. Абсолютна економія населення від зниження цін становить:

$$\pm \Delta q = \sum p_0 q_1 - \sum p_1 q_1 = 1170000 - 930000 = 240000 \text{ грн.}$$

Отриманий індекс фізичного обсягу продаж свідчить про зростання реалізованої продукції в звітному періоді проти базового на 23,2 %, за рахунок чого товарообіг ринку зріс на 215 тис. грн. ($\Delta q = 1170000 - 955000 = 215000$).

Загальний індекс цін засвідчує, що ціни на ринку в середньому знизились на 20,5%, а абсолютна економія населення в результаті зниження цін становить 240 тис. грн.

Агрегатну форму індексів застосовують також при аналізі собівартості продукції (виконаних робіт чи наданих послуг), продуктивності праці, трудомісткості тощо. Зокрема для розрахунку собівартості продукції застосовують наступні формули:

$$I_z = \frac{\sum z_1 q_1}{\sum z_0 q_1}, \quad (10.12).$$

де z_1, z_0 - собівартість одиниці продукції в звітному та базовому періодах. Наведемо приклад розрахунку індексу собівартості (табл. 10.2)

Таблиця 10.2

Дані для розрахунку індексу собівартості

| | Виготовлено автомобілів, од. (q) | | Собівартість одного автомобіля, тис. ум. од.(z) | |
|-----------------------------|---------------------------------------|---------------------------------------|--|---------------------------------------|
| | в базовому періоді, q ₀ | в звітному періоді. q ₁ | в базовому періоді, z ₀ | в звітному періоді, z ₁ |
| Спеціального призначення | 10 | 12 | 20,0 | 19,2 |
| Вантажні | 30 | 31 | 70,0 | 68,0 |
| Легкові | 156 | 160 | 15,0 | 14,2 |

Визначити:

4. Агрегатний індекс собівартості виготовлення автомобіля.
5. Абсолютну економію від зниження собівартості автомобіля.

Розв'язок:

1. Агрегатний індекс собівартості автомобіля складатиме:

$$I_z = \frac{\sum z_1 q_1}{\sum z_0 q_1} = \frac{19,2 \cdot 12 + 68,0 \cdot 31 + 14,2 \cdot 160}{20,0 \cdot 12 + 70,0 \cdot 31 + 15,0 \cdot 160} = \frac{4610}{4810} = 0,958.$$

2. Абсолютна економія від зниження собівартості продукції.

$$\sum z_0 q_1 - \sum z_1 q_1 = 4810 - 4610 = 200 \text{ грн.}$$

Одержаний індекс свідчить про зниження собівартості автомобіля на заводі в звітному періоді порівняно з базовим на 4,1%. Результатом такої роботи заводу забезпечили економію коштів на суму 200 тис. ум. од.

Індекс трудомісткості характеризує зміну трудових затрат на виробництво одиниці продукції, тобто:

$$I_t = \frac{\sum t_1 q_1}{\sum t_0 q_1}, \quad (10.13).$$

де t_0 , t_1 - трудові затрати на одиницю продукції в людино-годинах в базовому та звітному періодах.

Агрегатний трудовий індекс продуктивності праці розраховується за допомогою такої формули:

$$I_{\frac{1}{t}} = \frac{\sum t_0 q_1}{\sum t_1 q_1}, \quad (10.14).$$

де $t_1 q_1$ фактичні витрати часу на виготовлення одиниці продукції в звітному періоді; $t_0 q_1$ - загальні витрати часу на виготовлення продукції в звітному періоді при трудомісткості базового. Наведемо приклад розрахунку індексу трудомісткості (див. табл. 10.3).

Таблиця 10.3

Виробництво продукції та витрати часу

| Види продукції | Виготовлено продукції в звітному періоді, шт. | Витрати часу на виробництво одиниці продукції, людино-годин (t) | |
|----------------|---|---|----------------|
| | | базовий період | звітний період |
| Телефони | 540 | 43 | 40 |
| Комп'ютер | 250 | 90 | 92 |
| Телевізори | 170 | 145 | 125 |

Виходячи з умов приведеної задачі розрахуємо загальні індекси:

а) трудомісткості

$$I_t = \frac{\sum t_1 q_1}{\sum t_0 q_1} = \frac{40 \cdot 540 + 92 \cdot 250 + 125 \cdot 170}{43 \cdot 540 + 90 \cdot 250 + 145 \cdot 170} = \frac{65850}{70370} = 0,935;$$

б) продуктивності праці:

$$I_{\frac{1}{t}} = \frac{\sum t_0 q_1}{\sum t_1 q_1} = \frac{43 \cdot 540 + 90 \cdot 250 + 145 \cdot 170}{40 \cdot 540 + 92 \cdot 250 + 125 \cdot 170} = \frac{70370}{65850} = 1,068,$$

або $I_{\frac{1}{t}} = \frac{1}{I_t} = \frac{1}{0,935} = 1,068.$

Індекс трудомісткості засвідчує, що на заводі в звітному періоді зниження затрат праці на виробництво одиниці продукції становило 6,5%, завдяки чому продуктивність праці зросла на 6,8%. В цілому ж економія часу складає 30736 людино/годин.

10.3 Середньозважені індекси. Факторний аналіз

Інколи агрегатний індекс обчислюють за допомогою індивідуальних індексів. Це виникає тоді, коли по елементах, які включає індекс, невідомі окремі значення q , p , z , t . Натомість приводяться індивідуальні індекси i_q , i_p , i_z , i_t , а також добутки $p_1 q_1$, $p_0 q_0$, $z_1 q_1$. За таких умов використовують середні індекси. Обчислення середнього індексу на основі індивідуальних забезпечується шляхом розрахунку **середнього арифметичного** або **середнього гармонійного** з індивідуальних індексів. Складність для застосування середнього індекса полягає у встановленні ваг для індивідуальних індексів, які повинні бути підібрані таким чином, щоби забезпечувалась тотожність середніх індексів агрегатному.

Для фізичного обсягу продукції середній арифметичний індекс розраховують за допомогою формули:

$$I_q = \frac{\sum i_q \cdot q_0 p_0}{\sum q_0 p_0}, \quad (10.15).$$

де $i_q = \frac{q_1}{q_0}$ - індивідуальні індекси фізичного обсягу; $q_0 p_0$ - вартість продукції базового періоду в цінах цього ж періоду. Формула вищенаведеного середнього індексу тотожна агрегатному: $q_1 = i_q q_0$.

$$I_q = \frac{\sum \frac{q_1}{q_0} \cdot q_0 p_0}{\sum q_0 p_0} = \frac{\sum q_1 p_0}{\sum q_0 p_0}. \quad (10.16).$$

Наведемо приклад розрахунку середньоарифметичного індексу (табл. 10.4)

Таблиця 10.4

Дані для розрахунку середньоарифметичного індексу

| Види меблів | Зміна обсягу виробництва в квітні в порівнянні з березнем, % $i_q q_0$ | Виробництво продукції в березні, тис. грн. |
|-------------|--|--|
| Столи | + 12 | 20 |
| Дивани | +10 | 50 |
| Стільці | +15 | 30 |

Визначити зміну обсягу виробництва меблів у квітні порівняно з березнем (у %).

Розв'язок:

Виходячи з вихідних даних загальний індекс фізичного обсягу може бути розрахований як середній арифметичний, тобто

$$I_q = \frac{\sum i_q \cdot q_0 p_0}{\sum q_0 p_0} = \frac{1,12 \cdot 20 + 1,1 \cdot 50 + 1,15 \cdot 30}{20 + 50 + 30} = 1,119.$$

Одержаний результат засвідчує, що в цілому по підприємству виробництво меблів в квітні порівняно з березнем зросло на 11,9%.

Аналогічно проведемо перетворення агрегатного індексу ціни в середній гармонійний.

$$i_p = \frac{p_1}{p_0}, \quad p_0 = \frac{p_1}{i_q};$$

$$I_p = \frac{\sum p_1 q_1}{\sum \frac{p_1 q_q}{i_p}} - \text{середньогармонійний індекс ціни.} \quad (10.17).$$

Приклад. Визначити середню зміну цін на борошно в другому кварталі в порівнянні з першим на основі таких даних (табл.10.5).

Таблиця 10.5

Дані для розрахунку середнього гармонійного індексу

| Сорт борошна | Зростання цін, % | Продано в другому кварталі, тис. грн. |
|--------------------|------------------|---------------------------------------|
| Борошно I-го сорту | 12 | 50 |
| Борошно II-го | 15 | 70 |

Розв'язок: В даному випадку загальний індекс цін розраховується за допомогою середньогармонійного індексу:

$$I_p = \frac{\sum p_1 q_1}{\sum \frac{p_1 q_q}{i_p}} = \frac{50 + 70}{\frac{50}{1,12} + \frac{70}{1,15}} = \frac{120}{45 + 61} = \frac{120}{106} = 1,137.$$

Одержаний результат засвідчує, що ціни на борошно в звітному періоді в середньому зросли на 13,7%. Аналогічно розраховують середні індекси собівартості та продуктивності праці.

Статистичне дослідження явищ і процесів суспільного життя часто забезпечується завдяки використанню взаємозв'язку індексів. Система взаємозв'язаних індексів дає можливість вивчати зв'язки між явищами, а також виявити вплив окремих факторів на зміну складного явища.

Виходячи з суті економічних індексів, взаємозв'язок має місце між загальними індексами товарообігу, ціни та фізичного обсягу продукції, тобто:

$$I_{pq} = I_p \cdot I_q, \text{ або } \frac{\sum p_1 q_1}{\sum p_0 q_0} = \frac{\sum p_1 q_1}{\sum p_0 q_1} \cdot \frac{\sum p_0 q_1}{\sum p_0 q_0}. \quad (10.18).$$

У взаємозв'язку знаходяться між собою також абсолютні величини, котрі розраховані на основі системи взаємозв'язаних індексів. При цьому:

$$\Delta_{pq} = \Delta_p + \Delta_q \text{ або } \Delta_{pq} = \sum p_1 q_1 - \sum p_0 q_0 ; \quad (10.19).$$

$$\Delta_p = \sum p_1 q_1 - \sum p_0 q_1 ; \quad \Delta_q = \sum p_0 q_1 - \sum p_0 q_0 . \quad (10.20).$$

Статистична оцінка затрат базується на взаємозв'язку індексів собівартості та фізичного обсягу продукції, зваженого за собівартістю базового періоду, тобто

$$I_{zq} = I_z \cdot I_q , \quad (10.21).$$

звідси

$$\frac{\sum z_1 q_1}{\sum z_0 q_0} = \frac{\sum z_1 q_1}{\sum z_0 q_1} \cdot \frac{\sum z_0 q_1}{\sum z_0 q_0} . \quad (10.22).$$

Загальний індекс продуктивності праці (трудоий) дорівнює відношенню індексу фізичного обсягу продукції, зваженого за трудовими затратами базового періоду, до індексу трудових затрат:

$$I_{\frac{1}{t}} = I_q : I_t \text{ або } \frac{\sum t_0 q_1}{\sum t_1 q_1} = \frac{\sum q_1 t_0}{\sum q_0 t_0} : \frac{\sum t_1 q_1}{\sum t_0 q_0} \quad (10.23).$$

Вище нами розглядався взаємозв'язок індексів, в основі якого було покладено два чинники. Разом з тим, статистична оцінка взаємозв'язку може охоплювати дію трьох, чотирьох і більше факторів. Зазначене обумовлює потребу розкладання загального індексу на два, три і більше факторних індексів. Зокрема, індекс валового збору зернових культур залежить від величини урожайності (y), розміру посівних площ (n) і структури посівних площ (h). Загальний індекс валового збору дорівнює:

$$I_{в.з.} = \frac{\sum y_1 n_1 h_1}{\sum y_0 n_0 h_0} . \quad (10.24).$$

1. Індекс урожайності зернових обчислюють за формулою:

$$I_y = \frac{\sum y_1 n_0 h_0}{\sum y_0 n_0 h_0} . \quad (10.25).$$

2. Індекс розміру посівних площ дорівнює:

$$I_{n.n.} = \frac{\sum y_1 n_1 h_0}{\sum y_1 n_0 h_0} = \frac{\sum n_1}{\sum n_0} . \quad (10.26).$$

3. Індекс структури посівних площ дорівнює:

$$I_{c.n.} = \frac{\sum y_1 n_1 H_1}{\sum y_1 n_1 H_0}. \quad (10.27).$$

Взаємозв'язок чотирьох вищенаведених індексів можна виразити так:

$$\frac{\sum y_1 n_1 H_1}{\sum y_0 n_0 H_0} = \frac{\sum y_1 n_0 H_0}{\sum y_0 n_0 H_0} \cdot \frac{\sum y_1 n_1 H_0}{\sum y_1 n_0 H_0} \cdot \frac{\sum y_1 n_1 H_1}{\sum y_1 n_1 H_0}. \quad (10.28).$$

Позначимо факторні ознаки буквами a , b , c . Тоді систему взаємозв'язаних індексів можна записати так:

$$\frac{\sum a_1 b_1 c_1}{\sum a_0 b_0 c_0} = \frac{\sum a_1 b_0 c_0}{\sum a_0 b_0 c_0} \cdot \frac{\sum a_1 b_1 c_0}{\sum a_1 b_0 c_0} \cdot \frac{\sum a_1 b_1 c_1}{\sum a_1 b_1 c_0}. \quad (10.29).$$

Оцінка взаємозв'язку значною мірою залежить від послідовності факторів, які розміщені в формулі. Зокрема, індекс результативного показника буде дорівнювати добуткові факторних індексів лише при дотриманні певних вимог. Першим фактором - співмножником буде той, чисельник розрахункової формули якого є чисельником результативного показника; у наступного фактора - співмножника чисельником розрахункової формули є знаменником першого фактора, і так враховують усі фактори.

Правильність нашого твердження засвідчує формула взаємозв'язку трьох факторів, які нами позначено відповідно a , b , c . Отже,

$$I_{e.z.} = I_a \cdot I_b \cdot I_c. \quad (10.30).$$

Абсолютний вплив зміни будь-якого фактора на динаміку загального результату визначається як різниця між чисельником і знаменником індексу, який прийнятий для розрахунку.

Як нами зазначалося вище, індекси відображають зміну суспільних явищ не лише в часі, а й у просторі. Показники, що характеризують співвідношення рівнів складних економічних явищ у просторі, називаються територіальними індексами. При обчисленні територіальних індексів за базу порівняння можуть бути прийняті показники будь-якої з порівнюваних територій.

10.4 Індeksi змінного, постійного складу та структурних зрушень

Часто динаміка суспільних явищ може задаватися за допомогою середніх величин: середнім виробітком, середньою заробітною платою, середньою собівартістю тощо. На зміну динаміки середніх рівнів можуть впливати зміна індексованої величини в окремих елементах досліджуваного явища, так і структура явища.

Зокрема, середня продуктивність праці може зростати за рахунок підвищення її у робітників окремих спеціальностей і збільшення питомої ваги робітників з вищою продуктивністю праці у загальній чисельності працюючих.

Динаміка середньої урожайності зернових в районі залежить від зростання середньої урожайності у кожному господарстві і збільшення питомої ваги суб'єктів господарювання з вищою урожайністю зернових у загальній їх кількості.

З наведених прикладів видно, що на зміну середнього значення явища, що вивчається, впливає зміна динаміки середнього рівня і структура. Розрахунок впливу вищенаведених факторів на середню здійснюється за допомогою індексів змінного, постійного (фіксованого) складу та структурних зрушень.

Наведемо приклад побудови таких індексів.

Приводяться дані по двох паперових фабриках щодо обсягів виробництва та собівартість паперу за два періоди (табл. 10.6):

Таблиця 10.6

Розрахунок індексів середніх величин

| Фабрика | Виготовлено, тонн | | Собівартість 1 тонни, ум.о. | |
|---------|---------------------------|---------------------------|-----------------------------|---------------------------|
| | в базовому періоді, q_0 | в звітному періоді, q_1 | в базовому періоді, z_0 | в звітному періоді, z_1 |
| №1 | 50000 | 80000 | 150 | 135 |
| №2 | 60000 | 40000 | 250 | 230 |

Визначити: 1. Індивідуальні індекси собівартості 1 т паперу по кожній фабриці зокрема. 2. Загальні індекси собівартості паперу змінного, постійного складу та структурних зрушень.

Розв'язок:

1. Індивідуальні індекси собівартості паперу по кожній фабриці становлять:

а) фабрика № 1 $i_z = 135 : 150 = 0,90$;

б) фабрика № 2 $i_z = 230 : 250 = 0,92$.

Собівартість паперу на фабриці № 1 знизилась на 10, а на фабриці № 2 - на 8 відсотків.

Розрахунок середнього індексу собівартості змінного складу забезпечується завдяки визначенню середньої собівартості паперу по двох фабриках у звітному і базовому періодах. В звітному періоді середня собівартість 1 т паперу становила:

$$\bar{z}_1 = \frac{\sum z_1 q_1}{\sum q_1} = \frac{135 \cdot 80000 + 230 \cdot 40000}{80000 + 40000} = \frac{20000000}{120000} = 166,7 \text{ ум.о.}$$

Середня собівартість 1 т паперу в базовому періоді:

$$\bar{z}_0 = \frac{\sum z_0 q_0}{\sum q_0} = \frac{150 \cdot 50000 + 250 \cdot 60000}{50000 + 60000} = \frac{22500000}{110000} = 204,5 \text{ ум.о.}$$

Загальний індекс змінного складу дорівнює:

$$I_z^{z.c.} = \bar{z}_1 : \bar{z}_0 = 166,7 : 204,5 = 0,815,$$

тобто середня по двох фабриках собівартість 1 т паперу знизилась на 18,5 %.

Зниження середньої собівартості паперу є результатом не тільки зменшення собівартості паперу на кожній фабриці, а й впливу структурного фактора-збільшення випуску більш дешевого паперу на фабриці № 1, на якій собівартість однієї тонни паперу значно нижча, ніж на фабриці №2.

Зміна собівартості по двох фабриках разом лише за рахунок її коливань на окремих фабриках обчислюється за допомогою індексу постійного складу, тобто:

$$\bar{I}_z^{n.c.} = \frac{\sum z_1 q_1}{\sum q_1} : \frac{\sum z_0 q_1}{\sum q_1} = \frac{\sum z_1 q_1}{\sum z_0 q_1} = \frac{135 \cdot 80000 + 230 \cdot 40000}{150 \cdot 80000 + 250 \cdot 40000} = \frac{20000000}{22000000} = 0,909.$$

Індекс постійного складу показує, що в середньому по двох фабриках собівартість знизилась на 9,1%.

Поділивши індекс змінного складу на індекс постійного складу, одержимо індекс структурних зрушень, який засвідчує як змінилась середня величина індексованого показника за рахунок структурного фактора.

В нашому прикладі:

$$\bar{I}_z^{c.z.} = \bar{I}_z^{з.с.} : \bar{I}_z^{n.с.} = 0,815 : 0,909 = 0,897.$$

Одержаний результат засвідчує, що середня собівартість 1 т паперу знизилась на 10,3% за рахунок зміни питомої ваги фабрик у випуску продукції.

Систему взаємозв'язаних індексів для аналізу динаміки середніх показників можна подати так:

$$\frac{\sum x_1 f_1}{\sum f_1} : \frac{\sum x_0 f_0}{\sum f_0} = \left(\frac{\sum x_1 f_1}{\sum f_1} : \frac{\sum x_0 f_1}{\sum f_1} \right) \cdot \left(\frac{\sum x_0 f_1}{\sum f_1} : \frac{\sum x_0 f_0}{\sum f_0} \right) \text{ або } I_{з.с.} = I_{n.с.} \times I_{c.z.}, \quad (10.31).$$

де x_1 і x_2 - рівні осереднюваного показника, f_0 і f_1 - частоти осереднюваних показників.

Тестові завдання

1. За способом побудови загальні індекси поділяють на:

- а) агрегатні та середні;
- б) групові та зведені;
- в) планові та звітні;
- г) ланцюгові та базові.

2. За характером порівнянь (у часі, просторі) індекси поділяють на:

- а) планові та звітні;
- б) динамічні, територіальні та групові;
- в) базові та ланцюгові;
- г) загальні та індивідуальні.

3. За ступенем агрегованості інформації індекси поділяють на:

- а) індивідуальні та загальні;
- б) загальні та середні;
- в) індивідуальні та зведені;
- г) планові та фактичні.

4. Частка від ділення наступного базового індексу на попередній дорівнює:

- а) загальному ;
- б) індивідуальному;
- в) груповому;
- г) ланцюговому.

5. Загальний індекс складається з елементів:

- а) двох;
- б) трьох;
- в) чотирьох;
- г) одного.

6. Агрегатна форма індексу використовується для:

- а) узагальнюючої характеристики однорідної сукупності;
- б) узагальнюючої характеристики неоднорідної сукупності;
- в) розрахунку індивідуальних індексів;
- г) визначення тісноти зв'язку між індексованими величинами.

7. Величина загального індексу товарообігу залежить від:

- а) ціни та собівартості;
- б) обсягу продукції та витрат часу;
- в) ціни та кількості реалізованих товарів;
- г) виробітку та часу.

8. Величина загального індексу фізичного обсягу продукції залежить від:

- а) ціни та собівартості продукції;
- б) обсягу продукції та витрат часу;
- в) ціни та кількості реалізованих товарів;
- г) ціни та кількості виготовленої продукції.

9. Величина загального індексу ціни залежить від

- а) ціни та обсягу реалізованої продукції;
- б) ціни та собівартості;
- в) ціни та виробітку;
- г) ціни та витрат часу на виготовлення товару.

10. Величина загального індексу собівартості залежить від:

- а) ціни та затрат часу на виготовлення товару;

- б) рівня собівартості та обсягу виготовленого товару;
- в) рівня продуктивності праці;
- г) трудомісткості продукції.

11. Величина загального індексу трудомісткості залежить від:

- а) обсягу виготовленої продукції та ціни;
- б) обсягу реалізованої продукції та ціни;
- в) затрат часу та обсягу виготовленої продукції;
- г) собівартості та ціни виготовленої продукції.

12. Агрегатний трудовий індекс продуктивності праці складається з таких двох співмножників:

- а) ціни та затрат часу;
- б) ціни та фізичного обсягу виготовленої продукції;
- в) собівартості та обсягу виготовленої продукції;
- г) затрат часу та обсягу виготовленої продукції в звітному періоді.

13. Виходячи з суті економічних індексів взаємозв'язок має місце між

загальними індексами:

- а) товарообігу, ціни та фізичного обсягу продукції;
- б) ціни, фізичного обсягу продукції та собівартості ;
- в) продуктивності праці, товарообігу та собівартості;
- г) товарообігу, трудомісткості та затрат.

14. Статистична оцінка затрат базується на взаємозв'язку індексів:

- а) продуктивності праці та трудомісткості;
- б) собівартості та фізичного обсягу продукції;
- в) ціни та фізичного обсягу продукції;
- г) товарообігу та ціни.

15. Індекс продуктивності праці (трудовий) дорівнює відношенню індексу:

- а) трудових затрат до індексу фізичного обсягу продукції;
- б) середнього виробітку до індекса трудомісткості;
- в) фізичного обсягу продукції, зваженого за трудовими затратами базового періоду до індексу трудових затрат;
- г) собівартості до індексу фізичного обсягу продукції.

Типові задачі

Задача 1. Відомі такі дані про реалізацію бензину А-95 на трьох АЗС міста:

| № АЗС | Ціна за 1 л, грн. | | Реалізовано бензину, тис. л | |
|-------|-------------------|-------------|-----------------------------|-------------|
| | Базисний рік | Звітний рік | Базисний рік | Звітний рік |
| 1 | 4,0 | 5,3 | 150 | 200 |
| 2 | 4,2 | 5,5 | 220 | 350 |
| 3 | 4,2 | 5,2 | 340 | 390 |

Визначіть індивідуальні індекси цін, фізичного обсягу та товарообороту. Зробіть висновки.

Задача 2. Маємо дані про реалізацію товарів магазинами ТОВ «Кераміка». Визначіть індивідуальні індекси цін, фізичного обсягу та товарообороту. Зробіть висновки.

| Товар | Середня ціна одиниці товару, грн. | | Обсяг реалізованого товару, шт. | |
|--------------|-----------------------------------|-------------|---------------------------------|-------------|
| | Базисний рік | Звітний рік | Базисний рік | Звітний рік |
| Ванни чавуні | 2500 | 2600 | 50 | 52 |
| Змішувачі | 810 | 950 | 150 | 200 |
| Мийки | 260 | 300 | 190 | 210 |

Задача 3. За даними про обсяг виробництва продукції та її собівартість (табл. 6.8) визначіть індивідуальні індекси собівартості, фізичного обсягу продукції та витрат. Зробіть висновки.

| Продукція | Обсяг виробленої продукції, тис. шт. | | Собівартість одиниці, грн. | |
|-----------|--------------------------------------|-------------|----------------------------|-------------|
| | Базисний рік | Звітний рік | Базисний рік | Звітний рік |
| А | 80 | 90 | 22 | 20 |
| Б | 65 | 110 | 20 | 18 |
| В | 560 | 610 | 80 | 94 |

Задача 4. Маємо такі дані про реалізацію продукції (табл.). Визначіть загальні агрегатні індекси цін, фізичного обсягу, доходу (виручки). Покажіть взаємозв'язок між індексами. Розрахуйте абсолютну зміну обсягу доходу (виручки) у звітному періоді загалом та за рахунок зміни: а) цін; б) фізичного обсягу реалізації. Зробіть висновки.

| Продукція | Ціна за 1 шт., грн. | | Обсяг реалізованої продукції, тис. шт. | |
|-----------|---------------------|----------------|--|----------------|
| | Базисний період | Звітний період | Базисний період | Звітний період |
| AP-2 | 25 | 23 | 45 | 53 |
| CM-14 | 20 | 18 | 13 | 18 |

Задача 5. Відомі дані про реалізацію трьох видів продукції:

| Продукція | Ціна за 1 шт., грн. | | Обсяг реалізованої продукції, тис. шт. | |
|-----------|---------------------|-------------|--|-------------|
| | Базисний рік | Звітний рік | Базисний рік | Звітний рік |
| ЛБ-20 | 2,5 | 2,3 | 45 | 53 |
| ЛД-20 | 2,0 | 1,8 | 13 | 18 |
| ЛБ-40 | 14,0 | 14,5 | 10 | 14 |

Визначіть загальні індекси цін, фізичного обсягу, доходу (виручки). Покажіть взаємозв'язок між індексами. Розрахуйте абсолютну зміну обсягу доходу (виручки) у звітному періоді загалом та за рахунок зміни: а) цін; б) фізичного обсягу реалізації. Зробіть висновки.

Задача 6. За результатами обстеження одержано дані про реалізацію окремих товарів у торговельній мережі міста.

| Товар | Реалізовано товару, шт. | | Середня ціна одиниці товару, грн. | |
|----------------|-------------------------|-------------|-----------------------------------|-------------|
| | Базисний рік | Звітний рік | Базисний рік | Звітний рік |
| Холодильники | 50 | 54 | 2400 | 2500 |
| Світильники | 140 | 180 | 800 | 850 |
| Телевізори | 280 | 305 | 1200 | 1350 |
| Пральні машини | 310 | 330 | 1700 | 1950 |

Визначіть загальні агрегатні індекси цін, фізичного обсягу, доходу (виручки). Покажіть взаємозв'язок між індексами. Розрахуйте абсолютну зміну обсягу доходу (виручки) у звітному періоді загалом та за рахунок зміни: а) цін; б) фізичного обсягу реалізації. Зробіть висновки.

Задача 7. Відомі дані про реалізацію товарів мережею магазинів.

| Товарні групи | Реалізовано товарів | Зміна кількості реалізованих |
|---------------|---------------------|------------------------------|
|---------------|---------------------|------------------------------|

| | у базисному місяці, тис. грн. | товарів у звітному місяці порівняно з базисним, % |
|----------------|----------------------------------|--|
| Електротовари | 450 | – 6 |
| Побутова хімія | 750 | + 15 |
| Інструменти | 350 | + 8 |

Визначіть загальний індекс фізичного обсягу, а також загальний індекс цін, якщо відомо, що товарооборот у порівнянні з базисним місяцем зріс на 10%. Розрахуйте абсолютну зміну товарообороту, зумовлену зміною фізичного обсягу реалізації. Зробіть висновки.

Задача 8. Відомі такі дані по підприємству:

| Виріб | Витрати на виробництво продукції у звітному періоді, тис. грн. | Зміна собівартості виробництва одного виробу в звітному періоді порівняно з базисним, % |
|-------|---|---|
| А | 1550 | – 8 |
| Б | 940 | + 5 |
| В | 450 | Без змін |

На основі наведених даних визначіть загальний індекс собівартості продукції й абсолютну суму економії (перевитрат) від зміни собівартості продукції, а також загальний індекс витрат на виробництво за умови, що кількість готових виробів у звітному періоді зросла у 1,5 рази. Зробіть висновки.

Задача 9. Маємо дані про реалізацію товарів трьома магазинами (табл. 6.14). Обчисліть загальний індекс цін. Використовуючи взаємозв'язок індексів визначіть, на скільки процентів збільшилася кількість реалізованих товарів, якщо відомо, що товарооборот зріс на 12 %. Визначіть абсолютний розмір економії чи перевитрат населення від зміни цін. Зробіть висновки.

| № магазину | Обсяг товарообороту в звітному періоді, тис. грн. | Середній процент зміни цін у звітному періоді порівняно з базисним |
|------------|--|--|
| 26 | 985 | + 5 |
| 48 | 450 | Без змін |
| 56 | 340 | – 12 |

Задача 10. Відомі дані про виробництво однойменної продукції та її собівартість по трьох підприємствах (табл. 6.15).

Визначіть індекси собівартості змінного, постійного складу та структурних зрушень. Покажіть взаємозв'язок між цими індексами та зробіть висновки.

| № підприємств а | Обсяг виробництва продукції, тис. шт. | | Собівартість одиниці продукції, грн. | |
|--------------------|---------------------------------------|-------------|--------------------------------------|-------------|
| | Базисний рік | Звітний рік | Базисний рік | Звітний рік |
| 1 | 85 | 93 | 208 | 189 |
| 2 | 77 | 88 | 180 | 155 |
| 3 | 65 | 70 | 980 | 1020 |

Питання для самоконтролю

1. Дайте визначення індексу.
2. Види індивідуальних індексів.
3. Як класифікуються індекси?
4. Наведіть формули індивідуальних індексів.
5. Найважливіші елементи індексу.
6. Напишіть формули найважливіших загальних індексів.
7. Яким чином пов'язані між собою чисельник і знаменник індексу?
8. Дайте тлумачення агрегатної форми індексу.
9. В чому полягає різниця між індексом трудомісткості та індексом продуктивності праці?
10. Поясніть зміст середньоарифметичного та середньо гармонійного індексів.
11. Для чого використовують взаємозв'язок індексів?
12. Назвіть найважливіші види взаємозв'язків індексів.
13. Наведіть приклад взаємозв'язку абсолютних величин загальних індексів.
14. Як слід розуміти багатofакторний взаємозв'язок індексів?
15. Наведіть формули індексів собівартості змінного, постійного складу та структурних зрушень.

РОЗДІЛ 11. ВИБІРКОВИЙ МЕТОД

11.1 Суть і значення вибіркового спостереження

Вибірковий метод дослідження використовують у тих випадках, коли проведення суцільного спостереження унеможливлене або його вважають недоцільним. **Вибіркове спостереження** - найбільш поширений вид несцільного спостереження, при якому досліджується не вся сукупність, а лише її частина, що відібрана за встановленими правилами відбору, завдяки яким забезпечується репрезентативність сукупності в цілому.

Значення вибіркового спостереження полягає в наступному:

по-перше, воно є оперативнішим за суцільне, позаяк значно скорочує строки проведення робіт;

по-друге, воно дає значну економію, оскільки скорочується обсяг робіт, а також матеріальні і грошові затрати;

по-третє, результати вибіркового спостереження часто точніші, ніж суцільного, так як для його проведення можна підібрати більш кваліфікованих виконавців, простіше організувати контроль;

по-четверте, вибірковий метод застосовується при неможливості проведення суцільного спостереження (великий обсяг сукупності, дослідження пов'язане з псуванням матеріалів і т.п.).

Частина одиниць, яка відібрана для спостереження за певними встановленими правилами називається **вибірковою сукупністю**, а вся сукупність одиниць, із якої проводиться відбір, - **генеральною**.

Узагальнюючими показниками генеральної сукупності слугують: **середня** \bar{x} , **дисперсія** - σ^2 та **p - доля**, котра обчислюється шляхом ділення числа M одиниць, які мають досліджувану властивість, на всю чисельність генеральної сукупності, тобто

$$p = \frac{M}{N}, \quad (11.1).$$

де N - обсяг генеральної сукупності.

Узагальнюючими характеристиками для вибіркової сукупності слугують **вибіркова середня** (\tilde{x}), **дисперсія** (σ^2) та **доля** (w). Зокрема, доля розраховується за допомогою формули (11.2).

$$w = \frac{m}{n}, \quad (11.2).$$

де m - кількість одиниць ознаки, що мають дану властивість.

Різниця $\tilde{x} - \bar{x} = \Delta_x$ **називається похибкою репрезентативності вибіркової середньої**; $w - p = \Delta_w$ - **похибка частки**; $\sigma_s^2 - \sigma^2 = \Delta\sigma^2$ - **похибка дисперсії**.

Похибка може бути із знаком “плюс” або “мінус”. Середня похибка дорівнює середній квадратичній з усіх можливих конкретних похибок стохастичного процесу (11.3).

$$\bar{\Delta} = \sqrt{\frac{\sum \Delta^2 n_i}{n_i}}, \quad (11.3).$$

де n_i - кількість різномовірних можливих результатів вибірки, завдяки котрим проявляється дана конкретна похибка. Конкретна похибка - різниця між величиною вибіркової і генеральної середньої.

11.2 Похибки вибіркового спостереження. Способи відбору.

Результати вибіркового спостереження не завжди відповідають даним, які би точно відтворювали найважливіші характеристики генеральної сукупності. Причина невідповідності полягає, перш за все в обстежені не всієї сукупності, а лише її частини. Саме тому при проведенні вибіркового спостереження можуть виникати погрішності, похибки. Різниця між показниками вибіркової й генеральної сукупності називається **похибками вибірки**, котрі поділяються на **похибки реєстрації** і **похибки репрезентативності**, тобто представництва.

Похибки реєстрації виникають у результаті використання помилкових або неточних даних. Вони поділяються на **систематичні**, що зумовлені причинами, котрі діють в одному напрямку й тому неправильно відтворюють явище, що вивчається (наприклад, округлення років) і **випадкові**.

Перші, спричинені результатами порушення принципу випадковості відбору одиниць сукупності для статистичного спостереження (наприклад, розрахунок середньої успішності по групі на основі екзаменаційних оцінок студентів-відмінників).

Забезпечення репрезентативності вибірки в статистиці досягається завдяки дотриманню принципу випадковості відбору одиниць. Його суть полягає в тому, що можливість залучення у вибірку сукупність або навпаки вилучення одиниці сукупності з неї не може вплинути який-небудь інший фактор, крім випадкового. Розрізняють такі види відбору: **індивідуальний, груповий і комбінований**.

При **індивідуальному** відбирається тільки одна одиниця, а число разів повторюється відповідно до кількості відбору. При **груповому** кожний раз відбирається група одиниць, серій. **Комбінований** відбір передбачає поєднання індивідуального з груповим.

За схемами відбору вибірку поділяються на повторну та безповторну. За способами відбору виокремлюють просту випадкову, систематичну (механічну), районовану та серійну. **Власне випадковий відбір**, або випадкова вибірка здійснюється за допомогою жереба або таблиць випадкових чисел. В першому випадку всім елементам генеральної сукупності присвоюється порядковий номер і на кожний елемент заводиться жереб, пронумеровані кульки або картки, фішки, котрі поміщають у скриньку і перемішують, а потім витягують для зчитування інформації. Використання таблиць випадкових чисел полягає в проведенні відбору випадкових чисел (із спеціальних таблиць), які утворюють порядкові номери для відбору.

При **систематичному відборі** (механічна вибірка) вибір елементів здійснюється через рівні інтервали. Всі одиниці сукупності, що вивчаються, попередньо розміщують у певному порядку (наприклад, за алфавітом, табельним номером, тарифним розрядом тощо). Залежно від кількості одиниць генеральної сукупності, що має бути відібрана для вибірки може бути використана 5, 10, 20% і інше відсоткове значення вибірки. Зокрема, якщо вибірка 10% і проводиться з 100 чол., це означає, що із генеральної сукупності відбирається кожний десятий - 1; 11; 21 і т. д., якщо 5% із 100 чол., то відбирається кожний двадцятий. Механічна вибірка завжди неповторювана.

Районований (типовий) відбір - це спосіб формування вибірки з урахуванням структури генеральної сукупності, котру попередньо структуризують за певною суттєвою ознакою. Після цього з кожної групи простим випадковим або механічним способом відбирається певна кількість досліджуваних одиниць пропорційно питомій вазі групи в усій сукупності. Така вибірка може бути повторною і неповторною

Наприклад, для оцінювання загальної успішності по університету потрібно провести відбір 1500 студентів із 10000, які навчаються на чотирьох факультетах.

Таблиця 11.1

Розподіл студентів університету за факультетами

| Назва факультету | Число студентів, чол. | Питома вага студентів у загальній чисельності, % | Відібрано студентів для проведення вибіркового спостереження, чол. |
|------------------|-----------------------|--|--|
| Економічний | 1800 | 18 | $\frac{1500 \cdot 18}{100} = 270$ |
| Юридичний | 2400 | 24 | $\frac{1500 \cdot 24}{100} = 360$ |
| Філософський | 3200 | 32 | $\frac{1500 \cdot 32}{100} = 480$ |
| Математичний | 2600 | 26 | $\frac{1500 \cdot 26}{100} = 390$ |
| Всього: | 10000 | 100 | 1500 |

В даному прикладі відбір студентів проведено з урахуванням питомої ваги студентів кожного факультету в загальній їх кількості по університету. Використання зазначеного критерію дає можливість регулювати відбір відповідно до числа студентів, які навчалися на факультетах навчального закладу. Зокрема, на економічному факультеті, де навчається лише 1800 студентів, для проведення вибіркового спостереження було вибрано лише 270 студентів, на філософському з 3200 відповідно 480 студентів.

Серійний або гніздовий вибір полягає в тому, що в порядку проведення власне випадкової або механічної вибірки відбирають не одиниці, а певні серії (гнізда), всередині котрих проводять суцільне спостереження.

Об'єктивність результатів вибіркового спостереження значною мірою залежать від схем відбору. При повторному відборі кожна одиниця або серія, що потрапила у вибірку з генеральної сукупності, має шанс ще раз потрапити у вибірку. При цьому ймовірність потрапляння у вибірку сукупності усіх одиниць генеральної сукупності залишається однаковою. Неповторюваний відбір означає, що кожна відібрана одиниця (або серія) не повертається в генеральну сукупність і не може бути зареєстрована у вибірці повторно. Використання такої схеми відбору збільшує ймовірність для інших одиниць потрапити у вибірку.

В практиці проведення вибірки інколи використовують моментні спостереження. **Моментним** називається спостереження, при якому дослідженню підлягають усі елементи сукупності (суцільне спостереження) на певні моменти часу (вибіркове спостереження за часом). Наприклад, в цеху працювало 25 робітників. Перевірка їх роботи проводилась через кожні півгодини за 8-годинну зміну, тобто 16 раз. Число замірів становить $25 \times 16 = 400$. В 360 випадках робітники працювали, а в 40 - ні. Частка робітників, які працювали постійно за зміну склала 0,9 (360/400), або 90 % усього робочого часу.

В практиці статистичних досліджень інколи виникає потреба поєднати різні схеми вибіркового методу. Таке поєднання можливе в рамках **багатоступеневої** вибірки. Вибірка називається багатоступеневою, коли відбір сукупності проводили за ступенями, послідовними стадіями, причому кожен ступінь, стадія відбору має свою одиницю відбору. Залежно від кількості ступенів відбору вибірка може бути дво - три і більше ступеневою. Зокрема, відбираючи загальну кількість сімей для дослідження середньомісячного доходу, спочатку встановлюють загальну кількість сімей для відбору. Потім сім'ї поділяють на міські і сільські. Перший ступінь - загальне число сімей, другий – поділ їх на міські і сільські.

Повнота охоплення обсягу сукупності та схеми відбору одиниць на різних ступенях відрізняється. Наприклад, проведення вибіркового дослідження гранично допустимого рівня забрудненості атмосфери в цілому по країні може охоплювати такі послідовні ступені:

перша - регіони держави: охоплюються вибірково 40 % ($D_1 = 0,4$);

друга - райони в областях: охоплюються вибірково до 20 % їх кількості в тих областях, які відібрані на першому ступені ($D_2 = 0,2$);

третья - суб'єкти господарювання, котрі є найбільшими забруднювачами повітря: обстежують 5 % їх чисельності в тих районах, які потрапили до вибірки на другому ступені ($D_3 = 0.05$).

Частка вибіркової сукупності від генеральної становить: $D=0.4 \cdot 0.2 \cdot 0.05=0.004$. Отже обстеженню підлягає 0,4 % підприємств, які є найбільшими забруднювачами атмосфери.

Обстеження, при якому дослідженню підлягає не одна, а дві і більше ознак, які значно варіюють здійснюється за допомогою **багатофазної вибірки**. Вона характеризується тим, що на всіх стадіях вибірки зберігається одна і та ж одиниця відбору, котра проходить декілька фаз вибірових досліджень, які відрізняються між собою шириною програми досліджень і обсягом вибірки. Наприклад, вибірково досліджуються джерела доходів населення

району. На першій фазі, за короткою програмою (8 питань) обстежують 30 % населення; на другій, за ширшою програмою (19 питань) - 15 % генеральної сукупності і т.д.

11.3 Основні характеристики вибіркового спостереження

Об'єктивність результатів вибіркового спостереження досягається завдяки розв'язанню таких трьох основних завдань:

а) визначені із заданою ймовірністю можливі границі похибки репрезентативності « ε » (епсілон). За умови, що фактичні границі більші заданих, результатами такої вибірки користуватись не можна;

б) визначена ймовірність « p » того, що можливі границі похибки репрезентативності не перевищать заданих величин;

в) визначений необхідний обсяг вибірки « n ».

Ці завдання розв'язуються за допомогою теорем закону великих чисел. По суті вони зводяться до визначення однієї із величин: ε , p і n в теоремі П.Чебишева, за умови, що величина \bar{x} невідома. Змыст теореми фрагментарно формулюється так: при достатньо великому числі досліджень n можна з ймовірністю, якомога близькою до одиниці, стверджувати, що різниця між середньою \tilde{x} і \bar{x} за абсолютною величиною буде завжди менша від наперед заданого числа ε і т. д.

Згідно з теоремою $p\{\tilde{x} - \bar{x} < \varepsilon\} \geq 1 - h$, за умови, що n - достатньо велике число а, ε і h - будь-які малі додатні числа.

Залежно від постановки завдання, котре потребує розв'язання в теоремі задаються відомі або невідомі величини, які необхідно визначити. При розв'язанні першого, задаються n і h , або n і p . Завдання полягає в тому, щоби визначити $\Delta x = \tilde{x} - \bar{x}$ за умови, що величина \bar{x} невідома.

При розв'язанні другого задаються n і Δx . Необхідно визначити невідому величину P . В третьому завданні відомі ε і p , необхідно визначити величину n .

Застосування вибіркового методу передбачає поширення результатів на усю сукупність. Однією з найважливіших передумов поширення даних вибіркового спостереження на генеральну сукупність є визначення величини похибок між зведеними показниками вибіркової й генеральної сукупності. Відомо, що вибіркова середня і частка є змінними величинами, і тому можуть приймати різні результати в залежності від мінливості значень одиниць сукупності, що потрапили у вибірку. В зв'язку з цим розраховується середня похибка вибірки, шляхом добування квадратного кореня з дисперсії (11.4).

$$\mu = \sqrt{\frac{\sigma^2}{n}} \quad \text{або} \quad \mu = \frac{\sigma}{\sqrt{n}}. \quad (11.4).$$

Використання формул 11.4 дає можливість встановити залежність середньої похибки вибірки від ступеня мінливості ознаки та обсягу вибірки. Середня частка ознаки у вибірковій сукупності розраховується за допомогою формули (11.5).

$$\mu = \sqrt{\frac{\bar{w}(1-\bar{w})}{n}}, \quad (11.5).$$

де n - чисельність вибірки; σ^2 - дисперсія кількісної ознаки в генеральній сукупності; $\bar{w}(1-\bar{w})$ - дисперсія частки в генеральній сукупності. Використовуючи функцію нормального розподілу, розраховують ймовірність граничної похибки певного розміру. Вона являє собою значення відомого інтеграла ймовірностей Чебишева і Ляпунова (11.6).

$$F_t = \frac{1}{\sqrt{2\pi}} \int_0^t l^{-\frac{t^2}{2}} dt. \quad (11.6).$$

При цьому відомо, що ймовірність того, що в окремій вибірці похибка не перевищить 2μ становить 0,954. При імовірності 0,990 похибка буде становити $\pm 2,58 \mu$, при 0,997 – $\pm 3\mu$.

Гранична похибка вибірки для певної ймовірності визначається за формулою: $\Delta = t\mu$, де t — коефіцієнт довіри, якому відповідає ймовірність P .

У практиці статистичних досліджень інколи доводиться мати справу із сукупністю, чисельність елементів якої не перевищує 20.

Це мала вибірка. При обчисленні похибки вибірки в цьому випадку використовують дисперсію не генеральної, а вибіркової сукупності.

Оскільки $\sigma_r^2 = \sigma^2 \frac{n}{n-1}$, то у великих за обсягом вибіркових сукупностях ці дисперсії збігаються. За умови, що вибірка малочисельна, в формулу похибки вибірки вноситься поправка

$$\frac{n}{n-1}. \quad (11.7).$$

Для повторної вибірки

$$\mu = \sqrt{\frac{\sigma^2}{n-1}},$$

для неповторної

$$\mu = \sqrt{\frac{\sigma^2}{n-1} \left(1 - \frac{n}{N}\right)}. \quad (11.8).$$

Треба зазначити, що при малій вибірці діє особливий закон розподілу величини t , відкритий англійським вченим Стьюдентом. Суть його зводиться до того, що при визначенні ймовірності враховується не тільки коефіцієнт t , але й обсяг вибірки (табл. 11.2).

Таблиця 11.2

Розподіл імовірностей за Стьюдентом

| n/t | 5 | 7 | 10 | і т.д. | 20 |
|-----|-------|-------|-------|--------|-------|
| 1 | 0,626 | 0,644 | 0,656 | ... | 0,670 |
| 2 | 0,884 | 0,908 | 0,924 | ... | 0,940 |
| 3 | 0,960 | 0,976 | 0,984 | ... | 0,992 |

Порівняння розбіжностей ймовірностей P між нормальним розподілом і малою вибіркою засвідчує, що вони не є суттєвими.

При $t=1$ $P=0,683$ і $0,6740$; $t=2$ $P=0,954$ і $0,940$.

Гранична похибка вибірки Δx залежить від розміру величини t ; σ^2 ; n .

Таким чином, на основі рівності $\Delta \bar{x} = t \sqrt{\frac{\sigma^2}{n}}$ можуть бути розв'язані всі три завдання вибіркового дослідження, зокрема, якщо здаються:

$$1) n \text{ і } t, \text{ то } \Delta \bar{x} = t \sqrt{\frac{\sigma^2}{n}}; \quad (11.9).$$

$$2) n \text{ і } \Delta \bar{x}, \text{ то } t = \frac{\Delta \bar{x}}{\sqrt{\frac{\sigma^2}{n}}}; \quad (11.10).$$

$$3) \Delta \bar{x} \text{ і } t, \text{ то } n = \frac{t^2 \sigma^2}{\Delta^2 \bar{x}} \quad (11.11).$$

Гранична похибка для долі визначається таким чином:

$$\Delta_w = t \sqrt{\frac{w(1-w)}{n}}.$$

Відповідно

$$t = \frac{\Delta_w}{\sqrt{\frac{w(1-w)}{n}}} \text{ і } n = \frac{t^2 w(1-w)}{\Delta^2 w}. \quad (11.12).$$

Граничні похибки вибірки при власне випадковому і механічному відборі (вибірка неповторна) розраховують:

$$\text{для середньої} \quad \Delta \bar{x} = t \sqrt{\frac{\sigma^2}{n} \left(1 - \frac{n}{N}\right)}. \quad (11.13).$$

З наведеної формули знайдемо значення t і n :

$$t = \frac{\Delta \bar{x}}{\sqrt{\frac{\sigma^2}{n} \left(1 - \frac{n}{N}\right)}}; \quad n = \frac{t^2 \sigma^2 N}{\Delta^2 \bar{x} N + t^2 \sigma^2}. \quad (11.14).$$

Для долі

$$\Delta_w = t \sqrt{\frac{w(1-w)}{n} \left(1 - \frac{n}{N}\right)}. \quad (11.15).$$

Розрахуємо t і чисельність вибірки n :

$$t = \frac{\Delta w}{\sqrt{\frac{w(1-w)}{n} \left(1 - \frac{n}{N}\right)}}, \quad n = \frac{t^2 w(1-w)N}{\Delta^2 wN + t^2 w(1-w)}. \quad (11.16).$$

Приймаючи до уваги, що гранична похибка вибірки для ймовірності P є максимальним відхиленням розміру значень вибіркової оцінки від генеральної сукупності, можливі межі значень останньої визначають так:

для середньої $\bar{x} = \tilde{x} \pm \Delta x$, для долі $p = w \pm \Delta w$.

Граничні похибки при районованому відборі визначають за допомогою формул (11.17, 11.18).

а) повторний відбір:

$$\Delta x = t \sqrt{\frac{\sigma^2}{n}} \quad - \text{ для середньої}; \quad \Delta w = t \sqrt{\frac{w(1-w)}{n}} \quad - \text{ для доли}; \quad (11.17).$$

б) неповторюваний відбір;

$$\Delta x = t \sqrt{\frac{\sigma^2}{n} \left(1 - \frac{n}{N}\right)} \quad - \text{ для середньої}; \quad \Delta w = t \sqrt{\frac{w(1-w)}{n} \left(1 - \frac{n}{N}\right)} \quad - \text{ для доли}, \quad (11.18).$$

де σ^2 - середня дисперсія з внутрішньогрупових.

Розрахунки t і p для районованого відбору проводять аналогічно, як і при випадковій і механічній вибірці.

При серійному відборі формули розрахунку граничних похибок вибірки приведено в таблиці 11.3.

Таблиця 11.3

Граничні похибки вибірки при серійному відборі

| Схема відбору | Граничні похибки відбору | |
|-----------------------|---|---|
| | для середньої | для доли |
| Повторна вибірка | $\Delta x = t \sqrt{\frac{\delta^2}{S}}$ | $\Delta w = t \sqrt{\frac{w(1-w)}{S}}$ |
| Неповторювана вибірка | $\Delta x = t \sqrt{\frac{\delta^2}{S} \left(1 - \frac{s}{S}\right)}$ | $\Delta w = t \sqrt{\frac{w(1-w)}{n} \left(1 - \frac{s}{S}\right)}$ |

Де S - загальне число серій; δ^2 - міжгрупова дисперсія. Наведемо приклади розв'язання задач із використанням різних комбінацій вибіркового дослідження.

На одній із торгових баз проводилось вибірконе дослідження ваги цукру. Загальна кількість мішків становила 4000 штук. Вибірковому дослідженню було піддано 10 %, тобто 400 мішків. Результати вибіркового спостереження виявились такими: а) середня вага одного мішка цукру 50 кг (\bar{x}); б) середнє квадратичне відхилення $\sigma = 6$ кг; в) із загальної відібраної партії 100 мішків важили більше, ніж 50 кг ($w = \frac{100}{400} = 0,25$) Визначити:

1. Які можливі границі похибки середньої Δ_x і частки Δ_w , якщо їх потрібно гарантувати з ймовірністю 0,954.

Як нами зазначалось вище, при $p = 0,954$, $t = 2$. Вибірка проведена за схемою повторного відбору. Звідси:

$$\Delta_x = 2 \cdot \sqrt{\frac{6^2}{400}} = 2 \cdot \frac{6}{20} = \pm 0,6 \text{ кг};$$

$$\Delta_w = 2 \cdot \sqrt{\frac{0,25 \cdot 0,75}{400}} = 2 \cdot \frac{0,433}{20} = \pm 0,04.$$

З наведених розрахунків з ймовірністю 0,954 можна стверджувати:

а) можлива гранична похибка вибіркової середньої становить 600 грамів, а середня в генеральній сукупності, тобто в сукупності всієї партії мішків, знаходиться в межах

$$50 - 0,6 \leq \bar{x} \leq 50 + 0,6;$$

б) можлива гранична похибка частки мішків, вага яких є більшою, ніж 50 кг, дорівнює 0,04 і в генеральній сукупності частка таких мішків p знаходиться в межах

$$0,25 - 0,04 \leq P \leq 0,25 + 0,04. \quad 25\% - 4\% \leq p \leq 25\% + 4\%.$$

Одержані нами результати обчислень можна інтерпретувати так: у 954 випадках із 1000 вибірових досліджень можна очікувати, що середня вага мішка цукру коливається від 49 кг 400 грамів до 50 кг 600 грамів, а частка - в межах від 21 до 29 %.

2. Яка ймовірність того, що гранична похибка вибіркової середньої ваги 1 мішка цукру не буде більшою за 0,6 кг, а гранична похибка частки 100 мішків не перевищить 0,04 при обсязі вибірки $n = 400$. Виходячи із попередніх умов відомо, що $\sigma = 6$, а $w = 0,25$.

Використавши наведені вище формули, визначимо t :

$$t = \frac{0,6}{\sqrt{\frac{6^2}{400}}} = 0,6 : \frac{6}{20} = 2;$$

$$t = \frac{0,04}{\sqrt{\frac{0,25 \cdot 0,75}{400}}} = 0,04 : \frac{0,433}{20} = 2.$$

У нашій задачі для середньої та для долі $P(t) = 0,954$.

3. Якою повинна бути чисельність вибірки для того, щоб з ймовірністю 0,954 гарантувати, що границі можливої середньої похибки не перевищать 600 грамів, а можливої похибки частки - 0,04, якщо з попередніх даних відомо, що $\sigma = 6$, а $w = 0,25$.

Знову ж таки, використавши наведені вище формули для повторної механічної вибірки, визначимо n :

$$n = \frac{t^2 \sigma^2}{\Delta x^2} = \frac{2^2 \cdot 6^2}{0,6^2} = \frac{1,44}{0,36} = 400 ;$$

$$n = \frac{t^2 w(1-w)}{\Delta_w^2} = \frac{4 \cdot 0,25 \cdot 0,75}{0,04^2} = \frac{0,75}{0,0016} = 468 .$$

Одержані результати не збігаються, що є результатом заокруглень Δ_w . Разом з тим у даному випадку чисельність вибірки повинна становити не менше 400.

Дещо інакше шукають показники вибіркової сукупності при проведенні безповторного відбору. У формули $\Delta_x = t \sqrt{\frac{\sigma^2}{n}}$ і $\Delta_w = t \sqrt{\frac{w(1-w)}{n}}$ вносять обов'язково корегуючі коефіцієнти $\sqrt{1 - \frac{n}{N}}$. Такого типу формули наведені нами вище. Зазначимо, що корегуючий коефіцієнт завжди менший за одиницю. Зменшення співвідношення $\frac{n}{N}$ наближає цей коефіцієнт до одиниці. При 5 %-вій вибірці значення коефіцієнта рівне 0,95, при 2 %-вій вибірці - 0,98.

Прийmemo умовно, що в попередньому прикладі 10 % - відсоткова вибірка є неповторюваною. В результаті розрахунків одержимо такі відповіді:

$$\Delta x = 2 \cdot \sqrt{\frac{6^2}{400} \left(1 - \frac{400}{4000}\right)} = 2 \cdot \frac{6}{20} \cdot 0,95 = 0,63 ;$$

$$\Delta_w = 2 \cdot \sqrt{\frac{0,25 \cdot 0,75}{400} \left(1 - \frac{400}{4000}\right)} = 2 \cdot \frac{0,433}{20} \cdot 0,95 = 0,041 ;$$

$$t = \frac{0,6}{\sqrt{\frac{6^2}{400} \left(1 - \frac{400}{4000}\right)}} = 0,6 : \left(\frac{6}{20} \cdot 0,95\right) = 2,10 .$$

$$t = \frac{0,04}{\sqrt{\frac{0,25 \cdot 0,75}{400} \left(1 - \frac{400}{4000}\right)}} = 1,9 .$$

Коефіцієнт $t = 1,9$. Йому відповідає ймовірність 0,942, а $t=2,1$ $p=0,972$.

$$n = \frac{2^2 \cdot 6^2 \cdot 4000}{0,6^2 \cdot 4000 + 2^2 \cdot 6^2} = \frac{576000}{1584} = 364 ;$$

$$n = \frac{2^2 \cdot 0,25 \cdot 0,75 \cdot 4000}{0,04^2 \cdot 4000 + 2^2 \cdot 0,25 \cdot 0,75} = \frac{3000}{7,15} = 419 .$$

Аналогічно проводять розрахунки показників вибіркового спостереження при використанні районованого та серійного відбору.

11.4. Поширення результатів вибірки на генеральну сукупність

Поширення вибірових даних на генеральну сукупність здійснюють шляхом в порівняння характеристик генеральної сукупності з показниками вибірки. Найбільш широке використання в статистиці одержали два способи поширення вибірових даних:

- 1) прямого перерахунку;
- 2) поправочних коефіцієнтів.

При першому способі середні величини або частки, одержані в результаті дослідження вибіркової сукупності, множать на кількість одиниць генеральної сукупності.

Приклад (дані умовні). В результаті вибіркового обстеження депозитних рахунків 200 клієнтів (п) із загальної їх кількості 6000 (N) одного з комерційних банків встановлено з ймовірністю 0,954, що середній розмір вкладу на рахунку одного клієнта становив 5000 у.о., а можливі відхилення від середньої величини ± 500 у.о. Способом прямого перерахунку можна визначити, що загальна сума депозитних вкладів 6000 клієнтів банку становила $(5000 \pm 500) \cdot 6000 = 30000000 \pm 500$ умовних одиниць.

Часто результати вибіркового дослідження можуть використовуватися з метою уточнення даних суцільного спостереження. При цьому використовують спосіб поправочних коефіцієнтів. Його застосуванню передує порівняння даних вибіркового спостереження та генеральної сукупності, в результаті чого встановлюють відсоток недообліку. Коефіцієнти, що одержані за результатами порівняння, використовують для внесення змін в формуляри суцільного обліку.

Приклад. За результатами перепису населення було встановлено, що кількість дорослих осіб, які проживають в одному з

сільських населених пунктів становила 5000 чоловік. З метою уточнення даних перепису по адресах проживання було проведено вибіркоче обстеження. Згідно даних перепису адрес, які потрапили до вибірки, чисельність населення склала 150 осіб, а за даними вибіркового обстеження – 157. Таким чином, встановлено факт недообліку населення. Треба встановити відсоток недообліку (коефіцієнт). Виходячи з умови задачі він становить $4,7\% \left(\frac{7}{150} \cdot 100 \right)$.

Одержаний коефіцієнт поширюють на всю генеральну сукупність, тобто $5000 \cdot 1,047 = 5235$ осіб.

Тестові завдання

1. Узагальнюючим показником генеральної сукупності є:

- а) середня, дисперсія, доля;
- б) обсяг сукупності, дисперсія, частка;
- в) вибіркова середня, частка, обсяг сукупності;
- г) дисперсія, частка, доля.

2. Доля обчислюється як відношення:

- а) обсягу генеральної сукупності до вибіркової;
- б) числа M одиниць, які мають дану властивість, до всієї генеральної сукупності;
- в) вибіркової сукупності до генеральної;
- г) середньої до дисперсії.

3. Узагальнюючими характеристиками вибіркової сукупності є:

- а) середня, дисперсія, частка;
- б) середня, дисперсія, доля;
- в) вибіркова середня, дисперсія, доля;
- г) вибіркова сукупність, середня, дисперсія.

4. Середня похибка дорівнює:

- а) квадратичній з усіх можливих конкретних похибок стохастичного процесу;
- б) різниці між величиною вибіркової та генеральної середньої;
- в) відношенню кількості одиниць ознаки, що мають дану властивість до обсягу вибіркової сукупності;

г) середній квадратичній з усіх можливих конкретних похибок стохастичного процесу.

5. Похибка вибірки визначається як:

- а) різниця між характеристиками вибіркової та генеральної сукупності;
- б) різниця між характеристиками генеральної та вибіркової сукупності;
- в) відношення характеристик вибіркової до генеральної сукупності;
- г) відношення характеристик генеральної до вибіркової сукупності.

6. Похибки реєстрації виникають у результаті використання:

- а) лише даних звітності;
- б) неправильних або неточних даних;
- в) лише даних вибіркового спостереження;
- г) лише результатів інвентаризації.

7. Похибки реєстрації поділяють на:

- а) навмисні та ненавмисні;
- б) систематичні та випадкові;
- в) прості та випадкові;
- г) систематичні та районовані.

8. Розрізняють такі види відбору:

- а) систематичний, районований та серійний;
- б) механічний, випадковий, гніздовий;
- в) індивідуальний, груповий, комбінований;
- г) випадковий, систематичний, механічний.

9. За схемами відбору вибірку поділяють на:

- а) індивідуальну, групову;
- б) індивідуальну, комбіновану;
- в) повторну, безповторну;
- г) індивідуальну, групову, комбіновану.

10. За способами відбору вибірки поділяють на:

- а) повторні, безповторні;
- б) індивідуальні, групові, комбіновані;
- в) систематичні, випадкові;
- г) випадкові, механічні, районовані, серійні.

11. Власне випадковий відбір, або випадкова вибірка здійснюється:

- а) за допомогою жереба або таблиць випадкових чисел;
- б) через рівні інтервали;
- в) з урахуванням структури генеральної сукупності;
- г) з урахуванням структури вибіркової сукупності.

12. При систематичному відборі вибір елементів сукупності здійснюється:

- а) за допомогою жереба або таблиць випадкових чисел;
- б) через рівні інтервали;
- в) шляхом вибору серій ;
- г) з урахуванням структури вибіркової сукупності.

13. Районований(типовий) відбір – це спосіб формування вибірки з урахуванням:

- а) інтервалів вибірки;
- б) обсягу сукупності;
- в) структури генеральної сукупності;
- г) серій, всередині яких проводять суцільне спостереження.

14. Моментним називається спостереження, при якому дослідженню підлягають:

- а) усі елементи сукупності;
- б) частина елементів сукупності;
- в) більш як половина елементів генеральної сукупності;
- г) усі елементи сукупності на певні моменти часу.

15. Обстеження, при якому дослідженню підлягає не одна, а дві і більше ознак, які значно варіюють здійснюється за допомогою:

- а) багатофазної вибірки;
- б) багатоступеневої вибірки;
- в) моментної вибірки;
- г) механічної вибірки.

16. Середня похибка вибірки розраховується:

- а) добуванням кореня з дисперсії;
- б) добуванням квадратичного кореня з дисперсії;
- в) добуванням кубічного кореня з дисперсії;

г) добуванням кореня з середнього лінійного відхилення.

17. Поширення вибірових даних на генеральну сукупність полягає в порівнянні:

- а) середньої вибіркової та генеральної сукупності;
- б) дисперсії вибіркової та генеральної сукупності;
- в) характеристик генеральної сукупності з показниками вибіркового спостереження;
- г) долі вибіркової та генеральної сукупності.

Типові задачі

Задача 1. З партії виготовлених 1800 виробів МГ перевірено за допомогою систематичної вибірки 30% виробів, з яких 20% мали дефект.

Визначити:

- 1) долю виробів, які мали дефект;
- 2) границі, в яких знаходиться відсоток дефектних виробів у всій партії з ймовірністю 0,997.

Задача 2. В результаті аналізу 1000 прооб, відібраних у випадковому порядку, одержані такі дані щодо вологості озимої пшениці:

| Вологість зерна, % | 3-6 | 6-9 | 9-12 | 12-15 | 15-18 | 18-21 |
|--------------------|-----|-----|------|-------|-------|-------|
| Число проб | 50 | 60 | 55 | 101 | 111 | 100 |

Визначити:

- 1) середній відсоток вологості озимої пшениці та дисперсію;
- 2) граничну помилку вибірки з ймовірністю 0,997;
- 3) яка повинна бути чисельність вибірки, щоб гранична помилка зменшилась у 3 рази?

Задача 3. При підготовці вибіркового спостереження було проведено пробне обстеження 15 серій з метою оцінки варіації ознаки. Одержано такі результати:

| Номер серії | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|--|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| Середній розмір ознаки кожній серії, см ³ | 77 | 71 | 74 | 80 | 84 | 81 | 77 | 73 | 79 | 78 | 70 | 80 | 82 | 81 | 79 |

Скільки серій потрібно відібрати (вибірка неповторна) для суцільного обстеження, щоби помилка не перевищила 2 см^3 (з ймовірністю 0,683), якщо генеральна сукупність становить 6000 рівноважних серій.

Задача 4. За результатами 10%-го неповторного відбору отримано дані про час роботи ламп КГМ-12-50. На основі цих даних визначіть: а) з ймовірністю 0,997 межі питомої ваги ламп з часом роботи понад 1220 год. у генеральній сукупності; б) з ймовірністю 0,954 граничну помилку вибірки й межі, в яких можна очікувати середній час роботи ламп у генеральній сукупності. Зробіть висновки.

| Час роботи, год. | Кількість ламп, шт. |
|------------------|---------------------|
| До 1200 | 5 |
| 1200–1210 | 15 |
| 1210–1220 | 40 |
| 1220–1230 | 30 |
| 1230–1240 | 10 |
| Усього | 100 |

Задача 5. Для дослідження розміру вкладів населення проведено 12%-ну неповторну вибірку вкладників. На основі цих даних визначіть: а) з ймовірністю 0,997 граничну помилку вибірки й межі, в яких знаходиться середній розмір вкладу в генеральній сукупності; б) з ймовірністю 0,954 частку вкладників з розміром заощаджень понад 7000 грн. Зробіть висновки.

| Сума вкладу, грн. | Кількість вкладників, осіб |
|-------------------|----------------------------|
| До 3000 | 50 |
| 3000–5000 | 150 |
| 5000–7000 | 400 |
| 7000–9000 | 300 |
| Понад 9000 | 100 |
| Усього | 1000 |

Задача 6. При проведенні 2%-го безповторного відбору отримано результати про ресурс роботи електродвигунів. Виходячи з цих даних визначіть: а) імовірністю 0,997 межі питомої ваги електродвигунів з часом роботи понад 12200 год. у генеральній сукупності; б) з імовірністю 0,954 граничну помилку вибірки й межі, в яких можна очікувати середній час роботи електродвигуна. Зробіть висновки.

| Час роботи, год. | Кількість двигунів, шт. |
|------------------|-------------------------|
| До 12000 | 5 |
| 12000–12100 | 15 |
| 12100–12200 | 40 |
| 12200–12300 | 30 |
| 12300–12400 | 10 |
| Усього | 100 |

Задача 7. Для дослідження доходів зайнятого населення проведено 5%-ну безповторну вибірку і отримано результати, наведені у таблиці. На основі цих даних визначіть: а) з імовірністю 0,997 граничну помилку вибірки та межі, в яких знаходиться середній розмір доходу на одну особу в генеральній сукупності; б) з імовірністю 0,954 частку осіб з річним доходом до 10000 грн. Зробіть висновки.

| Рівень річного доходу, грн. | Кількість зайнятих, осіб |
|-----------------------------|--------------------------|
| До 9000 | 50 |
| 9000–10000 | 150 |
| 10000–11000 | 400 |
| 11000–12000 | 300 |
| Понад 12000 | 100 |
| Усього | 1000 |

Задача 8. При проведенні 8%-го безповторного вибіркового обстеження працівників, отримано дані, наведені у таблиці. На основі цих даних визначіть: а) з імовірністю 0,954 граничну помилку вибірки і межі, в яких можна очікувати середній виробіток

робітників підприємства; б) з імовірністю 0,997 частку робітників, які виробляють за зміну від 25 до 45 виробів. Зробіть висновки.

| Кількість виробів за зміну, шт. | Чисельність робітників, осіб |
|---------------------------------|------------------------------|
| До 25 | 5 |
| 25–30 | 10 |
| 30–35 | 20 |
| 35–40 | 45 |
| 40–45 | 15 |
| 45–50 | 3 |
| 50 і більше | 2 |
| Усього | 100 |

Задача 9. Площа, зайнята під посівами кукурудзи в господарстві складає 1500 га, середнє квадратичне відхилення врожайності 4 ц/га. Вибірка механічна, без повторна.

Визначити необхідний обсяг вибірки при розрахунку середньої врожайності кукурудзи з ймовірністю 0,954, при умові, що помилка середньої не перевищуватиме 0,5ц/га.

Задача 10. За даними вибіркового обстеження 17000 пасажирів приміських потягів встановлено, що середня тривалість поїздки пасажира складала 25 км, а середнє квадратичне відхилення – 5 км.

Визначити:

1) границі середньої тривалості поїздки пасажирів з ймовірністю 0,683;

2) як зміниться гранична помилка вибірки при ймовірності 0,954?

Задача 11. Станом на звітну дату у Тернопільській філії одного з комерційних банків нараховувалося 2500 відкритих депозитних рахунків. Скільки рахунків необхідно обстежити методом випадкового неповторного відбору, щоб з імовірністю 0,997 знайти межі середнього розміру депозиту за умови, що помилка репрезентативності не перевищить 10 грн., а середнє квадратичне відхилення 60 грн.? Зробіть висновки.

Задача 12. На 1 вересня звітного року чисельність професорсько-викладацького складу ВНЗ становила 2000 осіб. Для визначення

меж частки працівників, які отримують заробітну плату понад 5000 грн., необхідно провести неповторний відбір працівників ВНЗ. Дисперсія частки такого персоналу становить 0,1. Скільки працівників необхідно обстежити, щоб з імовірністю 0,954 помилка вибірки не перевищувала 2%? Зробіть висновки.

Питання для самоконтролю

1. Дайте визначення сукупності
2. Дати визначення репрезентативності.
3. Узагальнюючі показники вибіркової сукупності.
4. Похибки вибіркового спостереження.
5. Принципи забезпечення репрезентативності вибірки.
6. Схеми відбору.
7. Дайте тлумачення “багатофазна і багатоступенева вибірки”.
8. Середня та гранична похибка вибірки.
9. Основні характеристики малої похибки.
10. Способи поширення вибірових даних на генеральну сукупність.

РОЗДІЛ 12. ПОДАННЯ СТАТИСТИЧНИХ ДАНИХ: ТАБЛИЦІ, ГРАФІКИ, КАРТИ

12.1 Статистичні таблиці

Інколи результати статистичних досліджувань подають за допомогою статистичних таблиць, завдяки чому досягається наочність аналізу явищ і процесів суспільного життя.

Статистичними називаються таблиці, за допомогою яких наочно одержують зведену кількісну характеристику сукупності.

Значення статистичних таблиць полягає в тому, що за допомогою цифр найбільш раціонально, наочно і систематизовано подаються результати обробки статистичних матеріалів.

За змістом статистична таблиця відображає закінчену думку, що подана в цифрах. **Статистична таблиця** - це комбінація горизонтальних рядків і вертикальних граф. Їх перетин утворює клітини таблиці. Ліві, бічні і верхні клітини призначені для заголовків, а решта- для цифрових даних.

Сукупність горизонтальних рядків і вертикальних граф без цифрових даних називається макетом статистичної таблиці (рис.12.1)

| | | | | | |
|------------------------------------|----------------|---|---|---|---|
| Присудок | Заголовки граф | | | | |
| Підмет | | | | | |
| | 1 | 2 | 3 | 4 | 5 |
| Перелік (групи) одиниць сукупності | | | | | |

Рис.12.1 Макет статистичної таблиці

Складовою статистичної таблиці є підмет і присудок .
Статистичний підмет - це група або окремі одиниці сукупності, котрі досліджуються статистикою.

Присудком слугують показники, за допомогою яких характеризують об'єкт або групи досліджуваної сукупності. Підмет розміщується в лівій частині таблиці, в назвах горизонтальних рядків. Присудок - в правій частині вертикальних граф.

Залежно від побудови підмета розрізняють три види статистичних таблиць: **прості, групові та комбіновані**. **Прості**, будуються без групувань підмета. **Груповими** називають таблиці, в підметі яких міститься групування одиниць спостереження за однією ознакою.

Комбіновані таблиці - це такі, підмет яких містить групування одиниць спостереження за декількома ознаками одночасно.

Правила побудови статистичних таблиць:

- таблиця повинна бути невеликою за розміром, легко читатись і запам'ятовуватись;
- таблиця обов'язково має мати заголовок;
- якщо число показників у таблиці велике, їх нумерують;
- буквами позначаються графи підмета, цифрами графи присудка;
- якщо відомості про явище відсутні, то у відповідній клітинці проставляють (...), відсутність явища (-);
- кількість показників у межах однієї графи повинна наводитись з установленою точністю, тобто 0,1; 0,01; 0,001 і т.д;
- таблиці повинні бути замкнутими за підсумками;
- таблиця повинна мати номер та одиниці виміру;
 - необхідно розрізняти «разом» і «всього». «Разом» є підсумком для певної окремої сукупності, а «всього» - це «разом» для сукупності в цілому.

12.2 Статистичні графіки

Часто в практиці статистичних досліджень, для забезпечення наочності ряди розподілу зображають графічно. Найчастіше при цьому використовують діаграму казусів, полігон (дискретні ряди) і гістограму (інтервальні ряди) розподілу.

Діаграма казусів може бути побудована на основі даних емпіричного ряду розподілу. За характером розміщення крапок на графіку можна визначити величину варіації досліджуваної ознаки.

Наприклад, залишки депозитних вкладів 25 клієнтів банку на початок січня-місяця склали (євро).

5000, 5000, 6000, 8000, 8500, 4000, 3500, 10000, 15000, 7500, 6500, 10000 10000, 1000, 2000, 9000, 9500, 3500, 10000, 12000, 8000, 9000, 10000, 8000, 6000.

Побудуємо діаграму казусів, для чого в прямокутній системі координат на осі абсцис позначимо крапками значення, котрі відповідають порядковому номеру клієнта на початок місяця, а на осі ординат - розмір покладених на депозит грошових одиниць за встановленим масштабом (рис.12.2).

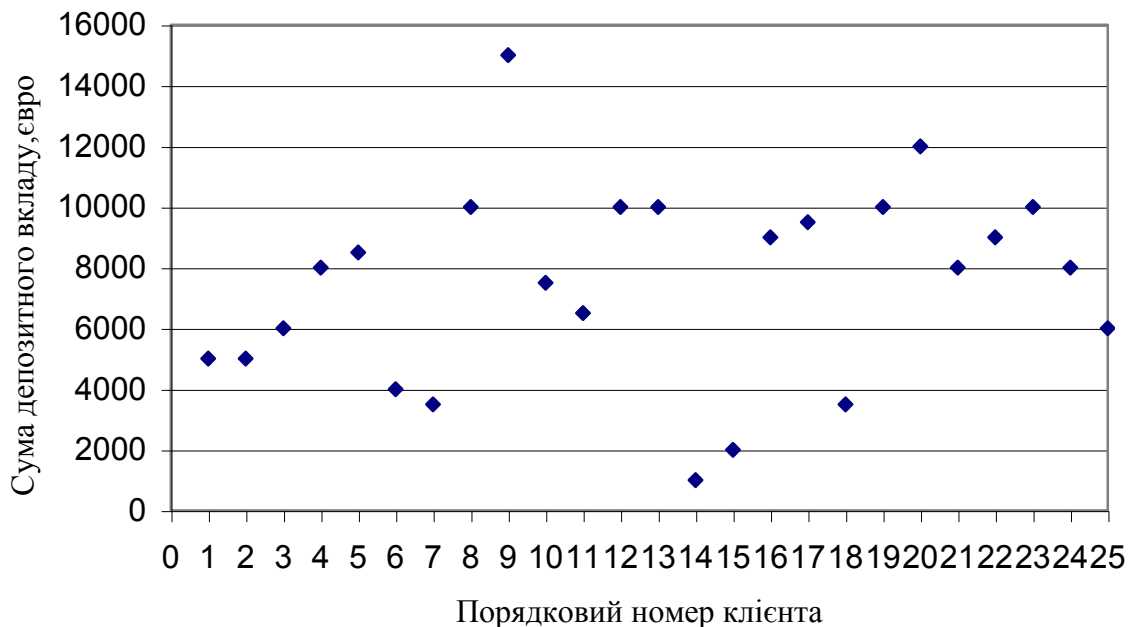


Рис. 12.2.1 Діаграма казусів депозитних вкладів клієнтів банку

Діаграма показує варіацію розмірів вкладів клієнтів банку в перший день січня-місяця аналізованого періоду.

Продовжуючи аналізувати ряд розподілу клієнтів за розміром вкладів, побудуємо на основі емпіричних даних ранговий ряд. Розмістимо розмір вкладів клієнтів у порядку зростання вкладених сум. В результаті одержимо огіву розміру вкладів клієнтів банку в перший день місяця (рис.12.3).

1000, 2000, 3500, 3500,4000, 5000, 5000, 6000, 6000,6500, 7500, 8000, 8000, 8000, 8500, 9000,9000,9500,10000,10000,10000,10000,10000, 12000,15000.

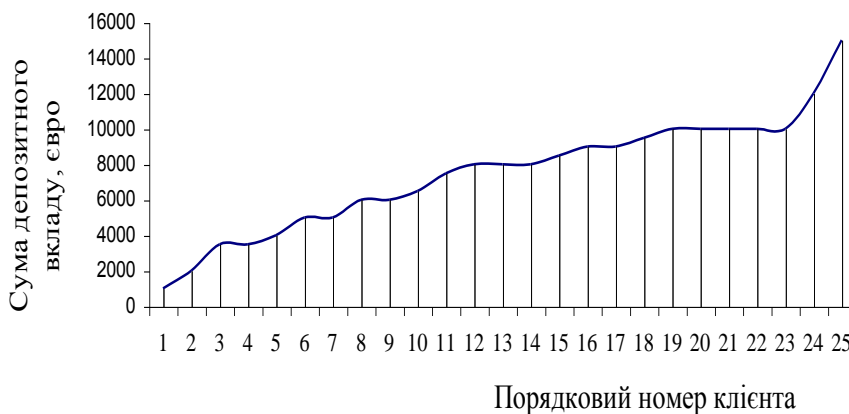


Рис. 12.3 Огіва розміру депозитних вкладів клієнтів банку

Побудований ранговий ряд характеризує розмір депозитних вкладів клієнтів банку на початок місяця. Мінімальний вклад склав 1000 максимальний –15000 євро. Побудована огіва характеризує також величину вкладу кожного клієнта по відношенню до мінімального та максимального розміру всіх інших вкладників.

Характеристика ряду розподілу може бути доповнена за допомогою полігону або гістограми.

Полігоном називають графік, на якому ряд розподілу будується за допомогою лінійної діаграми. При цьому на осі абсцис відкладають значення варіант у порядку зростання або спадання, а на осі ординат - частоти. Місце перетину цих значень позначають крапками і сполучають за допомогою ліній.

Наприклад, розмір місячної заробітної плати робітників арматурного заводу можна зобразити шляхом побудови полігону розподілу (рис.12.4).

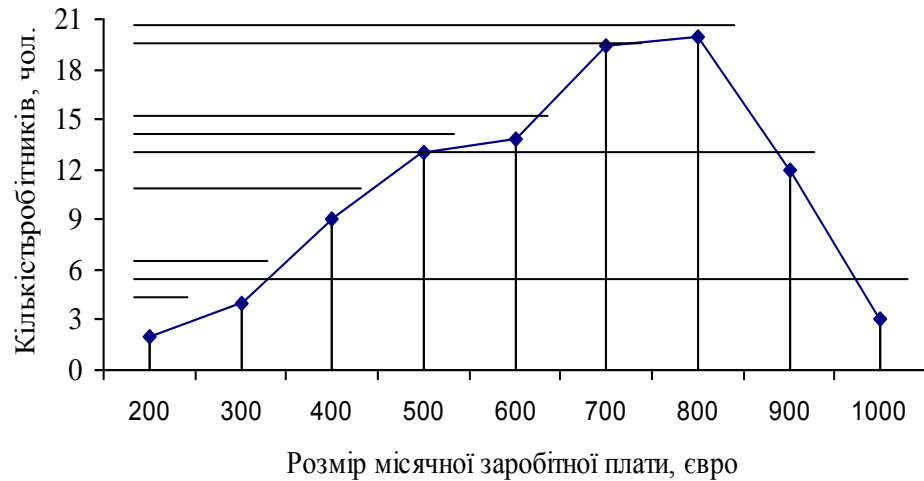


Рис. 12.4 Полігон розподілу робітників цеху за розміром зарплатної плати

Гістограмою називається графік, на якому ряд розподілу зображається за допомогою суміжних один із одним стовпчиків.

Наприклад, урожайність цукрових буряків у фермерських господарства регіону за минулий рік характеризується такими даними (таблиця 12.1)

Таблиця 12.1

Розподіл фермерських господарств за урожайністю цукрових буряків у минулому році

| Групи фермерських господарств за урожайністю цукрових буряків, ц/га | Число господарств | Накопичені частоти |
|---|-------------------|--------------------|
| 100-200 | 5 | 5 |
| 200 - 300 | 15 | 20 |
| 300 - 400 | 20 | 40 |
| 400 - 500 | 15 | 55 |
| 500 - 600 | 5 | 60 |
| 600 - 900 | 3 | 63 |
| Разом | 63 | |

Використавши дані наведеної таблиці побудуємо гістограму розподілу фермерських господарств за урожайністю цукрових буряків. Висоту стовпчиків визначимо пропорційно частотам. Ширина

стовпчиків при рівних інтервалах буде рівновеликою, а при нерівних - неоднаковою (рис.12.5).

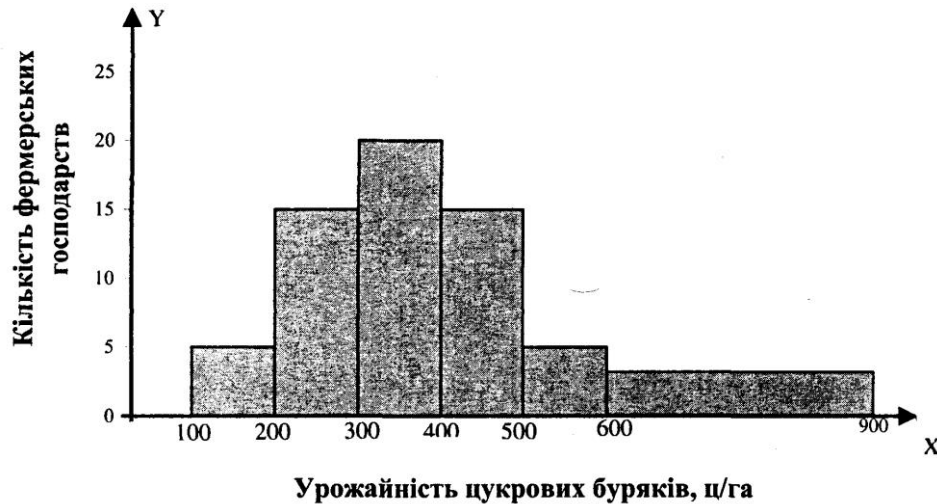


Рис.12.5 Гістограма розподілу фермерських господарств за урожайністю цукрових буряків

Для інтервальних рядів із нерівними інтервалами з метою нівелювання впливу величини інтервалу на розподіл сукупності - висоту стовпчиків встановлюють пропорційно не частотам, а щільності розподілу ознаки. **Щільність розподілу** - це кількість елементів сукупності, що припадає на одиницю ширини інтервалу групувальної ознаки. Наприклад, щільність розподілу урожайності цукрових буряків (табл.12.1) у групі від 300 до 400 ц/га становить:

$$\frac{20}{400 - 300} = 0,5.$$

Наочно розподіл значень досліджуваної ознаки ряду розподілу може бути зображено за допомогою **комуляти**. При цьому на осі абсцис відкладають значення варіант, а на осі ординат - накопичені частоти, котрі розраховують шляхом сумування частот даної групи (див. табл. 12.1).

Використавши дані, наведені в табл. 12.1, побудуємо комулятивну огіву (комулятивну криву) розподілу фермерських господарств за урожайністю цукрових буряків (рис.12.6).

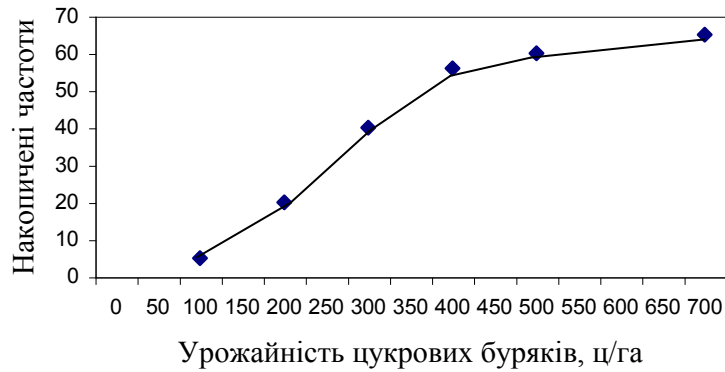


Рис. 12.6 Комулята розподілу фермерських господарств регіону за урожайністю цукрових буряків

Ступінь нерівномірності розподілу певного сумарного показника між одиницями окремих груп варіаційного ряду можна показати за допомогою кривої Лоренца, або кривої концентрації. Наприклад, щоб графічно зобразити нерівномірність розподілу міст регіону по числу жителів будуюмо квадрат 100x100. На осі абсцис відкладаємо значення нагромаджених підсумків відсотків міст, а на осі ординат комулятивні підсумки чисельності в них жителів. На крапках перетину будуюмо криву, котру називають кривою Лоренца. Наведемо приклад побудови кривої Лоренца. Для цього використаємо інформацію, наведену в таблиці 12.2.

Таблиця 12.2

Розподіл міст регіону за чисельністю жителів

| Групи міст за чисельністю жителів, чол. | Число міст у % до підсумку | Чисельність населення у % до підсумку | Комулятивні підсумки | |
|---|----------------------------|---------------------------------------|----------------------|-------------|
| | | | % міст | % населення |
| До 3000 | 49,4 | 4,6 | 49,4 | 4,6 |
| 3001 - 5000 | 26,5 | 5,2 | 75,9 | 9,8 |
| 5001 - 10000 | 10,1 | 14,6 | 86,0 | 24,4 |
| 10001 - 20000 | 6,3 | 18,8 | 92,3 | 43,2 |
| 20001 - 50000 | 4,5 | 30,1 | 96,8 | 73,3 |
| 50001 і більше | 3,2 | 26,7 | 100,0 | 100,0 |
| Разом | 100,0 | 100,0 | | |

На основі даних табл. 12.2 побудуємо криву Лоренца (рис. 12.7):

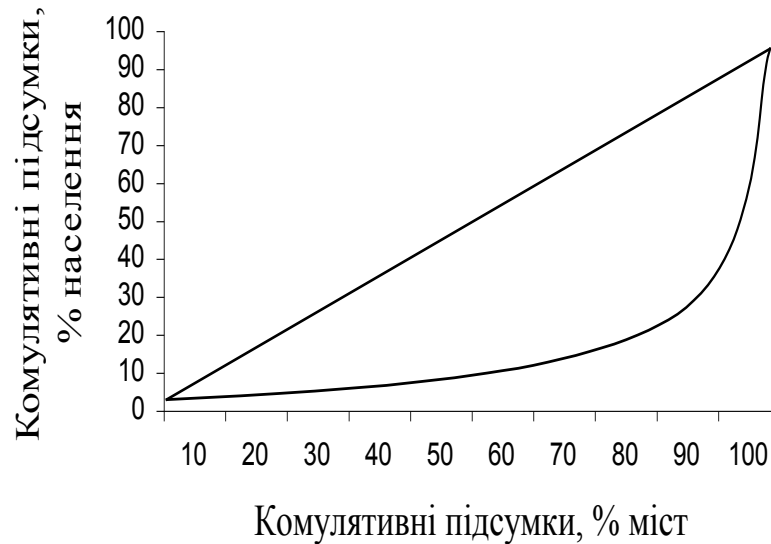


Рис. 12.7. Крива Лоренца

Побудований графік засвідчує, що розподіл чисельності жителів міст регіону обернений пропорційний до кількості міст. Зокрема, в містах із чисельністю жителів понад 50001 питома вага яких становить лише 3,2 %, проживає зверх четвертої частини (26,7 %) всього населення. Якби чисельність населення розподілялась рівномірно за кількістю міст, то концентрація була би відсутня, графік проходив би по діагоналі квадрата.

Порівняно висока вгнутість кривої на графіку засвідчує про досить високий рівень концентрації міських жителів у відносно невеликій кількості міст. Величина кривизни вказує також на нерівномірність розподілу ознак, які досліджуються.

Тестові завдання

1. Статистичними називаються таблиці за допомогою яких одержують:

- а) зведену кількісну характеристику сукупності;
- б) зведену якісну характеристику сукупності;
- в) зведену кількісну та якісну характеристику сукупності;
- г) цифрові дані щодо обсягу сукупності.

2. Статистичний підмет – це:

- а) група одиниць сукупності, що досліджуються;
- б) група або окремі одиниці сукупності, що досліджуються;
- в) показники, за допомогою яких характеризується об'єкт;
- г) показники, за допомогою яких характеризується об'єкт або група досліджуваної сукупності.

3. Статистичний присудок: - це:

- а) група одиниць сукупності, що досліджуються;
- б) група або окремі одиниці сукупності, що досліджуються;
- в) показники, за допомогою яких характеризується об'єкт або група досліджуваної сукупності;
- г) комбінація горизонтальних рядків і вертикальних граф.

4. Залежно від побудови підмета розрізняють такі види статистичних таблиць:

- а) цифрові;
- б) текстові;
- в) комплексні;
- г) прості, групові та комбінаційні.

5. Груповими називають таблиці:

- а) в підметі яких міститься групування одиниць спостереження за однією ознакою;
- б) підмет яких містить групування одиниць спостереження за декількома ознаками;
- в) підмет яких містить комбінацію горизонтальних рядків і вертикальних граф;
- г) в присудку яких міститься групування одиниць спостереження за двома ознаками.

6. Комбіновані таблиці – це такі:

- а) підмет яких містить групування одиниць спостереження за однією ознакою;
- б)) підмет яких містить групування одиниць спостереження за декількома ознаками;
- в) підмет яких містить комбінацію горизонтальних рядків вертикальних граф;
- г) які поєднують кількісні та якісні ознаки.

7. Полігоном називають графік, на якому ряд розподілу будується за допомогою:

- а) діаграми;
- б) гістограми;
- в) картограми;
- г) лінійної діаграми.

8. Гістограмою називається графік, на якому ряд розподілу зображається за допомогою:

- а) лінійної діаграми;
- б) стовпчикової діаграми;
- в) суміжних один з одним стовпчиків;
- г) картограми.

9. Щільність розподілу – це кількість елементів сукупності, що:

- а) припадає на одиницю ширини інтервалу групувальної ознаки;
- б) припадає на одиницю висоти інтервалу групувальної ознаки;
- в) міститься на одиницю масштабу графіка;
- г) задаються в умові задачі.

Питання для самоконтролю

1. Дайте визначення статистичного зведення .
2. Яка різниця між простим і складним зведенням?
3. Як Ви розумієте централізоване і децентралізоване зведення?
- 4 Дайте визначення статистичного групування.
5. Дайте визначення типологічного групування.
6. Яка різниця між структурним і аналітичним групуванням?
- 7 Які є види групування?

8. Порядок визначення групувальної ознаки
9. Порядок визначення інтервалу групування.
10. Порядок визначення числа груп
11. Яка різниця між закритими і відкритими інтервалами?
12. Яка різниця між простим і складним групуванням
13. Поясніть суть вторинного групування.
14. Дайте визначення ряду розподілу.
15. Яка різниця між дискретним та інтервальним рядом розподілу?
16. Дайте визначення полігону, гістограми, огіви.
17. Дайте визначення статистичної таблиці.
18. Основні правила побудови статистичних таблиць.

Змістовий модуль 3

9. Аналіз тенденцій розвитку та коливань

- 9.1 Загальна характеристика згладжування динамічного ряду.
- 9.2 Вирівняний динамічний ряд.
- 9.3 Аналітичне вирівнювання динамічних рядів.
- 9.4 Вирівнювання за рівнями прямої та кривої лінії.
- 9.5 Статистичне дослідження стійкості динамічних рядів.
- 9.6 Статистичне дослідження сезонності динамічних рядів.
- 9.7 Автокореляція.

10. Індексний метод

- 10.1 Використання індивідуальних індексів в статистиці.
- 10.2 Найважливіші агрегатні форми індексів.
- 10.3 Індексний метод аналізу виробництва продукції та індекси продуктивності праці.
- 10.4 Середні індекси.
- 10.5 Взаємозв'язки індексів із урахуванням двох чинників.
- 10.6 Взаємозв'язки індексів із урахуванням трьох і більше чинників.
- 10.7 Середні індекси постійного, змінного складу та структурних зрушень.

11. Вибірковий метод

- 11.1 Загальна характеристика, завдання та значення вибіркового методу дослідження.

- 11.2 Похибки вибіркового спостереження.
- 11.3 Способи відбору.
- 11.4 Характеристика схем відбору одиниць вибіркової сукупності.
- 11.5 Багатоступенева та багатофазна вибірка.
- 11.6 Три основні завдання забезпечення об'єктивності результатів вибіркового спостереження.
- 11.7 Найважливіші характеристики генеральної та вибіркової сукупності.
- 11.8 Повторна та неповторна вибірка.
- 11.9 Критерії оцінки достовірності результатів вибірки.

12. Подання статистичних даних: таблиці, графіки, карти.

- 12.1 Правила побудови статистичних таблиць.
- 12.2 Діаграма казусів, полігон, гістограма.
- 12.3 Крива Лоренца.

Завдання для самостійного вивчення

1. Вирівнювання динамічного ряду за допомогою укрупнення рівнів, середньої плинної та середнього абсолютного приросту.
2. Критерії використання функцій вимірювання динамічного ряду.
3. Вимірювання за рівнянням параболі другого порядку.
4. Лінійний тренд.
5. Взаємозв'язки індивідуальних індексів.
6. Методологічні засади встановлення індексованих величин та співвимірників.
7. Загальні індекси товарообігу, фізичного обсягу та ціни.
8. Загальні індекси собівартості, затрат та фізичного обсягу продукції.
9. Вартісні, натуральні та трудові індекси продуктивності праці.
10. Умови забезпечення репрезентативності вибірки.
11. Моментні спостереження вибірки.
12. Передумова поширення даних вибіркового спостереження на генеральну сукупність.
13. Оцінка ймовірності вибірки Стюдента.
14. Граничні похибки при серійному відборі.
15. Групові та комбіновані статистичні таблиці.

ДОДАТКИ

Критичні значення статистичних характеристик критеріїв
математичної статистики

Додаток 1

Квантили нормального розподілу

| $1 - \alpha$ | 0.800 | 0.900 | 0.950 | 0.975 |
|--------------|--------------|--------------|--------------|--------------|
| t | 0.84 | 1.28 | 1.64 | 1.96 |
| t | 1.28 | 1.64 | 1.96 | 2.24 |

Додаток 2

Квантили розподілу Стьюдента t

| k | $1 - \alpha$ | | | k | $1 - \alpha$ | | |
|----------|--------------|-------------|--------------|----------|--------------|-------------|--------------|
| | 0.90 | 0.95 | 0.975 | | 0.90 | 0.95 | 0.975 |
| 3 | 1.64 | 2.35 | 3.18 | 12 | 1.36 | 1.78 | 2.18 |
| 4 | 1.53 | 2.13 | 2.78 | 14 | 1.35 | 1.76 | 2.14 |
| 5 | 1.48 | 2.02 | 2.57 | 16 | 1.34 | 1.75 | 2.12 |
| 6 | 1.44 | 1.94 | 2.45 | 18 | 1.33 | 1.73 | 2.10 |
| 7 | 1.41 | 1.89 | 2.36 | 20 | 1.33 | 1.72 | 2.09 |
| 8 | 1.40 | 1.86 | 2.31 | 22 | 1.32 | 1.72 | 2.07 |
| 9 | 1.38 | 1.83 | 2.26 | 24 | 1.32 | 1.71 | 2.06 |
| 10 | 1.37 | 1.81 | 2.23 | 28 | 1.31 | 1.70 | 2.05 |
| 11 | 1.36 | 1.80 | 2.20 | ∞ | 1.28 | 1.64 | 1.96 |

Додаток 3

Квантили χ^2 -розподілу при $\alpha=0,05$

| k | 0.95 | k | 0.95 |
|----------|-------------|----------|-------------|
| 1 | 3.84 | 11 | 19.68 |
| 2 | 5.99 | 12 | 21.03 |
| 3 | 7.82 | 13 | 22.36 |
| 4 | 9.49 | 14 | 23.69 |
| 5 | 11.07 | 15 | 25.00 |
| 6 | 12.59 | 16 | 26.30 |
| 7 | 14.07 | 17 | 27.59 |
| 8 | 15.51 | 18 | 28.87 |
| 9 | 16.92 | 19 | 30.14 |
| 10 | 18.31 | 20 | 31.41 |

Критичні значення кореляційного відношення η^2
та коефіцієнта детермінації R^2 при $\alpha=0,05$

| k_2 | k_1 | | | | | | |
|-------|-------|-------|-------|-------|-------|-------|-------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| 5 | 0.569 | 0.699 | 0.764 | 0.806 | 0.835 | 0.854 | 0.872 |
| 6 | 0.500 | 0.632 | 0.704 | 0.751 | 0.785 | 0.811 | 0.831 |
| 7 | 0.444 | 0.575 | 0.651 | 0.702 | 0.739 | 0.768 | 0.791 |
| 8 | 0.399 | 0.527 | 0.604 | 0.657 | 0.697 | 0.729 | 0.754 |
| 9 | 0.362 | 0.488 | 0.563 | 0.618 | 0.659 | 0.692 | 0.719 |
| 10 | 0.332 | 0.451 | 0.527 | 0.582 | 0.624 | 0.659 | 0.687 |
| 12 | 0.283 | 0.394 | 0.466 | 0.521 | 0.564 | 0.600 | 0.630 |
| 14 | 0.247 | 0.348 | 0.417 | 0.471 | 0.514 | 0.550 | 0.580 |
| 16 | 0.219 | 0.312 | 0.378 | 0.429 | 0.477 | 0.507 | 0.538 |
| 18 | 0.197 | 0.283 | 0.345 | 0.394 | 0.435 | 0.470 | 0.501 |
| 20 | 0.179 | 0.259 | 0.318 | 0.364 | 0.404 | 0.432 | 0.468 |
| 22 | 0.164 | 0.238 | 0.294 | 0.339 | 0.377 | 0.410 | 0.439 |
| 24 | 0.151 | 0.221 | 0.273 | 0.316 | 0.353 | 0.385 | 0.414 |
| 26 | 0.140 | 0.206 | 0.256 | 0.297 | 0.332 | 0.363 | 0.391 |
| 28 | 0.130 | 0.193 | 0.240 | 0.279 | 0.314 | 0.344 | 0.371 |
| 30 | 0.122 | 0.182 | 0.227 | 0.264 | 0.297 | 0.326 | 0.353 |
| 40 | 0.093 | 0.139 | 0.176 | 0.207 | 0.234 | 0.259 | 0.282 |
| 50 | 0.075 | 0.113 | 0.143 | 0.170 | 0.194 | 0.216 | 0.235 |
| 60 | 0.063 | 0.095 | 0.121 | 0.144 | 0.165 | 0.184 | 0.202 |
| 80 | 0.047 | 0.072 | 0.093 | 0.110 | 0.127 | 0.142 | 0.156 |
| 100 | 0.038 | 0.058 | 0.075 | 0.090 | 0.103 | 0.116 | 0.128 |
| 120 | 0.032 | 0.049 | 0.063 | 0.080 | 0.087 | 0.098 | 0.109 |
| 200 | 0.019 | 0.030 | 0.038 | 0.046 | 0.053 | 0.060 | 0.067 |
| 400 | 0.010 | 0.015 | 0.019 | 0.023 | 0.027 | 0.031 | 0.034 |

Додаток 5

Критичні значення лінійного коефіцієнта кореляції при $\alpha=0,05$

| Обсяг вибірки | 5 | 6 | 7 | 8 | 9 | 10 | 12 | 14 | 16 |
|---------------|------|------|------|------|------|------|------|------|------|
| $r_{0,95}$ | 0.88 | 0.81 | 0.75 | 0.71 | 0.67 | 0.63 | 0.58 | 0.53 | 0.50 |

Додаток 6

Критичні значення коефіцієнта рангової кореляції Спірмена при
 $\alpha=0,05$

| Обсяг вибірки n | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|-------------------|------|------|------|------|------|------|------|------|
| $\rho_{0,95}$ | 0.90 | 0.83 | 0.71 | 0.64 | 0.60 | 0.56 | 0.53 | 0.50 |

Список використаної та рекомендованої літератури

1. Айрапетов А.М. Таблицы исчисления среднегодовых темпов роста, прироста и снижения. – М.: Статистика, 1967.
2. Банківська енциклопедія / За ред. А. М. Морозова. – К.: Фірма „Елітон”, 1993.
3. Бек В.Л. Теорія статистики: Навч. посіб. – К.: ЦУЛ, 2003.
4. Герасименко С.С., Головач А.В., Єріна А.М. Статистика: Підручник. – К.: КНЕУ, 2000.
5. Головач А.В., Єріна А.М., Козирев О.В. Статистика. – К.: Вища шк., 1993.
6. Гришин А.Ф. Статистика. – М.: Финансы и статистика, 2003.
7. Економічна енциклопедія: У 3 т. / Редкол.: С.В. Мочерний та ін. – К.: Академія, 2000.
8. Елисеева И. И., Юзбашев М.М. Общая теория статистики: Учебник. – М.: Финансы и статистика, 2004.
9. Ендрова М. Р., Ендрова М. В. Общая теория статистики. – М.: Юристъ, 2001.
10. Ефимова М.Р., Петрова Е.В., Румянцев В.Н. Общая теория статистики: Учебник. – М.: ИНФРА-М, 1997.
11. Єріна А. М., Пальян З. О. Теорія статистики: Практикум. – К.: Знання, 2004.
12. Єріна А. М. Статистичне моделювання та прогнозування: Навч. посіб. – К.: КНЕУ, 2001.
13. Закон України „Про державну статистику” // Голос України. – 2000. – 13 лип.; Відомості Верховної Ради України. – № 43. – К., 2000.
14. Закон України „Про інформацію” // Голос України. – 1992. – 13 листоп.
15. Захожай В.Б., Федорченко В.С. Теорія статистики. Навч. посіб. – К.: ЄУФІМП, 2000.
16. Кильдишев Г. С. Выборочное наблюдение. – М.: МЭСИ, 1965.
17. Крылов В. Н. Выборочный метод в статистике. – М.: Госстатиздат, 1957.
18. Кулинич О. І. Теорія статистики. – К.: Вища шк., 1992.

19. Лившиц Ф Д. Статистические таблицы. – М.: Госстатиздат, 1958.
20. Макроэкономическая статистика: Учеб. пособие / В. Н. Салин, В. Г. Медведев, С. И. Кудряшова, Е. П. Шпаковская. – М.: Дело, 2000.
21. Мармоза А. Т. Практикум з теорії статистики: Навч. посіб. – К.: Ельга Ніка-Центр, 2003.
22. Нечминов В. С. Избранные произведения. – М.: Наука, 1967. – Т.2.
23. Общая теория статистики/ Ф.Г. Долгушевский, В.С. Козлов, П. И. Полушин, Я. М. Эрлих. – М.: Статистика, 1967.
24. Освітньо-професійна програма підготовки бакалавра, спеціаліста і магістра напрямку 0501 „Економіка і підприємництво”: Галузевий стандарт вищої освіти. – К.: Міносвіти України, 2006.
25. Парфенцева Н. Міжнародні статистичні класифікації: Навч. посіб. – К.: Основи, 2000.
26. Пасхавер И. С. Средние величины в статистике. – М.: Статистика, 1979.
27. Попов І. І., Федорченко В. С. Теорія статистики. Практикум: Навч. посіб. – К.: КНТЕУ, 2001.
28. Практикум по теории статистики: Учеб. пособие / Под. Ред. Р.А. Шмойловой. – М.: Финансы и статистика, 2002.
29. Про заходи щодо розвитку державної статистики: Указ Президента України від 22.11.97 № 1299/97 // Статистика України. – 1998. – №1.
30. Програма реформування державної статистики на період до 2002 року: Постанова Кабінету Міністрів України від 27.06.98 № 971.
31. Рязов Н. Н. Общая теория статистики. – М.: Финансы и статистика, 1984.
32. Рязов Н. Н., Шор Ю. Л. Статистика в кредитных учреждениях. – М.: Финансы, 1968.
33. Социальная статистика: Учебник / Под ред. И. И. Елисеевой. – М.: Финансы и статистика, 1997.
34. Статистика: Учеб. пособие / Под ред. М. Р. Ефимовой. – М.: ИНФРА-М, 2003.

35. Статистичний щорічник України за 2000 рік. – К.: Техніка, 2001.
36. Статистичний щорічник України за 2006 рік. – К.: Консультант, 2007.
37. Статистичний щорічник України за 2007 рік. – К.: Консультант, 2008.
38. Теорія статистики: Навч. посіб. / П. Г. Вашків, П. І. Пастер, В. П. Сторожук, Є. І. Ткач. – К.: Либідь, 2001.
39. Толбатов Ю. А. Загальна теорія статистики засобами Excel: Посібник. – К.: Четверта хвиля, 1999.
40. Трінько Р. Теорія статистики: Навч. посіб. – Львів: Українські технології, 2003.
41. Уманець Т. В., Пігарев Ю. Б. Статистика: Навч. посіб. – К.: Вікар, 2003.
42. Урланис Б. Ц. Общая теория статистики: Курс лекций. – М.: Госкомстат, 1973.
43. Фінансова статистика: Навч. посібник / За кер. та наук. ред. В. Б. Захожая. – К.: МАУП, 2002.
44. Фінансова статистика (з основами теорії статистики): Підручник /А. В. Головач, В. Б. Захожай, Н. А. Головач, А. А. Шустіков. – К.: МАУП, 2005.
45. Яблочник А. Л. Общая теория статистики. – М.: Статистика, 1976.
46. Щурик М. В., Луцький І. М. Статистика: навч. посібник, Івано-Франківськ – 2004.