

**НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ЦИВІЛЬНОГО ЗАХИСТУ УКРАЇНИ
КАФЕДРА ПУБЛІЧНОГО АДМІНІСТРУВАННЯ У СФЕРІ ЦИВІЛЬНОГО ЗАХИСТУ**

Батир Ю.Г.

**СТАТИСТИЧНІ МЕТОДИ В ПУБЛІЧНОМУ
УПРАВЛІННІ ТА АДМІНІСТРУВАННІ**

КОНСПЕКТ ЛЕКЦІЙ

Харів 2019

Друкується за рішенням
Вченої ради факультету цивільного
захисту НУЦЗУ України
Протокол № 1 від 27.09.2019 р.

Рецензенти: д.держ.упр., доцент, професор кафедри публічного адміністрування у сфері цивільного захисту національного університету цивільного захисту України, В. Шведун;
д.держ.упр., професор, професор кафедри публічного адміністрування у сфері цивільного захисту національного університету цивільного захисту України, О.І. Крюков.

Статистичні методи в державному управлінні: конспект лекцій / Укладачі: Батир Ю.Г.. – Х.:НУЦЗУ, 2019. – 50 с.

Курс лекцій є складовою частиною методичного забезпечення при вивченні навчальної дисципліни «Статистичні методи в державному управлінні» магістерської підготовки за спеціальністю 074 «Публічне управління та адміністрування».

ВСТУП

Навчальна дисципліна «Статистичні методи в державному управлінні» є нормативною дисципліною, що включається в навчальні плани як дисципліна обов'язкового вибору.

Програма вивчення нормативної навчальної дисципліни «Статистичні методи в державному управлінні» складена відповідно до освітньо-професійних програм підготовки фахівців за спеціальністю 074 «Публічне управління і адміністрування».

Міждисциплінарні зв'язки: державне управління в економічній сфері, державне регулювання економіки, фінанси.

Предметом вивчення навчальної дисципліни є застосування статистичних методів у державному управлінні.

Мета викладання навчальної дисципліни – формування у слухачів теоретичних знань і практичних навичок статистичної оцінки соціально-економічних процесів.

Основними завданнями вивчення дисципліни є:

- ознайомити слухачів із сучасними методами економіко-статистичного аналізу державної політики;
- навчити слухачів застосовувати статистичний інструментарій при аналізі соціально-економічних та політичних процесів;
- заохочувати слухачів до використання різноманітних інформаційних джерел, необхідних для повного аналізу державної політики та науково обґрунтованих висновків і пропозицій стосовно неї.

Згідно з вимогами освітньо-професійної програми слухачі дисципліни повинні *знати та уміти*:

- теоретичні засади збирання, перевірки й опрацювання статистичної інформації, розробки статистичних формулярів;
- основи зведення та групування матеріалів статистичного спостереження, виявлення зв'язків між окремими явищами та процесами, встановлення його структури;
- методи обчислення узагальнюючих статистичних показників (абсолютних, відносних, середніх) та їх економічна інтерпретація;
- технології виявлення й оцінки взаємозв'язків між явищами і процесами суспільного життя;
- основи вивчення динаміки суспільних явищ, тенденцій та закономірностей їх розвитку;
- теоретичні засади аналізу складних суспільних явищ та виявлення дії окремих факторів на їх розвиток.
- застосовувати сучасні методи економіко-статистичного аналізу під час розгляду соціально-економічних та політичних процесів;
- обґрунтовано і конкретно використовувати абсолютні, відносні та середні показники з метою аналізу соціально-економічних та політичних процесів;
- проводити кількісний аналіз взаємозв'язків між соціально-економічними явищами на базі даних з різноманітних джерел інформації;
- шукати інформацію й уміти інтерпретувати її для проведення комплексного аналізу державної політики;
- аналізувати динаміку та прогнозувати розвиток соціально-економічних процесів мікро- і макрорівня;

ТЕМА 1. ВСТУП ДО ТЕОРІЇ ЕКОНОМІКО-СТАТИСТИЧНОГО АНАЛІЗУ. СТАТИСТИЧНЕ СПОСТЕРЕЖЕННЯ

1. Короткі відомості з історії статистики

Статистика як наукове знання виникла в XVII ст. майже одночасно в Європі й Англії. Проте форма виникнення і її зміст були різні: державоведення у Європі і політична арифметика в Англії, стали потім двома напрямками в статистичній науці.

Засновником політичної арифметики був **Вільям Петті** (1623 - 1687) - яскрава різнобічна особистість: лікар, винахідник копіювальної машини, доктор фізики, професор астрономії, основоположник трудової статистики, державний діяч, один із засновників Лондонського Королівського товариства (1662). Основні твори: «Трактат про податки і збори» (1662г.), «Політична анатомія Ірландії» (1672), «Різне про гроші» (1682г.), «Політична арифметика» (1690г.)

Одне з найважливіших досягнень політичних арифметиків - розуміння необхідності масових даних як умови отримання стійких, доказових висновків. Однак їм доводилося долати великі труднощі при доборі необхідних облікових матеріалів. Спеціальне статистичне спостереження тільки виникало, основна частина вихідних даних збиралася за допомогою матеріалів оподаткування.

Суттєво іншим був підхід державознавства. У Німеччині цей підхід набув вигляду так званого камерального (бюджетного) обліку. Статистика як державознавство базувалася на визнанні держави єдиним джерелом спостереження, єдиним апаратом спостереження були чиновники, адміністратори, керуючі. Представники цього напрямку недооцінювали математичні засоби пізнання. Вимірювальний характер статистики, принаймні на перших етапах, не розглядався як її відмінна риса. Кількісні оцінки трактувалися як окремий випадок загального опису. Звідси й інша назва цього напрямку - описова статистика.

Державознавство часто називалося описовою школою статистики. Одним із основоположників описової статистики був **Герман Конрінг** (1606 - 1681). Відомий німецький лікар, історик і державознавець, професор медицини і політики в Гельмштедте. У 1658 р. шведський король Карл X зробив Конрінга своїм радником і лейб-медиком.

Згідно з вченням Конрінга, державознавство - це опис сучасного стану держави крізь призму географічних, природних, етнографічних та екологічних аспектів, що мають важливе значення для її розвитку та місця на міжнародній арені.

Справжнім засновником теорії статистики є **Адольф Кетле** (1796 - 1874). Він був організатором міжнародних статистичних конгресів, які зіграли велику роль у розвитку практичної і теоретичної статистики. У його роботах досить глибоко розглянута статистика як знаряддя соціального пізнання, її предмет, теоретичний сенс основних методів статистичного дослідження тощо. А. Кетле розкрив статистичний характер закономірностей суспільного життя, які чітко проявляються лише в масі явищ. Його роботи показали значення статистики як могутнього знаряддя соціального пізнання.

2. Предмет та метод статистики. Теоретичні основи статистики, її зв'язок з іншими науками та місце серед них

Статистика як наука пов'язана з такими дисциплінами, як демографія, математика, економетрика, математична статистика, теорія вірогідності, економічна теорія, біологія, фізика, інформатика.

Сьогодні слово **статистика** використовується в декількох значеннях, а саме:

1. Як синонім слова «**відомості**» (**інформація**). Саме в цьому сенсі можна сказати: "статистика народжуваності і смертності" або "статистика злочинів".

2. Як **галузь знань**, яка об'єднує принципи та методи роботи з числовими даними,

що характеризують масові явища.

3. Як галузь **практичної діяльності**, спрямовану на збір, обробку й аналіз статистичних даних.

Отже, **статистика** - одна з суспільних наук, що має на меті збір, упорядкування, аналіз і зіставлення числового представлення фактів, що відносяться до найрізноманітніших явищ.

Предметом вивчення статистики як науки є кількісний бік масових соціально-економічних явищ. Статистика також вивчає вплив природних і технічних факторів на зміну кількісних характеристик. Явища і процеси протікають у житті суспільства вивчаються за допомогою статистичних показників.

Статистичний показник - це кількісна оцінка властивості досліджуваного явища.

Показники в статистиці бувають обліково-оціночні й аналітичні.

1. Обліково-оціночні – це кількісна оцінка певних явищ у конкретних умовах місця і часу. Наприклад, чисельність населення України склала на початок 1991 р. 62 млн.чол.

2. Аналітичні – це показники, що застосовуються для аналізу статистичної інформації і характеризують особливості розвитку досліджуваного явища, а саме: типовість ознаки, співвідношення частин до цілого, швидкість розвитку в часі тощо. В якості аналітичного показника застосовуються відносні і середні величини, показники варіації, динаміки тощо.

Статистика дозволяє виявити та виміряти закономірності розвитку соціально-економічних явищ і процесів, взаємозв'язок між ними. Пізнання закономірностей можливе лише в тому випадку, якщо вивчається не окремі явища, а сукупність явищ.

Статистична сукупність - це безліч одиниць досліджуваного явища, об'єднаних єдиною якісною основою.

Одиниця сукупності - це межа дроблення об'єкта дослідження, при якому зберігаються всі властивості досліджуваного процесу. Одиниці сукупності володіють певними властивостями (якостями) які прийнято називати **ознаками** (ознаки людини: вік, освіта, заняття, зріст, вага, сімейний стан і т.д.; ознаки підприємства: форма власності, спеціалізація (галузь), чисельність працівників, величина статутного фонду, економічна ефективність його діяльності тощо)

Статистика вивчає явища через їх ознак, а саме: чим більш однорідна сукупність, тим більше спільних ознак мають її одиниці і тим менше варіюються їх значення.

Ознаки розрізняються способами їх вимірювання та іншими особливостями, що впливають на прийоми статистичного вивчення. Це дає підставу для класифікації ознак.

Описові ознаки виражаються словесно, зокрема національність людини, різновид ґрунтів, матеріал стін будівлі. Описові ознаки поділяються на *номінальні* і *порядкові*.

- номінальні - це описові ознаки, за якими не можна ранжувати (упорядковувати) дані,
- порядкові - це ознаки, за якими можна ранжувати дані.

Кількісні ознаки виражаються числами (вік людини, площа ріллі, заробітна плата робітників, населення міста, дохід кооперативу і т.д.).

Первинні ознаки характеризують одиницю сукупності в цілому. Це абсолютні величини. Вони можуть бути виміряні, злічені, зважені й існують самі по собі незалежно від їх статистичного вивчення. Наприклад, площа ріллі, потужність двигунів на підприємстві, чисельність населення міста, кількість автомобілів, вироблених в країні.

Вторинні, або **розрахункові**, ознаки не вимірюються безпосередньо, а розраховуються. Вони є продуктами людської свідомості, результатом пізнання досліджуваного об'єкта (собівартість одиниці продукції, продуктивність праці, рентабельність, врожайність тощо). Вторинні ознаки являють собою співвідношення первинних ознак.

Прямі (безпосередні) ознаки - це безпосередні властивості об'єкта (вік людини, поголів'я корів на фермі, обсяг продукції заводу, чисельність його робітників).

Непрямі ознаки є властивостями, притаманними не самому об'єкту, а іншим сукупностям, що належать до об'єкта, які входять до нього. Наприклад, оплата праці робітників по відношенню до заводу. Це непрямая ознака заводу, але дуже важлива для того, хто збирається вступати на роботу і вибирає підприємство.

Практично поділ ознак на прямі і непрямі збігається з їх поділом на первинні та вторинні.

Ознаки розрізняються в статистиці і за характером їх варіації, тобто за відмінностями їх значень у різних одиницях сукупності. Виділяють, по-перше, **альтернативні** ознаки, які можуть приймати тільки два значення (стать людини, місце проживання (місто, село), ходова система трактора (гусеничний або колісний)). По-друге, **дискретні** ознаки, до них відносяться кількісні ознаки, які можуть приймати тільки цілочисельні значення, без проміжних значень між ними (це число членів сім'ї, кількість поверхів будівлі, кімнат у квартирі). По-третє, **безперервні**, точніше безперервно варіюючі, - це ознаки, що здатні приймати будь-які значення, звичайно, в певних межах.

Остання група – це моментні й інтервальні. **Моментні** ознаки характеризують досліджуваний об'єкт у певний момент часу, установлений планом статистичного дослідження. Вони характеризують наявність чого-небудь (наприклад, чисельність населення, вартість фондів, кількість худоби, розміри житлової площі).

До **інтервальних** відносяться ознаки, що характеризують результати процесів. Тому їх значення можуть виникати тільки за інтервал часу (рік, місяць, добу), але не на момент часу. Наприклад, число народжених, померлих, обсяг промислової продукції, надій молока, сума отриманого прибутку.

Моментні ознаки – це характеристики *стану*, а інтервальні – характеристики *процесу*. Різниця між *моментними* й *інтервальними* ознаками істотна при вивченні динаміки. Одиниці виміру моментних ознак відносяться тільки до характеристики властивостей об'єктів, а одиниці вимірювання інтервальних ознак містять ще й вказівку того відрізка часу, за який визначено значення ознаки.

3. Завдання статистики. Галузі статистичної науки

Статистика як наука виробила прийоми і способи вивчення масових суспільних явищ, які утворюють статистичну методологію.

Під **статистичною методологією** розуміється система прийомів, способів і методів, спрямованих на вивчення кількісних закономірностей, що виявляються в структурі, динаміці та взаємозв'язках соціально-економічних явищ.

Статистичне дослідження складається з трьох етапів:

1 етап: статистичне спостереження - науково організований збір відомостей про досліджуваних соціально-економічних процесах чи явищах. Характерним для цієї стадії є метод масових спостережень. Отримані в результаті статистичного спостереження дані є вихідним матеріалом для виконання наступних стадій статистичного дослідження.

2 етап: зведення й групування результатів спостереження. Даний етап статистичного дослідження являє собою комплекс послідовних дій щодо узагальнення конкретних фактів, що утворюють сукупність з метою виявлення типових рис і закономірностей. Найважливішим методом на даній стадії є метод групувань.

3 етап: аналіз отриманих узагальнюючих показників. Це заключна стадія статистичного дослідження, для якої характерне застосування узагальнюючих показників (абсолютних, відносних і середніх величин, індексного методу).

Виходячи з вище викладеного можна сформулювати такі завдання:

1. Загальні завдання:

1. Вивчення рівня і структури явища.
2. Вивчення взаємозв'язків явищ.
3. Вивчення динаміки явищі.

2. Прикладні завдання:

1. Удосконалення обліку та звітності, скорочення на цій основі документообігу.
2. Посилення роботи з контролю за достовірністю статистичної інформації, що надається підприємствам, установам та організаціям усіх галузей економіки та форм власності.
3. Підвищення своєчасності статистичної інформації.
4. Поглиблення аналітичних функцій в напрямку соціально-економічного розвитку країни.
5. Подальший розвиток і вдосконалення статистичної методології на основі все більш широкого впровадження ПЕОМ.

Отже, основна продукція статистики - це застосування узагальнюючих статистичних показників.

4. Статистичні спостереження. План статистичного спостереження. Форми, види і способи статистичних спостережень

Статистичне спостереження - планомірний науково обґрунтований збір даних або відомостей про соціально-економічні явища і процеси суспільного життя шляхом реєстрації, заздалегідь намічених істотних ознак з метою отримання в подальшому узагальнюючих характеристик цих явищ і процесів.

Розрізняють такі форми організації статистичного спостереження:

- статистична звітність;
- спеціально організовані статистичні обстеження (спостереження);
- реєстри.

1. **Статистична звітність** - це особлива форма організації збору даних державною статистикою про діяльність господарюючих суб'єктів через спеціально заповнювані документи-бланки, іменовані формами статистичної звітності.

Форма статистичної звітності - це спеціальний документ (здебільшого у формі бланку), що містить перелік певних показників, відомостей, що характеризують те чи інше явище, процес і результати їх розвитку (як позитивні, так і негативні), наданий до державних статистичних органів для подальшого узагальнення ними даних.

У відповідності з термінами подання звітність буває **добова (щоденна), тижнева, місячна, квартална, піврічна, річна**. Усі згадані вище види звітності, крім річної, об'єднують однією назвою - **поточна звітність**. Як правило, чим більше період, за який подається звітність, тим більше остання містить показників.

За способом передачі (подання) відомостей звітність поділяють на **електронну, телеграфну, телетайпну, поштову**.

2. **Спеціально організовані статистичні спостереження** - проводяться у вигляді переписів або спеціальних обстежень (вибіркових або суцільних). Їх застосовують у всіх випадках, коли необхідно отримати відомості, по яких відсутня звітність, коли потрібно уточнити або доповнити дані тієї чи іншої звітності або провести разове детальне, всебічне обстеження яких об'єктів.
3. **Регістр** - це поіменованій і постійно уточнюваний перелік тих чи інших одиниць спостереження, створений для безперервного тривалого статистичного спостереження за певною сукупністю. У реєстрі міститься інформація про кожну одиницю сукупності.

Етапи статичного спостереження:

1. Програмно-методологічна підготовка проведення спостереження.
2. Організаційна підготовка проведення спостереження.
3. Збір даних статистичного спостереження.
4. Контроль якості даних статистичного спостереження.
5. Вироблення висновків і пропозицій щодо удосконалення статистичного спостереження.

Статистичне спостереження має відповідати таким вимогам:

- Спостережуване явище повинне мати наукову або практичну цінність, висловлювати

певні соціально-економічні типи явищ.

- Безпосередній збір масових даних повинен забезпечити повноту фактів, що відносяться до розглянутого питання, так як явища перебувають у постійній зміні, розвитку.
- Для забезпечення достовірності статистичних даних необхідна ретельна та всебічна перевірка якості зібраних характеристик статистичного спостереження.
- Для створення найкращих умов для отримання об'єктивних матеріалів необхідна наукова організація статистичного спостереження.

Види і способи статистичного спостереження

У результаті того, що статистичне спостереження має різноманітні види необхідно навести таку класифікацію:

Поточне (безперервне) спостереження ведеться систематично, постійно, безперервно у міру виникнення, функціонування та розвитку явищ і процесів.

Перерване спостереження буває одноразовим і періодичним. Перерване спостереження, насамперед, відіграє роль моментальної інформації, що відбиває стан досліджуваного явища на певний момент часу, певну дату.

Періодичне спостереження - реєстрація проводиться через певні періоди, проміжки часу (5 років, 10 років).

Одночасне спостереження проводиться один раз для вирішення якого-небудь завдання або повторюється через невизначений проміжок часу.

При **суцільному** спостереженні реєстрації підлягають усі без винятку одиниці сукупності. При суцільному спостереженні ставиться завдання отримати відомості про всі одиниці досліджуваної сукупності. Даний вид спостережень вимагає великих витрат (фінансових і трудових) як при зборі інформації, так і при подальшій її обробці.

До **несуцільного** спостереження вдаються у тих випадках, коли фізично неможливо, важко або недоцільно здійснити суцільне спостереження, зокрема, коли спостереження спричиняє псування або знищення спостережуваної одиниці.

Спосіб основного масиву - обстеженню піддається основний масив, тобто та частина одиниць, яка має найбільшу вагу у досліджуваному явищі.

Вибіркове спостереження - це такий вид несучільного спостереження, при якому з усієї досліджуваної сукупності випадково, власне, навмання (шляхом жеребкування або іншим методом), відбирається певна кількість одиниць (вибіркова сукупність). Статистика-дослідника цікавлять ознаки, на підставі яких обчислюються необхідні вибіркові показники (середні величини, відносні тощо), які поширюють потім на вихідну основну сукупність.

Монографічне спостереження - докладний опис окремих одиниць сукупності з метою їх поглибленого вивчення.

Анкетне спостереження проводиться таким чином: організації (або установи), які поставили перед собою завдання вивчити те чи інше питання, складають і розсилають (або роздають) певному колу осіб (респондентам) особливі *анкети* з питаннями. Відтак, організація цього спостереження, що передбачає неповне повернення розісланих (або розданих) анкет, дає підставу відносити цей вид спостереження (анкетне) до несучільного.

Реєстрація необхідних відомостей при статистичному спостереженні може проводитися на основі різних джерел. При цьому виділяють такі способи статистичного спостереження:

1. **Безпосереднє спостереження** здійснюється шляхом реєстрації досліджуваних одиниць та їх ознак на підставі безпосереднього огляду, підрахунку, зважування тощо.

2. **Документальний спосіб** передбачає використання в якості статистичних відомостей різних документів первинного обліку.

3. **Опитування** – це спосіб статистичного спостереження, для якого джерелом виступають відомості, які дають опитуваній особі.

Розглянуті вище способи статистичного спостереження розрізняються не тільки за джерелами зібраних відомостей, але залежно від **організації їх збору**, особливо це стосується збору відомостей шляхом опитування. Так, опитування може бути здійснене такими

способами: експедиційним; способом самореєстрації; кореспондентським; явочним.

При **експедиційному способі** спеціально підготовлені працівники (реєстратори) відправляються до осіб (одиницям), від яких повинні бути отримані відомості, і на місці опитують їх усно, реєструючи відповіді в спеціальних бланках.

При **само реєстрації використовуються** переписні листи (бланки), які заповнюють самі респонденти (опитувані), а реєстратори роздають респондентам ці бланки, пояснюють правила їх заповнення і збирають заповнені опитувальні листи.

При **кореспондентському способі** відомості по заздалегідь визначеному колу показників надходять до органів, що ведуть спостереження, від спеціального штату добровільних кореспондентів, що працюють на місцях, звідки отримуються дані.

При **явочному способі** спостереження передбачається, що відомості з певного кола показників повинні повідомлятися в органи, що ведуть спостереження за цими явищами, в явочному порядку.

Програмно-методологічні й організаційні питання статистичного спостереження

Перш ніж приступити безпосередньо до проведення статистичного спостереження необхідно вирішити ряд питань, а саме:

Програмно-методологічні питання включають в себе:

- визначення мети, об'єкта й одиниці спостереження;
- розробку програми спостереження та статистичного формуляра, що містить її.

При плануванні будь-якого статистичного спостереження, перш за все необхідно точно сформулювати його **мету**, оскільки тільки конкретна мета, конкретні завдання дослідження визначають ті відомості, які повинні бути отримані в процесі статистичного спостереження.

Об'єкт спостереження - сукупність одиниць, відомості про стан яких повинні бути отримані. Визначити об'єкт спостереження означає точно встановити межі досліджуваної сукупності, тобто вирішити, що має бути обстежено або хто повинен бути обстежений у процесі статспостереження.

Одиницею спостереження в статистиці називають ту одиницю, той елемент об'єкта спостереження, який характеризується рядом ознак і до якого ведеться реєстрація цих ознак.

Програмою статистичного спостереження називається перелік показників, що підлягають вивченню. У програму спостереження повинні включатися тільки ті питання, які відповідають завданням дослідження, на які може бути отримані правдиві, достовірні відповіді. Питання повинні бути сформульовані таким чином, щоб їх зміст всюди і всіма розумілося однаково.

Статистичні формуляри, що містять програму і результати реєстрації, зустрічаються двох видів: **індивідуальні** (карткові) і **спискові**. У першому випадку формуляр (індивідуальний) заводиться на кожен одиницю спостереження окремо, тобто в кожному формулярі містяться відомості лише про одну одиницю спостереження. У другому випадку один формуляр (список) складається на кілька одиниць.

Організаційні питання статистичного спостереження включають у себе вирішення таких важливих моментів, як визначення: суб'єкта спостереження; місця і часу спостереження; організаційної форми, виду і способу спостереження.

Визначення **суб'єкта спостереження** зводиться до вирішення питання про те, хто буде здійснювати статистичне спостереження. Це можуть бути: 1) органи статистики; 2) можуть залучатися широкі кола громадськості. 3) може бути залучено саме населення, роздають анкети, а потім збирають їх вже заповненими (це так звана самореєстрація).

При плануванні статистичного спостереження необхідно вирішити ряд питань, а саме: про **час спостереження**, тобто визначити, коли воно буде проведено.

В обраному періоді важливо визначити певний момент, іменованій **критичним**, станом на який повинні реєструватися всі відомості. Від критичного моменту (крапки відліку) слід відрізнити **термін спостереження** або **час виробництва спостереження**, -це відрізок часу, протягом якого повинні бути зібрані відомості про явище.

Крім питання про час спостереження, вирішується також питання про **місце**

спостереження, тобто про те, де слід проводити спостереження. (При переписах населення місцем проведення важали місце проживання, окремі квартири).

До початку проведення статистичного спостереження має бути вирішено, якою буде **організаційна форма спостереження**, яке спостереження буде застосовано за охопленням одиниць (суцільне або несучільне; якщо несучільне, то який вигляд воно повинно мати: вибірковий, основний масив, анкетне спостереження), на підставі яких джерел буде проводитися реєстрація відомостей (шляхом безпосереднього спостереження на основі документів, шляхом опитування), а також яким **способом** буде організовано збір інформації (експедиційним, явочним або іншим).

5. Помилки та контроль статистичних спостережень

Ретельно розроблені і продумані питання статистичного спостереження – це запорука успіху в отриманні якомога більшої кількості достовірних даних про досліджуване явище або процес. При будь-якому спостереженні можуть виникнути помилки (похибки).

Помилками спостереження називають розбіжності між установленими статистичним спостереженням і дійсними значеннями досліджуваних величин.

1. Помилки реєстрації виникають через неправильне встановлення фактів у процесі спостереження або помилкового їх запису, або і того й іншого.

Випадковими помилками називають помилки, які утворюються внаслідок різних випадкових причин. При досить великій кількості спостережень у результаті дії закону великих чисел ці помилки більш-менш взаємно усуваються (погашаються).

Систематичні помилки виникають із певних постійних причин, зокрема у бік перебільшення чи применшення розміру даних, що призводить до серйозних перекручувань загальних результатів статистичного спостереження.

Навмисними помилками називаються помилки, причиною яких є свідоме спотворення даних. *Ненавмисними помилками* називають помилки, які носять випадковий, ненавмисний характер, наприклад, несправності, збої вимірювальних приладів.

2. *Помилки репрезентативності* виникають при несучільному спостереженні. Бувають випадковими і систематичними, навмисними і ненавмисними.

Випадкові помилки виникають в силу того, що сукупність відібраних на основі принципу випадковості одиниць спостереження відтворює не всю сукупність.

Систематичні помилки виникають унаслідок порушення принципу випадковості відбору одиниць досліджуваної сукупності, які повинні бути піддані спостереженню. Розміри цих помилок, як правило, не піддаються кількісному виміру.

Перевірка правильності зафіксованих у статистичних формулярах відомостей повинна проводитися з точки зору **логічного та арифметичного контролю**.

Арифметичний контроль передбачає перевірку результатів, перерахунок отриманих показників на предмет встановлення похибки. Такий контроль доцільно здійснювати за допомогою ЕОМ.

Логічний контроль має за мету визначити відповідність відповіді на поставлене питання або відповідність між відповідями на різні запитання програми.

Контрольні питання

1. Короткі відомості з історії статистики
2. Хто є засновником державоведення і політичної арифметики у Європі?
3. Предмет та метод статистики.
4. Теоретичні основи статистики, її зв'язок з іншими науками та місце серед них.
5. Що визначають програмно-методологічні питання плану спостереження?
6. Що є програмою спостереження?
7. Завдання статистики.
8. Галузі статистичної науки.
9. Статистичні спостереження.
10. План статистичного спостереження.
11. Форми, види і способи статистичних спостережень.

12. Помилки та контроль статистичних спостережень.

ТЕМА 2. ЗВЕДЕННЯ ТА ГРУПУВАННЯ СТАТИСТИЧНИХ ДАНИХ. ГРАФІЧНИЙ МЕТОД ЗОБРАЖЕННЯ СТАТИСТИЧНИХ ДАНИХ

1. Суть статистичного зведення та групування. Поняття про зведення, його організацію та види групувань

Завдання зведення і його зміст. Інформація про окремі одиниці сукупності, одержувана в процесі статистичного спостереження, характеризує їх з різних боків. Найважливішим етапом дослідження є систематизація первинних даних і отримання на цій основі зведеної характеристики об'єкта в цілому за допомогою узагальнюючих показників, що досягається шляхом зведення й групування первинного статистичного матеріалу.

Статистичні дані – це систематизація одиничних фактів, що дозволяє перейти до узагальнюючих показників для виявлення типових рис і закономірностей, властивих досліджуваному явищу в цілому.

Статистичні дані – це науково організована обробка матеріалів статистичного спостереження в цілях отримання узагальнюючих характеристик досліджуваного явища по ряду істотних для нього ознак

Статистична зведення розрізняється за низкою ознак:

| | |
|--|---|
| За складністю побудови | просте зведення - операції з підрахунку загальних підсумків за сукупністю одиниць спостереження |
| | складне зведення - комплекс операцій, що включають групування одиниць спостереження, підрахунок підсумків по кожній групі і по всьому об'єкту в цілому і представленні результатів у вигляді <i>статистичних таблиць</i> |
| За місцем проведення | централізоване , коли весь первинний матеріал надходить в одну організацію і піддається в ній обробці від початку до кінця |
| | децентралізоване , коли звіти зводяться місцевими статистичними органами, а підсумки надходять у Держкомстат України і там визначаються підсумкові показники в цілому по країні в цілому |
| За способом обробки матеріалів статистичного спостереження | механізоване - усі операції здійснюються з використанням ЕОМ |
| | ручне - усі основні операції здійснюються людиною (групою людей) власноруч |

Статистичне зведення проводиться за спеціально розробленою програмою, яка визначається завданнями статистичного дослідження. Зведення статистичної інформації, як правило, не обмежується отриманням загальних підсумків. Найчастіше вихідна інформація на цій стадії систематизується, утворюються окремі статистичні сукупності, тобто здійснюється статистична групування.

Отже, **статистичне групування** - це процес утворення однорідних груп на основі поділу статистичної сукупності на частини або об'єднання одиниць у групи сукупностей за певними, істотним для них ознаками.

2. Основні питання методології статистичних групувань. Вторинне групування

Групувальна ознака (підстава групування) - це ознаки, за якими проводиться розподіл одиниць спостереження сукупності на групи.

Відповідно до завдань групування поділяють на такі види: типологічне, структурне й аналітичне.

Типологічне групування - розбивка різнорідної сукупності на якісно однорідні групи і виявлення на цій основі економічних типів явищ.

Структурне групування - групування, яке призначене для вивчення складу однорідної сукупності за будь-якою варіюючою ознакою або кількома ознаками.

Аналітичне групування - групування, що виявляє взаємозв'язок між досліджуваними ознаками (факторними). **Факторними** називаються ознаки, під впливом яких змінюються інші ознаки, тобто вони і утворюють групу результативних ознак.

Узаємозв'язок виявляється в тому, що із зростанням значення факторної ознаки систематично зростає або знижується значення результативної ознаки.

Особливостями аналітичного групування є те, що одиниці групуються за факторною ознакою, кожна виділена група характеризується середніми значеннями результативної ознаки.

Усі розглянуті види групувань можуть бути побудовані за одним або кількома істотним ознаками і мають назву простого або складного групування.

Просте - це групування, в якій, групи утворені за однією ознакою.

Складне - це угруповання, в якій поділ сукупності на групи проводиться за двома і більше ознаками, об'єднаними в ціле.

Спочатку групи формуються за однією ознакою, потім групи діляться на підгрупи за іншою ознакою і т.д. Складні групування дають можливість вивчити одиниці сукупності одночасно за кількома ознаками.

Складне групування будується в такій послідовності: спочатку проводиться групування за атрибутивними ознаками, потім за кількісними.

Виокремлюють також **багатовимірне групування** - це групування, в якому поділ сукупності на групи проводиться за двома і більше ознаками, взятими одночасно. Розвиток багатовимірного групування вплинув на розробку кластерного аналізу.

Термін **кластерний аналіз** (перше він був уведений у 1939 р.) насправді включає в себе набір різних алгоритмів класифікації. Сама назва методу походить від того ж кореня, що й слово «клас», «класифікація». Англійське слово «the cluster» має таке значення: група, тобто об'єднання однорідних явищ. Кожна одиниця сукупності в кластерному аналізі розглядається як крапка в заданому просторі ознак. Значення кожної з ознак у даній одиниці служить її координатою в цьому «просторі» за аналогією з координатами точки в тривимірному просторі.

3. Статистичні таблиці, їх види, правила побудови та аналізу

Перед тим, як дати визначення статистичної таблиці, слід вказати принципи побудови статистичних групувань, етапи їх побудови.

Увесь процес побудови групування можна розбити на декілька етапів:

1 етап: Визначають складу групувальну ознаку. При цьому в основу групування можуть бути покладені як кількісні, так і якісний (атрибутивні) ознаки.

2 етап: Визначають кількість груп, на які треба розбити сукупність.

Якщо групування будується за якісною (атрибутивною) ознакою, то груп визначається стільки, скільки є градацій (стать людини, отже, виділяють 2 групи) .

Якщо групування будується за кількісною ознакою, то слід звернути увагу на число одиниць об'єкта. При невеликому обсязі не слід утворювати занадто велике число груп.

Визначення числа груп можна здійснити математичним шляхом за допомогою формули Стерджесс:

$$n = 1 + 3,322 \text{ Lg } N$$

n - число груп

N -число одиниць сукупності

3 етап: Визначають інтервал групування.

Інтервал - це значення варіюючої ознаки, що знаходиться в певних межах. Кожен інтервал має свою величину – **нижню** (найменше значення ознаки в інтервалі) і **верхню** (відповідно, найбільше значення ознаки в інтервалі).

Якщо варіація ознаки проявляється у вузьких межах і розподіл носить рівномірний характер, то будують групування з рівними інтервалами.

Величина рівного інтервалу визначається за формулою:

$$h = R / n = (X_{max} - X_{min}) / n$$

R - розмах сукупності

X_{max} і X_{min} - максимальне та мінімальне значення ознаки в сукупності

n - число груп

Якщо розмах варіації великий і значення варіюють нерівномірно, то використовують угруповання з нерівними інтервалами.

Інтервали групування можуть бути **відкритими** (указана тільки одна межа - верхня або нижня) і **закритими** (у яких є і верхня межа, і нижня межа).

Правила визначення інтервалу групування:

1. Якщо величина інтервалу має один знак до коми (наприклад, 0,7; 0,58; 2,359), то отримане значення слід округлити до десятих (тобто 0,7; 0,6; 2,4).

2. Якщо величина інтервалу має дві значущі цифри до коми і кілька після коми (наприклад, 11,2; 23,385), то це значення слід округлити до цілого числа (тобто 11; 23).

3. Якщо величина інтервалу являє собою тризначне число (наприклад, 123; 757), то це значення доцільно округлити до найближчого числа, кратного 10 (тобто 120; 760).

4. Якщо інтервали груп закриті і підставою групування служить безперервна ознака, то нижня межа формується за принципом «включно», а верхня - за принципом «винятково» (наприклад, якщо нижня межа i -групи дорівнює 50, а верхня - 100, то одиниця сукупності із значенням ознаки, рівним 100, потрапить у групу $i + 1$).

5. Якщо значення ознаки збігається з межами інтервалів, то можна використовувати відкриті інтервали, увівши слова «до», «менше» і «більше».

6. Якщо в підставі групування лежить дискретна ознака, то верхня межа i -го інтервалу дорівнює нижній межі $(i + 1)$ -го інтервалу, збільшеної на 1.

4 етап: Проведення групування, складання таблиці.

| | | | | |
|---|----|----|----|----|
| - | - | -- | . | - |
| . | -- | -- | - | |
| - | -- | - | -- | - |
| - | -- | - | - | -- |

Отже, **статистична таблиця** – це система рядків і стовпців, в яких у певній послідовності і зв'язку викладається статистична інформація про стан соціально-економічних явищ. У статистичній таблиці розрізняють підмет (об'єкт дослідження) і присудок (система показників, що характеризують об'єкт). Власне, у **підметі** вказується характеризується об'єкт - або одиниці сукупності, або група одиниць, або сукупності в цілому. А в **присудку** дається характеристика підмета в кількісній формі - у вигляді системи показників.

Підмет таблиці розташовується в лівій частині, присудок - в правій, але можуть бути й виключення. За характером підмета таблиці поділяються: **прості і комбіновані**.

При побудові статистичних таблиць користуються такими правилами:

1. Заголовок таблиці, назва рядків і граф мають бути чіткими, зрозумілими і по можливості лаконічними. Загальний заголовок повинен відображати основний зміст таблиці, час і місце, до яких належать показники

2. У заголовку повинні бути вказані одиниці виміру, якщо вони однакові для всіх клітин таблиці. Якщо ж вони різні, то у верхніх або бічних заголовках необхідно вказати, в яких одиницях наведено дані.

3. Числа в клітинах доцільніше (по можливості) округляти; округлення чисел в межах однієї і тієї ж графи або рядка слід проводити з однаковим ступенем точності.

4. Якщо дані запозичені, то під таблицею вказується джерело.

5. У разі необхідності можна давати примітку до таблиці, в якому розкривається методика розрахунку показників.

4. Ряди розподілу, їх суть, види, графічне зображення. Основні види графіків, правила їх побудови

Після визначення групової ознаки будується статистичний **ряд розподілу** – це послідовність чисел, які характеризують склад або структуру соціально-економічного явища залежно від групової ознаки.

Ряди розподілу відповідно до групових ознак, за якими їх було утворено, можуть бути:

- 1) *атрибутивними* (розподіл населення за статтю, національністю, розподіл підприємств за формою власності, розподіл злочинів за тяжкістю скоєного правопорушення тощо);
- 2) *варіаційними або дискретними* (розподіл сімей за кількістю дітей, розподіл міст за кількістю жителів, розподіл підприємств за кількістю працюючих);
- 3) *інтервальними* (розподіл зарплати працюючих в установі, розподіл підприємств за величиною основних фондів).

Статистичний ряд динаміки - це послідовність чисел, які характеризують зміну явища в часі.

Графічним зображенням дискретних варіаційних рядів є полігон розподілу частот (набір перпендикулярних відрізків, довжина відрізка відображає частоту відповідної ознаки), інтервальних – гістограма (набір суміжних стовпчиків, ширина яких відповідає ширині інтервалу, а висота – частоті).

Отже, **статистичний графік** - це креслення, на якому статистичні сукупності, що характеризуються певними показниками, описуються за допомогою умовних геометричних образів або знаків.

При побудові графічного зображення слід дотримуватися ряду вимог, а саме:

По-перше, у відповідності з метою використання вид графічного зображення.

По-друге, визначається поле графіка, - той простір, в якому розміщуються геометричні знаки.

По-третє, задаються масштабні орієнтири за допомогою масштабних шкал.

По-четверте, вибирається система координат, необхідна для розміщення геометричних знаків у полі графіка. Найбільш поширеною системою координат є система прямокутних координат. При цьому найкраще співвідношення масштабу по осях абсцис і ординат, рівне 1,62:1, називається «Золотим перетином».

По-п'яте, графік складається з графічного образу і допоміжних елементів.

Підсумовуючи, варто зазначити як переваги, так і недоліки графічного зображення.

Переваги: 1. Воно наочне й виразне. 2. Відразу видно межі зміни показника, порівняльна швидкість зміни. Недоліки: 1. Включають меншу кількість даних, ніж у таблиці. 2. На графіку показуються округлені дані, тобто загальна ситуація, але не деталі.

Відтак, доцільним є зважене й раціональне використання всіх способів статистичного аналізу.

Контрольні питання

1. Суть статистичного зведення та групування.
2. Поняття про зведення, його організацію та види групувань.
3. Основні питання методології статистичних групувань.
4. Вторинне групування.
5. Як здійснюється розподіл неоднорідної сукупності на якісно однорідні групи?
6. Статистичні таблиці, їх види, правила побудови та аналізу.
7. Ряди розподілу, їх суть, види, графічне зображення.
8. Основні види графіків, правила їх побудови.

ТЕМА 3. АБСОЛЮТНІ, ВІДНОСНІ ТА СЕРЕДНІ ВЕЛИЧИНИ

ПЛАН

1. Поняття та види статистичних показників.
2. Абсолютні статистичні показники.
3. Поняття та види відносних величин. Одиниці вимірювання статистичних показників.
4. Суть та умови використання середніх величин. Види середніх величин (середня арифметична, середня геометрична, середня квадратична, середня гармонійна, середня хронологічна).
5. Середня арифметична та її властивості. Порядкові середні (мода, медіана).

1. Поняття та види статистичних показників

Статистичні показники відображають кількісну сторону досліджуваної сукупності суспільних явищ, власне, представляють собою їх величину, виражену відповідною одиницею виміру.

Отже, статпоказник – це узагальнююча характеристика; міра як єдність якісного та кількісного відображення певної властивості явища чи процесу.

Статистичні показники, відображаючи економічні категорії, мають взаємозалежні кількісну й якісну сторони.

Якісний бік відображається у змісті показника безвідносно до конкретного розміру ознаки. Кількісний же бік статистичного показника – це його числове значення. Наприклад, обсяг роздрібного товарообігу магазину в поточному році склав 3.5 млн. грн.

Крім якісного та кількісного змісту сьатпоказника, виокремлюють також його логічну формулу, що розкриває його статистичну структуру.

Статистичні показники виконують ряд функцій, а саме:

Пізнавальна функція полягає в тому, що вони характеризують стан і розвиток досліджуваних явищ, напрям та інтенсивність процесів.

Управлінська функція полягає в тому, що на базі статистичних показників проводиться аналіз стану явища і приймається відповідне управлінське рішення.

Функція моніторингу передбачає те, що за допомогою статистичних показників проводиться постійне спостереження за соціально-економічними явищами і процесами.

Прогностична функція означає, що статистичні показники використовуються для передбачення (прогнозування) майбутнього.

Оціночна функція полягає в тому, що за допомогою статпоказників домогосподарства, суспільство і держава оцінюють діяльність підприємств, організацій, трудових колективів і урядів.

Рекламно-пропагандистська функція відзначається тим, що за допомогою статистичних показників втілюються в життя будь-які ідеї або ілюструються явища.

Усі статистичні показники поділяються на три групи: *абсолютні, відносні і середні*. Вихідною, первинною формою вираження статистичних показників є абсолютні величини.

Крім того, у залежності від способу обчислення статпоказники можуть бути *первинними* (визначаються за допомогою зведення та групування даних і подаються у формі абсолютних величин, наприклад, чисельність працюючих в установі, розмір заборгованості перед бюджетом тощо) і *похідними* (обчислюються на базі первинних показників і подаються у вигляді відносних чи середніх величин, наприклад, розмір середньомісячної зарплати, індекс споживчих цін, густина населення тощо).

За ознакою часу статпоказники можуть бути *інтервальними* (характеризують явище за день, місяць, квартал, рік) і *моментними* (характеризують явище на початок року, кінець місяця тощо).

2. Абсолютні статистичні показники

Абсолютні показники - це число одиниць за сукупністю в цілому або за її окремим групам, яке отримують у результаті підсумовування зареєстрованих значень ознак первинного статистичного матеріалу. Дані показники можуть бути отримані й розраховані на основі інших показників (наприклад, приріст банківських вкладів населення за період визначається як різниця вкладів на кінець і початок періоду).

Абсолютні величини отримують у результаті статистичного спостереження і зведення статистичної інформації. Практично статистична інформація формується з абсолютних величин.

Абсолютні величини як узагальнюючі показники характеризують або **чисельність сукупності** (чисельність економічно активного населення, кількість підприємств різних форм власності тощо), або **обсяг ознак сукупності** (розмір інвестицій, витрати на робочу силу тощо).

Будь-яка абсолютна величина завжди має свою одиницю виміру, притаманну тим чи іншим явищам.

За способом вираження розмірів абсолютні величини поділяються на такі:

Індивідуальні абсолютні показники, які отримують безпосередньо в процесі статистичного спостереження. Вони – це результат виміру, зважування, підрахунку й оцінки; характеризують розміри кількісних ознак окремих одиниць. Цей вид показника служить підставою при статистичного зведення і групування.

Зведені об'ємні показники характеризують обсяг ознаки або обсяг сукупності як в цілому, так і з якої-небудь його частини, їх (показники) отримують у результаті зведення й групування індивідуальних значень. Серед зведених абсолютних показників виділяють групові (підсумкові) та загальні (сумарні).

Абсолютні величини мають вагоме наукове та практичне значення, тому що їх використовують для розробки планів соціально-економічного розвитку регіонів і для контролю процесу їх (планів) виконання.

Отже, *абсолютні величини* – це завжди числа іменовані, що мають розмір. Залежно від соціально-економічної сутності досліджуваних явищ, їх фізичних властивостей, вони виражаються в натуральних, вартісних і трудових одиницях виміру. Хоча деякі вчені серед одиниць вимірювання абсолютних величин (АВ) виокремлюють натуральні, умовно-натуральні, вартісні і трудові. Вибір одиниці виміру АВ залежить від сутності явища, його фізичних і соціально-економічних властивостей, а також від цілей дослідження.

Натуральні показники характеризують фізичний обсяг, площу явища. Наприклад, виробництво е/енергії в Україні склало 79 млрд. кВт / ч. якщо для характеристики якогось явища однієї одиниці вимірювання недостатньо, використовують 2 одиниці. Наприклад, тонно-км, пасажиро-км.

Вартісні показники характеризують явища і процеси в грошовому вираженні (грн., тис., млн., \$ тощо) і дозволяють отримувати різні зведені показники.

Трудові показники дозволяють ураховувати як загальні витрати праці на підприємстві, так і трудомісткість окремих операцій технологічного процесу. (Чол-год, чол-день).

3. Поняття та види відносних величин. Одиниці вимірювання статистичних показників

Під час аналізу статистичної інформації важливе місце займають похідні узагальнюючі показники – *середні і відносні*.

Відносні показники – це результат ділення одного абсолютного показника на інший, що виражає співвідношення між кількісними характеристиками соціально-економічних явищ і процесів. Вони виражають співвідношення між кількісними характеристиками соціально-економічних процесів і явищ, а також інтенсивність розвитку досліджуваного явища в часі.

При розрахунку відносних показників (див. формулу нижче), абсолютний показник,

що знаходиться в чисельнику, одержуваного відношення, називається поточним або порівнюваним. Показник, з яким проводиться порівняння і який знаходиться в знаменнику, називається підставою або базою порівняння.

$$ОП = \frac{\text{текущий или сравняемый показатель}}{\text{основание или база сравнения}}$$

Відносні показники можуть виражатися в коефіцієнтах, бути іменованими числами. У разі якщо підстава приймається за 100, відносна величина виражається у відсотках (%), Якщо за 1000 – у проміле (‰), якщо за 10 000 – в продецимілле (‱).

Усі використовувані в статистиці відносні показники можна поділити на такі види:

1. Відносний показник (величина) динаміки (ОПД). Показує у скільки разів поточний рівень перевищує базовий.

$$ОПД = \frac{\text{текущий уровень}}{\text{предшествующий (базовый) уровень}}$$

2. Відносний показник (величина) структури (ОПВ). Розраховані величини, звані частками або питомими вагами, показують, якою часткою володіє або який питому вагу має та чи інша частина в загальному підсумку.

$$ОПС = \frac{\text{показатель характеризующий часть совокупности}}{\text{показатель по всей совокупности}}$$

3. Відносний показник (величина) координації (ОПК) отримують як співвідношення між частинами одного цілого.

$$ОПК = \frac{\text{количество единиц i - ой части совокупности}}{\text{количество единиц j - ой части совокупности}}$$

Характеризує у скільки разів дана частина більше базисної, або скільки% від неї становить, або скільки одиниць однієї частини припадає на 1 (10, 100, 1000, 10000) одиницю іншої частини.

4. Відносний показник (величина) інтенсивності (УПІ). Його отримують, зіставляючи різнойменні ознаки однієї сукупності, а також об'єкти двох зв'язаних між собою сукупностей. Прикладами такого роду показників можуть служити коефіцієнт народжуваності (кількість народжених у розрахунку на 1000 осіб населення), рівень зайнятості (відношення числа незайнятих до чисельності економічно активного населення). Найчастіше вони являють собою співвідношення різнойменних, але пов'язаних явищ, де в чисельнику знаходиться величина явища, а в знаменнику – обсяг того середовища, в якому відбувається розвиток досліджуваного явища. Найчастіше їх розраховують на 100 або 1000 одиниць.

$$ОПИ = \frac{\text{показатель характеризующий явление А}}{\text{показатель характеризующий среду распределения явление А}}$$

5. Відносний показник (величина) порівняння (ОГсер) отримують в результаті зіставлення однойменних абсолютних показників, що відносяться до різних сукупностей, різних об'єктів.

$$ОПСр = \frac{\text{показатель характеризующий объект А}}{\text{показатель характеризующий объект Б}}$$

6. Відносний показник (величина) плану (ОПП) або відносний показник планового завдання. Показує у скільки разів запланований обсяг виробництва перевищить досягнутий рівень чи скільки % від цього рівня складе.

$$\text{ОПП} = \frac{\text{уровень планируемый на } t + 1 \text{ период}}{\text{уровень достигнутый в } t - \text{ом периоде (или базовый)}}$$

7. Відносний показник реалізації плану (ОПРП) або відносний показник виконання плану. Відбиває фактичний обсяг виробництва в% або коефіцієнт у порівнянні з плановим рівнем.

$$\text{ОПРП} = \frac{\text{уровень достигнутый в } (t + 1) - \text{й период}}{\text{уровень планируемый на } (t + 1) - \text{й период}}$$

8. Відносний показник (величина) диференціації. Це результат зіставлення двох взаємопов'язаних структурних рядів. Він дозволяє виявити ступінь нерівномірності розподілу ознаки в сукупності, його диференціацію (Крива Лоренца).

$$\text{ОПДф} = \frac{\text{структура совокупности по числу едениц}}{\text{структура совокупности по размеру выбранного признака}}$$

Отже, відносними статистичними величинами є показники, що виражають кількісні співвідношення між соціально-економічними явищами та їх ознаками. Вони отримуться у результаті ділення однієї величини на іншу. Найчастіше відносні величини є відношенням двох абсолютних величин, наприклад, поділивши чисельність міського населення на чисельність усього населення країни (міського і сільського), одержимо відносну величину – частку міського населення.

4. Суть та умови використання середніх величин. Види середніх величин (середня арифметична, середня геометрична)

Серед узагальнюючих показників, що характеризують статистичну сукупність, велике значення мають середні величини.

Середні величини – це узагальнююча характеристика безлічі індивідуальних значень деякої кількісної ознаки.

Найважливіша властивість середньої величини полягає в тому, що вона відображає те спільне, що притаманне всім одиницям досліджуваної сукупності.

Визначити середню в багатьох випадках зручніше через вихідне співвідношення середньої або її логічну формулу:

$IBC = \text{сумарне значення або обсяг усередненої ознаки} / \text{число одиниць або обсяг сукупності.}$

Ознака, за якою знаходиться середня, називається осередненою і позначається \bar{X} ;

величина усередненої ознаки в кожній одиниці сукупності називається індивідуальним значенням ознаки або варіантами і позначається $x_1; x_2; x_3; \dots; x_n$;

частота – це повторюваність індивідуальних значень ознаки, яка позначається f ; $w_i = X_i * f_i$ – це обсяг частот.

Формули для середніх величин можуть бути отримані на основі таких середніх:

$$\bar{x} = \frac{\sum_{i=1}^n x_i^k f_i}{\sum_{i=1}^n f_i} \qquad \bar{x} = \frac{\sum_{i=1}^n x_i^k}{n}$$

Зважена середня

де x - ознака, за якою знаходиться середня (осереднена ознака);

k - ступінь середньої;

n - обсяг статистичної сукупності (число одиниць у сукупності);

f – частота або повторюваність індивідуальних значень ознаки.

Проста середня

Величини середніх, розраховані на основі одних і тих же індивідуальних значень, ознаки при різних значеннях ступеня (k), не однакові. Чим вище ступінь середньої, тим більше величина самої середньої.

Питання про вибір середньої вирішується в кожному окремому випадку, виходячи із завдань дослідження, матеріального змісту досліджуваного явища і наявності вихідної інформації. Він складається з декількох етапів:

1. Установлюється визначальний показник, тобто узагальнюючий показник сукупності, від якого залежить величина середньої.

2. Визначається математичний вираз для визначального показника.
3. Проводиться заміна індивідуальних значень середніми показниками.
4. Рішення рівняння середньої.

Види середніх величин:

- середня арифметична (використовується для осереднення прямих значень ознак шляхом їх підсумовування);
- середня геометрична (використовується для осереднення темпів зростання);
- середня квадратична (застосовується в тих випадках, якщо при заміні індивідуальних величин ознаки на середню величину необхідно зберегти незмінною суму квадратів вихідних величин);
- середня гармонійна (або гармонічна) використовується для осереднення обернених індивідуальних значень ознак шляхом їх підсумовування;
- середня хронологічна (використовується для осереднення моментних показників).

5. Середня арифметична та її властивості. Порядкові середні (мода, медіана)

Структурні середні, на відміну від статичних середніх, які значною мірою є абстрактною характеристикою сукупності, виступають як конкретні величини, що збігаються з цілком певними варіантами сукупності. Це робить їх незамінними при вирішенні низки практичних завдань.

У практичних розрахунках використовують декілька видів середніх величин. Використання того чи іншого виду середніх величин залежить від:

- 1) характеру індивідуальних значень ознаки (прямі, обернені, моментні, інтервальні, відносні);
- 2) характеру алгебраїчного зв'язку між індивідуальними значеннями ознаки та її загального обсягу (сума, добуток, ступінь).

Середня арифметична використовується для осереднення прямих значень ознак шляхом їх підсумовування (формула 1).

$$\bar{x} = \frac{\text{Обсяг}_\text{ознаки}}{\text{Чисельність}_\text{сукупності}}$$

Для не згрупованих даних використовується *середня арифметична проста*, яка обчислюється діленням загального обсягу значень ознаки на обсяг сукупності (формула 2):

$$X_{\text{сер}} = (x_1 + x_2 + \dots + x_n) / n = \sum x / n \quad \text{Формула 2}$$

Наприклад, статутний капітал акціонерної компанії сформований 6 засновниками. Розмір внеску кожного з них відповідно становив 8; 10; 12; 9; 6; 5 млн.грн. Середній внесок одного засновника розраховується так:

$$X_{\text{сер}} = \text{Сума внесків} / \text{Число засновників} = (8 + 10 + 12 + 9 + 6 + 5) / 6 = 50 / 6 \approx 8,3 \text{ млн.грн.}$$

Зважена середня арифметична використовується у тих випадках, коли значення ознаки подано у вигляді варіаційного ряду, в якому чисельність одиниць у варіантах неоднакова. Формула середньої арифметичної зваженої має вигляд (формула 4):

$$X_{\text{сер}} = (x_1 f_1 + x_2 f_2 + \dots + x_n f_n) / f_1 + f_2 + \dots + f_n = \sum x f / \sum f$$

З формули видно, що середня зважена принципово не відрізняється від середньої простої арифметичної. Тут додавання f разів варіанти x змінюється множенням її на кількість повторень (f).

Середня гармонічна — це обернена до середньої арифметичної із обернених значень ознак. Її обчислюють, коли необхідно осереднення обернених індивідуальних значень ознак шляхом їх підсумування (наприклад, у випадках визначення середніх витрат часу, праці, матеріалів на одиницю продукції тощо). У випадку розрахунку середньої гармонічної зваженої її обчислюють тоді, коли відомі дані про загальний обсяг ознаки ($z = xf$), а також

індивідуальні значення ознаки (x), невідома є частота (f). Формули середньої гармонічної - простої і зваженої — мають такий вигляд:

- для простої: $X_{сер} = n / \sum 1/x$ - для зваженої: $X_{сер} = \sum z / \sum z/x$

Середня квадратична використовується для визначення показників варіації (коливання) ознаки – дисперсії та середнього квадратичного відхилення. Обчислюється на основі квадратів відхилень індивідуальних значень ознаки від їх середньої величини. Формула середньої квадратичної має такий вигляд (формула 8):

- проста: $X_{сер} = \sqrt{x^2 / n}$ - зважена: $X_{сер} = \sqrt{x^2 f / \sum f}$

Середню геометричну застосовують у тих випадках, коли обсяг сукупності формується не сумою, а добутком індивідуальних значень ознак. Цей вид середньої використовується здебільшого для обчислення середніх коефіцієнтів (темтів) зростання в рядах динаміки. Так, у випадку однакових часових інтервалів між рівнями динамічного ряду середня геометрична проста має такий вигляд (формула 9):

$$k_{сер} = \sqrt[n]{k_1 * k_2 * \dots * k_n}$$

де $k = y_i / y_{i-1}$ - темпи зростання; y_i, y_{i-1} - відповідно розглядаємий та попередній рівні ряду; n — кількість інтервалів

Поряд з розглянутими середніми величинами в якості статистичних характеристик варіаційних рядів розподілу розраховуються структурні середні – *мода і медіана*.

Середніми величинами в статистичних рядах розподілу є мода і медіана, які відносяться до класу структурних {позиційних} середніх. Їх величини залежать лише від характеру частот, тобто від структури розподілу. На відміну від інших середніх, які залежать від усіх значень ознаки, мода і медіана не залежить від крайніх значень. Це особливо важливо для незакритих крайніх інтервалів варіаційних рядів розподілу.

Мода (Mo) — це значення варіанти, що найчастіше повторюється в ряду розподілу. Спосіб обчислення моди залежить від виду статистичного ряду. Для атрибутивних і дискретних рядів розподілу моду визначають візуально без будь-яких розрахунків за значенням варіанти з найбільшою частотою (часткою).

Медіаною (Me) називають варіанту, що ділить ранжирований (впорядкований за мірою зростання або зменшення) ряд на дві рівні за обсягом частини. Медіана для дискретного ряду з непарним числом варіант буде відповідати середній варіанті:

$$Me = xm-1,$$

де m - номер кратної варіанти першої половини ранжированного ряду.

Крім того, в аналізі закономірностей розподілу використовуються також такі характеристики, як кватилі та децилі. *Кватилі* — це варіанти, які поділяють обсяги сукупності на чотири рівні частини, *децилі* — на десять частин.

Контрольні питання

1. Поняття та види статистичних показників.
2. Абсолютні статистичні показники.
3. Поняття та види відносних величин.
4. Одиниці вимірювання статистичних показників.
5. Суть та умови використання середніх величин.
6. Види середніх величин (середня арифметична, середня геометрична, середня квадратична, середня гармонійна, середня хронологічна).
7. Середня арифметична та її властивості.
8. Дати визначення порядкових середніх: мода та медіана.

ТЕМА 4. СТАТИСТИЧНЕ ВИВЧЕННЯ ВАРІАЦІЇ ТА ФОРМИ РОЗПОДІЛУ

1. Суть варіації, її роль у вивченні економічних явищ. Характеристики варіацій. Значення вивчення варіації

Середня величина дає узагальнюючу характеристику всієї сукупності досліджуваного явища. Однак тільки ґрунтуючись на середній величині не можна сказати нічого про те, як окремі значення досліджуваної ознаки групуються навколо середньої. Так, у деяких випадках окремі значення дуже близькі до середньої, у цьому випадку значення середньої добре характеризують сукупність. В інших випадках окремі значення далекі від середньої (викиди) і тоді середня величина не відобразить всієї сукупності. Тому побудову середньої величини необхідно доповнювати вивченням показників варіації.

Варіація ознаки - це розходження індивідуальних значень ознаки всередині досліджуваної сукупності, тобто коливання, різноманіття, змінність величини ознаки в окремих одиниць сукупності.

Варіація виникає в результаті того, що індивідуальні значення складаються під впливом різноманітних факторів.

Показники варіації та способи їх розрахунку

У практичному аналізі оцінка розсіювання (коливання) значень ознаки може виявитися не менш важливою, ніж визначення середньої величини.

Оцінка розсіювання легко визначається за даними варіаційного ряду і характеризує кордон варіації досліджуваної ознаки.

$$R = X_{max} - X_{min}$$

де x_{max} і x_{min} - Найбільше і найменше значення варіюючої ознаки.

Варіація показує на скільки велике розходження між одиницями сукупності, власне, які мають найменше і найбільше значення ознаки. Показник заснований на крайніх значеннях варіюючої ознаки і не відображає відхилень усіх варіантів в ряду. Цей показник становить інтерес у тих випадках, коли важливо знати, яка амплітуда коливань значень ознаки, наприклад, якими є коливання ціни на даний товар протягом тижня або по різних регіонах в даний відрізок часу.

Однак цей показник не дає уявлення про характер варіаційного ряду, розташуванні варіантів навколо середньої величини і може сильно змінюватися, якщо додати або виключити крайні варіанти (коли ці значення аномальні для даної сукупності). У цих випадках розмах варіації дає викривлену амплітуду коливання проти нормальних її розмірів. Тому слід очистити сукупність від аномальних спостережень, перш ніж визначати розмах варіації.

Для оцінки коливальності значень ознаки щодо середньої використовуються характеристики розсіювання. Вони розрізняються за обраною формою середньої величини і способами оцінки відхилень від неї окремих варіантів. До таких показників відносяться:

середнє лінійне відхилення; дисперсія; середньоквадратичне відхилення; коефіцієнт варіації.

Середнє лінійне відхилення – середня арифметична з абсолютних значень відхилень окремих варіантів від їх середньої величини:

| | |
|---|--------------------------|
| $L = \frac{\sum x_i - \bar{x} }{n}$ | для НЕ згрупованих даних |
| $L = \frac{\sum x_i - \bar{x} f_i}{\sum f_i}$ | для згрупованих даних |

де x_i - значення ознаки в дискретному ряді або середина інтервалу в інтервальному розподілі; f_i - частота ознаки.

Середнє лінійне відхилення виражено в тих же одиницях виміру, що і варіанти або їх середня. Воно дає абсолютну міру варіації. Показує, на яку величину відхиляється ознака у досліджуваній сукупності від середньої величини.

Щоб уникнути рівності нулю, суми відхилень від середньої, використовують або абсолютні значення відхилень, або їх парні ступені, наприклад, квадрати. В останньому випадку міра варіації називається *дисперсією* і позначається D або σ^2 :

Дисперсія являє собою середній квадрат відхилень індивідуальних значень ознаки від їх середніх величин. Залежно від вихідних даних обчислюється за формулами:

| | |
|---|--------------------------|
| $\sigma^2 = \frac{\sum(x_i - \bar{x})^2}{n}$ | для НЕ згрупованих даних |
| $\sigma^2 = \frac{\sum(x_i - \bar{x})^2 f_i}{\sum f_i}$ | для згрупованих даних |

Розрахунок дисперсії може бути спрощений і представлений такою формулою.

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

Унаслідок підсумовування квадратів відхилень, дисперсія дає видозмінене уявлення про відхилення, вимірюючи їх у квадратних одиницях. Тому на основі дисперсії вводяться ще дві характеристики: *середнє квадратичне відхилення та коефіцієнт варіації*.

Середнє квадратичне відхилення вимірюється в тих же одиницях, що і варіююча ознака, і обчислюється шляхом вилучення квадратного кореня з дисперсії:

| | |
|--|--------------------------|
| $\sigma = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n}}$ | для НЕ згрупованих даних |
| $\sigma = \sqrt{\frac{\sum(x_i - \bar{x})^2 f_i}{\sum f_i}}$ | для згрупованих даних |

σ - Середнє квадратичне відхилення, як і середнє лінійне відхилення, показує, на скільки в середньому відхиляються конкретні варіанти ознаки від його середнього значення. Це величина, що часто використовується в якості одиниці вимірювання відхилень від середньої арифметичної. Відхилення вважається **нормованим**.

Коефіцієнт варіації характеризує міру варіації значень ознаки навколо середньої величини. Дає відносну оцінку варіації і виводиться шляхом зіставлення середнього лінійного або середнього квадратичного відхилення з середнім рівнем явища, а результат виражається у відсотках:

$$V = \frac{\sigma}{\bar{x}} \times 100$$

Чим коефіцієнт менше, тим однорідніше сукупність і навпаки, чим більше тим неоднорідна.

Оскільки коефіцієнти варіації дають відносну характеристику однорідності явищ і процесів, вони дозволяють порівнювати ступінь варіації різних ознак.

| | |
|------------------------------|---|
| Лінійний коефіцієнт варіації | $V_L = \frac{L}{\bar{x}} \times 100$ або $V_L = \frac{L}{M_e} \times 100$ |
| Коефіцієнт осциляції | $V_n = \frac{R}{\bar{x}} \times 100$ |

2. Коливання будь-якої ознаки як властивість статистичної сукупності. Основні та другорядні причини коливань

Якщо статистична сукупність розбита на групи за будь-якою ознакою і для цих груп відомі (або можуть бути знайдені) середній рівень і дисперсія, то нерідко при об'єднанні певних груп у сукупність потрібно оцінити варіації показників об'єднаної сукупності на основі показників окремих груп. При цьому необхідно враховувати, що варіація ознаки в цілому за сукупністю залежить як від варіації ознаки всередині кожної групи, так і від варіації групових середніх, тобто від міжгрупової варіації ознаки. Іншими словами, загальну дисперсію $\sigma^2_{\text{общ}}$, що характеризує варіацію ознаки під впливом всіх факторів, можна отримати на основі її складових - міжгрупової і внутрішньогрупової дисперсій (табл. 5).

$$t_{\text{i\u0304i\u0304}}^2 = \frac{\sum_{i=1}^k (x_i - \bar{x}_{\text{i\u0304i\u0304}})^2 n_i}{N}$$

Загальна дисперсія сукупності

Загальна дисперсія відображає варіацію ознаки за рахунок усіх умов (факторів), що діють в даній сукупності.

Варіацію між групами за рахунок ознаки-фактора, покладеної в основу групування, відображає **міжгрупова дисперсія**, яка обчислюється за відхиленнями групових середніх від загальної середньої:

$$\sigma^2 = \frac{\sum_{j=1}^m (\bar{x}_j - \bar{x}_{\text{i\u0304i\u0304}})^2 N_j}{N}$$

Варіацію усередині кожної групи досліджуваної сукупності відображає **приватна групова дисперсія**, яка обчислюється як середній квадрат відхилень значень ознаки x від приватної середньої \bar{x}_j .

$$t_1^2 = \frac{\sum_{i=1}^k (x_i - \bar{x}_1)^2 f_i}{N_1}, \quad t_2^2 = \frac{\sum_{i=1}^k (x_i - \bar{x}_2)^2 s_i}{N_2}, \quad \dots, \quad t_m^2 = \frac{\sum_{i=1}^k (x_i - \bar{x}_m)^2 t_i}{N_m}$$

Так, як досліджувана сукупність розбита на кілька груп, то для всієї сукупності внутрішньогрупову варіацію буде виражати **внутрішньогрупова дисперсія**, яка розраховується як середня арифметична з групових дисперсій:

$$\bar{x}^2 = \frac{\sum_{j=1}^m x_j^2 N_j}{N}$$

де k - число груп; f_j - число одиниць у j -ій групі; \bar{x}_j - Приватна середня по j -ій групі; \bar{x}_0 - Загальна середня по сукупності одиниць.

Таблиця 5

| Вид дисперсії | Формула для розрахунку | Характеристика |
|--|--|--|
| загальна дисперсія | $\sigma^2 = \frac{\sum_{i=1}^k (x_i - \bar{x}_0)^2 f_i}{\sum_{i=1}^k f_i}$ | відображає варіацію ознаки за рахунок усіх умов (факторів), що діють в даній сукупності |
| міжгрупова дисперсія | $\delta_x^2 = \frac{\sum_{j=1}^k (\bar{x}_j - \bar{x}_0)^2 f_j}{\sum_{j=1}^k f_j}$ | відображає варіацію між групами за рахунок ознаки-фактора, покладеного в основу групування |
| середня з внутрішньогрупових дисперсій | $\bar{\sigma}^2 = \frac{\sum_{j=1}^k \sigma_j^2 f_j}{\sum_{j=1}^k f_j}$ | відображає випадкову варіацію, зумовлену неврахованими факторами і не залежить від ознаки-фактора, покладеного в основу групування |

$$\sigma_j^2 = \frac{\sum_{i=1}^{n_j} (x_{ij} - \bar{x}_j)^2}{f_j}$$

Внутрішньогрупова дисперсія

Між представленими видами дисперсій існує певне співвідношення: загальна дисперсія дорівнює сумі дисперсій внутрішньогрупової (середньої з групових дисперсій) і міжгруповий (дисперсії приватних середніх), тобто:

| | | |
|----------------------|------------------------|--|
| Загальна дисперсія = | Міжгрупова дисперсія + | Середня з внутрішньогрупових дисперсій |
| $\sigma^2 =$ | $\delta_x^2 +$ | $\bar{\sigma}_i^2$ |

Ця вищевказана формула відома як **правило складання дисперсій**, його автором є Вільгельм Лексис (1837-1914), німецький статист і економіст.

Правило складання дисперсії дозволяє виявити залежність результату від визначальних чинників за допомогою співвідношення між групової дисперсії і загальної дисперсії. Це співвідношення називається *емпіричним коефіцієнтом детермінації*:

$$\eta^2 = \frac{\delta_x^2}{\sigma^2}$$

Показує, яка частка в загальній дисперсії припадає на дисперсію, обумовлену варіацією ознаки, покладеної в основу групування.

Корінь квадратний з емпіричного коефіцієнту детермінації називається *емпіричним кореляційним відношенням*:

$$\eta = \sqrt{\frac{\delta^2}{\sigma^2}}$$

Відображає вплив ознаки, покладеної в основу групування, на варіацію результативної ознаки. $\eta \in [0; 1]$. При цьому якщо $\eta = 0$, то групувальна ознака не впливає на результат. Якщо $\eta = 1$, то результативна ознака змінюється тільки в залежності від ознаки, покладеної в основу групування, а вплив інших факторів дорівнює нулю.

Для перевірки суттєвості зв'язку міжгрупової ознаки і варіацією досліджуваної ознаки часто використовується дисперсійне відношення або *F-критерій Фішера*:

$$F = \frac{\delta^2}{v_1} / \frac{\sigma^2}{v_2}$$

v_1 і v_2 - число ступенів свободи для порівнюваних дисперсій, при цьому: $v_1 = m - 1$; $v_2 = N - m$ - число груп, N - число спостережень

Отримане значення критерію називають фактичним (розрахунковим), і порівнюють з табличним (критичним) значенням, яке визначається за таблицею в залежності від ступенів свободи. Якщо $F_{\text{факт}} > F_{\text{табл}}$ наявність зв'язку доведено.

КОНТРОЛЬНІ ПИТАННЯ

1. Суть варіації, її роль у вивченні економічних явищ.
2. Характеристики варіацій. Значення вивчення варіації.
3. Показники варіації та способи їх розрахунку
4. Коливання будь-якої ознаки як властивість статистичної сукупності.
5. Основні і другорядні причини коливань.
6. Правило складання дисперсій

ТЕМА 5. ВИБІРКОВИЙ МЕТОД

1. Суть і переваги вибіркового обстеження. Його види

Найбільш широко поширеним видом несуцільного спостереження є вибіркоче, при якому обстежуються не всі одиниці досліджуваної сукупності, а лише певним чином відібрана їх частина.

Отже, **вибірковим спостереженням** є несуцільне спостереження, при якому статистичному обстеженню піддаються одиниці досліджуваної сукупності, відібрані випадковим способом.

Мета вибіркового спостереження полягає в тому, щоб за характеристиками відібраної частини одиниць (*вибірка*) можна було говорити про характеристики всієї сукупності (*генеральної*).

Основні **причини**, з яких у багатьох випадках вибіркового спостереження віддається йому **перевага**, порівняно із суцільним спостереженням, такі:

(Переваги вибіркового спостереження)

1. Досягнення більшої точності результатів обстеження завдяки скороченню помилок реєстрації (за рахунок роботи більш кваліфікованих учасників)
2. Економія трудових і грошових коштів і часу в результаті скорочення обсягу роботи
3. Можливість детального обстеження кожної одиниці спостереження за рахунок розширення програми спостереження
4. Зведення до мінімуму знищення та приведення в непридатність обстежуваних одиниць сукупності
5. Уточнення результатів суцільного спостереження

Уся сукупність одиниць, з якої здійснюється відбір, називається генеральною сукупністю, а одиниці, відібрані для безпосереднього спостереження, являють собою вибіркочу сукупність, або просто вибірку.

Відбір з генеральної сукупності проводиться таким чином, щоб на основі вибірки можна було отримати досить точне уявлення про основні параметри сукупності загалом. При цьому мова йде, як про точкову оцінку, в якості якої приймається відповідне значення середньої, тобто частки отриманої в результаті вибірки, так і про інтервальну оцінку.

У статистиці прийнято розрізняти **параметри** та **властивості** генеральної сукупності та їх оцінки за вибіркочими даними. З цією метою прийнята така система позначень:

генеральні параметри позначаються грецькими літерами, вибіркочі показники, які розглядаються як оцінки генеральних параметрів - латинськими літерами:

| Показники | Генеральні параметри | Вибіркові показники |
|--------------------------------------|--|--|
| Об'єм сукупності (число одиниць) | N | n |
| Середнє значення ознаки | $\bar{x} = \frac{\sum x_i}{N}$ | $\tilde{x} = \frac{\sum x_i}{n}$ |
| Частка одиниць, що володіють ознакою | $p = M / N$ M - число одиниць, що володіють певним ознакою в ГС | $w = m / n$ m - число одиниць, що володіють певною ознакою в НД |
| Частка одиниць, що не володіють | q | $1-w$ |

| | | |
|-----------------------------|--|--|
| ознакою | | |
| Дисперсія кількісної ознаки | $\sigma_{\bar{x}}^2 = \frac{\sum(x_i - \bar{x})^2}{N}$ | $\sigma_{\hat{x}}^2 = \frac{\sum(x_i - \hat{x})^2}{n}$ |
| Дисперсія частки | $\sigma_p^2 = p(1 - p)$ | $\sigma_w^2 = w(1 - w)$ |

У зв'язку з вищесказаним можна виділити такі **завдання вибіркового** спостереження:

- 1) визначення меж генеральної сукупності;
- 2) розробка програми спостереження й інструкцій;
- 3) визначення основи для проведення вибірки - списку одиниць генеральної сукупності, відомостей про їх розміщення тощо;
- 4) встановлення допустимого розміру похибки і визначення обсягу вибірки;
- 5) визначення виду вибіркового спостереження;
- 6) встановлення термінів проведення спостереження;
- 7) визначення потреби в кадрах для проведення вибіркового спостереження та їх підготовка;
- 8) оцінка точності та достовірності даних вибірки, визначення порядку їх розповсюдження на генеральну сукупність.

2. Схеми й етапи відбору

Основні етапи вибіркового спостереження:

- 1) визначення мети, завдань і складання програми спостереження;
- 2) аналіз інформаційних джерел, що використовуються для виділення генеральної сукупності об'єктів спостереження (основи вибірки);
- 3) формування генеральної сукупності для проведення вибіркового спостереження;
- 4) розробка методології формування вибіркової сукупності, що включає вибір способу відбору, визначення необхідного обсягу вибірки, етапів відбору одиниць з генеральної сукупності, планування і проведення пробної вибірки;
- 5) формування вибірки;
- 6) збір даних на основі розробленої програми;
- 7) аналіз отриманих результатів і розрахунок основних характеристик вибіркової сукупності;
- 8) розрахунок помилок вибірки і поширення її результатів на генеральну сукупність.

Методи та види відбору одиниць з генеральної сукупності

У теорії вибіркового спостереження розроблені різні методи і види відбору одиниць з генеральної сукупності повторний і неповторний.

Так, як неповторний відбір охоплює все нові й нові сукупності, а повторний відбір на всьому протязі одну і ту ж сукупність, неповторний відбір дає більш точні результати, ніж повторний.

За способом відбору (способом формування) вибірки одиниць з генеральної сукупності поширені такі види вибіркового спостереження:

- 1) проста випадкова вибірка (власне-випадкова);
- 2) типова (стратифікована);
- 3) серійна (гніздова);
- 4) механічна;
- 5) комбінована;

Проста випадкова вибірка (власне-випадкова) є відбір одиниць з генеральної сукупності шляхом випадкового відбору, але за умови ймовірності вибору будь-якої одиниці з генеральної сукупності. Відбір проводиться методом жеребкування або за таблицею випадкових чисел.

Типова (стратифікована) вибірка передбачає поділ неоднорідної генеральної сукупності на типологічні або районовані групи з якого-небудь суттєвою ознакою, після чого з кожної групи проводиться випадковий відбір одиниць.

Для **серійної** (гніздовий) вибірки характерно те, що генеральна сукупність спочатку розбивається на певні рівновеликі або нерівновеликі серії (одиниці всередині серій пов'язані за певною ознакою), з яких шляхом випадкового відбору відбираються серії і потім всередині відібраних серій проводиться суцільне спостереження.

Механічна вибірка представляє собою відбір одиниць через рівні проміжки (за алфавітом, через часові проміжки, за просторовим критерієм тощо). При проведенні механічного відбору генеральна сукупність розбивається на рівні за чисельністю групи, з яких потім відбирається по одній одиниці.

Комбінована вибірка заснована на поєднанні декількох способів вибірки.

3. Визначення середніх і граничних помилок при різних схемах відбору. Визначення обсягу вибірки. Способи поширення вибіркових даних на генеральну сукупність

При проведенні вибіркового спостереження не можна навіть теоретично отримати абсолютно точні дані, як при суцільному обстеженні. Обумовлено це тим, що спостереженню піддається не вся сукупність, а тільки її частину, тому при проведенні вибіркового спостереження неминуча деяка властива йому *похибка* (помилки). Прийнято розрізняти *середню і граничну помилки вибірки*.

Помилка репрезентативності - розбіжність між вибірковою характеристикою й характеристикою генеральної сукупності.

Теоретичним обґрунтуванням появи випадкових помилок вибірки є теорія ймовірностей та її граничні теореми. Сутність граничних теорем полягає в тому, що в масових явищах сукупний вплив випадкових причин на формування закономірностей і узагальнюючих характеристик буде малою величиною або практично не залежатиме від випадку.

Так, як випадкова помилка вибірки виникає в результаті випадкових відмінностей між кордонами вибіркової і генеральної сукупностей, при досить великому обсязі вибірки помилка буде мала. Цей висновок, що спирається на докази граничних теорем, дозволяє припускати, що характеристики вибіркового спостереження можуть включати характеристики генеральної сукупності.

Випадкові помилки можуть бути доведені до незначних розмірів, а головне, їх розміри та межі можна визначити з достатньою точністю на підставі *закона великих чисел*.

Середня помилка вибірки – це така розбіжність між середніми вибіркової і генеральної сукупностями, яке не перевищує σ_{\pm}

У математичній статистиці доводиться, що значення середньої помилки вибірки визначаються за такими формулами:

Формула для визначення величини середньої помилки вибірки для кількісної ознаки:

$$\mu_{\bar{x}} = \sqrt{\frac{\sigma_{\bar{x}}^2}{n}} = \frac{\sigma_{\bar{x}}}{\sqrt{n}}$$

| | | | | | | | | | | |
|--------------|-----|--------|--------|--------|--------|--------|--------|--------|--------|---------|
| t | 0 | 0,1 | 0,5 | 1 | 1,5 | 2 | 2,5 | 2,6 | 3 | 4 |
| $\varphi(t)$ | 0,1 | 0,0797 | 0,3829 | 0,6827 | 0,8664 | 0,9545 | 0,9876 | 0,9907 | 0,9973 | 0,99994 |

З вищесказаного випливає, що лише з певним ступенем ймовірності можна стверджувати, що показники генеральної сукупності та їх відхилення не перевищать величину $t\mu_{\bar{x}}$.

Отримана величина $t\mu_{\bar{x}}$ називається **граничною помилкою вибірки**.

Гранична помилка вибірки – це максимально можлива розбіжність вибіркової і генеральної *середніх*, тобто максимум помилки при заданій ймовірності її появи.

Гранична помилка вибірки для кількісної ознаки:

$$\Delta_{\bar{x}} = t\mu_{\bar{x}}$$

Гранична помилка вибірки для альтернативної ознаки:

$$\Delta_w = t\mu_w$$

У зв'язку з тим, що існують різні методи, види і способи відбору одиниць з генеральної сукупності формули для розрахунку середньої помилки вибірки також будуть різнитися.

Визначення необхідної (оптимальною) чисельності вибірки

При розробці програми вибіркового обстеження одним з найбільш складних є питання про те, скільки одиниць досліджуваної сукупності необхідно обстежити, тобто **обсяг вибірки**.

При цьому слід мати на увазі, що при будь-якому способі відбору гранична помилка вибірки обернено пропорційна кількості обстежених одиниць. Тобто середня помилка

вибірки $\mu_{\bar{x}} = \sqrt{\frac{\sigma_{\bar{x}}^2}{n}}$ пропорційна \sqrt{n} , при збільшенні чисельності вибірки в 4 рази, помилка зменшиться вдвічі. Збільшуючи n можна звести помилку до *min*. При $n \rightarrow N$, $\mu_{\bar{x}} \rightarrow 0$.

Так, як при проведенні вибіркового спостереження визначення характеристик вибірки в ряді випадків супроводжується руйнуванням обстежуваних зразків, то норми відбору повинні бути мінімальні, так само не слід забувати про основну перевагу несущільного спостереження (мінімум витрат і часу).

Підвищення відсотка вибірки веде до збільшення обсягу досліджуваної роботи. Якщо ж для вибірки взята недостатня кількість проб, то результати дослідження будуть мати великі похибки. Усе це необхідно враховувати при організації вибіркового обстеження.

Поширення результатів вибіркового спостереження на генеральну сукупність

На заключному етапі вибіркового обстеження вирішується питання про можливість поширення отриманих результатів на генеральну сукупність. При цьому враховуються дві основні обставини:

1. Наскільки адекватно представлена генеральна сукупність у вибірці, тобто чи не змінилася в результаті обстеження структура запланованої її основи, чи дотримані основні пропорції між типовими групами у вибірковій і генеральній сукупності.

Для відновлення вихідних пропорцій генеральної сукупності проводиться коригування вибірки або шляхом відсікання частини одиниць, частка яких у вибірці непропорційно велика порівняно з часткою у генеральній сукупності, або шляхом багаторазового використання результатів спостереження за одиницями тих груп, які недостатньо широко представлені у вибірці.

2. Який ступінь відповідності фактично отриманої відносної помилки вибірки запланованому її рівню. Фактичне значення відносної помилки визначається шляхом зіставлення абсолютної величини граничної помилки вибірки, отриманої в результаті обстеження, з середнім рівнем ознаки, розрахованим на основі вибірки.

Якщо вибірка адекватна генеральній сукупності і фактична відносна помилка вибірки незначним чином відрізняється від запланованого її рівня, то на основі проведеного обстеження можна оцінити межі, в яких знаходиться середнє значення досліджуваної ознаки (або частки) у генеральній сукупності, а також можна вказати його можливе значення для сукупності в цілому.

Загальне значення досліджуваного показника для сукупності в цілому визначається двома способами: **методом прямого рахунку і методом коефіцієнтів.**

Якщо в результаті обстеження отримані верхня і нижня межі досліджуваної ознаки в розрахунку на одиницю сукупності, тобто знайдені величини $\check{x} - \Delta_{\check{x}} \leq \bar{x} \leq \check{x} + \Delta_{\check{x}}$, то з відповідною ймовірністю можна знайти ці кордони для сукупності в цілому.

1. Суть і переваги вибіркового обстеження.
2. Види вибіркового обстеження.
3. Схеми й етапи відбору.
4. Що є метою вибіркового спостереження
5. Визначення середніх і граничних помилок при різних схемах відбору.
6. Що є «Тенденційні похибки»
7. Вибіркове спостереження
8. Визначення обсягу вибірки.
9. Способи поширення вибіркових даних на генеральну сукупність.

ТЕМА 6. СТАТИСТИЧНІ МЕТОДИ ВИМІРЮВАННЯ ВЗАЄМОЗВ'ЯЗКІВ

1. Види та форми взаємозв'язку між явищами

Усі явища навколишнього світу взаємопов'язані та взаємообумовлені, між ними існує причинно-наслідковий зв'язок.

Мета вимірювання взаємозв'язків – це виявлення причинних зв'язків і надання їм кількісної характеристики.

Ознаки, що характеризують причини й умови зв'язку, називають факторними (x), а ті, що характеризують наслідки зв'язку – результативними (y).

Типи взаємозв'язків:

1. *Функціональний* (при ньому кожному значенню фактора x відповідає одне або кілька чітко визначених значень y).

2. *Стохастичний* (при ньому кожному значенню фактора x відповідає певна ножина значень ознаки y , які утворюють умовний розподіл).

3. *Кореляційний* (різновид стохастичного зв'язку, при якому кожному значенню фактора x відповідає середнє значення результативної ознаки y_1).

Варто відзначити, що розвиток статистичної сукупності проявляється не тільки в кількісному зростанні або зменшенні елементів цієї системи, але й у зміні її структури.

Поняття структури дуже тісно переплітається з поняттям групування і класифікації.

Структура – це будова, форма організації системи, що складається з окремих елементів і зв'язків між ними.

Ієрархічною (деревовидною) структурою називається складна структура, утворена при послідовному дробленні системи на більш однорідні групи елементів. Вона складається з декількох рівнів («кроків» дроблення).

Основна перевага ієрархічної структури полягає у її великій інформаційній місткості (клас, підклас, група, підгрупа, вид, (різновиди), традиційність і звичність застосування, пристосованість до обробки інформації, а також можливість для створення при кодуванні об'єктів класифікації кодів, що несуть смислове навантаження).

Недоліком є слабка гнучкість її структури і заздалегідь установлений порядок ступенів розподілу, який не допускає включення при відсутності резервної ємності нових об'єктів класифікаційних угруповань і ознак. Унаслідок цього зміна хоча б однієї ознаки веде до перерозподілу багатьох класифікаційних угруповань.

Ієрархічна структура характеризується не тільки частками обсягу ознаки, але і додатковими показниками, а саме:

1. Характеристикою ступеня складності структури, тобто числом рівнів дроблення («порядок» структури).

2. Середнім порядком структури, тобто середнім номером рівня, зважених за часткам обсягу ознаки, дроблення яких завершилося на даному рівні. Ця величина характеризує середнє число дроблень обсягу ознаки.

3. Загальним числом кінцевих (тобто не дробящихся далі) гілок структури.

4. Середнім числом кінцевих гілок, що припадають на один рівень.

Баланс (фр. balance - буквально ваги, рівновага) - це особлива форма зіставлення структури однієї і тієї ж величини ознаки, що характеризується з двох різних сторін або у двох різних аспектах.

У найбільш загальній формі динамічний баланс складається з чотирьох складових: запас на початок періоду, дохід за період, витрати (втрати) за період, запас на кінець періоду.

$ДБ = \text{запас на початок періоду} + \text{дохід витрати} + \text{запас на кінець періоду}.$

Для аналітичних цілей кожна з чотирьох складових ділиться за різними класифікаційними ознаками на частини, групи або підгрупи.

Якщо загальний обсяг ознаки поділений за однією з групувальних ознак, а потім кожен

груповий і загальний обсяги знову поділені за іншою групувальною ознакою, то утворюється багатовимірна - двомірна структура з пересічними ознаками.

Двомірна структура дозволяє розрахувати п'ять видів структурних показників (часток). При трьох пересічних ознаках групування число різних видів структур досягає 19. У загальному вигляді при n взаємопроникаючих ознаках структура містить n -видів часток.

Для мети аналізу структури, а також порівняння двох (або більше) структур в динаміці використовується велика кількість статистичних методів і підходів які можна представити у вигляді такої схеми:

2. Правила, методи та прийоми вивчення взаємозв'язку

Для обґрунтованого прийняття управлінських рішень важливо мати об'єктивну статистику інформацію про взаємодію між різноманітними соціально-економічними явищами. Аналіз взаємозв'язків надає статданим системності, що у свою чергу дозволяє значно підвищити їх корисність і результативність.

Отже, у процесі статдослідження взаємозв'язків між соціально-економічними явищами та процесами потрібно відповісти на такі питання:

1. Чи існує взаємозв'язок між цими явищами взагалі і якщо так, то наскільки він суттєвий щодо його врахування при прийнятті управлінських рішень.
2. Якісна оцінка зв'язку, тобто чи це позитивна, чи негативна залежність між явищами, які аналізуються. Якщо явище має велику кількість позитивних впливів, то вон повинно стати об'єктом підвищеної уваги з боку державних органів (*ефект управлінського леверіджу*). Показовим прикладом на макрорівні державного управління є економічна сфера, що має позитивний вплив на ключові соціальні, політичні, культурні та інші чинники.
3. Визачення аналітичної (кількісної) залежності між явищами, які аналізуються. Це дозволяє прогнозувати розвиток об'єкта держ.управління і резервує час для реалізації дій з усунення негативних явищ у суспільстві.

Методи та прийоми вивчення взаємозв'язку детальніше розглянуті нижче.

3. Основи кореляційно-регресійного аналізу

Вивчення залежності варіації ознаки від навколишніх умов становить зміст теорії кореляційно-регресійного аналізу.

Основоположниками даної теорії є англійські біометрики Френсіс Гальтон і Карл Пірсон. У Росії їх ідеї отримали розвиток в працях А.А. Чупрова.

Основні ідеї теорії кореляції вперше висловив у своїх лекціях **Френсіс Гальтон** (1822-1911) в кінці 70-х років дев'ятнадцятого століття. Досліджуючи закономірності спадковості, він встановив, що кількісні ознаки батьків у нащадків зм'якшувалися, «регресували до середнім величинам за сукупністю», до «середнього стану». Цей зв'язок названа їм регресією. Обраний Ф. Гальтон термін закріпився за рівнянням, що дозволяє за величиною одного кореляційно пов'язаного ознаки обчислити середні величини іншої ознаки.

Карл Пірсон (1857-1936) став користуватися коефіцієнтом кореляції як вимірником щільності зв'язку. Спочатку ним була досліджена прямолінійна залежність, яка пов'язана із законом нормального розподілу, потім з'явилася необхідність дослідити та нелінійну залежність. Для вимірювання сили зв'язку нелінійних залежностей К. Пірсон запропонував кореляційне відношення. Після розробки методів аналізу взаємозв'язку двох змінних була запропонована теорія приватних і чистих коефіцієнтів кореляції, а також теорія множинної (сукупної, багатофакторної) кореляції.

В економічних дослідженнях використовується велика кількість різноманітних змінних, які можна розділити на такі групи:

Результуюча змінна y – це вихідна змінна (або ознака), що характеризує результат (на

підстві якого визначається ефективність) функціонування аналізованої соціально-економічної системи. Її значення формуються в процесі і всередині функціонування цієї системи під впливом ряду інших змінних і факторів, частина з яких піддається реєстрації та, певною мірою, регулювання та планування.

Пояснюючі змінні (незалежні змінні, вхід, факторний ознака) $X = (x_1, x_2, \dots, x_n)$ - це змінні (або ознаки), які піддаються реєстрації, що описують умови функціонування досліджуваної соціально-економічної системи і в істотній мірі визначають процес формування значень результуючої змінної.

Отже, одним із найважливіших понять у статистиці є поняття взаємозв'язків, які можуть бути **кореляційними** (імовірнісними) і **функціональними** (математичними, детермінованими).

Пр цьому завданнями кореляційно-регресійного аналізу є:

1. Виділення важливих факторів, що впливають на результативну ознаку, на базі заходів тісноти зв'язку факторів з результативним ознакою;
2. Опис впливу факторів за допомогою регресійного рівняння;
3. Оцінка статистичної значущості рівняння регресії і його параметрів;
4. Прогнозування (імітація) можливих значень результативної ознаки при задаються значеннях факторних ознак.

Вимірювання щільності зв'язку в разі парної кореляції

Існують кілька елементарних методів виявлення наявності кореляційного зв'язку:

Зіставлення двох паралельних рядів (ряду значень факторної ознаки і відповідних йому значень результативної ознаки). Значення факторної ознаки розташовують у зростаючому порядку і потім простежують напрямок зміни величини результативної ознаки.

Якщо із зростанням величини факторної ознаки спостерігається зростання результативної ознаки, то можна припустити прямий кореляційний зв'язок і навпаки (якщо x зростає, а y знижується), то зв'язок є зворотним.

Побудова аналітичного групування, де всі спостереження розбиваються на групи за величиною факторної ознаки, і по кожній групі обчислюється середнє значення результативної ознаки.

Передбачається, що всі інші причини носять випадковий характер, при визначенні середньої за групами взаємопогашаються, тобто дають у кожній групі один і той же результат. Недоліком цього методу є неоднозначність результатів, які залежать як від числа виділених груп, так і від установлення меж інтервалів.

Графічний метод. Для попереднього виявлення наявності зв'язку та розкриття її характеру застосовують графічний метод, при цьому використовуються дані про індивідуальні значення факторної ознаки і відповідних йому значень результативної ознаки. Побудований даним методом малюнок називається **полем кореляції**.

Найбільш поширеним методом виявлення і вимірювання кореляційного зв'язку є *розрахунок коефіцієнтів кореляції*.

У статистиці розрізняють такі варіанти залежностей:

1. **парна кореляція** - зв'язок між двома ознаками, один з яких результативний, а інший факторний;
2. **приватна кореляція** - залежність між результативною і одною факторною ознакою, при фіксованому значенні інших факторних ознак;
3. **множинна кореляція** - залежність результативної ознаки від декількох факторних ознак.
4. **канонічна кореляція** - залежність групи результативної ознаки від групи факторних ознак.

Змінні - Y і X можуть бути виміряні в різних шкалах, саме це визначає вибір відповідного

коефіцієнта кореляції. Співвідношення між типами шкал, в яких можуть бути виміряні аналізовані змінні відповідними заходами зв'язку, може бути представлено у вигляді таблиць.

Виділяють чотири типи вимірювальних шкал (або способів вимірювання)

1. **Номінативна**, номінальна або шкала **найменувань**. Вимірювання за даною шкалою полягає в привласненні якої-небудь властивості або ознаки певного позначення або символу (чисельного, літерного і т.п.), тобто при вимірюванні за цією шкалою здійснюється класифікація або розподіл об'єктів на непересічні класи, групи.
2. **Порядкова**, ординарна або рангова шкала. Зміна за цією шкалою поділяє всю сукупність виміряних ознак на такі множини, які пов'язані між собою відносинами типу «більше - менше», «вище - нижче», «сильніше - слабше». Тобто в порядковій шкалі всі ознаки розташовуються за рангом - від найбільшої до самої маленької або навпаки.
3. **Інтервальна** або шкала рівних інтервалів. За даною шкалою кожне з можливих значень виміряних величин стоїть від найближчого на рівній відстані. Головне поняття цієї шкали - інтервал, який можна визначити як частку або частину вимірюваної властивості між двома сусідніми позиціями на шкалі.
4. Шкала **рівних відносин** або шкала відносин. Особливістю цієї шкали є наявність твердо фіксованого нуля, який означає повну відсутність якої-небудь властивості чи ознаки.

4. Множинна регресія. Дисперсійний аналіз

Показники вимірювання парної лінійної кореляції

Для дослідження ступеня щільності зв'язку між якісними ознаками, кожен з яких представлений у вигляді альтернативних ознак, може бути використаний **коефіцієнт асоціації Д. Юла** або **коефіцієнт контингенції К. Пірсона**.

Коефіцієнт асоціації обчислюється за формулою:

$$K_a = \frac{ad - bc}{ad + bc}$$

Коефіцієнт контингенції:

$$K_k = \frac{ad - bc}{\sqrt{(a+b)(b+d)(a+c)(c+d)}}$$

Коефіцієнт контингенції за значенням завжди менше коефіцієнта асоціації. Зв'язок вважається досить значущою і підтвердженою, якщо $|K_a| > 0,5$ або $|K_k| > 0,3$

Для оцінки щільності зв'язку між альтернативними ознаками, приймаючими будь-яке число варіантів значень, застосовується **коефіцієнт взаємної відмінюваності (російською – спряженности) К. Пірсона і А.А.Чупрова**.

Коефіцієнт взаємної спряженості Пірсона визначається за формулою:

$$K_{\Pi} = \sqrt{\frac{\varphi^2}{\varphi^2 + 1}}$$

До показників ступеня щільності зв'язку відносять **коефіцієнт кореляції знаків**, який був запропонований німецьким вченим Г.Фехнером (1801-1887). Цей показник заснований на оцінці ступеня узгодженості напрямків відхилень індивідуальних значень факторного та результативного ознак від відповідних середніх. Для його розрахунку обчислюють середні значення результативного і факторного ознак, а потім проставляють знаки відхилень для всіх значень взаємопов'язаних пар ознак.

Якщо ввести позначення:

n_a - число збігів знаків відхилень індивідуальних величин від середньої,
 n_b - кількість розбіжностей знаків відхилень, то коефіцієнт Фехнера можна записати таким чином:

$$K_{\Phi} = \frac{n_a - n_b}{n_a + n_b}$$

Коефіцієнт Фехнера може приймати різні значення в межах від -1 до +1. Якщо знаки всіх відхилень співпадуть, то $n_b = 0$ і тоді показник дорівнюватиме 1, що свідчить про можливу наявність прямого зв'язку. Якщо ж знаки всіх відхилень будуть різними, тоді $n_a = 0$ і коефіцієнт Фехнера дорівнюватиме -1, що дає підставу припустити наявність зворотного зв'язку.

Як видно з наведеної формули для розрахунку коефіцієнта Фехнера, величина цього показника не залежить від величини відхилень факторної та результативної ознаки від відповідної середньої величини. Тому не можна говорити про ступінь щільності кореляційного зв'язку, а тим більше про оцінку її суттєвості на підставі тільки коефіцієнта Фехнера. При малому обсязі вихідної інформації коефіцієнт Фехнера практично вирішує ту ж задачу, яка ставиться при побудові групових і кореляційних таблиць, тобто відповідає на питання про наявність і напрямок кореляційного зв'язку між ознаками.

Більш досконалим показником, використовуваним для вимірювання щільності зв'язку як якісних, так і кількісних факторів, за умови, що їх значення можна проранжувати, є **ранговий коефіцієнт кореляції Спірмена**, який має такий вигляд:

$$\rho = 1 - \frac{6 \times \sum_{i=1}^n (R_i^X - R_i^Y)^2}{n(n^2 - 1)}$$

R_i^X, R_i^Y - ранги по результативному і факторному ознакою;

n - обсяг досліджуваної сукупності.

Ранжування - процедура впорядкування об'єктів вивчення, яка виконується на основі переваги.

Ранг - це порядковий номер одиниці сукупності в ранжированому ряді.

1. Коефіцієнт кореляції Спірмена може приймати значення від 0 до 1

Перевага коефіцієнта кореляції рангів полягає в тому, що на його основі оцінюється корельованість якісних ознак, що не мають точного кількісного виміру.

Коефіцієнт лінійної кореляції був запропонований англійським статистиком К. Пірсоном. Його інтерпретація така: відхилення ознаки-фактора від його середнього значення на величину свого середнього квадратичного відхилення в середньому по сукупності призводить до відхилення ознаки-результату від свого значення на r його середнього квадратичного відхилення.

Коефіцієнт кореляції є абстрактним показником, що характеризує тісноту зв'язку між змінними, якщо цей зв'язок лінійна. Однією з формул розрахунку показника є наступна:

$$r_{yx} = \frac{\overline{yx} - \bar{y}\bar{x}}{\sqrt{y^2 - (\bar{y})^2} \sqrt{x^2 - (\bar{x})^2}}$$

Коефіцієнт кореляції має такі властивості:

Приймає значення на відрізку [-1; 1]

0 - зв'язок між x і y відсутня;

(0-0,3] - зв'язок присутній але вона незначна;

(0,3-0,5] - помірний зв'язок;

(0,5-0,7] - середня зв'язок;

(0,7-0,99] - тісний зв'язок;

1 - зв'язок між x і y функціональна.

Наступний коефіцієнт - **коефіцієнт детермінації**, що дорівнює квадрату коефіцієнта кореляції r^2 , Виражений у відсотках і показує, який відсоток варіації результату ознаки пояснюється варіацією факторного ознаки.

Показники вимірювання множинної лінійної кореляції

Найважливішим показником інтенсивності зв'язку в багатофакторній системі є **множинний коефіцієнт детермінації**, що позначається латинською літерою R^2 .

Він вимірює загальну щільність зв'язку варіації результативної ознаки Y з варіацією всієї системи входять у модель факторів. Іншими словами, він характеризує питому вагу (відсоток) загальної дисперсії, що пояснюється рівнянням регресії (або зміною факторного показника x). Чим більше ця питома вага, тим більшою мірою варіація у пояснюється змінами комплексу змінних x , і, отже, зв'язок між ними є більш інтенсивним.

Значення коефіцієнта детермінації змінюється від 0 до 1 і ніяк не вказує на напрям зв'язку.

Величина множинного коефіцієнта детермінації може бути обчислена кількома способами, а саме:

- а) на основі матриці парних коефіцієнтів кореляції.
- б) на основі парних коефіцієнтів кореляції і бета коефіцієнтів:
- в) як кореляційне відношення, тобто відношення загальної варіації (загальна сума квадратів) до варіації залишків (суми квадратів залишків):.

$$R^2 = \frac{\sum(y_i - \bar{y})^2}{\sum(y_i - \hat{y}_i)^2} = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2}$$

\hat{y}_i - Вирівняні значення за рівнянням регресії;

y_i - індивідуальне значення результативного ознаки;

\bar{y} - Середнє значення результативної ознаки.

Математичні властивості множинного коефіцієнта детермінації такі, що величина такого коефіцієнта автоматично наближається до одиниці при числі факторів, що наближається до $n - 1$. При цьому неважливо, чи введені в модель реальні фактори, або будь-які варіюючі величини, що не відносяться до досліджуваної проблеми. Щоб виключити цю проблему з величини коефіцієнта детермінації, проводиться його коректування на втрату ступенів свободи, одержуваний показник називається скоригований коефіцієнт множинної детермінації:

$$\bar{R}^2 = 1 - (1 - R^2) \frac{n - 1}{n - m - 1}$$

На практиці показник використовується для виявлення найкращої моделі з числа можливих.

Показники вимірювання приватної лінійної кореляції

Окремі коефіцієнти кореляції характеризують щільність зв'язку між результатом і відповідним фактором при усуненні впливу інших факторів, включених у рівняння регресії.

Рівняння регресії У статистиці виділяють різні види регресійних моделей. парна, множинна, лінійна, не лінійна, позитивна (пряма), негативна (обратна), непосредствена.

Контрольні питання

1. Види та форми взаємозв'язку між явищами.
2. Правила, методи та прийоми вивчення взаємозв'язку.
3. Основи кореляційно-регресійного аналізу.
4. **Завдання кореляційно-регресійного аналізу**
5. Функціональний та стохастичний зв'язок
6. Коефіцієнт детермінації
7. Множинна регресія.
8. Дисперсійний аналіз.

ТЕМА 7. ІНДЕКСИ

1. Суть індексів та їх роль у економіко-статистичному аналізі. Методологічні принципи побудови індексів.

Важливе значення в статистичних дослідженнях має індексний метод. Отримані на основі цього методу показники використовуються для характеристики розвитку аналізованих показників у часі, на території, вивчення структури і взаємозв'язків, виявлення ролі факторів у зміні складних явищ.

У статистиці під **індексом** розуміється відносний показник, який виражає співвідношення величин якого-небудь явища в часі, в просторі або дає порівняння фактичних даних з будь-яким еталоном (план, прогноз, норматив тощо).

У міжнародній практиці індекси прийнято позначати символами i і I (початкова буква латинського слова index). Буквою « i » позначаються індивідуальні (приватні) індекси, буквою « I »-загальні індекси. Знак внизу праворуч означає період: 0 - базисний; 1 - звітний.

Використовуються певні символи для позначення індексованих показників:

q - кількість (обсяг) виробленої продукції (або кількість проданого товару) даного виду в натуральному вираженні;

p - ціна одиниці продукції або товару;

z - собівартість одиниці продукції;

t - витрати робочого часу (праці) на виробництво одиниці продукції даного виду, тобто трудомісткість одиниці виробу;

$\wedge T$ - загальні витрати робочого часу (праці) на виробництво продукції даного виду або чисельність працівників підприємства, фірми і т.д.;

$w = q / T$ - виробництво продукції даного виду в одиницю часу або в розрахунку на одного робітника, тобто рівень продуктивності праці у вартісному вираженні;

v - вироблення продукції в натуральному вираженні на одного робітника або в одиницю часу;

$F = zq$ - загальні витрати на виробництво продукції даного виду;

$Q = pq$ - загальна вартість виробленої продукції даного виду або товарообіг.

Всі статистичні індекси можна класифікувати за такими ознаками:

За ступенем охоплення явища індекси бувають індивідуальні та зведені.

1.1. **Індивідуальні індекси** служать для характеристики зміни окремих елементів складного явища.

1.2. **Загальні (зведені) індекси** використовуються для вимірювання динаміки складного явища, складові частини якого безпосередньо несумірні.

2. За базою порівняння всі індекси можна розділити на дві групи:

2.1. **Динамічні** (базисні і ланцюгові) відображають зміни явищ у часі.

2.2. **Індекси виконання плану**

2.3. **Територіальні** застосовуються для міжрегіональних порівнянь. Велике значення ці індекси мають як у міжнародній статистиці, так і під час оцінювання стану розвитку територій.

3. За видом ваги індекси бувають:

3.1. Індекси з постійними вагами (стандартними, звітного періоду, базисного періоду)

3.2. Індекси зі змінними вагами

4. За формою побудови розрізняються індекси агрегатні і середні зважені (арифметичні і гармонійні).

4.1. **Агрегатна форма** загальних індексів є основною формою економічних індексів.

4.2. **Середні зважені** (гармонійні і арифметичні) це похідні, вони виходять в результаті перетворення агрегатних індексів.

5. За складом явища можна виділити дві групи індексів: постійного складу і змінного

складу. Ділення індексів на ці дві групи використовується для аналізу динаміки середніх показників.

6. За характером обсягу дослідження загальні індекси поділяються на індекси кількісних (індекс фізичного обсягу продукції) і якісних показників (як ціна, собівартість, врожайність, продуктивність праці, заробітна плата). В основі такого поділу індексів лежить вид величини, що індексується.

7. За періодом обчислення індекси поділяються на річні, квартальні, місячні, тижневі.

За допомогою статистичних індексів вирішуються низка **завдань**, принципи визначення яких такі:

- вимір динаміки соціально-економічного явища за два і більше періодів часу;
- вимір динаміки середнього статистичного показника;
- вимірювання співвідношення показників у різних регіонах;
- визначення ступеня впливу змін значень одних показників на динаміку інших;
- перерахунок значень, наприклад, макроекономічних показників з фактичних цін у порівнянні запланованої ціни.

2. Середньоарифметичний та середньограмонійний індекси. Індексний метод вимірювання динаміки середнього рівня інтенсивного показника

Індивідуальний індекс - це характеризує динаміку рівня досліджуваного явища в часі за два порівнюваних періоду або виражає співвідношення окремих елементів сукупності. Індекси отримують в результаті порівняння однотоварного явищ (*i*).

Індивідуальні індекси являють собою відносні величини динаміки, виконання плану, порівняння, вони бувають: фізичного обсягу продукції, собівартості, цін, трудомісткості і тощо.

| | | |
|---|---|--|
| Індивідуальний індекс фізичного обсягу продукції | Індивідуальний індекс ціни | Індивідуальний індекс вартості продукції |
| $i_q = q_1/q_0$ | $i_p = p_1/p_0$ | $i_{pq} = p_1q_1/p_0q_0$ |
| q_1, q_0 - кількість продажу окремого товару в поточному і базисному періоді. | p_1, p_0 - ціна проданого товару в поточному і базисному періоді. | p_1q_1, p_0q_0 - вартість продукції в поточному і базисному періоді. |

В економічних розрахунках найчастіше використовуються **агрегатні індекси**, які характеризують зміну сукупності в цілому.

^ **Агрегатний індекс** - складний відносний показник, який характеризує середню змінну соціально-економічного явища, що складається з несумірних елементів.

^ **Загальний індекс** - характеризує узагальнюючі результати спільної змінної всіх одиниць, що утворюють статистичну сукупність.

Особливість однієї з цих форм індексу полягає в тому, що в агрегатній формі безпосередньо порівнюються дві суми однойменних показників. Чисельник і знаменник агрегатного індексу є суму творів двох величин, одна з яких змінюється (індексируемая величина), а інша залишається незмінною в чисельнику і знаменнику (вага індексу).

Індексуючою величиною називається ознака, зміна якого вивчається (ціна товарів, курс акцій, витрати робочого часу на виробництво продукції, кількість проданих товарів тощо).

Вага індексу - це величина, що служить для цілей порівняння індексованих величин.

Агрегатна формула загального індексу цін була вперше запропонована в 1864 р. німецьким ученим Е. Ласпейреса. Він пропонував будувати агрегатний індекс цін, прийнявши в якості ваг продукцію базисного періоду q_0 :

$$I_p^L = \frac{\sum p_1 q_0}{\sum p_0 q_0}$$

У такому вигляді, тобто побудований по продукції базисного періоду, цей індекс відомий як **індекс цін Ласпейреса**.

У 1874 р. інший німецький вчений Г. Пааше, запропонував будувати агрегатний індекс цін по продукції поточного періоду q_1 :

$$I_p^П = \frac{\sum p_1 q_1}{\sum p_0 q_1}$$

Такий індекс, тобто побудований по продукції поточного періоду, відомий як **індекс цін Пааше**.

На практиці використовуються формули індексів цін і Ласпейреса, і Пааше, хоча вони дають різні результати. (За значенням індекс Ласпейреса, як правило, більше індексу Пааше.). Кожен з цих індексів має свої особливості, яким віддається перевага в конкретних умовах використання.

На початку ХХ в. американський економіст І. **Фішер** запропонував замість формул індексів цін Ласпейреса і Пааше використовувати середню геометричну з них, тобто корінь квадратний з твору індексів цін Ласпейреса і Пааше:

$$I_p^Ф = \sqrt{\frac{\sum p_1 q_1}{\sum p_0 q_1} \times \frac{\sum p_1 q_0}{\sum p_0 q_0}} = \sqrt{I_p^L \times I_p^П}$$

Цей індекс Фішер назвав *«ідеальним»*, оскільки в ньому не віддається перевага ні продукції базисного періоду, ні продукції поточного періоду.

3. Середні індекси з індивідуальних (групових)

Загальні індекси можуть бути обчислені не тільки як агрегатні, але і як середні з індивідуальних або групових. Наприклад, якщо є дані про зміну цін на конкретні товари, то, природно, з таких індивідуальних індексів можуть бути розраховані загальні (зведені) індекси як середні величини, причому зважені.

Середній індекс - це індекс, обчислений як середня величина з індивідуальних індексів. Оскільки існує кілька форм (видів) середніх величин, то при розрахунку середніх індексів, насамперед, виникає питання про форму середньої і про вагу.

У статистичній практиці середні індекси розраховуються переважно у формі середнього арифметичного і середнього гармонійного індексів:

| | | |
|---|---|---|
| $\bar{I}_{\text{арифм}} = \frac{\sum if}{\sum f}$ | і | $\bar{I}_{\text{гарм}} = \frac{\sum M}{\sum \frac{M}{i}}$ |
|---|---|---|

де: i - індивідуальні індекси досліджуваного показника (величини, що індексується);

f і M - ваги індексу.

| | | |
|---|---|---|
| Середній арифметичний індекс фізичного обсягу | Середній арифметичний індекс цін | |
| $\bar{I}_q = \frac{\sum i_q p_0 q_0}{\sum p_0 q_0}$ | Ласпейреса | Пааше |
| | $\bar{I}_p^L = \frac{\sum i_p p_0 q_0}{\sum p_0 q_0}$ | $\bar{I}_p^П = \frac{\sum i_p p_1 q_0}{\sum p_1 q_0}$ |
| Середній гармонійний індекс обсягу | Середній гармонійний індекс цін | |
| $\bar{I}_q = \frac{\sum q_1 p_0}{\sum \frac{q_1 p_0}{i_q}}$ | $\bar{I}_p^L = \frac{\sum q_0 p_1}{\sum \frac{q_0 p_1}{i_p}}$ | $\bar{I}_p^П = \frac{\sum q_1 p_1}{\sum \frac{q_1 p_1}{i_p}}$ |

3. Індекси змінного та постійного складу. Індекс структури

При вивченні якісних показників часто доводиться розглядати зміну в часі (або просторі) середньої величини показника, що індексується, для певної однорідної сукупності.

Будучи зведеною характеристикою якісного показника, середня величина складається як під впливом значень показника у індивідуальних елементів (одиниць), з яких складається об'єкт, так і під впливом співвідношення їх ваг («структури» об'єкта).

Якщо будь-який якісний індексований показник позначити через x , а його ваги - через f , то динаміку середнього показника можна відобразити як за рахунок зміни обох факторів (x і f), так і за рахунок кожного фактора окремо. В результаті отримуємо три різних індексу:

- індекс змінного складу;
- індекс фіксованого складу;
- індекс структурних зрушень.

Методика розрахунку індексів

| ^ Індекс змінного складу ($I_{пс} = I_{фс} \cdot I_{сс}$) | Індекс фіксованого складу | Індекс структурних зрушень |
|---|---|--|
| $I_{пс} = \frac{\bar{x}_1}{\bar{x}_0} = \frac{\sum x_1 f_1 / \sum x_0 f_0}{\sum f_1 / \sum f_0}$ | $I_{фс} = \frac{\bar{x}_1}{\bar{x}_{усл}} = \frac{\sum x_1 f_1 / \sum x_0 f_1}{\sum f_1 / \sum f_1}$ | $I_{сс} = \frac{\bar{x}_{усл}}{\bar{x}_0} = \frac{\sum x_0 f_1 / \sum x_0 f_0}{\sum f_1 / \sum f_0}$ |
| Відображає динаміку середнього показника (для однорідної сукупності) за рахунок зміни величини, що індексується x у окремих елементів (частин цілого) і за рахунок зміни ваг f , якими зважуються окремі значення x . | Відображає динаміку середнього показника лише за рахунок зміни величини, що індексується x , при фіксуванні ваг на рівні, як правило, звітного періоду f_1 | Відображає динаміку середнього показника лише за рахунок зміни ваг f при фіксуванні величини, що індексується на рівні базисного періоду x_0 . |
| Характеризує динаміку середніх величин не тільки за рахунок зміни величини, що | Виключає вплив зміни структури (складу) сукупності на динаміку середніх величин, | - |

| | | |
|---|--|---|
| індексується у окремих елементів (частин цілого), але і за рахунок зміни питомої ваги цих частин в загальній сукупності, тобто зміни складу сукупності. | тобто він характеризує динаміку середніх величин, розрахованих для двох періодів по одній і тій же фіксованою структурі ваг. | |
| Цей індекс показує, як змінилася середня ціна певного виду товару, реалізована за різними цінами на різних ринках, за рахунок двох факторів: p - зміни цін на окремих ринках і q - зміни кількості (частки) товарів, реалізованих на різних ринках; | Цей індекс, усуваючи вплив структурного фактора на динаміку середніх цін, визначає середня зміна цін на даний товар на всіх ринках, тобто по всій сукупності реалізованої продукції; | Даний індекс характеризує зміну середньої ціни товару за рахунок структурного фактора, тобто зміни часток продукції |

Якщо в індексах середніх рівнів в якості ваг використовуються питомі ваги одиниць сукупності в загальній чисельності сукупності (частка $d = f \Sigma / f$), то система індексів може бути записана в наступному вигляді:

$$I_{pc} = \frac{\sum x_1 d_1}{\sum x_0 d_0} I_{\phi c} = \frac{\sum x_1 d_1}{\sum x_0 d_1} I_{cc} = \frac{\sum x_0 d_1}{\sum x_0 d_0}$$

Контрольні питання:

1. Суть індексів та їх роль у економіко-статистичному аналізі. Методологічні принципи побудови індексів.
2. Середньоарифметичний та середньогармонійний індекси. Індексний метод вимірювання динаміки середнього рівня інтенсивного показника.
3. Індеси змінного та постійного складу. Індекс структури.
4. Системи співзалежних індексів і визначення впливу окремих факторів.
5. Територіальні індекси.
6. індекс цін *Ласпейреса*.
7. Індивідуальний індекс
8. індекс цін Пааше.
9. індекс змінного складу;
10. індекс фіксованого складу;
11. індекс структурних зрушень
12. Методика розрахунку індексів

ТЕМА 8. СТАТИСТИЧНЕ ВИВЧЕННЯ ДИНАМІКИ

1. *Поняття про ряди динаміки, види та типи динамічних рядів*

Основна мета статистичного вивчення динаміки комерційної діяльності полягає у виявленні і вимірюванні закономірностей їх розвитку у часі. Це досягається за допомогою побудови і аналізу статистичних рядів динаміки.

Рядами динаміки називаються статистичні дані, що відображають розвиток досліджуваного явища в часі. У кожному ряду динаміки є два основних елементи: показник часу t ; відповідні їм рівні розвитку досліджуваного явища y . В якості показань часу в рядах динаміки виступають або певні дати (моменти) часу, або окремі періоди (роки, квартали, місяці, добу).

Рівні рядів динаміки відображають кількісну оцінку (міру) розвитку в часі досліджуваного явища. Вони можуть виражатися абсолютними, відносними або середніми величинами.

Залежно від характеру досліджуваного явища рівні рядів динаміки можуть відноситися або до певних дат (моментів) часу, або до окремих періодах. У відповідності з цим, ряди динаміки поділяються на моментні та інтервальні.

Моментні ряди динаміки відображають стан досліджуваних явищ на певні дати (моменти) часу.

Особливістю моментного ряду динаміки є те, що в його рівні можуть входити одні й ті ж одиниці досліджуваної сукупності. Так, основна частина персоналу фірми N, складова Облікова чисельність на 1.01.1994г., Що продовжує працювати протягом даного року, відображена в рівнях наступних періодів. Тому при підсумовуванні рівнів моментного ряду динаміки може виникнути повторний рахунок.

Інтервальні ряди динаміки відображають підсумки розвитку (функціонування) досліджуваних явищ за окремі періоди (інтервали) часу.

Особливістю інтервального ряду динаміки є те, що кожен його рівень складається з даних за більш короткі інтервали часу. Наприклад, підсумовуючи товарообіг за перші три місяці року, отримують його обсяг за I квартал, а сума товарообігу чотирьох кварталів дає обсяг товарообігу за рік і т.д.

Ряди динаміки можуть бути повними і неповними.

Повний ряд - ряд динаміки, в якому однойменні моменти часу або періоди часу строго слідує один за іншим в календарному порядку чи рівновіддалена один від одного.

Неповний ряд динаміки - ряд, в якому рівні зафіксовані в нерівновіддаленими моменти або періоди часу.

2. *Характеристика динамічних рядів і способи їх розрахунку та побудови*

Приведення рядів динаміки в порівнянний вид.

Ряди динаміки, що вивчають зміна статистичного показника, можуть охоплювати значний період часу, протягом якого можуть відбуватися події, що порушують порівнянність окремих рівнів ряду динаміки (зміна методології обліку, зміна цін і т.д.).

Для того, щоб аналіз ряду був об'єктивний, необхідно враховувати події, що призводять до непорівнянності рівнів ряду і використовувати прийоми обробки рядів для приведення їх у порівнянний вид.

Найбільш характерні випадки непорівнянності рівнів ряду динаміки: Територіальні зміни об'єкта дослідження, до якого відноситься досліджуваний показник (зміна меж міського району, перегляд адміністративного поділу області і т.д.). Різновеликі інтервали часу, до яких відноситься показник. Так, наприклад, в лютому - 28 днів, у березні - 31 день, аналізуючи зміни показника по місяцях, необхідно враховувати різницю в кількості днів.

Зміна дати обліку. Наприклад, чисельність поголів'я худоби в різні роки могла визначатися за станом на 1 січня або на 1 жовтня, що в даному випадку призводить до непорівнянності. Зміна методології обліку або розрахунку показника.

Зміна цін.

Зміна одиниць вимірювання.

Визначення середнього рівня ряду динаміки.

В якості узагальненої характеристики рівнів ряду динаміки служить середній рівень ряду динаміки \bar{y} . В залежності від типу ряду динаміки використовуються різні розрахункові формули.

Інтервальний ряд абсолютних величин з рівними періодами (інтервалами часу):

$$\bar{y} = \frac{\sum y_i}{n}$$

Моментний ряд з рівними інтервалами між датами:

$$\bar{y} = \frac{1}{2} y_1 + y_2 + \dots + y_{n-1} + \frac{1}{2} y_n$$

Моментний ряд з нерівними інтервалами між датами:

$$\bar{y} = \frac{\sum y_i t_i}{\sum t_i}$$

де y_i - Рівні ряду, що зберігаються без зміни протягом інтервалу часу t_i .

Показники зміни рівнів ряду динаміки.

Одним з найважливіших напрямків аналізу рядів динаміки є вивчення особливостей розвитку явища за окремі періоди часу.

З цією метою для динамічних рядів розраховують ряд показників:

K - темпи зростання;

Δy - Абсолютні прирости;

K_{Δ} - Темпи приросту.

Темп зростання - відносний показник, що виходить у результаті ділення двох рівнів одного ряду один на одного. Темпи зростання можуть розраховуватися як ланцюгові, коли кожен

рівень ряду порівнюється з попереднім йому рівнем: $K_{\delta} = \frac{y_i}{y_{i-1}}$, Або як базисні, коли всі

рівні ряду зіставляються з одним і тим же рівнем y_0 , Обраним за базу порівняння: $K_{\delta} = \frac{y_i}{y_0}$

. Темпи зростання можуть бути представлені у вигляді коефіцієнтів або у вигляді відсотків. Абсолютний приріст - різниця між двома рівнями ряду динаміки, має ту ж розмірність, що і рівні самого ряду динаміки. Абсолютні прирости можуть бути ланцюговими і базисними, в залежності від способу вибору бази для порівняння:

ланцюгової абсолютний приріст - $\Delta y_{\delta} = y_i - y_{i-1}$;

базисний абсолютний приріст - $\Delta y_{\delta} = y_i - y_0$.

Для відносної оцінки абсолютних приростів розраховуються показники темпів приросту.

Темп приросту - відносний показник, що показує на скільки відсотків один рівень ряду динаміки більше (або менше) іншого, прийнятого за базу для порівняння.

Базисні темпи приросту: $\left| K_{\dot{a}} = \frac{y_{\dot{a}}}{y_0} \right|$.

Ланцюгові темпи приросту: $\left| K_{\delta} = \frac{y_{\delta}}{y_{i-1}} \right|$.

$\left| y_{\dot{a}i} \right|$; $\left| y_{\delta i} \right|$ - Абсолютний базисний або ланцюгової приріст;

y_0 - Рівень ряду динаміки, обраний за базу для визначення базисних абсолютних приростів;

y_{i-1} - Рівень ряду динаміки, обраний за базу для визначення і-го ланцюгового абсолютного приросту.

Існує зв'язок між темпами зростання і приросту:

$\left| K = K - 1 \right|$ або $\left| K = K - 100\% \right|$ (якщо темпи зростання визначені у відсотках).

Якщо розділити абсолютний приріст (ланцюговий) на темп приросту (ланцюговий) за відповідний період, отримаємо показник, званий - абсолютне значення одного відсотка

приросту: $A = \frac{y_{\delta}}{K_{\delta}}$.

Визначення середнього абсолютного приросту,
середніх темпів зростання і приросту.

За показниками зміни рівнів ряду динаміки (абсолютні прирости, темпи зростання і приросту), отриманим в результаті аналізу вихідного ряду, можуть бути розраховані узагальнюючі показники у вигляді середніх величин - середній абсолютний приріст, середній темп зростання, середній темп приросту.

Середній абсолютний приріст може бути отриманий за однією з формул:

$$\left| \bar{y} = \frac{\sum y_{ui}}{n} \right| \text{ або } \left| \bar{y} = \frac{y_n - y_1}{n - 1} \right|$$

де n - число рівнів ряду динаміки;

y_1 - Перший рівень ряду динаміки;

y_n - Останній рівень ряду динаміки;

$\left| y_{ui} \right|$ - Ланцюгові абсолютні прирости.

Середній темп зростання можна визначити, користуючись формулами:

$$\bar{K} = \sqrt[n]{K_{1u} * K_{2u} * \dots * K_{nu}}$$

$$\bar{K} = \sqrt[n]{\frac{y_n}{y_0}}$$

$$\bar{K} = \sqrt[n]{\frac{K_{n\delta}}{K_{1\delta}}}$$

де n - число розрахованих ланцюгових або базисних темпів зростання;

y_0 - Рівень ряду, прийнятий за базу для порівняння;

y_n - Останній рівень ряду;

K_{iu} - Ланцюгові темпи зростання (у коефіцієнтах);

K_{16} - Перший базисний темп зростання;

K_{n6} - Останній базисний темп зростання.

Між темпами приросту K_i темпами зростання K існує співвідношення $K_i = K - 1$, аналогічне співвідношення вірно і для середніх величин.

Визначення в рядах динаміки загальної тенденції розвитку. Визначення рівнів ряду динаміки на протязі тривалого періоду часу обумовлено дією ряду факторів, які неоднорідні по силі та напрямку впливу, що чиниться на досліджуване явище. Розглядаючи динамічні ряди, намагаються розділити ці фактори на постійно діючі та надають визначальний вплив на рівні низки, що формують основну тенденцію розвитку, і випадкові фактори, що призводять до короткочасних змін рівнів ряду динаміки. Найбільш важлива при аналізі ряду динаміки його основна тенденція розвитку, але часто по одному лише зовнішньому вигляду ряду динаміки її встановити неможливо, тому використовують спеціальні методи обробки, що дозволяють показати основну тенденцію ряду. Методи обробки використовуються як прості, так і досить складні. Найпростіший спосіб обробки ряду динаміки, застосовуваний з метою встановлення закономірностей розвитку - метод укрупнення інтервалів. Суть методу в тому, щоб від інтервалів, або періодів часу, для яких визначено вихідні рівні ряду динаміки, перейти до більш тривалим періодам часу і подивитися, як рівні ряду змінюються в цьому випадку.

Інший спосіб визначення тенденції у низці динаміки - метод ковзних середніх. Суть методу полягає в тому, що фактичні рівні низки замінюються середніми рівнями, обчисленими за певним правилом, наприклад: $y_1, y_2, y_3, y_4, y_5, \dots, y_n$ -

Вихідні чи фактичні рівні ряду динаміки замінюються середніми рівнями:

$$y_1 = \frac{y_1 + y_2 + y_3 + y_4 + y_5}{5}$$

$$y_2 = \frac{y_2 + y_3 + y_4 + y_5 + y_6}{5}$$

$$y_3 = \frac{y_3 + y_4 + y_5 + y_6 + y_7}{5}$$

У результаті виходить згладжений ряд, що складається з ковзних п'ятизвенная середніх рівнів $y_1, y_2, y_3, y_4, y_5, \dots, y_{n-2}$. Між розташуванням рівнів y_i і \bar{y}_i встановлюється відповідність:

$$y_1, y_2, y_3, y_4, y_5, \dots, y_{n-3}, y_{n-2}, y_{n-1}, y_n$$

згладжений ряд коротший вихідного на число рівнів $\frac{k-1}{2}$, Де k - число рівнів, обраних для визначення середніх рівнів ряду.

Згладжування методом ковзних середніх можна виробляти за чотирма, п'яти або іншому числу рівнів ряду, використовуючи відповідні формули для усереднення вихідних рівнів.

Отримані при цьому середні рівні називаються чотириланкової ковзними середніми, п'ятизвенная ковзними середніми і т.д.

При згладжуванні ряду динаміки по парним числом рівнів виконується додаткова операція, звана центруванням, оскільки, при обчисленні ковзного середнього, наприклад по чотирьох

рівнях, $\bar{y}_1 = \frac{y_1 + y_2 + y_3 + y_4}{4}$ відноситься до тимчасової точці між моментами часу, коли були

зафіксовані фактичні рівні y_2 і y_3 . Схема обчислень і розташувань рівнів згладженого ряду стає складніше:

$y_1, y_2, y_3, y_4, y_5, y_6 \dots$ - Вихідні рівні;
 - $\bar{y}_1, \bar{y}_2, \bar{y}_3 \dots$ - Згладжені рівні;
 - $\bar{y}_{1ц}, \bar{y}_{2ц} \dots$ - Центровані згладжені рівні;

$$\bar{y}_{1ц} = \frac{\bar{y}_1 + \bar{y}_2}{2} \quad \bar{y}_{2ц} = \frac{\bar{y}_2 + \bar{y}_3}{2}$$

Метод ковзних середніх не дозволяє отримати чисельні оцінки для вираження основної тенденції у низці динаміки, даючи лише наочне графічне представлення.

Найбільш досконалим способом визначення тенденції розвитку в ряду динаміки є метод аналітичного вирівнювання. При цьому методі вихідні рівні ряду динаміки y_i замінюються теоретичними чи розрахунковими \hat{y}_i , які представляють із себе деяку досить просту математичну функцію часу, виражає загальну тенденцію розвитку ряду динаміки. Найчастіше в якості такої функції вибирають пряму, параболу, експоненту і ін

Наприклад, $\hat{y}_i = a_0 + a_1 t_i$,

де a_0, a_1 - Коефіцієнти, що визначаються в методі аналітичного вирівнювання;
 t_i - Моменти часу, для яких були отримані вихідні та відповідні теоретичні рівні ряду динаміки, що утворюють пряму, яка визначається коефіцієнтами a_0, a_1 .

Розрахунок коефіцієнтів a_0, a_1 ведеться на основі методу найменших квадратів:

$$\sum_{i=1}^n (y_i - \hat{y}_i)^2 \rightarrow \min$$

Якщо замість y_i підставити $a_0 + a_1 t_i$ (Або відповідне вираження для інших математичних функцій), отримаємо:

$$\sum_{i=1}^n (a_0 + a_1 t_i - y_i)^2 \rightarrow \min$$

Це функція двох змінних a_0, a_1 (Всі t_i і y_i відомі), яка за певних a_0, a_1 досягає мінімуму. З цього виразу на основі знань, отриманих у курсі вищої математики про екстремуми функції п змінних, отримують значення коефіцієнтів a_0, a_1 .

Для прямої:

$$a_0 = \frac{\sum y_i \sum t_i^2 - \sum t_i y_i \sum t_i}{n \sum t_i^2 - \sum t_i \sum t_i}$$

$$a_1 = \frac{n \sum y_i t_i - \sum t_i \sum y_i}{n \sum t_i^2 - \sum t_i \sum t_i}$$

де n - число моментів часу, для яких були отримані вихідні рівні ряду y_i .

Якщо замість абсолютного часу t_i вибрати умовний час таким чином, щоб $\sum t_i = 0$, То записані вирази для визначення a_0, a_1 спрощуються:

$$a_0 = \frac{\sum y_i}{n} \quad a_1 = \frac{\sum t_i y_i}{\sum t_i^2}$$

Визначення в рядах внутрігодової динаміки.

Багато процесів господарської діяльності, торгівлі, сільського господарства та інших сфер людської діяльності піддані сезонним змінам, наприклад, продаж морозива, споживання електроенергії, виробництво молока, цукру, продаж сільгосппродукції та ін

Для аналізу рядів динаміки, схильних до сезонних змін, використовуються спеціальні методи, що дозволяють встановити і описати особливості зміни рівнів ряду. Перш, ніж використовувати методи вивчення сезонності, необхідно підготувати дані, наведені в порівнянний вид, за декілька років спостереження по місяцях або кварталах. Зміни сезонних коливань здійснюється за допомогою індексів сезонності. В залежності від існуючих в ряду динаміки тенденцій використовуються різні правила побудови індексів.

1. Ряд динаміки не має загальної тенденції розвитку, або вона не велика.

$$\text{Індекс сезонності: } I_{s_i} = \frac{\bar{y}_i}{\bar{y}}$$

де \bar{y}_i - Середній рівень ряду, отриманий в результаті осереднення рівнів ряду за однойменні періоди часу (наприклад, середній рівень січня за всі роки спостереження);

\bar{y} - Загальний середній рівень ряду за весь час спостереження.

Висновок про наявність або відсутність у низці динаміки яскраво вираженої тенденції може проводитися, наприклад, за допомогою методу укрупнення інтервалів.

2. Ряд динаміки має загальну тенденцію, і вона визначена або методом ковзного середнього, або методом аналітичного вирівнювання.

$$\text{Індекс сезонності } \bar{I}_{s_i} = \left[\sum \frac{y_i}{y_i^*} \right] : n$$

де y_i - Вихідні рівні ряду; y_i^* - Рівні ряду, отримані в результаті визначення ковзних середніх для тих же періодів часу, що й вихідні рівні:

I - номер місяця або кварталу, для якого визначається індекс сезонності:

n - число років спостереження за процесом.

Контрольні питання

1. Поняття про ряди динаміки, види та типи динамічних рядів.
2. Характеристика динамічних рядів та способи їх розрахунку.
3. Основні прийоми обробки динамічних рядів.
4. Метод змикання рядів динаміки.
5. Приведення динамічних рядів до єдиної основи.
6. Основні прийоми аналізу рядів динаміки (методи визначення, основні тенденції розвитку).
7. Метод збільшення інтервалів.
8. Метод плинних середніх.
9. Метод аналітичного вирівнювання.
10. Метод сезонних коливань. Індекс сезонності.

ЛІТЕРАТУРА:

1. Айвазян С.А. Прикладная статистика: Исследование зависимостей [Электронный ресурс] / С. А. Айвазян, И. С. Енюков, Л. Д. Мешалкин. – Режим доступа: <http://uadoc.zavantag.com/text/1222235/index-4.html>.
2. Вашків П. Г. Статистика підприємства : навч. посіб. / П. Г. Вашків, П. І. Пастер, В. Й. Сидорчук ; за ред. П. Г. Вашків.
3. Єріна А. М. Теорія статистики : практикум / А. м. Єріна, З. О. Пальян. – К. : Т-во «Знання» ; КОО, 1997. – 325 с.
4. Ефимова М. Р. Общая теория статистики : учеб. / М. Р. Ефимова, Е. В. Петрова, В. И. Ремянцев. – М. : ИНФА-М, 1998. – 416 с.
5. Мармоза А. Т. Теория статистики / А. Т. Мармоза. – К. : Ельга ; Ника-Центр, 2003. – 392 с.
6. Мостовий Г. І. Статистика : практикум / Г. І. Мостовий, А. О. дегтяр, В. К. Горкавий. – Х. : ХарРІ НАДУ «Магістр», 2004. – 128 с.
7. Наумік К. Г. Регіональна соціально-економічна політика. Статистичні методи в державному управлінні : навч. посіб. / К. Г. Наумік, Т. О. Коваль. – Х. : Вид-во ХНЕУ, 2012. – 148 с.
8. Опря А. Т. Статистика (модульний варіант з програмованою формою контролю знань) : навч. посіб. – К. : Центр учбової літератури, 2012. – 448 с.
9. Саврас І. З. Статистичні методи в державному управлінні : навч. посіб. / І. З. Саврас. – Львів : ЛРІ НАДУ, 2010. – 132 с.
10. Статистика : підруч. / С. С. Герасименко [та ін.]. – К. : КНЕУ, 1998. – 468 с.
11. Теорія статистики : навч. посіб. / Г. І. Мостовий, А. О. дегтяр, В. К. Горкавий, В. В. Ярова ; за ред. проф. Г. І. Мостового. – Х. : ХарРІ НАДУ «Магістр», 2002. – 300 с.
12. Уманець Т. В. Статистика : навч. посіб. / Т. В. Уманець, Ю. Б. Пігарев. – К. : Вікар, 2003. – 623 с.

ЗМІСТ

| | Стор. |
|--|-----------|
| ВСТУП | 1 |
| Тема 1 Вступ до теорії економіко-статистичного аналізу. Статистичне спостереження..... | 2 |
| Тема 2 Зведення та групування статистичних даних. Графічний метод зображення..... | 9 |
| Тема 3 Абсолютні, відносні та середні величини Абсолютні, відносні та середні величини..... | 13 |
| Тема 4 Статистичне вивчення варіації та форми розподілу..... | 19 |
| Тема 5 Вибірковий метод..... | 24 |
| Тема 6 Статистичні методи вимірювання взаємозв'язків..... | 30 |
| Тема 7 Індекси..... | 36 |
| Тема 8 Статистичне вивчення динаміки..... | 41 |
| Література..... | 47 |

Навчальне видання

Батир Юрій Георгійович

«СТАТИСТИЧНІ МЕТОДИ В ДЕРЖАВНОМУ УПРАВЛІННІ»

Конспект лекцій

Підписано до друку 26.01. 2018р. формат 60X84/16
Ум. друк. арк. 2,0. Тираж 100 примірників.
Оригінал – макет підготував Ю.Г.Батир

Видавництво РВВ НУЦЗУ, 2018