

ЛЕКЦІЯ №4

СТАТИЧНІ ДАНІ І СТОХАСТИЧНА МОДЕЛЬ



- 1. Екологічні дані. Цілі і завдання збору статистичних даних
- 2 Зведення та групування статистичних даних
- 3. Статистичні показники
- 4. Середні характеристики динамічного ряду

1. Екологічні дані. Цілі і завдання збору статистичних даних

- Закономірності в еколого-географічних дослідженнях виражаються у вигляді зв'язків і залежностей еколого-географічних показників, математичних моделей їх поведінки. Такі залежності і моделі можуть бути отримані тільки шляхом обробки реальних статистичних даних, із врахуванням внутрішніх механізмів і випадкових чинників. Модель може бути отримана й апробована на основі аналізу статистичних даних, і зміни у поведінці останніх говорять про необхідність уточнення і розвитку моделі.

- Математична статистика (тобто теорія обробки і аналізу даних) і її застосування в еколого-географічних дослідженнях дає можливість будувати еколого-географічні моделі й оцінити їх параметри, перевіряти гіпотези про властивість цих показників і формах їх зв'язку, що у кінцевому результаті служить основою для еколого-географічного аналізу і прогнозування, створюючи можливість для прийняття обґрунтованих рішень.

- Статистичні дані в екології є основою для виявлення і обґрунтування емпіричних закономірностей. Без конкретних кількісних даних, що характеризують функціонування досліджуваних еколого-географічних об'єктів, не завжди можна визначити практичне значення застосованої моделі, навіть якщо метою є виявлення переважно якісних закономірностей.

- Метою збору еколого-географічних даних є отримання інформаційної бази для прийняття рішень. Природно, що аналіз даних і прийняття рішень проводиться на основі якої-небудь інтуїтивної (нереальної) або кількісної (реальної) моделі. Тому збирають необхідні дані, які потрібні для відповідної моделі.

- Існують різноманітні методи збору еколого-географічних даних: шляхом опитування, анкетування, отримання офіційної статистичної звітності тощо. У багатьох країнах існують статистичні органи, які займаються збиранням, обробкою, поширенням і публікацією важливих даних. Цією діяльністю займаються також багато спеціалізованих державних і приватних агентств.

- Першим етапом будь-якого дослідження є збирання інформації, а саме, статистичне спостереження.
- Статистичне спостереження – це спланована, науково організована реєстрація масових даних про соціально-економічні й екологічні явища та процеси.
- Статистичне спостереження може бути первинним або вторинним. Первинне – це реєстрація даних, що надходять безпосередньо від об'єкта, який їх продукує. Вторинне – збирання раніше зареєстрованих та оброблених даних.

Статистичні дані – це масові системні кількісні характеристики соціально-економічних явищ і процесів.

Статистичні дані мають відповідати певним вимогам:

- бути вірогідними;
- повними;
- своєчасними;
- порівнянними за часом або у просторі;
- доступними.

- *Підготовка спостереження*
починається із складання
плану спостереження –
сукупності програмно-
методологічних та
організаційних питань.

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

для чого проводиться обстеження (мета обстеження);

що обстежують (об'єкт обстеження);

складові частини об'єкта (одиниця сукупності);

джерело інформації (одиниця спостереження);

на які питання планується одержати відповіді (програма спостереження).



Мета спостереження – одержання статистичних даних, які є підставою для узагальненої характеристики стану та розвитку явища або процесу. Кінцевою метою є підготовка управлінських рішень та прийняття заходів.



Об'єкт спостереження – сукупність явища, що підлягають обстеженню. Чітке визначення суті та меж об'єкта дає можливість уникнути різного тлумачення результатів. Для цього застосовуються ценси – набори кількісних та якісних обмежувальних ознак.



Одиниця сукупності – це первинний елемент об'єкта, що є носієм ознак, які підлягають реєстрації.

- *Ознаковою* є властивість, що відбиває сутність, характер та особливість одиниці сукупності. За формою виразу ознаки можуть бути : описовими (атрибутивними) та кількісними. Кількісні ознаки можуть бути дискретними або неперервними.
- *Дискретні* – ознаки, що набувають лише окремих, ізольованих значень. Як правило, вони виникають в результаті лічби (кількість суб'єктів діяльності, чисельність акціонерів).
- *Неперервні* – ознаки, що набувають будь-яких значень у певних межах.
- Відомості про ознаки одиниць сукупності збираються від одиниць спостереження. *Одиниця спостереження* – це первинна одиниця, від якої одержують інформацію.

- Програма спостереження – це перелік запитань, на які слід одержати відповіді в результаті спостереження. Зміст та обсяг запитань залежить від мети спостереження та можливостей його проведення (грошових та трудових витрат терміну реєстрації) У програму спостереження також включаються: розробка статистичного інструментарію, визначення виду та способу обстеження.
- Статистичний інструментарій – це набір статистичних формулярів, інструкцій та роз'яснень щодо проведення спостереження.

- Статистичний формуляр – обліковий документ єдиного зразка, що містить адресну характеристику об'єкта спостереження та статистичні дані про нього. Статистичними формулярами є звіти, переписні та опитувальні листки, бланки документів, анкети. При складанні формулярів враховується не тільки зміст та інформативність ознак, а й можливість їх статистичної обробки. Остання забезпечується завдяки застосуванню системи шкал.

Шкала – це засіб упорядкування та кількісного вираження ознак. Використовуються такі види шкал:

номінальна – шкала найменувань, що встановлює відношення подібності елементів, за якою порядок розташування ознак значення не має;

порядкова (рангова) – шкала, що встановлює послідовність інтенсивності прояву ознаки;

метрична – кількісна шкала, в основу якої покладено результати безпосереднього вимірювання.

Другою складовою частиною плану спостереження є організаційні питання, що визначають:

хто проводить
спостереження
(органи та персонал);

де проводить (місце
спостереження);

за допомогою чого
(матеріально-
технічне
забезпечення);

спосіб забезпечення
точності результатів
(система контролю та
пробні обстеження);

коли проводить (час
та період
спостереження).

- Час спостереження (об'єктивний час) – це час, до якого відносяться дані спостереження. Якщо об'єктом спостереження є процес, то визначається інтервал часу, впродовж якого накопичуються дані. Якщо об'єктом спостереження є стан, то обирається критичний момент – момент часу, станом на який реєструються дані.
- Період спостереження (суб'єктивний час) – час, упродовж якого здійснюється реєстрація даних. Наприклад, суб'єктивний час перепису населення – 8 днів (з 12.01 по 20.01). Вибір часу спостереження здійснюється у найсприятливіший період (найменша міграція населення; спокійна соціально-політична ситуація).

- Контроль даних спостереження – це засіб попередження, виявлення та виправлення помилок спостереження, який полягає у перевірці даних на повноту та вірогідність. Повнота даних перевіряється візуально, а вірогідність – шляхом логічного та арифметичного контролю.
- Логічний контроль – це перевірка сумісності даних шляхом порівняння співзалежних ознак, яка встановлює тільки наявність помилки, а не її величину.
- Арифметичний контроль – перевірка зареєстрованих даних шляхом прямих або непрямих перерахунків. Залежно від причини виникнення виділяють помилки: реєстрації – виникають при будь-якому спостереженні внаслідок неправильного встановлення фактів, репрезентативності – властиві лише вибірковому спостереженню, виникають внаслідок несучільного характеру реєстрації або порушення принципів випадковості відбору.

За природою
виникнення
розрізняють
такі помилки
реєстрації:

випадкові – внаслідок збігу випадкових обставин, через неуважність; недбалість; некваліфікованість реєстратора або незосередженість респондента (їх дія у масі випадків врівноважується і на результати не впливає);

систематичні – внаслідок постійних спотворень в одному напрямі (їх дія призводить до зсуву результатів у бік збільшення або зменшення).

Систематичні помилки бувають ненавмисними і навмисними



Навмисні помилки
— наслідок
свідомого
викривлення
фактів з певною
метою
(погіршення або
прикрашання
дійсності).

Ненавмисні
помилки — через
необґрунтованість
програми
спостереження,
некомпетентність
реєстраторів,
неосвіченість
респондентів.

- Різноманітність сфер спостереження обумовлює застосування різних організаційних форм, видів та способів статистичного спостереження.

Використовуються
такі організаційні
форми
спостереження:

звітність;

спеціально організовані
спостереження;

реєстри.

- Звітність – форма спостереження, за якою кожен суб'єкт діяльності регулярно подає свої дані до державних органів статистики у вигляді документів (звітів) спеціально затвердженої форми.

Спеціально
організовані
спостереження
— форма
спостереження,
яка охоплює
сфери життя, що
не уловлюються
звітністю. До
них належать :

переписи,

обліки,

спеціальні обстеження,

опитування.

Перепис – суцільне спостереження масових явищ з метою визначення їх розміру та складу станом на певну дату. Переписи проводяться одночасно для всієї сукупності (території) за єдиною для всіх одиниць програмою, як правило, періодично з рівним інтервалом (перепис населення – кожні 10 років).

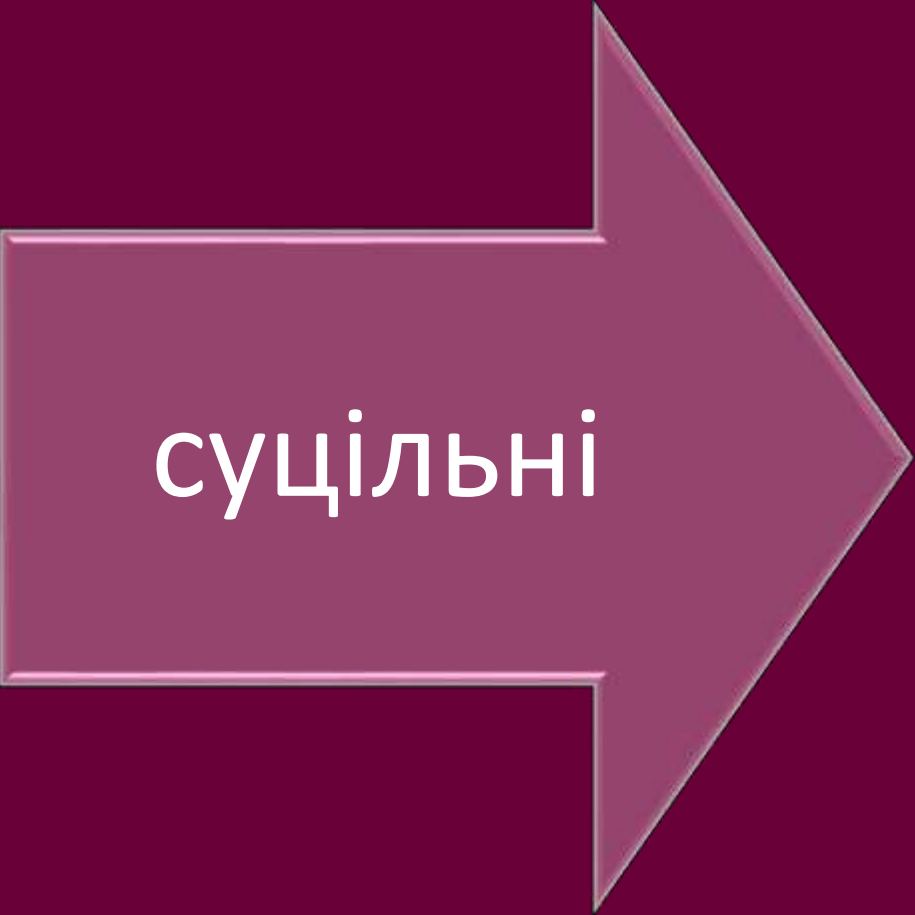
Обліки – суцільні спостереження масових явищ, що ґрунтуються на даних опитування, огляду та документальних записів (облік земельного фонду, облік об'єктів незавершеного будівництва).

Спеціальні обстеження – несуцільні спостереження окремих масових явищ за певною тематикою.

Опитування – частіше несуцільні спостереження думок, мотивів, оцінок, які реєструються зі слів респондентів. Винятком з цього є суцільне опитування всього населення – референдум (масове волевиявлення з принципових соціально-політичних чи екологічних питань).

- Статистичні реєстри – списки або переліки одиниць певного об'єкта спостереження із зазначенням необхідних ознак, які постійно оновлюються та поповнюються (реєстр населення – поіменний список мешканців регіону із зазначенням їх паспортних даних; реєстр суб'єктів діяльності; реєстр домашніх господарств; реєстр земельного фонду).
- Види спостереження розрізняють за ступенем охоплення одиниць та за часом реєстрації даних.

За ступенем охоплення одиниць
обстеження бувають:



суцільні



несуцільні

- *Суцільні обстеження* передбачають реєстрацію всіх без винятку одиниць сукупності (звітність, більшість переписів).

Несуцільні – обстеження за яких реєструються не всі одиниці сукупності, а лише їх певна частина

вибіркове спостереження

обстеження основного масиву

монографічне обстеження

анкетне спостереження

моніторинг

Вибіркове – спостереження, коли реєструється певна частина одиниць сукупності, відібрана у випадковому порядку.

Обстеження основного масиву – реєстрація даних щодо переважної частини одиниць сукупності, які є визначники для характеристики об'єкта спостереження (обстеження міст з найвищим рівнем атмосферного забруднення та ін.).

Монографічне обстеження – детальне обстеження окремих одиниць сукупності з метою їх досконального вивчення.

Анкетне спостереження, при якому не всі розповсюджені реєстраційні формуляри (анкети) повертаються з відповідями.

Моніторинг – спеціально організоване систематичне спостереження за станом певного середовища (моніторинг екологічного стану природних водойм).

За часом
реєстрації
даних
спостереження
поділяються на:

поточне

періодичне

одноразове

Поточне спостереження – систематична реєстрація фактів перебігу явища у міру його виникнення або стосовно безперервного процесу.

Періодичне – спостереження, що проводиться через певні, як правило, рівні, проміжки часу (перепис населення або виробничих потужностей).

Одноразове – спостереження, що проводиться у міру виникнення потреби дослідження явища або процесу.

Статистичне спостереження здійснюється такими способами:



Безпосередній облік – реєстрація фактів здійснюється особисто обліковцем шляхом підрахунку, вимірювання, оцінки, огляду.

Документальний облік – реєстрація фактів відбувається за даними, наведеними у документах первинного обліку.

Опитування – це спостереження, що здійснюється експедиційним шляхом, самореєстрацією, кореспондентським чи анкетним шляхами.

Експедиційне опитування – реєстрація фактів спеціально підготовленими обліковцями з одночасною перевіркою точності реєстрації.

Самореєстрація – реєстрація фактів самими респондентами після попереднього інструктажу.

Кореспондентське опитування – реєстрація фактів на місцях виникнення явищ добровільно обраними особами, які надсилають результати у відповідні інстанції.

Анкетне опитування – реєстрація думок, намірів і мотивів респондентів шляхом їх самостійного заповнення анкети.

2. Зведення та групування статистичних даних

- Статистичні зведення – другий етап дослідження масових суспільних явищ. Суть його полягає в класифікації та агрегуванні первинних статистичних даних.
- Існують загальноприйняті методологічні стандарти розподілу сукупностей на групи – чітко визначені групувальні ознаки та сформульовані вимоги щодо умов формування груп. Це класифікації.
- Для розв'язання конкретних аналітичних задач проводяться нестандартні групування за певними ознаками, що легко розпізнаються. Групування за однією ознакою називають простим, у разі поєднання двох і більше ознак – комбінаційним.

На групування у статистичному аналізі
покладаються певні функції, відповідно до
яких їх розрізняють

структурні

типологічні

аналітичні

- *Структурне групування* характеризує склад однорідної сукупності за певними ознаками, обсяги явища та вагомість окремих груп.
- *Типологічне групування* – це розподіл якісно неоднорідної сукупності на класи, соціально-економічні типи, однорідні групи.
- *Аналітичне групування* – виявлення наявності та на прямого зв'язку між двома ознаками, з яких одна представляє результат. У класичному варіанті аналітичного групування сукупність поділяється на групи за факторною ознакою, і в кожній групі визначається середній рівень результативної ознаки. За наявності зв'язку між факторною та результативною ознаками групові середні від групи до групи поступово змінюються – збільшуються або зменшуються.

- При формуванні груп постає питання про їх кількість та межі кожної з них. Кількість груп залежить від ступеня варіації групувальної ознаки та обсягу сукупності, у кожному окремому випадку її необхідно обґрунтувати. Якщо групувальна ознака атрибутивна, кількість груп певною мірою визначається кількістю найменувань ознаки.
- У процесі формування груп за варіативною ознакою – неперервною або дискретною, з широким діапазоном варіації – необхідно встановити інтервали груп та визначити межі кожного з них з такою точністю, щоб розподіл сукупності був однозначним. Інтервали бувають рівні та нерівні, відкриті та закриті.
- Рівні інтервали використовують за умови, що значення ознаки x у діапазоні варіації змінюється рівномірно.

- Наприклад, еродовані землі в межах області коливаються від 15–65%. При $m=5$ ширина інтервалу становить , а межі інтервалу відповідно 15–25; 25–35; 35–45; 45–55; 55 і більше. Позаяк межі інтервалів збігаються, то порядок віднесення до груп межових значень ознаки визначають слова останнього відкритого інтервалу “55 і більше”, тобто нижню межу закритого інтервалу слід вважати “включно”, а верхню – “виключно”.

- У разі, коли діапазон значень ознаки надто широкий і розподіл сукупності за цією ознакою нерівномірний, використовують нерівні інтервали. Наприклад, розподіл сумарного обсягу викидів по адміністративних районах області, тис. т на рік : до 3; 3–4,9; 5–9,9; 10–19,9; 20–49,9. Оскільки межі інтервалу не збігаються, то обидві межі (верхню і нижню) слід вважати “включно”.

- Невід'ємним елементом зведення та групування є статистична таблиця, в якій зведена інформація подається компактно, у зручній для порівняння та аналізу формі. За логічним змістом статистична таблиця розглядається як “статистичне речення”, підметом якого є об'єкт дослідження, а присудком – система показників, що характеризує об'єкт. Залежно від структури підмета статистичні таблиці поділяються на прості, групові та комбінаційні. Підметом простої таблиці є перелік елементів сукупності, територіальний або хронологічний ряд. У груповій таблиці підметом є групування за однією ознакою, у комбінаційній – за двома і більше ознаками.

- Складання статистичної таблиці має два етапи. На першому проектується макет таблиці, на другому таблиця заповнюється статистичними даними. Макет статистичної таблиці – це комбінація горизонтальних рядків та вертикальних граф, на перетині яких утворюються клітини. Ліві бічні та верхні клітини призначені для словесних заголовків – переліку складових підмета та системи показників присудка, решта – для числових даних. Основний зміст таблиці вказується у назві.

3. Статистичні показники

- Статистичний показник – це узагальнююча характеристика соціально-економічного явища чи процесу, в якій поєднуються якісна та кількісна визначеність останнього.
- Якісний зміст показника залежить від суті явища (процесу) і знаходить своє відображення у назві (народжуваність, прибутковість тощо).
- Кількісний бік явища представляють числа та його вимірник.

Показники розрізняють за:

способом
обчислення

ознакою
часу

своїми
функціями

За способом
обчислення
розрізняють:

первинні
показники

похідні
показники

- *Первинні* визначаються шляхом зведення та групування даних і подаються у формі абсолютних величин.
- *Похідні* показники обчислюються на базі первинних або вторинних показників і мають форму середніх чи відносних величин.

За ознакою
часу показники
поділяються на:

інтервальні

моментні

- *Інтервальні* характеризують явище за певний час (день, місяць, рік).
- До *моментних* відносять показники, які характеризують явище на певний момент часу.

- *Абсолютні статистичні величини* характеризують розміри соціально-економічних явищ – обсяги сукупності або обсяги значень певних ознак. Це іменовані числа. Залежно від конкретної задачі дослідження та характеру явища використовують натуральні, трудові та вартісні (грошові) одиниці вимірювання. Якщо виникає потреба звести воєдино кілька різновидностей однієї споживчої властивості, обсяги такого явища виражаються в умовно натуральних одиницях . Перерахунок в умовні одиниці здійснюють за допомогою спеціальних коефіцієнтів - сумірників. Наприклад, паливний баланс складається у тонах умовного палива. Еталоном слугує кам'яне вугілля, теплотворна спроможність якого становить 7000 кал на 1 кг. Калорійний коефіцієнт донецького вугілля –0,9; природного газу – 1,2 і т.д.

- *Відносні величини* характеризують кількісні співвідношення різнойменних чи однойменних показників. Будь-яка відносна величина представляє собою дріб, чисельником якого є порівняльна величина, а знаменником – база порівняння. Відносна величина показує, у скільки разів порівнянна величина більша базисної або яку частку вона становить відносно базисної, іноді скільки одиниць одної величини припадає на 100, на 1000, на 10000, на 100000 одиниць іншої. За аналітичною функцією виділяють відносні величини інтенсивності, динаміки, просторового порівняння, структури, координації.

- *Відносна величина інтенсивності* характеризує ступінь поширення явища у певному середовищі. Це іменована величина, у якій поєднуються одиниці вимірювання чисельника і знаменника. Наприклад, демографічні коефіцієнти (народжуваність, смертність) на 1000 чол. населення, забезпеченість лікарями на 10000 чол. населення. та ін.
- *Відносна величина динаміки* характеризує напрямок та інтенсивність зміни явища у часі, розраховується співвідношенням значень показника за два періоди чи моменти часу. При цьому базою порівняння може бути або попередній рівень, або рівень, більш віддалений у часі.

- У територіально-просторових порівняннях співвідносять однойменні показники, що характеризують різні об'єкти (підприємства, галузі) або території (міста, регіони, країни) і мають однакову часову визначеність. Інтерпретація цих величин залежить від бази порівняння.
- *Базою порівняння* може виступати певне еталонне значення показника (норматив, стандарт тощо). Відхилення відносної величини порівняння з еталоном від 1 або 100 % свідчить про порушення оптимального процесу.

- Відносна величина структури характеризує склад, структуру сукупності за тією чи іншою ознакою, обчислюється відношенням розміру складової частини до загального підсумку. Відносні величини структури називають частками, сума їх становить 1 або 100 %.
- На використанні часток ґрунтується порівняльний аналіз складу різних за обсягом сукупностей, оцінка структурних зрушень у часі. Різницю між частками називають процентними пунктами.
- Співвідношення між окремими складовими сукупності є відносними величинами координації.

4. Середні характеристики динамічного ряду

- Середня величина – це узагальнююча міра варіюючої ознаки, що характеризує її рівень у розрахунку на одиницю сукупності. Умовами застосування середніх величин є: наявність якісно однорідної сукупності та достатньо великий її обсяг.
- У статистичній практиці використовують декілька видів середніх : середнє арифметичне, середнє квадратичне і т.д.

- Середнє арифметичне – використовується для осереднення прямих значень ознак шляхом їх підсумування.
- Якщо дані незгруповані, використовується середня арифметична проста.

- Якщо моментів більше двох і інтервали між ними рівні, то середня обчислюється за формулою середньої хронологічної.
- Якщо дані згруповані, то використовують середньозважену арифметичну.
- Осередненню підлягають не тільки окремі значення варіант, а й їх групові середні, тоді вагою буде частота (частка) кожної групи
- Обчислена в такий спосіб середня з групових середніх називається загальною.

- Середня арифметична має певні математичні властивості, які розкривають її суть. Так, сума відхилень окремих варіант від середньої дорівнює нулю, а сума квадратів таких відхилень наближається до мінімуму. Ці дві властивості покладені в основу вивчення варіації ознак.
- Якщо окремі значення варіант збільшити (зменшити) на величину A або в k разів, то середня зміниться відповідно.
- Наприклад, якщо грошові внески громадян до ощадбанку скоригувати на рівень інфляції, що становить 1,2, то середній розмір внеску збільшиться відповідно в 1,2 рази.

- Середня не зміниться за пропорційної зміни усіх ваг, але її розмір зазнає змін при певних структурних зрушеннях.
- Наприклад, за незмінної курсової вартості акцій окремих емітентів середня вартість акцій може підвищуватись за рахунок збільшення частки “дорогих” акцій у загальній кількості їх продажу.

- Зазначені властивості середньої використовують у разі осереднення ознак порядкової (рангової) шкали. Варіанти ознак можна оцифровувати порядковими рангами $R_j = 1, 2, \dots, n$ або центрованими $R_o = R_j - 1/2(R_{\max} + R_{\min})$. Середній центрований бал відхиляється від середнього порядкового на величину $1/2(R_{\max} + R_{\min})$.
- Середній центрований бал набуває додатних або від'ємних значень і свідчить про позитивну чи негативну оцінку явища. Крім того, його використовують для порівняння оцінок різних явищ, оскільки він не залежить від розмірності шкали. Отже, рівень відношення до екологічної ситуації можна оцінити як задовільний, але поки що невисокий.

- Середня гармонічна використовується для осереднення обернених індивідуальних значень ознак шляхом їх підсумовування. Для незгрупованих даних це середня гармонічна простота. Якщо дані згруповані, то використовують *середню гармонічну зважену*.
- Середня геометрична визначається як добуток відносних величин динаміки x_i , які є кратним співвідношенням i -го значення показника до попереднього ($i-1$).

- *Приклад.* Кількість зареєстрованих екологічних злочинів за чотири роки зросла у 1,57 рази, у тому числі за перший рік – у 1,08, за другий – у 1,1, за третій – у 1,18, за четвертий – у 1,12 рази. Середньорічний темп зростання кількості зареєстрованих екологічних злочинів становить тобто число зареєстрованих екологічних злочинів зросло щорічно у середньому на 12 %.
- Якщо часові інтервали неоднакові, використовують середню геометричну зважену.

- Середня квадратична розглядається як характеристика варіації.
- Соціально-економічні явища надзвичайно складні та багатогранні. Будь-який показник відображає лише одну грань предмета пізнання. Комплексна характеристика останнього передбачає використання системи показників. Кожний показник системи має самостійне значення і водночас є складовою узагальнюючої властивості, що дає підстави для конструювання інтегральних оцінок явищ. Позаяк показники системи, як правило, різнойменні, то об'єднання їх в інтегральну оцінку передбачає стандартизацію – приведення до одного виду. При стандартизації індивідуальні значення показників замінюються рангами, балами, відносними величинами, стандартними відхиленнями тощо.

- Так, рейтингова оцінка передбачає, що кожний параметр оцінюється балами.
- При стандартизації за допомогою відносних величин базою порівняння може бути або еталонне значення (норма, стандарт) або середнє значення показника за сукупністю.